

Katarzyna Cheba

Akademia Rolnicza w Szczecinie

ZASTOSOWANIE METODY WSKAŹNIKÓW SEZONOWOŚCI W PROGNOZOWANIU DLA DANYCH DEKADOWYCH

1. Wstęp

W prognozowaniu sezonowych szeregów czasowych istnieje możliwość pośredniego wykorzystania metod właściwych dla danych bez wahań sezonowych. Można zastosować w tym celu prezentowaną w literaturze [*Prognostowanie ekonomiczne...*, 2003, s. 90] metodę wskaźników sezonowości w wariancie addytywnym lub multiplikatywnym. Polega ona na bezpośrednim prognozowaniu zmiennych, z których wyeliminowano wahania sezonowe. Następnie do prognoz dla danych oczyszczonych dodaje się składniki sezonowości lub mnoży się je przez wskaźniki sezonowości.

W przypadku modelu addytywnego wyznacza się bezwzględne (absolutne) wahania sezonowe (g_j).

$$g_j = \frac{1}{N} \sum_{i=1}^N e_{t_{ij}} \quad (j=1, 2, \dots, m), \quad (1)$$

gdzie: $\sum_{i=1}^N e_{t_{ij}}$ – suma różnic między wartościami szeregu empirycznego i teoretycznego
w podokresie j ,

N – liczba okresów (cykli), m – liczba faz w cyklu, (sezonów).

Dla modelu multiplikatywnego dla poszczególnych faz cyklu wyznacza się tzw. surowe wskaźniki sezonowości (w_j):

$$w_j = \frac{1}{N} \sum_{i=1}^N u_{t_{ij}} \quad (j=1, 2, \dots, m), \quad (2)$$

gdzie: $\sum_{i=1}^N u_{ij}$ – suma ilorazów wartości szeregu empirycznego i teoretycznego (trendu) w okresie j .

Następnie dokonuje się obliczenia oczyszczonych wskaźników lub składników w ten sposób, aby spełnić odpowiednie warunki:

$$\sum_{j=1}^m \hat{g}_j = 0,$$

$$\prod_{j=1}^m \hat{w}_j = m.$$

W celu otrzymania danych oczyszczonych od wartości szeregu odejmuje się składniki $\hat{g}_j$ lub dzieli się je przez wskaźniki sezonowości $\hat{w}_j$.

Prognozy obliczamy na podstawie wzorów:

a) dla modelu addytywnego: $\hat{y}_{tp} = Y_{tp}^* + \hat{g}_j$ (3)

b) dla modelu multiplikatywnego: $\hat{y}_{tp} = Y_{tp}^* \cdot \hat{w}_j$ (4)

2. Materiał badawczy

W prezentowanym opracowaniu wykorzystany został wariant multiplikatywny metody wskaźnikowej. Zastosowanie przedstawionej procedury oczyszczania szeregu z wahań sezonowych pozwoliło na wykorzystanie w prognozowaniu brakujących danych metod nieuwzględniających tych wahań.

Z badań przeprowadzonych w pracy [Szmuksta-Zawadzka, Zawadzki 2000] wynika, że w szeregach opisujących kształtowanie się dekadowych stanów depozytów bankowych mamy do czynienia z dwoma rodzajami sezonowości:

- roczną o cyklu 36-dekadowym,
- miesięczną o cyklu 3-dekadowym.

Podstawę informacyjną badań stanowiły dekadowe szeregi czasowe kształtowania się stanów depozytów a'vista zebrane w wybranych oddziałach banku X. Dane statystyczne zostały zaczerpnięte z banku danych Katedry Zastosowań Matematyki.

Prezentowane wyniki prognozowania dotyczą dwóch zmiennych $Y_1 - Y_2$, różniących się zarówno oceną współczynnika zmienności, jak i amplitudą wahań dla 3- i 36-dekadowych wskaźników sezonowości.

Oceny podstawowych charakterystyk opisowych badanych zmiennych zestawiono w tab. 1.

Tabela 1. Podstawowe charakterystyki opisowe dla zmiennych $Y_1 - Y_2$ (144 dekady)

Zmienna	Statystyki		
	średnia	odchylenie standardowe	współczynnik zmienności (%)
Y1	555,4	540,5	97,32
Y2	1209,9	929,8	76,85

Źródło: opracowanie własne.

Wyznaczając 3- i 36-dekadowe wskaźniki sezonowości, przyjęto wariant zakładający dostępność informacji obejmujących pierwsze dwa lata, czyli 72 dekady.

Tabela 2. Wskaźniki sezonowości o cyklu 3-dekadowym dla zmiennych $Y_1 - Y_2$

Zmienna	Dekada			Maksimum – minimum
	I	II	III	
Y_1 (72 dekady)	101,60	89,89	108,51	18,62
Y_2 (72 dekady)	104,99	93,76	104,99	11,23

Źródło: opracowanie własne.

Tabela 3. Różnice między maksymalnym i minimalnym odchyleniem sezonowym wskaźników o cyklu 36-dekadowym dla zmiennych $Y_1 - Y_2$

Zmienna	Maksimum	Minimum	Maksimum – minimum
Y_1 (72 dekady)	132,38	75,96	56,42
Y_2 (72 dekady)	113,67	87,51	26,16

Źródło: opracowanie własne.

Ze względu na występowanie widocznych wahań sezonowych o cyklu 36- i 3-dekadowym zdecydowano, że modelowanie predyktywne i prognozowanie inter- i ekstrapolacyjne stanów bankowych depozytów a'vista zostanie przeprowadzone także dla szeregów czasowych oczyszczonych z 36-dekadowymi wskaźnikami sezonowości (wariant A) i oczyszczonych z 3-dekadowymi wskaźnikami sezonowości (wariant B).

3. Zastosowane metody badawcze

Do budowy prognoz inter- i ekstrapolacyjnych dla oczyszczonych z wahań sezonowych zmiennych wykorzystano następujące metody: odcinkową, łuków I, II, wielomianową Lagrange'a, wielomianową Lagrange'a z węzłami rozłożonymi zgodnie z funkcją optymalizacyjną Czebyszewa, najmniejszych kwadratów oraz modele wyrównania

wykładniczego Browna i liniowy model Holta. Wszystkie wymienione tu metody mogą być wykorzystane do budowy prognoz interpolacyjnych. Natomiast do budowy prognoz ekstrapolacyjnych nie może być wykorzystana: metoda odcinkowa oraz metody łuków I i łuków II [Z badań ... 1979]. W przypadku tych procedur do budowy prognoz ekstrapolacyjnych wykorzystana została metoda najmniejszych kwadratów, dla której równania trendów były szacowane na podstawie szeregów uzupełnionych o oceny interpolowane otrzymane za pomocą wymienionych metod. Budowa prognoz na podstawie wybranych metod numerycznych była możliwa dzięki wykorzystaniu programu komputerowego *Prognostyk* pracującego w środowisku *Microsoft Windows*. Program ten umożliwia nie tylko wyznaczenie prognoz i obliczenie ich błędów, ale także ich prezentację graficzną.

4. Prezentacja i ocena wyników badań

Wyniki modelowania predyktywnego i prognozowania ekstrapolacyjnego badanych zmiennych dla pełnych szeregów czasowych (bez luk w danych), oczyszczonych z wahań sezonowych za pomocą 36- i 3-dekadowych wskaźników sezonowości przedstawiono w tab. 4.

Tabela 4. Zestawienie średnich względnych błędów prognoz ekstrapolacyjnych dla zmiennych $Y_1 - Y_2$ bez luk w danych

Zmienna	Błędy prognoz ekstrapolacyjnych (dla $h = 9$, w %)					
Y1 (A)	L-3WP	15,85	T2	5,67	P5	19,70
	L-3WC	8,28	T3	5,28	L2	17,25
	L-4WP	18,63	TW	19,66	K1	12,66
	L-4WC	15,10	TW2	>20	H3I	14,60
			TW3	9,68		
Y1 (B)	L-3WP	7,11	T2	4,88	P6	12,12
	L-3WC	4,08	T3	4,15	L2	5,26
	L-4WP	8,98	TW	19,13	K1	4,25
	L-4WC	8,31	TW2	>20	H4I	5,92
			TW3	8,05		
Y2 (A)	L-3WP	4,50	T2	3,55	P7	7,44
	L-3WC	3,61	T3	3,62	L2	4,89
	L-4WP	4,64	TW	5,41	K1	4,48
	L-4WC	4,61	TW2	8,06	H3I	4,72
			TW3	3,61		
Y2 (B)	L-3WP	3,14	T2	4,16	P8	4,00
	L-3WC	4,15	T3	7,88	L2	3,71
	L-4WP	2,43	TW	7,70	K1	6,21
	L-4WC	2,27	TW2	10,68	H5I	3,00
			TW3	7,42		

W tabelach 5-6 zestawiono wyniki prognozowania inter- i ekstrapolacyjnego w sytuacji występowania luk w danych. Analizie poddano dwa warianty luk zakładających występowanie brakujących informacji rozkładających się niesystematycznie. W wariantach tych założono dostępność danych w pierwszych dwóch latach. Natomiast w kolejnych dwóch latach w wariancie pierwszym luki obejmowały na przemian: pierwszą i drugą dekadę lub tylko dekadę pierwszą, w wariancie drugim zaś luki obejmowały trzy dekady i dotyczyły miesięcy nieparzystych.

Dla wszystkich rodzajów prognoz obliczono przeciętne względne błędy prognoz. W przypadku prognoz interpolacyjnych punktem odniesienia były realizacje zmiennych dla okresów, w których wystąpiły luki w danych. Natomiast dla prognoz ekstrapolacyjnych przeprowadzona została analiza *ex post* ich dokładności. Dla każdej zmiennej zaznaczone zostały predyktory charakteryzujące się najniższymi ocenami przeciętnych względnych błędów prognoz. Punktem odniesienia dla „jakości” modelowania są wyniki prognozowania dla modeli oszacowanych na podstawie 144 obserwacji (bez luk w danych).

We wszystkich prezentowanych tabelach użyto następujących oznaczeń:

O – metoda odcinkowa, Ł1 – metoda łuków I, Ł2 – metoda łuków II, L-3WP – metoda Lagrange’a z 3 węzłami rozłożonymi proporcjonalnie, L-4WP – metoda Lagrange’a z 4 węzłami rozłożonymi proporcjonalnie, L-3WC – metoda Lagrange’a z 3 węzłami rozłożonymi zgodnie z funkcją optymalizacyjną Czebyszewa, L-4WC – metoda Lagrange’a z 4 węzłami rozłożonymi zgodnie z funkcją optymalizacyjną Czebyszewa.

Dla predyktorów szacowanych MNK przyjęto oznaczenia: trend kwadratowy (T2), wielomian trzeciego stopnia (T3), trend wykładniczy o stałej stopie wzrostu (TW), wykładniczy o zmiennej stopie wzrostu (TW2), wykładniczy z wielomianem trzeciego stopnia (TW3). W przypadku metody najmniejszych kwadratów ograniczono się do zaprezentowania wyników prognozowania uzyskanych na podstawie pięciu wybranych modeli trendu. Nie uwzględniono w nich wyników otrzymanych na podstawie modelu liniowego (TL), dla którego oceny błędów prognoz inter- i ekstrapolacyjnych były znacznie wyższe od pozostałych, powyżej 20%.

Dla predyktorów wykładniczych przyjęto oznaczenia: P – prosty model Browna, L – liniowy model Browna, K – kwadratowy model Browna, H – dwuparametrowy liniowy model Holta. W przypadku tej klasy modeli po oznaczeniach modeli występują optymalne oceny stałej (stałych) wygładzania.

Analiza informacji zawartych w przedstawionych tabelach wskazuje, że efektywność prognozowania na podstawie szeregów z lukami w danych oraz pełnych szeregów (bez luk w danych) jest porównywalna. Co więcej, w wielu wypadkach oceny średnich błędów uzyskane na podstawie szeregów z lukami są nawet niższe od tych uzyskanych na podstawie pełnego szeregu.

Wszystkie z zastosowanych metod pozwoliły na uzyskanie najniższych ocen błędów prognoz inter- i ekstrapolacyjnych kształtujących się poniżej 10%. Przy czym oceny błędów otrzymane dla zmiennej Y_2 charakteryzującej się niższą oceną współczynnika zmienności, a także o zdecydowanie niższych różnicach między maksymalnym i mini-

Tabela 5. Zestawienie średnich błędów prognoz inter- I ekstrapolacyjnych dla zmiennej Y_i

Wariant	Błędy prognoz															
	interpolacyjnych								ekstrapolacyjnych ($h = 9$)							
Y_i (A1) I wariant luk	O	10,80	L-3WP	15,61	T2	13,35	P7	13,14	O	6,30	L-3WP	15,85	T2	8,08	P7	>20
	Ł1	13,32	L-3WC	16,26	T3	13,22	L2	11,11	Ł1	6,36	L-3WC	8,28	T3	6,26	L2	13,03
	Ł2	13,66	L-4WP	>20	TW	13,75	K1	11,23	Ł2	6,69	L-4WP	18,63	TW	14,22	K1	10,17
			L-4WC	15,76	TW2	13,60	H61	11,29			L-4WC	16,76	TW2	18,24	H61	16,80
					TW3	12,04							TW3	11,22		
Y_i (B1) I wariant luk	O	6,34	L-3WP	7,80	T2	8,54	P7	10,43	O	4,86	L-3WP	7,10	T2	7,58	P7	13,13
	Ł1	7,21	L-3WC	7,94	T3	8,82	L2	7,87	Ł1	4,80	L-3WC	4,08	T3	4,66	L2	5,83
	Ł2	8,29	L-4WP	17,71	TW	10,20	K1	7,83	Ł2	5,20	L-4WP	8,98	TW	12,75	K1	4,75
			L-4WC	9,09	TW2	10,14	H51	7,71			L-4WC	7,38	TW2	18,28	H51	6,29
					TW3	8,07							TW3	8,90		
Y_i (A2) II wariant luk	O	5,22	L-3WP	12,17	T2	9,76	P6	10,35	O	3,51	L-3WP	15,85	T2	6,83	P6	>20
	Ł1	7,51	L-3WC	13,27	T3	9,23	L2	8,34	Ł1	3,57	L-3WC	8,28	T3	5,52	L2	15,00
	Ł2	7,05	L-4WP	18,82	TW	10,08	K1	8,00	Ł2	3,45	L-4WP	18,63	TW	15,19	K1	10,26
			L-4WC	13,62	TW2	10,31	H51	8,52			L-4WC	15,10	TW2	>20	H51	17,65
					TW3	7,84							TW3	9,70		
Y_i (B2) II wariant luk	O	3,87	L-3WP	7,50	T2	8,23	P6	8,64	O	3,74	L-3WP	7,10	T2	5,16	P6	12,39
	Ł1	5,12	L-3WC	8,06	T3	7,67	L2	5,23	Ł1	4,97	L-3WC	4,08	T3	4,14	L2	4,59
	Ł2	4,94	L-4WP	19,73	TW	8,98	K1	5,81	Ł2	2,30	L-4WP	8,98	TW	16,53	K1	4,21
			L-4WC	13,79	TW2	9,64	H51	5,04			L-4WC	8,31	TW2	>20	H51	5,38
					TW3	7,12							TW3	6,57		

Źródło: obliczenia własne.

Tabela 6. Zestawienie średnich względnych błędów prognoz inter- i ekstrapolacyjnych dla zmiennej Y_2

Wariant	Błędy prognoz															
	interpolacyjnych							ekstrapolacyjnych ($h = 9$)								
Y_2 (A1) I wariant luk	O	4,17	L-3WP	5,06	T2	3,98	P7	6,04	O	3,56	L-3WP	4,50	T2	4,05	P7	7,64
	Ł1	4,61	L-3WC	4,07	T3	3,92	L2	4,35	Ł1	3,56	L-3WC	3,61	T3	3,50	L2	4,64
	Ł2	5,59	L-4WP	4,49	TW	4,22	K1	4,05	Ł2	3,36	L-4WP	4,64	TW	3,75	K1	4,01
			L-4WC	14,45	TW2	4,19	H61	4,35			L-4WC	3,46	TW2	6,34	H61	4,46
					TW3	3,81							TW3	4,02		4,46
Y_2 (B1) I wariant luk	O	3,86	L-3WP	4,67	T2	4,98	P7	5,56	O	2,15	L-3WP	3,14	T2	2,14	P7	4,06
	Ł1	4,03	L-3WC	4,56	T3	4,85	L2	4,08	Ł1	2,21	L-3WC	4,15	T3	5,01	L2	2,34
	Ł2	4,71	L-4WP	5,62	TW	5,22	K1	4,27	Ł2	2,20	L-4WP	2,43	TW	5,60	K1	2,63
			L-4WC	9,35	TW2	4,92	H51	4,24			L-4WC	4,98	TW2	8,05	H51	3,46
					TW3	4,86							TW3	4,13		
Y_2 (A2) II wariant luk	O	5,22	L-3WP	4,59	T2	4,12	P6	5,44	O	3,51	L-3WP	4,50	T2	3,53	P6	7,39
	Ł1	7,51	L-3WC	4,37	T3	3,94	L2	4,29	Ł1	3,57	L-3WC	3,61	T3	3,62	L2	4,45
	Ł2	7,05	L-4WP	4,29	TW	3,90	K1	4,03	Ł2	3,45	L-4WP	4,64	TW	5,79	K1	3,83
			L-4WC	3,97	TW2	3,86	H51	4,35			L-4WC	4,61	TW2	9,64	H51	4,50
					TW3	3,77							TW3	3,73		
Y_2 (B2) II wariant luk	O	3,87	L-3WP	5,00	T2	5,29	P6	4,83	O	3,74	L-3WP	3,14	T2	3,47	P6	4,02
	Ł1	5,12	L-3WC	5,10	T3	4,79	L2	3,22	Ł1	4,97	L-3WC	4,15	T3	6,37	L2	2,46
	Ł2	4,94	L-4WP	6,55	TW	4,66	K1	3,58	Ł2	2,20	L-4WP	2,43	TW	7,71	K1	2,66
			L-4WC	6,42	TW2	4,40	H51	4,35			L-4WC	2,27	TW2	10,60	H51	3,03
					TW3	4,76							TW3	5,28		

Źródło: obliczenia własne.

malnym odchyleniem sezonowym dla wskaźników zarówno o cyklu 36-, jak i 3-dekadowym były jeszcze o kilka punktów procentowych niższe i kształtowały się poniżej 5%. Przedstawione wyniki wskazują więc na istnienie zależności pomiędzy efektywnością prognozowania a natężeniem sezonowości, mimo że analizie poddano szeregi czasowe, które zostały oczyszczone z wahań sezonowych.

Analiza porównawcza wyników prognozowania dla obu rodzajów szeregów czasowych wskazuje, że w przeważającej większości analizowanych modeli niższe oceny prognoz inter- i ekstrapolacyjnych otrzymano dla szeregów czasowych oczyszczonych za pomocą 3-dekadowych wskaźników sezonowości.

Biorąc pod uwagę tylko te modele, dla których otrzymano najniższe oceny błędów, trudno jest jednoznacznie wskazać metodę, dla której oceny prognoz byłyby jednoznacznie najniższe. Otrzymane oceny błędów prognoz inter- i ekstrapolacyjnych charakteryzowały się bowiem stosunkowo niewielkim zróżnicowaniem.

Na szczególną uwagę zasługuje metoda wielomianowa Lagrange'a. Wyniki prognozowania ekstrapolacyjnego, uzyskane dzięki zastosowaniu tej procedury, okazały się identyczne z tymi otrzymanymi na podstawie pełnego szeregu. Widać w związku z tym wyraźnie brak wpływu rozmieszczenia luk w danych na dokładność oszacowanych prognoz.

W dalszej części otrzymane wyniki prognozowania inter- i ekstrapolacyjnego dla szeregów czasowych oczyszczonych z wahań sezonowych porównano z wynikami otrzymanymi dla szeregów czasowych nieoczyszczonych z tych wahań. Analiza porównawcza otrzymanych wyników przedstawiona została w tab. 7. Użyto w niej następujących oznaczeń: *s* – model ze stałą sezonowością, *z* – model z sezonowością zmienną.

Prezentowane w tabelach modele dla danych nieoczyszczonych z wahań sezonowych (wariant C) zostały wybrane spośród wielu oszacowanych wersji różniących się zestawem składowych harmonicznych. Kryterium wyboru prezentowanych wersji modeli była przede wszystkim minimalizacja ocen błędów prognoz ekstrapolacyjnych. Proces modelowania i prognozowania poprzedzono wyborem istotnych składowych harmonicznych dokonany metodą selekcji *a priori*.

Otrzymane oceny błędów prognoz inter- i ekstrapolacyjnych, zarówno dla danych nieoczyszczonych, jak i oczyszczonych z wahań sezonowych, kształtowały się na zbliżonym poziomie. W większości wypadków najniższe oceny błędów uzyskano dla szeregów nieoczyszczonych z wahań sezonowych. Natomiast oceny błędów dla szeregów oczyszczonych za pomocą wskaźników sezonowości o cyklu 36- i 3-dekadowym były wyższe maksymalnie o 1-2%.

Przedstawione wyniki potwierdziły przydatność proponowanej w literaturze metody polegającej na oczyszczaniu z wahań sezonowych zmiennych wykazujących te wahania, a następnie wykorzystaniu tak przygotowanych szeregów w prognozowaniu brakujących danych.

Tabela 7. Porównanie ocen błędów prognoz inter- i ekstrapolacyjnych

Zmienna	Wariant	Szereg	Błędy prognoz							
			interpolacyjnych		ekstrapolacyjnych					
					$h = 3$		$h = 6$		$h = 9$	
Y1	I	A1(36)	T3	13,22	T3	3,27	T3	3,15	T3	6,26
		B1(3)	T3	8,82	T3	5,60	T3	4,42	T3	4,66
		C	T3S	9,62	T3S	4,36	T3S	3,62	T3S	4,24
	II	A2(36)	T3	9,23	T3	3,83	T3	2,71	T3	5,52
		B2(3)	T3	7,67	T3	4,71	T3	4,08	T3	4,14
		C	T3Z	6,88	T3Z	2,74	T3Z	3,16	T3Z	3,66
	PEŁNY SZEREG	A1(36)	-	-	TW3	3,31	T2	2,75	T3	5,28
		B1(3)	-	-	T3	4,77	T3	4,05	T3	4,15
		C	-	-	T3Z	2,07	T3	2,71	T3	3,32
Y2	I	A1(36)	T3	3,92	T3	2,95	T3	4,12	T3	3,50
		B1(3)	T2	4,98	T2	3,47	T2	2,81	T2	2,14
		C	T2Z	3,65	T2Z	4,11	T2Z	2,77	T2Z	2,79
	II	A2(36)	T2	4,12	T2	3,21	T2	4,40	T2	3,53
		B2(3)	T2	5,29	T2	4,16	T2	3,95	T2	3,47
		C	T2Z	3,83	T2Z	2,49	T2Z	1,89	T2Z	2,18
	PEŁNY SZEREG	A2(36)	-	-	T2	3,05	T2	4,24	T2	3,55
		B2(3)	-	-	T2	4,72	T2	4,59	T2	4,16
		C	-	-	TW3S	1,94	TW3S	2,61	TW3S	2,46

Źródło: obliczenia własne.

Literatura

Szmuksta-Zawadzka M., Zawadzki J., *Prognozowanie brakujących informacji a modele oszczędne dla okresowych szeregów czasowych*, Prace Naukowe Akademii Ekonomicznej, AE, Kraków 2000.

Z badań nad metodami predykcji brakujących informacji, red. A. Zeliaś, Prace Naukowe Akademii Ekonomicznej nr 114, AE, Kraków 1979.

Prognozowanie ekonomiczne. Teoria, przykłady, zadania, red. A. Zeliaś, PWN, Warszawa 2003.

THE APPLICATION OF THE INDEX SEASONS METHOD IN FORECASTING DECADE DATA

Summary

The main objective of this paper was to examine the effectiveness of the inter- and extrapolating methods in forecasting real processes under the condition of missing information, as well as to find out the factors determining their effectiveness. The studies were based on a'vista deposits collected from two banks. The literature presents a variety of numeric methods used for estimating interpolating forecasts. This work suggests how extrapolating forecasts can be applied. The following methods were discussed: a segment method, an arc method I,II, Lagrange method, Chebyshev polynomials, Brown and Holt models and trigonometric polynomial model. The results of empirical studies indicate that econometric methods may be applied to forecast missing data.