Milda BURZAŁA

ZASTOSOWANIE UNORMOWANEJ WARIANCJI RÓŻNICY DO BADANIA PODOBIEŃSTWA PRZEBIEGU SZEREGÓW CZASOWYCH

Badanie podobieństwa (zróżnicowania) przebiegu szeregów czasowych nabiera szczególnego znaczenia przy ustalaniu związków przyczynowo-skutkowych oraz przy prognozowaniu z wykorzystaniem metod analogowych.

Okazuje się, że jeśli obserwujemy dużą nieregularność rozwoju badanych wielkości, niestabilne związki przyczynowo-skutkowe czy też dysponujemy zbyt krótkimi szeregami czasowymi, to metody badania podobieństwa przebiegu szeregów czasowych stwarzają możliwość prognoz przez analogie, głównie historyczne i przestrzenno-czasowe.

Przewidując przyszłe wartości zmiennych, możemy wykorzystywać podobieństwo poziomu lub podobieństwo kształtu albo też obydwa kryteria jednocześnie. Pierwsze z nich dotyczy osiągania tych samych wartości. Drugie, dużo częściej wykorzystywane, dotyczy podobieństwa kształtu krzywych czasowych.

Celem artykułu jest krótki przegląd stosowanych miar podobieństwa kształtu szeregów czasowych, ocena ich przydatności w badaniach empirycznych oraz pokazanie możliwości wykorzystania prostego miernika opartego na wariancji różnicy. Omawiany wskaźnik jest prosty w budowie, uzyskane oceny stopnia podobieństwa są porównywalne ze wskazaniami innych miar. Zróżnicowanie wyników jest w wielu przypadkach dużo większe, co ma zasadnicze znaczenie przy podejmowaniu decyzji, stawianych diagnozach i w konsekwencji dla trafności prognoz.

1. Wybrane miary podobieństwa

Proponowane w literaturze miary podobieństwa wykorzystują różne kryteria porównań. Wynika to z potrzeby odmiennych zastosowań. Przykładowo w analizach strukturalnych i przy badaniu opóźnień czasowych porównujemy pewne wektory liczb. W zależności od typu analizy reprezentują one jednak odrębne własności porównywanych wielkości. Stąd przed przystąpieniem do badań należy precyzyjnie określić cel prowadzonych porównań. Łatwiej wtedy wybrać odpowiednią miarę podobieństwa. Często bowiem pojawia się problem niejednakowych wskazań.

Wśród wybranych miar można wyróżnić pewne typy, analizujące podobieństwo pod kątem

a) występujących współzależności,
b) zgodności struktur,
c) przebiegu wartości funkcji,
d) wahań okresowych.

Najpopularniejszą miarą **współzależności** jest współczynnik korelacji liniowej. Jest on miarą unormowaną, informującą o kierunku i sile (natężeniu) liniowego związku między zmiennymi. Jak wiadomo, istnieje prosta zależność między współczynnikiem regresji b_{yx} zmiennej Y względem zmiennej X a współczynnikiem korelacji r[1]:

$$b_{yx} = r\frac{s_y}{s_x}. \tag{1}$$

Stąd określenie współczynnik korelacji liniowej.

Mówimy, że współczynnik korelacji jest tym większy, im mniejszy jest kąt rozwarcia między liniami regresji o współczynnikach b_{yx} i b_{xy} (silniejszy związek liniowy między zmiennymi). Jest on powszechnie stosowany w ustalaniu wskaźników wiodących dla przebiegu cyklu koniunkturalnego, m.in. w Organizacji Współpracy Gospodarczej i Rozwoju (OECD)[2]. Jednakże współczynnik korelacji dwóch szeregów czasowych przesuniętych względem siebie o 1 lub 2 okresy w celu określenia opóźnienia czasowego sygnalizuje bardzo małe zróżnicowanie, często dopiero na 3 lub 4 miejscu po przecinku, co wydaje się pewnym ograniczeniem tej miary. Zdarza się też sytuacja, że między

---

[1] K. Zając, *Zarys metod stastycznych*, PWE, Warszawa 1994, s. 241.

[2] O. Brunet, *Calculation of composite leading indicators: a comparison of two different methods*, Paper for presentation at the CIRET Conference in Paris 2000; OECD Leading Indicators and Business Cycles in Member Countries, Sources and Methods 1960-1985, 1987 January, no. 39.

zmiennymi nie ma korelacji liniowej ($r \Rightarrow 0$), a jednak istnieje wyraźna współzależność.

Do miar **zgodności struktur** można zaliczyć przekształcenie wskaźnika niedokładności informacyjnej Theila przedstawione przez I. Kudrycką[3]

$$P(y:x)=\frac{1}{1+I(y:x)}, \tag{2}$$

gdzie:

$$I(y:x)=\sum_{t=1}^{T} y_t^* \log \frac{y_t^*}{x_t^*} \tag{3}$$

jest niedokładnością informacyjną szeregu czasowego dla zmiennej Y i zmiennej X, przy czym

$$x_t^*=\frac{x_t}{\sum_{j=1}^{T} x_j}, \quad y_t^*=\frac{y_t}{\sum_{j=1}^{T} y_j}, \quad \sum_{t=1}^{T} x_t^*=1, \quad \sum_{t=1}^{T} y_t^*=1. \tag{4}$$

Sumowanie przekształconych zmiennych do jedności jest prawdziwe dla wartości dodatnich. Wcześniejsza transformacja, polegająca na przesunięciu umownego zera, jest konieczna w przypadku ujemnych stóp wzrostu (np. wszystkie wartości szeregu zwiększamy o 100). Im wartość $I(y:x)$ jest mniejsza, tym mniej informacji potrzeba do określenia jednej struktury na podstawie znajomości innej, a wartość miary $P(y:x) \rightarrow 1$. Niedokładność informacyjna $I(y:x)=0$, gdy frakcje x_t^*, y_t^* obydwu zmiennych są sobie równe. Jeśli różnice frakcji rosną, to rośnie niepewność przy określaniu struktury zmiennej Y na podstawie znajomości struktury zmiennej X, więcej informacji brakuje. Gdy $I(y:x) \Rightarrow \infty$, wtedy $P(y:x) \Rightarrow 0$.

Miary zgodności struktur nie ograniczają się do zależności liniowych, ale często dają niewielkie zróżnicowanie wyników. Pojawia się wtedy problem wyboru spośród przebiegów podobnych – tych „bardziej podobnych" i „mniej podobnych".

Miary **podobieństwa funkcji** umożliwiają wymagane zróżnicowanie otrzymywanych wyników. Są za to bardzo pracochłonne ze względu na dużą wrażliwość miar na gęstość podziału całego zbioru obserwacji. Przykładem

---

[3] J. Kudrycka, *Metody badania koniunktury na podstawie szeregów czasowych*, Ekonomista 1995, nr 4; pierwotnie wprowadzony do porównywania zgodności rozkładów, jednakże z powodzeniem mógłby być stosowany do badania zgodności struktur.

może tu być propozycja M. Cieślak i J. Jasińskiego[4]. Podobieństwo funkcji f określonej w przedziale $P^{(0)}=\langle a,b\rangle$ oraz funkcji g określonej w przedziale $P^{(k)}=\langle c,d\rangle$ dla podziału odpowiednio $a_1,\ldots,a_{n+1}$; $c_1,\ldots,c_{n+1}$, przy czym $b-a=d-c$, określa liczba

$$m=\frac{1}{n}\sum_{i=1}^{n} m_i, \qquad -1<m\leqslant 1, \tag{5}$$

gdzie:

$$m_i=\begin{cases} 1-\dfrac{2}{\pi}\alpha_i, & 0\leqslant\alpha_i\leqslant\dfrac{\pi}{2} \quad \text{dla przypadku I,} \\ -\dfrac{\alpha_i}{\pi}, & 0<\alpha_i\leqslant\pi \quad \text{dla przypadku II,} \end{cases} \tag{6}$$

przypadek I oznacza tę samą monotoniczność funkcji f i g, przypadek II – różną monotoniczność,

α_i – miara łukowa kąta zawartego między prostą przechodzącą przez punkty $(a_i,f(a_i))$; $(a_{i+1},f(a_{i+1}))$ i prostą przechodzącą przez punkty $(c_i, g(c_i))$; $(c_{i+1}\, g(c_{i+1}))$.

Miara przyjmuje wartość 1, gdy odpowiednie odcinki łamanej są do siebie równoległe ($m_i=1$ dla każdego i). Wartość miary m jest malejącą funkcją liniową kąta zawartego między odpowiednimi odcinkami łamanej. Zróżnicowanie i dokładność otrzymywanych wyników rosną w miarę zwiększania liczby punktów podziału i tym samym zmniejszania odległości między nimi.

Innym kryterium określenia podobieństw szeregów czasowych jest **struktura harmoniczna** wahań okresowych. Generalnie problem polega na oszacowaniu funkcji gęstości spektralnej $f(\omega)$, która umożliwia rozpoznanie i porównanie wpływu poszczególnych rodzajów wahań na powstanie wariancji procesu stochastycznego. Funkcję gęstości spektralnej wyrażoną w zależności od funkcji kowariancyjnej można przedstawić w postaci[5]:

$$f(\omega)=\frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty} K(\tau)\cos\omega\tau. \tag{7}$$

W celu porównania dwóch szeregów czasowych sporządza się wykres przedziału ufności dla stosunku dwóch funkcji gęstości spektralnej[6]

---

[4] *Prognozowanie gospodarcze – metody i zastosowania*, red. M. Cieślak, WN PWN, Warszawa 1997; patrz przykłady w: M. Cieślak, R. Jasiński, *Miara podobieństwa funkcji*, Przegląd Stastyczny 1979, nr 3/4, s. 172.

[5] L. Talaga, Z. Zieliński, *Analiza spektralna w modelowaniu ekonometrycznym*, PWN, Warszawa 1986, s. 66.

[6] Ibidem, s. 105 i n.

$$P(R_j / F_{\frac{\alpha}{2}, k_1, k_2} < P_j < R_j / F_{1-\frac{\alpha}{2}, k_1, k_2}) = 1 - \alpha, \tag{8}$$

gdzie:

$$P_j = \frac{f_1(\omega_j)}{f_2(\omega_j)}, \qquad R_j = \frac{\hat{f}_1(\omega_j)}{\hat{f}_2(\omega_j)}, \tag{9}$$

$\hat{f}_1(\omega_j), \hat{f}_2(\omega_j)$ – estymatory funkcji gęstości spektralnej,

F_{k_1, k_2} – statystyka o rozkładzie Snedecora o k_1 i k_2 stopniach swobody.

Jeżeli $P = 1$ będzie leżała wewnątrz przedziału ufności dla wszystkich częstości, to badane szeregi będą miały tę samą strukturę harmoniczną.

Obecnie nie ma żadnych problemów z estymacją funkcji gęstości spektralnej od strony technicznej i progamowej. Przeszkodą jest jednak dostępność dostatecznie długich szeregów czasowych, którymi powinniśmy dysponować przy tego rodzaju analizie. Bardzo rzadko statystyki państwowe dają nam do dyspozycji takie szeregi dla wielkości makroekonomicznych.

2. Postulaty stawiane wobec miary podobieństwa

Oznaczenia:

y, x – porównywane szeregi czasowe dla dwóch zmiennych Y i X obserwowanych w przedziale czasowym tej samej długości,

$\mu(y, x)$ – miara podobieństwa przebiegu dwóch szeregów czasowych y i x.

Wśród postulatów, które powinna spełniać miara podobieństwa najczęściej wymienia się:

1° unormowanie

$$\mu \in \langle a, b \rangle, \qquad -\infty < a, b < \infty, \tag{10}$$

umożliwiające ocenę stopnia zróżnicowania;

2° zwrotność, tzn. największe podobieństwo i jednocześnie najmniejsze zróżnicowanie w przypadku porównania szeregu czasowego z samym sobą

$$\mu(y, y) = g, \tag{11}$$

gdzie – w zależności od budowy miary: $g = a$ lub $g = b$, lub g stanowi środkową wartość przedziału $\langle a, b \rangle$;

3° symetryczność, czyli

$$\mu(y,x)=\mu(x,y); \tag{12}$$

4° brak reakcji na przesunięcie wszystkich wartości szeregu, tzn. jeśli

$$\forall t: y_t=x_t+c, \qquad c \text{ – stała przesunięcia,}$$

to

$$\mu(y,x)=\mu(y,y); \tag{13}$$

5° reakcja na zmianę monotoniczności.

3. Określenie miary opartej na wariancji różnicy

Załóżmy, że porównujemy przebieg dwóch szeregów czasowych x i y o tej samej liczbie obserwacji. Porównywane szeregi czasowe dotyczyć muszą **zmiennych jednoimiennych** wyrażonych w tych samych jednostkach miary z zachowaniem skali lub **różnoimiennych pozbawionych miana** w wyniku normalizacji.

Budujemy nowy szereg czasowy z, którego wartości z_t będą przedstawiały różnice kolejnych wartości szeregu x_t od kolejnych wartości szeregu y_t, tzn.

$$\forall t: z_t=y_t-x_t. \tag{14}$$

Dla szeregów podobnych dyspersja tak określonych różnic powinna być niewielka. Powszechnie stosowaną miarą dyspersji jest wariancja. Proponowana miara podobieństwa szeregów *UWR* to próba unormowania wariancji różnic.

Określmy miarę podobieństwa szeregów y i x jako:

$$UWR(y,x)=1-\frac{s_z^2}{s_y^2+s_x^2}, \tag{15}$$

gdzie:

s_z^2 – wariancja wartości szeregu z,
s_x^2 – wariancja szeregu x,
s_y^2 – wariancja szeregu y.

Miara tak zbudowana jest unormowana, przyjmuje wartości z przedziału $\langle -1, 1\rangle$, ponieważ

$$|\mathrm{cov}(y,x)| \leqslant \frac{s_y^2 + s_x^2}{2}. \tag{16}$$

Przedstawiona zależność (16) wynika ze wzoru na wariancję różnicy[7]

$$s_z^2 = s_y^2 + s_x^2 - 2\,\mathrm{cov}(y,x). \tag{17}$$

Podobną zależność można wyprowadzić dla wariancji sumy. Kowariancja może być dodatnia lub ujemna. Ponieważ wariancja musi przyjmować wartości nieujemne, zwiększenie lub zmniejszenie sumy wariancji możliwe jest tylko w granicach określonych przez (16).

Warunek 1° jest więc spełniony, podobnie 2° i 4°, bo otrzymujemy wariancję stałej wartości. W przypadku 3° przesunięcie szeregu różnic na osi współrzędnych nie ma wpływu na miarę rozproszenia. Wymagania dotyczące reakcji na zmianę monotoniczności również są zachowane.

Miara uwzględnia nie tylko wariancję zmiennych, ale również skalę podobieństwa i kierunek korelacji.

$UWR = 0$, gdy zmienne nie wykazują podobieństwa w zakresie współzmienności (kowariancja równa jest 0, zmienne X i Y są nieskorelowane).

$UWR \to 1$, gdy kowariancja jest „duża" i dodatnia (wartości szeregów x_t, y_t przyjmują jednocześnie duże wartości i jednocześnie małe wartości), dla szeregów identycznych $UWR = 1$.

$UWR \to -1$, gdy kowariancja jest „duża", ale ujemna (wartości szeregów zachowują się przeciwnie).

Pozostaje pytanie, jaka powinna być wartość miary UWR, aby przebiegi szeregów uznać za podobne pod względem kształtu. Można przyjąć umownie, że dla szeregów podobnych powinno zachodzić:

$$UWR > p, \tag{18}$$

gdzie:

p – ustalona empirycznie granica (np. 0,95).

W szczególnym przypadku porównywania dwóch szeregów uzyskanych w wyniku standaryzacji wyjściowych wartości unormowaną miarę zróżnicowania może stanowić sama wariancja różnicy.

---

[7] Wyprowadzenie wzoru znajduje się w: K. Zając, op. cit., s. 195.

4. Możliwości zastosowań i wnioski

Okazało się, że duże zróżnicowanie i „zdecydowane" wskazanie na największe podobieństwo kształtu daje wariancja szeregu różnic badanych wielkości. Nieunormowana wariancja różnicy mogłaby stanowić dobre kryterium wyboru szeregów najmniej zróżnicowanych z danego zbioru. Pozostaje jednak wtedy pytanie, czy stopień podobieństwa jest wystarczający?

Przedstawione podejście można stosować zarówno do szeregów czasowych, jak i danych przestrzennych.

Miara umożliwia badanie całego zbioru obserwacji, co nie wyklucza podziału na podzbiory.

Propozycję można wykorzystać przy ustalaniu opóźnień czasowych.

W badaniach przebiegu cyklu koniunkturalnego informacja o opóźnieniach czasowych umożliwia podział zmiennych na wskaźniki wyprzedzające, równoczesne i opóźnione. Jeżeli dwie wielkości ekonomiczne znajdują się w różnej fazie cyklu koniunkturalnego, to wariancja szeregu różnic jest na pewno większa od wariancji szeregu różnic odpowiadających tej samej fazie cyklu.

Przy prognozowaniu przez analogie warto wcześniej zastosować podobieństwo poziomu w celu wyeliminowania z listy wzorców pozycji charakteryzujących się dużo niższym poziomem rozwoju.

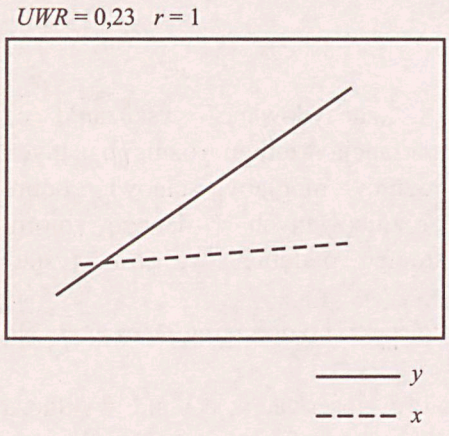
Uzyskane szeregi różnic mogą stanowić podstawę dalszych analiz w celu określenia wartości granicznych czy też wykrycia pewnych regularności obserwowanych odstępstw.

5. Przykłady

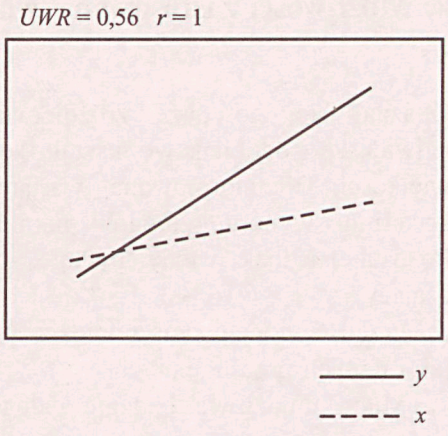
Wszystkie obliczenia zostały przeprowadzone przy użyciu arkusza kalkulacyjnego Excel. W celu uzyskania wartości *UWR* wykorzystano funkcję programową obliczenia wariancji.

Przykład 1

Pierwszy z przykładów obrazowo pokazuje różnice w odniesieniu do wyników wskazywanych przez współczynnik korelacji r i przedstawioną miarę *UWR*. Badamy podobieństwo przebiegu dwóch szeregów czasowych x i y.



Rys. 1



Rys. 2

Współczynnik korelacji liniowej nie spełnia roli miernika podobieństwa kształtu w przypadku tej samej postaci analitycznej trendów. Wartości obydwu szeregów czasowych x i y zmieniają się liniowo, ale w różnym tempie. Na rysunku 1 prosta związana z szeregiem y rośnie szybciej niż prosta dla szeregu x. Nie ma to wpływu na wartość współczynnika korelacji r, który wskazuje maksymalne „dodatnie podobieństwo", tym samym nie reaguje na zmianę tempa wzrostu wartości szeregu x przedstawioną na rysunku 2. Różnice w przebiegu krzywych czasowych uwzględnia wartość miary *UWR* równa 0,23 w pierwszym przypadku i 0,56 w drugim, gdzie różnica tempa wzrostu wartości szeregów x i y jest mniejsza niż na rysunku 1 (podobieństwo większe). Interpretacja wyników na podstawie miary *UWR* nie przedstawia problemów. Szeregi x i y na rysunku 2 są bardziej podobne niż szeregi x i y na rysunku 1.

Przykład 2

Podstawowym problemem prognozowania przez analogie jest wybór zmiennych wiodących i określenie opóźnień czasowych. W przykładzie 2 porównamy wartości i wskazania według wybranych miar podobieństwa: r, $UWR(y,x)$, $P(y,x)$ i m. Ze względu na krótkie szeregi wartości nie analizujemy ich struktury harmonicznej.

W tabeli 1 przedstawiono wartości standaryzowanej stopy bezrobocia (SUR) wyrażonej w % dla Polski (POL), Szwecji (SWE), Francji (FRA) i Portugalii (PRT). Przypuszczamy, że przebieg wskaźników makroekonomicznych dla Polski w latach dziewięćdziesiątych może być podobny do przebiegu

Tabela 1. Standaryzowana stopa bezrobocia SUR (w %)

Rok	FRA	SWE	PRT	POL
1982	7,7	3,3		
1983	8,1	3,7	7,8	
1984	9,7	3,3	8,5	
1985	10,1	2,9	8,7	
1986	10,2	2,7	8,4	
1987	10,4	2,2	6,9	
1988	9,8	1,8	5,5	
1989	9,3	1,5	4,9	
1990	9	1,7	4,6	
1991	9,5	3,1	4	
1992	10,4	5,6	4,2	
1993	11,7	9,1	5,7	14
1994	12,3	9,4	7	14,4
1995	11,7	8,8	7,3	13,3
1996	12,4	9,6	7,3	12,3
1997	12,3	9,9	6,8	11,2
1998	11,7	8,3	5,1	10,6

Źródło: *OECD Main economic indicators*, 1960-1998.

wskaźników makroekonomicznych innych krajów europejskich, które na drogę gospodarki rynkowej weszły wcześniej od nas. Wybór obiektów wzorcowych w porównaniach międzynarodowych jest problemem trudnym i skomplikowanym. Poprawne przeprowadzenie etapu weryfikacji wstępnej listy obiektów znacznie poszerzyłoby ramy tej pracy. Z tego powodu przyjęto tak ogólne kryterium doboru i pominięto problem porównań poziomu rozwoju gospodarczego, społecznego czy uwarunkowań kulturowych.

W tabelach 2-4 przedstawiono wyniki badania podobieństwa kształtu krzywej SUR dla Polski i wybranego wzorca według analizowanych wcześniej miar. Ze względu na krótki szereg danych dla Polski (sześć obserwacji) obliczenia przeprowadzono bez podziału na podzbiory. Pierwsza kolumna w tabelach 2-4 określa opóźnienie szeregu dla Polski względem wybranego kraju, tzn. opóźnienie = 0 oznacza badanie dwóch szeregów w tym samym okresie, lata 1993-1998, opóźnienie = 4 oznacza porównywanie szeregu dla Polski z lat 1993-1998 oraz wzorca z lat 1989-1994. Dla dostępnych szeregów czasowych przeanalizowano wszystkie możliwe opóźnienia. Dla zapewnienia

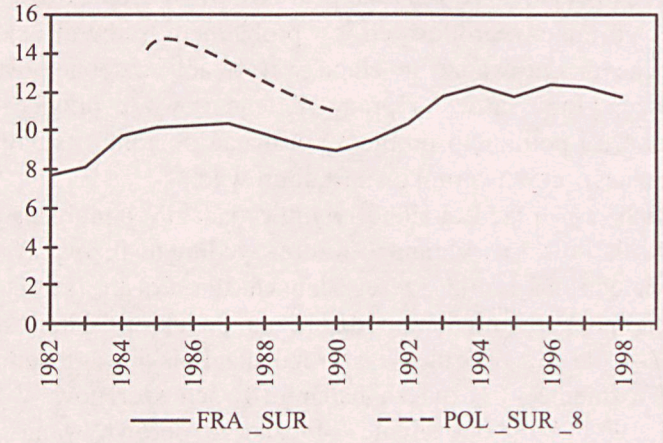
Tabela 2. Wartości miar podobieństwa kształtu szeregu SUR dla Polski i Francji

Opóźnienie	*r*	*UWR*	*P*	*m*
11	−0,8774	−0,8456	0,9895	−0,3477
10	−0,5163	−0,4333	0,9939	−0,2944
9	0,4855	0,2359	0,9979	0,6706
8	**0,9201**	**0,5834**	**0,9990**	**0,7577**
7	0,8588	0,5359	0,9988	0,7084
6	−0,0083	−0,0055	0,9967	−0,1901
5	−0,8036	−0,7300	0,9923	−0,3057
4	−0,9878	−0,9810	0,9888	−0,3621
3	−0,9077	−0,8995	0,9895	−0,3477
2	−0,7887	−0,7567	0,9918	−0,3574
1	−0,6236	−0,4941	0,9948	−0,3057
0	−0,0125	−0,0054	0,9972	−0,1901

Źródło: Obliczenia własne.

porównywalności wyników nie liczono wartości miary *m* w podzbiorach. W celu zachowania własności szeregu zrezygnowano również z wygładzania danych źródłowych.

W tabeli pogrubionym drukiem zaznaczono maksymalne wartości uzyskanych wskaźników podobieństwa. Dla porównywanych szeregów uzyskano jednakowe wskazanie opóźnienia czasowego (8 lat). Największe podobieństwo



Rys. 3. Stopa bezrobocia Francji i opóźniony szereg stopy bezrobocia dla Polski

szeregu czasowego dla stopy bezrobocia w Polsce w latach 1993-1998 obserwujemy z odpowiednim szeregiem dla Francji w latach 1985-1990.

Wyznaczenie poziomu granicznego miary, który gwarantowałby dostateczne podobieństwo dla celów prognostycznych należałoby przeprowadzać dla każdej miary osobno. Już pierwsze wyniki dają pogląd na uzyskiwane zróżnicowanie wyników. Tylko podobieństwo według P(POL, FRA) mierzone jest w skali od 0 do 1. Zarówno współczynnik korelacji $r=0{,}9201$, jak i miara P(POL, FRA)$=0{,}999$ wykazują bardzo duże podobieństwo kształtu. Miary m(POL, FRA)$=0{,}7577$ i UWR(POL, FRA)$=0{,}5834$ przyjmują znacznie niższe wartości. Opóźnienie szeregu dla Polski o kolejny rok powoduje spadek UWR(POL, FRA) o 0,3475, natomiast wartość m(POL, FRA) zmniejszyła się tylko o 0,011. Znaczący spadek wartości miary przy zwiększeniu opóźnienia o jeden okres ułatwia podjęcie decyzji i wskazania diagnostyczne. Przy zmianie o 0,011 pojawia się problem istotności zmian i niezawodności wykorzystywanej miary podobieństwa.

Tabela 3. Wartości miar podobieństwa kształtu szeregu SUR dla Polski i Szwecji

Opóźnienie	r	UWR	P	m
11	**0,9684**	0,5991	**0,9991**	0,7577
10	0,9557	**0,7209**	0,9962	**0,8510**
9	0,9635	0,7195	0,9943	0,8398
8	0,9178	0,5984	0,9954	0,7698
7	−0,0968	−0,0677	0,9827	−0,2155
6	−0,7744	−0,7743	0,9357	−0,3802
5	−0,9225	−0,7448	0,8795	−0,4989
4	−0,9875	−0,7175	0,8728	−0,4895
3	−0,8954	−0,6793	0,9073	−0,4948
2	−0,7417	−0,6425	0,9546	−0,4814
1	−0,5634	−0,5631	0,9851	−0,4162
0	0,1454	0,0960	0,9971	0,7208

Źródło: Obliczenia własne.

Analiza tych samych wartości miar podobieństwa w odniesieniu do szeregów Polski i Szwecji pokazuje, że wskazania nie zawsze muszą być identyczne. Z przeprowadzonych badań[8] wynika, że niejednokrotnie mogą to być różnice 3,

[8] Przedstawione wyniki stanowią fragment szerszych badań podobieństwa kształtu szeregów czasowych dla wybranych wskaźników makroekonomicznych.

a nawet 4 okresów. Pojawia się wtedy pytanie o trafność ocen podobieństwa, ich przydatności w wyznaczaniu wskaźników wiodących i prognozowaniu. Empiryczna weryfikacja wybranych miar podobieństwa ułatwia wybór odpowiedniego narzędzia badań.

Dla szeregów czasowych Polski i Szwecji porównywanych w tym samym przedziale, lata 1993-1998, otrzymaliśmy bardzo rozbieżne wskazania – niewielkie podobieństwo według współczynnika korelacji i jeszcze mniejsze według unormowanej wariancji różnic. Wartość miary $m(\text{POL, SWE}) = 0{,}7208$ przy opóźnieniu równym 0 niewiele różni się od miary $m(\text{POL, FRA}) = 0{,}7698$ przy opóźnieniu 8 okresów. Jeśli przyjmiemy, że poprzednio wskazanie potwierdzało duże podobieństwo w sensie współczynnika korelacji r, to obecnie niewiele niższej wartości m odpowiada niewielki współczynnik korelacji ($r = 0{,}1454$). Bardziej wiarygodne wskazania uzyskalibyśmy, licząc miarę m w podprzedziałach lub zwiększając częstotliwość obserwacji.

Tabela 4. Wartości miar podobieństwa kształtu szeregu SUR dla Polski i Portugalii

Opóźnienie	r	UWR	P	m
10	0,8055	0,7868	**0,9982**	0,8943
9	**0,9835**	**0,9808**	0,9974	**0,9829**
8	0,9642	0,9537	0,9955	0,9429
7	0,8719	0,8697	0,9943	0,9208
6	0,8101	0,7593	0,9973	0,9350
5	0,0565	0,0421	0,9942	−0,2028
4	−0,7359	−0,7012	0,9829	−0,3167
3	−0,9676	−0,9655	0,9746	−0,3477
2	−0,9361	−0,9361	0,9736	-0,3758
1	−0,6673	−0,6506	0,9836	−0,3427
0	0,3518	0,3102	0,9959	−0,1521

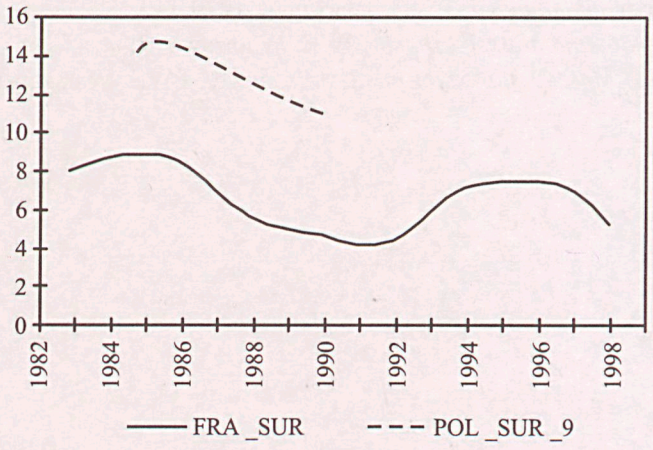
Źródło: Obliczenia własne.

Wyniki otrzymane dla $P(\text{POL, FRA})$, $P(\text{POL, SWE})$ i $P(\text{POL, PRT})$ wskazują na duże podobieństwo szeregów w większości analizowanych przedziałów. Wynika to z małych wartości mierników niedokładności informacyjnej $I(y, x)$. Przykładowo $I(\text{POL, PRT}) = 0{,}00414$ w latach 1993-1998 i $I(\text{POL, PRT}) =$

$=0{,}00176$ w latach 1983-1988. Niewielkie zróżnicowanie kolejnych wartości wynika z niewielkiego wpływu jednostkowych zmian wartości pierwotnych na zróżnicowanie frakcji x_t^* i y_t^*.

Podsumowanie

Na podstawie analizy wybranych wzorców można wysunąć wniosek, że szereg stopy bezrobocia dla Polski z lat 1993-1998 wykazuje największe podobieństwo z odpowiednim szeregiem dla Portugalii z lat 1984-1989 (rysunek 4). Prognozowanie przez analogie wzrostu bezrobocia w Polsce powinno być związane z analizą przebiegu zmian innych wskaźników makroekonomicznych i wymaga dalszych prac obliczeniowych.



Rys. 4. Stopa bezrobocia Portugalii i opóźniony szereg bezrobocia dla Polski

Przeprowadzone wnioskowanie potwierdza zasadność stosowania unormowanej wariancji różnicy do badania podobieństwa kształtu krzywych czasowych. Ocena trafności wyboru wymaga kontynuowania weryfikacji empirycznej.

Porównanie przedstawionych wyników nasuwa przypuszczenie, że miara $UWR(y,x)$ silniej niż inne wskaźniki reaguje na zmiany wartości szeregu.

Dla krzywych czasowych o bardzo podobnym przebiegu wybrane miary dają podobne wskazania. W miarę zmniejszania podobieństwa rośnie zróżnicowanie wartości przedstawionych miar. Wybór odpowiedniego miernika podobieństwa kształtu w konkretnych badaniach empirycznych powinien być jednym z ważniejszych etapów prognozowania analogowego.

Literatura

Brunet O., *Calculation of composite leading indicators: a comparison of two different methods*, Paper for presentation at the CIRET Conference in Paris 2000

Cieślak M., (red.) *Prognozowanie gospodarcze – metody i zastosowania*, WN PWN, Warszawa 1997

Cieślak M., Jasiński R., *Miara podobieństwa funkcji*, Przegląd Statystyczny 1979, nr 3/4.

Kudrycka I., *Metody badania koniunktury na podstawie szeregów czasowych*, Ekonomista 1995, nr 4

OECD Leading Indicators and Business Cycles in Member Countries, Sources and Methods 1960-1985, 1987 January, no. 39

Talaga L., Zieliński Z., *Analiza spektralna w modelowaniu ekonometrycznym*, PWN, Warszawa 1986

Theil H., *Zasady ekonometrii*, PWN, Warszawa 1979

Zając K., *Zarys metod statystycznych*, PWE, Warszawa 1994

Zeliaś A., (red.) *Metody statystyki międzynarodowej*, PWE, Warszawa 1988