

Marian BŁACHUTA

PIPS: WERSJA DYDAKTYCZNA PROGRAMU IDENTYFIKACJI I PRZEDYKCJI SZEREGÓW CZASOWYCH *

Streszczenie. W artykule przedstawiono program realizujący identyfikację i predykcję szeregów czasowych stacjonarnych i pewnej klasy szeregów niestacjonarnych, zawierających deterministyczne składowe wielomianowe. Dla celów dydaktycznych program umożliwia porównanie różnych estymatorów.

PIPS: A DIDACTICAL VERSION OF A PROGRAM FOR IDENTIFICATION AND PREDICTION OF TIME SERIES

Summary. In the paper, a computer program is presented for identification and prediction of both stationary and a class of nonstationary time series that contain deterministic polynomial components. For didactical reasons, the program enables a comparison of different estimators, including nonparametric ones of the Power Density Function to be done. The results are illustrated by example outcomes of the program.

*Niniejsza praca została wykonana w ramach Projektu Badawczego KBN nr 3 P403 010 06.

1. Wstęp

Zasadniczym przeznaczeniem programu jest demonstracja opracowanej przez autora (Błachuta, 1996) metody identyfikacji modeli szeregów czasowych oraz ich predykcji. Program umożliwia prezentację zarówno elementów składowych proponowanej metody, jak również innych zagadnień związanych z analizą szeregów czasowych, takich jak : nieparametryczne oceny funkcji korelacji gęstości widmowej mocy oraz porównanie różnych metod estymacji parametrów modeli ARMA.

2. Opis funkcji programu dla procesów stacjonarnych

W programie analizuje się procesy stacjonarne opisane modelem autoregresji i średniej ruchomej ARMA(p, q), (Box, Jenkins 1983):

$$\Phi(B)z(t) = \Theta(B)a(t),$$

gdzie:

B - operator przesunięcia wstecz,

$\Phi(B)$ - operator autoregresyjny rzędu p ,

$\Theta(B)$ - operator średniej ruchomej rzędu q ,

$a(t)$ - ciąg niezależnych zmiennych losowych o rozkładzie $N(0, \sigma^2)$,

$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$,

$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$.

Warunkiem koniecznym i wystarczającym stacjonarności procesu ARMA jest, aby wszystkie pierwiastki λ_i , ($i = 1, 2, \dots, p$) wielomianu $\Phi(\lambda)$ leżały na zewnątrz okręgu jednostkowego na płaszczyźnie zespolonej λ . Zakłada się również, że model jest odwracalny. Warunkiem koniecznym i wystarczającym odwracalności modelu ARMA jest, aby wszystkie pierwiastki λ_i , ($i = 1, 2, \dots, q$) wielomianu $\Theta(\lambda)$ leżały na zewnątrz okręgu jednostkowego na płaszczyźnie zespolonej λ .

2.1. Identyfikacja

Celem identyfikacji jest estymacja stopni p oraz q wielomianów AR i MA oraz wartości współczynników tych wielomianów (Hannan, 1980), (Hannan, Rissanen, 1982), (Hannan, Kavalieris, 1983). Procedura identyfikacji jest następująca:

Najpierw za pomocą przybliżonej metody największej wiarygodności dobierany jest model autoregresyjny wysokiego rzędu, dostatecznie dobrze przybliżający zachowanie się systemu.

Stopień m wielomianu AR dobiera się tak, aby zminimalizować wartość kryterium informacyjnego Akaike

$$AIC = \sigma_m^2 \exp(2m/T),$$

gdzie σ_m^2 jest wariancją dla danego m , zaś T jest długością realizacji. Wielomian AR jest użyty do wyznaczania ocen innowacji. Na podstawie ocen szeregu innowacji oraz szeregu oryginalnego dokonuje się metodą najmniejszych kwadratów dopasowania ciągu modeli ARMA(p, q) dla $p = q$. Liczebność ciągu modeli wynika z minimalizacji kryterium informacyjnego Schwartz

$$BIC = \sigma^2 \exp((p + q) \ln(T)/T).$$

Następnie dokonuje się dopasowania modeli ARMA(p, q) dla $p \neq q$, $p \leq p_0 + 1$, $q \leq q_0 + 1$, gdzie $p_0 = q_0$ są wartościami zapewniającymi minimum BIC przy jednakowych wartościach p i q .

Spośród wszystkich modeli wybierane są trzy modele o najmniejszych wartościach BIC.

Parametry wyselekcjonowanych modeli oraz wartości kryteriów informacyjnych są następnie poprawiane za pomocą numerycznej minimalizacji dokładnej funkcji największej wiarygodności uzyskanej na podstawie filtracji Kalmana. Model o najmniejszej wartości BIC jest uważany za właściwy model procesu.

Dla wyselekcjonowanej trójki modeli obliczane są testy istotności Boxa-Ljunga oraz Godolphina. Przyjmuje się, że wynik testu jest pozytywny, gdy wartość testu jest większa od 0.05.

Wyniki procesu identyfikacji są prezentowane na czterech kolejnych ekranach monitora.

Na pierwszym ekranie prezentuje się realizację szeregu czasowego, typ procesu (stopnie wielomianów AR i MA), współczynniki wielomianów, wariancja bieżącej realizacji szumu pobudzającego układ, ocena wariancji uzyskana z filtru Kalmana przy znajomości parametrów układu oraz wartości funkcji wiarygodności dla tej realizacji.

Podaje się również parametry wielomianu AR dopasowanego do szeregu oraz tabelę zawierającą wartości BIC dla modeli ARMA(p, q) oszacowanych metodą największej wiarygodności.

Na następnym ekranie prezentuje się parametry trzech wyselekcjonowanych modeli oraz wyniki ocen największej wiarygodności dla tych modeli. Następnie podaje się poziomy istotności testów białości zastosowane do szeregów reszt uzyskanych dla różnych modeli oraz dla szumu pobudzającego układ.

Na dwu następnych ekranach przedstawiono wykresy obrazujące kolejno przebiegi wartości funkcji autokorelacji (ACF) teoretycznej dla systemu generującego, oceny funkcji autokorelacji z próby oraz funkcji autokorelacji obliczonej dla wyselekcjonowanych uprzednio modeli. Podobnie wyświetlono przebiegi funkcji gęstości widmowych mocy (SPDF): teoretycznej dla systemu generującego, obliczonej metodą Blackmana-Turkey'a z próby na podstawie oceny funkcji korelacji oraz teoretycznych obliczonych dla wyestymowanych modeli.

Rysunek 1-4 przedstawia kopie wydruków uzyskanych przy identyfikacji systemu ARMA(2,1) o równaniu

$$(1 - 1.6B + 0.9B^2)y_t = (1 + 0.8B)e_t.$$

Na rys. 5-8 przedstawiono wyniki uzyskane przy identyfikacji systemu ARMA(2,1) o równaniu:

$$(1 - 0.8B + 0.5B^2)y_t = (1 + 0.5B + 0.2B^2)e_t.$$

Jako najlepszy zidentyfikowano model AR(3). Jak widać, model ten został zaakceptowany przez testy. Również porównanie ACF i SPDF dla różnych modeli wskazuje, iż model AR(3) należy uznać za właściwy.

2.2. Predykcja

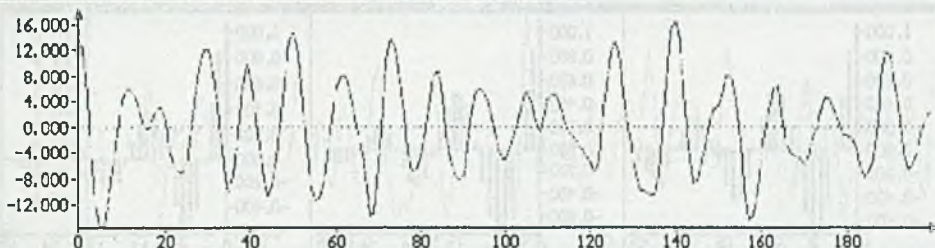
Predykcja może być dokonywana zarówno w oparciu o parametry systemu generującego szereg czasowy jak i na podstawie zidentyfikowanego modelu.

W przypadku estymacji parametrów modelu o zadanej strukturze, odpowiadającej systemowi generującemu, dokonuje się estymacji na podstawie dokładnej metody największej wiarygodności, wyświetla się parametry modelu, wartości BIC, wariancji, funkcji wiarygodności oraz 90 % przedziały ufności dla parametrów.

Prognozy oblicza się na 12 kroków w przód zarówno na podstawie modelu, jak i oryginalnego systemu podając również 90% przedziały ufności dla prognoz. Przedziały ufności dla prognoz dokonanych w oparciu o zidentyfikowany model uwzględniają rozrzut parametrów modelu oraz macierz dyspersji parametrów obliczaną jako odwrotność oceny macierzy informacyjnej uzyskanej z procedury estymacji parametrów.

Na rys. 9 przedstawiono przykładowe wyniki predykcji na podstawie modelu AR(2):

$$(1 - 1.5B + 0.7B^2)y_t = e_t.$$

IDENTYFIKACJA: system, BIC, model AR, kryteria informacyjne
 kłuczenowanu szeregu czasowy


System ARMA(2,1)	a[1]	a[2]	b[1]	Kryteria inform. szumu pobudzajacego uklad		
	-1.6000	0.9000	0.6000	Wariancja 0.9707		
BIC dla dopasowanych modeli ARMA(p,q) LTK	a			Ocena wariancji z filtru Kalmana	0.9976	
	0	1	2		1.0263	
	1	2	3			
	2	3				
Model AR(k)	a[1]	a[2]	a[3]	a[4]	a[5]	a[6]
dobranu HNA ug BIC	-2.3364	2.7172	-2.0093	1.3259	-0.6333	0.2460

Szereg czasowy, system, BIC, kryteria informacyjne, model AR
 zidentyfikowane modele, estymacja ML modeli, testy Box-Ljung, Godolphina
 funkcja autokorelacji: teoretyczna, z proby, z modeli
 funkcja gęstości widmowej: teoretyczna, z proby, z modeli

Esc - Wyjście D - Drukowanie Esc - przerwanie drukowania

Rys. 1

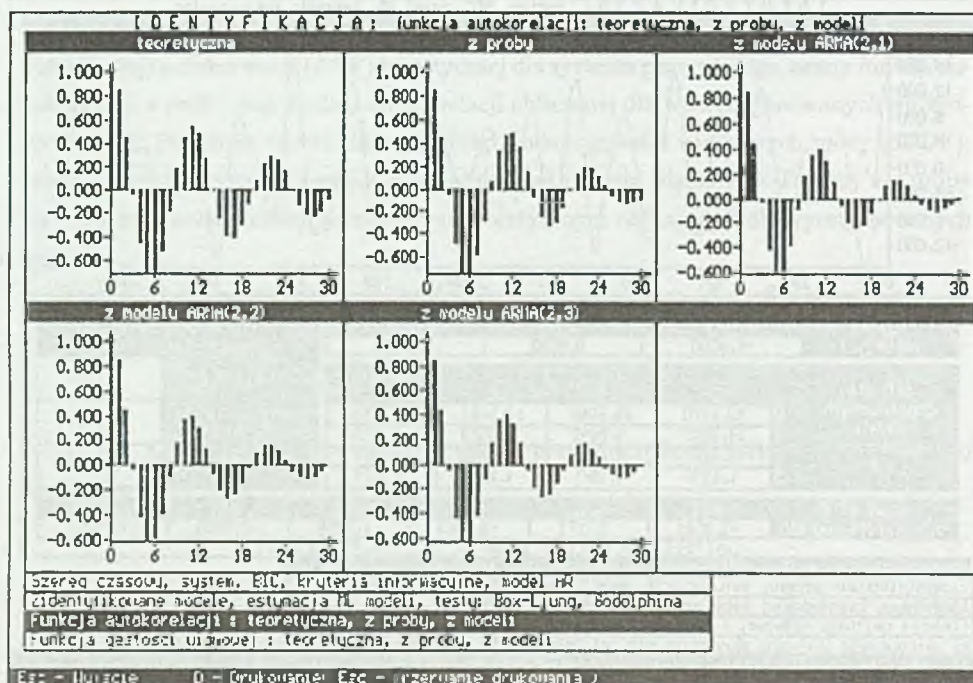
IDENTYFIKACJA: modele, estymacja ML, testy Box-Ljung, Godolphina

Wyselekcjonowane modele o najniższych wartościach BIC							
Model	a[1]	a[2]	Parametry			Kryteria informacyjne	
			b[1]	b[2]	b[3]	BIC	Wariancja LTK
ARMA(2,1)	-1.5275	0.9402	0.6932			1.1449	1.0257 1.0571
ARMA(2,2)	-1.5324	0.9434	0.6950	-0.0494		1.1763	1.0259 1.0580
ARMA(2,3)	-1.5430	0.9448	0.6761	-0.0681	-0.1602	1.2031	1.0225 1.0533
Estymacja ML wyselekcjonowanych modeli							
Model	a[1]	a[2]	Parametry			Kryteria informacyjne	
			b[1]	b[2]	b[3]	BIC	Wariancja LTK
ARMA(2,1)	-1.5901	0.8961	0.7596			1.1034	0.9824 1.0191
ARMA(2,2)	-1.5894	0.8973	0.7391	0.0548		1.1299	0.9783 1.0163
ARMA(2,3)	-1.6198	0.9038	0.7484	-0.1519	-0.2043	1.1336	0.9564 0.9930
Testy istotności dla wyselekcjonowanych modeli i szumu pobudzajacego uklad							
	Test Box-Ljung					Test Godolphina	
ARMA(2,1)	1.0596					0.0000	
ARMA(2,2)	1.0767					0.0000	
ARMA(2,3)	1.1064					0.0000	
szum pobudzajacy uklad	3.5656					0.0000	

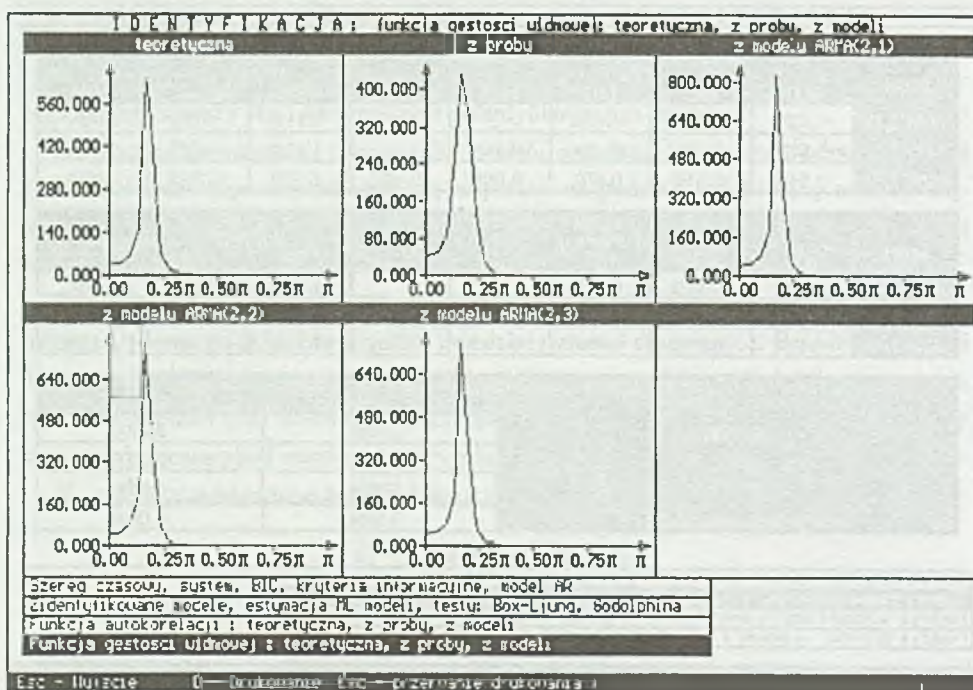
Szereg czasowy, system, BIC, kryteria informacyjne, model ML
 zidentyfikowane modele, estymacja ML modeli, testy Box-Ljung, Godolphina
 funkcja autokorelacji: teoretyczna, z proby, z modeli
 funkcja gęstości widmowej: teoretyczna, z proby, z modeli

Esc - Wyjście D - Drukowanie Esc - przerwanie drukowania

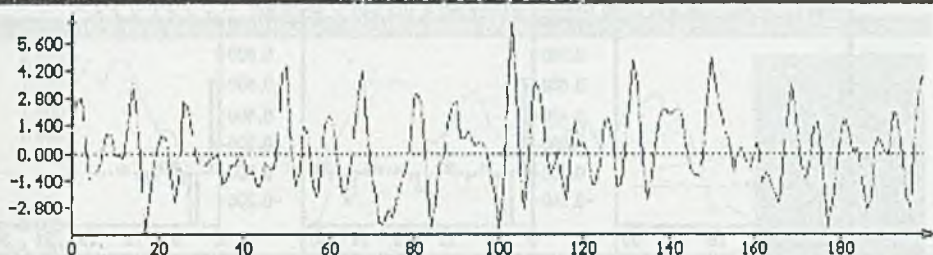
Rys. 2



Rys. 3



Rys. 4

IDENTYFIKACJA: system, BIC, model AR, kryteria informacyjne
 Nupenerowany szereg czasowy


System	a[1]	a[2]	b[1]	b[2]	Kryteria inform. szumu pobudzającego układ	
ARMA(2,2)	-0.8000	0.6000	0.6000	0.2000		
BIC dla dopasowanych modeli ARMA(p,q)	0	1	2	3	Wariancja Ocena wariancji z filtru Kalmana LIK	
	0	1.7771	1.2741	2.9887		1.0656
	1	2.2675	1.7503	1.3288		1.0657
	2	1.1970	1.1718	1.1913		1.2181
3	1.1655	1.2202	1.2170	1.2496	1.0764	
Model AR(3) dobraną MNW od AIC	a[1]	a[2]	a[3]			
	-1.2706	0.9567	-0.2305			

Szereg czasowy, system, BIC, kryteria informacyjne, model AR
 zidentyfikowane modele, estymacja ML modeli, testy: Box-Ljung, Godolphina
 funkcja autokorelacji: teoretyczna, z próbą, z modeli
 funkcja gęstości widmowej: teoretyczna, z próbą, z modeli

Esc - Wyjście 0 - (drukowanie: Esc - przerwanie drukowania)

Rys. 5

IDENTYFIKACJA: modele, estymacja ML, testy: Box-Ljung, Godolphina

Model	Wyselekcjonowane modele o najmniejszych wartościach BIC					Kryteria informacyjne		
	a[1]	a[2]	a[3]	b[1]	b[2]	BIC	Wariancja	LIK
ARMA(3,0)	-1.2456	0.9259	-0.2386			1.1655	1.0658	1.0765
ARMA(2,1)	-0.9393	0.5878		0.3374		1.1718	1.0713	1.0823
ARMA(2,2)	-0.8514	0.5740		0.4134	0.1325	1.1813	1.0606	1.0716

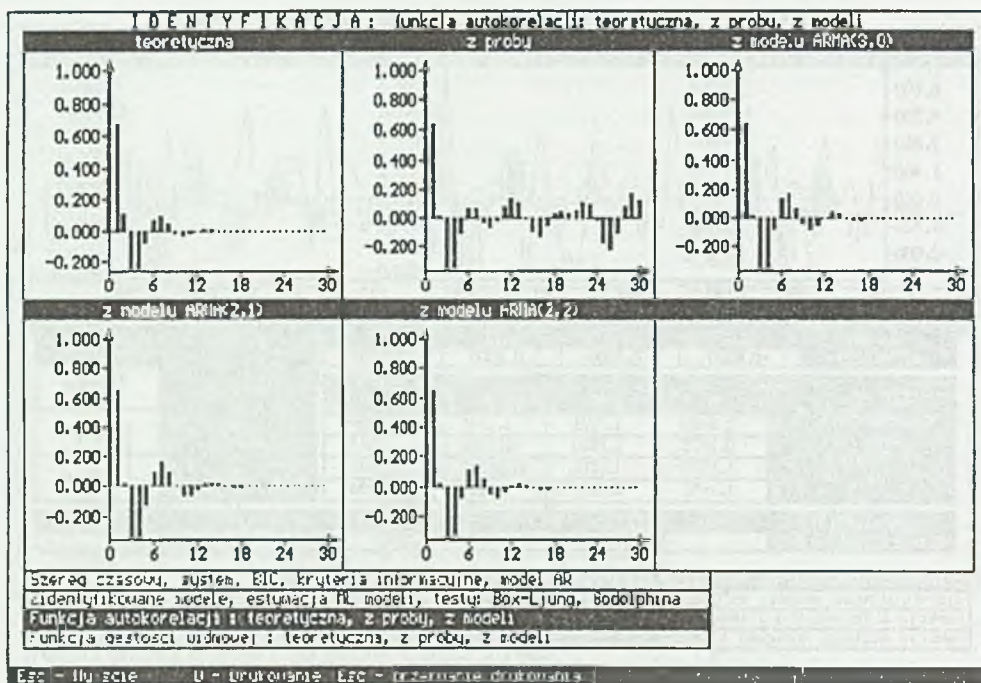
Model	Estymacja ML wyselekcjonowanych modeli					Kryteria informacyjne		
	a[1]	a[2]	a[3]	b[1]	b[2]	BIC	Wariancja	LIK
ARMA(3,0)	-1.2726	0.9570	-0.3324			1.1647	1.0647	1.0757
ARMA(2,1)	-0.9526	0.5169		0.2764		1.1694	1.0690	1.0800
ARMA(2,2)	-0.9360	0.5728		0.4469	0.1828	1.1816	1.0606	1.0718

	Testy istotności dla wyselekcjonowanych modeli i szumu pobudzającego układ	
	Test Box-Ljung	Test Godolphina
ARMA(3,0)	3.8174	0.0000
ARMA(2,1)	2.8409	0.0000
ARMA(2,2)	2.8059	0.0000
szum pobudzający układ	3.8684	0.0000

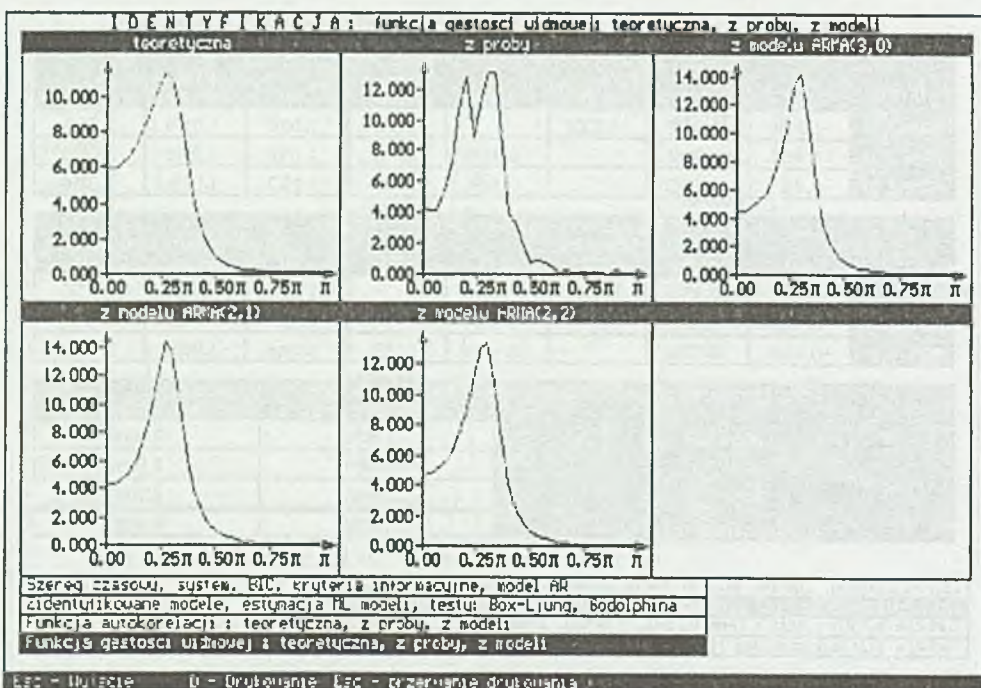
Szereg czasowy, system, BIC, kryteria informacyjne, model AR
 zidentyfikowane modele, estymacja ML modeli, testy: Box-Ljung, Godolphina
 funkcja autokorelacji: teoretyczna, z próbą, z modeli
 funkcja gęstości widmowej: teoretyczna, z próbą, z modeli

Esc - Wyjście 0 - (drukowanie: Esc - przerwanie drukowania)

Rys. 6

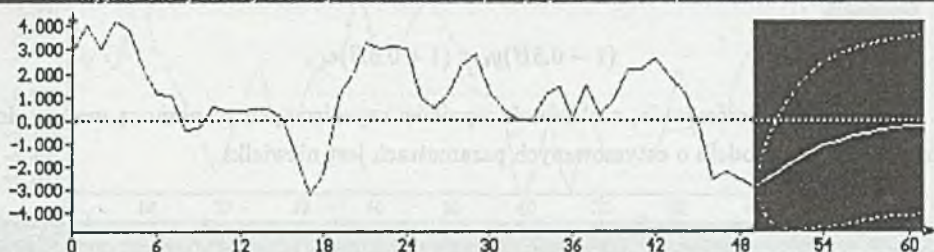


Rys. 7



Rys. 8

P R E D Y K C J A dokonana u oparciu o parametry systemu generującego szereg czasowy
 Wygenerowany szereg czasowy oraz 12-krokowa predykcja wraz z 90% przedziałem ufności



Szereg czasowy		Prognozy szeregu czasowego		
i	$u(i)$	i	$u(i)$	Przedział ufności
33	0,2178	50	-2,6124	-4,3436 .. -0,8812
33	0,8816	51	-2,1665	-4,7401 .. 0,4071
40	2,1644	52	-1,7561	-4,8231 .. 1,3169
41	2,1067	53	-1,4118	-4,7846 .. 1,9610
42	2,8857	54	-1,1315	-4,6874 .. 2,4243
43	1,7614	55	-0,9088	-4,5251 .. 2,7634
44	1,1359	56	-0,7248	-4,4651 .. 3,0154
45	-0,2015	57	-0,5798	-4,3680 .. 3,2051
46	-2,6028	58	-0,4640	-4,2774 .. 3,3495
47	-2,2219	59	-0,3712	-4,2027 .. 3,4601
48	-2,6201	60	-0,2869	-4,1400 .. 3,5461
49	-2,5466	61	-0,2376	-4,0880 .. 3,6129

PROCES STACJONARNY

Esc - Wyjście D - Drukowanie Esc - przerwanie drukowania Szeregi - menu Przejście między

Rys. 9

Na rys. 10, 11 przedstawiono wyniki predykcji uzyskane na podstawie modelu ARMA(1, 1):

$$(1 - 0.8B)y_t = (1 + 0.8B)e_t .$$

Jak widać z porównania, rozbieżność wyników uzyskiwanych za pomocą modelu dokładnego oraz modelu o estymowanych parametrach jest niewielka.

2.3. Model autoregresyjny

Porównano dwie metody dopasowania modeli autoregresyjnych: na podstawie algorytmu Levinsona-Durbina (Box, Jenkins, 1983), bazującego na funkcji autokowariancji oraz przybliżoną metodę największej wiarygodności.

Wyniki dla zwiększających się rzędów modelu AR przedstawiono w postaci tabulogramów, których wiersze zawierają współczynniki modelu. W ostatniej skrajnie prawej kolumnie zamieszczono wartości wariancji reszt. W celu wyróżnienia współczynniki wielomianu minimalizującego AIC zostały podświetlone.

Uzyskane w ten sposób modele AR są punktem wyjścia dla stosowanej procedury identyfikacji. Rysunki 12-13 przedstawiają wyniki dla realizacji procesu ARMA(2,1). Zwraca uwagę nieskuteczność algorytmu Levinsona-Durbina.

2.4. Oszacowanie wstępne

Podstawą szybkiej zbieżności newtonowskiego algorytmu minimalizacji funkcji są wartości startowe niezbyt odległe od punktu minimum. O powodzeniu i szybkości algorytmu estymacji parametrów metodą największej wiarygodności (Mélard, 1984) decyduje zatem dobre oszacowanie wstępne parametrów.

W programie porównuje się kilka metod oszacowań wstępnych:

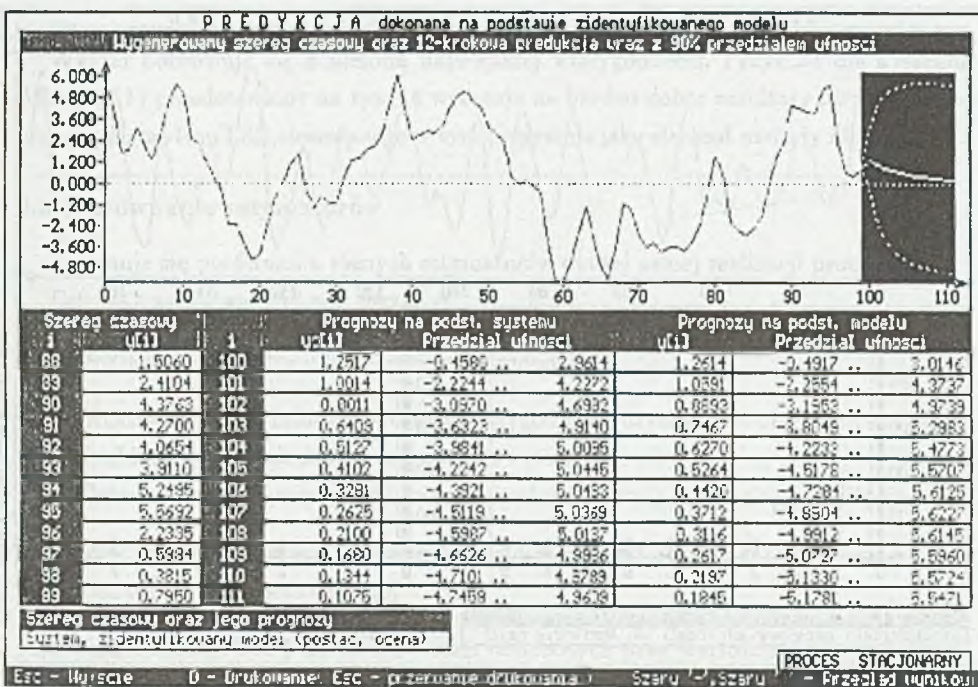
IACV - metoda funkcji korelacji procesu odwrotnego

B-J - metoda Boxa-Jenkinsa dopasowania do funkcji korelacji procesu

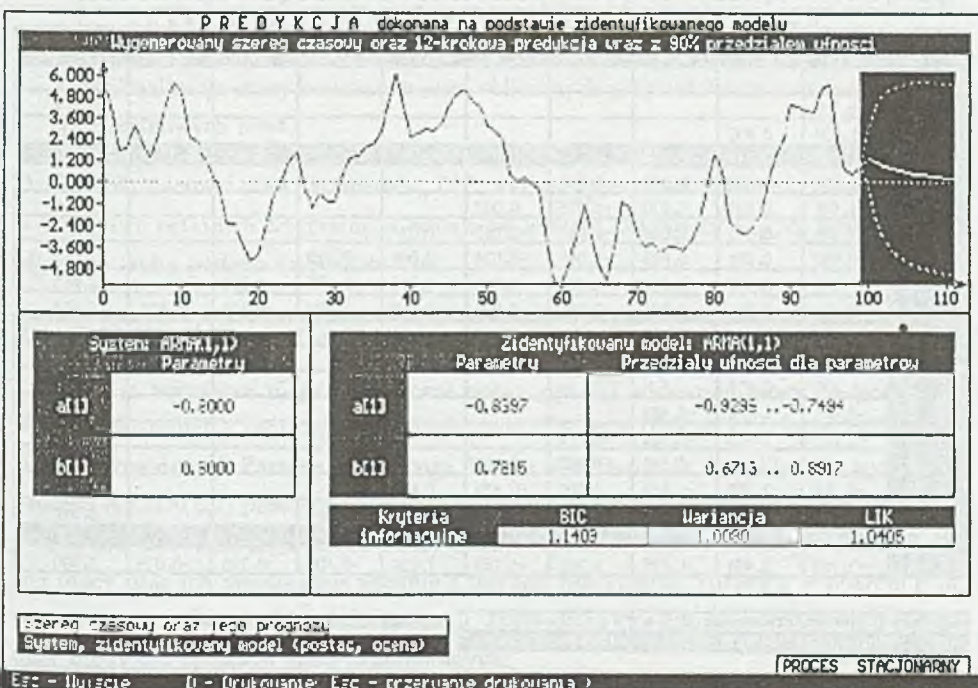
RPEM - rekursywna metoda błędu predykcji

LSA - metoda najmniejszych kwadratów na podstawie oceny reszt dla modelu AR uzyskanego z algorytmu Levinsona-Durbina

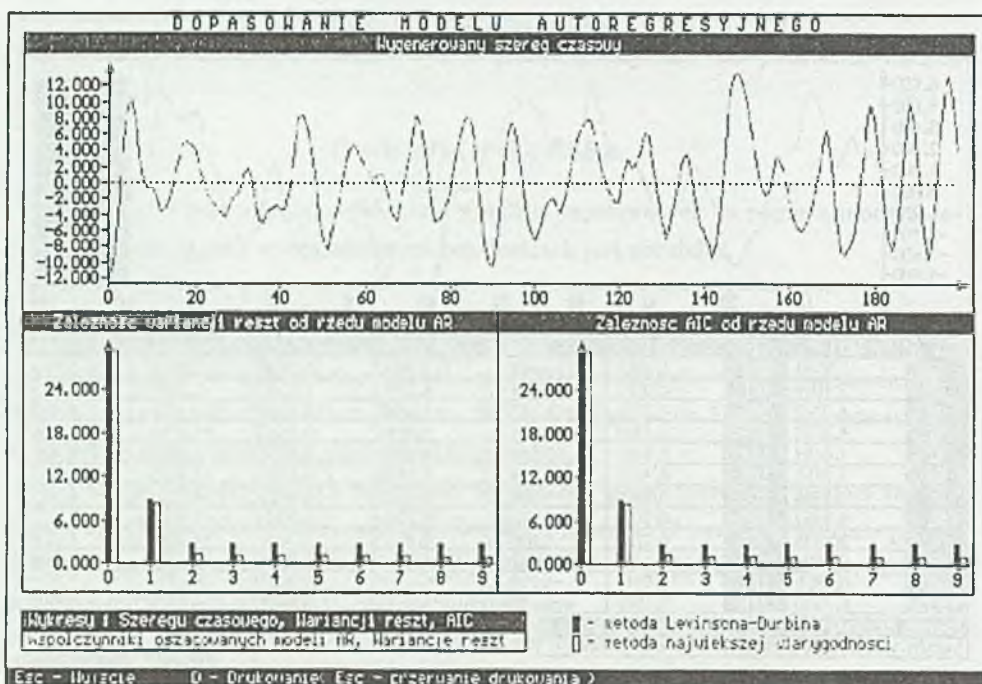
LSE - jak wyżej, lecz modelu autoregresywnego uzyskanego metodą największej wiarygodności.



Rys. 10



Rys. 11



Rys. 12

DOPASOWANIE MODELU AUTOREGRESYJNEGO

Rzad AR	Współczynniki modelu autoregresyjnego									Wariancje reszt
	a(1)	a(2)	a(3)	a(4)	a(5)	a(6)	a(7)	a(8)	a(9)	
0										23.416
1	-0.937									5.314
2	-1.534	0.833								2.701
3	-1.535	0.544	-0.073							2.686
4	-1.591	0.585	0.029	-0.064						2.675
5	-1.591	0.595	0.035	-0.075	0.007					2.675
6	-1.591	0.582	0.036	-0.042	-0.052	0.037				2.672
7	-1.590	0.591	0.035	-0.041	-0.034	0.005	0.020			2.670
8	-1.590	0.591	0.036	-0.041	-0.035	-0.014	0.055	-0.022		2.669
9	-1.591	0.582	0.036	-0.041	-0.036	-0.013	0.023	-0.055	0.021	2.668
0										23.416
1	-0.959									5.173
2	-1.499	0.931								1.340
3	-2.071	1.709	-0.486							1.020
4	-2.177	2.065	-0.923	0.214						1.053
5	-2.200	2.186	-1.194	0.503	-0.135					0.985
6	-2.215	2.357	-1.360	0.800	-0.435	0.138				0.967
7	-2.242	2.531	-1.499	1.033	-0.810	0.507	-0.169			0.941
8	-2.277	2.495	-1.666	1.248	-1.122	0.988	-0.632	0.239		0.901
9	-2.283	2.454	-1.695	1.281	-1.159	1.036	-0.703	0.275	-0.029	0.900

Wskazy: Szeregu czasowego, Wariancji reszt, AIC
Współczynniki oszacowanych modeli AR, Wariancje reszt

Współczynniki wielomianów minimalizujących AIC podświetlono

Esc - Wyjście 0 - Drukowanie: Esc - przerwanie drukowania)

Rys. 13

Wyniki porównuje się z metodą największej wiarygodności. Przykład dla systemu ARMA(2,1) przedstawiony na rys. 14 wskazuje na bardzo dobre rezultaty otrzymane za pomocą algorytmu LSE, stosowanego w tym programie jako element metody identyfikacji.

2.5. Porównanie estymatorów

Dokonuje się porównania różnych estymatorów dla tej samej realizacji procesu.

Estymatorami branyymi pod uwagę są:

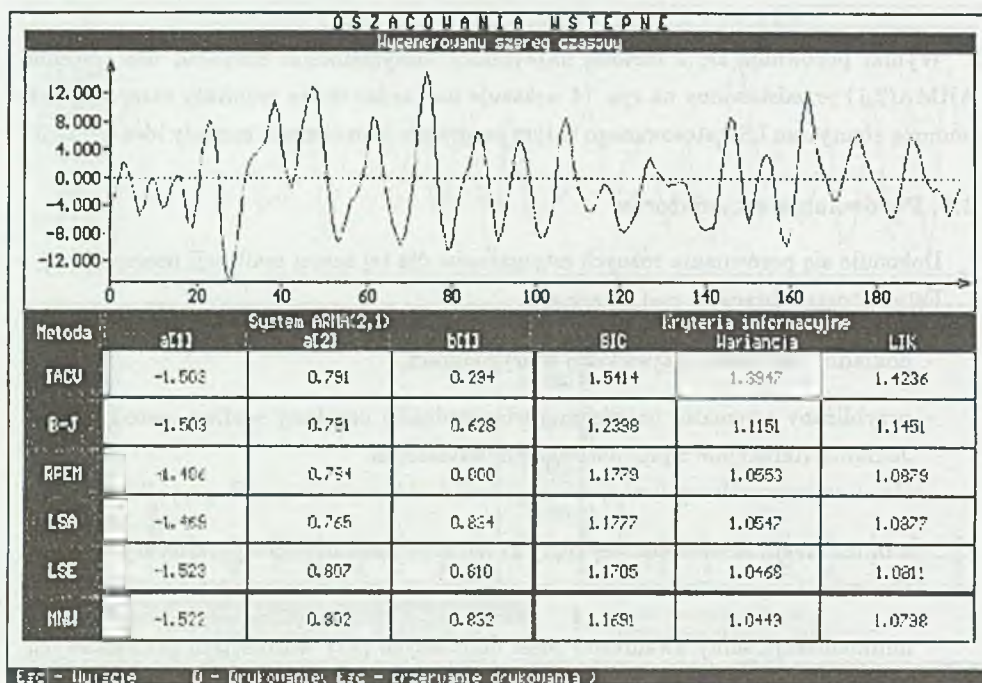
- dokładny estymator największej wiarygodności;
- przybliżony estymator największej wiarygodności uzyskany według metody Boxa-Jenkinsa iteracyjnie z prognozowaniem wstecznym.
Człon wyznacznikowy funkcji wiarygodności obliczany w sposób przybliżony;
- minimalizacja bezwarunkowej sumy kwadratów reszt uzyskanej jak wyżej.
Bez członu wyznacznikowego;
- minimalizacja sumy kwadratów reszt obliczonych przy wartościach początkowych ocenionych metodą najmniejszych kwadratów;
- minimalizacja sumy kwadratów reszt obliczonych z pominięciem pewnej ich początkowej liczby dla eliminacji wpływu warunku początkowego;
- minimalizacja sumy kwadratów reszt obliczonych przy założeniu zerowych wartości początkowych reszt.

Drukuje się wartości ocen parametrów, BIC, wariancję reszt, wartość funkcji wiarygodności. Dla dwu ostatnich estymatorów podaje się wartość minimalizowanych funkcji. Przykłady wydruku podano na rys. 16-17.

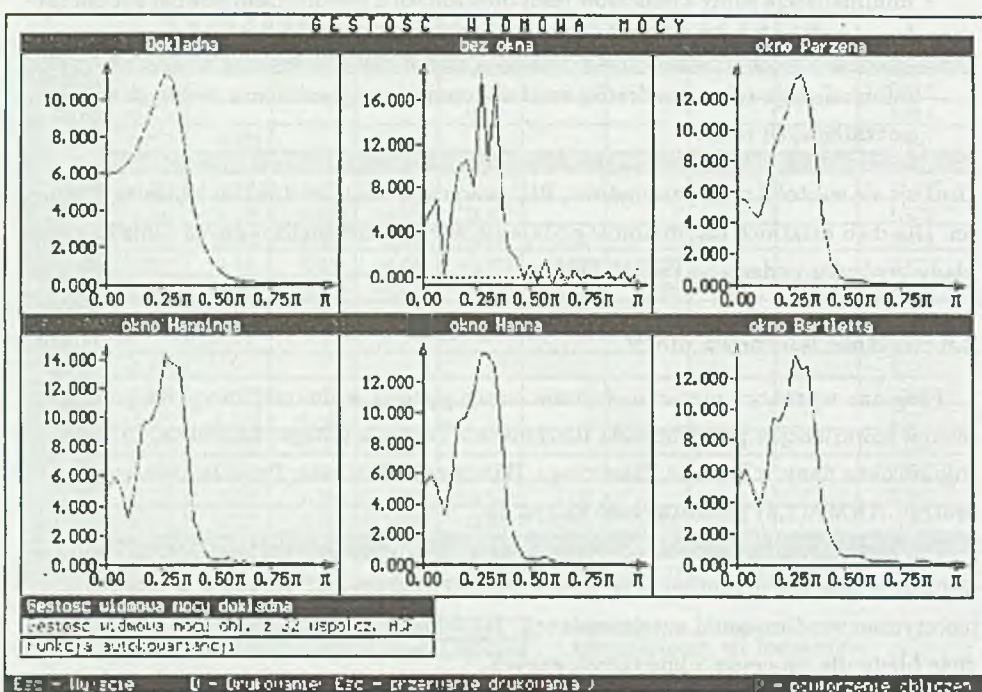
2.6. Gęstość widmowa mocy

Program wyznacza nieparametryczne oceny gęstości widmowej mocy na podstawie funkcji kowariancji z próby metodą Blackmana - Tuckey'a (Otnes, Enochson, 1978). Stosuje się okna danych Parzena, Hamminga, Hanna oraz Bartletta. Przykładowe wyniki dla procesu ARMA(2,2) przedstawiono na rys. 15.

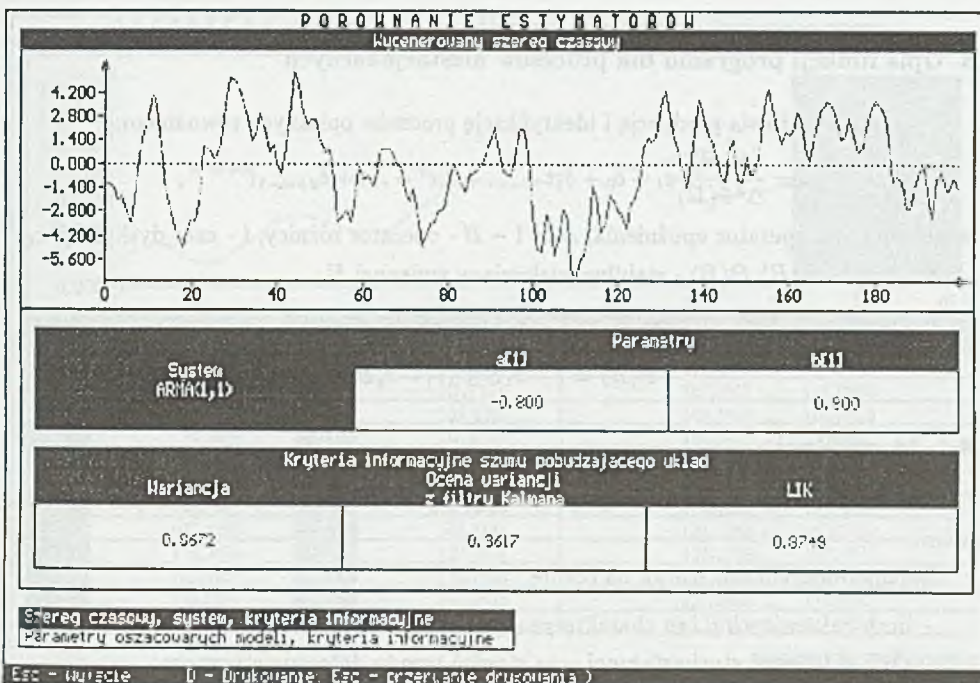
Dla zorientowania się o jakości oceny podaje się teoretyczne wartości gęstości widmowej mocy oraz ich przybliżenie zgodnie z metodą Blackmana- Tuckey'a w oparciu o 32 teoretyczne współczynniki autokowariancji. Jak widać z rys. 3, metoda może wprowadzać duże błędy dla procesów silnie skorelowanych.



Rys. 14



Rys. 15



Rys. 16

PORÓWNANIE ESTYMATORÓW

Dokładny estymator IIL	Model ARMA(1,1)		BIC	Kryteria informacyjne	
	a[1]	b[1]		Wariancja	LIK
	-0.7731	0.8176	0.9213	0.8607	0.8737

Przybl. e. ML 41 prognoz wstecznych	Model ARMA(1,1)		BIC	Kryteria informacyjne	
	a[1]	b[1]		Wariancja	LIK
	-0.7731	0.8176	0.9212	0.8606	0.8737

Bezwarunkowa suma kwadr.	Model ARMA(1,1)		BIC	Kryteria informacyjne	
	a[1]	b[1]		Wariancja	LIK
	-0.7818	0.8137	0.9213	0.8606	0.8738

Estymowany war. pocz.	Model ARMA(1,1)		BIC	Kryteria informacyjne	
	a[1]	b[1]		Wariancja	LIK
	-0.7820	0.8206	0.9068	0.8600	0.8600

Okienko 40 opuszcz. danych	Model ARMA(1,1)		BIC	Kryteria informacyjne	
	a[1]	b[1]		Wariancja	LIK
	-0.7633	0.8577	0.9365	0.8502	0.8502

Warunkowa suma kwadr.	Model ARMA(1,1)		BIC	Kryteria informacyjne	
	a[1]	b[1]		Wariancja	LIK
	-0.7811	0.8139	0.9142	0.8671	0.8671

Szereg czasowy, system, kryteria informacyjne
Parametry oszacowanych modeli, kryteria informacyjne

Esc - Wracajcie D - Drukowanie: Esc - przerwanie drukowania

Rys. 17

3. Opis funkcji programu dla procesów niestacjonarnych

Program umożliwia predykcję i identyfikację procesów opisanych równaniem:

$$y_t = \frac{\Theta(B)}{\Delta^d \Phi(B)} \varepsilon_t + \delta_0 + \delta_1 t + \dots + \delta_d t^d + \dots + \delta_{d+m-1} t^{d+m-1},$$

w którym: B - operator opóźnienia, $\Delta = 1 - B$ - operator różnicy, t - czas dyskretny, ε_t - n.i.d.(0, σ^2), $\Phi(B)$, $\Theta(B)$ - stabilne wielomiany zmiennej B ,

$$\Phi(B) = 1 - \phi_1 B + \dots - \phi_p B^p,$$

$$\Theta(B) = 1 - \theta_1 B + \dots - \theta_q B^q.$$

3.1. Identyfikacja

Zalóżmy, że dana jest realizacja $y(1) \dots y(T)$ procesu opisanego powyższym równaniem.

Zadanie identyfikacji polega na ocenie:

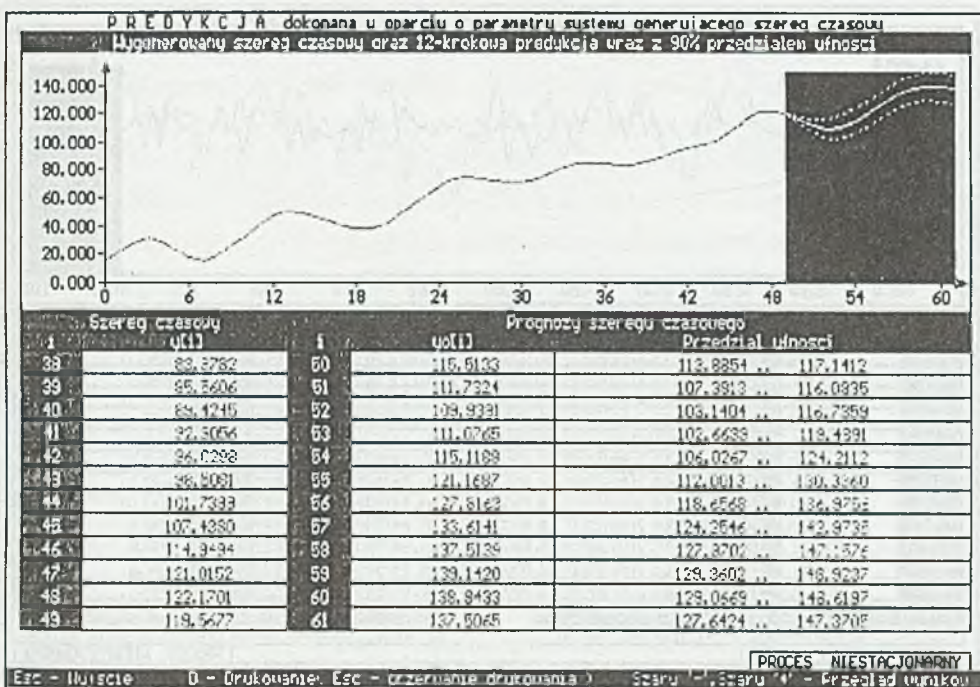
- liczb całkowitych d i m charakteryzujących rząd różnicy koniecznej dla stacjonaryzacji składowej stochastycznej oraz stopień trendu deterministycznego,
- liczb całkowitych p i q określających stopnie wielomianów $\Phi(B)$ i $\Theta(B)$,
- współczynników $\phi_1 \dots \phi_p, \theta_1 \dots \theta_q$.

Procedura identyfikacji jest wieloetapowa i polega na selekcjonowaniu modeli na podstawie kryteriów informacyjnych. W pierwszym etapie dokonuje się oceny liczb d i m , posługując się modelem autoregresyjnym składowej stochastycznej, określonym wielomianem $\Phi_n(B)$ stopnia n . W drugim etapie na podstawie modelu autoregresyjnego oraz zróżnicowanej realizacji z usuniętym trendem określa się strukturę modelu ARMA i jego parametry. W etapie trzecim po określeniu struktury i wstępnej ocenie parametrów dokonuje się ostatecznej łącznej oceny parametrów trendu i parametrów modelu ARMA za pomocą metody największej wiarygodności.

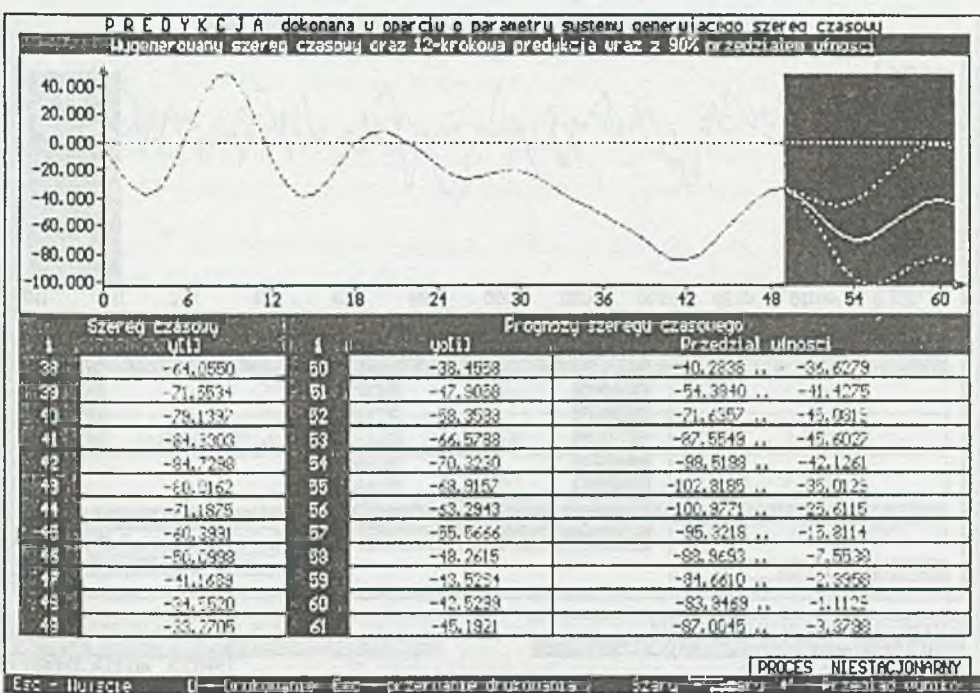
3.2. Predykcja

Podobnie jak dla układów stacjonarnych predykcja może być dokonana zarówno na podstawie równania systemu generującego, jak i na podstawie modelu o ustalonej strukturze i estymowanych parametrach (Harvey, 1981), (Harvey, McKenzie, 1982).

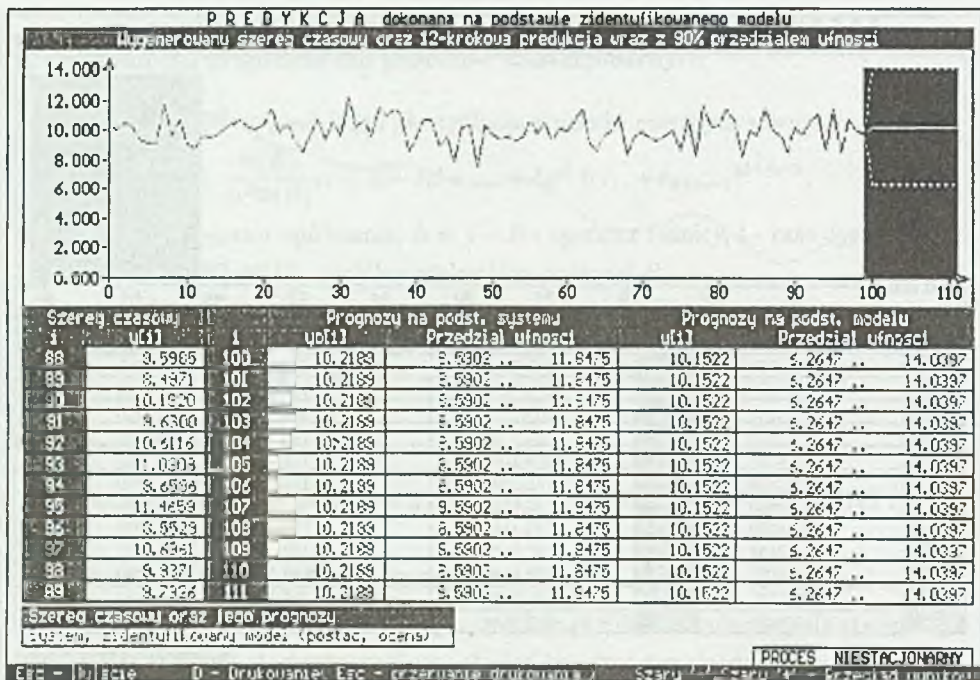
Na rys. 18 przedstawiono przykładowe prognozy dla systemu ARMA(2,1) z dodaną składową deterministyczną liniowo narastającą, zaś na rys. 19 prognozy dla systemu ARIMA(2,1,1).



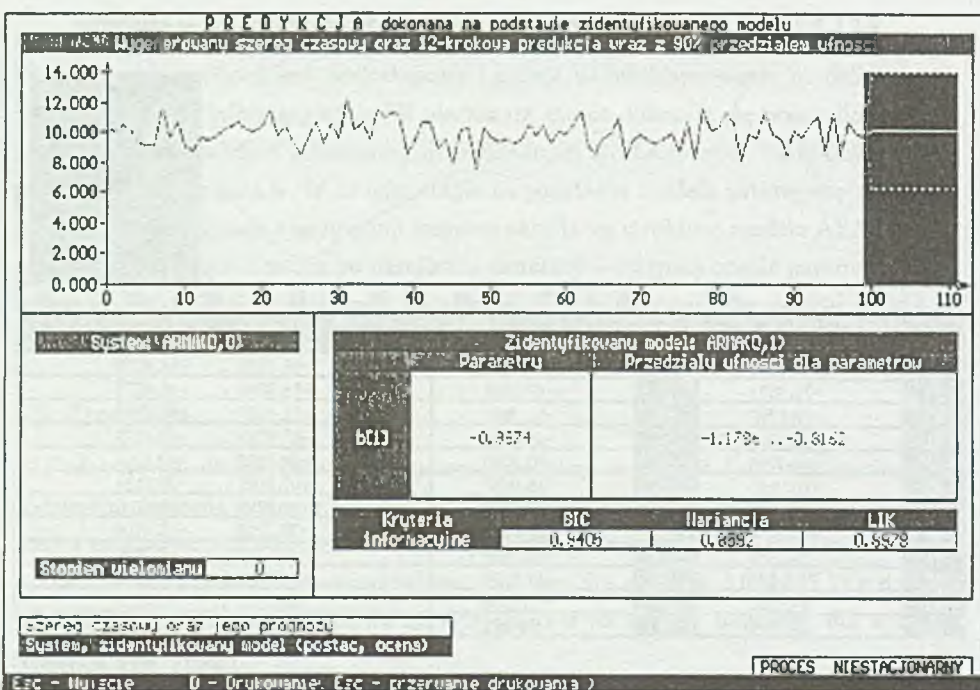
Rys. 18



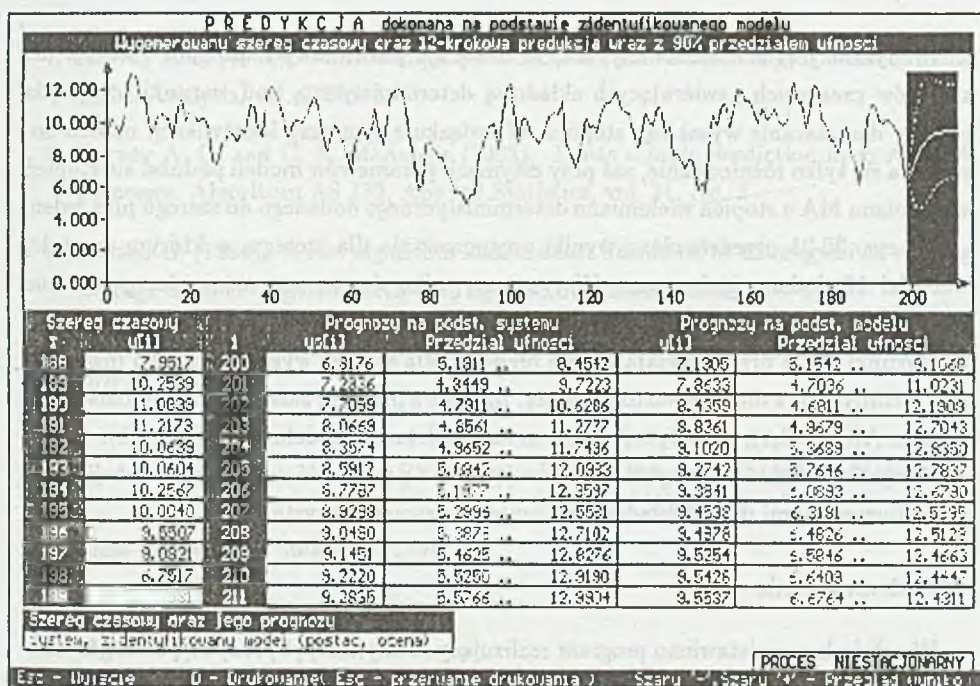
Rys. 19



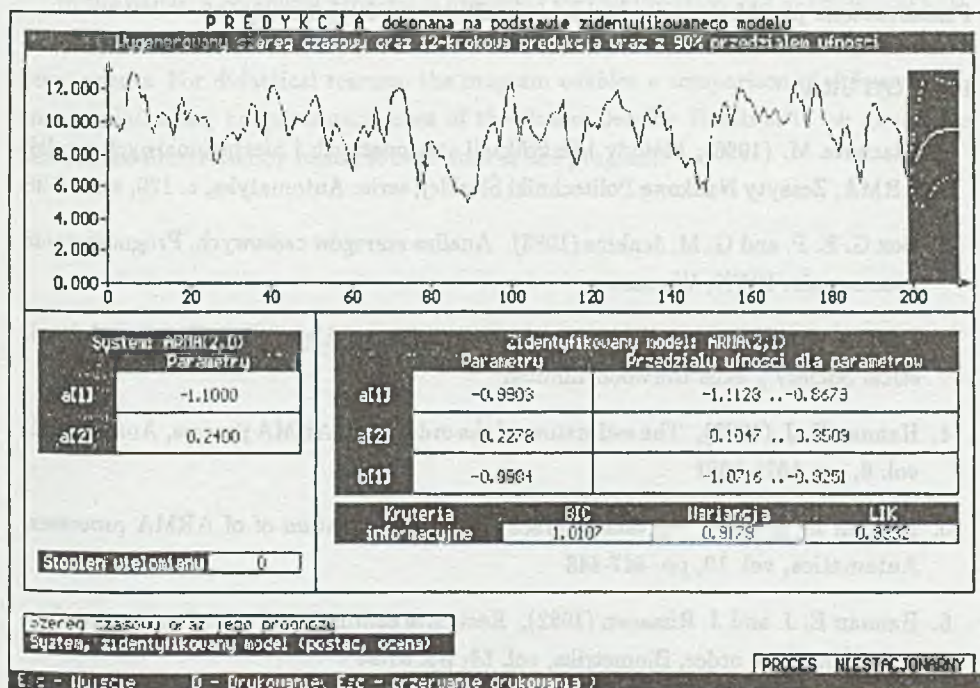
Rys. 20



Rys. 21



Rys. 22



Rys. 23

Predyktor jest skonstruowany tak, że dokonuje prawidłowych prognoz również dla szeregów czasowych zawierających składową deterministyczną pod warunkiem wzięcia różnicy dostatecznie wysokiego stopnia. W związku z tym przy identyfikacji modeli dopuszcza się tylko różnicowanie, zaś przy estymacji parametrów modeli podnosi się stopień wielomianu MA o stopień wielomianu deterministycznego dodanego do szeregu plus jeden.

Na rys. 20-21 przedstawiono wyniki prognozowania dla szeregu, w którym do stałej wartości 10 dodano biały szum. W tym przypadku do szeregu różnic dopasowywano model MA(1). Predykcji dokonano na podstawie modelu ARIMA(0,1,1).

Rysunki 22-23 przedstawiają wyniki predykcji dla szeregu wygenerowanego przez system ARMA(2,0) z dodaną wartością stałą. Zgodnie z przyjętą zasadą oceniono parametry modelu ARMA(2,1), a predykcji dokonano na podstawie modelu ARIMA(2,1,1).

Wszystkie omówione przykłady wskazują na dużą zgodność wyników predykcji z wynikami uzyskanymi przy dokładnej znajomości parametrów systemu.

4. Podsumowanie

W artykule przedstawiono program realizujący identyfikację i predykcję szeregów czasowych stacjonarnych i pewnej klasy szeregów niestacjonarnych, zawierających deterministyczne składowe wielomianowe według metody opracowanej przez Błachutę (1996). Przedstawione przykłady wskazują na dużą efektywność tej metody.

LITERATURA

1. Błachuta M. (1996), Metody identyfikacji stacjonarnych i niestacjonarnych modeli ARMA, Zeszyty Naukowe Politechniki Śląskiej, seria: Automatyka, z. 120, str. 15-38
2. Box G. E. P. and G. M. Jenkins (1983), Analiza szeregów czasowych. Prognozowanie i sterowanie. PWN, Warszawa
3. Griffiths P. and I. D. Hill (1985), Applied Statistics Algorithms. The Royal Statistical Society / Ellis Horwood Limited
4. Hannan E. J. (1988), The estimation of the order of an ARMA process, Ann. Statist, vol. 8, pp. 1071-1081
5. Hannan E. J. and L. Kavalieris (1983), Linear estimation of of ARMA processes, Automatica, vol. 19, pp. 447-448
6. Hannan E. J. and J. Rissanen (1982), Recursive estimation of mixed autoregressive-moving average order, Biometrika, vol. 69, pp. 81-94

7. Harvey A.C. (1981), Finite sample prediction and overdifferencing, *Journal of Time Series Analysis*, vol. 2, No. 4
8. Harvey A. C. and C. R. McKenzie (1982), Finite sample prediction form ARIMA processes, *Algorithm AS 182, Applied Statistics*, vol. 31, No. 2
9. M elard G. (1984), A fast algorithm for the exact likelihood of autoregressive- moving average models. *Algorithm AS 197, Applied Statistics*, vol. 33, No. 1
10. Otnes R. K. and L. Enochson (1978), *Analiza numeryczna szereg w czasowych*. WNT, Warszawa

Recenzent: Doc. dr hab. in . Zbigniew Nahorski
IBS PAN Warszawa

Wpłynęło do Redakcji dnia 15.05.1994

Abstract

In the paper, a computer program is presented for identification and prediction of both stationary and a class of nonstationary time series that contain deterministic polynomial components. For didactical reasons, the program enables a comparison of different estimators, including nonparametric ones of the Power Density Function to be done. The results are illustrated by example outcomes of the program.