International Journal of Engineering & Technology, 7 (4.40) (2018) 136-140

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Predicting Gross Domestic Product Using Weighted Exponential Moving Average on Phatsa Web Application

Marcel Bonar Kristanda[1]*, Seng Hansun[2]

[1] *Informatics Department, Universitas Multimedia Nusantara, Indonesia*
[2] *Informatics Department, Universitas Multimedia Nusantara, Indonesia*
**Corresponding author E-mail:marcel.bonar@umn.ac.id*

Abstract

Gross Domestic Product (GDP) has been known as multi criteria measurement of market value from goods and services across nation. Even there has been an improvement study on accuracy recently, this subject still opens a huge potential of future research and method development and implementation. At this study, we compared a new approach on Moving Average method called Weighted Exponentially Moving Average (WEMA) that created in 2014. We implement the method on Phatsa Web Application and evaluate the results by calculating its Mean Squared Error (MSE), Mean Absolute Scaled Error (MASE), and Mean Absolute Percentage Error (MAPE). Comparison study with the conventional moving average methods named Simple Moving Average (SMA), Weighted Moving Average (WMA) and Exponentially Moving Average then conducted to review the capability of WEMA in predicting Indonesia's GDP data. It is proven that WEMA can provide more accurate results rather than SMA and WMA, but it still has the same error rate value if compared to EMA with more computation time. In addition to the result, Czech's GDP data set is also tested and provides more information on how WEMA calculation with expanded data set, and more volatile value compared to Indonesia's GDP in Czech's GDP data set.

Keywords: *forecasting;gross domestic product;time-series;web application;weighted exponential moving average;*

1. Introduction

Gross Domestic Product (GDP) has been known as multi criteria measurement of market value from goods and services across the nation. Since 2010, government central bank of Indonesia, Bank Indonesia (BI), has been monitoring GDP in quarterly periods and regularly updated in Special Data Dissemination Standard (SDDS) [1]. Based on the data published, fluctuation in this time-series data is influenced by 17 business sectors as well as employment and unemployment rates. The importance of predicting GDP's value has been studied by researchers to indicate the future economic condition in a country. Several methods and models have been implemented in researches study to predict its value or its growth level or percentage[2]–[4]. Even there has been an improvement study on accuracy recently [3], this subject still opens a huge potential of future research and method development and implementation.

In 2014, Hansun [5] developed a new method in the time-series analysis that calculates prediction data with collaboration between two approaches: Weighted Moving Average (WMA) and Exponential Moving Average (EMA). This method has a different approach to Exponentially Weighted Moving Average (EWMA), where the main difference lays on base calculation of past data values that was precalculated with WMA. The proposed method called Weighted Exponential Moving Average has been tested in forecasting Jakarta Stock Exhange (JKSE) value the evaluation of Mean Absolute Percentage Error (MAPE) and Mean Square Error (MSE). Further studies then showed that it can be implemented on smoothing algorithm [6], foreign-currency value prediction [7] and even it was developed into a new method in the previous year[8].

Phatsa is a web-based application that created for forecasting generic data using conventional moving average methods[9]. Recent development for this tool targets the WEMA implementation on its calculation engine. It was evaluated in performance analysis of forex forecasting[10], and it can calculate three kinds of error rates, MSE, MAPE and Mean Absolute Squared Error (MASE).

Based on those three main ideas, this study wants to evaluate on how GDP prediction performs in Phatsa Web Application that implements WEMA for its calculation. Several factors in GDP growth value also calculated and evaluated using error rates in the application.

2. Weighted Exponential Moving Average

This section consists of the WEMA method calculation on data prediction, including step-by-step calculation of the method and on how it is related to WMA and EMA. Second part of this section is the explanation of error rate evaluations: MSE, MAPE and MASE with their respective formulas. The last part contains performance evaluation on Phatsa Web Application based on time consumed for the WEMA and error evaluations.

2.1. Calculation Method

First, given time-series data (x), WMA formula can be used to predicts data $(\bar{x}_{t+1})$ in a future time $(t+1)$ by adding periodic sequential (i.e. daily, weekly, monthly, yearly, etc) past data $(x_t, x_{t-1}, \dots, x_{t-s+2}, x_{t-s+1})$ in a time span (s) that multiplied by each weighting factor $(t, t-1, \dots, t-s+2, t-s+1)$, then divide it by the sum of weighting factors. In this study, weighted factors presented as their index value of an array, assuming the array of data already sorted from oldest data (index = 1) to the recent one (index = t). Formula that written on a study by Zhuang, et al. [11] then adopted to a single sensor (single type value) and index-based weighting value in Definition 2.1. The formula works well when $t \geq s > 1$, where $t, s \in \mathbb{N}$.

Definition 2.1: Weighted Moving Average Formula

$$\bar{x}_{t+1}^{WMA} = \frac{\sum_{n=t-s+1} \; x_n}{\sum t} \quad (1)$$

Second, the result from WMA calculation then converted to a base value $(\bar{x}_t^{WMA})$ for EMA formula in predicted data $(\bar{x}_{t+1})$ of a future time $(t+1)$ that can recursively calculated[12] as it can be written in the Definition 2.2.

Definition 2.2: Weighted Exponential Moving Average Formula

$$\bar{x}_1 = x_1, \quad (2)$$
$$for\ t > 0, \qquad \bar{x}_{t+1} = \alpha x_t + (1-\alpha)\bar{x}_t^{WMA}$$

In this formula, α represents exponential degree of decreasing data. This study use controlled iteration to find best α value ranging from 0, 0.01, 0.02, 0.03, …, 1 (101 loops) that has minimum MASE rate value on each loop.
The last step of this calculation is going back to WMA calculation in first step until every data point in the given time-series data set is calculated by Definition 2.2. It reaches its end when there are no previous data anymore $(t = last\ given\ period\ on\ data\ set)$.

2.2. Error Rates Evaluation

Mean Squared Error (MSE) evaluates the quality of predictor by adding every value difference between prediction results set $(\bar{x})$ and actual value set (x) ranging from starting index (k) to last calculated prediction index (t), powered by 2, then divide it by the number of data quantity in the given set (t). As given in the study of MSE[13], Definition 2.3 shows, the formula that works in this application where a user can set their preference about start index (k).

Definition 2.3: Mean Squared Error Formula

$$MSE_{(x,\bar{x})} = \frac{1}{t}\sum (x_i - \bar{x}_i)^2 \quad (3)$$

Other errors rate component, Mean Absolute Scaled Error (MASE) and Mean Absolute Percentage Error (MAPE) are the measurement rate that represent prediction accuracy for a forecasting method. In 2006, Hynderman[14] proposed the scaled error to be an alternative measurement, for any time-series situation, to relative measurement that has questionable ability of handling certain condition in time-series, such as out-of-sampling data. The idea of MASE (shown in Definition 2.4) is calculating the mean of differences between predicted data $(\bar{x}_i)$ and original data (x_i) from a starting index (k) in calculated set (from k to t) which divided by scaled sum of all differences between current original data (x_i) and the previous one (x_{i-1}) in the complete series set (from $j = 2$ to n).

Definition 2.4: Mean Absolute Scaled Error Formula

$$MASE_{(x,\bar{x})} = \frac{\qquad n-1 \sum_{j=2} \quad x_{j-1}}{\qquad} \quad (4)$$

Mean Absolute Percentage Error (MAPE) is formed as scale-independent and it has a drawback when original data is equal to zero [14]. The formula described in Definition 2.5 as the sum of error that divided by original data in a certain time, then multiplied by 100 to get the percentage.

Definition 2.5: Mean Absolute Percentage Error Formula

$$MAPE_{(x,\bar{x})} = \frac{\qquad x_i \qquad}{\qquad} * 100\% \quad (5)$$

2.3. Time Performance Calculation

Time performance is benchmarked in 3 types of form:

1. Calculating Average Time (CAT) is time consumed in mili-seconds for the application to compute the WEMA forecasting method.
2. Calculating Error Time (CET) is the time consumed in mili-seconds for the application to compute MSE, MASE and MAPE error evaluation.
3. Overall Calculating Time (OCT) is the time consumed from the beginning of moving average calculation until the application gets the error evaluation results.

The timer for this study is provided by CodeIgniter benchmark class[15] that runs on certain hardware and software specification on Phatsa Web Application[9]. This benchmarking is running on the server side, data visualization, UI rendered, and JavaScript execution in the Web Browser are excluded.

3. Experiment Methodology

The experiment on implementing WEMA in GDP prediction composed of several steps: collecting and formatting data, preparing moving average parameters (i.e. start index, span data, and prediction period), submitting original data to web application and finally comparing GDP overall value prediction with WEMA and other conventional methods (SMA, WMA, and EMA).

3.1. Data Collection and Formatting

Primary data on this study is taken from the official site of Bank Indonesia about Gross Domestic Product (GDP) reports that compiled into quarterly time-series[1]. The data starts from Q1 2010 to Q2 2018, mapped on exactly 34 time-series data of overall GDP. Figure 1 shows the format of tabular data that already converted from the original file into Phatsa Web Application compatible format.

Table 1: Indonesian GDP tabular data in Phatsa's Compatible Format

No.	Time-series Data	
	Period	Value
1	2010 Q1	1603771
2	2010 Q2	1704509
3	2010 Q3	1786196
4	2010 QIV	1769654
5	2011 QI	1834355
6	2011 QII	1928233
7	2011 QIII	2053745
8	2011 QIV	2015392
9	2012 QI	2061338
10	2012 Q2	2162036
11	2012 Q3	2223641
12	2012 Q4	2168687
13	2013 Q1	2235288
14	2013 Q2	2342589
15	2013 Q3	2491158
16	2013 Q4	2477097
17	2014 Q1	2506300
18	2014 Q2	2618947
19	2014 Q3	2746762
20	2014 Q4	2697695
21	2015 Q1	2728180
22	2015 Q2	2867948
23	2015 Q3	2990645

24	2015 Q4	2939558
25	2016 Q1	2929897
26	2016 Q2	3074804
27	2016 Q3	3206377
28	2016 Q4	3195694
29	2017 Q1	3228034
30	2017 Q2	3366585
31	2017 Q3	3503568
32	2017 Q4	3490608
33	2018 Q1	3506720
34	2018 Q2	3683896

This compatible data saved in a comma separated value (.csv) file format with the restriction of data in 'Value' cell must not contain more than one dot (.) as the dot represents floating number between 0 to 1. Value presented is in billions of Indonesian Rupiah (IDR).

Additionally, this study also tested the data of Czech Republic's GDP from Czech Statistical Office (CZSO)[16]. It is intended to provide a preview of the comparable results with other GDP data set.

3.2. Defining Experiment Parameters

There are two required parameters that needed to be set in Phatsa Web Application:

1. Span Data Parameter is the time window that represents how many previous data are needed to be calculated for the prediction.
2. Starting Index Parameter is the index of a certain period in data set that user wanted to start the forecasting calculation.

This rule is applied for these two parameters: Span Data ≤ Starting Index, considering any span data that could be more than starting index than the application can prompt an error since it will try to calculed minus the offset in the array (i.e. if span data is 10 while starting index is 1, then it will try to calculate index number -9 to 0 in the array of original data).

Based on th previous study on Phatsa Web Application[9], there is a parameter left that defined as a static number of 1 in this study: Period to Predict. This parameter represents how many prediction in the future needed to be calculated based on original data. Since it still needs to be reviewed about the significant effects of the parameter, in this study we only need 1 future data of GDP prediction in the next quarter.

Fig. 1: User Interface(UI) of Phatsa Web Application in Indonesia's GDP Prediction

3.3. Time-series Data Submission

After the data and parameters are set, then all of those items are submitted to Phatsa Web Application with steps written in a previous study[9]. Fig. 1 shows the results after WEMA method and error calculations worked with starting index and span data is equal to 15. The steps then iterate for span data and starting index (k) for given defined parameter in this study is Span Data = Starting Index in the number of 5, 10, 15, and 20 as well as each method SMA, WMA and EMA.

4. Prediction Results

All the prediction results tested on Indonesia's GDP Data on SMA, WMA, EMA and WEMA shown in Table 2, 3 ,4 and 5 respectively.

Table 2: GDP Data Prediction Evaluation Results with SMA

(k)	GDP Prediction with Simple Moving Average					
	MSE	MASE	MAPE (%)	CAT (ms)	CET (ms)	OCT (ms)
5	37052086404	2.79	6.83	0	0	0.1
10	114734210261	4.48	11.62	0	0	0.1
15	242962139457	6.19	16.20	0	0	0.1
20	417998665836	8.49	20.18	0.1	0,1	0.2

Table 3: GDP Data Prediction Evaluation Results with WMA

(k)	GDP Prediction with Weighted Moving Average					
	MSE	MASE	MAPE (%)	CAT (ms)	CET (ms)	OCT (ms)
5	24246991682	2.18	5.33	0	0	0.1
10	63327718505	3.28	8.50	0.1	0	0.1
15	126782667575	4.44	11.61	0.1	0	0.1
20	210908585193	6.00	14.25	0.1	0	0.1

Table 4: GDP Data Prediction Evaluation Results with EMA

(k)	GDP Prediction with Exponential Moving Average					
	MSE	MASE	MAPE (%)	CAT (ms)	CET (ms)	OCT (ms)
5	9071755418	1.22	3.00	1.5	0	1.6
10	9871876071	1.08	2.80	1.4	0	1.5
15	9471338298	0.99	2.55	1.3	0	1.3
20	10533664958	1.09	2.56	1.1	0	1.2

Table 5: GDP Data Prediction Evaluation Results with WEMA

(k)	GDP Prediction with Weighted Moving Average					
	MSE	MASE	MAPE	CAT	CET	OCT
5	9071755418	1.22	3.00	4.2	0	4.2
10	9366353807	1.08	2.80	5.5	0	5.6
15	9471338298	0.99	2.55	6.1	0	6.1
20	10533664958	1.09	2.56	5.9	0	5.9

This result shows that there are differences of error rates between SMA and WMA method on predicting GDP Data, in this case WMA has fewer error rate values than SMA. Between three conventional moving average methods, EMA shown the best result in predicting GDP data, and it has the same error rate values with WEMA method. However, WEMA method takes more calculating time than the EMA method. Moreover, if we take a look at differences between k = 15 and k = 20 in WEMA and EMA, we can find the error rates bounced back to a greater value while from k=5 to k=10, the error rates show decreasing calculation.

In accordance with this GDP prediction with WEMA method, we test the prediction on more data in Czech's GDP data set that has 94 data rows ranging quarterly from Q1 1995 to Q2 2018. It also has more volatile data rather than Indonesia's GDP data set Table 6 shows the prediction result with the same defined parameters. Fig.2 shows the result of Czech GDP data prediction on Phatsa Web Application when k = 15.

Fig. 2: User Interface(UI) of Phatsa Web Application in Czech's GDP Prediction

Table 6: Czech GDP Data Prediction Evaluation Results with WEMA

(k)	GDP Prediction with Weighted Moving Average					
	MSE	MASE	MAPE (%)	CAT (ms)	CET (ms)	OCT (ms)
5	2640875465	1.82	5.49	11.8	0.1	11.9
10	3650736539	1.69	5.91	18.3	0.1	18.4
15	3799697291	1.70	5.90	23.4	0.1	23.5
20	3947084672	1.70	5.89	27.9	0.1	28.0

As we compared the result in Table 5 and Table 6, we can see that WEMA method implemented in GDP prediction has relatively small percentage error based on MAPE error evaluation (ranging from 3% to 6 %). With the same defined parameters in the method and web application, it also shows that WEMA computation provide more error when GDP data quantity is expanded and more volatile data set.

5. Conclusion

WEMA implementaton on predicting GDP data set has shown its capability in this study. It has fewer error rates values compared to SMA and WMA, but has the same capability with EMA with more computation time. For further research, we took an assumption that it has a deal with the algorithm in finding the proper exponential degree (α) and also testing starting index and span data variation.

Additionally, we also find out that WEMA implementation in this matter also has an effect when it consumes expanded GDP data set with more volatile values.

Acknowledgement

This project was sponsored and fully funded by Indonesia Government, through the Research Grant under the scheme of 'Applied Product Research' (Penelitian Produk Terapan) from Ministry of Research, Technology and Higher Education of the Republic of Indonesia (Contract No. : 425/LPPM-UMN/IV/2017). This project then continued under the scheme of "National Strategic Research" (Penelitian Strategis Nasional-Institusi) from Ministry of Research, Technology and Higher Education of the Republic of Indonesia (Contract No. : 034/KM/PNT/2018).

References

[1] Bank Indonesia, "Economic and Financial Data for Indonesia," *Economic and Financial Data for Indonesia*, 21-Sep-2018. [Online]. Available: https://www.bi.go.id/sdds/. [Accessed: 23-Sep-2018].

[2] A. Barnett, H. Mumtaz, and K. Theodoridis, "Forecasting UK GDP growth and inflation under structural change. A comparison of models with time-varying parameters," *International Journal of Forecasting*, vol. 30, no. 1, pp. 129–143, 2014.

[3] H. C. Bjørnland, F. Ravazzolo, and L. A. Thorsrud, "Forecasting GDP with global components: This time is different," *International Journal of Forecasting*, vol. 33, no. 1, pp. 153–173, 2017.

[4] T. E. Clark and F. Ravazzolo, "Macroeconomic forecasting performance under alternative specifications of time-varying volatility," *Journal of Applied Econometrics*, vol. 30, no. 4, pp. 551–575, 2015.

[5] S. Hansun, "A Novel Research of New Moving Average Method in Time-series Analysis," *International Journal of New Media Technology*, vol. 1, no. 1, pp. 22–26, 2014.

[6] A. Ranjan and P. Jetley, “Evaluation of signal smoothing algorithms for stability of a quadrotor MAV,” in *Autonomous Robot Systems and Competitions (ICARSC), 2014 IEEE International Conference on*, 2014, pp. 309–314.

[7] S. Hansun, “WEMA versus B-WEMA Methods in Forex Forecasting,” in *Proceedings of the 9th International Conference on Machine Learning and Computing*, 2017, pp. 268–271.

[8] S. Hansun, “FX forecasting using B-WEMA: Variant of Brown’s Double Exponential Smoothing,” in *Informatics and Computing (ICIC), International Conference on*, 2016, pp. 262–266.

[9] M. B. Kristanda and S. Hansun, “Phatsa: A web-based application for forecasting using conventional moving average methods,” in *New Media Studies (CONMEDIA), 2017 4th International Conference on*, 2017, pp. 38–43.

[10] S. Hansun and M. B. Kristanda, “Performance analysis of conventional moving average methods in forex forecasting,” in *Smart Cities, Automation & Intelligent Computing Systems (ICON-SONICS), 2017 International Conference on*, 2017, pp. 11–17.

[11] Y. Zhuang, L. Chen, X. S. Wang, and J. Lian, “A weighted moving average-based approach for cleaning sensor data,” in *Distributed Computing Systems, 2007. ICDCS’07. 27th International Conference on*, 2007, pp. 38–38.

[12] I. Noda, “Adaptation of stepsize parameter to minimize exponential moving average of square error by newton’s method,” in *9th International Conference on Autonomous Agents and Multiagent Systems, pp. M-2–1*, 2010.

[13] Z. Wang and A. C. Bovik, “Mean squared error: Love it or leave it? A new look at signal fidelity measures,” *IEEE signal processing magazine*, vol. 26, no. 1, pp. 98–117, 2009.

[14] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” *International journal of forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[15] L. Ezell, *Practical CodeIgniter 3*. Leanpub, 2015.

[16] The Czech Statistical Office (CZSO), “Gross domestic product - time-series | CZSO,” *Gross domestic product - time-series | CZSO*, 31-Aug-2018. [Online]. Available: https://www.czso.cz/csu/czso/hdp_ts. [Accessed: 30-Sep-2018].