INSTITUTO POLITÉCNICO NACIONAL

ESCUELA SUPERIOR DE ECONOMÍA

SECCIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

STUDIES OF FUZZY TIME SERIES MODELS: AN APPLICATION TO THE FORECAST EXCHANGE RATE, MEXICAN PESO/ AMERICAN DOLLAR

T E S I S

QUE PARA OBTENER EL GRADO DE:

MAESTRA EN CIENCIAS ECONÓMICAS

(ECONOMÍA FINANCIERA)

P R E S E N T A

JOSÉ EDUARDO MEDINA REYES

MÉXICO, CIUDAD DE MÉXICO AGOSTO, 2019

INDEX

.

INDEX OF FIGURES AND TABLES

GLOSSARY

ARIMA: Autoregressive Integrated Moving Average Model.

Artificial Neural Network: An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain.

EGARCH: Exponential Generalized Autoregressive Conditional Heteroskedasticity Model.

FTS-Fuzzy ARIMA: Hybrid model between Fuzzy Time Series and Fuzzy ARIMA.

Fuzzy AR: Fuzzy Linear Regression with Gaussian or T-Student parameters.

Fuzzy ARIMA: Fuzzy Linear Regression applied to the ARIMA model.

Fuzzy Categories: The use of set membership as a key component of category theory can be generalized to fuzzy sets.

Fuzzy EGARCH: EGARCH with Gaussian parameters.

Fuzzy GARCH: GARCH with Gaussian parameters.

Fuzzy Linear Regression: Model of Linear Regression with possibility interval.

Fuzzy Logic: This is an extension of the case of multi-valued logic. And let determined degrees of membership.

Fuzzy Set: This is somewhat like sets whose elements have degrees of membership.

Fuzzy Theory: This theory is used in a wide range of domains in which information is incomplete or imprecise

Fuzzy Time Series: This is a model that lets us generate a forecast of time series assuming that it is a fuzzy set.

Fuzzy Trapezoidal NARNET: Hybrid model between NARNET and Fuzzy Time Series with Fuzzy Trapezoidal Set.

Fuzzy Triangular NARNET: Hybrid model between NARNET and Fuzzy Time Series with Fuzzy Triangular Set.

GARCH: Generalized Autoregressive Conditional Heteroskedasticity.

In-Sample test: Evaluation of the models when the model estimates the real values.

MAD: Mean Absolute Deviation.

Membership Function: A fuzzy number is a convex, normalized fuzzy set of real numbers whose membership function is at least segmentally continuous and has the functional value at least one element.

NARNET: Non-linear Autoregressive Neural Network.

Out-Sample test: Evaluation of the models when the method does not estimate the real values, if not use the parameters obtained in the sample test.

PARCH: Parsimonious Autoregressive Conditional Heteroskedasticity.

ABSTRACT

This research presents the comparison of the techniques modified using fuzzy logic called FTS-Fuzzy ARIMA, Fuzzy AR, Fuzzy GARCH, Fuzzy EGARCH, Fuzzy Triangular NARNET, and Fuzzy Trapezoidal NARNET, with respect to the volatility prediction techniques ARIMA, GARCH, EGARCH, and Nonlinear Autoregressive Neural Network, For this purpose, it applies to the time series of the foreign exchange market to forecast the foreign currency exchange rate, MX pesos against the US Dollar. In this work, the growth rate of the data from the time series is used in a daily format from January 2008 to December 2017; to perform the sample test is used in January 2018. We found that the models based on fuzzy theory have better estimates of volatility in financial time series. This allows developing new prediction methods based on the structure of fuzzy logic, it is also essential to establish an analysis of a greater number of data to try to discriminate if the effect of the error effect decreases.

RESUMEN

La presente investigación desarrolla una comparación de las técnicas modificadas con lógica difusa conocidas como FTS-Fuzzy ARIMA, Fuzzy AR, Fuzzy GARCH, Fuzzy EGARCH, Fuzzy triangular NARNET y Fuzzy trapezoidal NARNET con respecto a los métodos ARIMA, GARCH, EGARCH y Non-linear Autoregressive Neural Network. Para este propósito, se aplican al pronóstico del tipo de cambio, peso mexicano contra dólar americano. Este trabajo utiliza la tasa de crecimiento de la serie de tiempo asociadas al tipo de cambio con periodicidad diaria desde enero de 2008 a diciembre de 2017 y se anexa enero de 2018 para evaluar los modelos fuera de muestra. Los resultados muestran que los modelos FTS-Fuzzy ARIMA, Fuzzy GARCH, Fuzzy EGARCH, Fuzzy triangular NARNET y Fuzzy trapezoidal NARNET estiman mejor el comportamiento de la volatilidad de las variables del mercado cambiario en comparación con las técnicas tradicionales.

INTRODUCTION

This research presents the comparison of the techniques modified using fuzzy logic called FTS-Fuzzy ARIMA, Fuzzy AR, Fuzzy GARCH, Fuzzy EGARCH, Fuzzy Triangular NARNET, and Fuzzy Trapezoidal NARNET, with respect to the volatility prediction techniques ARIMA, GARCH, EGARCH, and Nonlinear Autoregressive Neural Network.

The increase in volatility in the markets causes that the forecast of the foreign exchange market variables be more and more complicated and at the same time are essential for taking decisions. Then, the models that try to predict these variables need more level of specialty. In this way, exist a number of analysis tools that allow understand and forecast this series, however, they have not been able to give accurate estimates.

The object of this investigation is to develop an analysis of econometric models based on fuzzy logic and its application to forecasts the exchange rate and identify if these methods generate better estimations and forecast than other techniques.

Therefore, the following questions arise:

1. Why and what is the best model for the forecast of the exchange rate under the concepts of fuzzy theory, fuzzy time series, and neural networks?
2. What are the advantages of fuzzy models to predict the exchange rate?
3. Why do Fuzzy Time Series Models capture better the changes and abnormalities of the exchange rate than the ARMA, Conditional Variance and Autoregressive Neural Networks models?

The hypothesis is that the models of the Fuzzy Time Series Theory combined with Neural Networks generate a better forecast of the behavior of the exchange rate compared with the ARIMA, EGARCH, PARCH and Autoregressive Neural Network models.

For this purpose, it applies to the time series of the foreign exchange market to forecast the foreign currency exchange rate, MX pesos against the US Dollar. In this work the growth rate of the data from the time series is used in a daily format from January 2008 to December 2017; to perform the sample test is used in January 2018.

Primarily, the existing tools can be classified into four groups García *et al*. (2013). The first one is the common methods, the Box-Jenkins methodology, as AR, MA, ARMA, ARIMA, ARFIMA, GARCH and other related; These assume that future value is a linear combination of the past values and its errors. The study of García *et al*. (2013) exhibited that these technics cannot capture the non-linear of the financial time series, and this makes it necessary to combine these processes with others to generate a better forecast.

The second category is the non-linear models such as Neuronal Networks, Candlestick Chart Analysis Expert Systems, Machine Learning, Fuzzy Time Series and other related. Zemke (1999) showed that these methods have until 64% of effectiveness being the Neuronal Networks, it shows better estimate results given the non-linearity of the financial time series. In support of this research other works of Neuronal Networks can be seen in Gómez-Ramos & Venegas-Martínez (2013) and Gómez-Ramos *et al*. (2016).

The third group is the hybrid models. The concept of hybrid refers to the combine various elements of different techniques. In this category, there are methods such as hybrid k-means and nonlinear autoregressive neural network models, hybrid models based on rough sets theory and genetic algorithms, A hybrid fuzzy time series and particle swarm optimization model and other related (Singh, 2017). These models have better forecast that the models of the first and second category García *et al*. (2013).

And finally, the fourth groups have models that use various variables. These models take a linear combination of the past values and their errors, and it generates an equation system for forecast future value, for example, VAR model, VEC model and other related Maysami *et al*. (2004). These models are not used for the present analysis because only we work with a single variable to make the forecast.

The methodology of this research is to generate through Fuzzy Time Series models and the combination with other models; new techniques for make forecasts and develops a comparison with ARIMA, EGARCH, PARCH and Autoregressive Neural Network. As well as, to identify which is the model that best estimates the behavior of the exchange rate.

Fuzzy Logic has its origins in the works of Zadeh (1965), who started from the fact to define the concept of the belong on certain sets, called the study of fuzzy sets. An important aspect developed by this author is the assignment levels of membership to an element x in a set A, this is known as membership function; which associates each point in x with a value in the interval $[0, 1]$ of the real numbers.

Furthermore, this study constitutes a generalization of classical logic, since if x is defined in A; it can only take the value of 0 or 1 (belong or not to the set). According to Zadeh (1965), a fuzzy set is a class of objects with a *continuum* that shows their degree of membership. On the other hand, Fuzzy logic is a multi-valued form of logic, it can take approximate reasoning and linguistic variables are used in the definition of sets (Coyaso *et al*., 2015).

Tanaka *et al*. (1982), for the first time, developed a description of this topic apply to linear regression. The main result of this investigation was the construction, from the Fuzzy Linear Function and the Linear Regression Model, of a model of the econometric analysis called the Fuzzy Linear Regression Model. This methodology is the basis for the development of various similar models about fuzzy regression, how Kim *et al*. (1998), Tanaka (1987), Chang (1997), Cheng *et al*. (1999a), Cheng *et al*. (1999b), Özelkan *et al*. (2000), Dunyak *et al*. (2000).

Another application of fuzzy logic was developed by Song and Chissom (1993a), they studied of fuzzy sets and their applications to decision making and essentially dynamic processes, in which observations are linguistic values. These are called Fuzzy Time Series, defining two types of series: time-invariant and time-variant.

Likewise, Song and Chissom (1993b) under the previous definition developed an application of this time-invariant methodology to model the enrollments of the University of Alabama, compared it with other existing methods for forecast this series. So, the advantage of this methodology is that it works in competitive scenarios; other works that start from the contributions of Song and Chissom are Chen (1996), Chen *et al*. (2004), Yu (2005a), Yu (2005b), Huarng (2005).

Tseng *et al*. (2001) from the ARIMA time series model and the fuzzy regression model developed a new methodology, called Fuzzy-ARIMA model. This model is applied to estimate the foreign exchange rate of Thailand Dollar to the American dollar. The importance of this method is that provide to the decision-makers the best and worst possible situations.

Popov *et al*. (2005) explained as to integrate the characteristics of Fuzzy Theory in the GARCH Models, it generated the model "Fuzzy generalization of Autoregressive Conditional Heteroscedasticity" (Fuzzy GARCH). They found that this technique generates better results than the basic GARCH model, especially to measure volatility in Financial Time Series. Hung (2009) developed a model similar called "Fuzzy Threshold Asymmetric Generalized Autoregressive Conditional Heteroskedasticity Model" and it found that the financial series studied with these model indicate that the propagation of volatility in financial markets is non-linear. Although, the method proposed by Hung is still very limited to estimate high volatility behaviors.

Tsaur (2012) made a method of fuzzy time series with Markov Chain, used to forecast the values through transferring fuzzy time-series data to the fuzzy logic group and use the obtained fuzzy logic group to derive a Markov chain in the transition matrix.

Other ways to analyze the Fuzzy Time Series were reviewed by Singh (2017), which exposed the main characteristics of the investigations are the determination of the length of intervals; the establishment of fuzzy logical relationships between different factors; and defuzzification in a Hybridize modeling. Also, he developed a review of the works that have been established with this methodology.

Dash *et al.* (2016) developed a new methodology, which takes the volatility of financial series as a Fuzzy process. In the EGARCH model, its structure of the variance equation is evaluated as a Gaussian membership function and with a Neural Network that learns of the Gaussian volatility let generates the fuzzy rules to forecast the GARCH process. It found that this hybrid model improves significantly with respect to models of less complexity.

Pal *et al*. (2017) made a model that used the Genetic Algorithm to decide the size, number of intervals and fuzzy relationships in a Fuzzy Time Series. This work improved the forecast with respect to previous investigations of this topic.

Namugaya *et al.* (2019) developed a comparison of TSK Fuzzy-GARCH and GARCH for forecast the Uganda Securities Exchange. This research used the ANFIS model to generate the fuzzy rules of the TSK Fuzzy-GARCH based in Hung (2009). The result is that the fuzzy model generates the best forecast of the compared methods. Other techniques of fuzzy theory applied to forecast time series can be seen in (Aliyev *et al.*, 2019), (Zhang *et al.*, 2019), (Vovan, 2019), (Chen *et al.*, 2019), (S.Atsalakis *et al.*, 2019), (Kristjanpoller *et al.*, 2018), (Lepskiy *et al.*, 2017). Continuing with the objective of this investigation, the literary revision allows identify the Fuzzy Time Series models as follows:

1. In which the time series is Fuzzy Time Series.
2. Hibrized 1: Linear and no-linear models with Fuzzy parameters.
3. Hibrized 2: Models of the hybrid 1 combined with the first category and other methods[1].

The chapter I analyze the Fuzzy Theory, Fuzzy Time Series and Neural Networks designed to generate the forecast of the exchange rate. First, it recognizes the theoretical framework of the exchange rate, mainly identifying the Purchasing Power Parity hypothesis. Moreover, basic notions of fuzzy set and artificial neural networks to make predictions are studies.

In Chapter II are estimates and evaluates the ARIMA, EGARCH, PARCH and Autoregressive Neural Network models. And its forecasts are used in the third section to realize a comparative with the Fuzzy models. The main characteristics of section II are that the models (ARIMA, EGARCH, PARCH, and NARNET) were developed with their theoretical frame and its application to estimate the exchange rate. Subsequently, the efficiency of these methods is compared by the Absolute Mean Error.

[1] Neuronal Networks, Candlestick Chart Analysis Expert Systems, Machine Learning, Markov Chain, etc.

Finally, chapter III comprehends the Fuzzy models that apply to the exchange rate and the evaluation of fuzzy models versus traditional techniques. The importance of this research is that develop the comparison of different types of models to forecast the exchange rate (Mexican Peso/ American Dollar) and it stands out that the methods of chapter III are own suggestion from the concepts revised in chapter I. These models were generated and design by the authors of the present research; the calculations were in Excel and MatLab.

CHAPTER I. EXCHANGE RATE AND FUZZY THEORY

"Understand the world is like understand oneself, an endless structure of processes that it only allows us to intuit what their path is"

In chapter I the analyze of the concepts and theoretical aspects related to this research is developed. First, it shows why the exchange rate has high volatility in its behavior. In the second instance, the key concepts to understand the fuzzy theory and neural networks are described. Therefore, we study the models that use Fuzzy Theory, Fuzzy Time Series and Neural Networks to generate the forecast of the exchange rate Mexican Peso against American Dollar. The objective of this section is to identify if those models are possible to answer the research questions and thereby corroborate or refute the hypothesis.

The objective of this chapter is to understand and define the main concepts of this research. Firstly, it recognizes the theoretical framework of the exchange rate, mainly identifying the Purchasing Power Parity hypothesis. Moreover, the principal notions of fuzzy set theory and its applications in the study of time series. But also, identify the properties of artificial neural networks to make predictions.

In addition, formalize models that allow modeling and adequately the forecast of the exchange rate and other variables of great relevance in the Mexican economy. Therefore, each concept must be defined beforehand, the Exchange Rate is understood as:

> *"Price in that a currency is exchanged for another, like gold or special drawing rights. These transactions are carried out in cash or in the future (spot market and futures market) in the currency markets, the price of one currency in terms of another. It is usually expressed in terms of the number of units of the national currency that must be delivered in exchange for a unit of foreign currency. The equivalence of the Mexican Peso, respect to foreign currency "INEGI (1998), pp.723.*

The value of the exchange rate is affected by economic, political and social conditions, economic growth of the country, the level of technology, the structure of the business organization, the competitiveness of the economy abroad, international trade, level of economic dependence on another economy and the financial market.

On the other hand, the political conditions of the country, such as the exchange rate regime, foreign trade policy, monetary policy, fiscal policy, government regime, economic model, etc. When properly the country's governors determine what the conditions under which the economy will be governed, the economic variables will be determined, among them the exchange rate.

Ultimately, the social conditions of an economy affect its behavior. Therefore, it is understood that there are social movements the impact of these is captured in the prices of the goods of the economy (the exchange rate is a price) and these fluctuate depending on the influence of that.

It is important to recognize the existence of the main situations that can impact on the determination of the exchange rate value. Although for the purposes of this study, it is not a priority to develop the effect of each one for the exchange rate, theoretically to formalize this variable in order to analyze the specific model for its forecast. For the purposes of this research, it is assumed that these impacts are recognized in history and it is enough to understand the dynamics of their behavior.

The study of exchange rate according to Corbae *et al.* (1988) is determined from their behavior either theoretically or empirically; they are built under the assumption of "*Purchasing Power Parity (PPP)"*. In this research, a series of stationarity and cointegration tests are used to corroborate the *PPP* hypothesis.

1.1. Theoretical Specification of the Exchange Rate.

The theoretical study of the exchange rate, begin since the *"Purchasing Power Parity",* it is understood as:

> *"The real parity in each moment between two nations is represented by the coefficient of purchasing power of one country and that of another" (Cassel, 1918) pp. 413*

It is understood that when this coefficient is equal to the unit, the different currencies of each country must have the same purchasing power and with it the capacity to acquire an equal basket of goods with exactitude. From Reinert *et al*. (2009), it is known that international arbitration would have to cause adjustments in prices; which ultimately should restore parity. In other words, it is understood that the condition of parity is that the exchange rate between two currencies must be equal to the quotient of the price level of both countries.

However, it is not only a matter of determining the concept but also the importance of it. Bearing in mind that the use of PPP has become an important reason for academic study, it should be specified that the term exchange rate matches the price level; so their behaviors according to this theory must be correlated.

1.1.1. Power Parity of Purchase.

As already mentioned in the previous section, the PPP describes the relationship between the exchange rate and the price level of the two economies. This is usually used as an indicator to assess the imbalances of the exchange rate, with respect to what the theory mentions.

Therefore, the existence of two aspects of this analysis is recognized as the Strong version and the weak hypothesis. Therefore, the strong PPP is denominated in the expression of foreign and national currency. On the other hand, weak PPP is understood as variations in price levels (Murias, 1998).

1.1.1.1. Strong Version.

This hypothesis of strong purchasing power parity according to Alessandria *et al.* (2011), part of a state in which the same baskets of goods should be able to be bought or sold at the same price in any place. This allows recognized the absolute PPP as;

$$s = p - p^*$$

Where s is the exchange rate how the domestic price of the foreign currency, p is the price index of the national economy, and p^* shows the corresponding price index of the foreign country. It is identified that the exchange rate according to this theory is determined by the difference in prices of both economies and this is what guarantees the balance in the transactions of the two economies.

Rivero (2011) shows that as greater is p, the greater must be s. In other words, a relatively depreciated exchange rate would be required to maintain the purchasing power of the national currency. On the other hand, as smaller is p relative to p^*, the smaller it must be s, requiring a relatively appreciated exchange rate to restore the purchasing power parity.

1.1.1.2. Weak Version.

In contrast to the previous section, the weak purchasing power parity admits that the exchange rate has a deviation from relative prices by a constant; that shows the existence of barriers to international trade or the cost of transporting merchandise so that certain restrictive situations such as those mentioned are reflected in this hypothesis (Rivero, 2011).

In this way, changes in the exchange rate would have to be compensated for by changes in the price indices of the national and foreign economies. So the weak version of the PPP can be identified from:

$$s = c + p - p^*$$

Where c is a constant that is added to the equation of the strong PPP, this hypothesis of the theory gives the purchasing power parity is less restrictive, as a result, prices are also reflected in s. It is said that the requirement is that the exchange rate of changes in the exchange rate is the same as the proportion of the dissimilarity in relative prices (Reinert *et al.*, 2009).

In this way, the weak version of the PPP mentions that if the variations of the national prices are greater than those of the foreigner; there will be an increase in the exchange rate in order to continue the purchasing power of the domestic currency. In contrast, if the changes in domestic prices are lower than foreign ones; the exchange rate must be reduced to preserve the purchasing power of the national currency.

1.1.1.3. Law of One Price.

Now, from the previous sections, it is recognized that the exchange rate is the regulatory component so that nations can carry out exchanges. Therefore, it is considered that the exchange is fair because the price is recognized as the balance of the international supply and demand of the good.

The Law of the single price that according to Reinert *et al.* (2009) is understood as the state in which the price of an identical good in different countries is equal after adjustment of the exchange rate.

$$P_i = sP_i^f$$

Where the equation P_i represents the price of good i in the domestic economy and P_i^f the price of good i in the economy and s the market exchange rate. It is understood that the exchange rate will vary depending on the movement of domestic and foreign prices.

This section highlights the great importance of the exchange rate. Since it allows reach a balance of exchange between bidders and demanders of different nations, thus achieving formalize a price to make transactions.

It should be understood that the PPP and the single price law are introduced to identify how theoretically the behavior of the exchange rate is. This is for the purpose of establishing where volatility and strong disturbances of this variable come from.

1.2. Fuzzy Theory, Sets, and Membership Function.

This section aims to develop the main concepts of fuzzy theory that impact the study of fuzzy time series. As well as, the incorporation of aspects is considered relevant for the present study. Returning to the concepts of Zadeh (1965) can be studied and address several definitions that allow developing in this analysis. Therefore, we began the theoretical study with a series of definitions.

Definition 1. Let X a space of points, with a generic element of X denoted by x. So, X = x. A fuzzy set A in X is characterized by a membership function $\mu_A(x)$, which associates each element of X with one and only one element of the real numbers in the interval $[0,1]$. Where, the value of $\mu_A(x)$ evaluated in x represents the degree of belonging of x in A.

$$A = \frac{\int_i^{\infty} \mu_A(x_i)}{x_i}$$

From definition 1, it is identified that two types of logic are developed from these types of sets. The first is when the x element can only take values from the edge of the interval. What tells us only, if it is a member in the case of $x = 1$ or if it is not a member in

$x = 0$, known as the classical logic. On the other hand, when $\mu_A(x)$ is allowed to associate values throughout the interval $[0,1]$, this is called how fuzzy logic.

Arise two key concepts for our study; the first is the fuzzy set, which implies the fuzziness of membership from x to A; and the second is the membership function, which tells us the level of belonging from x to A.

In relation to fuzzy sets, it proposes is to solve the problem of the ambiguity of various topics of human life. This defines the membership of certain situations within the object or fact in question, such as linguistic values.

1.2.1. Fuzzy Sets.

In order to better understand fuzzy sets; the concepts, operations, and derivations of this idea must be considered. Therefore, Zadeh (1965) defines the following:

Definition 2. Let A′ be the complement of a fuzzy set A and it is defined as:

$$\mu_{A'} = 1 - \mu_A$$

See figure 1 [c] and [f].

Definition 3. It is understood that the union of two fuzzy sets A and B with respect to the membership function $\mu_A(x)$ and $\mu_B(x)$ generates a third fuzzy set C, denoted as $C = A \cup B$ and its membership function is understood as follows:

$$\mu_C(x) = Max[\mu_A(x), \mu_B(x)], \quad x \in X$$

See the figure 1 [a] and [d].

Definition 4. The intersection of two fuzzy sets with respect to the membership function $\mu_A(x)$ and $\mu_B(x)$ generates a third fuzzy set C, such that $C = A \cap B$ and its functional form is determined by:

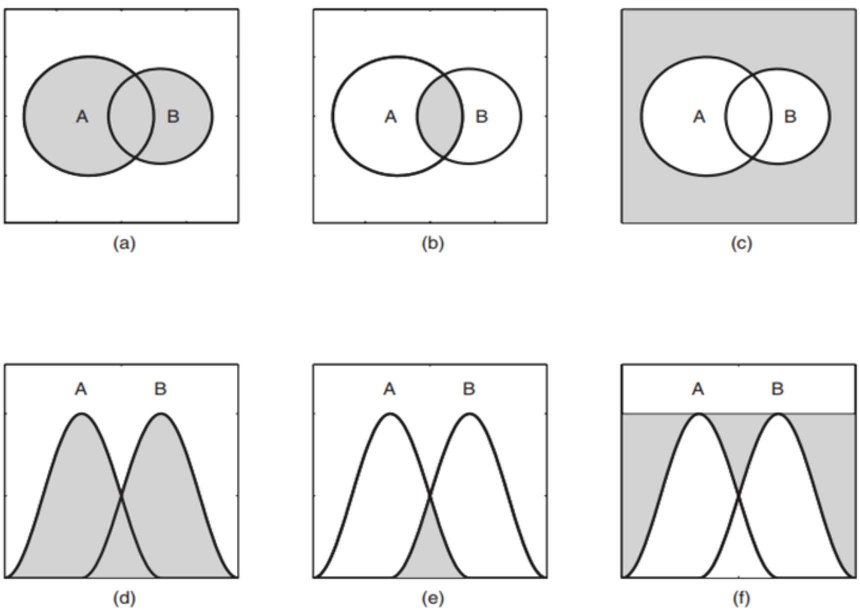
$$\mu_C(x) = Min[\mu_A(x), \mu_B(x)], \quad x \in X$$

See the figure 1 [b] and [d].

With respect to the three definitions mentioned above, it is recognized that from **definition 2**, the notion of fuzzy set and its elements start from the ambiguity of its membership.

Then the design of union of these sets, it is established that given the fuzziness in the union of these sets is considered that a membership function that manages to unite the functions creating another membership function defined in the union space. On the other hand, the perception of the intersection of fuzzy sets is specified in the third generated set; it is obtained by minimizing the membership functions.

Figure 1. Operations with fuzzy sets



Source: obtained from Jantzen (2007) pp.21.

Continuing with the analysis, it is said that from Terano *et al.* (1992) the main operations are developed from the following definitions:

Definition 5. Let A and B are two fuzzy sets, such that their algebraic sum is understood as:

$$A \boxplus B = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$$

Definition 6. Let A and B are two fuzzy sets, such that their algebraic product is conceived from:

$$A \cdot B \leftrightarrow \mu_{A \cdot B}(x) = \mu_A(x)\mu_B(x)$$

Definition 7. It is known as a fuzzy subset A in X, in its discrete form is:

$$A = \sum_{i=1}^{n} \mu_A\,(x_i)/x_i$$

And in its continuous expression:

$$A = \int_{i=1}^{n} \mu_A\,(x_i)/x_i$$

Definition 8. Let a normal fuzzy set be that which has in its membership function a value associated with 1. In such a way that:

$$\max_{x \in X} \mu_A(x) = 1$$

From this definition, it is intended to formalize various situations that are of interest for the present investigation. Like the construction of fuzzy modeling for time series.

1.2.2. Membership Function.

There are several membership functions that allow modeling different behaviors; these can include the whole fuzzy set or part of it. These are intended to formalize the most relevant for this research. For this, it is proposed to define each of them from Rutkowski (2004).

a) Singleton membership function.

Singleton is a membership function that takes the value of 1 only at one point $\bar{x}$ of the universe X and 0 in any other case.

$$\mu_{sing}(x) = \begin{cases} 1 & para\ x = \bar{x} \\ 0 & para\ x \neq \bar{x} \end{cases}$$

See the figure 2 [a].

b) Gaussian membership function.

A Gaussian membership function has two parameters. The first is $\bar{x}$ which represents the center of the function. And σ is the width of the function.

$$\mu_{gauss}(x) = e^{-\left(\frac{x-\bar{x}}{\sigma}\right)^2}$$

See the figure 2 [b].

c) Generalized Bell membership function.

A generalized Bell membership function is expressed by three parameters. And a measures the width of the function; c is the center and b is the slope of the function.

$$\mu_{bell}(x) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}}$$

See the figure 2 [c].

d) Sigmoid membership function.

A Sigmoid membership function is formed in two parameters, a is the slope of the function at the crossing point of $x = c$.

$$\mu_{sigm}(x) = \frac{1}{1 + e^{-a(x-c)}} \qquad (1.1)$$

See the figure 2 [d].

e) Triangular membership function.

A Triangular membership function is described by three parameters where a and c are the extremes of the function and b is the center, such that $a > b > c$.

$$\mu_{tri}(x)\begin{cases} 0 & x \leq a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ \dfrac{c-x}{c-b} & b \leq x \leq c \\ 0 & c \leq x \end{cases}$$

Another way to express the triangular function is.

$$\mu_{tri}(x) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right)$$

See the figure 2 [e].

f) Trapezoidal membership function.

A trapezoidal membership function is comprised of four parameters $\{a, b, c, d\}, a > b > c > d$.

$$\mu_{trap}(x)\begin{cases} 0 & x \leq a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ 1 & b \leq x \leq c \\ \dfrac{d-x}{d-c} & c \leq x \leq d \\ 0 & d \leq x \end{cases}$$

Another way to express the trapezoidal function is.

$$\mu_{trap}(x) = \max\left(\min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right)$$

See the figure 2 [f].

g) Z Membership function.

The Z membership function is expressed by two parameters a and b found at the ends of the slope of the curve.

$$\mu_z(x)\begin{cases} 1 & x \le a \\ 1-2\left(\frac{x-a}{b-a}\right)^2 & a \le x \le \frac{a+b}{2} \\ 2\left(\frac{x-b}{b-a}\right)^2 & \frac{a+b}{2} \le x \le b \\ 0 & b \le x \end{cases}$$

See the figure 2 [g].

h) Membership function Π.

A Π membership function is characterized by four parameters of which the two a and d are at the extremes; b and c at the breaking point of the curve $\{a, b, c, d\}, a > \ b > \ c > \ d$.

$$\mu_{pi}(x)\begin{cases} 0 & x \le a \\ 2\left(\frac{x-a}{b-a}\right)^2 & a \le x \le \frac{a+b}{2} \\ 1-2\left(\frac{x-b}{b-a}\right)^2 & \frac{a+b}{2} \le x \le b \\ 1 & b \le x \le c \\ 1-2\left(\frac{x-c}{d-c}\right)^2 & c \le x \le \frac{c+d}{2} \\ 2\left(\frac{x-d}{d-c}\right)^2 & \frac{c+d}{2} \le x \le d \\ 0 & x \ge d \end{cases}$$

See the figure 2 [h].

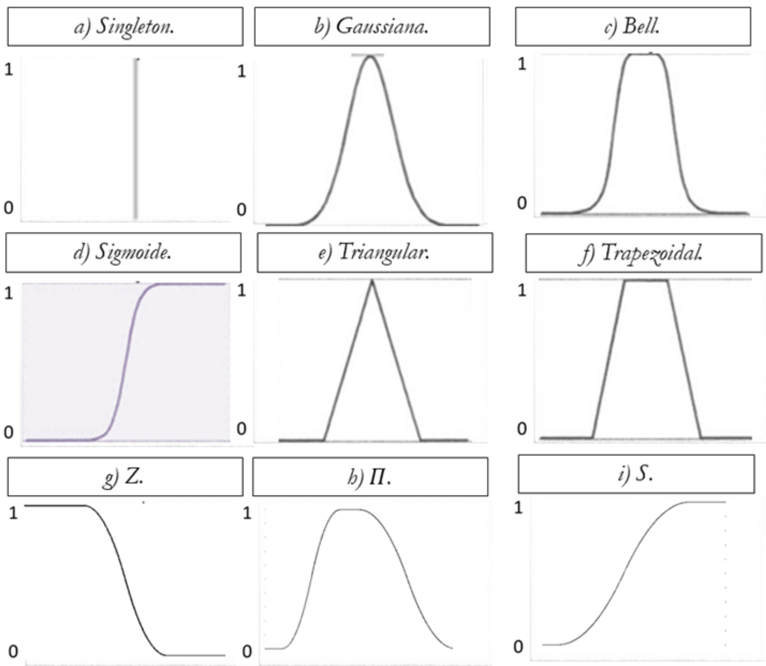
i) S Membership function.

The S Membership Function is represented by two parameters a and b located at the ends of the slope.

$$\mu_s(x)\begin{cases} 0 & x \leq a \\ 2\left(\frac{x-a}{b-a}\right)^2 & a \leq x \leq \frac{a+b}{2} \\ 1-2\left(\frac{x-b}{b-a}\right)^2 & \frac{a+b}{2} \leq x \leq b \\ 1 & b \leq x \end{cases}$$

See the figure 2 [i].

Figure 2. Various Membership Functions



Source: Own elaboration.

1.3. Neural Networks, Theory, and Models.

Artificial neural networks are a model that uses computational techniques to examine the information and develop modeling of it. Therefore, you can find patterns of interest and dependency structures in the data. When the neural network has learned the patterns and dependencies in the series, it can be used to generate various output data.

According to Smith (2000), networks can be used to learn to predict future events; classify information not accurate and understand volatility clusters. There are other applications but with respect to this research, we are interested in the three functions mentioned. Given that, it is intended to forecast information in environments of uncertainty and identify clusters of volatility in the market.

For this reason, the characteristics of this study methodology are formalized. Starting from the fact that this can be combined with the fuzzy theory to verify if together they allow improving the modeling of the exchange rate and its forecast. The objective of this section is to collect the relevant aspects of neural networks to give an overview of this theory and its implications.

The use of neural networks to predict or forecast the study variable from the existing information set of it and a series of explanatory variables. The main characteristic of this model is that it contains one or more hidden layers in which the study variable is transformed by the activation function. This represents an efficient method for modeling non-linear statistical processes.

1.3.1. Feed Forward.

This type of neural network feedforward is based on the structure of a Perceptron multilayer. The characteristic that identifies these models is that the interconnections of the different layers are unidirectional, which means that the information will never pass through the same neuron.

The intermediate segments are not related to external factors and are considered as hidden layers. Each Perceptron is a stage that only connected to the next step, so the neurons of the same phase are not connected.

López (2014), this type of network is characterized by creating a set of neurons that receive multivariate information, processes it and provides output information that can be multivariate. In relation to the information transaction between neurons is established by connection weights, specified as.

$$\sum_j \omega_{ij} x_j$$

Where ω_{ij} represent the weights of the neurons x_j to the receiver. To complete the process of this network, a threshold is attached to know if the stimulus has relevance or not. Therefore, unless the sum exceeds the determined barrier, the output of said information will be considered. So it can be re-expressed.

$$\sum_j \omega_{ij} x_j - m_i$$

As an annex, it is mentioned that the structure of this system is of the linear type, in which only the input layers can be observed and the information that will be used to predict together the output values given by the weights in the input neurons. The connections between the input variables are commonly known as "input neurons", the neurons in the hidden layers connected to the output phase are known as "synapses".

These types of artificial neural networks are commonly applied in economics and finance. In general, this network represents the way in which the human brain processes input information in a set of neurons that provide output information (McNelis, 2005).

The way of learning these belongs to the so-called "supervised learning" so that for pairs of input and output values are fed by several cycles. As such, the network learns from a relationship between input and output data.

1.3.2. Hopfield.

The Hopfield neuronal network is a type of network known as recurrent with a single layer of neurons. The activation function can be a continuous or discrete function; therefore, this network is classified as continuous and discrete.

The continuous "Hopfield" network consists of a set of strongly connected neurons. The number of feedback loops is equal to the number of neurons. The synaptic weights w_{ij} between the neuron i and the j are symmetric, this means that $w_{ij} = w_{ji}$ for all i, j. In addition, this type of network can be viewed from a dynamic perspective.

To understand the dynamic neural network Zhang (2018) departs as follows. Let $v_i(t)$ be the induced local field that assumes the activation function of the continuous "Hopfield" network of the hyperbolic tangent type, such that.

$$x = \varphi_i(v) = tanh(\frac{a_i v}{2}) = \frac{1 - e^{(-a_i v)}}{1 + e^{(-a_i v)}} \qquad a_i \in \mathbb{Z}_+$$

Where a_i refers to the gain of the neuron i. The slope of the activation function that originates in $\frac{a_i}{2}$, is given by.

$$x + x\, e^{(-a_i v)} = 1 - e^{(-a_i v)}$$

The "Hopfield" neural network continues from N interconnected neurons the dynamics of this network is defined by the following differential equation:

$$C_j \frac{d}{dt} v_j(t) + \frac{v_j(t)}{R_j} = \sum_{i=1}^{N} \omega_{ij} \varphi_i(v_i(t)) + I_j \qquad (j = 1, \dots, N)$$

Where C_j is the propagation capacity, R_j is the resistance to diffusion, ω_{ij} are symmetric synaptic conductors or weights, and I_j is the bias applied externally. The previous thing, it is necessary to add that within the study of these models its dynamics is motivated by the minimum of the function of effort that is denoted of the following way:

$$E = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_{ij} x_i x_j + \sum_{j=1}^{N} \frac{1}{R_j} \int_0^{x_i} \varphi_i^{-1}(x)\, dx - \sum_{j=1}^{N} I_j x_i$$

Where x_i is the output of the *j-th* neuron, $\omega_{ij} = \omega_{ji}$, I_j is the applied external bias, and R_j is the resistance to diffusion, the functions are defined as.

$$x_i = \varphi_i(v_i(t)) = \frac{1 - e^{(-a_i v)}}{1 + e^{(-a_i v)}}$$

And

$$\varphi_i^{-1}(x) = -\frac{1}{a_i} log(\frac{1-x}{1+x})$$

The observations that can be made to this neural network, the main one is that it is asymptotically stable in the sense of "Liapunov" and the fixed points of this model are the minimum of the energy function and on the contrary, it must also be fulfilled.

On the other hand, the discrete "Hopfield" neuronal network consists of the interconnection of a group of neurons with synchronization correspondence of only one unit of time. The amount of retro feeders equals the number of neurons. The synaptic weights ω_{ij} between the neurons i and j are symmetric, this means that ω_{ij} =ω_{ji}. In addition, there is no self-feedback of the network, ω_{ij} = 0, this means that the outputs of the neurons are synchronized per unit time with other neurons in the network.

The energy function of this network is developed from the interconnection of N neurons; the states of the network are defined as x $= (x_1, \dots, x_n)^T$. The synaptic weights are represented by the matrix $W = (\omega_{ij})_{i,j=1,\cdots,N.}$. Which is symmetric, $\omega_{ij} = \omega_{ji}$ and $\omega_{ij} = 0$ the induced local field v_i of the neuron i is:

$$v_i = \sum_{j=1}^{N} \omega_{ij} x_j + b_i, \qquad (i = 1, \dots, N)$$

Where b_i is the external bias applied to the neuron i and the state of the neuron x_j is denoted by:

$$x_i = sign(v_i) = \begin{cases} 1 & si\ v_i > 0 \\ -1 & si\ v_i < 0 \end{cases}$$

In the case that $v_i = 0$, it is most convenient for the neuron i to return to its previous state. The energy function of the discrete network is defined as:

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \omega_{ij} x_i x_j - \sum_{i=1}^{N} b_i x_i$$

This function varies with the states of the neurons in the network which grow from $-1\ to\ 1$, in such a way that the energy of the network decreases.

1.3.3. Backpropagation Algorithm.

When recognizing the classifications of the neural networks and the characteristics of the main ones, it is also necessary to develop the algorithm that they follow to feed the network with information. Therefore, it is suggested by Leung *et al.* (1991), that the backpropagation algorithm is one of the most common learning algorithms in artificial neural networks. This method develops an approximation to the overall minimization of the stationary phase method. This algorithm can be seen as a generalization of the algorithm adapted "least mean squares", which differentiates them, is that the last-mentioned method is proposed from adapting linear combinations and the Backpropagation.

This algorithm minimizes the cost function by the recursive alteration of the coefficients $\{\omega_{ij}^{(l)}, \theta_i^{(l)}\}$ based on the gradient search techniques. Thus, when finding the gradient vector of the cost function, the logarithm of Backpropagation is derived. In Smith (2000), the phases that this algorithm regularly follows when applied to a "multilayer perceptron" are considered. It is posed as follows:

I. The pattern of inputs x is randomly selected and entered through the network in the input stage.

The network of input and output of hidden stages in the network is calculated.

$$net_j^h = \sum_{i=1}^{N+1} \omega_{ij} x_i \, ; \qquad y_j = f(net_j^h)$$

II. Calculate the network of input and output of the final stage of neurons.

$$net_j^0 = \sum_{j=1}^{J+1} v_{kj} y_j \, ; \qquad z_k = f(net_k^0)$$

III. Update the weights in the exit stages (for all *k, j* pair).

$$v_{kj} \leftarrow \omega_{kj} + c\lambda(d_k - z_k) z_k (1 - z_k) y_j$$

IV. Update the weights in the hidden stages (for all *i, j* pair).

$$\omega_{ji} \leftarrow \omega_{ji} + c\lambda^2 y_j(1-y_j)x_i(\sum_{k=1}^{K}(d_k - z_k)z_k(1-z_k)v_{kj})$$

V. Update the error term:

$$E \leftarrow E + \sum_{k=1}^{K}(d_k - z_k)^2$$

Repeat phase 1 until you have the input patterns that are presented.

VI. If E is below the defined tolerance level (0.000001), stop. Restart the process with $E = 0$, and repeat the process.

1.3.4. Differential Neural Networks.

To continue with this analysis, we must point out the existence of two types of neural networks; the first is the discrete-type differential network, which has shown an excellent performance in financial and economic variables, however, several studies have found that this type of networks can be improved both in the accuracy and quantity of information what you need to model.

In Ortiz-Arango *et al.* (2013), it is suggested that dynamic or differential neural networks improve the capabilities of discrete-type neural networks. Since differential networks develop the study of time series continuously, and its structure of convergence is based on the stability theory proposed by Liapunov.

1.3.4.1. Structure of the Network State Variable.

From the studies of Poznyak *et al.* (1999), concepts about dynamic neural networks are developed to model nonlinear systems. And this is of the type:

$$\frac{d}{dt}x_t = \dot{x}_t = f(x_t, u_t, t) + \varepsilon_{1,t}; \qquad x_0 \text{ is known}$$

$$y_t = Cx_t + \varepsilon_{2,t}$$

Where x_t is the vector of states, y_t is the corresponding output, u_t is the internal supply C is the transformation matrix of the state at the output, and ε are the noises that can not be measured within the dynamics of the states and the departures of the states.

For the nominal dynamics which can be chosen by the theoretical results of the neural network, and the non-modeled dynamics that should be minimized during the training process. To do this, they proposed a parameter β which is adjusted to obtain the values perfectly the two dynamics of the differential neural network. From where a nominal dynamic is established as:

$$f(x_t, u_t, t) = A^0 x + W_1^0 \sigma(x) + W_2^0 \varphi(t) u$$

$$\beta = [W_1^0 , W_2^0]$$

And the activation functions are sigmoid of the type:

$$\sigma_j(t) = a_j \left(1 + b_j e^{\sum_{j=1}^{n} c_j x_j}\right)^{-1}$$

$$\varphi_{kl}(t) = a_j \left(1 + b_{kl} e^{\sum_{j=1}^{n} c_{kl} x_l}\right)^{-1}$$

$$j = 1 \ldots, n; \; k = 1, \ldots, n; \; l = 1, \ldots, m$$

Now the state identifier is considered from a differential neural network:

$$\dot{\hat{x}}_t = A\hat{x}_t + W_{1,t}\sigma(\hat{x}_t) + W_{2,t}\varphi(\hat{x}_t)\, u_t$$

Where $\dot{\hat{x}}_t$, is the state variable of the neural network $W_{1,t}\sigma(\hat{x}_t)$ and $W_{2,t}\varphi(\hat{x}_t)$ are the weight matrices described in the outputs and hidden connections. $\sigma(\hat{x}_t)$ is the field vector constructed for elements with incremental behavior as Sigmoid Function, $\varphi(\hat{x}_t)$ is the activation function in each element and finally u_t, is the differentiable input function of which the identifier requires that there be a positive defined Hurwitz Q matrix such that the Riccati equation has a positive solution $P = P^T > 0$ (Cabrera-Llanos *et al.*, 2012). And the form of this solution is as follows:

$$A^T P + PA + PRP + Q = 0$$

1.3.4.2. Learning Algorithm Based on the "Approach" of Liapunov.

Cabrera-Llanos *et al*. (2012), it knows that these model studies systems of the non-linear type. Such that, the identifiers of the neural network where the weights are adjusted by the laws of learning that follow:

$$\frac{d}{dt}W_{1,t} = -K_1 P\Delta_t \sigma(\hat{x}_t)^T$$
$$\frac{d}{dt}W_{2,t} = -K_2 P\Delta_t \gamma(\hat{u}_t)^T \varphi(\hat{x}_t)^T$$
$$\Delta_t = x_t - \hat{x}_t$$

So, when $t = 0$, $W_{1,0}$ and $W_{2,0}$, it has the matrices of initial weights. Where K_1 and K_2 represent a pair of positive definite matrices, P is the Riccati solution. In addition, it is assumed that within the dynamics developed by the weights these are bounded so that:

$$W_{1,t} \epsilon L_\infty$$
$$W_{2,t} \epsilon L_\infty$$

And these values converge to:

$$\lim_{t\to\infty} W_{1,t} = 0$$
$$\lim_{t\to\infty} W_{2,t} = 0$$

Therefore, it can be concluded that given these situations the identification process is asymptotically consistent such that:

$$\lim_{t\to\infty} \Delta_t = 0$$

Ortiz-Arango *et al*. (2016) if a system is assumed so that it is represented as follows:

$$\dot{\hat{x}}_t = A\hat{x}_t + W_{1,t}\sigma(\hat{x}_t) + W_{2,t}\varphi(\hat{x}_t)\gamma(\hat{u}_t)$$

And the selections of the Liapunov function by:

$$V_t = P\Delta_t^2 + \frac{1}{K_1}W_{1,t}^2 + \frac{1}{K_2}W_{2,t}^2$$

Where the condition for the system to be asymptotically stable is the Riccati solution:

$$\dot{V}_t \leq -Q_0 \Delta_t^2$$

1.4. Fuzzy Time Series Models.

From the theory of fuzzy sets developed by Zadeh (1965), the concepts of fuzzy sets and fuzzy relations have been used in combination with other theories. In this section, a series of models of time series and econometrics are presented under the assumption that some of their components can be described by fuzzy theory.

From this aspect, two situations are recognized. The first is when the models assume that the estimation parameters are described by a membership function or possibility distribution function, which allows the estimated variable to be described by the same membership function as the fuzzy parameters. In the second instance, reference is made to consider the dependent variable as fuzzy, determined by a series of fuzzy relationships in its past realizations.

1.4.1. Fuzzy Time Series.

This type of model has its origins in Song *et al*. (1993a), in which a new aspect of the study of time series is developed. Called like Fuzzy Time Series, which are based on the premise that the time series can be explained by fuzzy theory.

Definition 9. Let $Y\ (t)\ (t\ =\ \cdots, 0,1\ldots)$, a subset of the real numbers, the fuzzy universe of study, $\mu_i(t)\ \ (i = 1,2 \ldots)$, is defined as; $F(t)$ is a collection of $\mu_1(t), \mu_2(t), \cdots$ membership functions. Therefore, it is understood that $F\ (t)$ is a Fuzzy Time Series determined in $Y\ (t)\ (t\ =\ \cdots, 0,1\ldots)$.

In this definition, it is possible to understand the time series from a series of fuzzy sets. In other words, the time series in each of its values has a membership function.

Definition 10. The universe of study is understood as the space in which the fuzzy time series has all its membership functions defined. This has upper and lower level such that:

$$[L_{BD}, U_{BD}] \tag{1.2}$$

From the above, it is concluded that the time series is defined under a universe set. In which will be the possibilities distribution functions, these determined between the upper and lower limit of the universe set.

Definition 11. A fuzzy logic relation assumes that $F(t-1) = A_i$ and $F(t) = A_j$. The relation between $F\ (t-1)$ and $F\ (t)$ refers to the fuzzy relation, in such a way that:

$$A_i \rightarrow A_j \tag{1.2}$$

Definition 12. Let the If-Then rule be such that if $F\ (t)$ is caused by $F\ (t-1)$. Such that, the first model of $F\ (t)$ denoted as:

$$F(t) = F(t-1)°R(t, t-1) \tag{1.3}$$

Where, $R\ (t, t-1)$ represents the fuzzy logic relation (1.2) and the symbol "°" represents the composition operator Max-Min. From the above definitions, it is recognized that the forecast of the fuzzy time series is determined by the value one period back and the relation of fuzzy logic that has its membership functions of each performance (Singh, 2017).

From the previous definition, if $R\ (t, t-1)\ =\ R\ (t-1, t-2)\ \forall t$ then it is said that the fuzzy time series is invariant time. On the other hand, if $R\ (t, t-1)$ is dependent on the time, such that $R\ (t, t-1)\ \neq\ R\ (t-1, t-2)$, the Fuzzy Time Series will be variant time (Chen *et al*., 2004).

$$R(t, t-1) = R(t-1, t-2)\ \ \forall t; \quad \text{invariant time}$$
$$R(t, t-1) \neq R(t-1, t-2)\ \ \forall t; \quad \text{variant time}$$

Definition 13. The relationship of fuzzy logic in groups is understood as in which the following is assumed:

$$A_i \to A_{k1}$$
$$A_i \to A_{k2}$$
$$\cdots$$
$$A_i \to A_{km}$$

Or

$$A_i \to A_{k1}, A_{k2}, \dots, A_{km}$$

Definition 14. It is assumed that $F\ (t)$ depends on $F\ (t-1), F\ (t-2), \cdots, F\ (t-n)$ it is said that there is a higher-order model which is expressed as follows:

$$F(t-n), \dots, F(t-2), F(t-1) \to F(t)$$

It is expressed as the value of $F\ (t)$ is determined by the values that it previously took and their relations of fuzzy logic. This situation is important when the processed of the Fuzzy Time Series is of long memory.

The definitions in this section allow us to identify the elements under which it is assumed that the series of study time has and is described under fuzzy theory conditions. It is based on the assumption that the expert system that is needed to define the intervals to determine the relations of fuzzy logic is given by the market. Therefore, in the economic applications of the fuzzy time series, the market is considered to be the expert system that provides enough information to develop the If-Then" rules of the models.

1.4.2. Fuzzy Linear Regression.

The methodology that starts from the classic models of linear regression and time series, assuming its parameters and variable estimated as fuzzy. This indicates that both components of the models have a function of distribution of possibility. These models were developed from the work Tanaka *et al.* (1982), who develop from the Fuzzy Theory and the linear regression model a new model known as the Fuzzy Linear Regression Model.

Definition 15. Let a Fuzzy Function denote by:

$$f: X \to \mathfrak{F}(y); \quad \boldsymbol{Y} = f(x, \boldsymbol{A})$$

Where $\mathfrak{F}(y)$ is a set of all fuzzy subsets in Y, the fuzzy set Y is defined by the following membership function:

$$\mu_Y(y) = \begin{cases} \max\limits_{\{a|y=f(x,\boldsymbol{a})\}} \mu_A(a) & \{a|y = f(x,\boldsymbol{a})\} \neq \emptyset \\ 0 & otherwise \end{cases}$$

Where A is a fuzzy set determined by the parameters with its membership function $\mu_A(a)$. So this space includes all those parameters that are within the possibility distribution function. On the other hand, this membership function of the parameters can be re-expressed as, obtaining the least amount of parameters that manage to estimate the behavior of the study variable and thus minimize the fuzzy component of these models.

$$\mu_A(a) = \min_j \mu_{A_j}(a_j)$$

Now, by identifying which is the function associated with $\mu_A(a)$, which is primarily represented by a function of the triangular type. Since what is of primary interest to these models is to identify the width of the function and thus obtain extreme values of possibility (Tanaka *et al*., 1982). So the distribution function of possibility is denoted as:

$$\mu_{A_j}(a_j) = \begin{cases} 1 - \dfrac{|\alpha_j - a_j|}{c_j}, & \alpha_j - c_j \leq a_j \leq \alpha_j + c_j \\ 0, & otherwise \end{cases}$$

Where c represents the width and α is the center of the triangular function. Therefore, the fuzzy parameters in A can be represented as $A = (A_1, \cdots, A_n)$ which is denoted as:

$$\boldsymbol{A} = \{\alpha, c\}$$
$$\alpha = (\alpha_1, \cdots, \alpha_N)^t$$
$$c = (c_1, \cdots, c_n)^t$$

Therefore, the fuzzy linear regression model can be represented with $A = \{\alpha, c\}$ in the following way:

$$Y = A_1x_1 + \cdots + A_nx_n = Ax$$

So, according to Tanaka *et al.* (1982), it is shown[2] that the membership function associated with the dependent variable is represented in the following way:

$$\mu_Y(y) = \begin{cases} 1 - \dfrac{|y - x^t\alpha|}{c^t|x|}, & x \neq 0 \\ 1, & x = 0; y = 0 \\ 0, & x = 0; y \neq 0 \end{cases}$$

Now the model proposes the use as value α or center of the distribution function of possibility to the value of the parameter associated with the independent variables of the best linear regression model of the independent variable. Therefore, the value of the fuzzy parameters will oscillate between the value of α and the width of the function c. Therefore, once having the center values of the membership function the model will be denoted as:

$$Y^* = A_1^* x_1 + \cdots + A_n^* x_n = A^* x$$

Arrived at this instance, it is necessary to identify the values associated with the width of the membership function. Therefore, the following definition should be understood to structure the analysis for this component of the membership function.

Definition 16. Let the estimated degree of the estimated fuzzy regression $Y^* = A^* x$ be such that the information provided by the model is given by $Y_i = (y_i, e_i)$ that is measured by the following index $\overline{h_i}$ Where we seek to maximize h subject to $Y_i^h \subset Y_i^{*h}$, such that;

$$Y_i^h = \{y | \mu_{Y_i}(y) \geq h\}$$
$$Y_i^{*h} = \{y | \mu_{Y_i^*}(y) \geq h\}$$

What represents the h-level of the fuzzy set. The degree of estimation of the fuzzy linear regression model of all the information of the model $\mathrm{Y}_1, \cdots, \mathrm{Y}_\mathrm{N}$, is found by $\min_{\mathrm{j}}[\overline{\mathrm{h}_\mathrm{j}}]$. So that the fuzzy part of the model is given by:

$$J = c_1, \cdots, c_n$$

[2] See the demonstration in Tanaka *et al.* (1982), appendix.

Then the problem in question is given by obtaining the fuzzy parameters A_i^* such that the J value subject to $\overline{h_i} \geq H, \forall i$ is minimized. Where H is chosen by the expert system as mentioned in the previous section and represents the degree of estimation of the fuzzy linear regression model. So $\overline{h_i}$ can be obtained by:

$$\overline{h_i} = 1 - \frac{|y - x_i^t \alpha|}{\sum_j c_j |x_{ij}| - e_i}$$

From the last two equations, the problem of finding the fuzzy parameters is specified. So the following linear programming problem must be solved:

$$\min_{\alpha,c} J = c_1, \cdots, c_n$$

Subject to

$$\begin{gathered} c \geq 0 \\ \alpha^t x_i + (1 - H) \sum_j c_j |x_{ij}| \geq y_i + (1 - H) e_i \\ -\alpha^t x_i + (1 - H) \sum_j c_j |x_{ij}| \geq -y_i + (1 - H) e_i \\ i = 1, \cdots, N \end{gathered} \tag{1.4}$$

By solving the above minimization problem it finds the best values to indicate the width of the membership function for each parameter of the fuzzy linear regression. And together these denote the width of the triangular function associated with the estimated variable. Consequently, this model provides more information than the traditional linear regression model. In the sense that not only shows the average estimate but also denotes an estimation interval. Finally, it is recognized that for the following models of this section the same methodology was used to find the fuzzy parameters and the fuzzy estimation of the dependent variable.

1.4.3. Fuzzy ARIMA.

The Fuzzy ARIMA model born of the ARIMA model and this model lets obtain the center values of the possibility distribution function. And apply the Tanaka methodology to find the fuzzy parameters and thus the fuzzy estimate. The difference between this model and the fuzzy linear regression model is that in this the dependent variable is explained by its delayed values in p periods back, unlike the previous model where x are variables that explain the behavior of the dependent variable.

Therefore, in this model, concepts such as Autoregressive $AR\ (p)$ and Moving Average $MA\ (q)$ must be considered. This model should be considered taking as reference the main characteristics of the traditional ARIMA model, so it is assumed that the fuzzy part of the dependent variable is determined through the fuzzy parameters associated with the autoregressive segment and the moving average.

It starts from assuming a time series $\{Y_t\}$ generated by an ARIMA process (p, d, q). Such that, it is represented in the following way:

$$\phi_p(B)Y_t = \theta_q(B)\varepsilon_t$$

So that,

$$Y_t = (1 - B)^d(Z_t - \mu)$$

Where, $(1 - B)^d$ represents the polynomial of delays and Z_t are the observations $\phi_1 + \cdots + \phi_p$ and $\theta_1 + \cdots + \theta_q$. Such that, they assume that these coefficients are fuzzy and therefore have membership function.

$$Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + \varepsilon_t - \sum_{i=p+1}^{p+q} \theta_i \varepsilon_{t+p-i}$$

So the membership functions of each parameter according to Tseng *et al.* (2002). They are represented under the methodology of the fuzzy linear regression model, in the following way:

$$\mu_{\phi_i}(\phi_i) = \begin{cases} 1 - \dfrac{|\alpha_i - \phi_i|}{c_i}, & \alpha_i - c_i \leq \phi_i \leq \alpha_i + c_i \\ 0, & otherwise \end{cases}$$

$$
\mu_{\theta_i}(\theta_i) = \begin{cases} 1 - \dfrac{|\alpha_i - \theta_i|}{c_i}, & \alpha_i - c_i \leq \ \theta_i \leq \alpha_i + c_i \\ 0, & otherwise \end{cases}
$$

Where, $\mu_{\phi_i}(\phi_i)$ and $\mu_{\theta_i}(\theta_i)$ represents the possibility of distribution functions associated with the parameters of the model processes. And α is the center and c is the width of the membership function. Therefore, under the fuzzy linear regression model, the membership function associated with the dependent variable is denoted as follows:

$$
\mu_Y(Y_t) = \begin{cases} 1 - \dfrac{\left|Y_t - \sum_{i=1}^{p} \phi_i Y_{t-i} - \varepsilon_t + \sum_{i=p+1}^{p+q} \theta_i \varepsilon_{t+p-i}\right|}{\sum_{i=1}^{p} c_i |Y_{t-i}| + \sum_{i=p+1}^{p+q} c_i |\varepsilon_{t+p-i}|}, & x \neq 0 \\ 1, & \displaystyle\sum_{i=1}^{p} \phi_i Y_{t-i} - \varepsilon_t + \sum_{i=p+1}^{p+q} \theta_i \varepsilon_{t+p-i} = 0; Y_t = 0 \\ 0, & \displaystyle\sum_{i=1}^{p} \phi_i Y_{t-i} - \varepsilon_t + \sum_{i=p+1}^{p+q} \theta_i \varepsilon_{t+p-i} = 0; Y_t \neq 0 \end{cases}
$$

On the other hand, under the considerations previously made, the center and width values of each membership function have to be found; for the value of α, the variable Y_t is estimated under the ARIMA (p, d, q) and the one provided by the best ARIMA model is used as the average value. In such a way, that the fuzzy component of the model is represented in the following way:

$$
J = \sum_{i=1}^{p} \sum_{t=1}^{k} c_i \, |\varphi_{ii}| |Y_{t-i}| + \sum_{i=p+1}^{p+q} \sum_{t=1}^{k} c_i \, |\rho_{i-p}| |\varepsilon_{t+p-i}|
$$

Where, φ_{ii} represents the autocorrelation function and ρ_{i-p} the partial autocorrelation function. From which the optimal lags of the model are obtained and multiplied by the width of the membership function. Now, taking the value of α, this value is used to find the width of each of the membership functions and thus define the minimum number of fuzzy lags in the models. This is done by solving a linear programming problem, described as follows:

$$\min_c J = \sum_{i=1}^{p}\sum_{t=1}^{k} c_i\,|\varphi_{ii}||Y_{t-i}| + \sum_{i=p+1}^{p+q}\sum_{t=1}^{k} c_i\,|\rho_{i-p}||\varepsilon_{t+p-i}|$$

$$s.a \qquad c_i \geq 0,\ \ \forall i = 1,\cdots,p+d$$

$$\sum_{i=1}^{p}\phi_i Y_{t-i} + \varepsilon_t - \sum_{i=p+1}^{p+q}\theta_i\varepsilon_{t+p-i} + (1-H)\left(\sum_{i=1}^{p} c_i|Y_{t-i}| + \sum_{i=p+1}^{p+q} c_i|\varepsilon_{t+p-i}|\right) \geq Y_t$$

$$\sum_{i=1}^{p}\phi_i Y_{t-i} + \varepsilon_t - \sum_{i=p+1}^{p+q}\theta_i\varepsilon_{t+p-i} - (1-H)\left(\sum_{i=1}^{p} c_i|Y_{t-i}| + \sum_{i=p+1}^{p+q} c_i|\varepsilon_{t+p-i}|\right) \leq Y_t$$

$$t = 1,2,\cdots,k$$

When solving this problem, the optimal values for the fuzzy linear regression model are found. Such that, the fuzzy ARIMA model can be represented in the following way:

$$Y_t = \langle\alpha_1, c_1\rangle Y_{t-i} + \cdots + \langle\alpha_p, c_p\rangle Y_{t-p} + \varepsilon_t - \langle\alpha_{p+1}, c_{p+1}\rangle\varepsilon_{t-1} - \cdots - \langle\alpha_{p+d}, c_{p+d}\rangle\varepsilon_{t-d}$$

Where $\langle\alpha_1, c_1\rangle$ represent the values of each coefficient. Taking as center α and c is the extension of the model understood as the upper and lower limit of the fuzzy set associated with the coefficient. Therefore, under this methodology, it is possible to carry out a forecast of the time series, obtaining, as a result, a range of possibilities of the future values of the variable.

1.4.4. Gaussian Volatility Fuzzy GARCH.

The present model combines some definitions of fuzzy time series and the methodology developed by Tanaka *et al.* (1982). The GARCH (p,q) models are used to capture the behavior of the variance. It is based on the situation that the traditional model is used and it is assumed that its coefficients have a possibility distribution function.

Considering now the development of Almeida *et al*. (2014), it is assumed that the study variable y_t and the conditional variance σ_t^2 are determined by the first relationship of fuzzy logic. This expressed as:

$$R_l: Si\ x\ es\ F_l\ entonces\ y_{t,l}|x_t, \sigma_{t,l}^2 \sim N(u_l, \sigma_{t,l}^2)$$

$$y_{t,l} = \sigma_{t,l}\varepsilon_t$$
$$\sigma_{t,l}^2 = \alpha_{0,l} + \sum_{i=1}^{q} \alpha_{i,l}\, y_{t-1}^2 + \sum_{j=1}^{p} \beta_{j,l}\, \sigma_{t-j}^2$$
$$\sigma_{t-j}^2 = \sum_{l=1}^{L} g_{t-j,l}\sigma_{t-j,l}^2\,;\ \forall\, j = 1, \cdots, p$$

Where, R_l is the If-Then rule of the process$y_{t,l}$. Such that, the passed values of each rule are fuzzy sets such that the l-*th* rule is that it is fulfilled in the next equation and x is one of the independent variables. $F_l: X \rightarrow [0,1]$ It is a multidimensional fuzzy conjuncture defined in a continuous space X. The fuzzy model process is from the next equation, where $g_{t,l}$ is a normalized membership function such that:

$$g_{t,l} = \frac{\mu_{t,l}}{\sum_{l=1}^{L} \mu_{t,l}}$$

And the estimated value of this model must comply that:

$$y_{t,l}|\sigma_{t,l}^2, x_t \sim N(u_l, \sigma_{t,l}^2)$$

In the parameters of this model, each rule $l = 1, \cdots, L$, must fulfill these conditions:

$$\alpha_{0,l} > 0$$
$$\alpha_{i,l} \geq 0$$
$$\beta_{j,l} \geq 0$$
$$\sum_{i=1}^{q} \alpha_{i,l} + \sum_{j=1}^{p} \beta_{j,l} < 0$$
$$i = 1, \cdots q;\ \ j = 1, \cdots, p.$$

The model can be obtained through the maximum likelihood method, which takes the type and number of membership functions known as well. Starting from the following logarithmic function:

$$\ln \ell(y|I_{t-1}) = ln \prod_{t=t^*}^{T} \ell(y_t|x_t, \sigma_t^2) = \sum_{t=t^*}^{T} \ln \left(\sum_{l=1}^{L} g_{t,l} \xi(y_t; \mu_l; \sigma_{t,l}^2) \right)$$

Where $t^* = \max(p, q)$ and $\xi(\cdot)$ takes into account all the functions of the fuzzy GARCH model. The membership functions are of the next form:

$$\mu_{t,l}(x_t) = \prod_{k=1}^{n} e^{\left(-\frac{1}{2} \frac{(x_{kt} - c_{k,l}}{S_{k,l}^2} \right)}$$

In order to generate the estimate of this model, it starts to obtain the center values of the GARCH model. And after that, the following problem is resolved:

$$\min_{u_l, \theta_{g,l}, \theta_{\mu,l}} - \ln \ell(y|I_{t-1}) = - \sum_{t=t^*}^{T} \ln \left(\sum_{l=1}^{L} g_{t,l} \xi(y_t; \mu_l; \sigma_{t,l}^2) \right)$$

$$s.a\ c_l \leq c_{l+1}$$
$$\alpha_{0,l} > 0$$
$$\alpha_{i,l} \geq 0$$
$$\beta_{j,l} \geq 0$$
$$\sum_{i=1}^{q} \alpha_{i,l} + \sum_{j=1}^{p} \beta_{j,l} < 1$$
$$i = 1, \cdots q;\ \ j = 1, \cdots, p.$$

This method models the volatility and obtained a conventional forecast, not fuzzy output.

1.4.5. Fuzzy Time Series Model Combined with Artificial Neural Networks.

In this section, models that combine the theory of neural networks and fuzzy logic are synthetically developed. Starting from the assumption that the variable that feeds the network is a fuzzy time series and in the hidden layers of the network, there are membership functions that relate to the nodes of each phase. Specifically, multilayer Perceptron type neural networks are used and they arise from assuming relations of fuzzy logic through If-Then rules.

1.4.5.1. Adaptive Networks Based on Fuzzy Inference System (ANFIS).

According to Jang (1993), we must consider the variable x which generates an output variable z through a neural network. From which a fuzzy first-order model of the Sugeno type is formed, in which the If-Then rules are determined as follows:

$$R_1: if\ x\ is\ A_1\ then\ f_1 = p_1 x + r_1$$
$$R_2: if\ x\ is\ A_2\ then\ f_2 = p_2 x + r_2$$

In the first phase of the network, each node is concentrated with the respective membership function of the fuzzy subset. This means that x is the variable that feeds each node i and A_i is the fuzzy subset denoted as linguistic value.

$$O_i^1 = \mu_{A_i}(x)$$

Where O_i^1 is the membership function of A_i and expresses the degree of membership of x over the fuzzy subset. In the specific case of this membership function, the expression that will be used will be a function of the Bell-type, as expressed below.

$$\mu_{\mathrm{A_i}}(x) = \frac{1}{1 + \left|\frac{x - c_i}{a_i}\right|^{2b_i}}$$

In the second phase, in the neural network, each of the nodes is represented by a circular marking, which means that the node will be the multiplication of the signals and this product will be the output. Each node represents the weight of each of the rules, such as:

$$w_i = \mu_{\mathrm{A_i}}(x) \times \mu_{\mathrm{A_j}}(x), \quad i \neq j$$

In phase three, the node ratio of the $i - th$ fuzzy rule is calculated. This is to find the impact of each weight on the total weights of the network, such that it can be determined as follows:

$$\overline{w_i} = \frac{w_i}{w_1 + w_2}, \quad i = 1{,}2.$$

In other words, this phase is understood as the normalization of the weights. The fourth phase expresses each node i in such a way that it can be expressed as a normalized node with its function obtained from the If-Then rules, so that:

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i (p_i x + r_i)$$

Ultimately, in a single node, the sum of all the output signals of the previous phases is performed; the output variable z is represented as:

$$O_i^5 = \frac{\sum_i \overline{w}_i f_i}{\sum_i \overline{w}_i}$$

The learning algorithm of adaptive network-based on fuzzy inference, it is given by the rules "If-then" and the observation of the parameters of the models what can be expressed by a linear combination of general results so that it can be expressed as follows:

$$f = \frac{w_i}{w_1 + w_2} f_1 + \frac{w_i}{w_1 + w_2} f_2 = \overline{w_1}(p_1 x + r_1) + \overline{w_2}(p_2 x + r_2)$$

1.4.5.2. Fuzzy Neural Network Model.

From Wu *et al.* (2000), a model of neural networks is proposed, where the fundamental premise is that the information that feeds the network is a fuzzy set. In each phase of the network, the nodes can be expressed through membership functions. The latter, seeking to extract the belonging of the information about the study set.

The structure of the single network with an input phase where each input node represents a variable of the linguistic type. In other words, there is a fuzzy subset for each input node in the neural network. In the second stage, each node will be represented by a membership function, of the following form:

$$\mu_{\mathrm{ij}}(x_i) = e^{-\frac{(x_i - c_{ij})^2}{\sigma_j^2}}; \quad i = 1, \cdots, r; j = 1, \cdots, u.$$

Where, $\mu_{\text{it}}(x_i)$ is the membership function of the element x_i and σ_j^2 is the width measurement for the membership function with a center in c_{ij}. The variables that feed the network are r and the number of membership functions is u. In the third phase, each node represents a part of the fuzzy rule If, so that the $j - th$ rule R_j, its output is given by:

$$\varphi_j = e^{-\frac{\sum_{i=1}^{r}\left(x_i - c_{ij}\right)^2}{\sigma_j^2}}$$

$X = (x_1, \cdots, x_r)$ and C_j is the center of the $j - th$ is the unit of the radial basis function. In the fourth phase, the nodes are denoted as normalized nodes, as in the previous model. The number of normalized nodes N is equal to the number of nodes R.

$$\theta_j = \frac{\varphi_j}{\sum_{k=1}^{u} \varphi_k}$$

In the fifth phase, each node is represented by the output variables as the sum of the input signals to this phase, such that:

$$y(X) = \sum_{k=1}^{u} w_{2k} \cdot \varphi_k$$

Where y is the value of the output variable and w_{2k} is the weight of each rule. The weights are given by:

$$w_{2k} = k_{j0} + k_{j1}x_1 + \cdots + k_{jr}x_r$$

In conclusion, this model allows us to model the fuzzy time series using the components of a neural network. It is assumed that the membership functions are of Gaussian type which supports identifying the belongings of the input value on the fuzzy subset. It is recognized that it is possible to develop the forecast of the exchange rate through fuzzy theory and artificial neural networks. This is done through the implementation of several analysis models, which according to their study method are classified into models with fuzzy parameters and fuzzy time series. Therefore, the following concern is presented: Which of these models provides a better forecast of the exchange rate? This question is addressed in the following sections.

CHAPTER II. FORECAST OF EXCHANGE RATE WITH VARIOUS METHODOLOGIES

"You have to create original science, in all the orders of philosophy, chemistry, biology, sociology, and so on. After the original science will come to the industrial application of scientific principles, because always explodes next to the new fact the exploitation of it, that is, the application to the increase and comfort of life. In the end, the fruit of science applied to all the orders of human activity is wealth, well-being, the increase of population and military and political force".

Santiago Ramón y Cajal, The half science cause of ruin.

The main objective of this section is to develop econometric models for economic variables (ARMA, E-GARCH, PARCH and NEURONAL NETWORKS), which will be applied to the forecast of the exchange rate in the period it comprises; January 2, 2008, to December 29, 2017, in daily format (excluding non-working days, with a total of 2514 observations). For the out sample test 26 observations are added, January 2, 2018, to February 7, 2018.

The justification for this sample is that the behavior of the growth rate of the exchange rate before 2008 maintained a structure of behavior different from that observed after the same year. In addition, it is necessary to model the dynamics of the time series in order to comply with the objective of this section (prediction).

In this chapter, the main characteristics of each model are presented together with their application to the exchange rate. It starts from the theoretical aspects of each one and assuming that these are capable of optimally modeling the behavior of the exchange rate.

Subsequently, the efficiency of these is compared through various analysis criteria based on the variance and covariance of the errors, so that it is possible to identify which is the best model.

Because of the nature of the exchange rate (as a financial variable) and its importance for various situations in the economic sphere, it is necessary for modeling to recognize that the existence of political, social and international factors has an impact on the behavior of the series, an aspect that makes the forecast of this, in particular, an arduous task.

Consequently, trying to obtain the best possible study, we present a series of methodologies capable of explaining the dynamics of the behavior of the exchange rate. The statistic that will be used to compare is divided into evaluating the errors such as Absolute mean error and mean square error.

By using the methods of error evaluation (mean absolute error and root mean square error) seeks to give an overview of the range of linear and nonlinear models for modeling time series, concluding with the use of indicators to infer what the method that adapts better to the behavior of series with high volatility.

Finally, the main results obtained are shown, as well as the aspects that are relevant to understand the dynamics of the variable studied; giving a guideline, to begin with, the analysis of fuzzy time series models and dynamic neural networks for forecasting the exchange rate.

2.1. Time series, stationarity, and exchange rate.

In the first instance, the characteristics of the time series are identified: trend, seasonality, cycle, and irregular component; relevant aspects to model the behavior of variables, through a range of mathematical and statistical tools that allow us to intuit future values.

The first methods recognize the trend and seasonality, Autoregressive and Moving Average; the second ones, also annex the irregular part, diffuse models. Both methodologies are implemented to respond to the research hypothesis. In the beginning, linear and non-linear autoregressive models are studied (Chapter II) and then uncertainty models (Chapter III). Since the objective of this research does not require to make use of the models that study the cyclical component. Only the mentioned methodologies are used to find, what is the best model for the forecast of the exchange rate?

Before starting, we start from the concept of stationarity of the time series, Brockwell *et al*. (2006):

Let a time series$\{\mathbb{X}_t, t \in \mathcal{Z}\}$, with index $\mathcal{Z} = \{0, \pm 1, \pm 2, \dots\}$ is said to be stationary in the strict sense if the joint distribution of $(\mathbb{X}_{t1}, \dots, \mathbb{X}_{tk})'$ and $(\mathbb{X}_{t1+h}, \dots, \mathbb{X}_{tk+h})$ is the same for any $k \in \mathcal{Z}^+$ and for all $t_n, \dots, t_k,\ \ h \in \mathcal{Z}$. And it says stationary in a weak sense if:

i. $E(\mathbb{X}_t{}^2) < \infty \quad for\ all\ t \in \mathcal{Z}.$
ii. $E(\mathbb{X}_t) = m \quad for\ all\ t \in \mathcal{Z}.$
iii. $\gamma_{\mathbb{X}}(r, s) = \gamma_{\mathbb{X}}(r + t, s + t)\ \ for\ any\ r, s, t \in \mathcal{Z}.$

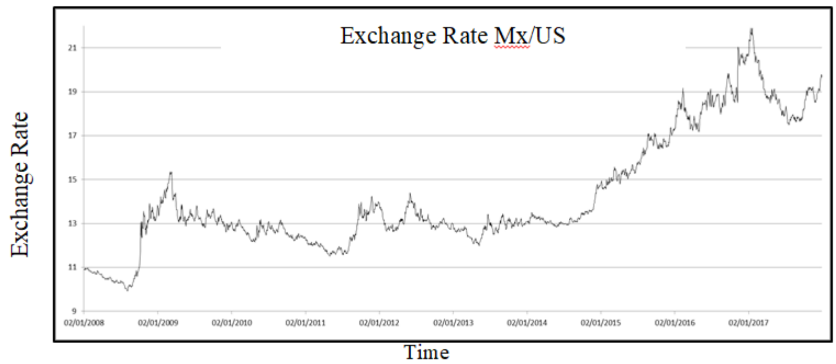
The previous one is understood as; there is stationarity in the strict sense if the time series has the same distribution throughout its history. And in a weak sense, it must be fulfilled that: the expected value of the variations of the series is finite so that the variability of the time series can be measured and intuited; the expected value of the variable is constant and the auto-covariance function for two embodiments only depends on the distance in time, not the variation between them.

It is emphasized that this condition must be fulfilled in order to successfully develop the study of the dynamics of the time series with certain probabilities of success. This as a consequence, if the time series does not comply with stationarity, finding its behavior dynamics becomes complicated and even impossible. Therefore, it is necessary to study a stationary process, and if not, apply the necessary transformations to obtain this process.

2.1.1. Stationarity of the exchange rate.

In this section, the study of the time series of the exchange rate is developed, it is worth mentioning that the series of time used to carry out the tests and the models of the present chapter is obtained from the official page of the Bank of Mexico and this is denominated as FIX; it is of daily temporality from January 2, 2008, to December 29, 2017, which includes 2514 observations only taking into account bank business days (Banco de México, 2018).

Figure 3. Time series of the exchange rate Mexican Peso against American Dollar, from 02/01/2008 to 12/29/2017 on a daily period



Source: Own elaboration in Excel.

Figure 3 shows the behavior of the exchange rate on a daily basis for 9 years. It is observed that there is a clear trend towards the depreciation of the Mexican peso against the US dollar, in addition to denoting significant leaps or aberrant values that indicate the high volatility of the time series.

The theory of time series indicates that in order to develop models it is necessary to work with variables that comply with stationarity. Therefore, various tests are performed to verify this property in the case study.

Table 1 presents three of the main tests for assessing stationarity in time series. The first, called "Dickey-Fuller Augmented", states that the exchange rate is a random walk, in other words, it is not stationary. On the other hand, the "Phillips-Perron" test indicates that the process is a unitary root, there is no stationarity and finally, the "KPSS" statistic shows that there is no empirical evidence to say that the time series is stationary. In conclusion, the time series in the sample period for the exchange rate does not meet the criterion of stationarity.

Table 1. Unit root test for the exchange rate MX/US

Unitary Root Test	**Probability (i) / Statistic (ii).**	**Interpretation.**
Dickey-Fuller Augmented (ADF).	0.8909 (i)	The exchange rate is a unit root.
Phillips-Perron (PP).	0.8837 (i)	The exchange rate is a unit root.
Kwiatkowski–Phillips–Schmidt–Shin (KPSS).	4.4429[3] (ii)	The exchange rate is non-stationary.

Source: Own elaboration in Eviews.

Therefore, a transformation must be made to the data in order to carry out the study. Given the previous results it is considered to develop the research with the growth rate, the transformation is established by means of the following equation:

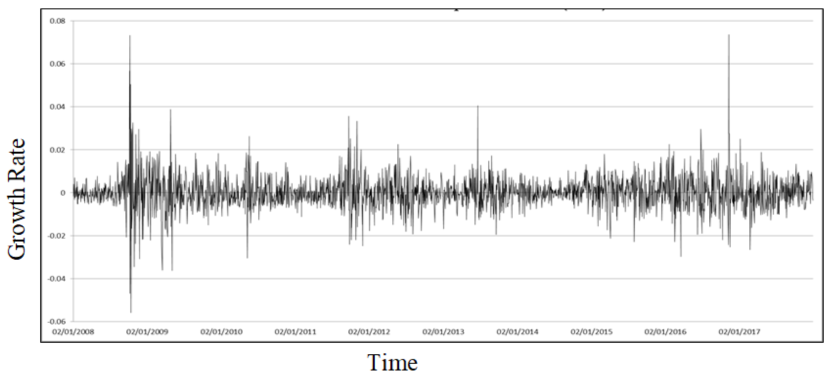
$$r = \ln\left(\frac{y_t}{y_{t-1}}\right)$$

Let r be the growth rate of the series, y_t the series of time in the current time and y_{t-1} the previous lag of the time series. This equation applies to the entire time series and so one observation is lost, the sample is 2513 realizations.

[3] The critical values are 0.739000, 0.463000 and 0.347000 at 90%, 95% and 99% respectively.

Figure 4 shows the growth rate of the exchange rate, denominating the time series under the acronym RTC (to simplify its use in the models). It is observed that the series does not have a marked tendency and its variability is around its average value, but it is not possible to observe an average variation. There are important volatility clusters, which can impact when trying to determine the dynamics of their behavior.

Figure 4. Time series of the growth rate MX/US, from 03/01/2001 to 12/29/2017



Source: Own elaboration in Excel.

Now, one must prove statistically that this series is stationary. Therefore, the growth rate of the exchange rate is evaluated with the stationarity tests used with the exchange rate. Table 2 shows the results of the unit root tests for the growth rate of the exchange rate. The ADF and PP tests indicate that there is no information to indicate that RTC is a unitary root, so it is not rejected is stationary. On the other hand, KPSS indicates that there is empirical evidence to say that the time series is a stationary process. So it is concluded that RTC is a series of stationary time fulfilling the requirements of the theory of time series.

In view of the above, the models are carried out taking into account the growth rate of the exchange rate as the main input or dependent variable. These are divided into two main sections, the linear and non-linear models, which are mentioned in the following sections.

Table 2. Unit root test for the time series of the exchange rate

Unitary Root Test	Probability (i) / Statistic (ii).	Interpretation.
Dickey-Fuller Augmented (ADF).	0.0001 (i)	The exchange rate is not a unit root.
Phillips-Perron (PP).	0.0001 (i)	The exchange rate is not a unit root.
Kwiatkowski–Phillips–Schmidt–Shin (KPSS).	0.07752[4] (ii)	The exchange rate is stationary.

Source: Own elaboration in Eviews.

2.2. ARMA.

In the first instance, the ARMA methodology is used, which seeks to decipher the dynamics of the time series through the Autoregressive and Moving Average structure; In other words, the behavior of the growth rate of the exchange rate is examined, making use of its past achievements, as well as the forecast errors that were made in previous stages (Pindick *et al.*, 2001).

[4] The critical values are 0.739000, 0.463000 and 0.347000 at 90%, 95% and 99% respectively.

As a consequence of the fact that the model is constructed from a transformation of the exchange rate, it is identified that the methodology is of the ARIMA (p, d, q). To begin with this analysis, we start by recognizing the mixed process as:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \cdots - \beta_q \; \varepsilon_{t-q}$$

Where it should be noted that y_t is a stochastic process indexed to time. Assuming, the information of their past achievements and the estimation errors of previous periods are sufficient to describe the dynamics of the time series. Being $\alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p}$ the existing information in the previous step of the process, that has relevance to understanding the current behavior.

On the other hand, $\beta_1 \varepsilon_{t-1} - \cdots - \beta_q \; \varepsilon_{t-q}$ show the forecast errors that were developed in previous stages and serve to provide the possible variation that has the estimated value with respect to the real one.

Before analyzing the application of this model, it should be noted that it must comply with the criterion of stationarity of time series. This is carried out under the condition that the process average is constant over time and this is given by:

$$\mu = \alpha_1 \mu + \cdots + \alpha_p \mu$$

$$\mu = \frac{1}{1 - \alpha_1 - \cdots - \alpha_p}$$

The last equation represents the necessary condition for stationarity of the process that is understood as: the sum of the parameters of the process $AR\ (p)$ is less than 1, $1 - \alpha_1 - \cdots - \alpha_p < 1$. This is the condition necessary and sufficient to comply with the theory of time series.

2.2.1. ARMA Structure of the Exchange Rate.

At this point, the specification of the time series is developed in relation to its application to the ARMA model. In the same way, it is intended to show the aspects that allow giving coherence and identity to the modeling that is intended to be carried out.

In the first instance, it is recognized that the exchange rate (FIX) is the determination in the market of the price at which a dollar valued in pesos will be paid. Knowing this, it is assumed that this valuation of the asset is made taking into account the existing information and to which it is accessed by the agents that carry out these transactions. In other words, the value of the exchange rate is indexed to the decision making of economic agents; therefore, it must be said that the past values have the quality of having enough information to estimate the current value of the series.

Therefore, it is said that the determination of the FIX is the average of the transactions carried out by economic agents as defined by the Bank of Mexico, which allows inferring that it is possible to identify the dynamics of the exchange rate based on its historical behavior. Infer their present and future realizations.

Based on the above, it is said that the history of the series recognizes all the information that was used. And with this, it is possible to take as an explanatory variable of the exchange rate to its past achievements.

Therefore, it is important to return to the $Ar(p)$ model to try to recognize the behavior of the exchange rate. In this aspect, it is understood that in performing this analysis in the same way it is necessary to return the forecast errors $Ma(q)$ as a complement to the information that cannot be measured by its previous values.

Now, knowing that this series shows great variability in its realizations; it is considered a mixed model to try to know the behavior of the exchange rate. In this case, an ARMA model (p, q) is used.

Before developing the model, it must be identified if the time series is stationary. Returning to what has been obtained in section (2.1.1); the model should be considered from the series in its steady-state and knowing that this series is non-stationary. And in the application, a transformation to the series is developed. Therefore, the model that will be applied is of the ARIMA (p, d, q).

Various ARIMA models were developed, which are not presented because it is not the objective of this section, only the one that gave the best information criteria will be shown

(Hannan-Quinn, Akaike and Schwartz). This model will be used for later stages of the present investigation with different purposes, for which the relevance of it is highlighted.

Table 3 shows the results obtained by modeling the behavior of the growth rate of the exchange rate under the ARIMA methodology, in which the valuation of the parameters and their level of statistical significance are developed. From what stands out, all the coefficients are statistically different from zero. Therefore, we proceed to review the stability condition of the model that was explained in the previous section.

Table 3. ARIMA model for the growth rate of the foreign Exchange rate

Variable.	Coefficient.	Statistical Significance[5].
AR(1)	0.043999	***
AR(2)	0.038673	***
AR(3)	-0.725395	***
MA(3)	0.669597	***

Source: Own elaboration in Eviews.

The condition in which this model was first arrived at must comply with the fact that the sum of the parameters of the $AR(p)$ process is less than 1, $(\alpha_1 + \cdots + \alpha_p) < 1$. And the result can be seen in table 2.3 in which the sum of the coefficients $AR(p)$ is less than 1, in this case, was -0.63. Knowing the above, we proceed to study the aspects of econometric nature that this study comprises.

Table 4. Tests performed on the ARIMA model for the growth rate of the exchange rate

Test / name.	Probability (i)/Statistic (ii).	Interpretation.
Heteroscedasticity / LM-ARCH.	0.0000 (i)	Heteroscedasticity
Autocorrelation / Durbin-Watson.	1.9808[6] (ii)	No Autocorrelation

[5] Statistical significance to the 99 % (***), 95 % (**) and 90 % (*).
[6]The value of this test must be near to 2 or exactly 2.

Normality/Jarque-Bera.	0.0000 (i)	No Normality

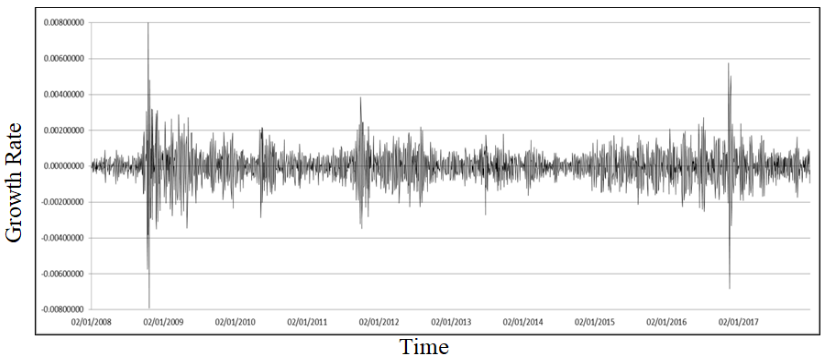
Source: Own elaboration in Eviews.

The results obtained when valuing the ARIMA model of the exchange rate are presented in table 4. The model complies with no first-order serial autocorrelation given the Durbin-Watson result. On the other hand, heteroscedasticity is identified in the result of the ARCH test and the residuals of the model are not normal, as indicated by Jarque-Bera.

It is recognized that the methodology does not efficiently capture the disturbances of the growth rate of the exchange rate, but even knowing this problem the modeling structure allows to develop forecast (objective of the use of the method).

Now let's analyze the estimated values of the ARIMA model which are in Figure 5, these show us two main situations: the first and the most important the ARIMA model if it manages to understand the behavior of the growth rate of the exchange rate modeling its changes and abnormalities. The situation indicates that there is information in the history of the time series that allows identifying it because of the behavior of the variable studied.

Figure 5 Values of the ARIMA model estimate for the growth rate of the exchange rate

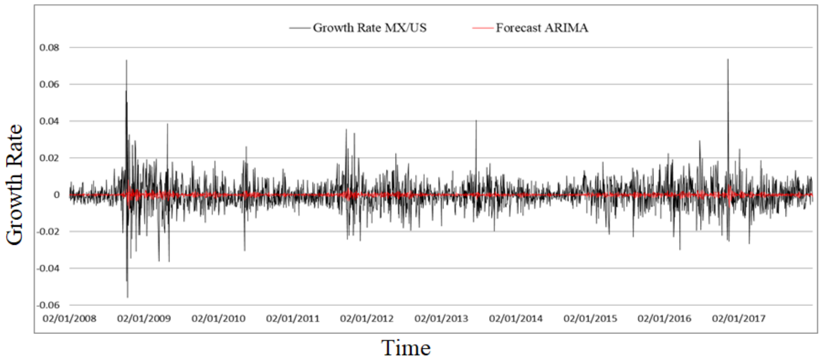


Source: Own elaboration in Eviews.

This is because figure 6 shows the comparison of the estimated values (red line) and the values of the growth rate of the exchange rate. Carrying out a detailed study of this comparison, it was found that the important volatility clusters the ARIMA model can only estimate a part of its behavior as shown in table 4.

Second, it has been found that this methodology does not recognize the magnitude of the aberrant behavior of this time series. The situation that is verified in the heteroscedasticity and normality tests, table 4. The red line denotes the estimate and the black line values the series.

Figure 6. Comparison of the forecast ARIMA model and the values of the growth rate MX/US



Source: Own elaboration in Eviews.

Therefore, it is understood that the forecast that this model can provide will not be the desired one, but it is admissible to perform evaluation tests to compare this result with those of the other models presented in the investigation. In addition to this model is expected that due to problems of Heteroskedasticity and non-normality of the residuals, the confidence interval of the forecast tends to grow, making it more complicated to find an accurate value of the estimate as to the realizations increase.

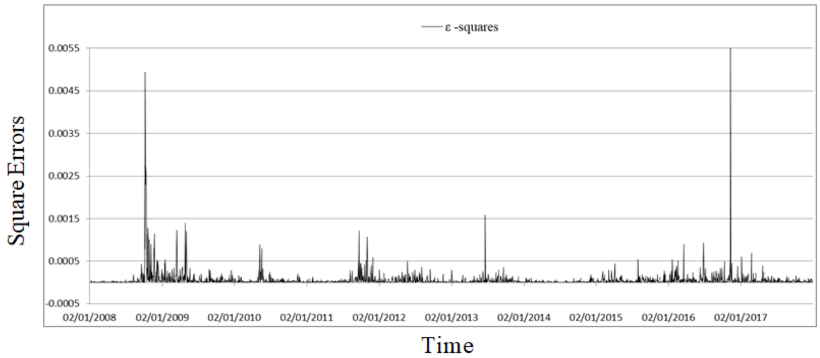
2.3. Models of conditional variance.

In this part of the present chapter, we study the models of the GARCH family (conditional variance) applied to the study variable because there is evidence of the presence of heteroscedasticity in the previous model (Table 4). Therefore, the main theoretical considerations of these models are shown and the reason for the use of certain models for forecasting is explained.

In the first instance, the conditional variance methodology comprises two main groups; the symmetry models in which the conditional variance is identified by similar magnitudes in various situations (there will be high volatility with negative information and low volatility with certainty environments). In the case of asymmetry models, the way to understand the conditional variance is assuming various magnitudes of the innovations, the impacts of situations of high and low volatility produce diverse situations (Karlsson, 2002).

Continuing, the LM-ARCH test (Table 4) shows the existence of ARCH processes in the ARIMA model, a situation whereby it is established that the model must and must be estimated using the conditional variance methodology. Now, in the first instance, they must observe the square errors of the ARMA model and, based on this, decide with which category of GARCH models to model the growth rate of the exchange rate.

Figure 7. Square errors of the ARIMA model



Source: Own elaboration in Eviews.

Considering the above, the two situations that characterize the GARCH models and studying the behavior of the squared errors of the ARIMA model (figure 7) seek to identify the most adequate methods to estimate the exchange rate.

In this case, the behavior of the squared residuals indicates that the ARMA methodology is not capable of understanding the variability in the realizations of the exchange rate. Therefore, the alternatives to achieve the study of the dynamics of the variable are the symmetric and asymmetric models of the GARCH family.

In this sense, figure 7 reveals that in the case of high and low volatility events, innovations tend to have different magnitudes, as can be seen in the points of greatest uncertainty in the sample in October 2008 and November 2016. Now, given the asymmetric behavior of errors what is understood; how the magnitudes of volatility before negative and positive events are different in time.

Finally, this result is shown by establishing two models to search and analyze the dynamics of the behavior of the exchange rate from the asymmetric methodology of the GARCH models. Two models of great relevance are estimated according to (Karlsson, 2002) in the case of financial variables: "Exponential Generalized Autoregressive Conditional Heteroskedasticity Model" (E-GARCH) and "Asymmetric Power Autoregressive Conditional Heteroskedasticity Model" (PARCH). It will be implemented in the exchange rate and later will be supported to meet the objective of the present investigation.

2.3.1. EGARCH.

The next point deals with the main characteristics of the E-GARCH model and the justification for its use as a methodology for forecasting the time of change. In addition, the evaluation tests of the model and the results obtained from the estimation will be presented.

In the first instance in Nelson (1991), the methodologies called "Exponential Generalized Autoregressive Conditional Heteroskedasticity Model" are shown for the first time; this is born because sometimes the construction of the non-negativity of the parameters of the GARCH model is necessary, this when the estimation of the parameters fall in the restriction that the sum of them is greater than one. Therefore, under this methodology, the problem of restriction is solved, making it possible to estimate the variance. The variance of the model is modeled with the following equation, where we find the characteristic conditions of an ARMA in variance except that the dependent variable is the natural logarithm of the variance.

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \, g(y_{t-i}) + \sum_{j=1}^{q} \beta_j \ln(\sigma_{t-j}^2)$$

The section, $\beta_j \ln(\sigma_{t-j}^2)$ specifies the GARCH part of the methodology plus the function that allows modeling the condition of asymmetry in the variance, defined as $g(y_{t-i})$. This can be represented as:

$$g(y_{t-i}) = \theta_1 y_t + \theta_2[|y_t| - E|y_t|]$$

Where the sign effect will be given by$\theta_1 y_t$, in other words, the impact of the information whether positive or negative will be given by this component. On the other hand, the magnitude of the effect will be given by $\theta_2[|y_t| - E|y_t|]$ this component denotes that both events of high or low volatility will influence the model (Karlsson, 2002).

What is intended is to understand the behavior of the asymmetry observed in figure 8, assuming that the sign effect denotes whether there is high or low volatility in the market, such that the effects of innovations tend to be positive and negative. Knowing this situation, the same model determines through the behavior of past innovations the magnitude of the perturbations before the environment under which the value of the exchange rate is valued.

2.3.1.1. GARCH Structure of the Exchange Rate.

Now, the previous section states that it is more viable to model under the criteria of conditional variance. In Figure 7, asymmetry is observed in square innovations, as already known innovations have different magnitudes; the estimation is developed from the asymmetric GARCH methodology. Now, the growth rate of the exchange rate is specified by the E-GARCH model.

Table 5. E-GARCH model for the growth rate of MX/US

Variable.	Coefficient.	Statistical significance[7].
AR(1)	0.063427	***
AR(4)	-0.043480	**
MA(8)	0.042681	**
C(4)	3.73E-07	***
C(5)	0.093158	***
C(6)	0.907075	***

[7] Statistical significance at 99% (***), 95% (**) and 90% (*).

Source: Own elaboration in Eviews.

Table 5. shows the coefficients associated with the estimation of the E-GARCH model. It is illustrated that the coefficients for the mean ARMA equation are statistically significant, of which it is highlighted that only the $AR(1)$ passes the test at 99% confidence while the other parameters only do it at 95%. On the other hand, the ARMA coefficients of the variance are all 99% significant, these denoted as $C(i)$ (See section 2.3.1 to identify the coefficients).

Table 6. Tests performed on the E-GARCH model for the growth rate of the exchange rate

Test/ name.	Probability (i) /Statistic (ii).	Interpretation.
Heteroskedasticity / LM-ARCH.	0.4474 (i)	No Heteroskedasticity.
Autocorrelation / Durbin-Watson.	2.027260 (ii)	No Autocorrelation.
Normality / Jarque-Bera.	0.0000 (i)	No Normality.

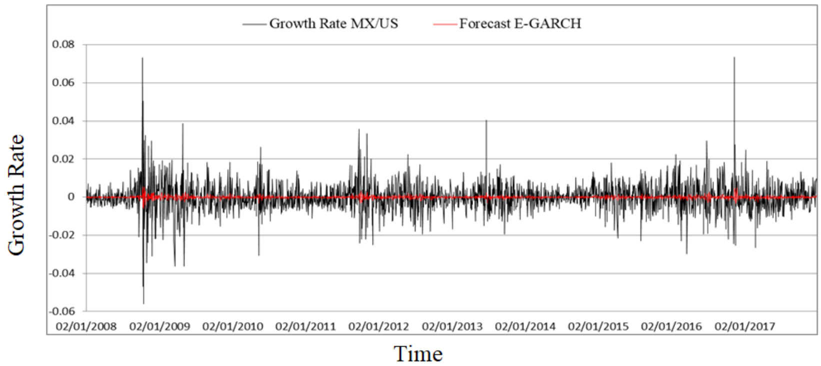
Source: Own elaboration in Eviews.

On the other hand, Table 6 contains the results of the tests carried out, showing the non-existence of ARCH processes, which denotes the non-heteroscedasticity of the residues. In addition, the non-existence of the first-order autocorrelation is maintained. The Jarque-Bera test shows the non-normality of the errors, so it is said that the distribution of the innovations does not have a normal behavior; a situation that indicates that innovations must follow the behavior of another distribution function.

Now, continuing with the main objective of this model for research, Figure 8 shows the comparison between the values of the growth rate of the exchange rate and the E-GARCH estimate. The red line represents the estimated values and the black line represents the growth rate of the exchange rate; it can be observed that when modeling the conditional variance assuming conditional variance of the exponential type, the adjustment of the model is below the values of the original series. As it is possible to observe the red line, it models the disturbances but not the magnitude; this implies that the magnitude effect of the function $g(\varepsilon_{t-i})$fails to understand the inertia of the time series.

Then, the adjustment that was found is not enough to recognize the changes and abnormalities of the modeled series. It is recognized that even considering the variance asymmetry modeling of the exponential type, this is not enough to correctly identify the volatility of the exchange rate.

Figure 8 Comparison of the forecast E-GARCH model and the values of the growth rate MX/US



Source: Own elaboration in Eviews.

Finally, the results show that the E-GARCH methodology has not allowed model all the changes and abnormalities that the growth rate of the exchange rate presents. The Situation shows that it motivates us to continue with the study of the asymmetric models of the GARCH family, and thus, to look for the best possible adjustment of the time series. Therefore, the PARCH model is proposed which will be modeled and explained in the next section.

On the other hand, the objective of the present investigation is to perform modeling of the type and make the forecast. The results shown by this methodology are important for the next stages of the present investigation. This is because they will be used to answer the research question, and thus corroborate or reject the hypothesis.

2.3.2. PARCH.

Consider, the model known as the "Asymmetric Power Autoregressive Conditional Heteroskedasticity Model" (PARCH). Which considers as a starting point the modeling of the conditional variance under a mathematical transformation; it seeks to linearize non-linear processes through the CES function. Returning to Karlsson (2002) describes the ARMA equation of variance as follows:

$$\sigma_{t-j}^{\delta} = \alpha_0 + \sum_{i=1}^{p} \alpha_i \left(|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i}\right)^{\delta} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^{\delta}$$

$$\alpha_0 > 0, \delta \geq 0$$

$$\alpha_i \geq 0 \quad i = 1, \cdots, p$$

$$\beta_j \geq 0 \quad j = 1, \cdots, q$$

$$-1 < \gamma_i < 1 \quad i = 1, \cdots, p$$

This model is developed by Ding *et al*. (1993) as a proposal for modeling the conditional variance of GARCH models. This methodology imposes an enhanced transformation of the process called conditional standard deviation, taking the absolute asymmetry of the residuals, it is sought to linearize the square errors. This allows specifying the ARMA equation of variance according to the method of analysis that identifies the best specification (for more details see Karlsson (2002)).

Recalling that the asymmetric methodology responds to negative and positive disturbances, such as it is observed in the financial series. In other words, we seek to model the behavior of a series of time that is affected by the good and bad news that exists in the economic environment. This version of the asymmetric models comprises 7 different forms of specification of the conditional variance (EGARCH, TARCH, etc.). But for the present investigation, the PARCH model is modeled assuming that this is the best way to find the dynamics of the conditional variance. In other words, an attempt is made to recognize the existence of biases in the information due to positive or negative news that affects the asset

2.3.2.1. PARCH structure of the Exchange Rate.

In this section, the growth rate of the exchange rate is estimated using the PARCH methodology. In the first instance, we have the coefficients of the equation of mean and variance, from which we seek to evaluate the statistical significance of the estimated parameters; secondly, for the residues, different hypothesis tests are developed that allow evaluating the efficiency of the model; and finally, the graphical results of the estimation of both the exchange rate and its growth rate are presented.

In table 7 shows the coefficients of the equation of mean and variance of the PARCH model; the mean ARMA coefficients are statistically significant at 99% confidence; also the parameters of the variance equation (are expressed as $C\ (i)$).

Table 7. Parameters of the PARCH model for the growth rate of the exchange rate

Variable.	Coefficient.	Statistical significance.
AR(1)	0.065165	***
MA(8)	0.065200	***
C(3)	6.97E-05	***
C(4)	0.085213	***
C(5)	-0.444076	***
C(6)	0.925306	***

Source: Own elaboration in Eviews.

Table 8 shows three hypothesis tests to evaluate model residuals. First, the LM-ARCH test is used to find the existence of processes that need to be introduced within the conditional variance; the probability value of the test indicates that there is no information in the residuals evidencing the presence of heteroscedasticity.

Table 8. Tests performed on the E-GARCH model for the growth rate of the exchange rate

Test / name.	Probability (i) /Statistic (ii).	Interpretation.
Heteroskedasticity / LM-ARCH.	0.8648 (i)	No Heteroskedasticity.
Autocorrelation	2.022979 (ii)	No Autocorrelation

/ Durbin-Watson.		
Normality / Jarque-Bera.	0.0000 (i)	No Normality

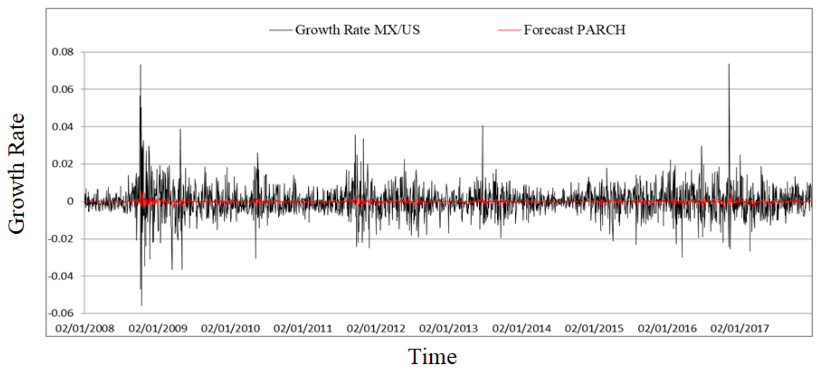
Source: Own elaboration in Eviews.

On the other hand, it is necessary to know if there is serial autocorrelation in the errors (which indicates that there is information in the mean equation that is not being taken into account), the Durbin-Watson statistic is used which indicates according to its result that there is no first-order autocorrelation in the errors.

As the last test, the normality of the waste is evaluated under the Jarque-Bera method, the probability value indicates that the waste is not distributed as normal. Same situation that was evidenced in the previous model, so it is concluded that innovations are not distributed Gaussian, which implies that these errors must follow the behavior of some other distribution, however for the purposes of this research, the result does not present great relevance since only the prediction of the series is of interest.

Figure 9 illustrates the behavior of the rate of growth of the exchange rate (black line) and the PARCH estimate (red line), as observed the adjustment of the model manages to understand the changes and abnormalities of the exchange rate. However, the magnitudes of these variations cannot be estimated; the methodology lacks efficiency to estimate the series of time studied.

Figure 9 Comparison of the forecast PARCH model and the values of the growth rate MX/US



Source: Own elaboration in Eviews.

Finally, the model presented in this section identifies the effects that positive and negative news has on the behavior of the time series, but is unable to distinguish the magnitude they may have. For the purposes of this investigation, this result allows the development of a prognosis under this methodology, to later compare it in subsequent sections.

2.4. Nonlinear Autoregressive Neural Network (NARNET).

Up to now, only linear methodologies (ARMA) and conditional variance (asymmetric models) have been used. However, in the present investigation, it is considered important to study also the non-linear methods, given that in the analysis of the forecast models of the financial series the use of models of artificial neural networks is indispensable.

In the first chapter of the present investigation a brief review of the main neural networks is made, as well as, of their components and forms of learning, but without presenting any forecast model applied to financial time series.

In this section, the "Non-linear Autoregressive Neural Network (NARNET)" methodology is developed in a theoretical manner, in addition to presenting its application to the estimation of the growth rate of the exchange rate. In an initial, recapitulating the existence of neural networks of the type "Single-layer" and "Multi-layer" that is simply understood as the existence of a single layer in the first case or more than one in the second case. In summary, the multi-layered structure with the learning algorithms of the feed type backward is used for this model, hoping to find the structure that best estimates the behavior of the exchange rate. The functional form of the NARNET according to Benmouiza *et al.* (2013), is expressed as follows:

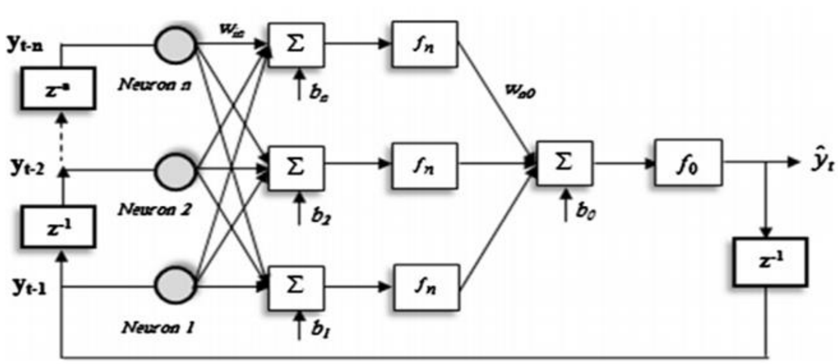
$$y_t = f(y_{t-1} + y_{t-2} + \cdots + y_{t-p})$$

Where it is emphasized that y_t is the estimate of the time series studied from the neural network and $f(y_{t-1} + y_{t-2} + \cdots + y_{t-p})$ is a non-linear function that seeks to understand the behavior of the variable analyzed. These elements express the input of information to the network and the result that will be obtained under the learning process in the hidden layers.

$$y = f(b + \sum_{i} w_i x_i)$$

What represents the output equation of the neurons, where b is the bias of the neurons, $w_i x_i$ the weight associated with the output values of the neurons and finally and is the output or estimate of the time series.

Figure 10 Structure of Nonlinear Autoregressive Neural Network (NARNET)



Source: Research of Benmouiza *et al.* (2013).

In figure 10, the structure followed by the neural network used in this section to forecast the growth rate of the exchange rate is exemplified in a general way. Now, it must be explained that the learning process of these networks is of the "Back-propagation" type, as shown above; if the network does not achieve stability it will return to learn again in previous stages as it is observed in the bottom line of the network diagram. It should be noted that the training algorithms used by this network will be those implemented in the MatLab program within their NTSTOOL; which are mainly the "Levenberg-Marquardt", "Bayesian Regularization," Scaled Conjugate Gradient ". These are based on the backward feeding structure (as in the network described above). For the purposes of the present investigation, only those mentioned above were used (for more information on the training algorithms see NTSTOOL, MatLab).

2.4.1. NARNET structure of the Exchange Rate.

In this section, the main results of the estimation of the Autoregressive Neural Network of the growth rate of the exchange rate are explained. It is important to note that the estimates and the generation of the neural networks were implemented in the MatLab program and the results shown are exclusively obtained from the same software. Several network structures were considered to look for the dynamics of the series of time studied, where the number of lags implemented to make the network and the number of neurons in the hidden layers or their size varied. The "Mean Quadratic Error" was used as an objective function for the neural network to search for stability and it was also used to identify the number of lags and neurons that best learned from the inertia of the time series.

In table 9, the various structures used are shown, as well as the results of the objective function of each of them. As can be seen, not only are they divided into the number of lags and neurons, but they are also characterized by the training algorithm that was implemented for their estimation.

Table 9. Comparison of the NARNET models: objective function "Mean Square Error"

Nonlinear Autoregressive Neural Networks				
Train Function.	Feedback Delays/ hidden Layer Size.			
	01 / 09	02 / 09	03 / 09	04 / 09
LM.(i)	5.54E-05	5.22E-05	5.21E-05	5.73E-05
BR.(ii)	5.24E-05	5.15E-05	5.22E-05	5.71E-05
SCG.(iii)	1.01E-04	6.53E-05	5.96E-05	7.21E-05
	01 / 10	02 / 10	03 / 10	04 / 10
LM.(i)	5.97E-05	5.55E-05	5.21E-05	5.90E-05
BR.(ii)	5.28E-05	5.19E-05	4.79E-05	**4.52E-05**
SCG.(iii)	7.17E-05	5.57E-05	8.25E-05	6.69E-05
	01 / 11	02 / 11	03 / 11	04 / 11
LM.(i)	5.83E-05	5.21E-05	5.07E-05	4.74E-05
BR.(ii)	5.58E-05	5.39E-05	4.82E-05	4.81E-05
SCG.(iii)	5.54E-05	6.41E-05	6.34E-05	5.59E-05
	01 / 12	02 / 12	03 / 12	04 / 12
LM.(i)	5.53E-05	5.41E-05	5.05E-05	4.87E-05
BR.(ii)	6.95E-05	5.14E-05	4.94E-05	4.81E-05
SCG.(iii)	6.25E-05	5.98E-05	6.93E-05	5.61E-05

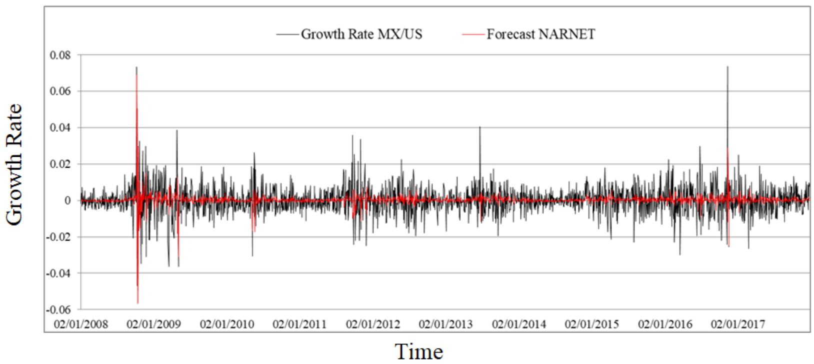
Training algorithm: (i) *Levenberg-Marquardt*, (ii) *Bayesian Regularization*, (iii) *Scaled Conjugate Gradient*.

Source: Own elaboration in MatLab.

The results obtained were that the best training was the "Bayesian Regularization", which is what allows in most cases the minimum value of the objective function. In addition, the structure with 4 lags and 10 neurons was the one with the lowest "Mean Quadratic Error". Therefore, to perform the comparative analysis (objective of this chapter) the structure mentioned with the training algorithm (BR) is used.

As figure 11 shows, the adjustment of the neural network is better in times of greater volatility, as observed in September-December 2008 and October-November 2017. In environments where there is economic stability the neural network decreases its learning capacity, therefore, the estimation in these segments has lower efficiency.

Figure 11. Comparison of the forecast NARNET and the growth rate of the exchange rate



Source: Own elaboration in MatLab.

The neural network presented better the estimation of the growth rate of the exchange rate, which means that the non-linear methodology shows better results than the non-linear ones. Therefore, the presented structure better recognizes the inertia of the time series than the other models presented in this section.

In conclusion, the presented methodology better identifies situations of greater uncertainty but instability environments; these are not estimated in a good way. Therefore, this model does not show the desired fit for the optimal prognosis of the study series; a situation that makes it necessary to review other forecasting methods and other structures of neural networks.

2.5. Valuation of forecasts.

In the previous sections, the results of the estimation of various models for the rate of growth of the exchange rate are shown, so we proceed to evaluate their efficiency. In the foreground, the analysis is carried out within the sample, it is sought to identify which is the model that minimizes the estimation errors in the study sample. And later, the efficiency of developing out-of-sample prognosis is evaluated (January 1 to February 7, 2018). In the sample, the "Absolute Mean Error" (AME) is used, which can be obtained by:

$$AME = \frac{\sum_{i=1}^{N} |y_{it} - \hat{y}_{it}|}{N} * 100 = \frac{\sum_{i=1}^{N} |\varepsilon_t|}{N} * 100$$

Where, y_t is the value of the exchange rate and $\hat{y}_t$ the estimate, so the difference represents the estimation errors ε_t, from which the absolute value is taken and the absolute average error is obtained; N represents the number of observations used for the calculation of the statistic. For its best visualization, it is multiplied by a hundred obtaining the AME.

Table 10. Comparison of Absolute Mean Error

In Sample	
Model	Daily AME
ARIMA	0.5299 %
E-GARCH	0.5274 %
PARCH	0.5281 %
NARNET	0.5110 %

Source: Own elaboration in Excel.

Table 10 shows the Comparison of the Absolute Mean Error Percentage of the exchange rate and its growth rate. It identifies that the model that best estimates the behavior of the series studied is the Autoregressive Neural Network; where the absolute average error is 0.51% per day for the growth rate, there is up to a pointless error than the linear models.

Taking the comparison from the exchange rate, the absolute daily average error of the network is up to 1.3% lower than the ARMA structure models. Within the linear models, the one that shows the best results is the E-GARCH, followed by the PARCH and the ARIMA.

On the other hand, it is necessary to do the evaluation outside the sample to identify which the forecast efficiency of each model. The Percentage of the Absolute Mean Error is used for 26 periods ahead, what is done is to develop the static forecast of each model in one step and the new value is used to obtain the next period.

Table 11. Comparison of the Absolute Mean Error

Out Sample	
Model	Daily AME
ARIMA	0.5982 %
E-GARCH	0.5925 %
PARCH	0.5922 %
NARNET	0.5915 %

Source: Own elaboration in Excel.

Table 11 shows the evaluation of the models outside the sample where it is identified that the network continues to be the best forecasting model. However, these results are not conclusive as a consequence of the fact that the error outside the sample is significantly high, reaching up to 74%.

2.6. Conclusion of Methodology Chapter.

In this chapter, the study of the forecast models of the exchange rate is developed. In the foreground, the study variable does not meet the criterion of stationarity situation that forces to work with a transformation of the same, in this case, the growth rate. Subsequently, the time series is estimated under the ARIMA method; in which errors present two problems, the first is that heteroscedasticity exists and the second is that innovations are not normal. This situation implies that the problem of the ability to models the behavior of the time series.

The ARMA results give guidelines to develop models of conditional variance and since the behavior of square innovations presents great variability on a different scale, the modeling is developed with the asymmetric GARCHs. In the EGARCH the sign effect correctly identifies the positive and negative signals but fails to estimate the magnitude of these. Then, the methodology of the Autoregressive Neural Network is estimated, which is evaluated with the mean square error, different estimation structures are compared with different learning algorithms. The results were that the best network was the one with 4 lags and 10 neurons with the Bayesian regularization learning algorithm.

Finally, each of the estimated models is compared where the methodology that presents the best fit within the sample was the neural network. However, the results outside the sample identify that the neural network loses efficiency when making predictions, so it is difficult to identify which is the best forecast model outside of the sample given that the least mean square error is the network but the one that has better Theil coefficient is the ARIMA. In conclusion, this section provides results that will be used to develop a comparison with the fuzzy models of the next chapter and dynamic neural networks. This tries to identify which is the model that best adjusts to changes and abnormalities of the exchange rate.

CHAPTER III. *MODELS OF FUZZY TIME SERIES AND FORECAST OF THE EXCHANGE RATE*

"In general, complexity and precision bear an inverse relation to one another in the sense that, as the complexity of a problem increases; the possibility of analyzing it in precise terms diminishes. Thus 'fuzzy thinking' may not be deplorable, after all, if it makes possible the solution of problems which are much too complex for precise analysis".

Lotfi A. Zadeh

This chapter presents several methods of Fuzzy Theory, applied to forecast the foreign exchange market. The next models are developed and designed by the authors of the present research; the calculations were made in Excel. First, it does develop the theoretical frame of each model and its applicability to prediction the exchange rate Mexican Peso/ American Dollar. Initially is described the Hybrid Fuzzy Time Series and Fuzzy ARIMA Model and its principal results, the Gaussian parameters Fuzzy GARCH is showed in section 3.2, the Gaussian parameters Fuzzy EGARCH is showed in section 3.2, the Triangular Fuzzy Nonlinear Autoregressive Neural Network and finally the Trapezoidal Fuzzy Nonlinear Autoregressive Neural Network. In addition in the last section is presented a Differential Neural Network and the comparison of each model of this thesis is analyzed.

3.1. A Hybrid Fuzzy Time Series and Fuzzy ARIMA Model.

The fuzzy time series model has shown various fields of application of which the foreign exchange market stands out. It worked to define the fuzzy logical relationships between different factors and defuzzification from the data of the time series. But in the case of the fuzzy set is not enough to model behaviors that come out of it, the fuzzy time series model loses forecasting capacity. For that, the universe of discourse (1.1) will be the growth rate of the time series data.

Assumption 3.1. The probability of passing from one fuzzy subset to another is equal to the probability existing in a previous step.

$$P_t(A_{t-1} \rightarrow A_t) = P_{t-1}(A_{t-2} \rightarrow A_{t-1}) \tag{3.1}$$

Where $\mathrm{P_t}$ is the probability in the actual time and $\mathrm{P_{t-1}}$ in a previous step. A_{t-1} is the fuzzy subset in a last step and A_t in the actual step, $A_{t-1} \rightarrow A_t$ this represents the change from the past period to the current one. And finally, $A_{t-2} \rightarrow A_{t-1}$ it is the change of the period two periods previous to the past.

Assumption 3.2. The fuzzy logical relationship of passing from one fuzzy subset to another is equal to the fuzzy logical relationship existing in a previous step.

$$R_t(A_{t-1} \rightarrow A_t) = R_{t-1}(A_{t-2} \rightarrow A_{t-1}) \tag{3.2}$$

The possibility matrix represents the probabilities of change of a fuzzy subset to another; this is obtained from the historical distribution of the data, the sum in each column is 1.

$$PM = \begin{pmatrix} P(A_1 \rightarrow A_1) & \cdots & P(A_1 \rightarrow A_k) \\ \vdots & \ddots & \vdots \\ P(A_k \rightarrow A_1) & \cdots & P(A_k \rightarrow A_k) \end{pmatrix} \tag{3.3}$$

The relationship matrix represents the relationship of a fuzzy subset to another; this is obtained from the historical distribution of data. The value in each element will be one if a fuzzy subset can go to another and zero in the case that it cannot go to another.

$$RM = \begin{pmatrix} R(A_1 \rightarrow A_1) & \cdots & R(A_1 \rightarrow A_k) \\ \vdots & \ddots & \vdots \\ R(A_k \rightarrow A_1) & \cdots & R(A_k \rightarrow A_k) \end{pmatrix} \tag{3.4}$$

The index t refers to the number of non-fuzzy data used for constructing the model. From the matrix of possibility and fuzzy relations, the **first fuzzy rule** is found.

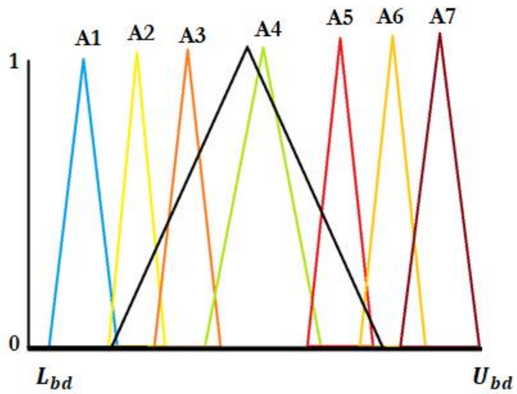
$$IF: \quad P(A_n \rightarrow A_m) > 0$$

AND

$$R(\mathrm{A_n} \rightarrow \mathrm{A_m}) > 1$$

THEN: there is the possibility of the forecast, in otherwise there not forecast. And the **second fuzzy rule** is, IF: There is the possibility of the forecast; THEN: the forecast is the maximum probability of the possibility matrix associated with the three estimated values in the Fuzzy-ARIMA model.

Figure 12. Memberships function of an FTS-fuzzy ARIMA model



Source: Own elaboration.

Figure 12 represents the membership function of an FTS-fuzzy ARIMA model, where the triangular function of black color is the forecast Fuzzy ARIMA and the others belong to each fuzzy subset of (1.1).

The proposed model has the next phases:

I. Define the fuzzy set U associated with the historical data, such as Song and Chissom (1993a) but taking the growth rate of the time series, not the time series.
II. Partition the universe U into several even lengthy intervals Chen and Hsu (2004), where the two intervals with the highest data concentration will be repartitioned.
III. Fuzzify the historical data, i.e., find out an equivalent fuzzy set to each growth rate.

IV. Obtain the values of the distribution of the fuzzy subsets and make the matrix of possibility and relationships.

V. Define the "If-then" rules for the process and Defuzzification of fuzzified time series values from the Tseng's model (2001) or Tanaka's model (1882)

VI. Obtain the forecasted outputs.

This model allows recognizing all possible situations that the variable can take and with the Fuzzy ARIMA shows the best, intermediate and worst. Then of the possible forecasts, the most probable one will be obtained.

3.1.1. Application to forecast exchange rate of MX peso to US dollar

The exchange rate is a variable of great importance because it determines some of the decisions of the agents of the economy. In addition, it is used to carry out national or international transactions. Within this point, the influence of trade, consumption, investment, and intermediate prices are recognized.

In the case of Mexico the exchange rate MX peso (Mexican peso) to US dollar (USD). It represents a variable of great importance because it allows identifying the status of the economy, as well as, influences the decision making of the agents. Therefore, understanding the behavior of this series and developing a more accurate forecast takes great relief for the Mexican economy and its agents.

In first, the series that is used for the present model is taking from the publications of the Bank of Mexico in specific the series "FIX" in daily temporality from January 02, 2008, to December 29, 2017, consists of 2514 observations. After are added 26 observations.

3.1.1.1. The forecast

Applying the hybrid fuzzy time series and fuzzy ARIMA method, are used 2514 observations to formulate the model and the next 26 observations to evaluate the performance of the model.

Phase I: Define the fuzzy set U associated with the historical data.

$$U = [-0.08, 0.08] \quad \textbf{(3.5)}$$

Phase II: Partition the universe (1.1) into several even lengthy intervals, where the two intervals with the highest data concentration will be repartitioned, then it haven eighteen fuzzy subsets.

Table 12. Fuzzy subsets of the exchange rate of MX peso to US dollar

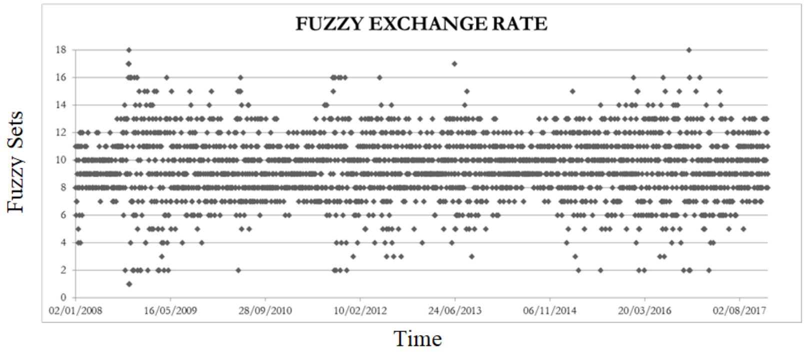
	L_{bd}	U_{bd}
A1	-0.06	-0.04
A2	-0.04	-0.02
A3	-0.02	-0.01714
A4	-0.01714	-0.01429
A5	-0.01429	-0.01143
A6	-0.01143	-0.00857
A7	-0.00857	-0.00571
A8	-0.00571	-0.00286
A9	-0.00286	0.00000
A10	0.00000	0.00286
A11	0.00286	0.00571
A12	0.00571	0.00857
A13	0.00857	0.01429
A14	0.01429	0.01714
A15	0.01714	0.02
A16	0.02	0.04
A17	0.04	0.06
A18	0.06	0.08

Source: Own elaboration in Excel.

Table 13 shows the fuzzy subsets of (1.1), where it is emphasized that unlike what is proposed by Song and Chissom (1993a) it does not have the same size because the two intervals with the highest data concentration were repartitioned in seven subsets every one.

Phase III. Fuzzify the historical data; find out an equivalent fuzzy subset to each diary rate growth of the exchange.

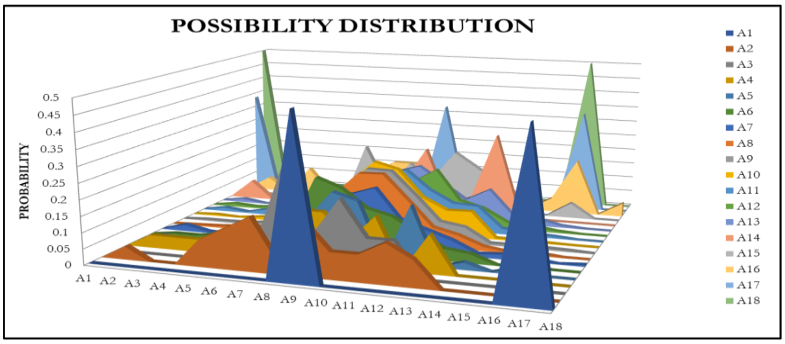
Figure 13. Fuzzy historical data of rate growth of the rate exchange MX/US



Source: Own Elaboration in Excel.

Phase IV. Obtain the values of the distribution of the fuzzy subsets and make the matrix of possibility and relationships.

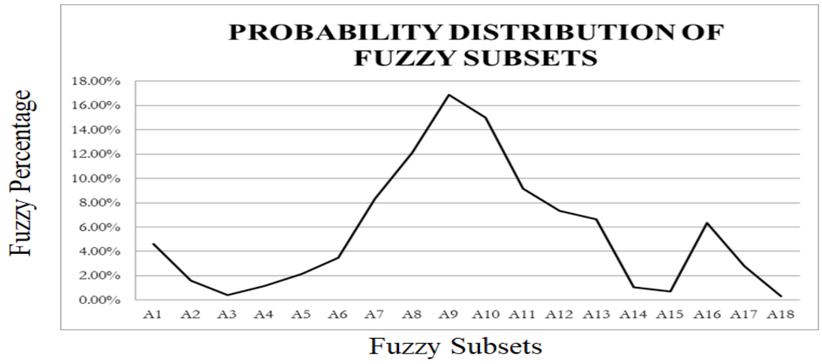
Figure 14. Possibility distributions of fuzzy subsets of diary rate growth of the exchange rate MX/US



Source: Own elaboration in Excel.

In this phase, all the information can get from the growth rate of the time series; it is assumed that the necessary expert system in fuzzy theory is the market that defines the value of the exchange rate.

Figure 15 the probability distribution of the fuzzy subset of the exchange rate MX/US



Source: Own elaboration in Excel.

In this case, figure 15 shows the possibility of moving from one fuzzy subset to another, for example, if in the period $F(t-1)$ the fuzzy exchange rate was A1 there is a probability of 50% change to fuzzy subset A9 and A17 in the next period $F(t)$, all the probabilities of change can be seen in (3.6).

$$PM = \begin{array}{cc} FTS & \begin{array}{cccccccccccccccccc} A_1 & A_2 & A_3 & A_4 & A_5 & A_6 & A_7 & A_8 & A_9 & A_{10} & A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} & A_{17} & A_{18} \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \\ A_7 \\ A_8 \\ A_9 \\ A_{10} \\ A_{11} \\ A_{12} \\ A_{13} \\ A_{14} \\ A_{15} \\ A_{16} \\ A_{17} \\ A_{18} \end{array} & \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0.041 & 0 & 0 & 0.083 & 0.125 & 0.166 & 0.041 & 0.166 & 0.083 & 0.083 & 0.12 & 0.083 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.111 & 0.33 & 0.11 & 0.22 & 0.11 & 0.11 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.04 & 0.04 & 0.04 & 0.08 & 0.04 & 0.08 & 0.16 & 0.16 & 0.08 & 0.16 & 0 & 0.12 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.022 & 0 & 0 & 0.045 & 0 & 0.2 & 0.11 & 0.20 & 0.18 & 0 & 0.18 & 0 & 0.022 & 0 & 0.02 & 0 & 0 \\ 0 & 0.009 & 0.009 & 0.019 & 0.019 & 0.028 & 0.06 & 0.22 & 0.2 & 0.14 & 0.12 & 0.076 & 0.038 & 0.028 & 0 & 0.009 & 0 & 0 \\ 0 & 0.02 & 0.005 & 0 & 0.02 & 0.046 & 0.1 & 0.18 & 0.16 & 0.19 & 0.09 & 0.072 & 0.046 & 0.01 & 0.01 & 0.01 & 0 & 0.005 \\ 0 & 0.002 & 0.005 & 0.01 & 0.013 & 0.04 & 0.075 & 0.142 & 0.21 & 0.22 & 0.14 & 0.061 & 0.043 & 0.013 & 0.002 & 0.008 & 0 & 0 \\ 0 & 0.004 & 0.004 & 0.006 & 0.014 & 0.036 & 0.073 & 0.148 & 0.21 & 0.21 & 0.14 & 0.063 & 0.05 & 0.006 & 0.006 & 0.006 & 0.004 & 0 \\ 0 & 0.004 & 0 & 0.008 & 0.014 & 0.037 & 0.043 & 0.16 & 0.22 & 0.20 & 0.12 & 0.083 & 0.08 & 0.004 & 0.004 & 0.002 & 0 & 0 \\ 0 & 0.012 & 0.003 & 0.018 & 0.009 & 0.037 & 0.08 & 0.135 & 0.17 & 0.19 & 0.12 & 0.089 & 0.08 & 0.01 & 0.009 & 0.006 & 0 & 0 \\ 0 & 0.01 & 0.005 & 0.01 & 0.026 & 0.041 & 0.093 & 0.145 & 0.16 & 0.13 & 0.18 & 0.083 & 0.06 & 0.02 & 0.01 & 0.005 & 0 & 0 \\ 0 & 0.011 & 0 & 0.005 & 0.023 & 0.087 & 0.075 & 0.11 & 0.15 & 0.17 & 0.12 & 0.069 & 0.11 & 0.02 & 0.02 & 0.005 & 0 & 0 \\ 0 & 0.062 & 0 & 0 & 0.031 & 0.062 & 0.031 & 0.093 & 0.09 & 0.21 & 0.03 & 0.062 & 0.28 & 0.031 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0.05 & 0.15 & 0.1 & 0.2 & 0.15 & 0.1 & 0 & 0 & 0.05 & 0 & 0 \\ 0 & 0.045 & 0 & 0.09 & 0 & 0.045 & 0.09 & 0.13 & 0.13 & 0 & 0 & 0.09 & 0.09 & 0 & 0.04 & 0.18 & 0 & 0.04 \\ 0.33 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.33 & 0 & 0 & 0 & 0 & 0 & 0.33 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 \end{bmatrix} \end{array} \qquad \textbf{(3.6)}$$

In addition, the fuzzy logical relationship between different fuzzy sets or themselves is taken into account with the purpose of identifying those fuzzy sets where there is no possibility of change. For example, if in the period $F(t-1)$ the fuzzy exchange rate was A1 then there is a fuzzy logical relationship with the fuzzy subset A9 and A17, all the relationship can be seen in (3.7).

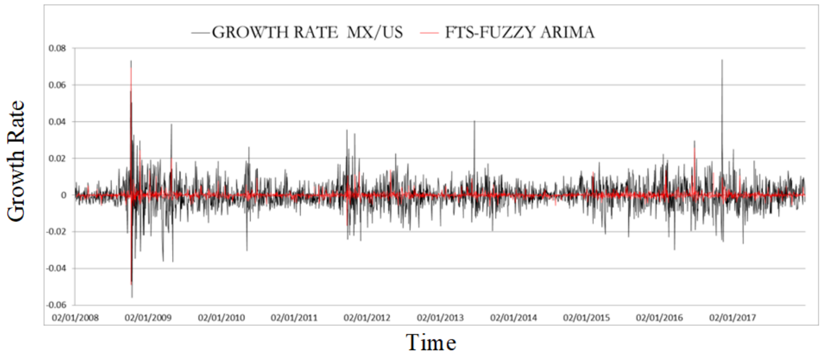
$$RM = \begin{array}{c} FTS \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \\ A_7 \\ A_8 \\ A_9 \\ A_{10} \\ A_{11} \\ A_{12} \\ A_{13} \\ A_{14} \\ A_{15} \\ A_{16} \\ A_{17} \\ A_{18} \end{array} \begin{array}{c} \begin{array}{cccccccccccccccccc} A_1 & A_2 & A_3 & A_4 & A_5 & A_6 & A_7 & A_8 & A_9 & A_{10} & A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} & A_{17} & A_{18} \end{array} \\ \left[\begin{array}{cccccccccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{array}\right] \end{array} \qquad (3.7)$$

Finally, the daily probability of the fuzzy growth rate of the exchange rate is shown in Figure 15. Where, the most probable fuzzy subsets are A8, A9 and A10 and the least likely subsets are A3 and A18. This is understood as if in period $F(t-1)$ there is any fuzzy subset the probability of change to another fuzzy subset in period $F(t)$ is shown by the probability distribution Figure 14.

Phase V. Define the "If-then" rules for the process and the Defuzzification of fuzzified time series values from Tseng's model (2001) or Tanaka's model (1982). The **first fuzzy rule** says that if (3.6) and (3.7) shows the possibility of the forecast then defuzzification is obtained from the **second fuzzy rule.**

Phase VI. Obtain the forecasted outputs; these values are obtained by combining the fuzzy time series and the fuzzy ARIMA method. Then, it considered the best and worst situations combined with the probabilities of change between fuzzy subsets which shows a forecast that takes into account the uncertainty and the memory of the time series (FTS-fuzzy ARIMA model).

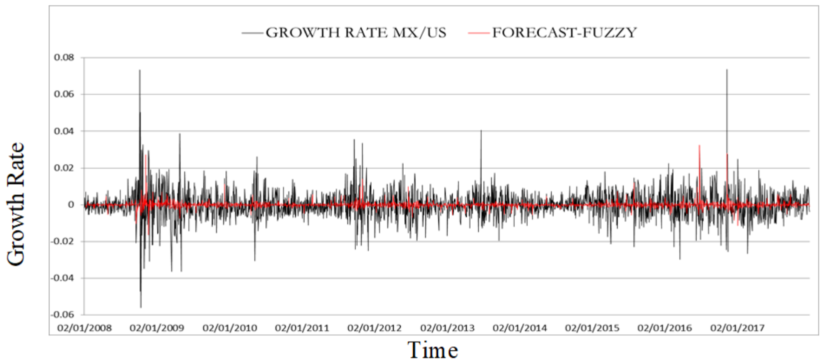
Figure 16. Forecast of FTS-fuzzy ARIMA Tanaka's model



Source: Own elaboration in Excel

Figure 16 shows the forecast of the proposed FTS-fuzzy ARIMA based in the Tanaka methodology, this is represented with the red line and the black line is the growth rate of the exchange rate. The estimated values denoted a better approximation to the time series analyzed in comparison with the traditional ARIMA model.

Figure 17 Forecast of FTS-fuzzy ARIMA Tseng's model



Source: Own elaboration in Excel

In figure 17 is represented the forecast FTS-fuzzy ARIMA of Tseng methodology, it is represented with the red line and the growth rate of the exchange rate with the black line. The estimated values denoted a better approximation to the time series analyzed in comparison with the traditional ARIMA model.

3.2. Gaussian Parameters Fuzzy GARCH

The Generalized Autoregressive Conditional Heteroskedasticity GARCH (p,q) models are used to capture the behavior of the variance in the finance time series (Bollerslev, 1986). This method is categorized as an asymmetric model because assumed that the conditional variance is caused by the magnitude, and not for the negative and positive information. For example, if there is bad news the volatility is higher than other cases and can this methodology not take this information to make the forecast. The GARCH model is represented as:

$$y_t = \sigma(t)\varepsilon(t)$$

$$\sigma^2(t) = \omega + \sum\nolimits_{i=1}^{p} \alpha_i\, y^2(t-i) + \sum\nolimits_{j=1}^{q} \beta_j\, \sigma^2(t-j)$$

Where y_t is a stochastic time series process determined by the $\sigma(t)$ volatility function and $\varepsilon(t)$ a white noise. The equation of the $\sigma^2(t)$ conditional variance is the function of the square lags of y_t and their own lags. This method must comply with the next conditions:

$$\begin{gathered} \omega > 0 \\ \alpha_i \geq 0 \\ \beta_j \geq 0 \\ \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1 \\ i = 1, \cdots p;\ \ j = 1, \cdots, q. \end{gathered}$$

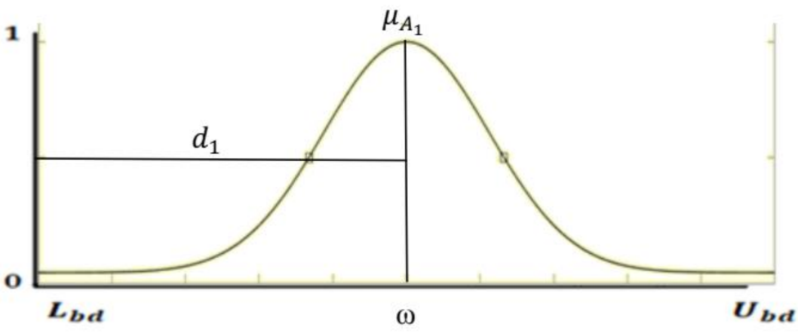
The importance of this model is because can make a forecast of the various exchange rate, but also, for is possible combined this technique with fuzzy theory and make a different forecast of the traditional GARCH.

The important increase in volatility in the Foreign Exchange market causes that the forecast this type of variables be more and more complicated. In this way, the Fuzzy EGARCH with the Gaussian Parameters model looks to improve the predictive values of the various exchange rates. For understanding this method it's necessary to know the traditional GARCH (p,q) model (Bollerslev, 1986).

Assumption 3.3. The membership function of the GARCH parameters (ω, α_i and β_j) is a Gaussian type. In this way, it can be expressed as:

$$\mu_{A_1}(\omega_k) = e^{-\left(\frac{\omega_k - \omega}{\delta_\omega}\right)^2} \quad \textbf{(3.8)}$$

Figure 18. Gaussian Membership Function of ω_k



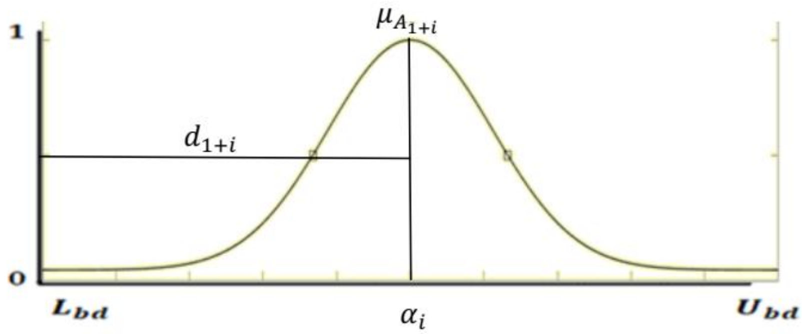
Source: Own elaboration in MatLab.

$$d_1 = 3 * \delta_\omega - \omega$$

Where $\mu_{A_1}(\omega_k)$ represents the membership function of the fuzzy set of the parameter ω_k, ω is the center and δ_ω is the variance. And d_1 is the distance where there is 99% of all parameters in this membership function. In this research is used the letter k how the subscript that defines an element of the membership function associated.

$$\mu_{A_{1+i}}(\alpha_{ik}) = e^{-\left(\frac{\alpha_{ik} - \alpha_i}{\delta_{\alpha_i}}\right)^2} \quad \textbf{(3.9)}$$

Figure 19. Gaussian Membership Function of α_i



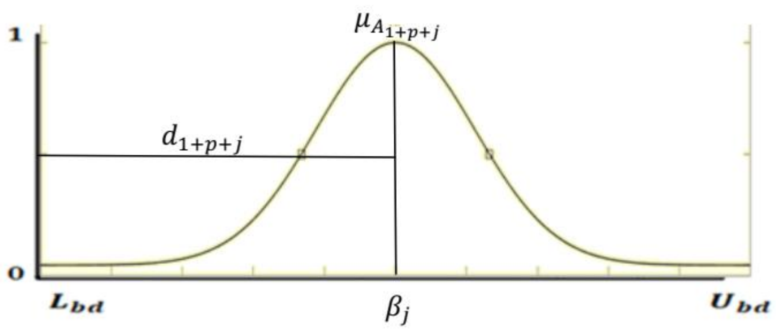
Source: Own elaboration in MatLab.

$$d_{1+i} = 3 * \delta_{\alpha_i} - \alpha_i$$

Where $\mu_{A_{1+i}}(\alpha_{ik})$ represents the membership function of the fuzzy set of the parameter α_{ik}, α_i is the center and δ_{α_i} is the variance. And d_{1+i} is the distance where there is 99% of all parameters in this membership function.

$$\mu_{A_{1+i+j}}(\beta_{jk}) = e^{-\left(\frac{\beta_{jk}-\beta_j}{\delta_{\beta_j}}\right)^2} \quad \textbf{(3.10)}$$

Figure 20 Gaussian Membership Function of β_j



Source: Own elaboration in MatLab.

$$d_{1+p+j} = 3 * \delta_{\beta_j} - \beta_j$$

Where $\mu_{A_{1+i+j}}(\beta_{jk})$ represents the membership function of the fuzzy set of the parameter β_{jk}, β_j is the center and δ_{β_j} is the variance. And d_{1+p+j} is the distance where there is 99% of all parameters in this membership function.

Assumption 3.4. If the $\sigma^2(t)$ conditional variance of a GARCH process is a fuzzy function. Then it can be expressed as:

$$\sigma^2(t) = \mu_{A_1}(\omega_k) + \sum_{i=1}^{p} \mu_{A_{1+i}}(\alpha_{ik})\, y^2(t-i) + \sum_{j=1}^{q} \mu_{A_{1+p+j}}(\beta_{jk})\, \sigma^2(t-j) \quad \textbf{(3.11)}$$

Where (3.11) represents the fuzzy conditional variance of the GARCH process. In this case is necessary to define an error minimization function, for the present research is taken the Mean Absolute Error (MAD) and can be expressed as follows:

$$\epsilon = \frac{E\left[y_t - \hat{y}_t\right]}{n}$$

Then the problem is finding the fuzzy GARCH parameters that can obtain resolving the next linear programming problem:

$$\text{Minimize} \quad \epsilon \tag{3.12}$$

$$\text{Subject to} \quad \begin{array}{ll} d_1 & > 0 \\ d_{1+i} & > 0 \\ d_{1+p+j} & > 0 \end{array}$$

When is originated the parameters that guarantee the minimum ϵ also are found the non-fuzzy forecast of the conditional variance. And the Fuzzy GARCH (p, q) model can be expressed as:

$$y_t = \sigma(t)\varepsilon(t)$$

$$\sigma^2(t) = \omega_k + \sum_{i=1}^{p} \alpha_{ik}\, y^2(t-i) + \sum_{j=1}^{q} \beta_{jk}\, \sigma^2(t-j) \tag{3.13}$$

This method must comply with the next conditions:

$$\begin{gathered} \omega_k > 0 \\ \alpha_{ik} \geq 0 \\ \beta_{jk} \geq 0 \\ \sum_{i=1}^{q} \alpha_{ik} + \sum_{j=1}^{p} \beta_{jk} < 1 \\ i = 1, \cdots p;\;\; j = 1, \cdots, q. \end{gathered}$$

The proposed model has the next phases:

I. Estimate the parameters of the GARCH (p, q) with the Maximum Likelihood method.

II. Use the results of the previous step how the center of the membership functions (3.8), (3.9) and (3.10).

III. Define the distance in the model d_n of each membership function and find the standard deviation as:

$$\delta_\varphi = \frac{d_n + \varphi}{3}$$

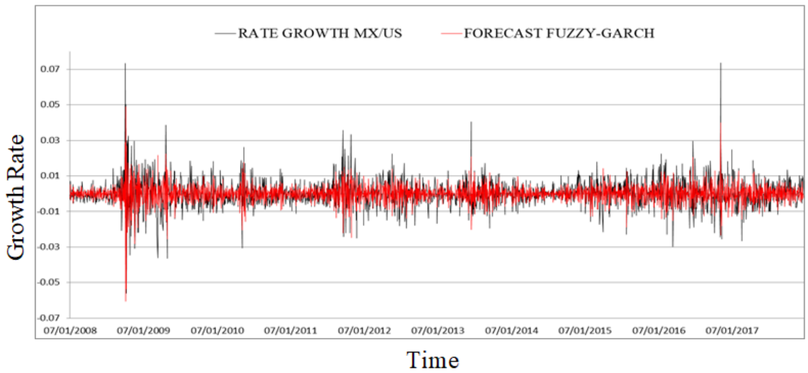
Where n is the lag associated with parameter φ and δ_φ is the standard deviation of the membership function.

- **IV.** Obtain the probability of the membership function with δ_φ, d_n and φ.
- **V.** Determine with the Gaussian function the parameter φ_k using the values φ, δ_φ and the probability of the last step.
- **VI.** Result of the linear programming problem (3.12).
- VII. Finally calculate the fuzzy GARCH (p, q) forecast.

3.2.1. Forecast Gaussian Parameters Fuzzy GARCH (p, q).

The application of the Gaussian parameter Fuzzy GARCH (1, 1) is developed in the present section. First is estimated the GARCH (1, 1) model through the Maximum Likelihood method and with it are obtained the center values of each Gaussian membership function. After is defined every possibility distribution associated with the parameters GARCH (1,1). And determined the Gaussian function the parameter φ_k using the values φ, δ_φ is the possible result in the linear programming problem (3.12) and obtain the Fuzzy GARCH (1,1) forecast (figure 21).

Figure 21. Forecast of Gaussian Parameters Fuzzy GARCH (1, 1)



Source: Own elaboration in Excel.

Figure 21 shows the forecast Fuzzy GARCH (1, 1) (red line) and the growth rate MX/US (black line). It lets us see that the estimated values are similar to the real values, in other words, the model gives us a good prognostic of the volatility present in the foreign exchange market. Up to now, this model is better than the traditional models ARIMA, EGARCH, PARCH and Nonlinear Autoregressive Neural Network.

The Symmetric Conditional Variance model joint with fuzzy theory shows that the fuzzy method is better than the original technique. This is because the fuzzy theory lets identify the uncertainty present in the financial time series, and also detect the moments where higher volatility exist.

3.3. Gaussian Parameters Fuzzy EGARCH $(\mathbf{p}, \mathbf{q})$.

From Nelson (1991) the Exponential Generalized Autoregressive Conditional Heteroskedasticity EGARCH (p, q) is categorized as to how asymmetric model because this method recognized the magnitude and the sign of the information that causes the volatility of the time series. In the important of this methodology is that the conditional variance is a linear combination of y_t and the logarithm of its own lags.

In the present section is developed the Fuzzy Exponential Generalized Autoregressive Conditional Heteroskedasticity Fuzzy EGARCH (p, q). This method can also be categorized how an asymmetric model because this method recognized the magnitude and the sign of the information that causes the volatility in financial time series. From the model of Nelson (1991), it is assumed that their parameters have a Gaussian membership function.

This model is developed because sometimes the construction of the non-negativity of the parameters of the EGARCH (p, q) model is necessary, this when the estimation of the parameters falls in the restriction that the sum of them is greater than one. Therefore, under this methodology, the problem of restriction is solved, making it possible to estimate the variance. The variance of the model is determined for the following equations:

$$y_t = \sigma(t)\varepsilon(t)$$

$$\ln\left(\sigma^2(t)\right) = \omega + \sum_{i=1}^{p} \alpha_i\, g(y_{t-i}) + \sum_{j=1}^{q} \beta_j\, \sigma^2(t-j) \qquad \textbf{(3.14)}$$

The section $\sum_{j=1}^{q} \beta_j \, \sigma^2(t-j)$ specifies the GARCH (p, q) part of the methodology plus the function that allows modeling the condition of asymmetry in the variance, defined as $g(y_{t-i})$.

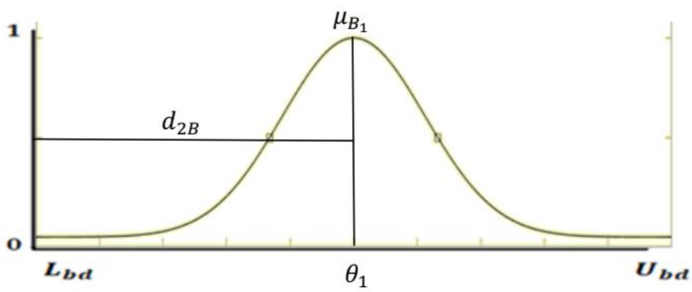
$$g(y_{t-i}) = \theta_1 y_{t-i} + \theta_2[|y_{t-i}| - E|y_{t-i}|] \tag{3.15}$$

Where the sign effect is $\theta_1 y_t$, in other words, the impact of positive or negative information. Moreover, the magnitude is $\theta_2[|y_t| - E|y_t|]$ this equation denotes that both events of high or low volatility have influence in the forecast of the financial time series (Karlsson, 2002).

Assumption 3.5. The membership function of the EGARCH (p, q) parameters ($\omega, \alpha_i, \beta_j, \theta_1$ and θ_2) is a Gaussian type. In this way, it can be expressed as (3.8), (3.9), (3.10) and:

$$\mu_{B_1}(\theta_{1k}) = e^{-\left(\frac{\theta_{1k} - \theta_1}{\delta_{\theta_1}}\right)^2} \tag{3.16}$$

Figure 22. Gaussian Membership Function of θ_1



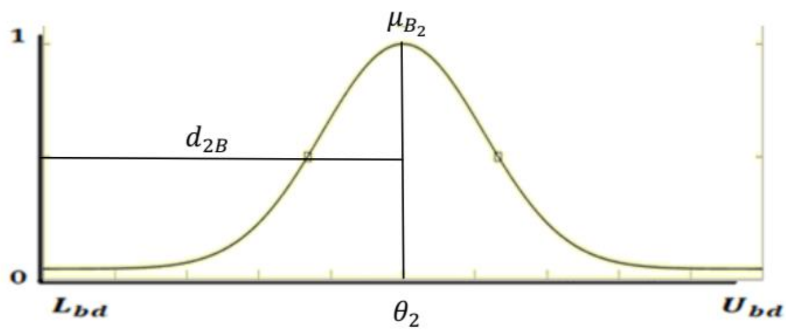
Source: Own elaboration in MatLab.

$$d_{1B} = 3 * \delta_{\theta_1} - \theta_1$$

Where $\mu_{B_1}(\theta_{1k})$ represents the membership function of the fuzzy set of the parameter θ_{1k}, θ_1 is the center and δ_{θ_1} is the variance. And d_{1B} is the distance where there is 99% of all parameters in this membership function.

$$\mu_{B_2}(\theta_{2k}) = e^{-\left(\frac{\theta_{2k}-\theta_2}{\delta_{\theta_2}}\right)^2} \quad \textbf{(3.17)}$$

Figure 23. Gaussian Membership Function of $\boldsymbol{\theta_2}$



Source: Own elaboration in MatLab.

$$d_{2B} = 3 * \delta_{\theta_2} - \theta_2$$

Where $\mu_{B_2}(\theta_{2k})$ represents the membership function of the fuzzy set of the parameter θ_{2k}, θ_2 is the center and δ_{θ_2} is the variance. And d_{2B} is the distance where there is 99% of all parameters in this membership function.

Assumption 3.6. If the $\sigma^2(t)$ conditional variance of an EGARCH (p, q) process is a fuzzy function. Then it can express as:

$$\sigma^2(t) = \mu_{A_1}(\omega_k) + \sum_{i=1}^{p} \mu_{A_{1+i}}(\alpha_{ik})\, y^2(t-i)$$
$$+ \sum_{j=1}^{q} \mu_{A_{1+p+j}}(\beta_{jk})\, \sigma^2(t-j) \qquad \textbf{(3.18)}$$

$$g(y_{t-i}) = \mu_{B_1}(\theta_{1k}) y_{t-i} + \mu_{B_2}(\theta_{2k})[|y_{t-i}| - E|y_{t-i}|]$$

Where (3.18) represents the fuzzy conditional variance of the EGARCH (p, q) process. In this case is necessary to define an error minimization function, for the present research is taken the Mean Absolute Error (MAD) and can be expressed as follows:

$$\epsilon = \frac{E\left[y_t - \hat{y}_t\right]}{n}$$

Then the problem is finding the fuzzy EGARCH (p, q) parameters that can obtain resolving the next linear programming problem:

$$\text{Minimize} \quad \epsilon$$

$$\text{Subject to} \quad \begin{array}{ll} d_1 & > 0 \\ d_{1+i} & > 0 \\ d_{1+p+j} & > 0 \\ d_{1B} & > 0 \\ d_{2B} & > 0 \end{array} \qquad \textbf{(3.19)}$$

When finding the parameters that guarantee the minimum ϵ also is found the non-fuzzy forecast of the conditional variance. And the Fuzzy EGARCH (p, q) model can be expressed as:

$$y_t = \sigma(t)\varepsilon(t)$$

$$\ln\left(\sigma^2(t)\right) = \omega_k + \sum_{i=1}^{p} \alpha_{ik}\, g(y_{t-i}) + \sum_{j=1}^{q} \beta_{jk}\, \sigma^2(t-j) \qquad \textbf{(3.20)}$$

$$g(y_{t-i}) = \theta_{1k} y_{t-i} + \theta_{2k}[|y_{t-i}| - E|y_{t-i}|]$$

This method must comply with the next conditions:

$$\omega_k > 0$$
$$\alpha_{ik} \geq 0$$
$$\beta_{jk} \geq 0$$
$$\theta_{1k} \geq 0$$
$$\theta_{2k} \geq 0$$
$$i = 1, \cdots p; \;\; j = 1, \cdots, q.$$

The proposed model has the next phases:

I. Estimate the parameters of the EGARCH (p, q) with the Maximum Likelihood method.

II. Use the results of the previous step how the center of the membership functions (3.8), (3.9), (3.10), (3.16) and (3.17).

III. Define the distance in the model d_n of each membership function and find the standard deviation as:

$$\delta_\varphi = \frac{d_n + \varphi}{3}$$

Where n is the lag associated with parameter φ and δ_φ is the standard deviation of the membership function.

IV. Obtain the probability of the membership function with δ_φ, d_n and φ.

V. Determine with the Gaussian function the parameter φ_k using the values φ, δ_φ and the probability of the last step.

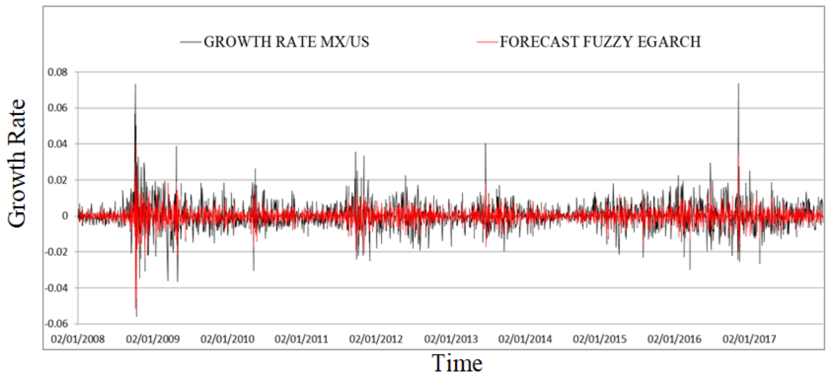
VI. Result of the linear programming problem (3.19).

VII. Finally calculate the fuzzy EGARCH (p, q) forecast.

3.3.1. Forecast Gaussian Parameters Fuzzy EGARCH $(1, 1)$.

The application of the Gaussian parameter Fuzzy EGARCH $(1, 1)$ is developed in the present section. First is estimated the EGARCH $(1, 1)$ model through the Maximum Likelihood method and with it are obtained the center values of each Gaussian membership function. After is defined every possibility distribution associated with the parameters EGARCH $(1, 1)$. And determined the Gaussian function the parameter φ_k using the values φ, δ_φ is the possible result in the linear programming problem (3.19) and obtain the Fuzzy EGARCH forecast (figure 24).

Figure 24 Forecast of Gaussian Parameters Fuzzy EGARCH (1, 1)



Source: Own elaboration in Excel.

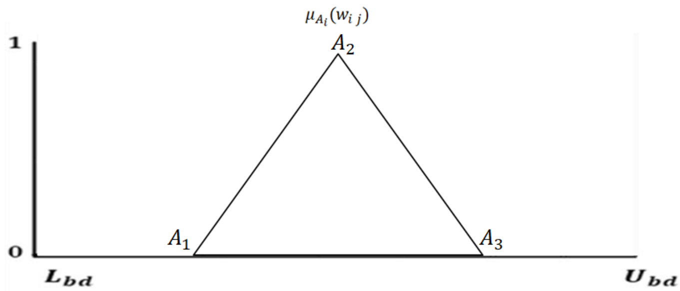
The hybrid method recognized the magnitude and the sign of the positive and negative information that cause volatility in the financial time series. Figure 24 shows the forecast Fuzzy EGARCH $(1, 1)$ (red line) and the growth rate MX/US (black line). It lets us see that the estimated values are similar to the real values, in other words, the proposed methodology presents a good prediction of the volatility in the foreign exchange market. This model is better than the traditional models ARIMA, EGARCH, PARCH and Nonlinear Autoregressive Neural Network.

The Asymmetric Conditional Variance model combined with fuzzy theory shows that the fuzzy method is better than the original technique. This is because the fuzzy theory lets recognize the uncertainty present in the financial time series, and also identify the moments where higher volatility exist.

3.4. Hybrid Fuzzy Nonlinear Autoregressive Neural Network: Fuzzy Triangular NARNET and Fuzzy Trapezoidal NARNET

The Fuzzy Autoregressive Neural Network is a fuzzy first-order model of the Sugeno-type, its If-Then rules in the input layer is determined as a Triangular Membership function.

Figure 25. Triangular membership function



Source: Own elaboration in MatLab

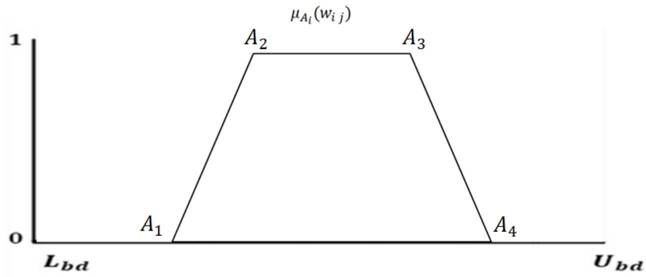
$$\begin{array}{l} R_1\colon if\ y_{t-i}\ is\ A_1\ then\ f_1 = w_{11}y_{t-1} + w_{12}y_{t-2} + \cdots + w_{n1}y_{t-n} \\ R_2\colon if\ y_{t-i}\ is\ A_2\ then\ f_2 = w_{21}y_{t-1} + w_{22}y_{t-2} + \cdots + w_{n2}y_{t-n} \\ R_3\colon if\ y_{t-i}\ is\ A_3\ then\ f_3 = w_{31}y_{t-1} + w_{32}y_{t-2} + \cdots + w_{n3}y_{t-n} \end{array} \tag{3.21}$$

Where R If-Then rules are the If-Then rules of the fuzzy time series, A is the triangular fuzzy subset of each function f that is the fuzzy Perceptron with their respective w weights. And the If-Then rules in the hidden layer are:

$$\begin{array}{l} R_1\colon if\ y_{t-i}\ is\ A_1\ then\ f_l = w_{k1}f_1 + w_{k2}f_1 \\ R_2\colon if\ y_{t-i}\ is\ A_2\ then\ f_l = w_{m1}f_2 + w_{m2}f_2 \\ R_3\colon if\ y_{t-i}\ is\ A_3\ then\ f_l = w_{m1}f_3 + w_{m2}f_3 \end{array} \tag{3.22}$$

In this case, A is the triangular fuzzy subset of each function f that is the fuzzy Perceptron with their respective w weights in the hidden layer of the neural network determinate as a Trapezoidal Membership function.

Figure 26. Trapezoidal Membership function



Source: Own elaboration in MatLab

$$\begin{array}{l} R_1: if\ y_{t-i}\ is\ A_1\ then\ f_1 = w_{11}y_{t-1} + w_{12}y_{t-2} + \cdots + w_{1n}y_{t-n} \\ R_2: if\ y_{t-i}\ is\ A_2\ then\ f_2 = w_{21}y_{t-1} + w_{22}y_{t-2} + \cdots + w_{2n}y_{t-n} \\ R_3: if\ y_{t-i}\ is\ A_3\ then\ f_3 = w_{31}y_{t-1} + w_{32}y_{t-2} + \cdots + w_{3n}y_{t-n} \\ R_4: if\ y_{t-i}\ is\ A_4\ then\ f_4 = w_{41}y_{t-1} + w_{42}y_{t-2} + \cdots + w_{4n}y_{t-n} \end{array} \quad \textbf{(3.23)}$$

Where R are the "If-Then" rules of the fuzzy time series A is the trapezoidal fuzzy subset of each function f that is the fuzzy Perceptron with their respective w weights. And the If-Then rules in the hidden layer are:

$$\begin{array}{l} R_1: if\ y_{t-i}\ is\ A_1\ then\ f_l = w_{k1}f_1 + w_{k2}f_1 \\ R_2: if\ y_{t-i}\ is\ A_2\ then\ f_l = w_{m1}f_2 + w_{m2}f_2 \\ R_3: if\ y_{t-i}\ is\ A_3\ then\ f_l = w_{m1}f_3 + w_{m2}f_3 \\ R_4: if\ y_{t-i}\ is\ A_4\ then\ f_l = w_{m1}f_4 + w_{m2}f_4 \end{array} \quad \textbf{(3.24)}$$

And A is the trapezoidal fuzzy subset of each function f that is the fuzzy Perceptron with their respective w weights in the hidden layer of the neural network. In the first phase of the network, each node is concentrated with the respective membership function of the fuzzy subset. This means that y_t is the variable that feeds each node i and A_i is the fuzzy subset denoted as linguistic value.

$$y_t = A_i$$

Where A_i is the fuzzy set of y_t and express the fuzzy time series. In the specific case, memberships function of triangular or trapezoidal type. And after for improve the learned function, the information is softened as follows, García *et al* (2013):

$$z_{t-1} = \frac{y_{t-1} - \mu(y_{t-1} + y_{t-2} + \cdots + y_{t-n})}{\sigma(y_{t-1} + y_{t-2} + \cdots + y_{t-n})} \quad \textbf{(3.25)}$$

In the second phase, in the neural network, each of the nodes is represented for a function determinate by the rules If-Then (12) or (14). This means that the node is the multiplication of the different signals and their sum is the output.

$$f = w_1 z_{t-1} + w_2 z_{t-2} + \cdots + w_n z_{t-n} \quad \textbf{(3.26)}$$

In the third phase, the output of the previous step is transformed through a sigmoidal function. Then the hidden layer is determinate by the rules If-Then (13) or (15).

$$f_s(f) = \frac{1}{1 + e^{-f}} \quad \textbf{(3.27)}$$

The last equation is a sigmoidal activation function. So, the hidden step is denoted by the next equation:

$$f_l = w_1 f_s + w_2 f_s \quad \textbf{(3.28)}$$

Ultimately, in a single node, the sum of all the output signals of the previous phases is performed; the output variable z is represented as:

$$\widehat{y_t} = f_l * \sigma(y_{t-1} + y_{t-2} + \cdots + y_{t-n}) + \mu(y_{t-1} + y_{t-2} + \cdots + y_{t-n}) \quad \textbf{(3.29)}$$

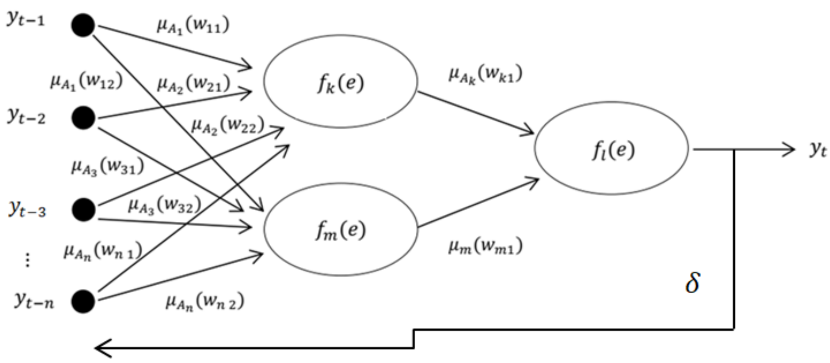
Where $\widehat{y_t}$ is the forecast of y_t, σ is the variance of the past values and μ represents the mean of the past realizations. The error function is defined as the "Absolute Mean Deviation", such as:

$$\epsilon = \frac{E\left|y_t - \hat{y}_t\right|}{n} \quad \textbf{(3.30)}$$

The learning algorithm of the adaptive network-based on fuzzy inference, it is given by the rules "If-then" and the observation of the parameters of the models what can be expressed by a linear combination of general results so that it can be expressed as follows:

$$f = \overline{w_1}(p_1 y_t + r_1) + \overline{w_2}(p_2 y_t + r_2)$$

Figure 27. Hybrid Fuzzy Nonlinear Autoregressive Neural Network



Source: Own elaboration.

3.4.1. Forecast the Hybrid Fuzzy Autoregressive Neural Network (Fuzzy NARNET).

The hybrid fuzzy autoregressive Neural Network method is applied with 2514 observations to formulate the model and the next 26 observations to evaluate the performance of the model.

First, it defined the membership function associated with system equations (12) and (13). This is denoted by the table 13 which shows the lower bound and the upper bound of each triangular membership function of the If-Then rules. For the fuzzy subset $\boldsymbol{A_1}$ the $\boldsymbol{L_{BD}}$ is -0.06 and the $\boldsymbol{U_{BD}}$ is -0.0085, the fuzzy subset $\boldsymbol{A_2}$ the $\boldsymbol{L_{BD}}$ is -0.0085 and the $\boldsymbol{U_{BD}}$ is 0.0085 and finally the fuzzy subset $\boldsymbol{A_3}$ is from 0.0085 to 0.08.

Table 13. Triangular fuzzy subsets

	$\boldsymbol{L_{BD}}$	$\boldsymbol{U_{BD}}$
A1	-0.06	-0.00857
A2	-0.00857	0.00857
A3	0.00857	0.08

Source: Own Elaboration in Excel.

Second, it expressed the membership function associated with system equations (14) and (15). The lower bound and the upper bound of each trapezoidal membership function of the If-Then rules are specified by table 14 which shows that the fuzzy subset $\boldsymbol{A_1}$ the $\boldsymbol{L_{BD}}$ is -0.06 to the $\boldsymbol{U_{BD}}$ is -0.014, the fuzzy subset $\boldsymbol{A_2}$ the $\boldsymbol{L_{BD}}$ is -0.014 and the $\boldsymbol{U_{BD}}$ is 0, the fuzzy subset $\boldsymbol{A_3}$ is from 0 to 0.017 and finally $\boldsymbol{A_4}$ the $\boldsymbol{L_{BD}}$ is 0.017 to the $\boldsymbol{U_{BD}}$ is 0.08.

Table 14. Trapezoidal fuzzy subsets

	$\boldsymbol{L_{BD}}$	$\boldsymbol{U_{BD}}$
A1	-0.06	-0.01
A2	-0.01429	0.00
A3	0.000	0.01714
A4	0.01714	0.08

Source: Own elaboration in Excel.

In this way are denoted the membership function associated with the hybrid fuzzy NARNET. In this case, it is important said that is estimated two neural networks, the first is in which are assumed that the sign fuzzy set is triangular and in the second one is a trapezoidal fuzzy set. These models are applied to estimate the behavior of the exchange rate MX/US.

In the first phase of the network, each node is concentrated with the respective membership function of the fuzzy subset. This means that $\boldsymbol{y_t}$ is the variable that feeds each node $\boldsymbol{i}$ and $\boldsymbol{A_i}$ is the fuzzy subset denoted as linguistic value. In the triangular situation is how follows:

$$y_t = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix}$$

And in the case of a trapezoidal membership function.

$$y_t = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix}$$

Where A_i is the fuzzy set of y_t and express the fuzzy time series. In specific, these models correspond to third-order NARNET and their information is softened as follows, García *et al* (2013):

$$z_{t-1} = \frac{y_{t-1} - \mu(y_{t-1} + y_{t-2} + y_{t-3})}{\sigma(y_{t-1} + y_{t-2} + y_{t-3})}$$

In the second phase, in the neural network, each of the nodes is represented for a function determinate by the rules If-Then (3.21) and (3.23). This means that the node is the multiplication of the different signals and their sum is the output.

$$f = w_1 z_{t-1} + w_2 z_{t-2} + w_3 z_{t-3}$$

In the third phase, the output of the previous step is transformed through a sigmoidal function. Then the hidden layer is determinate by the rules If-Then (3.22) or (3.24).

$$f_s(f) = \frac{1}{1 + e^{-f}}$$

The last equation is a sigmoidal activation function. So, the hidden step is denoted by the next equation:

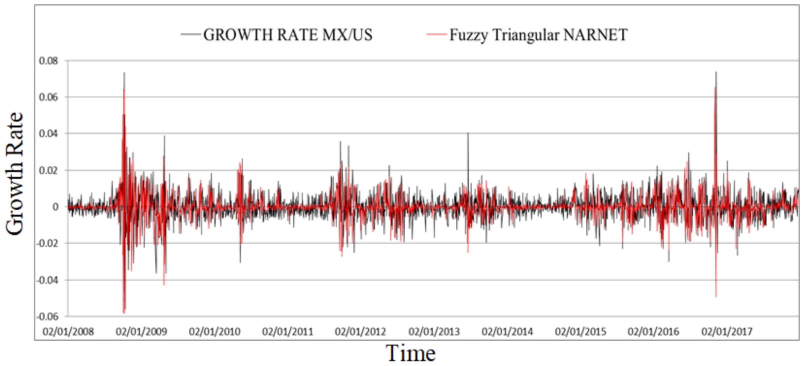
$$f_l = w_1 f_s + w_2 f_s$$

Ultimately, in a single node, the sum of all the output signals of the previous phases is performed; the output variable z is represented as:

$$\widehat{y_t} = f_l * \sigma(y_{t-1} + y_{t-2} + y_{t-3}) + \mu(y_{t-1} + y_{t-2} + y_{t-3})$$

Where $\widehat{y_t}$ is the forecast of y_t, σ is the variance of the past values and μ represents the mean of the past realizations in a third-order NARNET model. And the object is to minimize the equation (3.30) to understanding the behavior of the exchange rate through fuzzy signs.

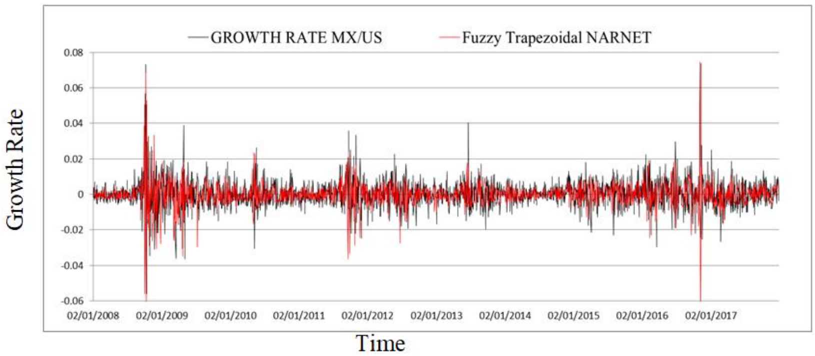
Figure 28. Forecast of fuzzy triangular NARNET model for the growth rate of the exchange rate of Mexican Peso versus American Dollar



Source: Own elaboration in Excel.

The results are showed in figure 28 for the case of the hybrid fuzzy triangular NARNET that the forecast of this model is represented by the red line and the black one is the growth rate of the exchange rate MX/US. In it is possible to look at a good estimate, because the red line is similar to the black and there are a few points where the model not can identify the uncertainty associated with the time series. This model takes in count the extreme values of the data on how a fuzzy subsets and only a center fuzzy subset.

Figure 29 Forecast of fuzzy trapezoidal NARNET model for the growth rate of the exchange rate of Mexican Peso versus American Dollar



Source: Own elaboration in Excel.

Now, figure 28 presents the estimate of the hybrid fuzzy trapezoidal NARNET with the red line and the black is the exchange rate. Both lines are similar or in other words, the model achieves forecast of a good form the movements that have the growth rate of the exchange rate MX/US. In specific, the model with trapezoidal membership function model fuzzy extreme values and two centers fuzzy subsets, this let denoted that there is a different Perceptron according to the measurement of the information uncertainty.

3.5. Fuzzy ARIMA Models.

The Fuzzy ARIMA model was developed by Tseng *et al.* (2001). This method based on a fuzzy linear regression model with a triangular membership function. Such that, the fuzzy ARIMA model can be represented in the following way:

$$Y_t = \langle \alpha_1, c_1 \rangle Y_{t-i} + \cdots + \langle \alpha_p, c_p \rangle Y_{t-p} + \varepsilon_t - \langle \alpha_{p+1}, c_{p+1} \rangle \varepsilon_{t-1} - \cdots - \langle \alpha_{p+d}, c_{p+d} \rangle \varepsilon_{t-d}$$

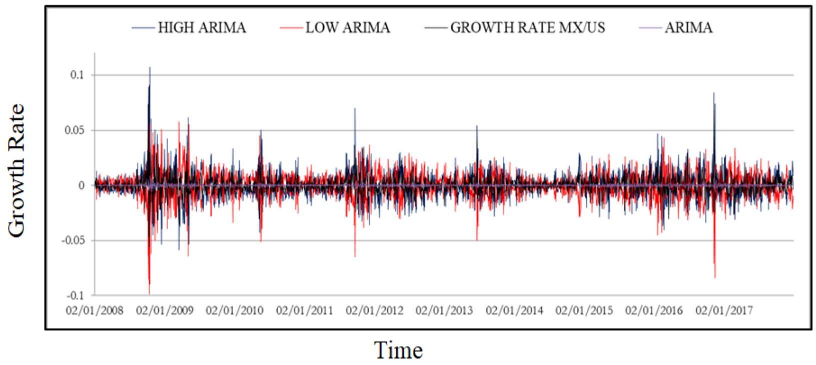
Where $\langle \alpha_1, c_1 \rangle$ represents the values of each coefficient. Taking as center α and c is the extension of the model understood as the upper and lower limit of the fuzzy set associated with the coefficient. Therefore, under this methodology, it is possible to carry out a forecast of the time series, obtaining, as a result, a range of possibilities of the future values of the variable.

This section developed a study of the different fuzzy ARIMA models and their application to forecasting the foreign exchange market. First is estimated the Tanaka's methodology to the growth rate MX/US, next the forecast the Tseng's methodology applied to the growth rate MX/US, after the Gaussian parameters FUZZY AR is obtained to forecast the growth rate MX/US, and finally the T-student parameters FUZZY AR is estimated and applied to the growth rate MX/US.

3.5.1. Tanaka Methodology.

The present model is understood how a method of crispy input, fuzzy parameters and fuzzy output with triangular membership functions. For that, there are tree forecast values the high ARIMA, low ARIMA, and the center ARIMA. The original paper Tanaka *et al.* (1982) says that the only values that must be taken are major or lower than the real values but for this research are taken all the estimated values.

Figure 30 Forecast Fuzzy ARIMA of Tanaka Methodology for the Growth Rate MX/US



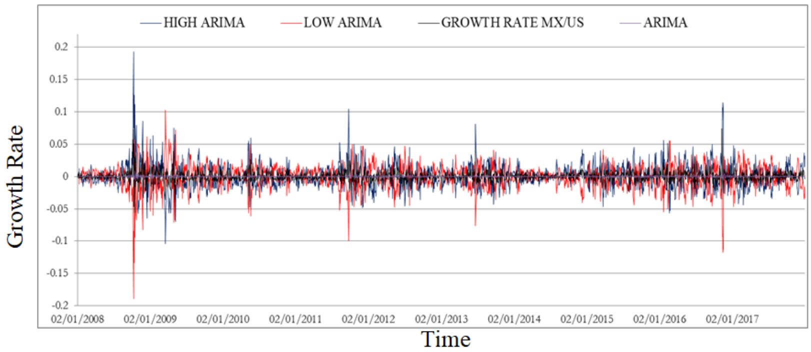
Source: Own elaboration in Excel.

Figure 30 shows the Fuzzy ARIMA of Tanaka Methodology, the blue line is the high ARIMA, the low ARIMA is the red line, the center ARIMA is the purple line, and finally, the black line is the growth rate MX/US. This model did not let see a possibility to carry out a forecast of the time series, this obtains a range of possibility of the future values of the variable. The Tanaka Methodology recognized the volatility of the exchange rate MX/US and also lets the decision-makers know the best and the worst possible values of the time series.

3.5.2. Tseng Methodology.

The present method is assumed how a crispy input, fuzzy parameters and fuzzy output with triangular membership functions. So that, exist tree forecast values the high ARIMA, low ARIMA, and the center ARIMA. The original research Tseng *et al.* (2001) say that the only values that must be taken those that are major or lower than the real values but for this research are taken all the estimated values.

Figure 31. Forecast Fuzzy ARIMA of Tseng Methodology for the Growth Rate MX/US



Source: Own elaboration in Excel.

Figure 31 shows the Fuzzy ARIMA of Tseng Methodology, the blue line is the high ARIMA, the low ARIMA is the red line, the center ARIMA is the purple line, and finally, the black line is the growth rate MX/US. This method is not let see a possibility to carry out the forecast of the time series; this obtains a range of possibility of the future values of the variable. The Tseng Methodology recognized the volatility of the exchange rate MX/US and also lets the decision-makers know the best and the worst possible values of the time series.

The difference between the Tanaka and Tseng methodology is how they have taken the h-level in the triangular membership function and how obtained the fuzzy outputs. Both models let see similar results but it is not possible to recognize what is the best model. For this situation, these models are combined with the Fuzzy Time Series method in section 3.1. The results are shown in the comparison section of this chapter.

3.5.3. Gaussian parameter FUZZY AR.

The Gaussian parameter FUZZY AR is developed of similar form of the Fuzzy GARCH but the difference is that the present model not estimated the conditional variance if not the AR process of the stochastic time series. It can see as follows:

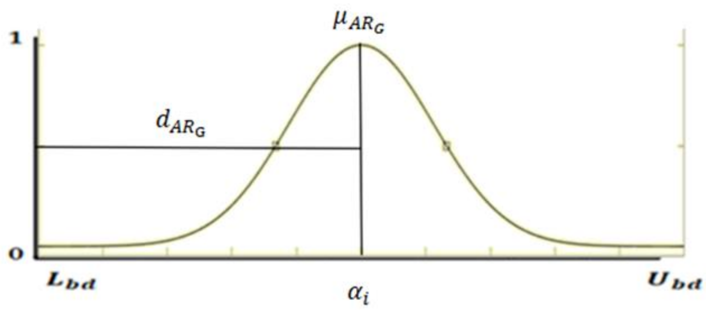
$$Y_t = \alpha_1 Y_{t-i} + \cdots + \alpha_p Y_{t-p} + \varepsilon_t$$

Where the letter Y_t is the time series analyzed and $\alpha_1 Y_{t-i} + \cdots + \alpha_p Y_{t-p}$ represents the autoregressive part of the variable and finally ε_t is a white noise of the process. Up to now, the last equation is a traditional AR model and this is combined with fuzzy theory assuming the follows:

Assumption 3.7. The membership function of the AR (p) parameters $(\alpha_1, \cdots, \alpha_p)$ is a Gaussian type. In this way, it can be expressed as:

$$\mu_{AR_G}(\alpha_{ik}) = e^{-\left(\frac{\alpha_{ik}-\alpha_i}{\delta_{\alpha_i}}\right)^2} \quad \textbf{(3.31)}$$

Figure 32. Gaussian Membership Function of α_i



Source: Own elaboration in MatLab.

$$d_{AR_G} = 3 * \delta_{\alpha_i} - \alpha_i$$

Assumption 3.8. If the time series Y_t of the AR (p) process is a fuzzy function, then it can express as:

$$Y_t = \mu_{AR_G}(\alpha_{1k})Y_{t-i} + \cdots + \mu_{AR_G}(\alpha_{pk})Y_{t-p} + \varepsilon_t \quad \textbf{(3.32)}$$

Where (3.32) represents the fuzzy input of the AR (p) process. In this case is necessary to define an error minimization function, for the present research is taken the Mean Absolute Error (MAD) and can be expressed as follows:

$$\epsilon = \frac{E[Y_t - \hat{Y}_t]}{n}$$

Then the problem is finding the fuzzy AR (p) parameter that can obtain resolving the next linear programming problem:

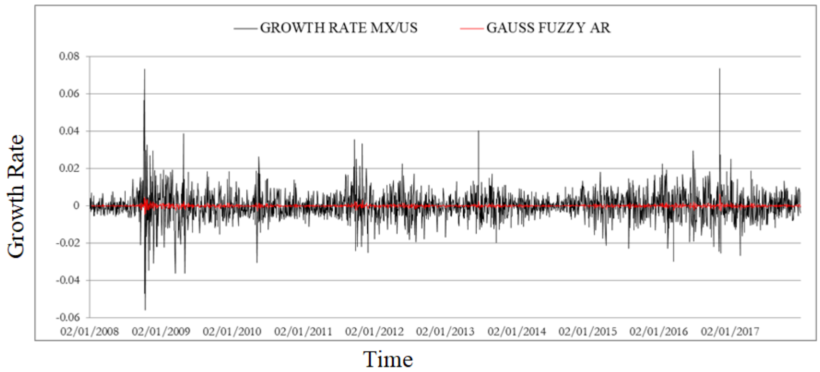
$$\begin{aligned} &\text{Minimize} \quad \epsilon \\ &\text{Subject to} \quad d_{AR_G} > 0 \end{aligned} \tag{3.33}$$

When finding the parameters that guarantee the minimum ϵ also is found the non-fuzzy forecast of the conditional variance. And the Fuzzy AR (p) model can be expressed as:

$$Y_t = \alpha_{1k} Y_{t-i} + \cdots + \alpha_{pk} Y_{t-p} + \varepsilon_t \tag{3.34}$$

Figure 13 shows the forecast Gaussian Fuzzy AR (6) applied to the growth rate of MX/US. The red line represents the forecast values and the black line the exchange rate, the red line has a bad accurate of the growth rate MX/US. For this model not is found a good result in the graphical comparison but until the moment is necessary to check the MAD of the model.

Figure 33 Forecast Gaussian Fuzzy AR for the Growth Rate MX/US

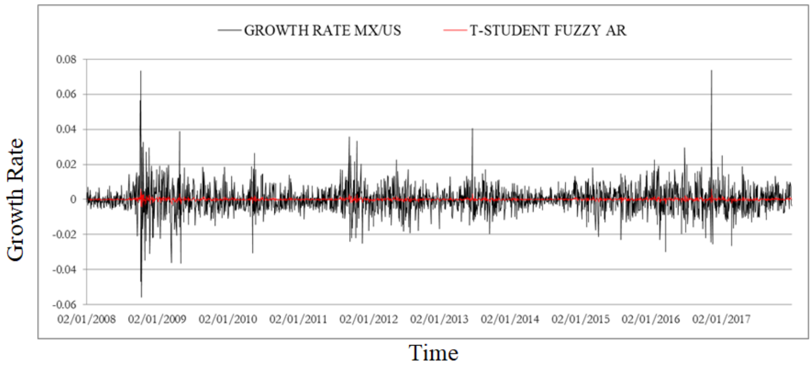


Source: Own elaboration in Excel.

3.5.4. T-student parameters FUZZY AR.

The present method is assumed how a crispy input, fuzzy parameters and crisp output with T-student membership functions. Figure 34 shows the forecast T-student Fuzzy AR(6) applied to the growth rate of MX/US. The red line represents the estimated values and the black one is the exchange rate, the red line does not have good accuracy of the growth rate MX/US. This model found that the graphical result is bad.

Figure 34 Forecast T-student Fuzzy AR for the Growth Rate MX/US



Source: Own elaboration in Excel.

The Fuzzy AR (6) with Gaussian or T-student parameters, not generate a better forecast of the growth rate MX/US in comparison with the traditional models (ARIMA, EGARCH, PARCH and Nonlinear Autoregressive Neural Network).

3.6. Benchmark Differential Neural Network.

The present section shows the forecast of the Differential Neural Network (DNN) for the Growth Rate MX/US. This forecast is developed in the MatLab program and the different simulations were made in Simulink Toolbox. The structure of the model is taken of Ortiz-Arango *et al.* (2016) and it is denoted as follows:

$$\dot{\hat{x}}_t = A\hat{x}_t + W_{1,t}\sigma(\hat{x}_t) + W_{2,t}\varphi(\hat{x}_t)\gamma(\hat{u}_t)$$

Cabrera-Llanos *et al*. (2012), it knows that this model studies system of the non-linear type. Such that, the identifiers of the neural network where the weights are adjusted by the laws of learning that follow:

$$\frac{d}{dt}W_{1,t} = -K_1 P\Delta_t\sigma(\hat{x}_t)$$
$$\frac{d}{dt}W_{2,t} = -K_2 P\Delta_t\gamma(\hat{u}_t)\varphi(\hat{x}_t)$$
$$\Delta_t = x_t - \hat{x}_t$$

Indeed, the differential neural network in the unidimensional case is represented. And the activation functions are:

$$\sigma(\hat{x}_t) = \left(\left(\frac{1}{a_1 - e^{-a_2 x_t}}\right) - a_3\right)$$

$$\varphi(\hat{x}_t) = \left(\left(\frac{1}{b_1 - e^{-b_2 x_t}}\right) - b_3\right)$$

Table 15. Values of the Differential Neural Network

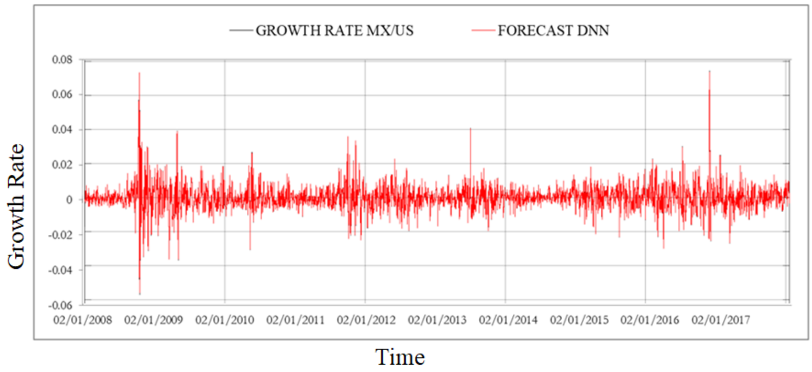
Element	A	a_1	a_2	a_3	b_2	b_2	b_3	K_1	P	K_2
Value	10	1	2	5	1	2	5	4	100, 200	3

Source: Own elaboration in MatLab.

The Differential Neural Network is estimated with 2514 states. First were developed the activation functions with different values and its results are compared for choosing the best estimate. The conclusion was that the best values are shown in the above equations. Where the Hurwitz value is -10 and the solution of Ricatti's matrix is 100 and 300.

Figure 35 shows the forecast values of the unidimensional Differential Neural Network (red line) in comparison with the Growth Rate MX/US (black line). The main result is that the DNN model is accurate in the forecast values of the Exchange Rate Mexican Peso against American Dollar. For instance, the advanced model has the empirical conditions for model the behavior of the intrinsic volatility of the financial time series.

Figure 35 Forecast Differential Neural Network for the Growth Rate MX/US



Source: Own elaboration in MatLab.

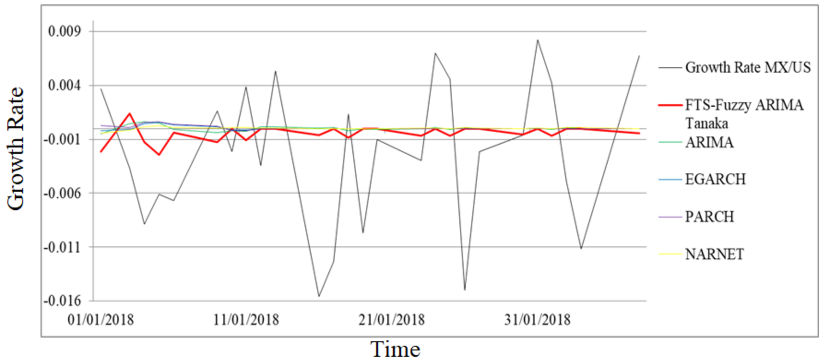
3.7. Comparison of the fuzzy models and traditional models.

In the present section are compared the FTS-Fuzzy ARIMA Tanaka's, FTS-Fuzzy ARIMA Tseng's, Fuzzy GARCH, Fuzzy EGARCH, Fuzzy Triangular Nonlinear Autoregressive Neural Network, Fuzzy Trapezoidal Nonlinear Autoregressive Neural Network, Fuzzy Gaussian AR and Fuzzy T-Student models and also the traditional models (ARIMA, EGARCH, PARCH and Nonlinear Autoregressive Neural Network). This comparison is developed with the Mean Absolute Deviation, first in the sample and after out sample.

The forecast of the exchange rate in the period it comprises; January 2, 2008, to December 29, 2017, in daily format (excluding non-working days, with a total of 2514 observations). For the out sample test 26 observations are added, January 2, 2018, to February

7, 2018. First is developed the graphical comparison of each proposed model with the traditional models out of the sample.

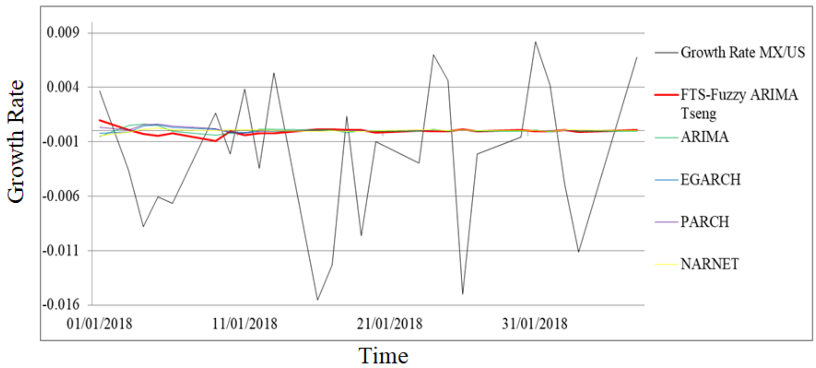
Figure 36. Benchmark Out sample of FTS-Fuzzy ARIMA Tanaka with the traditional models



Source: Own elaboration in Excel.

Figure 36 represents the FTS-Fuzzy ARIMA Tanaka (red line) model in comparison with the traditional models of chapter II. The black line is the Growth Rate MX/US and this denotes a great variability, the object is tried to understand this behavior with the traditional and fuzzy models.

Figure 37. Benchmark Out sample of FTS-Fuzzy ARIMA Tseng with the traditional models

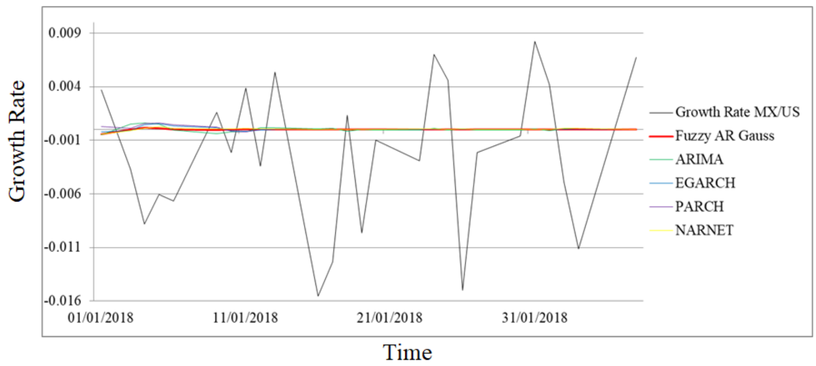


Source: Own elaboration in Excel.

The EGARCH model is the blue line but its forecast does not have great variability, the PARCH model is the purple line its estimated values out the sample the variability is not good for predict the time series analyzed. The ARIMA model is the green line it has a little variability but not enough for models the Growth Rate MX/US. And the NARNET is represented for the yellow line this forecast does not have a great variability how the exchange rate. In summary, the proposed model estimates better the volatility out sample of the Growth Rate MX/US than the traditional models.

Figure 37 shows the FTS-Fuzzy ARIMA Tseng model the red line in comparison with the traditional models and the black line is the Growth Rate MX/US. The EGARC H model is the blue line; the PARCH model is the purple line; the ARIMA model is the green line and the NARNET is the yellow line. The results denoted that the proposed model estimates not show great improvement out the sample of the estimate the Growth Rate MX/US than the traditional models.

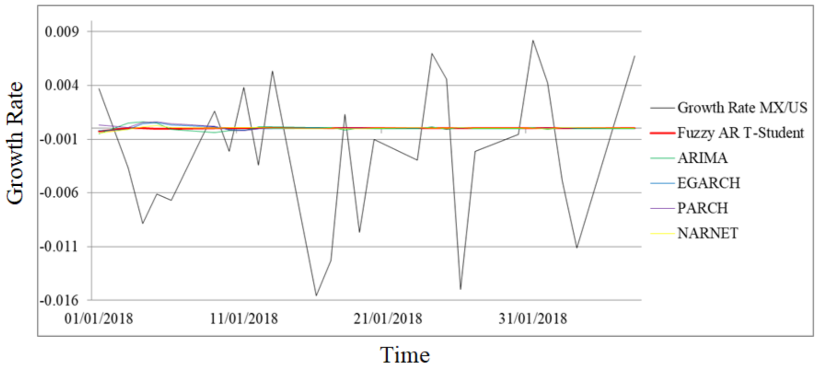
Figure 38. Benchmark Out-Sample of Fuzzy AR Gauss with the traditional models



Source: Own elaboration in Excel.

Figure 38 denoted the evaluation of the Fuzzy AR Gauss with the traditional models. In this case, the proposed model has a worse forecast that the traditional models. Figure 33 confirms this result because the Fuzzy AR Gauss does not generate a good forecast in the sample.

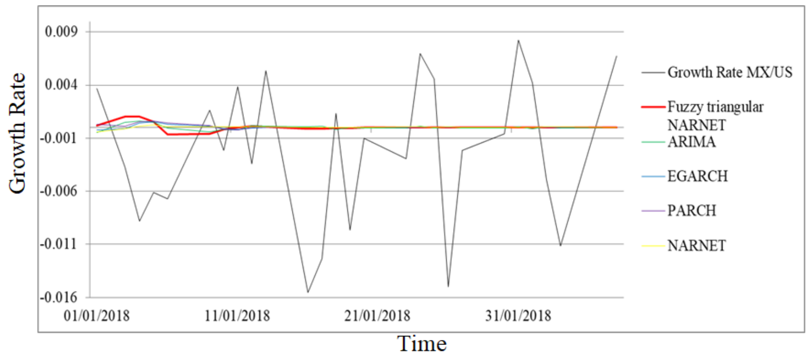
Figure 39. Benchmark Out sample of Fuzzy AR T-Student with the traditional models



Source: Own elaboration in Excel.

The contrast of the traditional models and the Fuzzy AR T-Student is represented in figure 39. The results show that the fuzzy model has a worse forecast that the traditional models.

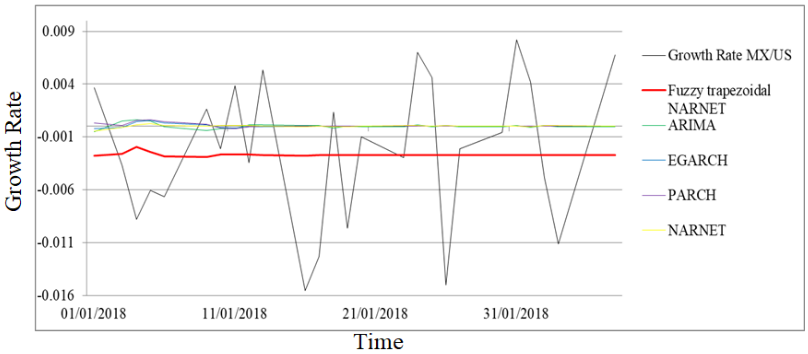
Figure 40. Benchmark Out sample of Fuzzy AR T-Student with the traditional models



Source: Own elaboration in Excel.

Figure 40 shows the comparison of the Fuzzy Triangular NARNET and the traditional models. The proposed model denoted more variability that the traditional models and this propagation last up five days. The forecast out the sample of the fuzzy model is better than the traditional models and the FTS-Fuzzy ARIMA Tseng, Fuzzy AR Gauss, and the Fuzzy AR T-Student.

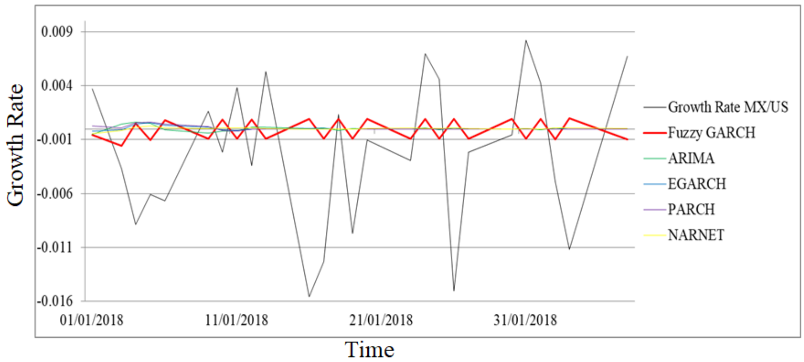
Figure 41. Benchmark Out sample of Fuzzy Trapezoidal NARNET with the traditional models



Source: Own elaboration in Excel.

The Fuzzy Trapezoidal NARNET shows a similar forecast that the traditional models but the difference is that the forecast out sample of the proposed model is displaced to down, it is occasioned by the fuzzy trapezoidal membership function because this can measurement the different phases of the volatility and in this case, the forecast is developed by the parameters of the fuzzy subset A_1. Figure 41 represents the forecast of the suggested model the red line and the traditional models, it denoted that the Fuzzy NARNET does not have great variability.

Figure 42 Benchmark Out sample of Fuzzy GARCH with the traditional models

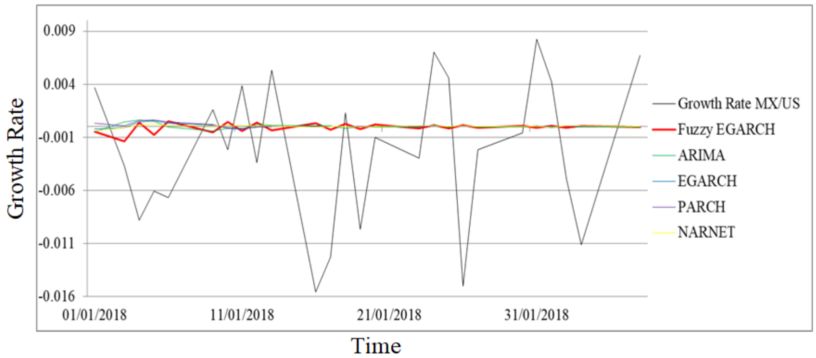


Source: Own elaboration in Excel.

Figure 42 shows the forecast out sample of the Fuzzy GARCH models and this is compared with the traditional models. The red line is the model proposed its values have significances variability in comparison with the benchmark models, however, although this model has a better forecast than the traditional methods, it not is enough for understanding the behavior of the growth rate MX/US black line.

Figure 43 describes the comparison of the Fuzzy EGARCH model red line. The black line is the Growth Rate MX/US and this denotes a great variability, the object is tried to understand this behavior with the traditional and fuzzy models. The EGARCH model is the blue line but its forecast does not have great variability, the PARCH model is the purple line its estimated values out the sample the variability is not good for predict the time series analyzed. The ARIMA model is the green line it has a little variability but not enough for models the Growth Rate MX/US. And the NARNET is represented for the yellow line this forecast does not have a great variability how the exchange rate. In summary, the proposed model estimates better the volatility out sample of the Growth Rate MX/US than the traditional models.

Figure 43. Benchmark Out sample of Fuzzy EGARCH with the traditional models

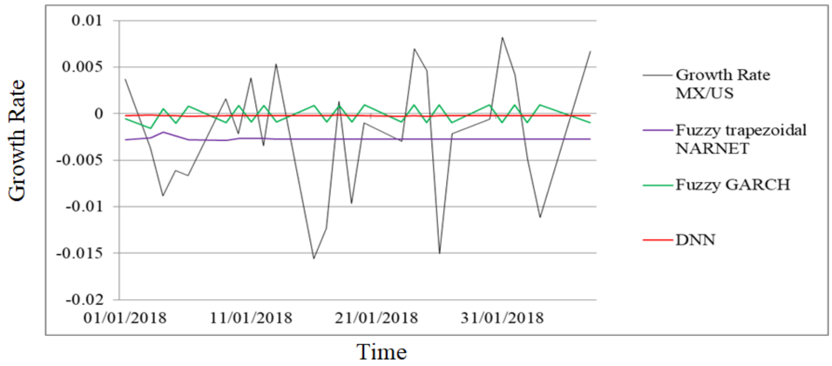


Source: Own elaboration in Excel.

Finally, figure 44 illustrates the comparison of the advanced model DNN with the Fuzzy GARCH and Fuzzy Trapezoidal NARNET out the sample. The main result is that the model DNN generates a forecast with lower variability than the fuzzy proposed model. Though, this situation not said anything about the efficiency of the prediction.

In this way, the importance of the graphical comparison of the fuzzy models with the traditional and advanced model was looked at the variability and the loss of predictive ability when the forecast is generated the periods out the sample. And the common result in each figure is that the fuzzy model has major variability than the benchmark models.

Figure 44. Benchmark Out-Sample of Fuzzy GARCH and Fuzzy Trapezoidal NARNET with DNN



Source: Own elaboration in Excel.

3.7.1. Evaluation of the Models: In-Sample and Out-Sample Test.

The next part is developing the evaluation of the forecast in the sample and out sample test for the model of January 2, 2018, to February 7, 2018; 26 days of the sample or five weeks plus one day. First, Figure 45 illustrates the contrast of the Prediction Percentage Error Test for the Fuzzy, the traditional and the advanced models. The subsection (a) and (b) represents the out sample test of the FTS-Fuzzy ARIMA Tseng's Model in the first day the error was of 0.7%, in the week one the mean error daily was of 0.95%, for the next week was of 1.2%, the third week has a mean error of 0.9%, for the four is 0.99% and the fifth is 1%. In the case of FTS-Fuzzy ARIMA Tseng's Model, the first day was of 1.6%, in the week one the mean error daily was of 0.9%, the second week was of 1.2%, the third week has a mean error of 1%, for the four is 1% and the fifth is 0.7%. These types of models not loosed the variability in the last samples. In (c) and (d) the Gauss Fuzzy AR had a mean error in the day one of 1.12%, in the first week 1.01%, for the next week was of 1%, the third week has a mean error of 1%, for the four is 1% and the fifth is 1%. And the T-Student Fuzzy AR the first day was of 1.07%, in the week one the mean error daily was of 0.99%, the second week was of 1%, the third week has a mean error of 1%, for the four is 1% and the fifth is 1%. These models loosed the variability in the second week.

In other cases the (e) and (f) the Fuzzy Triangular NARNET had a mean error in the first day of 0.98%, in the first week 1.1%, for the next week was of 1.2%, the third week has a mean error of 1%, for the four is 1% and the fifth is 1%. And the Fuzzy Trapezoidal

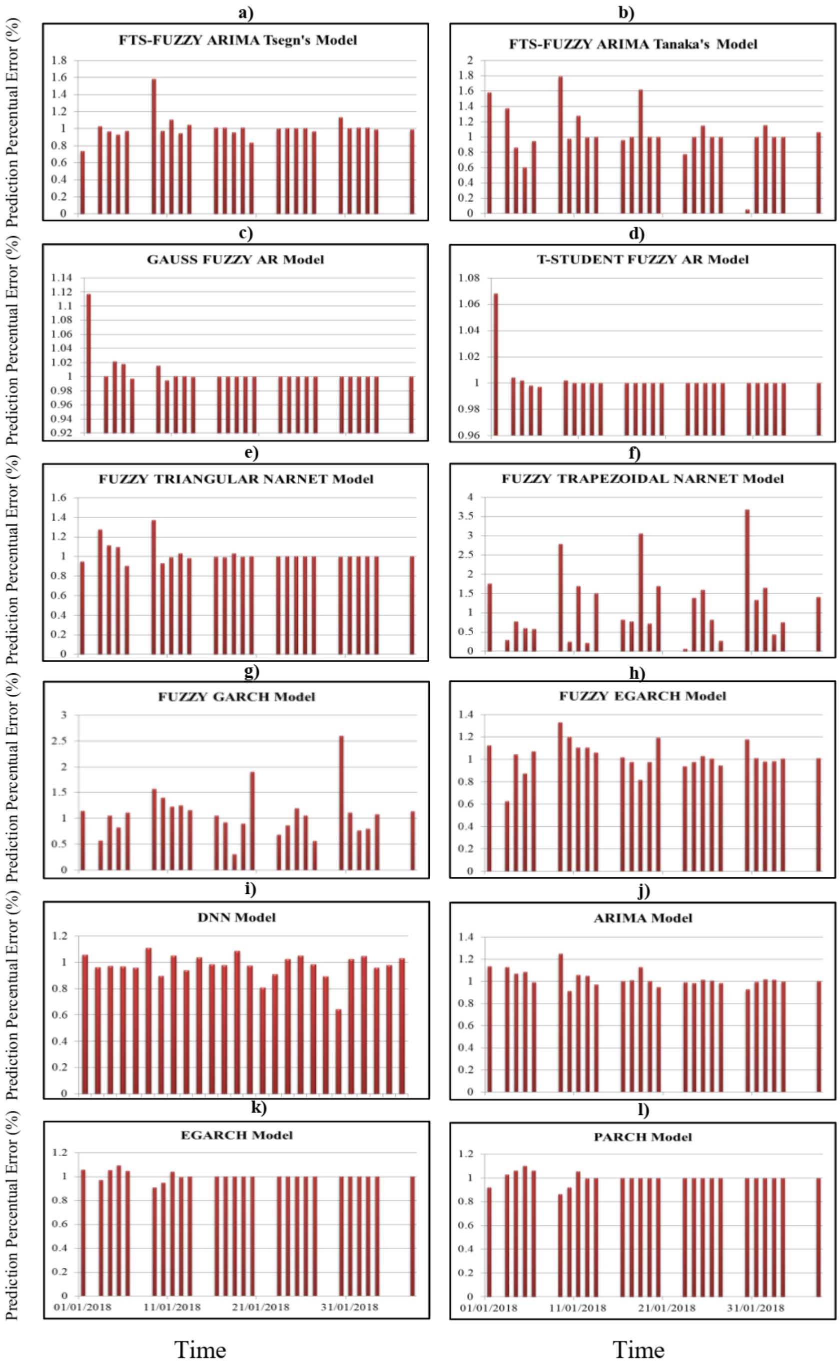
NARNET in the one day was a 1.6% of error, in the week one the mean error daily was of 0.7%, the second week was of 1.5%, the third week has a mean error of 1.8%, for the four is 0.9% and the fifth is 1.5%. These models have a great error but the variability is major than the traditional models.

Continuing, the (e) and (f) are the Fuzzy GARCH and the Fuzzy EGARCH respectively, these models have major variability in its forecast. But in some cases, the error is less than one for that and in others is greater than one. This situation reveals that the parameters of the fuzzy conditional variance models let generate forecasts of more days out of the sample.

The subsection (i) is the Differential Neural Network and this method is the most stable because the trajectory of the forecast does not lose efficiency when leaving the sample. And finally (j), (k) and (l) represent the traditional models these methods not generate a good forecast out the sample. The ARIMA has an error daily major of on, the PARCH loosed the variability in the second week and the EGARCH loosed the variability in the third week.

Figure 46 illustrates the out sample mean absolute deviation of each model in the first ten days of prediction, the values can be seen in table 3.6. The results show that the fuzzy models developed best out-sample test than the traditional models, it can be seen in the lower MAD of each day are of the fuzzy techniques. But the major MAD is of the fuzzy methods, these results are denoted in the last section in which the out-sample forecast of the suggested models generates major variability in their values.

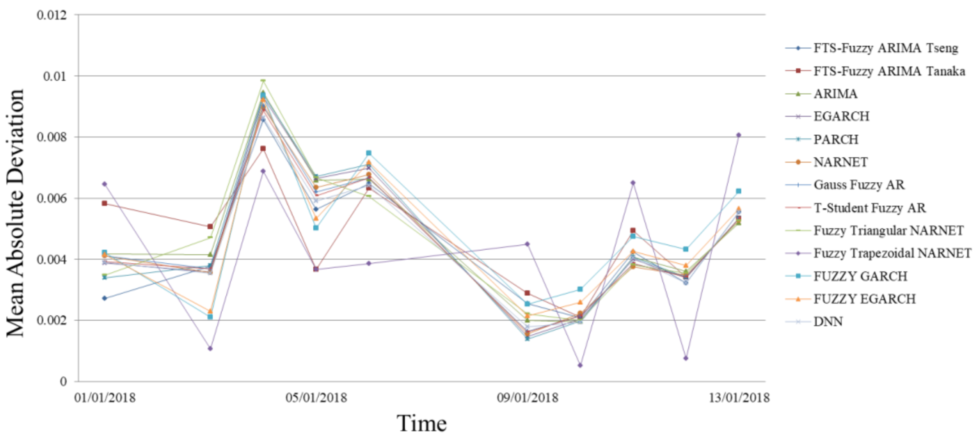
Figure 45. Comparisons of Prediction Percentage Error



Source: Own elaboration in Excel.

Table 16 is the comparison of the mean absolute deviation in the sample. The results are understood as the percentage of daily mean error at each model. Furthermore, the table is divided into traditional models and fuzzy models. Now, the traditional ARIMA model has an AME of 0.5299%; the FTS-Fussy ARIMA Tanaka 0.5157; the FTS-Fussy ARIMA Tseng 0.5199; Fuzzy AR Gauss 0.5298% and the Fuzzy AR T-Student 0.5295%. This result shows that the Fuzzy AR Gauss and T-Student do not generate an improvement in comparison with the traditional ARIMA, situation that already was identified in the lasts sections. But nevertheless, the FTS-Fuzzy ARIMA models are better than the ARIMA model both models have 0.0100% lower. These results said that the fuzzy models can improve the forecast or worsen the estimated values, for that is necessary to determine a good combination of the traditional and the fuzzy theory.

Figure 46. Out-Sample Mean Absolute Deviation Comparison



Source: Own elaboration in Excel.

The E-GARCH has an AME of 0.5274%; for the PARCH is 0.281%, these are the traditional variance conditional models and compared with the fuzzy models. The Fuzzy GARCH has an AME of 0.3928% and the Fuzzy EGARCH is 0.3965%. The test indicated that the hybrid Fuzzy with variance conditional models generate a better forecast that the traditional heteroskedastic methods. The difference is of 0.12% lower error of the fuzzy models in comparison with de basic methodologies. This combination resulted better that the Fuzzy theory with ARMA methods.

Table 16. Absolute Mean Deviation Comparison

In the Sample Test	
Model	Daily
ARIMA (traditional)	0.5299 %
E-GARCH (traditional)	0.5274 %
PACH (traditional)	0.5281 %
Nonlinear Autoregressive Neural Network (traditional)	0.5110 %
FTS-FUZZY ARIMA Tseng (fuzzy)	0.5199 %
FTS-FUZZY ARIMA Tanaka (fuzzy)	0.5157 %
FUZZY GARCH (fuzzy)	**0.3928 %**
FUZZY AR Gauss (fuzzy)	0.5298 %
FUZZY AR T-Student (fuzzy)	0.5295 %
FUZZY EGARCH (fuzzy)	0.3965 %
FUZZY Trapezoidal NARNET (fuzzy)	**0.3473 %**
FUZZY Triangular NARNET (fuzzy)	0.3954 %
Differential Neural Network	**0.0048%**

Source: Own elaboration in Excel.

And the Nonlinear Autoregressive Neural Network has an AME of 0.5110%; the Fuzzy Triangular NARNET has a 0.3954% and the Fuzzy Trapezoidal NANET has an AME of 0.3473%. The AME of the traditional NARNET is greater than the Fuzzy Networks. The hybrid models have until 0.17% lower error than the traditional network, this situation said that the best model of the present research is the Fuzzy Trapezoidal NARNET.

The estimation in-sample shows that the best model for forecast the Exchange Rate MX/US is the Differential Neural Network and the Fuzzy Trapezoidal NARNET followed by the Fuzzy GARCH. Then the dynamic of the volatility in the Exchange Rate can be modeled by the methods that combine Neural Networks with Fuzzy Theory and Variance Conditional techniques with Fuzzy Theory Models.

The principal conclusion is that hybrid fuzzy methods are better than traditional ARMA methods. In fact, the results illustrate until 0.20% of lower daily error for the fuzzy models and in the case of the Differential Neural Network of 0.50% lower error than the ARIMA, EGARCH, PARCH and NARNET methods.

Table 17. Absolute Mean Deviation comparison

Out Sample Test

Model	1 day	3 days	5 days	26 days
ARIMA (traditional)	0.4172 %	0.5929 %	0.6199 %	0.5982 %
E-GARCH (traditional)	0.3884 %	0.5584 %	0.6076 %	0.5925 %
PARCH (traditional)	**0.3387 %**	0.5523 %	0.6076 %	0.5922 %
Nonlinear Autoregressive Neural Network (traditional)	0.4111 %	0.5549 %	0.5954 %	0.5915 %
FTS-FUZZY ARIMA Tseng (fuzzy)	**0.2719 %**	**0.5025 %**	**0.5440 %**	**0.5873 %**
FTS-FUZZY ARIMA Tanaka (fuzzy)	0.5814 %	0.6160 %	0.5694 %	0.5971 %
FUZZY GARCH (fuzzy)	0.4225 %	0.5228 %	0.5637 %	0.5953 %
Gaussian FUZZY AR (fuzzy)	0.4111 %	0.5607 %	0.5935 %	0.5903 %
T-Student FUZZY AR (fuzzy)	0.3931 %	0.5494 %	0.5843 %	0.5884 %
FUZZY EGARCH (fuzzy)	0.4143 %	**0.5225 %**	**0.5635 %**	0.5915 %
FUZZY Trapezoidal NARNET (fuzzy)	**0.3486 %**	0.6007 %	0.6147 %	**0.5677 %**
FUZZY Triangular NARNET (fuzzy)	0.6463 %	**0.4802 %**	**0.4384 %**	0.5953 %
Differential Neural Network	0.3903 %	0.5356 %	0.5676 %	**0.5821 %**

Source: Own elaboration in Excel.

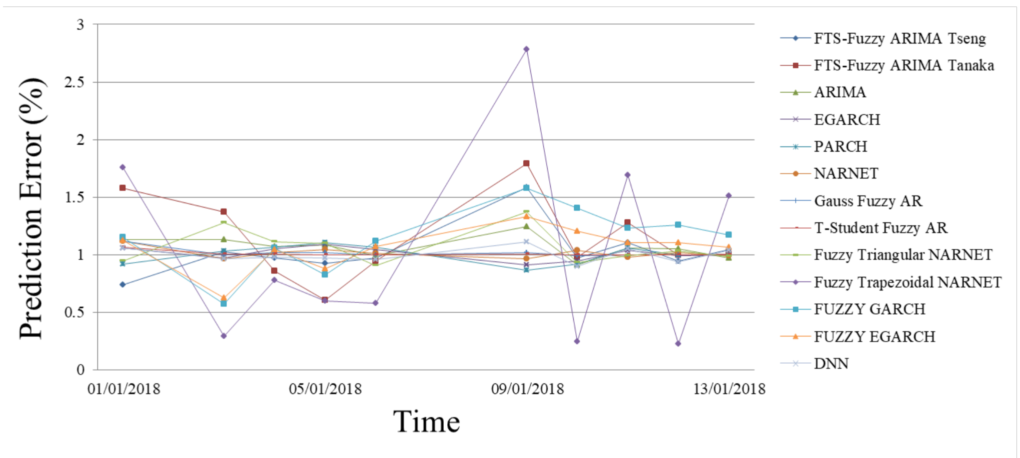
Now table 17denotes the comparisons of the Absolute Mean Deviation out-sample on different days. We found the best model in the day one of the forecasts was the FTS-FUZZY ARIMA Tseng with an error of 0.27% followed by the PARCH (0.33%) and the Fuzzy

Trapezoidal NARNET (0.34%), and the worst method was the Fuzzy Triangular NARNET (0.64%).

However, the models with minor errors in the first three days of prediction were the Fuzzy Triangular NARNET with 0.48%, the FTS-Fuzzy ARIMA Tseng (0.50%) and the Fuzzy EGARCH (0.52%). The methods with the major error were the FTS-Fuzzy ARIMA Tanaka (0.61%) and the Fuzzy Triangular NARNET (0.60%). For five days of forecast, the best model was Fuzzy Triangular NARNET (0.43%) followed by the FTS-Fuzzy ARIMA Tseng (0.54%) and the Fuzzy EGARCH (0.56%), the worst model is the ARIMA with an error of (0.61%). In summary, the evaluation of the first five days of the forecast is that the fuzzy hybrid models (FTS-Fuzzy ARIMA Tseng, Fuzzy Triangular NARNET and Fuzzy EGARCH) generate the value closets to the volatility of the exchange rate.

In 26 days the best models were the Fuzzy Trapezoidal NARNET (0.56% error), Differential Neural Network (0.582% error) and FTS-Fuzzy ARIMA Tseng (0.587 error); the test Absolute Mean Deviation denotes that the fuzzy models generate a better forecast than the traditional ARIMA, EGARCH, PARCH and NARNET.

Figure 47. Out-Sample Prediction Error Comparison



Source: Own elaboration in Excel.

Figure 47 denotes the percentage error of the prediction in the first 13 days, the results show that the fuzzy models developed best out-sample percentage test than the traditional models, it can be seen in the lower percentage error of each day are of the fuzzy techniques.

But the major error is of the fuzzy methods, these results are denoted in the last section in which the out-sample forecast of the suggested models generates major variability in their values.

Table 18. Percentage Prediction Error Comparison

Out Sample Test

Model	1 days	3 days	5 days	26 days
ARIMA (traditional)	1.1336 %	1.1109 %	1.0813 %	1.0255 %
E-GARCH (traditional)	1.0553 %	1.0261 %	1.0432 %	1.0041 %
PARCH (traditional)	**0.9205 %**	1.0041 %	1.0355 %	**1.0006 %**
Nonlinear Autoregressive Neural Network (traditional)	1.1171 %	1.0335 %	1.0316 %	1.0087 %
FTS-FUZZY ARIMA Tseng (fuzzy)	**0.7388 %**	**0.9123 %**	**0.9269 %**	1.0092 %
FTS-FUZZY ARIMA Tanaka (fuzzy)	1.5798 %	1.2717 %	1.0728 %	1.0454 %
FUZZY GARCH (fuzzy)	1.1482 %	**0.9261 %**	**0.9447 %**	1.0918 %
Gaussian FUZZY AR (fuzzy)	1.1172 %	1.0464 %	1.0309 %	1.0064 %
T-Student FUZZY AR (fuzzy)	1.0681 %	1.0248 %	1.0139 %	**1.0027 %**
FUZZY EGARCH (fuzzy)	1.1259 %	**0.9321 %**	0.9491 %	1.0236 %
FUZZY Trapezoidal NARNET (fuzzy)	**0.9473 %**	1.1120 %	1.0674 %	1.0256 %
FUZZY Triangular NARNET (fuzzy)	1.7563 %	0.9416 %	**0.8007 %**	1.1899 %
Differential Neural Network	1.0605 %	0.9999 %	0.9859 %	**0.9763 %**

Source: Own elaboration in Excel.

Finally, Table 18 illustrates the comparison of the Mean Percentage Prediction Error for the forecast on different days. For this test, in 26 days the best model is the DNN with 0.9763% of daily error out sample followed by the PARCH and T-Student Fuzzy AR. In this case, these methods lose variability and therefore the average errors tend to be lower than those with greater variability. Only the DNN follows a decreasing trend in how the time series. This test shows that the models no exists an important difference between the forecast of more rigorous models and the traditional methods in the out sample prediction.

The hypothesis of the present research is that the models of the Fuzzy Time Series Theory combined with Neural Networks generate a better forecast of the behavior of the exchange rate compared with the ARIMA model, EGARCH model, PARCH model and Autoregressive Neural Network. And the answer is that the hybrid models are better than the traditional models if the combinations are:

1. Fuzzy Time Series, Fuzzy ARIMA and ARMA models.

2. Conditional Variance methods and Fuzzy Theory.

3. Fuzzy Time Series Theory with Neural Networks.

Finally, the present chapter shows that the hybrids models have better forecasts that the models of the first and second categories, as said *García et al.* (2013). This research found that the hybrid fuzzy models generate the best forecast of the models used.

CONCLUSION AND RECOMMENDATIONS

We designed and developed the FTS-Fuzzy ARIMA, Gauss Fuzzy AR, Gaussian Fuzzy GARCH, Gaussian Fuzzy EGARCH, Fuzzy Triangular NARNET, and Fuzzy Trapezoidal NARNET Models, for forecast time series of the foreign exchange market. The results show that the methods of fuzzy methodology improvement the prediction compared to the cases in only the technique are used directly.

In this research, based on the Conditional Heteroskedasticity models we suggest new techniques (Fuzzy GARCH and Fuzzy EGARCH), and it is applied to estimate the foreign exchange MX pesos against EE.UU Dollar. The Fuzzy GARCH and Fuzzy EGARCH models show major effectiveness for forecast the behavior of the volatility in the exchange rates that the traditional methods.

Furthermore, the present search developed a new method of Autoregressive Neural Network combined with Sugeno-Type Fuzzy Theory, called Fuzzy Triangular NARNET and Fuzzy Trapezoidal NARNET. The first one improves the forecast because assume that the time series is a triangular fuzzy set, this situation recognizes the high volatility in the presence of good and bad news. On the other hand, the Fuzzy trapezoidal NARNET assumes that the time series are trapezoidal fuzzy set and identify the high and low volatility in the presence of good and bad news. These models generate a better forecast than the comparison of the NARNET model.

Otherwise, the new FTS-Fuzzy ARIMA assumes that the time series is a fuzzy time series and its behavior is captured by a fuzzy linear regression. In this situation, the proposed model generates a possibility forecast and generates the prediction value with major probability in the fuzzy interval. This method makes a better forecast that the traditional ARIMA and the Conditional Variance models.

The best models in this research were the Differential Neural Network, Fuzzy Trapezoidal NARNET and Fuzzy GARCH models In-Sample. And Out-Sample the best models were FTS-Fuzzy ARIMA Tseng, Fuzzy Triangular NARNET and Fuzzy EGARCH. We found that the models based on fuzzy theory have better estimates of volatility on financial time series because the objective function minimizes the error and generates fuzzy parameters that allow recognizing better the changes and abnormality of the exchange rate. Then the principal result is that the fuzzy logic achieves to identify the behavior of the volatility of the exchange rate better than the traditional models.

We found that the models based on fuzzy theory have better estimates of volatility in financial time series. This allows developing new prediction methods based on the structure of fuzzy logic, it is also necessary to establish an analysis of a greater number of data to try to discriminate if the effect of the error effect decreases.

Though results show that the In-Sample test the fuzzy models have an important improvement of the forecast in comparison with the traditional model when we used this model for analyzing the Out-Sample test the benefits of using these hybrid models are lower than the In-Sample test. In this case, the proposed methods are better than the traditional techniques In-Sample and Out-Sample test, however, the fuzzy models presented in this thesis need more tests with other exchange rates to better identify the results.

The recommendation is generated new prediction methods based on the structure of fuzzy logic, it is also necessary to establish an analysis of a greater number of data to try to discriminate if the effect of the error result decreases. In this situation, we propose analysis combinations as a fuzzy theory with different Neural Networks, Time series models, Limited Dependent Variable Models, and others.

BIBLIOGRAPHY

Alessandria, G., & Kaboski, J. P. (2011). Pricing to Market and the Failure of Absolute PPP. *American Economic Journal: Macroeconomics, 3*, pp. 91-127.

Aliyev, R., Salehi, S., & Aliyev, R. (2019). Development of Fuzzy Time Series Model for Hotel Occupancy Forecasting. *Sustainability*, 793.

Almeida, R. J., Basturk, N., Kaymak, U., & Sousa, J. M. (2013). Conditional Density Estimation Using Fuzzy GARCH Models. *Information Sciences, 267*, 173-181.

Arango, F. O., Llanos, A. I., & Herrera, F. L. (2013). Pronóstico de los índices accionarios DAX y S&P 500 con redes neuronales diferenciales. *Contaduría y Administración*, pp. 203-225.

Bai, E., Wong, W., Chu, W., Xia, M., & Pan, F. (2011). A heuristic time-invariant model for fuzzy time series forecasting. *Expert Systems with Applications, 38*, 2701–2707.

Banco de México. (2018). *Mercado Cambiario (Tipo de cambio)*. Retrieved 09 19, 2018, from Banco de México: http://www.banxico.org.mx/tipcamb/tipCamMIAction.do?idioma=sp

Bank of England. (2019). *Statistics, Exchanges Rates*. Retrieved from Bank of England: https://www.bankofengland.co.uk/statistics/exchange-rates

Benmouiza, K., & Cheknane, A. (2013). Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models. *Energy Conversion and Management*, pp. 561-569.

Billot, A. (1992). *Economic Theory of Fuzzy Equilibria An Axiomatic Analysis.* Paris: Springer.

Blanco, G. P. (2006). La Memoria Larga y la Cointegración Fraccional en las Series Temporales; una Aplicación a la PAridad de Poder de Compra en los Países del Cono Sur. 1-49.

Bojadziev, G., & Bojadziev, M. (2007). *Fuzzy Logic for Business, Finance, and Management.* Singapore: World Scientific Publishing Co. Pte. Ltd.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 307-327.

Brockwell, P. J., & Davis, R. A. (2006). Stationary Time Series. In P. J. Brockwell, & R. A. Davis, *Time Series: Theory and Methods* (pp. pp. 1-41). New York: Springer.

Cassel, G. (1918). Abnormal Deviations in International Exchanges. *The Economic Journal, 28*, pp. 413-415.

Chang, P. T. (1997). Fuzzy seasonality forecasting. *Fuzzy Sets and Systems, 90*(1), 1-10.

Chen, C.-H., Hong, T.-P., & Tseng, V. S. (2005). Analyzing time-series data by fuzzy data-mining technique. *Granular Computing, 2005 IEEE International Conference on*, 112-117.

Chen, S.-M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems, 81*, 311-319.

Chen, S.-M., & Hsu, C.-C. (2004). A New Method to Forecast Enrollments Using Fuzzy Time Series. *International Journal of Applied Science and Engineering*, 234-244.

Chen, S.-M., Xin-Yao, Z., & Gunawana, G. C. (2019). Fuzzy time series forecasting based on proportions of intervals and particle swarm optimization techniques. *Information Sciences*, 127-139.

Cheng, C., & Lee, E. S. (1999a). Nonparametric Fuzzy Regression--k-NN and Kernel Smoothing Techniques. *Computers and Mathematics with Applications, 38*, 239-251.

Cheng, C., & Lee, E. S. (1999b). Applying Fuzzy Adaptive Network to Fuzzy Regression Analysis. *Computers and Mathematics with Applications, 38*, 123-140.

Corbae, D., & Ouliaris, S. (1988). Cointegration and Tests of Purchasing Power Parity. *The Review of Economics and Statistics, 70*, pp. 508-511.

Coyaso, F. J., & Vermonden, A. (2015). Fuzzy Logic for Decision-Making and Personnel Selection. *Universidad & Empresa, 17*, 239-256.

Dalar, A. Z., & Eğrioğlu, E. (2018). Bootstrap Type-1 Fuzzy Functions Approach for Time Series Forecasting. *Trends and Perspectives in Linear Statistical Inference*, 69-87.

Dash, R., & Dash, P. (2016). An evolutionary hybrid Fuzzy Computationally Efficient EGARCH model for volatility prediction. *Applied Soft Computing*, 40-60.

Dash, R., Dash, P., & Bisoi, R. (2015). A differential harmony search based hybrid interval type2 fuzzy EGARCH model for stock market volatility prediction. *International Journal of Approximate Reasoning, 1*, 1-24.

Diamond, P. (1988). Fuzzy Least Squares. *Information Sciences, 46*, 141-157.

Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, pp. 83-106.

Dunyak, P. J., & Wunsch, D. (2000). Fuzzy regression by fuzzy number neural networks. *Fuzzy Sets and Systems*, 371-380.

Fang, Y., Lai, K. K., & Wang, S. (2008). *Fuzzy Portfolio Optimization Theory and Methods.* Berlin: Springer.

Flores, J. J., Calderón, F., & Lara, J. O. (2015). Aprendizaje de Modelos Difusos para Predicción de Series de Tiempo. *IZTATL COMPUTACIÓN, 8*, 10-26.

García, J., López, A. M., Romero, J., García, A. R., Camacho, C., Cantero, J. L., *et al*. (2002). Hojas de Cálculo para la Simulación de Redes Neuronales Artificiales. *Questtió, 26*, 289-305.

García, M. C., Jalal, A. M., Jalal, A. M., & López, J. M. (2013). Methods for Predicting Stock Indexes. *Ecos de Economía*(37), 51-82.

Gómez-Ramos, E. L., Venegas-Martínez, F., & Guerrero-Martínez, H. A. (2019). Relevancia del índice Dow Jones para AL: una aproximación con redes neuronales. *Quantitativa, Revista de Economía, 6*(1), 1-19.

Gómez-Ramos, E., & Venegas-Martínez, F. (2013). A Review of Artificial Neural Networks: How Well Do They Perform in Forecasting Time Series? *Analítika, Journal of Statistical Analysis, 6*(2), 7-15.

Gradojevic, N. (2007). Non-linear, hybrid exchange rate modeling and trading profitability in the foreign exchange market. *Journal of Economic Dynamics & Control, 31*, 557–574.

Granger, C. (1986). Developments in the Study of Cointegración Economic Variables. *Oxford Bulletin of Economics and Statistics*, 213-228.

Huarng, K., & Yu, H.-K. (2005). A Type 2 fuzzy time series model for stock index forecasting. *Physica A: Statistical Mechanics and its Applications, 353*, 445-462.

Hung, J.-C. (2009). A Fuzzy Asymmetric GARCH model applied to stock markets. *Information Sciences, 179*, 3930–3943.

Hung, J.-C. (2011a). Applying a combined fuzzy systems and GARCH model to adaptively forecast stock market volatility. *Applied Soft Computing, 11*, 3938–3945.

Hung, J.-C. (2011b). Adaptive Fuzzy-GARCH model applied to forecasting the volatility of stock markets using particle swarm optimization. *Information Sciences, 181*, 4673–4683.

Hung, J.-C. (2014). Robust Kalman filter based on a fuzzy GARCH model to forecast volatility using particle swarm optimization. *Soft Comput, 19*, 2861–2869.

INEGI. (1998). *Anuario estadístico del estado de Puebla 1998.* Puebla: INEGI.

Jang, J. R. (1993). ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on Systems*, pp.665-685.

Jantzen, J. (2007). Fuzzy Reasoning. In J. Jantzen, *Foundations of Fuzzy Control* (pp. pp. 13-46). Wiley.

Kahraman, C., Beskese, A., & Bozbura, F. T. (2006). Fuzzy Regression Approaches and Applications. In C. Kahraman, *Fuzzy Applications in Industrial Engineering* (pp. 589-612). Istanbul, Turkey: Springer-Verlag Berlin Heidelberg.

Karlsson, L. (2002). *GARCH-Modelling: Theoretical Survey, Model Implementation, and Robustness Analysis.* Kungl Tekniska Högskolan.: Karlsson, L. (2002). GARCH-Modelling: Theoretical Survey, Mo Unpublished master thesis.

Kim, B., & Bishu, R. R. (1998). Evaluation of fuzzy linear regression models by comparing membership functions. *Fuzzy Sets and Systems, 100*(1-3), 343-352.

Konings, L. S. (2010). La Conferencia de Bretton Woods. Estados Unidos y el dólar como Centro de la Economía Mundial. *Procesos Históricos, 18*, 72-81.

Kristjanpoller, W., & Michell, K. (2018). A stock market risk forecasting model through integration of switching regime, ANFIS and GARCH techniques. *Applied Soft Computing*, 106-116.

L.A.Zadeh. (1965). Fuzzy sets. *Information and Control, 8*, 338-353.

Lemma, T. A. (2018). Model Identification Using Neuro-Fuzzy Approach. In T. A. Lemma, & S. i. Intelligence (Ed.), *A Hybrid Approach for Power Plant Fault Diagnostics* (Vol. 743, pp. 37-74). Springer.

Lepskiy, A., & Suevalov, A. (2017). Application of Fuzzy Asymmetric GARCH-Models to Forecasting of Volatility of Russian Stock Market. *International Conference on Intelligent Information Technologies for Industry*, 286-294.

Leung, H., & Haykin, S. (1991). The complex backpropagation algorithm. *Transactions on Signal Processing*, pp. 2101-2104.

Li, D.-F. (2014). *Decision and Game Theory in Management with Intuitionistic Fuzzy Sets.* Republic of China: Springer.

Llanos, A. I., & Arango, F. O. (2012). Pronóstico del rendimiento del IPC (Índice de Precios y Cotizaciones) mediante el uso de redes neuronales diferenciales. *Contaduría y Administración*, pp. 63-81.

López, J. P. (2014). *Red Neuronal Feedforward como Estimador de Patrones de Corrientes en el Interior del Puerto de Manzanillo Sujeto a la Acción de Tsunamis.* Quintanaroo: Secretaría de Comunicaciones y Transportes.

Maysami, howe, & hamzah. (2004). Relationship between Macroeconomic Variables and Stock Market Indices: Cointegration Evidence from Stock Exchange of Singapore's All-S Sector Indices. *Journal Pengurusan*, 47-77.

McNelis, P. D. (2005). What Are Neural Networks? In P. D. McNelis, *Neural Networks in Finance: Gaining Predictive Edge in the Market* (pp. pp. 13-59.). Elsevier.

Morabito, F. C., & Versaci, M. (2002). Fuzzy Time Series for Forecasting Pollutants Concentration in the Air. In M. Marinaro, & R. Tagliaferri, *Neural Nets* (pp. 152-159). Vietri sul Mare, Italy: Springer.

Murias, P. G. (1998). La paridad de poder adquisitivo: concepto y evolución historica. *Estudios de economía aplicada., 9*, pp. 79-102.

Namugaya, J., Waititu, A. G., & Diongue, A. K. (2019). Forecasting Stock Returns Volatility on Uganda Securities Exchange Using TSK Fuzzy-GARCH and GARCH Models. *Reports on Economics and Finance*, 1-14.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica, 59*(2), pp. 347-370.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *The Econometric Society*, 347-370.

Ortiz-Arango, F., Cabrera-Llanos, A. I., & Venegas-Martínez, F. (2016). Euro Exchange Rate Forecasting with Differential Neural Networks with an Extended Tracking Procedure. *International Journal of Pure and Applied Mathematics., 107*(1), pp.87-109.

Ozawa, K., Watanabe, T., & Kanke, M. (1997). Fuzzy auto-regressive model and its applications. *1997 First International Conference on* (pp. 21-23). Adelaide, Australia: IEEE.

Özelkan, E. C., & Duckstein, L. (2000). Multi-objective fuzzy regression: a general framework. *Computers & Operations Research, 27*, 635-652.

Pal, S. S., & Kar, S. (2017). Fuzzy Time Series Model for Unequal Interval Length Using Genetic Algorithm. *Advances in Intelligent Systems and Computing*, 205-216.

Pindick, R. S., & Rubinfeld, D. L. (2001). Modelos lineales de series de tiempo. In R. S. Pindick, & D. L. Rubinfeld, *Econometría: Modelos y Pronósticos* (pp. pp 547-574). Ciudad de México: McGraw-Hill.

Popov, A., & Bykhanov, K. (2005). Modeling Volatility of Time Series Using Fuzzy GARCH Models. *Science and Technology*, 687-692.

Poznyak, A. S., Yu, W., Sanchez, E. N., & Perez, J. P. (1999). Nonlinear Adaptive Trajectory Tracking Using Dynamic Neural Networks. *IEEE Transactions On Neural Networks*, pp. 1402 - 1411.

Reinert, K. A., & Rajan, R. S. (2009). *Encyclopedia of the World Economy.* Woodstock, United Kingdom: Princeton University Press.

Rivero, S. S. (2011). Teorías del Tipo de Cambio. *Tendencias y Nuevos Desarrollos de la Teoría Económica, 858*, pp. 23-34.

Rutkowski, L. (2004). Elements of the Theory of Fuzzy Sets. In L. Rutkowski, *Flexible Neuro-Fuzzy Systems: Structures, Learning and Performance Evaluation* (pp. pp. 7-26). Boston: Kluwer Academic Publishers.

S.Atsalakis, G., G.Atsalaki, I., Pasiouras, F., & Zopounidis, C. (2019). Bitcoin price forecasting with neuro-fuzzy techniques. *European Journal of Operational Research*, 770-780.

S.R.Singh. (2007). A simple method of forecasting based on fuzzy time series. *Applied Mathematics and Computation, 186*, 330-339.

Singh, P. (2017). A brief review of modeling approaches based on fuzzy time series. *International Jornal Machine Learning & Cyber*, 397-420.

Singh, P. (2017). A brief review of modeling approaches based on fuzzy time series. *International Journal of Machine Learning and Cybernetics*, 397–420.

Singh, P. (2017). A brief review of modeling approaches based on fuzzy time series. *Internation Journal of Machine Learning and Cybernetics.*, 397–420.

Singh, P., & Borah, B. (2014). Forecasting stock index price based on M-factors fuzzy time series and particle swarm optimization. *International Journal of Approximate Reasoning, 55*, 812–833.

Smith, K. A. (2000). Neural Networks for Business: An Introduction. In K. A. Smith, & J. N. Gupta, *Neural networks in business: techniques and applications for the operations researcher* (pp. pp. 1-25.). Computers & Operations Research.

Song, Q., & Chissom, B. S. (1993a). Fuzzy time series and its models. *Fuzzy Sets and Systems, 54*, 269-277.

Song, Q., & Chissom, B. S. (1993b). Forecasting enrollments with fuzzy time series — Part I. *Fuzzy Sets and Systems, 54*, 1-9.

Song, Q., & Chissom, B. S. (1994). Forecasting enrollments with fuzzy time series — part II. *Fuzzy Sets and Systems, 62*, 1-8.

Sriboonchitta, S., Sriboonjit, J., Chokethaworn, K., Chaitip, P., & Chaiboonsri, C. (2011). Forecasting of International Tourists' Expenditure in Thailand: Using AFIMAX and ARFIMAX-GARCH Approach. *Thailand Econometric Society*, 220-232.

Sugeno, M., & Takagi, T. (1993). Fuzzy Identification of Systems and Its Applications to Modeling and Control. *Fuzzy Sets for Intelligent Systems*, 387-403.

Tanaka, H. (1987). Fuzzy Data Analysis by Possibilistic Linear. *Fuzzy Sets and Systems*, 363-375.

Tanaka, H., Asai, K., & Uejima, S. (1982). Linear Regression Analysis with Fuzzy Model. *IEEE Transactions On Systems, Man, and Cybernetics, 12*(6), 903-907.

Terano, T., Asai, K., & Sugeno, M. (1992). In T. Terano, K. Asai, & M. Sugeno, *Fuzzy Systems Theory and its Applications* (pp. 69-84). San Diego: Academic Press.

Tsai, C.-C., Wu, S.-J., & Ting, W.-H. (2002). On the effect of membership function in a high-order fuzzy time series. *Fuzzy Systems*, 12-17.

Tsaur, R.-C. (2012). A Fuzzy Time Series-Markov Chain Model with an Application to Forecast the Exchange Rate Between the Taiwan and US Dollar. *International Journal of Innovative Computing, Information, and Control, 8*, 4931{4942.

Tsaur, R.-C., Yang, J.-C. O., & Wang, H.-F. (2005). Fuzzy relation analysis in the fuzzy time series model. *Computers & Mathematics with Applications, 49*, 539-548.

Tseng, F.-M., & Tzeng, G.-H. (2002). A fuzzy seasonal ARIMA model for forecasting. *Fuzzy Sets and Systems, 126*, 367-376.

Tseng, F.-M., Tzeng, G.-H., Yu, H.-C., & Yuan, B. J. (2001). Fuzzy ARIMA model for forecasting the foreign exchange market. *Fuzzy Sets and Systems, 118*, 9-19.

Vovan, T. (2019). An improved fuzzy time series forecasting model using variations of data. *Fuzzy Optimization and Decision Making*, 151-173.

Wang, C.-C. (2011). A comparison study between the fuzzy time series model and the ARIMA model for forecasting Taiwan export. *Expert Systems with Applications, 38*, 9296–9304.

Wu, B., & Hsu, Y. Y. (2002). On Multivariate Fuzzy Time Series Analysis and forecasting. In P. Grzegorzewski, O. Hryniewiski, & M. A. Gil, *Soft Methods in Probability, Statistics, and Data Analysis* (pp. 363-372). Warsaw, Poland: Springer-Verlag.

Wu, S., & Er, M. J. (2000). Dynamic Fuzzy Neural Networks; A Novel Approach to Function Approximation. *IEEE Transactions on Systems, Man, and Cybernetics*, pp. 358-364.

Yabuuchi, Y., Watada, J., & Toyoura, Y. (2004). Fuzzy AR Model of Stock Price. *Scientiae Mathematicae Japonicae Online, 10*, 485-492.

Yen, K., Ghoshray, S., & Roig, G. (1999). A linear regression model using triangular fuzzy number coefficients. *Fuzzy Sets and Systems, 106*, 167-177.

Yu, H.-K. (2005a). Weighted fuzzy time series models for TAIEX forecasting. *Physica A: Statistical Mechanics and its Applications, 349*, 609-624.

Yu, H.-K. (2005b). A refined fuzzy time-series model for forecasting. *Physica A: Statistical Mechanics and its Applications, 346*, 567-681.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, pp. 338-353.

Zemke, S. (1999). Nonlinear index prediction. *Physica A: Statistical Mechanics and its Applications*, 177-183.

Zhang, H., & Watada, J. (2019). Building Fuzzy Levy-GJR-GARCH American Option Pricing Model. *International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making*, 197-209.

Zhang, Z. (2018). Artificial Neural Network. In Z. Zhang, *Multivariate Time Series Analysis in Climate and Environmental Research* (pp. pp. 1-36). Springer, Cham.