

Gold Price Prediction based on HP Filtering ARMA-GARCH Model

Yan Li, Ruiqian Wang

School of Business, Henan University, Kaifeng 475004, China.

Abstract

As a special precious metal, gold not only has three functions of currency, finance and commodity, but also has an important impact on the economic field under the impact of COVID-19 epidemic. Therefore, it is of great significance to predict the trend of gold price for individual investor sentiment relief and social and economic development. Based on previous studies, the prediction was made in this paper: first, the time series was decomposed into trend series and periodic fluctuation series by HP filtering. Then, according to the properties of different sequences, the trend element sequences were fitted by difference integrated moving average autoregressions (ARIMA) model, and the periodic fluctuation sequences were fitted by ARMA-GARCH model. Finally, the two prediction sequences are added and compared with the original sequence. The prediction results are satisfactory in terms of model accuracy and range.

Keywords

Gold price forecast; HP filter; Differential integrated Moving average autoregressive model (ARIMA); ARMA model - the family of GARCH model.

1. Introduction

Due to the impact of COVID-19 epidemic, the domestic futures market showed "black Swan" on the first day after the holiday, with many commodities falling by the daily limit, while gold and silver performed well and closed in the sunshine. However, with the continuous introduction of China's strong control measures, market confidence was enhanced and risk aversion eased, and gold and silver prices fluctuated widely. The market is still assessing the impact of the global public health emergency on the global economy. Investors' concerns have eased but not been completely eliminated. In the short term, the development of the COVID-19 epidemic has become an important factor influencing the future price of gold. Historically, gold because of its value and become the most extensive currency circulation, until today have not quit the stage of history, gold has the property of commodity, financial and monetary attribute, this means that not only can be used to make jewelry of gold, it is widely used in many aspects, play an important role in human society. Gold is an asset, because of its rarity and very precious, at the same time, the stability of the gold to make it easy to save, so gold is not only become the human material wealth, but also an important means to storing wealth, is also important investment way, such as gold T + D, gold ETF, gold stocks, paper gold bullion investment in the form. Therefore, the study of gold price prediction is of great significance not only for investors, but also for economic prosperity and social progress. At present, a large amount of literature has been accumulated on the quantitative research of gold price prediction at home and abroad, and various prediction models have been put forward.

The classical methods include GreyModel analysis, traditional regression analysis, time series analysis and GARCH family model analysis. These theories and methods effectively describe the running law of time series and deepen the understanding of it from different angles. There are many studies on gold price prediction in China. Jingwen Fei found that ARIMA model can make a short-term prediction on the trend of gold futures price, which can generally reflect the fluctuation of gold futures price [1]. However, the prediction error increases with the increase of the prediction time. Ding Lei and Guo Wanshan combined ARIMA and GARCH, constructed a new hybrid model of Arima-Garch, and predicted the gold price by arima-Garch in different distributions, and found that the new hybrid model was more practical in short-term prediction [2]. At the same time, Jing Zhigang and Shi

Guoliang used LS-SVM and ARIMA models to model and predict different trends, and reconstructed the results of gold price combination prediction [3]. Chen Peng also conducts an empirical study on the long memory of gold price, and USES Hurst index to prove that there is indeed a significant long memory in gold price [4]. The arfIMA-Garch model family is established after fractional difference of gold price yield, which reflects the volatility aggregation of gold yield series. Wu Nanlin et al. found through empirical analysis that the EGARCH-T model described the volatility characteristics of Shanghai gold futures well, and measured the risk loss of the gold futures market with the help of CVaR method [5]. Finally on this basis to the gold futures market investors and regulators to put forward risk control recommendations. Peng Xiaoshu et al., based on the heteroscedasisness of the international gold price, introduced the international oil price and the DOLLAR index as exogenous variables to make fitting prediction of the international gold price [6]. Zhang Pinyi et al. constructed a multi-input GA-BP neural network model and made nonlinear prediction for the gold price [7]. Yang Xinyuan et al. predicted the gold price based on artificial neural network algorithm [8]. Foreign studies on gold price prediction include: Farah Naz et al. used box-Jenkins method to predict the time series of gold price in India [9]. Peter Sephton et al. studied how the impact of oil price affects the gold price, which depends on the scale of the impact and the region where the system is located when the impact occurs [10]. Shahriar Shafiee et al. used the long-term trend reversal jump and DIP diffusion model to predict the price of natural resource commodities, indicating that the historical data of mineral commodities have three terms to show the price fluctuation: long-term trend reversal component, diffusion component and jump or fall component [11]. Larry A Sjaastad et al. studied the main exchange rates and prices of international commodity trade, and believed that, compared with other currencies, the appreciation or depreciation of European currencies had A strong impact on the price of gold [12].

On the basis of previous studies, this paper USES HP filtering to separate the trend sequence T from the closing price of gold in Shanghai Gold Exchange, and then makes an empirical analysis on the gold price by using ARIMA model. Finally, the trend series and the periodic fluctuation factors are added to obtain the prediction series. Meanwhile, the prediction series is compared with the original series. Through testing the prediction effect, it is found that the average error rate of ARMA(4,2) - garch (1,1) is less than 1%, and the prediction effect is very good.

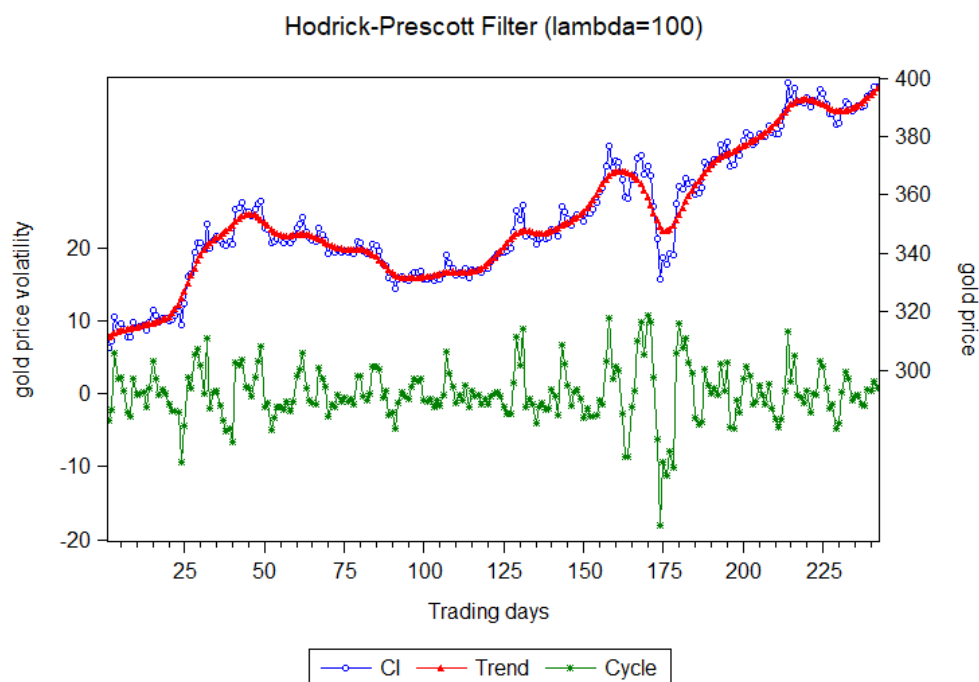


Fig. 1 HP filter decomposition diagram

2. Methodology and empirical findings

2.1 Sample selection

The data in this paper are from the CSMAR database. Gold Au9999 of Shanghai gold exchange is selected as the research object. The closing price of solstice on June 30, 2019 and June 29, 2020 is selected, and 242 data are excluded from holidays, It's in yuan per gram. The first 217 data are selected to build the model, and then the 25 data out of sample are predicted.

2.2 HP filtering decomposes the data

HP filtering relies on the flat index. In this paper $\lambda=100$ is selected to decompose the original time series CI into trend series T and periodic fluctuation series Y, and the decomposition results are shown in Figure 1.

It can be seen from Figure 1 that HP filter decomposed the trend sequence T with a relatively smooth trend and the periodic sequence Y with a fluctuation law. Then, differential integrated moving average autoregressive model (ARIMA) and GARCH family model were respectively used to fit and predict the decomposed data.

2.3 Forecast the trend series

Before the model regression of trend T, the stationarity test is required. If the sequence is not stationary, the autoregression modeling can be conducted after the sequence is stabilized. The unit root test results of trend T are as follows:

Null Hypothesis: T has a unit root Exogenous: Constant Lag Length: 5 (Automatic - based on SIC, maxlag=14)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.496730	0.8882
Test critical values:		
1% level	-3.457984	
5% level	-2.873596	
10% level	-2.573270	

Fig. 2 Trend T unit root test results

From figure 2, we can see that ADF of trend T equals to $-0.496730 > 5\%$ significance level critical value -2.873596 , which fails to pass the test, that is, there is a unit root (unstable), and the trend T needs to be stabilized. At this point, the first-order difference unit root test result of trend T is checked through the first-order difference (figure 3).

Null Hypothesis: D(T) has a unit root Exogenous: Constant Lag Length: 4 (Automatic - based on SIC, maxlag=14)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.406536	0.0004
Test critical values:		
1% level	-3.457984	
5% level	-2.873596	
10% level	-2.573270	

Fig. 3 Test results of trend T first order difference unit root

As shown in Fig. 3, $ADF = -4.406536$, less than the threshold value of 5% significance level, so D(T) sequence is a stationary sequence.

Modeling was continued on the basis of stationary fluctuation sequence D(T), and partial autocorrelation of the sequence was tested. The test results were shown in Figure 4. The probability values (P values) were all equal to 0.

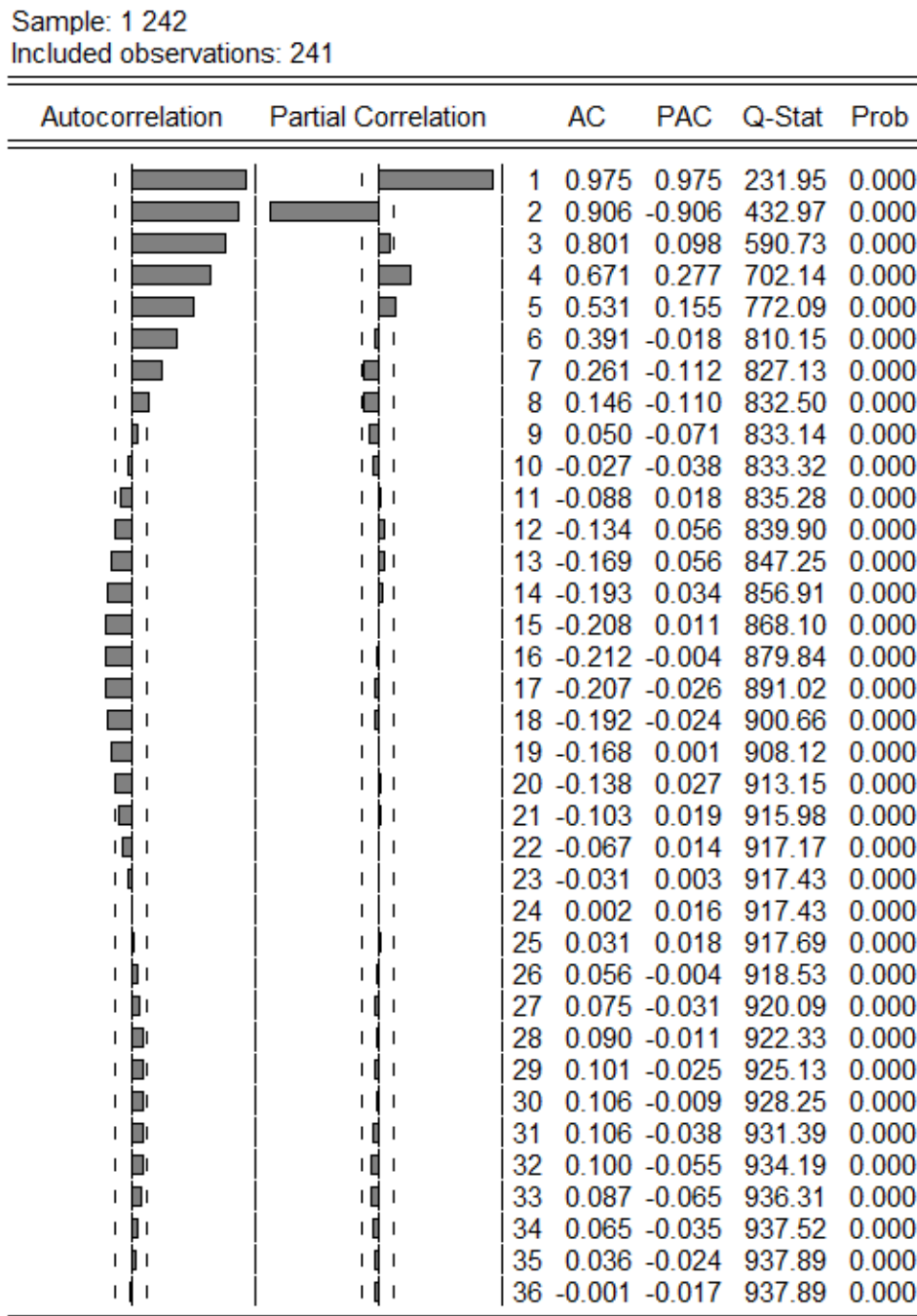


Fig. 4 D(T) Correlation test results

It can be seen from Figure 4 that the autocorrelation coefficient and partial autocorrelation coefficient of trend series D(T) are trailing, it can be seen from the figure 4 p and q, respectively, of the ARMA model can choose 1, 2, 3, 4, 5, and The order of the model is determined according to the minimum selection of AIC and SC criteria. The previous 217 data are modeling data, and the ARIMA(4,1,1) model is established for the trend sequence T after multiple fitting, Autoregressive model is established and the parameters can be seen in table 1, the fitting coefficient is $R^2 = 0.999355$, and p value and the T statistic are through inspection (figure 5), the fitting effect is good.

Table 1 Left and right ARIMA model AIC and SC values of order 5 and below

Model	AIC	SC
ARIMA (1,1,1)	-1.545333	-1.482828
ARIMA (1,1,2)	-2.349668	-2.271536
ARIMA (1,1,3)	-3.298647	-3.204889
ARIMA (1,1,4)	-3.659425	-3.550041
ARIMA (1,1,5)	-4.067899	-3.942888
ARIMA (2,1,1)	-3.652498	-3.574365
ARIMA (2,1,2)	-4.135094	-4.041336
ARIMA (2,1,3)	-4.268181	-4.158797
ARIMA (2,1,4)	-4.381519	-4.256509
ARIMA (2,1,5)	-4.427099	-4.286462
ARIMA (3,1,1)	-4.391581	-4.297824
ARIMA (3,1,2)	-4.443946	-4.334562
ARIMA (3,1,3)	-4.446718	-4.321707
ARIMA (3,1,4)	-4.457156	-4.316520
ARIMA (3,1,5)	-4.451359	-4.295097
ARIMA (4,1,1)	-4.490118	-4.380734
ARIMA (4,1,2)	-4.480918	-4.355908
ARIMA (4,1,3)	-4.471824	-4.331206
ARIMA (4,1,4)	-4.468250	-4.311987
ARIMA (4,1,5)	-4.459638	-4.287794
ARIMA (5,1,1)	-4.480926	-4.355916
ARIMA (5,1,2)	-4.472363	-4.331726
ARIMA (5,1,3)	-4.463966	-4.307703
ARIMA (5,1,4)	-4.477159	-4.305270
ARIMA (5,1,5)	-4.473395	-4.285879

Sample: 2 217
 Included observations: 216
 Convergence achieved after 17 iterations
 Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.346903	0.163450	2.122373	0.0350
AR(1)	3.424255	0.113339	30.21253	0.0000
AR(2)	-4.582276	0.312393	-14.66829	0.0000
AR(3)	2.842017	0.299837	9.478551	0.0000
AR(4)	-0.690612	0.099847	-6.916718	0.0000
MA(1)	-0.350912	0.144264	-2.432429	0.0158
SIGMASQ	0.000578	5.37E-05	10.76726	0.0000
R-squared	0.999355	Mean dependent var		0.374250
Adjusted R-squared	0.999336	S.D. dependent var		0.948830
S.E. of regression	0.024444	Akaike info criterion		-4.490118
Sum squared resid	0.124879	Schwarz criterion		-4.380734
Log likelihood	491.9327	Hannan-Quinn criter.		-4.445927
F-statistic	53956.28	Durbin-Watson stat		1.985457
Prob(F-statistic)	0.000000			
Inverted AR Roots	.91+.15i	.91-.15i	.80-.41i	.80+.41i
Inverted MA Roots	.35			

Fig. 5 ARIMA (4,1,1) model estimation results

On the basis of this model, the static prediction method is used to predict 25 data out of sample, and the prediction effect is shown in Figure 6.

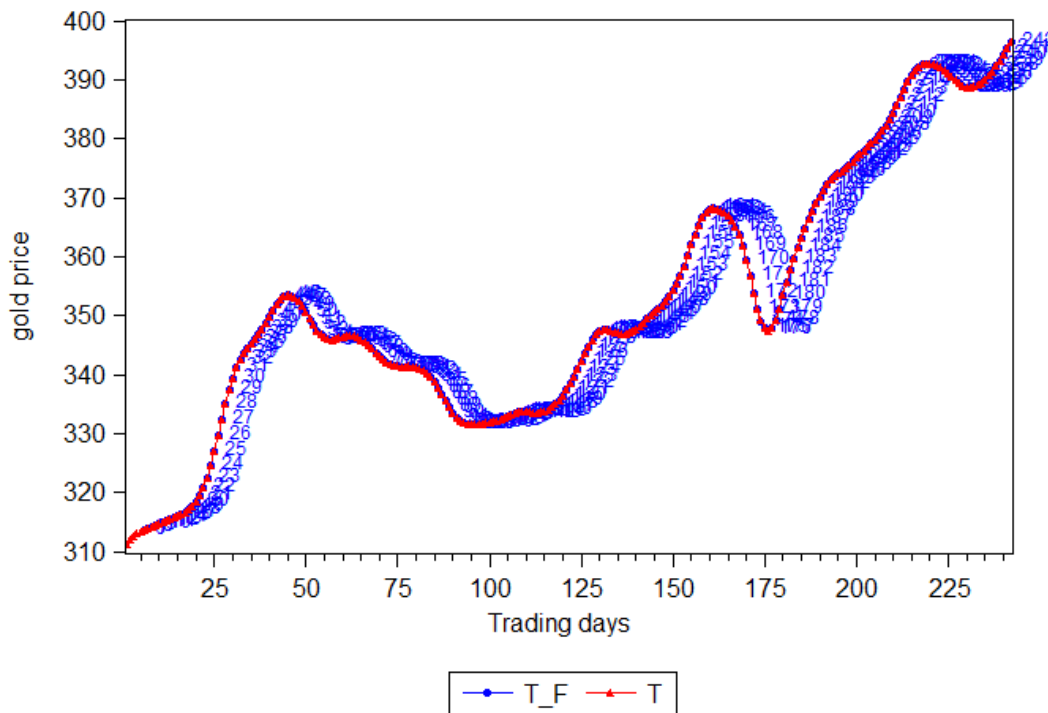


Fig. 6 Trend series T out of sample prediction

It can be seen from figure 6 that the prediction effect is good. The predicted value TF of ARIMA(4,1,1) model and the trend value T after HP filtering decomposition almost coincide. Thus, the random fluctuation sequence $Y_0 = CI-TF$ with period not predicted by the regression model is obtained, and then the GARCH family model was used to fit and predict Y_0 .

2.4 GARCH family model fitted and predicted periodic random fluctuation sequence Y_0

The ADF test of periodic random fluctuation sequence Y_0 is shown in Table 2.

Table 2 ADF test results of random fluctuation sequence Y_0

ADF statistic t-Statistical	P value	Test critical value /%	t-Statistical
-9.716933	0.0000	1	-2.575011
		5	-1.942205
		10	-1.615874

It can be seen from Table 2 that $ADF = -9.716933$, less than the critical value of 1% test level, so periodic random fluctuation series Y_0 is stationary series.

Correlation test was performed on the periodic random fluctuation sequence, as shown in Figure 7, It can be seen from Figure 7 that the periodic random fluctuation sequence Y_0 has autocorrelation and partial autocorrelation. Because the periodic random fluctuation sequence Y_0 is stationary but has autocorrelation, and the regression equation is an autoregressive moving average model (ARMA model). Therefore, multiple attempts should be made to determine the order of the model according to the minimum criteria of AIC SC, and ARMA(4,2) is finally determined.

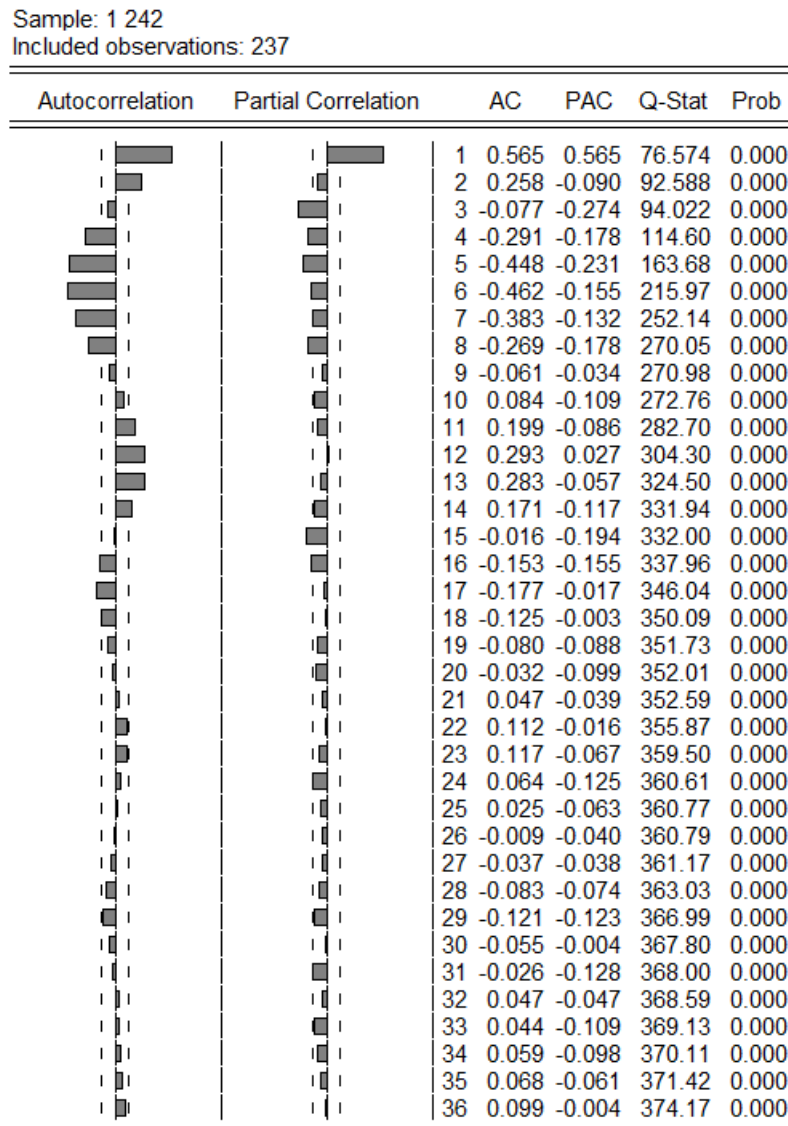


Fig. 7 Correlation test results of periodic random fluctuation sequence Y_0

LM test and ARCH test are carried out for the regression equation ARMA(4,2) of the periodic random fluctuation sequence Y_0 to verify whether the sequence Y_0 has ARCH effect. When the lag order $q=3$, The ARCH test results are shown in table 3.

Table 3 ARCH test results of periodic random fluctuation sequence Y_0

F-statistic	P value	Obs*R-squared	P value
8.971362	0.0000	24.47566	0.0000

According to Table 3, the P value of chi-square test is $0.0000 < 0.05$, that is, there is ARCH(3) effect in the residual series, and the null hypothesis is accepted when $q < 3$, and rejected when $q \geq 3$. It is proved that there is high order ARCH effect in periodic random fluctuation sequence Y_0 , so GARCH(p, q) model should be selected.

ARMA-GARCH family model was used to test the data. There were seven models ARMA(4,2)-GARCH(1,1) ARMA(4,2)-EGARCH(2,1) ARMA(4,2)-TARCH(2,1) ARMA(4,2)-GARCH-M(1,1) ARMA(4,2)-EGARCH-M(1,2) ARMA(4,2)-TARCH-M(1,2) ARMA(4,2)-CAR(1,1), The comparison table of AIC SC and HQ of the 7 models is shown in Table 4

Table 4 AIC SC HQ Comparison Table of 7 alternative models

Model	AIC	SC	HQC
ARMA(4,2)-GARCH(1,1)	4.803084	4.963543	4.867965
ARMA(4,2)-EGARCH(2,1)	4.867754	5.060304	4.945611
ARMA(4,2)-TARCH(2,1)	4.820588	4.997093	4.891958
ARMA(4,2)-GARCH-M(1,1)	4.831377	5.007881	4.902746
ARMA(4,2)-EGARCH-M(1,2)	4.899867	5.108463	4.984213
ARMA(4,2)-TARCH-M(1,2)	4.815856	5.024452	4.900202
ARMA(4,2)-CAR(1,1)	4.852326	5.044876	4.930184

3. Results and discussion

According to the minimum principle of AIC SC criterion, the ARMA(4,2) -garch (1,1) model was determined by multiple tests. The coefficient estimation results are shown in table 5 below.

Table 5 The Y coefficient estimation results of wave series under ARMA (4,2)-GARCH (1,1) model

Mean value equation and variance equation coefficient	Variables	Coefficient	t-Statistical	P value
Mean value equation	C	-0.00953	0.012656	-0.75273
	AR(1)	0.591026	0.075239	7.855334
	AR(2)	0.528696	0.087068	6.072217
	AR(3)	-0.47558	0.092441	-5.14463
	AR(4)	-0.11296	0.075931	-1.48767
	MA(1)	-0.30906	0.008996	-34.3571
	MA(2)	-0.69094	0.020038	-34.4804
Variance equation coefficient	C	0.497205	0.328507	1.513529
	RESID(-1) ²	0.077308	0.046185	1.673865
	GARCH(-1)	0.853805	0.087129	9.799344

In the above table, the sum of the square coefficient of RESID(-1)² for the lag period and the variance coefficient GARCH(-1) for the lag period is $0.931113 < 1$, meeting the parameter constraints; As the sum of the coefficients is close to 1, it indicates that the impact on the conditional variance of periodic series is long-lasting, that is, the impact will play an important role in all future predictions.

After the ARMA-GARCH model was used to fit, the ARCH effect of the fluctuating series was tested again. The test results showed that the F statistic was 0.318443, the Chi-square statistic was 0.321052, and the corresponding probability values (P values) were 0.5732, 0.5710 respectively, both of which were greater than 0.05. Therefore, the ARCH effect of the fluctuating series Y was considered to be eliminated.

The prediction series TF of trend series T by autoregression model and the prediction series Y_0F of periodic random fluctuation series Y_0 by GARCH model family are recombined, that is, $CIF = TF + Y_0F$ serves as the closing price prediction series of gold price Au9999, with a total of 25 values. The model error rate is based on the predicted value CIF minus the absolute value of the original sequence CI and then divided by the original sequence CI, It is calculated that the average error rate of model prediction is 0.32%, that is, the accuracy of HP filtering - ARIMA(4,1,1) - ARMA(4,2) - GARCH(1,1) model can reach 99.68%, and the prediction effect of the model is very good.

4. Conclusion

This article will be used in macroeconomic analysis of HP filter is applied to time series analysis, the original sequence is decomposed into two parts: the periodic fluctuation and trend sequences and then, according to sequence of different nature, different model fitting prediction, the trend of gold prices on the selected elements according to the nature of the sequence ARIMA model fitting prediction, elements of its cyclical fluctuations ARMA model - the family of GARCH model fitting forecast in the end, the trend of elements and the periodic fluctuations in together, get forecast sequence, compared with the original sequence evaluation at the same time The accuracy of the model selected in this paper can reach 99.68%, with high prediction accuracy and good prediction effect.

References

- [1] Fei Jingwen. Analysis and Forecast of China's gold Futures Price based on ARIMA Model [J]. Contemporary Economics, 2017, (09):148-150.
- [2] Ding lei, guo wanshan. Gold price prediction based on arima-garch hybrid model [J]. Journal of xuchang university, 2019, 38(06): 124-129. (in Chinese)
- [3] Jing zhigang, shi guoliang. Gold price prediction based on ls-svm ARIMA combination model based on wavelet analysis [J]. Gold, 2017, 38(05):5-8+14.
- [4] Chen Peng, LI Xingye. Prediction analysis of gold Price based on ArfIMA-Garch Model Family [J]. E-commerce, 2018, (05): 42-44.
- [5] Wu Nanlin, Song Xuan, Jing Yiyang. Risk measurement of Chinese gold futures market based on EGARCH Model [J]. New Economy, 2020,(Z2):20-31.
- [6] Peng xiaoshu, zhang desheng, wang ruoxing, Chen cong. GARCH prediction model of international gold price with exogenous variables [J]. Gold, 2011, 32(01):10-14.
- [7] Zhang Pinyi, LUO Chunyan, Liang Strontium. Simulation prediction of gold Price based on GA-BP neural Network Model [J]. Statistics and Decision-making, 2016, 34(17):158-161.
- [8] Yang Xinyuan, ZHAO Yingqi, WU Man, LIU Jinlu, XU Jie, LIU Ke, GAO Xiang. Research on Mathematical Model of Precious metal Price Prediction [J]. Modern Economic Information, 2016, (08): 104-106+108.
- [9] Naz, Farah and Ahmad, Zahid(2016).Forecasting of Indian Gold Prices Using Box Jenkins Methodology. Journal of Indian Studies, Vol.2, No.1,p.75 – 83.
- [10] Peter Sephton, Janelle Mann(2018).Gold and crude oil prices after the great moderation, Energy Economics, Volume 71, p. 273-281.
- [11]Shahriar Shafiee, Erkan Topal(2018). An overview of global gold market and gold price forecasting, Resources Policy, Volume 35, Issue 3, p.178-189,
- [12]Sjaastad, L. A., & Scacciavillani, F. (1996). The price of gold and the exchange rate. Journal of International Money and Finance, 15(6), 879-897.