



Następnie korzystamy z menu DANE ▷ WYBIERZ OBSERWACJE i wybieramy opcję WSZYSTKIE OBSERWACJE (wówczas wszystkie obserwacje są aktywne). Wreszcie wybieramy z menu ANALIZA opcję OPIS STATYSTYCZNY ▷ STATYSTYKI OPISOWE i obliczamy średnią wartość nowej zmiennej — 0,04059. Wariancja *ex post* prognoz (średnia kwadratów różnic pomiędzy wartościami rzeczywistymi i wartościami prognoz) wynosi więc 0,04059. Pierwiastek wariancji *ex post* prognoz wynosi 0,2015 punktu procentowego. Oznacza to, że wartości prognoz odchyłały się od rzeczywistych wartości przeciętnie o 0,2015 punktu procentowego. Stanowi to 4,72% średniej rzeczywistych wartości stopy bezrobocia w tym czasie — oznacza to, że jakość prognoz *ex post* można uznać za satysfakcjonującą (choć w każdym z 9 okresów prognozy nie doszycowują rzeczywistych wartości prognozowanej zmiennej). Innym wskaźnikiem względnym, który można tu rozpatrywać, jest stosunek pierwiastka oceny wariancji *ex post* prognoz do minimalnej spośród wartości prognoz. W tym przypadku pierwiastek wariancji *ex post* prognoz stanowi 4,79% minimalnej wartości prognoz w tym okresie.

7.4 Automatyczne stawianie prognoz

Opcja UTWÓRZ MODELE posiada również tryb automatycznego wyboru „najlepszego” modelu i postawienie na jego podstawie prognoz. Ma ona jeszcze jedną podstawową zaletę — w przypadku istnienia informacji o wartościach zmiennych niezależnych umożliwia wybór najlepszej kombinacji zmiennych niezależnych z odpowiednimi przekształceniami z uwzględnieniem niezbędnych opóźnień. Tryb automatyczny może być więc preferowany każdorazowo w sytuacji istnienia informacji o zmiennych niezależnych. Pamiętać jednak należy, że w sytuacji, gdy chcemy stawiać prognozy, niezbędne jest zadeklarowanie wartości zmiennych niezależnych na okresy, dla których stawiane będą prognozy. Z tego powodu na początku wprowadzamy zakładane lub prognozowane wartości zmiennych niezależnych na kolejne okresy.

W poniższym przykładzie nadal rozważamy dane dotyczące bezrobocia w USA z podrozdziału 7.3 (zawarte w pliku *gospodarka_USA.sav*). Ponieważ zdecydowaliśmy się na postawienie prognoz na 4 okresy, musimy zadeklarować wartości zmiennych niezależnych na kolejne 4 miesiące. Wartościami tymi mogą być przykładowo wartości prognoz zmiennych niezależnych. W naszym przypadku, dla uproszczenia deklarujemy wartości tych zmiennych w kolejnych 4 okresach takie jak w ostatnim okresie, dla którego były dostępne dane. Następnie — co ważne — po pierwsze, uaktualniamy zmienne definiujące datę, tak aby obejmowały okresy prognozowane, czyli miesiące do stycznia 2000 (DANE ▷ DEFINIUJ DATĘ I CZAS), a po drugie, jako nieaktywne zaznaczamy ostatnie 4 okresy (DANE ▷ WYBIERZ OBSERWACJE).

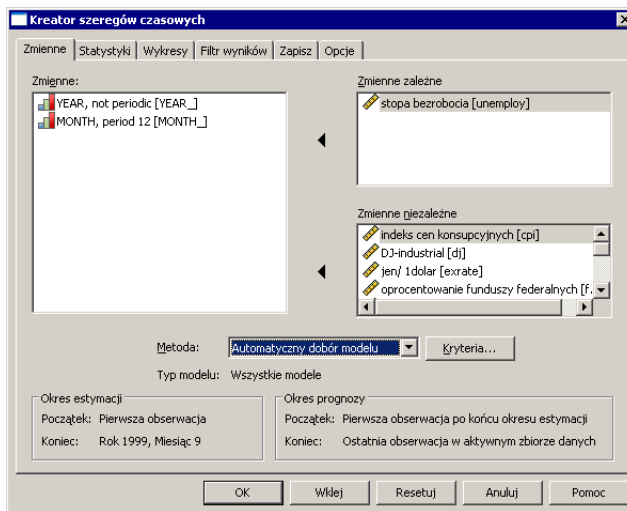
Arkusz z danymi w oknie edytora danych wygląda wówczas jak na rysunku 7.52.

	federal	indprod	discount	loanrate	sales	m3rate	m6rate	crude	unemploy	YEAR	MONTH	DATE
80	5,07	137,60	4,56	8,06	252760,0	4,72	4,87	21,30	4,20	1999	8	AUG 1999
81	5,22	137,64	4,75	8,25	252549,0	4,68	4,88	23,90	4,20	1999	9	SEP 1999
82	5,22	137,64	4,75	8,25	252549,0	4,68	4,88	23,90	.	1999	10	OCT 1999
83	5,22	137,64	4,75	8,25	252549,0	4,68	4,88	23,90	.	1999	11	NOV 1999
84	5,22	137,64	4,75	8,25	252549,0	4,68	4,88	23,90	.	1999	12	DEC 1999
85	5,22	137,64	4,75	8,25	252549,0	4,68	4,88	23,90	.	2000	1	JAN 2000

Rysunek 7.52
Okno edytora danych

Należy podkreślić, że w przypadku, gdy „najlepszy” model będzie uwzględnił zmienne niezależne, zadeklarowane na okresy przyszłe wartości zmiennych niezależnych będą bardzo silnie wpływać na wartość prognoz. Modele takie umożliwiają jednak rozważanie różnych wariantów, w zależności od poziomu różnych czynników. Od zakładanych wartości zmiennych niezależnych zależą wartości prognoz. Tak więc dla innych zakładanych wartości uzyskamy inne wartości prognoz. W ten sposób możemy analizować zachowanie się zmiennej zależnej w zależności od wartości zmiennych niezależnych, jakie wystąpią w przyszłości.

Zatem z menu ANALIZA wybieramy opcję SZEREGI CZASOWE ▷ UTWÓRZ MODELE, i deklarujemy wartości zmiennej zależnej i wartości zmiennych niezależnych a także automatyczny wybór modelu (rys. 7.53).



Rysunek 7.53
Okno dialogowe:
Kreator szeregów czasowych

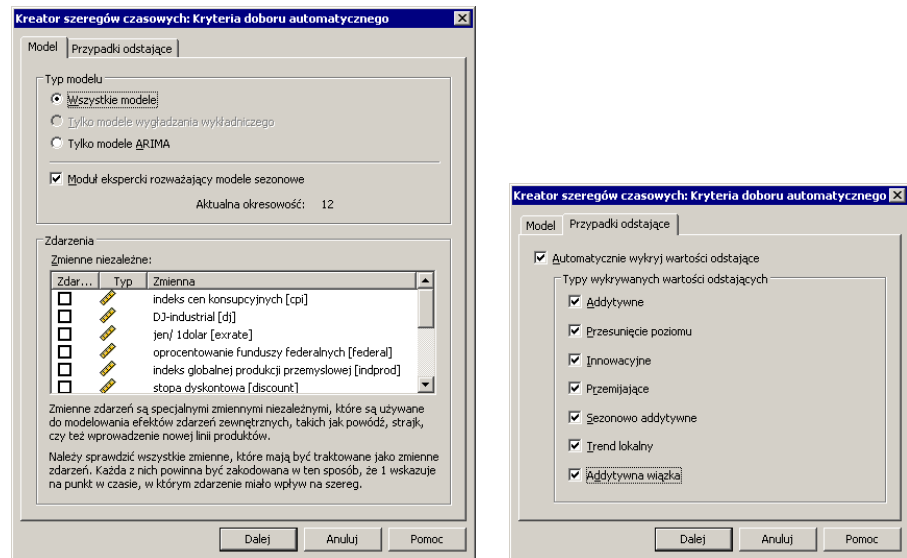
Po wyborze opcji KRYTERIA deklarujemy, aby wyboru „najlepszego” modelu dokonano spośród wszystkich modeli oraz by uwzględniono ewentualny wpływ czynników sezonowych (rys. 7.54a). W oknie tym istnieje ponadto możliwość zadeklarowania dostępnych zmiennych jako zmienne zerojedynkowe określające okresy czasu, w których wystąpiły zdarzenia mogące mieć wpływ na wartości badanej zmiennej (ZDARZENIA) — w analizowanym przypadku nie ma takich zmiennych.



Wybierając zakładkę PRZYPADKI ODSTAJĄCE, decydujemy się automatyczne wykrywanie (i uwzględnienie w modelu) wartości oddalonych — rys. 7.54b.

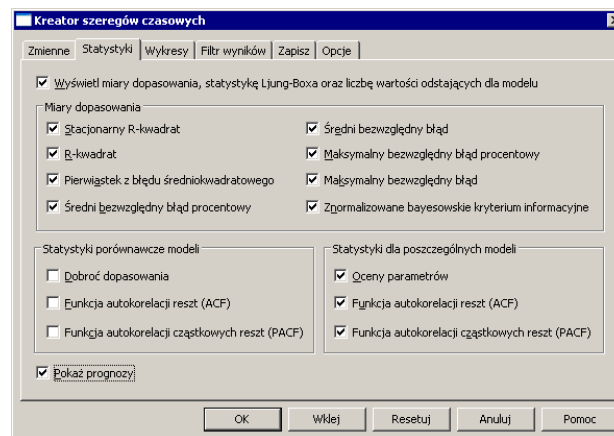
a b

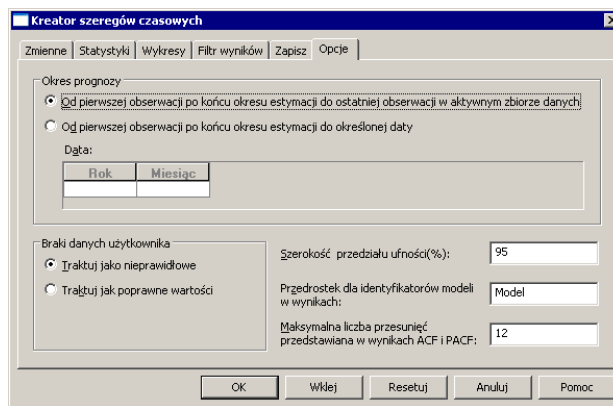
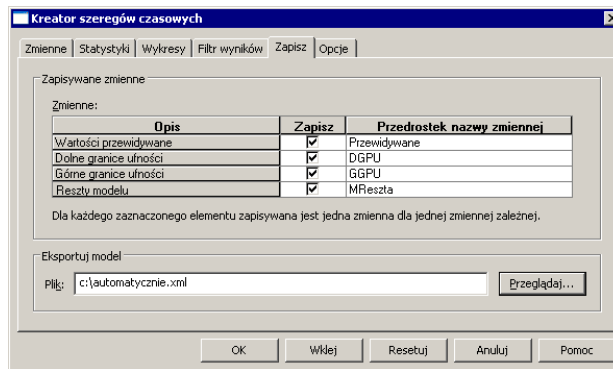
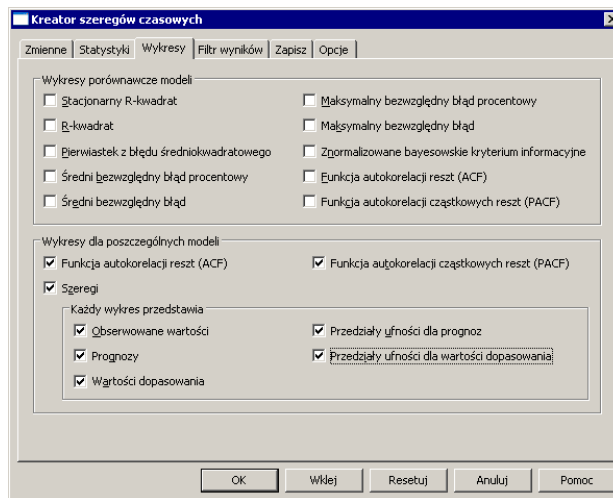
Rysunek 7.54
Kreator szeregów czasowych.
Opcja: Kryteria doboru automatycznego: a) Model;
b) Przypadki odstające



Klikając DALEJ wracamy do okna KREATOR SZEREGÓW CZASOWYCH. Wybierając kolejne zakładki deklarujemy opcje pokazane na rysunkach 7.55–7.56 (w zakładce FILTR WYNIKÓW pozostawiamy wybraną domyślnie pierwszą opcję od góry).

Rysunek 7.55
Kreator szeregów czasowych.
Zakładka: Statystyki





a
b
c

Rysunek 7.56
Kreator szeregów czasowych.
Zakładka:
a) Wykresy;
b) Zapisz;
c) Opcje

Po uruchomieniu procedury do arkusza z danymi zostają dopisane wartości prognoz, dolne i górne granice przedziałów ufności i wartości błędów *ex post* prognoz (rys. 7.57).



Rysunek 7.57
Okno edytora danych z prognozowanymi wartościami

	crude	unemploy	YEAR	MONTH	DATE	Przewidywane _unemploy _Model_1	DGPU _unemploy _Model_1	GGPU _unemploy _Model_1	MReszta _unemploy _Model_1
78	17,90	4,30	1999	6	JUN 1999	4,15	3,97	4,32	,15
79	20,10	4,30	1999	7	JUL 1999	4,22	4,05	4,39	,08
80	21,30	4,20	1999	8	AUG 1999	4,11	3,94	4,28	,09
81	23,90	4,20	1999	9	SEP 1999	4,20	4,03	4,37	,00
82	23,90	.	1999	10	OCT 1999	4,20	4,03	4,38	.
83	23,90	.	1999	11	NOV 1999	4,09	3,89	4,30	.
84	23,90	.	1999	12	DEC 1999	4,07	3,83	4,30	.
85	23,90	.	2000	1	JAN 2000	4,08	3,82	4,35	.

Prognozowane wielkości stopy bezrobocia na październik 1999, listopad 1999, grudzień 1999 oraz styczeń 2000 wynoszą odpowiednio: 4,20%; 4,09%; 4,07%; 4,08%. Dla prognozy na październik 1999 połowa szerokości przedziału ufności wynosi $\frac{1}{2}(4,38\% - 4,03\%) = 0,175$ punktu procentowego, co należy uznać za bardzo niską wartość w porównaniu z wartością prognozy. Jakość prognozy na ten okres należy uznać za satysfakcjonującą. Połowa szerokości przedziału ufności dla prognoz rośnie z okresu na okres, przyjmując wartości 0,205, 0,235 oraz 0,265 punktu procentowego.

Po naciśnięciu przycisku ZAKOŃCZ okazuje się, że „najlepszym” (pod względem wartości współczynnika NBIC) modelem jest model ARIMA(0, 1, 1) ze zmiennymi niezależnymi, uwzględniający występowanie wartości oddalonych (rys. 7.58–7.64).

Rysunek 7.58
ARIMA. Raport: Opis modelu

			TYP MODELU
Model ID	stopa bezrobocia	Model_1	ARIMA (0,1,1) (0,0,0)

		MODEL
		STOPA BEZROBOCIA-MODEL_1
LICZBA ZMIENNYCH DODATKOWYCH		3
STATYSTYKI DOPASOWANIA MODELU	Stacjonarny R-kwadrat	0,655
	R-kwadrat	0,991
	Pierwiastek błędu średniokwadratowego (RMSE)	0,086
	Średni bezwzględny błąd procentowy (MAPE)	1,226
	Średni błąd bezwzględny (MAE)	0,064
	Maksymalny bezwzględny błąd procentowy (MaxAPE)	3,559
	Maksymalny błąd bezwzględny (MaxAE)	0,153
Znormalizowane bayesowskie kryterium informacyjne (NBIC)		-3,993
LJUNG-BBOX Q (18)	Statystyki	22,486
	DF	17
	Istotność	0,167
LICZBA OBSERWACJI ODSTAJĄCYCH		3

Rysunek 7.59
ARIMA. Raport: Statystyki modelu

Z zestawienia na rysunku 7.59 dowiadujemy się m.in., że wartości rzeczywiste stopy bezrobocia odchylają się od wartości teoretycznych (wynikających z oszacowanego modelu) o przeciętnie 1,226%, a wartość największej reszty wyrażona w procentach to 3,559%. Jakość dopasowania modelu do danych empirycznych należy uznać za satysfakcjonującą. Ponadto obserwowany poziom istotności dla testu Ljunga-Boxa wynosi 0,167 i jest większy

od zakładanego poziomu istotności 0,05, co pozwala na przyjęcie hipotezy zerowej głoszącej brak autokorelacji składników losowych.

Zestawienie na rysunku 7.60 przedstawia m.in. wartości estymatorów parametrów oszacowanego modelu wraz ze średnimi błędami szacunku i obserwowanym poziomem istotności dla testu *t*-Studenta z hipotezą zerową głoszącą, że parametr modelu nie jest istotnie różny od zera.

				WARTOŚĆ OSZACOWANIA	BŁĄD STANDARDOWY	T	ISTOTNOŚĆ	
STOPA BEZROBOCIA- MODEL_1	STOPA BEZROBOCIA	Bez transformacji	Stała	-0,037	0,007	-5,432	0,000	
			Różnicowanie	1				
			Średnia ruchoma (MA)	Opóźnienie 1	0,343	0,112	3,059	0,003
	STOPA DYSKONTOWA	Bez transformacji	Minimalne opóźnienie	1				
			Licznik	Opóźnienie 0	0,467	0,155	3,016	0,004
			Różnicowanie		1			
			Mianownik	Opóźnienie 1	-0,689	0,067	-10,252	0,000
				Opóźnienie 2	-0,871	0,067	-13,011	0,000
			Minimalne opóźnienie	1				
	OPROCENTOWANIE POŻYCZEK DLA NAJLEPSZYCH KLIENTÓW	Bez transformacji	Licznik	Opóźnienie 0	-0,429	0,141	-3,043	0,003
			Różnicowanie		1			
			Mianownik	Opóźnienie 1	-0,632	0,114	-5,528	0,000
				Opóźnienie 2	-0,796	0,128	-6,225	0,000
	OPROCENTOWANIE 3-MIESIĘCZNYCH BONÓW SKARBOWYCH	Bez transformacji	Licznik	Opóźnienie 0	-0,153	0,063	-2,441	0,018
				Opóźnienie 2	0,165	0,064	2,587	0,012
			Różnicowanie		1			
			Mianownik	Opóźnienie 1	-0,481	0,078	-6,128	0,000
		Opóźnienie 2	-0,944	0,064	-14,666	0,000		

Rysunek 7.60
ARIMA. Raport:
Parametry modelu

Wartości obserwowanego poziomu istotności są mniejsze od zakładanego poziomu istotności 0,05. Można więc z prawdopodobieństwem $(1 - 0,05) = 0,95$ twierdzić, że każdy z parametrów jest istotnie różny od zera. Zauważmy również, że oszacowany model to model MA rzędu 1 dla szeregu czasowego po jednokrotnym obliczeniu różnic (czyli ARIMA (0, 1, 1)) bez transformacji stabilizujących wariancji (bez logarytmowania czy pierwiastkowania) i że w modelu uwzględniono stałą.

Do objaśniania stopy bezrobocia po jednokrotnym obliczeniu różnic wykorzystano następujące zmienne po jednokrotnym obliczeniu różnic (gdyż dla każdej ze zmiennych niezależnych zapisano RÓŻNICOWANIE: 1):

- wartości stopy dyskontowej opóźnione o 1 miesiąc (MINIMALNE OPÓŹNIENIE: 1, LICZNIK: 0);
- odchylenia wartości stopy dyskontowej od jej średniej opóźnione o 2 i 3 miesiące (MINIMALNE OPÓŹNIENIE: 1, MIANOWNIK: 1 i 2);
- wartości oprocentowania pożyczek dla najlepszych klientów bez opóźnień (LICZNIK: 0);
- odchylenia wartości oprocentowania pożyczek dla najlepszych klientów od jej średniej opóźnione o 1 i 2 miesiące (MIANOWNIK: 1 i 2);
- wartości oprocentowania 3-miesięcznych bonów skarbowych bez opóźnień i opóźnione o 2 miesiące (LICZNIK: 0 i 2);



- odchylenia wartości oprocentowania 3-miesięcznych bonów skarbowych od jej średniej opóźnione o 1 i 2 miesiące (MIANOWNIK: 1 i 2).

W modelu uwzględniono występowanie trzech wartości oddalonych (rys. 7.61).

				WARTOŚĆ OSZACOWANIA	BŁĄD STANDARDOWY	T	ISTOTNOŚĆ
STOPA BEZROBOCIA-MODEL_1	KWI 1995	INNOWACYJNY		0,358	0,095	3,762	0,000
	SIE 1996	WYGASAJĄCY	Magnituda	-0,377	0,084	-4,470	0,000
			Czynnik spadku	0,472	0,204	2,311	0,024
	KWI 1998	ADDYTYWNY		-0,304	0,077	-3,952	0,000

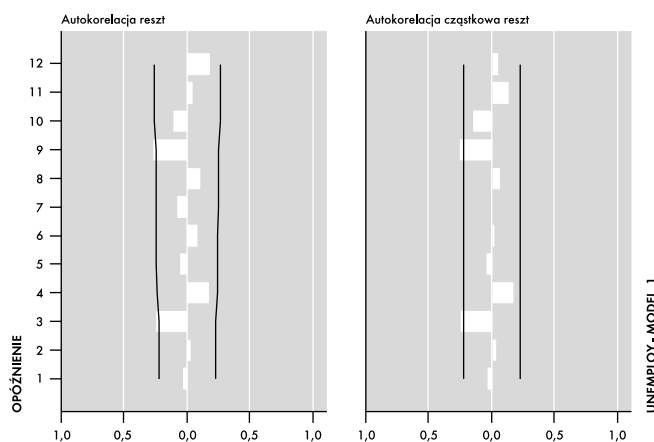
Rysunek 7.61
ARIMA. Raport:
Obserwacje odstające

Warto zwrócić uwagę, że w zestawieniu podano daty, w których obserwowano wartości oddalone, ich typ oraz fakt, że na poziomie istotności 0,05 zmienne opisujące wartości oddalone wpływają istotnie na zmienną zależną.

Wartości współczynników autokorelacji i autokorelacji cząstkowej reszt przedstawione zostały na rysunkach 7.62 i 7.63.

	MODEL				MODEL		
	STOPA BEZROBOCIA-MODEL_1				STOPA BEZROBOCIA-MODEL_1		
	FUNKCJA AUTOKORELACJI (ACF)	BŁĄD STANDARDOWY			CZĄSTKOWA FUNKCJA AUTOKORELACJI (ACF)	BŁĄD STANDARDOWY	
a	1	-0,033	0,114	1	-0,033	0,114	
b	2	0,032	0,114	2	0,030	0,114	
	3	-0,249	0,114	3	-0,248	0,114	
	4	0,170	0,121	4	0,165	0,114	
	5	-0,054	0,124	5	-0,043	0,114	
	6	0,080	0,124	6	0,014	0,114	
	7	-0,078	0,125	7	0,005	0,114	
	8	0,106	0,126	8	0,061	0,114	
	9	-0,267	0,127	9	-0,255	0,114	
	10	-0,109	0,134	10	-0,154	0,114	
	11	0,046	0,135	11	0,131	0,114	
	12	0,182	0,135	12	0,045	0,114	

Rysunek 7.62
ARIMA. Raport:
a) Autokorelacja reszt;
b) Autokorelacja cząstkowa reszt



Rysunek 7.63
ARIMA. Raport:
Wykres autokorelacji

Niskie wartości sugerują brak występowania autokorelacji składników losowych, co wcześniej sprawdzono już formalnie, wykorzystując test Ljunga–Boxa.

Na rysunkach 7.64 zaprezentowane są wartości prognoz wraz z wartościami końców przedziałów ufności.

MODEL		PAŹ 1999	LIS 1999	GRU 1999	STY 2000
STOPA BEZROBOCIA-MODEL_1	WARTOŚCI PROGNOZOWANE	4,20	4,09	4,07	4,08
	GÓRNA GRANICA PRZEDZIAŁU UFNOŚCI	4,38	4,30	4,30	4,35
	DOLNA GRANICA PRZEDZIAŁU UFNOŚCI	4,03	3,89	3,83	3,82

Dla każdego modelu, prognozowanie rozpoczyna się od obserwacji następującej po ostatniej ważnej wartości wyznaczonej w ramach wyspecyfikowanego do estymacji okresu, a kończy się, w zależności co następuje wcześniej, albo na ostatniej obserwacji, dla której – w celu obliczenia prognoz – dostępne są ważne wartości we wszystkich dodatkowych zmiennych (predyktorach), albo na obserwacji odpowiadającej dacie wyznaczającej koniec wyspecyfikowanego do prognozowania okresu.

