International Journal of Advances in Applied Sciences (IJAAS)
Vol. 1, No. 4, December 2012, pp. 153~158
ISSN: 2252-8814

ARIMA Model for Gold Bullion Coin Selling Prices Forecasting

Lazim Abdullah *
* Departement of Mathematics, University Malaysia Terengganu

Article Info

Article history:

Received Jul 22, 2012
Revised Oct 23, 2012
Accepted Nov 14, 2012

Keyword:

Forecasting
Stationary Estimation
Auto Regressive
Moving Average
Golg Prices

ABSTRACT

Time series forecasting is an active research area that has drawn considerable attention for applications in a variety of areas. Auto-Regressive Integrated Moving Average (ARIMA) models are one of the most important time series models used in financial market forecasting over the past three decades but not very often used to forecast gold prices. This paper attempts to address the forecasting of gold bullion coin selling prices. The forecasting models ARIMAs are applied to forecast the gold bullion coin prices. The result suggests that ARIMA (2, 1, 2) is the most suitable model to be used for forecasting gold bullion coin prices. Closer examination suggests that the gold bullion coin selling prices are in upward trends and could be considered as a worthy investment.

Corresponding Author:

Lazim Abdullah,
Department of Mathematics,
University Malaysia Terengganu,
21030 Kuala Terengganu, Malaysia.
Email: lazim_m@umt.edu.my

1. INTRODUCTION

Central bank of Malaysia has recognized that gold bullion coins are legal tender coins whose market price depends on their gold content rather than on their face value. Gold bullion coins are traded daily throughout the world and their price depends on the prevailing international gold price. Investors are encouraged to invest in gold bullion coins because its price depends on the international gold price and not very subjected to inflation. With investing in gold bullion, investors may reduce the risk of losing their cash such as in the case of a sudden slide in the stock market or increased inflation rate. Gold can be considered as an option to protect against any eventuality since it is pretty immune from national and regional economies. Most of investors would like to keep a portion of their total assets in gold because it is low-to-negative correlation with stocks and bonds thereby make it an excellent portfolio diversifier. Gold investors may depend on historical data of gold price to forecast future prices prior to making their investment decision. The main reason for forecasting is to minimize risk when making a decisive move.

Forecasting is a process in management to assist decision making. It is also described as the process of estimation in unknown future situations. In a more general term it is commonly known as prediction which refers to estimation of time series or longitudinal type data. The most popular model for this method is the Box-Jenkins model introduced by [1]. Box-Jenkins has suggested the time-series autoregressive integrated moving average (ARIMA) model for forecasting. Like any other such methods, it requires historical time series data on the variable under forecasting. It assumes that the future values of a time series have a clear and definite functional relationship with current, past values and white noise. Kumar et al. [2] stated that the ARIMA offers a good technique for predicting the magnitude of any variables. The model has been successfully tested in many forecasting. In fishery industries, Lloret, et al. [3] suggests ARIMA models as the most appropriate to forecast fishery landings in the Hellenic marine waters, since systematic biological time-series data sets from explanatory variables are lacking. This methodology has been used to model and forecast the landings and catch per unit effort of many fish and invertebrate species. In financial forecasting, Fang [4] combines two methods to develop the fuzzy ARIMA model based upon the works of time-series ARIMA (*p*,*d*,*q*) model and fuzzy regression model. He uses the new method Fuzzy ARIMA to forecast the foreign market exchange and get the accurate forecasting value in a short time period. Ediger and Akar [5] used ARIMA model and seasonal ARIMA methods to forecast primary energy demand on fossil fuel in Turkey starts in year 2005 to year 2020. Wood and Dasgupta [6] used regression model, ARIMA's model and neural network model to forecast the MSCI, Capital Market Index of United Stated Amerika. They found that the ARIMA model which was built on the percentage changes in 3-period moving average is performing better than the ARIMA model build on the index itself.

The idea that Box and Jerkin's ARIMA model has predictability in many bussiness activities including gold price is accepted in many researches in various countries. Selvanathan [7] reported that in Australia, the comparison between the forecasted London daily gold price resulted from the Economic Research Centre and ARIMA model has been done. The paper was claimed and proved that only simple ARIMA is very low cost and effectively enough to predict gold price. Additionally, Box-Jenkins' ARIMA is widely used to predict the future outcomes for economic or financial purposes. In this paper, we test the gold bullion coin selling prices data to ARIMA forcasting model. Specifically the data of gold bullion coin selling prices are employed to determine the best fit ARIMA model.

2. ARIMA MODEL

One of widely used time series models is ARIMA. We choose ARIMA models rather than the others such as Average Moving, Average Naïve, due to its flexibility that it can represent several different types of time series, i.e. pure autoregressive (AR), pure moving average (MA) and combined AR and MA (ARMA) series. The ARIMA model is denoted by ARIMA (*p*,*d*,*q*), where "*p*" stands for the order of the auto regressive process, "*d*" is the order of the data stationary and "*q*" is the order of the moving average process [8]. In ARIMA model, the future value of a variable is assumed to be a linear function of several past observations and random errors. That is, the underlying process that generate the time series has the form

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... - \theta_1 \varepsilon_{t-1} \quad (1)$$

where y_t and ε_t are the actual value and random error at time period *t*, respectively; $\o_i$ (i=1, 2,…, p) and θ_j (j=0, 1, 2,…, q) are model parameters. The integers p and q are often referred to as orders of the model. Random errors, ε_t , are assumed to be independently and identically distributed with a mean of zero and a constant variance of σ^2. Equation (1) entails several important special cases of the ARIMA family of models. If $q = 0$, then (1) becomes an AR model of order p. When $p = 0$, the model reduces to an MA model of order *q*. One central task of the ARIMA model building is to determine the appropriate model order (*p*, *q*) [9]. In ARIMA stages there is more detail step from choosing model until forecasting step; it is called Box-Jenkins methodology for forecasting. The Box-Jenkins methodology includes three iterative steps of the model identification, parameter estimation and diagnostic checking. This three-step model building process is typically repeated several times until a satisfactory model is finally selected. The final selected model can then be used for forecasting [9].

3. BUILDING ARIMA MODEL FOR GOLD BULLION COIN SELLING PRICES

To fit an ARIMA model, it requires a sufficiently large data set. In this study, we used the data for daily selling prices of Malaysia's own gold bullion coins, Kijang Emas Gold Bullion Coins for the year 2002-2007 [10]. As we have earlier stated that development of ARIMA model for any variable involves primarily three steps: identification, estimation and diagnostic checking. Each of these three steps is now explained for the gold bullion coin selling prices forecasting.

3.1 Model Identification

At the beginning step, the original data was plotted to observe their trends and stationarity. In a case of non stationary, time series has to be transformed to a stationary series before being modelled by the Box-Jenkins approach by taking the first differences of the non stationary time series values. The stationary series is the one whose values vary over time only around a constant mean and constant variance. Figure 1 shows the original data of gold bullion coin selling prices and Figure 2 depicts the first difference of the data series. These figures show that the data are non stationary time series.

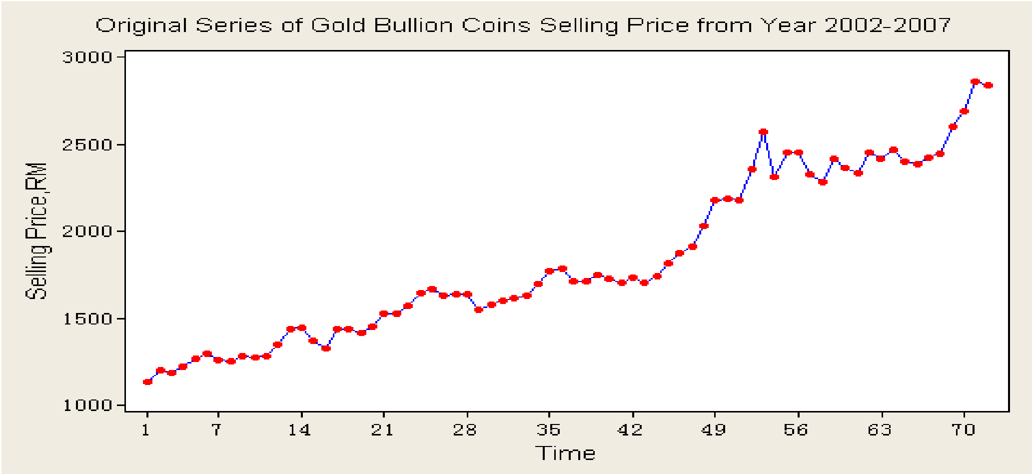


Figure 1. Original series of gold bullion coins selling prices

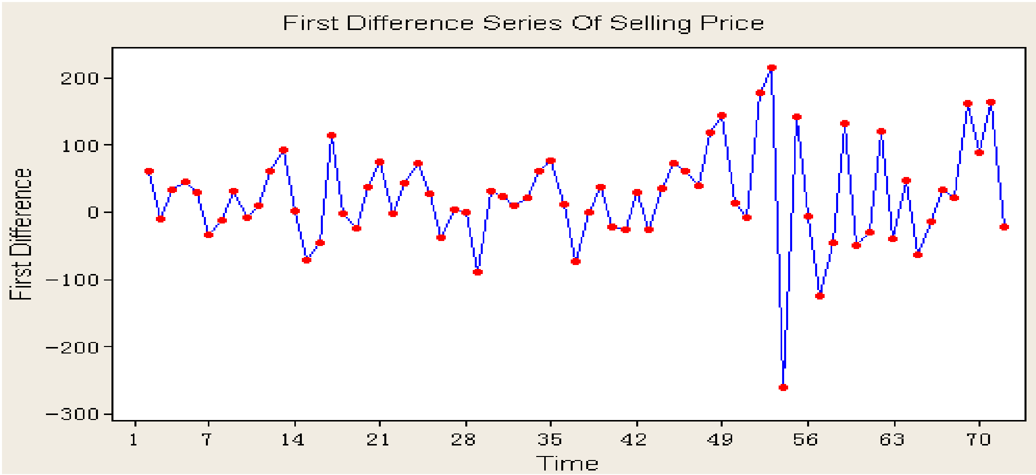


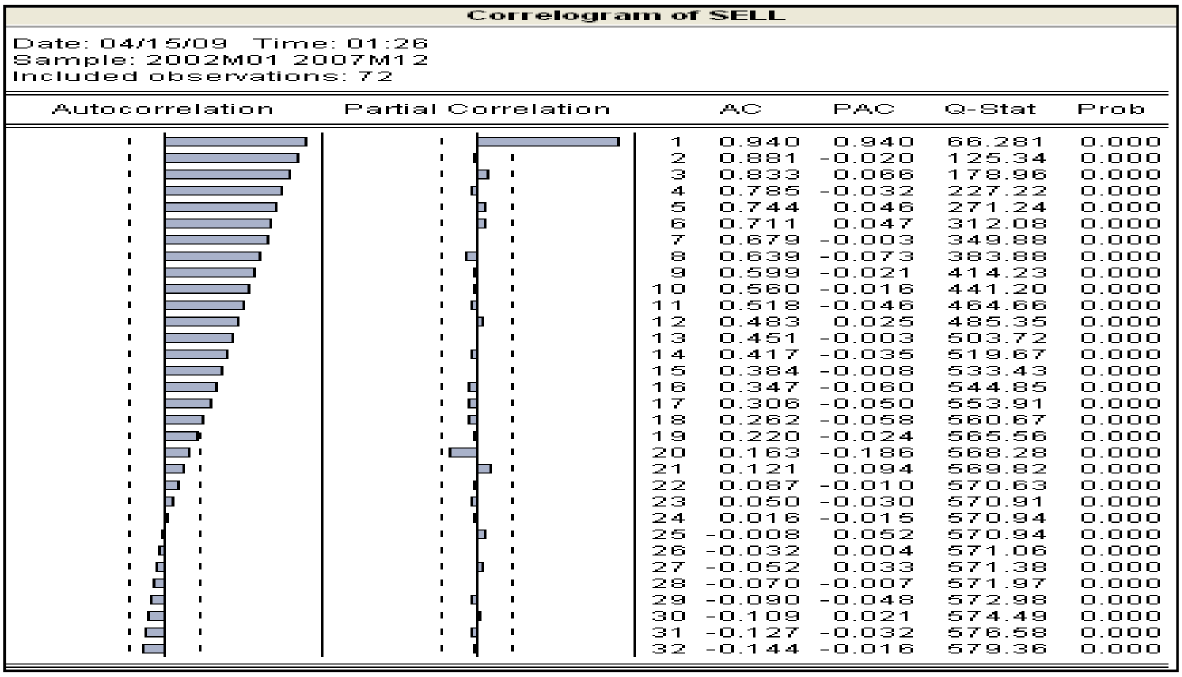
Figure 2. First difference series versus time.

The Autocorrelation (ACF) and Partial Autocorrelation Function (PACF) were also plotted as to collect more conclusive evidence on its stationary condition. The ACF and PACF graphs can be seen in Figure 3. Since the sample ACF values are large and decline rather slowly to zero, therefore, we can conclude that the series is not stationary.

First difference of the original series was then taken and the sample ACF and PACF were transformed. Figure 4 shows the sample ACF and PACF for selling price series in the first difference. From the ACF and PACF analyses we can conclude that the series now is stationary.

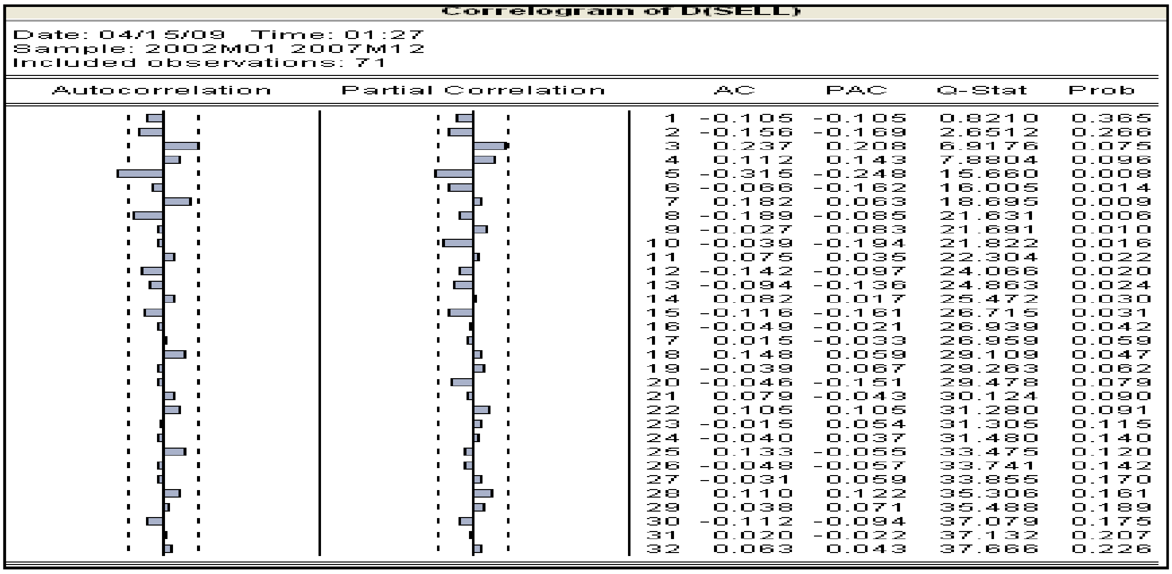
3.2 Parameter Estimation

In ARIMA model, stationary condition must have $|\phi_p| \prec 1$ and invertibility condition must have $|\theta_q| \prec 1$. Besides, ARIMA model also required $|t| \succ 1.96$ and $p - value \prec 0.05$. If not, the model is inadequacy. Table 1 shows ARIMA statistical result for the gold bullion coin selling prices.



	AC	PAC	Q-Stat	Prob
1	0.940	0.940	66.281	0.000
2	0.881	-0.020	125.34	0.000
3	0.833	0.066	178.96	0.000
4	0.785	-0.032	227.22	0.000
5	0.744	0.046	271.24	0.000
6	0.711	0.047	312.08	0.000
7	0.679	-0.003	349.88	0.000
8	0.639	-0.073	383.88	0.000
9	0.599	-0.021	414.23	0.000
10	0.560	-0.016	441.20	0.000
11	0.518	-0.046	464.66	0.000
12	0.483	0.025	485.35	0.000
13	0.451	-0.003	503.72	0.000
14	0.417	-0.035	519.67	0.000
15	0.384	-0.008	533.43	0.000
16	0.347	-0.060	544.85	0.000
17	0.306	-0.050	553.91	0.000
18	0.262	-0.058	560.67	0.000
19	0.220	-0.024	565.56	0.000
20	0.163	-0.186	568.28	0.000
21	0.121	0.094	569.82	0.000
22	0.087	-0.010	570.63	0.000
23	0.050	-0.030	570.91	0.000
24	0.016	-0.015	570.94	0.000
25	-0.008	0.052	570.94	0.000
26	-0.032	0.004	571.06	0.000
27	-0.052	0.033	571.38	0.000
28	-0.070	-0.007	571.97	0.000
29	-0.090	-0.048	572.98	0.000
30	-0.109	0.021	574.49	0.000
31	-0.127	-0.032	576.58	0.000
32	-0.144	-0.016	579.36	0.000

Figure 3. Sample ACF and PACF for original selling price data



	AC	PAC	Q-Stat	Prob
1	-0.105	-0.105	0.8210	0.365
2	-0.156	-0.169	2.6512	0.266
3	0.237	0.208	6.9176	0.075
4	0.112	0.143	7.8804	0.096
5	-0.315	-0.248	15.660	0.008
6	-0.066	-0.162	16.005	0.014
7	0.182	0.063	18.695	0.009
8	-0.189	-0.085	21.631	0.006
9	-0.027	0.083	21.691	0.010
10	-0.039	-0.194	21.822	0.016
11	0.075	0.035	22.304	0.022
12	-0.142	-0.097	24.066	0.020
13	-0.094	-0.136	24.863	0.024
14	0.082	0.017	25.472	0.030
15	-0.116	-0.161	26.715	0.031
16	-0.049	-0.021	26.939	0.042
17	0.015	-0.033	26.959	0.059
18	0.148	0.059	29.109	0.047
19	-0.039	0.067	29.263	0.062
20	-0.046	-0.151	29.478	0.079
21	0.079	-0.043	30.124	0.090
22	0.105	0.105	31.280	0.091
23	-0.015	0.054	31.305	0.115
24	-0.040	0.037	31.480	0.140
25	0.133	-0.055	33.475	0.120
26	-0.048	-0.057	33.741	0.142
27	-0.031	0.059	33.855	0.170
28	0.110	0.122	35.306	0.161
29	0.038	0.071	35.488	0.189
30	-0.112	-0.094	37.079	0.175
31	0.020	-0.022	37.132	0.207
32	0.063	0.043	37.666	0.226

Figure 4 Sample ACF and PACF for selling prices series in the first difference.

Based on these statistical results, selling price ARIMA models (1,1,0), (1,1,2), (1,1,3), (1,1,5), (2,1,1), (2,1,3), (2,1,5), (3,1,1), (3,1,2), and (3,1,4) are rejected because the models failed to fulfill the conditions.

3.3 Diagnostic Checking

In a well fitted model the residuals obtained are expected to have the property of *white noise*. Hence, model validation and diagnostic checking involved analyzing the residual for resemblance of white noise characteristic. More sophisticated technique of establishing the stationary condition of the residuals is to check the Ljung-Box Q statistic. This statistic is used to test the following hypotheses.

H_0 : Errors are random (white noise)

H_1 : Errors are nonrandom (not white noise)

Table 1. Statistical Results For Arima Models

ARIMA Model (p,d,q)	ϕ_p	t-test	p-value	θ_q	t-test	p-value
(1,1,1)	0.8183	15.8572	0.0000	-0.9973	-27.7354	0.0000
(1,1,2)	-0.6404	-1.6031	0.1137	-0.1932	-1.5163	0.1342
(1,1,3)	-0.0398	-0.1042	0.9174	0.3496	2.8004	0.0067
(1,1,4)	-0.5324	-3.0404	0.0034	0.5619	5.0976	0.0000
(1,1,5)	0.0436	0.2404	0.8108	-0.7274	-7.8280	0.0000
(2,1,1)	-0.2353	-1.8891	0.0633	0.3251	0.6240	0.5348
(2,1,2)	-0.9816	-32.4770	0.0000	0.9384	36.5795	0.0000
(2,1,3)	-0.9069	-13.7352	0.0000	0.0677	0.5070	0.6140
(2,1,4)	-0.6222	-4.7406	0.0000	0.5052	3.8728	0.0003
(2,1,5)	0.0428	0.2244	0.8232	-0.7105	-7.3508	0.0000
(3,1,1)	0.2691	1.8240	0.0729	-0.1955	-0.3548	0.7239
(3,1,2)	0.0954	0.6647	0.5087	0.8147	7.2564	0.0000
(3,1,3)	0.8734	10.7426	0.0000	-0.9277	-24.7472	0.0000
(3,1,4)	-0.1402	-0.5611	0.5768	0.5024	3.8848	0.0003

The summary of the various statistics obtained from fitting the model are tabulated in Table 2

Table 2. Summary Of Postmanteau Test Of Selling Price

Statistics	Model ARIMA				
	(1,1,1)	(1,1,4)	(2,1,2)	(2,1,4)	(3,1,3)
AIC	11.3869	11.4063	11.3380	11.3592	11.3512
MAPE	0.4002	0.1432	0.1642	0.0013	0.1482
Calculated Q	20.9	8.0	11.2	8.8	7.5
DF	12	7	7	6	6
Tabulated Q	21.03	14.07	15.51	12.59	12.59
Decision (5% Sig. level)	Accept H_0	Accept H_0	Accept H_0	Accept H_0	Accept H_0

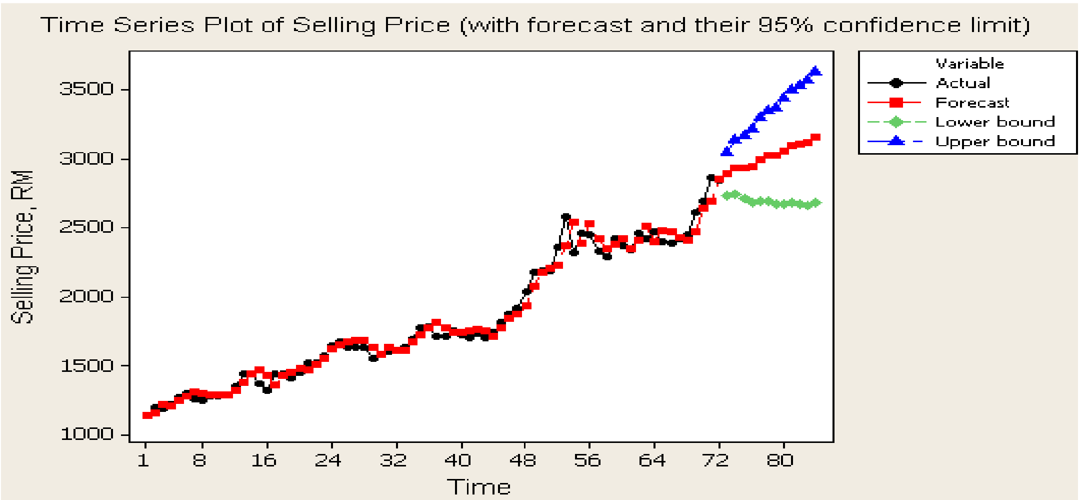


Figure 5. The ARIMA model result of selling price

Checking the values of the calculated Q and comparing against the tabulated values we can accept the null hypothesis that the errors for each of the models are white noise. Hence, the conclusion is that the models are well specified and adequate. In general, a MAPE within 10% is considered very well. However, we only need one fitted model among the five well-specified models. Thus, based on the smallest value of AIC, ARIMA (2,1,2) is the best for selling price data.

3.4 Forecasting

In this paper, ARIMA models are developed to forecast the selling prices of gold bullion coin . There are two kinds of forecasts: sample period forecasts and post-sample period forecasts. The former are used to develop confidence in the model and the latter to generate genuine forecasts for use in planning and other purposes. The gold bullion coin sellingprices for two kinds of forecasts are shown in Figure 5. Efficiency of the forecasting models is checked through percentage of errors. The ARIMA model (2, 1, 2) turns out to be less than 10 percent errors. This measure indicates that the forecasting of gold bullion coin selling prices inaccuracy is low.

4 CONCLUSION

ARIMA model offers a good technique for predicting any fluctuated variables. Its strength lies in the fact that the method is suitable for any time series with any pattern of change and it does not require the forecaster to choose a priori the value of any parameter. However the model has some limitations. One of the limitations is the model requires of a long time series. It is also very often to be called a 'Black Box' model as the mathematical procedures behind model building is not shown. Like any other methods, this technique also does not guarantee perfect forecasts. Nevertheless, it can be successfully used for forecasting long time series data. In this paper, we have developed the ARIMA model for gold bullion coin selling prices. It has been shown that the ARIMA (2,1,2) found to be the best fit model for selling prices of the gold bullion coins. From the forecasted prices that obtain from the developed model, it can be seen that forecasted prices has been trended upward over time.

REFERENCES

[1] G.E.P., and G. M. Jenkins. *Time series analysis: forecasting and control*, Holden Day, San Francisco, 1970.

[2] M., Kumar, A. Kumara, N.C., Mallik, C. and R.K. Shuklaa, "Surface flux modelling using ARIMA technique in humid subtropical monsoon area" , *Journal of Atmospheric and Solar-Terrestrial Physics*, Vol.71, pp.1293-1298, 2009.

[3] J.Lloret, , J.Lleonart, , I.Sole,. "Time series modeling of landings in North Mediterranean Sea", *ICES Journal of Marine Science*, Vol. 57, pp.171-184, 2000.

[4] F.M. Tseng, G.H. Tzeng, H. C. Yu, J.C. Yuan, , Fuzzy ARIMA model for forecasting the foreign exchange market, *Fuzzy Sets and Systems*, 118, 1-11, 2001.

[5] S.A Ediger, "ARIMA Forecasting of Primary Energy Demand By Fuel in Turkey", *Energy Policy*, Vol.35, pp.1-8. 2006.

[6] D. Wood, and B. Dasgupta, "Classifying Trend Movements In The MSCI U.S.A. Capital market Index – A Comparison of Regressions, ARIMA And Neural Network Method", *Computers & Operation Research*, Vol. 23, pp. 611-622, 1996.

[7] E.A., Selvanathan, "A Note on the Accuracy of Business Economists'Gold Price Forecasts, *Australian Journal of Management*, , Vol. 16, pp. 91-94, 1991.

[8] C.A. Wankhade, R. Mahalle, S., Gajbhiye, S. and Bodade V.M. "Use of the ARIMA Model for Forecasting Pigeon Pea Production in India", *International Review of Business and Finance*,. Vol. 2, pp.97-102., 2010.

[9] P.G. Zhang. "Time series forecasting using a hybrid ARIMA and neural network model", *Neurocomputing*. Vol. 50, pp. 159-175, 2003.

[10] Bank Negara Malaysia (2008). http://www.bnm.gov.my/index.php?ch=12&pg=141. Retrieved on 11 Nov 2008.

BIOGRAPHY OF AUTHOR

Mohd Lazim Abdullah is Associate Professor at the Department of Mathematics, Faculty of Science and Technology, University Malaysia Terengganu. He holds a B.Sc (Hons) in Mathematics from the University of Malaya, Kuala Lumpur in June 1984 and the M.Ed in Mathematics Education from University Sains Malaysia, Penang in 1999. He received his Ph.D. from the University Malaysia Terengganu, (Information Technology Development) in 2004. Currently, he is Head of E-learning Unit at the University Malaysia Terengganu. His research focuses on the mathematical theory of fuzzy sets and its applications in social ecology, environmental sciences, health sciences, and education. He is interested in the measurement of social indicators, an index of health and educational constructs using computational intelligence approaches. Analyzing data with computing software is an advantage to him because of his direct involvement in undergraduate teaching in the subject of biostatistics and data analysis. Currently he is a member of editorial boards to several international journals related with computing and information technology. Besides, he is also a reviewer for number of articles to local and international journals, member of scientific committees of several symposia and conferences at national and international levels. Dr Lazim is an associate member IEEE Computational Intelligence Society, a member of Malaysian Mathematical Society and a member of the International Society on Multiple Criteria Decision Making.