

MODELE HARMONICZNE ZE ZŁOŻONĄ SEZONOWOŚCIĄ W PROGNOZOWANIU SZEREGÓW CZASOWYCH Z LUKAMI SYSTEMATYCZNYMI

Maria Szmuksta-Zawadzka, Jan Zawadzki

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

e-mail: jan.zawadzki@zut.edu.pl

Streszczenie: W modelowaniu zmiennych ze złożoną sezonowością dla pełnych danych i danych z lukami niesystematycznymi mogą być wykorzystywane zarówno modele ze zmiennymi zero-jedynkowymi jak i modele harmoniczne. Natomiast w przypadku występowania luk systematycznych- jedynie oszczędne modele harmoniczne. W modelach tych każdy rodzaj wahań opisywany jest za pomocą odrębnych zestawów składowych sinuso- i kosinusoidalnych. Rozważania teoretyczne zostaną zilustrowane przykładem empirycznym.

Słowa kluczowe: modele harmoniczne, sezonowość złożona, brakujące dane

WPROWADZENIE TEORETYCZNE

Celem pracy jest wykazanie, że modele harmoniczne mogą być wykorzystywane w prognozowaniu zmiennych ze złożoną sezonowością w warunkach występowania systematycznych luk w danych. Złożoną sezonowość dla dziennych szeregów czasowych bez luk, uwzględniającą występowanie wahań o cyklu 12 miesięcznym oraz siedmiodniowym można opisać za pomocą odrębnych zespołów zmiennych zero-jedynkowych lub wielomianów trygonometrycznych (por. [Kufel 2010], [Szmuksta-Zawadzka, Zawadzki 2013]).

Model ze złożoną sezonowością z wielomianami trygonometrycznymi o cyklach 12-miesięcznym (rocznym) i 7-dniowym (tygodniowym) przyjmuje postać:

$$\begin{aligned}
 Y_t = & \alpha_1 t + \alpha_0 + \sum_{i=1}^6 (\alpha_{0i} M \sin \omega_i t_m + b_{0i} M \cos \omega_i t_m) \\
 & + \sum_{j=1}^3 (\alpha_{0j} D \sin \omega_j t + b_{0j} D \cos \omega_j t) + U_t
 \end{aligned} \tag{1}$$

gdzie:

$$\omega_i = \frac{2 \cdot \Pi \cdot i}{12} \quad (i = 1, 2, \dots, 6),$$

$$\omega_j = \frac{2 \cdot \Pi \cdot j}{7} \quad (j = 1, 2, 3),$$

$$t_m = 1, 2, \dots, 12, \dots, 12 \cdot r \quad (r - \text{liczba lat}),$$

$$t = 1, 2, \dots, n,$$

Symbole M i D poprzedzające zapisy funkcji trygonometrycznych odnoszą się odpowiednio do miesięcy i dni tygodnia.

W zapisie modelu (1) występują dwie zmienne oznaczające czas: t_m oraz t . Zmienna t_m przyjmuje takie same wartości dla wszystkich dni wchodzących w skład danego miesiąca. Takie jej zdefiniowanie oznacza, że wartości składowych harmonicznych w danym miesiącu będą identyczne. Tym samym więc model z wielomianem trygonometrycznym będzie równoważny modelowi ze zmiennymi zero-jedynkowymi. Równoważność ta polega na tym, że mając oceny parametrów jednego z nich można wyznaczyć parametry drugiego. Ponadto obydwa modele, dla pełnych danych, charakteryzują się identycznymi ocenami parametrów struktury stochastycznej oraz dają identyczne prognozy.

W szeregach czasowych z sezonowością niezłożoną mogą występować dwa rodzaje luk:

- niesystematyczne,
- systematyczne.

Występowanie luk niesystematycznych oznacza, że dostępna jest przynajmniej jedna obserwacja dla każdego podokresu.

Natomiast z lukami systematycznymi będziemy mieć do czynienia wtedy, gdy brak będzie danych przynajmniej o jednym podokresie.

W przypadku złożonej sezonowości możliwych kombinacji występowania luk jest więcej. Luki systematyczne lub niesystematyczne mogą występować bądź tylko dla jednego rodzaju wahań lub dla obu rodzajów wahań jednocześnie. Można też wyróżnić przypadek mieszany, w którym dla jednego rodzaju wahań występują luki systematyczne a drugiego niesystematyczne.

Występowanie systematycznych luk w danych (dziennych o cyklu tygodniowym i miesięcznych o cyklu rocznym), w zależności od ich liczby i układu może pociągać za sobą daleko idące następstwa polegające na tym, że:

- składowe harmoniczne będą ze sobą skorelowane

- niektóre nawet mogą być dokładnie współliniowe (oceny współczynników korelacji będą przyjmować wartości +1 lub -1),
- składowe harmoniczne mogą przyjmować stałe wartości tzn., że ich wariancje będą równe zeru,
- niektóre składowe mogą być linio zależne (tworzyć kombinacje liniowe).

Szerszego omówienia wymagają niektóre sytuacje wymienione w powyższych punktach. W przypadku dokładnej współliniowości dwóch lub więcej składowych szacowanych będzie tyle wersji modelu ile składowych jest dokładnie współliniowych. Wersje te będą nierozróżnialne z punktu widzenia ocen parametrów struktury stochastycznej charakteryzujących własności predyktywne (R^2 , SE, V_{SE}) a także innych statystyk (np. DW). Modele te mogą się różnić (i zazwyczaj będą) ocenami błędów prognoz inter- i (lub) ekstrapolacyjnych. Z podobną sytuacją mamy do czynienia wtedy, gdy wystąpi zjawisko liniowej zależności składowych.

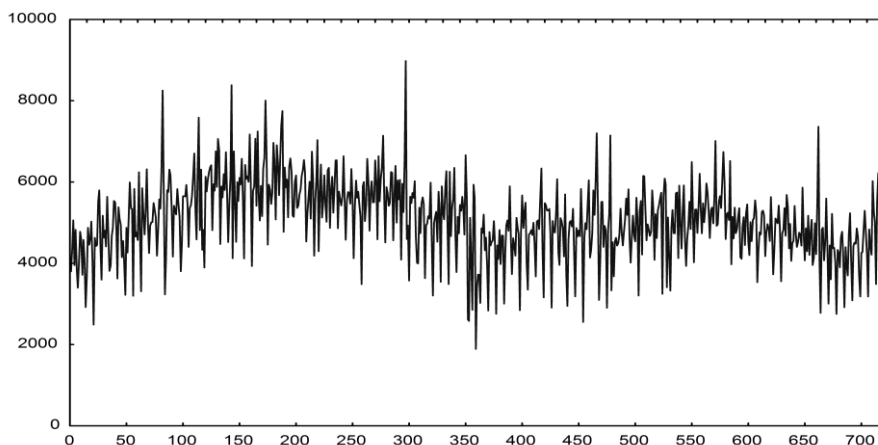
Do stwierdzenia występowania dokładnej współliniowości wystarczy obliczenie współczynników korelacji liniowej. Natomiast wystąpienie zjawiska liniowej zależności składowych harmonicznych nie jest widoczne bezpośrednio w momencie rozpoczęcia modelowania. Uwidacznia się ono w kolejnym etapie zastosowania procedury selekcji *a priori* do wyboru istotnych zmiennych, w tym przypadku składowych harmonicznych w ramach regresji krokowej. Z liniową zależnością będziemy mieć do czynienia wtedy, gdy w kolejnej iteracji otrzymamy co najmniej dwie takie same, co do wartości bezwzględnej, maksymalne oceny współczynników korelacji cząstkowej. Proces estymacji przebiega podobnie jak w przypadku współliniowości tzn. szacuje się tyle wersji modelu ile jest składowych liniowo zależnych. Może się ponadto zdarzyć, że w którejś z następnych iteracji liniową zależność będą tworzyć co najmniej dwie inne składowe. W takim przypadku łączna liczba szacowanych wersji będzie równa iloczynowi liczb składowych liniowo zależnych w kolejnych iteracjach.

W pracy [Szmuksta–Zawadzka, Zawadzki 2013] przedstawiono wykorzystanie modeli szeregu czasowego w prognozowaniu brakujących danych dziennych ze złożoną sezonowością dla luk niesystematycznych. Rozpatrywane były różne warianty występowania tego rodzaju luk równocześnie w określonych dniach i miesiącach. W niniejszej pracy dokonano uogólnienia rozważań na przypadek występowania luk systematycznych.

MODELOWANIE DZIENNEJ SPRZEDAŻY PALIW PŁYNNYCH

Egzemplifikacją przeprowadzonych wyżej rozważań o charakterze teoretycznym będzie przykład empiryczny. Dotyczyć on będzie dziennej sprzedaży paliw płynnych na stacji X w okresie od 7 stycznia pierwszego roku do 31 grudnia roku następnego. Rok trzeci będzie okresem empirycznej weryfikacji prognoz.

Rys. 1. Kształtowanie się dziennej wielkości sprzedaży paliw (w litrach)



Źródło: Szmuksta-Zawadzka i Zawadzki 2011

Rozpatrywany będzie jeden wariant występowania luk systematycznych. Zakładamy, że luki wystąpią w dwóch spośród 12 miesięcy (w sierpniu i październiku) oraz dwóch spośród siedmiu dni tygodnia (środa i piątek). Oznacza to, że luki wystąpią w 294 spośród 724 dni, obejmować więc będą ponad 40% długości szeregu. W szeregu dla pełnych danych składowe harmoniczne dla każdego rodzaju wahań z osobna są ze sobą nieskorelowane. Również nieskorelowane są składowe należące do różnych rodzajów wahań. Natomiast występowanie systematycznych luk w danych sprawiło, że większość bezwzględnych ocen współczynników korelacji składowych harmonicznych o cyklu tygodniowym (13 z 15) przyjęło wartości nie mniejsze od 0,12. Ich oceny mieściły się w przedziale od 0,01 (składowe DC1¹ i DC3) do 0,81 (składowe DS3 i DC1).

Dla składowych o cyklu rocznym w 17 na 54 możliwe kombinacje otrzymano bezwzględne oceny współczynników większe od 0,19. Ocenę najwyższą wynoszącą 0,45 otrzymano dla pary składowych MS1 i MS5. Następnie oceny równe co do wartości bezwzględnej 0,40, ale różniące się znakami, otrzymano dla par składowych: MS2 i MC3 oraz MS4 i MC3.

Występowanie stosunkowo dużej liczby ocen współczynników korelacji dla par składowych różniących się znacznie od zera, może oznaczać, że niektóre z nich mogą być liniowo zależne. W celu sprawdzenia, czy tego rodzaju sytuacja nie ma miejsca, do u wyboru statystycznie istotnych składowych harmonicznych wykorzystana zostanie metoda selekcji *a priori*. W procesie modelowania zostaną także uwzględnione, jako potencjalne zmienne „objaśniające”, zmienne zero-

¹ Symbole DS_j oraz DC_j oznaczają odpowiednio: $D \sin \omega_j, D \cos \omega_j$ a MS_i oraz MC_i: $M \sin \omega_i, M \cos \omega_i$

jedynkowe oznaczające występowanie świąt oraz dni około świątecznych (tzn. dnia poprzedzającego święto oraz dnia po nim następującego).

W Tabeli 1 zestawione zostały oceny współczynników korelacji cząstkowej otrzymane po 15 iteracji dla składowych o maksymalnych, co do wartości bezwzględnej, ich ocenach.

Tabela 1. Zmienne poza równaniem po 15 iteracji o maksymalnych ocenach współczynników korelacji cząstkowej

		Cząstk.	Semicz.		Minimum		
Zmienne	Beta dla	Kor.	Kor.	Teleran.	Toleran.	t(413)	poziom p
DS3	-0,3887	-0,1053	-0,0750	0,0372	0,0372	-2,1524	0,0319
DC2	0,1228	0,1053	0,0750	0,3724	0,3724	2,1524	0,0319
DC3	0,0811	0,1053	0,0750	0,8546	0,7412	2,1524	0,0319

Źródło: obliczenia własne

Oceny te wynoszące 0,1053, otrzymano dla składowych harmonicznych: DS3, DC2 i DC3. Dla pierwszej składowej otrzymano ocenę ujemną a dwóch pozostałych dodatnie.

Otrzymanie identycznych bezwzględnych ocen współczynników korelacji cząstkowej wskazuje, że składowe DS3, DC2 i DC3 są liniowo zależne. Oznacza to konieczność szacowania trzech wersji modelu, z których każda zawierać będzie tylko jedną z tych składowych.

Z kolei w Tabeli 2 zestawione zostały zmienne (składowe) charakteryzujące się maksymalnymi ocenami współczynników korelacji cząstkowej po 21 iteracji dla zmiennych, które dotąd nie weszły do modelu.

Tabela 2. Zmienne poza równaniem po 21 iteracji o maksymalnych ocenach współczynników korelacji cząstkowej

		Cząstk.	Semicz.		Minimum		
	Beta dla	Kor.	Kor.	Teleran.	Toleran.	t(407)	poziom p
MS2	-0,0411	-0,0459	-0,0321	0,6092	0,0370	-0,9272	0,3544
MS4	-0,0411	-0,0459	-0,0321	0,6090	0,0370	-0,9272	0,3544
DC2	-	-	-	0,0000	-	-	-
DC3	-	-	-	0,0000	-	-	-

Źródło: obliczenia własne

Z tabeli tej wynika, że maksymalne, co do wartości bezwzględnej, oceny współczynników korelacji cząstkowej otrzymano dla składowych: MS2 i MS4. Oznacza to, podobnie jak poprzednio, konieczność szacowania równań zawierających jedną z nich. Brak charakterystyk statystycznych dla zmiennych (składowych) DC2 i DC3 świadczy o tym, że tworzą one kombinację liniową i są

liniowo zależne. Z uwagi na to, że po 15-tej iteracji otrzymano identyczne oceny współczynników korelacji cząstkowej dla trzech składowych a po 21-szej dla dwóch, trzeba będzie szacować sześć równań zawierających po jednej składowej z każdej grupy. Równania te zawierać będą oczywiście te same pozostałe zmienne wybrane w wcześniejszych iteracjach.

W Tabeli 3 zestawione zostały oszacowania wybranych parametrów struktury stochastycznej, będące syntetycznymi miernikami ich własności predykcyjnych wyrażonych za pomocą współczynników determinacji (R^2) oraz odchyłeń standardowych składników losowych (SE) i współczynników zmienności losowej (V_{SE}) dla sześciu wersji modeli.

Tabela 3. Wyniki oszacowań modeli oszczędnych dla systematycznych luk danych

	DS3_MS2	DS3_MS4	DC2_MS2	DC2_MS4	DC3_MS2	DC3_MS4
R^2	0,4869	0,4869	0,4869	0,4869	0,4869	0,4869
SE	690,01	690,01	690,01	690,01	690,01	690,01
V_{SE}	0,1399	0,1399	0,1399	0,1399	0,1399	0,1399

Źródło: obliczenia własne

Równania te oznaczono podwójnymi symbolami oddzielonymi podkreślnikami, przy czym pierwszy symbol odnosi się do jednej z trzech składowych liniowo zależnych o cyklu tygodniowym (DS3, DC2 lub DC3) a drugi dla składowych zależnych liniowo o cyklu rocznym (MS2 lub MS4).

Z uwagi na to, że oceny syntetycznych mierników własności predykcyjnych są takie same, równania te są nierozróżnialne z tego punktu widzenia. W kolejnych dwóch wierszach zamieszczone zostały oceny błędów ogółem prognoz interpolacyjnych i ekstrapolacyjnych. Zgodnie z tym co powiedziano w poprzednim podrozdziale, mogą się one i zazwyczaj będą różnić ocenami błędów prognoz inter- i (lub) ekstrapolacyjnych.

WYNIKI PROGNOZOWANIA INTER- I EKSTRAPOLACYJNEGO

Na podstawie oszacowanych równań dla dni i miesięcy, w których wystąpiły luki zostały wyznaczone prognozy interpolacyjne. Z uwagi na to, że luki otrzymano przez „wymazanie” części danych z pełnego szeregu, możliwe było obliczenie średnich względnych błędów tych prognoz. W wierszu drugim Tabeli 4 zamieszczone zostały oceny błędów prognoz interpolacyjnych ogółem obliczone dla 294 dni. Pozostałe wiersze zawierają ich dekompozycję dla tych dni tygodnia i miesięcy, w których wystąpiły luki. W celach porównawczych podane zostały uśrednione oceny względnych błędów prognoz dla pozostałych dni i miesięcy nie zawierających luk.

Tabela 4. Oceny błędów prognoz interpolacyjnych

	DS3_MS2	DS3_MS4	DC2_MS2	DC2_MS4	DC3_MS2	DC3_MS4
Ogółem	0,2130	0,2186	0,1133	0,1189	0,1031	0,1087
Środa	>0,3000	>0,3000	0,1339	0,1346	0,1127	0,1146
Piątek	0,2114	0,2127	0,1172	0,1194	0,1093	0,1104
Poz. dni	0,0836	0,0989	0,0836	0,0989	0,0836	0,0989
Sierpień	0,1118	0,1188	0,0959	0,1129	0,0904	0,1080
Październik	0,1483	0,1679	0,1011	0,1107	0,0954	0,1045
Poz. m-ce	0,2736	0,2736	0,1244	0,1244	0,1100	0,1100

Źródło: obliczenia własne

Spośród sześciu oszacowanych modeli zdecydowanie niższymi ocenami błędów prognoz interpolacyjnych ogółem charakteryzują się modele z dodatnimi ocenami współczynników korelacji cząstkowej. Oceny błędów zawarte są w przedziale od 10,31% (DC3_MS24) do 11,89% (DC2_MS4). Natomiast znacznie wyższe oceny błędów prognoz, przynajmniej o 9 punktów procentowych, otrzymano na podstawie modeli o ujemnej ocenie współczynnika korelacji cząstkowej (DS3_MS2, DS3_MS4).

Obecnie analizie poddane zostaną oceny błędów dla dni i miesięcy z lukami. Minimalną ocenę błędu dla środy, wynoszącą 11,27%, otrzymano dla predyktora DC3_MS2 i tylko nieznacznie wyższą dla predyktora C3_MS3. Oceny błędów prognoz z pierwszą składową DC2 były wyższe o ok. 2 punkty procentowe. Zdecydowanie najwyższe oceny błędów, przekraczające 30 procent, otrzymano dla predyktorów o ujemnej ocenie współczynnika korelacji cząstkowej. W przypadku piątku najlepszym okazał się także predyktor DC3_MS2. Oceny te dla pozostałych trzech predyktorów o dodatnich ocenach współczynników korelacji cząstkowej były wyższe nie więcej niż o 1 punkt procentowy. Natomiast najwyższe oceny błędów, podobnie jak dla środy, otrzymano dla predyktorów o ujemnej ocenie tego współczynnika. Uśrednione oceny błędów dla pozostałych dni przyjęły wartość 8,36% dla predyktorów z drugą składową MS2 oraz 9,89% ze składową MS3.

Ocenę minimalną dla sierpnia, wynoszącą 9,04%, otrzymano także dla predyktora DC3_MS2. Oceny błędów prognoz otrzymanych na podstawie pozostałych modeli były nie więcej niż o 2,89 punktu wyższe, przy czym dotyczy to także modeli DS3_MS2 oraz DS3_MS4. Oceny błędów otrzymanych na podstawie predyktorów o dodatnich ocenach współczynników korelacji cząstkowej zawarte były w przedziale od 9,54% (DC3_MS2) do 11,17% (DC2_MWS4). Natomiast dla predyktorów o ujemnej ocenie tego współczynnika wynosiły one odpowiednio: 14,83% i 16,79%. Uśrednione oceny otrzymane dla tych

predyktorów wynosiły 27,36%. Dla predyktorów z pierwszą składową DC3 kształtowały się one na poziomie 11,00%, a ze składową DC2 12,44%.

Na podstawie oszacowanych równań zostały wyznaczone także prognozy ekstrapolacyjne a następnie przeprowadzona ich empiryczna weryfikacja tzn. obliczone zostały przeciętne względne błędy prognoz. Było to możliwe, ponieważ dla okresu empirycznej weryfikacji prognoz znane były realizacje zmiennej prognozowanej.

W Tabeli 5 zestawione zostały oceny błędów prognoz ekstrapolacyjnych ogółem oraz błędy prognoz zdezagregowanych na dni tygodnia i miesiące w których wystąpiły luki.

Z informacji zawartych w drugim wierszu wynika, że oceny błędów prognoz ekstrapolacyjnych ogółem są o ok. 5-6 punktów procentowych wyższe od ocen błędów prognoz interpolacyjnych, otrzymanych na podstawie predyktorów z dodatnimi ocenami współczynników korelacji cząstkowej. Natomiast są one o ok. 3 punkty niższe od ocen błędów prognoz interpolacyjnych otrzymanych na podstawie predyktorów z ujemną oceną tego współczynnika. Najniższą ocenę wynoszącą 15,94% charakteryzuje się predyktor DC3_MS2. Dla pozostałych trzech predyktorów z dodatnimi ocenami współczynników korelacji cząstkowej są one wyższe nie więcej niż o 0,57 punktu procentowego.

Tabela 5. Oceny błędów prognoz ekstrapolacyjnych

	DS3_MS2	DS3_MS4	DC2_MS2	DC2_MS4	DC3_MS2	DC3_MS4
Ogółem	0,1848	0,1877	0,1626	0,1651	0,1595	0,1618
Środa	0,2975	0,2978	0,1925	0,1922	0,1319	0,1353
Piątek	0,1717	0,1777	0,1202	0,1235	0,1583	0,1565
Poz. dni	0,1651	0,1680	0,1651	0,1680	0,1651	0,1680
Sierpień	0,1624	0,1897	0,1738	0,2152	0,1716	0,2130
Październik	0,2065	0,2123	0,1741	0,1607	0,1749	0,1590
Poz. m-ce	0,1874	0,1874	0,1612	0,1612	0,1528	0,1528

Źródło: obliczenia własne

Oceny wyższe o ok. 2-3 punkty procentowe otrzymano dla predyktorów z ujemnymi ocenami współczynników korelacji cząstkowej (DS3_MS2 i DS3_MS4). W wierszach od 3 do 5 zestawiono oceny błędów prognoz dla dni tygodnia z lukami i bez luk. W przypadku środy widoczne jest bardzo duże ich zróżnicowanie dla par predyktorów, różniących się pierwszą składową. Zdecydowanie najniższe oceny, kształtujące się w granicach 13 procent otrzymano dla predyktorów: DC3_MS2 oraz DC3_MS4. W przypadku drugiej pary z dodatnią oceną współczynnika korelacji cząstkowej (MC2_MS2 oraz MC2_MS4) oceny te są o około 6 punktów procentowych wyższe. Najwyższe oceny błędów, wynoszące ok. 30 procent otrzymano dla pary predyktorów

z ujemną oceną współczynnika korelacji cząstkowej (DS3_MS2 oraz DS3_MS4). Najniższe oceny dla piątku, wynoszące odpowiednio 12,01% oraz 12,34% otrzymano dla pary predyktorów z pierwszą składową DC2. W przypadku pary z pierwszą składową DC3 oceny te są o około 3 punkty procentowe wyższe. Oceny błędów otrzymane dla predyktorów o ujemnej ocenie współczynnika korelacji cząstkowej (z pierwszą składową DS3) są o dalsze 1-2 punkty procentowe wyższe. Uśrednione oceny błędów dla pozostałych dni tygodnia przyjęły wartość 16,51% dla predyktorów z drugą składową MS2 i 16,80% ze składową MS4.

Kolejne trzy wiersze zawierają oceny błędów prognoz dla miesięcy z lukami i uśrednione dla miesięcy bez luk. Dla sierpnia niższe oceny błędów dla par predyktorów różniących się pierwszą składową otrzymano dla tych, w których występowała druga składowa MS2. Ich oceny zawarte są w przedziale od 16,15% (dla predyktora DS3_MS2) do 17,40% dla predyktora DC2_DS2). Oceny błędów otrzymane dla predyktorów z drugą składową są w ramach poszczególnych par są o ok. 2,5 – 4 procent wyższe. Dla października widoczny był wpływ znaku oceny współczynników korelacji cząstkowej. Dla predyktorów z dodatnią jego oceną, błędy kształtowały się na poziomie od 15,90 % (DC3_MS4) do 17,49% (DC3_MS2). Dla predyktorów z oceną ujemną, przekraczają one 20 procent. Dla pozostałych miesięcy oceny błędów otrzymanych dla par predyktorów są takie same. Kształtują się one w przedziale od 15,28% dla predyktorów z pierwszą składową DC3 do 18,74% ze składową DS3.

WNIOSKI

Z przeprowadzonych w pracy rozważań można wyprowadzić następujące wnioski:

1. W procesie wyboru statystycznie istotnych składowych harmonicznych potwierdzona została zasadność stosowania procedury *selekcji a priori* w modelowaniu w warunkach występowania systematycznych luk w danych.
2. Jej zastosowanie pozwala zidentyfikować liniowo zależne składowe harmoniczne. Tworzą ją, dla przynajmniej jednej z kolejnych iteracji, składowe o maksymalnych bezwzględnych ocenach współczynników korelacji cząstkowej. Oznacza to konieczność szacowania liczby wersji modelu będącej iloczynem liczb: zmiennych liniowo zależnych i iteracji.
3. Oszacowane wersje charakteryzują się identycznymi ocenami syntetycznych ocen własności predykcyjnych (R^2 , Se , V_{Se}). Natomiast różnią się ocenami błędów prognoz inter- i ekstrapolacyjnych.
4. Z przykładu empirycznego, obejmującego występowanie luk systematycznych w dwóch wybranych dniach tygodnia i dwóch miesiącach, wynika, że minimalne oceny błędów ogółem prognoz interpolacyjnych są niższe niż prognoz ekstrapolacyjnych.
5. Istotnym czynnikiem determinującym dokładność prognoz okazał ujemny znak współczynnika korelacji cząstkowej dla jednej z liniowo zależnych składowych

o cyklu tygodniowym (DS3). Oceny błędów ogółem prognoz interpolacyjnych były blisko dwukrotnie wyższe od błędów otrzymanych dla predyktorów z dodatnimi jego wartościami. Dla prognoz ekstrapolacyjnych różnice w dokładności wynosiły 2-3 punkty procentowe.

Z przedstawionych wyżej wniosków wynika, że oszczędne modele harmoniczne szeregu czasowego ze złożoną sezonowością dla danych dziennych, w których wahania o cyklu rocznym i tygodniowym opisywanych za pomocą wielomianów trygonometrycznych, mogą być z powodzeniem wykorzystywane w prognozowaniu w warunkach występowania luk systematycznych obejmujących nawet ok. 40 % długości szeregu.

BIBLIOGRAFIA

- Kufel T. (2010) Ekonometryczna analiza cykliczności procesów gospodarczych o wysokiej częstotliwości obserwowania, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń.
- Szmuksta-Zawadzka M., Zawadzki J. (2011) The ex post analysis of the forecast accuracy of liquid fuel sales for daily data, „Econometrics”, Research Papers of Wrocław University of Economics Nr 32, s.194-204, Wrocław.
- Szmuksta-Zawadzka M., Zawadzki J. (2013) Wykorzystanie oszczędnych modeli harmonicznych z cyklicznością w prognozowaniu brakujących danych(w druku)

HARMONICAL MODELS WITH COMPLEX SEASONALITY IN FORECASTING TIME SERIES WITH SYSTEMATIC GAPS

Abstract: In the modeling of the variables with complex seasonality for complete time series and with unsystematic data gaps can be used both types of models: with dummy variables and harmonic models. However, in modeling variable with systematic gaps can be used only harmonic models. In these models, each type of fluctuation is described by separate sets of sine- and cosine component. Theoretical considerations are illustrated by an empirical example.

Keywords: harmonic models, complex seasonality, missing data