

**AGH**

**AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE**  
**Wydział Zarządzania**

## Projekt EFID - część 2b

Autorzy: *Patrycja Piła, Mateusz Strojek, Julia Szutka, Magdalena Wnuk*  
Kierunek studiów: *Informatyka i Ekonometria*  
Prowadzący: *dr hab. Tomasz Wójtowicz*

Kraków, 2025

## Spis treści

Polecenie zadania .....	3
Wstęp.....	3
Estymacja modeli.....	3
Stacjonarność szeregów czasowych .....	3
Użyte biblioteki do budowy modeli .....	4
Dopasowane modele ARMA .....	4
Walidacja modeli .....	4
Testy normalności i autokorelacji.....	5
Istotność statystyczna parametrów .....	7
Heteroskedastyczność.....	8
Prognoza .....	8
Wnioski .....	11

## Polecenie zadania

Do szeregów czasowych wykorzystanych w części 2A dopasować odpowiedni model ARMA (minimalizujący wybrane kryterium informacyjne). Zbadać poprawność wybranych modeli. Niezależnie od poprawności, na podstawie dopasowanych modeli wyznaczyć prognozę dziesięciu kolejnych obserwacji. Obliczyć odchylenia stand. błędów tych prognoz i wyznaczyć przedziały ufności dla prognozowanych obserwacji. Przedstawić graficznie wartości prognoz, odchylenia standardowe błędów prognoz i szerokość przedziałów ufności. Skomentować i porównać uzyskane wyniki.

## Wstęp

Analiza tygodniowych logarytmicznych stóp zwrotu oraz ich charakterystyk statystycznych pozwala ocenić efektywność inwestycji, zmienność i ryzyko na rynku, a także wykryć ewentualne wzorce, takie jak efekt dnia tygodnia.

Celem analizy jest modelowanie logarytmicznych stóp zwrotu wyodrębnionych dla poniedziałków i śród za pomocą modeli ARMA, analogicznie do podejścia zastosowanego w projekcie 2A. Sprawdzono czy modele nie są wadliwe, badając m.in. reszty modeli. Na podstawie dopasowanych modeli przeprowadzono prognozę 10 kolejnych obserwacji oraz wyznaczono odpowiadające im przedziały ufności. Otrzymane prognozy porównano z rzeczywistymi wartościami w celu oceny jakości dopasowania modeli.

## Estymacja modeli

### Stacjonarność szeregów czasowych

W celu weryfikacji stacjonarności szeregów czasowych, przeprowadzono rozszerzony test Dickeya-Fullera (test ADF) dla tygodniowych stóp zwrotu logarytmicznych, obliczonych osobno dla poniedziałków oraz śród. Hipotezy testu są następujące:

- $H_0$ : szereg posiada pierwiastek jednostkowy, czyli jest niestacjonarny,
- $H_1$ : szereg nie posiada pierwiastka jednostkowego, czyli jest stacjonarny

Wyniki testu ADF kształtują się następująco:

Tabela 1. Wyniki testu ADF

Dzień tygodnia	Statystyka ADF	P-wartość	Wniosek
Poniedziałek	-14,7445	2,5407e-27	Szereg jest stacjonarny
Środa	-22,3986	0,0	Szereg jest stacjonarny

Poniedziałkowe, jak i środowe logarytmiczne stopy zwrotu okazały się stacjonarnymi szeregami czasowymi na podstawie testu ADF. Oznacza to, że spełniony jest podstawowy warunek do zastosowania modelu ARMA.

## Użyte biblioteki do budowy modeli

Analogicznie jak w projekcie 2A, w badaniu wykorzystano funkcję `arma_order_select_ic` w celu automatycznego doboru optymalnych parametrów dla modelu ARMA. Funkcja ta przeprowadza systematyczne przeszukiwanie zadanego zakresu wartości parametrów, oceniając każdą ich kombinację przy użyciu konkretnego kryterium informacyjnego. Skorzystano z kryterium Akaike (AIC), które na tle pozostałych kryteriów każe złożoność modeli łagodniej.

## Dopasowane modele ARMA

### Poniedziałek:

- Model: ARMA (2, 3)
- AIC: -2390,874
- BIC: -2361,110
- HQIC: -2379,213

Dla poniedziałkowych stóp zwrotu najlepiej dopasowanym modelem okazał się ARMA(2, 3), czyli model autoregresyjny z dwoma opóźnieniami oraz częścią średniej ruchomej rzędu trzeciego.

### Środa:

- Model: ARMA (0,0)
- AIC: -2468,769
- BIC: -2460,265
- HQIC: -2465,438

Dla środowych stóp zwrotu najlepszym modelem okazał się ARMA(0,0). Wskazuje to na to, że seria zachowuje się jak biały szum.

## Walidacja modeli

Po oszacowaniu parametrów modelu, zweryfikowano jego poprawność. Wykorzystano następujące kroki:

- Sprawdzono rozkład reszt (MNW wymaga, aby losowe składniki miały rozkład normalny),
- Wykonano testy autokorelacji,
- Zweryfikowano istotność statystyczną parametrów dla największych opóźnień,
- Przeprowadzono testy na homoskedastyczność reszt.

Rezultaty przeprowadzonej weryfikacji mogą wskazywać na potrzebę modyfikacji modelu lub poszukiwania nowego:

- Jeżeli parametry przy maksymalnych opóźnieniach okażą się nieistotne, zaleca się wybór prostszego modelu.
- W przypadku stwierdzenia autokorelacji, należy podnieść rząd opóźnień w modelu.

## Testy normalności i autokorelacji

Aby zweryfikować, czy rozkład reszt można uznać za normalny, zastosowano dwa testy: Shapiro-Wilka oraz Jarque-Bera. We wszystkich przypadkach hipotezy były następujące:

- $H_0$ : rozkład badanej cechy jest rozkładem normalnym,
- $H_1$ : rozkład badanej cechy nie jest rozkładem normalnym.

Test Shapiro-Wilka polega na porównaniu uporządkowanych wartości próby z wartościami oczekiwanymi przy założeniu rozkładu normalnego. Jest on szczególnie skuteczny w wykrywaniu odchyień od normalności w przypadku niewielkich prób. Z kolei test Jarque-Bera ocenia, czy skośność i kurtoza badanej zmiennej odpowiadają wartościom typowym dla rozkładu normalnego.

Tabela 2. Wyniki testów normalności reszt modeli

Dzień	Statystyka Shapiro-Wilka	Wartość p Shapiro-Wilka	Statystyka Jarque-Bera	Wartość p Jarque-Bera
Poniedziałek	0,8976	3,9013e-18	1202,8012	6,5319e-262
Środa	0,9009	7,7125e-18	1127,4384	1,5130e-245

Wyniki wszystkich przeprowadzonych testów wskazują zgodnie, że reszty nie podążają za rozkładem normalnym. Wartości p dla wszystkich testów wyniosły wyraźnie poniżej 0,05.

W celu weryfikacji, czy w resztach modelu pozostaje autokorelacja, przeprowadzono test Ljung-Boxa. Test ten sprawdza, czy zestaw współczynników autokorelacji dla kolejnych opóźnień różni się istotnie od zera, czyli czy w danych występuje autokorelacja. Hipotezy mają postać:

- $H_0$ : brak autokorelacji do rzędu  $m$ ,
- $H_1$ : istotna autokorelacja do rzędu  $m$ .

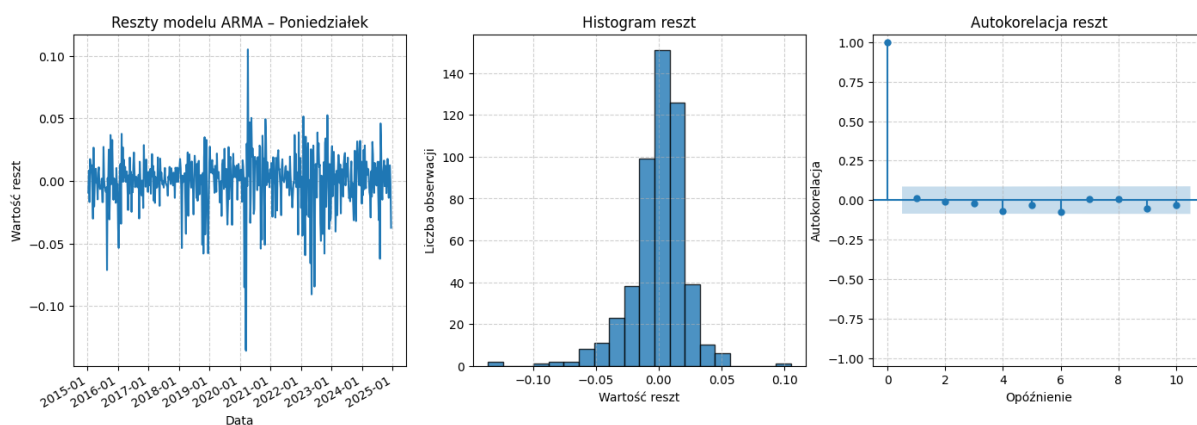
Tabela 3. Wyniki testów Ljung-Boxa dla reszt modeli ARMA

Opóźnienie	Poniedziałek		Środa	
	Statystyka	Wartość p	Statystyka	Wartość p
1	0,006768	0,934436	0,068233	0,793927
2	0,044456	0,978017	0,095206	0,953512
3	0,046492	0,997371	0,279863	0,963770
4	0,122631	0,998195	2,923582	0,570694
5	0,194719	0,999170	3,427204	0,634430
6	0,269938	0,999630	6,455143	0,374179
7	0,458867	0,999584	6,485410	0,484345
8	1,596824	0,990981	6,502001	0,591186
9	1,732028	0,995032	8,137170	0,520384
10	1,732029	0,998012	8,674078	0,563289

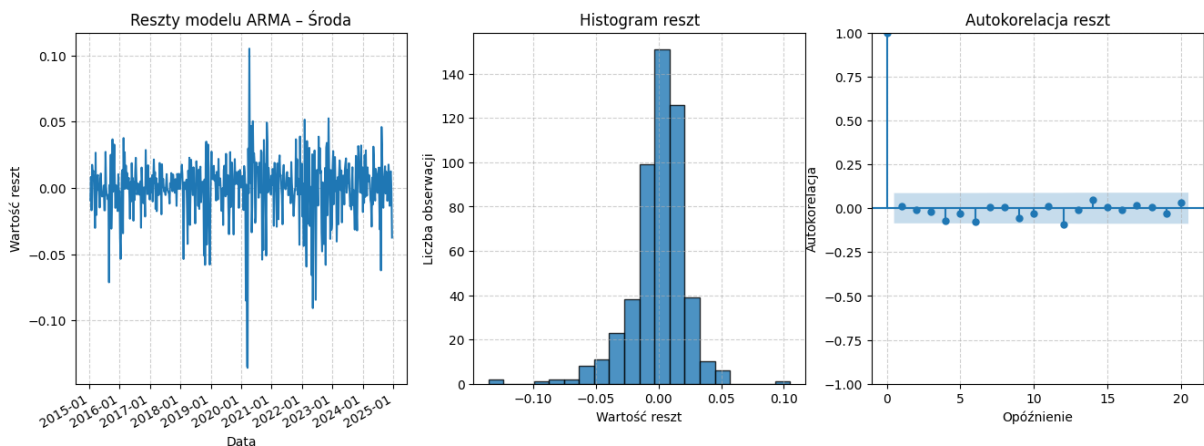
Na podstawie wyników testu Ljung-Boxa można stwierdzić, że reszty modeli nie wykazują autokorelacji. Wszystkie testy Ljung-Boxa, przeprowadzone dla kolejnych rzędów opóźnień, zwróciły wartości p większe niż 0,05, co potwierdza brak istotnych zależności pomiędzy resztami.

Wnioski otrzymane przy pomocy testów normalności oraz testu Ljung-Boxa potwierdzają histogramy reszt oraz wykresy ACF.

Rysunek 1. Histogram reszt i ACF dla poniedziałków



Rysunek 2. Histogram reszt i ACF dla śród



Dla obu histogramów reszt modeli zauważalne jest to, że występuje skośność lewostronna; ekstremalne wartości ujemne reszt pojawiają się częściej niż dodatnie. Dość ciekawym zjawiskiem jest bardzo duże odchylenie reszt dla obu modeli ARMA na początku 2020 roku. Bardzo możliwe, że jest to spowodowane wybuchem pandemii COVID-19. Ponadto, wykresy ACF ewidentnie wskazują na brak autokorelacji; autokorelacje dla wszystkich opóźnień znajdują się w przedziale ufności lub na jego granicy.

### Istotność statystyczna parametrów

Zbudowano dwa modele ARMA, przy czym model dla śród jest zwykłym białym szumem. Dlatego analizę istotności statystycznej najpóźniejszych opóźnień przeprowadzono jedynie dla modelu ARMA(2,3), który został zbudowany dla poniedziałkowych stóp zwrotu.

Tabela 4. Istotność parametrów modelu ARMA dla poniedziałków

Zmienna	Wartość współczynnika	Błąd standardowy	Statystyka z	P wartość	Dolna granica przedziału ufn.	Górna granica przedziału ufn.
Stała	0,0021	0,001	2,444	0,015	0,000	0,004
ar.L1	-0,0749	0,172	-0,436	0,663	-0,412	0,262
ar.L2	0,5914	0,190	3,108	0,002	0,218	0,964
ma.L1	0,0617	0,178	0,347	0,729	-0,287	0,410
ma.L2	-0,6315	0,200	-3,158	0,002	-1,023	-0,240
ma.L3	-0,1179	0,048	-2,475	0,013	-0,211	-0,025

Wartość p dla współczynnika przy drugim opóźnieniu części AR oraz dla współczynnika przy trzecim opóźnieniu MA okazała się niższa niż 0,05. Oznacza to, że obie zmienne są statystycznie istotne i mają istotny wpływ na model. Zatem kryterium istotności ostatniego opóźnienia zostało spełnione. Jednakże, ze względu na stwierdzony brak normalności rozkładu reszt, oszacowane błędy standardowe oraz wartości p mogą być nieprawidłowe (co potencjalnie prowadzi do przekłamania w ocenie istotności).

## Heteroskedastyczność

W celu sprawdzenia występowania efektu heteroskedastyczności autoregresyjnej (ARCH) w resztach modeli przeprowadzono test ARCH-LM (Engle'a), wykorzystując funkcję `het_arch` z pakietu `statsmodels`. Hipotezy są następujące:

- $H_0$ : reszty są homoskedastyczne,
- $H_1$ : w resztach występuje efekt ARCH, czyli autokorelacja w wariancji reszt jest obecna.

Tabela 5. Wyniki testu Engle'a

Dzień tygodnia	Statystyka	P-wartość	Wniosek
Poniedziałek	120,4629	2.5043e-24	Występuje efekt ARCH
Środa	81,5232	4.0279e-16	Występuje efekt ARCH

Uzyskane wartości p są znacznie mniejsze od przyjętego poziomu istotności, co prowadzi do odrzucenia hipotezy zerowej. Oznacza to, że w resztach modeli występuje istotny efekt heteroskedastyczności typu ARCH.

## Prognoza

Wykresy na przedstawiają ponad 500 historycznych obserwacji oraz prognozę na przyszłe poniedziałki i środy z 95-procentowym przedziałem ufności. Tabela poniżej natomiast porównuje prognozy punktowe z wartościami rzeczywistymi dla 10 ostatnich okresów, weryfikując jednocześnie, czy wartości rzeczywiste mieszczą się w oszacowanych przedziałach ufności.

Rysunek 3. Prognoza modelu ARMA dla poniedziałku

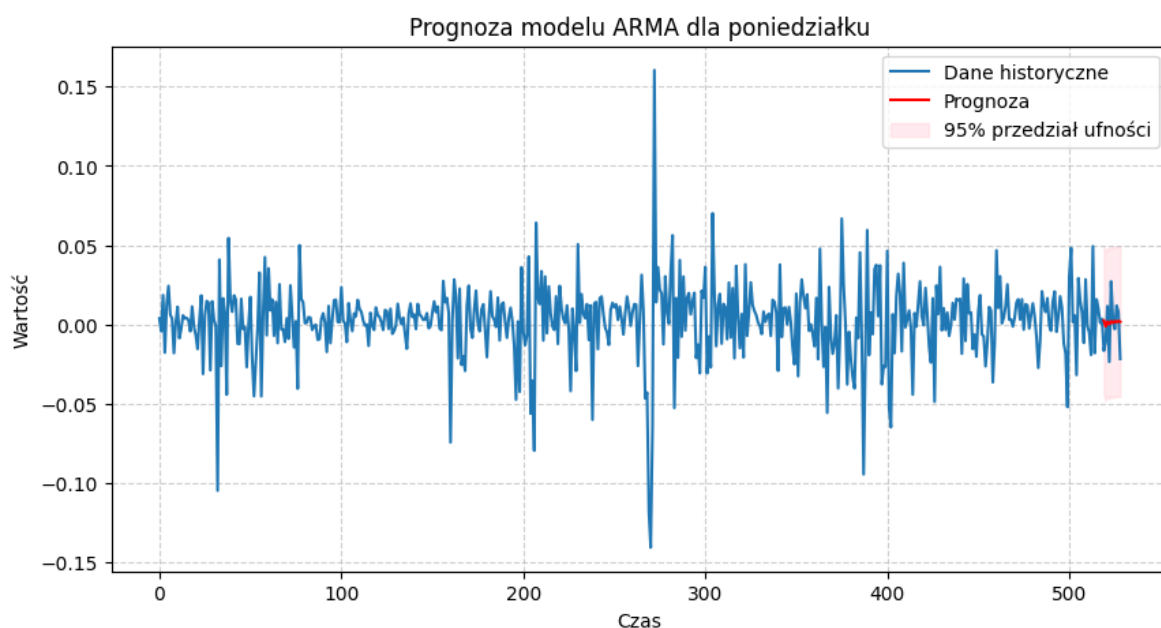




Tabela 6. Wartości zaprognozowanych poniedziałkowych stóp zwrotu oraz błędy predykcji

Obserwacja	Prognoza	Rzeczywista wartość	Błąd	Błąd standardowy	Dolny przedział ufn.	Górny przedział ufn.
1	0,002429	-0,016602	-0,019031	0,023853	-0,044322	0,049180
2	-0,001238	-0,011301	-0,010063	0,023855	-0,047992	0,045517
3	0,001692	0,011520	0,009828	0,023873	-0,045099	0,048482
4	0,000158	-0,023564	-0,023722	0,024052	-0,046983	0,047299
5	0,002005	0,027119	0,025114	0,024054	-0,045141	0,049151
6	0,000960	0,002601	0,001642	0,024115	-0,046305	0,048224
7	0,002130	-0,002950	-0,005080	0,024115	-0,045134	0,049395
8	0,001424	0,011918	0,010493	0,024136	-0,045881	0,048730
9	0,002170	0,007912	0,005742	0,024136	-0,045136	0,049475
10	0,001696	-0,021720	-0,023417	0,024143	-0,045623	0,049016

We wszystkich 10 testowanych przypadkach rzeczywista stopa zwrotu zmieściła się w 95-procentowym przedziale ufności. Model ma trudności z trafieniem w kierunek zmian stóp zwrotu (50% skuteczność). Prognoza zakłada, że stopy zwrotu nie będą się zmieniać w sposób gwałtowny. Wszystkie predykcje oscylują w okolicy zera.

Rysunek 4. Prognoza modelu ARMA dla środy

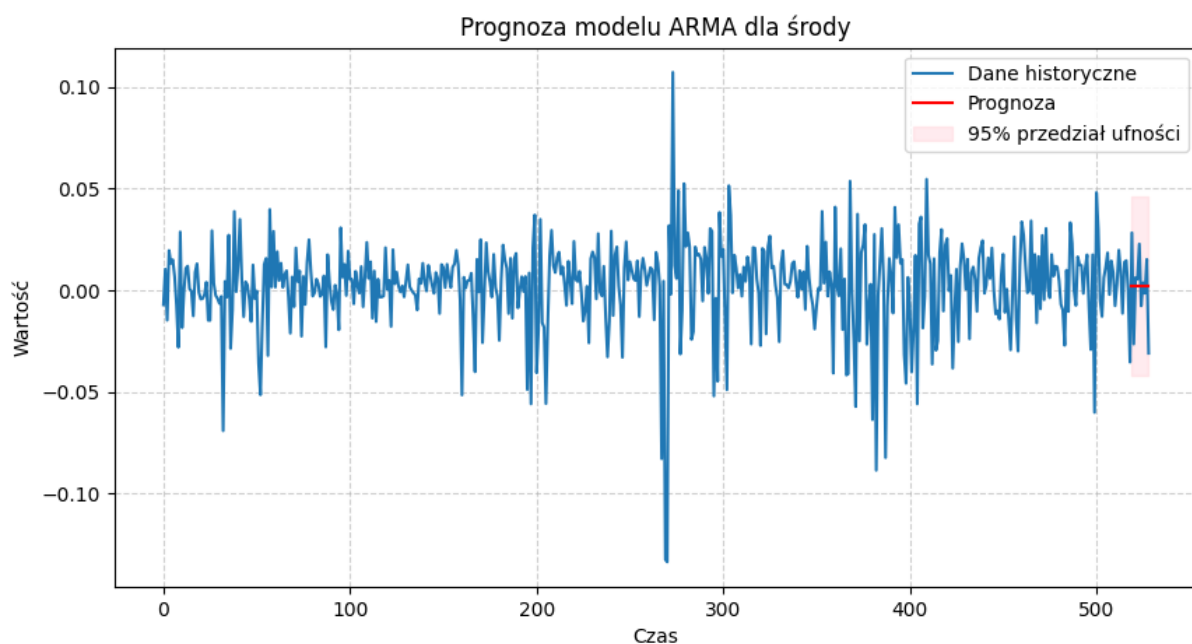


Tabela 7. Wartości zaprognozowanych śródownych stóp zwrotu oraz błędy predykcji

Obserwacja	Prognoza	Rzeczywista wartość	Błąd	Błąd standardowy	Dolny przedział ufn.	Górny przedział ufn.
1	0,002046	0,028188	0,026143	0,022340	-0,041739	0,045830
2	0,002046	-0,026577	-0,028622	0,022340	-0,041739	0,045830
3	0,002046	0,006207	0,004161	0,022340	-0,041739	0,045830
4	0,002046	0,005335	0,003290	0,022340	-0,041739	0,045830
5	0,002046	0,022676	0,020630	0,022340	-0,041739	0,045830
6	0,002046	-0,007762	-0,009808	0,022340	-0,041739	0,045830
7	0,002046	0,003664	0,001619	0,022340	-0,041739	0,045830
8	0,002046	-0,001570	-0,003616	0,022340	-0,041739	0,045830
9	0,002046	0,015117	0,013071	0,022340	-0,041739	0,045830
10	0,002046	-0,031091	-0,033137	0,022340	-0,041739	0,045830

Tak samo jak dla poniedziałkowych predykcji, we wszystkich 10 testowanych przypadkach rzeczywista stopa zwrotu zmieściła się w 95-procentowym przedziale ufności oraz model miał trudności z trafieniem w kierunek zmian stóp zwrotu (40% skuteczność). Prognoza dla modelu ARMA(0, 0) jest taka sama dla wszystkich horyzontów czasowych, ponieważ model ten traktuje szereg jako biały szum i przewiduje jedynie jego stałą średnią.

Dla pełnej oceny jakości predykcyjnej modelu, poza standardową analizą prognoz i ich błędów, obliczono również zagregowane miary błędu. W tym celu oszacowano MAE, RMSE oraz MAPE. Wyniki tych metryk kształtują się następująco:

Tabela 8. Wyniki miar dokładności modeli

Dzień	MAE	RMSE	MAPE	Średnia szerokość przedziału ufn.
Poniedziałek	0,013413	0,015727	98,60%	0,094208
Środa	0,014410	0,018268	101,40%	0,087570

Modele osiągają porównywalny poziom dokładności w sensie średniego błędu (MAE). Wyższy RMSE w środę wskazuje na to, że duże błędy prognozy w tym dniu są częstsze lub bardziej ekstremalne niż w poniedziałek. Metryka MAPE jest podobna dla obu predykcji tygodniowych stóp zwrotu, choć znowu nieco lepszy okazał się być model dla poniedziałków. Pomimo delikatnych różnic w MAE, RMSE i MAPE, średnia szerokość przedziału ufności była mniejsza w środę, co świadczy o niższej niepewności w predykcji.

# Wnioski

Na podstawie przeprowadzonej estymacji i walidacji modeli ARMA dla tygodniowych logarytmicznych stóp zwrotu, można sformułować następujące wnioski:

- Oba szeregi czasowe są stacjonarne (potwierdzone testem ADF). Optymalny model dla poniedziałków to ARMA(2,3), a dla śród ARMA(0,0).
- Modele pomyślnie przeszły test autokorelacji reszt (test Ljung-Boxa) oraz test istotności statystycznej kluczowych opóźnień (dla ARMA(2,3)).
- Głównymi niedoskonałościami modeli są brak normalności reszt (potwierdzony testami Shapiro-Wilka i Jarque-Bera) oraz obecność heteroskedastyczności autoregresyjnej (efekt ARCH), zidentyfikowanej za pomocą testu Engle'a.
- Modele są poprawnie skalibrowane pod względem niepewności; wszystkie predykcje zmieściły się w 95-procentowych przedziałach ufności dla obu dni tygodnia. Prognozy punktowe oscylują wokół zera, zakładając brak gwałtownych zmian. Miary dokładności modeli zwróciły podobne rezultaty, bardzo nieznacznie wskazując, że model ARMA(2, 3) dla poniedziałków lepiej przewiduje przyszłe stopy zwrotu.