Application of Portfolio Price Forecasting Based on ARIMA-GARCH Model

Jiayu Wang(✉)

Jilin University, Changchun, China
1471947151@qq.com

Abstract. Gold and bitcoins as a typical liquid asset, gold and bitcoin as two types of investment options are favored by many investors in the market, but how to make a portfolio investment, the biggest return in the limited principal conditions, is a matter many investors are concerned about. Based on the historical price data of gold and bitcoin, the cointegration relationship between gold and bitcoin prices was derived, and the arima-garch model was established to make more accurate predictions about the price fluctuations of gold and bitcoin [1]. In order to maximize total returns, and the volatility of the bitcoin market and the volatility of bitcoin prices, and the importance of gold to the world economy, it is important to predict the price of gold and bitcoin and its volatility [2].

Keywords: ARIMA · GARCH · Volatile assets · To predict · Quantitative trading

1 Introduction

ARIMA model is essentially the purpose of predicting the trend of future data by using the correlation and partial self-related functions to simulate the random changes of the data. GARCH model based on the which introduces the lag phase into conditional variance, and they use the information they get to help determine the pricing and determine which assets may provide higher returns, and predict the return of current investments to help asset allocation, hedging, risk management, and portfolio optimization decisions. Based on the ARIMA-GARCH model, this paper makes a more accurate prediction of the price fluctuations of gold and bitcoin [3].

2 Forecasting Model Based on ARIMA-GARCH Model

The autoregressive integrated moving average (ARIMA) model is often used to forecast different types of time series. The model has high reliability and validity in financial time series forecasting, even better than the most popular artificial neural network techniques. As can be seen from the provided data, this is for the prediction of variables about time series, we build ARIMA model for prediction, but ARIMA model requires the data to be a smooth series, we first perform ADF test on the data and find that the p-value of the

K. Hemachandran et al. (Eds.): ICAID 2023, AHIS 9, pp. 296–303, 2023.
https://doi.org/10.2991/978-94-6463-222-4_31

Table 1. ADF Inspection Form

ADF Inspection Form							
Variables	Difference order	T	P	AIC	Threshold		
					1%	5%	10%
Value	0	-0.238	0.934	29168.936	-3.434	-2.863	-2.568
	1	-8.535	0.000***	29151.816	-3.434	-2.863	-2.568
	2	-15.723	0.000***	29187.267	-3.434	-2.863	-2.568
	2	-15.723	0.000***	29187.267	-3.434	-2.863	-2.568

Note: ***, **, * represent 1%, 5%, 10% significance levels, respectively

original data is <0.05, the original data does not satisfy the smoothness test, we perform difference transformation on it to reduce the p-value and change it to a smooth time sequences. In Table 1 Only tests on the BCHAIN-MKPRU.csv data are shown here, and the same for the LBMA- GOLD.csv data [4].

The results of this series test show that based on the field Value: at the difference of order 0, the significance P-value is 0.934, which does not present significance at the level, and the original hypothesis cannot be rejected, and the series is an unsteady time series. At the difference of order 1, the significance P-value is 0.000***, the level of significance is presented, the original hypothesis is rejected, and the series is a smooth time series. At the difference of order 2, the significance P-value is 0.000***, which presents significance at the level, and the original hypothesis is rejected, and the series is a smooth time series. The first-order difference transformation is performed and then the ACF and PACF tests are performed.

In Fig. 1 the horizontal axis represents the number of delays and the vertical axis represents the autocorrelation coefficient. Blue is the autocorrelation coefficient, green is the upper bound of the ACF 95% confidence interval, and yellow is the lower bound of the ACF 95% confidence interval. In Fig. 2 the horizontal axis represents the number of delays and the vertical axis represents the autocorrelation coefficient. The blue is the

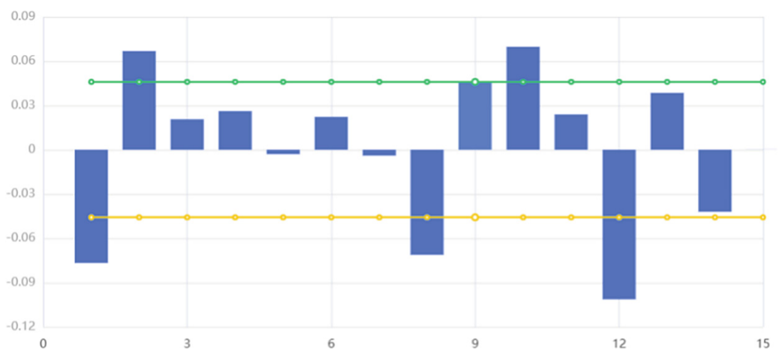

Fig. 1. Final Differential Data Autocorrelation Plot (ACF)

autocorrelation coefficient, the green is the upper bound of the PACF 95% confidence interval, and the yellow is the lower bound of the PACF 95% confidence interval. In Fig. 2, After the analysis of the truncated and trailing tails for ACF and PACF, we choose the ARIMA (0, 1, 2) model for further testing of this model.

Table 2 shows the results of this model test, including sample size, degrees of freedom, Q-statistic and goodness-of-fit of the information criterion model.

- ARIMA model requires that the residuals of the model are not autocorrelated, i.e., the model residuals are white noise, check the model test table, and test the model white noise according to the p-value of the Q statistic (p-value greater than 0.1 is white noise).
- AIC and BIC values for multiple analysis model comparisons according to the information criterion (lower is better).

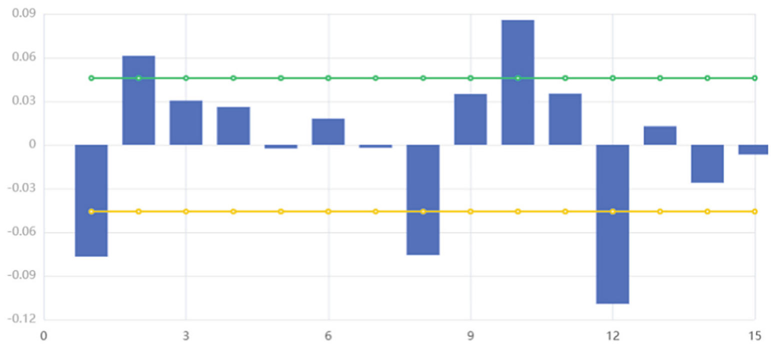

Fig. 2. The Final Differential Data Bias Autocorrelation Plot (PACF)

Table 2. ARIMA Model (0, 1, 2) Test Form

item	Symbols	value
Number of samples	Df Residuals	1822
	N	1826
Q-statistic	Q6(p-value)	0.005(0.944)
	Q12(p-value)	3.367(0.762)
	Q18(p-value)	51.396(0.000***)
	Q24(p-value)	57.011(0.000***)
	Q30(p-value)	99.005(0.000***)
Information Guidelines	AIC	29601.348
	BIC	29623.385
Goodness of fit	R^2	0.997

Note: ***, **, * represent 1%, 5%, 10% significance levels, respectively

- R2 represents the degree of fit of the time series, and the closer to 1 the better the effect.

Model results for the ARIMA model (0, 1, 2) test table, based on the field: Value, from the analysis of the Q statistic results can be obtained: Q6 does not present significance at the level, and the hypothesis that the residuals of the model are white noise series cannot be rejected, while the goodness-of-fit R2 of the model is 0.997, the model performs excellent, and the model basically meets the requirements.

3 Optimization of the Prediction Model based on the GARCH Model

In Fig. 3, Financial time series are characterized by "clustered" and "persistent" return volatility, and their time series errors do not always satisfy the homoscedasticity assumption. By looking at the data variance and difference plots [5].

In Fig. 4, In the previous model building we differenced the series and passed the smooth white noise test. However, the white noise test on the residual series is found to be a non-white noise series.

Fig. 3. The Timing Diagram of the Original Data After 1st Order Differencing

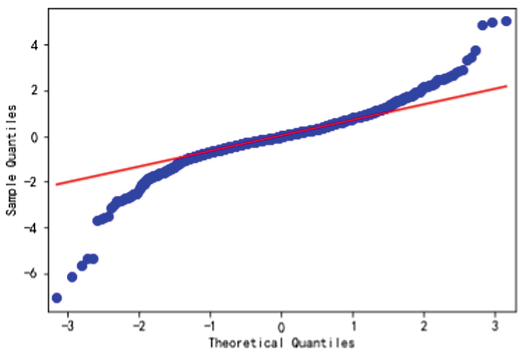

Fig. 4. Residual Plot

Further analysis of the residual series 7: the residuals seem to contain more information, and we consider an ARCH effect. We find that certain sets of variables have different variances from the rest and we consider the set to be heteroskedastic, and since financial series are series without any trend and seasonal effects, some insights are given through previous work and we try to build a combined ARIMA- GARCH model to fit forecasts to the data of bitcoin and gold separately [6], which does not easily ignore the heteroskedasticity of the random error difference term in the autoregressive function and thus does not seriously underestimate the variance of the residuals, and at the same time can improve the fitting accuracy of the model. As shown in formula 0.1, The GARCH model is the ARMA model applied to the variance of the time series, and it contains an autoregressive term and a moving average term. If the time-series data y_t can be expressed as 1.

$$y_t = \sigma_t \omega_t \tag{1}$$

ω_t is Gaussian white noise with mean 0 and variance 1. Here σ_t.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{2}$$

$$\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, \quad \sum_{i=1}^{\max(m,s)} \left(\alpha_i + \beta_j\right) < 1 \tag{3}$$

As shown in formula 0.2, We combine the ARIMA model and the GARCH model to forecast [7]. Finally, we examine the normalized residuals of the GARCH model, As shown in formula 0.3, as well as the squared normalized residuals, and they both pass the white noise test, which means that our modeling is truly complete [8].

4 Prediction Results of the Data by ARIMA-GARCH Model

The predicted results are shown in the Fig. 6 shows the time on the horizontal axis and the value on the vertical axis, the blue line indicates the predicted data and the green line indicates the actual data as shown in the Fig. 5.

The combined ARIMA-GARCH model is closer to the actual values in terms of accuracy, indicating that the model has good characteristics for short-term forecasting. In the long term, the accuracy of the ARIMA-GARCH model prediction decreases due to the emergence of un- expected events, but it is still better than the ARIMA model. The horizontal axis is time, the vertical axis is VALUE, green is predicted data, and blue is real data [9] as shown in the Fig. 6.

It can be seen that the forecast and the actual situation are basically the same, but locally the forecast data are generally higher than the actual data, but the trend of change is the same as the actual data, which has little impact on the subsequent decision. To make our scheme more visual, the flowchart of our ARIMA-GARCH modeling is given in Fig. 7.

Fig. 5. Bitcoin Value Forecast Chart

Fig. 6. Gold Value Forecast Chart

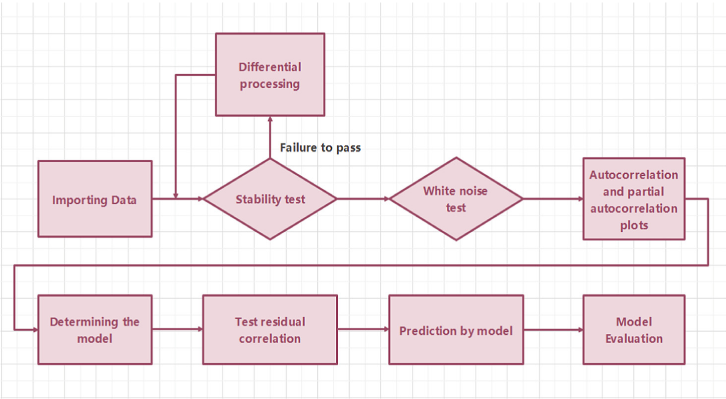

Fig. 7. Algorithm to ARIMA-GARCH Model

5 Conclusion

ARIMA-GARCH model e main strength are its great scalability and applicability, the low requirements of our model in terms of data characteristics, and the fact that our paired trading model can be used to combine not only bitcoin and gold, but also two other stocks. For the establishment of the ARIMA-GARCH model, we conducted a rigorous analysis at each step to ensure the correctness and rigor of the model [10]. Based on the scientific and reasonable financial decision model of paired trading, we boldly and innovatively assume that a small period of time satisfies the mean reversion Omstein-Uhlenbeck regression process, which is found to be basically consistent with the actual situation after analysis. Our model effectively achieves all the objectives. Not only is it fast and capable of han- dling large amounts of data, but it is also flexible and practical, achieving better results than basic predictive and decision models [11].

References

1. DING Lei, GUO Wanshan, Prediction of Gold Price Based on ARIMA-GRACH Mixed Model, JOURNAL OF XUCHANG UNIVERSITY, No. 6, 2019.
2. CHEN Yuanqiao, The optimal analytical solution of signaling mechanism for pairs trading strategy, 5, 2015.
3. Ganapathy Vidyamurthy, “Pairs Trading: Quantitative Methods and Analysis”, Wiley.
4. SUN Jianquan, HAN Botang, SUN Shulei, Pricing Warrants on Stocks Driven by Ornstein-Uhlenback Process, CHINESE JOURNAL OF MANAGEMENT SCIENCE, 10, 2006.
5. “GARCH(p,q) model” https://blog.csdn.net/weixin_43171270/article/details/102975924
6. DUAN Gege, Bitcoin Price Prediction Based on ARIMA Model, MARKETING MANAGEMENT REVIEW, No.1, 2021.
7. Raghavender K V, Premchand P. Accurate and reliable detection of DDoS attacks based on ARIMA-SWGARCH model [J]. International Journal of Information and Computer Security, 2021, 14(2):118.
8. Setyowibowo S, As’ad M, Sujito E F, et al. Forecasting of Daily Gold Price using ARIMA-GARCH Hybrid Model [J]. Jurnal Ekonomi Pembangunan, 2021, 19(2): 257-270.
9. Xiang Y. Comprehensive trading strategy of gold and bitcoin based on ARIMA-GARCH model [C]//Second International Symposium on Computer Technology and Information Science (ISCTIS 2022). SPIE, 2022, 12474: 297-302.
10. Wang L, Zhang Z. Research on shanghai copper futures price forecast based on X12-ARIMA-GARCH family models[C]//2020 International Conference on Computer Information and Big Data Applications (CIBDA). IEEE, 2020: 304-308.
11. Sahed A, Mekidiche M, Kahoui H. Fuzzy auto-regressive integrated moving average (FARIMA) model for forecasting the gold prices[J]. Journal of Smart Economic Growth, 2020, 5(1): 1-13.