Overview and Comparison of Short-term Interval Models for Financial Time Series Forecasting

M. Khashei, F. Mokhatab Rafiei & M. Bijari

Mehdi Khashei, Ph.D student of Industrial Engineering, Isfahan University of Technology Isfahan, Iran
Farimah Mokhatab Rafiei, Assistant professor of Industrial Engineering, Isfahan University of Technology Isfahan, Iran
Mehdi Bijari, Associated professor of Industrial Engineering, Isfahan University of Technology Isfahan, Iran

KEYWORDS

Artificial Neural Networks (ANNs), Auto-Regressive Integrated Moving Average (ARIMA), Time series forecasting, Hybrid forecasts, Interval models, Exchange rate

ABSTRACT

In recent years, various time series models have been proposed for financial markets forecasting. In each case, the accuracy of time series forecasting models are fundamental to make decision and hence the research for improving the effectiveness of forecasting models have been carried on. Many researchers have compared different time series models together in order to determine more efficient one in financial markets. In this paper, the performance of four interval time series models including autoregressive integrated moving average (ARIMA), fuzzy autoregressive integrated moving average (FARIMA), hybrid ANNs and fuzzy (FANN) and Improved FARIMA models are compared together. Empirical results of exchange rate forecasting indicate that the FANN model is more satisfactory than other those models. Therefore, it can be a suitable alternative model for interval forecasting of financial time series.

1. Introduction

Exchange rate is one of the most effective variables in financial environments and its changes can be very important for economic decision makers [1]. Several investigations have been accomplished in the field of exchange rate forecasting [2-5] that number of these investigations represents the mentioned issue importance. Nowadays, despite the numerous financial time series models available, accurate forecasts of exchange rate are not easy task [6-11]. Several different models have been suggested for time series forecasting, which are generally categorized in to linear and nonlinear models. One of the most important and widely used linear time series models are autoregressive integrated moving average (ARIMA) models that have enjoyed fruitful applications in forecasting social, economic, engineering, foreign exchange, and stock problems. Second class of time series models are nonlinear models. Artificial neural networks are one of these models that are able to approximate various nonlinearities in the data and are flexible computing frameworks for modelling a broad range of nonlinear problems.

One significant advantage of ANNs over than other nonlinear classes is that ANNs are universal approximators which can approximate a large class of functions with a high degree of accuracy. No prior assumption of the model form is required in the model building process. Instead, the network model is largely determined by the characteristics of the data. Commonly used neural networks include multi-layer perceptrons (MLPs), radial basis functions (RBFs), probabilistic neural networks (PNNs), and general regression neural networks (GRNNs) [12]. Single hidden layer feedforward network is the most widely used model form for time series forecasting [13]. Forecasting accuracy is one of the most important factors to choose the forecasting model; therefore,
several researchers have compared different time series models together in order to determine more accurate once. Ture compared the performance of four different time series models to forecast the hepatitis A virus infection [14]. Taylor et al. compared the univariate models for forecasting electricity demand [15]. Kima [16] forecasted the international tourist flows to Australia for comparison between the direct and indirect models. Cho also compared the three different approaches to tourist arrival forecasting [17]. Some other researches in this field are as follows: Weatherforda [18] to hotel revenue management forecasting, Smith [19] to traffic flow forecasting, and Sfetsos [20] to mean hourly wind speed time series forecasting.

As similar, various researches have been also done in the financial fields. Alon compared the performance of artificial neural networks and traditional models to aggregate retail sales forecasting [21]. Meade [22] compared the accuracy of short term foreign exchange forecasting models. Leunga et al. [23] compared the classification and level estimation models to forecasting the stock indices. Lisi also compared the neural networks and chaotic models for exchange rate prediction [24]. In this paper, the performance of four different interval time series models is compared for financial markets forecasting. The rest of the paper is organized as follows. In the next section, concepts of four used time-series models are briefly reviewed. Empirical results from forecasting the exchange rate (US dollar/ Iran rial) are reported in Section 3. The performance of each model is compared together in section 4, and finally the conclusions are discussed.

2. Time Series Forecasting Models

There are several different approaches for time series modelling. Interval models are a special class of the quantitative forecasting models. In interval models, an interval is calculated as optimum forecast of independent variable. In this section, four interval models are briefly reviewed.

2-1. The Auto-Regressive Integrated Moving Average Model

In an autoregressive integrated moving average model, the future value of a variable is assumed to be a linear function of several past observations and random errors. That is, the underlying process, generating the time series has the following form:

\[
y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \ldots - \theta_q \epsilon_{t-q}
\]

(1)

where \( y_t \) and \( \epsilon_t \) are the actual value and random error at time period \( t \), respectively; \( \phi_i (i = 1, 2, \ldots, p) \) and \( \theta_j (j = 1, 2, \ldots, q) \) are model parameters. \( p \) and \( q \) are integers and often referred to as orders of the model.

Random errors, \( \epsilon_t \), are assumed to be independently and identically distributed with a mean of zero and a constant variance of \( \sigma^2 \).

The Box-Jenkins [25] methodology includes three iterative steps of model identification, parameter estimation and diagnostic checking. The basic idea of model identification is that if a time series is generated from an ARIMA process, it should have some theoretical autocorrelation properties. By matching the empirical autocorrelation patterns with the theoretical ones, it is often possible to identify one or several potential models for the given time series. Box and Jenkins [25] proposed to use the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the sample data as the basic tools to identify the order of the ARIMA model.

Once a tentative model is specified, estimation of the model parameters is straightforward. The parameters are estimated such that an overall measure of errors is minimized. This can be done with a nonlinear optimization procedure. The last step of model building is the diagnostic checking of model adequacy. This is basically to check if the model assumptions about the errors, \( \epsilon_t \), are satisfied.

Several diagnostic statistics and plots of the residuals can be used to examine the goodness of fit of the tentatively entertained model to the historical data. If the model is not adequate, a new tentative model should be identified, which is again followed by the steps of parameter estimation and model verification.

Diagnostic information may help suggest alternative model(s). This three-step model building process is typically repeated several times until a satisfactory model is finally selected.

2-2. The Fuzzy Auto-Regressive Integrated Moving Average

The parameter of ARIMA(p,d,q), \( \phi_1, \phi_2, \ldots, \phi_p \) and \( \theta_1, \theta_2, \ldots, \theta_q \) are crisp. Instead of using crisp, fuzzy parameters, \( \tilde{\phi}_1, \tilde{\phi}_2, \ldots, \tilde{\phi}_p \) and \( \tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_q \), in the form of triangular fuzzy numbers are used in fuzzy autoregressive integrated moving average models [26]. A fuzzy ARIMA(p,d,q) model is described by a fuzzy function with a fuzzy parameter:

\[
\tilde{\Phi}_p (B) W_t = \tilde{\theta}_q (B) \epsilon_t
\]

(2)

\[
W_t = (1 - B)^d (Z_t - \mu)
\]

(3)

\[
W_t = \tilde{\phi}_p W_{t-1} + \tilde{\phi}_2 W_{t-2} + \ldots + \tilde{\phi}_p W_{t-p} + \epsilon_t - \tilde{\theta}_q \epsilon_{t-1} - \tilde{\theta}_2 \epsilon_{t-2} - \ldots - \tilde{\theta}_q \epsilon_{t-q}
\]

(4)

where \( \{Z_t\} \) are observations, \( \tilde{\phi}_1, \tilde{\phi}_2, \ldots, \tilde{\phi}_p \) and \( \tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_q \), are fuzzy numbers. Equation (4) is modified as:
\[ \tilde{W}_t = \tilde{\beta}_1 W_{t-1} + \tilde{\beta}_2 W_{t-2} + \ldots + \tilde{\beta}_p W_{t-p} + a_t \]
\[- \tilde{\beta}_{p+1} a_{t-1} - \tilde{\beta}_{p+2} a_{t-2} - \ldots - \tilde{\beta}_{p+q} a_{t-q} \quad (5) \]

Fuzzy parameters in form of triangular fuzzy numbers are used:

\[ \mu_{\tilde{\beta}}(\beta_i) = \begin{cases} \frac{|a_i - c_i|}{c_i} & \text{if } a_i - c_i \leq \beta_i \leq a_i + c_i, \\ 0 & \text{otherwise}, \end{cases} \quad (6) \]

where \( \mu_{\tilde{\beta}}(\beta_i) \) is the membership function of the fuzzy set that represents parameter \( \beta_i \). \( a_i \) is the centre of the fuzzy number, and \( c_i \) is the width or spread around the centre of the fuzzy number. Using fuzzy parameters \( \beta_i \) in the form of triangular fuzzy numbers and applying the extension principle, it becomes clear [27] that the membership of \( W \) in (5) is given as (7).

\[ \mu_w(W_i) = \begin{cases} 1 - \sum_{i=1}^{p} \alpha_i W_{t-i} - a_i + \sum_{i=p+1}^{p+q} \alpha_i a_{t+p-i} \\ \sum_{i=1}^{p} c_i |W_{t-i}| + \sum_{i=p+1}^{p+q} c_i |a_{t+p-i}| \end{cases} \]

for \( W_i \neq 0, \quad a_i \neq 0 \)

\[ \mu_w(W_i) = 0 \quad \text{otherwise} \]

Simultaneously, \( Z_t \) represents the \( t \)th observation, and \( h \)-level is the threshold value representing the degree to which the model should be satisfied by all the data points \( y_1, y_2, \ldots, y_k \) to a certain \( h \)-level.

\[ \mu_y(y_t) \geq h \quad \text{for } t = 1, 2, \ldots, k \quad (8) \]

The index \( t \) refers to the number of nonfuzzy data used for constructing the model. On the other hand, the fuzziness \( S \) included in the model is defined by:

\[ S = \sum_{i=1}^{p} \sum_{i=1}^{k} c_i \| W_{t-i} \| + \sum_{i=p+1}^{p+q} \sum_{i=1}^{k} c_i \| a_{t+p-i} \| \quad (9) \]

Where \( \rho_{t-p} \) is the autocorrelation coefficient of time lag \( t-p \), \( \varphi_i \) is the partial autocorrelation coefficient of time lag \( i \). The weight of \( c_i \) depends on the relation of time lag \( i \) and the present observation, where the \( p \) of AR (p) is derived by PACF and the \( q \) of MA (q) is derived by ACF.

Next, the problem of finding the fuzzy ARIMA parameters was formulated as a linear programming problem as (10). At last, according to the Ishibuchi and Tanaka [28] opinion, the data around the model’s upper bound and lower bound is deleted when the fuzzy ARIMA model has outliers with wide spread, and then reformulating the fuzzy regression model.

\[ \text{Minimize } S = \sum_{i=1}^{p} \sum_{i=1}^{k} c_i \| W_{t-i} \| + \sum_{i=p+1}^{p+q} \sum_{i=1}^{k} c_i \| a_{t+p-i} \| \]

\[ \sum_{i=1}^{p} \alpha_i W_{t-i} + a_t - \sum_{i=1}^{p+q} \alpha_i a_{t+p-i} + (1 + h) \left( \sum_{i=1}^{p} c_i |W_{t-i}| + \sum_{i=p+1}^{p+q} c_i |a_{t+p-i}| \right) \geq W_t, \quad t = 1, 2, \ldots, k \quad (10) \]

subject to

\[ \sum_{i=1}^{p} \alpha_i W_{t-i} + a_t - \sum_{i=1}^{p+q} \alpha_i a_{t+p-i} + (1 + h) \left( \sum_{i=1}^{p} c_i |W_{t-i}| + \sum_{i=p+1}^{p+q} c_i |a_{t+p-i}| \right) \leq W_t, \quad t = 1, 2, \ldots, k \]

\[ c_i \geq 0 \quad \text{for } i = 1, 2, \ldots, p + q \]

2-3. The Improved FARIMA Model with Probabilistic Neural Networks (PNNs)

Forecasting interval of the fuzzy autoregressive integrated moving average models is extended in some specific data conditions. According to the Ishibuchi and Tanaka opinion, forecasting interval can be too wide, when training data set includes the significant difference or outlying case.

In improved model, the abilities of the probabilistic neural networks (PNNs) [29] is used in order to recognize more probability spaces in forecasting interval of FARIMA model. Technically, PNN is a classifier and is able to deduce the class/group of a given input vector after the training process is completed. PNN is conceptually built on the Bayesian model of classification which, given enough data, is capable of classifying a sample with the maximum probability of success [30]. The procedure of improved model is as follows:

Phase I: Fitting the FARIMA model using the available observations. The result of phase I is:

\[ \tilde{W}_t = (\alpha_1, c_1) W_{t-1} + \ldots + (\alpha_p, c_p) W_{t-p} + a_t \]
\[- (\alpha_{p+1}, c_{p+1}) a_{t-1} - \ldots - (\alpha_{p+q}, c_{p+q}) a_{t-q}. \quad (11) \]
where \( W_i = (I-B)^p (Z_i - \mu) \). \( \alpha_i \) is the centre of the fuzzy number, and \( c_i \) is the width or spread around the centre of the fuzzy number. Then, the obtained interval of FARIMA model is divided to \( n \) equal sections for using in probabilistic neural network. The subinterval which includes the real value or \( n-1 \) other subintervals are considered as target data. Other information -results of FARIMA and time series data- is considered as training data.

**Phase II:** Designing and training a network to recognize more probability spaces in forecasted interval of FARIMA model. The result of this phase is a interval with \( 1/n \) width and confidence coefficient \( \alpha \). (\( \alpha \) is the performance of PNN in the test data).

### 2.4. The Hybrid Artificial Neural Networks and Fuzzy Logic

A hybrid FANN model can be described by a fuzzy function as follows [31]:

\[
\tilde{y}_t = f(\tilde{b}_0 + \sum_{j=1}^{q} \tilde{w}_{j} \cdot g(\tilde{b}_0 + \sum_{i=0}^{p} \tilde{w}_{i,j} \cdot y_{t-i})) \tag{12}
\]

where \( y_t \) are observations, \( \tilde{w}_{j}, \tilde{w}_{i,j}, \tilde{b}_0, \tilde{b}_0 \) are fuzzy numbers. Equation (12) is modified as:

\[
\tilde{y}_t = f(\tilde{b}_0 + \sum_{j=1}^{q} \tilde{w}_{j} \cdot \tilde{X}_{t,j}) = f(\sum_{j=0}^{q} \tilde{w}_{j} \cdot \tilde{X}_{t,j}) \tag{13}
\]

where \( \tilde{X}_{t,j} = g(\tilde{b}_0 + \sum_{i=0}^{p} \tilde{w}_{i,j} \cdot y_{t-i}) \). Fuzzy parameters in the form of triangular fuzzy numbers are used as (14),

\[
\mu_{\tilde{w}_{i,j}}(w_{i,j}) = \begin{cases} 
\frac{1}{b_{i,j} - a_{i,j}} (w_{i,j} - a_{i,j}) & \text{if } a_{i,j} \leq w_{i,j} \leq b_{i,j}, \\
\frac{1}{c_{i,j} - b_{i,j}} (w_{i,j} - b_{i,j}) & \text{if } b_{i,j} \leq w_{i,j} \leq c_{i,j}, \\
0 & \text{otherwise},
\end{cases} \tag{14}
\]

where \( \mu_{\tilde{w}_{i,j}}(w_{i,j}) \) is the membership function of the fuzzy set that represents parameter \( w_{i,j} \). Applying the extension principle, it becomes clear that the membership of \( \tilde{X}_{t,j} = g(\sum_{i=0}^{p} \tilde{w}_{i,j} \cdot y_{t-i}) \) in (13) is given as (15) [31].

Now, considering the fuzzy parameters \( w_{j} \) as (16)

\[
\mu_{\tilde{w}_{i,j}}(w_{j}) = \begin{cases} 
\frac{1}{e_{j} - d_{j}} (w_{j} - d_{j}) & \text{if } d_{j} \leq w_{j} \leq e_{j}, \\
\frac{1}{e_{j} - f_{j}} (w_{j} - f_{j}) & \text{if } e_{j} \leq w_{j} \leq f_{j}, \\
0 & \text{otherwise},
\end{cases} \tag{16}
\]

For the membership function of the three fuzzy sets \( \tilde{y}_{t} = f(\tilde{b}_0 + \sum_{j=1}^{q} \tilde{w}_{j} \cdot \tilde{X}_{t,j}) \) is given as (17),

\[
\mu_{\tilde{y}_{t}}(y_{t}) = \begin{cases} 
\frac{-B_j + \left[ B_1 \left( \frac{B_2}{2A_1} \right)^2 - C_j - \frac{f^{-1}(y_{t})}{A_j} \right]^{1/2}}{2A_1} & \text{if } C_1 \leq f^{-1}(y_{t}) \leq C_3, \\
\frac{B_2}{2A_2} + \left[ B_1 \left( \frac{B_2}{2A_2} \right)^2 - C_2 - \frac{f^{-1}(y_{t})}{A_2} \right]^{1/2} & \text{if } C_3 \leq f^{-1}(y_{t}) \leq C_2, \\
0 & \text{otherwise},
\end{cases} \tag{17}
\]

Where

\[
A_j = \sum_{j=0}^{n} \left( e_{j} - d_{j} \right) g \left( \sum_{i=0}^{p} b_{i,j}y_{t-i} \right) - g \left( \sum_{i=0}^{p} a_{i,j}y_{t-i} \right).
\]
3-1. Autoregressive Integrated Moving Average Model

Using the Eviey package software, the best-fitted model is ARIMA(2,1,0). The actual values and 95% confidence interval of the ARIMA model are given in Table 1.

<table>
<thead>
<tr>
<th>Date</th>
<th>Actual value</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>10- Dec</td>
<td>9082</td>
<td>9074</td>
<td>9090</td>
</tr>
<tr>
<td>11- Dec</td>
<td>9083</td>
<td>9075</td>
<td>9091</td>
</tr>
<tr>
<td>12- Dec</td>
<td>9083</td>
<td>9075</td>
<td>9090</td>
</tr>
<tr>
<td>13- Dec</td>
<td>9082</td>
<td>9074</td>
<td>9090</td>
</tr>
<tr>
<td>14- Dec</td>
<td>9081</td>
<td>9073</td>
<td>9089</td>
</tr>
<tr>
<td>15- Dec</td>
<td>9082</td>
<td>9074</td>
<td>9090</td>
</tr>
<tr>
<td>16- Dec</td>
<td>9082</td>
<td>9074</td>
<td>9090</td>
</tr>
</tbody>
</table>

3-2. Fuzzy Autoregressive Integrated Moving Average Model

Setting \((a_0, a_1, a_2) = (9060.05, 0.607, 0.421)\), the fuzzy parameters are obtained by (10) (with \(h=0\)). The results after deleting the outlier data are given in Table 2.

<table>
<thead>
<tr>
<th>Date</th>
<th>Actual value</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>10- Dec</td>
<td>9082</td>
<td>9081</td>
<td>9085</td>
</tr>
<tr>
<td>11- Dec</td>
<td>9083</td>
<td>9080</td>
<td>9084</td>
</tr>
<tr>
<td>12- Dec</td>
<td>9083</td>
<td>9081</td>
<td>9085</td>
</tr>
<tr>
<td>13- Dec</td>
<td>9082</td>
<td>9081</td>
<td>9085</td>
</tr>
<tr>
<td>14- Dec</td>
<td>9081</td>
<td>9080</td>
<td>9084</td>
</tr>
<tr>
<td>15- Dec</td>
<td>9082</td>
<td>9079</td>
<td>9083</td>
</tr>
<tr>
<td>16- Dec</td>
<td>9082</td>
<td>9080</td>
<td>9084</td>
</tr>
</tbody>
</table>

3-3. The Improved FARIMA Model with PNNs

In improved model, the probabilistic neural network is used after the FARIMA model. The best fitted network is a network with five input neurons and one output neuron. The structure of designed network is given in Fig. 2.

![Fig. 1. The structure of designed network](image-url)
where
Var1: Forecasting lower bond of time series in time \( t \) \((L_t)\)

Var2: Forecasting upper bond of time series in time \( t \) \((U_t)\)

Var3: Forecasting value of time series in time \( t \) \((Z_t)\)

Var4: Difference between forecasting value of time series in time \( t \) \& time \( t-1 \) \((Z_t-Z_{t-1})\)

Var5: Difference between forecasting upper bond (lower bond) of time series in time \( t \) \& time \( t-1 \) \((U_t-U_{t-1})\)

Obtained results of upper and lower bound with improved model with 100% confidence coefficient are given in Table 3.

<table>
<thead>
<tr>
<th>Date</th>
<th>Actual value</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>10- Dec</td>
<td>9082</td>
<td>9081</td>
<td>9084</td>
</tr>
<tr>
<td>11- Dec</td>
<td>9083</td>
<td>9081</td>
<td>9084</td>
</tr>
<tr>
<td>12- Dec</td>
<td>9083</td>
<td>9081</td>
<td>9084</td>
</tr>
<tr>
<td>13- Dec</td>
<td>9082</td>
<td>9081</td>
<td>9084</td>
</tr>
<tr>
<td>14- Dec</td>
<td>9081</td>
<td>9080</td>
<td>9083</td>
</tr>
<tr>
<td>15- Dec</td>
<td>9082</td>
<td>9079</td>
<td>9082</td>
</tr>
<tr>
<td>16- Dec</td>
<td>9083</td>
<td>9081</td>
<td>9084</td>
</tr>
</tbody>
</table>

4. Comparison the Performance of Models

In this section, based on the empirical results of this example, the predictive capabilities of the aforementioned models are compared together. The information of forecasted interval width and related performance of each model is given in Table 5.

<table>
<thead>
<tr>
<th>Model</th>
<th>Interval width</th>
<th>ARIMA</th>
<th>FARIMA</th>
<th>PNN/FARIMA</th>
<th>FANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>16.2</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FARIMA</td>
<td>4.2</td>
<td>74.1%</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PNN/FARIMA</td>
<td>3.1</td>
<td>80.9%</td>
<td>26.2%</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>FANN</td>
<td>2.5</td>
<td>84.6%</td>
<td>40.5%</td>
<td>19.4%</td>
<td>0</td>
</tr>
</tbody>
</table>

According to the above results, the autoregressive integrated moving average model has the lowest performance and the hybrid artificial neural networks and fuzzy logic model has the better performance than other models in exchange rate forecasting.

5. Conclusions

The foreign exchange markets are among the most important and the largest financial markets in the world with trading taking place twenty-four hours a day around the globe and trillions of dollars of different currencies transacted each day. Being able to accurately forecast the movements of exchange rates can result in considerable improvement in the overall profitability of the multinational financial firm, especially for firms, conducting substantial currency transfers in the course of business. However, predicting currency movements has always been a problematic task for academic researchers and despite the paramount modelling effort registered in the last three decades, it is widely recognized that exchange rates are extremely difficult to forecast. That is the reason why research on improving the effectiveness of time series models has been never witnessed a halt.

In this paper the performance of four different interval time series models (Auto-Regressive Integrated Moving Average (ARIMA), Fuzzy Auto-Regressive Integrated Moving Average (FARIMA), Hybrid ANNs and Fuzzy, Improved FARIMA) are compared together in exchange rate forecasting. Empirical results of exchange rate forecasting indicate that the hybrid ANNs and fuzzy model is more satisfactory than other those models. Therefore, it can be used as a suitable alternative model for interval forecasting in financial markets.

References


