

Henryk Gurgul*, Marcin Suder*

Ekonometryczna analiza fraktalnych właściwości struktury przepływu gazu w wybranych stacjach I stopnia

1. Wprowadzenie

Teoria chaosu znana też pod nazwą teorii złożoności albo teorii systemów złożonych została najpierw zdefiniowana w ramach matematyki i fizyki oraz zastosowana do różnych zagadnień w tych dyscyplinach nauki. Obecnie ma zastosowanie w programowaniu, mikrobiologii, inżynierii (szczególnie robotyce), finansach, meteorologii, filozofii, polityce, badaniach demograficznych i psychologii (por. [17], [21], [22], [23], [36]). Ta dziedzina badań dotyczy głównie nieliniowych systemów występujących w przyrodzie, szczególnie w naszym najbliższym środowisku. Jest ona stosowana między innymi w modelach prognostycznych, w gospodarce (np. do opisu kursów papierów wartościowych), do opisu zachowań ludzi, a więc psychologii tłumu, wyjaśniania korków ulicznych i tak dalej.

Znany w literaturze przedmiotu efekt motyla jest dobrym przykładem mogącym służyć do przedstawienia zakresu teorii złożoności i jej zastosowań. Czasem jest tak, że jednokrotne uderzenie skrzydeł motyla w Argentynie mogłoby wywołać tornado lub inne znaczne zmiany pogody np. w Kalifornii. Oczywiście tej sytuacji nie należy brać dosłownie, ale przedstawia ona dobrze główny problem teorii chaosu. Mała odchyłka lub mała zmiana warunków początkowych może istotnie zmienić dotychczasową trajektorię, po której dotąd poruszał się system,

* AGH Akademia Górniczo-Hutnicza w Krakowie, Wydział Zarządzania, Samodzielna Pracownia Zastosowań Matematyki w Ekonomii, email: henryk.gurgul@gmail.com, m_suder@wp.pl

a w konsekwencji spowodować przejście systemu do całkowicie odmiennego od przewidywanego stanu końcowego.

Jak to już zostało przedstawione w pracy Gurgula i Sudera (por. [15]), tak rynki finansowe, jak i pojedyncze papiery wartościowe mogą być uważane za chaotyczne, jeśli nie istnieją dla nich liniowe lub inne jednoznaczne reguły opisu.

Kurs akcji nie przedstawia bowiem niczego innego jak tylko wypadkową decyzji wszystkich uczestników rynku. Przebieg krzywych ilustrujących rozmaite dane gospodarcze daje się opisać za pomocą tzw. fraktali.

Znaczenie deterministycznych systemów chaotycznych wynika z dwóch istotnych faktów. Przede wszystkim systemy te nie dają się prognozować w długiej perspektywie czasowej. Jak już podkreśliśmy wyżej, konsekwencje każdego błędu pomiaru warunków początkowych mogą być niezwykle dramatyczne, narastają bowiem wykładniczo w czasie. Wysoka wrażliwość trajektorii na warunki początkowe zwykle uwidacznia się w postaci dodatniego tzw. wykładnika Lapunowa (por. [9], [28]). Wykładnik ten jest wykładniczą stopą wzrostu błędów. Inna charakterystyka systemów dynamicznych bazuje na teorii informacji. Nawiązuje ona do entropii w sensie Kołmogorowa–Sinaia (por. [20], [35]). Należy ją rozumieć jako miarę błędu prognozy stanu systemu, której horyzont wynosi jeden. Poza tym jeśli system dynamiczny jest chaotyczny, lecz przy tym deterministyczny, to możliwa jest krótkoterminowa prognoza. Taki szereg czasowy generowany przez ten system może być podstawą prognoz w krótkim horyzoncie czasowym. Istnieje wiele problemów nierozwiązanych, jeśli chodzi o wyniki testów dotyczących istnienia nieliniowości w szeregach czasowych. Własności chaotyczne przepływu gazu w rozdzielniach nie były dotąd badane w Polsce.

Głównym celem artykułu jest wykazanie, że nowe metody ekonometryczne bazujące na teorii chaosu mogą być wykorzystane do opisu struktury przepływu gazu w wybranych stacjach I stopnia. Rozpoznanie fraktalnych własności szeregów czasowych przepływu gazu może być pomocne w usprawnieniu i optymalizacji zarządzania systemami przesyłowymi gazu, co może zahamować wzrost kosztów przesyłu windujących cenę gazu dla odbiorców finalnych.

W następnym rozdziale omówimy krótko literaturę, głównie ekonometryczną, dotyczącą tematyki wykrywania i analizy chaosu, a także wskażemy kilka pozycji literatury poświęconych zagadnieniu kosztów przesyłu gazu. Rozdział 3 obejmuje głównie opis danych. Przedstawiono w nim w szczególności ich statystyki opisowe. W tym rozdziale zawarto też przegląd i opis stosowanych w pracy metod obliczeniowych.

Rozdział 4 poświęcony jest prezentacji wyników uzyskanych analiz oraz ich interpretacji. Podsumowanie i wnioski kończą pracę.

2. Przegląd literatury

Urodzony w Warszawie profesor matematyki Benoit B. Mandelbrot zajmuje się już od dawna teorią chaosu, szczególnie jeśli chodzi o rynki finansowe. Jest on odkrywcą tak zwanej „geometrii fraktalnej”. W swoich publikacjach [26] przedstawił wyniki badań i wyjaśnił ich sens w sposób przystępny także dla osób niebędących matematykami. Mandelbrot wyjaśnił, że jego modele nie są właściwym narzędziem (przynajmniej na obecnym etapie) do prognozowania kursów akcji czy walut, ale już teraz mogą być efektywnie zastosowane do minimalizacji ryzyka funduszu, depozytu i w innych celach. Jego zdaniem ryzyka na rynkach finansowych są nieprawidłowo szacowane. Są one o wiele większe niż się na ogół myśli. Przy tym można je znacznie lepiej oszacować niż wielu badaczy sądzi.

Niektóre towarzystwa ubezpieczeniowe i przedsiębiorstwa w swoich pracach nad możliwymi scenariuszami kształtowania się w przyszłości zjawisk takich jak liczba wypadków czy ścieżka wzrostu gospodarczego stosują już teraz fraktalną geometrię Mandelbrota.

Opis podstawowych metod teorii chaosu w odniesieniu do ekonomicznych szeregów czasowych można znaleźć np. w pracach Docknera, Prskawetza i Feichtingera [6]; Elsnera [10]; Franka *et al.* [11]; Hsieh [16] oraz Petersa [29], [30].

Nowe trendy zastosowań teorii fraktali w naukach technicznych prezentują Véhel i Lutton [37]. W swojej książce Kudrewicz [24] opisuje konstrukcję fraktali i ich zastosowania w nauce i technice, a także elementy teorii wymiaru.

Identyfikacja struktur chaotycznych w niektórych szeregach czasowych o dużej wariancji może być utrudniona przez obecność dużych składników losowych występujących w tych systemach dynamicznych. Metodyka testowania nieliniowych odstępstw od zachowania typowego dla białego szumu została opisana w pracy Scheinkmana i LeBarona [33]. Autorzy podkreślili, że jednym z możliwych podejść jest testowanie szeregu czasowego za pomocą estymatorów entropii i wymiaru w szeregu wyjściowym oraz w szeregu reszt z modelu liniowego w połączeniu z testem BDS (Brock, Dechert i Scheinkman) (zob. [1], [2], [3]). Można też szacować największy wykładnik Lapunowa i testować jego dodatniość.

Ważną rolę w teorii chaosu odgrywa tzw. atraktor. Jest to ukryte, trudne do zaobserwowania uporządkowanie badanego procesu. Poszukiwanie atraktorów jest bardzo istotne w wielu dziedzinach nauki. Bowiern znajomość atraktora umożliwia wywieranie wpływu na przebieg procesu.

Z atraktorami są ściśle związane tzw. wykładniki Lapunowa reprezentujące średnie wartości całego atraktora. Ich wersja globalna nie jest źródłem informacji o lokalnym tempie odbiegania trajektorii od warunków początkowych. W pewnych przypadkach mogą być one nawet ujemne. Ten ostatni fakt można by uważać za symptom zmniejszania się odległości pomiędzy trajektoriami w pewnych

obszarach atraktora. Dla ekonometryka jest to wskazówką, że istnieje – przynajmniej w pewnym zakresie – możliwość budowy krótkoterminowych prognoz. Generalnie jednak wykładnicze rozprzestrzenianie się błędów uniemożliwia prognozy. Dynamika stopy rozbieżności jest reprezentowana przez **efektywne wykładniki Lapunowa** ([14]), **lokalne wykładniki Lapunowa** (zob. [39]) oraz **prognozowalne wzorce** (zob. [7]).

Dla układów stochastycznych trudno przypisać wykładnikom Lapunowa konkretną interpretację. Jest to skutek nakładania się składowych deterministycznych i losowych w tego typu układach dynamicznych. W tym przypadku pojawiają się trudności w określeniu stopnia wrażliwości trajektorii rozwiązań na warunki początkowe (zob. [27]). Nakładanie się obu składowych jest źródłem tzw. **szumu dynamicznego**, któremu jest przeciwstawiany tzw. **szum pomiarowy** niewpływający na dynamikę systemu. Określenie **chaos szumowy** dotyczy układów chaotycznych zakłócanych **szumem dynamicznym**. W praktyce trudno ustalić za pomocą jakiego modelu układ dynamiczny może być najlepiej opisany (ze składową deterministyczną czy stochastyczną). Dlatego podchodzi się do modelowania zazwyczaj bardziej ogólnie, zakładając nieliniowe układy dynamiczne z szumem dynamicznym (zob. [5]).

W przypadku wielu układów dynamicznych obserwuje się niewielką korelację warunkowych średnich, które odpowiadają składowym deterministycznym. Istotna korelacja dotyczy składowych losowych. Własności składowych losowych, a zwłaszcza ich wariancji warunkowych, są głównymi przedmiotami badań. Chodzi tu o niejednorodność wariancji zwaną też z angielskiego heteroskedastycznością.

Chociaż wykładniki Lapunowa zostały określone tylko dla układów deterministycznych, to procedura szacowania wszystkich lub przynajmniej największego z nich (por. [8], [12], [18], [32] oraz [38]) ma miejsce także w przypadku, gdy szereg czasowy ma charakter białego szumu. Autorzy zwracają uwagę, że pewne procedury mogą prowadzić do tzw. pozornie dodatnich współczynników. Oprócz wartości punktowych wykładnika Lapunowa szacuje się także przedziały ufności dla tego współczynnika.

Schittenkopf *et al.* [[34] wykorzystali technikę bootstrapu w szacowaniu wykładników Lapunowa. Uzyskali zadowalające wyniki. Wprowadzili do literatury pewną miarę wrażliwości opartą na odległości pomiędzy trajektoriami, co może być uznane za rozszerzenie pojęcia wykładników Lapunowa. Ograniczeniem tej miary jest jednak to, że jest ona w odróżnieniu od wykładników Lapunowa miarą lokalną zależną od stanu.

W ostatnim dziesięcioleciu, kiedy nastąpiła demonopolizacja rynku gazowego w Polsce, zaczęto dużo większą uwagę zwracać na kwestię odpowiedniego zarządzania dystrybucją gazu, a szczególnie na optymalizację przepływów. Wiąże się to z wysoką i stale szybko rosnącą ceną tego surowca zużywanego w procesach

produkcyjnych, a przede wszystkim przez konsumentów indywidualnych. W cenie gazu dla finalnych odbiorców bardzo istotną pozycję stanowią koszty jego przesyłu. Dlatego racjonalizacja przepływów i właściwa ich struktura mogłyby zmniejszyć koszty przesyłu, a zatem i cenę detaliczną tego strategicznego surowca, który wykorzystują tak przedsiębiorstwa, jak i setki tysięcy gospodarstw domowych w Polsce. Firmy zajmujące się dostawą gazu poszukiwały optymalnych rozwiązań w tym zakresie. Efektem tych działań było również zainteresowanie naukowców tą problematyką. Z zagadnieniami zarządzania dystrybucją gazu można się zapoznać m.in. na podstawie lektury prac [4], [19] oraz [31].

Metody badania własności fraktalnych były stosowane w odniesieniu przede wszystkim do szeregów finansowych. Autorzy w niniejszej pracy podjęli próbę wykorzystania wybranych metod ekonometrycznych z zakresu analizy fraktalnej do badania własności przepływu gazu na wybranych stacjach I stopnia z Małopolski i Podkarpacia.

3. Dane i metodologia

3.1. Opis danych

W celu porównania badanych własności dla stacji o różnej charakterystyce do analizy wybrano pięć stacji I stopnia: Solec-Zdrój, Czechówka, zespół stacji Głogów i Miłocin oraz stację związaną z zakładem Huta Sendzimira. Poniżej został zamieszczony krótki opis analizowanych stacji.

Solec-Zdrój

Jest to mała stacja redukcyjno-pomiarowa o niewielkim znaczeniu dla systemu przesyłowego, której zadaniem jest pomiar oraz redukcja gazu. Stacja zasilana siecią dystrybucyjną średniego ciśnienia, z której zasilani są odbiorcy komunalni oraz kilka domów uzdrowiskowych. Brak połączeń na sieci średniego ciśnienia z innym punktem wyjścia powoduje, że stacja stanowi jedyne źródło zasilania dla odbiorców.

Czechówka

Obecnie jest to tzw. wirtualny punkt wyjścia z systemu przesyłowego, który docelowo zostanie zastąpiony przez aktualnie budowaną stację pomiarową. Z tego punktu jest zasilany gazociąg dystrybucyjny wysokiego ciśnienia DN300, na którym znajduje się 11 stacji redukcyjno-pomiarowych. Jest to ważny punkt w systemie ze względu na znaczne ilości przesyłanego gazu oraz zasilanie odbiorców na bardzo dużym obszarze.

Zespół stacji Głogów Małopolski i Miłocin

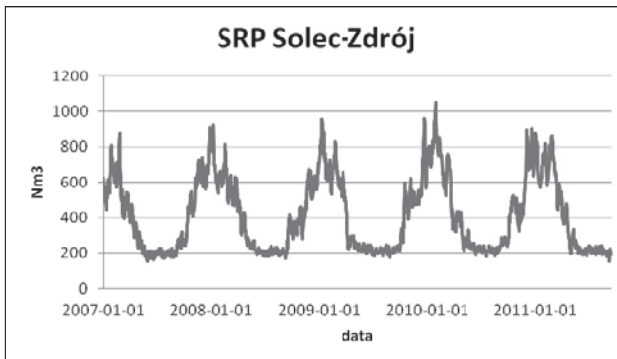
Zespół dwóch stacji redukcyjno-pomiarowych zasilających obszar sieci dystrybucyjnej średniego ciśnienia, do której są podłączeni odbiorcy przemysłowi oraz komunalni. Stacje są zasilane z dwóch niezależnie pracujących gazociągów wysokiego ciśnienia, stacja Głogów Małopolski z gazociągu o średnicy DN700, a stacja Miłocin z gazociągu DN400. Ze względu na połączenia hydrauliczne sieci średniego ciśnienia w sezonie letnim stacja Miłocin może zostać wyłączona, a całe zasilanie sieci średnioprężnej może być realizowane przez stację Głogów Małopolski. Wyłączenie stacji Głogów Małopolski nie jest możliwe nawet w sezonie letnim ze względu na konieczność zasilania czterech odbiorców przemysłowych.

ArcelorMittal Poland S.A. Oddział w Krakowie (dawna nazwa Huta Sendzimir)

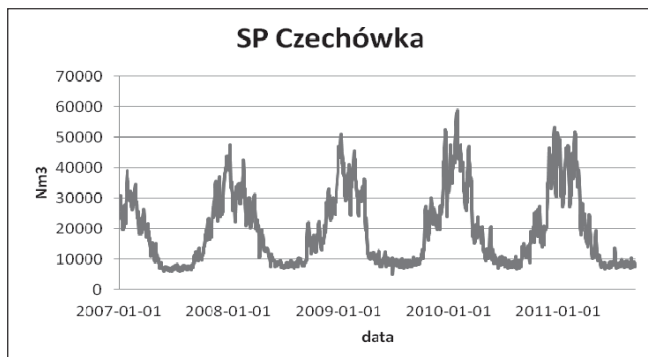
Huta pod względem rocznego zużycia gazu jest jednym z największych odbiorców przemysłowych zasilanych z sieci Oddziału Tarnów. Odbiorca ten charakteryzuje się bardzo dużą zmiennością poboru gazu przez co ma znaczący wpływ na pracę systemu przesyłowego. Jest to odbiorca końcowy, którego zakład jest zasilany bezpośrednio ze stacji pomiarowej na wysokim ciśnieniu. Ze względu na duże znaczenie stacja jest zasilana z dwóch niezależnie pracujących gazociągów przesyłowych.

Do badania wykorzystano dane dzienne wyrażone w jednostce Nm^3 (normalny metr sześcienny). Jest to jednostka określająca objętość gazu o ciśnieniu 1 atmosfery i temperaturze 1 stopnia Celsjusza. Analizowane szeregi dotyczą przepływu na poszczególnych stacjach od stycznia 2007 do września 2011.

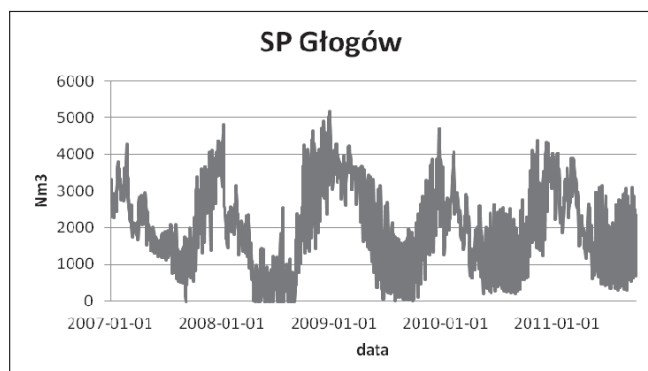
Wstępną charakterystykę danych będącą równocześnie pierwszym etapem analizy porównawczej badanych szeregów czasowych rozpoczęto od wykresów tych szeregów oraz wyznaczenia podstawowych statystyk i przeprowadzenia testów normalności. Dane przedstawione na rysunkach 1–5 oraz statystyki zamieszczone w tabeli 1 zostały przeskalowane w celu zachowania tajemnicy handlowej.



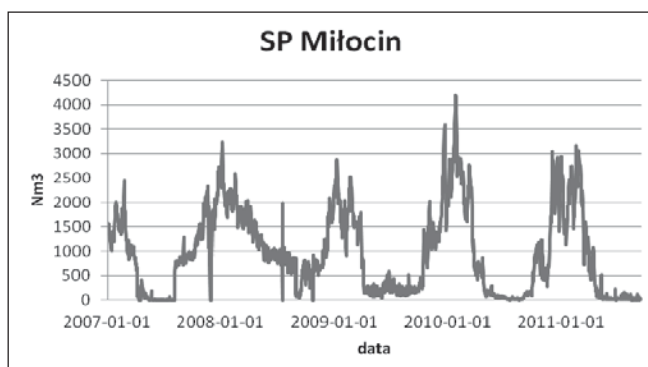
Rys. 1. Wykres dziennego przepływu gazu na stacji Solec-Zdrój



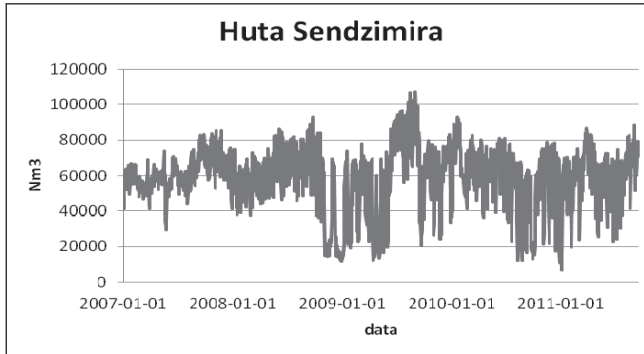
Rys. 2. Wykres dziennego przepływu gazu na stacji Czechówka



Rys. 3. Wykres dziennego przepływu gazu na stacji Głogów



Rys. 4. Wykres dziennego przepływu gazu na stacji Miłocin



Rys. 5. Wykres dziennego przepływu gazu na stacji Huta Sendzimira

Przedstawione w tabeli 1 podstawowe statystyki dla poszczególnych stacji potwierdzają ich różnorodność. Istotną różnicę widać w średniej wielkości dziennego przepływu gazu. W przypadku stacji Solec kształtuje się on na poziomie $495,8 \text{ Nm}^3$, a dla stacji Huta Sendzimira $71\,509 \text{ Nm}^3$. Dzienny przepływ również kształtuje się na zróżnicowanym poziomie: od najmniejszego dla stacji Huta Sendzimira (30%) do najwyższego na stacji Miłocin (90%). Dane pochodzące z pierwszych czterech stacji wskazują na podobne właściwości w zakresie asymetrii i rozproszenia.

Tabela 1
Podstawowe statystyki wielkości przepływu gazu

stacja	średnia	odchylenie standardowe	współczynnik zmienności	minimum	maksimum	skośność	kurtoza
Solec-Zdrój	495,8	252,9	51,02%	187,4	1266,2	0,1222	-0,1777
Czechówka	23504,2	14436,5	61,42%	5765,6	70517	0,1717	-0,0632
Głogów	2576,3	1307,5	50,75%	0	6222,6	0,0161	-0,1350
Miłocin	1163,6	1042,8	89,63%	0	5047	0,1403	-0,0646
Huta Sendzimira	71509	21889,6	30,61%	8425	128757	-0,1498	0,0878

Źródło: obliczenia własne

Wyniki testów normalności wskazują na brak dobrego dopasowania rozkładu normalnego do wielkości przepływu gazu na poszczególnych stacjach (por. tab. 2).

Tabela 2
Wyniki testów normalności badanych szeregów

Stacja	Solec-Zdrój		Czechówka		Głogów		Miłocin		Huta Sendzimira	
	Statystyka	<i>p-value</i>	Statystyka	<i>p-value</i>	Statystyka	<i>p-value</i>	Statystyka	<i>p-value</i>	Statystyka	<i>p-value</i>
chi-kwadrat	1856,06	0,0	1587,8	0,0	194,7	1,1E-16	781,3	0,0	287,65	0,0
Shapiro-Wilka	0,85923	0,0	0,84	0,0	0,956	0,0	0,892	0,0	0,93	0,0
skośności Z-score	5,5504	2,85E-8	7,7	1,28E-14	1,903	0,0570	6,997	2,62E-12	2,874	0,004
kurtozy Z-score	-10,645	0,0	0,048	0,96	-9,39	0,0	1,482	0,138	-7,96	1,7E-15

Źródło: obliczenia własne

3.2. Metodologia

Głównym celem artykułu jest wykorzystanie wybranych metod statystycznych do identyfikacji w badanych szeregach czasowych występowania struktur nieliniowych oraz chaosu deterministycznego. Autorzy wybrali cieszące się dużą popularnością tak wśród teoretyków, jak i praktyków metody statystyczne, tj. analizę R/S , test BDS , wymiar korelacyjny oraz największy wykładnik Lapunowa.

Analiza R/S

Głównym celem techniki nieparametrycznej, którą jest analiza przeskalowanego rozstępu (ang. *rescaled range*, ozn. R/S) jest wyznaczenie tzw. wykładnika Hursta. Tę technikę istotnie rozwinęli Mandelbrot [26] i Lo [25].

Model Hursta można przedstawić wzorem:

$$\left(\frac{R}{S}\right)_n = c \cdot n^H \quad (1)$$

gdzie S oznacza odchylenie standardowe przyrostów badanego szeregu, c – jest stałą dodatnią, n – jest liczbę elementów szeregu czasowego. Parametr H jest wykładnikiem Hursta. W przypadku procesu (szeregu czasowego) czysto losowego, czyli białego szumu, wraz ze wzrostem n wartość H z równania (1) powinna być coraz bliższa 0,5. Jeśli tak nie jest, to szereg czasowy może wykazywać tzw. długą pamięć (ang. *long memory*).

Jeśli wartość estymatora H przekracza 0,5, to badany szereg czasowy ma własność persystencji, co w szczególności oznacza, że średnia jest osiągana przy większej liczbie obserwacji niż w przypadku procesu o przyrostach niezależnych, a więc mamy do czynienia z długą pamięcią (długoterminową zależnością).

Jeśli wartość estymatora H jest mniejsza od 0,5, to szereg czasowy nazywa się antypersystentnym. Proces szybciej osiąga stan równowagi, niż ma to miejsce w przypadku procesu o przyrostach niezależnych. Mówi się, że w takich szeregach występuje zjawisko „powrotu do średniej” (ang. *mean reverting process*).

Wiele szeregów czasowych opisujących zjawiska przyrodnicze, techniczne i ekonomiczne wykazuje persystencję i ma długą pamięć.

Typowa analiza R/S dla szeregu czasowego przebiega w następujących etapach:

1. Wyjściowy szereg dzieli się na podszeregi czasowe liczące po n obserwacji.
2. Następnie przeprowadza się centrowanie każdego podszeregu, odejmując od każdej jego wartości średnią arytmetyczną, co daje szereg: z_1, z_2, \dots, z_N .
3. Wyznacza się skumulowane szeregi czasowe, według wzoru: $y_i = \sum_{j=1}^i z_j$.

4. W kolejnym kroku oblicza się rozstępy skumulowanych szeregów czasowych, zgodnie ze wzorem: $R_t = \max(y_1, y_2, \dots, y_t) - \min(y_1, y_2, \dots, y_t)$.
5. Dzieląc rozstęp przez odchylenie standardowe szeregu, otrzymuje się przeskalowane rozstępy (dla każdego skumulowanego szeregu czasowego).

Powtarzając 1–5 dla różnych długości N szeregu czasowego otrzymuje się zależność wielkości R/S od długości N . Wartość wykładnika Hursta, po zlogarytmowaniu (1), jest współczynnikiem kierunkowym prostej:

$$\ln(R/S) = \ln(c) + H \ln(n) \quad (2)$$

Wykładnik Hursta określa tangens kąta nachylenia prostej regresji otrzymanej w skali podwójnie logarytmicznej $(R/S)_n$ względem $\ln(n)$.

W badaniu należy zwracać uwagę na liczbę dzielników liczby określającej długość szeregu. Liczba N winna mieć jak najwięcej dzielników będących liczbami naturalnymi. Niekiedy wymaga to odjęcia kilku najstarszych danych. Występowanie w szeregach cykli, czyli procesów o skończonej pamięci, może zniekształcać wyniki wyznaczone za pomocą opisanego algorytmu. Dlatego wykres $\ln(R/S)_n$ względem $\ln(n)$ nie powinien zawierać punktów $\ln n^*$, na których prosta istotnie zmienia swoje nachylenie. Jeśli znajdzie się taki punkt, to wtedy do oszacowania wartości H trzeba ograniczyć się tylko do wartości $(R/S)_n$ dla $n < n^*$. Jeśli dla $n \geq n^*$ współczynnik regresji jest bliski wartości 0,5 lub nieistotny, to pamięć w systemie jest skończona i szereg czasowy został wygenerowany przez biały szum. Wartość n^* interpretuje się jako średnią długość cyklu nieokresowego.

Badanie istotności wartości wykładnika Hursta przeprowadza się przez jej porównanie z wartością teoretyczną podaną przez Anisa i Lloyda w postaci

$$E(R/S)_n = \begin{cases} \frac{n - \frac{1}{2}}{n} \cdot \frac{\Gamma(\frac{n-1}{2})}{\sqrt{\pi} \cdot \Gamma(\frac{n}{2})} \cdot \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}} & \text{dla } n \leq 300 \\ \frac{n - \frac{1}{2}}{n} \cdot \sqrt{\frac{2}{\pi \cdot n}} \cdot \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}} & \text{dla } n > 300 \end{cases}$$

Jeśli zachodzi

$$|H - E(H)| \geq \sigma \quad (3)$$

to wykładnik Hursta jest istotny. $E(H)$ jest wartością oczekiwaną wykładnika Hursta otrzymaną z regresji $E(R/S)_n$, zaś σ jest odchyleniem standardowym H , wynoszącym $\sigma = \sqrt{1/N}$ dla szeregu czasowego o długości N .

Analiza R/S pozwala sprawdzić, czy badany szereg jest losowy, czy też nie. Można też za jej pomocą określić długość cykli, jeśli wykazuje je szereg czasowy. Chodzi tu nie tylko o regularne cykle okresowe, ale również cykle nieokresowe.

Wymiar korelacyjny, test BDS, największy wykładnik Lapunowa

Wymiar korelacyjny może być wyznaczony nie tylko za pomocą współczynnika Hursta, ale i za pomocą nowszej bardziej dokładnej metody zaproponowanej przez Grassbergera i Procaccia [13]. Jej istota polega na odtwarzaniu wielowymiarowego atraktora na podstawie jednowymiarowego szeregu czasowego. Dane wyższych wymiarów otrzymuje się jako przesunięte dane jednowymiarowego szeregu czasowego. W procedurze tej konieczne jest określenie wymiaru atraktora poprzez zwiększanie wymiaru szeregu danych, zwanego wymiarem zanurzenia. Dla każdego z otrzymywanych szeregów danych określa się wymiar korelacyjny. Jeśli na którymś etapie tej iteracyjnej procedury zwiększeniu wymiaru danych nie towarzyszy dalszy wzrost wymiaru korelacyjnego, to jest to właśnie poszukiwany wymiar.

Rozważmy szereg czasowy x_1, x_2, \dots, x_T . Niech $X_{t,m}$ oznacza wektor m -wymiarowy postaci $X_{t,m} = (x_t, x_{t+1}, \dots, x_{t+m-1})$.

Wymiar korelacyjny określa się, wyznaczając tzw. całkę korelacyjną, według wzoru:

$$C_{m,N}(\varepsilon) = \frac{2}{N(N-1)} \cdot \sum_{1 \leq t < s \leq N} H(\varepsilon - \|X_{t,m} - X_{s,m}\|) \quad (4)$$

przy czym $N = T - m + 1$, $H(x)$ jest funkcją Heaviside'a określoną wzorem:

$$H = \begin{cases} 0 & \text{dla } x < 0 \\ 1 & \text{dla } x \geq 0 \end{cases}$$

natomiast $\| \cdot \|$ oznacza dowolną normę w przestrzeni m -wymiarowej (wartość tej całki nie zależy od wyboru normy).

Całkę tę określa prawdopodobieństwo tego, że dwa losowo wybrane elementy ciągu są oddalone (w sensie wybranej normy) mniej niż o wielkość ε . Wartość wymiaru korelacyjnego jest granicą iterowaną:

$$D_2 = \lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \frac{\ln C_{m,N}(\varepsilon)}{\ln \varepsilon} \quad (5)$$

Współczynnik kierunkowy regresji zmiennej $\ln C_{m,N}(\varepsilon)$ względem $\ln \varepsilon$ przy zastosowaniu niewielkich wartości ε określa przybliżoną wartość wymiaru korelacyjnego.

Wymiar korelacyjny pozwala na ustalenie minimalnej liczby zmiennych dynamicznych potrzebnych do opisanie układu. Wyznacza on zatem dolną granicę liczby możliwych stopni swobody.

Test *BDS* bazujący na pojęciu całki korelacyjnej pozwala wykrywać chaos deterministyczny. Za jego pomocą bada się hipotezę zerową, która głosi, że analizowany szereg czasowy jest ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie czyli i.i.d. (ang. *independent and identically distributed*). Test *BDS* pozwala na wykrycie zależności nieliniowych w badanym szeregu czasowym, bo mogą one świadczyć o występowaniu chaosu.

Statystyka testowa wyraża się wzorem:

$$BDS = \sqrt{N} \left[C_{m,N}(\varepsilon) - (C_{1,N}(\varepsilon))^m \right] \quad (6)$$

Ma ona asymptotyczny rozkład normalny o średniej 0 i wariancji zależnej od parametrów ε oraz m .

Jak to już podkreślono we wstępie, największy wykładnik Lapunowa jest miarą wrażliwości przebiegu trajektorii obserwacji na warunki początkowe, czyli szybkości rozchodzenia się bliskich sobie trajektorii układu.

Wykładnik ten można wyznaczyć, znając wymiar korelacyjny, tzw. średni czas orbitalny oraz opóźnienie czasowe. W tym celu na atraktorze, którego wymiar korelacyjny został oszacowany, należy wybrać dwa punkty odległe od siebie o co najmniej jeden okres orbitalny. Odległość między tymi punktami oznacza się przez $L(t_i)$. Następnie oblicza się odległość wybranych punktów po upływie pewnego czasu. Oznaczamy ją przez $L'(t_{i+1})$. Wówczas największy wykładnik Lapunowa wyraża się wzorem:

$$L_1 = \frac{1}{t} \sum_{i=1}^m \log_2 \frac{L'(t_{i+1})}{L(t_i)} \quad (7)$$

gdzie m oznacza liczbę okresów, natomiast t okres badania.

Badany szereg czasowy jest chaotyczny, jeśli największy wykładnik Lapunowa ma wartość dodatnią. Poza tym znajomość tego wykładnika pozwala na określenie czasu zaniku informacji o systemie. Ten fakt ma istotny wpływ na możliwości i jakość predykcji na podstawie badanego szeregu czasowego.

4. Wyniki badań

Prezentację wyników badań zaczniemy od analizy *R/S* i estymacji współczynnika Hursta.

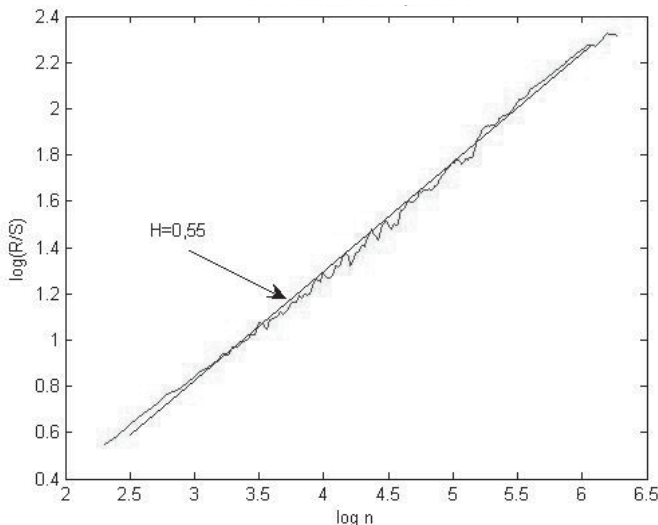
4.1. Wyniki analizy *R/S* i estymacji współczynnika Hursta

Próbie wykrycia cykli oraz określenia ich długości rozpoczęto od wykreślenia wartości $\log(R/S)_n$ w zależności od $\log(n)$ dla $n = 10, \dots, N/2$. Zagęszczenie

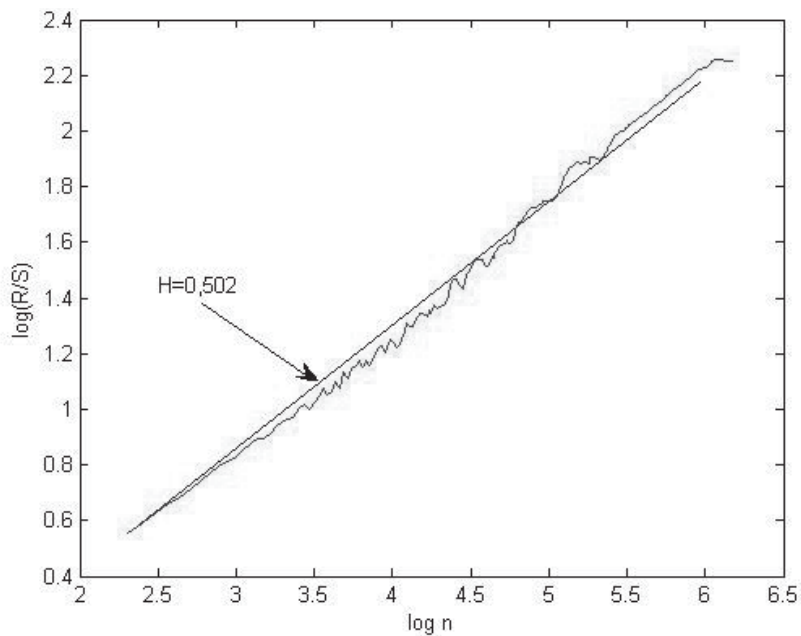
punktów krzywej przez użycie do wyznaczenia wartości $(R/S)_n$, wszystkich liczb naturalnych n z przedziału $[10, N/2]$ pozwoli na dokładniejsze określenie wartości $\log(n^*)$ (o ile istnieje), w której następuje istotna zmiana nachylenia krzywej. Na wykresach zaznaczono również proste regresji dla $n \in [10, n^*]$. Współczynnik kierunkowy prostej uzyskanej w wyniku regresji liniowej dla $n \in [10, n^*]$ jest jednocześnie wartością wykładnika Hursta uzyskanego za pomocą zmodyfikowanego algorytmu. Zmiana nachylenia badanych wykresów i jego ustalenie się na poziomie zbliżonym do 0,5 oznacza, że badany proces traci pamięć i staje się losowy. Punkt, w którym krzywa się „odgina”, uznajemy za wartość określającą długość pamięci procesu i jednocześnie za średnią długość cyklu nieokresowego.

Wyniki analizy R/S zilustrowane na rysunkach 6–10 oraz w tabeli 3 wskazują, iż w czterech przypadkach wykładnik Hursta jest istotnie różny od wartości teoretycznej. Mianowicie dla stacji: Czechówka, Głogów oraz Miłocin wykładnik Hursta jest istotnie większy od wartości 0,5, co oznacza, że te szeregi wykazują własności persystentne, czyli są szeregami z długą pamięcią. Dla danych pochodzących ze stacji Huta Sendzimir wykładnik Hursta jest istotnie mniejszy od 0,5, co wskazuje na proces powracający do średniej. Natomiast wyniki dla stacji Solec-Zdrój sugerują, że mamy do czynienia z procesem czysto losowym.

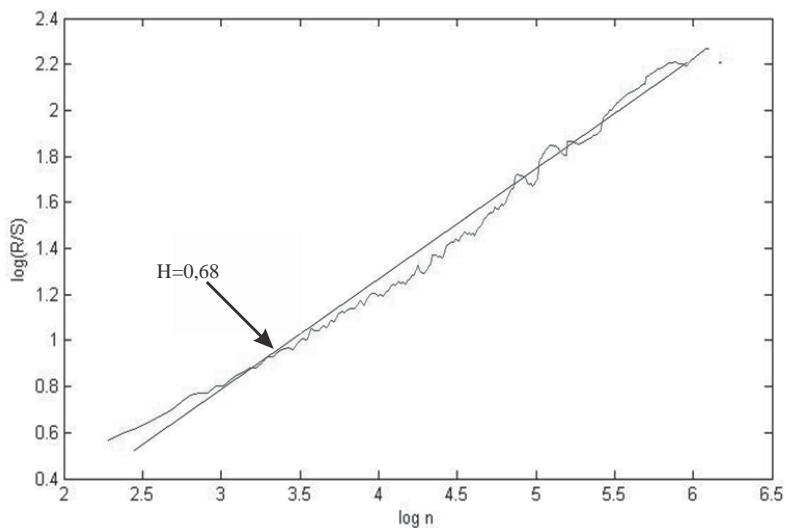
We wszystkich przypadkach trudno jest jednak wskazać punkt załamania się krzywej na wykresach przedstawionych wyżej, co utrudnia wyznaczenie długości cykli na podstawie analizy R/S .



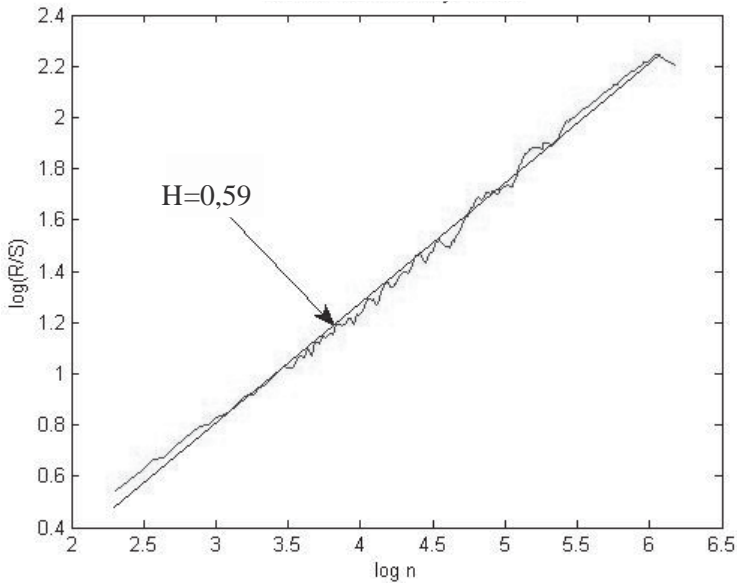
Rys. 6. Analiza R/S dla dziennego przepływu gazu na stacji Solec-Zdrój



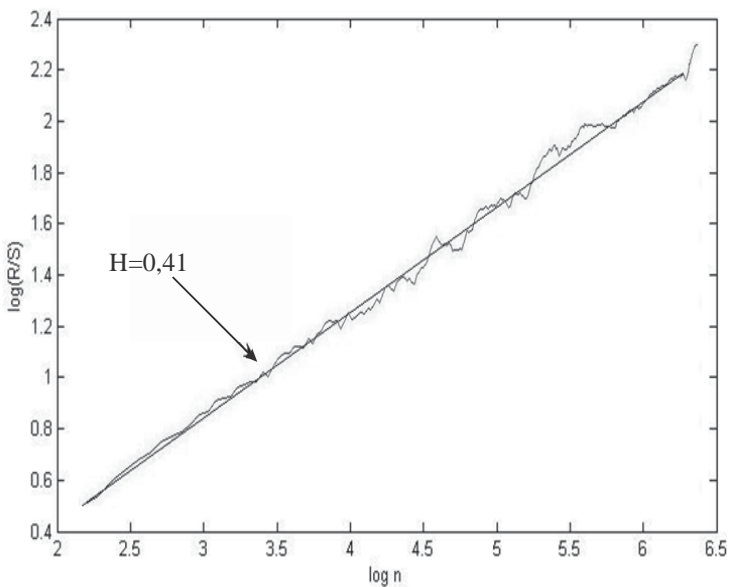
Rys. 7. Analiza R/S dla dziennego przepływu gazu na stacji Czechówka



Rys. 8. Analiza R/S dla dziennego przepływu gazu na stacji Głogów



Rys. 9. Analiza R/S dla dziennego przepływu gazu na stacji Miłocin



Rys. 10. Analiza R/S dla dziennego przepływu gazu na stacji Huta Sendzimir

Tabela 3
Wartości empiryczna i teoretyczna wykładnika Hursta

Stacja	Wartość empiryczna wykładnika Hursta (H)	Wartość teoretyczna (E(H))	$ E(H)-H $
Solec-Zdrój	0,55	0,56	0,01
Czechówka	0,502	0,56	0,058*
Głogów	0,68	0,65	0,03*
Miłocin	0,59	0,56	0,03*
Huta Sendzimira	0,41	0,56	0,15*

Źródło: obliczenia własne

Uwaga: * oznacza istotność wykładnika Hursta. Wartość krytyczna dla testu wynosi $\sigma = 0,024$.

4.2. Wyniki testu BDS, estymatory wymiaru korelacyjnego i największego wykładnika Lapunowa

Wyniki badań empirycznych zamieszczono w tabeli 4.

Tabela 4
Wyniki testu BDS dla badanych stacji (do 2 miejsc po przecinku)

Stacja	2	3	4	5	6	7	8	9	10
Solec-Zdrój	120,2	133,19	148,8	170,9	201,0	241,3	294,8	365,7	459,8
Czechówka	76,3	80,9	87,11	96,0	108,1	123,8	144,0	169,6	202,2
Głogów	76,82	81,14	85,102	90,91	99	110,67	129,48	152,09	179,58
Miłocin	87,09	92,928	100,02	110,2	124,14	142,37	165,85	195,77	233,81
Huta Sendzimira	62	65,13	68,11	72,26	77,67	84,44	93,22	104,5	119,08

Źródło: obliczenia własne

Uwaga: * oznacza, że nie ma podstaw do odrzucenia hipotezy, że dane są realizacją procesu i.i.d. przy poziomie istotności 1%. Wartość krytyczna dla testu wynosi 2,58.

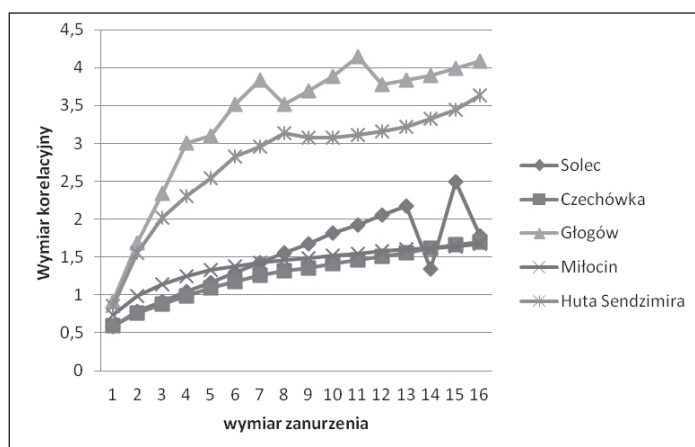
Tabela 5
Wyniki estymacji wymiaru korelacyjnego w zależności od wymiaru zanurzenia (maks. 16)

Stacja/ wymiar zanurzenia	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Solec-Zdrój	0,58	0,79	0,92	1,05	1,17	1,3	1,43	1,56	1,68	1,82	1,93	2,06	2,17	1,34	2,5	1,78
Czechówka	0,6	0,76	0,88	0,99	1,09	1,18	1,26	1,32	1,36	1,41	1,46	1,51	1,56	1,61	1,66	1,7
Głogów	0,92	1,69	2,34	3,01	3,1	3,51	3,84	3,52	3,69	3,88	4,15	3,78	3,84	3,9	3,99	4,08
Milocin	0,74	0,99	1,14	1,25	1,33	1,38	1,42	1,46	1,49	1,52	1,55	1,58	1,6	1,62	1,64	1,67
Huta Sendzimira	0,86	1,56	2,02	2,31	2,54	2,83	2,96	3,13	3,08	3,08	3,11	3,16	3,22	3,32	3,44	3,64

Źródło: obliczenia własne

Analizując wyniki testu *BDS* zamieszczone w tabeli 4, stwierdzamy, iż we wszystkich przypadkach mamy podstawy do odrzucenia hipotezy, że szereg jest realizacją procesu i.i.d. Może to być przesłanką do stwierdzenia, iż mamy do czynienia z procesami generowanymi przez układy chaotyczne.

Przy analizie rezultatów obliczeń wymiaru korelacyjnego dla wielkości przepływu gazu (por tab. 5, rys. 11) widać wyraźną zbieżność wymiaru korelacyjnego przy zwiększaniu wymiaru zanurzenia dla wszystkich stacji. W przypadku stacji Czechówka oraz Miłocin zauważamy, że ograniczeniem górnym (całkowitym) jest 2, dla stacji Solec-Zdrój 3, natomiast dla stacji Głogów i Huta Sendzimira wartość ta wynosi 4. Liczby te określają zatem wymiar korelacyjny badanych szeregów, a jednocześnie są oszacowaniem z dołu wymiaru atraktora i określają liczbę stopni swobody układu. Zatem podane wartości określają minimalną liczbę zmiennych potrzebnych do opisu układu. Uzyskane wyniki sugerują, że liczba zmiennych potrzebnych do opisu badanych układów jest niewielka. Stwarza to możliwość ich wyznaczenia oraz wskazania zależności między nimi.



Rys. 11. Zmiany wartości wymiaru korelacyjnego dla poszczególnych stacji w zależności od wymiaru zanurzenia

Wyniki uzyskane z estymacji największego wykładnika Lapunowa (por. tab. 6) sugerują, że w pierwszych czterech przypadkach mamy do czynienia z szeregami czasowymi będącymi realizacjami procesów o charakterze chaotycznym. Jak już wspominaliśmy, wykładnik Lapunowa mówi nam, w jakim tempie tracimy informację o układzie. Przykładowo dla stacji Solec-Zdrój wartość wykładnika 0,0023 oznacza, że w ciągu każdego dnia tracimy 0,0023 bita informacji o układzie.

Zatem całkowita utrata informacji nastąpi po $1/0,0023 = 434$ dniach. W przypadku szeregu czasowego związanego ze stacją Huta Sendzimira mamy do czynienia z procesem powracającym do średniej, czyli takim, który nie jest generowany przez układ chaotyczny.

Tabela 6

Wyniki estymacji największego wykładnika Lapunowa dla badanych szeregów czasowych

Indeks	Wykładnik Lapunowa	Czas utraty informacji o systemie
Solec-Zdrój	0,0023	434
Czechówka	0,006	166
Głogów	0,07	14
Miłocin	0,034	30
Huta Sendzimira	-0,19	-

Źródło: obliczenia własne

6. Wnioski

W pracy wykazano przydatność metod ekonometrycznych do analizy szeregów czasowych przepływu gazu w wybranych stacjach I stopnia.

Wyniki przeprowadzonych badań nad występowaniem chaosu deterministycznego w szeregach czasowych wielkości przepływów na wybranych stacjach są niejednoznaczne.

Z jednej strony zbieżność wymiarów korelacyjnych oraz niska wartość ich ograniczenia górnego świadczy o występowaniu w systemach fraktali niskowymiarowych. Potwierdzają tę tezę również wartości największego wykładnika Lapunowa, które w czterech przypadkach są dodatnie. Z drugiej strony wartości wykładnika Hursta, pomimo iż w czterech przypadkach okazały się istotne, są bliskie 0,5, co świadczyłoby przeciwko tezie o występowaniu chaosu deterministycznego w badanych procesach.

Badania dodatkowo wykazały, iż dla stacji Huta Sendzimira wyznaczone wielkości (wykładnik Hursta, wykładnik Lapunowa) różnią się dość istotnie od tych samych miar dla pozostałych stacji. Ten fakt może być efektem tego, iż ta stacja obsługuje tylko jednego odbiorcę, który planuje z wyprzedzeniem swoje działania.

Uzyskane wstępne wyniki wymagają zatem dalszych badań i analiz, z wykorzystaniem danych dla jeszcze innych stacji, w celu potwierdzenia diagnozy od-

nośnie do chaotycznej struktury wielkości przepływu gazu na stacjach I stopnia. Definitywne potwierdzenie hipotezy o chaotycznej strukturze szeregów czasowych przepływu gazu dałoby możliwość zastosowania teorii chaosu w zarządzaniu systemami stacji I stopnia.

Literatura

- [1] Brock W.A., *Distinguishing random and deterministic systems: Abridged version*, „Journal of Economic Theory”, 1986, vol. 40, s. 168–195.
- [2] Brock W.A., Dechert W.D., Scheinkman J.A., *A test for independence based on the correlation dimension*, Social Systems Research Institute, University of Wisconsin–Madison, 1987, Working paper no. 8702.
- [3] Brock W.A., Hsieh D., LeBaron B., *A Test of Nonlinear Dynamics, Chaos, and Instability*, Cambridge: MIT Press, 1991.
- [4] Bytnar K., Kogut K., *Technika komputerowa w bezpiecznym zarządzaniu pracą sieci gazowych*, Nafta Gaz / Instytut Górnictwa Naftowego i Gazownictwa, Instytut Technologii Nafty, Stowarzyszenie Inżynierów i Techników, 2007, s. 139–144.
- [5] Chan K.S., Tong H., *A note on noisy chaos*, „Journal of the Royal Statistical Society B”, 1994, Vol. 56, s. 301–311.
- [6] Dockner E.J., Prskawetz A., Feichtinger G., *Non-linear dynamics and predictability in the Austrian stock market*, w: C. Heij *et. al.* (eds.), *System Dynamics in Economic and Financial Models*. John Wiley Press, 1997, s. 45–62.
- [7] Doerner R., Huebinger B., Martienssen W., *Predictability portraits for chaotic motions*, *Chaos, Solitons & Fractals*, 1991, Vol. 1, s. 553–571.
- [8] Eckmann J.P., Kamphorst S.O., Ruelle D., Scheinkman J., *Lyapunov exponents for stock returns*, w: *The Economy as an Evolving Complex System*. SFI Studies in the Science of Complexity. Reading, MA: Addison-Wesley, 1988, s. 301–304.
- [9] Eckmann J.P., Ruelle D., *Ergodic theory of chaos and strange attractors*, „Reviews of Modern Physics”, 1985, Vol. 57, s. 617–656.
- [10] Elsner J., *Chaos und Zufall am deutschen Aktienmarkt*, Heidelberg, Germany: Physica-Verlag, 1996.
- [11] Frank M., Gencay R., Stengos T., *International chaos?*, „European Economic Review”, 1988, Vol. 32, s. 1569–1584.
- [12] Gencay R., Dechert W.D., *An algorithm for the Lyapunov exponents of an n-dimensional unknown dynamical system*, „Physica D”, 1992, Vol. 59, s. 142–157.
- [13] Grassberger P., Procaccia I., *Characterization of strange attractors*, „Physical Review Letters”, 1983, Vol. 50, s. 346–349.

- [14] Grassberger P., Badii R., Politi A., *Scaling laws for hyperbolic and nonhyperbolic attractors*, „Journal of Statistical Physics”, 1988, Vol. 51, s. 135–178.
- [15] Gurgul H., Suder M., *Nieliniowa dynamika indeksów giełdowych WIG20 i ATX: analiza porównawcza*, „Ekonomia Menadżerska”, 2010, Vol. 7, s. 103–120.
- [16] Hsieh D.A., *Chaos and nonlinear dynamics: Applications to financial markets*, „Journal of Finance”, 1991, Vol. 46, s. 1839–1877.
- [17] Hristu-Varsakelis D., Kyrtsov C., *Evidence for nonlinear asymmetric causality in US inflation, metal and stock returns*, „Discrete Dynamics in Nature and Society”, 2008, Vol. 2008, Article ID 138547, 7 pages, doi: 10.1155/2008/138547.
- [18] Kadtke J.B., Brush J., Holzfuss J., *Global dynamical equations and Lyapunov exponents from noisy chaotic time series*, „International Journal of Bifurcation and Chaos”, 1993, Vol. 3, s. 607–616.
- [19] Kogut K., *Możliwości wykorzystania sieci neuronowych w analizie pracy sieci przesyłowej gazu ziemnego*, „Nowoczesne Gazownictwo”, 2004, Vol. 3, s. 5–8.
- [20] Kolmogorov A.N., *A metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces*, Doklady Akademii Nauk SSSR, 1958, Vol. 119, s. 861–864.
- [21] Kyrtsov C., Labys W., *Evidence for chaotic dependence between US inflation and commodity prices*, „Journal of Macroeconomics”, 2006, Vol. 28 (1), s. 256–266.
- [22] Kyrtsov C., Labys W., *Detecting positive feedback in multivariate time series: the case of metal prices and US inflation*, „Physica A”, 2007, Vol. 377 (1), s. 227–229.
- [23] Kyrtsov C., Vorlow C., *Complex dynamics in macroeconomics: A novel approach*, w: Diebolt C., Kyrtsov C. (eds.), *New Trends in Macroeconomics*, Springer-Verlag, Berlin, 2005.
- [24] Kudrewicz J., *Fraktale i chaos*, WNT, Warszawa, 2009.
- [25] Lo A.W., *Long term memory in stock market prices*, „Econometrica”, 1991, Vol. 59, s. 1279–1313.
- [26] Mandelbrot B., *Statistical Methodology for Non-Periodic Cycles: from the covariance to R/S analysis*, „Annals of Economic and Social Measurement”, 1972, Vol. 1.
- [27] Nychka D., Ellner S., Gallant A.R., McCaffrey D., *Finding chaos in noisy systems*, „Journal of the Royal Statistical Society B”, 1992, Vol. 54, s. 399–426.
- [28] Oseledec V.I., *A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems*, Transactions of the Moscow Mathematical Society, 1968, Vol. 19, s. 197–231.

- [29] Peters E.E., *Chaos and Order in the Capital Markets*, New York, Wiley, 1991.
- [30] Peters E.E., *Teoria chaosu a rynki kapitałowe*, WIG Press, Warszawa, 1997.
- [31] Polak D., *SmartGaz – nowoczesny symulator do zarządzania sieciami gazowniczymi*, I Krakowska Konferencja Młodych Uczonych, 2006, s. 135–144.
- [32] Sano M., Sawada Y., *Measurement of the Lyapunov spectrum from a chaotic time series*, „Physical Review Letters”, 1985, Vol. 55, s. 1082–1085.
- [33] Scheinkman J.A., LeBaron B., *Nonlinear dynamics and stock returns*, „Journal of Business”, 1989, Vol. 62, s. 311–337.
- [34] Schittenkopf C., Dorffner G., Dockner E.J., *On Nonlinear, Stochastic Dynamics in Economic and Financial Time Series*, „Studies in Nonlinear Dynamics and Econometrics”, 2000, Vol. 4 (3), s. 101–121.
- [35] Sinai Y.G., *On the concept of entropy for a dynamic system*, Doklady Akademii Nauk SSSR, 1959, Vol. 124, s. 768–771.
- [36] Sneyers R., *Climate Chaotic Instability: Statistical Determination and Theoretical Background*, „Environmetrics”, 1997, Vol. 8 (5), s. 517–532.
- [37] Véhel J.L., Lutton E., *Fractals in engineering: new trends in theory and applications*, Springer, London, 2005.
- [38] Wolf A., Swift J.B., Swinney H.L., Vastano J.A., *Determining Lyapunov exponents from a time series*, „Physica D”, 1985, Vol. 16, s. 285–317.
- [39] Wolff R.C., *Local Lyapunov exponents: Looking closely at chaos*, „Journal of the Royal Statistical Society B”, 1992, Vol. 54, s. 353–371.