

## [WDEC] Wspomaganie Decyzji

### Laboratorium 11A

### ARIMA

### Modele prognozowania

#### Czynności wstępne

##### Stworzenie biblioteki Lab11A

```
libname Lab11A '/folders/myfolders/Lab11A';
```

#### Generacja danych

1) Wykorzystując data step systemu sas wygeneruj tabelę z danymi o charakterystyce procesu ARIMA (2,3).

- model  $u = 5.0 + 0.14*u_1 + 0.71*u_2 + a - 0.5*a_1 + 0.7*a_2 + 0.2*a_3$ ;
- współczynnik  $a$  powinien być obliczony z wykorzystaniem rozkładu normalnego o średniej 0 i wariancji 0.4.
- ilość wygenerowanych wartości: 2000
- ilość punktów startowych 100

```
data Lab11A.dane;  
    u1=0; u2=0; a1 = 0; a2=0; a3 = 0;  
  
    do i = -100 to 2000;  
        a = 0 + sqrt(0.4)*rannor( 32565 );  
        u = 5 + 0.14*u1 + 0.71*u2 + a - 0.5*a1 + 0.7*a2 +
```

```

0.2*a3;
    if i > 0 then output;
    u2 = u1; u1 = u; a3 = a2; a2 = a1; a1 = a;
end;
run;

```

- 2) Oblicz następujące statystyki dla wygenerowanych danych: średnia, min, max, oraz odchylenie standardowe (PROC MEANS).

```

proc means data = Lab11A.dane N MEAN MIN MAX STDDEV;
    title "Statystyki wygenerowanych danych";
run;

```

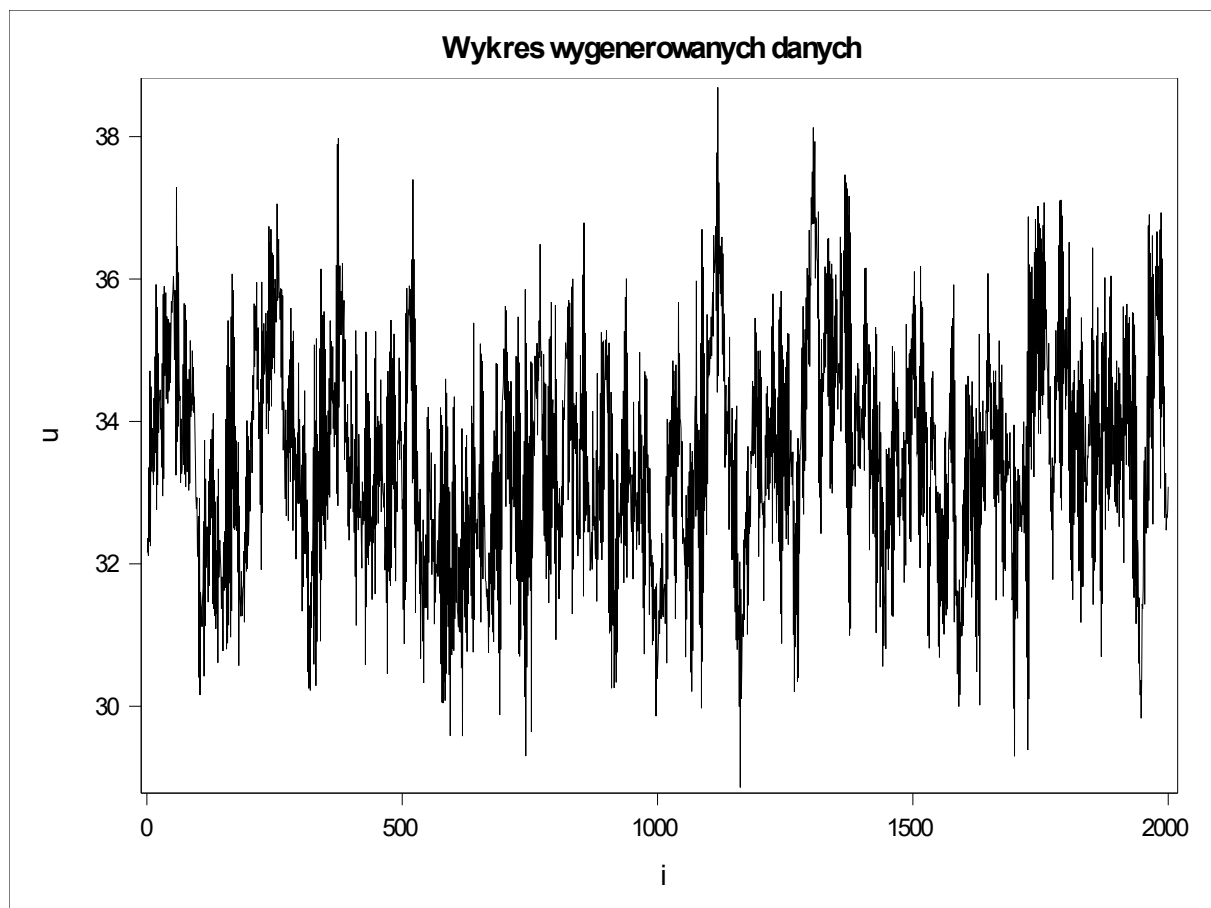
Zmienna	N	Średnia	Minimum	Maksimum	Odch. std.
u1	2000	33.4695496	28.8591425	38.6922481	1.5770764
u2	2000	33.4691776	28.8591425	38.6922481	1.5773334
a1	2000	0.0147600	-1.9512978	2.0906115	0.6427077
a2	2000	0.0145809	-1.9512978	2.0906115	0.6427395
a3	2000	0.0144135	-1.9512978	2.0906115	0.6428490
i	2000	1000.50	1.0000000	2000.00	577.4945887
a	2000	0.0154910	-1.9512978	2.0906115	0.6425208
u	2000	33.4700534	28.8591425	38.6922481	1.5767913

- 3) Narysuj wykres wygenerowanych danych.

```

proc sgplot data = Lab11A.dane;
    title "Wykres wygenerowanych danych";
    series x = i y = u;
run;

```



### Obliczenia modelu

- 1) Dla wygenerowanych danych zbuduj model ARMA(1,1) i oblicz prognozę na podstawie tego modelu.

```
/* Zbudowanie modelu ARMA(1,1) i obliczenie prognozy na podstawie  
tego modelu */  
proc arima data=Lab11A.dane;  
    title "Zbudowanie modelu ARMA(1,1) i obliczenie prognozy na  
podstawie tego modelu";
```

```

identify var=u scan;
run;

estimate p=1 q=1 /*printall*/ plot method=ML;
run;

outlier /*ALPHA=0.05 TYPE=additive*/;
run;

forecast id = i out=Lab11A.arma_1_1;
run;

/* Wykres obliczonej prognozy */
proc sgplot data=LAB11A.arma_1_1;
    title "Wykres prognozy na podstawie modelu ARMA(1,1)";
    series x=i y=FORECAST;
    series x=i y=u;
run;

```

***Zbudowanie modelu ARMA(1,1) i obliczenie prognozy na podstawie tego modelu***

### ***Procedura ARIMA***

Nazwa zmiennej = u	
Średnia szeregu roboczego	33.47005
Odchylenie standardowe	1.576397
Liczba obserwacji	2000

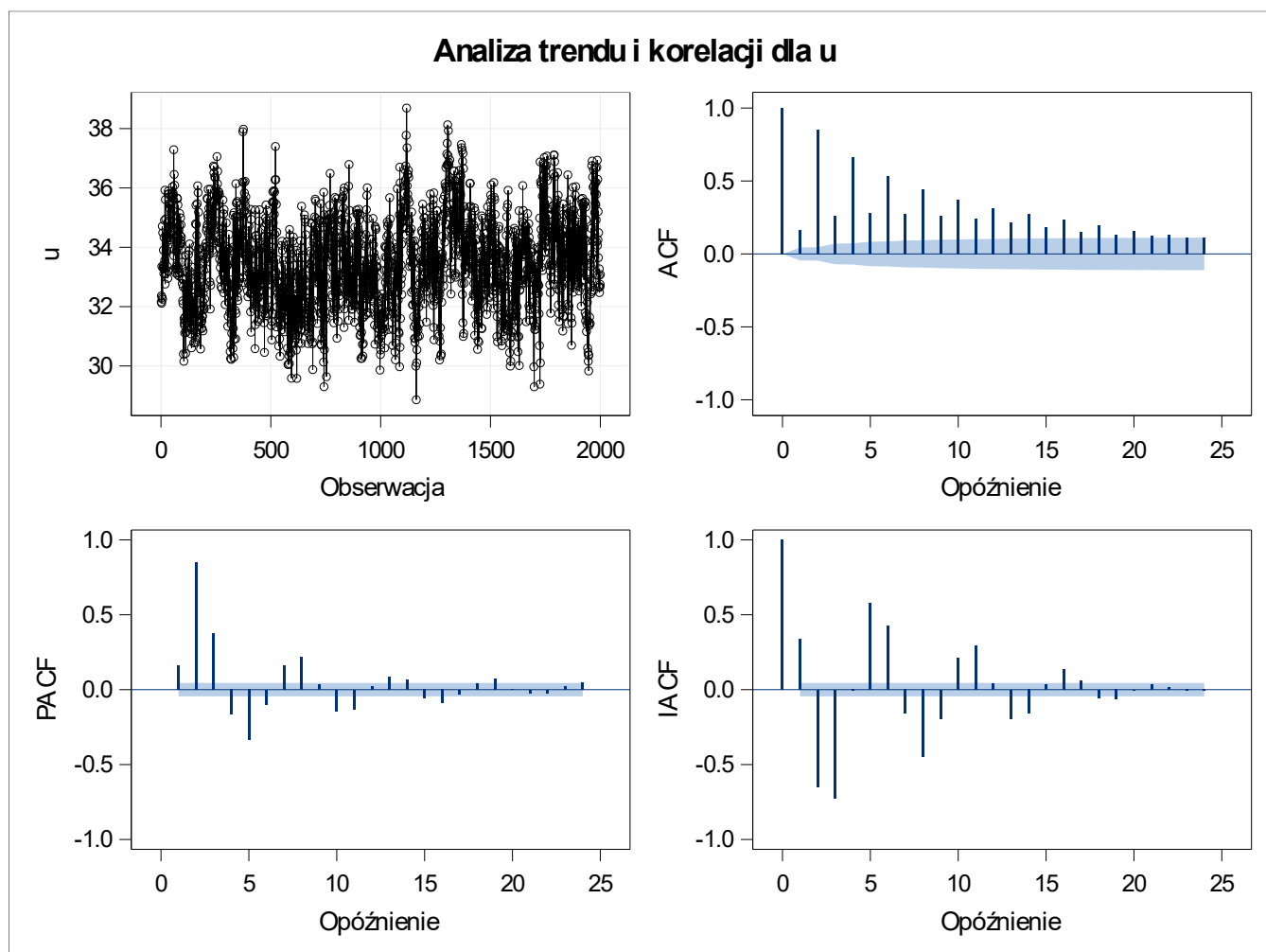
Kontrola autokorelacji względem białego szumu									
Opóźnienie do	Chi-kwadrat	DF	Pr. > chi-kw.	Autokorelacje					
6	3251.87	6	<.0001	0.161	0.854	0.260	0.663	0.277	0.531
12	4514.91	12	<.0001	0.273	0.440	0.258	0.370	0.241	0.313
18	5063.04	18	<.0001	0.215	0.272	0.182	0.236	0.153	0.198
24	5266.07	24	<.0001	0.133	0.160	0.123	0.129	0.114	0.111

Oceny kwadratowej korelacji kanonicznej						
Opóźnienia	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	0.0260	0.7299	0.0675	0.4408	0.0769	0.2825
AR 1	0.7233	0.5913	0.3197	0.2220	0.1368	0.0932
AR 2	0.1439	0.0897	0.0054	0.0004	<.0001	0.0011
AR 3	0.0277	0.1381	0.0010	<.0001	0.0003	0.0007
AR 4	0.1149	0.0472	0.0002	0.0008	0.0007	0.0007
AR 5	0.0110	0.0325	0.0009	0.0008	0.0003	0.0002

Chi-kwadrat SCAN [1] Wartości prawdopodobieństwa						
Opóźnienia	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
AR 1	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
AR 2	<.0001	<.0001	0.0124	0.5063	0.8653	0.2493
AR 3	<.0001	<.0001	0.3050	0.9921	0.5582	0.4443
AR 4	<.0001	<.0001	0.6572	0.4361	0.4394	0.4758
AR 5	<.0001	<.0001	0.3415	0.4355	0.5992	0.6063

Testy wyboru porządku próbnego ARMA(p+d,q)	
SCAN	
p+d	q
3	2
2	3

*(Poziom istotności 5% )*



Metoda największej wiarygodności					
Parametr	Ocena	Błąd standardowy	Wartość t	Przybl. pr. >  t	Opóźnienie
MU	33.45694	0.18647	179.42	<.0001	0
MA1,1	0.78348	0.01831	42.79	<.0001	1
AR1,1	0.96731	0.0074406	130.00	<.0001	1

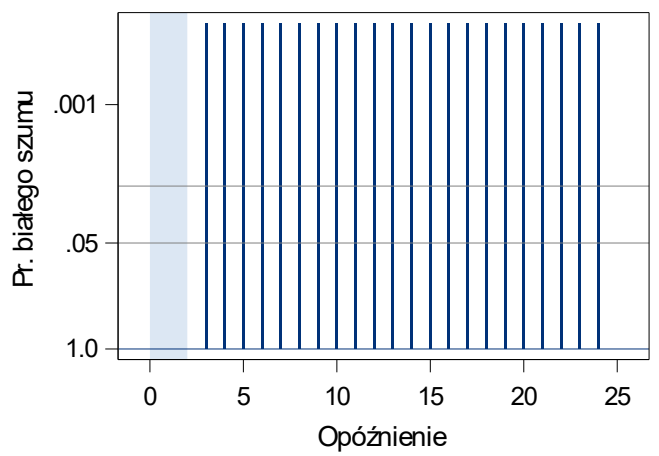
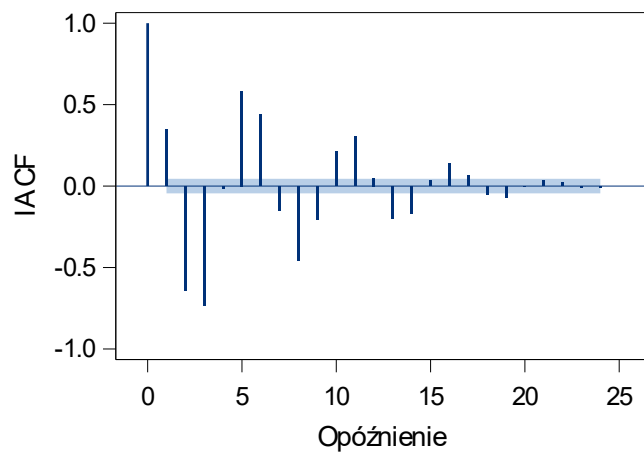
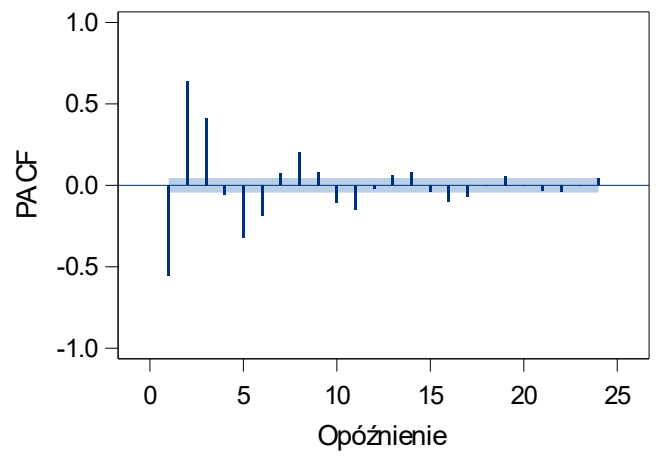
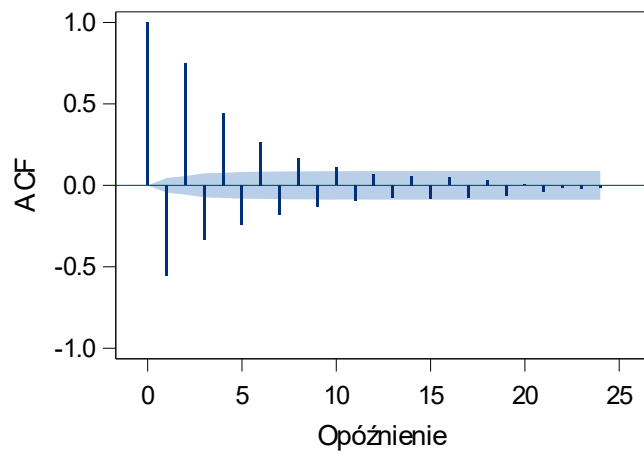
Ocena stałej	1.093815
Ocena wariancji	1.62675
Ocena bł. std.	1.275441
AIC	6652.779

<b>SBC</b>	6669.581
<b>Liczba reszt</b>	2000

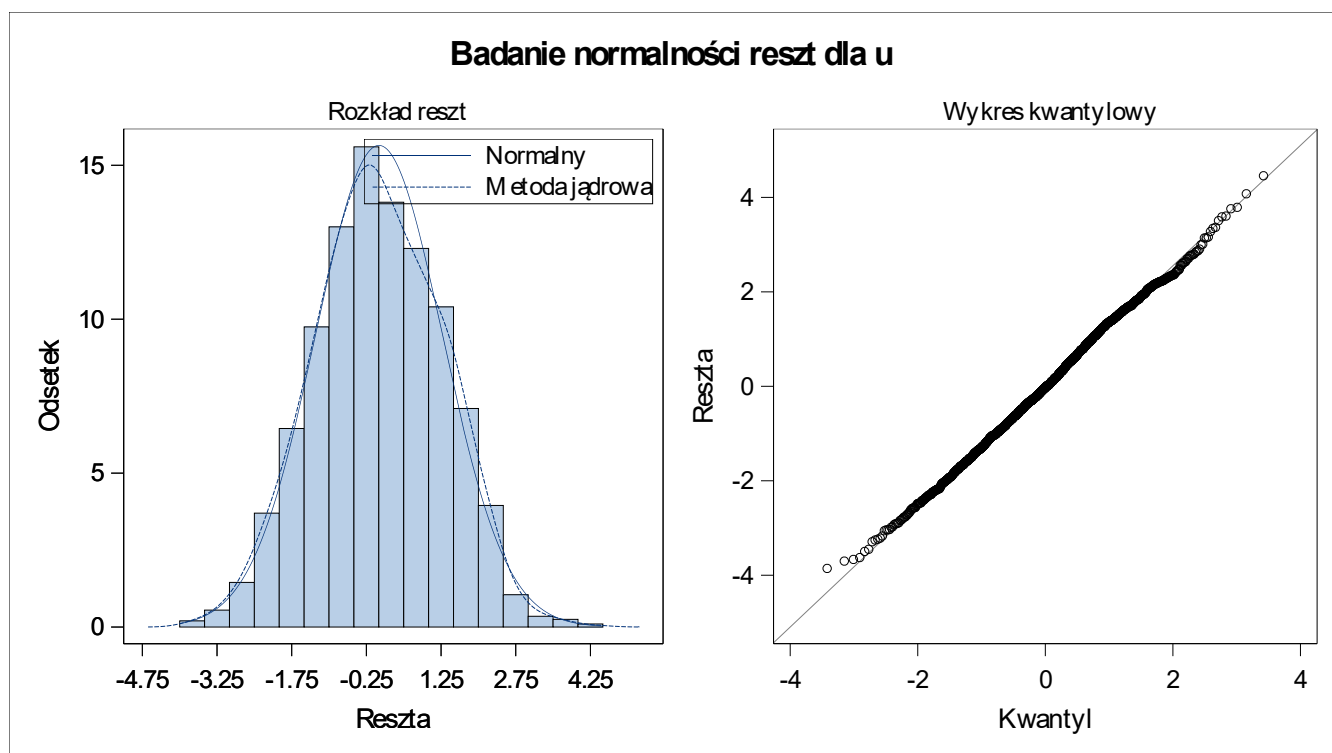
<b>Korelacje ocen parametrów</b>			
<b>Parametr</b>	<b>MU</b>	<b>MA1,1</b>	<b>AR1,1</b>
<b>MU</b>	1.000	0.001	0.001
<b>MA1,1</b>	0.001	1.000	0.653
<b>AR1,1</b>	0.001	0.653	1.000

<b>Sprawdzenie autokorelacji reszt</b>									
<b>Opóźnienie do</b>	<b>Chi-kwadrat</b>	<b>DF</b>	<b>Pr. &gt; chi-kw.</b>	<b>Autokorelacje</b>					
<b>6</b>	2623.27	4	<.0001	-0.555	0.749	-0.335	0.445	-0.241	0.266
<b>12</b>	2834.58	10	<.0001	-0.179	0.169	-0.135	0.109	-0.097	0.071
<b>18</b>	2886.23	16	<.0001	-0.080	0.056	-0.081	0.048	-0.080	0.031
<b>24</b>	2900.73	22	<.0001	-0.066	0.006	-0.043	-0.013	-0.023	-0.014
<b>30</b>	2906.56	28	<.0001	-0.018	-0.008	-0.017	-0.022	0.001	-0.041
<b>36</b>	2917.91	34	<.0001	0.004	-0.023	-0.015	0.021	-0.039	0.053
<b>42</b>	2924.38	40	<.0001	-0.033	0.039	-0.012	0.021	0.000	0.001
<b>48</b>	2926.46	46	<.0001	0.012	-0.009	0.013	-0.018	0.011	-0.013

### Badanie korelacji reszt dla u







Model dla zmiennej u	
Średnia estymowana	33.45694

Współczynniki autoregresyjne	
Współczynnik 1:	1 - 0.96731 B**(1)

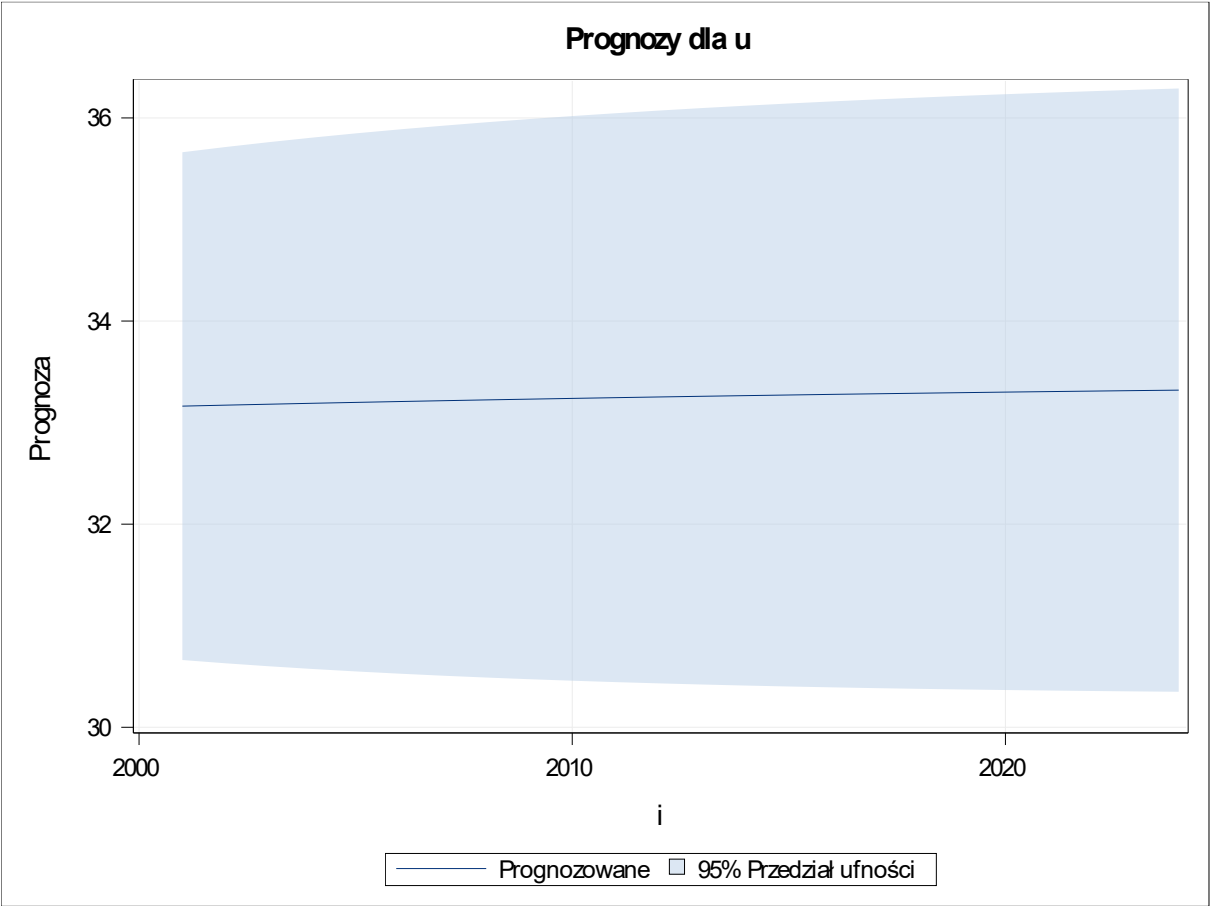
Współczynniki średniej ruchomej	
Współczynnik 1:	1 - 0.78348 B**(1)

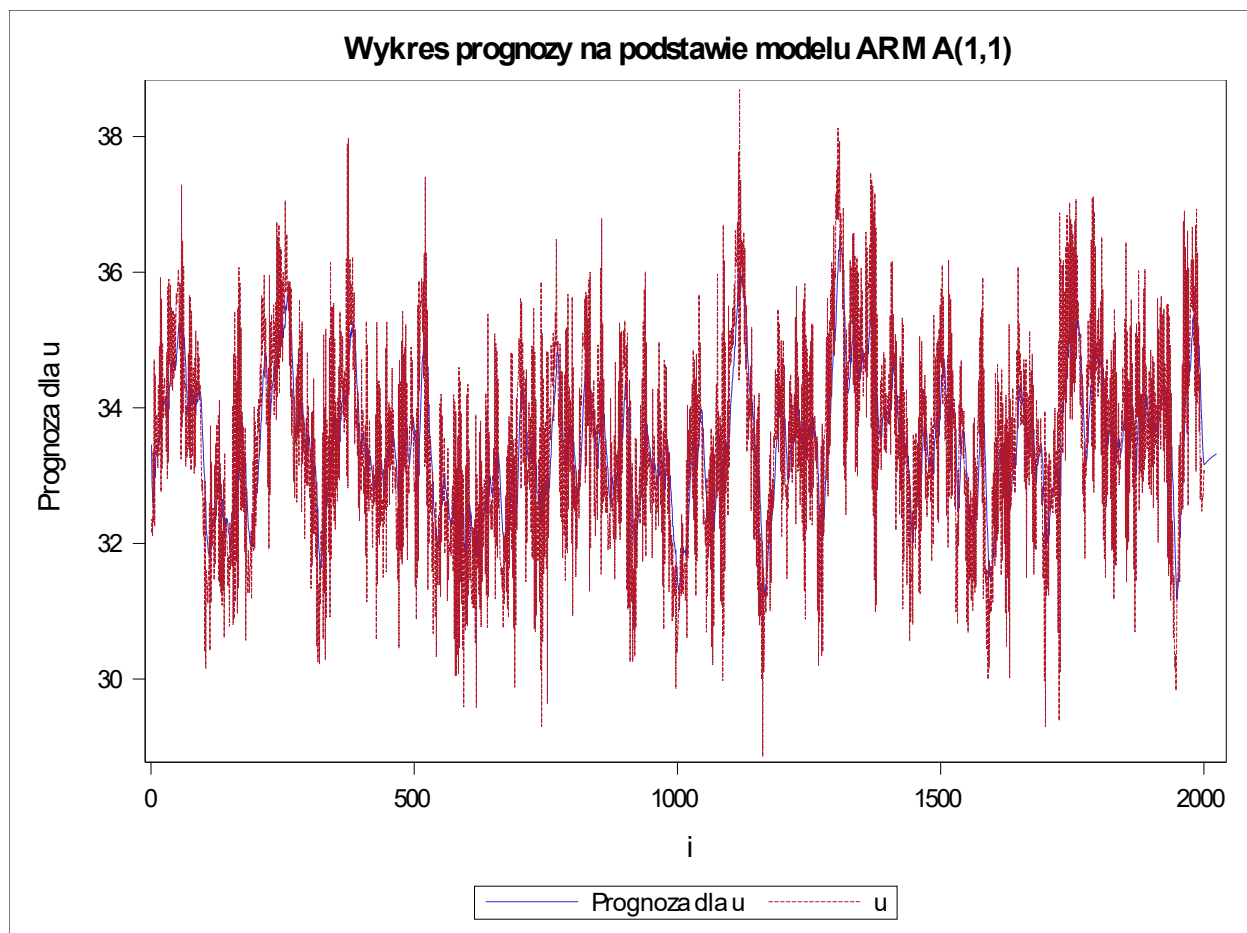
Podsumowanie wykrywania wartości odstających	
Maksymalna liczba poszukiwanych	5
Liczba znalezionych	5
Używana istotność	0.05

Informacje o wartościach odstających				
Obs.	Typ	Ocena	Chi-kwadrat	Przybl. prawd.>chi-kw.
1725	Addytywnie	-4.40295	11.53	0.0007
1727	Addytywnie	-4.36585	11.37	0.0007
1087	Addytywnie	3.81302	8.70	0.0032
753	Addytywnie	-3.80007	8.65	0.0033
742	Addytywnie	-3.72978	8.42	0.0037

Prognozy dla zmiennej u				
Obs.	Prognoza	Błąd std.	Przedział ufności 95%	
2001	33.1622	1.2754	30.6623	35.6620
2002	33.1718	1.2968	30.6301	35.7135
2003	33.1811	1.3165	30.6008	35.7614
2004	33.1901	1.3346	30.5743	35.8060
2005	33.1989	1.3514	30.5501	35.8476
2006	33.2073	1.3669	30.5282	35.8864
2007	33.2155	1.3813	30.5082	35.9227
2008	33.2234	1.3946	30.4901	35.9566
2009	33.2310	1.4069	30.4736	35.9884
2010	33.2384	1.4183	30.4586	36.0182
2011	33.2455	1.4289	30.4449	36.0462
2012	33.2524	1.4388	30.4325	36.0724
2013	33.2591	1.4479	30.4212	36.0970
2014	33.2656	1.4565	30.4110	36.1202
2015	33.2718	1.4644	30.4017	36.1420
2016	33.2779	1.4718	30.3933	36.1625
2017	33.2838	1.4787	30.3856	36.1819
2018	33.2894	1.4851	30.3788	36.2001
2019	33.2949	1.4910	30.3725	36.2172
2020	33.3002	1.4966	30.3669	36.2334
2021	33.3053	1.5018	30.3619	36.2487
2022	33.3103	1.5066	30.3574	36.2632

Prognozy dla zmiennej u				
Obs.	Prognoza	Błąd std.	Przedział ufności 95%	
2023	33.3151	1.5111	30.3533	36.2768
2024	33.3197	1.5153	30.3497	36.2897





- 2) Dla wygenerowanych danych zbuduj model ARMA(2,3) i oblicz prognozę na podstawie tego modelu.

```
/* Zbudowanie modelu ARMA(2,3) i obliczenie prognozy na podstawie
tego modelu */
proc arima data=Lab11a.dane;
    title "Zbudowanie modelu ARMA(2,3) i obliczenie prognozy na
podstawie tego modelu";

    identify var=u;
    run;

    estimate p=2 q=3 /*printall*/ plot method=ML;
    run;
```

```

outlier /*ALPHA=0.05 TYPE=additive*/;
run;

forecast id = i out=Lab11A.arma_2_3;
run;

/* Wykres obliczonej prognozy */
proc sgplot data = Lab11A.arma_2_3;
    title "Wykres prognozy na podstawie modelu ARMA(2,3)";
    series x=i y=FORECAST;
    series x=i y=u;
run;

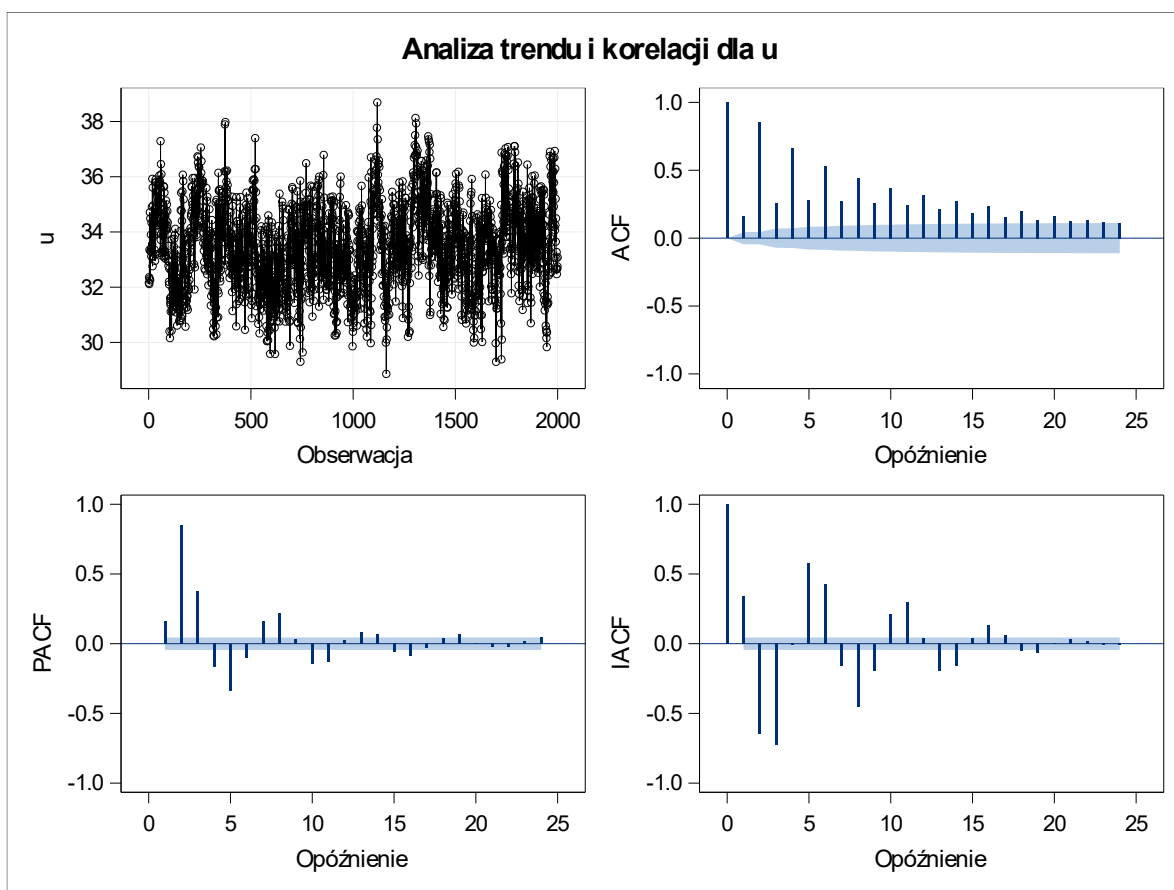
```

***Zbudowanie modelu ARMA(2,3) i obliczenie prognozy na podstawie tego modelu***

***Procedura ARIMA***

Nazwa zmiennej = u	
Średnia szeregu roboczego	33.47005
Odchylenie standardowe	1.576397
Liczba obserwacji	2000

Kontrola autokorelacji względem białego szumu									
Opóźnienie do	Chi-kwadrat	DF	Pr. > chi-kw.	Autokorelacje					
6	3251.87	6	<.0001	0.161	0.854	0.260	0.663	0.277	0.531
12	4514.91	12	<.0001	0.273	0.440	0.258	0.370	0.241	0.313
18	5063.04	18	<.0001	0.215	0.272	0.182	0.236	0.153	0.198
24	5266.07	24	<.0001	0.133	0.160	0.123	0.129	0.114	0.111



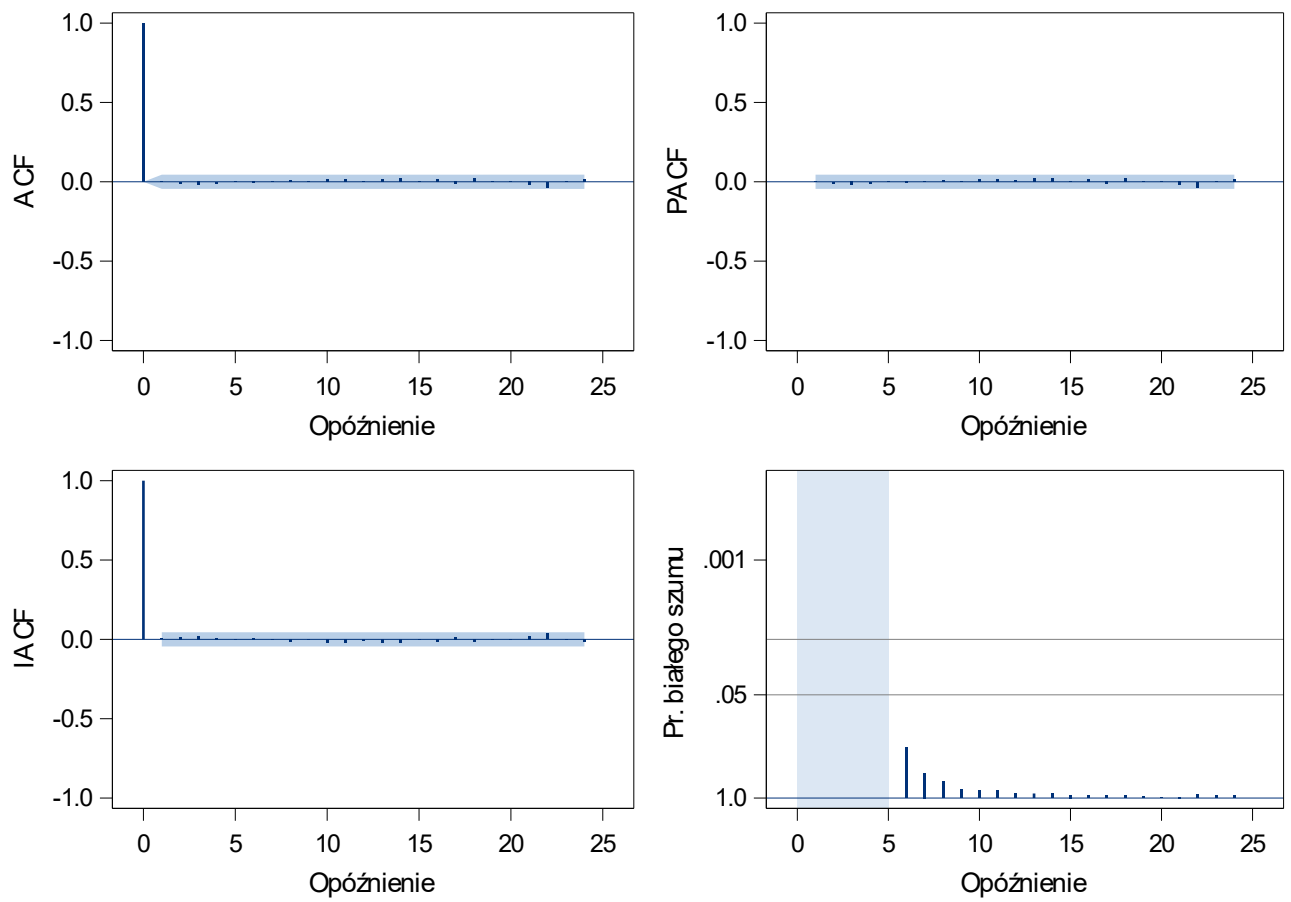
Metoda największej wiarygodności					
Parametr	Ocena	Błąd standardowy	Wartość t	Przybl. pr. >  t	Opóźnienie
MU	33.45887	0.13655	245.03	<.0001	0
MA1,1	0.50672	0.03349	15.13	<.0001	1
MA1,2	-0.67492	0.02856	-23.63	<.0001	2
MA1,3	-0.22152	0.03309	-6.69	<.0001	3
AR1,1	0.11980	0.02353	5.09	<.0001	1
AR1,2	0.73518	0.01916	38.36	<.0001	2

Ocena stałej	4.852126
Ocena wariancji	0.412188
Ocena bł. std.	0.642019
AIC	3915.173
SBC	3948.779
Liczba reszt	2000

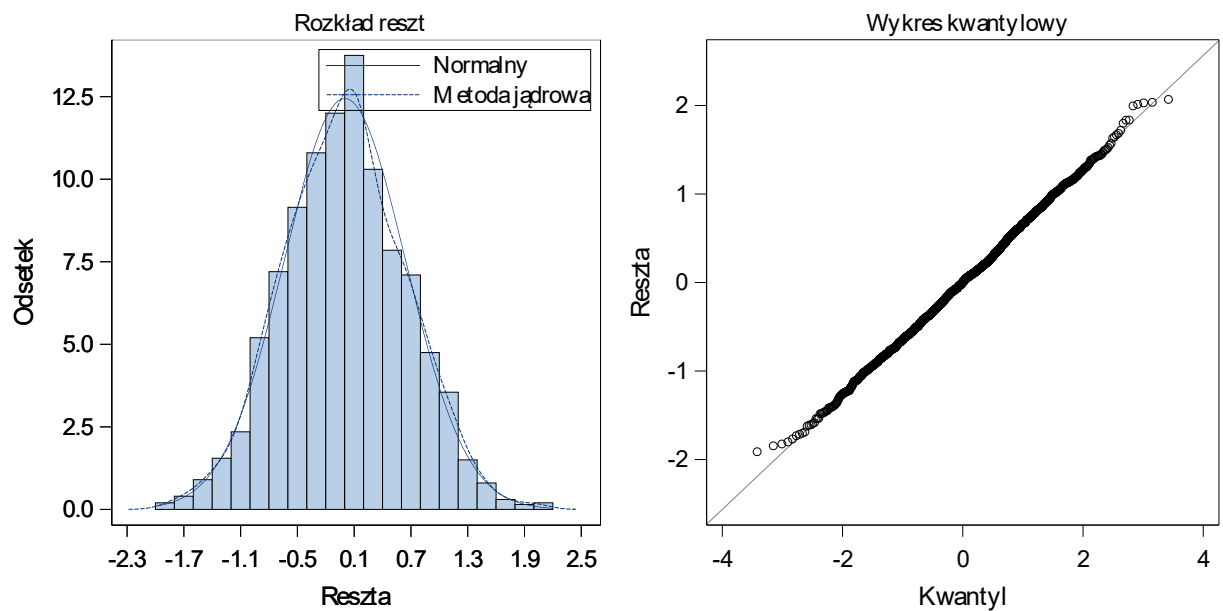
Korelacje ocen parametrów						
Parametr	MU	MA1,1	MA1,2	MA1,3	AR1,1	AR1,2
MU	1.000	-0.001	0.001	-0.001	-0.002	-0.001
MA1,1	-0.001	1.000	-0.896	0.925	0.755	-0.581
MA1,2	0.001	-0.896	1.000	-0.894	-0.711	0.607
MA1,3	-0.001	0.925	-0.894	1.000	0.746	-0.573
AR1,1	-0.002	0.755	-0.711	0.746	1.000	-0.686
AR1,2	-0.001	-0.581	0.607	-0.573	-0.686	1.000

Sprawdzenie autokorelacji reszt									
Opóźnienie do	Chi-kwadrat	DF	Pr. > chi-kw.	Autokorelacje					
6	1.45	1	0.2283	-0.002	-0.011	-0.020	-0.010	0.006	-0.008
12	3.08	7	0.8773	-0.001	0.012	-0.001	0.017	0.018	0.006
18	6.37	13	0.9318	0.019	0.020	-0.002	0.016	-0.015	0.019
24	10.48	19	0.9401	-0.003	-0.001	-0.020	-0.036	0.005	0.017
30	18.72	25	0.8105	0.032	0.024	-0.023	0.004	0.014	-0.042
36	29.76	31	0.5296	0.003	-0.025	0.000	0.005	-0.017	0.067
42	34.05	37	0.6079	0.000	0.007	0.023	0.017	-0.020	-0.029
48	38.95	43	0.6476	0.025	0.028	0.006	-0.021	0.008	-0.020

### Badanie korelacji reszt dla u



### Badanie normalności reszt dla u





Model dla zmiennej u	
Średnia estymowana	33.45887

Współczynniki autoregresyjne	
Współczynnik 1:	$1 - 0.1198 B^{**}(1) - 0.73518 B^{**}(2)$

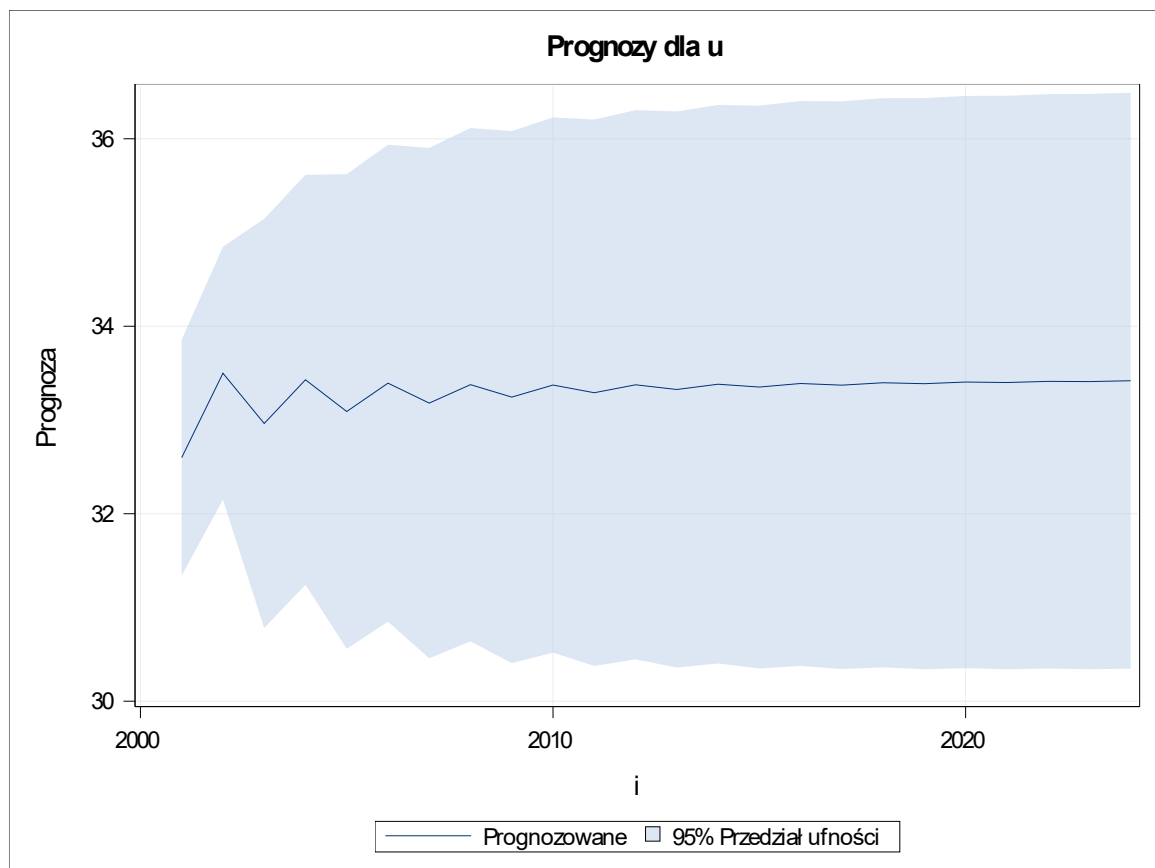
Współczynniki średniej ruchomej	
Współczynnik 1:	$1 - 0.50672 B^{**}(1) + 0.67492 B^{**}(2) + 0.22152 B^{**}(3)$

Podsumowanie wykrywania wartości odstających	
Maksymalna liczba poszukiwanych	5
Liczba znalezionych	5
Używana istotność	0.05

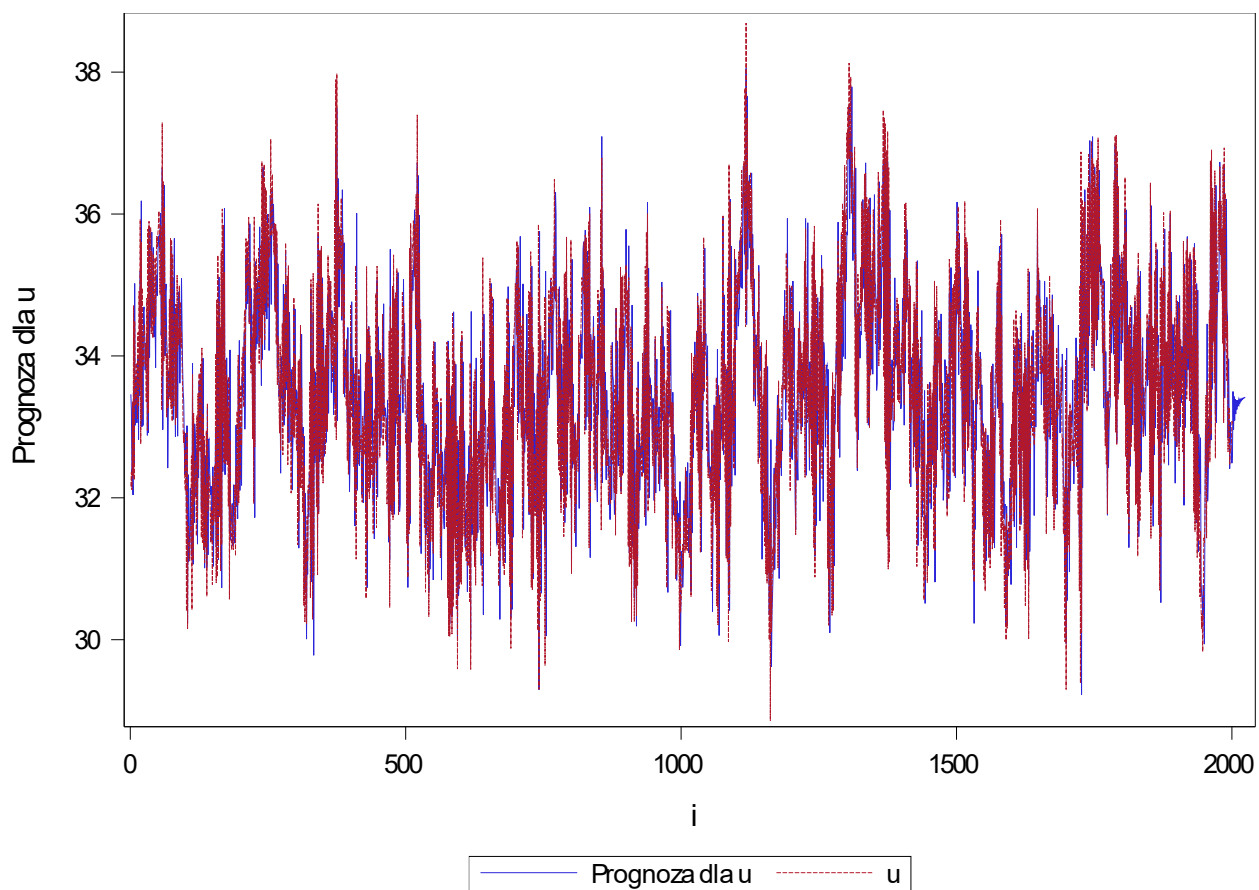
Informacje o wartościach odstających				
Obs.	Typ	Ocena	Chi-kwadrat	Przybl. prawd.>chi-kw.
168	Addytywnie	0.66307	9.77	0.0018
331	Addytywnie	-0.64370	9.25	0.0024
1155	Addytywnie	0.63708	9.04	0.0026
1185	Addytywnie	0.70900	11.25	0.0008
1246	Przesunięcie	0.47397	9.18	0.0024

Prognozy dla zmiennej u				
Obs.	Prognoza	Błąd std.	Przedział ufności 95%	
2001	32.5985	0.6420	31.3401	33.8568
2002	33.4991	0.6884	32.1499	34.8484
2003	32.9627	1.1138	30.7797	35.1456
2004	33.4290	1.1156	31.2424	35.6156
2005	33.0905	1.2919	30.5584	35.6226
2006	33.3928	1.2980	30.8488	35.9368
2007	33.1801	1.3888	30.4582	35.9021
2008	33.3769	1.3970	30.6388	36.1150

<b>2009</b>	33.2441	1.4481	30.4059	36.0824
<b>2010</b>	33.3729	1.4566	30.5180	36.2277
<b>2011</b>	33.2907	1.4870	30.3763	36.2051
<b>2012</b>	33.3755	1.4946	30.4461	36.3050
<b>2013</b>	33.3252	1.5135	30.3589	36.2916
<b>2014</b>	33.3816	1.5199	30.4026	36.3606
<b>2015</b>	33.3514	1.5320	30.3488	36.3540
<b>2016</b>	33.3892	1.5371	30.3765	36.4018
<b>2017</b>	33.3715	1.5450	30.3433	36.3996
<b>2018</b>	33.3972	1.5490	30.3612	36.4331
<b>2019</b>	33.3872	1.5543	30.3409	36.4336
<b>2020</b>	33.4049	1.5573	30.3527	36.4572
<b>2021</b>	33.3997	1.5609	30.3405	36.4590
<b>2022</b>	33.4121	1.5631	30.3484	36.4758
<b>2023</b>	33.4098	1.5656	30.3413	36.4783
<b>2024</b>	33.4186	1.5673	30.3468	36.4904



Wykres prognozy na podstawie modelu ARM A(2,3)



- 3) Oblicz średni błąd kwadratowy (MSE) dla modelu ARMA (1,1) i modelu ARMA(2,3) i porównaj te modele z punktu widzenia jakości prognozy.

```
/* ARMA(1,1) */
data Lab11A.arma_1_1;
    set Lab11A.arma_1_1;
    mean_squared_error = RESIDUAL*RESIDUAL;
run;

proc means data = Lab11A.arma_1_1;
    title "Statystyki MSE dla modelu ARMA(1,1)";
    var mean_squared_error;
run;
/* Wynik (MSE) odczytujemy z kolumny Średnia - wynosi ok. 1.6247144
*/
```

```
/* ARMA(2,3) */
data Lab11A.arma_2_3;
    set Lab11A.arma_2_3;
    mean_squared_error = RESIDUAL*RESIDUAL;
run;

proc means data = Lab11A.arma_2_3;
    title "Statystyki MSE dla modelu ARMA(2,3)";
    var mean_squared_error;
run;
/* Wynik (MSE) odczytujemy z kolumny Średnia - wynosi ok. 0.4123600
*/
```

### ***Statystyki MSE dla modelu ARMA(1,1)***

#### ***Procedura MEANS***

Zmienna analizowana: mean_squared_error				
N	Średnia	Odch. std.	Minimum	Maksimum
2000	1.6247144	2.1598536	9.504714E-7	19.8747191

## ***Statystyki MSE dla modelu ARMA(2,3)***

### ***Procedura MEANS***

<b>Zmienna analizowana: mean_squared_error</b>				
<b>N</b>	<b>Średnia</b>	<b>Odch. std.</b>	<b>Minimum</b>	<b>Maksimum</b>
2000	0.4123600	0.5708542	5.1943472E-8	4.2800881

### Porównanie modeli z punktu widzenia jakości prognozy

Analizując średni błąd kwadratowy (MSE) dla modelu ARMA(1,1) oraz ARMA(2,3), widzimy, że dla modelu ARMA(1,1) wyniósł ok. 1.6247144, natomiast dla modelu ARMA(2,3) wyniósł ok. 0.4123600. Zatem model ARMA(2,3) ma mniejszą wartość MSE, a co za tym idzie, jest lepiej dopasowany do naszych danych niż model ARMA(1,1). Wartości liczbowe odczytałem z kolumn Średnia z dwóch ostatnich tabel.

Dalej, analizując wykresy korelacji reszt dla zmiennej  $u$ , możemy dostrzec, że w przypadku modelu ARMA(1,1) reszty wykazują charakter cykliczny, co świadczy o słabym dopasowaniu do naszych danych. Dla modelu ARMA(2,3) możemy dostrzec, że reszty są bardzo niewielkie (praktycznie zerowe), co świadczy o dobrym dopasowaniu tego modelu do naszych danych.

