

Krótkoterminowe prognozowanie chaotycznych szeregów czasowych

1. Wstęp

Pozornie przypadkowy charakter chaotycznych szeregów czasowych może prowadzić do błędnego wniosku, że są one nieregularne i w efekcie nieprzewidywalne. Jednakże w istocie istnieje w nich ukryty porządek, dzięki któremu możliwe jest efektywne prognozowanie ich ewolucji w krótkim horyzoncie czasowym.

W niniejszej pracy przedstawiono wybrane metody stosowane do krótkoterminowego prognozowania szeregów chaotycznych; liniową i kwadratową lokalną aproksymację wielomianową oraz metodę Najbliższych Sąsiadów wraz z autorskim pomysłem pewnej jej modyfikacji. Metody te nie wymagają identyfikacji wzoru analitycznego funkcji generującej system ani estymacji jej parametrów. Dokładność wyznaczonych prognoz zależy od wielu czynników, w tym liczby obserwacji, wymiaru atraktora i wykładników Lapunowa systemu, horyzontu prognozy oraz poziomu zakłóceń losowych.

Skuteczność prezentowanych metod zweryfikowano w zastosowaniu do wygenerowanych znanych szeregów chaotycznych o różnej długości. Dokonano próby krótkookresowego prognozowania rzeczywistych ekonomicznych szeregów czasowych indeksu WIG oraz jego stóp zmian na podstawie dziennych obserwacji z okresu 03.10.1994-20.08.2001 r.

2. Chaotyczne szeregi czasowe

Rozważmy nieliniowy system dynamiczny (X, f) , gdzie $X \subset R^n$ jest przestrzenią stanów, zaś $f : X \rightarrow X$ jest odwzorowaniem generującym ich ewolucję w czasie. W systemach z czasem ciągłym odwzorowanie f zwykle definiuje się poprzez równanie różniczkowe zwyczajne w postaci normalnej:

$$\frac{dS}{dt} = f(S), \quad (2.1)$$

dla dowolnego $S \in X$, zaś w przypadku systemu dynamicznego z czasem dyskretnym przez zależność rekurencyjną:

$$S_{t+1} = f(S_t), \quad t = 0, 1, \dots \quad (2.2)$$

gdzie $S_t, S_{t+1} \in X$ są stanami systemu odpowiednio w momencie t i $t+1$. Operacyjna definicja chaosu głosi, że dysypatywny system dynamiczny (X, f)

jest chaotyczny, gdy w spektrum Lapunowa atraktora tego systemu największy wykładnik Lapunowa jest dodatni (np. Zawadzki [16], Lorenz [14])¹.

Przedmiotem analizy naturalnych, w tym i ekonomicznych systemów są szeregi czasowe będące jednowymiarowymi obrazami stanów systemu, tzn.

$$x_t = h(S_t), t=0,1, \dots$$

gdzie $h: X \rightarrow R$ jest pewnym nieznanym odwzorowaniem nazywanym urządzeniem pomiarowym, funkcją agregującą lub obserwabłą (Zawadzki [16]). Na bazie twierdzenia Takensa o zanurzaniu stworzono metodę opóźnień, która umożliwia rekonstrukcję przestrzeni stanów badanego systemu w oparciu o szereg (x_t) , a przez to jest punktem wyjścia wielu metod identyfikacji chaosu deterministycznego w procesach rzeczywistych (Takens [15]). W metodzie opóźnień przedmiotem badania jest ciąg wektorów o m współrzędnych postaci:

$$\hat{x}_t^m = (x_t, x_{t-lag}, \dots, x_{t-(m-1)lag}),$$

dla każdego naturalnego $t \in \langle (m-1) \cdot lag + 1, N \rangle$, utworzonych z szeregu czasowego składającego się z N obserwacji. Wektory \hat{x}_t^m nazywane są m -historiami lub wektorami opóźnień, liczba m – wymiarem zanurzenia, zaś lag – opóźnieniem czasowym. Z twierdzenia Takensa wynika (por. Jimenez i in. [12]), że dla dowolnych $t, T \in N$ istnieje funkcja $g_T: R^m \rightarrow R$, dla której:

$$x_{t+T} = g_T(x_t, x_{t-lag}, \dots, x_{t-(m-1)lag}) \equiv g_T(\hat{x}_t^m). \quad (2.3)$$

Równość (2.3) opisuje zależność deterministyczną pomiędzy kolejnymi obserwacjami szeregu chaotycznego, której źródłem jest funkcyjny związek między stanami generującego go systemu. Determinizm oznacza, że przebieg procesu zarówno w przyszłości, jak i w przeszłości jest jednoznacznie określony przez stan tego procesu w chwili obecnej (Arnold [1]). Skoro system chaotyczny z definicji jest deterministyczny, wydawać by się mogło, że nic nie stoi na przeszkodzie, aby znając prawa ewolucji i wartości stanów przeszłych, prognozować niemal bezbłędnie jego przyszłość. Jednakże w praktyce nie jest to możliwe. Analizując systemy naturalne nie jesteśmy w stanie dokładnie określić wartości zmiennych. Nieuniknione błędy pomiaru oraz zaokrąglenia wykonywane przez komputer wpływają na utratę precyzji obliczeń. Systemy chaotyczne cechują się dużą wrażliwością na zmianę warunków początkowych, która powoduje, że wszelkie, nawet wydawałoby się, nieistotnie małe niedokładności powiększają się w kolejnych iteracjach w tempie wykładniczym. W konsekwencji również przewidywalność systemu zmniejsza się wykładniczo wraz z wydłużaniem się horyzontu prognozy.

W przypadku naturalnych procesów nieznaną jest funkcja g_T z równania (2.3) i należy liczyć się z faktem, że potencjalnie może mieć ona bardzo skomplikowany nieliniowy charakter. Warto jednak podkreślić, że zaprezentowane w artykule metody prognozowania nie wymagają identyfikacji jej wzoru analitycznego, opierają się natomiast na wykorzystaniu samego faktu występowania związku deterministycznego (2.3). Efektywność przedstawionych

¹ Inne definicje chaosu można znaleźć m.in. w: Li, Yorke [13], Garrido [11], Devaney [5], Eckmann, Ruelle [7].

metod zależy od długości analizowanego szeregu, wymiaru atraktora i entropii systemu, horyzontu prognozy T oraz wielkości zakłóceń losowych (por. Farmer, Sidorowich [8]).

3. Metody krótkoterminowego prognozowania szeregów chaotycznych

3.1. Metoda Analogowa (Najbliższych Sąsiadów)

Punktem wyjścia metody analogowej, często zwanej w literaturze przedmiotu metodą Najbliższych Sąsiadów - NS jest wyznaczenie spośród m -historii (\hat{x}_t^m) , k najbliższych (w sensie zadanej m -wymiarowej metryki) sąsiadów wektora \hat{x}_N^m , gdzie N jest długością badanego szeregu. Oznaczmy znalezione m -historie przez $\hat{x}_{l_1}^m, \hat{x}_{l_2}^m, \dots, \hat{x}_{l_k}^m$. Idea metody NS oparta jest na założeniu, że w systemach deterministycznych, zbliżone do siebie m -historie powinny w podobny sposób ewoluować w czasie. Mówiąc ściśle, przyjmuje się, że z założenia $\hat{x}_N^m \approx \hat{x}_{l_i}^m$ wynika, że również $g_T(\hat{x}_N^m) \approx g_T(\hat{x}_{l_i}^m)$, co w konsekwencji oznacza, że $x_{N+T} \approx x_{l_i+T}$. W metodzie NS za prognozę \tilde{x}_{N+T} przyjmuje się średnią ważoną liczb $\{x_{l_1+T}, x_{l_2+T}, \dots, x_{l_k+T}\}$, z wagami dobranymi tak, by bliżsi sąsiedzi mieli na nią większy wpływ, tzn.:

$$\tilde{x}_{N+T} = \sum_{i=1}^k w(\|\hat{x}_N^m - \hat{x}_{l_i}^m\|) \cdot x_{l_i+T}, \quad (2.4)$$

gdzie $\|\cdot\|$ oznacza dowolną, przyjętą normę w R^m , zaś $w: R \rightarrow R$ jest dowolną funkcją malejącą spełniającą warunki:

1. $w(\|\hat{x}_N^m - \hat{x}_{l_i}^m\|) > 0$, dla każdego $i=1, 2, \dots, k$
2. $\sum_{i=1}^k w(\|\hat{x}_N^m - \hat{x}_{l_i}^m\|) = 1$.

Przykładowo, proponuje się przyjęcie następujących wag „wykładniczych” (Finkenstädt, Kuhbier [9]):

$$w(\|\hat{x}_N^m - \hat{x}_{l_i}^m\|) = \frac{e^{-\|\hat{x}_N^m - \hat{x}_{l_i}^m\|}}{\sum_{i=1}^k e^{-\|\hat{x}_N^m - \hat{x}_{l_i}^m\|}}. \quad (2.5)$$

Warto podkreślić, że w metodzie Najbliższych Sąsiadów nie dokonuje się aproksymacji funkcji g_T , co oczywiście wpływa pozytywnie na efektywność czasową obliczeń. Prostota obliczeń, a nie dokładność otrzymanych prognoz, jest zatem główną zaletą metody.

Praktycznym problemem, który pojawia się w trakcie obliczeń komputerowych, przy wagach „wykładniczych” jest zaokrąglanie do zera poszczególnych wartości e^{-x} dla dużych x . W konsekwencji suma wag może dać wartość różną od jedności. Problem ten można ominąć przekształcając oryginalny szereg, poprzez podzielenie go przez ustaloną odpowiednio dużą liczbę. Otrzymana prognoza teoretycznie powinna być niezależna od przyjętego

dzielnika, jednakże w praktyce, ze względu na dokonywane przez komputer zaokrąglenia, mogą istnieć różnice w otrzymywanych rezultatach.

W niniejszej pracy zaproponowano pewną modyfikację metody NS, nazwaną metodą rangową - NSR. Jej pierwszym etapem jest uporządkowanie k najbliższych (w sensie zadanej m -wymiarowej metryki $\| \cdot \|$) sąsiadów wektora \hat{x}_N^m i oznaczenie ich przez $\hat{x}_{l_1}^m, \hat{x}_{l_2}^m, \dots, \hat{x}_{l_k}^m$, gdzie $\hat{x}_{l_1}^m$ i $\hat{x}_{l_k}^m$ są odpowiednio najbliższym i najdalszym sąsiadem \hat{x}_N^m . Następnie wektorom $\hat{x}_{l_i}^m$ przyporządkowuje się wagi $R(i) = \frac{2 \cdot (k+1-i)}{k \cdot (k+1)}$, $i=1,2, \dots, k$, w oparciu o które wyznacza się prognozę:

$$\tilde{x}_{N+T} = \sum_{i=1}^k R(i) \cdot x_{l_i+T} \quad (2.6)$$

Wagi $R(i)$ nie są wprost funkcjami odległości $\|\hat{x}_N^m - \hat{x}_{l_i}^m\|$, lecz zależą od tego, którym w kolejności najbliższym sąsiadem wektora \hat{x}_N^m jest wektor $\hat{x}_{l_i}^m$. Są one tak dobrane, aby $R(1) > R(2) > \dots > R(k) > 0$, $\sum_{i=1}^k R(i) = 1$ oraz by różnica $R(i) - R(i+1)$ była stała i wynosiła $R(k)$ dla każdego i .

Zaletami metody NSR są mniejsza złożoność obliczeniowa oraz niewystępowanie wspomnianego wcześniej problemu zaokrąglania do zera. Oczywiście metoda rangowa jest pewnym uproszczeniem metody NS, lecz jak wykazały wyniki badań, zazwyczaj prowadziła do prognoz nawet nieco lepszych niż w przypadku wag „wykładniczych”.

3.2. Lokalna aproksymacja liniowa

W oparciu o zależność (2.3) prognozowanie szeregu w momencie $N+T$ sprowadza się do wyznaczenia wartości $\tilde{x}_{N+T} = \tilde{g}_T(x_N, x_{N-lag}, \dots, x_{N-(m-1)lag})$, gdzie \tilde{g}_T jest aproksymantą funkcji g_T . Za \tilde{g}_T można przyjąć funkcję liniową postaci $\tilde{g}_T(x_1, x_2, \dots, x_m) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m$, gdzie m jest ustalonym wymiarem zanurzenia. Estymacja parametrów α_i odbywa się za pomocą KMNK w oparciu o pary postaci $(\hat{x}_{l_i}^m, x_{l_i+T})$, $i=1,2, \dots, k$, gdzie $\{\hat{x}_{l_i}^m\}$ jest k elementowym zbiorem m -historii będących najbliższymi sąsiadami wektora \hat{x}_N^m . Nietrudno zauważyć, że gdy $k=N-(m-1)lag$ tzn. przy uwzględnieniu wszystkich dostępnych m -historii, przeprowadzona procedura jest regresją liniową.

3.3. Lokalna aproksymacja kwadratowa

Metoda lokalnej aproksymacji kwadratowej jest naturalnym uogólnieniem aproksymacji liniowej, poprzez przyjęcie za \tilde{g}_T funkcji wielomianowej m zmiennych stopnia drugiego:

$$\tilde{g}_T(x_1, x_2, \dots, x_m) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m + \sum_{1 \leq i < j \leq m} \alpha_{i,j} x_i x_j$$

Oczekuje się wyższości lokalnej aproksymacji kwadratowej nad aproksymacją wielomianową stopnia pierwszego (liniową). Jednakże badania wskazują, że różnica ta zdaje się maleć wraz z poziomem komplikacji systemu i długością szeregu (Castagli [4]).

4. Krótkoterminowe prognozowanie ewolucji wybranych szeregów czasowych

4.1. Prognozowanie dla $T=1$

Celem przeprowadzonego badania było zweryfikowanie możliwości efektywnego prognozowania szeregów czasowych przy użyciu zaprezentowanych metod. W pierwszej kolejności prognozowaniu poddano znane szeregi chaotyczne; generowane przez odwzorowanie Henona, odwzorowanie logistyczne i system Lorenza. System Lorenza jest przykładem ciągłego chaotycznego układu dynamicznego. Z kolei odwzorowania logistyczne i Henona są reprezentatywnymi przykładami, odpowiednio, jedno- i dwuwymiarowych systemów z czasem dyskretnym. Dla każdego systemu badaniu poddano szeregi składające się z 320 oraz 1715 obserwacji. Następnie wyznaczono prognozy dla poziomów indeksu WIG oraz logarytmicznych stóp jego zmian w oparciu o dane z okresu 03.10.1994-20.08.2001. Dodatkowo każdy z analizowanych szeregów poddano prognozowaniu przy zastosowaniu dopasowanego modelu ARIMA.

Prognozowanie przebiegało według następującej procedury szczegółowej: Każdy z analizowanych szeregów długości N został podzielony na dwie części, składające się odpowiednio z N_1 i N_2 obserwacji ($N_1 + N_2 = N$). Dla każdej obserwacji z drugiej próbki wyznaczono prognozę \tilde{x}_{N_1+i} dla $i=1, 2, \dots, N_2$. Do tego celu wykorzystano k najbliższych sąsiadów wektora $\hat{x}_{N_1+i-1}^m = (x_{N_1+i-1}, \dots, x_{N_1+i-1-(m-1) \cdot lag})$ skonstruowanych w oparciu o obserwacje (x_t) dla $t=1, 2, \dots, N_1+i-1$. Do oceny dokładności prognozy wykorzystano bezwzględny błąd predykcji ex-post zadany wzorem:

$$\sigma = \sqrt{\frac{1}{N_2} \cdot \sum_{t=N_1+1}^{N_1+N_2} (x_t - \tilde{x}_t)^2},$$

oraz względny błąd prognozy postaci:

$$\sigma' = \frac{\sigma}{\sigma_x} \cdot 100\%,$$

gdzie σ_x jest odchyleniem standardowym szeregu (x_t) dla $t=1, 2, \dots, N_1$.

Przedstawione metody wymagają doboru trzech parametrów: wymiaru zanurzenia – m , opóźnienia czasowego – lag oraz liczby najbliższych sąsiadów – k . W niniejszej pracy dokonano prognoz dla różnych wartości parametrów; $lag=1, 2, \dots, 5$, $m=1, 2, \dots, 15$ (za wyjątkiem aproksymacji kwadratowej, gdzie ze względu na złożoność obliczeniową przyjęto $m=1, 2, 3, 4$) i dla wszystkich

potencjalnie możliwych k . W celu znalezienia najbliższych sąsiadów zastosowano metrykę euklidesową. Do wyznaczenia prognoz metodami aproksymacji wielomianowej została wykorzystana funkcja Excela *reglinw()*.

W tabelach 1-8 zaprezentowano otrzymane wartości parametrów dające najdokładniejsze prognozy dla analizowanych szeregów czasowych oraz odpowiadające im błędy prognoz σ oraz σ' .

Odwzorowanie Henona

Odwzorowanie Henona generuje system chaotyczny (R^2, H) , gdzie $H: R^2 \rightarrow R^2$ określone jest następująco:

$$H(x_t, y_t) = (x_{t+1}, y_{t+1}) = (1 - 1,4x_t^2 + y_t; 0,3x_t).$$

Badaniu poddano szeregi złożone z pierwszych współrzędnych stanów systemu Henona dla $(x_0, y_0) = (0,9; 0,9)$.

Tabela 1. Najmniejsze błędy prognozy dla szeregu Henona przy $N_1=300, N_2=20$

Metoda	σ	σ'	Optymalne parametry
Metoda NSR	0,027	3,66%	$m=2, lag=1, k=2$
Metoda NS	0,028	3,80%	$m=2, lag=1, k=2$
Liniowa aproksymacja	0,002	0,24%	$m=4, lag=1, k=5$
Aproksymacja kwadratowa	$4,3 \cdot 10^{-15}$	$5,9 \cdot 10^{-13}\%$	$m=2, lag=1, k=281$
ARIMA	0,639	88,04%	ARMA(4,2)

Zródło: obliczenia własne

Tabela 2. Najmniejsze błędy prognozy dla szeregu Henona przy $N_1=1650, N_2=65$

Metoda	σ	σ'	Optymalne parametry
Metoda NSR	0,007	0,89%	$m=2, lag=1, k=6$
Metoda NS	0,008	1,02%	$m=2, lag=1, k=1$
Liniowa aproksymacja	$2,8 \cdot 10^{-5}$	0,0038%	$m=4, lag=1, k=7$
Aproksymacja kwadratowa	$9,5 \cdot 10^{-15}$	$1,3 \cdot 10^{-12}\%$	$m=2, lag=1, k=1603$
ARIMA	0,642	87,39%	ARMA(2,6)

Zródło: obliczenia własne

Odwzorowanie logistyczne

Odwzorowanie logistyczne generuje system chaotyczny $((0,1), f)$, gdzie $f(x_t) \equiv x_{t+1} = 4 \cdot x_t \cdot (1 - x_t)$. Prognozowaniu poddano szeregi wygenerowane dla $x_0=0,7$.

Tabela 3. Najmniejsze błędy prognozy dla odwzorowania logistycznego przy $N_1=300, N_2=20$

Metoda	σ	σ'	Optymalne parametry
Metoda NSR	0,003	0,86%	$m=1, k=3$
Metoda NS	0,003	0,92%	$m=1, k=3$
Liniowa aproksymacja	$2,4 \cdot 10^{-5}$	0,007%	$m=2, lag=1, k=7$
Aproksymacja kwadratowa	$1,1 \cdot 10^{-14}$	$3 \cdot 10^{-12}\%$	$m=2, lag=4, k=146$
ARIMA	0,360	102,54%	Biały szum

Źródło: obliczenia własne

Tabela 4. Najmniejsze błędy prognozy dla odwzorowania logistycznego przy $N_1=1650$, $N_2=65$

Metoda	σ	σ'	Optymalne parametry
Metoda NSR	0,001	0,18%	m=1, k=2
Metoda NS	0,001	0,21%	m=1, k=2
Liniowa aproksymacja	$2,5 \cdot 10^{-7}$	$7 \cdot 10^{-5}\%$	m=2, lag=1, k=6
Aproksymacja kwadratowa	$2,7 \cdot 10^{-14}$	$7,6 \cdot 10^{-12}\%$	m=1, k=1649
ARIMA	0,330	93,40%	Biały szum

Źródło: obliczenia własne

System Lorenza:

System Lorenza jest zdefiniowany przez następujący układ równań różniczkowych:

$$\frac{dx}{dt} = 16 \cdot (y - x)$$

$$\frac{dy}{dt} = -x \cdot z + 45,92 \cdot x - y$$

$$\frac{dz}{dt} = x \cdot y - 4 \cdot z$$

Badaniu poddano szeregi postaci $x_t = x(t \cdot 0,01)$, dla $t=1,2, \dots$ przy zadanych warunkach początkowych $(x(0), y(0), z(0)) = (1,1,1)$.

Tabela 5. Najmniejsze błędy prognozy dla szeregu Lorenza przy $N_1=300$, $N_2=20$

Metoda	σ	σ'	Optymalne parametry
Metoda NSR	0,154	1,705%	m=2, lag=1, k=1
Metoda NS	0,154	1,705%	m=2, lag=1, k=1
Liniowa aproksymacja	0,001	0,0096%	m=11, lag=3, k=18
Aproksymacja kwadratowa	0,001	0,0091%	m=4, lag=1, k=98
ARIMA	0,004	0,042%	ARMA(4,5)

Źródło: obliczenia własne

Tabela 6. Najmniejsze błędy prognozy dla szeregu Lorenza przy $N_1=1650$, $N_2=65$

Metoda	σ	σ'	Optymalne parametry
Metoda NSR	0,088	0,690%	m=10, lag=3, k=3
Metoda NS	0,090	0,704%	m=8, lag=3, k=3
Liniowa aproksymacja	0,001	0,0086%	m=14, lag=1, k=29
Aproksymacja kwadratowa	0,001	0,0083%	m=4, lag=1, k=23
ARIMA	0,004	0,034%	ARMA(5,4)

Źródło: obliczenia własne

Uwaga: w celu wyznaczenia prognozy metodą NS szereg został podzielony przez 10.

WIG

W badaniu uwzględniono 1716 dziennych obserwacji poziomu indeksu WIG z okresu 03.10.1994-20.08.2001.

Tabela 7. Najmniejsze błędy prognozy dla szeregu WIG przy $N_1=1650$, $N_2=66$

Metoda	σ	σ'	Optymalne parametry
Metoda NSR	177,55	4,61%	$m=11$, $lag=3$, $k=1$
Metoda NS	177,55	4,61%	$m=11$, $lag=3$, $k=1$
Liniowa aproksymacja	169,68	4,40%	$m=4$, $lag=5$, $k=47$
Aproksymacja kwadratowa	175,70	4,56%	$m=3$, $lag=4$, $k=628$
ARIMA	177,85	4,62%	Błądzenie przypadkowe

Zródło: obliczenia własne

Uwaga: w celu wyznaczenia prognozy metodą NS szereg został podzielony przez 1 000.

Logarytmiczne stopy zmian WIG

Podstawą badania był ten sam okres, jak w przypadku szeregu poziomów indeksu, co dało 1715 logarytmicznych stóp zmian WIG.

Tabela 8. Najmniejsze błędy prognozy dla szeregu stóp zmian WIG przy $N_1=1650$, $N_2=65$

Metoda	σ	σ'	Optymalne parametry
Metoda NSR	0,0123	64,03%	$m=14$, $lag=2$, $k=24$
Metoda NS	0,0122	63,26%	$m=14$, $lag=2$, $k=12$
Liniowa aproksymacja	0,0124	64,48%	$m=10$, $lag=3$, $k=96$
Aproksymacja kwadratowa	0,0124	64,38%	$m=2$, $lag=3$, $k=93$
ARIMA	0,0131	67,79%	Biały szum

Zródło: obliczenia własne

Wyniki badań wskazują, że zaprezentowane metody umożliwiają skuteczne prognozowanie ewolucji wygenerowanych chaotycznych szeregów czasowych. Najlepsza spośród nich jest lokalna aproksymacja kwadratowa, jednakże jest ona jednocześnie najbardziej czasochłonna. Bardzo duża dokładność prognozy, jaką daje ta metoda w zastosowaniu do przebadanych obu systemów dyskretnych wynika z faktu, że są one generowane przez funkcje wielomianowe stopnia drugiego (dla funkcji logistycznej $m=1$, dla odwzorowania Henona $m=2$, $lag=1$). Otrzymany błąd prognozy jest więc w tych przypadkach efektem niedokładności estymacji ich współczynników.

Gorsze, choć również zadowalające, rezultaty osiągnięto przy zastosowaniu metody Najbliższych Sąsiadów, której istotną zaletą pozostaje dużo mniejsza złożoność obliczeniowa. Spośród rozważonych wariantów metody, nieco lepsze wyniki zdaje się dawać przyjęcie zaproponowanych w niniejszej pracy rang.

Prognozowanie oparte na identyfikacji procesu ARIMA dla chaotycznych systemów dyskretnych okazało się być mniej dokładne od pozostałych metod, natomiast w przypadku systemu Lorenza dało prognozy gorsze jedynie od metod aproksymacji wielomianowej. W zastosowaniu do rzeczywistych danych, wszystkie przedstawione metody dały zbliżone rezultaty.

4.2. Prognozowanie dla $T>1$ metodą iteracyjną

Wyznaczenie prognozy \tilde{x}_{N_1+T} dla dłuższego horyzontu czasowego ($T>1$) może odbywać się według dwóch alternatywnych procedur. Pierwsza polega na bezpośrednim oszacowaniu \tilde{g}_T i obliczeniu $\tilde{x}_{N_1+T} = \tilde{g}_T(x_{N_1}, x_{N_1-lag}, \dots, x_{N_1-(m-1)lag})$. Druga jest metodą iteracyjną polegającą na wyznaczeniu kolejnych prognoz $\tilde{x}_{N_1+1}, \tilde{x}_{N_1+2}, \dots, \tilde{x}_{N_1+T}$ w oparciu o kolejno szacowane \tilde{g}_1 . W metodzie iteracyjnej predyktorem \tilde{g}_T jest T -krotne złożenie funkcji \tilde{g}_1 , tzn. $\tilde{g}_T = (\tilde{g}_1)^T$. Dokładność metody bezpośredniej opisuje zależność:

$$\sigma'(\tilde{g}_T) = e^{T(s+1)h} \sigma'(\tilde{g}_1),$$

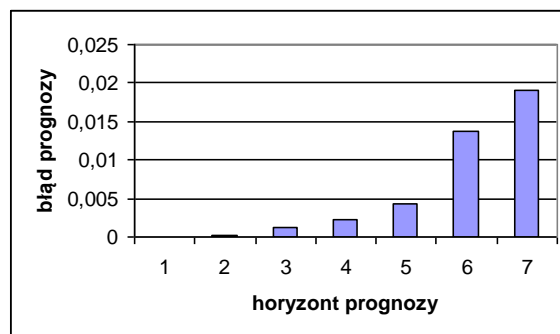
natomiast metody iteracyjnej:

$$\sigma'((\tilde{g}_1)^T) = e^{T\lambda} \sigma'(\tilde{g}_1),$$

gdzie h jest entropią systemu, s jest stopniem wielomianu aproksymującego, natomiast λ jest największym wykładnikiem Lapunowa systemu (Castagli [4]). Błąd predykcji metody iteracyjnej jest mniejszy niż metody bezpośredniej, ponieważ entropia systemu jest sumą dodatnich wykładników Lapunowa.

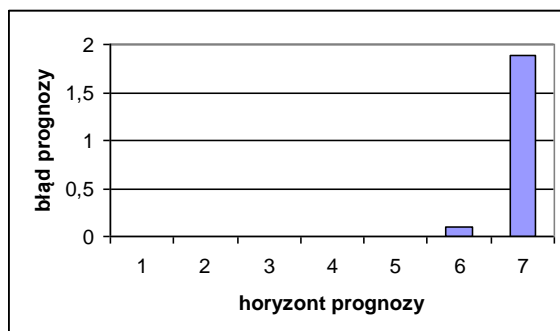
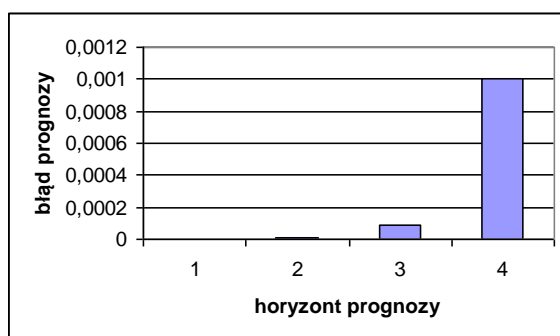
Przeprowadzone badanie polegało na prognozowaniu ewolucji szeregów czasowych dla horyzontu prognozy $T>1$ metodą iteracyjną. Dla każdego z rozważanych szeregów zastosowano liniową aproksymację wielomianową, z przyjętymi wartościami parametrów m , lag oraz k , zgodnie z tabelami 2,4,6,7 i 8. Na rys. 1-5 przedstawiono błędy prognoz σ w zależności od wartości horyzontu T .

Rys.1 Odwzorowanie Henona ($N_1=1650, N_2=65$).



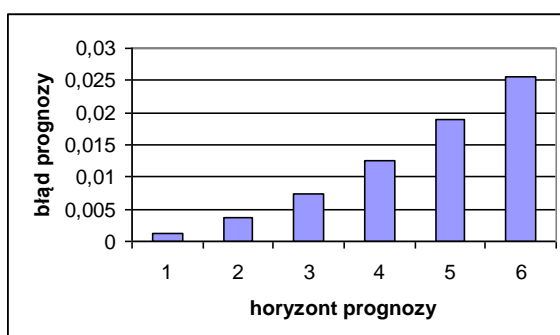
Źródło: obliczenia własne.

Rys.2 Odwzorowanie logistyczne ($N_1=1650, N_2=65$).



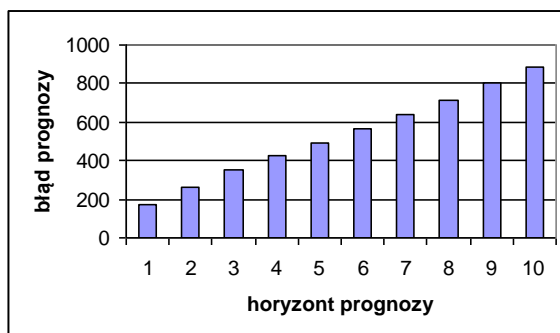
Źródło: obliczenia własne.

Rys.3 Odwzorowanie Lorenza ($N_1=1650$, $N_2=65$).



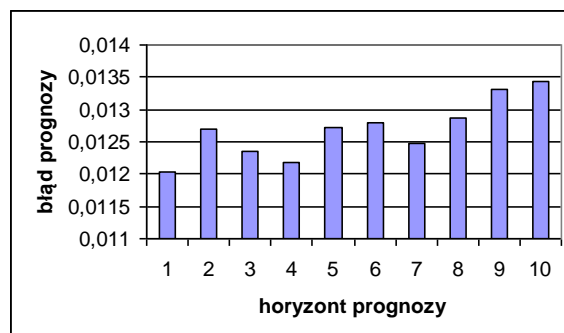
Źródło: obliczenia własne.

Rys.4 Szereg poziomów WIG ($N_1=1650$, $N_2=66$).



Źródło: obliczenia własne.

Rys.5 Logarytmiczne stopy zmian WIG ($N_1=1650$, $N_2=65$).



Źródło: obliczenia własne.

Zaprezentowane na wykresach 1-3 rezultaty krótkookresowego prognozowania szeregów chaotycznych potwierdzają wykładnicze powiększanie się błędów predykcji wraz z wydłużaniem horyzontu prognozy. Efekt ten jest nieobserwowany w przypadku szeregów giełdowych (rys. 4 i 5), co może stanowić sygnał, że nie są one generowane przez chaotyczny system dynamiczny.

Na podstawie przeprowadzonych badań należy stwierdzić, że w zastosowaniu do szeregów poziomów WIG oraz ich logarytmicznych stóp zwrotu, wątpliwa jest wyższość opisanych metod nad prognozowaniem opartym na modelach ARIMA. Powody takiego stanu rzeczy mogą być następujące:

- analizowane szeregi nie są chaotyczne,
- chaotyczny system generujący ma atraktor o wysokim wymiarze lub dużą entropię,
- badane szeregi są zbyt krótkie,
- obserwacje szeregów zostały wygenerowane przez stany systemu nie znajdujące się w obszarze przyciągania atraktora,
- w systemie obecny jest wysoki poziom zakłóceń losowych,
- omawiane metody nie są w stanie wykorzystać pewnych typów istniejących w szeregach zależności deterministycznych.

5. Zakończenie

Chaotyczne szeregi czasowe różnią się od szeregów prawdziwie losowych możliwością krótkoterminowego prognozowania ich ewolucji.

W pracy przedstawiono wybrane metody prognozowania – dwa warianty metody Najbliższych Sąsiadów oraz lokalną aproksymację wielomianową. Ich skuteczność zweryfikowano w oparciu o wygenerowane znane szeregi chaotyczne. Wyniki badań wskazują, że możliwe jest bardzo dokładne prognozowanie ewolucji tego typu szeregów w krótkim horyzoncie czasowym.

Zaprezentowane algorytmy różnią się złożonością obliczeniową. Najprostsze i w efekcie najbardziej efektywne czasowo są oba warianty metody Najbliższych Sąsiadów, z których nieco lepszy okazał się być zaproponowany w niniejszej pracy wariant rangowy. Dla szeregów chaotycznych najdokładniejsze prognozy otrzymywano stosując aproksymację wielomianową stopnia drugiego, zaś

wszystkie przedstawione metody dały lepsze wyniki niż odpowiednie modele ARIMA.

Zastosowanie zaprezentowanych metod do prognozowania szeregów czasowych indeksu WIG oraz jego stóp zmian na podstawie dziennych obserwacji z okresu 03.10.1994-20.08.2001 r. dało porównywalne wyniki do modeli ARIMA. W przebadanych rzeczywistych szeregach czasowych niezauważalny był efekt wykładniczego powiększania się błędu prognozy wraz ze wzrostem horyzontu czasowego, obserwowany w przypadku szeregów chaotycznych.

Wyniki badań wskazują, że przedstawione metody są bardzo skutecznym narzędziem prognozowania pewnych rodzajów szeregów czasowych, co potencjalnie może czynić je użytecznymi również w analizie niektórych procesów naturalnych w tym ekonomicznych.

Streszczenie

Chaotyczne szeregi czasowe różnią się od prawdziwie losowych obecnością trudno identyfikowalnego determinizmu, dzięki któremu możliwe jest ich prognozowanie. Ze względu na wykładniczą wrażliwość na zmianę warunków początkowych mogą być one efektywnie prognozowane jedynie w krótkim horyzoncie czasowym.

W pracy przedstawiono wybrane metody krótkoterminowego prognozowania chaosu: dwa warianty metody Najbliższych Sąsiadów oraz lokalną liniową i kwadratową aproksymację wielomianową. Ich skuteczność zweryfikowano w oparciu o wygenerowane szeregi chaotyczne różnej długości. Z przeprowadzonych badań wynika, że możliwe jest bardzo dokładne prognozowanie ewolucji analizowanych szeregów w krótkim horyzoncie czasowym. Zaprezentowane metody zastosowane do szeregów indeksu WIG oraz jego stóp zmian dały prognozy o podobnym stopniu dokładności, co prognozowanie wykorzystujące modele ARIMA.

Literatura

1. Arnold V.I., *Równania różniczkowe zwyczajne*, PWN, Warszawa (1975).
2. Badel A.E., Guégan D., Mercier L., Michel O., Comparison of Several Methods to Predict Chaotic Time Series, *IEEE-ICASSP'97*, Munich (Germany) (1997).
3. Brock W.A., Distinguishing Random and Deterministic Systems: Abridged Version, *Journal of Economic Theory* 40, 168-195 (1986).
4. Casdagli M., Nonlinear Prediction of Chaotic Time Series, *Physica D* 35, 335-356 (1989).
5. Devaney R.L., *An Introduction to Chaotic Dynamical Systems*, Addison-Wesley Publishing Company, Inc., Redwood City, CA (1987).
6. Diebold F.X., Nason J.A., Nonparametric Exchange Rate Prediction?, *Journal of International Economics* 28, 315-332 (1990).
7. Eckmann J.P., Ruelle D., Ergodic Theory of Chaos and Strange Attractors, *Reviews of Modern Physics*, 57, 617-656 (1985).

8. Farmer J.D., Sidorowich J.J., Predicting Chaotic Time Series, *Physical Review Letters* 59, 845-848 (1987).
9. Finkenstädt B., Kuhbier P., Forecasting nonlinear economic time series: A simple test to accompany the nearest neighbor approach, *Empirical economics* 20, 243-263 (1995).
10. Frank M., Stengos T., Chaotic dynamics in economic time series, *Journal of Economic Surveys*, 2 (2), 103-133 (1988).
11. Garrido L., Dynamical systems and chaos, Lectures Notes in Physics, nr 179, Springer-Verlag, Berlin, Heidelberg, New York (1983).
12. Jimenez J., Moreno J.A., Ruggeri G.J., Forecasting on chaotic time series: A local optimal linear-reconstruction method, *Physical Review A*, vol. 45, no. 6, 3553-3558 (1992).
13. Li T.Y., Yorke J.A., Period Three Implies Chaos, *American Mathematical Monthly*, 82, 985-992 (1975).
14. Lorenz H-W, *Nonlinear Dynamical Economics and Chaotic Motion*, Springer – Verlag Berlin Heidelberg (1989).
15. Takens F., Distinguishing deterministic and random systems, (G.Borenblatt, G. Iooss and D. Joseph, Eds), w: “Nonlinear Dynamics and Turbulence”, 315-333, Pitman, Boston (1985).
16. Zawadzki H., Chaotyczne systemy dynamiczne, Wydawnictwo Akademii Ekonomicznej im. Karola Adameckiego, Katowice (1996).