

Modelling and predicting stock returns: Empirical evidence from Vietnam

Cao Tan Binh*, Le Mong Huyen, Pham Nguyen Dinh Tuan

Faculty of Economics and Accounting, Quy Nhon University, Vietnam

Received: 21/03/2020; Accepted: 21/04/2020

ABSTRACT

In this paper, we analyze properties of the VN-Index's daily closing price data series from January 3, 2001 to March 13, 2020 by using the estimation method of time series model. The analysis and testing results show that shocks cause a persistent effect, leading to stationariness and long-term memory of this time series; and the ARFIMA (1,0.2,4)-FIGARCH (1,0.16,1) model was chosen to predict returns of above data series.

Keywords: *ARFIMA-FIGARCH, long-term memory, closing price, prediction.*

**Corresponding author.*

Email: caotanbinh@qnu.edu.vn

Mô hình và dự báo lợi nhuận cổ phiếu: Bằng chứng thực nghiệm từ Việt Nam

Cao Tấn Bình*, Lê Mộng Huyền, Phạm Nguyễn Đình Tuấn

Khoa Kinh tế & Kế toán, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 21/03/2020; Ngày nhận đăng: 21/04/2020

TÓM TẮT

Trong bài báo này, chúng tôi phân tích tính chất của chuỗi dữ liệu giá đóng cửa theo ngày của chỉ số VN-Index từ 3/1/2001 đến 13/3/2020 bằng phương pháp mô hình chuỗi thời gian. Kết quả phân tích và kiểm định cho thấy các cú sốc gây ra hiệu ứng dai dẳng, dẫn đến chuỗi dữ liệu này không dừng và có trí nhớ lâu; và mô hình ARFIMA(1,0.2,4)-FIGARCH(1,0.16,1) được lựa chọn để dự báo chuỗi tỷ suất lợi nhuận của chuỗi dữ liệu trên.

Từ khóa: ARFIMA-FIGARCH, trí nhớ lâu, giá đóng cửa, dự báo.

1. GIỚI THIỆU

Bài báo của Tim Bollerslev và Hans Ole Mikkelsen xuất bản năm 1996 về mô hình FIGARCH đã đặt nền tảng cho những mô hình có trí nhớ lâu, và được áp dụng một cách có hiệu quả trong việc kiểm định và dự báo các quá trình kinh tế và tài chính.¹ Trong những năm qua, các nhà nghiên cứu đã tiếp tục thử nghiệm đối với các mô hình chuỗi thời gian khác nhau để mô hình hóa dữ liệu của thế giới thực. Gần đây, có nhiều nghiên cứu được thực hiện nhằm khai thác những ứng dụng của mô hình lai ARFIMA-FIGARCH. Michel Beine, Sébastien Laurent và Christelle Lecourt đã ước lượng các mô hình FIGARCH cho bốn tỷ giá dựa trên đồng đôla Mỹ (DEM, FRF, YEN và GBP) và cuối cùng nhận thấy rằng DEM, FRF, YEN thỏa mãn mô hình ARFIMA-FIGARCH, ngoại trừ GBP.² Năm 2002, Richard T. Baillie và cộng sự đã kiểm chứng tính trí nhớ lâu của các cú sốc tài chính.³ Năm 2003, các công trình^{4,5} xem xét tính hiệu quả của thị trường chứng

khoán Nhật Bản đối với mã cổ phiếu Nikkei 225 và tỷ giá hối đoái cũng dựa trên đồng đôla Mỹ bởi mô hình ARFIMA-FIGARCH. Năm 2004, Wilfredo Palma và Mauricio Zavallos đã phân tích cấu trúc tự tương quan của chuỗi thời gian bậc hai;⁶ Jurgen A. Doornik and Marius Ooms đã sử dụng mô hình ARFIMA suy diễn và dự báo lạm phát của Mỹ và Anh.⁷ Vấn đề lạm phát thuộc các nước khu vực châu Âu tiếp tục được nghiên cứu vào năm 2005, và tính trí nhớ lâu lại xuất hiện trong mô hình của Christian Conrad và Menelaos Karanasos.⁸ Năm 2007, nhiều công trình khảo sát về tính không dừng và trí nhớ lâu của chuỗi thời gian và áp dụng phân tích cho một số mã cổ phiếu quốc tế như SP 500 (Mỹ), Nikkei (Nhật Bản), PSI 20 (Bồ Đào Nha), CAC 40 (Pháp), DAX 30 (Đức), FTSE 100 (Anh), IBEX 35 (Tây Ban Nha) và MIB 30 (Ý).⁹⁻¹⁰ Năm 2009, công trình của P. Bagavathi Sivakumar và Dr. V. P. Mohandas sử dụng mô hình ARFIMA-FIGARCH để dự đoán lợi tức cổ phiếu của thị trường cổ phiếu

*Tác giả liên hệ chính.

Email: caotanbinh@qnu.edu.vn

Ấn Độ¹¹. Trong khoảng gần một thập niên trở lại đây, hàng loạt các công trình được công bố trên các tạp chí uy tín, tiếp tục sử dụng mô hình lai ARFIMA-FIGARCH để phân tích và dự đoán các quá trình chuỗi thời gian trong kinh tế và tài chính.^{12-22,25,27}

Ở Việt Nam, trong những năm gần đây, một số tác giả đã sử dụng mô hình ARIMA hoặc mô hình GARCH trong các công trình nghiên cứu của mình: Tác giả Võ Xuân Vinh và cộng sự với công trình "Volatility in stock return series of Vietnam stock market",¹⁷ bài báo nghiên cứu các đặc điểm của sự biến động lợi nhuận của VNIndex qua việc sử dụng mô hình GARCH và nghiên cứu sự hiện diện của các điểm gãy cấu trúc trong phương sai của chuỗi lợi nhuận đó thông qua việc sử dụng thuật toán ICSS (iterated cumulative sums of squares). Sử dụng dữ liệu trong một khoảng thời gian dài, mô hình GARCH và GARCH-M tỏ ra hiệu quả trong việc mô tả các đặc điểm của lợi nhuận chứng khoán hàng ngày. Kết quả nghiên cứu có ý nghĩa quan trọng đối với các nhà đầu tư trong việc ra các quyết định liên quan đến việc định giá cổ phiếu, quản lý danh mục đầu tư, phòng ngừa rủi ro và dự báo. Ngoài ra, nó cũng có ích cho các nhà làm chính sách trong việc thực hiện và ban hành các chính sách tài chính; tác giả Ngô Văn Toàn và cộng sự với công trình "Dự báo giá vàng Việt Nam sử dụng mô hình GARCH",²³ bài báo đã sử dụng các mô hình ARIMA và mô hình GARCH kết hợp với tiêu chuẩn AIC để tìm ra mô hình phù hợp, và nghiên cứu kết luận rằng GARCH(1,2) là mô hình thích hợp hơn để dự báo; tác giả Nguyễn Ngọc Thạch và cộng sự với công trình "Mô hình GARCH trong dự báo sự biến động của giá dầu thế giới",²⁴ nghiên cứu này dự báo sự biến động của giá dầu thế giới. Với dữ liệu chuỗi giá dầu thô WTI, nhóm tác giả thực hiện ước lượng các dạng mô hình GARCH (1,1), EGARCH (1,1), GJR – GARCH (1,1) theo bốn quy luật phân phối khác nhau. Kết quả nghiên cứu cho thấy, mô hình EGARCH (1,1) theo quy luật phân

phối Student-t cho kết quả dự báo tốt nhất. Ngoài ra, kết quả nghiên cứu cũng chỉ ra, mức độ biến động của giá dầu thô trong tương lai có thể được dự báo bằng mức độ biến động giá của nguyên liệu này trong quá khứ đồng thời các cú sốc tăng giảm giá trên thị trường dầu thô có ảnh hưởng tương đối nhỏ đến biến động của giá dầu thô; tác giả Bùi Hữu Phước và cộng sự với công trình "Asset Price Volatility of Listed Companies in the Vietnam Stock Market",²⁶ nghiên cứu đo độ biến động giá cổ phiếu của một số công ty trên thị trường chứng khoán Việt Nam. Các tác giả đã sử dụng các mô hình AR, MA và ARMA kết hợp với ARCH và GARCH để ước lượng giá trị rủi ro (VaR) và các kết quả ước lượng thu được có độ chính xác một cách tương đối.

Ngoài lời giới thiệu, bài báo gồm các nội dung sau đây: Phần 2 trình bày một số mô hình chuỗi thời gian, trong đó nhấn mạnh các mô hình ARIMA-GARCH và ARFIMA-FIGARCH. Phần 3 áp dụng hai mô hình này cho dữ liệu hàng ngày về giá đóng cửa VN-Index và bình luận về kết quả. Cuối cùng, đưa ra một số kết luận và nhận xét ở phần 4.

2. CÁC MÔ HÌNH CHUỖI THỜI GIAN

2.1. Mô hình ARIMA-GARCH

Khi mô hình hóa tính biến động của một quá trình ngẫu nhiên theo thời gian, gọi là chuỗi thời gian, các mô hình tự hồi quy AR và mô hình trung bình trượt MA thường được sử dụng. Việc kết hợp hai mô hình trên và hệ số tích hợp d sẽ tạo thành mô hình ARIMA nếu bậc tích hợp d là số nguyên. Mô hình ARIMA được Box và Jenkins đề xuất vào năm 1970.

Giả sử ε_t là nhiễu trắng Gaussian với phương sai 1. Khi đó a_t là một quá trình ARCH(q) nếu $a_t = \sigma_t \varepsilon_t$, với $\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i a_{t-i}^2$. Quá trình ARCH(q) là không tương quan, có kỳ vọng và phương sai không điều kiện không đổi nhưng phương sai có điều kiện thay đổi. Một hạn chế của các mô hình ARCH(q) là quá trình

độ lệch chuẩn có điều kiện có dao động tần số cao với biến động cao một cách đột ngột.

Mô hình phương sai có điều kiện thay đổi tự hồi quy tổng quát GARCH được giới thiệu bởi Engle và Bollerslev vào năm 1980. Mô hình GARCH(p,q) có dạng $a_t = \sigma_t \varepsilon_t$, trong đó

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2.$$

Các giá trị quá khứ của quá trình σ_t được đưa trở lại vào giá trị hiện tại, sai số chuẩn có điều kiện này có thể tiếp tục dai dẳng hơn quá trình ARCH của những biến động cao hay thấp. Quá trình a_t là không tương quan với kỳ vọng và phương sai dừng và a_t^2 có hàm tự tương quan ACF giống như quá trình ARMA. Các mô hình GARCH là tổng quát của các mô hình ARCH, cho phép mô tả một cách rộng rãi các biến động dai dẳng hơn. Cho a_t là quá trình GARCH(p_G, q_G) và xem a_t như nhiễu trong mô hình ARIMA(p_A, d, q_A). Ta sẽ gọi mô hình như thế là mô hình ARIMA(p_A, d, q_A)-GARCH(p_G, q_G).

2.2. Mô hình ARFIMA-FIGARCH

Mô hình ARIMA được sử dụng để phân tích và dự báo đối với các chuỗi thời gian dừng có bậc tích hợp nguyên. Nếu mô hình ARIMA có bậc tích hợp không nguyên, ta được mô hình bình quân trượt tích hợp phân thứ ARFIMA, còn được gọi là mô hình trí nhớ lâu. Mô hình ARFIMA(p,d,q) lần đầu tiên được phát triển bởi Hosking, Granger và Joyeux vào năm 1980, có dạng $\Phi(B)(1-B)^d X_t = \Theta(B)\varepsilon_t$ với $\Phi(B)$ và $\Theta(B)$ là đa thức của toán tử dịch chuyển B bậc p và q tương ứng, ε_t là nhiễu trắng. Nếu tham số sai phân phân thứ $d \in [0, 0.5)$, chuỗi X dừng hiệp phương sai nhưng hệ số tự tương quan giảm về 0 rất chậm. Nếu $d \in [0.5, 1)$, chuỗi X không dừng nhưng vẫn đảo chiều về giá trị trung bình, cho rằng bất kỳ cú sốc nào ảnh hưởng đến quá trình cuối cùng sẽ biến mất trong thời gian dài. Nếu $d \geq 1$ chuỗi không dừng và không đảo chiều về giá trị trung bình với ảnh hưởng của những cú sốc kéo dài mãi mãi.

Mô hình FIGARCH được Baillie và cộng sự giới thiệu vào năm 1996, có dạng

$\varphi(B)(1-B)^d \varepsilon_t^2 = \omega + [1 - \beta(B)]v_t$, trong đó tham số sai phân phân thứ $d \in [0, 1]$. Các mô hình GARCH và IGARCH là trường đặc biệt của mô hình FIGARCH khi $d = 0$ và $d = 1$ tương ứng. Nếu $d \in (0, 1)$, mô hình FIGARCH được gọi là có trí nhớ lâu, hệ số tự tương quan giảm rất chậm về 0 dưới tác động của các cú sốc.

Chuỗi dữ liệu thời gian $\{X_t, t = 1, 2, \dots\}$ được gọi là tuân theo mô hình ARFIMA với hiệu ứng FIGARCH, ký hiệu

$X_t \sim ARFIMA(p, d, q) - FIGARCH(p_1, d_1, q_1)$, nếu

$$\left(1 - \sum_{i=1}^p \phi_i B^i\right)(1-B)^d X_t = \left(1 + \sum_{i=1}^q \theta_i B^i\right) \varepsilon_t \quad \text{và}$$

$$\varphi(B)(1-B)^{d_1} \varepsilon_t^2 = \omega + [1 - \beta(B)]v_t$$

trong đó $\varepsilon_t \sim N(0, \sigma_t^2)$, $v_t = \varepsilon_t^2 - \sigma_t^2$, $B^i X_t = X_{t-i}$

$$\text{và } (1-B)^d \approx 1 - dB + \frac{d(d-1)}{2!} B^2.$$

Các nhà kinh tế lượng, điển hình là Baillie, Bollerslev và Mikkelsen, đã kiểm chứng và nhận thấy rằng phần lớn các hiện tượng trong kinh tế và tài chính với tần số cao thường thỏa mãn các mô hình ARFIMA với hiệu ứng FIGARCH.

3. KẾT QUẢ THỰC NGHIỆM

3.1. Phân tích và dự báo bằng mô hình ARIMA-GARCH

Xét chuỗi dữ liệu giá đóng cửa theo ngày của chỉ số chứng khoán Việt Nam VN-Index từ ngày 3/1/2001 đến ngày 13/3/2020. Hình bên dưới mô tả biến động của giá đóng cửa Close.



Hình 1. Đồ thị giá đóng cửa

Kết quả Bảng 1 cho thấy chuỗi Close không dừng.

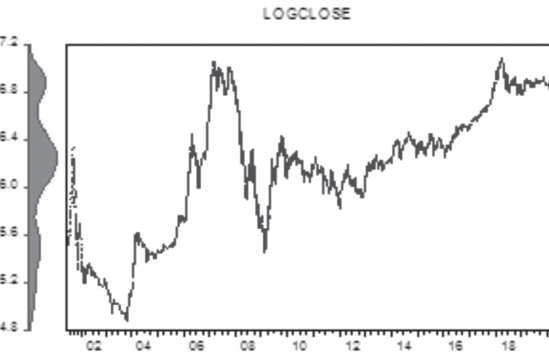
Bảng 1. Kiểm tra tính dừng của chuỗi Close

Null Hypothesis: CLOSE has a unit root
Exogenous: Constant
Lag Length: 4 (Automatic - based on SIC, maxlag=20)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.571952	0.4969
Test critical values:		
1% level	-3.431572	
5% level	-2.861965	
10% level	-2.567039	

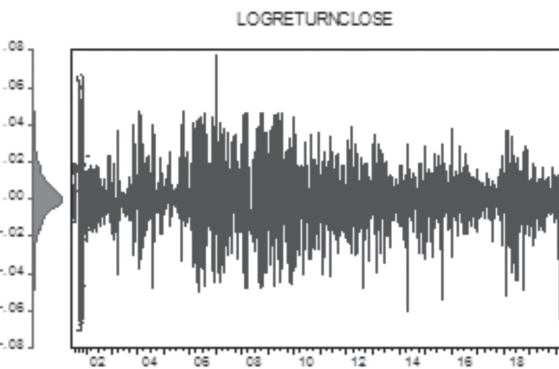
*MacKinnon (1996) one-sided p-values.

Hình 2 mô tả quá trình biến động của chuỗi Logclose bằng cách lấy logarit của chuỗi Close, và nhận thấy chuỗi này cũng có tính chất tương tự như chuỗi Close.



Hình 2. Đồ thị chuỗi dữ liệu logarit giá đóng cửa của Close

Để nghiên cứu chuỗi dữ liệu giá đóng cửa, ta thường xét chuỗi logarit của tỷ suất lợi nhuận, ký hiệu Logreturnclose, bằng cách lấy sai phân bậc 1 của chuỗi Logclose như Hình 3 bên dưới.



Hình 3. Đồ thị chuỗi dữ liệu sai phân bậc 1 của Logclose

Kiểm định nghiệm đơn vị của Dickey-Fuller ở Bảng 2 chỉ ra tính dừng của chuỗi Logreturnclose.

Bảng 2. Kiểm tra tính dừng của chuỗi Logreturnclose

Null Hypothesis: LOGRETURNCLOSE has a unit root
Exogenous: Constant
Lag Length: 4 (Automatic – based on SIC, maxlag=20)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-25.32669	0.0000
Test critical values:		
1% level	-3.431572	
5% level	-2.861965	
10% level	-2.567039	

*MacKinnon (1996) one-sided p-values.

Bây giờ, ta tìm bậc cho mô hình ARIMA(p,d,q) đối với chuỗi dừng Logreturnclose bằng cách mô tả hàm tự tương quan ACF và hàm tương quan riêng phần PACF của chuỗi này.

Sample: 1/03/2001 3/13/2020
Included observations: 4659

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob.
		1 0.265	0.265	326.46	0.000
		2 0.045	-0.027	335.88	0.000
		3 0.016	0.012	337.13	0.000
		4 0.080	0.079	367.06	0.000
		5 0.091	0.054	406.09	0.000
		6 0.070	0.034	429.24	0.000
		7 0.038	0.011	435.83	0.000
		8 0.020	0.003	437.78	0.000
		9 0.010	-0.006	438.23	0.000
		10 0.031	0.021	442.72	0.000
		11 0.039	0.019	449.83	0.000
		12 0.011	-0.011	450.39	0.000
		13 0.030	0.029	454.52	0.000
		14 0.053	0.038	467.65	0.000
		15 0.049	0.020	478.68	0.000
		16 0.030	0.009	483.01	0.000
		17 0.023	0.009	485.53	0.000
		18 0.019	0.002	487.26	0.000
		19 0.028	0.012	490.81	0.000
		20 0.034	0.015	496.22	0.000

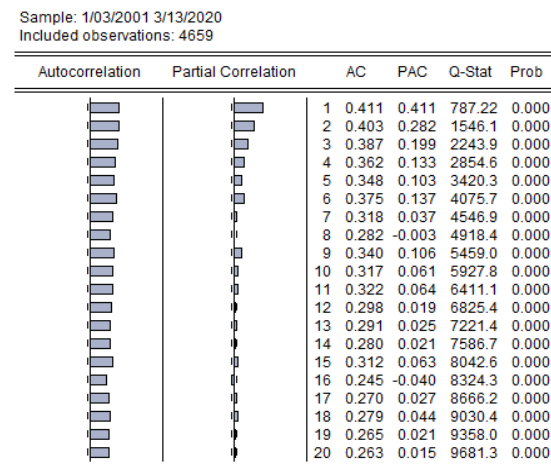
Hình 4. Tính chất của hàm tự tương quan và tương quan riêng phần

Nhìn vào Hình 4, các trường hợp có thể xảy ra cho mô hình ARIMA(p,0,q) hay ARMA(p,q) của chuỗi Logreturnclose được mô tả bằng bảng ước lượng bên dưới với mức ý nghĩa 5%.

Bảng 3. Ước lượng các hệ số của các mô hình ARIMA

ARIMA	(1,0,1)	(1,0,4)	(1,0,5)	(1,0,6)	(4,0,1)	(4,0,4)	(4,0,5)	(4,0,6)	(5,0,1)	(5,0,4)	(5,0,5)	(5,0,6)
ϕ_1	0.17*	0.90*	0.90*	0.93*	0.93*	0.4	1.91*	1.36	1.13*	1.53*	2.41*	1.19
ϕ_2	-	-	-	-	-0.21*	0.32	-1.34*	-0.27	-0.26*	-0.61	-2.58*	0.04
ϕ_3	-	-	-	-	0.01	-0.13	0.3	-0.48	0.02	-0.26	1.50*	-0.67
ϕ_4	-	-	-	-	0.07*	0.22	0.1	0.35	0.08*	0.42	-0.57*	0.32
ϕ_5	-	-	-	-	-	-	-	-	-0.03*	-0.12*	0.20*	0.069
θ_1	0.10*	-0.64*	-0.64*	-0.67*	-0.66*	-0.14	-1.65*	-1.09	-0.87*	-1.27*	-2.15*	-0.92
θ_2		-0.19*	-0.20*	-0.21*	-	-0.38	0.87*	-0.05	-	0.25	1.99*	-0.32
θ_3		-0.04*	-0.04*	-0.04*	-	0.02	-0.03	0.49	-	0.35	-0.92*	0.61
θ_4		0.06*	0.06*	0.07*	-	-0.14	-0.05	-0.14	-	-0.26	0.35*	-0.08
θ_5		-	0.002	0.02	-	-	-0.09*	-0.08*	-	-	-0.19*	-0.12
θ_6		-	-	-0.03*	-	-	-	-	-	-	-	-0.06
R ²	0.0701	0.0793	0.0791	0.0797	0.0796	0.0794	0.0805	0.0802	0.0798	0.0803	0.0806	0.0824
AIC	-5.64	-5.66	-5.66	-5.66	-5.66	-5.66	-5.66	-5.66	-5.66	-5.66	-5.66	-5.66
SIC	-5.64	-5.65	-5.65	-5.65	-5.65	-5.64	-5.64	-5.64	-5.65	-5.64	-5.64	-5.64

Căn cứ vào cột Prob., hệ số Adjusted R-squared, tiêu chuẩn AIC và SIC, mô hình ARIMA(1,0,4), ARIMA(5,0,5) là phù hợp cho chuỗi dữ liệu Logreturnclose. Tiếp theo, ta kiểm tra hiệu ứng ARCH và GARCH cho phần dư trong mô hình ARIMA(1,0,4).



Hình 5. Tính chất hàm tự tương quan và tương quan riêng phần của phần dư

Dựa vào Hình 5, nhận thấy mô hình ARIMA(1,0,4) có hiệu ứng ARCH tối đa bậc 6. Khi chạy mô hình ARIMA(1,0,4) với hiệu ứng ARCH có bậc từ 4 trở lên, có ít nhất các hệ số

ước lượng không có ý nghĩa thống kê. Do đó, ta chọn mô hình ARIMA(1,0,4) với hiệu ứng ARCH(3) để dự đoán cho chuỗi Logreturnclose.

Bảng 4. Hồi quy mô hình ARIMA(1,0,4)-ARCH(3)

Dependent Variable: LOGRETURNCLOSE
Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)
Sample (adjusted): 1/05/2001 3/13/2020
Included observations: 4659 after adjustments
Failure to improve likelihood (non-zero gradients) after 49 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(7) + C(8)*RESID(-1)^2 + C(9)*RESID(-2)^2 + C(10)*RESID(-3)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000869	0.205281	0.004232	0.9966
AR(1)	0.999995	1.26E-05	79677.43	0.0000
MA(1)	-0.763693	0.018798	-40.62562	0.0000
MA(2)	-0.207163	0.020389	-10.16074	0.0000
MA(3)	-0.039913	0.016419	-2.430975	0.0151
MA(4)	0.040757	0.012176	3.347367	0.0008

Variance Equation				
C	4.37E-05	1.37E-06	31.85579	0.0000
RESID(-1)^2	0.331102	0.023703	13.96855	0.0000
RESID(-2)^2	0.347392	0.022321	15.56320	0.0000
RESID(-3)^2	0.197346	0.015949	12.37347	0.0000

Ngoài ra, mô hình ARIMA(1,0,4) còn có hiệu ứng GARCH(1,1). Tuy nhiên hệ số hồi quy của MA(3) và MA(4) không có ý nghĩa thống kê.

Bảng 5. Hồi quy mô hình ARIMA(1,0,4)-GARCH(1,1)

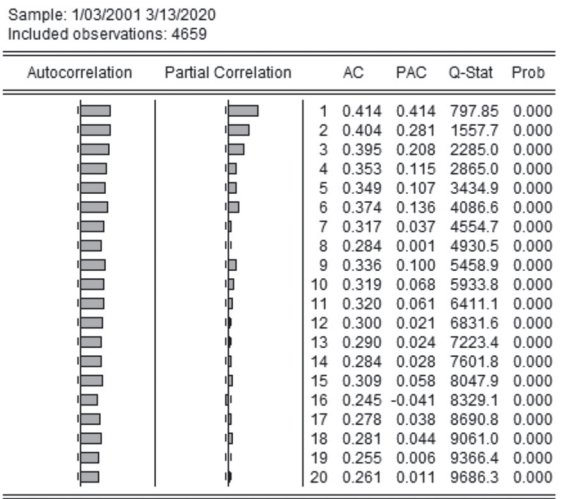
Dependent Variable: LOGRETURNCLOSE
Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)
Sample (adjusted): 1/05/2001 3/13/2020
Included observations: 4659 after adjustments
Convergence achieved after 57 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(7) + C(8)*RESID(-1)^2 + C(9)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	4.33E-05	0.000162	0.267555	0.7890
AR(1)	0.867253	0.074119	11.70084	0.0000
MA(1)	-0.662163	0.076685	-8.634806	0.0000
MA(2)	-0.164648	0.023920	-6.883405	0.0000
MA(3)	0.003075	0.020307	0.151451	0.8796
MA(4)	0.014632	0.018363	0.796826	0.4256

Variance Equation

C	2.87E-06	2.99E-07	9.593598	0.0000
RESID(-1)^2	0.191556	0.011459	16.71600	0.0000
GARCH(-1)	0.808853	0.008965	90.22681	0.0000

Kiểm tra hiệu ứng ARCH và GARCH cho phần dư trong mô hình ARIMA(5,0,5).



Hình 6. Tính chất hàm tự tương quan và tương quan riêng phần của phần dư của ARIMA(5,0,5)

Dựa vào kết quả Hình 6, nhận thấy mô hình ARIMA(5,0,5) có hiệu ứng ARCH tối đa bậc 6. Khi chạy mô hình ARIMA(5,0,5) với hiệu ứng ARCH có bậc 1 đến 6, có ít nhất các hệ số ước lượng không có ý nghĩa thống kê. Ngoài ra, mô hình ARIMA(5,0,5) cũng có hiệu ứng GARCH(1,1). Tuy nhiên hệ số hồi quy của AR(5), MA(2), MA(4) và MA(5) không có ý nghĩa thống kê, được thể hiện trong Bảng 5.

Bảng 6. Hồi quy mô hình ARIMA(5,0,5)-GARCH(1,1)

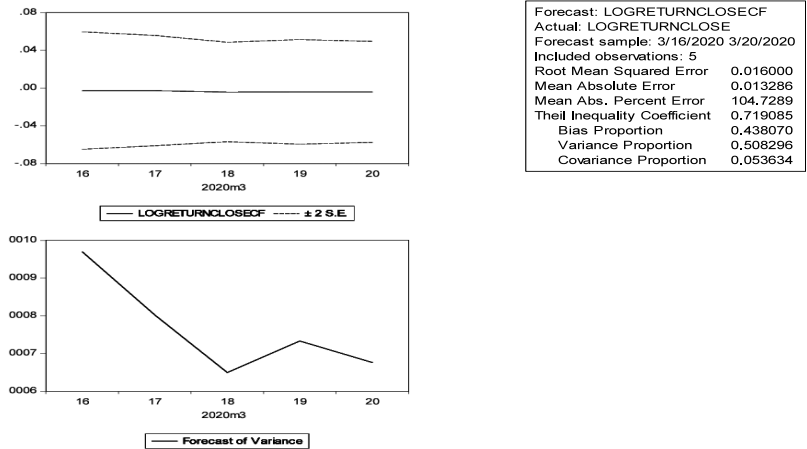
Dependent Variable: LOGRETURNCLOSE
Method: ML ARCH - Normal distribution (OPG - BHHH / Marquardt steps)
Sample (adjusted): 1/05/2001 3/13/2020
Included observations: 4659 after adjustments
Convergence achieved after 59 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(12) + C(13)*RESID(-1)^2 + C(14)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.14E-05	0.000294	-0.038737	0.9691
AR(1)	1.167588	0.351823	3.318684	0.0009
AR(2)	-0.597143	0.287518	-2.076888	0.0378
AR(3)	1.165938	0.174656	6.675617	0.0000
AR(4)	-0.725067	0.351804	-2.060996	0.0393
AR(5)	-0.013672	0.129294	-0.105746	0.9158
MA(1)	-0.961892	0.352770	-2.726683	0.0064
MA(2)	0.371823	0.226727	1.639961	0.1010
MA(3)	-1.042776	0.138013	-7.555648	0.0000
MA(4)	0.508597	0.325661	1.561737	0.1183
MA(5)	0.129903	0.095947	1.353915	0.1758

Variance Equation

C	2.84E-06	3.03E-07	9.373940	0.0000
RESID(-1)^2	0.192185	0.011525	16.67521	0.0000
GARCH(-1)	0.808577	0.009039	89.45234	0.0000

Kết luận, mô hình ARIMA(1,0,4)-ARCH(3) phù hợp để dự đoán cho chuỗi thời gian Logreturnclose. Hình 7 minh họa kết quả dự báo trong mẫu từ ngày 16/03/2020 đến ngày 20/03/2020 cho chuỗi Logreturnclose.



Hình 7. Kết quả dự báo cho chuỗi Logreturnclose

Sử dụng công thức logarit tỷ suất lợi nhuận $r_t = \log\left(\frac{Close_t}{Close_{t-1}}\right)$, ta có kết quả dự báo sau đây:

Bảng 7. Kết quả dự báo của mô hình ARIMA(1,0,4)-ARCH(3)

Ngày	Giá đóng cửa dự báo	Giá đóng cửa thực tế	MSE
16/03/2020	759.70	747.86	492.67
17/03/2020	757.64	745.78	
18/03/2020	754.45	747.66	
19/03/2020	751.38	725.94	
20/03/2020	748.32	709.73	

Bảng 8. Ước lượng các hệ số của các mô hình ARFIMA-FIGARCH

ARFIMA-FIGARCH	(1,d ₁ ,1)/(1,d ₂ ,1)	(1,d ₁ ,4)/(1,d ₂ ,1)	(4,d ₁ ,1)/(1,d ₂ ,1)	(4,d ₁ ,4)/(1,d ₂ ,1)
μ	-0.00001	-0.0003	-0.0003	-0.0003
ϕ_1	-0.13	0.91*	0.94*	0.57
ϕ_2	-	-	-0.1*	0.42
ϕ_3	-	-	0.05*	-0.15
ϕ_4	-	-	0.03	0.06
θ_1	0.27*	-0.91*	-0.95*	-0.59
θ_2	-	-0.11*	-	-0.53
θ_3	-	0.05*	-	1.17
θ_4	-	0.03*	-	0.005
d_1	0.06*	0.2*	0.21*	0.23*
ω	0.00017*	0.001*	1.39*	0.00014*
$\hat{\phi}$	0.90*	0.90*	0.90*	0.90*
β	0.98*	0.97*	0.98*	0.98*
d_2	0.15*	0.16*	0.16*	0.16*
R ²	0.071	0.073	0.073	0.072
AIC	-6.15	-6.15	-6.15	-6.15
BIC	-6.14	-6.13	-6.13	-6.13

Từ kết quả phân tích trên, kết luận rằng mô hình ARFIMA(1,0.2,4)-FIGARCH(1,0.16,1) là phù hợp để dự báo cho chuỗi thời gian Logreturnclose với mức ý nghĩa 10%, mặc dù mô hình ARIMA(1,0,4)-GARCH(1,1) được phân tích ở Bảng 7 là không phù hợp cho chuỗi dữ liệu. Điều này khẳng định tính hiệu lực của mô hình ARFIMA-FIGARCH khi áp dụng mô hình này cho quá trình trong kinh tế và tài chính có tần số cao và chịu tác động của các cú sốc thị trường. Khi đó hệ phương trình

$$\begin{cases} \left(1 - \sum_{i=1}^p \phi_i B^i\right) (1-B)^d (X_t - \mu) = \left(1 + \sum_{i=1}^q \theta_i B^i\right) \varepsilon_t \\ \varphi(B)(1-B)^{d_1} \varepsilon_t^2 = \omega + [1 - \beta(B)](\varepsilon_t^2 - \sigma_t^2) \end{cases}$$

trở thành

$$\begin{cases} X_t = \mu + (\phi_1 + d_1)(X_{t-1} - \mu) - \phi_1 d_1 (X_{t-2} - \mu) \\ \quad + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \theta_4 \varepsilon_{t-4} \\ \varepsilon_t^2 \approx \varphi \varepsilon_{t-1}^2 - d_2 \varepsilon_{t-2}^2 - \omega - \beta(\varepsilon_{t-1}^2 - \sigma_{t-1}^2) + \sigma_{t-1}^2 \end{cases}$$

Kết quả dự báo của mô hình ARFIMA(1,0.2,4)-FIGARCH(1,0.16,1) được cho bởi Bảng 8.

Bảng 9. Kết quả dự báo của mô hình ARFIMA(1,0.2,4)-FIGARCH(1,0.16,1)

Ngày	Giá đóng cửa dự báo	Giá đóng cửa thực tế	MSE
16/03/2020	736.83	747.86	371.97
17/03/2020	738.43	745.78	
18/03/2020	708.99	747.66	
19/03/2020	720.92	725.94	
20/03/2020	696.94	709.73	

4. KẾT LUẬN

Bài báo khảo sát tính chất của chuỗi dữ liệu giá đóng cửa theo ngày của chỉ số VN-Index từ 3/1/2001 đến 13/3/2020. Kết quả phân tích cho thấy chuỗi dữ liệu này thỏa mãn hai mô hình chuỗi thời gian là ARMA(1,4)-ARCH(3) và ARFIMA(1,0.2,4)-FIGARCH(1,0.16,1). Tuy nhiên, kết quả dự báo cho thấy trung bình bình

phương sai số MSE của mô hình ARMA(1,4)-ARCH(3) bằng 492.67, lớn hơn nhiều so với trung bình bình phương sai số MSE của mô hình ARFIMA(1,0.2,4)-FIGARCH(1,0.16,1). Do đó, chúng tôi đề xuất mô hình ARFIMA(1,0.2,4) với hiệu ứng FIGARCH(1,0.16,1) là mô hình phù hợp nhất để dự báo chuỗi dữ liệu nghiên cứu, và đưa ra kết quả dự báo với độ tin cậy 90% cho 5 ngày tiếp theo. Kết quả dự báo cho thấy giá đóng cửa của chỉ số VN-Index tiếp tục lao dốc từ 736.83 điểm xuống 696.94 điểm sau 5 ngày từ 16/03/2020 đến 20/03/2020. Điều này phù hợp với tính chất trí nhớ lâu của mô hình ARFIMA-FIGARCH và biến động thực tế của VN-Index. Một giải thích cho nguyên nhân này là do tình hình dịch bệnh Covid-19 đang diễn biến phức tạp, tạo cú sốc mạnh làm chỉ số VN-Index tiếp tục rớt giá cho đến khi tình hình dịch bệnh được kiểm soát hoàn toàn và nền kinh tế phục hồi trở lại. Chúng tôi dự định sẽ xem xét việc đưa thêm biến giả cú sốc dịch bệnh vào mô hình cho những nghiên cứu tiếp theo trong thời gian tới.

Lời cảm ơn: Nghiên cứu này được thực hiện trong khuôn khổ đề tài khoa học công nghệ cấp cơ sở của Trường Đại học Quy Nhơn với mã số T2019.636.31.

TÀI LIỆU THAM KHẢO

1. Bollerslev, T. and H.O. Mikkelsen. Modeling and pricing long memory in stock market volatility. *Journal of econometrics*, **1996**, 73(1), 151-184.

2. Beine, M., S. Laurent, and C. Lecourt. Accounting for conditional leptokurtosis and closing days effects in FIGARCH models of daily exchange rates. *Applied Financial Economics*, **2002**, 12(8), 589-600.

3. Baillie, R.T., Y.W. Han, and T.-G. Kwon. Further long memory properties of inflationary shocks. *Southern Economic Journal*, **2002**, 3(68), 496-510.

4. Nagayasu, J. The efficiency of the Japanese equity market. *International Finance Review*, **2003**, 4, 155-171.
5. Beine, M. and S. Laurent. Central bank interventions and jumps in double long memory models of daily exchange rates. *Journal of Empirical Finance*, **2003**, 10(5), 641-660.
6. Palma, W. and M. Zavallos. Analysis of the correlation structure of square time series. *Journal of Time Series Analysis*, **2004**, 25(4), 529-550.
7. Doornik, J.A. and M. Ooms. Inference and forecasting for ARFIMA models with an application to US and UK inflation. *Studies in Nonlinear Dynamics & Econometrics*, **2004**, 8(2), 1-25.
8. Conrad, C. and M. Karanasos. Dual long memory in inflation dynamics across countries of the Euro area and the link between inflation uncertainty and macroeconomic performance. *Studies in Nonlinear Dynamics & Econometrics*, **2005**, 9(4), 1-38.
9. Dionisio, A., R. Menezes, and D.A. Mendes. On the integrated behaviour of non-stationary volatility in stock markets. *Physica A: Statistical Mechanics and its Applications*, **2007**, 382(1), 58-65.
10. Kasman, A. and E. Torun. Long memory in the Turkish stock market return and volatility. *Central Bank Review*, **2007**, 2(2), 13-27.
11. Sivakumar, P. B., & Mohandas, V. P. *Modeling and predicting stock returns using the ARFIMA-FIGARCH*. In 2009 World Congress on Nature & Biologically Inspired Computing (NaBIC), IEEE, 2009.
12. Chokethaworn, K., et al. International tourist arrivals in Thailand: Forecasting with ARFIMA-FIGARCH approach. *Economics: Annals of the University of Petrosani*, **2010**, 10(2), 75-84.
13. Wiphatthananthakul, C. and S. Sriboonchitta. ARFIMA-FIGARCH and ARFIMA-FIAPARCH on Thailand volatility index. *International Review of Applied Financial Issues and Economics*, **2010**, 2(2), 280 - 294.
14. Tan, S. and M. Khan. Long memory features in return and volatility of the Malaysian stock market. *Economics Bulletin*, **2010**, 30(4), 3267-3281.
15. Mabrouk, S. and C. Aloui. One-day-ahead value-at-risk estimations with dual long-memory models: Evidence from the Tunisian stock market. *International Journal of Financial Services Management*, **2010**, 4(2), 77-94.
16. Sivakumar, P.B. and V. Mohandas. Performance Analysis of Hybrid Forecasting models with Traditional ARIMA models—A Case Study on Financial Time Series Data. *International Journal of Computer Information Systems and Industrial Management Applications*, **2010**, 2, 187-211.
17. Vo, X. and T. Nguyen. Volatility in stock return series of Vietnam stock market. *Science & Technology Development*, **2011**, Q3(14), 5-21.
18. Chen, J.-H. and J.F. Diaz. Long memory and shifts in the returns of green and non-green Exchange-Traded Funds (ETFs). *International Journal of Humanities and Social Science Invention*, **2013**, 2(10), 29-32.
19. Turkyilmaz, S. and M. Balibey. Long memory behavior in the returns of Pakistan stock market: Arfima-Figarch models. *International Journal of Economics and Financial Issues*, **2014**, 4(2), 400-410.
20. Aye, G.C., et al. Predicting BRICS stock returns using ARFIMA models. *Applied Financial Economics*, **2014**, 24(17), 1159-1166.
21. Ahamed, N., M. Kalita, and A.K. Tiwari. Testing the long-memory features in return and volatility of NSE index. *Theoretical Economics Letters*, **2015**, 5(03), 431-440.
22. Rahmani, M., *Volatility Modelling Using Long-Memory-GARCH Models, Applications of S&P/TSX Composite Index*, PhD Thesis, University of Ottawa, 2016.

23. Toàn, N.V., N.P. Quốc, and N.H. Thạch. Dự báo giá vàng Việt Nam sử dụng mô hình Garch. *Tạp chí Khoa học Trường Đại học An Giang*, **2016**, 2(10), 32-39.
24. Thạch, N.N. and L.H. Anh. Mô hình GARCH trong dự báo sự biến động của giá dầu thế giới. *Tạp chí Công nghệ Ngân hàng*, **2017**, 129, 38 - 47.
25. Masa, A.S. and J.F.T. Diaz. Long-memory Modelling and Forecasting of the Returns and Volatility of Exchange-traded Notes (ETNs). *Margin: The Journal of Applied Economic Research*, **2017**, 11(1), 23-53.
26. Huu Phuoc, B., P. Thi Thu Hong, and N. Van Toan. Asset Price Volatility of Listed Companies in the Vietnam Stock Market. *Asian Journal of Economics and Banking (AJEB)*, **2018**, 2(2), 203-219.
27. P. T. Hong, C. T. Binh. A note on exponential stability of non-autonomous linear stochastic differential delay equations driven by a fractional Brownian motion with Hurst index $> \frac{1}{2}$. *Statistics and Probability Letters*, **2018**, 138, 127-136.