

**Zbigniew Tarapata\***

## **Analiza możliwości wykorzystania wybranych modeli wygładzania wykładniczego do prognozowania wartości WIG-u**

**Wydział Cybernetyki Wojskowej Akademii Technicznej w Warszawie**

### Streszczenie

W artykule przedstawiono analizę porównawczą możliwości wykorzystania różnych modeli wygładzania wykładniczego do prognozowania wartości Warszawskiego Indeksu Giełdowego. Zastosowano dwie grupy modeli wygładzania wykładniczego: do prognozowania na podstawie niestacjonarnych oraz stacjonarnych szeregów czasowych. Z pierwszej grupy przebadano modele: Holta i Wintersa (w wersji addytywnej i multiplikatywnej), wykorzystując do doboru optymalnych parametrów tych modeli symulacyjną metodę zaproponowaną w [8]. Za podstawę w metodzie symulacyjnej posłużyły trzy algorytmy stochastycznego poszukiwania minimum funkcji (minimalizacji błędu standardowego prognozy): z losowaniem punktów próbnych, z wyznaczaniem kierunków poprawy oraz z adaptacją rozkładu prawdopodobieństwa kierunku poprawy. W grupie modeli do prognozowania na podstawie stacjonarnych szeregów czasowych podano wyniki dotyczące WIG-u dla modeli Browna: prostego, klasycznego oraz zmodyfikowanej wersji klasycznego. Porównano modele wygładzania ze względu na minimalizację standardowego błędu prognozy oraz szybkość osiągnięcia zadanej dokładności wartości funkcji błędu prognozy. Przedstawiono również porównanie szybkości zbieżności algorytmów stochastycznego doboru parametrów modeli wygładzania w zależności od rodzaju modelu.

### Wstęp

W pracy [8] przedstawiono wyniki badań dotyczących symulacyjnych metod doboru optymalnych parametrów w modelach wygładzania wykładniczego Browna: prostym, klasycznym oraz zmodyfikowanej wersji klasycznego zaproponowanej w [7]. W celu zweryfikowania właściwości prognostycznych prezentowanych modeli oraz oszacowania „dobroci” zaproponowanych algorytmów stochastycznego doboru najlepszych parametrów modeli posłużono się przykładem prognozowania zachowania się wartości WIG-u. Ponieważ modele zaproponowane w [8] (poza zmodyfikowanym modelem Browna) mogą być w zasadzie stosowane do prognozowania zjawisk, które charakteryzują się stacjonarnością, a szereg czasowy opisujący zachowanie się wartości WIG-u cechuje się dużą niestacjonarnością, tzn. wartość średnia indeksu ulega znacznym

---

\* Autor pracuje również jako wykładowca w Wyższej Szkole Ekonomicznej w Warszawie.

zmianom w czasie, więc wydaje się, że najodpowiedniejszymi modelami do prognozowania wartości WIG-u powinny być modele prognostyczne dla niestacjonarnych szeregów czasowych (oczywiście jest to tylko hipoteza, którą będziemy chcieli zweryfikować w trakcie badań). Przedstawicielami takich modeli są modele: Holta i Wintersa. Zaprezentujemy wyniki prognoz dla WIG-u na podstawie tych modeli. Ponadto dokonamy porównania własności prognostycznych tych modeli, jak również algorytmów stochastycznego doboru najlepszych parametrów pod kątem szybkości ich zbieżności do rozwiązania najlepszego. Biorąc pod uwagę powyższe, artykuł ten można traktować jako rozszerzenie i uzupełnienie pracy [8].

Przypomnimy obecnie postacie modeli Holta i Wintersa.

### Modele wygładzania wykładniczego Holta i Wintersa

Liniowy model Holta wygładzania wykładniczego ([3], [6], [9]) służy do wygładzania szeregu czasowego, w którym występują i tendencja rozwojowa, i wahania przypadkowe. Badania Holta były finansowane przez Oddział Badawczy Marynarki Wojennej Stanów Zjednoczonych; rozwinął on niezależnie modele wyrównywania wykładniczego dla procesów stabilnych, procesów z trendami liniowymi i dla danych sezonowych.

Model Holta jest bardziej elastyczny od modeli Browna ([1], [2]) ze względu na występowanie w nim dwóch parametrów:  $\alpha$  i  $\beta$ . Można go opisać za pomocą następujących równań [4]:

$$F_{t-1}(\alpha, \beta) = \alpha \cdot y_{t-1} + (1 - \alpha) \cdot (F_{t-2}(\alpha, \beta) + S_{t-2}(\alpha, \beta)) \quad (1)$$

$$S_{t-1}(\alpha, \beta) = \beta \cdot (F_{t-1}(\alpha, \beta) - F_{t-2}(\alpha, \beta)) + (1 - \beta) \cdot S_{t-2}(\alpha, \beta) \quad (2)$$

gdzie:

$F_{t-1}(\cdot, \cdot)$  - odpowiednik wygładzonej wartości otrzymanej z prostego modelu wygładzania wykładniczego;

$S_{t-1}(\cdot, \cdot)$  - wygładzona wartość przyrostu trendu na moment  $t-1$ ;

$\alpha, \beta$  - parametry modelu o wartościach z przedziału  $[0, 1]$ .

Prognoza na chwilę  $t$  jest wyznaczana następująco:

$$y_t^*(\alpha, \beta) = F_t(\alpha, \beta) + (t - T) \cdot S_T(\alpha, \beta), \quad t > T \quad (3)$$

gdzie  $2 \leq T \leq n$  oznacza liczbę chwil z szeregu czasowego, które bierzemy pod uwagę do budowy prognozy  $y_t^*$ . Najczęściej jest tak, że  $T=t-1$ . Mówimy wtedy o prognozach budowanych z jednookresowym wyprzedzeniem. Różnica między równaniem dla prostego modelu Browna a (1) polega na dodaniu do członu  $F_{t-1}$  (przyjmowanego w prostym modelu wygładzania wykładniczego za prognozę  $y_t^*$ ) oceny przyrostu trendu w momencie  $t-1$  ( $S_{t-1}$ ). Reguła zastosowana przy konstrukcji równania (2) jest taka sama jak w przypadku prostego modelu Browna. Przyjmowana za najnowszy przyrost trendu różnica  $F_{t-2} - F_{t-1}$  jest ważona przez parametr  $\beta$ , poprzednia zaś ocena przyrostu trendu  $S_{t-1}$  - przez  $(1-\beta)$ .

Do budowy modelu Holta potrzebne są początkowe wartości  $F_1$  i  $S_1$ . W literaturze (np. [6]) można znaleźć wiele propozycji rozwiązania tego problemu. Do naszych badań przyjmujemy jedną ze standardowych propozycji:  $F_1=y_1$ ,  $S_1=y_2-y_1$ .

Podobnie jak w modelu Browna podstawowy problem, który dotyczy modelu Holta można zdefiniować następująco:

znaleźć taką parę  $(\alpha^*, \beta^*)$ , dla której zachodzi

$$s(\alpha^*, \beta^*) = \min_{\substack{\alpha \in [0,1] \\ \beta \in [0,1]}} s(\alpha, \beta) \quad (4)$$

gdzie:

$$s(\alpha, \beta) = \sqrt{\frac{1}{n} \cdot \sum_{t=1}^n (y_t - y_t^*(\alpha, \beta))^2} \quad (5)$$

Problem ten można rozwiązać jedną ze standardowych metod optymalizacji nieliniowej (por. [10]). Można go również sprowadzić, po pewnych modyfikacjach funkcji celu, do równoważnego problemu programowania liniowego (por. [8]).

Model Wintersa wykładniczego ([3], [6], [9]) może być stosowany w przypadku szeregów czasowych zawierających tendencję rozwojową, wahania sezonowe oraz wahania przypadkowe. Znane są postacie addytywnej i multiplikatywnej wersji tego modelu.

Wersja addytywna modelu może być opisana za pomocą następujących równań [4]:

$$F_{t-1}(\alpha, \beta, \gamma) = \alpha \cdot (y_{t-1} - C_{t-1-r}(\alpha, \beta, \gamma)) + (1 - \alpha) \cdot (F_{t-2}(\alpha, \beta, \gamma) - S_{t-2}(\alpha, \beta, \gamma)) \quad (6)$$

$$S_{t-1}(\alpha, \beta, \gamma) = \beta \cdot (F_{t-1}(\alpha, \beta, \gamma) - F_{t-2}(\alpha, \beta, \gamma)) + (1 - \beta) \cdot S_{t-2}(\alpha, \beta, \gamma) \quad (7)$$

$$C_{t-1}(\alpha, \beta, \gamma) = \gamma \cdot (y_{t-1} - F_{t-1}(\alpha, \beta, \gamma)) + (1 - \gamma) \cdot C_{t-1-r}(\alpha, \beta, \gamma) \quad (8)$$

a wersja multiplikatywna modelu [4]:

$$F_{t-1}(\alpha, \beta, \gamma) = \alpha \cdot \frac{y_{t-1}}{C_{t-1-r}(\alpha, \beta, \gamma)} + (1 - \alpha) \cdot (F_{t-2}(\alpha, \beta, \gamma) + S_{t-2}(\alpha, \beta, \gamma)) \quad (9)$$

$$S_{t-1}(\alpha, \beta, \gamma) = \beta \cdot (F_{t-1}(\alpha, \beta, \gamma) - F_{t-2}(\alpha, \beta, \gamma)) + (1 - \beta) \cdot S_{t-2}(\alpha, \beta, \gamma) \quad (10)$$

$$C_{t-1}(\alpha, \beta, \gamma) = \gamma \cdot \frac{y_{t-1}}{F_{t-1}(\alpha, \beta, \gamma)} + (1 - \gamma) \cdot C_{t-1-r}(\alpha, \beta, \gamma) \quad (11)$$

gdzie:

$F_{t-1}(\cdot; \cdot; \cdot)$  – odpowiednik wygładzonej wartości otrzymanej z prostego modelu wykładniczego (ocena wartości średniej);

$S_{t-1}(\cdot; \cdot; \cdot)$  – wygładzona wartość przyrostu trendu na moment  $t-1$ ;

$C_{t-1}(\cdot; \cdot; \cdot)$  – ocena wskaźnika sezonowości na moment  $t-1$ ;

$y_{t-1}$  – wartość zmiennej w szeregu czasowym w chwili  $t-1$ ;

$r$  – liczba chwil tworzących cykl sezonowy (tzw. liczba faz w cyklu),  $1 \leq r \leq n$ ;

$\alpha, \beta, \gamma$  – parametry modelu o wartościach z przedziału  $[0, 1]$ .

Zauważmy, że we wzorach (6)-(11)  $t$  może przyjmować wyłącznie wartości  $r+2 \leq t \leq n$ .

Prognoza  $y_t^*$  na chwilę  $t$  jest wyznaczana następująco:

- w modelu addytywnym:

$$y_t^*(\alpha, \beta, \gamma) = F_T(\alpha, \beta, \gamma) + (t - T) \cdot S_T(\alpha, \beta, \gamma) + C_{t-r}(\alpha, \beta, \gamma), \quad t > T \quad (12)$$

- w modelu multiplikatywnym:

$$y_t^*(\alpha, \beta, \gamma) = (F_T(\alpha, \beta, \gamma) + (t - T) \cdot S_T(\alpha, \beta, \gamma)) \cdot C_{t-r}(\alpha, \beta, \gamma), \quad t > T \quad (13)$$

gdzie  $r+1 \leq T \leq n$  oznacza liczbę chwil z szeregu czasowego, które bierzemy pod uwagę do budowy prognozy  $y_t^*$ .

W literaturze można spotkać wiele propozycji dotyczących doboru wartości parametrów  $\alpha, \beta, \gamma$ . Np. w pracy [11] podano następujące wskazówki co do ich wartości: jeżeli poszczególne składowe modelu adaptacyjnego zmieniają się szybko, to parametry modelu należy przyjąć na poziomie bliskim jedności; gdy składowe są względnie stabilne, to wówczas parametry modelu powinny być bliskie zeru. Inne propozycje można znaleźć m.in. w [3], [5], [9]. Oczywiście takie podejścia nie gwarantują nam dobrania najlepszej kombinacji współczynników. Dlatego też zastosujemy metody symulacyjnego doboru ich wartości (reprezentowane przez algorytmy stochastyczne) minimalizujące standardowy błąd prognozy.

Do budowy modelu Wintersa potrzebne są początkowe wartości  $F_{r+1}, S_{r+1}$  i  $C_1, \dots, C_r$ . W literaturze (np. [5], [6]) można znaleźć wiele propozycji rozwiązania tego problemu. Do naszych badań przyjmujemy jedną ze standardowych propozycji:

- $F_{r+1} = y_{r+1}$ ;
- $S_{r+1} = 0$ ;
- dla  $i$ -tego wyrazu w cyklu,  $i = \overline{1, r}$  :

$$C_i = \begin{cases} y_i - \frac{1}{r} \cdot \sum_{k=1}^r y_k & , \text{ dla wersji addytywnej modelu} \\ \frac{y_i}{\frac{1}{r} \cdot \sum_{k=1}^r y_k} & , \text{ dla wersji multiplikatywnej modelu} \end{cases} \quad (14)$$

Podobnie jak to było w modelach Browna i Holta podstawowy problem, który dotyczy modeli Wintersa można zdefiniować następująco:

znaleźć taką trójkę  $(\alpha^*, \beta^*, \gamma^*)$ , dla której zachodzi

$$s(\alpha^*, \beta^*, \gamma^*) = \min_{\substack{\alpha \in [0,1] \\ \beta \in [0,1] \\ \gamma \in [0,1]}} s(\alpha, \beta, \gamma) \quad (15)$$

gdzie:

$$s(\alpha, \beta, \gamma) = \sqrt{\frac{1}{n} \cdot \sum_{t=1}^n (y_t - y_t^*(\alpha, \beta, \gamma))^2} \quad (16)$$

oraz  $y_t^*(\alpha, \beta, \gamma)$  w (16) opisane jest przez (12) dla modelu addytywnego lub przez (13) dla modelu multiplikatywnego.

Problem ten, tak jak dwa poprzednie, można rozwiązać jedną ze standardowych metod optymalizacji nieliniowej (por. [10]). Można go również sprowadzić, po pewnych modyfikacjach funkcji celu, do równoważnego problemu programowania liniowego (por. [8]). Ponadto w modelach Wintersa pojawia się dodatkowy problem doboru wartości długości  $r$  cyklu. Wartość ta w zasadzie powinna wynikać z szeregu czasowego (np. gdy występuje sezonowość tygodniowa, kwartalna, itp.), bo przecież pod tym kątem dobieramy model: jeżeli występuje sezonowość, to model Wintersa powinien być dobrym narzędziem prognozowania, jeżeli tej sezonowości nie widać, to nie powinno się używać modeli Wintersa. Jednakże czasami możemy chcieć zastosować ten model nawet dla szeregów czasowych, w których nie występuje wyraźna sezonowość. Wówczas właśnie pojawia się problem doboru najlepszej wartości  $r$ . Odniesiemy się do tego problemu przy analizie wyników badań.

### **Idea stochastycznych algorytmów doboru najlepszych parametrów w modelach Holta i Wintersa**

We wspomnianej wcześniej pracy [8] przedstawiono szeroki opis algorytmów stochastycznego doboru optymalnych parametrów modeli prognostycznych wraz z uzasadnieniem ich wyboru, więc nasze rozważania dotyczące zastosowania tych algorytmów do wyznaczania parametrów modeli Holta i Wintersa ograniczymy do niezbędnego minimum, przedstawiając jedynie modyfikacje tych algorytmów, które z racji charakteru modeli Holta i Wintersa mają trochę inną postać niż w [8].

#### Algorytm 1 (z losowaniem punktów próbnych)

Ogólny schemat tego algorytmu przedstawia się następująco:

$$x^{n+1} = \begin{cases} \xi^n, & \text{gdy } f(\xi^n) < f(x^n) \\ x^n, & \text{w przeciwnym przypadku} \end{cases} \quad (17)$$

gdzie:

$x^{n+1}$  - kolejne przybliżenie (w  $n+1$ -szym losowaniu) rozwiązania  $x \in S$ ,

$$x = \begin{cases} (\alpha, \beta), & \text{dla modelu Holta} \\ (\alpha, \beta, \gamma), & \text{dla modeli Wintersa} \end{cases} \quad (18)$$

$$S = \begin{cases} [0,1] \times [0,1], & \text{dla modelu Holta} \\ [0,1] \times [0,1] \times [0,1], & \text{dla modeli Wintersa} \end{cases} \quad (19)$$

$\xi^n$  - punkt losowy w zbiorze  $S$ ,  $\xi^n \sim U(S)$ ;

$f(\cdot)$  - minimalizowana wartość funkcji celu,

$$f(\cdot) = \begin{cases} (5), & \text{dla modelu Holta} \\ (16), & \text{dla modeli Wintersa} \end{cases} \quad (20)$$

Przypomnijmy, że np. zapis:  $\xi^n \sim U(S)$  oznacza, iż zmienna losowa  $\xi^n$  ma rozkład równomierny na zbiorze  $S$ .

Algorytm 2 (z losowaniem kierunku poprawy)

Ogólny schemat tego algorytmu przedstawia się następująco:

$$x^{n+1} = \begin{cases} x^n + a_n \cdot \xi^n, & \text{gdy } f(x^n + a_n \cdot \xi^n) < f(x^n) \\ x^n, & \text{w przeciwnym przypadku} \end{cases} \quad (21)$$

gdzie:

$x^{n+1}$  - kolejne przybliżenie (w  $n+1$ -szym losowaniu) rozwiązania  $x \in S$ ,  $x$  – określone przez (18),  $S$  – określone przez (19);

$\xi^n$  - punkt losowy (kierunek poprawy) w zbiorze  $Z$ ,

$$Z = \begin{cases} \{-e_1, e_1, -e_2, e_2\}, & \text{dla modelu Holta} \\ \{-e_1, e_1, -e_2, e_2, -e_3, e_3\}, & \text{dla modeli Wintersa} \end{cases} \quad (22)$$

gdzie wielkości  $e_1, e_2, e_3$  oznaczają wersory poszczególnych osi układu współrzędnych,  $\xi^n \sim U(Z)$ ;

$a_n$  - długość „kroku”, zmienna losowa o rozkładzie równomiernym losowanym ze zbioru  $T$ ,

$$T = \begin{cases} [0, 1 - x^n], & \text{gdy } \xi^n \in \{e_1, e_2, e_3\} \\ [0, x^n], & \text{gdy } \xi^n \in \{-e_1, -e_2, -e_3\} \end{cases} \quad (23)$$

oraz  $a_n \sim U(T)$ ;

$f(\cdot)$  - funkcja określona przez (20);

Komentarza wymaga interpretacja kierunku poprawy. Np. dla modelu Holta, jeżeli wylosujemy  $-e_2$ , to oznacza to, że będziemy zmniejszać (bo „-”) wartość drugiego parametru (bo „numer”=2, czyli  $\beta$ ), itp.

Algorytm 3 (z adaptacją rozkładu prawdopodobieństwa kierunku poprawy)

Idea tego algorytmu jest podobna jak algorytmu 3 zaprezentowanego w [8]. W tym celu weźmy pod uwagę algorytm 2 i niech

$$Z = \{e'_1, e'_2, \dots, e'_{2 \cdot K}\} \quad (24)$$

będzie zbiorem rozważanych tam kierunków, przy czym  $e'_j = e_j$  dla  $j = \overline{1, K}$  oraz

$e'_j = -e_{j-K}$  dla  $j = \overline{K+1, 2 \cdot K}$ , gdzie  $K$  oznacza liczbę parametrów modelu (tzn.

$K=2$  dla modelu Holta oraz  $K=3$  dla modeli Wintersa). Niech  $(p_j^{(n)})_{1 \leq j \leq 2 \cdot K}$  będzie

rozkładem prawdopodobieństwa, według którego w  $n$ -tym kroku iteracyjnym losujemy jeden z kierunków  $e'_j$ ,  $j = \overline{1, 2 \cdot K}$ . Niech

$$p_j^{(1)} = \frac{1}{2 \cdot K}, \quad j = \overline{1, 2 \cdot K} \quad (25)$$

i rozważmy następujący sposób modyfikacji rozkładu  $(p_j^{(n)})_{1 \leq j \leq 2 \cdot K}$ : jeżeli na  $n$ -tym

etapie obliczeń wylosowaliśmy kierunek  $e'_{j(n)}$ , to przyjmujemy:

$$p_{j(n)}^{(n+1)} = p_{j(n)}^{(n)} - \theta \cdot (1 - p_{j(n)}^{(n)}) \cdot \text{sgn} [f(x^n + a_n \cdot e'_{j(n)}) - f(x^n)] \quad (26)$$

oraz dla  $i = \overline{1, 2 \cdot K}$ ,  $i \neq j(n)$ :

$$p_i^{(n+1)} = p_i^{(n)} + \frac{\theta}{2 \cdot K - 1} \cdot (1 - p_{j(n)}^{(n)}) \cdot \text{sgn} [f(x^n + a_n \cdot e'_{j(n)}) - f(x^n)] \quad (27)$$

gdzie  $\theta \in [0, 1]$  jest parametrem algorytmu (dla  $\theta=0$  mamy nieadaptacyjny algorytm 2). Długość „kroku”  $a_n$  w kolejnych iteracjach losujemy z rozkładem równomiernym ze zbioru:

$$T(n) = \begin{cases} [0, \delta \cdot (1 - x^n)], & \text{gdy } j(n) \in \{1, \dots, K\} \\ [0, \delta \cdot x^n], & \text{gdy } j(n) \in \{K + 1, \dots, 2 \cdot K\} \end{cases} \quad (28)$$

przy czym, jeżeli na  $n$ -tym i  $n-1$ -szym etapach obliczeń wylosowaliśmy ten sam kierunek  $\xi^{n-1} = e'_{j(n-1)} = \xi^n = e'_{j(n)}$ , to  $\delta \in (0, 1)$ ; w przeciwnym przypadku  $\delta = 1$ . Możemy obecnie podać ogólny schemat algorytmu 3 :

$$x^{n+1} = \begin{cases} x^n + a_n \cdot \xi^n, & \text{gdy } f(x^n + a_n \cdot \xi^n) < f(x^n) \\ x^n, & \text{w przeciwnym przypadku} \end{cases} \quad (29)$$

gdzie:

- $x^{n+1}$  - kolejne przybliżenie (w  $n+1$ -szym losowaniu) rozwiązania  $x \in S$ ,  $x$  – określone przez (18),  $S$  – określone przez (19);
- $\xi^n$  - punkt losowy (kierunek poprawy) w zbiorze  $Z$  określonym przez (24) losowany z rozkładem prawdopodobieństwa określonym przez (26)÷(27);
- $a_n$  - wartość długości „kroku” będąca zmienną losową o rozkładzie równomiernym na zbiorze  $T(n)$  określonym przez (28),  $a_n \sim U(T(n))$ ;
- $f(\cdot)$  - funkcja określona przez (20);

## Analiza wyników badań

Przedstawimy obecnie wyniki badań mających na celu porównanie możliwości prognozowania wartości Warszawskiego Indeksu Giełdowego za pomocą modeli wygładzania wykładniczego. Ponadto porównamy zbieżność prezentowanych algorytmów oraz dokładność rozwiązań, które z nich otrzymujemy. W tym celu musimy zdefiniować pewne miary za pomocą których będziemy porównywać zarówno modele prognozowania, jak i algorytmy służące do doboru optymalnego zestawu parametrów modeli. Będą one uogólnieniem miar zaproponowanych w [8].

Pojedynczy eksperyment symulacyjny polegał na wylosowaniu  $N=500$  punktów losowych. Wykonywano  $M=30$  eksperymentów symulacyjnych. Oznaczmy przez  $x_{i,j,k,l}^N$  wartość  $l$ -tej współrzędnej ( $l = \overline{1, K}$ ) wektora współczynników (18) odpowiedniego modelu otrzymaną po  $N$  losowaniach w  $i$ -tym ( $i = \overline{1, M}$ ) eksperymencie symulacyjnym, dla  $j$ -tego ( $j = \overline{1, 3}$ ) algorytmu oraz  $k$ -tej metody wygładzania (dla  $k=1$  mamy model Holta,  $k=2$  – model Wintersa addytywny,  $k=3$  –

model Wintersa multiplikatywny). Estymator punktowy wartości oczekiwanej poszczególnych składowych wektora  $x$  dla  $j$ -tego algorytmu oraz  $k$ -tego modelu wygładzania po  $N$  losowaniach punktów próbnych wyznaczmy następująco:

$$\tilde{x}_{j,k,l}^N = \frac{1}{M} \cdot \sum_{i=1}^M x_{i,j,k,l}^N, \quad j = \overline{1,3}, \quad k = \overline{1,3}, \quad l = \overline{1,K} \quad (30)$$

Oczywiście mamy:  $\tilde{x}_{j,k}^N = (\tilde{x}_{j,k,l}^N)_{l=\overline{1,K}}, \quad j = \overline{1,3}, \quad k = \overline{1,3}$ .

Zależność (30) przedstawia oszacowanie wartości  $l$ -tego parametru dokonane przez  $j$ -ty algorytm zastosowany dla  $k$ -tego modelu wygładzania przy  $N$  losowaniach punktów próbnych.

Oznaczmy przez  $\mathbf{N} = \{1, \dots, N\}$  zbiór numerów wylosowanych punktów próbnych. Z kolei przez  $n_{i,j,k}$  oznaczmy następującą wielkość:

$$n_{i,j,k} = \min \left\{ n \in \mathbf{N} : \frac{f(x_{i,j,k}^n) - f(x_k^*)}{f(x_k^*)} \leq 0.01 \right\} \quad (31)$$

gdzie  $x_{i,j,k}^n$  oznacza wektor wartości parametrów otrzymany do  $n$ -tego losowania w  $i$ -tym eksperymencie symulacyjnym ( $i = \overline{1, M}$ ), dla  $j$ -tego algorytmu ( $j = \overline{1,3}$ ) i  $k$ -tego modelu wygładzania (dla  $k=1$  mamy model Holta,  $k=2$  – model Wintersa addytywny,  $k=3$  – model Wintersa multiplikatywny), a  $x_k^*$  oznacza wektor optymalnych wartości parametrów dla  $k$ -tego modelu wygładzania. Ponieważ wektor  $x_{i,j,k}^n$  jest punktem losowym w przestrzeni  $S$  (por. (19)), więc  $n_{i,j,k}$  jest  $i$ -tą realizacją zmiennej losowej  $n_{j,k}$ . Przeprowadzając  $M$  eksperymentów symulacyjnych otrzymamy pewien rozkład empiryczny zmiennej  $n_{j,k}$ , który może być reprezentowany np. za pomocą dystrybuanty empirycznej  $F_{j,k}^e$ . Wyznaczmy estymator punktowy  $\hat{n}_{j,k}^{0.9}$  kwantyla rzędu 0.9 zmiennej losowej  $n_{j,k}$ . Jest to taka liczba, która spełnia warunek:

$$F_{j,k}^e(\hat{n}_{j,k}^{0.9}) = P(n_{j,k} < \hat{n}_{j,k}^{0.9}) = 0.9 \quad (32)$$

Interpretacja liczby  $\hat{n}_{j,k}^{0.9}$  (dla  $j$ -tego algorytmu i  $k$ -tego modelu wygładzania) jest następująca: jest to minimalna liczba losowań punktów próbnych potrzebnych do tego, aby z prawdopodobieństwem 0.9 można było stwierdzić, że  $f(x_{i,j,k}^{\hat{n}_{j,k}^{0.9}})$  różni się od  $f(x^*)$  nie więcej niż o jeden procent. Podobnie jak w [8] dla algorytmu 3 przyjęto  $\theta=0.5$  (por. (26), (27)).

Przejdziemy obecnie do interpretacji wyników.

W Tabeli 1 przedstawiono wartości  $\tilde{x}_{j,k}^N, x^*, f(\tilde{x}_{j,k}^N), f(x^*), \hat{n}_{j,k}^{0.9}$  dla rozpatrywanych modeli Holta i Wintersa oraz trzech algorytmów (dla WIG-u w okresie X.1994÷XI.1999). Z Tabeli 1 wynika, że najlepsze prognozy uzyskujemy dla modeli Wintersa przy  $r=1$ . Zauważmy przy tym, że w modelach Wintersa niewielki wpływ na wartość błędu prognozy miał parametr  $\gamma$ . Jest to zrozumiałe, gdyż parametr  $\gamma$  „odpowiada” za sezonowość, która w tych modelach w zasadzie



nie istniała (długość cyklu sezonowego  $r$  wynosiła 1). Świadczy o tym chociażby fakt dużej zmienności wartości WIG-u.

Na **Wykresie 1** zaprezentowano przykład wygładzenia szeregu czasowego za pomocą modeli Holta i Wintersa dotyczący WIG-u długookresowego, dla następujących wartości parametrów: dla modelu Wintersa -  $\alpha=0.08$ ,  $\beta=0.4$ ,  $\gamma=0.01$ ; dla modelu Holta -  $\alpha=0.08$ ,  $\beta=0.07$ . Wartości standardowego błędu prognozy dla obu tych modeli wyniosły: dla modelu Wintersa -  $s(\alpha=0.08, \beta=0.4, \gamma=0.01) \approx 808.6$ ; dla modelu Holta -  $s(\alpha=0.08, \beta=0.07) \approx 745$ . Widać, że wartości te są dużo większe od wartości błędu dla optymalnych zestawów parametrów (por. z  $f(x^*)$  w **Tabeli 1**). Jest to dość oczywiste, gdyż z reguły im większy stopień wygładzenia szeregu czasowego, tym większe błędy prognoz.

Na **Wykresie 2** przedstawiono zmiany wartości indeksu WIG we wrześniu 1999r. oraz prognozy (liczone z jednookresowym wyprzedzeniem) wartości WIG-u na podstawie modeli Holta i Wintersa. Zauważmy, że w tym przypadku również najlepsze prognozy otrzymaliśmy z modeli Wintersa, przy czym tym razem mieliśmy sezonowość 4-ro okresową ( $r=4$ ), którą łatwo jest wyodrębnić z szeregu czasowego zaprezentowanego na tym wykresie. Z **Wykresu 2** nie wynika, żeby model Wintersa był zdecydowanie lepszy od modelu Holta (co wynika z **Tabeli 2**, por. wartości  $f(x^*)$  dla modelu Holta i Wintersa, różnica ponad 15%). Być może miał na to wpływ fakt, że prognozy w modelu Wintersa mogły być liczone dopiero od piątego (tzn.  $r+1$ -szego) okresu.

W **Tabeli 2** przedstawiono wartości  $\tilde{x}_{j,k}^N$ ,  $x^*$ ,  $f(\tilde{x}_{j,k}^N)$ ,  $f(x^*)$ ,  $\hat{n}_{j,k}^{0.9}$  dla rozpatrywanych modeli oraz algorytmów (dla WIG-u na wrzesień 1999).

Na **Wykresie 3** przedstawiono porównanie własności prognostycznych sześciu modeli wygładzania wykładniczego na podstawie wartości standardowego błędu prognozy dla WIG-u krótkookresowego (wrzesień 1999) i długookresowego (X.1994÷XI.1999). Zauważmy, że zdecydowanie najlepsze prognozy uzyskano z modeli Wintersa. Innym modelem, który może się równać ze wspomnianymi jest zmodyfikowany model Browna [7]. Jest on o tyle korzystniejszy od modelu Wintersa, że ma dużo prostszą postać, jest modelem jednoparametrowym i - w konsekwencji - jest dużo szybszy od modelu Wintersa. Wystarczy powołać się na **Wykres 5** (por. też **Wykres 4**), z którego wynika, że średnia wartość kwantyla  $\hat{n}_{j,k}^{0.9}$  (por. interpretację przy wzorze (32)) dla zmodyfikowanego modelu Browna (dla trzech algorytmów i WIG-u krótko- oraz długo-okresowego) wyniosła  $105/6=17.5$  natomiast dla modeli Wintersa wartość ta jest większa niż  $1394/6 \approx 232$ .

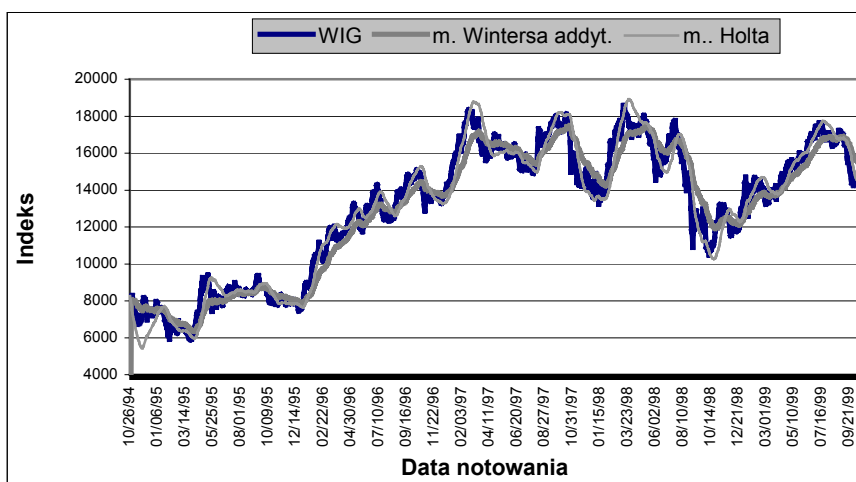
**Wykres 4** przedstawia porównanie szybkości zbieżności różnych metod prognozowania dla średniej wartości kwantyli  $\hat{n}_{j,k}^{0.9}$ . Widać wyraźną zależność między liczbą parametrów modelu i szybkością jego zbieżności (im mniej parametrów, tym szybkość zbieżności większa).

Na **Wykresie 5** zaprezentowano szybkość zbieżności algorytmów, mierzoną wartością kwantyla  $\hat{n}_{j,k}^{0.9}$ , dla różnych modeli prognozowania. Wynika z niego, że najlepszym algorytmem, bez względu na model prognozowania, okazał się *algorytm 3*. Nieznacznie ustępuje mu *algorytm 1* i dość znacznie „odstaje” od nich *algorytm 2*. Częściowo ten fakt wyjaśniono w [8]. Szczególnie widoczna jest ta różnica dla modeli wieloparametrowych (Holta i Wintersa).

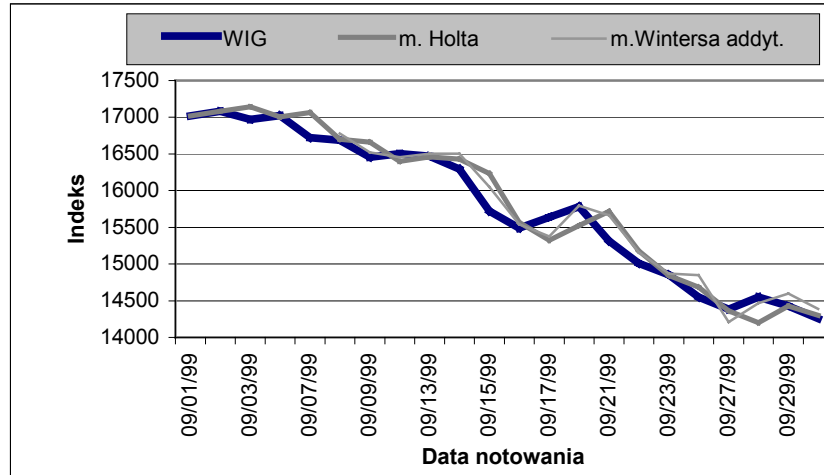
**Tabela 1**  
**Oszacowane wartości współczynników modeli Holta i Wintersa oraz odpowiadające im wartości funkcji  $f$  błędu standardowego prognozy dla WIG-u (X.1994-XI.1999)**

| Rodzaj modelu<br>$j$                        | Rodzaj algorytmu<br>$k$ | $\tilde{x}_{j,k}^N$     | $x^*$                   | $f(\tilde{x}_{j,k}^N)$ | $f(x^*)$ | $\hat{n}_{j,k}^{0.9}$ |
|---|-------------------------|-------------------------|-------------------------|------------------------|----------|-----------------------|
| Holta ( $j=1$ )                             | 1                       | 0,99<br>0,046           | 1<br>0,035              | 247,63                 | 247,03   | 215                   |
|   | 2                       | 0,954<br>0,05           |                         | 249,63                 |          | >500                  |
|   | 3                       | 0,98<br>0,05            |                         | 248,21                 |          | 395                   |
| Wintersa addytywny<br>( $j=2$ )<br>$r=1$    | 1                       | 0,71<br>0,698<br>0,225  | 0,697<br>0,731<br>0     | 240,37                 | 239,01   | 21                    |
|   | 2                       | 0,511<br>0,569<br>0,638 |                         | 240,59                 |          | 18                    |
|   | 3                       | 0,456<br>0,705<br>0,803 |                         | 239,13                 |          | 16                    |
| Wintersa multiplikat.<br>( $j=3$ )<br>$r=1$ | 1                       | 0,71<br>0,698<br>0,225  | 0,503<br>0,732<br>0,683 | 240,31                 | 238,94   | 21                    |
|   | 2                       | 0,511<br>0,569<br>0,638 |                         | 240,50                 |          | 18                    |
|   | 3                       | 0,456<br>0,705<br>0,803 |                         | 238,99                 |          | 16                    |

**Wykres 1**  
**Przykład częściowego wygładzenia szeregu czasowego na podstawie modeli Holta i Wintersa (wersja addytywna,  $r=1$ ) dla WIG-u (X.1994-XI.1999)**



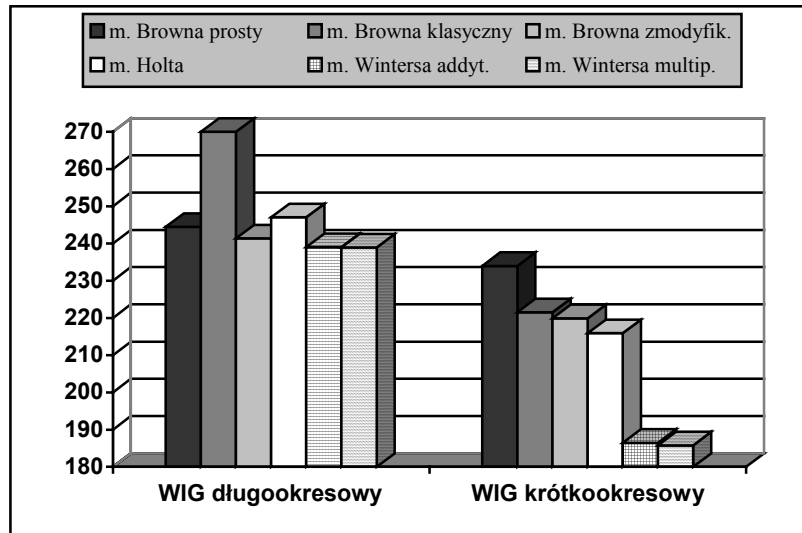
**Wykres 2**  
**Wartości historyczne i najlepsze prognozy na podstawie modeli Holta i Wintersa**  
**(wersja addytywna,  $r=4$ ) dla WIG-u na wrzesień 1999r.**



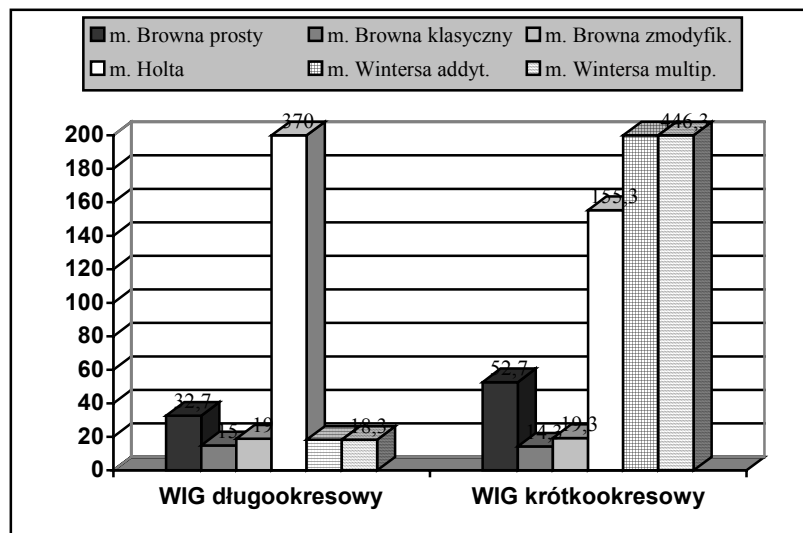
**Tabela 2**  
**Oszacowane wartości współczynników modeli Holta i Wintersa oraz odpowiadające im wartości funkcji  $f$  błędu standardowego prognozy dla WIG-u na wrzesień 1999r.**

| Rodzaj modelu $j$                        | Rodzaj algorytmu $k$ | $\tilde{x}_{j,k}^N$     | $x^*$             | $f(\tilde{x}_{j,k}^N)$ | $f(x^*)$ | $\hat{n}_{j,k}^{0.9}$ |
|--|----------------------|-------------------------|-------------------|------------------------|----------|-----------------------|
| Holta ( $j=1$ )                          | 1                    | 0,971<br>0,172          | 1<br>0,173        | 216,80                 | 215,92   | 96                    |
|  | 2                    | 0,994<br>0,211          |                   | 216,62                 |          | 258                   |
|  | 3                    | 0,999<br>0,203          |                   | 216,30                 |          | 112                   |
| Wintersa addytywny ( $j=2$ )<br>$r=4$    | 1                    | 0,772<br>0,93<br>0,209  | 0,8<br>1<br>0,097 | 191,78                 | 186,47   | >500                  |
|  | 2                    | 0,822<br>0,972<br>0,446 |                   | 190,37                 |          | >500                  |
|  | 3                    | 0,784<br>0,998<br>0,307 |                   | 187,56                 |          | 339                   |
| Wintersa multiplikat. ( $j=3$ )<br>$r=4$ | 1                    | 0,772<br>0,93<br>0,209  | 0,8<br>1<br>0,127 | 191,13                 | 185,78   | >500                  |
|  | 2                    | 0,822<br>0,972<br>0,446 |                   | 189,43                 |          | >500                  |
|  | 3                    | 0,784<br>0,998<br>0,307 |                   | 186,56                 |          | 339                   |

**Wykres 3**  
**Porównanie własności prognostycznych różnych modeli wykładniczego<sup>1</sup>**  
**dla WIG-u, na podstawie wartości standardowego błędu prognozy (oś Y)**  
**dla optymalnych wartości parametrów modeli**

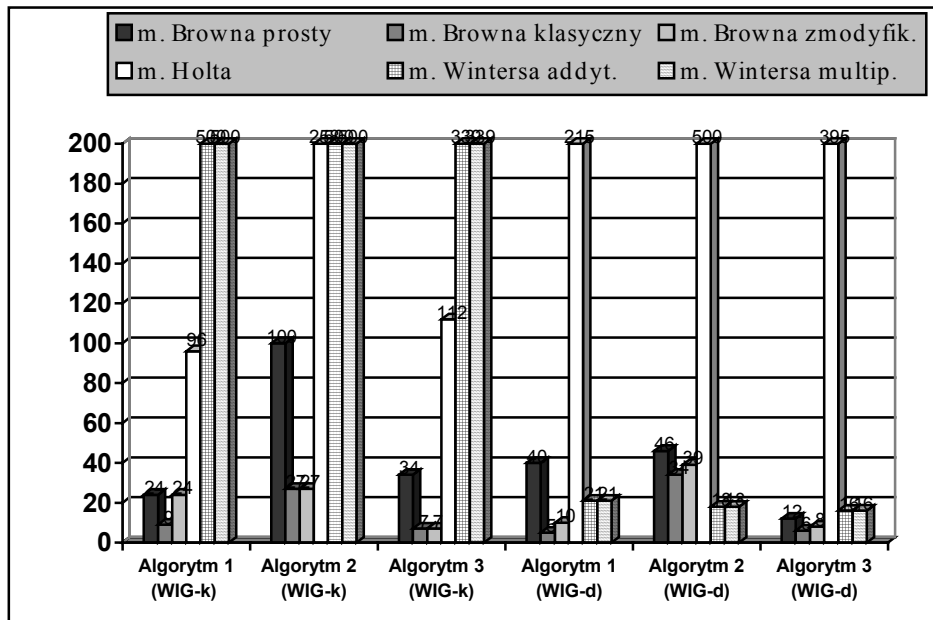


**Wykres 4**  
**Porównanie szybkości zbieżności metod prognozowania<sup>1</sup> dla WIG-u**  
**na podstawie średniej wartości kwantyla  $\hat{n}_{j,k}^{0.9}$  (oś Y)**



<sup>1</sup> Wyniki dla modeli Browna zaczerpnięto z [8]

**Wykres 5**  
**Porównanie szybkości zbieżności algorytmów dla WIG-u**  
**na podstawie wartości kwantyla  $\hat{n}_{j,k}^{0.9}$  (oś Y)<sup>2</sup>**  
**(k - WIG krótkookresowy, d - WIG długookresowy)**



### Uwagi i wnioski

Z wyników badań zaprezentowanych w poprzednim rozdziale wynika, że tylko częściowo potwierdziła się hipoteza postawiona na początku tego artykułu, iż modele Holta i Wintersa będą dawać lepsze (tzn. dokładniejsze) prognozy przy niestacjonarnym szeregu czasowym (jaki tworzy wartość WIG-u w okresie X.1994÷XI.1999) niż modele Browna, dla których szczegółowe wyniki zaprezentowano w [8]. Hipotezę tę potwierdziły jedynie modele Wintersa (przy czym różnica w wynikach dla addytywnej i multiplikatywnej wersji tego modelu była pomijalnie mała). Model Holta okazał się stosunkowo słaby (jak na stopień skomplikowania, w porównaniu z modelami Browna) zarówno ze względu na szybkość zbieżności (por. **Wykres 4**), jak i minimalizację błędu prognozy (por. **Wykres 1**). Bardzo dobre własności (ze względu na szybkość zbieżności, jak i wielkość błędu prognozy) potwierdził zmodyfikowany model Browna. Dodajmy jeszcze, że model ten ma bardzo prostą postać a fakt, że jest to model tylko z jednym parametrem znacznie ułatwia i przyspiesza prognozowanie. Przedstawione wyniki badań nie dyskryminują jednakże żadnego z modeli, przynajmniej z punktu widzenia wartości błędu prognozy. Zauważmy bowiem, że

<sup>2</sup> Wyniki dla modeli Browna zaczerpnięto z [8]

procentowe różnice między wielkościami błędów prognoz dla poszczególnych modeli w stosunku do rzeczywistych wartości WIG-u nie są większe niż ok. 2%. Gdyby jednak chcieć wskazać model wygładzania wykładniczego, który był stosunkowo najlepszy (z punktu widzenia zarówno szybkości zbieżności, jak i wartości błędu prognozy), to należałoby wskazać na zmodyfikowany model Browna. Wprawdzie szybkość zbieżności klasycznego modelu Browna jest nieznacznie większa niż dla jego zmodyfikowanej wersji (por. **Wykres 4**), ale klasyczny model Browna jest gorszy ze względu na wartość błędu prognozy (por. **Wykres 3**). Poza tym posiada bardziej skomplikowaną postać. Podsumowując, można zaryzykować stwierdzenie, że zaprezentowane modele wygładzania wykładniczego dość dobrze będą sobie radzić również w przyszłości z prognozowaniem wartości WIG-u. Częściowym potwierdzeniem tego przypuszczenia mogą być wyniki zaprezentowane w niniejszym artykule.

### Literatura

- [1]. Brown R.G.: Statistical Forecasting for Inventory Control. *McGrow Hill*, New York 1959.
- [2]. Brown R.G.: Smoothing, Forecasting and Prediction of Discrete Time Series. *McGrow Hill*, New York 1963.
- [3]. Chatfield Ch.: Holt-Winters Forecasting: Some Practical Issues. *The Statistician*, vol. 37 (1986), pp. 129-140.
- [4]. Cieślak M. (red.): Prognozowanie gospodarcze. *PWN*, Warszawa 1997.
- [5]. Gardner, E. S., Jr.: Exponential smoothing: The state of the art. *Journal of Forecasting*, 4 (1985), 1-28.
- [6]. Makridakis S., Wheelwright S.C.: Forecasting Methods and Applications. *John Wiley & Sons*, New York 1989.
- [7]. Pawłowski Z.: Prognozy ekonometryczne. *PWN*, Warszawa 1973.
- [8]. Tarapata Z.: Symulacyjna metoda doboru optymalnych parametrów w prognostycznych modelach wygładzania wykładniczego. *Zeszyty Naukowe Wyższej Szkoły Ekonomicznej*, Warszawa 2000. (w druku)
- [9]. Yar M., Chatfield Ch.: Prediction Intervals for the Holt-Winters Forecasting Procedure. *International Journal of Forecasting*, vol. 6 (1990), pp.127-137.
- [10]. Zangwill W. I.: Programowanie nieliniowe. *WNT*, Warszawa 1974.
- [11]. Zawadzki J.: Ekonometryczne metody prognozowania w przedsiębiorstwie. *Wydawnictwo Uniwersytetu Szczecińskiego*, Szczecin 1989.