International Journal of Computational Intelligence Systems, Vol. 6, No. 5 (September, 2013), 954-968

# Hybrid Fuzzy Auto-Regressive Integrated Moving Average (FARIMAH) Model for Forecasting the Foreign Exchange Markets

**Mehdi Khashei***
*Department of Industrial Engineering, Isfahan University of Technology*
*Isfahan 84156-83111, Iran*

**Farimah Mokhatab Rafiei, Mehdi Bijari**
*Department of Industrial Engineering, Isfahan University of Technology*
*Isfahan 84156-83111, Iran*
*E-mail: Farimah@cc.iut.ac.ir , Bijari@cc.iut.ac.ir*
*www.iut.ac.ir*

Received 7 April 2012

Accepted 18 February 2013

**Abstract**

Improving forecasting especially time series forecasting accuracy is an important yet often difficult task facing forecasters. Fuzzy autoregressive integrated moving average (FARIMA) models are the fuzzy improved version of the autoregressive integrated moving average (ARIMA) models, proposed in order to overcome limitations of the traditional ARIMA models; especially data limitation, and yield more accurate results. However, the forecasted interval of the FARIMA models may be very wide in some specific Circumstances. For instance, when data has high volatility or includes a significant difference or outliers. In this paper, a new hybrid model of FARIMA models is proposed by combining with probabilistic neural classifiers, called FARIMAH, in order to yield a more general and more accurate model than FARIMA models for financial forecasting in incomplete data situations. The main idea of the proposed model is based on this fact that the distribution of the actual values in the forecasted interval by FARIMA is not uniform. Thus, by detecting the spaces with more probability for actual values using the probabilistic classifier, narrower interval than traditional FARIMA models can be obtained. Empirical results of exchange rate markets forecasting indicate that the proposed model exhibit effectively improved forecasting accuracy, so it can be used as an alternative model to exchange rate forecasting, especially when the scant data made available over a short span of time.

*Keywords*: Fuzzy autoregressive integrated moving average (FARIMA); probabilistic neural classifiers; Time series forecasting; Foreign exchange markets; Fuzzy hybrid models.

## 1. Introduction

Time series forecasting is an important area of forecasting in which past observations of the same variable are analyzed in order to develop a model describing the underlying relationship. The model is then used to extrapolate the time series into the future. This modeling approach is particularly useful when little knowledge is available on the underlying data generating process or when there is no satisfactory explanatory model that relates the dependent variable to other explanatory variables. Several different

*Corresponding author.
*Tel.: +98-311-3912550-1; Fax: +98-311-3915526*
*E-mail address: Khashei@in.iut.ac.ir (M. Khashei).*

approaches have been proposed to time series forecasting. One of the most popular and widely used time series models are autoregressive integrated moving average (ARIMA) models that have enjoyed fruitful applications in forecasting problems. The popularity of the ARIMA model is due to its statistical properties as well as the well-known Box–Jenkins methodology [1] in the model building process. In addition, ARIMA models can implement various exponential smoothing models. Although ARIMA models have the advantage of accurate forecasting in a short time period and easy to implement, these models have some limitations that detract from their popularity for financial time series forecasting, such as data limitation.

The autoregressive integrated moving average models need the large amount of historical data (at least 50 and preferably 100 observations or more) in order to yield desired results [2]. However, in our society today, due to factors of uncertainty from the integral environment and rapid development of new technology, we usually have to forecast future situations using little data in a short span of time. The historical data must be less than what the ARIMA model requires which limits its application. The fuzzy regression is an interval-forecasting model that suitable for the condition of little attainable historical data. However, the performance of the fuzzy regression models is not always satisfactory. In addition, these models do not include the concepts of the Box-Jenkins models for time series forecasting.

In order to fulfill the limitations of the fuzzy regression and the autoregressive integrated moving average models and also to yield more accurate results, the fuzzy autoregressive integrated moving average (FARIMA) is proposed by Tseng *et al.* [3]. This model is formulated based on the basic concepts of the ARIMA model and Tanaka fuzzy regression that combine the advantages of the fuzzy regression and ARIMA models. In FARIMA models, instead of using crisp parameters, fuzzy parameters, in the form of triangular fuzzy numbers are used. By using the fuzzy parameters, the requirement of historical data would be reduced. Their results of foreign exchange markets forecasting indicate that the FARIMA model not only can make good forecasts but also provides the decision makers with the best and worst possible situations [3]. However, the forecasting interval of the FARIMA models may also be very wide if data includes a significant difference or outliers or when data has the high volatility; and hence, it is not wise to apply them blindly to any type of data.

Using hybrid models or combining several models has become a common practice in order to overcome the limitations of components models and improve the forecasting accuracy. Many researches in time series forecasting have been argued that predictive performance improves in combined models. Typically, this is done because the underlying process cannot easily be determined. The motivation for using hybrid models comes from the assumption that either one cannot identify the true data generating process or that a single model may not be sufficient to identify all the characteristics of the time series. In pioneering work on combined forecasts, Bates and Granger showed that a linear combination of forecasts would give a smaller error variance than any of the individual methods. Since then, the studies on this topic have expanded dramatically [4].

In recent years, more hybrid forecasting models have been developed, integrating autoregressive integrated moving average (ARIMA), artificial neural networks (ANNs), and fuzzy models together in order to improve the prediction accuracy and overcome the deficiencies of the single models. Andres *et al.* [5] proposed a strategy for constructing a hybrid model, which combines the fuzzy clustering and the multivariate adaptive regression splines (MARS) in order to use their theoretical advantages of these models for bankruptcy forecasting, especially when the information applied for forecasting is drawn from company financial statements. Lin and Cobourn [6] combined Takagi–Sugeno fuzzy system and a nonlinear regression (NLR) model for daily ground-level ozone predictions. Chang *et al.* [7] developed a hybrid model by integrating fuzzy rule base (FRB), self-organization maps (SOMs), and Genetic Algorithms (GAs) to forecast the future sales of a printed circuit board factory. Teoh *et al.* [8] proposed a hybrid model based on multi-order fuzzy time series, which employs rough sets theory in order to mine fuzzy logical relationship from time series and an adaptive expectation model to adjust forecasting results, to improve forecasting accuracy.

Kim and Shin [9] investigated the effectiveness of a hybrid approach based on the artificial neural networks for time series properties, such as the adaptive time delay neural networks (ATNNs) and the time delay neural networks (TDNNs), with the genetic algorithms in detecting temporal patterns for stock market prediction tasks. Khashei *et al.* [10] based on the basic concepts of multilayer perceptrons (MLPs), proposed a new hybrid model for financial time series forecasting using fuzzy regression models in order to overcome the data limitation of the multilayer perceptrons and yield more accurate results, especially in incomplete data situations. Li and Su [11] introduced a hybrid model, integrating genetic algorithm and hierarchical adaptive network-based fuzzy inference system (HANFIS) in which GA optimizes the structure and number of fuzzy if-then rules in a hierarchical ANFIS by finding the best parameter values of a subtractive clustering method.

Pai [12] proposed the hybrid ellipsoidal fuzzy system (HEFST) model to forecast regional electricity loads in Taiwan. Azadeh *et al.* [13] presented a hybrid algorithm based on fuzzy linear regression (FLR) and fuzzy cognitive map (FCM) to deal with the problem of forecasting and optimization of housing market fluctuations. Huang *et al.* [14] presented a new forecasting model based on two computational methods, fuzzy time series and particle swarm optimization for academic enrollments. Yadav and Srinivasan [15] introduced a hybrid method using smooth transition autoregressive (STAR), feed-forward neural network (FFNN), and self-organizing map (SOM) for short-term load prediction. Yu *et al.* [16] proposed a novel nonlinear ensemble forecasting model integrating generalized linear auto regression (GLAR) with back-propagation neural network (BPNN) in order to obtain accurate prediction in foreign exchange market. Khashei and Bijari [17] proposed a novel hybrid model of artificial neural networks (ANNs), based on the basic concepts of the Box-Jenkins methodology for autoregressive integrated moving average (ARIMA (p,d,q)) models, called ANN(p,d,q) model, in order to overcome the linear limitation of traditional neural networks and yield more accurate results.

Amin-Naseri and Soroush [18] presented a hybrid model of feed forward neural networks for daily electrical peak load forecasting using self-organizing maps (SOMs). Lin and Wu [19], in similar work, proposed a hybrid neural network model to forecast the typhoon rainfall using the self-organizing maps and the multilayer perceptrons. Ismail *et al.* [20] proposed a hybrid artificial intelligence model combining the least square support vector machine (LSSVM) and self-organizing maps for time series forecasting. Khashei and Bijari [21] introduced a new class of hybrid models by combining time series models such as autoregressive moving average (ARMA) and feed forward neural networks (FFNNs) and probabilistic neural networks (PNNs) for time series forecasting. Hajizadeh *et al.* [22] proposed a hybrid models based on Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and Artificial Neural Networks to forecast the volatility of S&P 500 index. In this model, the estimates of volatility obtained by a GARCH model are fed forward to a neural network.

Ince and Trafalis [23] proposed a two-stage hybrid model which incorporates parametric techniques such as autoregressive integrated moving average, vector autoregressive (VAR) and co-integration techniques, and nonparametric techniques such as support vector regression (SVR) and artificial neural networks for exchange rate prediction. Pham *et al.* [24] presented an improvement of hybrid of nonlinear autoregressive with exogenous input (NARX) model and autoregressive moving average model for long-term machine state forecasting based on vibration data. Shafie-khah *et al.* [25], based on wavelet transform, autoregressive integrated moving average, and radial basis function neural networks (RBFN), proposed a novel hybrid model to forecast electricity price.

In this paper, the probabilistic neural classifiers are applied in order to construct an improved model of the fuzzy autoregressive integrated moving average models with higher forecasting accuracy, called FARIMAH. In the proposed model, a probabilistic neural network (PNN) is used to determine the spaces of the forecasted interval by FARIMA in which the probability of existing actual values is higher. Then, according to the achieved results by PNN, the spaces that have lower existing probability are deleted from obtained interval by FARIMA. In order to show the applicability and effectiveness of the proposed model, it is applied to foreign exchange rate markets forecasting and its performance is compared with the traditional fuzzy autoregressive integrated moving average models.

The rest of the paper is organized as follows. In the next section, the basic concepts of the fuzzy autoregressive integrated moving average (FARIMA) models are briefly reviewed. In section 3, the probabilistic neural networks (PNNs), which are chosen as classifier model, are briefly reviewed. In section 4, the formulation of the hybrid proposed model is presented. In section 5, the proposed model is applied to foreign exchange rate markets forecasting and its performance is compared with other those models. Conclusions will be the final section of the paper.

## 2. The Fuzzy Autoregressive Integrated Moving Average (FARIMA) model

For more than half a century, the autoregressive integrated moving average (ARIMA) models have dominated many areas of time series forecasting. In an ARIMA (p,d,q) model, the future value of a variable is assumed to be a linear function of several past observations and random errors [2]. That is, the underlying process that generates the time series with the mean $\mu$ has the form:

$$\phi(B)\nabla^d (y_t - \mu) = \theta(B)a_t \tag{1}$$

where, $y_t$ and $a_t$ are the actual value and random error at time period *t*, respectively; $\phi(B) = 1 - \sum_{i=1}^{p} \varphi_i B^i$, and $\theta(B) = 1 - \sum_{j=1}^{q} \theta_j B^j$ are polynomials in B of degree p and q, $\phi_i$ $(i = 1,2,...,p)$ and $\theta_j$ $(j = 1,2,...,q)$ are model parameters, $\nabla = (1 - B)$, *B* is the backward shift operator, p and q are integers and often referred to as orders of the model, and d is an integer and often referred to as order of differencing. Random errors, $a_t$, are assumed to be independently and identically distributed with a mean of zero and a constant variance of $\sigma^2$.

However, the parameters of the autoregressive integrated moving average, $\varphi_1, \varphi_2, ...., \varphi_p$ and $\theta_1, \theta_2, ...., \theta_q$ are crisp. In the fuzzy autoregressive integrated moving average models [3], Instead of using these crisp parameters, fuzzy parameters, $\tilde{\varphi}_1, \tilde{\varphi}_2, ...., \tilde{\varphi}_p$ and $\tilde{\theta}_1, \tilde{\theta}_2, ...., \tilde{\theta}_q$, in the form of triangular fuzzy numbers are used. A fuzzy ARIMA model is described by a fuzzy function with a fuzzy parameter as follows:

$$\tilde{\Phi}_p(B) W_t = \tilde{\theta}_q(B) a_t \tag{2}$$

$$W_t = (1 - B)^d (Z_t - \mu) \tag{3}$$

$$\begin{aligned} \tilde{W}_t &= \tilde{\phi}_1 W_{t-1} + \tilde{\phi}_2 W_{t-2} + .... + \tilde{\phi}_p W_{t-p} + a_t \\ &\quad - \tilde{\theta}_{p+1} a_{t-1} - \tilde{\theta}_{p+2} a_{t-2} - ... - \tilde{\theta}_{p+q} a_{t-q} \end{aligned} \tag{4}$$

where $\{Z_t\}$ are observations, $\tilde{\varphi}_1, \tilde{\varphi}_2, ...., \tilde{\varphi}_p$ and $\tilde{\theta}_1, \tilde{\theta}_2, ...., \tilde{\theta}_q$, are fuzzy numbers. Eq. (4) is modified as:

$$\begin{aligned} \tilde{W}_t &= \tilde{\beta}_1 W_{t-1} + \tilde{\beta}_2 W_{t-2} + .... + \tilde{\beta}_p W_{t-p} + a_t \\ &\quad - \tilde{\beta}_{p+1} a_{t-1} - \tilde{\beta}_{p+2} a_{t-2} - ... - \tilde{\beta}_{p+q} a_{t-q} \end{aligned} \tag{5}$$

Fuzzy parameters in the form of triangular fuzzy numbers are used as follows:

$$\mu_{\beta_i}(\beta_i) = \begin{cases} 1 - \dfrac{|\alpha_i - \beta_i|}{c_i} & if \quad \alpha_i - c_i \le \beta_i \le \alpha_i + c_i, \\ 0 & otherwise, \end{cases} \tag{6}$$

where $\mu_{\tilde{\beta}}(\beta_i)$ is the membership function of the fuzzy set that represents parameter $\beta_i$, $\alpha_i$ is the center of the fuzzy number, and $c_i$ is the width or spread around the center of the fuzzy number. Using fuzzy parameters $\beta_i$ in the form of triangular fuzzy numbers and applying the extension principle, the membership of *W* in Eq. (5) is given as:

$$\mu_{\tilde{w}}(W_t) = \begin{cases} 1 - \dfrac{\left| W_t - \sum_{i=1}^{p} \alpha_i W_{t-i} - a_t + \sum_{i=p+1}^{p+q} \alpha_i a_{t+p-i} \right|}{\sum_{i=1}^{p} c_i |W_{t-i}| + \sum_{i=p+1}^{p+q} c_i |a_{t+p-i}|} \\ \qquad for \quad W_t \ne 0, \quad a_t \ne 0 \\ 0 \qquad otherwise \end{cases} \tag{7}$$

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 $y_1, y_2, ...., y_k$ to a certain *h*-level. A choice of the *h*-level value influences the widths *c* of the fuzzy parameters:

$$\mu_y(y_t) \ge h \qquad for \quad t = 1,2,...,k \tag{8}$$

The index *t* refers to the number of non-fuzzy data used for constructing the model. On the other hand, the fuzziness *S* included in the model is defined by

$$S = \sum_{i=1}^{p} \sum_{t=1}^{k} c_i |\varphi_{ii}| |W_{t-i}| + \sum_{i=p+1}^{p+q} \sum_{t=1}^{k} c_i |\rho_{i-p}| |a_{t+p-i}| \tag{9}$$

where $\rho_{i-p}$ is the autocorrelation coefficient of time lag *i-p*, $\varphi_{ii}$ 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:

$$\begin{aligned}
&\textit{Minimize} \quad S=\sum_{i=1}^{p}\sum_{t=1}^{k}c_i\left|\phi_{ii}\right|\left|W_{t-i}\right|+\sum_{i=p+1}^{p+q}\sum_{t=1}^{k}c_i\left|\rho_{i-p}\right|\left|a_{t+p-i}\right| \\
&\qquad \sum_{i=1}^{p}\alpha_i W_{t-i}+a_t-\sum_{i=p+1}^{p+q}\alpha_i a_{t+p-i}+(1+h)\left(\sum_{i+1}^{p}c_i\left|W_{t-i}\right|+\sum_{i=p+1}^{p+q}c_i\left|a_{t+p-i}\right|\right)\ge W_t \quad t=1,2,..,k \\
&\textit{subject.to} \quad \sum_{i=1}^{p}\alpha_i W_{t-i}+a_t-\sum_{i=p+1}^{p+q}\alpha_i a_{t+p-i}+(1+h)\left(\sum_{i+1}^{p}c_i\left|W_{t-i}\right|+\sum_{i=p+1}^{p+q}c_i\left|a_{t+p-i}\right|\right)\le W_t \quad t=1,2,..,k \\
&\qquad c_i\ge 0 \qquad for\, i=1,2,...,p+q
\end{aligned} \tag{10}$$

At last, according to the Ishibuchi and Tanaka [26] 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.

## 3. Probabilistic Neural Networks (PNNs)

The probabilistic neural network (PNN) is a Bayes–Parzen classifier [27] that is often an excellent pattern classifier in practice. The foundation of the approach is well known decades ago (1960s), however, the method was not of a widespread use because of the lack of sufficient computation power until recently [28]. Donald Specht [29] first introduced the probabilistic neural networks in 1990, who demonstrated how the Bayes–Parzen classifier could be broken up into a large number of simple processes implemented in a multilayer neural network each of which could be run independently in parallel.

Because the probabilistic neural network is primarily based on Bayes–Parzen classification, it is of interest to discuss briefly both Bayes theorem for conditional probability and Parzen's method for estimating probability density function of random variables. In order to understand Bayes' theorem, consider a sample $x=[x_1,x_2,...,x_n]$ taken from a collection of samples belonging to a number of distinct populations $(1,2,...,k,...K)$. Assuming that the (prior) probability that a sample belongs to the *k*th population (class) is $h_k$, the cost associated with misclassifying that sample is $l_k$, and that the true probability density function of all populations $f_1(x),f_2(x),...,f_k(x),...,f_K(x)$ are known, Bayes theorem classifies an unknown sample into the ith population [30] if

$$h_i l_i f_i(x) > h_j l_j f_j(x) \qquad \forall j\ne i, \quad j=1,2,...K. \tag{11}$$

The density function $f_k(x)$ corresponds to the concentration of class $k$ examples around the unknown example. As seen from Eq. (11), Bayes' theorem favors a class that has high density approximately the unknown sample, or if the cost of misclassification or prior probability is high.

The biggest problem with the Bayes' classification approach lies in the fact that the probability density function $f_k(x)$ is not usually known. In nearly all standard statistical classification algorithms, some knowledge regarding the underlying distribution of the population of all random variables used in classification should be known or reasonably assumed. Most often, normal (Gaussian) distribution is assumed; however, the assumption of normality cannot always be safely justified [31]. When the distribution is not known (which is often the case) and the true distribution deviates considerably from the assumed one, the traditional statistical methods normally run into major classification problems resulting in high misclassification rate. There is a need to derive an estimate of $f_k(x)$, from the training set composed of the training example, rather than just assume normal distribution. The resulting distribution will be a multivariate probability density function (PDF) that combines all the explanatory random variables.

In order to derive such distribution estimator from a set of training examples, the Parzen's method [32] is usually used. The univariate case of PDF was proposed by Parzen [32] and then was extended to the multivariate case by Cacoullos [33]. The multivariate PDF estimator, $g(x)$, may be expressed as:

$$g(x_1, x_2, ..., x_n) = \frac{1}{N\,\sigma_1\sigma_2...\sigma_n} \times \sum_{i=1}^{N} W\left(\frac{x_1 - x_{1,i}}{\sigma_1}, \frac{x_2 - x_{2,i}}{\sigma_2}, ..., \frac{x_n - x_{n,i}}{\sigma_n}\right), \quad (12)$$

where $\sigma_1, \sigma_2, ... \sigma_n$ are the smoothing parameters representing standard deviation (also called window or kernel width) around the mean of *n* random variables $x_1, x_2, ..., x_n$, $W$ is a weighting function to be selected with specific characteristics [27, 29], and $N$ is the total number of training examples. Now, if all smoothing parameters are assumed equal (i.e., $\sigma_1 = \sigma_2 = .... = \sigma_n = \sigma$) and a bell-shaped Gaussian function is used for $W$, a reduced form of Eq. (12) is as follows [34]:

$$g(x) = \frac{1}{(2\pi)^{n/2}\,\sigma^n} \times {}^{1}\!/_{N} \sum_{i=1}^{N} exp\left[-\frac{\|(x - x_i)\|^2}{2\sigma^2}\right], \quad (13)$$

where $x$ is the vector of random variables (explanatory variables), and $x_i$ is the *i*th training vector. Eq. (13) represents the average of the multivariate distributions where each distribution is centered at one distinct training example. It is worth mentioning that the assumption of a Gaussian weighting function does not imply that the overall PDF will be Gaussian (normal), however, other weighting functions such as the reciprocal function $\left(w(r) = {}^{1}\!/_{1+r^2}\right)$ may be used. As the sample size, $N$, increases, the Parzen's PDF estimator asymptotically approaches the true underlying density function.

Regarding the network's operation based on the aforementioned mathematics, consider the simple network architecture in Fig. 1 with *n* input nodes in the input layer, two population classes (classes 1 and 2), $N_1$ training examples belonging to class 1, and $N_2$ examples in class 2. The pattern layer is designed to contain one neuron for each training case available and the neurons are split into the two classes. The summation layer contains one neuron for each class. The output layer contains one neuron that operates trivial threshold discrimination; it simply retains the maximum of the two summation neurons [35].

The probabilistic neural network executes a training case by first presenting it to all pattern layer neurons. Each neuron in the pattern layer computes a distance measure between the presented input vector and the training example represented by that pattern neuron. The probabilistic neural network then subjects this distance measure to the Parzen window (weighting function, $W$) and yields the activation of each neuron in the pattern layer. Subsequently, the activation from each class is fed to the corresponding summation layer neuron, which adds all the results in a particular class together. The activation of each summation neuron is executed by applying the remaining part of the Parzen's estimator equation (e.g., the constant multiplier in Eq. (13)) to obtain the estimated probability density function value of population of a particular class [28].

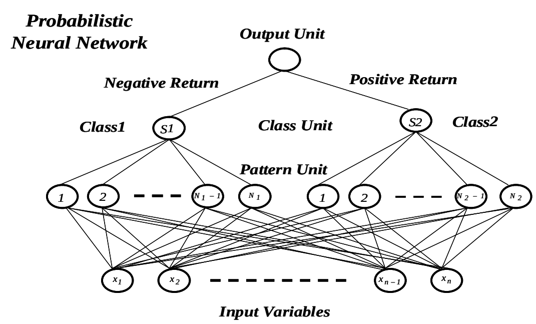


Fig. 1. A simple probabilistic neural network.

If the misclassification cost and prior probabilities are equal between the two classes, and the classes are mutually exclusive (i.e., no case can be classified into more than one class) and exhaustive (i.e., the training set covers all classes fairly), the activation of the summation neurons will be equal to the posterior probability of each class. The results from the two summation neurons are then compared and the largest is fed forward to the output neuron to yield the computed class and the probability that this example will belong to that class.

The most important parameter that needs to be determined to obtain an optimal probabilistic neural network is the smoothing parameters ($\sigma_1, \sigma_2, ... \sigma_n$) of the random variables [36]. A straightforward procedure involves selecting an arbitrary value of σ's, training the network, and testing it on a test (validation) set of examples. This procedure is repeated for other σ's and the set of σ's that produces the least misclassification rate (percentage of examples that were misclassified) is chosen. A better and more efficient procedure for searching for the optimal smoothing parameter of random variables and classes is proposed by Masters [27]. This procedure prevents any bias in the network to the correctly classified examples, and thus will be followed in this study. Other details on the mathematics as well as advanced variations of probabilistic neural networks are given in Specht [29] and Masters [27].

## 4. Formulation of proposed model (FARIMAH)

Although autoregressive integrated moving average models have the advantages of accurate forecasting over a short period and ease of implementation, they have data limitation. ARIMA models require at least fifty, or preferably one hundred and higher data in order to yield desired results. However, in real situations, due to uncertainty resulting from the integral environment and rapid development of new technology, future situations must be forecasted using small data sets over a short span of time. Efficient forecasting methods are, therefore, needed today that can achieve their objectives in situations with small quantities of historical data available [10].

Tseng *et al.* [3] proposed the fuzzy autoregressive integrated moving average (FARIMA) models in order to combine the advantages of the fuzzy regression and ARIMA models and also to simultaneously fulfill the limitations of these models. The FARIMA models not only can overcome the limitations of their components but also can provide better performance than ARIMA models. Despite all advantages cited for the FARIMA models, the forecasted interval of these models is extended in some specific data conditions. For instance, when data has high volatility or includes a significant difference or outliers. In additional, since in the basis model of FARIMA model, ARIMA, the future value of a variable is assumed to be a linear function of several past observations and random errors, the approximation of FARIMA models may be totally inappropriate if the underlying mechanism is nonlinear. However, real world systems are often nonlinear [2].

The main purpose of the proposed model is to overcome two aforementioned limitations of the FARIMA models using unique distinction ability of nonlinear probabilistic classifiers and yield a more general and more accurate model than traditional FARIMA models in financial incomplete data situations. In the proposed model, the probabilistic neural networks (PNNs) are applied as nonlinear classifiers in order to determine more probability spaces for actual values in the forecasted interval by FARIMA model and also existing nonlinear patterns in the time series. There are a number of appealing features, which justify our adoption of this type of neural networks in this study. First, training of probabilistic neural networks is rapid, enabling us to develop a frequently updated training scheme. Essentially, the network is re-trained each time the data set is updated and thus the most current information can be reflected in estimation. Second, the logic of probabilistic neural network is able to extenuate the effects of outliers and questionable data points and thereby reduces extra effort on scrutinizing training data. Third and the most important, probabilistic neural networks are conceptually built on the Bayesian method of classification which given enough data, is capable of classifying a sample with the maximum probability of success [35].

In our proposed model, based on the literature of hybrid models [37], a time series $\{y_t\}$ is considered to be composed of a linear autocorrelation structure ($L_t$) and a nonlinear component ($N_t$). Therefore, in the first stage of the proposed model, an autoregressive integrated moving average (ARIMA) is initially applied in order to model the linear component and to generate the residuals ($e_t$).

$$\begin{aligned} L_t &= \left[\sum_{i=1}^{p} \varphi_i \, z_{t-i} - \sum_{j=1}^{q} \theta_j \, \varepsilon_{t-j}\right] + e_t \\ &= \sum_{k=1}^{p+q} \alpha_k \, x_k + e_t = \hat{L}_t + e_t , \end{aligned} \tag{14}$$

where, $x_k$ is equal to the $z_{t-k}$ for $k = 1,2,...,p$; and equal to the $\varepsilon_{t+p-k}$ for $k = p+1, p+2,..., p+q$, $\hat{L}_t$ is the estimation of the linear component at time *t*, $\theta_1, \theta_2,....,\theta_q$ and $\varphi_1, \varphi_2,....,\varphi_p$ are the parameters of the ARIMA, and $e_t$ is the residual of the ARIMA at time t. Since the ARIMA cannot capture nonlinear structures, achieved residuals of this stage will contain all nonlinear structures, so they can be used as nonlinear component of time series. The results of the first stage are the optimum solution of the ARIMA parameters, $\alpha^* = \left(\alpha_1^*, \alpha_2^*,....,\alpha_{p+q}^*\right)$, the estimation of the linear component, $\hat{L}_t$, and the nonlinear component, $N_t$, which are used in the next stages.

In the second stage and after linear modeling, obtained linear parameters according to the basic concepts of FARIMA, $\alpha^* = \left(\alpha_1^*, \alpha_2^*, ...., \alpha_{p+q}^*\right)$, are considered in the form of triangular fuzzy numbers as Eq. (6). Then the minimal fuzziness of these parameters is determined using the same criterion as in the Eq. (10).

In the third stage, based on the Ishibuchi and Tanaka [26] opinion, the data around the model's upper and lower boundaries are deleted and then the model is reformulated. The result of this stage is a fuzzy model as follows:

$$\tilde{z}_t = \langle \alpha_1, c_1 \rangle z_{t-1} + ... + \langle \alpha_p, c_p \rangle z_{t-p} - \langle \alpha_{p+1}, c_{p+1} \rangle e_{t-1} - .... - \langle \alpha_{p+q}, c_{p+q} \rangle e_{t-q} \quad (15)$$

where $\alpha_i$ is the center of the fuzzy number, and $c_i$ is the width or spread around the center of the fuzzy number. Finally, in the fourth stage, a probabilistic neural network is designed and trained in order to determine more probability spaces for actual values. For this purpose, the forecasted interval in the previous stage is first divided in to $n$ equal subintervals. Then each $k$ consecutive subinterval is considered as a class with assigned numbers "*class=1, 2,…, n-k+1*". Where $n$ and $k$ are integer, $n>1$, and $k<n$. Afterwards, the assigned number(s) of the mentioned subinterval(s) that consists of the actual value is considered as target value(s) of the probabilistic neural network. The probabilistic neural network is then trained by considering the target values and a subset of effective variables as output and input values, respectively. The effective variables on the target value of the mentioned probabilistic neural network are as follows:

i) Lags 1 until $p$th of the time series at time $t$ $\left(z_{t-1}, z_{t-2}, ..., z_{t-p}\right)$.

ii) Lags 1 until $q$th of the residuals of ARIMA at time $t$ $\left(e_{t-1}, e_{t-2}, ..., e_{t-q}\right)$.

iii) Estimated value of the time series by ARIMA at time $t$ $(\hat{y}_t)$.

iv) Lags 1 until $r$th of the estimated values of the ARIMA at time t $\left(\hat{z}_{t-1}, \hat{z}_{t-2}, ..., \hat{z}_{t-r}\right)$.

v) Estimated lower and upper bounds of the time series by FARIMA at time $t$ $\left(Lo_t, Up_t\right)$.

vi) Lags 1 until $s$th of the estimated lower and upper bounds at time $t$ $\left(Lo_{t-1}, Up_{t-1}, ..., Lo_{t-s}, Up_{t-s}\right)$.

where $p$, $q$, $r$ and s are integer. The result of this stage is an interval with $k/n$ width of FARIMA and confidence degree $\alpha$ ($\alpha$ is the distinction ability of the PNN in the test data). The flowchart of the hybrid proposed model is shown in Fig.2. It must be noted that $\alpha$ generally has a straight relationship to the number of the subintervals. Therefore, based on the conditions of the under-study problem, both width or confidence degree of the final interval can be considered as decision-making criterion. For instance, using proposed model, we can determine the maximum confidence degree of the forecasted interval for a given width or the minimum width of the forecasted interval for a given confidence degree.

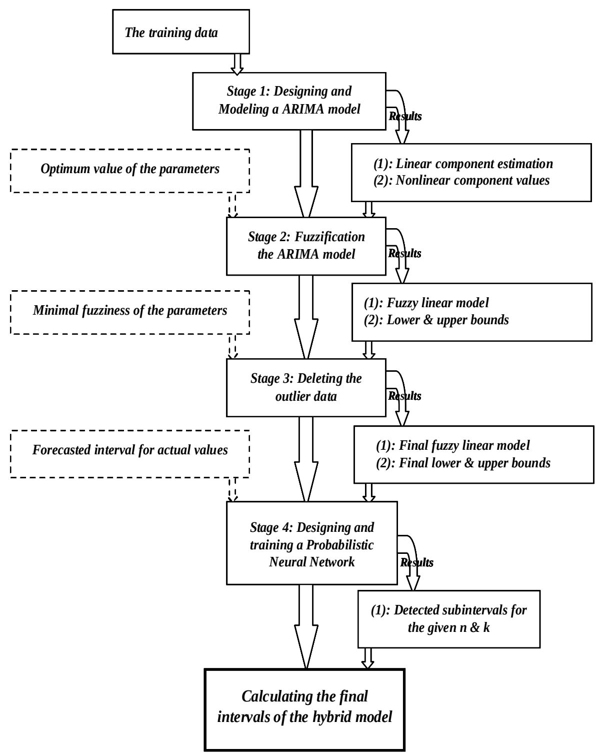


Fig. 2. The flowchart of the hybrid proposed model.

## 5. Application of the proposed model to financial markets forecasts

In order to demonstrate the appropriateness and effectiveness of the proposed model, the following applications in exchange rate ─*the United States dollar, British pound, and Euro all against the Iran rial*─ forecasting have been considered. In the next section, the process of the proposed hybrid models is illustrated, as an example for forecasting the United States dollar against the Iran rial exchange rate.

### 5.1. *The exchange rate (US dollar/ Iran rial) forecasts*

The information used in this investigation consists of 42 daily observations of the exchange rate of United States dollar against Iran rial from *5 Nov to 16 Dec*, 2006 that are shown in Fig. 3. As in the previous works, applying the hybrid method, 35 observations (five weeks) are first used to formulate the model and the last seven observations (last week) are used to evaluate the performance of the proposed model [10].

***Stage I:*** *fitting the ARIMA model:* Using the Eviews package software, the best-fitted model is ARIMA (2, 1, 0) as follows. The fitted values by ARIMA model are plotted in Fig. 4.

$$Z_t = 9060.05 + 0.607\,Z_{t-1} + 0.421 Z_{t-2} + a_t. \tag{16}$$

***Stage II:*** *determining the minimal fuzziness:* Setting $(\alpha_0, \alpha_1, \alpha_2) = (9060.05, 0.607, 0.421)$, the fuzzy parameters are calculated using Eq. (10) (with *h=0*) as follows. The obtained upper and lower bounds in this stage are plotted in Fig. 5.

$$\begin{aligned} \tilde{Z}_t = 9060.05 &+ \langle 0.607, 0.00028 \rangle Z_{t-1} \\ &+ \langle 0.421, 0.00 \rangle Z_{t-2}. \end{aligned} \tag{17}$$

It can be seen from Fig. 5 that the actual values located in the fuzzy intervals, however, the thread of fuzzy intervals are too wide, especially when the macro-economic environment is smooth. Therefore, the method of Ishibuchi and Tanaka is applied in the next stage in order to resolve this problem and provide a narrower interval for the decision maker.

***Stage III:*** *Deleting the outliers:* It is known from the aforementioned results that the observation of *15 Nov* is located at the upper bound (outlier); therefore, the LP constrained equation that is produced by this observation is first deleted and then the stage II is renewed, with *h=0*. The obtained upper and lower bounds in this stage are plotted in Fig. 6. These results for test data set before and after deleting the outlier data are also given in Table 1.

Table 1. Obtained results in stage III for test data before and after deleting outlier data*

| *Date* | *Actual* | *Before deleting* | | *After deleting* | |
|---|---|---|---|---|---|
| | | *Lower bound* | *Upper bound* | *Lower bound* | *Upper bound* |
| *10-Dec.* | *9082* | *9080* | *9085* | *9081* | *9085* |
| *11- Dec.* | *9083* | *9080* | *9085* | *9080* | *9084* |
| *12- Dec.* | *9083* | *9080* | *9085* | *9081* | *9085* |
| *13- Dec.* | *9082* | *9080* | *9085* | *9081* | *9085* |
| *14- Dec.* | *9081* | *9080* | *9085* | *9080* | *9084* |
| *15- Dec.* | *9082* | *9079* | *9084* | *9079* | *9083* |
| *16-Dec.* | *9082* | *9079* | *9084* | *9080* | *9084* |

*Note*: The upper and lower bounds values are rounded.

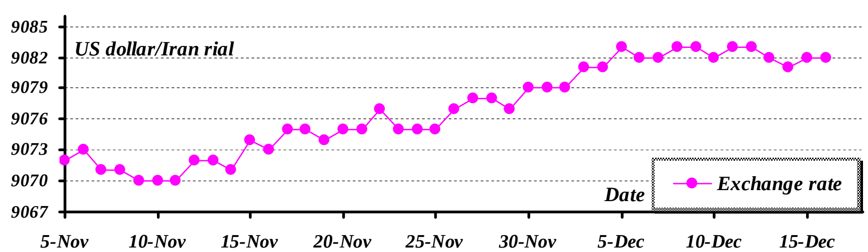


Fig. 3 Exchange Rate (US dollar/ Iran rial) from *5 Nov to 16 Dec 2006*.

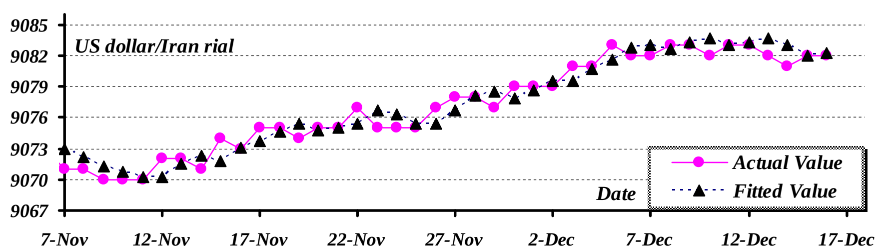


Fig. 4 ARIMA fitted values for US dollar against Iran rial exchange rate.

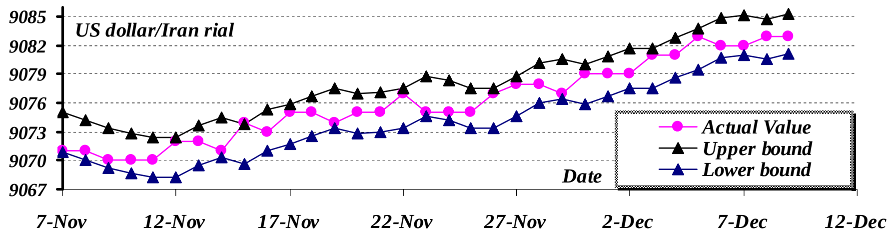


Fig. 5 Upper and lower bounds obtained for US dollar against Iran rial.

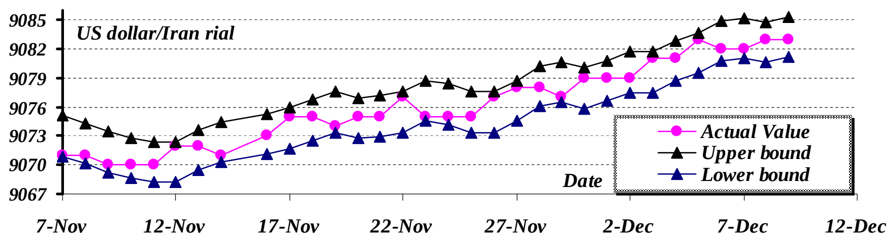


Fig. 6 Upper and lower bounds obtained for US dollar against Iran rial (after deleting).

***Stage IV:*** In the fourth stage, the probabilistic neural network is applied in order to diagnose more probability spaces for the actual values in the forecasted interval in the previous stage. Similar to the FARIMA model, the first 35 observations (training sample) are applied in order to train the designed network and the last seven observations (test sample) are applied in order to test the performance of the network. The number of the subinterval is considered equal to two. In the other hand, the forecasted interval by FARIMA is divided to two equal subintervals ($n = 2$). In order to obtain the optimal architecture of the probabilistic neural network, based on the concepts of artificial neural networks design [38] and using Constructive algorithm in *MATLAB7* package software, different network architectures are evaluated in order to compare the PNN performance. The best-fitted network, which is selected, and therefore, the architecture that present the best forecasting accuracy with the test data, is composed of five inputs and one output neuron. The architecture of the designed probabilistic neural network is shown in Fig. 7. The obtained results of the probabilistic neural network for training data set are also given in Table 2.

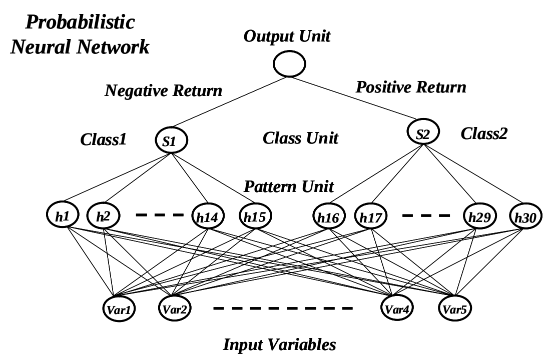


Fig. 7 The architecture of the designed PNN.

where

Var 1: First lag of the time series at time $t$ $(z_{t-1})$.

Var 2: Estimated value of the time series by ARIMA at time $t$ $(\hat{y}_t)$.

Var 3: First lag of the estimated values of the ARIMA at time $t$ $(\hat{z}_{t-1})$.

Var 4: Estimated lower and upper bounds of the time series by FARIMA at time $t$ $(Lo_t, Up_t)$.

Var 5: First lag of the estimated lower and upper bounds at time $t$ $(Lo_{t-1}, Up_{t-1})$.

Table 2. Detected subintervals for actual values in the forecasted interval by FARIMA for training data set.

| *Date* | *Subintervals* | |
|---|---|---|
| | ***Actual*** | ***Detected*** |
| ***7-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***8-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***9-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***10-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***11-November*** | *Upper Subinterval* | *Lower Subinterval* |
| ***12-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***13-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***14-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***16-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***17-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***18-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***19-November*** | *Upper Subinterval* | *Lower Subinterval* |
| ***20-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***21-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***22-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***23-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***24-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***25-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***26-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***27-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***28-November*** | *Lower Subinterval* | *Upper Subinterval* |
| ***29-November*** | *Lower Subinterval* | *Lower Subinterval* |
| ***30-November*** | *Upper Subinterval* | *Upper Subinterval* |
| ***1-December*** | *Upper Subinterval* | *Upper Subinterval* |
| ***2-December*** | *Upper Subinterval* | *Upper Subinterval* |
| ***3-December*** | *Upper Subinterval* | *Upper Subinterval* |
| ***4-December*** | *Upper Subinterval* | *Upper Subinterval* |
| ***5-December*** | *Upper Subinterval* | *Upper Subinterval* |
| ***6-December*** | *Upper Subinterval* | *Lower Subinterval* |
| ***7-December*** | *Upper Subinterval* | *Lower Subinterval* |
| ***8-December*** | *Upper Subinterval* | *Upper Subinterval* |
| ***9-December*** | *Upper Subinterval* | *Upper Subinterval* |

As can be seen in Table 2, the probabilistic neural network can approximately determined the 85% of subinterval correctly. Now, based on the achieved subintervals by probabilistic neural network and achieved lower and upper bounds by fuzzy autoregressive integrated moving average, the lower and upper bounds of the hybrid proposed model are calculated. These results for training data set are presented in Table 3 and Fig 8. In additional, the results of the probabilistic and proposed model for test data set are given in Table 4.

Table 3. Detected subintervals for actual values in the forecasted interval by FARIMA for training data set*.

| ***Date*** | ***Actual*** | ***Lower bound*** | ***Upper bound*** |
|---|---|---|---|
| ***7- November*** | *9071* | *9070* | *9073* |
| ***8- November*** | *9071* | *9069* | *9071* |
| ***9- November*** | *9070* | *9069* | *9071* |
| ***10- November*** | *9070* | *9068* | *9070* |
| ***11- November*** | *9070* | *9070* | *9072* |
| ***12- November*** | *9072* | *9070* | *9072* |
| ***13- November*** | *9072* | *9072* | *9074* |
| ***14- November*** | *9071* | *9070* | *9072* |
| ***16- November*** | *9073* | *9071* | *9073* |
| ***17- November*** | *9075* | *9074* | *9076* |
| ***18- November*** | *9075* | *9075* | *9077* |
| ***19- November*** | *9074* | *9075* | *9077* |
| ***20- November*** | *9075* | *9075* | *9077* |
| ***21- November*** | *9075* | *9075* | *9077* |
| ***22- November*** | *9077* | *9075* | *9077* |
| ***23- November*** | *9075* | *9074* | *9076* |
| ***24- November*** | *9075* | *9073* | *9075* |
| ***25- November*** | *9075* | *9073* | *9075* |
| ***26- November*** | *9077* | *9075* | *9077* |
| ***27- November*** | *9078* | *9077* | *9079* |
| ***28- November*** | *9078* | *9075* | *9077* |
| ***29- November*** | *9077* | *9076* | *9078* |
| ***30- November*** | *9079* | *9078* | *9080* |
| ***1- December*** | *9079* | *9079* | *9081* |
| ***2- December*** | *9079* | *9079* | *9081* |
| ***3- December*** | *9081* | *9079* | *9081* |
| ***4- December*** | *9081* | *9081* | *9083* |
| ***5- December*** | *9083* | *9081* | *9083* |
| ***6- December*** | *9082* | *9083* | *9085* |
| ***7- December*** | *9082* | *9082* | *9085* |
| ***8- December*** | *9083* | *9082* | *9084* |
| ***9- December*** | *9083* | *9083* | *9085* |

*Note*: The upper and lower bounds values are rounded.

Now, if suppose that our goal is to provide a interval with confidence coefficient of 100%; in the other hands, if we want to yield the narrowest interval with confidence coefficient of 100%, then $n$ and $k$ will be respectively equal to 5 and 3 and achieved interval of the proposed model will be a interval with 2.5 width. The lower and upper bounds of the proposed model for test data set with $n = 5$, $k = 3$, and $\alpha = 100\%$ are given in Table 5.

Table 4 Obtained results of the probabilistic neural network and proposed model for test data set in stage IV*

| Date | Actual | Subintervals | | Bounds | |
|---|---|---|---|---|---|
| | | Actual | Detected | Lower bound | Upper bound |
| 10-Dec. | 9082 | Lower | Lower | 9081 | 9083 |
| 11- Dec. | 9083 | Upper | Upper | 9082 | 9085 |
| 12- Dec. | 9083 | Upper | Upper | 9083 | 9085 |
| 13- Dec. | 9082 | Lower | Lower | 9081 | 9083 |
| 14- Dec. | 9081 | Upper | Lower | 9082 | 9085 |
| 15- Dec. | 9082 | Upper | Upper | 9081 | 9084 |
| 16-Dec. | 9082 | Upper | Upper | 9082 | 9084 |

*Note*: The upper and lower bound values are rounded.

In addition, if suppose that our goal is to provide a interval with 1.7 width; in the other hands, if we decide to obtain the maximum confidence of a interval with 1.7 width, then *n* and *k* will be respectively equal to 5 and 2 and achieved interval of the proposed model will be a interval with confidence coefficient of 57%. The lower and upper bounds of the proposed model for test data set with $n=5$, $k=3$, and $\alpha=57\%$ are also given in Table 5.

Table 5. Obtained lower and upper bounds by the proposed model for test data set*

| Date | Actual | Bounds (n=5, k=3, and α=100%) | | Bounds (n=5, k=2, and α=57%) | |
|---|---|---|---|---|---|
| | | Lower bound | Upper bound | Lower bound | Upper bound |
| 10-Dec. | 9082 | 9081 | 9084 | 9083 | 9085 |
| 11- Dec. | 9083 | 9082 | 9085 | 9082 | 9084 |
| 12- Dec. | 9083 | 9083 | 9085 | 9081 | 9083 |
| 13- Dec. | 9082 | 9081 | 9084 | 9083 | 9085 |
| 14- Dec. | 9081 | 9080 | 9083 | 9082 | 9084 |
| 15- Dec. | 9082 | 9081 | 9084 | 9081 | 9083 |
| 16-Dec. | 9082 | 9081 | 9084 | 9082 | 9084 |

*Note*: The upper and lower bound values are rounded.

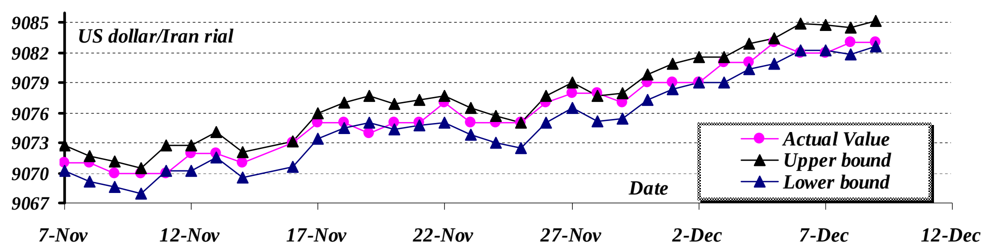


Fig. 8 Upper and lower bounds obtained by proposed model for training data (width =2.6 & α = 85%).

### 5.2. *Comparison with other forecasting models*

In this section, the predictive capability of the proposed model in both interval and point estimation cases is compared with some other fuzzy and nonfuzzy forecasting models in these fields, using three exchange rate data sets including the United States dollar, British pound, and Euro all against the Iran rial. The considered fuzzy and nonfuzzy interval forecasting models in this study are respectively including the fuzzy autoregressive integrated moving average (FARIMA) and classic autoregressive integrated moving average (ARIMA). In addition, the autoregressive integrated moving average (ARIMA), the multilayer perceptrons (MLPs) and Chen's fuzzy time series (first-order) [39], Chen's fuzzy time series (high-order) [40], Yu's fuzzy time series [41], and adaptive neuro-fuzzy inference systems (ANFIS) [42] are respectively considered as nonfuzzy and fuzzy models in the field of the point estimation. The width of the forecasted interval, and MAE (Mean Absolute Error) and MSE (Mean Squared Error) are respectively employed as performance indicators in order to measure forecasting performance in the interval and point estimation cases. The MAE and MSE are respectively computed from the following equations:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|e_i| \tag{18}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 \tag{19}$$

Based on the results obtained from these cases studied (Table 6), the predictive capabilities of the proposed model are rather encouraging and the possible interval by the proposed model with 100% confidence is narrower than the possible interval of the fuzzy autoregressive integrated moving average (FARIMA). The width of the forecasted interval in the proposed model is 2.5, 14.1, and 7.0 in the US dollar, Euro, and British pound exchange rate forecasting cases, indicating a 40.5%, 40.5%, and 39.76% improvement upon the possible interval of the FARIMA, respectively. Moreover, the width of the forecasted interval by the proposed model is narrower than obtained interval by ARIMA (95% Confidence Interval) model.

In addition, according to the numerical results (Table 7 and 8), the MAE and MSE of the proposed model are lower than the FARIMA for all aforementioned exchange rate cases. For example in terms of MSE, the percentage improvements of the proposed model over the FARIMA, are 25.63%, 22.04%, and 64.17%, in the US dollar, Euro, and British pound exchange rate forecasting cases, respectively. Similarity, the MAE and MSE of the proposed model are lower than Chen's fuzzy time series (first-order and second-order), and Yu's fuzzy time series, autoregressive integrated moving average (ARIMA), multilayer perceptrons (MLPs), and adaptive neuro-fuzzy inference systems (ANFIS) in all cases.

Table 6. Comparison of forecasted interval widths by the proposed model compared to other models (interval estimation).

| **Model** | **Cases** | **Forecasted interval width** | **Improvement percentage** ARIMA (95% Confidence) | Fuzzy ARIMA | Proposed Model (α=100%) |
|---|---|---|---|---|---|
| ***ARIMA (95% Confidence Interval)*** | US dollar /Iran rial | 16.2 | 0.0 | - | - |
| ***Fuzzy ARIMA*** | | 4.2 | 74.1 | 0.0 | - |
| ***Proposed Model (α=100%)*** | | 2.5 | 84.6 | 40.5 | 0.0 |
| ***ARIMA (95% Confidence Interval)*** | Euro / Iran rial | 66.9 | 0.0 | - | - |
| ***Fuzzy ARIMA*** | | 23.7 | 64.6 | 0.0 | - |
| ***Proposed Model (α=100%)*** | | 14.1 | 78.9 | 40.5 | 0.0 |
| ***ARIMA (95% Confidence Interval)*** | B. pound /Iran rial | 56.8 | 0.0 | - | - |
| ***Fuzzy ARIMA*** | | 11.6 | 79.6 | 0.0 | - |
| ***Proposed Model (α=100%)*** | | 7.0 | 87.7 | 39.7 | 0.0 |

Table 7. Comparison of the performance of the proposed model with those of other forecasting models (point estimation).

| **Model** | **US dollar / Iran rial** MAE | MSE | **Euro / Iran rial** MAE | MSE | **British pound / Iran rial** MAE | MSE |
|---|---|---|---|---|---|---|
| ***Autoregressive Integrated Moving Average (ARIMA)*** | 0.924 | 1.24 | 56.44 | 3443 | 19.772 | 849.5 |
| ***Chen's fuzzy time series (first-order)*** | 0.750 | 0.777 | 48.60 | 3324 | 29.047 | 1232.7 |
| ***Chen's fuzzy time series (second-order)*** | 0.750 | 0.777 | 48.60 | 3324 | 28.011 | 1110.3 |
| ***Yu's fuzzy time series*** | 0.750 | 0.777 | 40.27 | 1783 | 25.964 | 917.3 |
| ***Multilayer perceptrons (MLPs)*** | 0.692 | 0.686 | 43.1 | 2173 | 18.112 | 415.7 |
| ***Adaptive Neuro-Fuzzy Inference Systems (ANFIS)*** | 0.625 | 0.547 | 40.89 | 1979 | 19.429 | 763.8 |
| ***Proposed Model (FARIMAH)*** | 0.619 | 0.407 | 34.8 | 1542.8 | 15.662 | 273.7 |

Table 8. Improvement percentage of the proposed model in comparison with those of other forecasting models in point estimation.

| **Model** | **US dollar / Iran rial** MAE | MSE | **Euro / Iran rial** MAE | MSE | **British pound / Iran rial** MAE | MSE |
|---|---|---|---|---|---|---|
| ***Autoregressive Integrated Moving Average (ARIMA)*** | 33.01% | 67.19% | 38.34% | 55.19% | 20.79% | 67.78% |
| ***Chen's fuzzy time series (first-order)*** | 17.47% | 47.64% | 28.40% | 53.59% | 46.08% | 77.80% |
| ***Chen's fuzzy time series (second-order)*** | 17.47% | 47.64% | 28.40% | 53.59% | 44.09% | 75.35% |
| ***Yu's fuzzy time series*** | 17.47% | 47.64% | 13.58% | 13.47% | 39.68% | 70.16% |
| ***Multilayer perceptrons (MLPs)*** | 10.55% | 40.70% | 19.26% | 29.00% | 13.53% | 34.16% |
| ***Adaptive Neuro-Fuzzy Inference Systems (ANFIS)*** | 0.96% | 25.63% | 14.89% | 22.04% | 19.39% | 64.17% |

## 6. Conclusions

Time series forecasting has been an active research area for the last few decades. Improving forecasting especially time series forecasting accuracy is an important yet often difficult task facing forecasters. Despite the numerous time series models available, the research for improving the effectiveness of forecasting models has been never stopped. Several large-scale forecasting competitions with a large number of commonly used time series forecasting models conclude that combining several models or using hybrid models can be an effective way to improve forecasting performance. Additionally, because of the possible unstable or changing patterns in the data, using the hybrid method can reduce the model uncertainty, which typically occurred in statistical inference and time series forecasting.

In this paper, a hybrid model of the fuzzy autoregressive moving average (FARIMA) models is proposed using the probabilistic neural networks (PNNs), namely FARIMAH, in order to yield more accurate results. The main idea of the proposed model is based on this fact that the distribution of the actual values in the forecasted interval by FARIMA is not uniform. Therefore, in proposed model, a probabilistic neural network is applied in order to determine the spaces of FARIMA interval that probability of existing actual vales in which is higher. Then the spaces that have lower existing probability are deleted from obtained interval by FARIMA, according to the achieved results by the probabilistic neural network. Empirical results of exchange rate forecasting indicate that the proposed model exhibit effectively improved forecasting accuracy, so it can be used as an alternative model to exchange rate forecasting, especially when the scant data made available over a short span of time. In addition, the proposed model based on the opinion of decision maker(s) can provide an interval with minimum width for a given confidence degree or maximum confidence degree for a given width.

## Acknowledgements

*The authors wish to express their gratitude to anonymous reviewers, Dr. G. A. Raissi ardali, and Dr. A. K Tavakoli, Industrial Engineering Department, Isfahan University of Technology, for their insightful and constructive comments, which helped us to improve this paper greatly.*

## References

1. P. Box and G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, (Holden-day Inc, San Francisco, CA, 1976).
2. M. Khashei, M. Bijari and GH A. Raissi, Improvement of Auto-Regressive Integrated Moving Average Models Using Fuzzy Logic and Artificial Neural Networks (ANNs), *Neurocomputing* **72**, (2009) 956- 967.
3. F.M. Tseng, G.H. Tzeng, H.C. Yu and B.J.C. Yuan, Fuzzy ARIMA model for forecasting the foreign exchange market, *Fuzzy Sets and Systems* **118**, (2001) 9-19.
4. M. Khashei and M. Bijari, A novel hybridization of artificial neural networks and ARIMA models for time series forecasting, *Applied Soft Computing*, **11**, (2011) 2664-2675.
5. J. Andres, P. Lorca, F. Juez and F. Sanchez-Lasheras, Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS), *Expert Systems with Applications* **38**, (2011) 1866–1875.
6. Y. Lin and W.G. Cobourn, Fuzzy system models combined with nonlinear regression for daily ground-level ozone predictions, *Atmospheric Environment* **41**, (2007) 3502–3513.
7. P. Chang, C. Liu and Y. Wang, A hybrid model by clustering and evolving fuzzy rules for sales decision supports in printed circuit board industry, *Decision Support Systems* **42**, (2006) 1254–1269.
8. H. Teoh, T. Chen, C. Cheng and H. Chu, A hybrid multi-order fuzzy time series for forecasting stock markets, *Expert Systems with Applications* **36**, (2009) 7888–7897.
9. H. Kim and K. Shin, A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets, *Applied Soft Computing* **7**, (2007) 569–576.
10. M. Khashei, S. R. Hejazi and M. Bijari, A new hybrid artificial neural networks and fuzzy regression model for time series forecasting, *Fuzzy Sets and Systems* 159, (2008) 769 – 786.
11. K. Li and H. Su, Forecasting building energy consumption with hybrid genetic algorithm-hierarchical adaptive network-based fuzzy inference system, *Energy and Buildings* **42**, (2010) 2070–2076.
12. P.F. Pai and C.S. Lin, A hybrid ARIMA and support vector machines model in stock price forecasting, *Omega* **33**, (2005) 497– 505.
13. A. Azadeh, B. Ziaei and M. Moghaddam, A hybrid fuzzy regression-fuzzy cognitive map algorithm for forecasting and optimization of housing market fluctuations, *Expert Systems with Applications* **39**, (2012) 298–315.
14. Y. Huang, S. Horng, M. He, P. Fan, T. Kao, M. Khan, J. Lai and I. Kuo, A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization, *Expert Systems with Applications* **38**, (2011) 8014–8023.

15. V. Yadav and D. Srinivasan, A SOM-based hybrid linear-neural model for short-term load forecasting, *Neurocomputing* **74**, (2011) 2874–2885.
16. L. Yu, S. Wang and K.K. Lai, A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates, *Computers and Operations Research* **32**, (2005) 2523–2541.
17. M. Khashei and M. Bijari, An artificial neural network (p, d, q) model for time series forecasting, *Expert Systems with Applications* **37**, (2010) 479–489.
18. M.R. Amin-Naseri and A.R. Soroush, Combined use of unsupervised and supervised learning for daily peak load forecasting, *Energy Conversion and Management*, **49**, (2008) 1302–1308.
19. G. Lin and M. Wu, A hybrid neural network model for typhoon-rainfall forecasting, *Journal of Hydrology*, **375**, (2009) 450–458.
20. S. Ismail, A. Shabri and R. Samsudin, A hybrid model of self-organizing maps (SOM) and least square support vector machine (LSSVM) for time-series forecasting, *Expert Systems with Applications* **38**, (2011) 10574–10578.
21. M. Khashei and M. Bijari A new class of hybrid models for time series forecasting, *Expert Systems with Applications* **39**, (2012) 4344– 4357.
22. E. Hajizadeh, A. Seifi, M.H. Fazel Zarandi and I.B. Turksen, A hybrid modeling approach for forecasting the volatility of S&P 500 index return, *Expert Systems with Applications* **39**, (2012) 431– 436.
23. H. Ince and T. Trafalis, A hybrid model for exchange rate prediction, *Decision Support Systems* **42**, (2006) 1054–1062.
24. H. Pham, V. Tran and B. S. Yang, A hybrid of nonlinear autoregressive model with exogenous input and autoregressive moving average model for long-term machine state forecasting, *Expert Systems with Applications* **37**, (2010) 3310–3317.
25. M. Shafie-khah, M. Parsa Moghaddam and M.K. Sheikh-El-Eslami, Price forecasting of day-ahead electricity markets using a hybrid forecast method, *Energy Conversion and Management* **52**, (2011) 2165–2169.
26. H. Ishibuchi and H. Tanaka, Interval regression analysis based on mixed 0-1 integer programming problem, *J. Japan Soc. Ind. Eng* **40** (5), (1988) 312-319.
27. T. Masters, *Advanced Algorithms for Neural Networks*, (Wiley, New York, 1995).
28. M. Hajmeer and I. Basheer, Comparison of logistic regression and neural network-based classifiers for bacterial growth, *Food Microbiology* **20**, (2003) 43–55.
29. D. Specht, Probabilistic neural networks, *Neural Networks* **3**, (1990) 109– 118.
30. C. Tsai, On detecting nonlinear patterns in discriminant problems, *Information Sciences* **176**, (2006) 772–798.
31. M. Hajmeer and I. Basheer, A probabilistic neural network approach for modeling and classification of bacterial growth/no-growth data, *Journal of Microbiological Methods* **51**, (2002) 217– 226.
32. E. Parzen, On estimation of a probability density function and mode, *Annals of Mathematical Statistics* **36**, (1962) 1065–1076.
33. T. Cacoullos, Estimation of multivariate density, *Annals of the Institute of Statistical Mathematics* **18** (2), (1966) 179– 189.
34. S.S. Ge, Y. Yang and T.H. Lee, Hand gesture recognition and tracking based on distributed locally linear embedding, *Image and Vision Computing* **26**, (2008) 1607–1620.
35. A. Chen, M. Leung and H. Daouk, Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index, *Computers & Operations Research* **30**, (2003) 901–923.
36. C.X. Xue, X.Y. Zhang, M.C. Liu, Z.D. Hu and B.T. Fan, Study of probabilistic neural networks to classify the active compounds in medicinal plants, *Journal of Pharmaceutical and Biomedical Analysis* **38**, (2005) 497–507.
37. M. Khashei and M. Bijari, A New Hybrid methodology for Nonlinear Time Series Forecasting, *Modelling and Simulation in Engineering* **11**, (2011) 2664-2675.
38. M. Khashei, Forecasting the Isfahan Steel Company production price in Tehran Metals Exchange using Artificial Neural Networks (ANNs), *Master of Science Thesis*, Isfahan University of Technology, 2005.
39. S.M. Chen, Forecasting enrollments based on fuzzy time series, *Fuzzy Sets and Systems*, **81**(3), (1996)311-319.
40. S.M. Chen and N.Y. Chung, Forecasting enrollments using high-order fuzzy time series and genetic algorithms, *International Journal of Intelligence Systems* **21**, (2006) 485–501.
41. H.K. Yu, Weighted fuzzy time-series models for TAIEX forecasting, *Physica A* **349**, (2004) 609–624.
42. J.S.R. Jang, ANFIS: Adaptive-network-based fuzzy inference system, *IEEE Trans Syst, Man, Cybernet* **23**, (1993) 665–85.


Co-published by Atlantis Press and Taylor & Francis