

Sugerowany przypis: Chybalski F., *Propozycja modelu prognostycznego dla wartości jednostek rozrachunkowych OFE*, „Przegląd Statystyczny”, nr 3/2006, Dom Wydawniczy Elipsa, Warszawa 2006, s. 73-82

Propozycja modelu prognostycznego dla wartości jednostek rozrachunkowych OFE

1. Wstęp

Od 1999 roku, po przeprowadzeniu reformy systemu emerytalnego, funkcjonują w Polsce otwarte fundusze emerytalne, tworzące II filar systemu emerytalnego. Instytucje te realizują kapitałowe ubezpieczenia emerytalne poprzez gromadzenie i inwestowanie odprowadzanych do nich składek. Na koniec 2005 roku w OFE zgromadzonych było ponad 80 mld złotych. Kwota ta świadczy o ogromnej skali działalności tych funduszy.

Aktywa OFE zmieniają się pod wpływem trzech czynników:

- napływu składek emerytalnych,
- zysków z dokonywanych inwestycji,
- wypłat świadczeń emerytalnych.

Trzeci z wymienionych czynników zacznie oddziaływać na wartość aktywów funduszy dopiero w 2009 roku, gdyż wtedy zostaną wypłacone pierwsze emerytury z II filaru. Obecnie ma miejsce napływ składek i ich inwestowanie.

Miarą, która pozwala dokonać analizy zmian wartości aktywów funduszy, spowodowanych efektami działalności inwestycyjnej, jest wartość jednostki rozrachunkowej JR. Jednostkę rozrachunkową można uznać za odpowiednik jednostki uczestnictwa w funduszu inwestycyjnym. Jej wartość to relacja aktywów netto funduszu do liczby jednostek rozrachunkowych. „Zmiana wartości jednostki rozrachunkowej odpowiada tym względnym zmianom wartości aktywów netto funduszu, które są spowodowane zyskiem lub stratą z inwestycji” ([3], s. 11). Nowo napływające składki powodują wzrost liczby jednostek rozrachunkowych jednocześnie nie wpływając na jej wartość. Wynika to z faktu, iż nowo napływające składki są przeliczane na jednostki rozrachunkowe według ich bieżącej wartości. Znajomość wartości jednostki rozrachunkowej dla danego OFE w momencie początkowym i końcowym analizowanego okresu pozwala obliczyć stopę zwrotu tego OFE, a także miary charakteryzujące zmienność stóp zwrotu, np. odchylenie standardowe lub semiodchylenie standardowe, czy też asymetrię stóp zwrotu.

W toku prowadzonych przez autora badań podjęto już próbę prognozowania wartości jednostek rozrachunkowych OFE. Ponieważ zauważono, że portfele inwestycyjne OFE różnią się pod względem struktury nie znacznie, można było przypuszczać, iż także szeregi czasowe wartości JR funduszy są do siebie podobne (zob. rys. 1). Zaproponowano więc sposobów prognozowania wartości JR, wykorzystujący fakt tego podobieństwa i bazujący na modelach ARIMA oraz modelach regresji. Modelem ARIMA wyznaczono prognozy średniej wartości jednostki rozrachunkowej dla całego rynku, gdyż zauważono, iż zmienna ta jest silnie skorelowana z wartościami jednostek rozrachunkowych poszczególnych OFE. Zmienną tę potraktowano jako zmienną objaśniającą w modelach regresji dla wartości jednostek rozrachunkowych poszczególnych OFE. Wyznaczono prognozy na trzy miesiące wprzód, których średni absolutny błąd ex post nie przekroczył w przypadku żadnego OFE 3,7%. Prognozy te charakteryzowały się więc wysoką trafnością (zob. [1]).

Celem niniejszego artykułu jest zaproponowanie innego prognostycznego modelu szeregu czasowego dla wartości jednostek rozrachunkowych OFE oraz porównanie go pod względem trafności uzyskanych prognoz z modelami wyrównywania wykładniczego i funkcjami trendu.

2. Wyznaczenie prognoz wartości jednostek rozrachunkowych OFE oraz ocena ich trafności

Wartość jednostek rozrachunkowych otwartych funduszy emerytalnych w okresie od stycznia 2000 do lipca 2005 charakteryzowała się wzrostową tendencją rozwojową, co potwierdza wykres (zob. rys. 1).

Rys.1

Zaproponowano następujący model prognostyczny dla wartości jednostek rozrachunkowych:

$$y_T^* = y_n + (T - n) \sum_{i=T-k}^{T-1} w_{i-T+k+1} \Delta_i \quad (1)$$

gdzie:

y_T^* – prognoza na okres T ,

n – długość szeregu czasowego (liczba obserwacji, które wykorzystano do budowy modelu)

k – liczba przyrostów wykorzystana do oszacowania pojedynczej prognozy,

Δ_i – przyrost wartości zmiennej prognozowanej w okresie od momentu $i - 1$ do momentu i , obliczony według wzoru:

$$\Delta_i = y_i - y_{i-1} \quad (2)$$

$w_{i-T+k+1}$ – parametr przyporządkowany przyrostowi Δ_i .

Powyższy model nazwano modelem szeregu czasowego z kombinacją liniową przyrostów zmiennej prognozowanej, gdyż prognozą jest suma ostatniej obserwacji oraz pomnożonej przez liczbę okresów w przód, na które wyznaczana jest prognoza, kombinacji liniowej k ostatnich przyrostów. Ostatnia obserwacja pochodzi z okresu najbliższego okresowi prognozowanemu i przez to zawiera najnowsze informacje o zmiennej prognozowanej. W szeregu zachodzą jednak zmiany i są one odzwierciedlane przez przyrosty. k ostatnich przyrostów zostaje pomnożonych przez parametry (skalary) i dodanych do siebie, tworząc w ten sposób kombinację liniową. Jeżeli szereg czasowy zawiera składową systematyczną w postaci stałego poziomu, przyrosty te powinny być w przybliżeniu równe 0 i podobnie ich kombinacja liniowa powinna dać w przybliżeniu wartość 0. Jeżeli szereg czasowy ma składową systematyczną w postaci rosnącej tendencji rozwojowej, wartości przyrostów będą w znaczącej większości większe od 0 i podobnie będzie z ich kombinacją liniową. Jeżeli natomiast tendencja będzie malejąca, przeważająca liczba przyrostów będzie ujemna i ujemną wartość będzie miała także ich kombinacja liniowa. Można przypuszczać, iż model ten będzie pozwalał na wyznaczanie trafnych prognoz w przypadku szeregów czasowych ze stałym poziomem oraz szeregów czasowych zawierających tendencję rozwojową zbliżoną do liniowej, jak to ma miejsce w przypadku wartości JR funduszy emerytalnych (zob. rys. 1). Natomiast ewentualne wykorzystanie prezentowanego modelu do prognozowania szeregów czasowych z nieliniową tendencją rozwojową wymaga przetestowania go na innych danych, co stanowi przedmiot odrębnych badań prowadzonych przez autora.

Przykładowo prognoza na okres $T = 20$, wyznaczona na podstawie szeregu czasowego o długości $n = 18$ przy liczbie przyrostów $k = 3$, jest równa:

$$\begin{aligned} y_{20}^* &= y_{18} + (20 - 18)(w_{17-20+3+1} \cdot \Delta_{17} + w_{18-20+3+1} \cdot \Delta_{18} + w_{19-20+3+1} \cdot \Delta_{19}) = \\ &= 2(w_1 \cdot \Delta_{17} + w_2 \cdot \Delta_{18} + w_3 \cdot \Delta_{19}) \end{aligned}$$

Liczba przyrostów k jest ustalana arbitralnie przez prognostę. Wartości parametrów natomiast są dobierane tak, aby minimalizowany był średni błąd ex post prognoz wygaśłych. Może to być zarówno średni absolutny błąd procentowy prognoz ($MAPE$), jak i średni kwadratowy błąd prognoz ex post (MSE) lub jego pierwiastek ($RMSE$). Jeżeli ustalimy z góry

zbyt wysoką wartość stałej wygładzania k , w procesie optymalizacji parametrów, parametry przyporządkowane najstarszym obserwacjom, czyli opatrzone najmniejszymi indeksami, będą równe 0. Wtedy faktyczna wartość stałej wygładzania będzie pomniejszona o liczbę m pierwszych parametrów, których wartość została ustalona w procesie optymalizacji jako 0. Autor próbował nadać różne ograniczenia na wartości parametrów, np. ich sumę równą 1 lub ich wartość niemniejszą od 0. To jednak sprawiało, iż uzyskiwane prognozy charakteryzowały się mniejszą trafnością.

Prezentowany model wiąże się z „mechanicznym podejściem” do prognozowania. Wymaga on przyjęcia od prognosty postawy pasywnej wobec przyszłości, zakładającej, że zmienna będzie się rozwijać w okresie prognozowanym zgodnie z wykrytą w przyszłości prawidłowością.

Wyznaczając prognozy wartości jednostki rozrachunkowej otwartych funduszy emerytalnych, posłużono się danymi miesięcznymi z okresu od stycznia 2000 do lipca 2005. Na podstawie tych szeregów wyznaczono prognozy na sześć miesięcy wprzód. Oprócz zaprezentowanej metody, predykcji wartości JR dokonano także za pomocą modeli wygładzania wykładniczego oraz funkcji trendu. Dobierając model wygładzania wykładniczego, ze względu na wykres prognozowanej zmiennej, dokonywano wyboru pomiędzy liniowym modelem Holta a modelem wygładzania wykładniczego dla trendu wykładniczego. Ostatecznie prognozy na sześć okresów w przód wyznaczono tym modelem, dla którego otrzymano niższą wartość średniego absolutnego błędu ex post prognoz wygasłych. We wszystkich przypadkach okazało się, że lepszy jest liniowy model Holta, za wyjątkiem OFE PZU, w przypadku którego niższy średni błąd ex post prognoz wygasłych otrzymano stosując model wygładzania wykładniczego z trendem wykładniczym.

Zastosowane modele wygładzania wykładniczego są wyrażone wzorami (zob. [4], [5], [6]):

- model liniowy Holta:

$$F_{t-1} = \alpha y_{t-1} + (1-\alpha)(F_{t-2} + S_{t-2}) \quad (3)$$

$$S_{t-1} = \beta(F_{t-1} - F_{t-2}) + (1-\beta)S_{t-2} \quad (4)$$

$$y_T^* = F_n + (T - n)S_n \quad (5)$$

- model wygładzania wykładniczego z trendem wykładniczym:

$$F_{t-1} = \alpha y_{t-1} + (1-\alpha)(F_{t-2} S_{t-2}) \quad (6)$$

$$S_{t-1} = \beta(F_{t-1}/F_{t-2}) + (1-\beta)S_{t-2} \quad (7)$$

$$y_T^* = F_n S_n^{T-n} \quad (8)$$

gdzie:

F_{t-1} – wygładzona wartość zmiennej prognozowanej na okres lub moment $t-1$,

S_{t-1} – wygładzona wartość przyrostu trendu na okres lub moment $t-1$,

α, β – parametry wygładzania o wartościach z przedziału $[0, 1]$,

y_T^* - prognoza zmiennej prognozowanej na okres lub moment T .

Parametry wygładzania są dobierane tak, aby minimalizowany był średni błąd ex post prognoz wygasłych (*MAPE*, *MSE* lub *RMSE*).

Natomiast przy doborze funkcji trendu, dla każdego OFE oszacowano funkcję liniową oraz funkcję wykładniczą i do wyznaczenia prognoz użyto tej spośród nich, która była lepiej dopasowana do wartości empirycznych. Dopasowanie to mierzono za pomocą standardowego błędu oceny modelu s . Okazało się, że w przypadku każdego funduszu lepszym dopasowaniem charakteryzowała się funkcja wykładnicza. Ogólne wzory szacowanych funkcji są postaci:

- liniowa funkcja trendu:

$$y_t = at + b \quad (9)$$

- wykładnicza funkcja trendu:

$$y_t = ab^t \quad (10)$$

Stosując proponowany model prognostyczny przyjęto liczbę parametrów $k = 10$, a w procesie szacowania wartości parametrów posłużono się kryterium minimalizacji średniego kwadratowego błędu prognoz ex post (*RMSE*). Dla przykładu poniżej przedstawiono wektor parametrów oszacowanego modelu dla wartości jednostki rozrachunkowej AIG OFE:

$$w_{AIG} = \begin{bmatrix} 0,2358 \\ 0,1334 \\ 0,3091 \\ -0,0761 \\ 0,1558 \\ 0,0231 \\ 0,0332 \\ -0,0549 \\ 0,2768 \\ 0,1200 \end{bmatrix}$$

Dla powyższego wektora parametrów otrzymano wartość błędu *RMSE* prognoz wygasłych równą 0,27 zł oraz błąd *MAPE* równy 1,33%.

W tab. 1 przedstawiono miary charakteryzujące jakość oszacowanych modeli i funkcji. W przypadku modelu szeregu czasowego z kombinacją liniową przyrostów oraz modeli wygładzania wykładniczego posłużono się następującymi błędami ex post prognoz wygasłych: średnim błędem procentowym (*MPE*), średnim absolutnym błędem procentowym (*MAPE*) oraz pierwiastkiem błędu średniokwadratowego (*RMSE*). W przypadku funkcji trendu oszacowano współczynnik determinacji R^2 , standardowy błąd oceny modelu s oraz współczynnik wyrazistości w , zwanym też współczynnikiem zmienności losowej.

Tabela 1

Wartości miar, znajdujące się w tab. 1 pozwalają uznać oszacowane modele prognostyczne za dobre. W przypadku proponowanego modelu prognostycznego, wartość średniego absolutnego błędu ex post prognoz wygasłych była dla każdego funduszu znacząco niższa od 2%. Błąd ten był także niski dla modeli wygładzania wykładniczego, ale w przypadku wszystkich OFE jego wartość była niższa, gdy prognozy szacowano modelem szeregu czasowego z kombinacją liniową przyrostów. Oszacowane wykładnicze funkcje trendu charakteryzowały się dobrym dopasowaniem do danych empirycznych, gdyż współczynnik wyrazistości w dla każdego funduszu był niższy od 6%. Wysokie są też współczynniki determinacji oszacowanych modeli, nie niższe niż 0,93.

Miarą, za pomocą której można próbować porównać jakość wszystkich oszacowanych modeli, jest pierwiastek średniego kwadratowego błędu prognoz ex post (*RMSE*) dla proponowanego modelu prognostycznego oraz modeli wygładzania wykładniczego, a dla

funkcji trendu standardowy błąd oceny modelu s , bowiem te dwie miary są porównywalne ([2], s. 49). Należy jednak pamiętać, iż w przypadku modelu szeregu czasowego z kombinacją liniową przyrostów oraz modeli wygładzania wykładniczego błąd $RMSE$ jest liczony dla prognoz i wartość teoretyczna, czyli wynikająca z modelu, na dany okres jest szacowana bez wykorzystania informacji o wartości empirycznej zmiennej prognozowanej z tego okresu. Natomiast standardowy błąd oceny modelu jest standardowym odchyleniem reszt modelu, będących różnicą pomiędzy wartościami empirycznymi a wartościami teoretycznymi zmiennej. Wartość teoretyczna zmiennej na dany okres obliczana jest z modelu, przy którego szacowaniu wykorzystano informację o wartości empirycznej zmiennej z tego samego okresu. Stosując mimo to tę miarę można zauważyć, iż najlepsze dopasowanie do danych empirycznych uzyskano w przypadku proponowanego modelu prognostycznego, a najgorsze w przypadku wykładniczych funkcji trendu.

Każdym z oszacowanych modeli wyznaczono prognozy na sześć okresów wpród, na miesiące od sierpnia 2005 do stycznia 2006. Trafność wyznaczonych prognoz oceniono za pomocą średniego procentowego błędu prognoz *ex post* MPE oraz średniego absolutnego błędu prognoz *ex post* $MAPE$. Obliczenia zawarto w tab. 2.

Tabela 2

Z tabeli 2 wynika, że zdecydowanie najwyższą trafnością charakteryzują się prognozy wartości jednostki rozrachunkowej OFE, wyznaczone za pomocą proponowanego modelu szeregu czasowego z kombinacją liniową przyrostów, dla których błąd $MAPE$ zawierał się w przedziale od 1,05% do 1,88%. Dwie pozostałe metody, chociaż także pozwoliły na wyznaczenie dobrych prognoz, to obliczone dla nich błędy $MAPE$ zawierały się w przedziale od 1,92% do 11,61%. Prognozy wyznaczone zaproponowanym modelem charakteryzowały się także zdecydowanie niższym stopniem obciążoności, aniżeli prognozy wyznaczone modelami wygładzania wykładniczego lub funkcjami trendu, o czym świadczą wartości błędów MPE . Błąd ten tylko w przypadku jednego funduszu (OFE Bankowy) przekroczył wartość 0,25%, można więc uznać wyznaczone prognozy za nieobciążone.

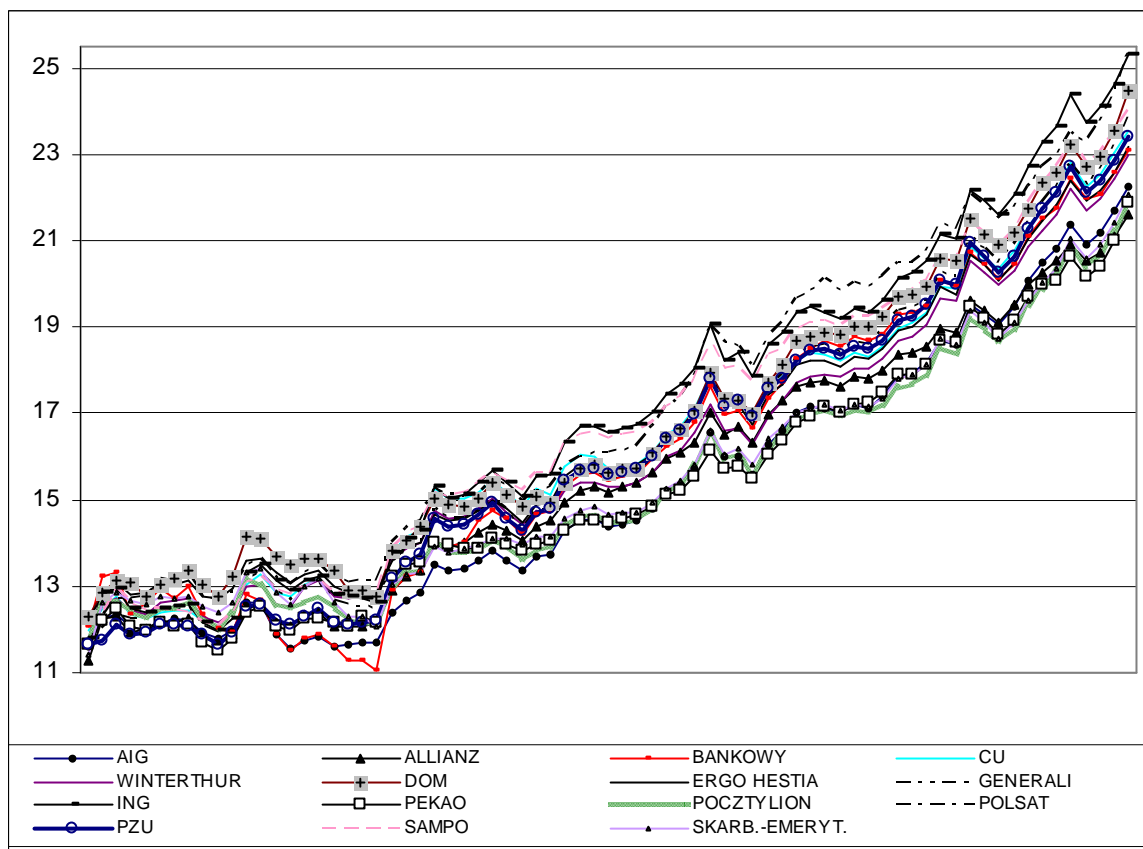
3. Podsumowanie

Zaproponowany model prognostyczny charakteryzuje się bardzo dużą elastycznością i szybko reaguje na zmiany w kształtowaniu się wartości prognozowanej zmiennej. Brak ograniczeń dotyczących parametrów tego modelu pozwala na taką kombinację liniową k

ostatnich przyrostów, aby średni błąd ex post prognoz wygasłych został zminimalizowany. Zaletą tego modelu jest także jego prostota i łatwość szacowania, a uzyskane prognozy charakteryzują się wysoką trafnością. Zaznaczyć jednak należy, iż niniejszy artykuł miał na celu zastosowanie proponowanego modelu wyłącznie do prognozowania wartości jednostek rozrachunkowych otwartych funduszy emerytalnych i sformułowanych na tej podstawie wniosków dotyczących jego własności nie można zbyt uogólniać. Aby ocenić możliwości szerszego stosowania prezentowanego modelu oraz określić jego wady i zalety, należy przetestować go na wielu innych szeregach czasowych, charakteryzujących się stałym poziomem i tendencją rozwojową, zarówno liniową, jak i nieliniową. Stanowi to przedmiot odrębnych badań prowadzonych przez autora.

Literatura

- [1] Chybalski F., *Prognozowanie efektów działalności inwestycyjnej otwartych funduszy emerytalnych w Polsce*, „Zarządzanie finansami firm – teoria i praktyka”, tom 1, Prace Naukowe Akademii Ekonomicznej we Wrocławiu, Wydawnictwo AE, Wrocław 2004
- [2] Cieślak M. (red.), *Prognozowanie gospodarcze. Metody i zastosowanie*, PWN, Warszawa 2001
- [3] Dzik W., *O wyższości IRR nad JR*, Parkiet, 17 stycznia 2001, s. 11
- [4] Hyndman R. J., Koehler A. B., Snyder R. D., Grose S., *A state space framework for automatic forecasting using exponential smoothing methods*, “International Journal of Forecasting”, 3 / 2002, s. 439 – 454
- [5] Makridakis S, Wheelwright S.C., *Forecasting methods for management*, John Wiley & Sons 1989
- [6] Snyder R. D., Koehler A. B., Hyndman R. J., Ord J. K., *Exponential smoothing for inventory control: means of variance of lead demand*, Working paper, 3 / 2002, Monash University, Australia



Rys. 1. Wartość jednostek rozrachunkowych OFE (w zł) w okresie od stycznia 2000 do lipca 2005

Źródło: Komisja Nadzoru Ubezpieczeń i Funduszy Emerytalnych

Tabela 1. Miary charakteryzujące jakość oszacowanych modeli prognostycznych

| OFE | Model szeregu czasowego z kombinacją liniową przyrostów – błędy prognoz wygasłych | | | Model wygładzania wykładniczego – błędy prognoz wygasłych | | | Funkcja trendu – miary dopasowania modelu | | |
|-----------|---|--------------------|---------------------|---|--------------------|---------------------|--|---------------|--------------|
| | <i>MPE</i> (%) | <i>MAPE</i> (%) | <i>RMSE</i> (zł) | <i>MPE</i> (%) | <i>MAPE</i> (%) | <i>RMSE</i> (zł) | R^2 | <i>s</i> (zł) | <i>w</i> (%) |
| AIG | 0,19 | 1,33 | 0,27 | -0,11 | 1,75 | 0,32 | 0,94 | 0,61 | 4,14 |
| Allianz | 0,02 | 1,28 | 0,28 | -0,08 | 1,54 | 0,30 | 0,98 | 0,42 | 2,76 |
| Bankowy | 0,53 | 1,86 | 0,39 | -0,15 | 2,30 | 0,45 | 0,93 | 0,80 | 5,18 |
| CU | 0,16 | 1,58 | 0,32 | -0,09 | 1,78 | 0,35 | 0,98 | 0,44 | 2,78 |
| Dom | 0,03 | 1,64 | 0,33 | -0,10 | 1,85 | 0,36 | 0,95 | 0,59 | 3,66 |
| Ergo H. | 0,09 | 1,53 | 0,31 | -0,09 | 1,65 | 0,32 | 0,96 | 0,49 | 3,07 |
| Generali | 0,06 | 1,68 | 0,34 | -0,10 | 1,87 | 0,36 | 0,98 | 0,46 | 2,90 |
| ING | 0,17 | 1,78 | 0,38 | -0,12 | 2,02 | 0,41 | 0,97 | 0,53 | 3,20 |
| Pekao | 0,23 | 1,31 | 0,27 | -0,89 | 1,61 | 0,30 | 0,96 | 0,48 | 3,23 |
| Pocztyl. | 0,10 | 1,59 | 0,30 | -0,09 | 1,81 | 0,33 | 0,95 | 0,52 | 3,53 |
| Polsat | 0,19 | 1,50 | 0,33 | -0,07 | 1,67 | 0,34 | 0,97 | 0,51 | 3,10 |
| PZU | 0,22 | 1,57 | 0,33 | 0,16 | 1,67 | 0,33 | 0,98 | 0,45 | 2,88 |
| Sampo | 0,23 | 1,44 | 0,31 | -0,10 | 1,64 | 0,33 | 0,98 | 0,44 | 2,73 |
| Sk-Em | 0,12 | 1,65 | 0,31 | -0,08 | 1,77 | 0,34 | 0,94 | 0,53 | 3,56 |
| Winterth. | 0,13 | 1,53 | 0,30 | -0,09 | 1,71 | 0,33 | 0,97 | 0,46 | 2,97 |

Źródło: obliczenia własne

Tabela 2. Ocena trafności prognoz

| OFE | Model szeregu czasowego z kombinacją liniową przyrostów | | Model wygładzania wykładniczego | | Funkcja trendu | |
|-----------|---|-----------------|---------------------------------|-----------------|----------------|-----------------|
| | <i>MPE</i> (%) | <i>MAPE</i> (%) | <i>MPE</i> (%) | <i>MAPE</i> (%) | <i>MPE</i> (%) | <i>MAPE</i> (%) |
| AIG | 1,43 | 1,64 | 2,00 | 2,00 | 7,22 | 4,86 |
| Allianz | 0,15 | 1,05 | 3,03 | 3,03 | 1,57 | 1,57 |
| Bankowy | 0,88 | 1,28 | -2,98 | 2,98 | 3,58 | 3,58 |
| CU | 0,63 | 1,26 | -4,18 | 4,18 | 4,86 | 4,86 |
| Dom | -0,03 | 1,50 | -6,93 | 6,93 | 6,41 | 6,41 |
| Ergo H. | 0,00 | 1,24 | -2,70 | 2,70 | 5,59 | 5,59 |
| Generali | 0,36 | 1,29 | -4,89 | 4,89 | 4,51 | 4,51 |
| ING | 0,68 | 1,40 | -11,61 | 11,61 | 4,16 | 4,16 |
| Pekao | -1,27 | 1,88 | 4,38 | 4,38 | 4,82 | 4,82 |
| oczty. | 0,91 | 1,46 | 4,88 | 4,88 | 7,33 | 7,33 |
| Polsat | 1,22 | 1,54 | -9,02 | 9,02 | 2,86 | 2,86 |
| PZU | 0,23 | 1,30 | -4,28 | 4,28 | 2,99 | 2,99 |
| Sampo | 0,22 | 1,19 | -6,79 | 6,79 | 2,33 | 2,33 |
| Sk-Em | 1,53 | 1,69 | 4,18 | 4,18 | 7,46 | 7,46 |
| Winterth. | 0,47 | 1,23 | -1,68 | 1,92 | 5,72 | 5,72 |

Źródło: obliczenia własne