Financial Time Series Forecasting Using Empirical Mode Decomposition and FNN: A Study on Selected Foreign Exchange Rates

P. Nanthakumaran[#1], C. D. Tilakaratne[#2]

***Abstract*— The exchange rate, an economic indicator of the country is the relative price of one country's currency in terms of another country's currency. The stability of the exchange rate is important for a stable economic growth. Exchange rate series are non-linear and non-stationary. The fluctuations in the forecasting exchange rate are very important to the economy of the country. Researchers have proposed many hybrid machine learning models to get a more accurate forecast. This study proposes a hybrid forecasting model using Empirical Mode Decomposition (EMD) and Feedforward Neural Network (FNN) for foreign exchange rates forecasting and comparing its performance with widely used Non-linear Autoregressive (NAR) and Support Vector Regression (SVR) models. EMD is used to decompose the original non-linear and non-stationary series into several Intrinsic Mode Functions (IMFs) and one residual. The hybrid model is then used to forecast the exchange rate with IMFs and residual obtained as inputs. Empirical results obtained from forecasting daily exchange rates of Sri Lankan Rupees to Euro and Yen showed that the proposed EMD-FNN model outperforms NAR and SVR models without time series decomposition.**

***Key words* – Exchange rates, Hybrid FNN, IMFs, EMD**

I. Introduction

In international trading between two countries where one country has to make payments to another country in different currency, exchange rates are used. Exchange rates influence the trade balance, inflation, decisions of investors on foreign investments, worker remittances and the reserve position of a country. Thus, knowing the future values of exchange rates, a financial time series is vital.

Financial time series highly fluctuate and forecasting financial time series is a challenging task. There are two major challenges associated with forecasting financial time series namely non-stationarity and the statistical property of the time series changing with time and non-linearity. The issue of non-linearity in forecasting financial time series can be addressed by using machine learning models NAR and SVR to some extent. It was further attempted to improve the accuracy of forecasting the financial time series by introducing EMD (for eg. [14]).

Empirical Mode Decomposition (EMD) is used to decompose the original time series into a finite set of nearly orthogonal oscillating components, called intrinsic mode functions (IMFs) thus addressing both the issues in forecasting financial time series.

This research assesses the efficiency and the accuracy of the new hybrid NAR model proposed in literature and compares its performance with the machine learning models, NAR with SCG learning algorithm and SVR with Gaussian Radial Basis Kernel function without time series decomposition. Previous studies that are relevant to this research are presented in the Section II. The methodology used was briefly discussed under the Section III. The results obtained were presented in the Section IV followed by the Section V which compares the performance of models. Finally, the Section VI presents the conclusions of the study.

II. Related Work

Initially, an empirical analysis was conducted by [1] to explore the main characteristics of stochastic behaviour of LKR and concluded that the common behaviour of LKR on the currency of any country has a property of non-linearity, non-stationary with stochastic trend and non-normality. There is a vast number of methods and methodologies used for forecasting purposes in various fields. Of all the methods and methodologies used in literature, three classes of models namely, stochastic models, Neural Network based models and Support Vector Regression based models were identified by [2] as the most popular ones for modelling financial time series. According to [3], [4], stochastic models such as random walk process and GARCH models could be used to represent time series. However, according to reference [5], GARCH models failed to capture the variations in the exchange rate series and concluded inefficient in forecasting exchange rates. The failure of stochastic models to provide a better forecast when the currency market was influenced by random events, motivated researchers to turn into machine learning models. Studies [6] and [7] showed evidence for the improvement in forecasting accuracy when exchange rates were modelled using Neural Networks. Authors of [8] and [9] convinced researchers that SVM had some predictive power and could be used to forecast financial time series and showed evidence for the increase in accuracy of financial time series forecasting. Our previous study [5] modelled three selected exchange rates using NAR and SVR. The results suggested that SVR outperforms other two modelling approaches by means of values accuracy and directional accuracy.

For many years, machine learning models are used more extensively in forecasting chaotic time series. Researchers are nowadays interested in developing hybrid neural network models that overcomes the shortcomings of the benchmark models by combining several techniques to increase the accuracy and practicability of neural network models. Some of these concepts which were found to perform better were

Manuscript received on 15th Dec 2017. Recommended by Dr. M. G. N. A. S. Fernando on 12th June 2018.
This paper is an extended version of the paper "A Comparison of Accuracy of Forecasting Models: A Study on Selected Foreign Exchange Rates" presented at the ICTer 2017.

Ms. P. Nanthakumaran and Dr. C. D. Tilakaratne are from the Department of Statistics, University of Colombo. (*prathananthan@gmail.com, cdt@stat.cmb.ac.lk*).

text mining techniques embedded in decision algorithms [10], deep learning algorithms [11], [12], hybrid modelling approaches [13] etc. Recently, [14] argued that the performance of models based on innovative concepts can be enhanced by considering the impact of dynamic environmental conditions. They proposed an intelligent system supported with a hybrid approach which combines multiple algorithms such as EMD in signal processing to form the final approach with FNN with simple back propagation learning algorithm and showed evidence for the improvement in forecasting accuracy.

In this study, we improved the forecasting accuracy of the same exchange rates that we used in [5] by building a hybrid model by combining EMD and FNN and compared the performance of the new model with the NAR and SVR models. The proposed model decomposes the original series into IMFs and a residual using EMD and then it is decided to use the decomposed series as an input for the hybrid model.

III. Methodology

A. Dataset

Daily exchange rates of USD, EUR and JPY to LKR from 2nd July, 2012 to 31st August, 2016 (1008 trade days) obtained from official website of Central Bank of Sri Lanka, were selected for the study since they are the most commonly trading countries with Sri Lanka. However, it was found that the exchange rate of USD/LKR was fixed during 2013 to 2015, thus modelling and forecasting USD/LKR would not be appropriate and removed from the analysis. The dataset was divided as shown in Table I to avoid overfitting of machine learning models and for comparison purposes.

TABLE I
Dataset for analysis

	Training Set	**Validation Set**	**Testing Set**
Dataset	02-07-2012 to 29-04-2016 (925 observations)	03-05-2016 to 30-06-2016 (42 observations)	01-07-2016 to 31-08-2016 (41 observations)

Financial time series are chaotic in nature and hence to identify the characteristics, preliminary analysis was carried and the characteristics of the series were examined. Then, to forecast the selected foreign exchange rate series, machine learning models namely NAR and SVR models and EMD-FNN model were fitted. The forecasting accuracy of the model fitted was compared using MSE and (Directional Accuracy) DA and the most accurate model for forecasting selected exchange rates was identified.

B. Neural Network Based Models

NAR is a recurrent dynamic neural network with feedback connections enclosing several layers of the network based on linear AR models, which are commonly used for modelling time series. The network is designed in such a way that the next value of the output is based on the previous values of the output signals. Since the output is again a feedback to the input, the network can be considered as a recurrent network. If at least one nonlinear activation function is used in this network it can be considered as a non-linear AR model. The NAR neural network can be designed either as a parallel architecture or series- parallel architecture. In parallel architecture, the output is fed back as the input to FNN whereas in series parallel architecture, true value is used instead of feeding back the estimated output as in [15]. One of the most impressive features of neural networks is the ability to learn the non-linear complex patterns in the input and target as described in [16]. For this study, SCG learning algorithm was used to train the data as it uses conjugate direction to capture the pattern and is found to be fast and very effective.

C. Support Vector Regression Based Models

Support vector regression is an extension of Support Vector Machine (SVM). According to [17], SVM use linear model to implement non-linear boundary classifier through some non-linear mapping into a higher dimensional feature space. A linear model constructed in the feature space can represent a non-linear decision boundary in the original space. An optimal separating hyperplane is constructed in the feature space termed as “maximum margin hyperplane” which provides the maximum separation among decision classes. When optimising the hyperplane, the training samples that lie on the hyperplane called support vectors are considered for the optimisation problem while other samples are considered irrelevant for defining binary class boundaries. This concept is modified in SVR to fit a regression line.

D. Empirical Mode Decomposition

Empirical Mode Decomposition (EMD), a comparatively recent development in signal processing methods, makes use of the Hilbert transform [15] to decompose a time series into a number of components each with slowly varying amplitude and phase. It is an adaptive method that decomposes the signals into components that separate phenomena occurring on different time scales. Each component of the EMD is called an Intrinsic Mode Function (IMF). Two characteristics of IMFs noted by [18] that resemble a generalised Fourier decomposition are:

An IMF only has one zero between successive extrema.

An IMF must have zero local mean.

The computation of IMFs that satisfies the criteria mentioned above and sum to create the signal explained by [18] are:

1. The highest frequency of IMF is determined by first fitting a cubic spline through all local maxima to create an upper envelope.
2. Similarly, a lower envelope is constructed by fitting a cubic spline through all local minima.
3. Envelopes together form the candidate for the time-varying amplitude of IMF
4. It makes sure that the component function has a negligible local mean. If the average of two envelopes does not fall uniformly within the threshold value of zero, the mean is subtracted from the envelope. This step is repeated until the second criterion is satisfied.
5. The first IMF is obtained as a result of the inner loop and the outer loop gives the residue.
6. Steps through 1 to 5 are repeated until the residue obtained is a constant or a monotone.

This procedure decomposes the series as follows:

$$y_t = \sum_{all\ i} IMF_i + Residue \quad (1)$$

where y_t is the time series to be decomposed. EMD outputs IMFs of a very close visual match to the original signal components, with nearly indistinguishable amplitudes and frequencies thus addressing the issue of non-stationarity and multi-scaling behaviour of financial time series.

E. EMD-FNN models

The proposed hybrid ANN model is a combination of EMD and ANN models. EMD is used to decompose the original series into locally orthogonal set. These IMFs and residual are fed into the FNN with SBP learning algorithm as input and the forecasted value is obtained as the output.

Fig 1 explains the procedure followed in fitting the proposed EMD-FNN forecasting model for the exchange rates series.

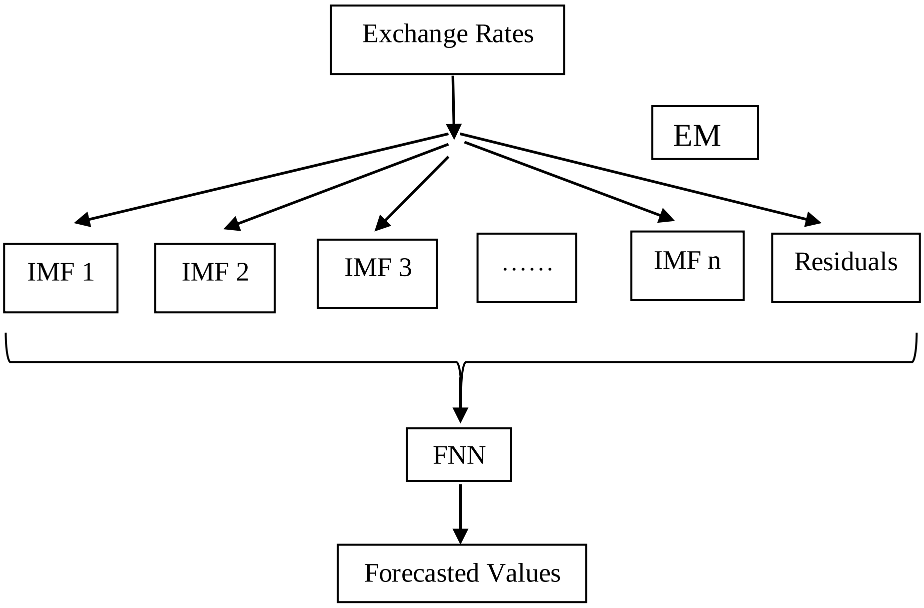

Fig1. The proposed EMD-FNN forecasting model for exchange rates

F. Performance criteria

The value forecasting accuracy and the directional forecasting accuracy were measured using Mean Square Error (MSE) and Directional Accuracy (DA) respectively.

$$MSE = \sum_{i=1}^{n}(y_t - \hat{y}_t)^2 / n \quad (2)$$

where y_t is the actual value, $\hat{y}_t$ is the forecasted value and n is the number of observations.

$$DA = \frac{1}{N}\sum_t f(t) \quad (3)$$

where,

$$f(t) = \begin{cases} 1 & if\ sign\ of\ actual\ and\ forecasted\ value\ is\ same \\ 0 & otherwise \end{cases}$$

The model with high DA representing high directional accuracy and low MSE representing fitted value close to actual value was chosen as the best model for forecasting.

IV. RESULTS

The results of the descriptive analysis and the model fitting procedure are presented in this section.

A. *Preliminary Analysis*

Time plots of the exchange rates are examined in order to identify their behaviour. Figure 2 and 3 depicts the time plot and correlogram of EUR/LKR series respectively.

According to Fig 2, EUR/LKR shows random fluctuations without any seasonality during the study period.

Fig. 3 reveals that the ACF of the original series are highly significant for more lags and does not decline quickly to zero.

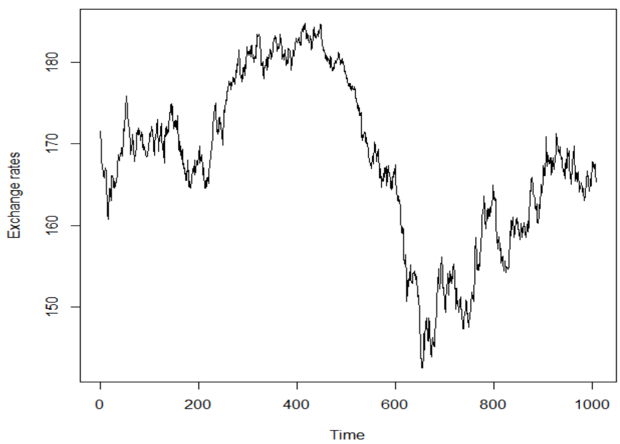

Fig 2 Time plot of EUR/LKR

Also, the partial autocorrelations of the original series are not significant for lower lags except for lag 1. Thus, EUR/LKR series might consist of a stochastic trend and is not stationary.

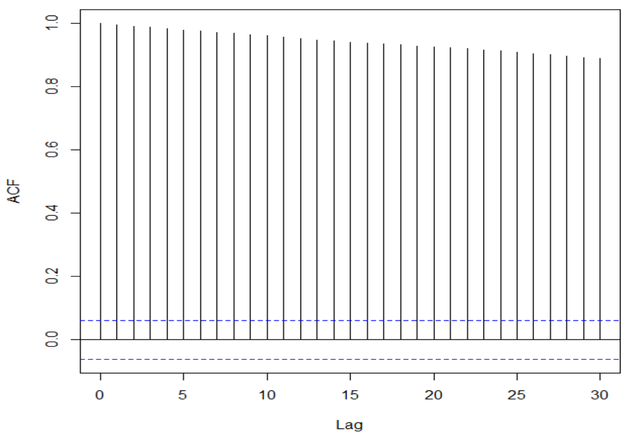

Fig. 3 ACF plot of EUR/LKR

The results obtained from the unit root test of EUR/LKR indicated evidence for the presence of stochastic trend without any drift or deterministic trend in the original series. The 1st differenced series did not show any stochastic trend or drift and became stationary. Thus, $\nabla y_{t_{EU}} \sim I(0)$ implies $y_{t_{EU}} \sim I(1)$ where $y_{t_{EU}}$ is the exchange rate of EUR/LKR on time t.

The return series of EUR/LKR was obtained using the equation,

$$r_{t_{EU}} = y_{t_{EU}} - y_{t-1_{EU}} \tag{4}$$

Fig. 4 represents the return series of EUR/LKR and shows that high periods of volatility are followed by low periods of volatility. Return series indicate the presence of volatility clusters.

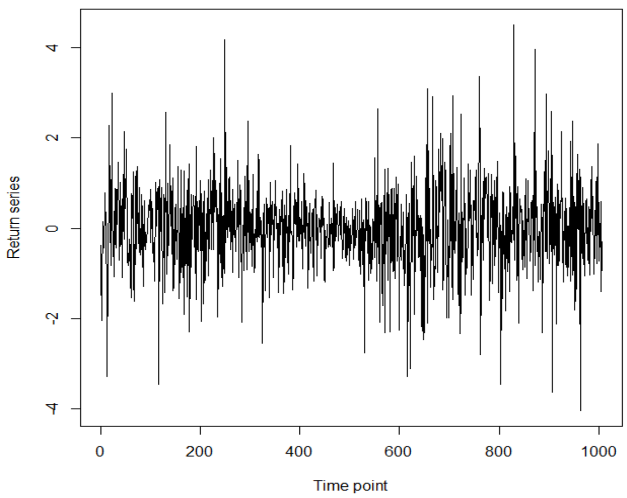

Fig 4 Return series of EUR/LKR

Figure 5 and 6 depicts the time plot and correlogram of JPY/LKR series respectively.

According to Fig 5, JPY/LKR shows random fluctuations without any seasonality during the study period.

Fig. 6 shows that the ACF of JPY/LKR series are highly significant for more lag values and does not decline quickly to zero whereas PACF of JPY/LKR are not significant

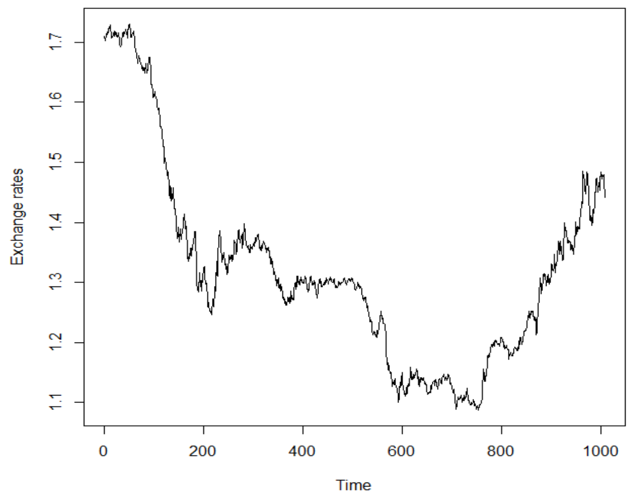

Fig. 5 Time plot of JPY/LKR

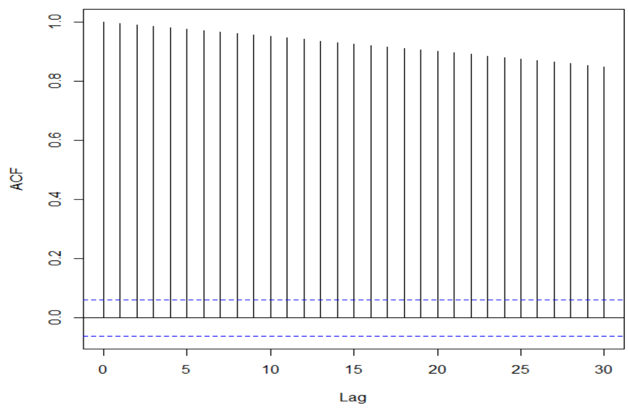

Fig. 6 ACF plot of JPY/LKR

PACF of JPY/LKR are not significant for lower lags except for lag 1. Thus, it can be assumed that the original series has trend and not stationary.

The results obtained from unit root test of JPY/LKR provide evidence for the presence of stochastic trend and deterministic trend without any drift in the original series. The 1st differenced series shows deterministic trend with drift but no stochastic trend. 2nd differenced series does not show any deterministic trend and is stationary. Thus, $\nabla^2 y_{t_{Japan}} \sim I(0)$ implies $y_{t_{Japan}} \sim I(1)$ + quadratic deterministic trend + drift where $y_{t_{Japan}}$ is the exchange rate of JPY/LKR on time t.

The return series of JPY/LKR was obtained using the equation,

$$r_{t_{JP}} = y_{t_{JP}} - y_{t-1_{JP}}, \tag{5}$$

Fig. 7 represents the return series of JPY/LKR and shows that the variance of JPY/LKR returns is not constant across time and is highly volatile and indicated the presence of volatility clusters.

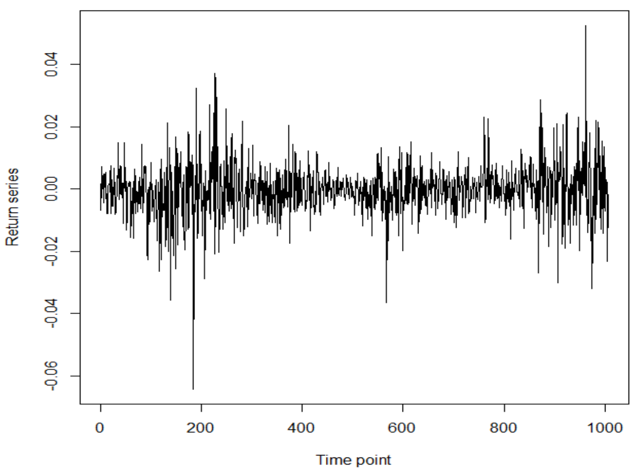

Fig 7 Return series of JPY/LKR

$$r_t = y_t - y_{t-1} \tag{6}$$

The autoregressive terms included in the model were selected by fixing default values to the parameters and changing the number of lags included. The MSE values and the adequacy of the model were compared for many iterations. The one that gave minimum MSE with adequacy for most of the iterations was chosen as the optimal lag. When the number of hidden layers included in the model was increased from one, the MSE value did not decrease much. Thus, one hidden layer was included in the model. The number of nodes in the hidden layer and the learning algorithm parameters sigma and lambda were found using trial and error method. The model was repeatedly fitted with the same set of parameters for each of the 20 iterations. Each time the model trained, MSE values of training and validation sets kept changing and the adequacy of the model was fluctuating between adequate and inadequate. A model is considered to be adequate if the squared correlation is not significant for lower lag values. Different combinations of the parameters were tried and different MSE values and adequacy were obtained for each iteration. However, when the optimal set of parameters was used, the MSE value and the adequacy did not change drastically for different iterations. The models thus obtained have been discussed below.

The architecture of the designed NAR for forecasting EUR/LKR included 8 autoregressive terms with one hidden layer made up of 14 nodes and one output layer made up of one node. *'mapminmax'* function was used to pre-process the data in the input layer and *'tansig'*, and *'purelin'* transfer functions were used in the hidden and output layer respectively. The optimal learning algorithm parameters selected were 1e-10 for sigma and 1e-7 for lambda. The designed NAR for EUR/LKR used SCG as the learning algorithm to train the dataset and MSE as a measure to check the performance of the network.

Fig 8 depicts the residual plot of the designed NAR model for EUR/LKR.

According to Fig. 8, the designed NAR described the dataset well with error ranging from -4 to +4 through zero.

Also, there was no significant autocorrelation for lower lag values among the residuals and cross-correlation

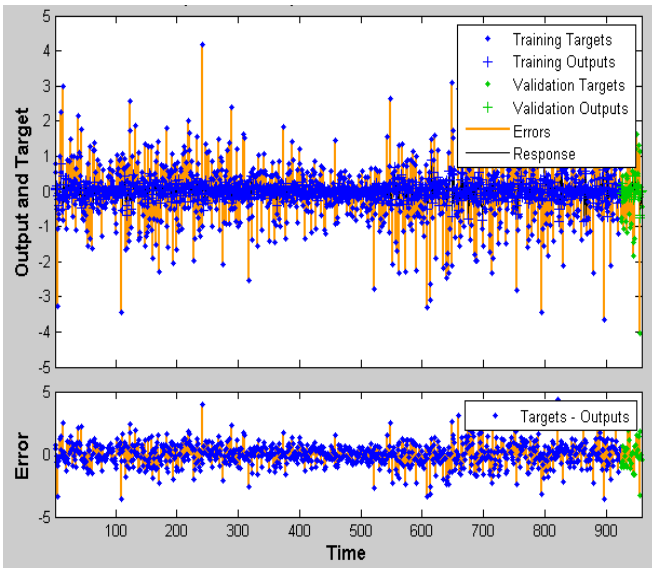

Fig 8 Residual plot

between the residual and input terms. Thus, the NAR model fitted was adequate.

Architecture of the designed NAR for forecasting JPY/LKR included 12 autoregressive terms with one hidden layer of 18 nodes and one output layer of one node. *'mapminmax'* function was used to pre-process the data in the input layer and , and *'purelin'* transfer functions were used in the hidden and output layer respectively. The optimal learning algorithm parameters included were 1e-10 for sigma and 1e-10 for lambda. The designed NAR used SCG as a learning algorithm to train the dataset with MSE used as a measure to check the performance of the network.

Fig 9 depicts the residual plot of the designed NAR model for JPY/LKR.

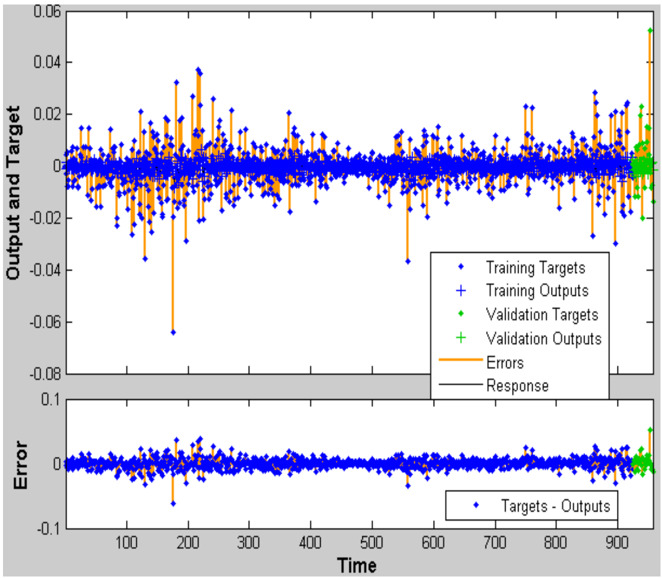

Fig 9 Residual plot

Fig. 9 shows that the designed NAR has described the dataset well with error ranging from -0.1 to +0.1 through zero. Also, there was no significant autocorrelation for lower lag values among the residuals and no significant cross-correlation between the residual and input terms. Thus, the NAR model fitted was adequate.

B. SVR Models

The SVR model was fitted in return for the exchange rate obtained by,

$$r_t = y_t - y_{t-1} \tag{7}$$

on training dataset with the function $'ksvm()'$ in R programming. The parameters in the model were lag values, epsilon, cost and sigma corresponding to the selected kernel function and the Gaussian radial basis kernel function.

Initially, by allowing default values for epsilon, sigma and cost, the lag values were changed from 3 to 9 and the best lag value was chosen as the one with minimum MSE in the training set. Then, the cost and sigma values were fixed and epsilon value was changed between 0 and 1. The epsilon value with the minimum MSE value in the validation set was selected. Then, the selected epsilon value and default sigma value were fixed and cost was changed between 0 and 2^{10}. The cost value with the minimum MSE value in the validation set was selected. Again, the selected epsilon value and cost value were fixed and sigma was changed between 0 and 1 and the combination with minimum MSE value in the validation set was selected. The process was repeated until a minimum MSE value was obtained in the validation set. Different combinations of values were obtained with different MSE values for training and validation set. The set of values with the minimum MSE in training and validation set was chosen as the optimal set of values. SVR thus designed were discussed below.

TABLE II
SELECTION OF PARAMETERS FOR SVR OF EUR/LKR RETURNS

No.	Cost	Epsilon	Sigma	Training MSE	Validation MSE
1	1024	0.1	0.95	0.0089	1.8037
2	1024	0.7	0.1	0.5047	4.2350
3	1024	0.95	0.95	0.4587	1.5403
4	1024	0.7	0.95	0.3017	1.6331
5	1	0.95	0.95	0.5867	1.3980
6	1	0.7	0.95	0.5101	1.3920
7	1	0.09	0.9	0.4057	1.3909
8	1	0.1	0.95	0.3940	1.3933

Six autoregressive terms were included in the model as the MSE value was minimum when lag value was six. Table II shows a few combinations of values, their corresponding minimum MSE values for SVR model in returns for EUR/LKR. The optimal set chosen were epsilon = 0.09, sigma = 0.9, cost = 1 as these combinations of parameter values gave a minimum MSE value in both training as well as validation set. Thus, the SVR model was designed including 6 autoregressive terms with an insensitive loss of 0.09, cost of 1 and Gaussian radial basis kernel function with sigma of 0.9.

Fig 10 and 11 illustrates the comparison of the fitted value and the actual value of the training dataset and the validation dataset of EUR/LKR respectively.

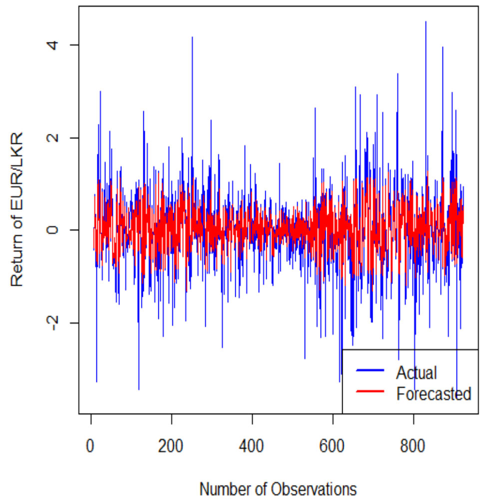

Fig 10 Plot of fitted vs actual value in training set

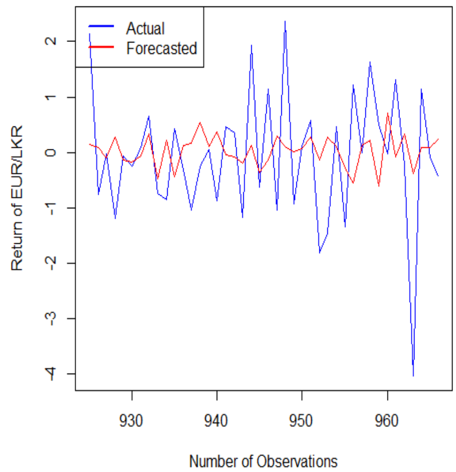

Fig 11 Plot of fitted vs actual value in validation set

The designed SVR model was well fitted with the training set (Fig. 10). Also, the model captured the variation in the validation set but underestimated in few cases (Fig. 11). Autocorrelation of residuals in the training set as well as in the validation set were not significant for lower lag values. The autocorrelation of squared residuals in training set and validation set were not significant for lower lag values. Hence, all the patterns in the training as well as validation set were captured by the model. Thus, the designed model was adequate.

TABLE III
SELECTION OF PARAMETERS FOR SVR OF JPY/LKR RETURNS

No.	Cost	Epsilon	Sigma	Training MSE	Validation MSE
1	1024	0.1	0.9	6.91E-06	0.000215
2	1024	0.9	0.5	3.39E-05	0.000229
3	1024	0.3	0.1	2.34E-05	0.000995
4	1024	0.9	0.9	3.00E-05	0.000276
5	1024	0.9	0.5	3.39E-05	0.000229
6	20	0.1	0.55	9.65E-06	0.000214
7	20	0.1	0.5	1.09E-05	0.000225
8	20	0.9	0.5	3.16E-05	0.000289

Six autoregressive terms were included in the model as the MSE value was minimum when lag value was 6. Table III shows a few combinations of values and their corresponding minimum MSE values for SVR model in return for JPY/LKR. The optimal set chosen was epsilon = 0.1, sigma = 0.55, cost = 20 as these combinations of parameter values gave a minimum MSE value in both training as well as validation set. Thus, SVR model for JPY/LKR series was designed by including 6 autoregressive terms with an insensitive loss of 0.1, cost of 20 and Gaussian radial basis kernel function with sigma of 0.55.

Fig 12 and 13 illustrates the comparison between the fitted value and the actual value in the training dataset and the validation dataset of JPY/LKR respectively.

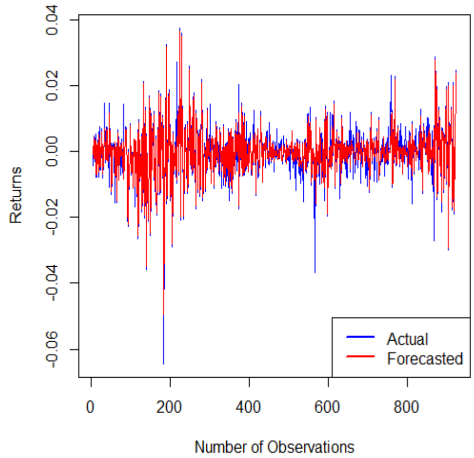

Fig 12 Plot of fitted vs actual value in training set

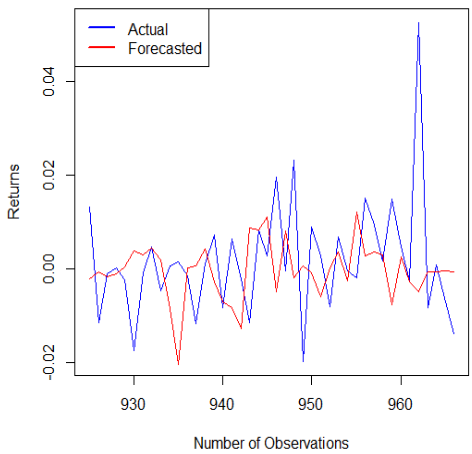

Fig 13 Plot of fitted vs actual value in validation set

The designed SVR model is fitted with the training set well (Fig. 12). The model also captured the variations in validation set with a few exceptions (Fig. 13). The autocorrelation of errors and autocorrelation of squared errors were not significant for lower lags in the training as well as validation set. Hence, all the patterns in the series were captured by the model. Thus, the designed SVR model was adequate.

D. *Empirical Mode Decomposition of Exchange Rates*

The exchange rate of EUR/LKR was decomposed into 10 IMFs using EMD (Fig. 14)

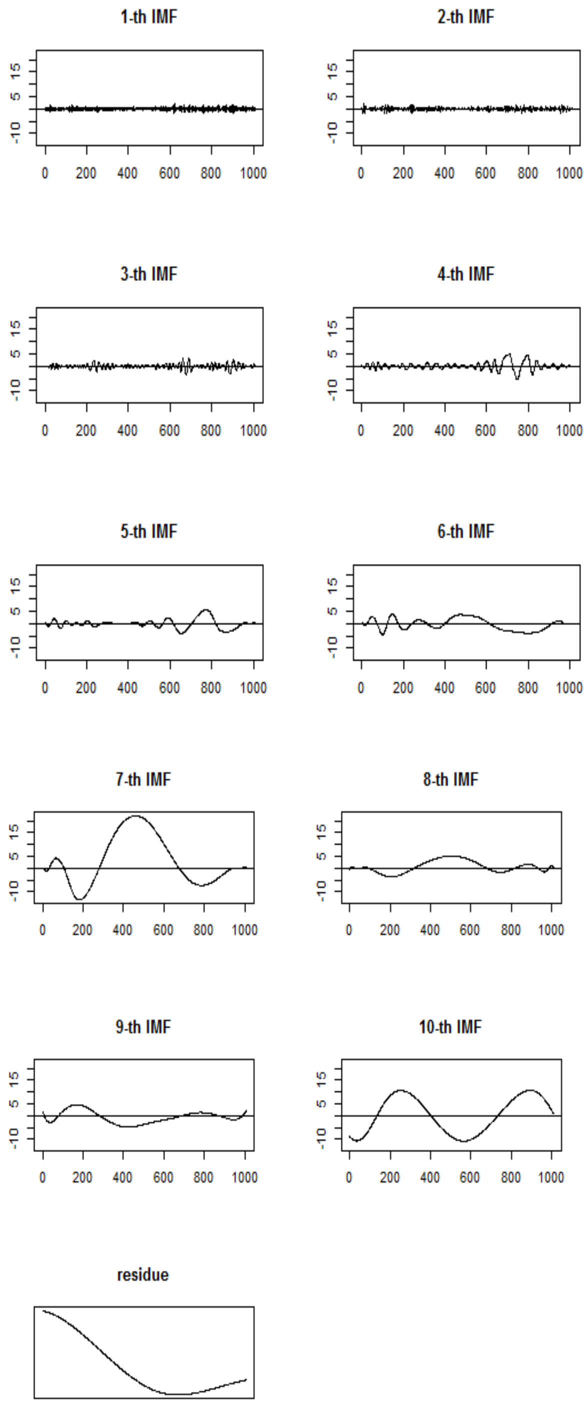

Fig.14 IMFs obtained for EUR/LKR series

The exchange rate of JPY/LKR was decomposed into 7 IMFs using EMD (Fig. 15).

E. *EMD-FNN Model*

The identified IMFs and residual series as the inputs were fed into FNN trained with simple back propagation as learning algorithm using R-programming. The parameters to be optimised are number of successive past observations of a time series to be included, number of hidden layers and number of nodes in the hidden layers.

The optimal value for the above-mentioned parameters was identified by trial and error method.

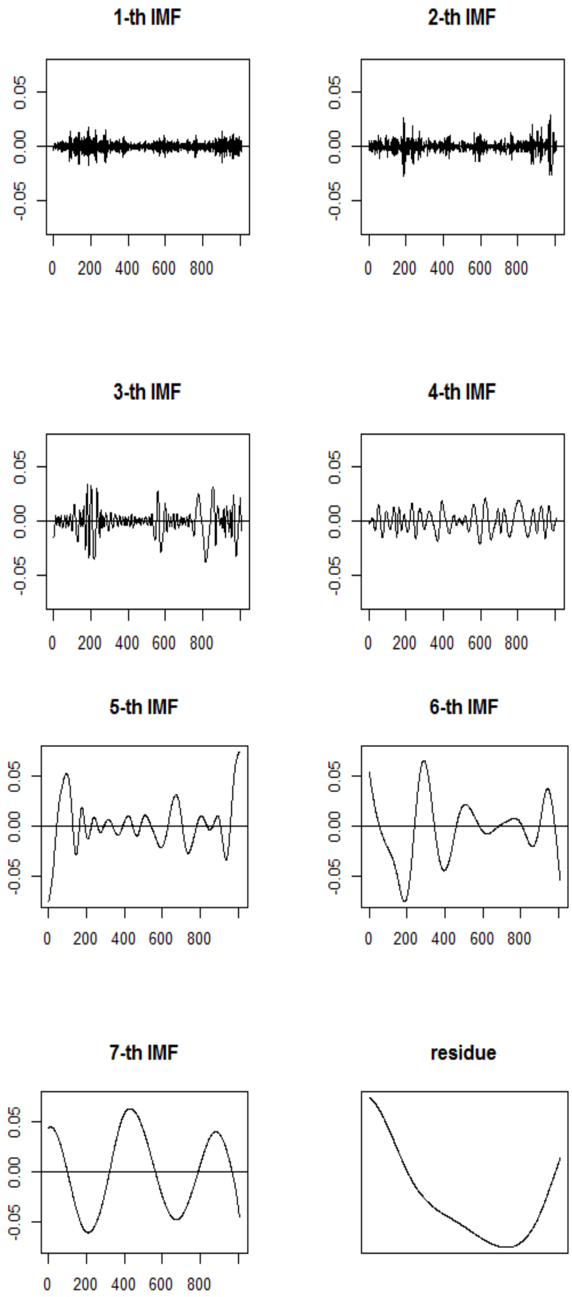

Fig. 15 IMFs obtained for JPY/LKR series

The model with minimum MSE value for both training and validation set was selected as the one model with optimal parameter values.

First, the model for forecasting EUR/LKR was trained for 20 iterations each by allowing one lag term and inserting one hidden layer with varying number of nodes and two hidden layers with different number of nodes. Then, the model was trained for 20 iterations each by allowing two lag terms and inserting one hidden layer with different number of nodes and two hidden layers with different number of nodes. The procedure was repeated allowing 3 lag terms and it was found that the increase in the number of lag terms did not show any significant increase in the model performance. Hence, one was selected as the optimal value for number of lags to be considered.

Next, to find out the number of hidden layers and number of nodes in each layer, the same approach was followed. The FNN model with one lag term as input with one hidden layer was fitted with different number of nodes each with 10 iterations. The minimum MSE values obtained for different number of nodes included were shown in Table IV.

TABLE IV
SELECTION OF NUMBER OF NODES IN HIDDEN LAYER 1

No. of nodes in Hidden Layer 1	Training MSE	Validation MSE
1	0.645157	0.796348
2	0.578376	2.33221
3	0.59515	6.249432
4	0.577441	4.512481
5	0.570462	4.872282
6	0.590265	14.43854
7	0.600194	6.80946
8	0.613668	12.14485
9	0.573512	0.920267
10	0.543195	2.10157

Thus, one node was selected as the optimal number of nodes in the hidden layer 1. Next, the procedure was followed until the optimal number of hidden layers and their corresponding numbers of nodes were found. Tables V and VI show the MSE values obtained for training and validation datasets of EUR/LKR in the training process of EMD-FNN model.

TABLE V
SELECTION OF NUMBER OF NODES IN HIDDEN LAYER 2

No. of nodes in hidden layer 1	No. of nodes in hidden layer 2	Training MSE	Validation MSE
1	1	0.628482	0.849926
	2	0.687404	0.80997
	3	0.560023	0.732678
	4	0.558333	0.73448
	5	0.568006	0.640553
	6	0.587987	0.868986
	7	0.561575	0.715382
	8	0.588804	0.912007
	9	0.569004	0.695933
	10	0.67631	0.80328

As the procedure was repeated further including hidden layers, the performance in the neural network did not vary significantly. Thus, three was selected as the optimal number of hidden layers with one node in hidden layer 1, five nodes in hidden layer 2 and four nodes in hidden layer 3.

Fig 16 and 17 compares the actual value and forecasted value obtained from EMD-FNN model corresponding to the training set and the validation set of EUR/LKR respectively.

TABLE VI
SELECTION OF NUMBER OF NODES IN HIDDEN LAYER 3

No. of nodes in hidden layer 1	No. of nodes in hidden layer 2	No. of nodes in hidden layer 3	Training MSE	Validation MSE
1	5	1	0.557888	0.706314
		2	0.606907	0.678023
		3	0.574716	0.679122
		4	0.561239	0.638103
		5	0.616014	0.688559
		6	0.569479	0.681645
		7	0.567103	0.66495
		8	0.565119	0.646873
		9	0.571301	0.634567
		10	0.572476	0.685405

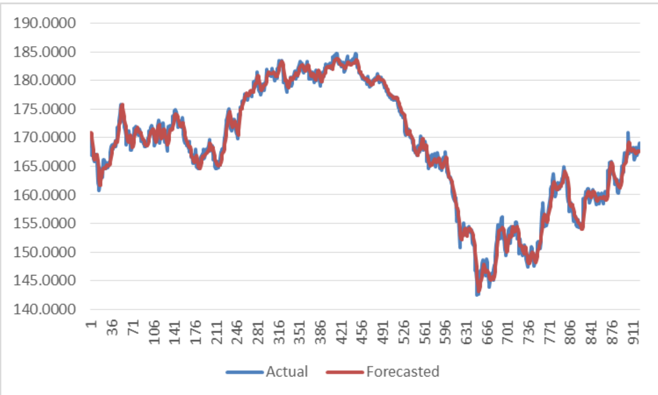

Fig 16 Plot of fitted vs actual values in the training set

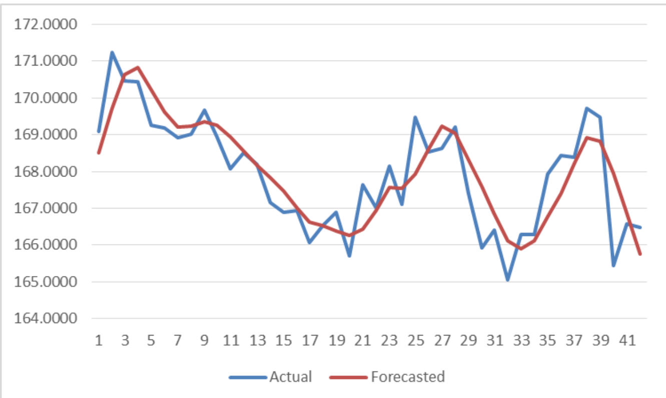

Fig 17 Plot of fitted vs actual values in the validation set

The designed EMD-FNN model for EUR/LKR captured the variation in the series very well and forecasted the training set (Fig.16) and the validation set (Fig. 17) well. The autocorrelation of errors and autocorrelation of squared errors were not significant for lower lags in the training as well as validation set. Hence, all the pattern in the series was captured by the model. Thus, the designed Hybrid ANN model for EUR/LKR was adequate.

Now, the model for forecasting JPY/LKR was trained for 20 iterations each by allowing one successive past observation and inserting one hidden layer with different number of nodes and two hidden layers with different number of nodes. Then, the model was trained for 20 iteration each by allowing two successive past observations and inserting one hidden layer with different number of nodes and two hidden layers with different number of nodes. The procedure was repeated allowing 3 successive past observations and it was found that the increase in the number of successive past observations did not show any significant increase in the model performance. Hence, the model with one successive past observation was selected as the optimal model.

Next, to find out the number of hidden layers and number of nodes in each layer, the same approach was followed. The ANN model with one lag term as input with one hidden layer was fitted with different number of nodes each with 20 iterations. The minimum MSE values obtained for different number of nodes included were shown in Table VII.

TABLE VII
SELECTION OF NUMBER OF NODES IN HIDDEN LAYER 1

No. of nodes in hidden layer 1	Training MSE	Validation MSE
1	5.9771E-05	0.000105
2	5.6537E-05	0.000107
3	4.6142E-05	0.000104
4	5.1956E-05	0.000132
5	5.3601E-05	0.000122
6	4.9831E-05	0.000126
7	5.0953E-05	0.000101
8	5.1252E-05	0.000127
9	5.496E-05	0.000131
10	5.2047E-05	0.000125

Thus, three nodes were selected as the optimal number of nodes in the hidden layer 1. Next, the procedure was followed until the optimal number of hidden layers and their corresponding numbers of nodes were found. Table VIII and IX show the MSE values obtained for training and validation datasets of JPY/LKR in the training process of EMD-FNN.

As the procedure was repeated further including hidden layers, the performance in the neural network did not vary significantly. Thus, two was selected as the optimal number of hidden layers with three nodes in hidden layer 1 and one nodes in hidden layer 2.

Fig 18 and 19 compares the actual value and forecasted value obtained from EMD-FNN model corresponding to the training set and the validation set of JPY/LKR respectively.

The designed EMD-FNN model for JPY/LKR captured the variation in the series very well and forecasted exchange rates corresponding to the training set (Fig.18) and the validation set (Fig. 19) well. The autocorrelation of errors and autocorrelation of squared errors were not significant for lower lags in the training as well as validation set. Hence, all the patterns in the series were captured by the model.

Thus, the designed Hybrid ANN model for JPY/LKR was adequate.

TABLE VIII
SELECTION OF NUMBER OF NODES IN HIDDEN LAYER 2

No. of nodes in hidden layer 1	No. of nodes in hidden layer 2	Training MSE	Validation MSE
3	1	4.39E-05	9.67E-05
	2	5.55E-05	0.000101
	3	5.57E-05	0.000117
	4	5.4E-05	0.000116
	5	4.84E-05	0.000132
	6	6.05E-05	0.000129
	7	5.34E-05	0.000126
	8	7E-05	9.04E-05
	9	5.86E-05	0.000109
	10	5.58E-05	0.000112

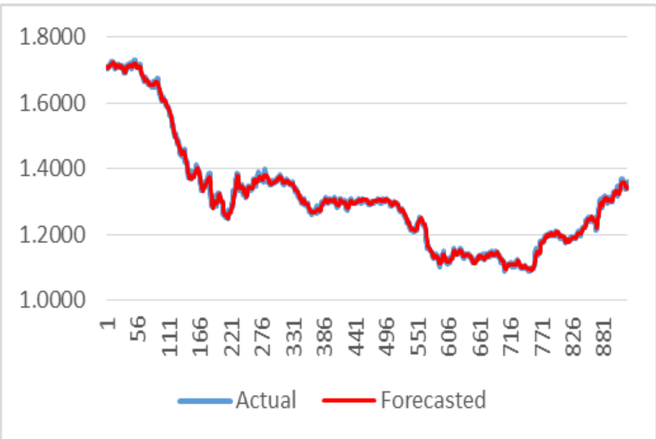

Fig 18 Plot of fitted vs actual in the training set

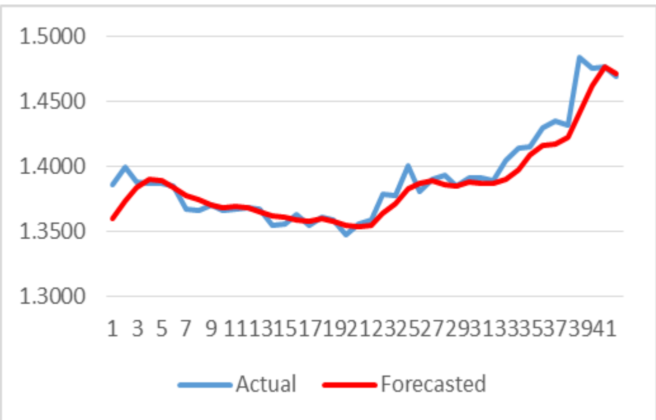

Fig 19 Plot of fitted vs actual in the validation set

F. *COMPARISON OF FITTED MODELS*

The forecasting accuracy of the models was compared using the test dataset. A multistep ahead forecast for 41 time points were obtained and the values and directional forecasts were compared using MSE and DA respectively.

Table X compares the forecasting accuracy of models fitted for the selected exchange rates.

TABLE X
COMPARISON OF FORECASTING ACCURACY OF MODELS FITTED FOR SELECTED EXCHANGE RATES

Exchange Rates	Model	Mean Square Error	Directional Accuracy
EUR/LKR	NAR Model	1.7279	60.98%
	SVR Model	2.8784	65.85%
	Hybrid (EMD-FNN) Model	0.3378	67.5%
JPY/LKR	NAR Model	0.0377	60.98%
	SVR Model	0.0005	73.17%
	Hybrid (EMD-FNN) Model	0.0002	77.5%

According to Table X, hybrid ANN model provided the best directional accuracy of about 67.5% and provided a good value forecast (minimum MSE) when considering EUR/LKR. Table X also showed that hybrid ANN model provided the best directional accuracy of about 77.5% and provided a good value forecast for JPY/LKR.

Fig 20 and 21 visualise and compare the actual value and the forecasted values obtained from different models fitted namely NAR model, SVR model and EMD-FNN model for EUR/LKR and JPY/LKR respectively.

According to Fig. 20, the forecasted values of EMD-FNN models were close to the actual values than that of other machine learning models when considering EUR/LKR. According to Fig. 21, the forecasted values of hybrid ANN models are close to the actual values of the testing period, than those of other machine learning models for JPY/LKR.

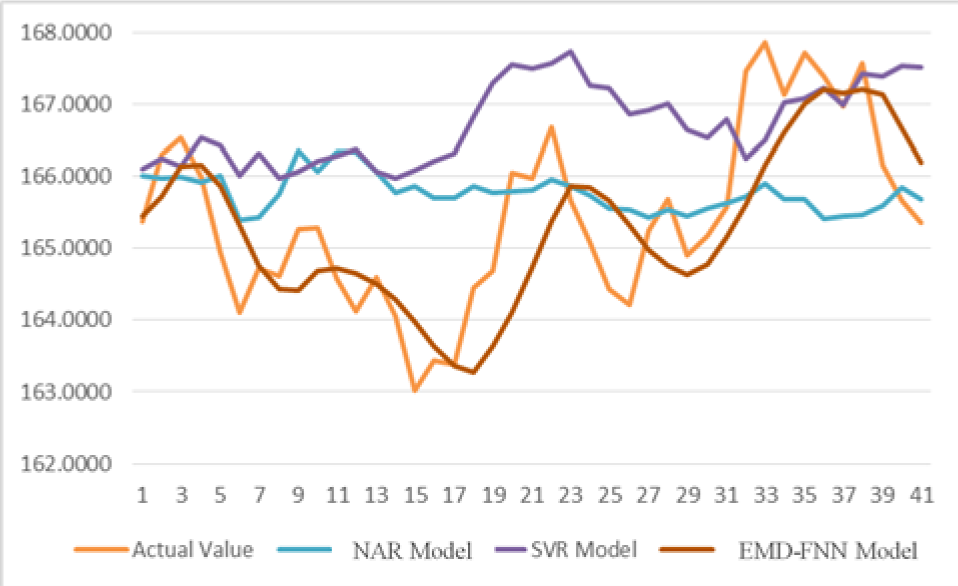

Fig 20 Graphical comparison of forecasting accuracy of different models for EUR/LKR

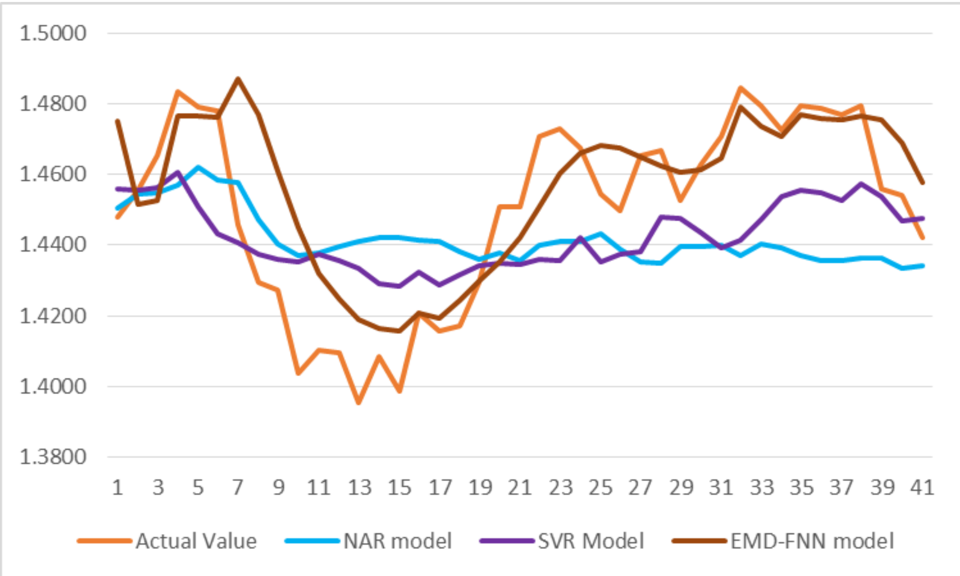

Fig 21 Graphical comparison of forecasting accuracy of models for JPY/LKR

V. Conclusions

The conclusions obtained were based on the datasets selected for the study. All models performed well in forecasting the daily exchange rates selected. However, the proposed hybrid EMD-FNN model gave a better value forecast as well as a directional accuracy for both datasets.

The results might differ for different datasets. Also, according to [19], the country had adopted floating exchange rate system since 2001 allowing an independent adjustment of exchange rate depending on the market forces. However, in practice, they have deviated from this stated policy and intervened to stabilise the exchange rate as in [20]. These findings could be useful to domestic as well as foreign users. Further, the forecasting ability can be improved by introducing evolutionary neural networks that have significant predictive and market timing ability.

Acknowledgment

Authors are very much indebted to Central Bank, Sri Lanka for providing the reliable dataset.

References

[1] S. Sivarajasingham, and A.M.M. Mustafa, (2011). "An empirical Analysis of stochastic behaviour of Sri Lankan exchange rate". [Online]. Available: http://ir.lib.seu.ac.lk/handle/123456789/1099

[2] R. Adhikari, and R. K.Agrawal, An Introductory Study on Time Series Modelling and Forecasting, Saarbrücken, Germany: LAMBERT Academic Publishing 2013.

[3] D. E.Allen, M.McAleer, S. Peiris and A. K. Singh, Non-linear Time Series and Neural-Network Models of Exchange Rates between the US Dollar and Major Currencies. Risks 4(7), 1-14. doi:10.3390/risks4010007, 2016

[4] A.Thavaneswaran, S. S. Appadoo and S. Peiris, "Forecasting volatility," *Statistics and Probability Lett.,* vol. 75, pp. 1-10, Jul. 2005.

[5] P. Nanthakumaran, and C. D. Tilakaratne. "A Comparison of Accuracy of Forecasting Models: A Study on Selected Foreign Exchange Rates", Proc. ICTer, 2017, pp. 324 – 331.

[6] A. Emam, and H. Min, The artificial neural network for forecasting foreign exchange rates. International Journal of Services and Operations Management 5(6), 740-757. doi: 10.1504/IJSOM.2009.026772, 2009

[7] J. Kamruzzaman, and R. Sarker, "Forecasting of currency exchange rates using ANN: A case study". [Online]. Available: https://www.researchgate.net/publication/224746570

[8] D. Z. Cao, S. L. Pang, and Y. H. Bai. "Forecasting exchange rate using support vector machines," in Proc. of ICMLC, 2005, p. 3448

[9] L. J. Cao, and F. E. H. Tay, "Support Vector Machine With Adaptive Parameters in Financial Time Series Forecasting," *IEEE TRANSCATIONS ON NEURAL NETWORKS* vol. 14(6), pp.1506-1518, Nov. 2003

[10] Al Nasseri, A. Tucker, and S. de Cesare. "Quantifying stocktwits semantic terms'trading behaviour in financial markets: An effective application of decision tree algorithms" [Online]. Available: http://www.sciencedirect.com/science.

[11] W. Cao, L. Hu, and L.Cao. "Deep modelling complex couplings within financial markets" [Online]. Citeseer. 2015. Available: http://www.staff.it.uts.edu.au

[12] M. F. Dixon, D. Klabjan, and J. H. Bang, "Classification-based financial markets prediction using deep neural networks" [Online]. 2016. Available: http: //www.arxiv.org.

[13] J. Chai, et al. "A hybrid leaset square support vector machine model with parameters optimization for stock forecasting" [Online]. 2015. Available: https://www.hindawi.com.

[14] H. A. Pathberiya, C. D. Tilakaratne, and L. L. Hansen. "An Intelligent System for Forex Trading: Hybrid ANN with GARCH and Intrinsic Mode FUnctions" in Proc. IntelliSys, 2017, pp. 436 – 445.

[15] MATLAB2012a help

[16] L. Jacobson, (2014). Introduction to Artificial Neural Network Part 2 – Learning. [Online]. Available: http://www.theprojectspot.com/tutorial-post-

[17] M. Alamili, "Exchange Rate Prediction using Support Vector Machines: A comparison with Artificial Neural Networks" (Unpublished Master's thesis). Delft University of Technology, Delft, Netherlands, Jan. 2011.

[18] S. Tolwinski, "The Hilbert Transform and Empirical Mode Decomposition as Tools for Data Analysis" [Online]. Available: http://math.arizona.edu/~flaschka/Topmatter/527files/termpapers/suz.pdf

[19] (2016) The CBSL website. [online]. Available: http://www.cbsl.gov.lk/

[20] L. Somaratne, (2016). Daily Financial Times home page. [Online]. Available: http://www.ft.lk/article-