KONAKA
KONFERENSI AKADEMIK

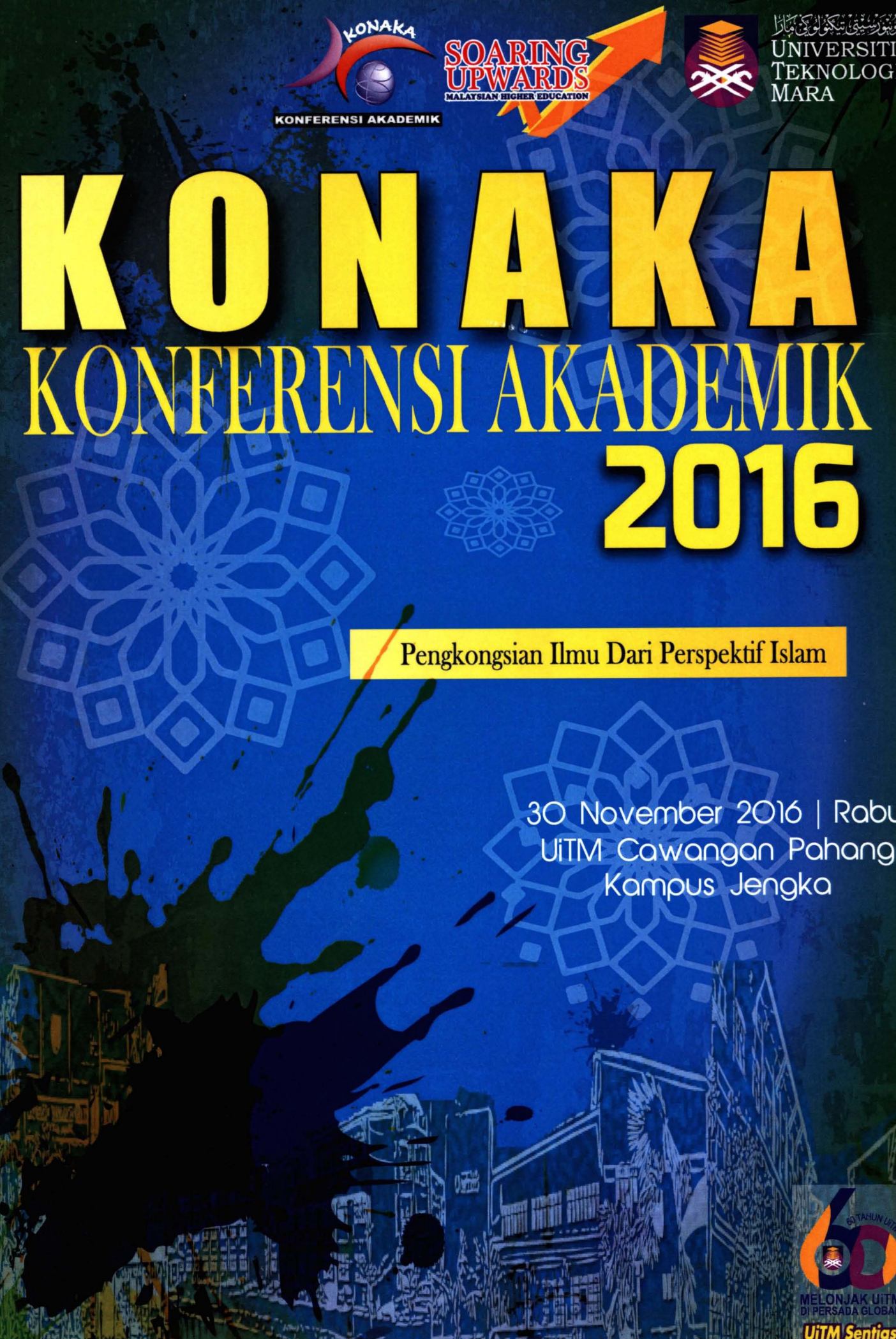
SOARING UPWARDS
MALAYSIAN HIGHER EDUCATION

UNIVERSITI TEKNOLOGI MARA

KONAKA
KONFERENSI AKADEMIK 2016

Pengkongsian Ilmu Dari Perspektif Islam

30 November 2016 | Rabu
UiTM Cawangan Pahang
Kampus Jengka



60 TAHUN UiTM
MELONJAK UiTM DI PERSADA GLOBAL
UiTM Sentiasa

Estimating Malaysia Gold Prices Via Nonlinear Prediction Model

Syazwani Zainal Abidin[1*], Nur Fatihah Haron[2], Siti Fudzla A'ini Mokhtar[3]

[1-3] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Pahang, Kampus Raub, 27600 Raub, Pahang, Malaysia
syazwani_za91@yahoo.com, fatihah3288@gmail.com, sitifudzlaaini@gmail.com[3]
*Corresponding Author

Abstract: This paper focuses on the development of the price of gold. Gold has been considered a safe return investment because of its characteristic to hedge against inflation. Gold prices prediction is an essential tool especially to the investors who have problems for making decisions whether they are supposed to buy or sell the gold because they do not have a specific method to determine exactly when the gold prices will increase or decrease at the current situation. Therefore, the purpose of this study is to search and find the possible chaotic behavior of gold prices by Nonlinear Prediction (NLP) method using theory of chaos involving daily gold prices data in Malaysia from January 2016 to June 2016. Nonlinear prediction method involves the reconstruction of phase space and local linear approximation approach. The phase space reconstruction involves the reconstruction of one-dimensional data into a multidimensional phase space to reveal the dynamics of the system. Revenue of phase space reconstruction was used to predict thedaily gold prices. A comparison of prediction performance based on Root Mean Square Error (RMSE) was employed to compare prediction performance for NLP and Linear Least Square Method (LLSM). Prediction performance comparisons shows that the prediction results using NLP method are better. Thus, NLP method is recommended by using many variables to construct the dimensions. Hence, this paper provides a summary of how price of gold can benefits to the investors to provide a better view of the gold prices movement and gold investment.

Keywords: Gold Prices, Linear Least Square Method, Local Linear Approximation, Nonlinear Prediction Method, Phasespace

1. Introduction

Gold is one of the world most precious metals that cannot be changed by normal chemical and has a symbol of wealth (Yang and Dou, 2014). In transition element, gold is represented with symbol *Au* and atomic number 79. Gold is soft, dense, malleable and ductile metal with a bright yellow color. Gold does not react with air, water and most chemicals so the shine of gold never fades. Commonly, most metals occur in free elemental (native) form, as nuggets or grains, in rocks, in veins and in alluvial deposits. Less commonly, it occurs in minerals as gold compounds like with tellurium as calaverite, sylvanite, or krennerite. Gold has various uses such as to make jewelry, as a component in electronics and computers, medical appliances, aerospace and others.

"Kijang Emas" is the official Malaysian gold bullion coin minted by the Royal Mint of Malaysia. Malaysia is the 12th country in the world to issue its own gold bullion coin and was launched on 17th July 2001. Now, Kijang Emas has joined the ranks of other international gold bullion coins. This gold bullion coin which come in three sizes which are 1 ounce, ½ ounce and ¼ ounce. In addition, it has 999.9 of gold purity which is 24 carats. In Malaysia, Kelantan became the first state to launch gold dinar coins and was introduced on the 20th September 2006 as Kelantan Gold Dinar. The features on its face are the date of production, Kelantan's state crest, its weight and purity. The gold coins can be bought and sold at the Kelantan Corporation Berhad and all eight Ar-Rahnu Islamic pawn shops in the state. However, this type of gold only use in Kelantan (Mohamed, 2010).

According to Goodman (1956), and Solt and Swamson (1981), gold acts as a store of wealth, medium of exchange and unit of value, possessing similar characteristics to money. Gold also can be classified as a monetary asset and a commodity. Gold is also providing high

liquidity, due to the fact that gold can be readily sold or bought anytime the holders want. Hillier, Draper and Faff. (2006), stated that total annual gold production is sold out or cleared in every 2.5 days by London Bullion Market Associated (LBAM). Similarly with Khaemasunun (2007), claims that gold investment commonly used or serve as currency depreciation and a hedge against inflation. Tharmmaphornphilas, Lohasiriwat and Vannasetta (2012) agreed with Khaemasunun (2007) and stated that investors hold or make gold to diverse the portfolio for managing risk intention and act as currency devaluation. In addition, governments apply gold as a relative standard for currency equivalent.

There are many existing models were used in predicting the gold prices. The models which has been used are artificial intelligence models, multiple linear regression models (Ismail et al., 2009), jump-and-dip diffusion (Shafiee & Topal, 2010), varying-coefficient regression model (Zhang et al., 2011), data mining methods (Mehmed, 2011), system dynamics model (Tharmmaphornphilas et al., 2012), back propagation neural networks (Parisi et al., 2008).

This study will contribute an additional knowledge to existing research on the gold prices in Malaysia in order to assist the beginner investors in predicting the future gold prices in Malaysia. First and foremost, investors are encouraged to invest in Kijang Emas because its price depend on the international gold price and not very subjected to inflation. Two simple mathematical methods are used to predict the future prices of gold in Malaysia which are Nonlinear Prediction method (NLP) and Linear Least Square method (LLSM). This study were used the secondary data on gold price collected from Bank Negara Malaysia (BNM) from January till June 2016. The data from January till May 2016 were used to obtain the equation of the models. Hence, to check the accuracy of NLP and LLSM models, the data of June 2016 were used.

Linear Least Square Method (LLSM) is one of the simple mathematical methods that can be used for predicting the future value of gold and also will be used to find the line that minimizes the sum of squares errors.In addition, this method is used to find the best linear relationship between two variables. According to Weisstein (2016), LLSM is the simplest and most commonly applied form of linear regression and provides a solution to the problem of finding the best fitting straight line through a set of points.

Kocak et al. (2004) is study about increasing the time series embedding dimension. The result shows that multivariate time series embedding dimension has low RMSE. On the other hand, they stated that a univariate time series is suffient for the proper embedding. Applicable to a short time series, stable to noise, computationally efficient, and it does not contain any purposely introduced parameters are some advantages of NLP study by Meng et al., (2010). Meng et al., (2010) claimed that the superiority of NLP is usually interpreted as an indication for nonlinear determinism in the data, and method based on predictions power and proposed to differentiate deterministic chaos from noise. From their investigations, their results shows that NLP to be helpful, non-invasive tools to assess directional coupling mechanisms from bivariate analysis of the short-term spontaneous fluctuations of the heart rate and arterial pressure during normal and diseased conditions.

The NLP method was applied by Adenan and Noorani (2013) in order to ensure proper management of water resources that can be optimally distribute water to consumers in Tanjung Tualang, Perak. They extended the NLP methods into two steps which are reconstruction of phase space and prediction using local linear approximation method. In this study, they compare prediction performance based on Correlation Coefficient (CC) and Root Mean Square Error (RMSE) with the other two methods which are ARIMA and Support Vector Machine (SVM). Thus, NLP method is recommended to develop an efficient water management system to optimize the allocation of water resources because this method provides lowest RMSE compared to the others two methods.

2. Methods

This chapter will discuss the implementation of both methods which are NLP and LLSM used in finding the most effective method.

2.1 Nonlinear Prediction Method

Nonlinear prediction method involves two steps which are reconstruction of phase space and prediction of gold prices.

2.1.1 Reconstruction of Phase Space

The first step is the reconstruction of the phase using deterministic data to reveal the dynamics of time series by refereeing at the trajectories in the phase space. Phase space reconstruction is a method to reconstruct the whole, multidimensional dynamics of a system made from one component only (Dietrich, 2008). Attractor of a system can be shown on the trajectories of a system. An attractor is the trajectories focus on a particular sub-space. Observations on the plot attractor in the phase space can give information about chaotic behavior. Chaotic behavior is the entire trajectories lie on an attractor and nearby trajectories diverge (Chatrath & Adrangi, 2001). Reconstruction of the phase space involves reconstruction from a single variable to the m- dimensional phase.

A scalar time series $x(t)$ forms a one-dimensional time series where as;

$$\{x_i\} = \{x_1, x_2, x_3, \ldots, x_N\},$$

where N is the total number of time series. From that scalar time series, m-dimensional signal can be formed. Attractors are reconstructed from time series by the method of time delays. According to the Kostelich and Swinney (1989) and Adenan and Noorani (2013), for a single variable sequence $x_1, x_2, x_3, \ldots, x_N$, a delay construction in m dimensions can be formed by the vector Y_t given as in equation below;

$$Y_t = \{x_t, x_{t+\tau}, x_{t+2\tau}, \ldots, x_{t+(m-1)\tau}\}, \quad (1)$$

where t=1,2,…,L and L=1,2,…,N-(m-1) τ.

From equation (1), the reconstructions of phase space were;

$$\begin{aligned} Y_1 &= \{x_1, x_{1+\tau}, x_{1+2\tau}, \ldots, x_{1+(m-1)\tau}\} \\ Y_2 &= \{x_2, x_{2+\tau}, x_{2+2\tau}, \ldots, x_{2+(m-1)\tau}\} \\ &\vdots \\ Y_L &= \{x_L, x_{L+\tau}, x_{L+2\tau}, \ldots, x_{L+(m-1)\tau}\} \end{aligned}$$

τ is an appropriate time delay (predetermined value) and m is a chosen embedding dimension. In this study, the time delay τ=1 was used because the reconstructed attractor lies nearly on a straight line (Kostelich & Swinney, 1989).

Meanwhile, m-dimensional is varied from 2 to 10 (m=2, 3, 4, …, 10) in order to find the best set of dimensions that can provide a good predictions results. For m> 10, the computed value of the information dimension is relatively constant (Kostelich & Swinney, 1989). RMSE is an average error magnitude in the model performance evaluations (Chai & Draxler, 2014). RMSE for various dimensions was used to determine the m optimal.
MSE for be defined as :

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_j - x_i)^2}{n}} \quad (2)$$

where x_j is the actual observed value at the j^{th} point and x_i is the fitted value at the i^{th} point.

2.1.2 Prediction of Gold Prices

The second step in nonlinear prediction method is prediction of gold prices. Since the gold prices represent the non-linear function, so it is hard to make prediction on future gold prices. The prediction is done on the phase space using Local Linear Approximation method.

Definition 1.0: Local linear approximation

The function $f(x_0)+f'(x_0)(x-x_0)$ is called the ***local linear approximation*** to f at x_0. The function is good approximation to $f(x)$ if x is close to x_0 and the closer the two points are the better approximation becomes.

To predict Y_{t+1}, the nearest neighbor(s) to Y_t was searched. The Euclidean distance between Y_t and the vectors before Y_i (i = 1, 2, …, t-1) is the error between the data (x_i,y_i) to a curve $y=Ax+B$ was calculated by using excel.

Let, Y_M be the minimum distance to the nearest neighbor. In other words, Y_M is the minimum of the function. Value Y_M and Y_{M+1} are used to satisfy a linear equation $Y_{M+1} = AY_M + B$. The value of constant A and B calculated using the least square method. Thus the equation for this model is $Y_{t+1} = AY_t + B$.

2.2 Linear Least Square Method

LLSM is one of the methods that were applied to estimate the gold prices.

$$y_i = a_0 x + a_1 \tag{3}$$

where;

y_i : dependent variable represents the predicted price

a_0 : y – intercept

a_1 : slope value of the linear line

x_i : independent variable represent the time (i.e: $i = 1,2,3,...,n$)

2.3 Judgment of Models Accuracy

The errors of hybrid models without vitality and with vitality can be calculated by using MAPE. The mathematical formula for MAPE is as follows;

$$\varepsilon_{mape} = \frac{1}{M}\sum_{j=1}^{M}\frac{\left|X_p - X_A\right|}{X_p} \tag{4}$$

where;

X_p : the prediction values of gold price

X_A : the actual values of gold price

According to Gundalia and Dholakia (2013) state that the scale of judgment of model accuracy based on MAPE as in Table 1 below.

Table 1. A Scale of Judgment of Forecast Accurac (Lewis)

MAPE	Judgment of Forecast Accuracy
< 10%	Highly accurate
11% to 20%	Good forecast
21% to 50%	Reasonable forecast
>51%	Inaccurate forecast

3. Results and Discussion

3.1 Nonlinear Prediction Method

Stage 1: Reconstruction of the Phase Space

Consider the stationary time series $\{x_i\} = \{x_1, x_2, \ldots, x_N\}$. When two dimensional phase, the time series become $Y_t = \{x_t, x_{t+1}\}$. The data of gold prices from January 2016 to June 2016 were transformed into two dimensional spaces. The table 3.2 shows the two dimensional time series of the data.

Table 2. Two dimensional time series data

t	$Y_t = \{x_t, x_{t+1}\}$	(x_t, x_{t+1})
1	$\{x_1, x_2\}$	(4859,4973)
2	$\{x_2, x_3\}$	(4973,4984)
3	$\{x_3, x_4\}$	(4984, 5124)
4	$\{x_4, x_5\}$	(5124, 5176)
5	$\{x_5, x_6\}$	(5176, 5195)
6	$\{x_6, x_7\}$	(5195, 5116)
⋮	⋮	⋮
119	$\{x_{119}, x_{120}\}$	(5249, 5284)

Hence by using Microsoft Excel, the data showed the chaotic behavior in two dimensional phase space.

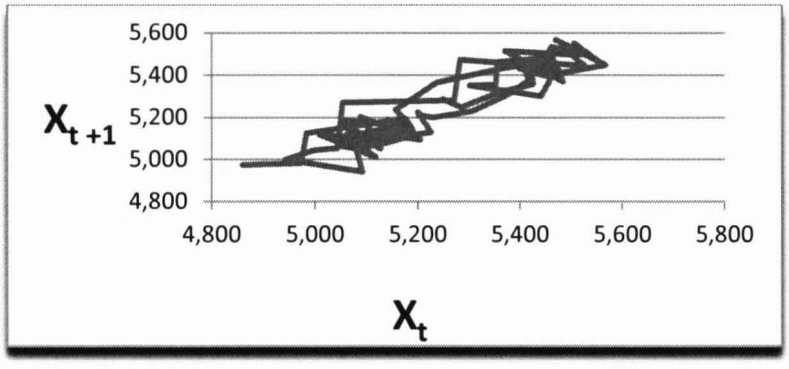


Fig. 1 2-Dimensional phase space

Stage 2: Determine the Optimal Embedding Dimension

The key to perform the reconstruction of the phase space from a real time series is to determine the optimal embedding dimension, m. The phase space reconstruction of Y_t when $\tau = 1$ were shown in Appendix B. In this study, m=2, 3, …,10 was used to find the best

dimension by calculating the RMSE. Below are some of the calculations for embedding dimensions. The steps of obtaining m_{opt} are as below:

To determine the RMSE of 2-dimensional embedding phase space (m=2) can be calculated as below. Hence, for (5094,5094), (5094,5106);

$$\text{RMSE} = \sqrt{\frac{(x_j - x_i)^2 + (y_j - y_i)^2}{n}} \tag{5}$$

$$\text{RMSE} = \sqrt{\frac{(5094-5094)^2 + (5106-5094)^2}{101}} = 1.194045 \tag{6}$$

The following RMSE for multi-dimension ($m = 2,...,5$) were shown in Table 3 below.

Table 3. RMSE for multi-dimension phase space

m	RMSE
2	1.194045
3	1.834758
4	3.011529
5	3.646509

From Table 3, the lowest RMSE value was 1.194045 where the optimum embedding dimension was 2 (m_{opt}= 2).

Stage 3: Estimation of the Gold Prices Using NLP

Since the optimal embedding dimension is 2, the minimum distance to the nearest neighbor was at the two dimensional phase space. The distance in two spaces is the distance in those spaces. The least square method was used to find the constant value of A and B. Mathematically, the NLP model can be formed as below.

$$Y_{t+1} = AY_t + B \tag{7}$$

In order to find the minimum of the function, Y_M the sum of the squares of the errors was determined. The least square criteria were used to find the line that minimizes the sum of the squares errors, S and solving the values for A and B.

After finding the value of A and B, the model obtained for Nonlinear Prediction when $A = 0.8979$ and $B = 542.81$ are as follow;

$$Y_{t+1} = 0.8979Y_t + 542.81 \tag{8}$$

Equation (8) were used to determine the future value of daily gold prices in June 2016 by using NLP method. Hence, the following values was obtained as shown in Table 3.

Table 4. Gold price prediction using NLP method

Date	t	Y_t	Predicted value (Y_{t+1})
31/05/2016	0	5289	5292
01/06/2016	1	5330	5329
02/06/2016	2	5348	5345
⋮	⋮	⋮	⋮
30/06/2016	21	5670	5634

Below are the some of the calculation for the gold price using Linear Least Square Method. After finding the value of a_0 and a_1 by using Microsoft Excel, the model for Linear Least Square Method when $a_0 = 2.0413$ and $a_1 = 5169.8$ are as below:

$$y_i = a_0 x_i + a_1 \tag{9}$$

Table 5. Gold price prediction using LLSM

Date	t	Predicted value
31/05/2016	1	4859
01/06/2016	2	4973
02/06/2016	3	4984
⋮	⋮	⋮
30/06/2016	22	5289

Table 6. Percentage of MAPE obtained using both methods

Type of methods	MAPE
NLP	1.12%
LLSM	6.16%

Table 6 showed that MAPE for NLP gold prices for June 2016 is 1.12% where else the MAPE for gold prices by using LLSM is 6.16%. Its indicates that the MAPE for gold prices with NLP is highly accurate since the errors are less than 10% according to a Lewis scale of judgment of forecast accuracy in Table 1.

4. Conclusion

The calculation that had been derived from the data of gold prices shows that the predicted future value of gold increased and decreased inconsistently. Furthermore, the finding shows that by using Mean Absolute Percentage Error (MAPE), it proves that Nonlinear Prediction (NLP) is the selected predicting model compared to Linear Least Square Model (LLSM) in order to predict the daily price of gold. As a result, it can conclude that it is better for investor to use NLP to predict the future value of gold for type 1 ounce (1 oz.).

5. References

Adenan, N. H., & Noorani, M. S. M. (2013). River flow prediction using nonlinear prediction method. *International Journal of Mathematical, Computational, Physical and Quantum Engineering, 7*, 1065-1068.

Bank Negara Malaysia (BNM). (2013). *Kijang Emas Prices*. Retrieved from May 20, 2014 from http://www.bnm.gov.my/index.php?ch=statistic&pg=statistic_kijangemas.

Chatrath, A., Adrangi, B., & Shank, T. (2001). Nonlinear dependence in gold and silver futures: Is it chaos?. *The American Economist*, 25-32.

Dietrich, J. P. (2008). Phase space reconstruction using the frequency domain: A generalization of actual methods. *Doctoral dissertation, Master's thesis*, University of Potsdam.

Goodman, B. (1956). The price of gold and international liquidity. *The Journal of Finance*, *11*(1), 15-28.

Gundalia, M. J., & Dholakia, M. B. (2013). Modelling pan evaporation using mean air temperature and mean pan evaporation relationship in middle south Saurashtra Region. *International Journal of Water Resources and Environmental Engineering*, *5*(11), 622-629.

Hillier, D., Draper, P., & Faff, R. (2006). Do precious metals shine? An investment perspective. *Financial Analysts Journal*, 98-106.

Ismail, Z., Yahaya, A., & Shabri, A. (2009). Forecasting Gold Prices Using Multiple Linear Regression Method. *American journal of Applied Sciences, 6,* 8, 1509-1514.

Khaemasunun, P. (2007). *Forecasting Thai Gold Prices*. Thailand: College of Innovation, Thammasat University.

Koçak, K., Şaylan, L., & Eitzinger, J. (2004). Nonlinear prediction of near-surface temperature via univariate and multivariate time series embedding. *Ecological modelling*, *173*(1), 1-7.

Kostelich, E. J., & Swinney, H. L. (1989). Practical considerations in estimating dimension from time series data. *Physica Scripta*, *40*(3), 436.

Mehmed, K. (2011). *Data Mining: Concepts, Models, Methods and Algorithms*. New Jersey: John Wiley.

Meng, Q., Zhou, W., Chen, Y., & Zhou, J. (2010, August). Feature analysis of epileptic EEG using nonlinear prediction method. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE* (pp. 3998-4001). IEEE.

Mohamed, S. N. (2010). Decision making for gold investment using fuzzy analytical hierarchy process (AHP) and fuzzy technique for order preference by similarity to ideal solution. *Thesis for Master of Applied Mathematics,* Universiti Teknologi Mara, Shah Alam, Selangor, Malaysia.

Parisi, A. Parisi, F., & Diaz, D. (2008). Forecasting gold price changes: Rolling and recursive neural network models. *Journal of Multi Finance Management, 18,* 477-487.

Shafiee, S. & Topal, E. (2010). An overview of global gold market and gold price forecasting. *Resources Policy, 35*: 178-189. http://dx.doi.org/10.1016/j.resourpol.2010.05.004.

Solt, M. E., & Swanson, P. J. (1981). On the Efficiency of the Markets for Gold and Silver. *Journal of Business, 54,* 453-478.

Tharmmaphornphilas, W., Lohasiriwat, H., & Vannasetta, P. (2012). Gold price modeling using system dynamics. *Engineering Journal*, *16*(5), 57-68.

Weisstein, Eric. W. (2016). *Least Squares Fitting*. Retrived from: http://matworld.wolfram.com/LeastSquaresFitting.html.

Yang, J., & Dou, W. (2014). A new Prediction method of Gold price: EMD-PSO-SVM. *Journal of Software*, *9*(1), 195-202.

Zhang, W. W., Li, G. R., & Xue, L. G. (2011). Profile inference on partially linear varying-coefficient errors-in-variables models under restricted condition. *Computational Statistics and Data Analysis*, *55*, 3027–3040.