

Jan ZAWADZKI

## MODELE HYBRYDOWE W PROGNOZOWANIU BRAKUJĄCYCH DANYCH W SZEREGACH O BARDZO WYSOKIEJ CZĘSTOTLIWOŚCI OBSERWOWANIA

### HYBRID MODELS IN THE FORECASTING OF MISSING DATA IN SERIES WITH A VERY HIGH OBSERVING FREQUENCY

Katedra Zastosowań Matematyki w Ekonomii, Zachodniopomorski Uniwersytet Technologiczny  
 w Szczecinie, ul. Klemensa Janickiego 31, 71-270 Szczecin, ORCID: 0000-0002-2996-3180  
 e-mail: jzawadzki@zut.edu.pl

**Summary.** The paper presents the application of single and double hybrid additive and multiplicative models in forecasting missing data in high frequency time series with cyclical fluctuations with for unsystematic gaps. Complex seasonal fluctuations with annual, weekly and daily cycles will overlap the trend in an additive or multiplicative manner. Fluctuations with even cycle lengths (12-month and 24-hour) were described using regular hierarchical models. The demand for electricity in hourly periods will be modelled and forecasted.

**Słowa kluczowe:** dane o wysokiej częstotliwości, sezonowość złożona, modele hybrydowe.  
**Key words:** high frequency data, complex seasonality, hybrid models.

#### WSTĘP

W modelowaniu zmiennych o wysokich lub bardzo wysokich częstotliwościach mogą być wykorzystywane modele szeregu czasowego z podwójnie (dla danych dziennych) lub potrójnie (dla danych godzinnych) złożonymi wahaniami sezonowymi w postaci addytywnej lub multiplikatywnej. Model multiplikatywny najczęściej przyjmuje postać modelu wykładniczego. Składowe modelu addytywnego (1) występują w wykładniku potęgi logarytmu naturalnego. Po obustronnym zloga-rytmowaniu otrzymuje się model liniowy względem logarytmów zmiennej prognozowanej ( $\ln Y_t$ ).

Przykładowy zapis modelu addytywnego z liniowym trendem i potrójnie złożonymi wahaniami sezonowymi (Kufel 2010; Szmuksta-Zawadzka i Zawadzki 2011, 2014):

$$Y_t = \alpha_1 t + \alpha_0 + \sum_{i=1}^{12} b_{0i} M_{it} + \sum_{j=1}^7 c_{0j} D_{jt} + \sum_{k=1}^{24} d_{0k} G_{kt} + U_t \quad (1)$$

przy warunkach:

$$\sum_{i=1}^{12} b_{0i} = \sum_{j=1}^7 c_{0j} = \sum_{k=1}^{24} d_{0k} = 0$$

gdzie:

$M_{it}$  – miesiące,

$D_{jt}$  – dni tygodnia,

$G_{kt}$  – godziny w cyklu dobowym.

Na parametry opisujące wahania sezonowe o cyklach: rocznym, tygodniowym i dobowym nałożony jest warunek sumowalności do zera. W celu uniknięcia współliniowości parametrów uwzględnia się je, wprowadzając w jednym z podokresów każdego cyklu wartości równe minus jeden. Wartości parametrów:  $b_{0i}$ ,  $c_{0j}$  i  $d_{0k}$  dla poszczególnych okresów składowych interpretuje się jako odchylenia od odpowiednich średnich. W podejściu równoważnym warunki te można ominąć, wprowadzając zamiast elementów minus jeden zera. Wtedy parametry interpretuje się jako odchylenia od określonego podokresu.

W modelowaniu i prognozowaniu zmiennych na podstawie danych w postaci szeregów czasowych mogą być wykorzystywane także modele hierarchiczne. Podstawy teoretyczne budowy i estymacji takich modeli i przykłady empiryczne podają m.in.: Little i Rubin (1987); Raudenbush i Bryk (2002); Szmuksta-Zawadzka i Zawadzki (2002); Junger i de Leon (2015).

W przypadku szeregów dla danych godzinnych ze złożonymi wahaniami sezonowymi do opisu wahań o parzystej długości cykli, w naszym przypadku o cyklach rocznym i dobowym, mogą być wykorzystywane regularne modele hierarchiczne. Modele, w których przynajmniej jeden z dwóch wymienionych wyżej rodzajów wahań opisywany jest za pomocą modeli hierarchicznych, nazywać będziemy modelami hybrydowymi. Trzeci rodzaj wahań złożonych – wahania 7-dniowe (tygodniowe), o nieparzystej długości cyklu, opisywane będą za pomocą zmiennych zero-jedynkowych  $D_{jt}$ .

Z definicji zamieszczonej w pracy Szmuksty-Zawadzkiej i Zawadzkiego (2002) wynika, że regularnymi modelami hierarchicznymi nazywane są modele, dla których podzielniki  $p_i$  długości cyklu wahań okresowych (sezonowych)  $m$ , będącego liczbą parzystą, spełniają jednocześnie dwa warunki:

$$2 \leq p_i \leq \frac{m}{2} \text{ oraz } \prod_i p_i = m \quad (2)$$

Dla danych miesięcznych o cyklu rocznym ( $m = 12$ ) oraz danych dekadowych (Szmuksta-Zawadzka i Zawadzki 2004) o takiej samej długości cyklu ( $m = 36$ ) będą to modele pojedyncze. W przypadku danych godzinnych za pomocą modeli hierarchicznych mogą być opisywane wahania o cyklu rocznym ( $m = 12$ ) oraz wahania o cyklu dobowym ( $m = 24$ ). Jeżeli za ich pomocą opisywany jest jeden z wymienionych wyżej rodzajów wahań okresowych, to będą to modele pojedyncze. Natomiast jeżeli opisywane będą jednocześnie obydwa rodzaje wahań, to będą to modele podwójne.

Modele hierarchiczne dla wahań o cyklu 12-miesięcznym i 24-godzinnym oznaczane będą odpowiednio jako HM i HG. Liczby występujące po tych symbolach oznaczać będą kolejne podzielniki długości cyklu wahań.

Liczba regularnych modeli hierarchicznych dla danego cyklu wahań jest równa liczbie permutacji oraz permutacji z powtórzeniami podzielników  $p_i$ . Dla danych miesięcznych o cyklu rocznym ( $m = 12$ ) wynosi ona 7, w tym:

- 4 modele dwustopniowe (HM26, HM34, HM43, HM62);
- 3 modele trójstopniowe (HM223, HM232, HM322).

Dla danych o cyklu dobowym ( $m = 24$ ) ich liczba wynosi 19, w tym:

- 6 modeli dwustopniowych (HG2,12, HG38, HG46, H64, HG83, HG12,2),

- 9 modeli trzystopniowych (HG226, HG234, HG243, HG262, HG324, HG342, HG423, HG432, HG622),
- 4 modele czterostopniowe (HG2223, HG2232, HG2322, HG3222).

Zapis ogólny czterostopniowego modelu addytywnego z liniowym trendem dla cyklu dobowego jest następujący:

$$Y_{srltk} = \alpha_1 t + \alpha_0 + \sum_{s=1}^{p_1} h_{0s} G_{st} + \sum_{r=1}^{p_2} h_{0sr} G_{srt} + \sum_{l=1}^{p_3} h_{0srl} G_{srlt} + \sum_{k=1}^{p_4} h_{0srk} G_{srlkt} + U_{srlkt} \quad (3)$$

przy warunkach:

$$\sum_{s=1}^{p_1} h_{0s} = \sum_{r=1}^{p_2} h_{0sr} = \sum_{l=1}^{p_3} h_{0srl} = \sum_{k=1}^{p_4} h_{0srk} = 0$$

$$\prod_i p_i = m$$

Jak wynika z informacji podanych wyżej, dla cyklu 12-miesięcznego szacuje się 7 pojedynczych modeli hybrydowych, a dla cyklu 24-godzinnego – 19. Natomiast liczba szacowanych podwójnych modeli hybrydowych równa jest iloczynowi liczb modeli pojedynczych i wynosi  $7 \cdot 19 = 133$ . Zatem dla jednej postaci analitycznej szacowanych jest łącznie 159 modeli hybrydowych. Do tego należy doliczyć jeden model klasyczny (z trzema zespołami zmiennych zero-jedynkowych) dany równaniem (1). Będzie on punktem odniesienia dla modeli hybrydowych.

Zasadniczą zaletą modeli hierarchicznych jest to, że dla wahań o długości cyklu  $m$  maksymalna liczba szacowanych parametrów wynosi nie więcej niż  $m/2$ , wobec  $m-1$  parametrów w modelach klasycznych.

Liczba szacowanych parametrów w modelach hierarchicznych jest sumą dzielników długości cyklu wahań pomniejszoną o ich liczbę. Przykładowo w modelach: HG2,12 i HG12,2 zamiast 23 parametrów modelu klasycznego szacuje się  $2 + 12 - 2 = 12$  oraz  $12 + 2 - 2 = 12$ . Natomiast w modelach 4-stopniowych: HG2223, HG2232, HG2322, HG3222 liczba ta wynosi 5. W przypadku danych o wysokiej częstotliwości zwiększenie liczby stopni swobody nie jest tak istotne jak w przypadku szeregów dla danych: miesięcznych, dekadowych czy dziennych. Jednak modele hierarchiczne, których parametry są uśrednionymi parametrami klasycznych modeli szeregu czasowego ze zmiennymi zero-jedynkowymi, „wyrównują” obserwacje, w których występują luki w danych lub będące skutkiem zakłóceń losowych.

W literaturze postuluje się (zob. np. Szmuksta-Zawadzka i Zawadzki 2014), aby kryterium wyboru modeli dla celów prognozowania były nie parametry charakteryzujące jakość modeli, a mierniki o minimalnych ocenach błędów względnych prognoz interpolacyjnych lub ekstrapolacyjnych (*ex post*):  $MAPE_I$  i  $MAPE_E$  dane wzorami:

$$MAPE_I = \frac{1}{z} \sum_{t \in I_{ep}} \frac{|Y_t - Y_{tp}^{(I)}|}{Y_t} \quad (4)$$

$$MAPE_E = \frac{1}{\tau} \sum_{t=n+1}^{\tau} \frac{|Y_t - Y_{tp}^{(E)}|}{Y_t} \quad (5)$$

gdzie:

$Y_t$  – wielkości zaobserwowane,

$Y_{tp}^{(I)}$  – prognozy wyznaczone dla kryterium minimalizującego błędy prognoz interpolacyjnych,

$Y_{tp}^{(E)}$  – prognozy wyznaczone dla kryterium minimalizującego błędy prognoz ekstrapolacyjnych,

$z$  – liczba prognoz interpolacyjnych (luk w danych),

$t \in I_{ep}$  –  $t$  należy do przedziału empirycznej weryfikacji prognoz,

$\tau$  – liczba prognoz ekstrapolacyjnych.

Przesłanką przemawiającą za wyborem tych mierników jest to, że dość często modelom o najlepszych własnościach predykcyjnych nie odpowiadają minimalne oceny błędów prognoz. Niejako przy okazji wyznaczania wymienionych wyżej mierników, stanowiących kryterium wyboru najlepszych modeli hybrydowych, liczone będą przeciętne względne błędy prognoz „konkurencyjnych” – tzn. dla prognoz interpolacyjnych błędy prognoz ekstrapolacyjnych  $MAPE_{I-E}$ , a dla ekstrapolacyjnych względne błędy prognoz  $MAPE_{E-I}$ . Oblicza się je ze wzorów:

$$MAPE_{I-E} = \frac{1}{\tau} \sum_{t=n+1}^{\tau} \frac{|Y_t - Y_{tp}^{(I)}|}{Y_t} \quad (6)$$

$$MAPE_{E-I} = \frac{1}{z} \sum_{t \in I_{ep}} \frac{|Y_t - Y_{tp}^{(E)}|}{Y_t} \quad (7)$$

Oznaczenia są takie same jak dla równań (4) i (5).

Wyznaczenie mierników danych wzorami (6) i (7) pozwoli określić różnice, jeżeli błędy uzyskane dla miernika głównego ( $MAPE_I$  i  $MAPE_E$ ) są mniejsze dla tego samego rodzaju prognoz od błędów otrzymanych dla drugiego miernika.

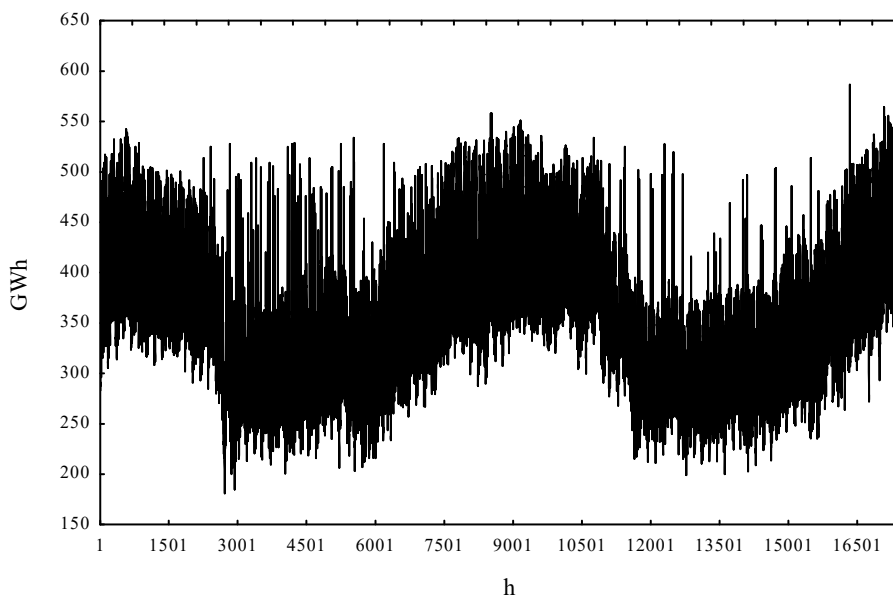
## PRZEDMIOT I ZAKRES BADAŃ EMPIRYCZNYCH

Modelowaniu, a następnie prognozowaniu poddano zapotrzebowanie na energię elektryczną w okresach godzinnych w aglomeracji A.

Dane statystyczne o kształtowaniu się zapotrzebowania na moc (w MWh) pochodzą z banku danych Katedry Zastosowań Matematyki w Ekonomii ZUT. Szereg wyjściowy (bez luk) obejmował okres dwóch lat, tj. 17 520 obserwacji. Trzeci rok był okresem empirycznej weryfikacji prognoz.

Kształtowanie się zmiennej w okresach estymacyjnym i prognozowanym zostało przedstawione w postaci graficznej na ryc. 1.

W pracy Szmuksty-Zawadzkiej i Zawadzkiego (2014), poświęconej zastosowaniu modeli hybrydowych szeregu czasowego w prognozowaniu zapotrzebowania na energię elektryczną w aglomeracji A, rozpatrywany był jeden wariant luk niesystematycznych obejmujący: 2 miesiące, 2 tygodnie i 8 okresów godzinnych.



Ryc. 1. Zapotrzebowanie na moc energetyczną (GWh) w okresach godzinnych w aglomeracji A  
Źródło: Bank danych Katedry Zastosowań Matematyki w Ekonomii.

Niniejsza praca stanowi rozwinięcie badań przywołanych wyżej o hybrydowe modele przyczynowo-skutkowe, przy czym rozpatrywany będzie jeden wariant niesystematycznych luk w danych. Luki będą występowały w okresie od 1 lipca pierwszego roku do 30 czerwca drugiego roku: w ośmiu okresach godzinnych (2, 6, 8, 10, 14, 16, 18, 22), dwóch dniach tygodnia (wtorek, czwartek) oraz dwóch miesiący (sierpień, październik).

Szacowane będą dwa rodzaje modeli hybrydowych. Pierwszy rodzaj to pojedyncze oraz podwójne modele hybrydowe szeregu czasowego, będące kombinacjami modeli pojedynczych. Natomiast drugi rodzaj to modele typu przyczynowego. Zawierać one będą dodatkowo: zmienną endogeniczną opóźnioną o 24 godziny, zmienne zero-jedynkowe opisujące występowanie świąt oraz jednego dnia przedświątecznego (Wielkiej Soboty). Wprowadzenie opóźnionej zmiennej prognozowanej wynikało z faktu 24-godzinnego wyprzedzenia dotyczącego składania zamówień na moc elektryczną.

Dla obydwu rodzajów modeli szacowane były dwie postaci analityczne – liniowa i wykładnicza, a ściślej: postać zlogarytmowana zmiennej prognozowanej za pomocą logarytmu naturalnego ( $\ln Y_t$ ). Na podstawie predyktorów opartych na oszacowanych równaniach dla okresów, w których wystąpiły luki, zostały wyznaczone prognozy interpolacyjne. Natomiast dla okresu empirycznej weryfikacji prognoz, wybiegających poza przedział czasowy „próby”, wyznaczonych zostało 8760 prognoz ekstrapolacyjnych *ex post*.

## WYNIKI MODELOWANIA I PROGNOZOWANIA MODELI HYBRYDOWYCH SZEREGU CZASOWEGO

W rozdziale tym przedstawione zostaną wyniki modelowania i prognozowania inter- oraz ekstrapolacyjnego. Najpierw przeprowadzona zostanie analiza wyników dla pojedynczych i podwójnych modeli hybrydowych szeregu czasowego, a następnie dla modeli hybrydowych typu przyczynowego zawierających dodatkowo: opóźnioną o 24 godziny zmienną prognozowaną i zmienne zero-jedynkowe oznaczające święta i Wielką Sobotę.

Do prognozowania zostały wybrane modele hybrydowe charakteryzujące się minimalnymi ocenami względnych błędów prognoz interpolacyjnych ( $MAPE_I$ ) i prognoz ekstrapolacyjnych ( $MAPE_E$ ). Już ze wstępnego porównania struktury najlepszych modeli hybrydowych wynika, że są to najczęściej różne modele. Dlatego analiza będzie prowadzona osobno dla predyktorów charakteryzujących się minimalnymi ocenami wymienionych wyżej mierników.

Przed przejściem do prezentacji i analizy wyników modelowania oraz prognozowania inter- i ekstrapolacyjnego prognoz otrzymanych, na podstawie hybrydowych modeli szeregu czasowego, przedstawione zostaną wyniki estymacji parametrów najlepszego podwójnego liniowego modelu hybrydowego, wraz z interpretacją dla kryterium  $MAPE_I$ . Model ten w tab. 1 został oznaczony jako LI\_HG432\_HM62. W szacowaniu parametrów wykorzystano wspomniane we Wstępie podejście alternatywne polegające na zastąpieniu w przypadku ostatnich podokresów wartości minus jeden zerami, przy czym w celu uniknięcia współliniowości szacuje się o jeden parametr mniej. Oceny parametrów interpretowane są wtedy jako różnice (odchylenia od ostatniego podokresu).

Wyniki oszacowania parametrów oraz statystyk testu t Studenta zawiera tab. 1.

Tabela 1. Oceny parametrów i wartości statystyk t podwójnego liniowego modelu hybrydowego LI\_H432\_HM62

Znaczenie zmiennych	Oceny parametrów	Statystyki t
Wyraz wolny	445,30	250,350
t	8,39E-06	0,108
Godziny 1–6	-84,54	-91,400
Godziny 7–12	-7,24	-7,852
Godziny 13–18	3,51	3,688
Pierwszy okres dwugodzinny w okresie sześciogodzinnym	4,00	4,918
Drugi okres dwugodzinny w okresie sześciogodzinnym	5,42	6,642
Pierwsza godzina w okresie dwugodzinnym	2,09	3,065
Styczeń–luty	10,81	7,964
Marzec–kwiecień	-32,77	-25,320
Maj–czerwiec	-106,70	-85,570
Lipiec–sierpień	-97,43	-83,330
Wrzesień–październik	-68,83	-59,480
Pierwszy miesiąc w okresie dwumiesięcznym	2,00	2,997
Sobota	-23,84	-19,100
Niedziela	-54,64	-43,930
Poniedziałek	-8,45	-6,790
Wtorek	-3,72	-2,982
Środa	-1,91	-1,533
Czwartek	-1,04	-0,836

Pierwszy podzielnik trójstopniowego modelu hierarchicznego HG432, opisującego wahania o cyklu 24-godzinnym, równy:  $24 : 4 = 6$  wyodrębnia 4 okresy 6-godzinne. Ujemne oceny parametrów dla pierwszego i drugiego okresu informują o tym, że zapotrzebowanie na moc w godzinach 1–6 i 7–12 było przeciętnie niższe, niż w godzinach 18–24, odpowiednio o 84,54 i 7,24 GWh, natomiast w godzinach 13–18 – wyższe o 3,51 GWh. Drugi stopień hierarchii wyznacza podzielnik  $6 : 3 = 2$ , który każdy okres 6-godzinny dzieli na 3 okresy 2-godzinne. Dodatkowo oceny parametrów dla pierwszego i drugiego okresu oznaczają, że zapotrzebowanie na moc było o 5,42 i 2,09 GWh większe niż w trzecim okresie dwugodzinnym. Trzeci podzielnik

(2 : 2–1) dzieli każdy okres dwugodzinny na okresy godzinne. Dodatnia ocena parametru, wynosząca 2,09 GWh, informuje, że zapotrzebowanie na moc w godzinach nieparzystych jest większe o tę wartość.

Obecnie przechodzimy do interpretacji parametrów dwustopniowego modelu hierarchicznego, drugiej części składowej modelu hybrydowego dla cyklu 12-miesięcznego HM62. Pierwszy dzielnik:  $12 : 6 = 2$  wyznacza 6 okresów dwumiesięcznych obejmujących kolejne miesiące. Każda z pięciu zamieszczonych w tablicy ocen parametrów będzie interpretowana jako przeciętne odchylenie zapotrzebowania na moc w listopadzie i grudniu. Dodatnia ocena dla pierwszego okresu oznacza, że zapotrzebowanie na moc w miesiącach styczeń–luty było o 10,81 GWh wyższe niż w okresie szóstym. Ponieważ dla pozostałych czterech okresów dwumiesięcznych otrzymano oceny ujemne, popyt na moc był w tych okresach niższy. Ich wartości bezwzględne wahały się w przedziale od 32,77 GWh dla okresu drugiego (marzec–kwiecień) do 106,76 GWh dla okresu trzeciego (maj–czerwiec). Drugi dzielnik:  $2 : 2 = 1$  delimituje podział każdego okresu dwumiesięcznego na 2 okresy miesięczne. Ocena parametru równa 2,0 GWh informuje o większym zapotrzebowaniu w miesiącach nieparzystych. Zwraca uwagę fakt, że w obu modelach wszystkie parametry są statystycznie istotne.

W ostatnich sześciu wierszach podane zostały oceny parametrów ostatniej składowej części systematycznej klasycznego modelu szeregu czasowego dotyczącego dni tygodnia. Z ich kolejności wynika, że Nowy Rok w pierwszym roku przedziału czasowego „próby” przypadał w sobotę. Tym samym więc punktem odniesienia w prowadzonych porównaniach będzie piątek. Wszystkie oceny parametrów przyjęły wartości ujemne. Oznacza to, że zapotrzebowanie w tych dniach było niższe niż w piątek. Największe różnice dotyczące wartości bezwzględnej, wynoszące 54,64 i 23,84 GWh, otrzymano odpowiednio dla niedzieli i soboty. Dla poniedziałku mieściły się one w przedziale od 1,04 GWh do 8,45 GWh.

W tabeli 2 zestawiono charakterystyki wybranych parametrów struktury stochastycznej ( $R^2$ ,  $S_e$ ,  $V_{Se}$ ) dla najlepszych pojedynczych i podwójnych modeli hybrydowych szeregu czasowego w postaci liniowej i wykładniczej, charakteryzujących się najniższymi ocenami błędów prognoz interpolacyjnych ( $MAPE_I$ ).

Tabela 2. Oceny parametrów struktury stochastycznej i błędy prognoz modeli hybrydowych dla kryterium  $MAPE_I$

Model	$R^2$ [%]	$S_e$ [GWh]	$V_{Se}$ [%]	$MAPE_I$ [%]	$MAPE_{LE}$ [%]
LI_0/1_HM62	77,90	34,24	9,369	6,839	9,749
WI_0/1_HM62	79,42	33,82	9,184	6,821	9,233
LI_HG432_0/1	75,30	36,18	9,901	6,469	9,856
WI_HG432_0/1	76,61	36,04	9,785	6,391	9,374
LI_HG432_HM62	72,39	38,25	10,466	7,076	10,501
WI_HG432_HM62	73,95	38,02	10,324	7,014	10,019
L_0/1_0/1	80,82	31,91	8,732	6,228	9,076
W_0/1_0/1	82,07	31,57	8,572	6,194	8,554

Oszacowane modele hybrydowe oznaczone będą symbolami literowo-liczbowymi. Pierwsza litera oznacza postać analityczną – liniową (L) lub wykładniczą (W). Druga litera odnosi się do kryterium wyboru najlepszego równania minimalizującego błędy prognoz: I – interpolacyjnych ( $MAPE_I$ ), E – ekstrapolacyjnych ( $MAPE_E$ ). Symbolami HG i HM oznaczono

odpowiednio modele hierarchiczne o cyklu dobowym i rocznym; liczby po nich występujące są kolejnymi podzielnikami długości cyklu wahań. W pojedynczych modelach hybrydowych występują symbole:  $\_0/1$  lub  $0/1\_$  oznaczające, że dany rodzaj wahań opisywany jest za pomocą zmiennych zero-jedynkowych. W ostatniej kolumnie tab. 1 podane zostały odpowiadające im oceny błędów prognoz ekstrapolacyjnych ( $MAPE_{LE}$ ). W ostatnich dwóch wierszach te same wielkości podano dla liniowej i wykładniczej postaci modeli klasycznych oznaczonych jako:  $L\_0/1$ ,  $0/1$  lub  $W\_0/1$ ,  $0/1$ .

Z informacji zawartych w kolumnie drugiej, dotyczących kształtowania się współczynników determinacji ( $R^2$ ), wynika, że wśród modeli klasycznych wyższą ocenę, wynoszącą 82,07%, otrzymano dla modelu wykładniczego ( $WI\_0/1\_0/1$ ). Z modeli hybrydowych, w których przynajmniej jeden rodzaj wahań opisywany był za pomocą modeli hierarchicznych, najlepszy okazał się model w postaci wykładniczej  $WI\_0/1\_HM62$ . W modelu tym wahania o cyklu rocznym zostały opisane za pomocą modelu hierarchicznego o podzielnikach 6 i 2 miesiące. Uzyskana dla tego modelu ocena współczynnika determinacji, wynosząca 79,42%, jest o 2,85 punktu procentowego (p.p.) niższa od oceny lepszego z modeli klasycznych.

Spośród modeli hybrydowych opisujących dwa rodzaje wahań lepszy okazał się model wykładniczy  $WI\_HG122\_HM62$  ( $R^2 = 73,95\%$ ). W przypadku modelu  $LI\_HG432\_HM62$  ocena jest niższa o ok. 1,5 p.p. Oceny te dotyczące pozostałych modeli z jednym rodzajem wahań, opisywanych za pomocą modeli hierarchicznych, mieszczą się w przedziale od 75,30% w przypadku modelu  $LI\_HG432\_0/1$  do 77,90% w przypadku modelu  $LI\_0/1\_HM62$ . Zwraca uwagę fakt, że składowe hierarchiczne z modeli pojedynczych (o cyklu dobowym lub rocznym) występują w modelach z dwoma ich rodzajami.

Oceny odchyłeń standardowych dotyczące modeli hybrydowych zawarte są w przedziale od 33,82 GWh ( $WI\_HG122\_0/1$ ) do 38,25 GWh ( $LI\_HG432\_HM62$ ). Ocena minimalna jest o 0,58 GWh wyższa od otrzymanej dla wykładniczej postaci modelu klasycznego. Natomiast oceny współczynników zmienności losowej dla modeli hybrydowych mieszczą się w przedziale od 9,184% do 10,466%; otrzymano je dla modeli o minimalnych i maksymalnych ocenach współczynników determinacji. Porównanie ocen parametrów opisujących własności predyktywne równań w postaci liniowej i wykładniczej dla tego samego rodzaju modeli wskazuje, że korzystniejsze wartości przyjmują one w przypadku modeli wykładniczych. Spostrzeżenie to odnosi się także do modeli klasycznych. Oceny błędów prognoz interpolacyjnych ( $MAPE_I$ ), otrzymane dla modeli hybrydowych, mieszczą się w przedziale od 6,391% w przypadku modelu  $WI\_HG432\_0/1$  do 7,076% w przypadku predyktora opartego na modelu  $LI\_HG432\_HM62$ . Ocena minimalna jest o około 0,2 p.p. wyższa od oceny minimalnej otrzymanej dla wykładniczej postaci modelu klasycznego ( $W\_0/1\_0/1$ ). Niższe oceny błędów prognoz interpolacyjnych, podobnie jak charakterystyk własności prognostycznych, otrzymano dla predyktorów wykładniczych. Błędy prognoz ekstrapolacyjnych ( $MAPE_{LE}$ ) dla predyktorów hybrydowych przyjmują wartości 9,233% (w przypadku modelu  $WI\_0/1\_HM62$ ) i 10,501% (w przypadku modelu  $LI\_HG432\_HM62$ ). Ocena minimalna jest o ok. 0,7 p.p. wyższa od otrzymanej dla modelu klasycznego  $W\_0/1\_0/1$ .

Z informacji zawartych w tab. 1 wynika ponadto, że wśród modeli hybrydowych najniższe oceny błędów prognoz interpolacyjnych otrzymano dla modeli, w których za pomocą modeli hierarchicznych opisywane były wahania o cyklu 24-godzinnym, a najwyższe dla predyktorów hierarchicznych z obydwoma rodzajami wahań.



W tabeli 3 zestawiono modele hybrydowe charakteryzujące się minimalnymi ocenami błędów prognoz ekstrapolacyjnych ( $MAPE_E$ ). Punktem odniesienia dla tych modeli będą oceny błędów tego rodzaju prognoz otrzymanych na podstawie modeli klasycznych ze zmiennymi zero-jedynkowymi (L\_0/1\_0/1 i W\_0/1\_0/1). Z porównania struktury modeli hybrydowych wynika, że jedynie w dwóch na sześć przypadków modele o minimalnych ocenach błędów prognoz inter- i ekstrapolacyjnych są takie same. Odnosi się to do równań, w których jeden rodzaj wahań opisany jest za pomocą modeli hierarchicznych i dotyczy modeli: LE\_0/1\_H62 oraz WE\_HG122\_0/1. W pozostałych czterech przypadkach są to modele hierarchiczne o różnej strukturze.

Tabela 3. Oceny parametrów struktury stochastycznej i błędy prognoz modeli hybrydowych dla kryterium  $MAPE_E$

Model	R <sup>2</sup> [%]	S <sub>e</sub> [GWh]	V <sub>Se</sub> [%]	MAPE <sub>I</sub> [%]	MAPE <sub>LE</sub> [%]
LE_0/1_HM62	77,90	34,24	9,369	9,749	6,839
WE_0/1_HM62	79,42	33,82	9,184	9,233	6,821
LE_HG12,2_0/1	79,85	32,70	8,946	9,259	6,566
WE_HG12,2_0/1	81,06	32,44	8,808	8,775	6,658
LE_HG12,2_HM62	76,93	34,97	9,569	9,915	7,151
WE_HG12,2_HM62	78,40	34,63	9,403	9,433	7,262

Najwyższe i najniższe oceny współczynników determinacji, wynoszące odpowiednio 81,06% i 76,93%, otrzymano dla modeli WE\_HG12,2\_0/1 i LE\_HG12,2\_HM62. Ocena maksymalna jest o ok. 1,8 p.p. wyższa od takiej samej oceny otrzymanej dla kryterium  $MAPE_I$  i jednocześnie o 1,01 p.p. niższa od oceny wykładniczej postaci modelu klasycznego (W\_0/1\_0/1). Oceny odchyłeń standardowych dla modeli hybrydowych kształtują się w przedziale od 32,44 GWh w przypadku modelu LE\_HG12,2\_0/1 do 34,97 GWh w przypadku modelu WE\_HG12,2\_HM62. Dla tych samych modeli otrzymano minimalną i maksymalną ocenę współczynników zmienności losowej – odpowiednio 8,808% i 9,569%.

Dla liniowej postaci modelu, w których sezonowość 24-godzinna opisywana była za pomocą modelu hierarchicznego LE\_HG12,2\_01, ocena tego współczynnika była tylko o około 0,14 p.p. wyższa od oceny minimalnej.

Oceny błędów prognoz ekstrapolacyjnych ( $MAPE_E$ ) nie przekroczyły 10%. Najniższą ich ocenę, wynoszącą 8,775%, otrzymano dla predyktora opartego na modelu WE\_HG12,2\_01. Oceną najwyższą (9,915%) charakteryzował się model liniowy z dwoma składowymi hierarchicznymi: LE\_HG12,2\_HM62. Dla wykładniczej postaci tego modelu otrzymano ocenę niższą prawie o 0,5 p.p. Dla dwóch postaci analitycznych modeli hybrydowych we wszystkich przypadkach oceny niższe otrzymano dla postaci wykładniczej.

Natomiast błędy prognoz interpolacyjnych ( $MAPE_{E,I}$ ) zawarte były w przedziale od 6,566% dla predyktora LE\_HG12,2\_0/1 do 7,262% dla predyktora WE\_HG12,2\_HM62.

## WYNIKI MODELOWANIA I PROGNOZOWANIA PRZYCZYNOWYCH MODELI HYBRYDOWYCH

Wynikiem modelowania i prognozowania dla równań hybrydowych o charakterze przyczynowo-skutkowymi różnią się od wyników równań hybrydowych szeregu czasowego tym, że zawierają dodatkowo opóźnioną o 24 godziny zmienną endogeniczną oraz zmienne zero-

-jedynekowe oznaczające dni świąteczne i Wielką Sobotę. Wprowadzenie opóźnionej zmiennej prognozowanej spowodowane było tym, że z takim wyprzedzeniem składane są zapotrzebowania na moc przez dystrybutorów energii.

W tabeli 4 zestawione zostały, w takim samym układzie jak poprzednio, modele hybrydowe charakteryzujące się minimalnymi ocenami błędów prognoz inter- i ekstrapolacyjnych. Litera P na początku symbolu oznacza model typu przyczynowego. Punktem odniesienia będą także oceny błędów otrzymane na podstawie predyktorów klasycznych – dla prognoz interpolacyjnych predyktora w postaci liniowej, a dla prognoz ekstrapolacyjnych predyktora w postaci wykładniczej.

Tabela 4. Oceny parametrów struktury stochastycznej i błędy prognoz przyczynowych modeli hybrydowych dla kryterium  $MAPE_t$

Model	$R^2$ [%]	Se [GWh]	$V_{Se}$ [%]	$MAPE_t$ [%]	$MAPE_{LE}$ [%]
PLI_0/1_HM62	84,04	29,10	7,963	5,536	6,944
PWI_0/1_HM62	84,61	29,24	7,940	5,616	6,705
PLI_HG432_0/1	83,44	29,63	8,107	5,406	6,840
PWI_HG432_0/1	83,79	30,00	8,145	5,431	6,622
PLI_HG432_HM62	82,93	30,08	8,231	5,543	6,880
PWI_HG432_HM62	83,30	30,21	8,276	5,591	6,672
PL_0/1_0/1	84,85	28,36	7,760	5,374	6,906
PW_0/1_0/1	85,41	28,48	7,734	5,435	6,647

Oceny współczynników determinacji dla modeli hybrydowych, charakteryzujących się najniższymi ocenami błędów prognoz dla kryterium  $MAPE_t$ , zawarte były w przedziale od 82,93% w przypadku modelu PLI\_HG432\_HM62 do 84,61% w przypadku modelu PWI\_0/1\_HM62. Z dwóch postaci analitycznych oceny współczynników determinacji przyjmowały nieco wyższe wartości dla postaci wykładniczej – różnice te nie przekraczały 0,63 p.p. Lepszy z modeli klasycznych – model wykładniczy – charakteryzował się oceną około 0,8 p.p. wyższą od oceny maksymalnej dla modeli hybrydowych.

Modele hybrydowe z dwoma rodzajami modeli hierarchicznych są „sumami” modeli z jednym rodzajem. Oceny współczynników  $R^2$  otrzymane dla tych równań były tylko nieznacznie niższe od oceny współczynników dotyczących modeli, w których wahania o cyklu 24-godzinnym były opisywane za pomocą tego rodzaju modeli – różnice te nie przekraczają 0,35 p.p. W przypadku modeli hybrydowych z wahaniami o cyklu rocznym, opisywanymi za pomocą modeli hierarchicznych, różnice te są wyższe – kształtują się w granicach 2 p.p.

Oceny odchyłeń standardowych składników losowych zawarte są w przedziale od 29,10 GWh dla modelu PLI\_0/1\_HM62 do 30,21 GWh dla modelu PWI\_HG432\_HM62. Ich oceny dotyczące modeli, w których wahania o cyklu rocznym opisywane są za pomocą modeli hierarchicznych, są nieznacznie niższe niż dotyczące modeli z dwoma rodzajami wahań – różnice te nie przekraczają 0,45 GWh. Natomiast dla modeli klasycznych są o ok. 1,7–1,8 GWh niższe.

Oceny współczynników zmienności losowej dla modeli hybrydowych zawarte są w przedziale od 7,940% dla modelu PWI\_0/1\_HM62 do 8,276% dla modelu PWI\_HG38\_HM43. Pomędzy ocenami omawianych współczynników zachodzi podobna relacja jak w przypadku ocen odchyłeń standardowych składników losowych. Ich oceny dla modeli klasycznych (PL\_0/1\_0/1 i PL\_0/1\_0/1) są o ok. 0,4 p.p. niższe od ocen otrzymanych dla najlepszych modeli hybrydowych.

Z porównania ocen błędów prognoz interpolacyjnych w postaci liniowej i wykładniczej, o strukturze hierarchicznej tego samego rodzaju, wynika, że są one nieco wyższe dla postaci wykładniczej. Największą różnicę, wynoszącą 0,080 p.p., otrzymano dla modeli, w których wahania o cyklu rocznym opisywane są za pomocą modeli hierarchicznych.

Ocenę minimalną błędu prognoz interpolacyjnych dla modeli hybrydowych, wynoszącą 5,406%, otrzymano dla predyktora opartego na modelu PLI\_HG432\_0/1, przy czym jest ona tylko o 0,033 p.p. wyższa od otrzymanej dla liniowej postaci modelu klasycznego (L\_0/1\_0/1). Oceny te dla predyktorów, w których wahania o cyklu 24-godzinnym są opisywane za pomocą modelu hierarchicznego, oraz dla modeli z dwoma rodzajami wahań są bardzo zbliżone – różniły się one nie więcej niż o 0,044 p.p.

Błędy prognoz ekstrapolacyjnych ( $MAPE_{LE}$ ) charakteryzowały się także stosunkowo niewielkim zróżnicowaniem. W przypadku modeli hybrydowych przyjęły one wartości od 6,622% dla modelu PWI\_HG432\_0/1 do 6,944% dla modelu PLI\_01\_HM62. Minimalna ocena tego błędu jest niższa o 0,025 p.p. od otrzymanej dla wykładniczej postaci modelu klasycznego (PW\_0/1\_0/1).

W tabeli 5 zestawiono modele hybrydowe charakteryzujące się minimalnymi ocenami błędów prognoz ekstrapolacyjnych ( $MAPE_E$ ). Punktem odniesienia w porównaniach własności predykcyjnych oraz ocen obu rodzajów prognoz będzie model klasyczny typu przyczynowego ze zmiennymi zero-jedynkowymi, o niższej ocenie błędu tego rodzaju (PW\_0/1\_0/1).

Tabela 5. Oceny parametrów struktury stochastycznej i błędy prognoz przyczynowych modeli hybrydowych dla kryterium  $MAPE_E$

Model	$R^2$ [%]	Se [GWh]	$V_{Se}$ [%]	$MAPE_I$ [%]	$MAPE_{LE}$ [%]
PLE_0/1_HM232	80,37	32,28	8,832	6,835	6,112
PWE_0/1_HM43	83,55	30,24	8,211	6,675	5,772
PLE_HG2322_0/1	82,11	30,80	8,428	6,677	5,599
PWE_HG432_0/1	83,79	30,00	8,145	6,622	5,431
PLE_HG2322_HM43	81,37	31,43	8,599	6,643	5,719
PWE_HG2322_HM322	81,18	32,32	8,775	6,608	5,965

Oceny współczynnika determinacji najniższą i najwyższą, które wynoszą odpowiednio 80,37% i 83,79%, otrzymano dla modeli o postaci PLE\_0/1\_HM232 i PWE\_HG432\_0/1. Ocena maksymalna jest o 0,82 p.p. niższa od oceny maksymalnej otrzymanej dla kryterium  $MAPE_I$  i jednocześnie o 1,62 p.p. niższa od otrzymanej dla wykładniczej postaci modelu klasycznego. Oceny odchyłeń standardowych składników losowych dla modeli hybrydowych kształtują się w przedziale od 30,00 GWh w przypadku modelu PWE\_HG432\_0/1 do 32,32 GWh w przypadku modelu PWE\_HG2322\_HM322. Minimalną ocenę współczynnika zmienności losowej (8,145%) otrzymano dla modelu o minimalnej ocenie odchylenia standardowego, a maksymalną (8,832%) – dla modelu PLE\_0/1\_HM232.

Oceny błędów prognoz ekstrapolacyjnych ( $MAPE_E$ ) charakteryzowały się niewielkim zróżnicowaniem – zawarte były w przedziale od 6,608% dla modelu PWE\_HG2322\_HM322 do 6,835% dla modelu o najniższej ocenie współczynnika determinacji – PLE\_0/1\_HM232. Minimalną ocenę błędów prognoz interpolacyjnych ( $MAPE_{E,I}$ ), wynoszącą 5,431%, otrzymano

dla predyktora PWE\_HG432\_0/1, a maksymalną (5,965%) – dla wykładniczej postaci hybrydowego predyktora podwójnego. Dla dwóch postaci analitycznych modeli hybrydowych we wszystkich przypadkach oceny niższe otrzymano dla postaci wykładniczej.

### PORÓWNIANIA STRUKTUR, WŁASNOŚCI PREDYKTYWNYCH I DOKŁADNOŚCI PROGNOZ DLA HYBRYDOWYCH MODELI KLASYCZNYCH SZEREGU CZASOWEGO I MODELI PRZYCZYNOWYCH

Różnice własności predyktywnych oraz różnice ocen błędów prognoz inter- i ekstrapolacyjnych, otrzymane na podstawie modeli hybrydowych ze zmienną opóźnioną i dniami świątecznymi oraz bez tych zmiennych, będą wyrażone, poza  $\Delta S_e$ , w punktach procentowych. Informacje, o których napisano wyżej dla kryterium  $MAPE_t$ , zamieszczone zostały w tab. 6.

Tabela 6. Różnice parametrów:  $R^2$ ,  $Se$ ,  $V_{Se}$  oraz  $MAPE_t$  i  $MAPE_{t,E}$  hybrydowych modeli przyczynowych i szeregu czasowego

Hybrydowe modele przyczynowe	Hybrydowe modele szeregu czasowego	$\Delta R^2$ [p.p.]	$\Delta Se$ [GWh]	$\Delta V_{Se}$ [p.p.]	$\Delta MAPE_t$ [p.p.]	$\Delta MAPE_{t,E}$ [p.p.]
PLI_0/1_HM62	LI_0/1_HM62	6,14	-5,14	-1,406	-1,303	-2,804
PWI_0/1_HM62	WI_0/1_HM62	5,20	-4,58	-1,244	-1,205	-2,528
PLI_HG432_0/1	LI_HG432_0/1	8,14	-6,56	-1,794	-1,063	-3,017
PWI_HG432_0/1	WI_HG432_0/1	7,18	-6,04	-1,640	-0,960	-2,752
PLI_HG432_HM62	LI_HG432_HM62	10,53	-8,17	-2,235	-1,534	-3,621
PWI_HG432_HM62	WI_HG432_HM62	9,35	-7,81	-2,048	-1,424	-3,346
PL_0/1_0/1	L_0/1_0/1	4,03	-3,55	-0,971	-0,855	-2,170
PW_0/1_0/1	W_0/1_0/1	3,34	-3,09	-0,838	-0,759	-1,907

Z porównania struktur przyczynowych modeli hybrydowych i hybrydowych modeli szeregu czasowego wynika, że we wszystkich sześciu przypadkach otrzymano równania o takiej samej strukturze hierarchicznej.

W kolumnie trzeciej i następnych podane zostały różnice ocen: współczynników determinacji, odchyłeń standardowych, współczynników zmienności losowej oraz błędów prognoz  $MAPE_t$  i  $MAPE_{t,E}$ . W przypadku współczynników determinacji modele przyczynowe charakteryzują się ich wyższymi ocenami. Największe różnice, wynoszące 10,53 i 9,35 p.p., otrzymano dla modeli hybrydowych z podwójną strukturą hierarchiczną, a następnie dla modeli, w których wahania o cyklu 24-godzinnym są opisane za pomocą modeli hierarchicznych. Najmniejsze różnice, wynoszące: 3,34 i 4,03 p.p., otrzymano dla modeli klasycznych – odpowiednio o postaci wykładniczej i liniowej. Ponadto ujemne znaki różnic parametrów struktury stochastycznej ( $Se$  i  $V_{Se}$ ) informują o lepszych własnościach predyktywnych równań przyczynowych. Bez-względne wartości różnic odchyłeń standardowych składników losowych dla modeli hybrydowych mieszczą się w przedziale od 4,58 GWh (dla wykładniczej postaci modeli hybrydowych z sezonowością roczną opisywaną za pomocą modeli hierarchicznych) do 8,17 GWh (dla liniowej postaci podwójnych modeli hybrydowych). Dla tych samych równań otrzymano odpowiednio minimalną i maksymalną ocenę współczynników zmienności losowej. Ich wartości bezwzględne wyniosły odpowiednio 1,244 i 2,235 p.p. Dla obu postaci modeli klasycznych otrzymano niższe bezwzględne oceny tych parametrów, w porównaniu z minimalnymi ocenami

otrzymanymi dla modeli hybrydowych Różnice mierników dokładności prognoz  $MAPE_I$  i  $MAPE_{LE}$  przyjęły także ujemne wartości. Oznacza to, że modele przyczynowe charakteryzują się nie tylko lepszymi własnościami predykcyjnymi, lecz także większą dokładnością obu rodzajów prognoz.

Minimalną bezwzględną ocenę różnicy mierników  $MAPE_I$ , wynoszącą 0,960 p.p., otrzymano dla pojedynczego predyktora hybrydowego, w którym za pomocą modelu hierarchicznego opisywane były wahania o cyklu dobowym. Dla modelu o minimalnej ocenie różnicy parametrów  $V_{Se}$  błąd ten był o 0,109 p.p. większy. Świadczy to o tym, że nie zawsze model o najlepszych własnościach daje minimalne błędy prognoz. Różnicę maksymalną (1,534 p.p.) otrzymano dla predyktora o maksymalnej wartości różnicy  $V_{Se}$ . Bezwzględne oceny różnic miernika dokładności prognoz ekstrapolacyjnych  $MAPE_{LE}$  są o ok. 1,3–2,5 p.p. wyższe od błędów prognoz interpolacyjnych. Bezwzględne wartości minimalną i maksymalną różnic, wynoszące odpowiednio 2,528 i 3,621 p.p., otrzymano dla predyktorów o minimalnej i maksymalnej różnicy parametru  $V_{Se}$ .

Różnice bezwzględnych wartości mierników dokładności obu rodzajów prognoz, otrzymanych na podstawie predyktorów klasycznych, przyjęły mniejsze wartości niż w przypadku predyktorów hybrydowych.

W tabeli 7 zestawione zostały różnice wartości tych samych parametrów oraz błędów prognoz dla modeli hybrydowych charakteryzujących się minimalnymi ocenami błędów prognoz ekstrapolacyjnych ( $MAPE_E$ ). Z porównania struktur pojedynczych i podwójnych modeli hybrydowych przyczynowo-opisowych i szeregu czasowego, opisujących wahania tego rodzaju, wynika, że we wszystkich sześciu przypadkach różnią się one podzielnikami długości cykli wahań. Przykładowo w podanym w pierwszym wierszu modelu przyczynowym podzielniki cyklu rocznego wynoszą: 2, 3, 2, a w modelu szeregu czasowego – 6, 2. Jest to sytuacja odmienna od otrzymanej dla kryterium  $MAPE_I$ , w przypadku którego struktury te były identyczne. Z informacji zawartych w trzecim wierszu wynika, że różnice ocen współczynników determinacji dla tego samego rodzaju modeli hybrydowych zawarte są w przedziale od 2,26 p.p. (w przypadku modelu liniowego o hierarchicznych składowych opisujących wahania o cyklu dobowym) do 4,43 p.p. (w przypadku liniowego modelu z dwiema składowymi hierarchicznymi).

Tabela 7. Różnice parametrów:  $R^2$ ,  $Se$ ,  $V_{Se}$  oraz  $MAPE_E$  i  $MAPE_{E,I}$  hybrydowych modeli przyczynowych i szeregu czasowego

Hybrydowe modele przyczynowe	Hybrydowe modele szeregu czasowego	$\Delta R^2$ [p.p.]	$\Delta Se$ [GWh]	$\Delta V_{Se}$ [p.p.]	$\Delta MAPE_E$ [p.p.]	$\Delta MAPE_{E,I}$ [p.p.]
PLE_0/1_HM232	LE_0/1_HM62	2,46	-1,96	-0,537	-2,914	-0,727
PWE_0/1_HM43	WE_0/1_HM62	4,13	-3,58	-0,973	-2,558	-1,049
PLE_HG2322_0/1	LE_HG12,2_0/1	2,26	-1,89	-0,518	-2,581	-0,968
PWE_HG432_0/1	WE_HG12,2_0/1	2,74	-2,44	-0,662	-2,153	-1,227
PLE_HG2322_HM43	LE_HG12,2_HM62	4,43	-3,55	-0,970	-3,272	-1,432
PWE_HG2322_HM322	WE_HG12,2_HM62	2,78	-2,31	-0,628	-2,825	-1,297

Dla kryterium  $MAPE_E$ , podobnie jak dla  $MAPE_I$ , otrzymano ujemne znaki różnic parametrów struktury stochastycznej ( $Se$  i  $V_{Se}$ ). Oznacza to, że modele przyczynowe mają lepsze własności predykcyjne. Najniższą bezwzględną ocenę różnicy parametru  $Se$  (1,89 GWh) otrzymano dla

pojedynczego predyktora liniowego z wahaniami o cyklu dobowym opisywanymi za pomocą modelu hierarchicznego. Natomiast różnicę największą (3,58 GWh) otrzymano dla predyktora wykładniczego, w którym za pomocą modelu hierarchicznego opisywane były wahania o cyklu rocznym. Dla wymienionych wyżej predyktorów otrzymano taką samą relację współczynników zmienności losowej. Należy przy tym zauważyć, że bezwzględne różnice tych parametrów są znacznie mniejsze niż w przypadku modeli otrzymanych na podstawie kryterium  $MAPE_I$ .

Różnice mierników dokładności prognoz  $MAPE_E$  i  $MAPE_{E,I}$  przyjęły także ujemne wartości. Oznacza to, że modele przyczynowe, podobnie jak w przypadku poprzedniego kryterium, charakteryzują się zarówno lepszymi własnościami predyktywnymi, jak i większą dokładnością obu rodzajów prognoz.

Minimalną bezwzględną ocenę różnicy mierników  $MAPE_E$ , wynoszącą 2,153 p.p., otrzymano dla pojedynczego wykładniczego predyktora hybrydowego, w którym za pomocą modelu hierarchicznego opisywane były wahania o cyklu dobowym, natomiast różnicę maksymalną wynoszącą 3,272 p.p. otrzymano dla liniowej postaci podwójnego predyktora hybrydowego. W obu przypadkach były to inne predyktory od tych, które charakteryzowały się minimalną bądź maksymalną wartością różnicy ocen parametrów  $V_{se}$ .

Ze wstępnego porównania ocen średnich względnych błędów prognoz, zarówno ekstrapolacyjnych ( $MAPE_E$ ), jak i interpolacyjnych ( $MAPE_{E,I}$ ), wynika, że dla tych samych postaci modeli hybrydowych we wszystkich przypadkach przyjęły one mniejsze wartości niż dla predyktorów szeregu czasowego. W przypadku prognoz ekstrapolacyjnych różnice ocen mierników  $MAPE_E$  zawierały się w przedziale od 2,153 p.p. (w przypadku postaci wykładniczej pojedynczego modelu hybrydowego z wahaniami o cyklu dobowym) do 3,272 p.p. (w przypadku liniowej postaci podwójnego modelu hybrydowego). W przypadku modeli klasycznych ze zmiennymi zero-jedynkowymi w postaci liniowej (L\_0/1\_0/1) i wykładniczej (W\_0/1\_0/1) otrzymano różnice ponaddwukrotnie większe – wynosiły one odpowiednio 2,91 i 2,56 p.p.

Dla prognoz interpolacyjnych różnice ocen mierników  $MAPE_I$  zawarte są w przedziale od 0,96 p.p. (w przypadku postaci wykładniczej pojedynczego modelu hybrydowego z wahaniami o cyklu dobowym) do 1,53 p.p. (w przypadku liniowej postaci podwójnego modelu hybrydowego). Dla modeli klasycznych ze zmiennymi zero-jedynkowymi w postaci liniowej (L\_0/1\_0/1) i w postaci wykładniczej (W\_0/1\_0/1) oceny błędów różniły się odpowiednio o 0,85 i 0,76 p.p.

## WNIOSKI

1. Z uwagi na różnice struktur najlepszych modeli hybrydowych za właściwe należy uznać przyjęcie dwóch kryteriów wyboru polegających na minimalizacji błędów prognoz interpolacyjnych i ekstrapolacyjnych.
2. Modele hybrydowe typu przyczynowego charakteryzowały się lepszymi własnościami prognostycznymi i niższymi ocenami błędów prognoz inter- i ekstrapolacyjnych, przy czym różnice w dokładności były większe w przypadku prognoz ekstrapolacyjnych.
3. W trzech przypadkach na cztery przypadki (po dwa na każde kryterium) lepsze okazały się pojedyncze modele hybrydowe z modelami hierarchicznymi opisującymi wahania o cyklu 24-godzinnym.
4. Także w trzech przypadkach na cztery przypadki niższe oceny błędów prognoz otrzymano dla postaci wykładniczej predyktorów hybrydowych.

5. Błędy prognozy otrzymanych na podstawie najlepszych predyktorów hybrydowych szeregu czasowego, charakteryzujących się najniższymi ocenami średnich błędów prognoz interpolacyjnych, były o ok. 0,2 p.p. wyższe od odpowiednich błędów prognoz otrzymanych na podstawie predyktorów klasycznych.
6. W przypadku modeli przyczynowo-opisowych dla prognoz interpolacyjnych błąd był wyższy o 0,07 p.p., a dla prognoz ekstrapolacyjnych – niższy o 0,39 p.p.

## PIŚMIENNICTWO

- Junger W.L., de Leon A.P.** 2015. Imputation of missing data in time series for air pollutants. *Atmosph. Environ.* 102, 96–104.
- Kufel T.** 2010. Ekonometryczna analiza cykliczności procesów gospodarczych o wysokiej częstotliwości obserwacji. Toruń, Wydaw. Nauk. UMK.
- Little R.J., Rubin D.B.** 1987. *Statistical analysis with missing data*. Chicago, Am. Market. Assoc.
- Raudenbush S.W., Bryk A.S.** 2010. *Hierarchical linear models. Applications and data analysis methods*. 2 nd. ed. London, Sage Publications.
- Szmuksta-Zawadzka M., Zawadzki J.** 2002. Hierarchiczne modele szeregów czasowych z wahaniami sezonowymi. Budowa. Estymacja. Prognozowanie. Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych. Kraków, Wydaw. AE, 193–204.
- Szmuksta-Zawadzka M., Zawadzki J.** 2004. On hierarchic models for decade data with seasonal fluctuations. *Dynamic Econom. Models* 6, 65–71.
- Szmuksta-Zawadzka M., Zawadzki J.** 2011. Modelowanie i prognozowanie zapotrzebowania na energię elektryczną w mikroskali. *Pr. Nauk. UE Wroc.* 165, 152–161.
- Szmuksta-Zawadzka M., Zawadzki J.** 2014. Hierarchiczne modele w prognozowaniu zmiennych o wysokiej częstotliwości obserwowania w warunkach braku pełnej informacji. *Ekonometria* 4(46), 72–84.

**Streszczenie.** W pracy przedstawiono wykorzystanie pojedynczych i podwójnych addytywnych i multiplikatywnych modeli hybrydowych w prognozowaniu brakujących danych w szeregach czasowych, z wahaniami cyklicznymi o wysokiej częstotliwości obserwowania dla luk niesystematycznych. Wahania złożone w cyklach: rocznym, tygodniowym i dobowym nakładają się na trend w sposób addytywny lub multiplikatywny. Wahania o parzystej długości cykli (12-miesięczne i 24-godzinne) opisywano za pomocą regularnych modeli hierarchicznych. Modelowaniu i prognozowaniu poddano zapotrzebowanie na energię elektryczną w okresach godzinnych.

