5th International Electrical Engineering Congress, Pattaya, Thailand, 8-10 March 2017

The Variable Forgetting Factor-based Local Average Model Algorithm for Prediction of Financial Time Series

Prakit Intachai
The Electrical Engineering Graduate Program,
Faculty of Engineering,
Mahanakorn University of Technology,
140 Cheumsamphan Rd., Nong-chok, Bangkok 10530,
Thailand
prakit2554@gmail.com

Peerapol Yuvapoositanon
Centre of Embedded Systems and Digital Signal Processing,
Faculty of Engineering,
Mahanakorn University of Technology,
140 Cheumsamphan Rd., Nong-chok, Bangkok 10530,
Thailand
peerapol.y@gmail.com

***Abstract*—In this paper, we propose a variable forgetting factor-based local average model for estimation of future values of financial time series. The forgetting factor is applied to the existing local average model to govern the weights of past records for the estimation of the future records. By using the trend direction from the turning points of the financial time series, the value of the forgetting factor can be estimated. The results of performance comparison between the proposed variable forgetting factor-based local average model and the original local average model on the actual time series derived from the stocks listed in the Stock Exchange of Thailand are shown. The results suggest that the proposed method offers consistent less prediction errors than the existing method.**

Keywords—forgetting factor, local average model, turning point, prediction method, financial time series.

I. INTRODUCTION

Searching for accurate prediction techniques of financial time series has long been one of the main activities of technical analysts as well as academia. Researchers have long analyzed the financial time series by means of some representatives on the movement of stocks in the financial market [1-5]. Specifically in [5], the method of turning points is used for the representative of the time series. This method relies on the characteristics of duration of time. The size (magnitude) of real data records at the highest and lowest positions represents the time series. In [2, 4], the calculation of turning point from the searching minimum and maximum point is proposed and the researches [4] tested with the time series. In [1], the turning points are searched by the scoping of the time interval and the price percentage in the financial time series.

In [5,6], the prediction method by the Local Average Model (LAM) with the turning points (TPs) is proposed. The averaged price and time on the financial time series are used for prediction of price in the future. Turning points are used in trending the financial time series mainly because of the non-stationary nature of the data records. In the original LAM, the prediction of future price is done using by averaging all the turning points.

However, there is a hypothesis that recent records may be more representative to the future price prediction than the historical ones. If that is the case, the recent data records can have more impact in the prediction process and should have more weights when averaging in LAM than the past ones. Traditionally, this task can be carried out by assigning larger weights to the more recent data with the forgetting factor λ as in the recursive least squares (RLS) algorithm [8-10]. However, since the financial time series is invariably non-stationary in nature, the forgetting factor λ cannot be fixed but must be variable with respect to the ever-changing trends of the data. The challenging problem is how to determine λ for each time series with different trend characteristics to obtain the possible minimum prediction errors.

In this paper, we propose a novel LAM by introducing the variable forgetting factor λ into the existing LAM. We also offer an approach to adapt the value of the forgetting factor λ by means of evaluating the changing trends of the turning points. Since there is two kinds of trends, i.e., rising and falling, we explain how to adjust the forgetting factor according to the trend direction. Collectively, the new LAM method of determining trend directions and estimating of adaptable λ is therefore called the variable forgetting factor-based local average model (VFF-LAM) algorithm for the prediction of future values of financial time series.

The paper is organized as follows. In Section II, we present the definition of turning points. The VFF-LAM and the method of how to adapt λ according to the data records are presented in Section III. In Section IV, the empirical results from the estimation and adaptation of the forgetting factor to the actual financial time series are shown. The comparison between the estimation with the adaptation forgetting factor by the errors of the prediction is performed. Finally, conclusions are provided in Section V.

II. THE DEFINITION OF TURNING POINTS

As a data set is recorded, the addition of noise or error signal into the time series data is inevitable. Upon removal of the noise, however, the trend of the data can be revealed and

978-1-5090-4666-9/17/$31.00 ©2017 IEEE

we can search the points of the changing direction of data in the time series. These points are defined as the turning points (TPs). The definition of TP involves the relying magnitudes (minimum or maximum of the data set) and times for the calculation of the representative data.

Fig. 1 shows an illustration of turning points whereas a 16 record of time series data is assumed. For example, referring to Fig. 1, TP_1 is the turning point that occurs between x_1 and x_5. The time interval z_1 of TP_1 can be calculated from $x_1 - x_5$ and it has four records of data. The time interval between z_1 to z_4 can help the calculation at the time window of the prediction data and this time is defined as T. The turning points play a crucial role in the prediction of future values by VFF-LAM.

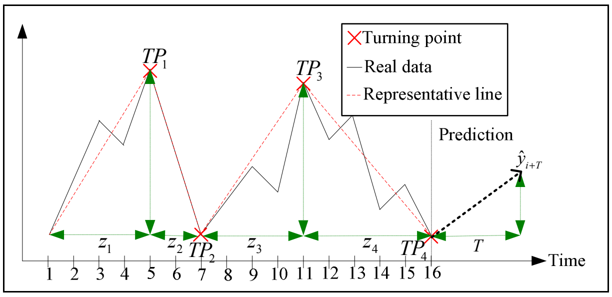

Fig. 1. An illustration of Turning Points (TPs) and the representative trend lines.

III. The Variable Forgetting Factor-based LAM

We now introduce the forgetting factor λ to the original LAM to govern the weights of past records for the estimation of the future records y_{i+T} as

$$\hat{y}_{i+T} = \overline{y}_{i(K)+T} = \frac{1}{K}\left(y_{i(1)} + \sum_{j=2}^{K-1} (\beta+\lambda)^{j-1} y_{i(j)} \right), \qquad (1)$$

where λ can be less than or equal to one. In Algorithm 1, the estimation of λ from the financial time series by the relative start and end points from the TP is described. The value of λ depends on how the trend slopes. When the trend slopes steeply, it means that past records are significant in predicting the future records. In such case, the forgetting factor should have a large value. When the trend slopes slightly or has no slope, past records can be negligible. The forgetting factor in such case can be of small value. Since λ can be adapted depending on the time series, this algorithm is then called the variable forgetting factor-based LAM (VFF-LAM) algorithm. The definition of θ and the designated points P_1, P_2 and P_3 used in Algorithm 1 is explained for a time series example in Fig.2. Also shown in Fig.2 is an example of how to depict the trend.

The parameter β in (1) functions as the LAM indicator for the trend direction. We define two kinds of trend directions: rising and falling trends. The rising trend has the start point lower than the end point and the falling trend has the start point higher than the end point. For the rising trend, $\beta = 1$ and the falling trend is $\beta = -1$. This valuation of β is therefore used in (1) and detailed in Algorithm 1. Fig.2 is therefore an example of a falling trend because the start point P_1 is higher than the end point P_3.

By comparing the angle between θ with 90, the ratio of the trend from the base angle at 0 degree to θ in the time series can be calculated. The estimate of λ of the data records is calculated by

$$\lambda_{EST} = \alpha_{EST}\left(\theta/90\right), \qquad (2)$$

where the scaling parameter α_{EST} is to be defined empirically. The details of how α_{EST} is defined is described in Section IV.

Input:	TP
Define:	PS // The start point
	PE // The end point
	β // The variable of trend in the time series
	$\Delta\theta$ // The ratio of the switching trend
1:	Normalize the price and time
2:	Draw the line between the start and end point
3:	If ($PS < PE$)
4:	$\beta = 1$ and Draw the base line between the start and end point at minimum point of PS
5:	else
6.	if ($PS > PE$)
7:	$\beta = -1$ and Draw the base line between the start and end point at minimum point of PE
8:	end if
9:	Calculate θ in degree format
10:	Calculate $\Delta\theta = \theta/90$
11:	Estimate $\lambda_{EST} = \alpha_{EST}\Delta\theta$
12:	Calculate $\hat{y}_{i+T} = \frac{1}{K}\left(y_{i(1)} + \sum_{j=2}^{K-1} (\beta+\lambda)^{j-1} y_{i(j)} \right)$

Algorithm 1: Pseudocode of the VFF-LAM algorithm.

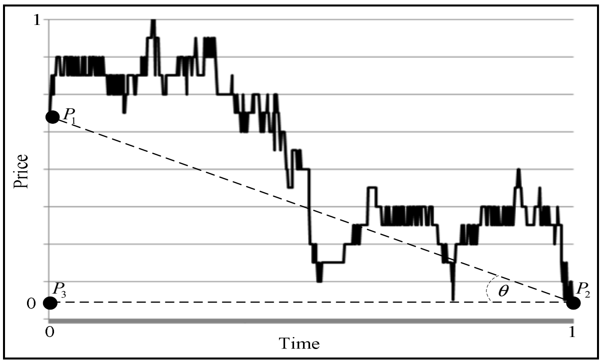

Fig. 2. The definition of the angle θ and the designated points P_1, P_2 and P_3 and the trend between the start and end points.

It should be noted that the only difference between VFF-LAM and LAM is the inclusion of the forgetting factor λ to the past records in VFF-LAM whereas LAM is simply the average or mean of the time series. When $\beta+\lambda=1$, the VFF-LAM is reverted to the original LAM.

IV. THE EMPIRICAL RESULTS OF THE VFF-LAM ALGORITHM

In this Section, we tested the performance of the VFF-LAM algorithm as compared to that of the original LAM with the actual financial data. We used the error as the performance metric, where it can be calculated from the ratio of the difference of the actual price (y_{i+T}) and the estimated price ($\hat{y}_{i+T}$) to y_{i+T}

$$\text{Error} = \left(y_{i+T} - \hat{y}_{i+T} \right) / \, y_{i+T} . \tag{3}$$

The error (3) needs λ_{EST} of (2) for calculating $\hat{y}_{i+T}$. In order to calculate λ_{EST}, we must define α_{EST} empirically. In Table I, we tested for α_{EST} by 10 stocks extracted from the stock exchange market of Thailand during 20 to 28 November 2016. We estimated the minimum and maximum α, i.e., $\alpha_{\min}$ and $\alpha_{\max}$ respectively, at the error levels of 5%, 10% and 20%. The average of $\alpha_{\min}$ and $\alpha_{\max}$ for each stack at designated error is defined as α_{av}. In Table I, the results of α_{av} at 5%, 10% and 20% error levels are 0.09818, 0.09818 and 0.0997 respectively. Therefore, α_{av} for of all error levels is 0.0986. So it is safe to assume that α_{EST} can be set around 0.1. Hence, we used $\alpha_{EST}=0.1$ for all the calculation of λ_{EST} in (2). For the estimation for other stocks, α_{EST} can be recalculated to meet the performance requirement at the defined error levels. We also tested to find the optimal value of λ at the minimum error attained for of each stock, i.e., λ_{OPT}. The optimum λ_{OPT} of each stock is used to compare with that of λ_{EST} in Table II. Note that the inclusion of λ_{OPT} is for the comparison purpose only since the λ_{OPT} must is associated with the minimum error which is difficult to obtain in real time.

In Table II, we compare the performance of LAM and VFF-LAM at the error level of 10% for the same previous 10 stocks. For VFF-LAM, the errors derived by using λ_{EST} and λ_{OPT} are compared. From Table II, for all of the tested stocks, it is shown that the errors of estimation by VFF-LAM with λ_{EST} are significantly lower than that of LAM. Obviously, the best performance of VFF-LAM is obtained by using λ_{OPT}. Four of the 10 stocks, i.e., KBANK, BEC, ESSO and CPF, are selected for plotting as they are differed in their trends as shown in Fig.3. In Fig.3 (a), the trend of the KBANK stock is shown where the start point at P_1, the end point P_2 and the baseline point P_3 are depicted. The line direction of $\overline{P_1P_2}$ is the rising trend because the start point is less than the end point. The lines $\overline{P_2P_1}$ and $\overline{P_1P_3}$ form the angle θ of 22 degrees. From (2), with $\theta=22$ and $\alpha_{EST}=0.1$, λ_{EST} is then 0.024. The error of VFF-LAM in this case is 4.2%. When calculated VFF-LAM with λ_{OPT} which, by inspection from Fig.3 (c), is 0.021. The error in this best-scenario case is 1.15%. For LAM, the error is 32.20% which is significantly higher than that obtained by VFF-LAM with either λ_{EST} or λ_{OPT}.

TABLE I. THE RESULT OF THE ESTIMATE FORGETTING FACTOR

Stock	Error <= 5%		Error <= 10%		Error <= 20%	
	α_{min}	α_{max}	α_{min}	α_{max}	α_{min}	α_{max}
KBANK	0.0988	0.1012	0.098	0.102	0.095	0.108
BEC	0.0938	0.0968	0.091	0.1	0.089	0.112
ESSO	0.093	0.096	0.091	0.099	0.084	0.108
CPF	0.095	0.1035	0.09	0.112	0.084	0.122
CBG	0.098	0.103	0.095	0.105	0.092	0.112
BTS	0.0975	0.0995	0.092	0.102	0.084	0.11
MAJOR	0.094	0.096	0.09	0.098	0.082	0.108
BAY	0.098	0.1	0.092	0.102	0.088	0.11
PTT	0.0975	0.099	0.095	0.105	0.092	0.112
SCB	0.099	0.104	0.096	0.106	0.088	0.114
α_{av}	0.09818		0.09805		0.0997	

TABLE II. THE SUMMARY OF THE PREDICTION ERRORS OF VFF-LAM AND LAM

Stock	Error from LAM	Error from VFF-LAM (λ_{EST})	Error from VFF-LAM (λ_{OPT})	λ_{EST}	λ_{OPT}
KBANK	32.20%	4.20%	1.15%	0.024	0.021
BEC	67.85%	7.40%	1.85%	0.044	0.05
ESSO	62.64%	9.54%	1.24%	0.02	0.023
CPF	53.35%	9.40%	1.35%	0.034	0.03
CBG	23.25%	7.64%	1.02%	0.034	0.03
BTS	56.75%	5.25%	1.45%	0.04	0.036
MAJOR	35.46%	9.25%	0.80%	0.024	0.04
BAY	49.96%	5.80%	1.22%	0.044	0.034
PTT	38.30%	6.56%	1.56%	0.044	0.038
SCB	23.25%	6.26%	1.76%	0.024	0.03

From Fig.3, it is noted that all the error plots of VFF-LAM for the other stocks also have the minimum points at particular values of α and λ. This pattern suggests that VFF-LAM has the capability of reducing the prediction errors for financial time series to the optimum points of α and λ.

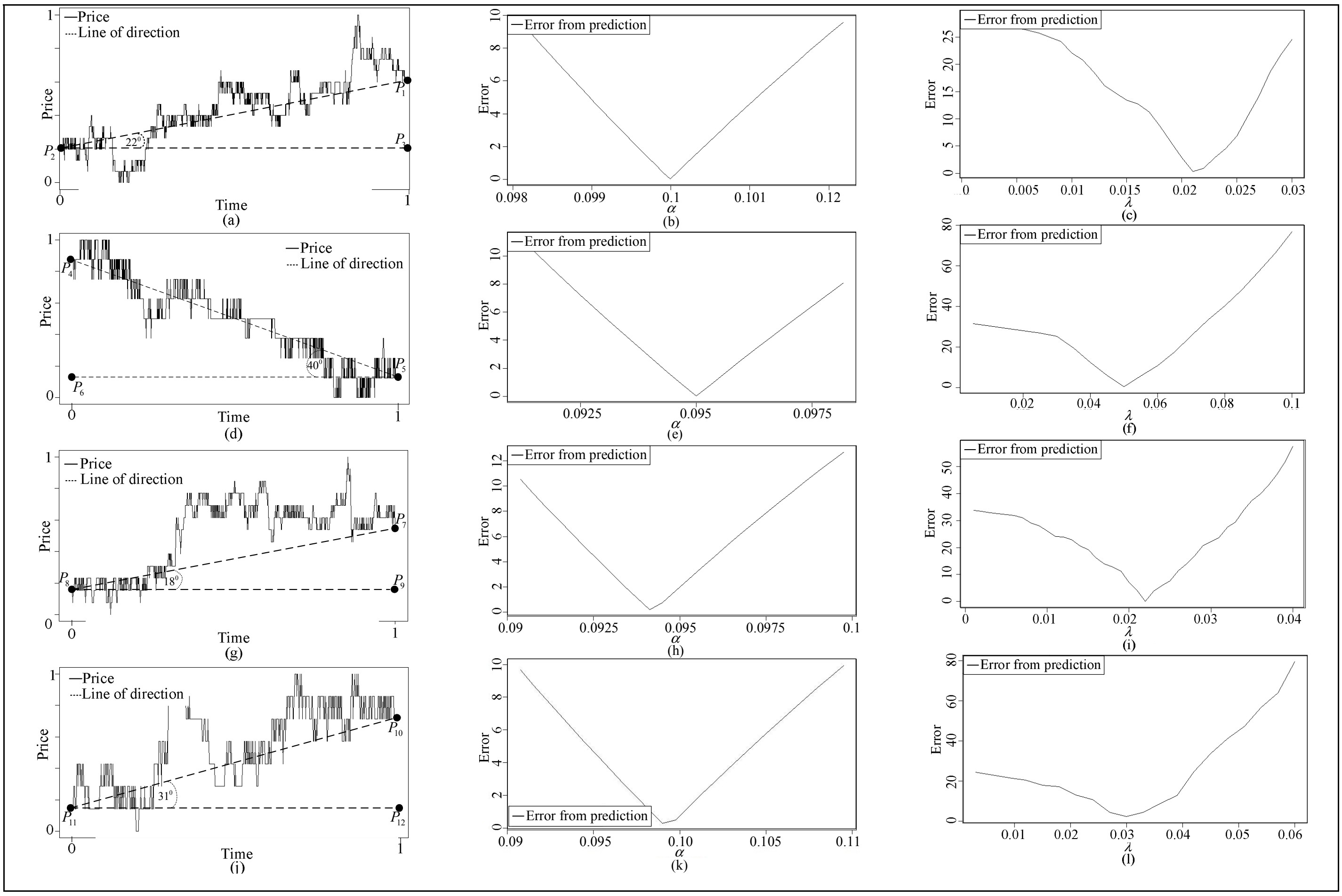

Fig. 3. The data records and trends and the angles θ , Errors in percent w.r.t α and Errors in percent w.r.t λ of KBANK in (a),(b),(c), of BEC in (d),(e),(f), of ESSO in (g),(h),(i) and in CPF in (j),(k),(l) respectively.

V. CONCLUSIONS

In this paper, we propose a novel VFF-LAM model by introducing the variable forgetting factor λ into the existing LAM. An approach to adapt the value of the forgetting factor λ by means of evaluating the two kinds of turning point trends, i.e., rising and falling, is provided. The performance of the proposed VFF-LAM in prediction of future values of financial time series was tested as compared to the existing LAM. The results show that the prediction errors obtained by VFF-LAM are significantly lower than those of LAM. Also the pattern of prediction errors plotted against the values of α and λ suggests that VFF-LAM has the capability of reducing the prediction errors for financial time series to the optimum points of α and λ. Future research can be directed to determine the optimization techniques of the forgetting factor λ and it related parameter α.

REFERENCES

[1] D. Bao, and Z. Yang, "Intelligent stock trading system by turning point confirming and probabilistic reasoning," Expert Systems with Applications, vol. 34, no. 1, 2008, pp.620-627.

[2] J. Yin, Y.-W. Si, and Z. Gong," Financial time series segmentation based on Turning Points," Proceedings International Conference on System Science and Engineering, IEEE, 2011, pp.394-399.

[3] T.-c. Fu, F.-l. Chung, R. Luk, and C.-m. Ng, "Representing financial time series based on data point importance," Eng. Appl. Artif. Intell., vol. 21, no. 2, 2008, pp.277-300.

[4] V. Vo, J. Luo and B. Vo, "A turning points method for stream time series prediction," International Journal of Innovative Computing, International and Control, ICIC, Volume 9, Number 10, 2013, pp.3965-3980.

[5] D. Kugiumtzis and I. Vlachos, "Turning Point Prediction of Oscillating Time Series using Local Dynamic Regression Models," Proceedings of NOLTA Conference : arXiv:0809.2229, 2008.

[6] D. Kugiumtzis, "Local prediction of turning points of oscillating time series," Physical Review E, 78(3), 036206, 2008.

[7] K. Nishiyama, "A new approach to introducing a forgetting factor into the normalized least mean square algorithm," Journal of Signal Processing, ELSEVIER, Vol.114, 2015, pp.19-23.

[8] C. Paleologu, J. Benesty and S. Ciochina, "A Robust Variable Forgetting Factor Recursive Least-Squares Algorithm for System Identification," Signal Processing Letters, Vol. 15, IEEE, 2008, pp.597-600.

[9] K. Nishikawa and J. Hasegawa, "Variable Forgetting Factor Algorithm Based on Sparsity for RLS-type Adaptive Algorithms," Proceedings of the Second APSIPA Annual Summit and Conference, 2010, pp.173-176.

[10] S.M. Jung and P. Park, "Variable Forgetting Factor Recursive Total Least Squares Algorithm for FIR Adaptive filtering," International Conference on Electronics Engineering and Informatics, ICEEI, Vol.49, 2012, pp.170-174.