Hybridization of Ensemble Kalman Filter and Non-linear Auto-regressive Neural Network for Financial Forecasting

Said Jadid Abdulkadir[1], Suet-Peng Yong[1], Maran Marimuthu[2], and Fong-Woon Lai[2]

[1] Department of Computer and Information Sciences
[2] Department of Management and Humanities,
Universiti Teknologi Petronas,
31750, Tronoh, Malaysia
said_g02156@utp.edu.my,
{yongsuetpeng,maran.marimuthu,laifongwoon}@petronas.com.my

Abstract. Financial data is characterized as non-linear, chaotic in nature and volatile thus making the process of forecasting cumbersome. Therefore, a successful forecasting model must be able to capture long-term dependencies from the past chaotic data. In this study, a novel hybrid model, called UKF-NARX, consists of unscented kalman filter and non-linear auto-regressive network with exogenous input trained with bayesian regulation algorithm is modelled for chaotic financial forecasting. The proposed hybrid model is compared with commonly used Elman-NARX and static forecasting model employed by financial analysts. Experimental results on Bursa Malaysia KLCI data show that the proposed hybrid model outperforms the other two commonly used models.

Keywords: chaotic time-series, ensemble model, non-linear autoregressive network, financial forecasting.

1 Introduction

Forecasting is a dynamic process and perplexing task in the financial division. It helps financial market analysts to evade stock trading losses and obtain huge profits by coming up with promising business policies. Hence, financial companies can make precise forecasts by planning some required interventions to meet their business performance targets [14]. Furthermore, stock trading companies are usually scrutinized by short and long term investors while concerning the expectations from shareholders. Stockholders may also like to analyze their investments by comparing the analysis of forecasting companies.

An example of financial time-series forecasting is the stock prices in the share market which are characterized by non-linearity, noisy, chaotic in nature and volatile thus making the process of forecasting cumbersome. The goal of financial

R. Prasath et al. (Eds.): MIKE 2014, LNAI 8891, pp. 72–81, 2014.

forecasters is to innovate numerous techniques that can forecast effectively by following legal trade strategies and avoiding losses. The general idea of successful stock market prediction to achieve best results is by using minimum required input data and the least complex stock market model [4]. The intricate nature of stock market forecasting has led to the need for further improvements in the use of intelligent forecasting techniques to drastically decrease the dangers of inaccurate decision making.

Financial controllers who adhere to the ideas of an efficient market hypothesis and random walk theories disbelieve that stock market can be predicted [15]. Nevertheless, fanatics of technical and fundamental analysis have shown numerous ways to counter the claim by adherents of random walk theory and efficient market hypothesis. Therefore, numerous approaches for tackling the chaotic nature of forecasting have been suggested. However, new improvements in the area of soft computing through the use of computational intelligence have offered new ideas in forecasting chaotic data in stock market and also modelling its non-linearity.

In this paper, deriving from computational intelligence method, a hybrid neural network model consisting of unscented kalman filter and parallel non-linear autoregressive neural network is developed to enhance the performance of financial forecasting based on Kuala Lumpur Composite Index (KLCI) data. The remainder of the paper is organized as follows: section 2 outline some related work in financial forecasting and its development; in section 3, the proposed hybrid neural network model is discussed; while section 4 addresses the experimental setup and performance analysis followed by the conclusion in the last section.

2 Related Works

Financial analysts have employed the use of static forecasting models which are a sequence of one-step ahead forecasts made at different points in time. Static forecasting uses actual rather than forecast value for the lagged variable and which can be done only if there are actual data available [5].

Computational Intelligence forecasting techniques such as fuzzy logic, genetic algorithms (GA) and artificial neural networks (ANN) are the most famous used techniques to cope with problems that have not been solved by complex mathematical systems. ANN applications have been widely used for forecasting in a variety of areas [9], [11], [13], [16], [21]. ANN was used for the solution of numerous financial problems [9]. It is also used in forecasting of financial markets, particularly forecasting of stock market indexes which are considered to be a barometer of the markets in many countries [21,19]. However, the problem of over-fitting [7] arises when a model describes noise instead of the underlying relationship, hence affecting the accuracy of forecasting. The forecast of Kuala Lumpur Composite Index (KLCI) has been investigated using ANN [21], fuzzy logic [2] and artificial neural fuzzy inference system (ANFIS) [23]. However, ANFIS has strong computational complexity restrictions and translates prior knowledge into network topology hence being sensitive to the number of input variables.

The advances of ANNs over the last few years is its ability of easily allowing more than one model to be combined with itself with multiple training, which is also referred to as hybrid modelling [7]. This technique has huge advantages because each part of the model performs and captures patterns within the data applied hence increasing the forecasting ability of each model inside the when being hybridized. A number of studies have employed the use of hybrid modelling in financial forecasting, for instance, a hybrid model consisting of neural networks and support vector machines (SVM) [10], a radial basis function neural network model hybridized with SVM [18], and another hybrid forecasting model developed from the integration of generalized linear auto-regression (GLAR) and neural networks (ANN) [22]. Apart from that, a hybrid neural network and fuzzy regression model has also been used in foreign exchange rate forecasting [8].

Elman and NARX network [3] have been hybridized for chaotic forecasting too, the hybrid model minimized the problem of vanishing gradients in recurrent networks, but did not consider the over fitting problem. In this paper, a novel hybrid model is proposed for financial forecasting by addressing the main issues of vanishing gradient [12] and over fitting [7] in recurrent neural networks. Its performance is compared to the recently developed Elman-NARX model.

3 Proposed Model

A novel hybrid model consisting of unscented Kalman filter and non-linear auto-regressive network with exogenous input is proposed to enhance multi-step-ahead forecasting of chaotic financial data. The hybrid model addresses the problem of vanishing gradient [12] experienced in network training by employing the use of bayesian regulation in training of the hybrid model and also the problem of over-fitting [7] by filtering the chaotic data before forecasting. The function of UKF is to create a better forecasting model by filtering the chaotic data because tiny errors in noise form [1] will be amplified hence affecting the forecasting performance of non-linear auto-regressive input network.

3.1 Unscented Kalman Filter (UKF)

Unscented Kalman filter addresses the approximation issues of extended Kalman filter [17]. In UKF, a minimal number of sigma points are selected that captures the mean and covariance of the state distribution which are obtained using a Gaussian random variable.

The random variable undergoes the process of non-linear unscented transformation. Assuming X has mean $\bar{X}$ and covariance P_k, each sigma point is propagated through the non-linear process model:

$$X_k^{f,j} = f(X_{k-1}^j) \tag{1}$$

The transformed points are used to compute the mean and covariance of the forecast value of X_k:

$$X_k^f = \sum_{j=0}^{2n} w^j X_k^{f,j} \tag{2}$$

$$P_k^f = \sum_{j=0}^{2n} w^j (X_k^{f,j} - X_k^f)(X_k^{f,j} - X_k^f)^T + Q_{k-1} \tag{3}$$

We propagate then the sigma points through the non-linear observation model:

$$Z_{k-1}^{f,j} = h(X_{k-1}^j) \tag{4}$$

With the resulted transformed observations, their mean and covariance (innovation covariance) is computed:

$$Z_{k-1}^f = \sum_{j=0}^{2n} w^j Z_{k-1}^{f,j} \tag{5}$$

$$Cov(\tilde{Z}_{k-1}^f) = \sum_{j=0}^{2n} w^j (Z_k^{f,j} - Z_k^f)(Z_{k-1}^{f,j} - Z_{k-1}^f)^T + R_k \tag{6}$$

The cross covariance between $\tilde{X}_k^f$ and $\tilde{Z}_{k-1}^f$ is:

$$Cov(\tilde{X}_{k-1}^f, \tilde{Z}_{k-1}^f) = \sum_{j=0}^{2n} w^j (X_k^{f,j} - Z_k^f)(Z_{k-1}^{f,j} - Z_{k-1}^f)^T + R_k \tag{7}$$

The information obtained form the time update step is combined with the measurement step Z_k. The gain K_k is given by:

$$K_k = Cov(\tilde{X}_k^f, \tilde{Z}_{k-1}^f) Cov^{-1}(\tilde{Z}_{k-1}^f) \tag{8}$$

The posterior covariance is updated from the following formula:

$$P_k = P_k^f - K_k Cov(\tilde{Z}_{k-1}^f) K_k^T \tag{9}$$

In our work for this paper, Q_k and R_k are the process and measurement noise which are set up as 1 and 0.001 respectively.

3.2 Recurrent Network

Non-linear auto-regressive network with exogenous input (NARX) can be easily applied for prediction of time series data with the embedded input reconstruction of the network [20]. Hence, the filtered dataset is then fed into the non-linear auto-regressive with exogenous input model in parallel mode as shown in Figure 1, which is mathematically expressed as:

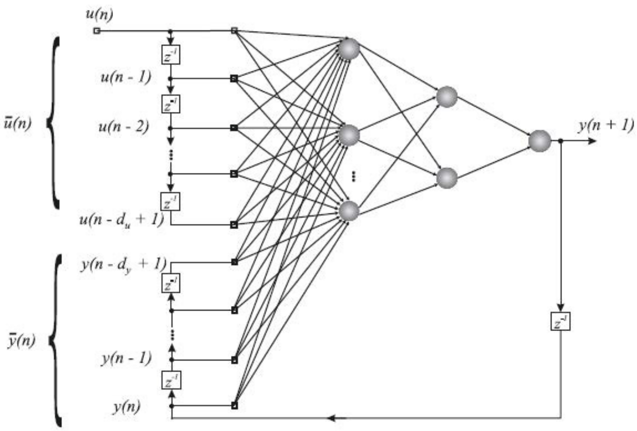


Fig. 1. Parallel-NARX recurrent network architecture (z^{-1} = unit time delay)

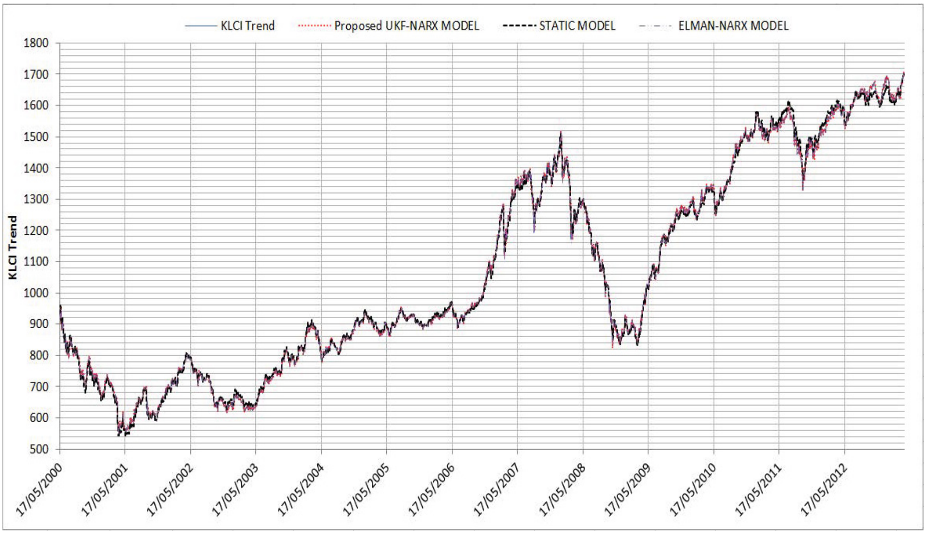


Fig. 2. Graph showing 12 years forecasted KLCI Trend output

$$y(n+1) = f[y(n), ..., y(n-d_y+1); u(n-k), u(n-k+1), ..., u(n-d_u-k+1)] \quad (10)$$

where $u(n)\epsilon\mathbb{R}$ and $y(n)\epsilon\mathbb{R}$ denote, respectively, the input and output of the model at discrete time step n, while $d_u \geq 1$ and $d_y \geq 1$, $d_u \leq d_y$, are the input-memory and output-memory orders, respectively. The parameter $k(\geq 0)$ is a delay term assumed to be zero hence referred to as the process dead-time.

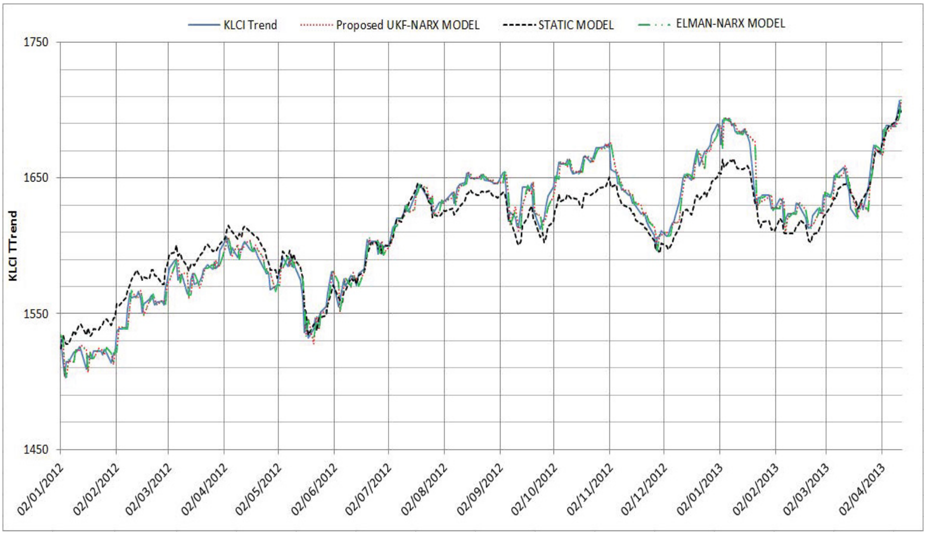


Fig. 3. Graph showing 1.5 years forecasted sample from the total forecasted period of 12 years

When gradient algorithms are used for training, the value decreases to zero as the number of time steps increases. This problematic outcome is commonly referred to as the vanishing gradient problem that results in reduced network performance on standard neural network models [12]. In the proposed model, Bayesian regulation is used as a training algorithm to adjust the parameters of the network so as to move the equilibrium in a way that will result in an output that is close as possible to the target output [6].

4 Results and Discussions

In this paper, Kuala Lumpur Composite Index (KLCI) traded in FTSE Bursa Malaysia that serves as an example of daily financial indexes, is used as real life financial time series dataset for our case study. The transaction date was from 12 April, 1988 to 12 April, 2013 with a total of 6524 daily samples over a period of 25 years. 3156 daily data points were used as training data and the remaining 3368 points are used for testing of the proposed hybrid model which translates to around 12 years of forecasting.

Determining chaos in time series analysis is a very crucial step to differentiate between chaotic and random time series. Largest Lyapunov Exponent (LLE)[3] is used to determine if KLCI data is chaotic or not, a negative LLE means the time series is not chaotic and a positive LLE shows the existence of chaos in the tested time series. The following equation is used to obtain the lyapunov exponent:

$$\lambda = \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} ln \mid \frac{dR}{dx}(x(t)) \mid \tag{11}$$

where the range of N is the size of the dataset which is 6524 samples in this case, R is the initial starting point that is used for differentiation in the lyapunov exponent formula.

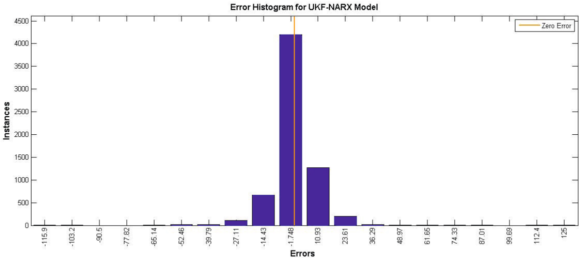
UKF with $Q_k = 1$ and $R_k = 0.001$ is used for filtering the KLCI data which had a +3.8 LLE value, hence showing chaotic characteristics. The filtered outputs are fed into the NARX network trained using Bayesian regulation algorithm. The NARX network for both experiments in parallel mode was set up with 10 neurons in the hidden layer, input delay $d_y = 3$ and feedback delay $d_n = 4$. Figure 2 shows the total forecasted period of 12 years using three different models and an in-depth graphical representation for a period of one and half years is shown in Figure 3.

Three commonly used performance metrics are employed to evaluate the forecasting accuracy in different aspects. Those metrics are Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The obtained forecasting errors from the three models are shown in Table 1. The results show that the proposed UKF-NARX model outperforms other commonly used forecasting models by having the least value of error i.e. 0.5641, 5.7331, 8.3878 for MAPE, MAE and RMSE respectively. Figure 4 shows the histograms of the forecasting errors from each model respectively.

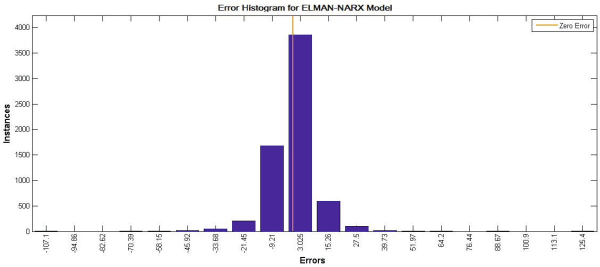
Table 1. Model comparison for forecasted KLCI trend

FORECASTING MODEL	**MAPE**	**MAE**	**RMSE**
Static Financial Model	0.8769	9.0408	11.4886
ELMAN-NARX Model	0.7654	6.3147	9.9556
UKF-NARX Model	**0.5641**	**5.7331**	**8.3878**

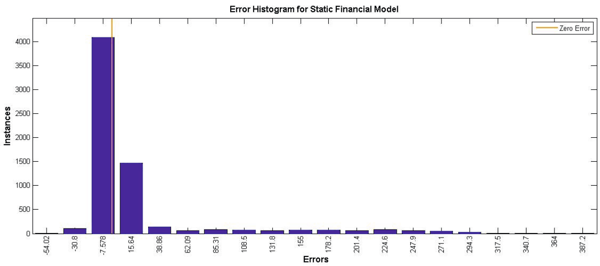
The regression (R) value is obtained by measuring the correlation between the forecasted and target output in the testing phase, the importance of this value is to check the model success in forecasting the dependant variable within the KLCI sample. The proposed UKF-NARX model had a regression value of 0.9985 depicted from Figure 5a; Elman-NARX model had a value of 0.9911 as shown in Figures 5b while the static model had a regression value of 0.982 which is shown in Figure 5c. All the three models had an accepted regression value which is closer to the value of 1 translating to a close model relationship with almost a perfect fit.



(a) The proposed UKF-NARX model

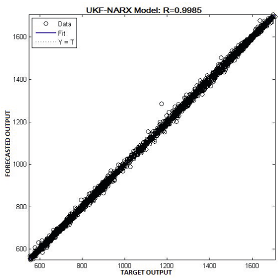


(b) Elman-NARX model

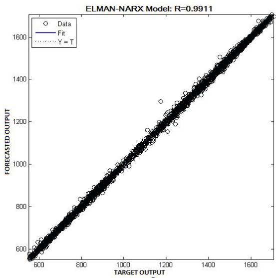


(c) Static forecasting model

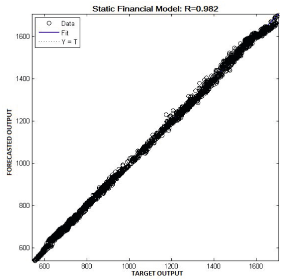
Fig. 4. Error histograms for the three comparing models



(a) UKF-NARX model



(b) Elman-NARX model



(c) Static forecasting model

Fig. 5. Regression analysis for the comparing models

5 Conclusion

In this study, a novel hybrid model which consists of Unscented Kalman Filter and Non-linear Auto-regressive Neural Networkis (UKF-NARX) was proposed for multi-step-ahead chaotic forecasting of the KLCI trend. The experimental results showed that the UKF-NARX hybrid model outperformed other commonly used models in terms of accuracy and regression value. It should be noted that Elman-NARX and static forecasting models can also be used for forecasting because the error and regression values obtained are within the accepted range. However, in terms of model enhancement, the proposed UKF-NARX model is better for financial forecasting for a period of 12 years, as shown is our investigations.

Future research may further explore the selection parameter settings for input and feedback delays in the proposed model and the forecasting effect of increasing the period to a longer period if possible. Furthermore, the daily KLCI time series trend has not been applied for multi-step-ahead forecasting with a forecasting horizon of over 10 years, hence no model comparison was reported in the literature based on the forecasting horizon.

References

1. Jadid Abdulkadir, S., Yong, S.-P.: Unscented kalman filter for noisy multivariate financial time-series data. In: Ramanna, S., Lingras, P., Sombattheera, C., Krishna, A. (eds.) MIWAI 2013. LNCS, vol. 8271, pp. 87–96. Springer, Heidelberg (2013)
2. Abdullah, L., Ling, C.Y.: A fuzzy time series model for kuala lumpur composite index forecasting. In: 2011 4th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO), pp. 1–5. IEEE (2011)
3. Ardalani-Farsa, M., Zolfaghari, S.: Chaotic time series prediction with residual analysis method using hybrid elman–narx neural networks. Neurocomputing 73(13), 2540–2553 (2010)
4. Atsalakis, G.S., Valavanis, K.P.: Surveying stock market forecasting techniques–part ii: Soft computing methods. Expert Systems with Applications 36(3), 5932–5941 (2009)
5. Klose, C., Pircher, M., Sharma, S.: Univariate time series forecasting. In: Term Paper, pp. 1–46. University of Vienna - Department of Economics (2004)
6. Fan, S., Yan, T.: Two-phase air–water slug flow measurement in horizontal pipe using conductance probes and neural network. IEEE Transactions on Instrumentation and Measurement (2014)
7. Hansen, L.K., Salamon, P.: Neural network ensembles. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 993–1001 (1990)
8. Khashei, M., Mokhatab, R.F., Bijari, M., Hejazi, S.R.: A hybrid computational intelligence model for foreign exchange rate forecasting. J. Ind. Eng. Int 7(15), 15–29 (2011)
9. Kumar, K., Bhattacharya, S.: Artificial neural network vs linear discriminant analysis in credit ratings forecast: a comparative study of prediction performances. Review of Accounting and Finance 5(3), 216–227 (2006)

10. Lai, K.K., Yu, L., Wang, S.-Y., Wei, H.: A novel nonlinear neural network ensemble model for financial time series forecasting. In: Alexandrov, V.N., van Albada, G.D., Sloot, P.M.A., Dongarra, J. (eds.) ICCS 2006. LNCS, vol. 3991, pp. 790–793. Springer, Heidelberg (2006)
11. Ludwig, R.S., Piovoso, M.J.: A comparison of machine-learning classifiers for selecting money managers. Intelligent Systems in Accounting, Finance and Management 13(3), 151–164 (2005)
12. Mahmoud, S., Lotfi, A., Langensiepen, C.: Behavioural pattern identification and prediction in intelligent environments. Applied Soft Computing 13(4), 1813–1822 (2013)
13. Rana, M., Koprinska, I., Khosravi, A., Agelidis, V.G.: Prediction intervals for electricity load forecasting using neural networks. In: The 2013 International Joint Conference on Neural Networks (IJCNN), pp. 1–8. IEEE (2013)
14. Swanson, R.A.: Analysis for improving performance: Tools for diagnosing organizations and documenting workplace expertise. Berrett-Koehler Publishers (2007)
15. Timmermann, A., Granger, C.W.: Efficient market hypothesis and forecasting. International Journal of Forecasting 20(1), 15–27 (2004)
16. Ukil, A., Jordaan, J.A.: A new approach to load forecasting: Using semi-parametric method and neural networks. In: King, I., Wang, J., Chan, L.-W., Wang, D. (eds.) ICONIP 2006. LNCS, vol. 4233, pp. 974–983. Springer, Heidelberg (2006)
17. Wan, E.A., Van Der Merwe, R.: The unscented kalman filter. Kalman Filtering and Neural Networks, 221–280 (2001)
18. Wang, D., Li, Y.: A novel nonlinear rbf neural network ensemble model for financial time series forecasting. In: 2010 Third International Workshop on Advanced Computational Intelligence (IWACI), pp. 86–90. IEEE (2010)
19. Ward, S., Sherald, M.: The neural network financial wizards. Technical Analysis of Stocks and Commodities, 50–55 (1995)
20. Xie, H., Tang, H., Liao, Y.H.: Time series prediction based on narx neural networks: An advanced approach. In: International Conference on Machine Learning and Cybernetics, vol. 3, pp. 1275–1279. IEEE (2009)
21. Yao, J., Tan, C.L., Poh, H.L.: Neural networks for technical analysis: a study on klci. International journal of theoretical and applied finance 2(02), 221–241 (1999)
22. Yu, L., Wang, S., Lai, K.K.: A novel nonlinear ensemble forecasting model incorporating glar and ann for foreign exchange rates. Computers & Operations Research 32(10), 2523–2541 (2005)
23. Yunos, Z.M., Shamsuddin, S.M., Sallehuddin, R.: Data modeling for kuala lumpur composite index with anfis. In: Second Asia International Conference on Modeling & Simulation, AICMS 2008, pp. 609–614. IEEE (2008)