Research Journal of Finance and Accounting www.iiste.org
ISSN 2222-1697 (Paper) ISSN 2222-2847 (Online)
Vol.5, No.7, 2014

A Hybrid Gold's Returns Prediction Model Based on Empirical Mode Decomposition

M. Khalid[1,2*], Mariam Sultana[1], Faheem Zaidi[1], Javed Khan[1]
1: Department of Mathematical Sciences, Federal Urdu University Arts,Science & Technology, University Road, Gulshan-e-Iqbal, Karachi-75300, Pakistan,
2: E-mail of corresponding author: khalidsiddiqui@fuuast.edu.pk

Abstract
Consumers have produced extraordinary levels of demand of Gold since the beginning of the financial crisis in 2008 and investment in small coins and bars striking a record high. Since the previous decade, the prices have reached the sky, but the demand for gold remains firm. With such an enormous need for gold coming from whole over the globe, forecast gold prices are of great interest. The main aim of this study is to forecast the price of gold returns, making use of Autoregressive (AR), Empirical Mode Decomposition Autoregressive (EMDAR) and hybrid Empirical Mode Decomposition Autoregressive Neural Network (EMDARNN). The daily data consists of 4837 observations starting from Jan 1995 to June 2013, has been used in this research. After assessing the accuracy of these models by mean absolute error and mean square error, it turns out that hybrid Empirical Mode Decomposition Autoregressive Neural Network excels all the other methods and produces better forecasting with high precision.
Keywords: Gold Price, Autoregressive, Empirical Mode decomposition, Artificial Neural Network

1. Introduction
Gold has been preferred by mankind in several sectors like jewelry, electronics etc.. The price and production behavior of gold differs from most other mineral commodities. Governments hold gold as a standard for currency equivalents. Investors use gold reserves as a hedge against inflation. It is observed normally that the demand and supply of gold do not coincide change in other financial assets (WGC, 2009).For example,in 2008, when the prices of other commodities fell by approximately 40%, the price of gold is increased by 6%. Due to such unique usage, it is not surprising that there will be growth in demand of gold in the future.
The objective of this research is to predict the monetary value of Pakistani Gold returns using Autoregressive (AR), Empirical Mode decomposition Autoregressive (EMDAR) and hybrid Empirical Mode Decomposition Autoregressive Neural Network (EMDARNN). The predicted values are then evaluated by error tests. This study lays a firm ground for the analysis of the problem of forecasting the price of gold returns. Section 2, shows a brief literature review, section 3 discussed methodology used, section 4 shows data description and setup, section 5 demonstrates detailed discussion and results and finally in section 6, the paper is concluded.

2. Literature Review
Various mathematical methods are applied to predicting the gold price. Ball, Torous & Tschoegl (1985), Bertus & Stanhouse (2001) and Hammoudeh, Malik & McAleer (2011) have been analyzed dynamic properties and futures prices of gold spot. A range of different and complex methods used in this respect is mentioned in literature. [Shafiee S. & Topal E. (2010) and Bhar, R. & Hamori, S. (2004)]
Earlier, the authors of this paper, applied wavelet scheme on the Pakistan gold market to predict gold price returns [Khalid & et al 2014]. Pakistan is the eight biggest gold market country in the world. The annual import of gold is approximately 127 tones. In Pakistan, like other countries in the region, gold is the most reliable mean of investment, which offers better returns than fixed deposits. Therefore, as a next step, we will use other techniques on Pakistan Gold market to improve the accuracy of forecasting

3. Methodology
The Time series model is often used to analyze the behavior of any process over a certain time span. It has its applications in weather forecasting, sales forecasting, etc. Time series models are one of the most effective methods of forecasting in the uncertain future decision making. The estimated results obtained from these models have encouraged organizations to develop forecasting techniques to be better disposed to face the seemingly doubtful future. In this study, we will use the hybrid model with the help of these time series methodologies

3.1 Autoregressive Model (AR)
In Autoregression models,the current value of a time series is expressed by a finite linear collection of previous values and by a shock $x_t = A_1 x_{t-1} + A_2 x_{t-2} + \dots\dots + A_p x_{t-v} + \mu_t$, where A_1 to A_p are the autoregression

parameters, μ_t is the white noise and p is the model order. In terms of deviations $\breve{Z} = Z - \mu$, it can be written as $\breve{Z}_t = \phi_1 \breve{Z}_{t-1} + \phi_2 \breve{Z}_{t-2} + \phi_3 \breve{Z}_{t-3} + ...+ \phi_p \breve{Z}_{t-p} + a_1$ with $\mu, \phi_1, \phi_2, \phi_3, ...\phi_p, a_t$ are unknown parameters to be approximated from the observation data. By using the autoregressive operator $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - ....- \phi_p B^p$, the autoregressive model can also express in the form $\phi(B)\breve{Z}_t = a_t$

3.2 Empirical Mode Decomposition (EMD)

Haung 1998 proposed a method of non-linear signal transformation method known as an Empirical Mode Decomposition algorithm Its work is to decompose a non-stationary time series into a sum of intrinsic mode function (IMF). This algorithm is based on constructing smooth envelopes described by local maxima and minima of a sequence and subsequent subtraction of the mean of these envelopes from the primary sequence. This method considers all local extrema which are further attached by cubic spline lines to produce both envelopes, i.e. the upper and the lower envelopes.

The mean produced by the two envelopes is then subtracted from the initial sequence. Hence the whole procedure helps in providing a required empirical function in the first approximation. An intrinsic mode function (IMF) extraction from the EMD shall satisfy only the following requirements.

(1) Number of IMF extrema should be equal to number of zero-crossings or difference should not exceed more than one;
(2) At any point of an IMF the mean value of the envelope defined by the local maxima and the envelope defined by the local minima shall be zero.

The second IMF is obtained by subtracting the previously extracted IMF from the original signal and hence repetition of the above explained methods, and can be continued till all desired IMFs are obtained. When the residue contains no more than two extrema, the sifting procedure stops. The final IMF are obtained when the same operation is applied to the residue signal till the properties of IMF are satisfied.

3.3 Artificial Neural Network (ANN)

Artificial Neural Networks (ANN) have been used to classify, recognize patterns and feature extraction in different fields (Widow et al., 1994). Since they are able to learn and generalize from previous events to recognize future unseen events (Kecman, 2001), therefore also widely utilized for financial forecasting (Sharda, 1994). An Artificial Neural Network (ANN) is a highly interconnected network of several simple processing units called neurons. These neurons are similar to the biological neurons in the human brain. Neurons with analogous features in an ANN are put together in groups called strata. The neurons in one layer are bonded to those in the adjacent strata, but not to those in the same stratum. The intensity of the association between the two neurons in adjacent layers is represented by what is recognized as a 'connection strength' or 'weight'. An ANN generally has three layers, an input layer, a hidden layer and an output layer. The signal produce from the inputs $x_1, x_2, x_3, ...x_n$ that are unidirectional towards the output signal flow(O).This is given by

$$O = f(net) = f\left(\sum_{j=1}^{n} w_j x_j\right)$$

where w_j is the weight vector, and the function $f(net)$ is an activation function. The variable net is defined as a scalar product of the weight and input vectors by

$$net = w^T x = w_1 x_1 + w_2 x_2 + ...+ w_n x_n$$

Where w^T is the transpose of a matrix w.

The output O is computed as

$$O = f(net) = \begin{cases} 1 & if\ w^T x \geq \theta \\ 0 & otherwise \end{cases}$$

Here θ is termed as the threshold level; such a node is called a linear threshold unit. The neurons' internal activity model is given below

$$v_k = \sum_{j=1}^{p} w_{kj} x_j$$

Hence the Neurons' output y_k is the outcome of some activation function on the value v_k

4. Data Description and Setup

The daily data consists of 4837 observations starting from Jan 1995 to June 2013, has been used in this research which is taken from the Karachi Gold Market (http://www.goldrates.pk). To begin with the transformation of the daily gold price into gold returns by using the formula

$$\text{Re}\,turns = \frac{p_{t+1}}{p_t} - 1 = \gamma_{t+1}$$

Then, these returns are directly predicted by Autoregressive model. A hybrid model based on Empirical Mode Decomposition is also applied to forecast the value of gold returns. The hybrid model is illustrated in Figure 1

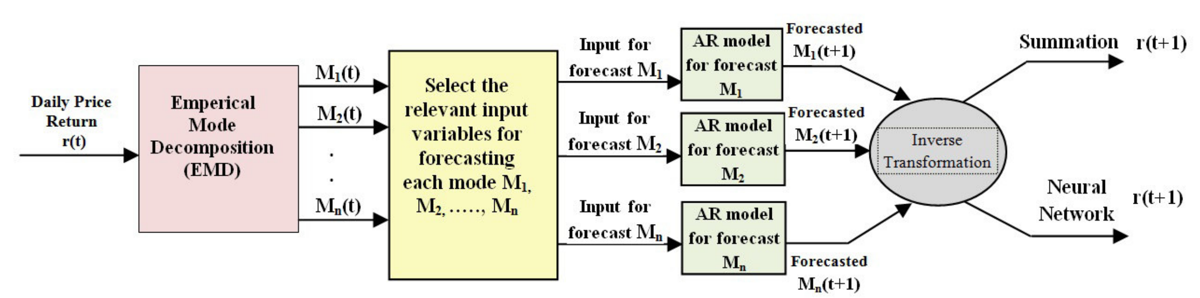

Figure1: Shows Hybrid model based on Empirical Mode Decomposition

In this study, two hidden nodes are used with sigmoid function "Tanh(x)" when we forecasted the gold returns by Artificial Neural Network.

5. Discussion of Results

The forecasting ability of these models is accessed using Mean Square Error(MSE) and Mean Absolute Error (MAE) which are given by

$$\varepsilon_{MSE} = \frac{1}{N}\sum_{1}^{N}(y_{real} - y_{forecast})^2$$

$$\varepsilon_{MAE} = \frac{1}{N}\sum_{1}^{N}\left|y_{real} - y_{forecast}\right|$$

Where y_{real} shows the real gold price, $y_{forecast}$ is the predicted gold price, $\bar{y}$ is the mean of y_{real}, and N is the number of data points.

These errors can be calculated for above mentioned models and comparison are shown in table 1. The results show that hybrid Empirical Mode Decomposition Autoregressive Neural Network (EMDARNN), when trained with sufficient data, proper inputs and with proper architecture, can predict the gold price returns very well.

It is also observed that MAE and MSE, calculates using EMDARNN model,are very less than other methods therefore its prediction is with high precision

	AR	EDM+AR	EDM+NN+AR
MSE	0.0001434	0.0000599	0.0001133
MAE	0.0077263	0.0052588	0.0072913

Table 1: Error analysis of daily gold price returns; Prediction by different forecasted method

Fig 2,3,4 show the regression analysis of all above discussed methods with the original values of the gold price returns. Graphical representation and error analysis, both support that hybrid Empirical Mode Decomposition Autoregressive Neural Network (EMDARNN) model is better than any other model.

6. Conclusion

In this paper, the monetary value of Pakistani gold returns is discussed using Autoregressive (AR), Empirical

Mode Decomposition Autoregressive (EMDAR) and hybrid Empirical Mode Decomposition Autoregressive Neural Network (EMDARNN). From the above, it is noted that the outcomes of the hybrid Empirical Mode Decomposition Autoregressive Neural Network (EMDARNN) model not only beats the rest of the two models, but also out performs the benchmark model.

Acknowledgement:
The authors are pleased to acknowledge the contribution made by Ms Erica Kartha.

References

Ball C., Torous W. and Tschoegl A. (1985), “The degree of price resolution: the case of the gold market", Journal of Futures Markets Vol. 5, pp. 29-43.
Bertus M. and Stanhouse B. (2001), “Rationale speculative bubbles in the gold futures market: an application of dynamic factor analysis", Journal of Futures Markets Vol. 21, pp. 79-108.
Hammoudeh S., Malik F. and McAleer M. (2011), “Risk management of precious metals", Quarterly Review of Economics and Finance Vol. 51, pp. 435-441.
Shafiee S. and Topal E. (2010), “An overview of global gold market and gold price forecasting", Resources Policy Vol. 35(3), pp. 178-189.
Bhar, R. and Hamori, S., (2004) "Information Flow between Price Change and Trading Volume in Gold Futures Contracts," with Ramaprasad Bhar, International Journal of Business and Economics, Vol. 3, pp. 45-56.
Bertus, M., and Stanhouse, B., (2001) “Rational speculative bubbles in the gold futures market: An application of dynamic factor analysis”, Journal of Futures Market, Vol 21, pp. 79-108.
M.Khalid,Mariam Sultana & Faheem Zaidi,(2014) “Forecasting Gold Price: Evidence from Pakistan Market.”, Research Journal of Finance and Accounting, Vol. 5(3), pp.70-74
Huang, N.E., Shen, Z., Long, S., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C.,Tung, C. C., Liu, H. H., (1998) “The empirical Mode Decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”. Procedure of the Royal Society of London,Vol . 454, 903-995
Kecman, V., (2001), “Learning and soft computing” The MIT Press Cambridge, Massachusetts London, England.
Sharda, R., (1994). “Neural networks for the MS/OR analyst: An application bibliography”,Interfaces, Vol. 24 (2), 116–130.
Widrow, B., Rumelhart, D.E., Lehr, M.A., (1994). “Neural networks: Applications in industry, business and science.” Communications of the ACM ,Vol. 37 (3), 93–105
WGC, 2009. Gold investments digest. www.gold.org: World Gold Council.

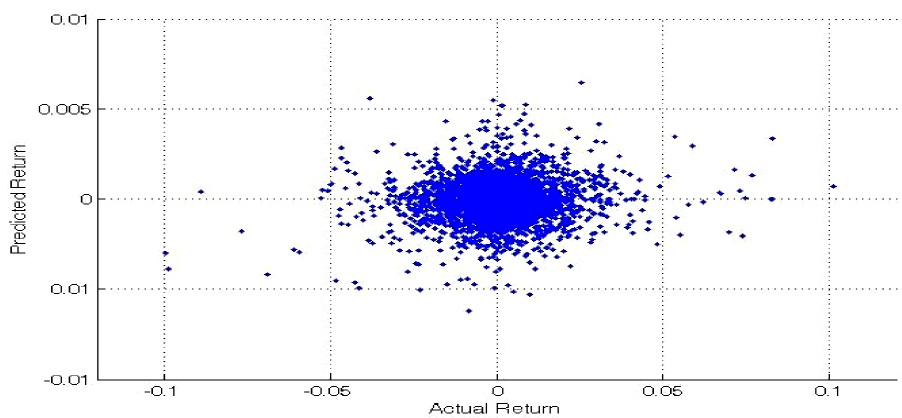

Figure 2: Shows regression between actual price returns and predicted price daily returns of gold by Autoregressive

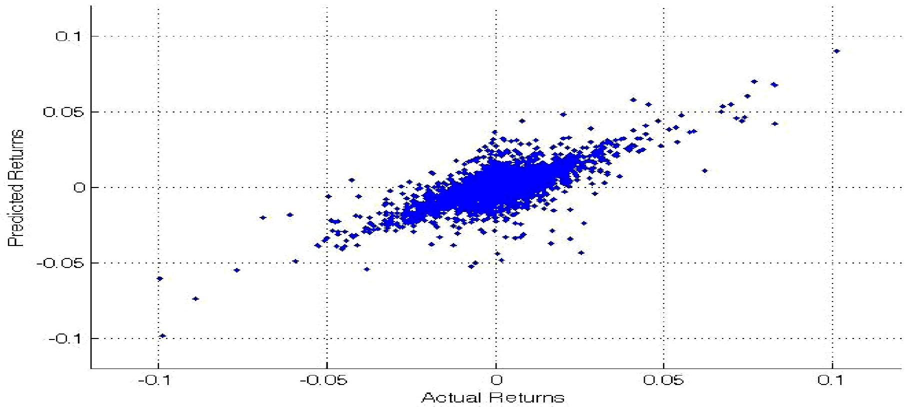

Figure 3: Shows regression between actual price returns and predicted price daily returns of gold by Empirical Mode Decomposition Autoregressive

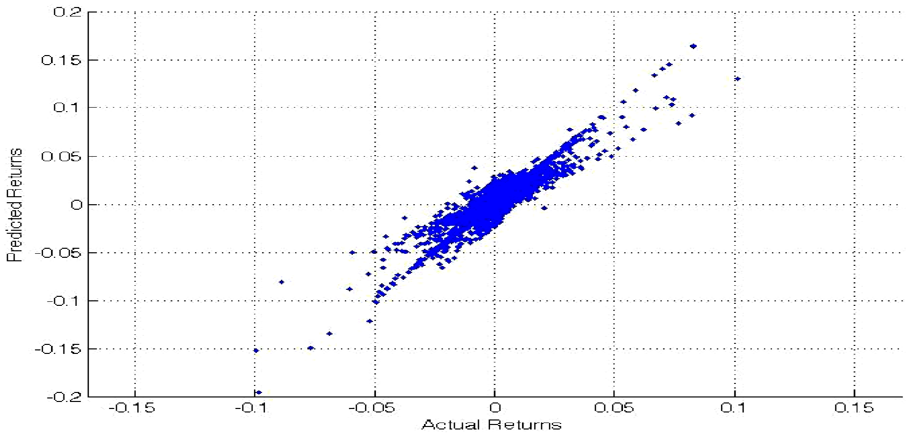

Figure 4: Shows regression between actual price returns and predicted price daily returns of gold by hybrid Empirical Mode Decomposition Autoregressive Neural Network (EMDARNN)

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Recent conferences: http://www.iiste.org/conference/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar