



Marek Nowiński

Uniwersytet Ekonomiczny we Wrocławiu
Wydział Zarządzania, Informatyki i Finansów
Katedra Badań Operacyjnych
marek.nowinski@ue.wroc.pl

NOWOCZESNA ANALIZA WIZUALNA EKONOMICZNYCH SZEREGÓW CZASOWYCH

Streszczenie: Artykuł przedstawia metody zaawansowanej analizy wizualnej, która nie polega jedynie na badaniu podstawowych, statystycznych własności szeregów czasowych, ale przede wszystkim na próbie wykrycia pewnej złożonej, ukrytej w nich struktury. Uzyskanie takich informacji nie jest możliwe na podstawie podstawowych badań statystycznych danych szeregu czasowego ani jego wykresu w postaci pierwotnej, który jest w rzeczywistości kompletnie nieczytelny. Współczesna analiza szeregu czasowego często przypomina przepuszczenie danych szeregu przez pewien *pryzmat* oraz przedstawienie wyników na odpowiednio skonstruowanym wykresie w celu wizualnej identyfikacji jego określonych własności. Istnieją również metody dwustopniowe, które dodatkowo zawierają analizę ilościową i możliwości szacowania specjalnych wskaźników uzyskanych wyłącznie na podstawie takich wykresów, które same niosą wiele przydatnych informacji i mogą stanowić wskazówki przy wyborze innych metod i narzędzi badawczych (por. np. metody *Recurrence Quantification Analysis* lub *Artificial Insymmetrised Patterns*).

Metody te pozwalają na odróżnienie badanego szeregu czasowego od losowego szumu, wykrycie zakłóconych procesów deterministycznych, ocenę rodzaju zależności w nim występujących, określenie stopnia stacjonarności, determinizmu i rekurencji. Mogą także pomóc w doborze metod pozwalających wykryć w danych elementy nieliniowości (a nawet chaosu deterministycznego). Inną zaletą tego podejścia jest możliwość ujawnienia w danych cykli okresowych o różnych długościach (co pozwala na bardziej skuteczne stosowanie modeli ARIMA lub wyrównywania wykładniczego, gdzie okres składnika cyklicznego musi być znany i może być wykorzystywany w pewnych modelach teoretycznych, np. średnich ruchomych lub autokorelacji).

Takie własności nowoczesnej wersji metod wizualnej analizy szeregów czasowych musiały wzbudzić zainteresowanie badaczy skomplikowanych zjawisk i procesów ekonomicznych, którzy próbują je wykorzystywać do pogłębionej analizy nieliniowej, a także do efektywnego modelowania i prób prognozowania tych procesów. Jest to również powód przedstawienia urozmaiconego przeglądu tych metod w niniejszym artykule.

Słowa kluczowe: złożone procesy dynamiczne, szeregi czasowe, wizualna analiza nieliniowa.

Wprowadzenie

Współczesne metody zaawansowanej analizy wizualnej polegają nie tylko na badaniu podstawowych własności szeregów czasowych, ale także na próbach wykrycia pewnych ukrytych w nich struktur deterministycznych. Te złożone struktury bardzo często nie mogą być ujawnione na podstawie wykresu szeregu czasowego w jego pierwotnej postaci ani nawet typowych badań statystycznych.

Nowoczesna analiza szeregów czasowych może do złudzenia przypominać efekt przepuszczenia pierwotnych danych szeregu przez pewien typ *pryzmatu*, który ujawnia jego specyficzne własności. Ich wizualna identyfikacja może być zrealizowana na odpowiednio skonstruowanym wykresie. Następnym etapem powinna być analiza ilościowa tych wykresów, pozwalająca na oszacowanie wartości pewnych wskaźników, stanowiących cenne wskazówki przy wyborze dalszych narzędzi badawczych analizy nieliniowej.

Metody te pozwalają na odróżnienie badanego szeregu czasowego od losowego szumu, wykrycie zakłóconych procesów deterministycznych, ocenę rodzaju zależności w nim występujących, określenie stopnia stacjonarności, determinizmu i rekurencji. Mogą także pomóc w doborze metod pozwalających wykryć w danych elementy nieliniowości (a nawet chaosu deterministycznego), a także ujawnić w danych pewne wzorce zachowań i cykle okresowe o różnych długościach, co pozwala na bardziej skuteczne stosowanie bardziej zaawansowanych metod wygładzania analizowanych szeregów, co ułatwia wiarygodne modelowanie, a nawet prognozowanie badanych procesów dynamicznych. Celem niniejszego opracowania jest prezentacja podstawowych własności wizualnych metod nieliniowej analizy szeregów czasowych oraz wskazanie możliwości ich stosowania do modelowania i prognozowania zjawisk gospodarczych i finansowych.

1. Metoda ilościowej analizy rekurencyjnej RQA^1

Wykresy rekurencyjne opierają się na spostrzeżeniu, że w naszym otoczeniu występuje wiele procesów obrazujących różnorodne zależności typu okresowego (np. sezonowe) lub zjawiska nieregularne, ale zbliżone do okresowych (np. powtarzalność pewnych wzorców zachowań pogodowych, społecznych czy ekonomicznych). Dawno już zauważono, że rekurencja, rozumiana jako bliska odległość przestrzenna stanów procesu odległych od siebie w skali czasowej,

¹ Opracowano opierając się częściowo na wcześniejszej monografii autora [Nowiński, 2007] oraz publikacji Nowiński [2008].

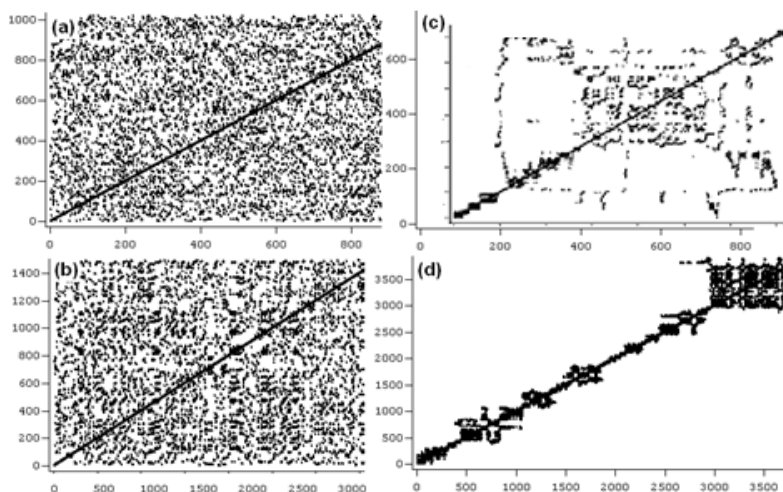
jest oczywistą cechą dynamicznych układów deterministycznych. Analiza rekurencyjna może pomóc w wykrywaniu niestacjonarności, zwłaszcza dla dużych wymiarów układu i dla danych z dużą ilością szumu, co dotyczy też ekonomicznych szeregów czasowych [Hołyst, Urbanowicz, 2000]. Metody te nie są uwarunkowane stacjonarnością procesu i nie narzucają wymogu specjalnego filtrowania danych. Pozwalają wykryć korelacje nawet wtedy, gdy jest to trudne przy użyciu klasycznej funkcji autokorelacyjnej. Nie wymagają założeń o postaci zależności funkcyjnych generujących badane sygnały ani założeń o typie rozkładu, a wykresy rekurencyjne zachowują własności niezmienników charakteryzujących zachowanie generujących je układów dynamicznych [Cao, Cai, 2000; Bradley, Mantilla, 2001]. Metody rekurencyjne początkowo nie zyskały znacznej popularności, gdyż tworzone przez nie wykresy nie są łatwe do interpretacji.

Zazwyczaj ekonomiczne i finansowe szeregi czasowe są badane na gruncie analizy procesów stochastycznych, ale także w powiązaniu z teorią nieliniowych układów dynamicznych generujących skalarne szeregi czasowe obserwacji, które stanowią numeryczny opis zachowania się badanego zjawiska. Inne znane metody oparte na analizie topologicznej pokazały, że dane finansowe nie mogą być traktowane jako tworzone przez procesy w żaden sposób uporządkowane, całkowicie nieprzewidywalne i przez to niepoddające się racjonalnej analizie. Rozwinięcie tego podejścia w postaci metody wykresów rekurencyjnych *RP* oraz ilościowej analizy rekurencyjnej *RQA* powinno pomóc w ujawnieniu objawów determinizmu oraz umożliwić badanie zmian dynamiki procesów generowanych przez określone układy dynamiczne, a w szczególności wykrywaniu kryzysów i załamań tendencji rozwojowych, które mogą być zbliżone do fizycznych przejść fazowych². Zadanie to jest szczególnie ważne, gdy mamy do dyspozycji jedynie niestacjonarne i silnie zniekształcone (zazsumione) dane w postaci jednowymiarowych szeregów czasowych x_t . Zmiany stanu układu dynamicznego mogą być odzwierciedlone w zróżnicowaniu ciągu wektorów stanu umieszczonych w zrekonstruowanej przestrzeni fazowej. Twierdzenie Takensa [Takens, 1993] zapewnia, że pod pewnymi warunkami każdy punkt pierwotnej przestrzeni stanów może być odtworzony dla ustalonych wartości wymiaru zanurzenia³ m i opóźnienia τ na podstawie wektorów zanurzenia $\tilde{x}_t = (x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau})$ w pewnej m -wymiarowej przestrzeni R^m .

² Przejście fazowe (ang. *phase transition*) – nagła zmiana zachowania układu dynamicznego. Typowy przykład to zmiana stanu, w którym występuje wyraźny brak ciągłości procesu w pewnej krytycznej temperaturze (punkt przejścia fazowego). Oznacza też przejście od stanu samoorganizacji do chaosu (przejście typu porządek – nieporządek).

³ Wymiar zanurzenia – kolejne m -ciągi opóźnionych czasowo punktów szeregu (ewentualnie z zastosowaną separacją czasową między obserwacjami) traktuje się jako punkty m -wymiarowej

Wykresy *RP* (ang. *Recurrence Plots*) umożliwiają badanie trajektorii procesu w wielowymiarowej przestrzeni stanów na podstawie ich reprezentacji w dwuwymiarowej przestrzeni rekurencji. Taka reprezentacja to macierz $N \times N$, gdzie N to liczba wektorów zanurzenia x_i (o wymiarze m) uzyskanych dla opóźnionych czasowo o okres τ obserwacji sygnału pierwotnego. W macierzy tej na pozycji (i, j) ⁴ znajduje się liczba oznaczająca euklidesowską odległość między wektorami x_i oraz x_j . Istnieją dwa typy wykresów rekurencyjnych – z wartością progową i bez niej. Oczywiście wykresy z wartością progową są symetryczne względem głównej przekątnej. Na wykresie rekurencyjnym tego typu punkt (i, j) jest zaznaczany tylko wtedy, gdy wektor x_i znajduje się dostatecznie blisko wektora x_j , czyli bliżej niż wyznaczona z góry wartość progowa. Wartość tę ustala się najczęściej na poziomie zależnym od wartości obserwacji (najczęściej od 0,05 do 0,2 odchylenia standardowego danych pierwotnych). Na wykresach bez wartości progowej promienia otoczenia piksel leżący na pozycji (i, j) jest kolorowany zgodnie z odległością od swego sąsiada według pewnej ustalonej skali kolorystycznej. Zazwyczaj ciemne kolory odpowiadają dużym odległościom, zaś jasne – małym.



Rys. 1. Wykresy rekurencyjne dla różnego typu procesów stochastycznych: (a) biały szum, (b) proces chaotyczny Henona, (c) ceny ropy naftowej⁵ oraz (d) giełdowy szereg czasowy DJIA⁶

przestrzeni zanurzenia, gdzie m to właśnie wymiar zanurzenia szeregu czasowego, procesu lub układu dynamicznego reprezentowanego przez skalarny ciąg obserwacji.

⁴ Gdzie oś pozioma reprezentuje indeks czasu, zaś oś pionowa przesunięcie czasowe.

⁵ Średnie ceny ropy naftowej (\$/baryłkę) na rynku amerykańskim – dane miesięczne z lat 1946-2012, 916 obserwacji, źródło: Wall Street Journal.

⁶ Wartości wskaźnika *Dow Jones Industrial Average* z giełdy *NYSE* – dane dzienne zamknięcia z lat 1986-2012, 4897 obserwacji, źródło: Yahoo! Finance Inc.

Pewna homogeniczność wzorców występujących na określonych fragmentach wykresu (zwarte prostokąty) oznacza stacjonarność określonego przedziału ciągu danych. Układy niestacjonarne powodują zmiany rozrzutu punktów rekurencyjnych, co objawia się w postaci rozjaśnionych fragmentów na wykresie.

Podstawowe wnioski, które można wyciągnąć na podstawie wizualnej analizy wykresów rekurencyjnych, są następujące:

- Wykresy homogeniczne, o jednolitym charakterze rozrzutu punktów, bez wyraźnych wzorców i struktur cechują procesy czysto losowe.
- Procesy zbliżone do okresowych charakteryzują się występowaniem przekątniowych, równoległych lub struktur zbliżonych do szachownicy, gdzie wzajemna odległość tego typu linii odpowiada w przybliżeniu wielkości okresu lub pseudocyklu.
- Nagłe zmiany dynamiki procesu oraz występowanie ekstremalnych wartości szeregu czasowego powodują pojawienie się jasnych obszarów lub pasów na wykresie, co dostarcza prostych możliwości identyfikacji tzw. rzadkich zdarzeń w ewolucji badanego układu.
- Linie diagonalne na wykresie występują wtedy, gdy trajektoria procesu odwiedza ten sam obszar przestrzeni fazowej w różnych okresach, czyli potwierdza występowanie zachowań rekurencyjnych. Długość tych linii przesądza o prawdopodobieństwie wystąpienia chaosu deterministycznego.
- Linie diagonalne prostopadłe do głównej przekątnej są zazwyczaj objawem źle dobranego wymiaru zanurzenia szeregu.
- Linie pionowe lub poziome oznaczają okresy, w których stan procesu prawie się nie zmienia lub zmiany te są bardzo powolne. Gwałtowne pojawienie się białych obszarów w pobliżu tych linii może sugerować proces z objawami zjawisk intermitencji, czyli nagłych, gwałtownych i krótkotrwałych zaburzeń jego przebiegu.
- Zaniki koloru w prawym dolnym i lewym górnym wierzchołku wykresu rekurencyjnego oznaczają niestacjonarność procesu związaną z występowaniem dryfu, czyli pewnym trendem interpretowanym jako powolna zmiana parametrów procesu.

Zbilut i Webber [1992] wprowadzili nowe narzędzia ilościowe dostarczające dodatkowych informacji na podstawie wykresów rekurencyjnych. Była to tzw. metoda rekurencyjnej analizy ilościowej *RQA* (ang. *Recurrence Quantification Analysis*).

Nowe miary złożoności procesu opierają się na analizie struktur charakterystycznych wykresu i mogą być wykorzystane do wykrywania elementów deterministycznych w pozornie losowych procesach dynamicznych oraz do precyzo-

wania stopnia oraz rodzaju determinizmu występującego w danych. Najważniejsze z tych miar to:

- Stopień rekurencji *REC* (uwaga – pomija się wszystkie punkty głównej przekątnej!)

$$REC = \frac{\text{liczba punktów rekurencyjnych}}{\text{liczba wszystkich punktów wykresu}}$$

- Stopień determinizmu *DET*:

$$DET = \frac{\text{liczba punktów w odcinkach diagonalnych (> 2)}}{\text{liczba wszystkich punktów rekurencyjnych}}$$

- Entropia informacyjna Shannona *ENT*:

$$ENT = \frac{1}{N} \sum_{i=1}^N p_i \log_2 \frac{1}{p_i}$$

gdzie N to długość ciągu danych (liczba wektorów zanurzenia), zaś p_i to prawdopodobieństwo (traktowane częstościowo) trafienia na linię diagonalną o długości i w zbiorze wszystkich punktów rekurencyjnych. Im mniejsza entropia, tym lepsze dopasowanie rekonstruowanego atraktora.

- Entropia przestrzenno-czasowa *STE* (ang. *spatio-temporal entropy*) – mierzy stopień strukturyzacji wykresu rekurencyjnego zarówno w dziedzinie przestrzeni, jak i czasu. Porównuje ona ogólne rozmieszczenie punktów na całym wykresie z ich rozkładem na każdej przekątnej, co daje wyobrażenie o stopniu ustrukturyzowania otrzymanego obrazu. Oznacza to, że wskaźnik 100% odpowiada właśnie całkowitemu brakowi jakichkolwiek struktur na wykresie rekurencyjnym (idealny proces białego szumu), zaś 0% to pełna struktura odpowiadająca determinizmowi i całkowitej prognozowalności procesu. Entropia ta zmienia się dla różnych wymiarów zanurzenia, co pozwala wykorzystać ją także do doboru optymalnej wartości tego parametru.

Powyższe 4 wskaźniki w sposób ilościowy opisują strukturę deterministyczną oraz stopień złożoności badanych procesów. Ponieważ metoda *RQA* jest w zasadzie niezależna od takich ograniczeń, jak wielkość zbioru, stacjonarność danych oraz założenia dotyczące ich rozkładów statystycznych, może być ona stosowana do opisu układów charakteryzujących się różnymi od homeostatycznych zachowaniami przejściowymi oraz nagłymi zmianami stanu (analiza dynamiki molekularnej, analiza zapisu sygnału mowy ludzkiej, kardiografia, a także badania ekonomicznych i finansowych szeregów czasowych).

Rezultaty przedstawione w tab. 1 pokazują, że między procesami losowymi a rzeczywistymi procesami gospodarczymi zachodzą wyraźne różnice zarówno

co do stopnia determinizmu, jak i wskaźnika entropii przestrzenno-czasowej *STE*, co pozwala na ich łatwe rozróżnienie i ustalenie obecności silnej podstawy deterministycznej w ekonomicznych szeregach czasowych. Bardzo duże wartości współczynnika *DET* wynikają z przyjęcia stosunkowo dużej wartości promienia sąsiedztwa, co jednakże nie zniekształca wzajemnych zależności między stopniem determinizmu badanych procesów. Entropia informacyjna Shannona *ENT* oraz wskaźnik *REC* mierzą siłę rekurencji badanych szeregów czasowych, co w pewien sposób pozwala sprawdzić stopień ich chaotyczności (a więc i prognozowalności), gdyż w przybliżeniu oceniają prawdopodobieństwo, że zbliżone stany układu będą się powtarzać w przyszłości. Ponadto mniejsza entropia oznacza lepszą ocenę jakości dopasowania w procesie zanurzenia.

Tabela 1. Wartości wskaźników analizy *RQA* dla przykładowych procesów (dla optymalnie dobranych wielkości wymiaru zanurzenia i opóźnienia czasowego)

Proces	Stopień rekurencji <i>REC</i>	Entropia informacyjna Shannona <i>ENT</i>	Stopień determinizmu <i>DET</i>	Entropia przestrzenno-czasowa <i>STE</i>
BSZUM - losowy	1,0	11,86	0,6	88
HENON- chaos	2,7	3,41	35	50
LORENZ	3,4	5,88	56	42
ROSSLER	3,3	5,40	42	44
CPI⁷ - ekonomiczne	1,2	8,94	69	16
DJIA	2,6	7,55	67	40
RNAFT	2,4	6,61	72	50
SREBRO⁸	2,9	7,05	65	48
US-DEM⁹	1,5	6,45	78	31

Widzimy wyraźnie, że procesy chaotyczne charakteryzują się dość niskim poziomem entropii Shannona i wysokim stopniem rekurencji. Wskaźniki stopnia rekurencji dla procesów ekonomicznych są również dość wysokie, co daje nam nadzieję na możliwość wykorzystania ich własności w procedurach prognozy stycznych. Natomiast znaczące spadki wartości wskaźnika determinizmu *DET* szeregów finansowych połączone z istotnym wzrostem wartości średniej (i jednocześnie znacznymi wahaniami konkretnych pomiarów) wartości entropii informacyjnej Shannona *ENT* mogą oznaczać rozpoczęcie okresu kryzysu ekonomicznego (bessy giełdowej, pęknięcia bańki spekulacyjnej).

⁷ Miesięczne wskaźniki poziomu cen towarów i usług konsumpcyjnych (*Consumer Price Index*) w USA – z lat 1913-2012, 1512 obserwacji, źródło: *U.S. Department Of Labor*.

⁸ Ceny srebra (\$/uncję) na rynku *NY Comex Silver* – dane dzienne z lat 1981-2006, 4500 obserwacji, źródło: *Norman's Historical Data of New York Commodity Exchange*.

⁹ Kurs wymiany walut US/DEM – dane dzienne z lat 1984-2012, 5096 obserwacji, źródło: *FXHistory - OANDA Corp*.

Wykresy rekurencyjne są dobrym narzędziem służącym do odróżniania szeregów chaotycznych od procesów zbliżonych do procesów autoregresyjnych, dla których zawodzi procedura ustalania ich wymiaru korelacyjnego, gdyż procesy regresyjne tego typu mogą zniekształcić wartość tego wymiaru. Ale wykresy rekurencyjne dla obu typów tych procesów wyglądają zupełnie inaczej (dla procesów autoregresyjnych uwidaczniają one duże skupiska punktów rekurencyjnych w pasie wokół głównej przekątnej, co jest objawem występowania silnych korelacji dynamicznych). Po ujawnieniu na wykresach rekurencyjnych potencjalnych chaotycznych własności badanego procesu możemy próbować je wykorzystać (i jednocześnie potwierdzić) za pomocą prostej procedury predykcji krótkoterminowej. Przyszłe krótkie ciągi stanów układu mogą być prognozowane na podstawie wyselekcjonowanych podzbiorów danych, które reprezentują wzorce zachowań zbliżone do wzorca bezpośrednio poprzedzającego okres prognozowany.

Powyższe wyniki stanowią właściwie tylko wstęp do znacznie bardziej rozbudowanych badań procesów tego typu z wykorzystaniem narzędzi analizy wizualnej wykresów rekurencyjnych i ilościowej analizy *RQA*, które aktualnie są przeprowadzane przez autora, ale wymagają długotrwałych i pracochłonnych obliczeń, gromadzenia wyników i gruntownej ich interpretacji. Mamy tu przede wszystkim na myśli badania zmienności wykresów i wskaźników poziomu rekurencji i determinizmu w czasie, co może pozwolić na ustalenie wielkości korelacji między zmianami właściwości układu a wpływem znanych czynników zewnętrznych. Niewątpliwie potrzeba dalszych pogłębionych badań, aby w sposób ilościowy można było ocenić te możliwości.

Przez badania różnych wykresów *RP* oraz analizę *RQA* można też wykryć zmiany stopnia korelacji między poszczególnymi procesami gospodarczymi podczas roku. Jedną z podstawowych reguł dobrej strategii inwestycyjnej jest dobór portfela aktywów tak, by nie wykazywały one tendencji do zmian w jednakowym kierunku w czasie. Dywersyfikacja portfela pomaga inwestorowi w minimalizowaniu ryzyka podczas nieoczekiwanych fluktuacji na rynkach walutowych, zwłaszcza poprzez wykorzystanie lokat wzajemnie nieskorelowanych. Powyższe techniki analityczne mogą pomóc w rozwiązywaniu problemów związanych z analizą, inwestowaniem i prognozowaniem na międzynarodowych rynkach walutowych i giełdach towarowych.

Wyniki dotychczas przeprowadzonych badań (z wykorzystaniem analizy wizualnej opartej na badaniu wykresów rekurencyjnych i śledzeniu zmian wartości istotnych wskaźników analizy *RQA*) pozwalają mieć nadzieję, że w niektórych przypadkach będzie można z pewnym opóźnieniem od pierwszych, jeszcze niezauważalnych momentów początku potencjalnego, silnego kryzysu uzyskać

pewne ostrzeżenia mogące złagodzić jego przyszłe skutki. Ostrzeżenia te mogą się pojawić wystarczająco wcześnie przed rzeczywistym wystąpieniem zjawiska kryzysowego, by poprzez systematyczne monitorowanie procesów giełdowych umożliwić wykrycie potencjalnych zmian stanu odpowiadających im układów dynamicznych i ujawnić objawy zachowań krytycznych.

2. Metoda wzorców ze sztucznie wprowadzoną symetrią AIP

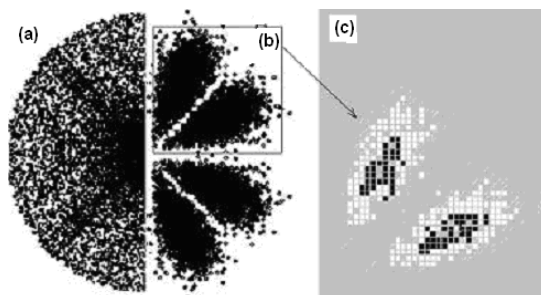
Metoda AIP (ang. *Artificial Insymmetrised Patterns*) to prosta metoda graficzna, która pozwala ocenić bieżące i przeszłe stany badanego procesu, a także w pewien sposób ostrzegać o nadchodzących zjawiskach nietypowych w postaci nagłych spadków wartości, załamania tendencji rozwojowych oraz ewentualnych etapach gwałtownego przyspieszenia tempa zmian, także korzystnych, w analizowanych szeregach czasowych. Może być ona wykorzystywana do wykrywania, charakteryzowania i porównywania wzorców dynamiki procesów, które są połączeniem zależności typu deterministycznego i stochastycznego, często o bardzo dużym wymiarze, a także zawierających elementy zachowań chaotycznych. Metodę tą wprowadzono pod koniec ubiegłego wieku [Pickover, 1990]. Oto krótki opis tej metody. Mając dany szereg czasowy x_t kolejno:

- Ustalamy wartości parametrów s oraz τ , czyli zakresu skali oraz przesunięcia czasowego obserwacji szeregu.
- Przeskalowujemy dane do zakresu $[0, s]$ za pomocą prostego przekształcenia:

$$x_t \rightarrow s \frac{x_t - x_{\min}}{x_{\max} - x_{\min}}$$

co pozwala na usytuowanie niskich wartości obserwacji w pobliżu zera, zaś wysokich blisko wartości s (jako typową wartość s najczęściej przyjmuje się 2π).

- Wykreślamy dwuwymiarowe odwzorowanie powrotne w biegunowym układzie współrzędnych $\{x_t, x_{t+\tau}\}$, gdzie x_t oznacza odległość punktu od zera, a $x_{t+\tau}$ kąt, pod którym zaznacza się punkt na płaszczyźnie.
- Wymuszamy symetrię wykresu poprzez k -krotny obrót o pewien wybrany kąt ϕ , czyli wykreślamy $\{x_t, x_{t+\tau} + k\phi\}$.
- Stosujemy odbicie lustrzane $\{x_t, \pm x_{t+\tau} + k\phi\}$.



Rys. 2. Fragmenty wykresów *AIP*: (a) dla procesu losowego o rozkładzie jednostajnym, (b) procesu o rozkładzie normalnym oraz (c) ilustracja gęstości punktów na fragmencie wykresu (b)

Wykres przedstawiony na rys. 2 jest nazywany wykresem *AIP* lub wykresem *SDP* (ang. *Symmetrised Dot Pattern*) i umożliwia ilustrację obrazów korelacji lokalnych. Kolor szary oznacza brak punktów, biały średnią gęstość, zaś czarny dużą gęstość punktów na wykresie. Widzimy wyraźnie, że dla szumu o rozkładzie jednostajnym punkty pokrywają prawie równomiernie całą powierzchnię wykresu i nie wyróżniamy żadnych wzorców, gdyż nie występują tu żadne składniki deterministyczne ani systematyczne. Dla szumu o rozkładzie $N(0,1)$ kontury wykresu są rozmazane, punkty częściowo skoncentrowane w okolicach zera i bardzo rzadkie na obrzeżach. Tak więc procesy zawierające więcej składowych deterministycznych (łatwiejsze w prognozowaniu) powinny być zwarte, o małym rozproszeniu punktów.

Zdecydowane różnice występują między sztucznie wygenerowanym procesem deterministycznym a tym samym procesem zakłóconym przez dodany szum losowy. Rozmazanie krawędzi i poszerzenie obszaru zajmowanego przez punkty wykresu to objawy zwiększonego stopnia losowości, ale z jednoznacznie widoczną podstawą zależności ściśle deterministycznej.

Na podstawie wykresów *AIP* można również uzyskać pewne charakterystyki ilościowe dotyczące badanych procesów, obliczone na podstawie miar gęstości i rozrzutu punktów. Jedną z takich statystyk może być entropia informacyjna Shannona *ENT1*, którą w tym przypadku wyznacza się poprzez podzielenie obszaru wykresu *AIP* na m^2 identycznych kwadratów oraz obliczenie liczby punktów wpadających do każdego z nich [Shannon, 1948]. Dla takiego podziału wylicza się częstościowe prawdopodobieństwo p_i , $i = 1, \dots, m^2$ znalezienia się w każdym kwadracie.

Następnie siatkę podziału zagęszcza się dwukrotnie, uzyskując $4m^2$ nowych, mniejszych kwadratów i powtarza się procedurę obliczania liczby punk-

tów w każdym z nich. W ten sposób możemy obliczyć graniczną wartość entropii Shannona według wzoru:

$$ENT1 = \lim_{i \rightarrow M} \sum_{p_i > 0} p_i \log_2 \frac{1}{p_i}$$

Oczywiście dla idealnego procesu losowego po kolejnym podziale liczba punktów w każdym kwadracie powinna się zmniejszać czterokrotnie i w prawie każdym z nich powinny zostawać jakieś punkty. Natomiast dla procesów o dużym udziale czynników deterministycznych będzie coraz więcej kwadratów, w których nie będzie żadnych punktów, gdyż koncentrują się one tylko w małej części wykresu. Tak więc duże wartości entropii (mierzącej stopień braku uporządkowania w systemie) będą oznaczać procesy losowe, zaś małe – deterministyczne. Liczne eksperymenty numeryczne pokazały, że jako wartość progową entropii $ENT1$ można przyjąć liczbę 4. Dla wartości $ENT1$ powyżej tej granicy można uznać, że badany proces jest losowy. Im wartość $ENT1$ jest mniejsza od 4, tym więcej jest w procesie składników deterministycznych, a mniej czynników czysto losowych.

Dzięki wykresom AIP można również zaobserwować i bardziej dokładnie zbadać specyficzne własności niektórych ekonomicznych szeregów czasowych, takie jak:

- tzw. prawostronne grube ogony rozkładu, czyli zwiększone częstotliwości obserwacji dla krańców przedziału wartości, mogą oznaczać okres występowania nagłych i dużych zmian kursów akcji;
- znacząco duża kurtoza, skupienie wartości szeregu wokół średniej, co minimalizuje możliwości znacznych zarobków w krótkim okresie;
- niezależność diagramu od zmieniającej się wartości opóźnienia czasowego τ , co może oznaczać częściowe „zamrożenie” rynku i sygnalizować zbliżający się krach giełdowy.

Szeregi czasowe kursów akcji notowanych na giełdach światowych dostarczyły wartości entropii Shannona mieszczących się w przedziale $[0,38;3,33]$, co świadczy o dużym zróżnicowaniu ich losowości, a zatem i prognozowalności [Najman, 2003]. Inne ekonomiczne szeregi czasowe także znacząco się różnią pod względem własności, ale wykresy AIP zawsze pozwalają na odróżnienie ich zarówno od procesów losowych, jak i czysto deterministycznych.

Podsumowanie

Szybki rozwój techniki komputerowej w połączeniu ze wzrostem zainteresowania naukowców badaniem zjawisk złożonej dynamiki nieliniowej (w tym ich analizy wizualnej) spowodował, że szeroko zakrojone badania w tej dziedzi-

nie przeniosły się z nauk eksperymentalnych i przyrodniczych (fizyka, biologia, medycyna) także do nauk ekonomicznych. Metody wizualne zaprezentowane powyżej bardzo zainteresowały badaczy złożonych zjawisk i procesów ekonomicznych, którzy w ostatnim czasie starają się je wykorzystywać do pogłębionej analizy nieliniowej, a także wykazać ich przydatność w modelowaniu i prognozowaniu finansowych szeregów czasowych (w postaci kursów akcji giełdowych i kursów wymiany walut), a także innych szeregów ekonomicznych (np. cen metali szlachetnych i surowców strategicznych na rynkach światowych, danych o koniunkturze gospodarczej, wskaźników makroekonomicznych itp.) oraz do prób skutecznego prognozowania (np. zużycia prądu elektrycznego przy jego dużej zmienności) i przewidywania sytuacji kryzysowych w gospodarce.

Zainteresowanie to musiało spowodować powstanie wielu narzędzi informatycznych wspomagających skomplikowane obliczenia związane z analizą nieliniową zjawisk opisywanych przez bardzo liczne szeregi czasowe danych (np. dla *danych tickowych* z giełdy Forex lub często próbkowanych danych giełdowych *typu intraday*). Najczęściej były to nowo opracowywane pakiety programów [np. Kantz, Schreiber, 2004], ale tworzono też oprogramowane w istniejących środowiskach do obliczeń statycznych (np. Xplore, Matlab lub R). Do najważniejszego oprogramowania w tej dziedzinie można współcześnie zaliczyć:

TISEAN ver. 2007 (*Nonlinear Time Series Analysis Software*), autorzy: Hegger, Kantz, Schreiber, [<http://www.mpipks-dresden.mpg.de/~tisean/>];

Dataplore – ixellence GmbH – *Analysis of Signals and Time Series Data*, [<http://www.mpipks-dresden.mpg.de/~tisean/>];

CRP Toolbox, autor N. Marwan, [<http://tocsy.pik-potsdam.de/CRPtoolbox/>];

RQA Software (*Recurrence Quantification Analysis*), autor C.L.Jr Weber [<http://homepages.luc.edu/~cwebber>].

Literatura

Bradley E., Mantilla R. (2001), *Recurrence Plots and Unstable Periodic Orbits*, Univ. of Colorado Tech Report CU-CS 919-01.

Cao J., Cai H. (2000), *On the Structure and Quantification of Recurrence Plots*, "Physica Letters A", 270, s. 75-87.

Hołyst J.A., Urbanowicz K. (2000), *Chaos Control in Economical Model by Time-delayed Feedback Method*, "Physica A", 287, s. 587-598.

Kantz H., Schreiber T. (2004), *Nonlinear Time Series Analysis*, Cambridge University Press.

- Najman K. (2003), *Diagramy ze sztucznie wprowadzoną symetrią jako narzędzie pomiaru natężenia składników losowych w szeregach czasowych stóp zwrotu*, PN AE Wrocław, nr 991, s. 423-439.
- Nowiński M. (2007), *Nieliniowa dynamika szeregów czasowych w badaniach ekonomicznych*, Wydawnictwo AE, Wrocław.
- Pickover C.A. (1990), *Computer, Patterns, Chaos and Beauty*, St. Martin's Press, New York.
- Shannon C.E. (1948), *A Mathematical Theory of Communication*, "Bell System Techn. J.", 27, s. 379-423.
- Takens F. (1993), *Detecting Nonlinearities in Stationary Time Series*, "Int. J. of Bifurcation and Chaos", 3, s. 241-56.
- Zbilut J.P., Webber C.L. (1992), *Embeddings and Delays as Derived from Quantification of Recurrence Plots*, "Phys.Lett. A", 171, s. 199-203.

MODERN VISUAL ANALYSIS OF ECONOMIC TIME SERIES

Summary: The paper presents an advanced visual analysis method that does not rely only on examining of basic statistical properties of the time series, but also on attempting to detect a complex structure, hidden in the original dynamic process. Obtaining such information is not possible on the basis of the basic survey of time series data, or its graph in the original form, which is in fact completely illegible. Contemporary time series analysis of time series data is often reminiscent of passing it through a prism, and presenting the results on a properly constructed plots for visual identification of its specific properties. There are also two-step methods, which include both the possibility of quantitative analysis and estimation of specific indicators derived on the basis of such plots. They can carry a lot of useful information and provide us the guidance for the further selection of proper research of nonlinear analysis methods and tools (see eg. the methods of *Recurrence Quantification Analysis* and *Artificial Insymmetrised Patterns*).

These methods allow to distinguish the analyzed time series from the random noise, to detect real deterministic processes biased with noise, to assess the type of dependencies in time series data, and to determine the degree of stationarity, determinism or recurrence. They can also help in the selection of methods to detect the nonlinearity in the data (or even deterministic chaos). Another advantage of this approach is the possibility of disclosure in data periodic cycles with periods of different lengths (for more efficient use of ARIMA and exponential smoothing models, where the period of the cyclical component must be known and can be used in certain theoretical models, eg. moving average or autocorrelation).

Such properties of the modern version of visual time series analysis caused the interest of scientists trying to understand the complex economic phenomena and processes. They are trying to use it for in-depth nonlinear analysis, economic modeling and attempts of effective forecasting of these processes. This is the main reason for presenting a review of these methods in this paper.

Keywords: complex dynamic processes, time series, visual nonlinear analysis.