Trường Đại học Kinh tế Quốc dân

Khoa Toán kinh tế

…ꕥ📖ꕥ…

BÀI TẬP LỚN

PHÂN TÍCH CHUỖI THỜI GIAN TRONG TÀI CHÍNH

Sinh viên thực hiện: Nhóm 8

1. Trịnh Thu Huyền
2. Trương Ngọc Thùy Trang
3. Phạm Diệu Linh

Lớp học phần: TOKT1147(223)_01

Giảng viên hướng dẫn: Th.S Bùi Dương Hải

Hà Nội – 2024

THÀNH VIÊN NHÓM

STT	Họ và tên	Mã sinh viên	Đóng góp
1	Trịnh Thu Huyền	11212771	33.33%
2	Trương Ngọc Thùy Trang	11217004	33.33%
3	Phạm Diệu Linh	11213377	33.33%

MỤC LỤC

PHỤ LỤC 3
A. GIỚI THIỆU CHUNG 4
B. BÀI CÁ NHÂN 5
I. Chuỗi PVS – Trương Ngọc Thùy Trang 5
1. Giới thiệu 5
2. Phân tích với số liệu theo quý – Chuỗi doanh thu thuần 5
3. Phân tích với số liệu theo ngày – Chuỗi giá đóng cửa và chuỗi log_return 8
4. Sử dụng ARCH/GARCH để dự báo rủi ro cho 10 ngày đầu năm 2024 12
II. Chuỗi PVG – Trịnh Thu Huyền 14
1. Giới thiệu 14
2. Phân tích với số liệu theo quý – Chuỗi doanh thu thuần 14
3. Phân tích với số liệu theo ngày – Chuỗi giá đóng cửa và chuỗi log_return 16
4. Sử dụng ARCH/GARCH để dự báo rủi ro cho 10 ngày đầu năm 2024 22
III. Chuỗi PVC – Phạm Diệu Linh 24
1. Giới thiệu 24
2. Phân tích với số liệu theo quý – Chuỗi doanh thu thuần 24
3. Phân tích với số liệu theo ngày – Chuỗi giá đóng cửa và chuỗi log_return 27
4. Mô hình ARCH/GARCH dự báo rủi ro 33
C. PHÂN TÍCH TỔNG HỢP 36
I. Tổng quan 36
1. Sự biến động chung 36
2. Sự tương quan 36
II. Phân tích đồng tích hợp 36
1. Lý thuyết 36
2. Kiểm định đồng tích hợp giữa các chuỗi giá 37
III. Phân tích VAR 38
1. Xây dựng VAR giữa các chuỗi log-return, kiểm định đánh giá 38
2. Dự báo các chuỗi log-return từ mô hình VAR 39
3.Hàm phản ứng và phân rã phương sai 40
IV. Kết luận 42

PHỤ LỤC

Bảng	Tên bảng	Trang
1	Các mô hình dự báo cho chuỗi doanh thu PVS	6
2	Kiểm định tính dừng cho chuỗi giá đóng cửa PVS	9
3	Kiểm định tính dừng cho chuỗi log-return PVS	9
4	Dự báo giá cổ phiếu PVS trong 10 phiên tiếp theo	11
5	Dự báo giá trị volatility cho 10 ngày đầu năm 2024	13
6	Các mô hình dự báo cho chuỗi doanh thu PVG	14
7	Kiểm định tính dừng cho chuỗi giá đóng cửa PVG	17
8	Kiểm định tính dừng cho chuỗi sai phân	18
9	Dự báo giá cổ phiếu PVG trong 10 phiên tiếp theo	21
10	Dự báo giá trị volatility cho 10 ngày đầu năm 2024	23
11	Các mô hình dự báo cho chuỗi doanh thu PVC	25
12	Kiểm định tính dừng chuỗi giá PVC	27
13	Kiểm định tính dừng chuỗi sai phân PVC	28
14	Kiểm định tính dừng chuỗi log_return	29
15	Kiểm định tính nhiễu trắng của phần dư	31
16	Dự báo chuỗi log – return trong 10 ngày đầu tiên tại năm 2024 với 2 mô hình	32
17	Chuỗi giá dự báo PVC trong 10 ngày đầu tiên của năm 2024	32
18	Dự báo sự biến động (volatility) của chuỗi log – return	35
19	So sánh kết quả dự báo	39

A. GIỚI THIỆU CHUNG

Theo số liệu thống kê của BP, Việt Nam là quốc gia đứng thứ 28 trên tổng số 52 nước trên thế giới có tài nguyên dầu khí. Việc gia nhập ngành dầu khí Việt Nam không hề dễ dàng do những rào cản khắt khe, đòi hỏi vốn đầu tư lớn và được sự chấp thuận của Nhà nước. Trong đó, rào cản về chính sách là nguyên nhân chính hạn chế các doanh nghiệp gia nhập ngành dầu khí Việt Nam. Nhiều doanh nghiệp gần như độc quyền trong lĩnh vực hoạt động của mình, điều này mang lại lợi thế lớn.

Hoạt động kinh doanh của ngành dầu khí liên quan mật thiết đến dầu thô và các sản phẩm chế biến từ dầu thô, do đó, biến động của giá dầu thô thế giới sẽ tác động trực tiếp đến hoạt động kinh doanh của các doanh nghiệp hoạt động trong ngành. Điều này có nghĩa là giá dầu thô tăng sẽ có tác động tích cực và ngược lại.

Các công ty trong ngành dầu khí Việt Nam tham gia vào tất cả các khâu từ đầu vào (khai thác, thăm dò) đến đầu ra (chế biến, phân phối). Dầu thô và khí ẩm là nguyên liệu thô được khai thác ngoài khơi, chế biến trong đất liền và phân phối cho khách hàng.

Các mô hình hoạt động của các doanh nghiệp trong ngành dầu khí có sự liên kết chặt chẽ, tạo thành chuỗi giá trị hoàn chỉnh. Sau khi dầu thô được chế biến để tạo thành xăng, dầu thành phẩm sẽ được đưa đi phân phối. Doanh nghiệp vừa chuyên môn hóa hoạt động kinh doanh của mình, vừa hỗ trợ các doanh nghiệp khác trong ngành, góp phần vào sự phát triển bền vững của ngành dầu khí Việt Nam. Khâu khai thác là khâu đầu tiên trong chuỗi giá trị của ngành dầu khí và có vai trò quan trọng nhất.

B. BÀI CÁ NHÂN

I. Chuỗi PVS – Trương Ngọc Thùy Trang

1. Giới thiệu

Mã PVS trên sàn chứng khoán HNX, đại diện cho Tổng Công ty cổ phần Dịch vụ Kỹ thuật Dầu khí Việt Nam. Được thành lập từ năm 2000, PVS đã có một quá trình phát triển ấn tượng và là một trong những đơn vị hàng đầu trong lĩnh vực cung cấp dịch vụ kỹ thuật và thiết bị cho ngành dầu khí ở Việt Nam. Mã cổ phiếu này thường thu hút sự quan tâm của các nhà đầu tư và nhà phân tích thị trường do vai trò quan trọng của ngành dầu khí trong nền kinh tế và sự phát triển của PVS trong lĩnh vực này.

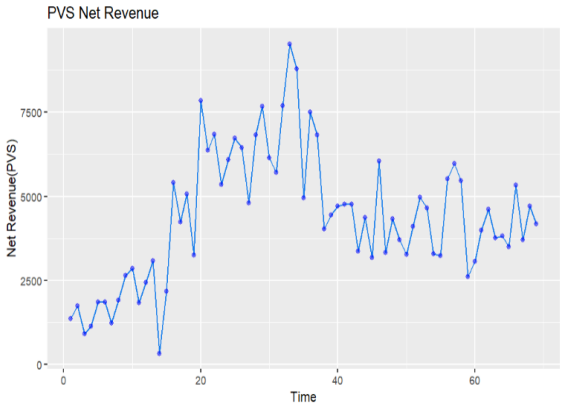
2. Phân tích với số liệu theo quý – Chuỗi doanh thu thuần

Doanh thu thuần là chuỗi dữ liệu theo thời gian phụ thuộc vào nhiều yếu tố như giá cả, cầu hàng hoá, lạm phát,… Nhưng trong phần này, chỉ sử dụng phương pháp phân tích doanh thu thuần qua phân tích xu hướng của bản thân nó theo thời gian và tác động của tính mùa vụ đến giá trị của doanh thu.

Số liệu: chuỗi doanh thu thuần theo quý của Tổng Công ty cổ phần Dịch vụ Kỹ thuật Dầu khí Việt Nam từ quý 3 năm 2006 đến hết quý 4 năm 2023 với mục tiêu dự báo cho 4 quý năm 2024.

Đơn vị: Tỷ VNĐ

Nhìn vào đồ thị, giai đoạn từ quý 3 năm 2006 đến quý 4 năm 2009 doanh thu của Tổng công ty luôn chỉ giữ dưới mức 3100 tỷ đồng, trong đó có quý 4 năm 2009 tụt xuống chỉ còn hơn 300 tỷ đồng. Thời điểm này diễn ra cuộc khủng hoảng tài chính toàn cầu 2009, ngành dầu khí là một trong những ngày cũng chịu ảnh hưởng vậy nên có thể lý giải vì sao doanh thu trong quý 4 năm 2009 của công ty bị sụt giảm. Sau giai đoạn này, doanh thu có đà tăng trưởng đạt đỉnh vào quý 3 năm 2014 với hơn 9500 tỷ đồng, sau đó biến động mạnh trong năm 2015 do giá dầu thế giới không ổn định chịu ảnh hưởng từ các sự kiện chính trị thế giới. Bắt đầu từ năm 2016 đến quý 3 năm 2023 doanh thu ổn định dần dao động trong khoảng 2500 đến 6100 tỷ đồng.



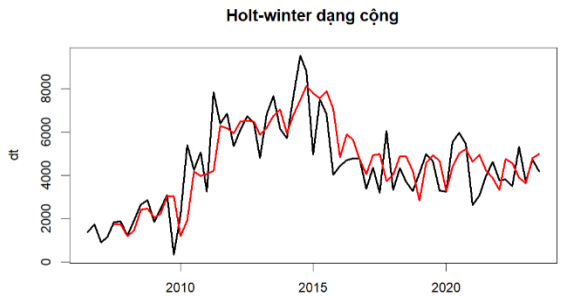
Doanh thu PVS giai đoạn quý 3/2006 – quý 4/2023

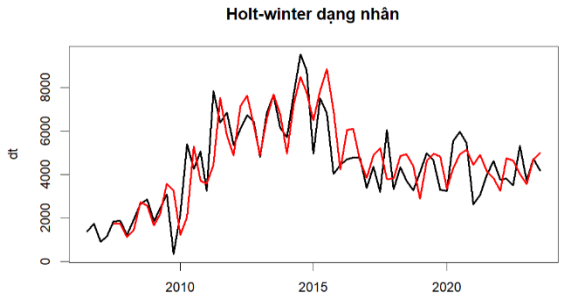
Bảng 1: Các mô hình dự báo cho chuỗi doanh thu PVS

STT	Mô hình	R_sq	MAPE	RMSE	Dự báo
1.	MH linear-linear: $\widehat{dt} = 3557.28 + 23.04t$	0.058	0.602	1882.567	2024Q1 = 5193.12 2024Q2 = 5216.16 2024Q3 = 5239.2 2024Q4 = 5262.24
2.	MH linear – log: $\widehat{dt} = 1154.3 + 978.4ln(t)$	0.208	0.499	1726.18	2024Q1 = 5324.906 2024Q2 = 5338.59 2024Q3 = 5352.086 2024Q4 =5365.397
3.	MH log-linear: $\widehat{ln(dt)} = 7.857 + 0.011t$	0.154	0.495	2028.29	2024Q1 = 5642.034 2024Q2 = 5704.439 2024Q3 = 5767.535 2024Q4 = 5831.328
4.	MH log-log: $\widehat{ln(dt)} = 7.037 + 0.370ln(t)$	0.337	0.427	1858.013	2024Q1 = 5509.293 2024Q2 = 5537.877 2024Q3 = 5566.212 2024Q4 =5594.303
5.	Hồi quy với biến giả mùa: $\widehat{dt} = 3459.4 + 1384.9s_2 + 1221s_3 + 1033.1s_4$	0.076	0.644	1864.462	2024Q1 = 3459.354 2024Q2 = 4844.287 2024Q3 = 4680.403 2024Q4 = 4492.498
6.	Xu thế thời gian tuyến tính và mùa vụ dạng cộng: $\widehat{dt} = 2664.68 + 22.71t + 1362.23s_2 + 1221.05s_3 + 1010.44s_4$	0.132	0.581	1807.176	2024Q1 = 4276.738 2024Q2 = 5661.67 2024Q3 = 5543.197 2024Q4 = 5355.291
7.	Xu thế thời gian tuyến tính và mùa vụ dạng nhân:	0.107	0.589	1833.009	2024Q1 = 3845.654 2024Q2 = 5751.576 2024Q3 = 5660.291 2024Q4 = 5549.096

	$\widehat{dt} = 3574.217 + 3.823t + 26.418t * s_2 + 24.753t * s_3 + 22.864t * s_4$				
8.	Xu thế thời gian không tuyến tính và mùa vụ dạng cộng: $\widehat{ln(dt)} = 7.644 + 0.012t + 0.32s_2 + 0.29s_3 + 0.199s_4$	0.201	0.481	1967.891	2024Q1 = 4661.322 2024Q2 = 6495.336 2024Q3 = 6415.186 2024Q4 = 5764.243
9.	Xu thế thời gian không tuyến tính và mùa vụ dạng nhân: $\widehat{ln(dt)} = 7.861 + 0.007t + 0.006t * s_2 + 0.005t * s_3 + 0.004 * s_4$	0.178	0.486	1997.795	2024Q1 = 4332.25 2024Q2 = 6610.237 2024Q3 = 6420.844 2024Q4 = 6053.156
10.	Holt-Winter dạng cộng: $\widehat{dt^F_{T+h}} = 4153.824 + 85.147 * h + seas$		0.314	1219.323	2024Q1 = 3347.003 2024Q2 = 4446.852 2024Q3 = 4621.542 2024Q4 = 4622.137
11.	Holt-winter dạng nhân: $\widehat{dt^F_{T+h}} = (4438.052 + 85.147 * h) * seas$		0.328	1199.647	2024Q1 = 3338.211 2024Q2 = 4305.617 2024Q3 = 4467.286 2024Q4 = 4666.302

Nhìn chung, các mô hình đều có hệ số xác định R khá là thấp, nghĩa là mô hình giải thích được ít sự thay đổi của doanh thu. Mô hình được đề xuất là hai mô hình Holt-winter dạng cộng và dạng nhân. Đây là hai mô hình có chỉ số RMSE và MAPE thấp nhất.

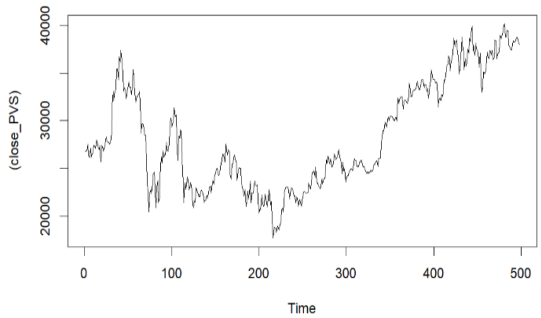




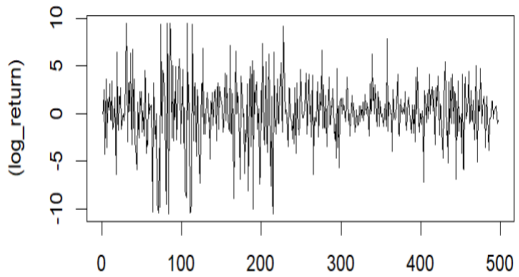
So sánh giá trị mô hình Holt-winter với thực tế

3. Phân tích với số liệu theo ngày – Chuỗi giá đóng cửa và chuỗi log_return

Số liệu: Giá đóng cửa của mã chứng khoán PVS từ năm 2022 – 2023 (số liệu tính theo ngày, bao gồm 498 quan sát, đơn vị: VNĐ)



Đồ thị chuỗi giá PVS



Đồ thị chuỗi log_return

Nhận xét: Chuỗi giá PVS có xu thế , Chuỗi log_return có thể dừng. Giá đóng cửa cổ phiếu PVS luôn giữ ở mức trên 16000 đồng một cổ phiếu, vào cuối năm 2022 có xu hướng giảm nhưng trong năm 2023 đã tăng trở lại. Với chuỗi log-return có trung bình xấp xỉ 0.05 và biến động trong khoảng từ [-11;11], biến động mạnh trong năm 2022 và giảm hơn ở năm 2023. Sự biến động mạnh và giảm trong năm 2022 là do tình hình thị trường dầu khí toàn cầu có biến động.

*** Kiểm định tính dừng**

Sử dụng kiểm định Dicky-Fuller, có giả thuyết:

$$\begin{cases} H_0: Chuỗi\ có\ nghiệm\ đơn\ vị\ (Chuỗi\ không\ dừng) \\ H_1: Chuỗi\ không\ có\ nghiệm\ đơn\ vị\ (Chuỗi\ dừng) \end{cases}$$

- Chuỗi giá đóng cửa PVS

Bảng 2: Kiểm định tính dừng cho chuỗi giá đóng cửa PVS

+ Kiểm định tính dừng xu thế:
 	
$P - value_{intercept} = 0.07 > 0.05$
 	
$P - value_{tt} = 0.08 > 0.05$
 	
$\lvert\tau_{stat}\rvert = 2.09 < \lvert\tau_{\alpha}\rvert = 3.42$
 	
⇒ Với mức ý nghĩa 5%, chưa có cơ sở để bác bỏ H_0 và chuỗi giá PVS không dừng xu thế
	`lm(formula = z.diff ~ z.lag.1 + 1 + tt)`
	
	`Residuals:`
	`  Min  1Q  Median  3Q  Max`
	`-2984.1  -424.9  -17.0  496.0  2959.5`
	
	`Coefficients:`
	`  Estimate Std. Error t value Pr(>|t|)`
	`(Intercept) 380.368574 211.006206  1.803  0.0721 .`
	`z.lag.1  -0.017705  0.008471  -2.090  0.0371 *`
	`tt  0.577075  0.336738  1.714  0.0872 .`
	`---`
	`Signif. codes:`
	`0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1`
	
	`Residual standard error: 906.4 on 494 degrees of freedom`
	`Multiple R-squared:  0.009724,  Adjusted R-squared:  0.005715`
	`F-statistic: 2.425 on 2 and 494 DF,  p-value: 0.08949`
	
	`Value of test-statistic is: -2.09 1.7198 2.4255`
	
	`Critical values for test statistics:`
	`  1pct  5pct 10pct`
	`tau3 -3.98 -3.42 -3.13`
	`phi2  6.15  4.71  4.05`
	`phi3  8.34  6.30  5.36`

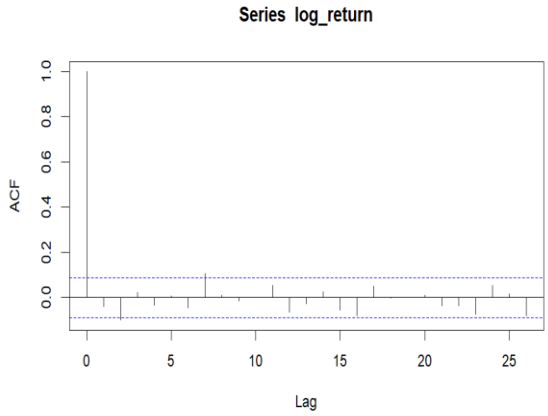
+ Kiểm định có hệ số chặn:
 	
$P - value_{intercept} = 0.144 > 0.05$
 	
$\lvert\tau_{stat}\rvert = 1.381 < \lvert\tau_{\alpha}\rvert = 2.87$
 	
⇒ Với mức ý nghĩa 5%, chưa có cơ sở để bác bỏ H_0 do đó chuỗi giá PVS không dừng
	`lm(formula = z.diff ~ z.lag.1 + 1)`
	
	`Residuals:`
	`  Min  1Q  Median  3Q  Max`
	`-3118.20  -436.40  18.51  499.12  2830.93`
	
	`Coefficients:`
	`  Estimate Std. Error t value Pr(>|t|)`
	`(Intercept) 301.954371 206.387897  1.463  0.144`
	`z.lag.1  -0.009863  0.007143  -1.381  0.168`
	
	`Residual standard error: 908.2 on 495 degrees of freedom`
	`Multiple R-squared:  0.003837,  Adjusted R-squared:  0.001825`
	`F-statistic: 1.907 on 1 and 495 DF,  p-value: 0.168`
	
	`Value of test-statistic is: -1.3808 1.1069`
	
	`Critical values for test statistics:`
	`  1pct  5pct 10pct`
	`tau2 -3.44 -2.87 -2.57`
	`phi1  6.47  4.61  3.79`

- Chuỗi log_return của giá đóng cửa PVS

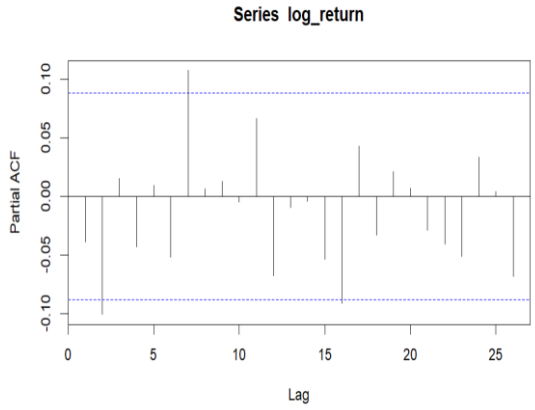
Bảng 3: Kiểm định tính dừng cho chuỗi log-return PVS

$P - value_{intercept} = 0.739 > 0.05$
 	
$\lvert\tau_{stat}\rvert = 23.1086 > \lvert\tau_{\alpha}\rvert = 2.87$
 	
⇒Với mức ý nghĩa 5% bác bỏ H_0 do đó chuỗi log-return là chuỗi dừng
	`lm(formula = z.diff ~ z.lag.1 + 1)`
	
	`Residuals:`
	`  Min  1Q  Median  3Q  Max`
	`-10.9453  -1.5373  0.0498  1.9407  9.5446`
	
	`Coefficients:`
	`  Estimate Std. Error t value Pr(>|t|)`
	`(Intercept)  0.07336  0.15157  0.484  0.629`
	`z.lag.1  -1.03898  0.04496 -23.109  <2e-16 ***`
	`---`
	`Signif. codes:`
	`0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1`
	
	`Residual standard error: 3.375 on 494 degrees of freedom`
	`Multiple R-squared:  0.5195,  Adjusted R-squared:  0.5185`
	`F-statistic:  534 on 1 and 494 DF,  p-value: < 2.2e-16`
	
	`Value of test-statistic is: -23.1086 267.0039`
	
	`Critical values for test statistics:`
	`  1pct  5pct 10pct`
	`tau2 -3.44 -2.87 -2.57`
	`phi1  6.47  4.61  3.79`

*** Xác định mô hình ARIMA**



Hình 3.1.1c: ACF của chuỗi log_return



Hình 3.1.1d: PACF của chuỗi log_return

Nhìn vào đồ thị ACF và PACF của chuỗi log_return, ta đề xuất các mô hình ARIMA sau:

ARIMA(2,0,2)	`sigma^2 estimated as 11.06:  log likelihood = -1302.55,  aic = 2617.1`
	`Training set error measures:`
	`ME RMSE MAE MPE MAPE MASE ACF1`
	`Training set -0.0001137222 3.325965 2.371678 NaN Inf 0.6584576 0.0005260892`

ARIMA(2,0,7)	`sigma^2 estimated as 10.98:  log likelihood = -1300.79,  aic = 2623.58`
	`Training set error measures:`
	`ME RMSE MAE MPE MAPE MASE ACF1`
	`Training set 0.0009045628 3.314028 2.37298 NaN Inf 0.6588191 -6.795708e-05`

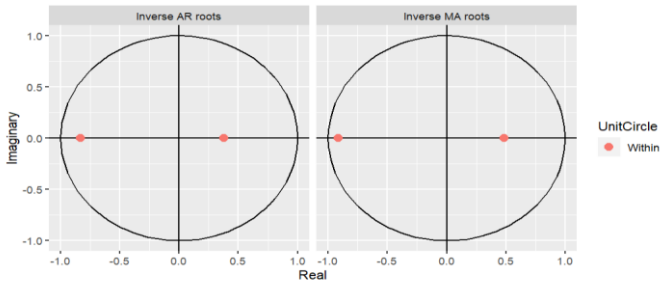
ARIMA(7,0,2)	`sigma^2 estimated as 11.01:  log likelihood = -1301.3,  aic = 2624.59`
	`Training set error measures:`
	`ME RMSE MAE MPE MAPE MASE ACF1`
	`Training set 0.0005937971 3.317482 2.368232 NaN Inf 0.657501 -0.002891654`

ARIMA(7,0,7)	`sigma^2 estimated as 10.61:  log likelihood = -1294.32,  aic = 2620.64`
	`Training set error measures:`
	`ME RMSE MAE MPE MAPE MASE ACF1`
	`Training set -0.0002437924 3.25701 2.35025 NaN Inf 0.6525087 0.0006317499`

So sánh chỉ số AIC của 4 mô hình trên thì được mô hình ARIMA(2,0,2) cho kết quả nhỏ nhất. Nhưng khi so sánh chỉ số RMSE thì mô hình ARIMA(7,0,7) lại là mô hình tốt nhất. Ở phần sau sẽ đánh giá kết quả dự báo để quyết định mô hình nào nên được sử dụng.

*** Đánh giá mô hình ARIMA**

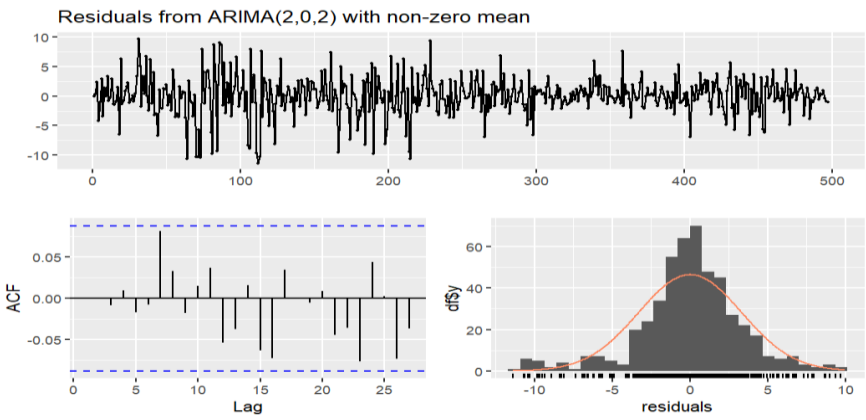
- Kiểm tra tính dừng thông qua nghiệm nghịch đảo



Nhận xét: Mô hình có các nghiệm nghịch đảo nằm trong vòng tròn đơn vị

- Kiểm định tính nhiễu trắng và tự tương quan của phần dư

```
        Ljung-Box test

data:  Residuals from ARIMA(2,0,2) with non-zero mean
Q* = 4.4042, df = 6, p-value = 0.6222
```



Kết quả cho thấy: phần dư không tự tương quan và là nhiễu trắng

*** Dự báo**

Bảng 4: Dự báo giá cổ phiếu PVS trong 10 phiên tiếp theo

Mô hình	ARIMA(2,0,2)	ARIMA(2,0,7)	ARIMA(7,0,2)	ARIMA(7,0,7)
Giá trị dự báo	40706.9	42112.58	40753.74	29367.11
	49989.31	50331.81	50190.64	29671.69
	37557.82	42340.81	39461.58	48488.31
	45460.16	47779.87	46965	40448.88
	37519.85	39852.43	38142.9	19268.29

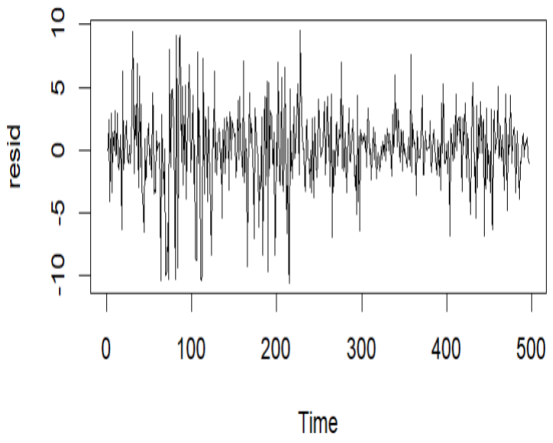
	42241.99	36886.52	38914.64	47143.68
	37102.25	33782.43	35975.13	56125.3
	41812.31	39108.4	40294.78	33919.92
	38021.17	39297.84	38705	38483.56
	40784.91	39435.11	39533.86	74076.18
MAPE	0.087	0.095	0.08	0.301

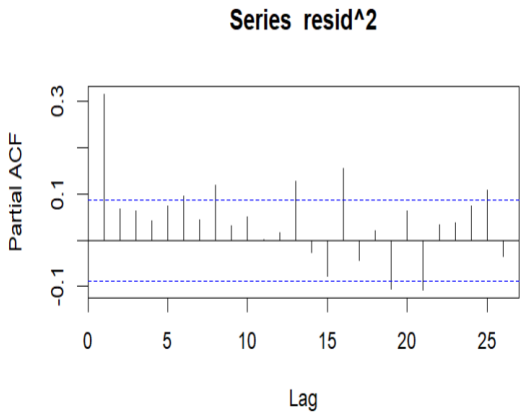
Từ bảng 4, ta thấy giá trị dự báo cho 10 ngày tiếp theo giá cổ phiếu PVS của mô hình ARIMA(7,0,2) cho sai số dự báo MAPE nhỏ nhất. Vậy nên ta sẽ sử dụng mô hình ARIMA(7,0,2) để thực hiện phân tích rủi ro.

4. Sử dụng ARCH/GARCH để dự báo rủi ro cho 10 ngày đầu năm 2024

ARCH là viết tắt của "AutoRegressive Conditional Heteroskedasticity". Nó là một lớp mô hình trong phân tích chuỗi thời gian và dùng để mô hình hóa biến động không đồng nhất (heteroskedasticity) trong dữ liệu. Cụ thể, các mô hình ARCH được sử dụng để phân tích và dự đoán biến động trong phương sai của dữ liệu thời gian, đặc biệt là trong tài chính và kinh tế.

Ở trên đã chọn ra được mô hình ARIMA(7,0,2) của chuỗi log-return có sai số dự báo MAPE nhỏ nhất nên sẽ sử dụng mô hình này để áp dụng ARCH phân tích biến động của phương sai.





Đây là đồ thị của phần dư và PACF của bình phương phần dư được lấy từ mô hình ARIMA(7,0,2). Sau khi kiểm định thì có ARCH ở bậc 1.

```
        ARCH LM-test; Null hypothesis: no ARCH effects

data:  resid
Chi-squared = 49.798, df = 1, p-value = 1.704e-12
```

Mô hình ARCH(1) : $\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + v_t = 8.39231 + 0.21981 u_{t-1}^2 + v_t$

Bảng 5: Dự báo giá trị volatility cho 10 ngày đầu năm 2024

02/01	**03/01**	**04/01**	**05/01**	**08/01**	**09/01**	**10/01**	**11/01**	**12/01**	**15/01**
16.47	10.57	6.79	4.36	2.81	1.807	1.176	0.776	0.529	0.384

II. Chuỗi PVG – Trịnh Thu Huyền

1. Giới thiệu

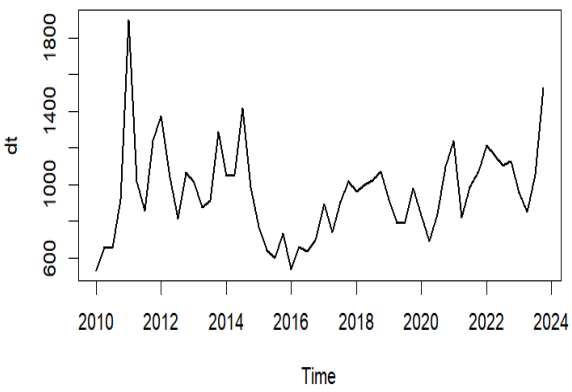
PVG là mã chứng khoán của công ty cổ phần kinh doanh LPG Việt Nam được niêm yết trên Sở Giao dịch Chứng khoán thành phố Hà Nội (HNX). Công ty là một thành viên của Tập đoàn Dầu khí Việt Nam, có hoạt động sản xuất kinh doanh trải dài từ Đà Nẵng trở ra các tỉnh phía Bắc. Ở phần này, chúng ta sẽ đi vào phân tích và dự báo về giá đóng cửa, lợi suất và doanh thu 10 phiên đầu tiên của mã chứng khoán PVG trong năm 2024.

2. Phân tích với số liệu theo quý – Chuỗi doanh thu thuần

Số liệu: chuỗi doanh thu thuần theo quý của Công ty cổ phần kinh doanh LPG Việt Nam từ quý 1 năm 2010 đến hết quý 4 năm 2023 với mục tiêu dự báo cho 4 quý năm 2024.

Đơn vị: Tỷ VNĐ

Nhìn vào đồ thị, doanh thu của công ty khá biến động trong giai đoạn 2010-2023. Năm 2010, do ảnh hưởng của khủng hoảng tài chính toàn cầu năm 2009, doanh thu chỉ dao động trong khoảng từ 500 đến dưới 1000 tỷ đồng. Đến quý 1 năm 2011, doanh thu đạt gần 2000 tỷ đồng, gấp 4 lần so với cùng kì năm trước với nguyên nhân là do công ty có khoản doanh thu bất thường từ việc thoái toàn bộ vốn tại CTCP Phân phối khối áp thấp dầu khí (PGD). Từ năm 2011-2014, doanh thu dao động khá ổn định trong khoảng 800-1400 tỷ, nhưng sau giai đoạn này, doanh thu giảm mạnh vào 2 năm 2015 và 2016, cụ thể chạm đáy vào quý 1 năm 2016 khi doanh thu chỉ còn 540 tỷ đồng do một lần nữa bị tác động bởi khủng hoảng tài chính diễn ra năm 2015. Kể từ đó đến nay, doanh thu công ty ổn định dần trong khoảng 700-1500 tỷ.



Doanh thu PVG quý 1/2010 đến quý 4/2023

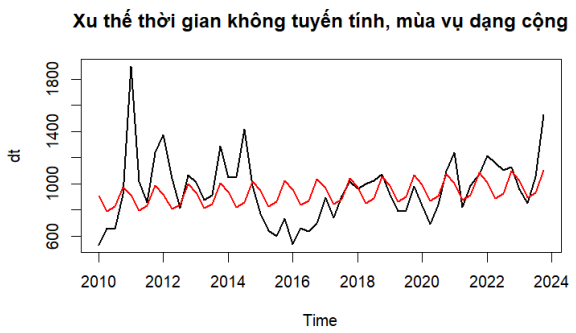
Bảng 6: Các mô hình dự báo cho chuỗi doanh thu PVG

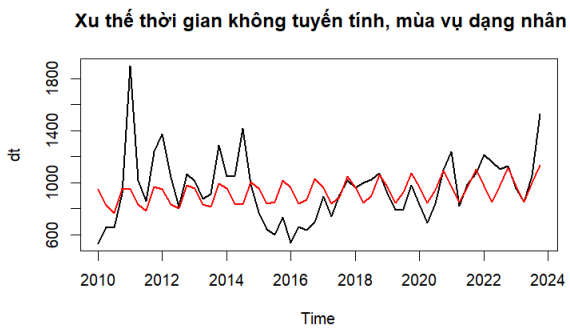
STT	Mô hình	R-sq	RMSE	MAPE	Dự báo cho 4 quý năm 2024
1	MH linear-linear: $\widehat{DT} = 912.763 + 1.534t$	0.98%	249.81	0.2070	Q1 = 1000.201 Q2 = 1000.735 Q3 = 1003.269 Q4 = 1004.803

2	MH linear – log: $\widehat{DT} = 873.19 + 27.15\ln(t)$	0.92%	249.89	0.2068	Q1 = 982.959 Q2 = 983.431 Q3 = 983.895 Q4 = 984.351
3	MH log-linear: $\ln\widehat{(DT)} = 6.76 + 0.002494t$	2.496%	251.73	0.2001	Q1 = 984.524 Q2 = 986.982 Q3 = 989.447 Q4 = 991.917
4	MH log-log: $ln\widehat{(DT)} = 6.67874 + 0.04941ln(t)$	2.961%	252.28	0.1991	Q1 = 971.175 Q2 = 972.010 Q3 = 972.832 Q4 = 973.639
5	Hồi quy với biến giả mùa: $\widehat{DT} = 1013.15 - 156.79s_2 - 114.41s_3 + 45.56s_4$	10.72%	237.22	0.1939	Q1 = 1013.15 Q2 = 856.360 Q3 = 898.740 Q4 = 1058.71
6	Xu thế thời gian tuyến tính và mùa vụ dạng cộng: $\widehat{DT} = 973.611 + 1.465t - 158.25s_2 - 117.339s_3 + 41.162s_4$	11.6%	236.04	0.1918	Q1 = 1057.116 Q2 = 900.3310 Q3 = 942.7070 Q4 = 1102.673
7	Xu thế thời gian tuyến tính và mùa vụ dạng nhân: $\widehat{DT} = 1074.274 - 2.264t + 2.585t * s_2 + 6.337t * s_3 + 5.991t * s_4$	14.4%	232.27	0.1869	Q1 = 945.2260 Q2 = 1092.892 Q3 = 1314.581 Q4 = 1297.894
8	Xu thế thời gian không tuyến tính và mùa vụ dạng cộng: $ln\widehat{(DT)} = 6.804 + 0.0023t - 0.136s_2 - 0.097s_3 + 0.071s_4$	12.65%	238.09	0.1836	Q1 = 1027.722 Q2 = 899.1051 Q3 = 937.0160 Q4 = 1110.982

9	Xu thế thời gian không tuyến tính và mùa vụ dạng nhân: $\widehat{ln(DT)} = 6.853 + 0.00056t - 0.000041t * s_2 + 0.0045t * s_3 + 0.0028t * s_4$	14.14%	235.16	0.1779	Q1 = 977.423 Q2 = 975.648 Q3 = 1276.07 Q4 = 1158.17
10	Holt-Winter dạng cộng: $\widehat{DT^F_{T+h}} = 1505.510 + 26.845 * h + seas$		809.26	0.6878	Q1 = 1232.387 Q2 = 1188.061 Q3 = 1336.042 Q4 = 1610.933
11	Holt-winter dạng nhân: $\widehat{DT^F_{T+h}} = (1851.283 + 40.610 * h) * seas$		741.87	0.6308	Q1 = 1360.588 Q2 = 1233.890 Q3 = 1382.999 Q4 = 1661.661

Nhận xét: Nhìn vào kết quả dự báo từ các mô hình, ta thấy hầu hết hệ số R bình phương rất thấp. Vì vậy, dựa vào chỉ số RMSE và MAPE để đề xuất 2 mô hình có MAPE thấp nhất là: Mô hình xu thế thời gian không tuyến tính và mùa vụ dạng cộng, mô hình xu thế thời gian không tuyến tính và mùa vụ dạng nhân.



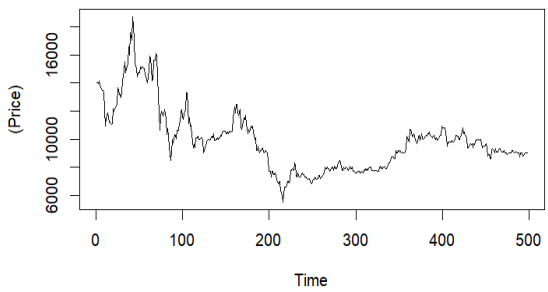


Đồ thị của 2 mô hình cho thấy kết quả dự báo khá sát trong giai đoạn 2017-2023 nhưng lệch nhiều vào giai đoạn 2010-2016 do mô hình không dự đoán trước được cú shock trong giai đoạn này.

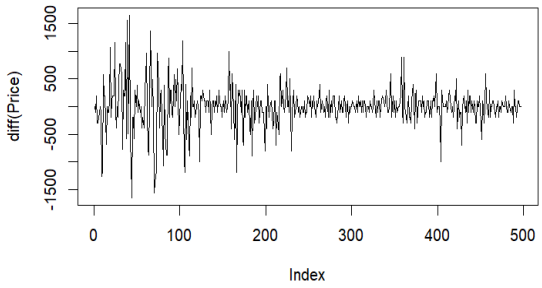
3. Phân tích với số liệu theo ngày – Chuỗi giá đóng cửa và chuỗi log_return

Số liệu: Giá đóng cửa của mã chứng khoán PVG từ năm 2022 – 2023 (số liệu tính theo ngày, bao gồm 498 quan sát, đơn vị: VNĐ)

3.1. Chuỗi giá đóng cửa



Đồ thị chuỗi sai giá PVG



Đồ thị sai phân chuõi giá PVG

Nhận xét: Chuỗi giá PVG chưa thể kết luận, chuỗi sai phân ΔPVG có thể dừng.

Kiểm định tính dừng

Sử dụng kiểm định Dicky-Fuller, có giả thuyết:

$$\begin{cases} H_0: Chuỗi\ có\ nghiệm\ đơn\ vị\ (Chuỗi\ không\ dừng) \\ H_1: Chuỗi\ không\ có\ nghiệm\ đơn\ vị\ (Chuỗi\ dừng) \end{cases}$$

$\tau = \frac{\hat{\delta}}{Se(\hat{\delta})}$; Nếu $|\tau| > |\tau_\alpha|$ thì bác bỏ H_0

Bảng 7: Kiểm định tính dừng cho chuỗi giá đóng cửa PVG

<u>*Có xu thế và có hệ số chặn:*</u>
$P - value_{intercept} = 0.03 < 0.05$
$\rightarrow$ *Hệ số chặn có ý nghĩa thống kê*
$P - value_{tt} = 0.36 > 0.05$
$\rightarrow$ *Xu thế không có ý nghĩa thống kê*
$\lvert\tau_{stat}\rvert = 2.56 < \lvert\tau_\alpha\rvert = 3.42$
$\Rightarrow$ Với mức ý nghĩa 5%, chưa bác bỏ H_0 do đó chuỗi giá PVG không dừng xu thế.
	`Test regression trend`
	
	`Call:`
	`lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)`
	
	`Residuals:`
	`  Min  1Q  Median  3Q  Max`
	`-1425.34  -143.72  12.54  145.24  1861.41`
	
	`Coefficients:`
	`  Estimate Std. Error t value Pr(>|t|)`
	`(Intercept) 249.437420 113.091456  2.206  0.02787 *`
	`z.lag.1  -0.022810  0.008915  -2.559  0.01080 *`
	`tt  -0.127367  0.141309  -0.901  0.36785`
	`z.diff.lag  0.117260  0.044741  2.621  0.00904 **`
	`---`
	`Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1`
	
	`Residual standard error: 377.7 on 492 degrees of freedom`
	`Multiple R-squared:  0.0251,  Adjusted R-squared:  0.01916`
	`F-statistic: 4.223 on 3 and 492 DF,  p-value: 0.005794`
	
	`Value of test-statistic is: -2.5587 2.3929 3.4457`
	
	`Critical values for test statistics:`
	`  1pct  5pct 10pct`
	`tau3 -3.98 -3.42 -3.13`
	`phi2  6.15  4.71  4.05`
	`phi3  8.34  6.30  5.36`

<u>*Có hệ số chặn:*</u> $P - value_{intercept} = 0.02 < 0.05$ $\rightarrow$ *Hệ số chặn có ý nghĩa thống kê* $\lvert\tau_{stat}\rvert = 2.46 < \lvert\tau_{\alpha}\rvert = 3.42$ $\Rightarrow$ Với mức ý nghĩa 5%, chưa bác bỏ H_o do đó chuỗi giá PVG không dừng.	Test regression drift Call: lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag) Residuals: Min 1Q Median 3Q Max -1429.30 -139.67 -5.47 141.90 1855.24 Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) 174.141752 76.215823 2.285 0.0227 * z.lag.1 -0.018434 0.007475 -2.466 0.0140 * z.diff.lag 0.114168 0.044600 2.560 0.0108 * --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 377.7 on 493 degrees of freedom Multiple R-squared: 0.02349, Adjusted R-squared: 0.01953 F-statistic: 5.931 on 2 and 493 DF, p-value: 0.00285 Value of test-statistic is: -2.466 3.1844 Critical values for test statistics: 1pct 5pct 10pct tau2 -3.44 -2.87 -2.57 phi1 6.47 4.61 3.79

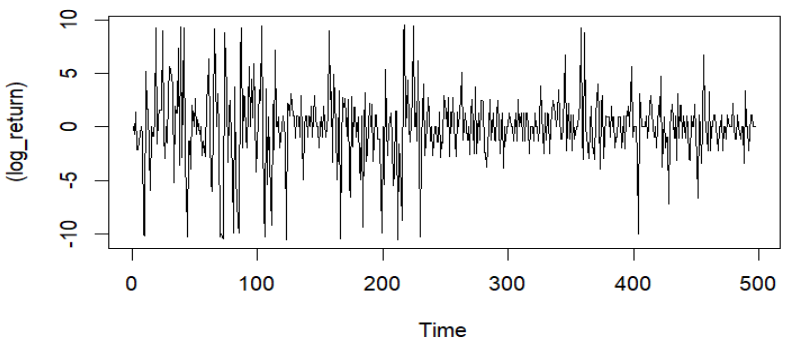
Bảng 8: Kiểm định tính dừng cho chuỗi sai phân

<u>*Có xu thế và có hệ số chặn*</u> $P - value_{intercept} = 0.4 > 0.05$ $\rightarrow$ *Hệ số chặn không có ý nghĩa thống* $P - value_{tt} = 0.57 > 0.05$ $\rightarrow$ *Xu thế không có ý nghĩa thống kê* $\lvert\tau_{stat}\rvert = 14.77 > \lvert\tau_{\alpha}\rvert = 3.42$ $\Rightarrow$ Với mức ý nghĩa 5%, bác bỏ H_o do đó chuỗi sai phân dừng .	Call: lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag) Residuals: Min 1Q Median 3Q Max -1531.24 -125.26 6.17 135.63 1729.32 Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) -25.736315 34.417483 -0.748 0.455 z.lag.1 -0.891256 0.060335 -14.772 <2e-16 *** tt 0.067622 0.119815 0.564 0.573 z.diff.lag -0.002767 0.045127 -0.061 0.951 --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 380.6 on 491 degrees of freedom Multiple R-squared: 0.4469, Adjusted R-squared: 0.4435 F-statistic: 132.2 on 3 and 491 DF, p-value: < 2.2e-16 Value of test-statistic is: -14.7719 72.7365 109.1046 Critical values for test statistics: 1pct 5pct 10pct tau3 -3.98 -3.42 -3.13 phi2 6.15 4.71 4.05 phi3 8.34 6.30 5.36
<u>*Có hệ số chặn:*</u> $P - value_{intercept} = 0.6 > 0.05$ $\rightarrow$ *Hệ số chặn không có ý nghĩa thống* $\lvert\tau_{stat}\rvert = 14.77 > \lvert\tau_{\alpha}\rvert = 2.87$ $\Rightarrow$ Với mức ý nghĩa 5%, bác bỏ H_o do đó chuỗi sai phân dừng .	Call: lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag) Residuals: Min 1Q Median 3Q Max -1543.68 -128.73 9.57 136.24 1714.46 Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) -8.884720 17.106966 -0.519 0.604 z.lag.1 -0.889923 0.060247 -14.771 <2e-16 *** z.diff.lag -0.003431 0.045081 -0.076 0.939 --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' Residual standard error: 380.4 on 492 degrees of freedom Multiple R-squared: 0.4465, Adjusted R-squared: 0.4443 F-statistic: 198.5 on 2 and 492 DF, p-value: < 2.2e-16 Value of test-statistic is: -14.7714 109.0965 Critical values for test statistics: 1pct 5pct 10pct tau2 -3.44 -2.87 -2.57 phi1 6.47 4.61 3.79

<u>*Có hệ số chặn:*</u> $\lvert\tau_{stat}\rvert = 14.77 > \lvert\tau_{\alpha}\rvert = 1.95$ $\Rightarrow$ Với mức ý nghĩa 5%, bác bỏ H_o do đó chuỗi sai phân dừng quanh 0.	Call: lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag) Residuals: Min 1Q Median 3Q Max -1552.8 -137.6 0.8 127.3 1705.0 Coefficients: Estimate Std. Error t value Pr(>|t|) z.lag.1 -0.888799 0.060163 -14.773 <2e-16 *** z.diff.lag -0.003993 0.045034 -0.089 0.929 --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 380.1 on 493 degrees of freedom Multiple R-squared: 0.4462, Adjusted R-squared: 0.444 F-statistic: 198.6 on 2 and 493 DF, p-value: < 2.2e-16 Value of test-statistic is: -14.7732 Critical values for test statistics: 1pct 5pct 10pct tau1 -2.58 -1.95 -1.62

Vậy chuỗi giá PVG không dừng, chuỗi sai phân dừng hay chuỗi giá dừng sai phân bậc 1 ~ I(1)

3.2. Chuỗi log_return của giá đóng cửa PVG



Đồ thị chuỗi log_return

Nhận xét: Chuỗi log_return có thể dừng

*** Kiểm định tính dừng**

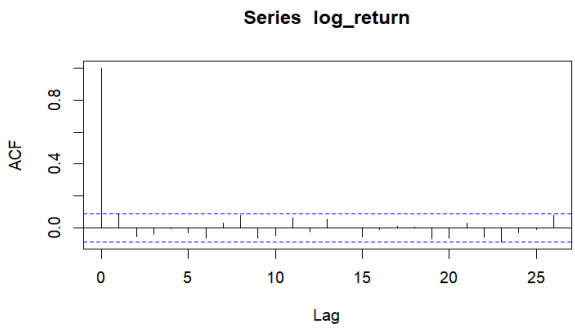
```
Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-10.7105  -1.3700   0.1118   1.5275  10.3108

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.2569037  0.3091140  -0.831    0.406
z.lag.1     -0.9671649  0.0608795 -15.887   <2e-16 ***
tt           0.0006878  0.0010761   0.639    0.523
z.diff.lag   0.0591912  0.0450492   1.314    0.189
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.418 on 491 degrees of freedom
Multiple R-squared:  0.4585,	Adjusted R-squared:  0.4552
F-statistic: 138.6 on 3 and 491 DF,  p-value: < 2.2e-16


Value of test-statistic is: -15.8865 84.1274 126.191

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.98 -3.42 -3.13
phi2  6.15  4.71  4.05
phi3  8.34  6.30  5.36
```

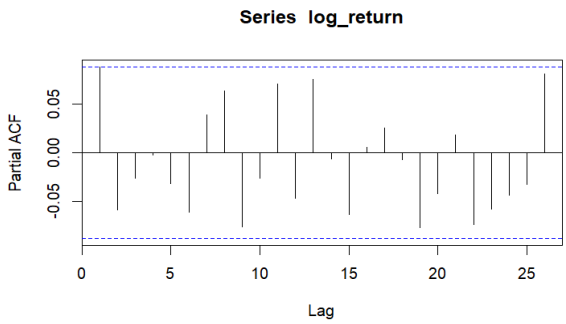
$P - value_{intercept} = 0.406 > 0.05$

$|\tau_{stat}| = 15.8865 > |\tau_{\alpha}| = 3.42$

Với mức ý nghĩa 5% bác bỏ H_0, chuỗi log-return là chuỗi dừng

*** Xác định mô hình ARIMA**





Nhìn vào đồ thị ACF và PACF của chuỗi log_return, ta đề xuất các mô hình ARIMA sau:

ARIMA(1,0,1)

```
sigma^2 estimated as 11.57:  log likelihood = -1313.66,  aic = 2635.32

Training set error measures:
                       ME     RMSE      MAE MPE MAPE      MASE
Training set -0.0001612047 3.401549 2.359601 NaN  Inf 0.7088299
                     ACF1
Training set -0.003034044
```

ARIMA(0,0,1)

```
sigma^2 estimated as 11.58:  log likelihood = -1313.93,  aic = 2633.86

Training set error measures:
                       ME     RMSE      MAE MPE MAPE      MASE
Training set -1.741811e-05 3.403401 2.361602 NaN  Inf 0.7094309
                     ACF1
Training set -0.004057558
```

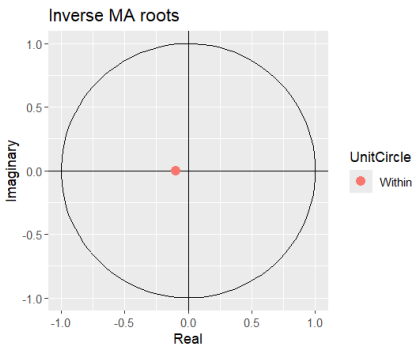
ARIMA(1,0,0)

```
sigma^2 estimated as 11.59:  log likelihood = -1314.15,  aic = 2634.3

Training set error measures:
                      ME     RMSE     MAE MPE MAPE      MASE
Training set 7.42617e-05 3.404904 2.35717 NaN  Inf 0.7080997
                    ACF1
Training set 0.005285847
```

Để lựa chọn mô hình, ta sử dụng tiêu chuẩn AIC để đánh giá. AIC càng nhỏ thì mô hình càng phù hợp. Do vậy, dựa vào chỉ số AIC của 3 mô hình trên thì ARIMA(0,0,1) cho AIC nhỏ nhất.

*** Kiểm định mô hình ARIMA(0,0,1)**

- Kiểm tra tính dừng thông qua nghiệm nghịch đảo



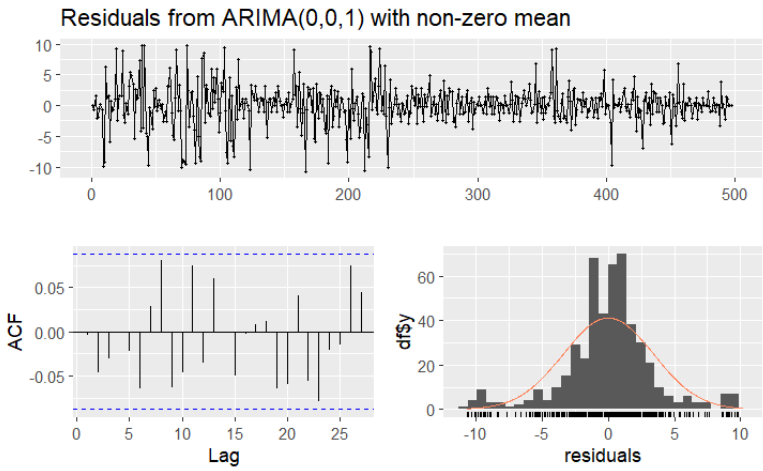
Nhận xét: mô hình có nghiệm nghịch đảo nằm trong vòng tròn đơn vị do đó mô hình cho kết quả dừng.

- Kiểm định tính nhiễu trắng và tự tương quan của phần dư

$$\begin{cases} H_0: u_t \text{ không có tự tương quan} \\ H_1: u_t \text{ có tự tương quan} \end{cases}$$

```
        Ljung-Box test

data:  Residuals from ARIMA(0,0,1) with non-zero mean
Q* = 10.846, df = 9, p-value = 0.2864

Model df: 1.   Total lags used: 10
```

$P - value = 0.2864 > 0.05 \Rightarrow Với\ ý\ nghĩa\ 5\%,$ Chưa bác bỏ H_o, u_t không có tự tương quan.



ACF của phần dư không có tự tương quan ở các bậc ⇒ u_t là nhiễu trắng.

Vậy phần dư không có tự tương quan và nhiều trắng.

*** Dự báo**

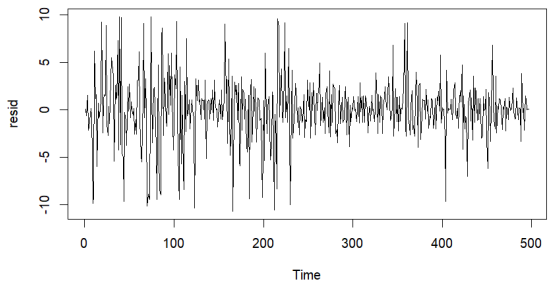
Sử dụng 3 mô hình ARIMA của chuỗi log_return đã đề xuất ở trên, ta dự báo chuỗi log_return từ đó dự báo giá cổ phiếu PVG trong 10 phiên tiếp theo để đánh giá mô hình.

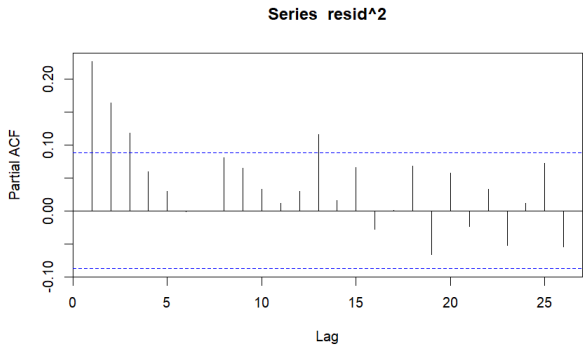
Bảng 9: Dự báo giá cổ phiếu PVG trong 10 phiên tiếp theo

Mô hình	ARIMA(1,0,1)	ARIMA(0,0,1)	ARIMA(1,0,0)
Giá trị dự báo	8267.825	8293.41	8292.657
	8220.279	8228.622	8233.463
	8232.236	8228.622	8228.302
	8320.657	8320.051	8319.271
	8321.425	8320.051	8319.231
	8229.789	8228.622	8227.808
	8229.838	8228.622	8227.807
	8321.268	8320.051	8319.228
	8321.271	8320.051	8319.228
	8229.828	8228.622	8227.807
MAPE	0.0919	0.09170	0.09172

Từ bảng trên, giá trị dự báo cho 10 ngày tiếp theo giá cổ phiếu PVG của mô hình ARIMA(0,0,1) từ chuỗi log_return cho sai số dự báo MAPE nhỏ nhất. Đây cũng là mô hình có chỉ số AIC nhỏ nhất.

4. Sử dụng ARCH/GARCH để dự báo rủi ro cho 10 ngày đầu năm 2024





Nhìn vào đồ thị PACF của u_t^2 ta thấy trễ bậc 1. Lấy phần dư từ mô hình ARIMA(0,0,1)

Thực hiện kiểm định tự động ARCH(1)

$$\text{Cặp giả thuyết: } \begin{cases} H_0\text{: Không có hiệu ứng ARCH} \\ H_1\text{: Có hiệu ứng ARCH} \end{cases}$$

```
        ARCH LM-test; Null hypothesis: no ARCH effects

data:  resid
Chi-squared = 25.531, df = 1, p-value = 4.354e-07
```

$p_{value} = 4.354 * 10^{-7} < 0.05. Vậy\ với\ ý\ nghĩa\ 5\%, bác\ bỏ\ H_0\ do$ đó có $ARCH(1)$

Mô hình ARCH(1) : $\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + v_t = 7.968 + 0.352 u_{t-1}^2 + v_t$

Bảng 10: Dự báo giá trị volatility cho 10 ngày đầu năm 2024

02/01	**03/01**	**04/01**	**05/01**	**08/01**	**09/01**	**10/01**	**11/01**	**12/01**	**15/01**
2.807	3.279	3.435	3.491	3.510	3.518	3.520	3.521	3.522	3.522

III. Chuỗi PVC – Phạm Diệu Linh

1. Giới thiệu

Mã PVC là mã chứng khoán của Tổng công ty Hoá chất và Dịch vụ Dầu khí (PVChem), được thành lập năm 1990 với tiền thân là Công ty Dung dịch khoan và Hoá phẩm Dầu khí (DMC), hoạt động chủ đạo là dịch vụ hoá kỹ thuật/dịch vụ dung dịch khoan.

Mã PVC được niêm yết trên sàn HaNoi Stock Exchange (HNX). Dữ liệu giá đóng cửa của PVC được lấy từ trang investing.com.

2. Phân tích với số liệu theo quý – Chuỗi doanh thu thuần

Số liệu: chuỗi doanh thu thuần theo quý của Công ty từ quý 1 năm 2008 đến hết quý 4 năm 2023 với mục tiêu dự báo cho 4 quý năm 2024.

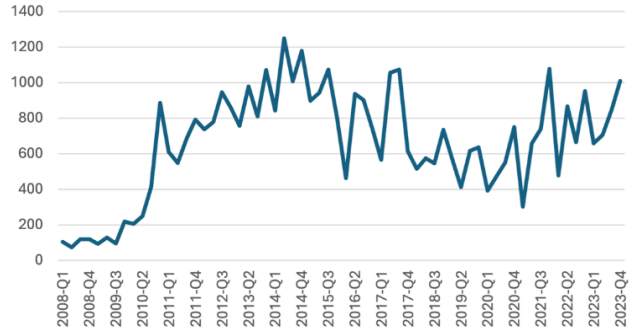
Đơn vị: Tỷ VNĐ

Đồ thị là dữ liệu doanh thu thuần từ Quý 1 – 2001 đến Quý 4 – 2023, bao gồm 64 quan sát. Đồ thị cho thấy chuỗi có biến động phức tạp và có xu thế không rõ ràng.

Cuối năm 2015, doanh thu thuần giảm mạnh. Điều này có thể được lý giải vì diễn biến phức tạp trên biển Đông tiềm ẩn nhiều rủi ro tới nền kinh tế, xã hội cũng như các doanh nghiệp. Bên cạnh đó, giá dầu giảm trong thời gian trước cũng khiến cho ngân sách của các chương trình khoan bị cắt giảm, làm gia tăng áp lực về việc giảm giá dịch vụ, dẫn tới doanh thu và lợi nhuận bị ảnh hưởng.

Đầu năm 2021, doanh thu thuần giảm. Điều này có thể được lý giải vì tình hình kinh tế/xã hội đang có nhiều bất ổn bởi tác động của đại dịch Covid – 19. Covid – 19 khiến cước vận tải biển tăng cao dẫn tới giá hoá chất và chi phí đầu vào tăng.

Hình 2.1. Đồ thị doanh thu thuần của PVC (Tr.Đ)

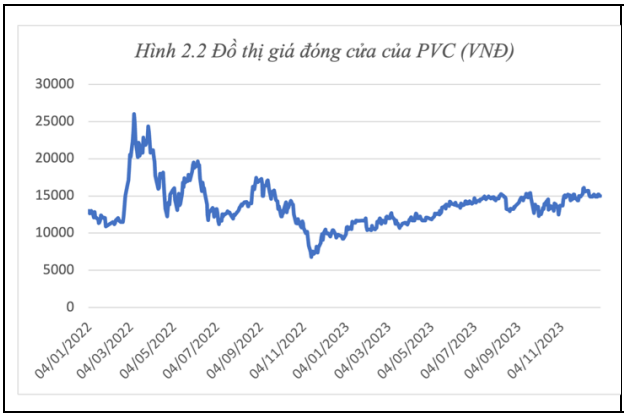


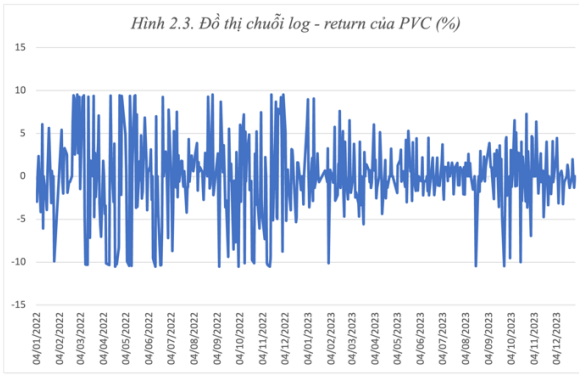
Bảng 11: Các mô hình dự báo cho chuỗi doanh thu PVC

STT	Mô hình	R^2	MAPE RMSE	Dự báo cho 2024
1	Mô hình linear – linear: $\widehat{NetR}_t = 463.615 + 6.097t$	0.1398	0.7419019 279.3839	Q1: 832.92
				Q2: 839.017
				Q3: 845.114
				Q4: 851.211
2	Mô hình linear – log: $\widehat{NetR}_t = 44.54 + 192.54ln(t)$	0.3291	0.5212637 246.7307	Q1: 848.277
				Q2: 851.216
				Q3: 854.112
				Q4: 856.964
3	Mô hình log – linear: $ln(\widehat{NetR}_t) = 5.691 + 0.0193t$	0.2586	0.5874515 318.6732	Q1: 1038.466
				Q2: 1058.703
				Q3: 1079.335
				Q4: 1100.368
4	Mô hình log – log: $ln(\widehat{NetR}_t) = 4.475 + 0.575\ln(t)$	0.5417	0.4428559 287.8414	Q1: 968.039
				Q2: 976.575
				Q3:985.056
				Q4: 993.483
5	Mô hình hồi quy với biến mùa vụ: $\widehat{NetR}_t = 512.31 + 152.26 * s2 + 180.46 * s3 + 265.14 * s4$	0.1011	0.8604948 285.604	Q1: 512.31
				Q2: 664.57
				Q3: 692.77
				Q4: 777.45
6	Mô hình tuyến tính thời gian, mùa vụ (dạng cộng):	0.2279	0.6995167 264.6985	Q1: 710.093
				Q2: 862.351
				Q3: 890.557

	$\widehat{NetR}_t = 331.988 + 5.817 * t + 146.441 * s2_t + 168.83 * s3_t + 247.689 * s4_t$			Q4: 975.233
7	Mô hình tuyến tính thời gian, mùa vụ (dạng nhân): $\widehat{NetR}_t = 412.126 + 3.232 * t + 43.692 * s2_t + 73.386 * s3_t + 119.451 * s4_t + 3.292 * t * s2_t + 3.049 * t * s3_t + 4 * t * s4_t$	0.2367	0.7018448 263.1852	Q1: 622.206 Q2: 886.402 Q3: 906.339 Q4: 1023.353
8	Mô hình không tuyến tính thời gian, mùa vụ (dạng công): $\ln(\widehat{NetR}_t) = 5.486 + 0.0188 * t + 0.203 * s2_t + 0.268 * s3_t + 0.41 * s4_t$	0.3028	0.5587274 304.2151	Q1: 818.931 Q2: 1022.29 Q3: 1111.649 Q4: 1305.577
9	Mô hình không tuyến tính thời gian, mùa vụ (dạng nhân): $\ln(\widehat{NetR}_t) = 5.573 + 0.016 * t + 0.008 * s2_t + 0.147 * s3_t + 0.376 * s4_t + 0.006 * t * s2_t + 0.0039 * t * s3_t + 0.0012 * s4_t$	0.3067	0.5594324 305.4841	Q1: 744.714 Q2: 1133.426 Q3: 1156.669 Q4: 1234.721
10	Mô hình Holt – Winter dạng cộng: $\widehat{NetR}_{t+k} = 757.933 + 5.723 * k + s_{seas}$ (s1 = -147.135981, s2 = 39.581645, s3 = 79.522125, s4 = 246.300615)		0.2866476 221.216	Q1: 616.52 Q2: 808.96 Q3: 854.624 Q4: 1027.126
11	Mô hình Holt – Winter dạng nhân: $\widehat{Net}_{t+k} = (824.246 + 5.5723 * k) * s_{seas}$ (s1 = 0.8007359, s2 = 1.0257075, s3 = 1.1045451, s4 = 1.1346116)		0.3127251 236.5392	Q1: 664.465 Q2: 856.866 Q3: 928.881 Q4: 960.489

3. Phân tích với số liệu theo ngày – Chuỗi giá đóng cửa và chuỗi log_return





Dựa vào đồ thị trên, ta có thể dự đoán chuỗi giá đóng cửa của PVC không phải một chuỗi dừng, tuy nhiên chuỗi log – return của PVC có khả năng lại là một chuỗi dừng. Để kiểm định giá cổ phiếu PVC có phải một chuỗi dừng hay không, ta sử dụng kiểm định Dickey – Fuller.

3.1. Kiểm định tính dừng

Kiểm định Dickey – Fuller:

- Không có hệ số chặn:

$$\Delta Y_t = \delta Y_{t-1} + u_t$$

$$\begin{cases} H_0: \delta = 0 \\ H_1: \delta < 0 \end{cases} \Longrightarrow \begin{cases} H_0: \text{Có nghiệm đơn vị} \\ H_1: \text{Không có nghiệm đơn vị} \end{cases}$$

- Có hệ số chặn:

$$\Delta Y_t = \mu + \delta Y_{t-1} + u_t$$

$$\begin{cases} H_0: \delta = 0 \\ H_1: \delta < 0 \end{cases} \Longrightarrow \begin{cases} H_0: \text{Có nghiệm đơn vị quanh } \mu \\ H_1: \text{Dừng quanh } \mu \end{cases}$$

- Có hệ số chặn và xu thế:

$$\Delta Y_t = \mu + \delta Y_{t-1} + \beta t + u_t$$

$$\begin{cases} H_0: \delta = 0 \\ H_1: \delta < 0 \end{cases} \Longrightarrow \begin{cases} H_0: \text{Có nghiệm đơn vị quanh xu thế} \\ H_1: \text{Dừng xu thế} \end{cases}$$

$\tau = \frac{\hat{\delta}}{Se(\hat{\delta})}$; Nếu $|\tau| > |\tau_\alpha|$ thì bác bỏ H_0

Bảng 12: Kiểm định tính dừng chuỗi giá PVC

<u>Có hệ số chặn và xu thế:</u> $P - value_{intercept} < 0.05$ → Intercept có ý nghĩa thống kê $P - value_{tt} > 0.05$ → Xu thế không có ý nghĩa thống kê	> summary(ur.df(Price,type = "trend",lag = 0)) Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) 362.400411 157.311489 2.304 0.0217 * z.lag.1 -0.024624 0.009974 -2.469 0.0139 * tt -0.080426 0.203268 -0.396 0.6925

	Value of test-statistic is: -2.4689 2.0397 3.0498 Critical values for test statistics: 1pct 5pct 10pct tau3 -3.98 -3.42 -3.13
<u>Có hệ số chặn:</u> $P - value_{intercept} < 0.05$ → Intercept có ý nghĩa thống kê $\lvert\tau_{stat}\rvert < \lvert\tau_{\alpha}\rvert$ ($\tau_{stat} = -2.4399$) → Chuỗi có nghiệm đơn vị → Chuỗi không dừng	> summary(ur.df(Price,type = "drift",lag=0)) Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) 332.293711 137.567385 2.415 0.0161 * z.lag.1 -0.023891 0.009791 -2.440 0.0150 * Value of test-statistic is: -2.4399 2.9863 Critical values for test statistics: 1pct 5pct 10pct tau2 -3.44 -2.87 -2.57

<u>Kết luận:</u> Chuỗi giá cổ phiếu PVC là chuỗi không dừng, có hệ số chặn. Suy ra, phải thực hiện kiểm định ADF cho sai phân bậc 1 của chuỗi giá PVC.

Bảng 13: Kiểm định tính dừng chuỗi sai phân PVC

<u>Có hệ số chặn và xu thế:</u> $P - value_{intercept} > 0.05$ → Intercept không có ý nghĩa thống kê $P - value_{tt} > 0.05$ → Xu thế không có ý nghĩa thống kê	> summary(ur.df(diff(Price),type = "trend",lag = 0)) Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) 3.785375 57.508951 0.066 0.948 z.lag.1 -0.892310 0.044760 -19.936 <2e-16 *** tt 0.002189 0.200521 0.011 0.991 Value of test-statistic is: -19.9356 132.4765 198.7144 Critical values for test statistics: 1pct 5pct 10pct tau3 -3.98 -3.42 -3.13
<u>Có hệ số chặn:</u> $P - value_{intercept} > 0.05$ → Intercept không có ý nghĩa thống kê	> summary(ur.df(diff(Price),type = "drift",lag = 0)) Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) 4.32944 28.68249 0.151 0.88

	z.lag.1 -0.89231 0.04471 -19.956 <2e-16 *** Value of test-statistic is: -19.9558 199.1177 Critical values for test statistics: 1pct 5pct 10pct tau2 -3.44 -2.87 -2.57
<u>Không có hệ số chặn:</u> $\lvert\tau_{stat}\rvert < \lvert\tau_{\alpha}\rvert$ ($\tau_{stat} = -19.975$) → Chuỗi không có nghiệm đơn vị → Chuỗi dừng	> summary(ur.df(diff(Price),type = "none",lag = 0)) Coefficients: Estimate Std. Error t value Pr(>|t|) z.lag.1 -0.89227 0.04467 -19.98 <2e-16 *** Value of test-statistic is: -19.975 Critical values for test statistics: 1pct 5pct 10pct tau1 -2.58 -1.95 -1.62

<u>Kết luận:</u> Sai phân bậc 1 của chuỗi giá PVC dừng ở mọi mức ý nghĩa, chuỗi không có hệ số chặn.

Bảng 14: Kiểm định tính dừng chuỗi log_return

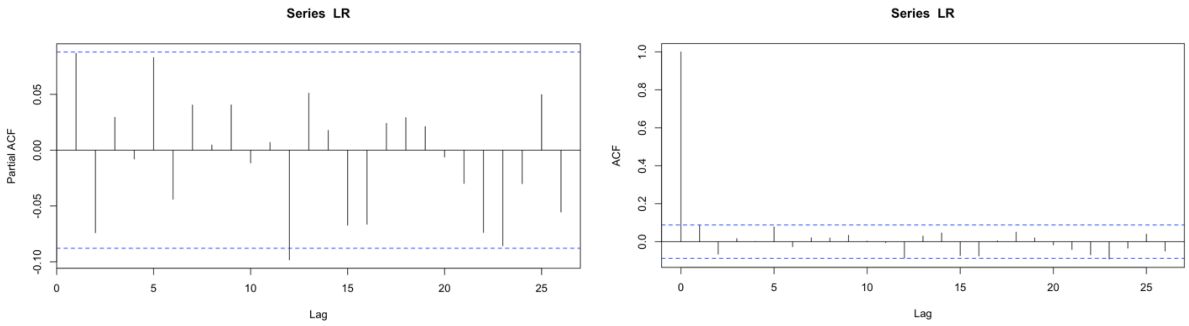
<u>Có hệ số chặn và xu thế:</u> $P - value_{intercept} > 0.05$ → Intercept không có ý nghĩa thống kê $P - value_{tt} > 0.05$ → Xu thế không có ý nghĩa thống kê	> summary(ur.df(LR,type="trend",lag=0)) Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) 8.111e-03 3.948e-01 0.021 0.984 z.lag.1 -9.132e-01 4.485e-02 -20.362 <2e-16 *** tt 9.632e-05 1.377e-03 0.070 0.944 Value of test-statistic is: -20.3615 138.1987 207.2976 Critical values for test statistics: 1pct 5pct 10pct tau3 -3.98 -3.42 -3.13
<u>Có hệ số chặn:</u> $P - value_{intercept} > 0.05$ → Intercept không có ý nghĩa thống kê	> summary(ur.df(LR,type="drift",lag=0)) Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) 0.03204 0.19691 0.163 0.871 z.lag.1 -0.91314 0.04480 -20.382 <2e-16 ***

	Value of test-statistic is: -20.382 207.7141 Critical values for test statistics: 1pct 5pct 10pct tau2 -3.44 -2.87 -2.57
<u>Không có hệ số chặn:</u> $\lvert\tau_{stat}\rvert < \lvert\tau_{\alpha}\rvert$ ($\tau_{stat} = -19.975$) → Chuỗi không có nghiệm đơn vị → Chuỗi dừng	> summary(ur.df(LR,type="none",lag=0)) Coefficients: Estimate Std. Error t value Pr(>|t|) z.lag.1 -0.91309 0.04476 -20.4 <2e-16 *** Value of test-statistic is: -20.4015 Critical values for test statistics: 1pct 5pct 10pct tau1 -2.58 -1.95 -1.62

<u>Kết luận:</u> Chuỗi log – return của PVC dừng ở mọi mức ý nghĩa, chuỗi không có hệ số chặn. Ta sử dụng chuỗi log – return để tiếp tục thực hiện các kiểm định và đánh giá còn lại.

3.2. Xác định bậc cho mô hình ARIMA(p,d,q)

Ta xác định bậc của AR(p) và MA(q) bằng cách sử dụng hàm PACF() và ACF() trên chuỗi log – return của PVC:



Ta thấy tại độ trễ 1 và 5, biểu đồ PACF có giá trị nằm ngoài khoảng tin cậy; tại độ trễ 1, 2 và 5, biểu đồ ACF có giá trị nằm ngoài khoảng tin cậy.

Như vậy, kết hợp bậc của p và q và giá trị d = 0 do chuỗi log – return là chuỗi dừng, ta có một số kịch bản như sau:

- ARIMA(1,0,1)
- ARIMA(1,0,2)
- ARIMA(1,0,5)
- ARIMA(5,0,1)
- ARIMA(5,0,2)
- ARIMA(5,0,5)

* ***Thực hiện các mô hình ARIMA***

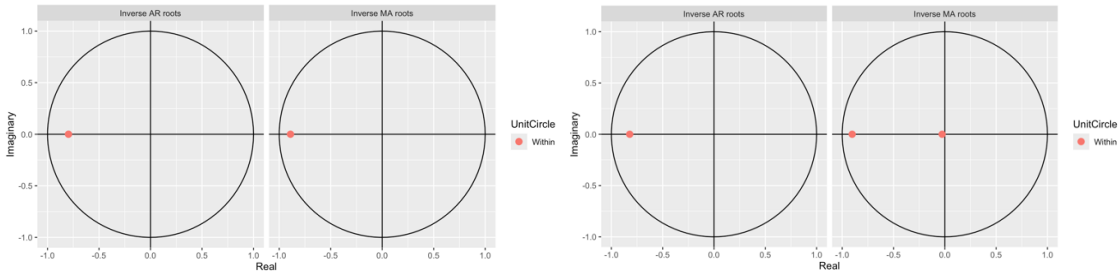
Một trong những tiêu chí thường được sử dụng để lựa chọn mô hình ARIMA đó là chỉ số AIC (Akaike Information Criteria). Tiêu chí thông tin này là một công cụ ước tính lỗi dự báo. Giá trị của AIC càng nhỏ thì mô hình của chúng ta càng phù hợp.

Suy ra, ta lựa chọn 2 mô hình ARIMA(1,0,1) và ARIMA(1,0,2) để tiếp tục đánh gía và kiểm định.

STT	Mô hình	AIC
1	ARIMA(1,0,1)	2877.18
2	ARIMA(1,0,2)	2879.01
3	ARIMA(1,0,5)	2883.36
4	ARIMA(5,0,1)	2883.46
5	ARIMA(5,0,2)	2885.05
6	ARIMA(5,0,5)	2885.48

* ***Đánh giá mô hình***

- *Xem tính dừng qua nghiệm nghịch đảo*



Từ hình vẽ ta thấy nghiệm nghịch đảo nằm trong vòng tròn đơn vị.

$\Longrightarrow$ Suy ra, mô hình cho kết quả dừng.

- *Kiểm định tính nhiễu trắng của phần dư*

Ta tiến hành kiểm định phần dư u_t của mô hình có phải một nhiễu trắng hay không bằng cách sử dụng kiểm định Ljung – Box.

Cặp giả thuyết: $\begin{cases} H_0\text{: Không tồn tại hiện tượng tự tương quan} \\ H_1\text{: Tồn tại hiện tượng tự tương quan} \end{cases}$

Bảng 15: Kiểm định tính nhiễu trắng của phần dư

STT	Mô hình	Kết quả kiểm định	Kết luận
1	ARIMA(1,0,1)	$P - Value = 0.9331$	Với $P - Value > 0.05$, chưa bác bỏ H_0. Suy ra, không tồn tại hiện tượng tự tương

2	ARIMA(1,0,2)	$P - Value$ = 0.8988	quan trong phần dư của mô hình. Vì vậy, phần dư của mô hình là nhiễu trắng.

3.2. Dự báo giá cổ phiếu

Bảng 16: Dự báo chuỗi log – return trong 10 ngày đầu tiên tại năm 2024 với 2 mô hình

STT	**ARIMA(1,0,1)**		**ARIMA(1,0,2)**	
	Reality	**Forecast**	**Reality**	**Forecast**
499	-0.6688988	0.002814945	-0.6688988	-0.002887461
500	0.6688988	0.048322645	0.6688988	0.051187000
501	0.0000000	0.012075882	0.0000000	0.007386738
502	0.0000000	0.040946333	0.0000000	0.043363224
503	-0.6688988	0.017951095	-0.6688988	0.013813000
504	-2.0339684	0.036266742	-2.0339684	0.038084846
505	0.0000000	0.021678377	0.0000000	0.018148533
506	0.6825965	0.033297971	0.6825965	0.034523743
507	-2.0619287	0.024042995	-2.0619287	0.021073537
508	0.0000000	0.031414559	0.0000000	0.032121214
	RMSE = 1.018337		**RMSE** = 1.017194	

Suy ra chuỗi giá dự báo PVC trong 10 ngày đầu tiên của năm 2024 như sau:

Bảng 17: Chuỗi giá dự báo PVC trong 10 ngày đầu tiên của năm 2024

STT	**ARIMA(1,0,1)**		**ARIMA(1,0,2)**	
	Reality	**Forecast**	**Reality**	**Forecast**
499	15000	15042.28	15000	14956.75
500	14900	15787.02	14900	15742.27
501	15000	15978.81	15000	15858.99
502	15000	16646.67	15000	16561.81

503	14900	16948.19	14900	16792.17
504	14600	17574.13	14600	17444.03
505	14600	17959.27	14600	17763.51
506	14700	18567.34	14700	18387.48
507	14400	19019.17	14400	18779.08
508	14400	19626.13	14400	19392.08
	MAPE = 0.1755091		**MAPE** = 0.165477	

Kết luận: Sai số giữa giá trị dự báo so với giá trị thực tế ở mô hình ARIMA(1,0,1) bằng 17,55% giá trị thực tế, ở mô hình ARIMA(1,0,2) bằng 16,55% giá trị thực tế. Chỉ số này càng nhỏ thì mô hình dự báo càng khớp với giá trị thực tế. Vì vậy, mô hình ARIMA(1,0,2) sẽ cho kết quả dự báo đáng tin cậy hơn so với mô hình ARIMA(1,0,1).

4. Mô hình ARCH/GARCH dự báo rủi ro

Mô hình GARCH kết hợp các thành phần auto regressive (AR) và moving average (MA) để mô hình hóa sự phụ thuộc của phương sai có điều kiện vào các giá trị đã qua xử lý cũng như các giá trị độ biến động trước đó. Một mô hình GARCH(p, q) có thể được biểu diễn như sau:

$$\sigma_t^2 = \omega + \beta_1\sigma_{t-1}^2 + \cdots + \beta_p\sigma_{t-p}^2 + \alpha_1 u_{t-1}^2 + \alpha_q u_{t-q}^2 + v_t$$

Khi $\left(\sum_i \beta_i + \sum_j \alpha_j\right) < 1$: Phương sai dừng, tính được phương sai không điều kiện

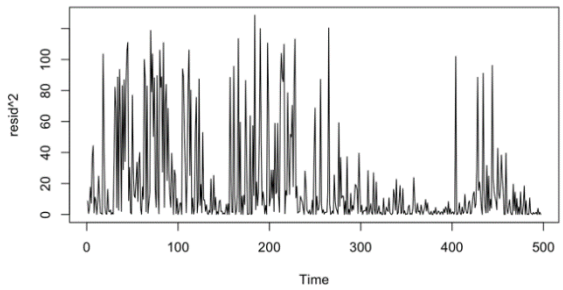
$$\sigma^2 = \frac{\alpha_0}{1 - \left(\sum_i \beta_i + \sum_j \alpha_j\right)}$$

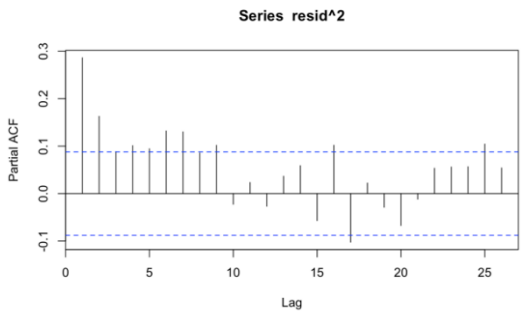
Kiểm định hiệu ứng ARCH – LM thực hiện kiểm tra xem các giá trị trong chuỗi dữ liệu có biến đổi không đồng nhất qua thời gian hay không. Phương pháp này sử dụng một mô hình auto regressive để ước lượng và tách ra thành phần không đồng nhất, sau đó kiểm tra tính chất ngẫu nhiên của phần dư (residuals) để xác định sự tồn tại của hiệu ứng ARCH. Mô hình ARCH(q) có thể được biểu diễn như sau:

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_q u_{t-q}^2 + v_t$$

a. Tính volatility và xem bậc tự hồi quy

Để ước lượng mô hình ARCH(q) cho chuỗi log – return (chuỗi dừng), phải tìm được bậc q thông qua việc tính volatility cho chuỗi và xem bậc tự hồi quy của phần dư (residuals) bằng PACF:





b. Thực hiện hồi quy phụ kiểm định ARCH(1):

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + v_t$$

Cặp giả thuyết: $\begin{cases} H_0: \alpha_1 = 0 \\ H_1: \alpha_1 \neq 0 \end{cases} \Rightarrow \begin{cases} H_0: \text{Không có ARCH}(1) \\ H_1: \text{Có ARCH}(1) \end{cases}$

Kết quả kiểm định	Kết luận
Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) 18.8352 1.3049 14.435 < 2e-16 *** L(resid) -1.2243 0.3005 -4.074 5.39e-05 ***	Vì – 1.2243 < 0, suy ra có ARCH(1).

Phương trình residuals của ARCH(1): $u_t^2 = 18.8352 - 1.2243u_{t-1}^2 + v_t$

c. Thực hiện kiểm định tự động ARCH(1):

Cặp giả thuyết: $\begin{cases} H_0: \text{Không có hiệu ứng ARCH} \\ H_1: \text{Có hiệu ứng ARCH} \end{cases}$

Kết quả kiểm định	Kết luận
ARCH LM-test; Null hypothesis: no ARCH effects data: resid Chi-squared = 40.694, df = 1, p-value = 1.78e-10	Vì P – Value < 0.05, bác bỏ H_0, suy ra có ARCH(1).

d. Ước lượng ARCH(1)/GARCH(0,1):

- Kiểm định Jacque – bera:
  Cặp giả thuyết: $\begin{cases} H_0: \text{Không có phân phối chuẩn} \\ H_1: \text{Có phân phối chuẩn} \end{cases}$

- Kiểm định Ljung – box:

Cặp giả thuyết: $\begin{cases} H_0: \text{Không có hiện tượng tự tương quan} \\ H_1: \text{Có hiện tượng tự tương quan.} \end{cases}$

Kết quả kiểm định	Kết luận
Coefficient(s): Estimate Std. Error t value Pr(>|t|) a0 13.06438 0.92600 14.108 < 2e-16 *** a1 0.31062 0.08794 3.532 0.000412 *** --- Diagnostic Tests: Jarque Bera Test data: Residuals X-squared = 12.323, df = 2, p-value = 0.002109 Box-Ljung test data: Squared.Residuals X-squared = 0.33183, df = 1, p-value = 0.5646	**Kiểm định Jacque – bera:** + P – Value < 0.05, bác bỏ H_0. Suy ra, phần dư của mô hình GARCH(0,1) là phân phối chuẩn. **Kiểm định Ljung – box:** + P – Value > 0.05, chưa bác bỏ H_0. Suy ra, phần dư của mô hình GARCH(0,1) không xuất hiện hiện tượng tự tương quan.

Phương trình GARCH(0,1): $\sigma_t^2 = 13.06438 + 0.31062u_{t-1}^2 + v_t$

Ta có: $\sigma^2 = \frac{13.06438}{1-0.31062} = 18.95 \geq 1 \Rightarrow$ Phương sai không dừng

e. Dự báo sự biến động (volatility) của chuỗi log – return:

Bảng 18: Dự báo sự biến động (volatility) của chuỗi log – return

STT	499	500	501	502	503	504	505	506	507	508
Volatility	3.561	4.105	4.27	4.323	4.341	4.346	4.348	4.349	4.349	4.349

C. PHÂN TÍCH TỔNG HỢP

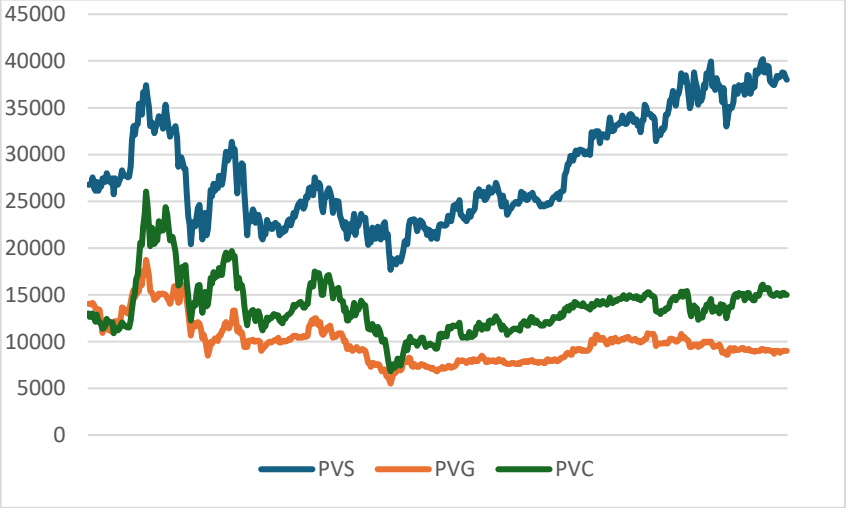
I. Tổng quan

1. Sự biến động chung

Trong những năm gần đây, ba mã cổ phiếu PVS, PVG, và PVC đã trải qua sự biến động đáng chú ý trên thị trường chứng khoán. Cổ phiếu PVS của Công ty Cổ phần Dịch vụ Kỹ thuật Dầu khí thường phản ánh tình hình của ngành công nghiệp dầu khí, và biến động của giá dầu thế giới thường có ảnh hưởng lớn đến giá cổ phiếu này. Trong những năm gần đây, biến động của giá dầu và tình hình kinh doanh của PVS có thể đã gây ra những biến động đáng kể trên thị trường chứng khoán. Cổ phiếu PVG của Công ty Cổ phần Gas Petrolimex cũng đã trải qua sự biến động trong bối cảnh thị trường năng lượng và khí đốt tự nhiên, cũng như các yếu tố liên quan đến quản lý và kinh doanh của công ty. Cổ phiếu PVC của Tổng Công ty Cổ phần Dầu Việt Nam cũng không ngoại lệ, và biến động của giá dầu thế giới và các yếu tố nội bộ của công ty đều có thể đã ảnh hưởng đến giá cổ phiếu này trong những năm gần đây. Tổng thể, biến động của ba mã cổ phiếu PVS, PVG, và PVC thường phản ánh sự biến động của ngành công nghiệp dầu khí và năng lượng, cũng như các yếu tố nội bộ của từng công ty trong thời gian gần đây.

2. Sự tương quan

Vì đều thuộc lĩnh vực dầu khí nên cả ba mã cổ phiếu PVS, PVG, PVC có sự tương quan cao và đều tương quan dương. Điều này thể hiện rằng khi một trong ba mã cổ phiếu này tăng lên thì hai mã còn lại cũng có xu hướng sẽ tăng. Số liệu thực tế chỉ ra rằng PVC có mối tương quan dương trên 50% với PVS và trên 77% với PVG. Trên đồ thị thể hiện xu hướng lên xuống khá tương đồng nhau của ba mã cổ phiếu này, đều biến động trong năm 2022 và tăng trong năm 2023, tuy nhiên chỉ có mã PVS là tăng vượt mức so với đầu năm 2022 còn PVG và PVC thì tăng chậm, phục hồi dần về mức giá đầu năm 2022.



II. Phân tích đồng tích hợp

1. Lý thuyết

Đồng tích hợp là hiện tượng hai hay nhiều chuỗi thời gian có tồn tại một số nguyên nhân gốc làm chúng có xu hướng biến động tương đồng nhau về dài hạn ngay cả khi nếu nhìn riêng lẻ thì chúng chỉ là bước ngẫu nhiên.

- Kiểm định với tiêu chí "Trace"

$$\begin{cases} H_0: \ r = 0: \ \textit{không có mối quan hệ đồng tích hợp} \\ H_1: \ r \neq 0: \ \textit{có ít nhất một mối quan hệ đồng tích hợp} \end{cases}$$

- Kiểm định với tiêu chí "Eigen Value"

Tiêu chí eigen value thường được sử dụng trong phương pháp kiểm định Johansen để kiểm tra đồng tích hợp giữa các biến trong một hệ thống. Đây là một phương pháp thống kê dựa trên việc kiểm tra eigen value của ma trận cointegration.

Trong kiểm định Johansen, eigen value được sắp xếp theo thứ tự giảm dần và kiểm tra xem số lượng eigen value lớn hơn không (hoặc lớn hơn một ngưỡng cụ thể) là bao nhiêu. Số lượng này cho biết có bao nhiêu vector cointegration tồn tại trong hệ thống biến. Nếu số lượng eigen value lớn hơn không, điều này ngụ ý rằng hệ thống biến có một hoặc nhiều vector cointegration và các biến trong hệ thống này di chuyển cùng nhau theo thời gian.

2. Kiểm định đồng tích hợp giữa các chuỗi giá

+ Kiểm định với tiêu chí Trace: Thống kê kiểm định: LR_{tr}= 25.82; Giá trị tới hạn mức 5% bằng 28.71 => Chưa bác bỏ Ho nên không có mối quan hệ đồng tích hợp giữa 3 chuỗi giá.	Test type: trace statistic , with linear trend Eigenvalues (lambda): [1] 0.030198791 0.019559627 0.001645195 Values of teststatistic and critical values of test: test 10pct 5pct 1pct r <= 2 | 0.82 6.50 8.18 11.65 r <= 1 | 10.61 15.66 17.95 23.52 r = 0 | 25.82 28.71 31.52 37.22
+ Kiểm định với tiêu chí Eigen Value: Thống kê kiểm định: LR_{tr}= 15.21; Giá trị tới hạn mức 5% bằng 21.07 => Chưa bác bỏ Ho nên không có mối quan hệ đồng tích hợp giữa 3 chuỗi giá.	Test type: maximal eigenvalue statistic (lambda max) , with linear trend Eigenvalues (lambda): [1] 0.030198791 0.019559627 0.001645195 Values of teststatistic and critical values of test: test 10pct 5pct 1pct r <= 2 | 0.82 6.50 8.18 11.65 r <= 1 | 9.80 12.91 14.90 19.19 r = 0 | 15.21 18.90 21.07 25.75

III. Phân tích VAR

1. Xây dựng VAR giữa các chuỗi log-return, kiểm định đánh giá

Vector Autoregressive Model (VAR): Mô hình dùng để đánh giá và phân tích ảnh hưởng các biến tác động qua lại với nhau, không còn biến phụ thuộc và độc lập riêng biệt. Ở phần này, ta xây dựng mô hình VAR cho chuỗi log_return của 3 mã cổ phiếu PVS, PVG và PVC.Trong phần kiểm định tính dừng ở trên thì 3 chuỗi này là chuỗi dừng.

- Lựa chọn bậc trễ cho mô hinh VAR:

```
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     1      1      1      1
```

Kết quả trên cho thấy phần mềm R chọn bậc trễ cho mô hình VAR là bậc 1. Vậy sẽ phân tích và dự báo trên mô hình VAR với trễ bậc 1.

- Ước lượng mô hình VAR(1)

Ta được kết quả ước lượng như sau:

$$\boldsymbol{r_{PVS} = 0.07462 - 0.09121 r_{t-1}^{PVS} - 0.01082 r_{t-1}^{PVG} + 0.05368 r_{t-1}^{PVC} + u_{1t}}$$
$$\boldsymbol{r_{PVG} = -0.0883 + 0.01546 r_{t-1}^{PVG} - 0.0414 r_{t-1}^{PVS} + 0.11842 r_{t-1}^{PVC} + u_{2t}}$$
$$\boldsymbol{r_{PVC} = 0.04475 + 0.11826 r_{t-1}^{PVC} - 0.10211 r_{t-1}^{PVS} + 0.06944 r_{t-1}^{PVG} + u_{3t}}$$

Sau khi ước lượng , tiến hành kiểm định mô hình dựa trên hiện tượng tự tương quan của phần dư . Cặp giả thuyết:

$$\begin{cases} H_o: \textit{Phần dư không có tự tương quan} \\ H_1: \textit{Phần dư có tự tương quan} \end{cases}$$

Kết quả kiểm định: $p_{value} = 0.002864 < 0.05$ $\Rightarrow$ Bác bỏ H_o do đó phần dư có tự tương quan. Vì vậy việc sử dụng mô hình VAR(1) để dự báo thì cho kết quả không đáng tin cậy.	Portmanteau Test (asymptotic) data: Residuals of VAR object var1 Chi-squared = 184.83, df = 135, p-value = 0.002864

2. Dự báo các chuỗi log-return từ mô hình VAR

```
$rPVS
            fcst
 [1,] 0.14634656
 [2,] 0.06859736
 [3,] 0.07142732
 [4,] 0.07104520
 [5,] 0.07102134
 [6,] 0.07101682
 [7,] 0.07101608
 [8,] 0.07101597
 [9,] 0.07101595
[10,] 0.07101595
```

```
$rPVG
             fcst
 [1,] -0.05627374
 [2,] -0.08095148
 [3,] -0.08810349
 [4,] -0.08877545
 [5,] -0.08891558
 [6,] -0.08893488
 [7,] -0.08893800
 [8,] -0.08893848
 [9,] -0.08893855
[10,] -0.08893856
```

```
$rPVC
            fcst
 [1,] 0.12504487
 [2,] 0.04068766
 [3,] 0.03693638
 [4,] 0.03570713
 [5,] 0.03555410
 [6,] 0.03552871
 [7,] 0.03552483
 [8,] 0.03552423
 [9,] 0.03552414
[10,] 0.03552412
```

Bảng 19: So sánh kết quả dự báo

	PVS			**PVG**			**PVC**		
2024	**VAR**	**ARIMA**	**Thực tế**	**VAR**	**ARIMA**	**Thực tế**	**VAR**	**ARIMA**	**Thực tế**
01/02	43988.7	40753.74	37900	8507.523	8293.41	9000	16997.99	14956.75	14900
01/03	40912.39	50190.64	38200	8300.146	8228.622	9000	15622.9	15742.27	15000
01/04	41135.74	39461.58	38300	8332.562	8228.622	9100	15564.41	15858.99	15000
01/05	41012.66	46965	38200	8326.964	8320.051	9100	15545.28	16561.81	15000
01/08	40474.88	38142.9	37700	8234.305	8320.051	9000	15439.29	16792.17	14900
01/09	39508.45	38914.64	36800	8234.146	8228.622	9000	15128.04	17444.03	14600
01/10	39186.34	35975.13	36500	8234.121	8228.622	9100	15127.99	17763.51	14600
01/11	39937.86	40294.78	37200	8325.607	8320.051	9100	15127.98	18387.48	14700
01/12	39.508.42	38705	36800	8234.116	8320.051	9000	14920.74	18779.08	14400
01/15	39508.42	39533.86	36800	8142.626	8228.622	8900	14920.74	19392.08	14400
MAPE	0.075	0.08		0.099	0.0917		0.044	0.165	

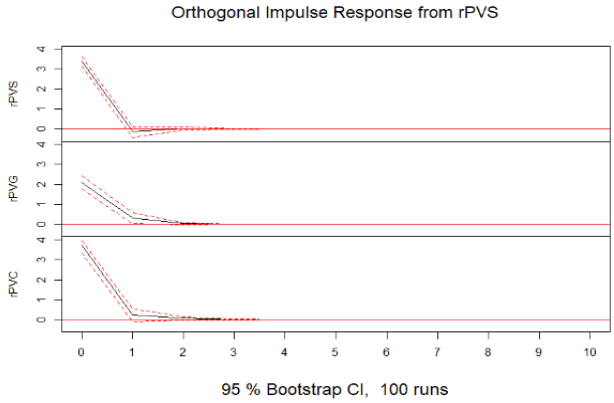
Từ bảng so sánh kết quả dự báo trên, có thể thấy rằng kết quả sai số dự báo MAPE ở mô hình VAR cho ba chuỗi đều tốt hơn khi dùng mô hình ARIMA để dự báo. Vì mô hình VAR được

thiết kế để xử lý chuỗi thời gian đa biến, các biến thường liên quan đến nhau, vây nên khi áp dụng VAR để dự báo giá 10 ngày đầu năm 2024 cho ba chuỗi giá cổ phiêu cùng nhóm ngành dầu khí này sẽ cho sai số thấp hơn.

3.Hàm phản ứng và phân rã phương sai

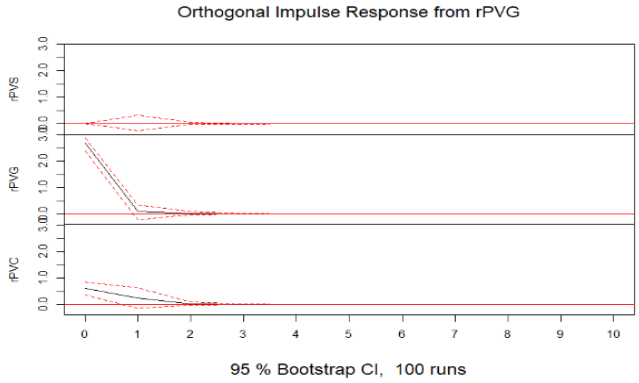
Hàm phản ứng

Khi phân tích hàm phản ứng, cho ta biết khi shock với một biến thì các biến còn lại trong mô hình VAR sẽ phản ứng thế nào.



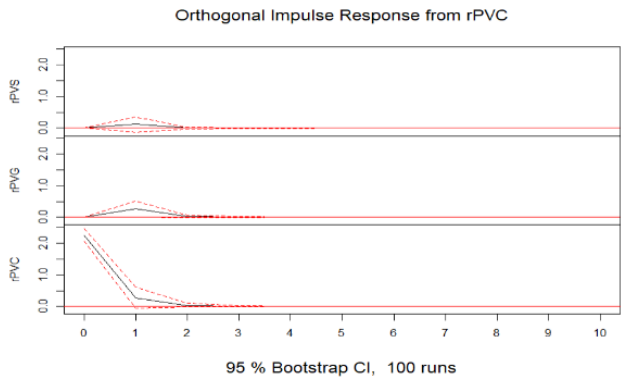
Đồ thị hàm phản ứng của chuỗi log-return mã PVS

Ta thấy log-return của PVS đều ảnh hưởng bởi shock của PVG, PVC và cả chính nó. Nhưng tác động của cú shock đó cũng chỉ tồn tại khoảng một kỳ sau và tắt dần.



Đồ thị hàm phản ứng của chuỗi log-return mã PVG

Với chuỗi log-return của mã PVG, nó chỉ phản ứng với shock của chính nó và log-return của PVC một kỳ sau. Còn với log-return của PVG thì gần như là không phản ứng.



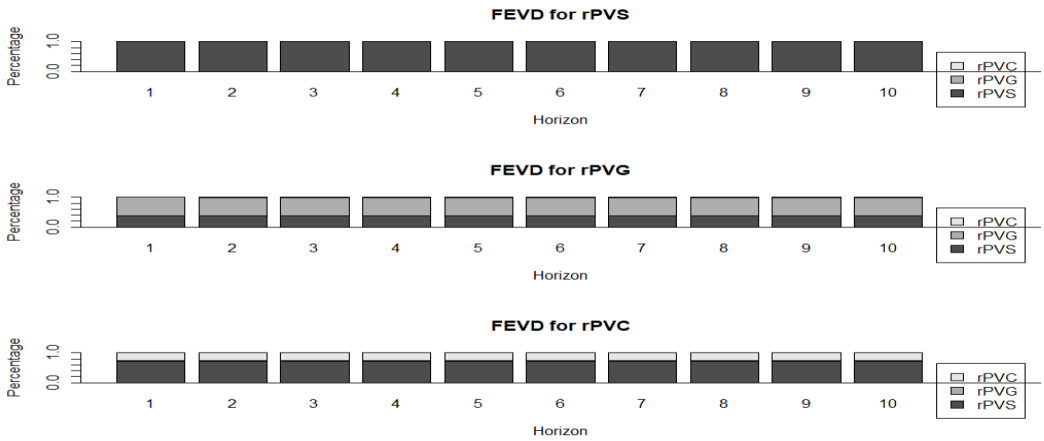
Đồ thị hàm phản ứng của chuỗi log-return mã PVC

Như hai chuỗi trên, log-return của PVC cũng phản ứng với shock của chính nó một kỳ trước. Với các shock một kỳ sau của hai chuỗi log-return còn lại thì có phản ứng nhưng không mạnh và đều mất đi sau hai kỳ với biên độ dao động hẹp.

Do đó dù có chịu ảnh hưởng bởi cú shock thì ba chuỗi log-return của ba mã cổ phiếu cũng chỉ ảnh hưởng trong vòng 1 kỳ hay tương đương với 1 ngày vì đây là số liệu được lấy theo ngày.

Phân rã phương sai

Tiếp theo ta sẽ phân tích phân rã phương sai của mô hình. Phân rã phương sai của sai số dự báo cho ta biết sự biến động một biến thời gian được giải thích bởi bao nhiêu phân trăm sự biến động của biến khác, giúp trong việc tìm ra cách để cải thiện khả năng dự báo của một chuỗi bất kỳ.



Biểu đồ phân rã phương sai

Theo đồ thị, sự biến động của rPVS gần như chỉ do mình chuỗi này giải thích.

Với sự biến động của rPVG được giải thích bởi trên 37.9% sự biến động rPVS và hơn 61.09% sự biến động rPVG trong 10 kỳ. Và sự biến động của rPVC chỉ giải thích được 0.6% sự biến động của rPVG.

Còn với rPVC, sự biến động của nó lại được giải thích bởi hơn 71.8% sau một kỳ và hơn 71.38% từ 9 kỳ còn lại của sự biến động rPVS. Còn lại là được giải thích bởi 26.2% từ rPVC và hơn 2% từ rPVG trong 10 kỳ.

IV. Kết luận

Sau khi phân tích 3 cổ phiếu PVD, PVS, PVC thuộc nhóm ngành dầu khí dựa trên 2 chuỗi doanh thu thuần và giá đóng cửa của cổ phiếu, ta có những nhận xét sau đây:

- Thứ nhất, vì đều bị ảnh hưởng bởi sự biến động của nền kinh tế, nên trong các giai đoạn nền kinh tế gặp khủng hoảng doanh thu thuần của cả 3 công ty đều có xu hướng giảm mạnh, tuy nhiên trong những năm gần đây, khi ngành dầu khí ít bị tác động một cách đáng kể bởi các yếu tố bên ngoài, doanh thu thuần đều duy trì ở một ngưỡng ổn định.

- Thứ hai, về dự báo giá cổ phiếu với mô hình ARIMA, cả 3 mô hình đều cho kết quả lệch so với thực tế. Để giải thích cho điều này, có thể do mô hình ARIMA chưa bao gồm đủ yếu tố tác động với sự thay đổi về giá của 3 công ty trên, hơn nữa trong mô hình còn bao gồm sai số, phần sai số này phụ thuộc vào rất nhiều các yếu tố bên ngoài tính kinh tế.

- Thứ ba, 3 chuỗi PVC, PVS và PVD không đồng tích hợp với nhau. Việc 3 chuỗi này không đồng tích hợp là điều hợp lý, vì giá cổ phiếu bị ảnh hưởng bởi các yếu tố kỹ thuật khác nhau, bên cạnh đó, các yếu tố như tâm lý thị trường hay thông tin không chính thức cũng có thể tạo ra các biến động không đồng nhất giữa các chuỗi giá cổ phiếu.

- Thứ tư, sau khi áp dụng mô hình VAR, 3 chuỗi giá cổ phiếu cho sai số thấp hơn so với mô hình ARIMA, vì mô hình VAR vốn được thiết kế để xử lý chuỗi thời gian đa biến, tuy nhiên, như đã kiểm định ở trên, mô hình VAR được lựa chọn cũng không cho kết quả đáng tin cậy.

- Thứ năm, sau khi phân rã phương sai và kiểm tra hàm phản ứng, ta có thể nhận xét rằng cả 3 chuỗi giá cổ phiếu thường chỉ phản ứng với shock của chính nó, còn đối với shock của 2 chuỗi cổ phiếu còn lại trong mô hình, nếu có xuất hiện phản ứng cũng chỉ tồn tại trong vòng 1 kỳ và sớm tắt.