

Maciej Oesterreich

WYKORZYSTANIE MODELU HOLTA-WINTERSA ORAZ METOD BOOTSTRAPOWYCH W PROGNOZOWANIU NA PODSTAWIE ZMIENNYCH EKONOMICZNYCH Z WAHANIAMI SEZONOWYMI

APPLICATION OF HOLT-WINTERS MODEL AND BOOTSTRAP METHODS IN FORECASTING ECONOMIC VARIABLES WITH SEASONAL FLUCTUATIONS

Katedra Zastosowań Matematyki w Ekonomii, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
ul. Klemensa Janickiego 31, 71-270 Szczecin, e-mail: moesterreich@zut.edu.pl

Summary. This paper presents bootstrap algorithm, that helps increase the accuracy of the forecasts created with additive Holt-Winters model for seasonal time series. In the process of forecasting was used the overlapping blocks bootstrap method (Kunsh method) with various blocks lengths. In calculations was used "forecast" and "tseries" modules from statistical program R.

Słowa kluczowe: bootstrap, Holt-Winters, prognozy, szeregi czasowe z sezonowością.
Key words: bootstrap, forecast, Holt-Winters, seasonal time series.

WSTĘP

Modele wyrównywania wykładniczego należą do nieklasycznych modeli szeregu czasowego najczęściej wykorzystywanych w prognozowaniu zmiennych z wahaniami sezonowymi. Związane jest to m.in. z ich elastycznością, tzn. z szybką reakcją na zmiany trendu oraz wahań okresowych. Dodatkową zaletą jest dostępność programów obliczeniowych, pozwalających na szybkie wyznaczenie optymalnych wartości stałych wyrównania i budowę prognoz. Niekiedy jednak dokładność prognoz dotyczących tych zmiennych jest niewystarczająca. Zwiększenie ich efektywności jest możliwe dzięki wykorzystaniu technik i metod symulacyjnych, w tym metod bootstrapowych.

W niniejszej pracy przedstawiono zastosowanie tej metody do szacowania parametrów sezonowego modelu Holta-Wintersa oraz do budowy prognoz. W literaturze przedmiotu nie znaleziono opisu tego rodzaju podejścia.

W analizie wykorzystano addytywny model Holta-Wintersa w postaci:

$$m_t = \alpha(Y_t - C_{t-m}) + (1 - \alpha)(m_{t-1} + \delta_{t-1}) \quad (1)$$

$$\delta_t = \beta(m_t - m_{t-1}) + (1 - \beta)\delta_{t-1} \quad (2)$$

$$C_t = \gamma(Y_t - m_t) + (1 - \gamma)C_{t-m} \quad (3)$$

gdzie:

m_t – operator rzędu pierwszego,

δ_t – współczynnik kierunkowy trendu,

C_t – składnik sezonowy

Predyktor, oparty na modelu opisanym równaniami (1)–(3), przyjmuje postać:

$$\hat{Y} = m_t + \delta_t h + C_{t-m+h} \quad (4)$$

gdzie:

m – długość cyklu wahań sezonowych.

h – horyzont prognozy.

Procedura budowy próby bootstrapowej nie jest skomplikowana. Zakładamy, że badamy próbę $X = (X_1, X_2, X_3, \dots, X_n)$, której realizacjami są $x = (x_1, x_2, x_3, \dots, x_n)$, pochodzącą z populacji X o nieznanym lub częściowo znanym rozkładzie F . Próba bootstrapowa powstaje poprzez zastosowanie schematu losowania ze zwracaniem z próby pierwotnej X (Efron 1979). W przypadku badania szeregów czasowych wykorzystuje się różne modyfikacje pierwotnej metody bootstrapowej. Jedną z nich jest bootstrap blokowy, polegający na tym, że zamiast pojedynczych elementów losuje się bloki X_{t+1}, \dots, X_{T+1} . Ma to na celu zachowanie struktury szeregu czasowego w poszczególnych blokach (Bühlmann 2002). W literaturze przedmiotu opisano kilka odmian tej metody, głównie metodę Kunscha z nakładającymi się blokami i metodę Carlsteina z nienakładającymi się blokami (Lahiri 1999). Zaletą pierwszej jest większa liczba bloków, a więc większe prawdopodobieństwo wylosowania niektórych realizacji szeregu czasowego (Kozarski 2004).

ALGORYTM OBLICZENIOWY

Procedura, zmierzająca do wyznaczenia prognoz i zbadania ich empirycznej dokładności, obejmuje 4 etapy; są to:

1. Eliminacja trendu z badanego szeregu poprzez podzielenie wartości rzeczywistych przez wartości teoretyczne, uzyskane z jednego z 4 modeli:

– linowego

$$\hat{Y}_t = \alpha_0 + \alpha_1 t \quad (5)$$

– kwadratowego

$$\hat{Y}_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 \quad (6)$$

– wykładniczego stopnia I

$$\hat{Y}_t = e^{\alpha_0 + \alpha_1 t} \quad (7)$$

– wykładniczego stopnia II

$$\hat{Y}_t = e^{\alpha_0 + \alpha_1 t + \alpha_2 t^2} \quad (8)$$

2. Budowa próby bootstrapowej, na podstawie szeregu oczyszczonego z trendu, metodą blokową (MMB), o zadanej liczbie powtórzeń (w tym wypadku 1000) i wielkości bloku ($l = 3, 4, 6, 12$).

3. Obliczenie dla poszczególnych prób bootstrapowych stałych wyrównania α , β i γ addytywnego modelu Holta-Wintersa, na podstawie modelu (4). Ostateczne parametry, z których budowane były prognozy, obliczono jako przeciętne dla wszystkich replikacji:

$$\hat{\alpha} = \frac{\sum_{r=1}^B \alpha_r}{B}, \quad \hat{\beta} = \frac{\sum_{r=1}^B \beta_r}{B}, \quad \hat{\gamma} = \frac{\sum_{r=1}^B \gamma_r}{B}, \quad r = 1, 2, 3, \dots, B$$

gdzie:

$\alpha_r, \beta_r, \gamma_r$ – parametry modelu Holta-Wintersa dla r -tej próby bootstrapowej,
 B – liczba prób bootstrapowych.

4. Budowa prognoz ekstrapolacyjnych na podstawie parametrów obliczonych w etapie 3, przemożenie ich przez prognozy trendu uzyskane w etapie 1 dla przedziału empirycznej weryfikacji prognoz, a następnie porównanie ich z realizacjami.

PRZYKŁAD EMPIRYCZNY

Zastosowanie omówionego wyżej algorytmu zostanie zilustrowane przykładem empirycznym, dotyczącym prognozowania wybranych wyrobów przemysłowych w Polsce według miesięcy. Modelowaniu i prognozowaniu poddane zostały następujące zmienne:

$Y1$ – produkcja serów podpuszczkowych, w tys. t;

$Y2$ – produkcja piwa, w tys. hl;

$Y3$ – produkcja oleju opałowego, w tys. t;

$Y4$ – produkcja nawozów azotowych, w tys. t;

$Y5$ – produkcja samochodów osobowych, w szt.

Dane zaczerpnięto z Biuletynów Statystycznych Głównego Urzędu Statystycznego z lat 2002–2009. Dane z lat 2002–2008 wykorzystane zostały do oszacowania modeli. Natomiast rok 2009 był okresem empirycznej weryfikacji prognoz.

Do eliminacji trendu z analizowanych zmiennych wykorzystano następujące modele:

– kwadratowy – $Y2$;

– wykładniczy I stopnia – $Y3, Y4$;

– wykładniczy II stopnia – $Y1, Y5$.

Postać trendu spośród 4 wymienionych wyżej wybrano oddzielnie dla każdej ze zmiennych. Kierowano się przy tym stopniem opisu, istotnością parametrów oraz kryterium AIC.

W obliczeniach wykorzystano środowisko statystyczne R . Do obliczenia parametrów modelu Holta-Wintersa wykorzystano procedurę HoltWinters z modułu „stats”, a do prognoz – funkcję forecast z modułu „forecast”. Próby bootstrapowe zostały wygenerowane za pomocą funkcji tsbootstrap z pakietu „tseries”.

W tabeli 1 przedstawiono średnie względne błędy prognoz oszacowanych za pomocą standardowego modelu Holta-Wintersa oraz przedstawionego powyżej algorytmu.

Z informacji zawartych w tab. 1 wynika, że w większości przypadków odnotowano wzrost efektywności prognoz wyznaczonych metodą bootstrapową. Zazwyczaj zmniejszenie średnich względnych błędów prognoz wynosiło 2–3 punkty procentowe. Wzrost efektywności dotyczył różnych długości horyzontu prognozy i różnych wielkości bloków. W pojedynczych przypadkach niższymi ocenami błędów charakteryzowały się prognozy otrzymane na podstawie klasycznego modelu Holta-Wintersa. Dla zmiennych $Y4$ i $Y5$ różnice ocen błędów między prognozami, uzyskanymi za pomocą metod bootstrapowych i standardowego modelu Holta-Wintersa, były większe niż w przypadku zmiennych $Y1$ – $Y3$ i kształtowały się w granicach 4–5 punktów procentowych.

Z tabeli 1 wynika także, że wpływ długości bloków na dokładność prognoz był niewielki. Natomiast długość horyzontu prognoz odgrywała istotną rolę. Oceny błędów prognoz dość często malały wraz z jego wydłużaniem.

Tabela 1. Średnie względne błędy prognoz ekstrapolacyjnych dla wybranych zmiennych [%]

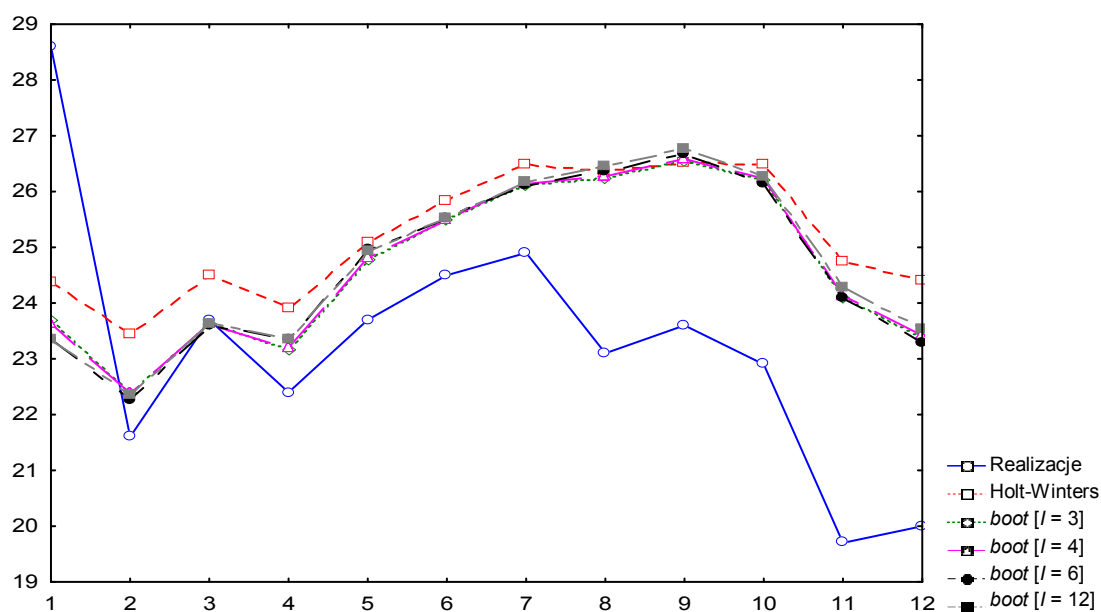
Zmienna	Metoda	$l = 3$				$l = 4$				$l = 6$				$l = 12$			
		$h = 3$	$h = 6$	$h = 9$	$h = 12$	$h = 3$	$h = 6$	$h = 9$	$h = 12$	$h = 3$	$h = 6$	$h = 9$	$h = 12$	$h = 3$	$h = 6$	$h = 9$	$h = 12$
Y1	<i>normal</i>	8,91	7,45	8,62	11,73	8,91	7,45	8,62	11,73	8,91	7,45	8,62	11,73	8,91	7,45	8,62	11,73
	<i>boot</i>	7,07	5,62	7,28	10,06	7,10	5,66	7,28	9,99	7,35	5,93	7,50	10,03	7,36	5,77	7,37	10,03
Y2	<i>normal</i>	19,81	14,67	13,52	14,68	19,81	14,67	13,52	14,68	19,81	14,67	13,52	14,68	19,81	14,67	13,52	14,68
	<i>boot</i>	17,25	13,19	15,30	16,29	17,52	13,04	15,04	16,06	16,51	13,66	16,00	16,88	16,69	12,64	14,65	15,53
Y3	<i>normal</i>	10,41	8,97	11,12	10,32	10,41	8,97	11,12	10,32	10,41	8,97	11,12	10,32	10,41	8,97	11,12	10,32
	<i>boot</i>	7,90	5,36	10,88	10,08	7,64	5,12	10,83	10,06	7,56	5,51	11,40	10,49	7,59	5,35	11,15	10,32
Y4	<i>normal</i>	12,62	12,65	9,43	9,88	12,62	12,65	9,43	9,88	12,62	12,65	9,43	9,88	12,62	12,65	9,43	9,88
	<i>boot</i>	9,37	7,56	5,80	5,91	9,17	7,48	5,62	5,80	8,78	7,16	5,62	5,66	8,74	7,30	6,83	6,43
Y5	<i>normal</i>	13,70	11,57	9,49	10,49	13,70	11,57	9,49	10,49	13,70	11,57	9,49	10,49	13,70	11,57	9,49	10,49
	<i>boot</i>	14,47	8,66	8,53	6,76	15,34	8,89	9,17	7,23	17,31	9,98	10,99	8,88	16,92	12,37	14,22	13,81

Pogrubioną czcionką wyróżniono oceny błędów prognoz, które są niższe od uzyskanych standardową metodą Holta-Wintersa.

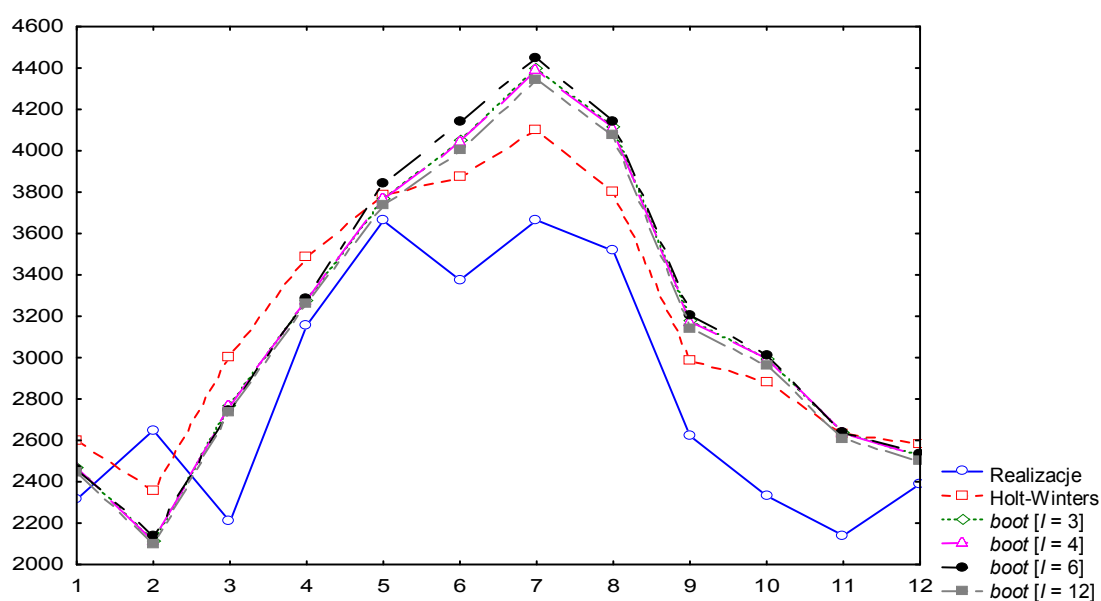
Z zaprezentowanych w tab. 1 wyników oraz z dokładnej analizy zmiennych wynika również, że:

- wzrost dokładności prognoz nastąpił dla zmiennych, których przebieg był zakłócony lub niestały. W przypadku zmiennych, których wahania sezonowe nie były zakłócone, lepsze prognozy uzyskano, stosując standardową metodę Holta-Wintersa;
- proponowana procedura może być stosowana zwłaszcza w przypadku prognoz krótkookresowych o horyzoncie 3–6-miesięcznym.

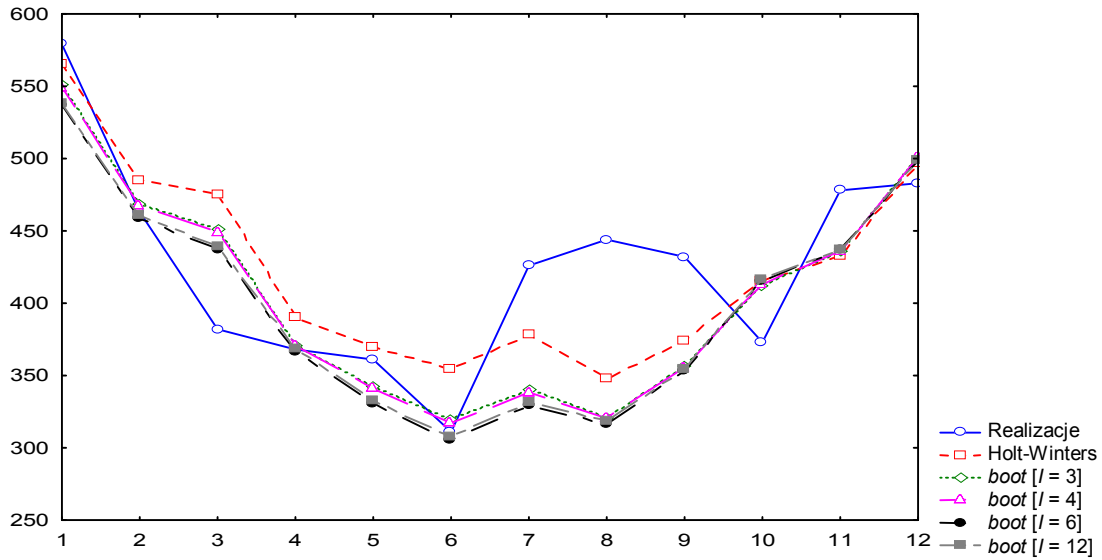
Kształtowanie się prognoz ekstrapolacyjnych dla analizowanych zmiennych przedstawiono na rys. 1–5.



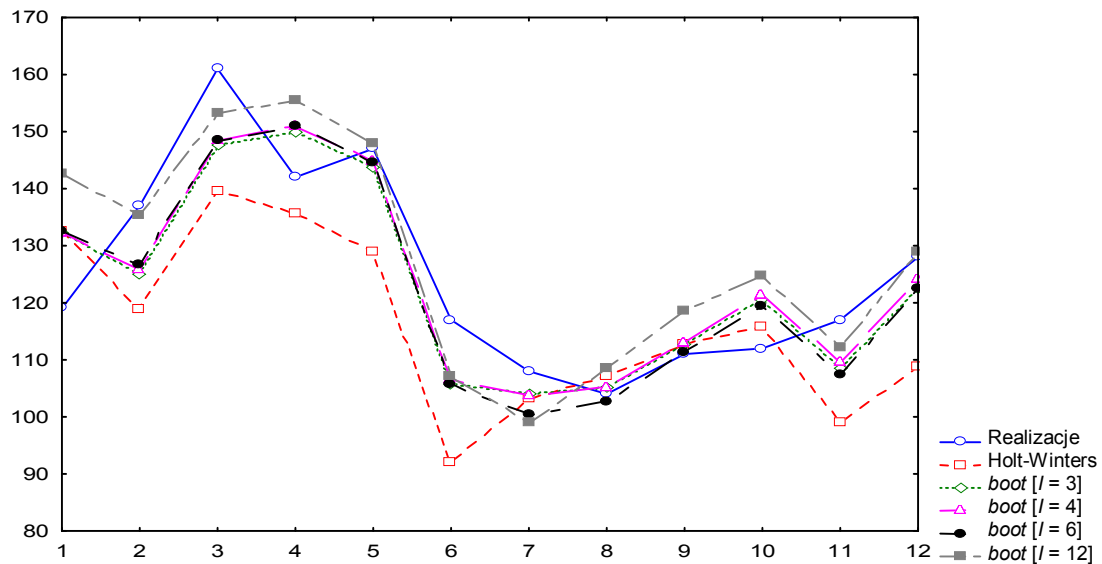
Rys. 1. Prognozy ekstrapolacyjne produkcji serów podpuszczkowych (Y_1)



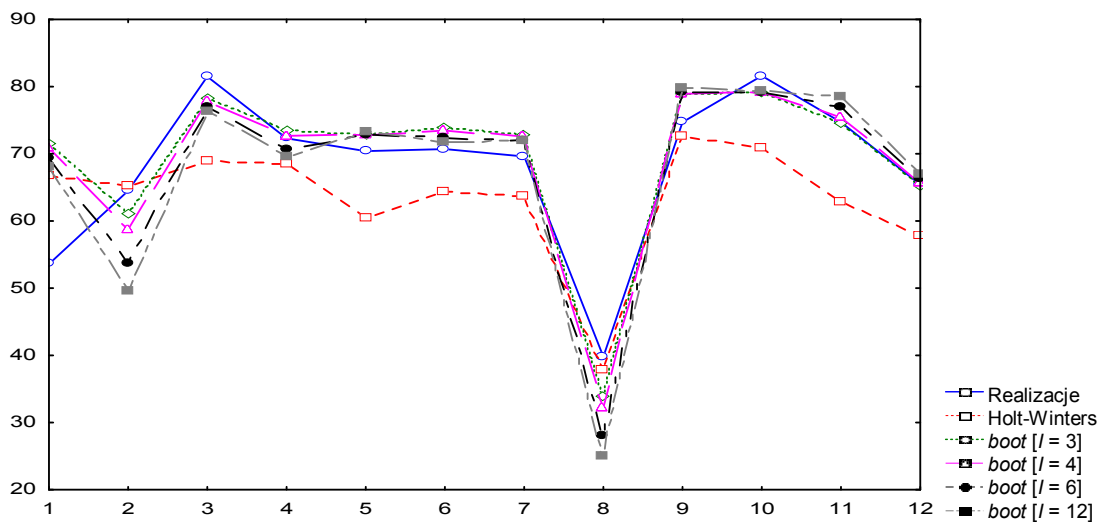
Rys. 2. Prognozy ekstrapolacyjne produkcji piwa (Y_2)



Rys. 3. Prognozy esktrapolacyjne produkcji oleju opałowego (Y3)



Rys. 4. Prognozy esktrapolacyjne produkcji nawozów azotowych (Y4)



Rys. 5. Prognozy esktrapolacyjne produkcji samochodów osobowych (Y5)

PODSUMOWANIE

W artykule przedstawiono zastosowanie metody bootstrapowej do prognozowania za pomocą metody Holta-Wintersa wybranych zmiennych wykazujących wahania sezonowe. Zaproponowany algorytm obliczeniowy oparto na metodzie bootstrapu blokowego. Z przeprowadzonych badań empirycznych wynika, że dla części zmiennych uzyskano zwiększenie dokładności prognoz ekstrapolacyjnych. Dotyczy to zwłaszcza prognoz z 3–6-miesięcznym wyprzedzeniem.

Pewną niedogodnością, związaną ze stosowaniem omawianej metody, jest znaczna czasochłonność, wynikająca z konieczności 1000-krotnego generowania próby bootstrapowej dla każdej zmiennej. Jednak perspektywa otrzymania dokładniejszych prognoz decyduje o wykorzystaniu prezentowanego algorytmu.

PIŚMIENNICTWO

Bühlmann P. 2002. Bootstraps for time series. *Statist. Sci.* 17 (1), 52–72.

Efron B. 1979. Bootstrap methods – another look at the jackknife. *The Ann. Stat.* 7 (1), 1–26.

Kozarski R. Metody bootstrapowe w analizie szeregów czasowych, <http://akson.sgh.waw.pl/~rkozar/boottimes.pdf>, dostęp dn. 12.10.2010 r.

Lahiri S.N. 1999. Theoretical comparisons of block bootstrap methods. *The Ann. Stat.* 27 (1), 386–404.

