FORECASTING THE EXCHANGE RATE SERIES WITH ANN: THE CASE OF TURKEY

Doç. Dr. Cem KADILAR* Prof. Dr. Muammer ŞİMŞEK**

Araş. Gör. Çağdaş Hakan ALADAĞ***

Abstract

As it is possible to model both linear and nonlinear structures in time series by using Artificial Neural Network (ANN), it is suitable to apply this method to the chaotic series having nonlinear component. Therefore, in this study, we propose to employ ANN method for high volatility Turkish TL/US dollar exchange rate series and the results show that ANN method has the best forecasting accuracy with respect to time series models, such as seasonal ARIMA and ARCH models. The suggestions about the details of the usage of ANN method are also made for the exchange rate of Turkey.

Keywords: Activation function, ARIMA, ARCH, Artificial neural network, Chaotic series, Exchange rate, Forecasting, Time series

Jel Classification: C220, C450, C530, F310, G170

Özet


Anahtar Kelimeler: Aktivasyon fonksiyonu, ARIMA, ARCH, Döviz kuru, Kaotik seriler, Öngörü, Yapay sinir ağları, Zaman serileri

Jel Sınıflaması: C220, C450, C530, F310, G170

* Adres: Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü
E-Mail: kadilar@hacettepe.edu.tr

** Adres: Cumhuriyet Üniversitesi, İİBF, İktisat Bölümü
E-Mail: msimsek@cumhuriyet.edu.tr

*** Adres: Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü
E-Mail: aladag@hacettepe.edu.tr
1. Introduction

Along with the beginning of the application of the flexible exchange rate system, prediction of exchange rate movements gained importance. Changes in exchange rates have an effect on imports, volume of external trade, balance of payments, inflation and public debt. In this respect, exchange rate is an important factor in an economy as both an important indicator of total demand and also in the application of financial policies.

However, exchange rates may not always be completely predictable. This is because movements in exchange rates have highly varying, chaotic and noisy structures. This characteristic makes exchange rates difficult to predict. For this reason, estimating exchange rate movements is always a difficult and important topic in academic frameworks and business life and, therefore, this has been one of the main concerns of academics and other researchers in multi-national financing.

Efforts to create a greater understanding of exchange rate movements have brought together many approaches. Studies in this area first started with low frequency macroeconomic models. Later, studies focused on high frequency models which were based on microeconomic variables. This was due to some researchers realizing that it was not possible to predict exchange rates with classical models (Meese and Rogoff, 1983). One of the most important reasons for this is the inadequacy of traditional exchange rate models (Yu et al, 2007). In the prediction of exchange rates, in which practical views are as important as theoretical views, many methods and techniques have been tried with the random walk model.

For this reason, because most econometric methods do not produce more significant predictions than the Random Walk (RW) model, a necessary tendency occurred in relation to exchange rate estimations. Recent studies, which have improved the performance of simple RW model, have obtained some evidence that non-linear models create better prediction results.

By the development of ANN, researchers and investors are hoping that they can solve the mystery of exchange rate predictions. It has been proved that the ANN model, which is a type of non-linear model, is a strong alternative in the prediction of exchange rates. ANN is a very suitable method to find correct solutions especially in a situation which has complex, noisy, irrelevant or partial information.

The main reason for the choice of ANN as a vehicle for the prediction of an exchange rate is that it possesses some important and attractive characteristics. The first of these, ANN, as opposed to prediction methods based on classical models, is a method which has few limiting hypotheses and can be easily adapted to the types of data. This characteristic of ANN is a highly desired situation in cases where basic data is not obtained in some of the financial predictions to be carried out (Qi and Zhang, 2001). Secondly, ANN can be generalized. Thirdly, ANN has a general functional structure (Hornik et al, 1989). Furthermore, ANN can be classified as non-linear models (Zhang et al, 1998). However, alongside these, it is too early to state whether ANN will be able to solve all of the problems which no method has hitherto been sufficiently successful in terms of solving the prediction of exchange rates.

The present exchange rate models have been developed by using various linear and non-linear models. Models based on linear framework include a simple effective market approach such as Cornell (1977), Hsieh (1984). Examples of simple random walk approach can be seen in Giddy and Duffey (1975); Hakkio and Rush (1986). Approaches based on non-linear
foundations are Dornbusch (1976); Frankel (1979); Meese and Rogoff (1983); Baillie and McMahon (1989); Hsieh (1989a); Meese and Rose (1990); Meese and Rose (1991); Mark (1995); Clark and McDonald (1998). Lin and Chen (1998); Ma and Kanas (2000); Coakley and Fuertes (2001). Keller (1989); Chuang (1993); Palma and Chan (1997); Brooks (1997) Parikh and Williams (1998); Baharumshah and Liew (2001) can be given as examples of time series approaches. For the ARCH (autoregressive conditional heteroskedasticity) approach, Engle (1982); Bollerslev (1986); Hsieh (1989b) are the basic studies in literature. In order to test whether there has been an important non-linear relationship in recent years, many studies have been carried out with a new technique called as ANN. Verkooijen (1996); Plasmans et al. (1998); Zhang and Hu (1998); Franses and Homelen (1998); Hu et al. (1999); Gradojevic and Yang (2000) can be shown as examples for the studies using ANN.

The aim of this study is to obtain the accurate predictions of the TL/USD exchange rate series using the artificial neural network method which is an alternative approach to models in time series analysis. For this reason, in the following section ANN method is given briefly and in Section 3 the Box-Jenkins, ARCH methods in time series analysis and ANN method are applied to exchange rate series. We also compare the obtained results from these methods. The fourth section concludes this study.

2. Artificial Neural Networks

“What is an artificial neural network?” is the first question that should be answered. Picton (1994) answered this question by separating this question into two parts. The first part is why it is called as an artificial neural network. The reason of it can be explained that artificial neural network is a network of interconnected elements which are inspired from studies of biological nervous systems. In other words, artificial neural network is an attempt at creating machines that work in a similar way to the human brain by building these machines using components that behave like biological neurons. The second part is what an artificial neural network does. The function of an artificial neural network is to produce an output pattern using an input pattern.

In forecasting, artificial neural networks (ANN) are mathematical models that imitate biological neural networks. Researchers have used artificial neural networks methodology to forecast many nonlinear time series events in the literature (Ture and Kurt, 2006). The methodology consists of some elements. Determining the elements of the artificial neural network’s issue, which affect the forecasting performance of artificial neural networks, should be considered carefully. Elements of an artificial neural network are generally known as network architecture, learning algorithm, and activation function. Egrioglu et al. (2008) described these elements as given below.

One critical decision is to determine the appropriate architecture, that is, the number of layers, number of nodes in each layers, and the number of arcs which interconnect with the nodes (Zurada, 1992). Feedforward neural network (FNN) has been used in many studies for the forecasting process of the series. Determining architecture depends on the basic problem. Since, in the literature, there are not general rules for determining the best architecture, many architectures should be examined for the correct results. Fig. 1 depicts the broad FNN architecture that has single hidden layer and single output. Other important architectures include direct connections from input nodes to output nodes, as shown in Fig. 2.

Learning of ANN for a specific task is equivalent to finding the values all of the weights such that the desired output is generated by the corresponding input. Various training
algorithms have been used for the determining optimal weights values. The most popularly used training method is the back propagation algorithm, presented in Smith (2002). As Cichocki and Unbehauen (1993) mention in the back propagation algorithm, learning of the artificial neural network consists of adjusting all weights such as the error measure between the desired output and actual output.

![Fig. 1. A broad FNN architecture](image)

Another element of ANN is the activation function. It determines the relationship between inputs and outputs of a node and a network. In general, the activation function introduces a degree of the nonlinearity that is valuable for the most ANN applications. The well known activation functions are logistic, hyperbolic tangent, sine (or cosine) and the linear functions. Among them, logistic transfer function is the most popular one (Zhang et al., 1998).

![Fig. 2. A direct connected FNN architecture](image)
3. Forecasting exchange rate series by time series analysis and neural network

For the exchange rate series, we take the weekly rates of TL/USD series between the period January 3, 2005 and January 28, 2008. We obtain this data set having total 160 observations from Central Bank of Turkey. In this study, we examine the log differences of the levels of this series. In other words, we are interested in the series calculated by

\[ y_t = \log(r_t) - \log(r_{t-1}), \]

where \( r_t \) is the exchange rate series. Here the series, \( y_t \), can be defined as the change of the exchange rate series in the logarithmic transformation from the period to period. The graph of this transformed series is given in Fig. 3.

\[ \begin{array}{c}
\begin{array}{c}
\text{Fig. 3. The graph of the series}
\end{array}
\end{array} \]

3.1 Box-Jenkins and ARCH Models

Box and Jenkins (1976) introduced the ARIMA \((p,d,q)\) and the seasonal ARIMA \((p,d,q)(P,D,Q)_s\) models and many researchers have applied these models to forecast economic time series for years. To determine the convenient ARIMA models for the exchange rate series, we obtain ACF (autocorrelation function) and partial ACF graphs, as shown in Fig. 4.
From Fig. 4 and after many trials of ARIMA models we see that the suitable model is the moving average (MA) model. Therefore, from Fig. 4, it is clear that the parameter of MA, $q$, should be 1 and the parameter of AR (autoregressive), $p$, should be 0 (for details of Box-Jenkins methodology, see Gaynor and Kirkpatrick, 1994). Note that $d$ and $D$ should also be 0 as we do not take any difference and seasonal difference of the series. From further analysis of Fig. 4, we also get suspicious about the seasonality in the series as there is an important correlation at the lag 6 in the ACF and partial ACF graphs. Depending on these prior information and according to significance of the parameters in the model and the values of Akaike and Schwarz criteria, we find the best model as ARIMA (0,0,1) (0,0,1)$_6$ after many trials of seasonal ARIMA models. This model for the exchange rate series is obtained by

$$y_t = -0.001 - 0.219 e_{t-1} + 0.187 e_{t-6} + e_t,$$

where $e_t$ is the error series. When we analyze the error series, we see that the error series has no autocorrelation. We reach this result by using Box-Pierce Test, given in Table 1. We see that all Box-Pierce statistics at all lags are smaller than the related critical values of chi-square table. However, by using the ARCH-LM (Autoregressive Conditional Heteroskedasticity − Lagrange Multiplier) Test, whose results are given in Table 2, we observe that there is heteroskedasticity problem in the model. From Table 2, we observe that the values of ARCH-LM statistics are bigger than the related critical values of $F$ distribution table. It means that the null hypothesis of homoskedasticity is rejected.
Table 1
The results of Box-Pierce Test

<table>
<thead>
<tr>
<th>lag</th>
<th>Box-Pierce</th>
<th>lag</th>
<th>Box-Pierce</th>
<th>lag</th>
<th>Box-Pierce</th>
<th>lag</th>
<th>Box-Pierce</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0798</td>
<td>6</td>
<td>1.2245</td>
<td>11</td>
<td>6.3233</td>
<td>16</td>
<td>11.017</td>
</tr>
<tr>
<td>2</td>
<td>0.1694</td>
<td>7</td>
<td>5.5296</td>
<td>12</td>
<td>8.7587</td>
<td>17</td>
<td>11.918</td>
</tr>
<tr>
<td>3</td>
<td>0.5635</td>
<td>8</td>
<td>5.6532</td>
<td>13</td>
<td>10.184</td>
<td>18</td>
<td>12.083</td>
</tr>
<tr>
<td>4</td>
<td>0.6454</td>
<td>9</td>
<td>5.8042</td>
<td>14</td>
<td>10.219</td>
<td>19</td>
<td>12.473</td>
</tr>
<tr>
<td>5</td>
<td>1.2173</td>
<td>10</td>
<td>5.8104</td>
<td>15</td>
<td>11.004</td>
<td>20</td>
<td>13.547</td>
</tr>
</tbody>
</table>

Note: All values of Box-Pierce statistics do not exceed the critical value of chi-square distribution table for 5% significance level.

Table 2
The results of ARCH-LM Test for seasonal ARIMA model

<table>
<thead>
<tr>
<th>lag</th>
<th>ARCH-LM</th>
<th>lag</th>
<th>ARCH-LM</th>
<th>lag</th>
<th>ARCH-LM</th>
<th>lag</th>
<th>ARCH-LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.801*</td>
<td>6</td>
<td>3.634</td>
<td>11</td>
<td>2.857</td>
<td>16</td>
<td>2.019</td>
</tr>
<tr>
<td>2</td>
<td>9.139*</td>
<td>7</td>
<td>3.639</td>
<td>12</td>
<td>2.598</td>
<td>17</td>
<td>1.895</td>
</tr>
<tr>
<td>3</td>
<td>6.166*</td>
<td>8</td>
<td>3.393</td>
<td>13</td>
<td>2.486</td>
<td>18</td>
<td>1.779</td>
</tr>
<tr>
<td>4</td>
<td>4.570*</td>
<td>9</td>
<td>3.258</td>
<td>14</td>
<td>2.287</td>
<td>19</td>
<td>1.656</td>
</tr>
<tr>
<td>5</td>
<td>3.942*</td>
<td>10</td>
<td>3.012</td>
<td>15</td>
<td>2.101</td>
<td>20</td>
<td>1.596</td>
</tr>
</tbody>
</table>

Note: All values of ARCH-LM statistics exceed the critical value of F distribution table at 1% significance level (*), at 5% significance level (**), or at 5% significance level (***)

To overcome the heteroskedasticity problem in the determined seasonal ARIMA model, we decide to integrate an ARCH(1) model to this ARIMA model by,

\[
y_t = -0.001 - 0.021e_{t-1} + 0.086e_{t-6} + e_t, \\
e_t^2 = 0.00008 - 0.90081e_{t-1}^2 + u_t,
\]

where \(u_t\) is the error series of the ARCH model. We would like to remark that we have also tried the GARCH (1,1) model but the parameter of the GARCH term is found as insignificant. When we analyze the error series of the ARCH model, we find that there is no autocorrelation and heteroskedasticity problem in the model. The results of ARCH-LM Test are given in Table 3.

Table 3
The results of ARCH-LM Test for ARCH model

<table>
<thead>
<tr>
<th>lag</th>
<th>ARCH-LM</th>
<th>lag</th>
<th>ARCH-LM</th>
<th>lag</th>
<th>ARCH-LM</th>
<th>lag</th>
<th>ARCH-LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0696</td>
<td>6</td>
<td>0.2628</td>
<td>11</td>
<td>0.5252</td>
<td>16</td>
<td>0.4370</td>
</tr>
<tr>
<td>2</td>
<td>0.0318</td>
<td>7</td>
<td>0.6209</td>
<td>12</td>
<td>0.4754</td>
<td>17</td>
<td>0.4244</td>
</tr>
<tr>
<td>3</td>
<td>0.1445</td>
<td>8</td>
<td>0.5511</td>
<td>13</td>
<td>0.4385</td>
<td>18</td>
<td>0.4367</td>
</tr>
<tr>
<td>4</td>
<td>0.1216</td>
<td>9</td>
<td>0.5412</td>
<td>14</td>
<td>0.4004</td>
<td>19</td>
<td>0.5897</td>
</tr>
<tr>
<td>5</td>
<td>0.2543</td>
<td>10</td>
<td>0.5389</td>
<td>15</td>
<td>0.3778</td>
<td>20</td>
<td>0.5970</td>
</tr>
</tbody>
</table>

Note: All values of ARCH-LM statistics do not exceed the critical value of F distribution table at 5% significance level.
In the literature, the various performance criteria are used to gauge the forecasting accuracy in the time series analysis. In order to measure the performance of the seasonal ARIMA and ARCH models, the root mean square errors (RMSE), the mean absolute percentage error (MAPE) and the mean absolute error (MAE) of these models are computed by

\[
RMSE = \left( \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n} \right)^{1/2},
\]

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|,
\]

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|,
\]

respectively. Here, \(y_i\) is the actual value; \(\hat{y}_i\) is the predicted value, and \(n\) is the number of observations.

When we estimate the first 144 in-sample data of the exchange rate series and forecast the last 16 out-of-sample data by the seasonal ARIMA (0,0,1) \((0,0,1)_6\) and ARCH (1) models, the RMSE values of these models for in-sample data, shown by TRMSE, and for out-of-sample, shown by FRMSE are given in Table 4. Also MAPE and MAE values of these models for test data are shown in the table.

Table 4.
Performance criteria values of the models for in-sample and out-of-sample data

<table>
<thead>
<tr>
<th></th>
<th>TRMSE</th>
<th>FRMSE</th>
<th>MAPE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>0.01500</td>
<td>0.01533</td>
<td>1.23272</td>
<td>0.01148</td>
</tr>
<tr>
<td>ARCH</td>
<td>0.01551</td>
<td>0.01448</td>
<td>0.97445</td>
<td>0.01089</td>
</tr>
</tbody>
</table>

From Table 4, we observe that the better forecasts are obtained by the aid of ARCH model, whereas the estimation of the sample data by ARIMA model is slightly better than ARCH model for exchange rate series.

3.2 Artificial neural network

Based on the architecture design and the activation function, four different ANN models are used in order to obtain forecasts of the exchange rate series. These models are constructed by shifting the activation function used in the output neuron and by changing the architecture structure. Two different activation functions are used in the output neuron. The first one is the logistic activation function obtained by,

\[
f(y) = \left(1 + \exp(-0.8 \, y)\right)^{-1},
\]
where the slope parameter is -0.8 and the other used activation function is the linear activation function expressed by,

\[ f(y) = y. \]

While the ANN models were being constructed, two architecture structures were employed. Both the architecture structure which does not have direct connections between input and output neurons and the other structure which includes direct connections were used in the analysis. As mentioned, these architecture designs are illustrated in Fig. 1 and Fig. 2, respectively. Hence, four different ANN models are constructed based on these elements. Employed ANN models are given in Table 5.

**Table 5**

<table>
<thead>
<tr>
<th>Models</th>
<th>Direct Connections</th>
<th>Activation Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>Logistic</td>
</tr>
<tr>
<td>2</td>
<td>Yes</td>
<td>Logistic</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>Linear</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>Linear</td>
</tr>
</tbody>
</table>

To determine the best architectures for each of four ANN models, many numbers of neurons in the input and the hidden layers are examined. For each model, the number of neurons in input layer changes 1 through 14, the number of neurons in hidden layer alters 1 through 14 again, and by this way 14x14=196 architectures are examined. As there are 4 models, totally 4x196=784 different architectures are examined in this analysis. For each ANN model, the best architecture which has the best forecasting accuracy is chosen based on a performance criterion, such as RMSE.

The examined data are split into training and test data sets. First 144 observations are used for training and the last 14 values of the data are used for the purpose of test. Calculated RMSE values for the training data set and test data are given in the notations of TRMSE and FRMSE, respectively. The best architectures for models are determined based on FRMSE values. It means that the architecture, which has the lowest FRMSE value, is chosen as the best architecture among all examined ones.

Various learning algorithms have been used to train the network in the literature. The Back Propagation Algorithm is the most preferred one in forecasting studies. In this study, the Back Propagation Algorithm, in which the learning parameter is updated at each iteration, is used to find the best values of the weights.

By using these elements of ANN, the exchange rate series, \( y_t \), is examined by four ANN models. Beside the RMSE criterion, MAPE and MAE values of examined ANN models are also computed for the test data. The obtained results for these models are given in Table 6.
Table 6
Performance criteria values of the neural networks for training and test data

<table>
<thead>
<tr>
<th>Models</th>
<th>Best Architecture</th>
<th>TRMSE</th>
<th>FRMSE</th>
<th>MAPE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11-1-1</td>
<td>0.0147</td>
<td>0.0138</td>
<td>2.11</td>
<td>0.011</td>
</tr>
<tr>
<td>2</td>
<td>14-1-1</td>
<td>0.0133</td>
<td>0.0133</td>
<td>0.79</td>
<td>0.01</td>
</tr>
<tr>
<td>3</td>
<td>7-5-1</td>
<td>0.0152</td>
<td>0.0143</td>
<td>1.01</td>
<td>0.01</td>
</tr>
<tr>
<td>4</td>
<td>7-1-1</td>
<td>0.0154</td>
<td>0.0150</td>
<td>1.43</td>
<td>0.01</td>
</tr>
</tbody>
</table>

As the second model has the lowest FRMSE value 0.01329 with 14-1-1 architecture, the best forecasting accuracy is obtained from the second model in which logistic activation function is used in output neuron and direct connections between input and output neurons exist. Furthermore, the second ANN model gives the best forecasts in terms of all computed performance criteria. In other words, when the second model is used, 14-1-1 architecture, which has 14 neurons in input layer, one neuron in hidden layer, and one neuron in output layer, produces the best forecasts.

4. Conclusion

In this study, we use neural network as an alternative forecasting method to seasonal ARIMA and ARCH models in the prediction of weekly Turkish Lira / US dollar exchange rate. From Tables 4 and 6, it is obviously seen that neural network has superior both in-sample and out-of-sample forecast than seasonal ARIMA and ARCH models by using the RMSE measure. It is certainly inferred that ANN performs better than other models in terms of accurate predictions for Turkish Lira / US dollar exchange rate volatility.

As Panda and Narasimhan (2007) mention, it is no doubt that the better forecasting by using artificial neural network may help the policy makers to conduct a suitable monetary policy which will in turn achieve its desired objectives and higher economic activity. This may also help the policy makers in extracting useful information about the economic and financial conditions.

From further analysis of Table 6, we can deduce that using logistic activation function in output neuron produces more accurate forecasts than those obtained from linear activation function. Models 1 and 2, in which logistic activation function is employed, generate lower FRMSE values than those from Models 3 and 4 in which linear activation function is used. For the Turkish Lira / US dollar exchange rate series, we recommend that the logistic activation function in output neuron should be preferred. In addition, according to Table 6, when logistic activation function is employed, using more input neurons produces accurate results. Consequently, it is clearly seen that the selection of activation function in the output neuron is more effective on the forecasting performance of ANN. This study shows that logistic activation function is determined as logistic function for the Turkish Lira / US dollar exchange rate series by conducting elaborate analysis with ANN and the best forecasts are obtained based on this decision.
Table 7.
ANN and conventional methods’ results

<table>
<thead>
<tr>
<th></th>
<th>FRMSE</th>
<th>FMAPE</th>
<th>FMAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>0.01533</td>
<td>1.23272</td>
<td>0.01148</td>
</tr>
<tr>
<td>ARCH</td>
<td>0.01448</td>
<td>0.97445</td>
<td>0.01089</td>
</tr>
<tr>
<td>ANN</td>
<td>0.01329</td>
<td>0.78822</td>
<td>0.00968</td>
</tr>
</tbody>
</table>

For the comparison, we summarize all obtained results from ANN method and from conventional methods ARIMA, GARCH in Table 7. According to Table 7, it is clearly seen that the ANN method produces the most accurate forecasts in terms of RMSE, MAPE and MAE performance criteria. In the literature, when ANN methods are used to forecast time series, the combinations of components for the method are not generally tried in detail. However, different components should be applied to data since ANN is a method which depends on the data and because of this, different results can be obtained for different time series. In this study, we used various ANN models to analyze the Turkish Lira / US dollar exchange rate series by utilizing different components of ANN. It is seen that using different ANN models, including different components, leads minimum error in forecasts. Furthermore, it is shown that the best ANN model gives more accurate forecasts than those obtained from ARIMA and ARCH models.
References


