

Maciej Oesterreich

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: maciej.oesterreich@zut.edu.pl

**SYMULACYJNE BADANIE WPLYWU LICZBY
I ROZMIESZCZENIA LUK NA DOKŁADNOŚĆ
PROGNOZ W SZEREGU CZASOWYM
DLA DANYCH DZIENNYCH**

**THE SIMULATION ANALYSIS OF IMPACT
OF NUMBER AND DISTRIBUTION OF GAPS
ON ACCURACY OF FORECASTS
IN DAILY TIME SERIES**

DOI: 10.15611/ekt.2017.1.05

JEL Classification: C13, C22

Streszczenie: W pracy przedstawione zostaną wyniki symulacyjnej analizy wpływu występowania luk na dokładność prognoz inter- i ekstrapolacyjnych dla szeregów charakteryzujących się wysoką częstotliwością obserwowania oraz złożoną sezonowością. Do budowy prognoz wykorzystano klasyczne modele szeregu czasowego, w którym wahania sezonowe o cyklu zarówno tygodniowym, jak i rocznym były opisane za pomocą zmiennych zero-jedynkowych. Zmienną, którą poddano analizie, była dzienna sprzedaż w litrach paliw płynnych na stacji paliw X w latach 2012-2014, przy czym lata 2012-2013 stanowiły przedział czasowy próby, a rok 2014 był okresem empirycznej weryfikacji prognoz. Rozpatrywanych było sześć wariantów luk różniących się odsetkami, z których każdy zawierał po dziesięć tysięcy wariantów luk. Obliczenia zostały wykonane z wykorzystaniem pakietu *R* oraz Statistica 12.

Słowa kluczowe: zmienne o wysokiej częstotliwości obserwowania, złożone wahania sezonowe, luki, prognozowanie, modele szeregu czasowego.

Summary: In the paper will be presented the simulation analysis of the impact of the occurrence of gaps on the accuracy of inter- and extrapolative forecasts for the high-frequency time series with complex seasonality. To forecasts construction were used a classic time series model, in which both weekly and annual seasonality was described by dummy variables. The analyzed variable was daily sale of liquid fuels in liters in petrol station X in years 2012-2014. Data in years 2012-2013 were used in a model construction and a year 2014 was a period of empirical verification of forecasts. There were examined six different variants of gaps, each of which contains 10,000 combinations of gaps. Calculations were made using R statistical environment and Statsoft Statistica12.

Keywords: high frequency time series, complex seasonal fluctuations, gaps, forecasting, time series models.

1. Wstęp

Jednym z klasycznych warunków modelowania ekonometrycznego dla danych zarówno przekrojowych, jak i w postaci szeregów czasowych jest kompletność danych statystycznych. Okazuje się jednak, że występowanie luk w szeregu czasowym z wahaniami sezonowymi w wielu przypadkach utrudnia, ale nie uniemożliwia modelowania i prognozowania zmiennych ekonomicznych. Analiza wpływu występowania luk na dokładność prognoz inter- i ekstrapolacyjnych dla danych miesięcznych o sezonowości rocznej została w pracy [Oesterreich 2012] (luki niesystematyczne) oraz [Oesterreich 2014a; 2014b] (luki systematyczne).

Celem niniejszej pracy będzie rozszerzenie badań na dzienne szeregi czasowe charakteryzujące się występowaniem złożonych wahań sezonowym o cyklu tygodniowym i rocznym o silnym natężeniu.

Wykorzystanie metod symulacyjnych polegało na wygenerowaniu dla każdego z sześciu wariantów, różniących się liczbą (odsetkiem) luk, po dziesięć tysięcy kombinacji (układów) luk. Dane te posłużyły następnie do oszacowania modeli i wyznaczenia prognoz i prognoz inter- i ekstrapolacyjnych oraz obliczania średnich względnych błędów prognoz (zob. [Oesterreich 2012]).

Do budowy prognoz zostały wykorzystane predyktory oparte na klasycznym modelu szeregu czasowego, w którym złożona sezonowość opisana została za pomocą zmiennych zero-jedynkowych [Zeliaś i in. 2003, s. 89-90]. Obliczenia wykonano z wykorzystaniem pakietu R.

2. Metodyka badań

Problem konstrukcji modelu dla danych o wysokiej częstotliwości obserwowania został przedstawiony m.in. w pracach [Kufel 2010] oraz [Szmuksta-Zawadzka, Zawadzki 2011; 2014a], natomiast prognozowanie brakujących danych oraz efektywność wykorzystania różnych klas predyktorów w pracach [Szmuksta-Zawadzka, Zawadzki 2012; 2014b].

Jednym z ważniejszych zagadnień stojących przed badaczem w przypadku modelowania tego rodzaju szeregów jest opis złożonej sezonowości. Można do tego celu wykorzystać modele szeregu czasowego, w których wahania sezonowe są opisane zarówno za pomocą zmiennych zero-jedynkowych, jak i za pomocą wielomianu trygonometrycznego. W niniejszym artykule zdecydowano się wykorzystać predyktory klasyczne, w których zarówno tygodniowa, jak i roczna sezonowość była opisana za pomocą zmiennych zero-jedynkowych. Zapis analityczny modelu w postaci addytywnej jest następujący:

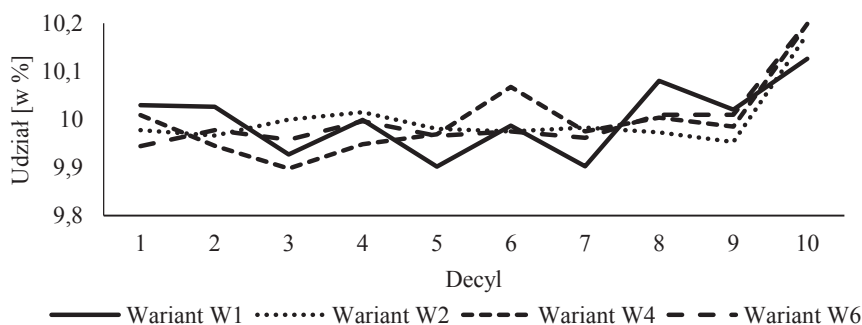
$$Y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^7 D_i Q_{idt} + \sum_{j=1}^{12} m_j Q_{jmt} + \varepsilon_t, \quad t = 1, 2, \dots, n, \quad (1)$$

gdzie: Q_{idt} – zmienna przyjmująca 1 dla i -tego dnia tygodnia oraz 0 dla pozostałych;
 Q_{jmt} – zmienna przyjmująca 1 dla j -tego miesiąca roku oraz 0 dla pozostałych.

Analizie poddano sześć wariantów po dziesięć tysięcy luk w każdym z nich. Udział luk w okresie „próby” wynosił odpowiednio:

- W1 – 5%, to jest ok. 46 luk,
- W2 – 10%, to jest ok. 91 luk,
- W3 – 15%, to jest ok. 137 luk,
- W4 – 20%, to jest ok. 182 luki,
- W5 – 25%, to jest 228 luk,
- W6 – 30%, to jest ok. 274 luki.

Dla każdego z wariantów wylosowano za pomocą algorytmu zbudowanego w programie *R* dziesięć tysięcy układów luk. Na rysunku 1 przedstawiano informacje o odsetku luk przypadającym na kolejne decyle przedziału estymacyjnego w wybranych wariantach.



Rys. 1. Udziały liczby luk w poszczególnych decylach* przedziału czasowego próby dla wariantów W1, W2, W4 oraz W6

*1 decyl \approx 91 obserwacji

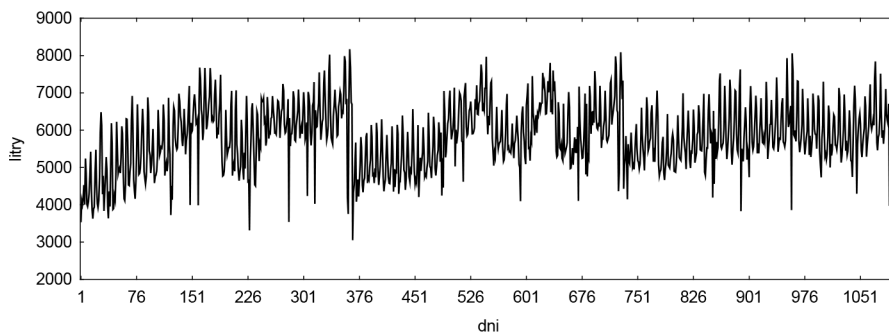
Źródło: opracowanie własne.

Z informacji zamieszczonych na rys. 1 wynika, że odsetki luk w kolejnych decylach przedziału estymacyjnego dla czterech wybranych wariantów luk oscylowały na poziomie ok. 10%, co oznacza, że możemy mówić o praktycznie jego jednorodnym rozkładzie. Jedynie dla decylu dziesiątego, dla wszystkich analizowanych wariantów, widzimy nieznaczne, nie większe od 0,2 punktu procentowego (p.p.), odchylenie *in plus* w liczbie luk.

3. Przykład empiryczny

Modelowaniu i prognozowaniu poddano dzienną sprzedaż paliw płynnych na stacji paliw X w latach 2012-2014 (w litrach). Obejmowała ona łączną sprzedaż benzyny bezołowiowej oraz Pb98, a także oleju napędowego. Dane statystyczne za okres od 1 stycznia 2012 r. do 31 czerwca 2014 r. (912 obserwacji) stanowiły przedział

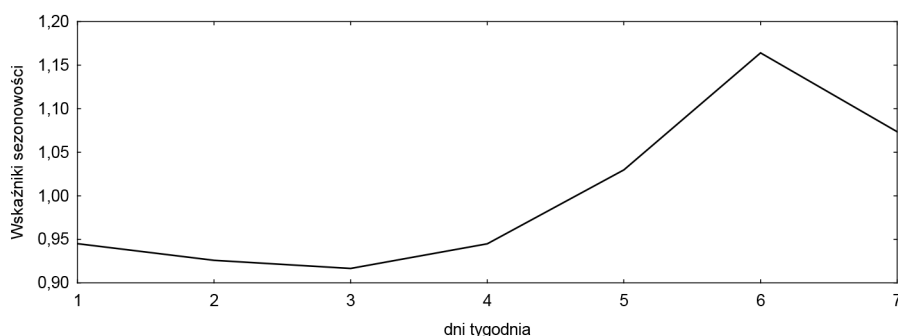
czasowy „próby”. Okresem empirycznej weryfikacji prognoz było drugie półrocze 2014 roku (184 obserwacje). Na rysunku 2 przedstawiano kształtowanie się zmiennej prognozowanej.



Rys. 2. Sprzedaż paliw płynnych na stacji paliw X w okresie od 1.01.2012-31.12.2014

Źródło: opracowanie własne.

Na podstawie szeregu czasowego próby oszacowano wskaźniki sezonowości dla poszczególnych dni tygodnia oraz miesięcy przedstawione w postaci graficznej na rys. 3 i 4.

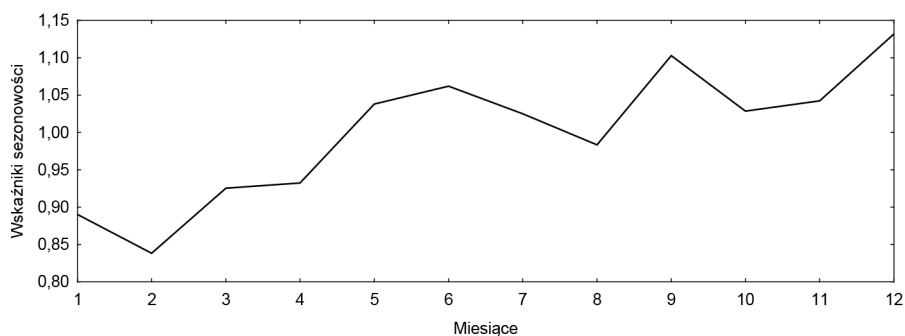


Rys. 3. Kształtowanie się wskaźników sezonowości dla dni tygodnia

Źródło: opracowanie własne.

Z analizy rysunku 3 wynika, że najwyższą ocenę wskaźnik sezonowości dla dni tygodnia przyjął w sobotę (116,4%), natomiast najniższy poziom odnotowano dla środy (91,66%). Różnica pomiędzy skrajnymi jego wartościami wyniosła 24,76 p.p.

Z rysunku 4 wynika, że w kształtowaniu się wskaźników sezonowości dla miesięcy występują trzy lokalne maksima sezonowe w: czerwcu (106,19%), we wrześniu (110,28%) i w grudniu (113,19%). Minimalną ocenę wskaźnik sezonowości



Rys. 4. Kształtowanie się wskaźników sezonowości dla miesięcy

Źródło: opracowanie własne.

przyjął w lutym (83,83%). Różnica pomiędzy ocenami wskaźników dla miesięcy o skrajnych wartościach wynosiła 29,37 p.p. i była ok. 5 p.p. wyższa niż dla dni tygodnia.

Z rysunku 1 wynika, że analizowana zmienna może charakteryzować się dużą liczbą obserwacji nietypowych, które wpływają na kształtowanie się wartości oszacowanych parametrów, a co za tym idzie – na własności predyktywne takiego modelu. Do weryfikacji hipotezy o występowaniu tego rodzaju obserwacji wykorzystane zostały reszty standaryzowane [Dittmann i in. 2009, s. 48]:

$$e_{t(stand)} = \frac{e_t}{S}, \quad (2)$$

gdzie:

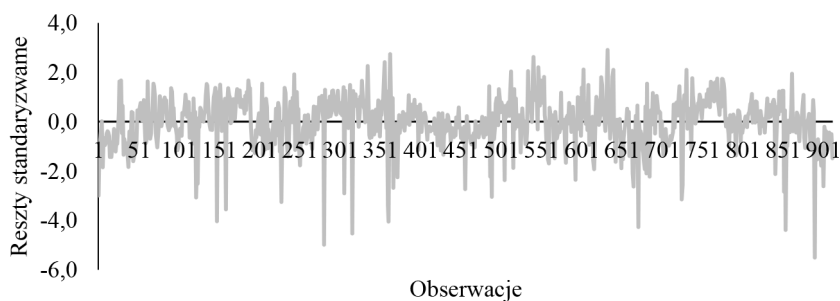
$$S = \sqrt{\frac{\sum_t^n e_t^2}{n - m - 1}} \quad (3)$$

gdzie: e_t – reszta modelu, $e_{t(stand)}$ – reszty standaryzowane, S – odchylenie standardowe składnika losowego, n – długość przedziału czasowego „próby”, m – liczba zmiennych objaśniających w modelu.

Za obserwacje nietypowe uważa się te, dla których reszty standaryzowane przyjmują wartości większe od 3 lub mniejsze od -3 .

Na rysunku 5 przedstawiano kształtowanie się reszt standaryzowanych oszacowanych dla modelu (1).

W tabeli 1 zestawione zostały informacje o liczbie obserwacji, dla których wartości reszt standaryzowanych były mniejsze od -3 lub większe od 3, a także o liczbie przypadków, gdy ich wartości zawierały się w przedziałach $(-3; -2) \cup (2; 3)$



Rys. 5. Kształtowanie się wartości reszt standaryzowanych

Źródło: opracowanie własne.

Tabela 1. Liczba reszt standaryzowanych o wartościach większych od 1,5 i mniejszych od $-1,5$

Przedział	Liczba przypadków	% obserwacji	% wariancji Y
$(-\infty; -3) \cup (3; \infty)$	13	1,425	4,684
$(-3; -2) \cup (2; 3)$	28	3,070	6,729
Razem	41	4,495	11,413

Źródło: opracowanie własne.

Z informacji zawartych w tab. 1 wynika, że stwierdzono tylko 13 obserwacji, dla których wartości reszt standaryzowanych były mniejsze od -3 i większe od 3 . Stanowiły one niecałe 1,5% wszystkich obserwacji, ale odpowiadały prawie za ok. 4,7% wariancji zmiennej zależnej. Wartości reszt z przedziału $(-3; -2) \cup (2; 3)$ odnotowano tylko dla 28 obserwacji, jednak stanowiły one aż ok. 6,7% wariancji.

Jak wynika z tab. 1, 41 (tj. 4,495%) obserwacji, dla których reszty standaryzowane były mniejsze od -2 i większe od 2 , odpowiadały aż za ok. 11,4% wariancji zmiennej prognozowanej.

Aby w sposób jednoznaczny określić wpływ 41 obserwacji, o których mowa w tab. 1, zbudowano dwa modele, korzystając ze wzoru (1) – pierwszy (I) na podstawie szeregu pełnego oraz drugi (II) na podstawie szeregu, z którego usunięto ww. obserwacje. Informacje o ocenach parametrów strukturalnych, odpowiadających im błędach szacunku parametrów, ocenach statystyk *t*-Studenta, a także wartościach parametrów struktury stochastycznej, dobroci dopasowania oraz oceny średnich względnych błędów MAPE prognoz ekstrapolacyjnych przedstawiono w tab. 2.

Z informacji zawartych w tab. 3 wynika, że model zbudowany na podstawie szeregu, z którego usunięto 41 obserwacji (tj. ok. 4,495%), charakteryzował się wyraźnie niższymi wartościami średnich błędów szacunku parametrów. Mniejsza (o jeden) była też w jego przypadku liczba parametrów nieistotnych statystycznie. Model ten charakteryzował się także aż o ok. 20% wyższym poziomem dopasowania

Tabela 2. Oceny parametrów struktury stochastycznej, średnie błędy szacunku parametrów, wartości statystyk *t*-Studenta, parametry struktury stochastycznej, dobroci dopasowania oraz oceny średnich względnych błędów MAPE prognoz ekstrapolacyjnych

Parametry	Model I*			Model II**		
	<i>b</i>	Bł. std.	<i>t</i> (893)	<i>b</i>	Bł. std.	<i>t</i> (893)
W. wolny	6724,560	90,821	74,042	6871,340	72,711	94,502
<i>t</i>	0,325	0,071	4,582	0,320	0,055	5,744
PN	-674,062	68,206	-9,883	-820,490	53,820	-15,245
WT	-787,646	68,364	-11,521	-907,630	54,179	-16,752
ŚR	-855,856	68,374	-12,517	-929,980	53,987	-17,226
CZW	-686,283	68,382	-10,036	-729,840	54,296	-13,442
PT	-182,509	68,372	-2,669	-291,810	54,301	-5,374
SO	536,862	68,343	7,855	438,580	53,728	8,163
STYCZEŃ	-1520,483	91,157	-16,680	-1530,010	72,901	-20,987
LUTY	-1303,560	92,635	-14,072	-1362,830	73,545	-18,531
MARZEC	-1157,539	90,750	-12,755	-1199,750	72,278	-16,599
KWIECIEŃ	-950,461	91,209	-10,421	-993,910	72,647	-13,681
MAJ	-554,836	90,572	-6,126	-464,760	73,439	-6,329
CZERWIEC	-147,363	91,103	-1,618	-173,100	73,378	-2,359
LIPIEC	-587,499	99,731	-5,891	-645,240	78,799	-8,188
SIERPIEŃ	-831,507	99,560	-8,352	-856,110	79,624	-10,752
WRZESIEŃ	56,108	100,165	0,560	-32,150	79,413	-0,405
PAŹDZIERNIK	-527,211	99,259	-5,311	-512,120	79,739	-6,423
LISTOPAD	-302,485	100,016	-3,024	-220,140	81,052	-2,716
R	0,792			0,868		
R ²	0,628			0,753		
F	83,605 (<i>p</i> = 0,000)			144,280 (<i>p</i> = 0,000)		
SE	551,921			422,6		
MAPE E [w %]	8,380			8,758		

* Model zbudowany na podstawie pełnego szeregu; ** model zbudowany na podstawie szeregu, z którego usunięto 41 obserwacji.

Źródło: opracowanie własne.

mierzonym współczynnikiem determinacji R^2 oraz o 23,43% mniejszym odchyleniem standardowym składnika losowego. Jednoznacznie wskazuje to na negatywny wpływ obserwacji, dla których wartości reszt standaryzowanych przyjmowały wartości mniejsze od -2 lub większe od 2 , na własności opisowe modelu. Analiza średnich względnych błędów prognoz ekstrapolacyjnych zbudowanych na podstawie obydwu modeli wskazuje jednak, że model zbudowany na podstawie szeregu

„oczyszczonego” charakteryzuje się ocenami błędów wyższymi o ok. 0,378 p.p. (tj. 4,32%). Oznacza to, że pomimo lepszych własności opisowych, model zbudowany na podstawie szeregu, z którego usunięto obserwacje, dla których wartości reszt standaryzowanych przyjmowały wartości mniejsze od -2 lub większe od 2 , charakteryzował się gorszymi własnościami predyktywnymi.

W tabeli 3 przedstawiano informacje o średnich wartościach wybranych parametrów struktury stochastycznej oraz dopasowania modeli oraz przeciętnych średnich względnych błędach prognoz inter- i ekstrapolacyjnych.

Tablica 3. Średnie oceny parametrów struktury stochastycznej i dobroci dopasowania modeli oraz średnie błędy MAPE prognoz inter- i ekstrapolacyjnych

	R^2	\overline{SE}	\overline{V}_s [w %]	$\overline{MAPE} I$ [w %]	$\overline{MAPE} E$ [w %]
Pełny	0,63	551,92	9,54	–	8,380
W1	0,63 (0,0062)	551,83 (5,0619)	9,54 (0,0903)	7,519	8,388
W2	0,63 (0,0089)	551,85 (7,3465)	9,54 (0,1306)	7,512	8,393
W3	0,63 (0,0113)	551,81 (9,2870)	9,54 (0,1652)	7,519	8,400
W4	0,63 (0,0133)	551,9 (10,9813)	9,54 (0,1959)	7,513	8,408
W5	0,63 (0,0154)	551,87 (12,6940)	9,54 (0,2265)	7,521	8,414
W6	0,63 (0,0177)	551,79 (14,6109)	9,54 (0,2604)	7,534	8,427

* W nawiasach podano informację o odchyleniu standardowym dla poszczególnych miar.

Źródło: opracowanie własne.

Z informacji zawartych w tab. 3 wynika, że oszacowane modele, niezależnie od odsetka luk, charakteryzowały się przeciętnymi ocenami parametrów struktury stochastycznej i dobroci dopasowania praktycznie identycznymi jak w przypadku modelu dla danych pełnych. Poziom przeciętnego współczynnika determinacji, niezależnie od odsetka luk, oscylował wokół 0,63, przy średnim poziomie odchylenia standardowego składnika losowego wynoszącego 551,9. Przeciętny poziom współczynnika zmienności losowej wynosił przy tym 9,54%.

Także przeciętne błędy MAPE prognoz inter- i ekstrapolacyjnych różniły się maksymalnie o ok. 0,05 p.p.

Z porównania informacji zamieszczonych w tab. 3 oraz uśrednionych wyników dla danych miesięcznych, dotyczących współczynników determinacji oraz ocen błędów prognoz inter- i ekstrapolacyjnych, zamieszczonych w pracach [Oesterreich 2012] (luki niesystematyczne) oraz [Oesterreich 2014a; 2014b] (luki systematyczne), wynika, że kierunek zmian ich wartości był zbliżony do tych zaprezentowanych w pierwszej z prac – oceny obu rodzajów błędów wzrastały wraz ze wzrostem liczby luk. Odmienna była jednak skala tych zmian, które w przypadku danych miesięcznych były większe.

Ze względu na charakterystyki rozkładów błędów prognoz, a zwłaszcza ich kształt, bardziej właściwa w tym przypadku jest analiza miar pozycyjnych. W tabeli 4 przedstawiono informację o wybranych statystykach opisowych rozkładów błędów MAPE prognoz interpolacyjnych.

Tablica 4. Wybrane statystyki opisowe rozkładów błędów MAPE prognoz interpolacyjnych

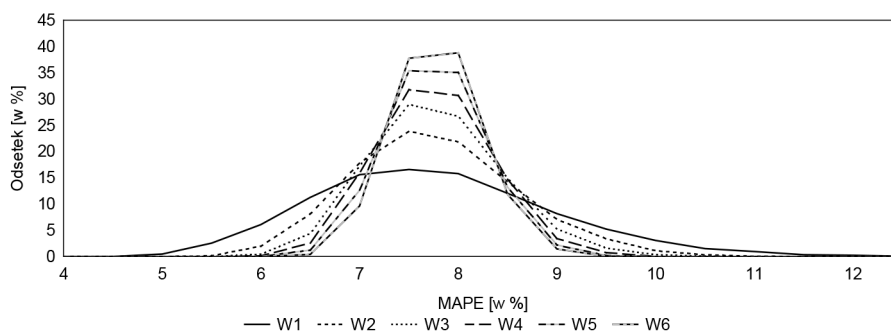
Wariant	Średnia	Q1	Me	Q3	Skośność	S	Min	Max
W1	7,519	6,660	7,421	8,260	0,530	1,214	4,036	13,342
W2	7,512	6,920	7,459	8,038	0,385	0,833	4,907	11,298
W3	7,519	7,053	7,481	7,940	0,363	0,661	5,466	11,105
W4	7,513	7,125	7,493	7,879	0,217	0,557	5,748	9,654
W5	7,521	7,193	7,510	7,831	0,200	0,478	5,954	9,545
W6	7,534	7,244	7,525	7,819	0,147	0,424	6,138	9,317

Źródło: opracowanie własne.

Z informacji zawartych w tab. 4 wynika, że analizowane rozkłady błędów MAPE prognoz interpolacyjnych charakteryzowały się wyraźną asymetrią prawostronną, której siła malała wraz ze wzrostem odsetka luk.

Zmniejszeniu, z 1,6% do 0,574%, ulegała też wartość rozstępu międzykwartylowego, który stanowił od 15,7% (wariant W3) do 19,3% (wariant W4) rozstępu klasycznego. Zmniejszające się różnice pomiędzy maksymalnymi a minimalnymi ocenami błędów w analizowanych wariantach wpływały na wartość odchylenia standardowego, które malało wraz ze wzrostem liczby luk.

Na rysunku 6 przedstawiono kształtowanie rozkładów błędów prognoz interpolacyjnych według wariantów luk.



Rys. 6. Kształtowanie się rozkładów błędów MAPE prognoz interpolacyjnych według wariantów luk

Źródło: opracowanie własne.

Na przedstawionym rysunku widać powolne przesuwanie się ku prawej stronie osi odciętych wykresów rozkładów błędów MAPE prognoz interpolacyjnych, którego powodem był wzrost odsetka luk w kolejnych wariantach. Odpowiada to wzrostowi ocen mediany w tab. 4. „Wysmuklanie” poszczególnych krzywych jest efektem zmniejszania się rozstępu międzykwartylowego oraz klasycznego.

Tak samo co do kierunku zachowywały się rozkłady błędów prognoz interpolacyjnych obliczonych na podstawie danych miesięcznych zawierających luki niesystematyczne dla wszystkich analizowanych modeli (por. [Oesterreich 2012]). Przedstawione na rysunku 6 rozkłady błędów różniły się natomiast wyraźnie kształtem rozkładów otrzymanych dla danych miesięcznych zawierających luki systematyczne – występowanie przypadków o skarnie wysokich ocenach błędów (por. [Oesterreich 2014a; 2014b]).

W tabeli 5 przedstawiono informacje o wybranych statystykach opisowych rozkładów błędów MAPE prognoz ekstrapolacyjnych.

Tablica 5. Wybrane statystyki opisowe rozkładów błędów MAPE prognoz ekstrapolacyjnych

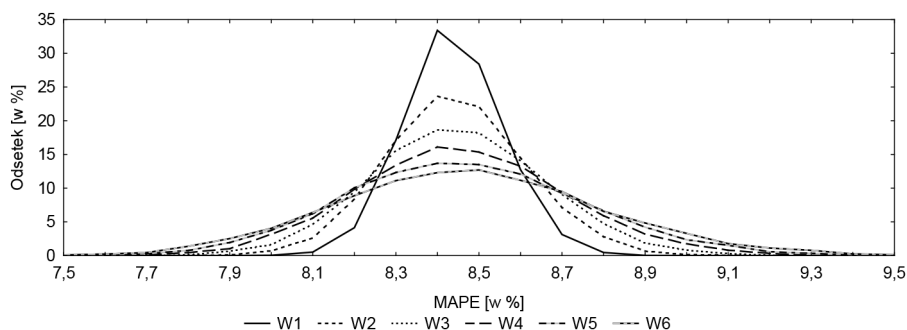
Wariant	Średnia	Q1	Me	Q3	Skośność	S	Min	Max
W1	8,388	8,311	8,385	8,463	0,095	0,115	7,917	8,881
W2	8,393	8,281	8,389	8,502	0,134	0,166	7,848	9,099
W3	8,400	8,259	8,395	8,538	0,135	0,209	7,627	9,293
W4	8,408	8,239	8,401	8,571	0,121	0,250	7,370	9,311
W5	8,414	8,216	8,406	8,602	0,186	0,287	7,474	9,646
W6	8,427	8,209	8,419	8,635	0,184	0,323	7,353	9,709

Źródło: opracowanie własne.

Z informacji zawartych w tab. 5 wynika, że rozkłady charakteryzowały się asymetrią prawostronną, której siła rosła wraz ze wzrostem odsetka luk. Wzrastała również ocena mediany, jak i wartości średniej błędów MAPE prognoz ekstrapolacyjnych, przy czym różnice między skrajnymi wariantami były niewielkie i wynosiły odpowiednio: 0,034 p.p. i 0,039 p.p. Także różnice pomiędzy wartością mediany a wartością średnią błędu w ramach danego wariantu luk były niewielkie i nie przekraczały 0,009 p.p.

Wzrostowi ulegały również oceny rozstępu międzykwartylowego oraz klasycznego. Udział pierwszego z nich w drugim wynosił od 15,8% (wariant W1) do 18,1% (wariant W6). Skutkiem zwiększania się różnicy pomiędzy skrajnymi wartościami błędów w poszczególnych wariantach był ponaddwukrotny wzrost (z 0,115% – wariant W1; do 0,323% – wariant W6) oceny odchylenia standardowego błędów prognoz.

Z analizy rysunku 7 wynika przede wszystkim, że rozkłady błędów MAPE prognoz ekstrapolacyjnych ulegały „spłaszczeniu”, czego skutkiem było zwiększanie się ocen rozstępu międzykwartylowego oraz klasycznego. Kierunek zmian kształ-



Rys. 7. Kształtowanie się rozkładów błędów MAPE prognoz interpolacyjnych według wariantów luk
Źródło: opracowanie własne.

tu rozkładów dla poszczególnych wariantów luk był podobny do tego uzyskanego dla danych miesięcznych zawierających luki niesystematyczne (por. [Oesterreich 2012]).

4. Podsumowanie

Z badań przeprowadzonych w pracy wyprowadzić można następujące wnioski:

- Wzrost liczby luk powodował nieznaczny wzrost wielkości przeciętnych błędów prognoz inter- i ekstrapolacyjnych.
- Przeciętne błędy prognoz ekstrapolacyjnych były jedynie nieznacznie wyższe dla szeregu z lukami (wszystkie warianty) od prognoz dla szeregu pełnego.
- Empiryczne rozkłady błędów prognoz interpolacyjnych wraz ze wzrostem odsetka luk ulegały „wysmukleniu”, tzn. wzrastał odsetek obserwacji w przedziale dominanty oraz malał rozstęp. Rozkłady błędów prognoz ekstrapolacyjnych ulegały „spłaszczeniu”, tzn. malał odsetek obserwacji w przedziale dominanty oraz wzrastał rozstęp. W obu przypadkach następowało przesunięcie rozkładów w prawą stronę osi rzędnych oraz zmniejszanie się współczynnika skośności.
- Odnosząc się do wyników wcześniejszych badań dla danych miesięcznych, należy podkreślić, że prawidłowości w kształtowaniu się rozkładów błędów prognoz inter- i ekstrapolacyjnych dla danych dziennych były zbliżone.
- Dalsze badania o charakterze symulacyjnym powinny koncentrować się na szeregach o wyraźnie zaznaczonych trendach i silnych wahaniami sezonowych o charakterze multiplikatywnym.

Literatura

- Dittmann P., Dittmann I., Szabela-Pasierbińska E., Szpulak A., 2009, *Prognozowanie w zarządzaniu przedsiębiorstwem*, Wolters Kluwer Polska, Kraków.
- Kufel T., 2010, *Ekonometryczna analiza cykliczności procesów gospodarczych o wysokiej częstotliwości obserwowania*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń.
- Oesterreich M., 2012, *Symulacyjne badanie wpływu częstości występowania luk niesystematycznych w szeregach czasowych na dokładność prognoz*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, seria Ekonometria, nr 38, Wrocław.
- Oesterreich M., 2014a, *Kombinacje liniowe składowych harmonicznych a dokładność prognoz w sezonowych szeregach czasowych z lukami systematycznymi*, Folia Pomeranae Universitatis Technologiae Stetinensis, seria Oeconomica 313(76)3, Szczecin.
- Oesterreich M., 2014b, *Badanie wpływu częstości występowania luk systematycznych w mikroekonomicznym szeregu czasowym na dokładność prognoz*, [w:] *Problemy społeczno-ekonomiczne w relacjach międzynarodowych: analiza modelowa rozwoju regionów*, red. W. Szkutnik, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.
- Szmuksta-Zawadzka M., Zawadzki J., 2011, *Modelowanie i prognozowanie zapotrzebowania na energię elektryczną w mikroskali, Społeczna rola statystyki 2011*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 165, Wrocław.
- Szmuksta-Zawadzka M., Zawadzki J., 2012, *Z badań nad metodami prognozowania na podstawie niekompletnych szeregów czasowych z wahaniami okresowymi (sezonowymi)*, Przegląd Statystyczny – Statistical Review, t. 59, numer specjalny: 1.
- Szmuksta-Zawadzka M., Zawadzki J., 2014a, *Regularne modele hierarchiczne dla szeregów czasowych o wysokiej częstotliwości*, [w:] *Problemy społeczno-ekonomiczne w relacjach międzynarodowych: analiza modelowa rozwoju regionów*, red. W. Szkutnik, Wydawnictwo Uniwersytetu Ekonomicznego, Katowice.
- Szmuksta-Zawadzka M., Zawadzki J., 2014b, *Wykorzystanie oszczędnych modeli harmonicznych z cyklicznością w prognozowaniu brakujących danych*, [w:] *Modelowanie i prognozowanie zjawisk społeczno-gospodarczych. Teoria i praktyka*, red. B. Pawełek, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.
- Zeliaś A., Pawełek B., Wanat S., 2003, *Prognozowanie ekonomiczne. Teoria, przykłady, zadania*, PWN, Warszawa.