

Analiza i prognozowanie cen złota za pomocą modeli ARIMA

Data: 17-07-2021

Autor: Mateusz Smolnik

Analiza i Prognozowanie Szeregów Czasowych

1. Opis danych

W niniejszym projekcie zostanie przeprowadzona analiza oraz próba prognozy cen złota przy wykorzystaniu modeli ARIMA.

Złoto jest jednym z najpopularniejszych metali szlachetnych wybieranych jako inwestycja alternatywna. Jest ono również jednym z najstarszych środków lokowania kapitału. Już od zarania dziejów ludzkości metal ten był wykorzystywany jako środek tezauryzacji. Obecnie nie ma państwa, którego waluta opierałby się na złocie. Natomiast wiele krajów posiada rezerwy państwowe w złocie. Złoto zawdzięcza swoją popularność również swojemu zastosowaniu w przemyśle, gdzie może zostać wykorzystane przy produkcji układów scalonych, tranzystorów a również w różnych działach medycyny np. stomatologii. Jednakże badanie dotyczy jedynie cen złota inwestycyjnego.

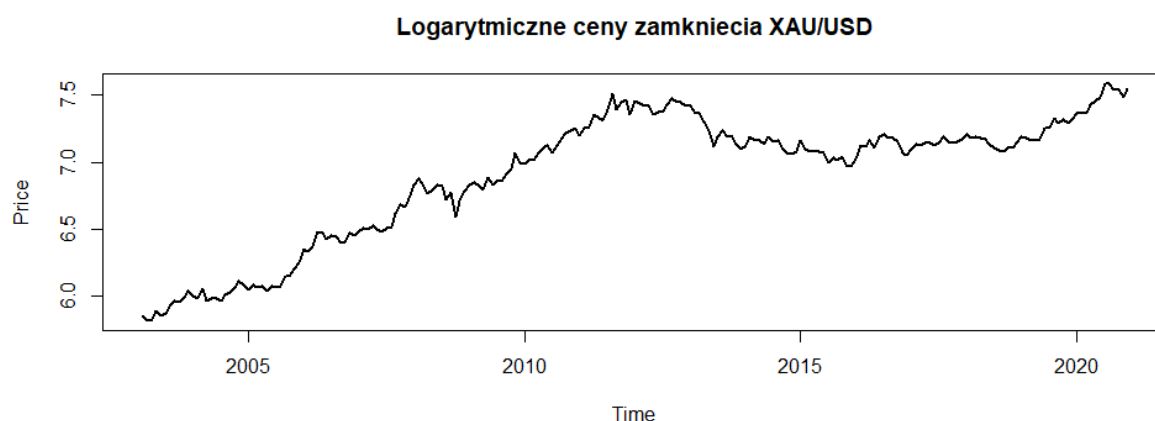
W analizie zostały uwzględnione dane 215 logarytmicznych cen zamknięcia w interwale miesięcznym. Ceny są podane w dolarze amerykańskim za jedną uncję złota. Pochodzą one z okresu od lutego 2003 roku do grudnia 2020 roku.

Dane do badania zostały pobrane z serwisu <https://stoq.com>. (Data dostępu: 17.07.2021r.)

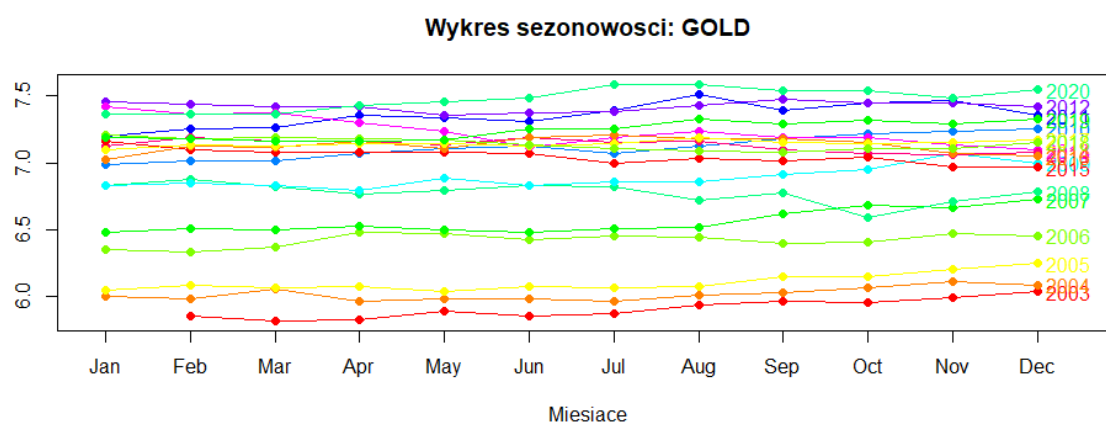
2. Wstępna analiza danych

Przed przystąpieniem do właściwej analizy oraz budowy modeli została wykonana wstępna analiza danych. W pierwszej kolejności przeanalizowany został wykres logarytmicznych cen zamknięcia (wykres 1). Można na nim zaobserwować, że analizowany szereg nie jest stacjonarny, a dane charakteryzują się trendem. Nie widać natomiast, aby dane wykazywały wahania sezonowe. Jednak aby to potwierdzić został również wygenerowany wykres sezonowości, który przedstawiono na wykresie 2. Na tym wykresie również nie widać szczególnych zależności między danymi miesięcznymi, więc przyjęte zostało, że badane dane nie zawierają w sobie sezonowości.

Wykres 1 Logarytmiczne ceny zamknięcia XAU/USD w interwale miesięcznym



Wykres 2 Wykres sezonowości cen zamknięcia XAU/USD



W kolejnym kroku przeanalizowane zostały statystyki opisowe dla danych z całego okresu, które zostały zaprezentowane w tabeli 1. Ujemna wartość skośności oznacza, że dane charakteryzują się lewostronną asymetrią rozkładu, co oznacza, że większość wyników znajduje się powyżej średniej. Zauważyć można również, że wartość kurtozy jest ujemna co świadczy o tym, że wartości cechy są mniej skoncentrowane w porównaniu z rozkładem normalnym

Tabela 1 Statystyki opisowe

vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
1	215	6.9	0.48	7.1	6.95	0.38	5.81	7.59	1.77	-0.84	-0.53	0.03

3. Właściwa analiza danych i budowa modeli

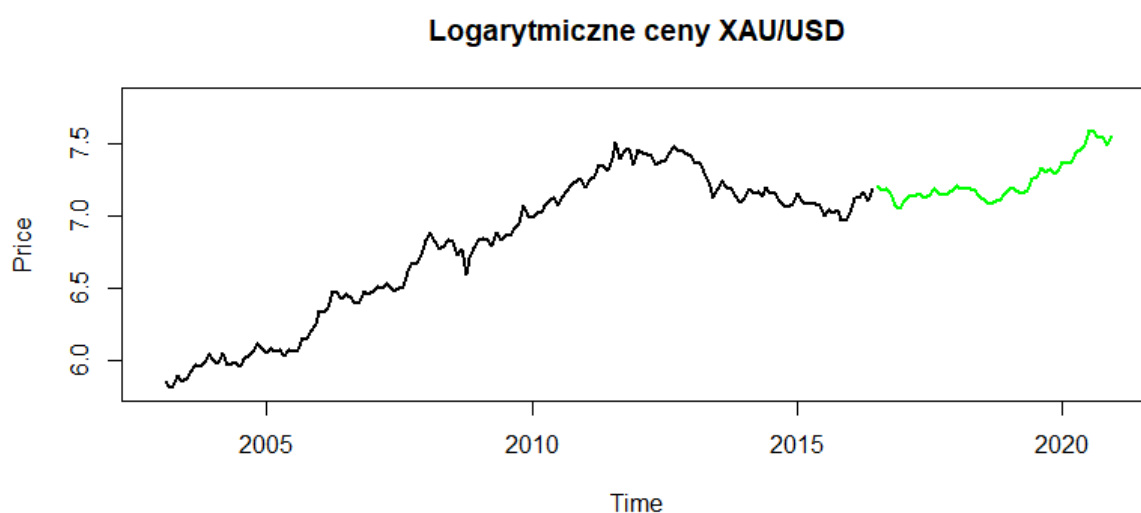
W dalszej części badania dane zostały podzielone na zbiór uczący oraz zbiór testowy. Podział został dokonany w poniższy sposób:

- Zbiór uczący – 75% danych, obserwacje od 1 do 161, okres od lutego 2003 roku do czerwca 2016 roku,

- Zbiór testowy – 25% danych, obserwacje od 162 do 215, okres od lipca 2016 roku do grudnia 2020 roku.

Na wykresie 3 zostały zaprezentowane dane z podziałem na zbiór uczący (kolor czarny) oraz zbiór testowy (kolor zielony).

Wykres 3 Logarytmiczne ceny XAU/USD z podziałem na zbiór uczący oraz testowy

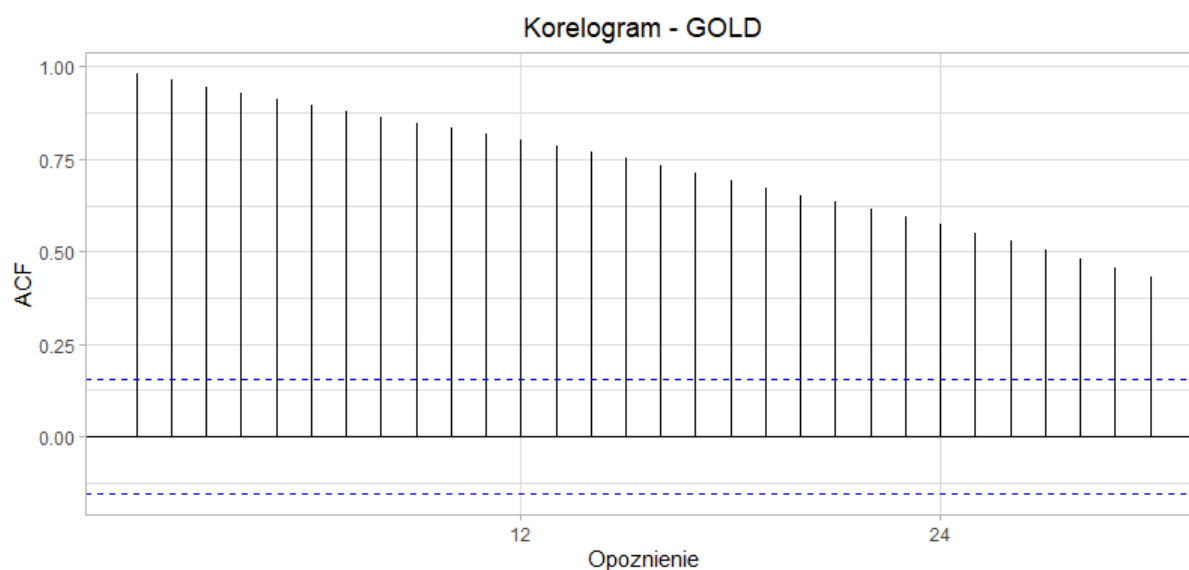


Od tego momentu dalsze analizy prowadzące do budowy modelu będą przeprowadzane na zbiorze uczącym.

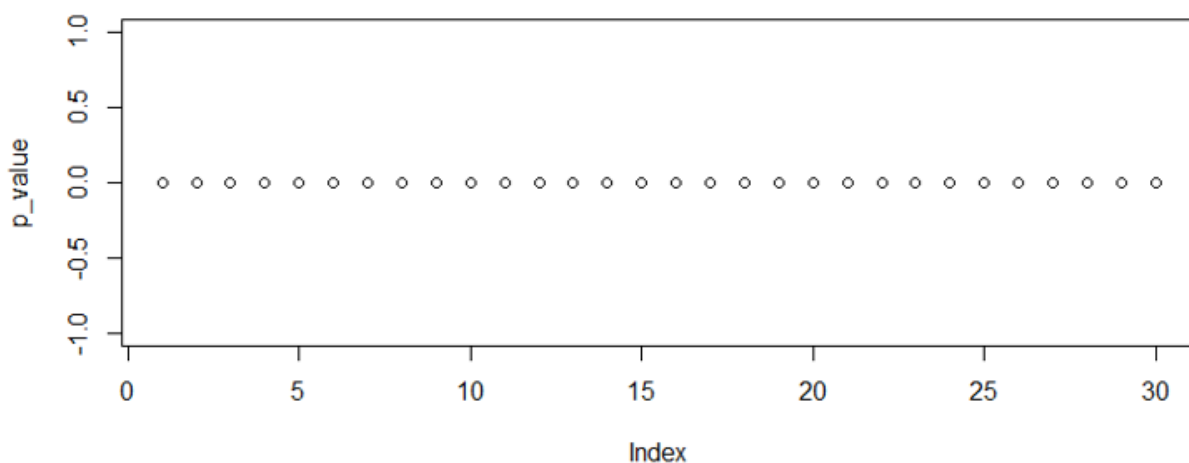
3.1. Badanie autokorelacji

Na wykresie 4 został zaprezentowany korelogram dla rzędu opóźnień od 1 do 30. Zauważyć można, że dla wszystkich rzędów opóźnień mamy istotną autokorelację. Potwierdza to również test Ljung-Boxa na autokorelację, którego hipoteza zerowa mówi o brak autokorelacji, wartości p dla wszystkich rzędów opóźnień są mniejsze od 0,05, więc hipoteza zerowa została odrzucona. Wyniki testu zostały zaprezentowane na wykresie 5 oraz rysunku 1.

Wykres 4 Korelogram dla rzędów opóźnień od 1 do 30



Wykres 5 Wartości p testu Ljung-Boxa dla danych opóźnień



Rysunek 1 Wynik testu Ljung-Boxa dla opóźnienia 30

```
Box-Ljung test
data:  gold_train
X-squared = 2937.6, df = 30, p-value < 2.2e-16
```

3.2. Badanie stacjonarności

W celu potwierdzenia zaobserwowanej wcześniej na wykresach niestacjonarności danych został przeprowadzony test pierwiastka jednostkowego Dickeya-Fullera, którego H_0 zakłada, że szereg jest niestacjonarny. Ponieważ wartości p testu ADF były większe niż 0,05 to nie było podstaw, aby odrzucić hipotezę zerową, świadczy to o tym, że badany szereg jest niestacjonarny, więc konieczne będzie różnicowanie szeregu w celu uzyskanie stacjonarnego szeregu czasowego.

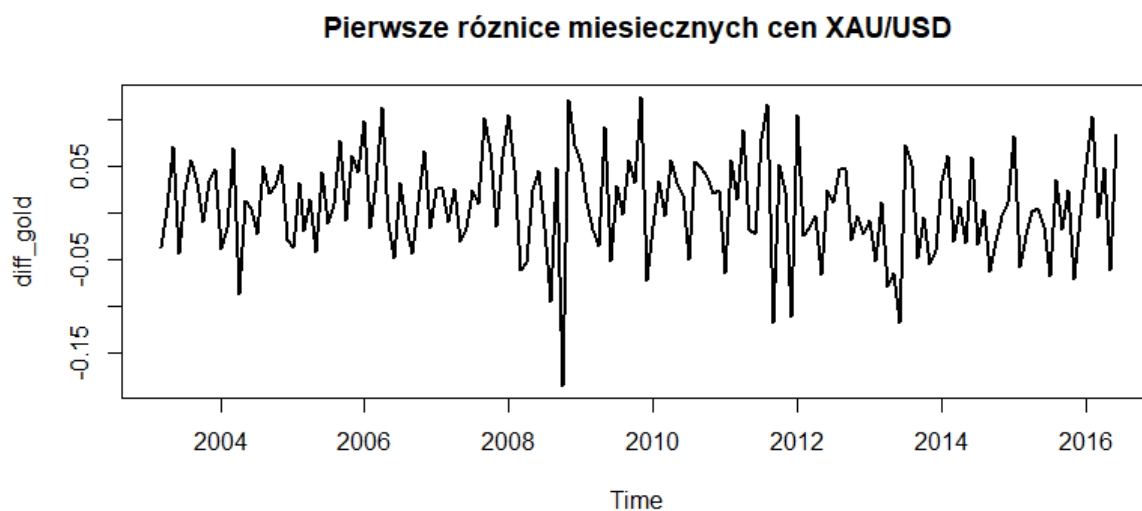
Rysunek 2 Wyniki testu ADF na stacjonarność

Type 1: no drift no trend			
	lag	ADF	p.value
[1,]	0	1.86	0.984
[2,]	1	2.19	0.990
[3,]	2	2.21	0.990
[4,]	3	2.03	0.990
[5,]	4	2.01	0.989
Type 2: with drift no trend			
	lag	ADF	p.value
[1,]	0	-1.76	0.424
[2,]	1	-2.01	0.324
[3,]	2	-2.09	0.294
[4,]	3	-1.90	0.367
[5,]	4	-2.04	0.314
Type 3: with drift and trend			
	lag	ADF	p.value
[1,]	0	-1.092	0.921
[2,]	1	-0.852	0.955
[3,]	2	-0.803	0.959
[4,]	3	-0.746	0.964
[5,]	4	-0.793	0.960

3.3 Usuwanie niestacjonarności

W celu usunięcia niestacjonarności szeregu zostały utworzone jego pierwsze różnice. Na wykresie 6 został zaprezentowany szereg czasowy pierwszych różnic. Można zauważyć, że dane przedstawione na wykresie wyglądają na szereg stacjonarny. W celu potwierdzenia tej hipotezy zostanie przeprowadzony test ADF dla pierwszych różnic cen.

Wykres 6 Pierwsze różnice miesięcznych cen XAU/USD



3.4. Badanie stacjonarności dla pierwszych różnic

Test ADF dla pierwszych różnic cen wykazał, że dane te są szeregiem stacjonarnym. Wyniki testu zostały przedstawione na rysunku 3. Wartości p są mniejsze od 0,05 co daje podstawy do odrzucenia hipotezy zerowej. Nie będzie dokonywane dalsze różnicowanie szeregu. Oznacza to, że stopień integracji modelu ARIMA będzie wynosił 1 – ARIMA(p,1,q).

Rysunek 3 Wyniki testu ADF dla pierwszych różnic cen

Augmented Dickey-Fuller Test			
alternative: stationary			
Type 1: no drift no trend			
	lag	ADF	p.value
[1,]	0	-14.23	0.01
[2,]	1	-9.33	0.01
[3,]	2	-7.05	0.01
[4,]	3	-5.84	0.01
[5,]	4	-4.88	0.01
Type 2: with drift no trend			
	lag	ADF	p.value
[1,]	0	-14.62	0.01
[2,]	1	-9.76	0.01
[3,]	2	-7.46	0.01
[4,]	3	-6.30	0.01
[5,]	4	-5.33	0.01
Type 3: with drift and trend			
	lag	ADF	p.value
[1,]	0	-14.84	0.01
[2,]	1	-10.03	0.01
[3,]	2	-7.71	0.01
[4,]	3	-6.62	0.01
[5,]	4	-5.67	0.01

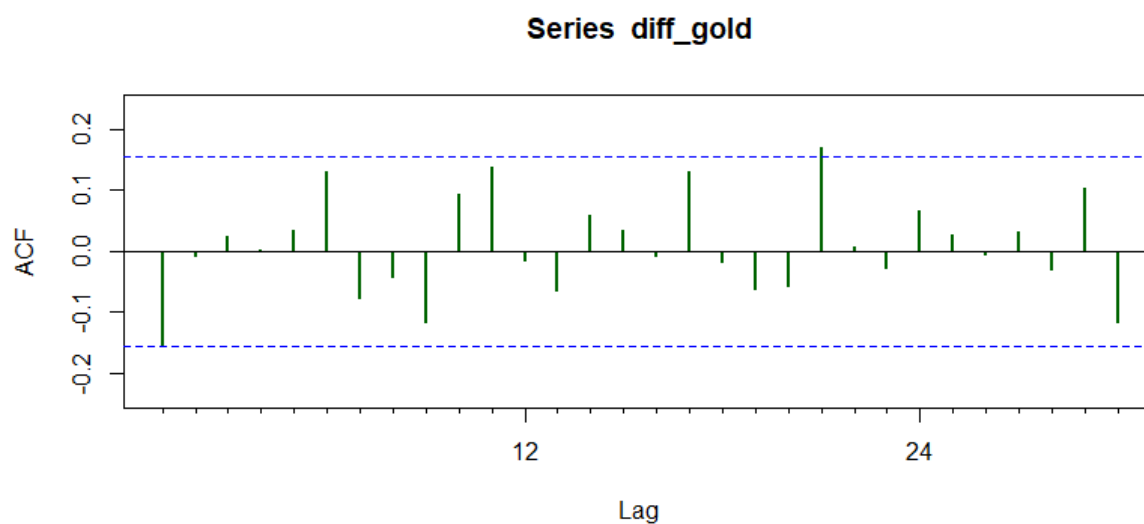
3.5. Badanie autokorelacji dla pierwszych różnic – ACF i PACF

Obserwując wykresy autokorelacji (wykres 7) oraz autokorelacji cząstkowej (wykres 8) można zauważyć, że pierwsze opóźnienie znajduje się na granicy 95% przedziału ufności zarówno na wykresie PACF jak i ACF. Natomiast test Ljung-Boxa wskazuje na istotną korelację tego opóźnienia, wyniki testu zostały przedstawione na wykresie 9 oraz w tabeli 2. Wykres ACF wskazuje również na zależność 21 opóźnienia, wykres PACF 11. Jednak w przypadku obu tych opóźnień test Ljung-Boxa przyjmuje wartości większe od 0,05. Oznacza to, że nie występuje istotna autokorelacja.

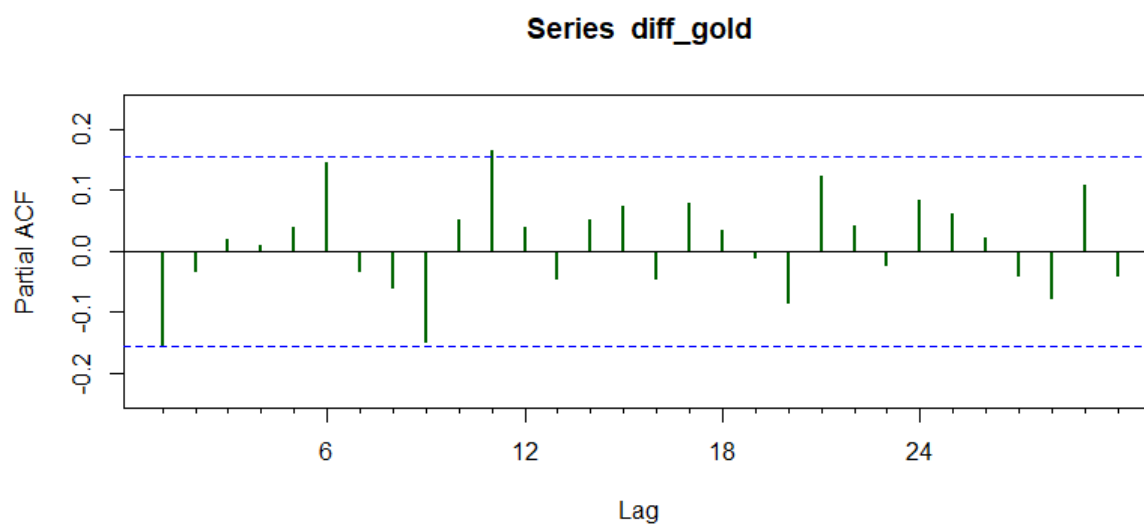
Na podstawie wykresów ACF i PACF zostaną przyporządkowane wstępne parametry modelu ARIMA(p,1,q). Wykres ACF sugeruje proces średniej ruchomej MA(1), natomiast wykres PACF proces autoregresyjny AR(1). Otrzymujemy więc wstępny model – ARIMA(1,1,1).

Wybór właściwego modelu zostanie wykonany poprzez minimalizację kryteriów informacyjnych.

Wykres 7 Autokorelacja pierwszych różnic



Wykres 8 Autokorelacja cząstkowa pierwszych różnic



Wykres 9 Wartości p testu Ljung-Boxa dla danych opóźnień

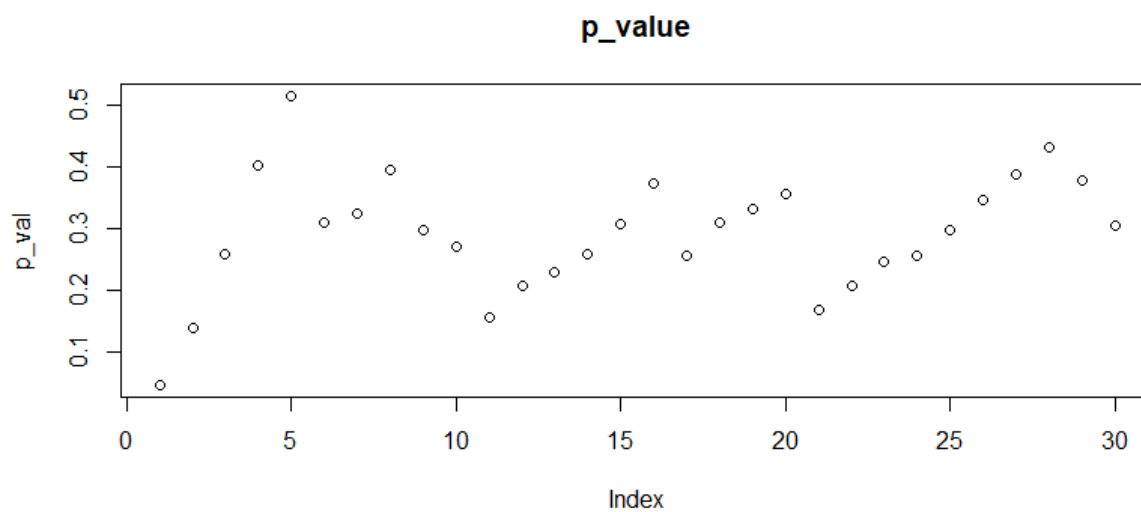


Tabela 2 Wartości p testu Ljung Boxa dla danych opóźnień

	p_val.gold_diff
1	0.04777206
2	0.14039205
3	0.25879161
4	0.40256807
5	0.51584204
6	0.31044362
7	0.32519774
8	0.39532203
9	0.29791291
10	0.27063587
11	0.15688429
12	0.20858803
13	0.23119468
14	0.25877209
15	0.30898754
16	0.37416838
17	0.25785078
18	0.31166268
19	0.33162908
20	0.35760045
21	0.16913237
22	0.20886951
23	0.24720963
24	0.25787242
25	0.29857980
26	0.34822013
27	0.38979160
28	0.43298199
29	0.37798570
30	0.30597206

Rysunek 4 Wynik testu Ljung-Boxa dla opóźnienia 30

```
Box-Ljung test

data: diff_gold
X-squared = 33.388, df = 30, p-value = 0.306
```

3.6. Wybór modelu na podstawie kryteriów informacyjnych

Porównanie wyestymowanych modeli za pomocą kryteriów informacyjnych zostało przedstawione w tabeli 3 oraz tabeli 4. Zarówno kryterium informacyjne Akaikego jak i Bayesowskie kryterium informacyjne Schwarza wskazują na model ARIMA(0,1,1) z dryfem. Postać modelu została zaprezentowana w tabeli 5.

Tabela 3 Porównanie Kryterium informacyjnego Akaikego

	df	AIC
arima111.drift	4	-484.3131
arima011.drift	3	-486.3118
arima110.drift	3	-486.1962
arima111	3	-480.9309
arima011	2	-482.8379
arima110	2	-482.9303
arima112	4	-483.4181
arima211	4	-482.9909
arima212	5	-482.9415

Tabela 4 Porównanie Bayesowskiego kryterium informacyjnego Schwarza

	df	BIC
arima111.drift	4	-472.0124
arima011.drift	3	-477.0863
arima110.drift	3	-476.9707
arima111	3	-471.7054
arima011	2	-476.6876
arima110	2	-476.7799
arima112	4	-471.1174
arima211	4	-470.6902
arima212	5	-467.5656

Tabela 5 Model ARIMA(0,1,1) z dryftem

```

Series: gold_train
ARIMA(0,1,1) with drift

Coefficients:
      ma1    drift
      -0.1603  0.0083
s.e.    0.0776  0.0035

sigma^2 estimated as 0.002733:  log likelihood=246.16
AIC=-486.31  AICc=-486.16  BIC=-477.09

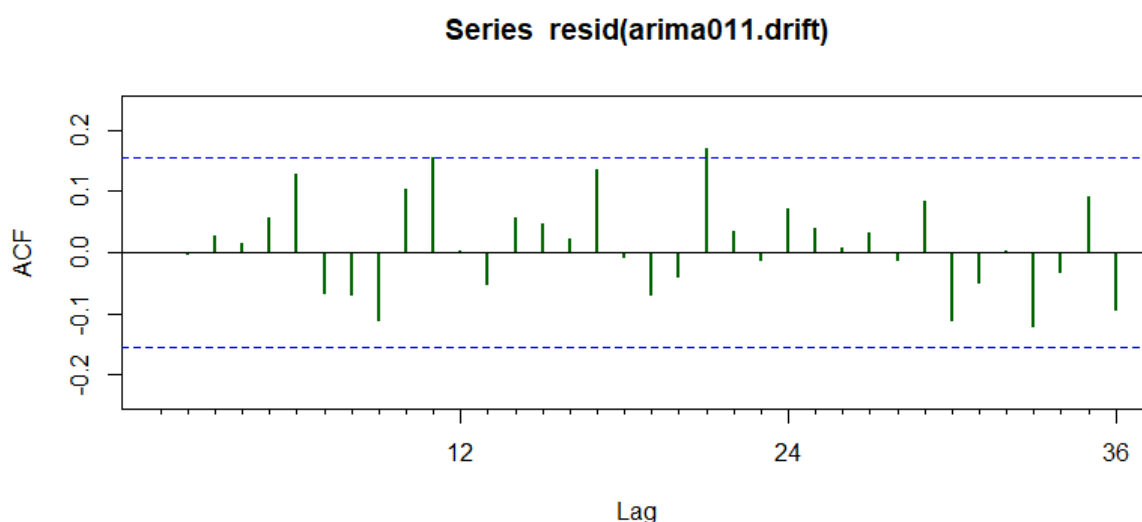
Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set -1.191722e-05  0.05178938  0.04100551  0.002365879  0.6022474  0.2575511
              ACF1
Training set -0.001390931

```

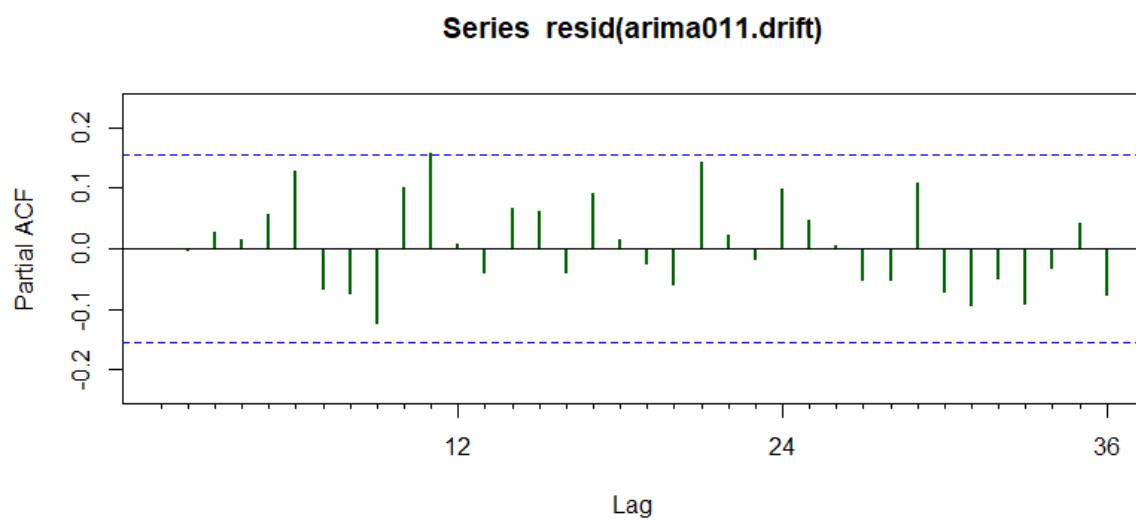
4. Ocena modelu

Aby ocenić czy wyestymowany model jest poprawny zostało sprawdzone czy jego reszty są białym szumem. Wykres ACF (wykres 10) oraz PACF (wykres 11) wskazują na to, że reszty są białym szumem, a wartości p testu Ljung-Boxa przedstawione na wykresie 12 i w tabeli 6 wskazują, że nie ma autokorelacji w resztach. Oznacza to, że proces został wyestymowany poprawnie i możemy dokonać prognozowania za pomocą tego modelu.

Wykres 10 ACF dla reszt modelu



Wykres 11 PACF dla reszt modelu



Wykres 12 Wartości p testu Ljung Boxa dla danych opóźnień

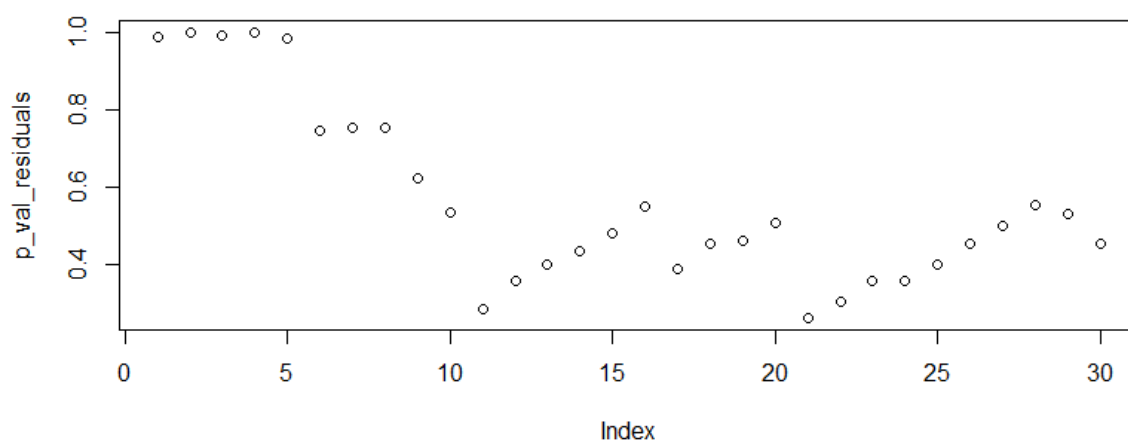


Tabela 6 Wartości p testu Ljung Boxa dla danych opóźnień

	p_val_residuals
1	0.9857875
2	0.9995721
3	0.9898836
4	0.9973109
5	0.9842757
6	0.7460981
7	0.7540266
8	0.7536495
9	0.6231327
10	0.5366551
11	0.2854481
12	0.3601977
13	0.4014969
14	0.4358970
15	0.4816373
16	0.5490545
17	0.3896429
18	0.4556030
19	0.4631990
20	0.5083967
21	0.2638778
22	0.3051701
23	0.3576180
24	0.3605922
25	0.3998928
26	0.4547341
27	0.4996901
28	0.5528423
29	0.5313999
30	0.4556836

Rysunek 5 Wartości p testu Ljung Boxa dla danych opóźnień

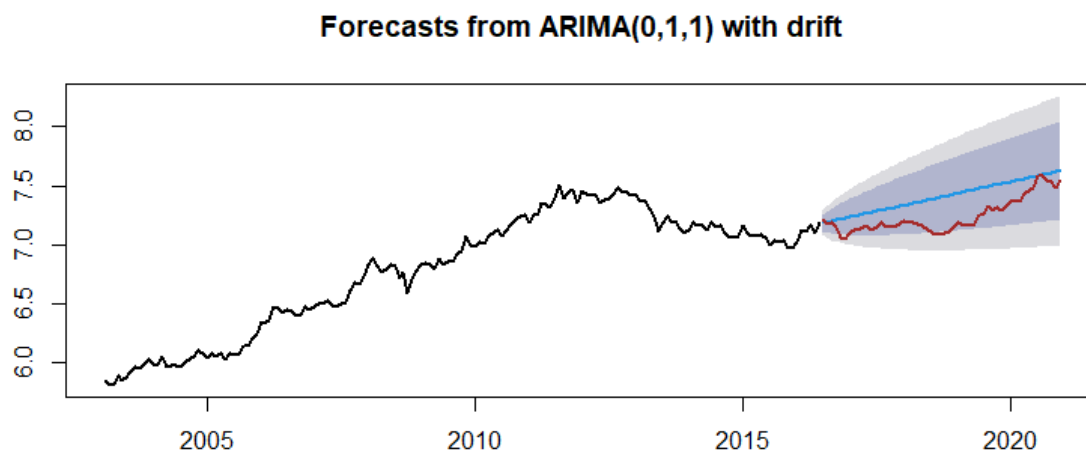
```
Box-Ljung test

data: resid(arima011.drift)
X-squared = 30.195, df = 30, p-value = 0.4557
```

5. Prognoza

Na wykresie 13 została zaprezentowana prognoza cen złota za pomocą modelu ARIMA wraz z rzeczywistymi wartościami (kolor brązowy na wykresie). Prognoza została wykonana dla 54 okresów. Można zauważyć, że prognoza z modelu ARIMA(0,1,1) z dryftem najlepiej sprawdziła się w krótkim okresie około 3 miesięcy, natomiast wartości prognozy długoterminowej w porównaniu z wartościami rzeczywistymi różnią się znacząco.

Wykres 13 Prognoza logarytmicznych cen złota za uncje



6. Ocena prognozy

Na rysunku 6 zostały przedstawione błędy prognozy. Zauważyć można, że najmniejszymi błędami na zbiorze uczącym charakteryzował się wybrany model ARIMA(0,1,1) z dryftem. Jednak model ten wykazał się dosyć wysokimi błędami na zbiorze testowym, co świadczy o raczej słabym dopasowaniu modelu prognostycznego.

Rysunek 6 Błędy prognozy dla zbioru treningowego i testowego

```
> accuracy(forecast_arima011.drift, gold_test)[,kryteria]
              MAE      RMSE      MAPE      MASE
Training set 0.04100551 0.05178938 0.6022474 0.2575511
Test set     0.16644837 0.18624222 2.3108006 1.0454438
> accuracy(forecast_arima111.drift, gold_test)[,kryteria]
              MAE      RMSE      MAPE      MASE
Training set 0.04100791 0.05178916 0.6022767 0.2575662
Test set     0.16640253 0.18619515 2.3101692 1.0451559
> accuracy(forecast_arima110.drift, gold_test)[,kryteria]
              MAE      RMSE      MAPE      MASE
Training set 0.04108645 0.05180828 0.603415 0.2580594
Test set     0.16657907 0.18633612 2.312624 1.0462648
> accuracy(forecast_arima011,gold_test)[,kryteria]
              MAE      RMSE      MAPE      MASE
Training set 0.04161536 0.0526846 0.6118131 0.2613815
Test set     0.10610891 0.1586695 1.4374616 0.6664584
> accuracy(forecast_arima111,gold_test)[,kryteria]
              MAE      RMSE      MAPE      MASE
Training set 0.04163284 0.05266904 0.6120154 0.2614913
Test set     0.10614048 0.15870966 1.4378822 0.6666567
> accuracy(forecast_arima110,gold_test)[,kryteria]
              MAE      RMSE      MAPE      MASE
Training set 0.04163158 0.05266917 0.6119999 0.2614833
Test set     0.10613773 0.15870203 1.4378476 0.6666394
```

7. Wnioski końcowe

Jako model końcowy został wybrany model ARIMA(0,1,1) z dryftem. Charakteryzował się on najniższymi wartościami kryteriów informacyjnych oraz błędów predykcji na zbiorze treningowym. Prognoza na podstawie tego modelu dobrze dopasowała się do danych testowych w krótkim okresie około 3 miesięcy. Natomiast model był niezbyt skuteczny w długoterminowym prognozowaniu wartości zbioru testowego o czym też świadczą dość wysokie błędy prognozy. Było to również widoczne na wykresie, gdzie wartości prognozy znacząco odstawały od wartości rzeczywistych.

Mimo, że modele nie uwzględniające dryftu wykazały się mniejszymi błędami prognozy to wciąż błędy te były stosunkowo wysokie, a ponadto nie mamy pewności czy modele przeszłyby diagnostykę pozwalającą na prognozowanie przy ich użyciu. Dlatego kolejnym krokiem w celu poprawy jakości modelu mogłaby być próba dopasowania modeli TBATS lub sieci neuronowych.