RESEARCH ARTICLE OPEN ACCESS

Implementing Method of Ensemble Empirical Mode Decomposition And Recurrent Neural Network For Gold Price Forecasting

Sri Herawati[1], Firmansyah Adiputra[2], M. Latif[3], Aeri Rachmad[4]
[1,2,3,4] *Faculty of Engineering – University of Trunojoyo Madura, Indonesia*
Corresponding author: Sri Herawati

ABSTRACT
Gold becomes one of long-term investment options and is used as a value protection against inflation or declining other assets. These gold price fluctuations tend to be nonlinear and uncertain. Most researchers and business practitioners fail to produce consistent pricing analyses, due to the complexity of the dynamic and volatile gold market. One method that can accommodate gold price fluctuations is using Ensemble Empirical Mode Decomposition (EEMD). Furthermore, the results of the gold price analysis can be used in forecasting. Forecasting fluctuations in gold prices are needed by importers, investors, and society to reduce risks and to help in making decision. The forecasting which has been done is the integration between EEMD and Feed-forward Neural Network (FNN) with good forecasting results. However, the use of FNN is less flexible for the use of free parameters, such as the type of activation function, initial initialization, number of input neurons, and output neurons. The setting of flexible free parameters can affect the performance of neural networks and improve forecasting accuracy. One way to overcome the weaknesses of FNN in the use of free parameters is, it can use the Recurrent Neural Network (RNN). The trial in this study is using monthly data of world gold price. The results proves that the performance of EEMD-RNN method forecasting is better than EEMD-FNN.
Keywords: Forecasting gold prices, Ensemble Empirical Mode Decomposition (EEMD), Recurrent Neural Network (RNN).

I. INTRODUCTION

Gold is one of the precious metals which is a long-term investment choice. In addition, gold is useful as a value protection against inflation [1] or decreasing other assets, such as: stocks, bonds, and foreign currencies [2]. The benchmark of domestic gold price refers to the gold price of the world which is divided into two types, namely gold fix and spot price. The fluctuation of gold price is influenced by two macroeconomic variables namely oil price and inflation rate [3]. The price of oil and gold has a high correlation of about 85%, while the price of gold and the rate of inflation has a less significant correlation of about -9% [4].

Gold price fluctuations are very important to understand, because they have non-linear and uncertain characteristics [5]. Most researchers and business practitioners fail to produce consistent pricing analyzes, due to the complexity of the dynamic and unstable gold market. One method that can accommodate gold price fluctuations is using Ensemble Empirical Mode Decomposition (EEMD). EEMD is a modification of Empirical Mode Decomposition (EMD) with the addition of white noise [6]. EEMD is considered more accurate and effective in analyzing the data because it is empirical, intuitive, direct and adaptive [7]. Furthermore, the results of the gold price analysis can be used in forecasting. Forecasting fluctuations in gold prices are needed by importers, investors, and the public to reduce risks and assist decision making.

Forecasting the price of gold has been done by some previous research. Forecasting uses the method of Multiple Linear Regression (MLR) to study the relationship between the price of gold with economic factors [8]. The macroeconomic effect on gold prices was investigated using asymmetric GARCH model [3]. These forecasting methods can provide good results when the price series tends to be linear. However, fluctuations in the price of real gold tend to have nonlinear and uncertain characteristics. Thus, forecasting performance results in poor forecasting and cannot capture hidden patterns of nonlinear when using traditional statistical and econometric methods [9].

To overcome the problem of gold price characteristics that tend to be nonlinear and uncertain, some forecasting are using artificial neural networks. For example, forecasting integrates EEMD and Feedforward Neural Network (FNN) neural networks with good forecasting results [10]. However, the use of FNN has a less flexible drawback in the use of free parameters.

One way to overcome the weakness of FNN in the use of free parameters is by using Recurrent Neural Networks (RNN). RNN is a more flexible type of neural network with the addition of feedback (feedback) from the output back to the input. The advantages of RNN also have internal memory from previous inputs that are adaptive, so RNN has better computing capabilities and faster convergence than FNN [11]. Therefore, the proposed research is to implement EEMD and RNN to improve the accuracy of gold price forecasting.

II. SYSTEM DESIGN

The research method by integrating EEMD and RNN is done as in Figure 1.

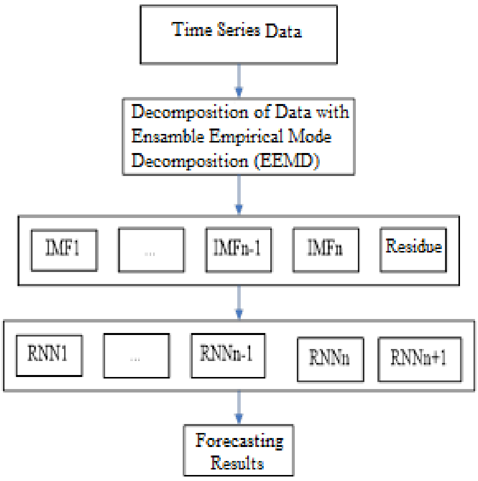

Figure 1. Research Methods

Based on Figure 1, it can be explained as follows:

a. Time series data

Time series data can be used to estimate future events, as the pattern of data changes in some past periods can be repeated in the present. This study was using daily data of world gold prices which were obtained from the site of The London Bullion Market Association (LBMA). LBMA is a trade association in determining the gold standard price recognized worldwide.

b. Ensemble Empirical Mode Decomposition(EEMD)

The data decomposition step with EEMD is done with EMD modification. Basically the process of modification is related to the decomposition of gold prices into several IMF and residues through the addition of white noise.

c. Recurrent Neural Network (RNN)

Recurrent neural networks (RNNs) is one type of artificial neural network. RNN has two types: Elman Network and Hopfield Network [12]. Elman Network has feedback connections to learn, recognize, and create network patterns. While Hopfield Network has one or more target vectors as memory that will be called when there is a vector similar to the network. The RNN architecture is shown in Figure 2.

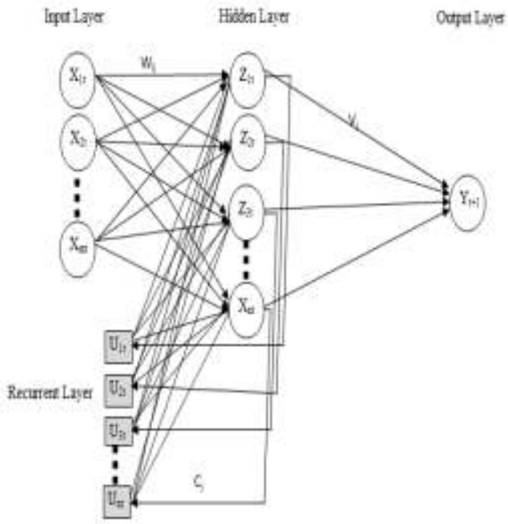

Figure 2. RNN Architecture

In Figure 2 it can be seen that the RNN has four layers, namely the input layer, the hidden layer, the recurrent layer, and the output layer [13]. All neurons for RNN can be calculated with equations (1), (2), and (3).

$$net_{jt}(k) = \sum_{i=1}^{n} w_{ij} x_{it}(k-1) + \sum_{j=}^{m} \quad (1)$$

$$u_{jt}(k) = z_{jt}(k-1), i = 1,2,\ldots,n, j = 1 \quad (2)$$

$$z_{jt}(k) = f_H(net_{jt}(k)) = f_H(\sum_{i=1}^{n} w_{ij} x_{it}(k) + \sum_{j=1}^{m} c_j u_{jt}(k \quad (3)$$

Where xit denotes the set of vector input neurons at time t, y t + 1 represents the output of network at time t + 1, zjt represents the output of the hidden layer neuron, and ujt represents the output of the recurene layer neuron. Wij denotes the weights of connecting nodes i in the neurons of the input layer to node j in the hidden layer. Cj, vj is the weight that connects node j in the hidden layer neuron to the recurrent layer node and the output layer.

d. Forecasting Results

The result of forecasting is obtained from the error performance function in RNN. This performance function can use Mean Square Error (MSE) and Root Mean Squared Error (RMSE). The equations for MSE and RMSE are shown in equation (4) and equation (5). At denotes actual data at time t, Ft is forecasting data at time t, and n represents the amount of data.

$$\mathrm{MSE} = \frac{1}{n} \sum_{t=1}^{n} (F_t - A_t)^2 \quad (4)$$

$$\mathrm{RMSE} = \sqrt{\sum_{t=1}^{n} (F_t - A} \quad (5)$$

III. RESULT AND DISCUSSIONS

The trial data uses the monthly gold price data from The London Bullion Market Association

(LBMA) website. This study used data consisting of 240 months (January 1997 to December 2016). Train data were using the train data as much as 168 and the test data as much as 72.

Trials used performance comparisons between forecasting methods that integrate EEMD and FNN (EEMD-FNN) with EEMD and RNN (EEMD-RNN). The performance comparison of forecasting results from both methods was based on MSE and RMSE values which were generated for various data patterns representing the architecture of artificial neural networks used.

Implementation of trials was performed to find the best forecasting results by determining artificial neural network through variations of input data patterns (number of neurons in input layer, number of neurons in hidden layer, and number of neurons in output layer). The test steps were implemented as follows:

- Data retrieval time for gold prices was using the data from the LMBA website at http://www.lmba.org.uk.
- Decomposition of gold price data was using EEMD with specified threshold and tolerance values [value, miner2, tolerance] = [0.05, 0.5, 0.05]. In this study, the number of ensemble members used the iteration of 100 and the standard deviation of white noise of 0.2. Implementation of decomposition was using algorithm in sub chapter 2.1.1. The decomposition process produced five IMFs and one residue as in Figure 3.
- After decomposition, the next step was to normalize the data into the range between 0 and 1. Normalization is required to get data in the same range. Subsequently, normalized results data were used for EEMD-RNN training and testing. Trials were performed using variations of input data patterns to get the best forecasting results.
- The output of EEMD-RNN for all IMF and residues was using Adaptive Linear Neural Network (Adaline). The merger is done to get forecasting results.
- The last stage was to normalize the data. Dennisalisasi important data was done for data forecasting results in the same value with actual data.

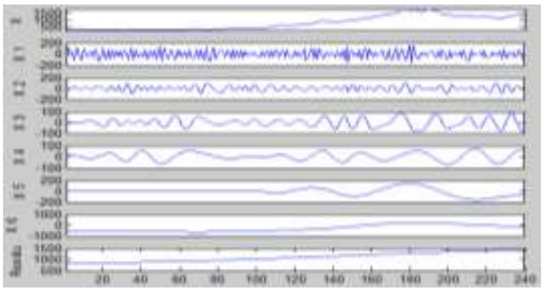

Figure 3 Data decomposition Monthly gold price using EEMD

Forecasting results with variations of input data patterns for gold prices can be shown in Table 1, and Table 2. Then, the results from forecasting that integrate EEMD and RNN were compared with EEMD and FNN integration as shown in Table 3.

Table 1 Comparison of Variations of Number of Neurons with Three input data patterns

Data Patterns	EEMD-RNN	
	MSE	RMSE
3-3-1	0,007975	0,089301
3-4-1	0,007131	0,084445
3-5-1	0,029968	0,173112
3-6-1	0,042545	0,206265
3-7-1	0,028268	0,168130
3-8-1	0,025653	0,160165
3-9-1	0,006059	0,077843
3-10-1	0,006311	0,079442

Table 2 Comparison of Variations of Number of Neurons with Data Pattern Input Six

Data Patterns	EEMD-RNN	
	MSE	RMSE
6-3-1	0,019386	0,139232
6-4-1	0,011663	0,107996
6-5-1	0,033725	0,183643
6-6-1	0,013751	0,117263
6-7-1	0,010089	0,100443
6-8-1	0,011095	0,105332
6-9-1	0,003460	0,058826
6-10-1	0,008697	0,093257

Table 3. Comparison of Forecasting Results using EEMD-RNN and EEMD-FNN

Data Patterns	EEMD-RNN		EEMD-FNN	
	MSE	RMSE	MSE	RMSE
3-9-1	0,006059	0,077843	0,023797	0,154262
6-9-1	0,003460	0,058826	0,009474	0,097336
9-9-1	0,009474	0,097336	0,00668	0,081729

From Table 1 it can be seen that there are table cells are grayed out for trials of three input data patterns. The gray-colored value represents the cell that has the smallest MSE and RMSE values with the number of hidden neurons as many as nine. This data pattern yields MSE of 0.006059 and RMSE of 0.077843. Meanwhile, based on Table 2 it can be seen that the number of hidden layer neurons which are suitable for network architecture with six data patterns are nine neurons. This data pattern produces MSE of 0.003460 and RMSE of 0.058826.

Table 3 shows that the EEMD-RNN method produces better MSE and RMSE than the EMD-FNN method for the three data patterns tested. Table 5.3 shows that the best performance for gold price data is obtained for data patterns 6-9-1 with MSE and RMSE values respectively of 0.003460 and 0.058826. Thus, based on the experimental data, pattern data input is using data patterns 6-9-1 (six input neurons, nine hidden neurons, and one output neuron). The comparison graph of EEMD-RNN forecasting results with actual data can be seen in Figure 4, whereas the EEMD-FNN order results with actual data can be seen in Figure 5. Then, Figure 6 shows how close the data forecasting process results with the actual data for both methods compared to actual data.

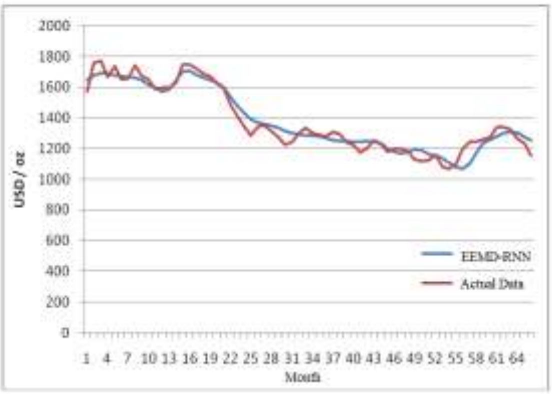

Figure 4. Graph of comparison of forecasting results for EEMD-RNN with actual data of gold price

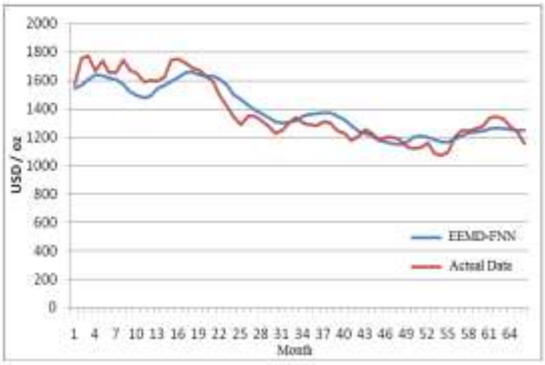

Figure 5. Graph of comparison of forecasting results for EEMD-FNN with actual gold price data

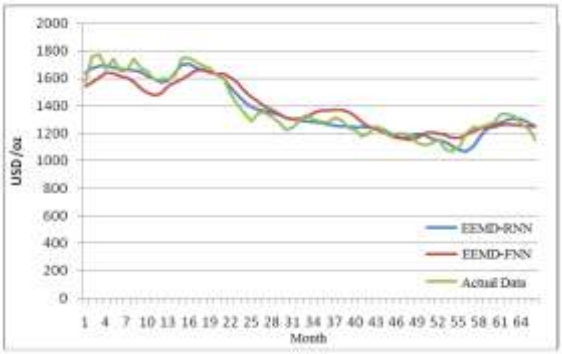

Figure 6. Comparison of actual data denagan forecasting data using gold price data (EEMD-RNN and EEMD-FNN)

IV. CONCLUSION

Based on the results and trial analysis, it can be concluded as follows:

a. From the experimental results in comparing several data patterns variations of neural network input, it can be concluded that the best performance by using data patterns with the number of hidden neurons is nine. This can be from the smallest MSE and RMSE values for data patterns with the number of input neurons 3, 6, and 9.

b. From the result of comparison test with other peramlan method, it is concluded that the forecasting method that integrates EEMD method with RNN (EEMD-RNN) in the research gives better result compared with EEMD and FNN (EMD-FNN) combined method. This is based on the smaller value of MSE and RMSE. The best MSE and RMSE values were 0.004560 and 0.058826 respectively.

REFERENCES

[1]. Beckmann, J., & Czudaj, R. (2013). Gold as an Inflation Hedge in a Time-Varying Coefficient Framework. *North American Journal Of Economics and Finance , 24*, 208-222.

[2]. Gangopadhyay, K., Jangir, A., & Sensarma, R. (2016). Forecasting the Price of Gold: An Error Correction Approach. *28*, 6-12.

[3]. Tully, E., & Lucey, B. M. (2007). A Power GARCH Examination of Gold Market. *Reseach in International Business and Finance , 21* (2), 316-325.

[4]. Shafiee, S., & Topal, E. (2010). An Overview of Gobal Gold Market and Gold Price Forecasting. *Resources Policy , 35*, 178-189.

[5]. Hui, Y., & Wei, D. (2012). Prediction of Gold Price Based on WT-SVR and EMD-SVR Model. *Eighth International Conference on Computational Intelligence and Security (CIS)*, (hal. 415-419).

[6]. Latif, M., & Herawati, S. (2016). The Application of EEMD and Neural Network

Based on Polak-Ribiere Conjugate Gradient Algorithm for Crude Oil Prices Forecasting. *Matec Web of Conferences*, *58*.

[7]. Xian, L., He, K., & Lai, K. (2016). Gold Price Analysis based on Ensemble Empirical Mode Decomposition and Independent Component Analysis. *Physica A* , *454*, 11-23.

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

[9]. Yu, L., Wang, S., & Keung, K. (2008). Forecasting Crude Oil Price With an EMD-Based Neural Network Ensemble Learning Paradigm. *Energy Economics* , *30*, 2623-2635.

[10]. Herawati, S., & Djunaidy, A. (2014). Forecasting Crude Oil Price Using Combined Ensemble Method Empirical Mode Decomposition (EEMD) and Artificial Neural Network . *Jurnal Simantec* , *4* (1), 60-69.

[11]. Cacciola, M., Megali, G., Pellicano, D., & Morabito, F. (2012). Elman Neural Network for Characterizing Voids in Welded Strips : a Study. *Neural Computing and Applications* , *21* (5), 869-875.

[12]. Demuth, H., & Beale, M. (2002). *Neural Network Toolbox.* The Mathwork Inc, Massachusets.

[13]. Wang, J., & Wang, J. (2016). Forecasting Energy Market Indices With Recurrent Neural Networks : Case Study of Crude Oil Price Fluctuations. *Energy* , *102*, 365-374.

Sri Herawati. "Implementing Method of Ensemble Empirical Mode Decomposition And Recurrent Neural Network For Gold Price Forecasting" International Journal of Engineering Research and Applications (IJERA), vol. XX, no. XX, 2017, pp. XX–XX.