UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO COPPEAD DE ADMINISTRAÇÃO

RICARDO FELÍCIO DE SOUZA

INVESTIGATING THE USE OF FUZZY INFERENCE SYSTEMS: AN APPLICATION IN THE BEAUTY INDUSTRY'S DEMAND FORECAST

Rio de Janeiro
2018

RICARDO FELÍCIO DE SOUZA

INVESTIGATING THE USE OF FUZZY INFERENCE SYSTEMS: AN APPLICATION IN THE BEAUTY INDUSTRY'S DEMAND FORECAST

Master's dissertation presented to the Instituto Coppead de Administração, Universidade Federal do Rio de Janeiro, as part of the mandatory requirements in order to obtain the degree of Master in Business Administration (M.Sc.).

SUPERVISOR: Peter Fernandes Wanke, D.Sc.

Rio de Janeiro
2018

CIP - Catalogação na Publicação

Fs719i

Felício De Souza, Ricardo
INVESTIGATING THE USE OF FUZZY INFERENCE SYSTEMS: AN APPLICATION IN THE BEAUTY INDUSTRY'S DEMAND FORECAST / Ricardo Felício De Souza. -- Rio de Janeiro, 2018.
69 f.

Orientador: Peter Fernandes Wanke.
Dissertação (mestrado) - Universidade Federal do Rio de Janeiro, Instituto COPPEAD de Administração, Programa de Pós-Graduação em Administração, 2018.

1. Computação. 2. Sistemas computacionais. 3. Redes neurais artificiais. 4. Administração - Teses. I. Fernandes Wanke, Peter, orient. II. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a).

RICARDO FELÍCIO DE SOUZA

INVESTIGATING THE USE OF FUZZY INFERENCE SYSTEMS: AN APPLICATION IN THE BEAUTY INDUSTRY'S DEMAND FORECAST

Master's dissertation presented to the Instituto Coppead de Administração, Universidade Federal do Rio de Janeiro, as part of the mandatory requirements in order to obtain the degree of Master in Business Administration (M.Sc.).

Approved by:

Peter Fernandes Wanke

Peter Fernandes Wanke, D.Sc
(COPPEAD/UFRJ)

Otavio Henrique dos Santos Figueiredo

Otavio Henrique dos Santos Figueiredo, D.Sc
(COPPEAD/UFRJ)

Virgilio José Martins Ferreira Filho

Virgilio José Martins Ferreira Filho, D.Sc
(COPPE/UFRJ)

Rio de Janeiro

2018

ACKNOWLEDGEMENTS

First of all, I thank god for giving me capacity and motivation to elaborate such an interesting and applied research.

This work would not be possible without the support of my entire family, but in special, my brother, Thiago Felício, whose strong statistic and programming knowledges were my north in many situations.

I am grateful to my supervisor, Peter Wanke, who granted me with the opportunity and the directions to accomplish this rewarding study, and to introduce Bruna Cunha, who I thank for using your professional influence to make this work feasible.

Finally, I would also like to thank all my 2016 COPPEAD cohort for such incredible 2 years' experience, and the COPPEAD staff for helping all of us in all sort of problems.

ABSTRACT

SOUZA, Ricardo Felício de. Title: INVESTIGATING THE USE OF FUZZY INFERENCE SYSTEMS: AN APPLICATION IN THE BEAUTY INDUSTRY'S DEMAND FORECAST. 2018. 69p. Dissertation project (Master in Business Administration) - Instituto COPPEAD de Administração, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2017

This study proposes 4 different forecasting tools based on fuzzy inference systems that consist on different combinations of classic and soft computing models. The forecasting tools were tested with 27 products of the nail polish line of a worldwide beauty company and the results were compared to the company's forecasts, that comprise qualitative decisions. The results were analyzed by the mean absolute percentage error and the percentage better dimensions, so it was possible to determine the characteristics and conditions that makes each model the fittest for each situation. The main takeaways were that low kurtosis, negatively skewed demand series and longer forecast horizons favors the fuzzy model. These results suggest that the fuzzy forecasting tool should be prioritized in the longer-term forecasts and be also considered over the qualitative decision for series with less extreme values.

Keywords: forecasting, soft computing, fuzzy inference system, neural network, genetic algorithm.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABIHPEC – Associação Brasileira Industria de Higiene Pessoal Perfumaria e Cosméticos
ANN – Artificial Neural Network
ANVISA – Agência Nacional de Vigilância Sanitária
APE – Absolute Percentage Error
ARIMA – Autoregressive Integrate Moving Average
CV – Coefficient of Variation
GA – Genetic Algorithm
MAPE – Mean Absolute Percentage Error
MA – Moving Average
NN – Neural Network
PB – Percentage Better
SARIMA – Seasonal Autoregressive Moving Average

SUMMARY

1. INTRODUCTION

1.1. Research motivation

Forecasting future events has always been a paradigm for the humankind. From weather forecast to human behavior, almost everything that substantially impacts in human lives, people desire to predict. To deal with the future uncertainty, lots of efforts were spent in order to have the best prediction, and, as a result, there is a huge range of forecasts types and styles. From qualitative to quantitative, from mystical methods, as divinations, card games and crystal balls, to complex calculations based on historic data, like neural networks and autoregressive methods. Consequently, the high number of forecasts solutions makes it a huge field of study, with a constant potential to be developed.

Being accurate on forecasting is fully attached to a comparative and competitive advantage among peers. It happens because the right planning can drive to success or failure, by creating opportunities to companies optimize the resources needed to complete a task. In the business environment, to deal with future uncertainty and working on mitigating it risks, companies are forced to figure out ways and methods to forecast many variables that impacts in its operations. One of the main drivers among those variables is the customer (SLACK; BRANDON-JONES; JOHNSTON, 2013).

The “customer”, in this case, can simply be translated into the demand, what drives the company’s existence. Sub-planning the demand, and consequently its operations, may drive a company to lose clients, and to harm its brand, since there will be lower products available than it should. It also can drive the company to lose gains of scale opportunities and expose the company to emergency costs, like extra hours or excessive logistics costs to properly supply its demands. However, over-planning the service level increase the company’s cost, which can also jeopardize the whole operations, harming the company’s competitiveness.

Forecasts are, undoubtedly, necessary to help managers to make decisions about resourcing the organization for the future, and helps driving many policies inside the whole company. It is important to notice that the demand forecast impact each sector in a different way (WANKE, 2010). For example, the marketing department, needs information to prepare a new product or invest in marketing campaign, the production department needs it to prepare its capacity, and the finance department is impacted, by the whole stock in the company, translated into working capital in balance sheets. In addition to that, It may have also different impact in different industry sectors: in the service industry for example, it is critical, since it is impossible to stock services; On the other side, in the manufacturing, it is more complex, where the demand planning is designed over several stages (WANKE, 2010).

This study is based on the Brazilian subsidiary of a French multinational company that is inserted in the cosmetic and personal care industry. Forecasting in this environment is highly complex, since the cosmetic and personal care industry is highly innovative, where new products and new trends are constantly being developed and established (ABIHPEC, 2015). Like the fashion industry, it is hard to perform forecasts due to its dynamic and competitive profile, which scenarios constantly change. Brazil is the 4th consumer market in world, and the Brazilian industry is composed by 2642 companies registered by ANVISA, where 75% of the total revenues are represented by 20 companies (ABIHPEC, 2017).

The dynamic environment is a challenge for classic forecasting methods, which depends on defined historic patterns as trend, and seasonality. For dynamic cases, soft computing tools may have a better fit, since gaps between the classic mathematic models and reality are reduced, by incorporating human behavior on computational models. This study aims to revise the mechanism developed by Vroman, Happiette and Rabenasolo (1998), where a fuzzy inference system is used to give weights to determined variables, and apply to a forecast combiner, as modified by Yesil, Kaya and Siradag (2012). The final objective is to explore the tools proposed with different forecast methods, which includes classic forecast methods as the moving average, ARIMA, and exponential smoothing, and also includes soft computing tools, as neural networks, fuzzy logic and genetic algorithm.

1.2. Study structure

This study is organized in five chapters. The first chapter presents the study's objectives and its relevance. The second chapter introduces the literature review around the main themes included in the study. Firstly, the demand forecast problem and, as part of it, the forecast techniques used in this research: moving average, ARIMA and SARIMA, exponential smoothing and neural networks. Secondly, the fuzzy logic and its forecasting applications are reviewed. Thirdly, the genetic algorithm theme is explored and then, its applications in the fuzzy inference systems. The third chapter presents the applied methodology, discussing how the different models were structured and describing the software tools used. The fourth chapter introduces the models' evaluation including the exploration of the data sets, the evaluation methodology, and the results discussion. Finally, the fifth chapter brings the study's conclusion, its limitations and suggestions for future researches.

2. LITERATURE REVIEW

2.1. Time series forecast

A time series is a sequence of observations taken sequentially in time. Stock prices, temperature recordings, monthly sales, are all examples of time series. An intrinsic feature of a time series is that, typically, adjacent observations are dependent (BOX; JENKINS; REINSEL, 2008). This interdependence is of fundamental importance for general interest, for example, for forecasting issues, where this dependency is extrapolated to the future. However, to understand this relationship it is necessary to develop models to deeply investigate the time series, which means not only analyzing it, but also building the model, identifying it, fitting it, and, finally, checking it (BOX; JENKINS; REINSEL, 2008).

In the last decades the quantitative methods have been deeply investigated and developed as seen in De Gooijer and Hyndman (2006) review. It means that there are several models to better understand the time series and provide a robust forecast. However, it's important to notice that certain forecast methods have better accuracy in certain circumstances (HOGARTH, R, M; MAKRIDAKIS, 1981), what makes choosing it a difficult task, since there is not a universal best model.

This section provides information about the models further used in this study: firstly, the simplest one, the moving average model. Then, the autoregressive models and the exponential smoothing model are presented. Finally, the artificial neural network (ANN) model is presented.

2.1.1. Moving Average model

The moving-average forecast technique consists on forecast the next period's demand by calculating the average of the n previous periods. Since n is constant, for the next forecast, a new period is inserted, and the older one is not considered anymore in the calculation. The moving average formula is disclosed below:

$$F_t = \frac{\sum_{i=1}^{n} A_{t-i}}{n} \tag{1}$$

Where F_t is the forecast for the time t, and A_t is the actual value for the time t.

It’s important to notice that identic weights are given to all n values included in the calculation. It means that the impact of the inclusion of new values in n depends on n size (WANKE, 2010). According to Slack, Brandon-Jones and Johnston (2013), the value of n can be set at any level, but is usually in the range 4 to 7.

Moving average is one of the most popular methods used in practice for short-term forecasting, due to its simplicity. This method should be applied only in time series without trend and seasonality, due to its probability of getting unsatisfactory results (WANKE, 2010). However, it might work reasonably well when the trend in data changes direction frequently (YESIL; KAYA; SIRADAG, 2012), which might fit in the Brazilian beauty industry, due to its competitive and innovative profile.

2.1.2. ARIMA and Seasonal ARIMA models

Autoregressive moving average models use past errors correlation pattern to extrapolate predictions. Yule (1927), in his investigation on disturbances periodicity suggested that this pattern could be repeated in the future like a pendulum, in his words. Since then, this technique has been extensively developed and applied for different kinds of problems, from specific to general ones. Consequently, several authors influenced the ARIMA improvement along its history.

The first models were designed to face simpler linear and stationary problems. However, the improvement enabled the model to be fitted into many other complex situations, including non-stationary and multivariate models.

Box and Jenkins (1976) integrated the literature available by the time in a book, which had an enormous impact on the theory and practice of modern time series analysis and forecasting (DE GOOIJER; HYNDMAN, 2006). In this book, it was proposed a three-stage iterative cycle for time series identification, estimation and verification, today commonly known as the Box-Jenkins approach. However, it is important to verify the series stationarity and seasonality conditions before following the steps (WANG, 2011; ZHANG, 2003). Unit root tests, as the Augmented Dickey-Fuller test (DICKEY; FULLER, 1981), or the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) (KWIATKOWSKI et al., 1992) test are tools to perform these tasks.

The general form of ARIMA (p,q,d) model is described below (BOX; JENKINS; REINSEL, 2008)

$$\varphi(B) = \phi(B)\nabla^d z_t = \theta_0 + \theta(B)a_t \tag{2.1}$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \tag{2.2}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \tag{2.3}$$

Where p is the autoregressive order, q the moving average order, d is the number of differencing operations. $\phi(B)$ is the autoregressive operator; $\varphi(B) = \phi(B)\nabla^d$ is the generalized autoregressive operator; $\theta(B)$ is the moving average operator; When $d = 0$, the model represents a stationary process.

The original model was then improved, and later it was able to consider seasonal pattern. To illustrate, the general form of Seasonal ARIMA (p, d, q) (P, D, Q) model is described below (BOX; JENKINS; REINSEL, 2008)

$$\varphi(B)\phi(B^s)(1 - B^s)^D(1 - B)^d z_t = \theta(B)\Theta(B^s)a_t \tag{3.1}$$

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p \tag{3.2}$$

$$\phi(B^s) = 1 - \phi_1 B^s - \phi_2 B^{s2} - \cdots - \phi_P B^{sP} \tag{3.3}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \tag{3.4}$$

$$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{s2} - \cdots - \Theta_Q B^{sQ} \tag{3.5}$$

Where p is the autoregressive order, q the moving average order, d is the number of differencing operations, and P, D and Q are the corresponding seasonal orders.

A common obstacle for many people in using Autoregressive Integrated Moving Average (ARIMA) models for forecasting is that the order selection process is usually considered subjective and difficult to apply. (HYNDMAN; KHANDAKAR, 2008). However there is already an extensive literature on this area (E. J. HANNAN; RISSANEN, 1982; GÓMEZ, 1998; HYNDMAN; KHANDAKAR, 2008; MÉLARD; PASTEELS, 2000), proposing algorithms for model identification.

This study uses the model presented by Hyndman and Khandakar (2008) which works on the software R Studio. This algorithm selects the appropriate model order by using unit root tests and the Akaike's Information Criterion (AIC), which is based on the likelihood penalization (HYNDMAN; KHANDAKAR, 2008)

$$AIC = -2\,log(L) + 2(p + q + P + Q + k) \tag{4}$$

In this algorithm, firstly, unit roots tests are performed where d and D are found. Then, using these values, iterations in the models are done considering values possible for p, q, P and Q, where the objective is to achieve the minimum AIC. Whenever a model with lower AIC is found, it becomes the new “current” model and the procedure is repeated, until the optimization is finished.

2.1.3. Exponential Smoothing model: Holt-Winters Additive

Exponential smoothing models have this designation due to unequal weights given to past values (MAKRIDAKIS, WHEELWRIGHT; HYNDMAN, 1998). According to Gardner (2006) this model was firstly originated during World War II for military purposes. Nowadays, this model is widely used because of its simplicity, and low cost in relation to its relatively good accuracy (MAKRIDAKIS, WHEELWRIGHT; HYNDMAN, 1998). The general simple model is shown below (GARDNER, 2006)

$$S_t = \alpha A_t + (1-\alpha)S_{t-1} \tag{5}$$

Where:

α : Smoothing parameter for the level of the series

S_t : The forecast for time t and also smoothed level of the series

A_t: Observed value of the time series in period t

Few decades after World War II, this model started being adapted to be applied in demand forecasting, where real situations usually have issues with trends and seasonality (MAKRIDAKIS; HIBON, 2000). Its original author, Brown (1959), and Holt (1957), in independent studies were the pioneers to adapt this model to these real world behaviors. Winters (1960) tested Holt’s methods with empirical data, when it became known as the Holt-Winters forecast system. Since then, there are several variants of the original model, which considers different behaviors (multiplicative and additive) in trend, seasonality and errors. The general form of the exponential smoothing method considering additive seasonality and trend is (GARDNER, 2006):

$$S_t = \alpha(A_t - I_{t-p}) + (1-\alpha)(S_{t-1} + T_{t-1}) \tag{6.1}$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1} \tag{6.2}$$

$$I_t = \delta(A_t - S_t) + (1-\delta)I_{t-p} \tag{6.3}$$

$$X_{t+k} = S_t + \mathrm{k} \cdot T_t + I_{t-p+k} \tag{6.4}$$

Where:

α : Smoothing parameter for the level of the series

γ: Smoothing parameter for the trend

δ: Smoothing parameter for seasonal indices

S_t : The forecast for t periods and also smoothed level of the series, computed after X_t is observed.

T_t : Smoothed additive trend at the end of period t

I_t: Smoothed seasonal index at the end of period t

A_t: Observed value of the time series in period t

X_t: Estimated value of the time series in period t

An issue in this methodology, also seen in ARIMA models, is to fit the best model and parameters for each situation. This study uses the algorithm presented by Hyndman and Khandakar (2008) to automatically achieve the best model to be fitted in the problem's data. This algorithm is programed in the package forecast in R. Firstly, for each data series all appropriate models are applied, optimizing the parameters of the model in each case (HYNDMAN; KHANDAKAR, 2008). Then, in order to select a model, the program uses the likelihood methodology proposed by Ord, Koehler and Snyder (1997) and the Akaike's Information Criterion (AIC) to define the best model to be fitted in the respective data.

2.1.4. Artificial Neural Networks model

Artificial neural network is a widely used forecasting model that has been reviewed and tested by the last decades and have shown impressive results. Accurate results of this forecasting model can be found in energy demand (DARBELLAY; SLAMA, 2000; HIPPERT; PEDREIRA; SOUZA, 2001), dairy products demand (JACOBS; ZANINI; COSTA, 2015), airline passengers, (NAM; SCHAEFER, 1995; FARAWAY; CHATFIELD, 1998) financial time series and stock market (CHEN; LEUNG; DAOUK, 2003; KIMOTO et al., 1990) and many other areas, such as engineering, social, etc.

Successful ANN researches in the forecasting field started appearing in the late 80s (LAPEDES; FARBER, 1987) and an extensive exploration by the academy took place since the 90s, what can be identified by many literature reviews (ADYA; COLLOPY, 1998; DASE; PAWAR, 2010; DE GOOIJER; HYNDMAN, 2006).

The idea behind the ANN model is that inputs, or dependent variables, get filtered through one or more hidden layers each of which consisting on hidden units, or nodes, before they reach the output variable (DE GOOIJER; HYNDMAN, 2006). Single hidden layer feed forward network is the most widely used neural network model form for time series modeling and forecasting (ZHANG; PATUWO; HU, 1998). In this standard model, the relationship between the output (y_t) and the inputs ($y_{t-1}, \dots, y_{t-p}$) has the following mathematical representation:

$$y_t = w_0 + \sum_{j=1}^{q} w_j \cdot g\left(w_{0.j} + \sum_{i=}^{p} w_{i.j} \cdot y_{t-i}\right) + \varepsilon_t \tag{7}$$

Where $w_{i.j}(i = 0,1,2, \dots, p, j = 1,2, \dots, q)$ and $w_j(j = 0,1,2, \dots, q)$ are model weight parameters; p is the number of input nodes, and q is the number of hidden nodes. In other words, the feed-forward neural network is fitted with lagged values of y as inputs and a single hidden layer with size nodes q. In fact, then, the model performs a nonlinear functional mapping from past observations to the future.

$$y_t = f(y_{t-1}, \dots, y_{t-p}, w) + \varepsilon_t \tag{8}$$

Where w is a vector of all parameters and $f()$ is a function determined by the network structure and connection weights. In other words, the neural network is equivalent to a nonlinear auto-regressive model (KHASHEI; BIJARI, 2010).

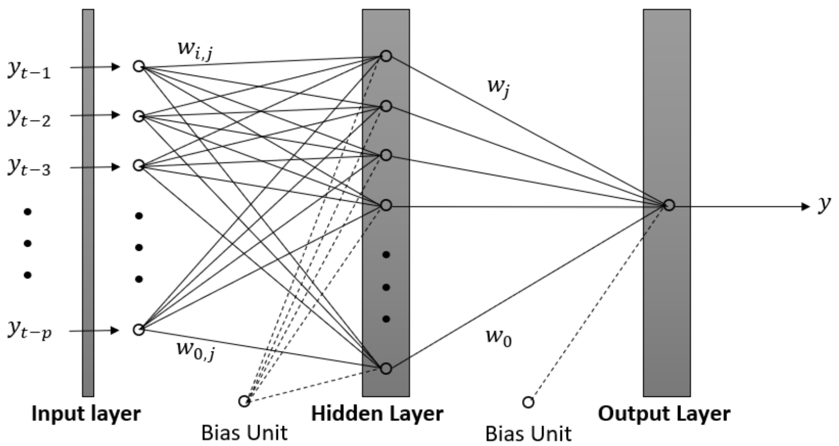

Figure 1. Neural network structure
Source: Adapted from khashei and Bijari (2010)

The model applied in this study is the default single hidden layered, where the number of nodes in this layer is half of input nodes plus 1. A total number of networks are fitted, each with random starting weights, and then averaged when computing forecasts. This algorithm is programed in the package forecast in R (HYNDMAN; KHANDAKAR, 2008)

The great advantage between the ANN and classic forecast models is the flexible non-linear modeling capability (ZHANG, 2003), pattern usually found in real world time series (ZHANG; PATUWO; HU, 1998). The main difference between ANN-based forecasts and traditional model-based ones is that few priori assumptions are taken before the model is applied. What happens is that the ANN model is largely determined by the characteristics of data, from where the model “learns” or “trains. In addition to that, ANNs are universal approximators which can approximate a large class of functions with a high degree of accuracy (ZHANG, 2003), what makes this model a robust tool, with high adaptability to many complex situations.

2.2. Fuzzy Logic

Firstly published by Zadeh (1965), professor of Computing Science in the University of California, the fuzzy set theory is based on the principles developed by Lukasiewicz (1878-1956), which combines the Boolean logic principles, published in 1847 (based on the Aristotelian logic), and multivalued membership levels in sets. Different from the classic logic based on the Aristotelian theory, which considers crisp values, as 0 or 1 and true or false, the fuzzy logic considers values inside 0 to 1 interval, in which a value can be “half true” or “half false”, for example.

Since there are gaps between the classic mathematic models and empirical interpretations from reality (ROSS, 2004; ZIMMERMANN, 2010), the theory become relevant, because it is capable to tolerate imprecision, uncertainty, and partial truth to achieve tractability and robustness on simulating human decision-making behavior (KO; TIWARI; MEHNEN, 2010)

The fuzzy set theory is based on recognition that certain sets have imprecise boundaries in which the transition from membership to non-membership in a subset of a reference set is gradual rather than abrupt (KO; TIWARI; MEHNEN, 2010; ROSS, 2004). It happens because natural languages are imprecise in the sense that everything is a matter of degree, which depends on perceptions (NOVÁK, 2005; ZADEH, 2008), as seen in the words “young”, “middle aged” and “old” for example. In the fuzzy set theory, the natural language sentences are labels in fuzzy sets, which represent the values of the linguistic variables. In this way, as the example, “young”, “middle aged” and “old” are values of the linguistic variable “age” shown in Figure 2.

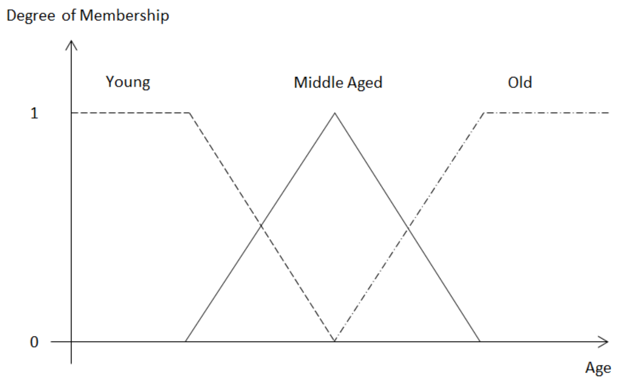

Figure 2. Membership function example
Source: From the author

What happens is that a person in one determined age can be 10% in the "young" group and 90% in the "middle aged" group, according to the membership function behavior, being partially in two groups.

The Fuzzy logic enables the uncertainty, inherent of human behavior, to be incorporated on computational models, reducing the gap between theory and reality (ZADEH, 2008). There are several successful applications of the fuzzy logic (MAIERS; SHERIF, 1985; TÜRKŞEN, 2009), in computing and engineering (AMMERLAAN; WRIGHT, 2004; GEERING, 1998; SUHAIL; KHAN, 2005), supply chain applications (GEN; TSUJIMURA; ZHENG, 1997; HSIEH, 2002; WANG, 2009), machining (SINGHAL et al., 2016), forecasting (FRANTTI; MAHOMEN, 2001; JACQUIN; SHAMSELDIN, 2009; JARRETT; PLOUFFE, 2011; TSENG; TZENG, 2002), and many other areas.

2.2.1. Forecasting and Fuzzy Logic

The most popular combination of forecasting and fuzzy logic was firstly proposed by Song and Chissom (1993a, 1993b), with the fuzzy time series theory. Fuzzy time series are composed by expressions instead of numbers, as a common time series is. In their study, to exemplify the fuzzy time series, the authors cite weather conditions (e.g. "cold", "cool", "good", "very hot") and mood (e.g. "very very good", "not too bad"), where linguistic variables are used to describe them. The variation interval of this kind of time series become more realistic when approached by a fuzzy set definition than to have assigned a number of these linguistic variables (SONG; CHISSOM, 1993a). To propose the forecast mechanism the authors assume that there is a causal relationship between the observations at time t and those at previous times. In other words, the modeling process is, in essence, to develop fuzzy relations among the observations at different times of interests. (SONG; CHISSOM, 1993a). Song and Chissom (1993b) then applied the theory on a numeric time series, by changing the numbers to linguistic variables to forecast university enrollments. The historic data was transformed into fuzzy values, through membership functions, and a fuzzy time series was set. Based on that time series, relationships among historical values were defined, and, based on these relationships, the forecast was estimated. This technique was widespread since then, what can be seen in many studies (CHEN, 1996; CHENG; CHANG; YEH, 2006; LEE et al., 2012; POULSEN, 2009; SINGH, 2007; UNION, 2008; WANG, 2011).

The original model was then adapted, and mixed with other soft computing methods, as the neural network. Short term load forecast (STLF) is a deep field of study in this area, where hybrid fuzzy models were developed using variate combination methods (BAKIRTZIS et al., 1995; DARBELLAY; SLAMA, 2000; HO; HSU; YANG, 1992; HOLMUKHE et al., 2010; KIM et al., 1995). According to Srinivasan and Lee (1995) the methods can be classified in four groups: fuzzy logic system at the output stage of the neural network forecaster to manipulate the output, fuzzy logic at the input stage of a neural network to preprocess the inputs, integrated fuzzy neural network to create a fuzzy rule base from the historical training data, and separate fuzzy logic and neural network forecasters to forecast different components of the load.

The forecasting use of fuzzy inference system has now approaches that are still based into the transformation of the numeric series into fuzzy values, but in many of them it’s no longer necessary a relationship determination to achieve a final forecast. In these situations, the fuzzy inference system is used as a support tool that adapts the forecast model along the data. Vroman, Happiette and Rabenasolo, (1998) are pioneers in studying fuzzy based hybrid forecasting methods in the textile industry. The authors develop a hybrid forecasting tool by using the fuzzy system to dynamically adapt the holt-winters parameters, which is outperformed by the new tool. Yesil, Kaya and Siradag (2012), apply a hybrid fuzzy model to forecasting in the fast fashion industry. In their study they develop a fuzzy logic tool to combine other statistical models to achieve a final forecast.

2.3. Genetic Algorithms optimization

Genetic algorithms are optimization methods based on the theory of populations’ evolution. Firstly introduced by Holland (1975) this methodology follows the idea proposed by Charles Darwin in “On the Origin of Species” where the natural selection is set based on the survivor of the fittest. According to Darwin, the higher the capacity of an individual in fitting an environment, the higher his probabilities of surviving and generating descendants. What happens is that genes from the adapted individuals will spread to an increasing number of individuals in each successive generation, evolving the species to become more and more suited to their environment.

Based on this idea, genetic algorithms methods also works with populations of individuals, where each individual represents a possible solution to an optimization problem. In this way, each individual has its own fitting, based on its optimization capacity among others. Individuals with higher fit in relation to others are enabled to reproduce by cross breeding with others in the population of possible solutions, on the other hand, individuals with lower fit dies without transferring characteristics to the next generation. This procedure is repeated n times, where good characteristics are spread throughout the population, being mixed and exchanged with other good characteristics (BEASLEY; BULL; MARTIN, 1993). This improves the population fit by exploring promising characteristics, and the resulting population tends to converge to an optimal solution to the problem. A canonical genetic algorithm model is shown below:

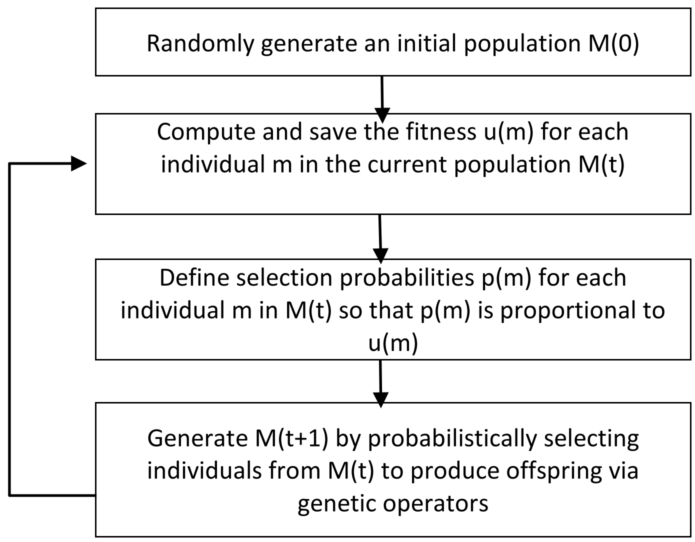

Figure 3. A canonical genetic algorithm
Source: Adapted from De Jong (1992)

Although the GA is a powerful optimization tool, it does have certain weaknesses in comparison to other optimization techniques. There are several optimization methods, where, in many cases are better than genetic algorithm, which are slow and, in mostly simpler problems, are still evaluating the first generations when other methods already achieved the final result (LACERDA; CARVALHO, 1999). In addition to that, due to the randomness of the GA operation, it is difficult to predict its performance, a factor that is crucial for hard-deadline, real-time applications (TANG et al., 1996).

However, real problems are complex and, to achieve good results, flexible tools are needed. Genetic algorithms can fit in these situations due to many factors: It's not necessary a full mathematic understanding of the considered problem; simultaneous researches are performed along the research space, since it is done with a population instead of a point, what avoids local maximums/minimums; it accepts a large number of variables to be optimized; it tolerates incomplete data and noise, it is relatively easy to be implemented, it is flexible to work with arbitrary restrictions and optimize multiple functions with conflicting objectives (LACERDA; CARVALHO, 1999).

The most traditional genetic algorithm research field is concentrated in the numerical function optimization (BEASLEY; BULL; MARTIN, 1993), where it have been shown to be able to outperform conventional optimization techniques on difficult, discontinuous, multimodal, noisy functions (DE JONG, 1975). However it has been done researches and applications in many other fields as image processing, combinatorial optimization, design and machine learning (BEASLEY; BULL; MARTIN, 1993).

2.3.1. GA and Fuzzy Logic's Membership Function

Together, GAs and Fuzzy Logic Controllers possess the capabilities necessary to produce powerful, efficient, and robust adaptive control systems (KARR, 1993). Karr (1991) and Thrift (1991) were pioneers to use GA in determination of membership functions and other Fuzzy Logic Controllers parameters. According to Karr (1993), "Such controllers are more suitable than past control systems for recognizing, quantifying, and adapting to changes in the problem environment".

Since then this technique has been widely applied in many other situations (AL-ADWAN et al., 2013; ALCALÁ-FDEZ et al., 2009; HERRERA; LOZANO; VERDEGAY, 2005; KISSI et al., 2003; LIU et al., 2001; PEDRYCZ, 1995; SEPTEM RIZA et al., 2014; SHIMOJIMA; FUKUDA; HASEGAWA, 1995). Herrera, Lozano and Verdegay (2005) states that the performance of a fuzzy logic controller depends on its control rules and membership functions, and, consequently, "it is very important to adjust these parameters to the process to be controlled". Karr (1991) states that GAs appear to be effective, versatile, and straightforward enough to locate high-performance membership functions in complex control problems. What happens is that due to the optimization complexity of the fuzzy system and, in addition to that, it needs to be flexible in many situations, GA is a usefull tool to improve the fuzzy inference system's results.

3. MODELING

3.1. Structure

The model's input is a time series, containing sales historic data, and its output is the final forecast for the periods $t+1$ and $t+3$. The model's structure is composed by two main units, the **forecasting unit**, mainly represented by the generation of classical forecast methods and the **combining unit**, represented by the fuzzy inference system, what can be identified in the figure 4.

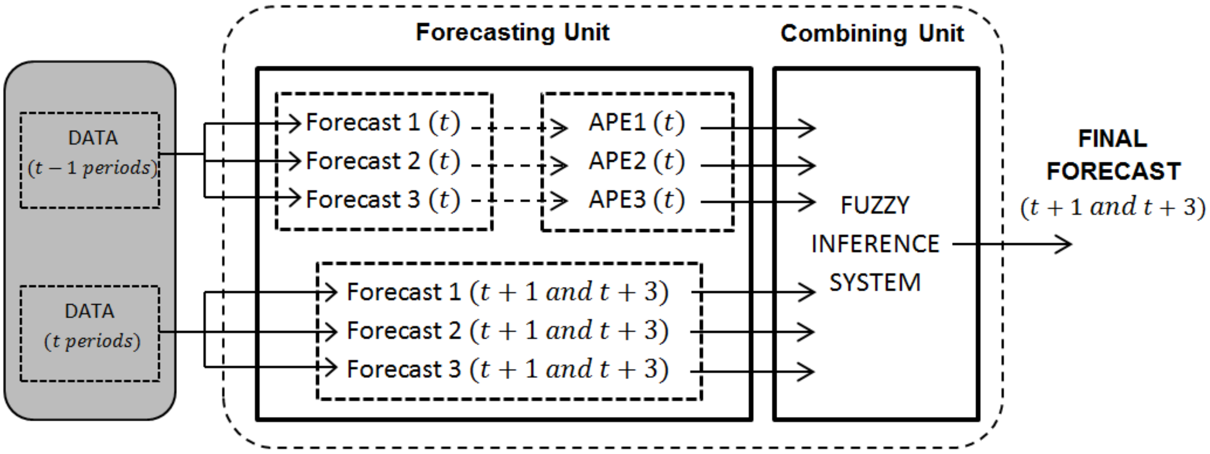

Figure 4. Fuzzy Forecast System
Source: From the author

3.1.1. Forecasting Unit

The forecasting unit's input is a time series provided by the company, which is composed by the monthly sales of nail polish. Based on this data, this part of the system has three tasks: (1) to generate forecasts for the actual month; (2) to calculate the absolute percentage error (APE) of the actual month's forecast; and, finally, (3) to generate forecasts for the next month and for three months ahead.

Since there are four forecasts models in this study, it was set two different systems, resulted from two combinations of four different forecasting models: the

SYSTEM A, which is composed by the Moving Average (5 periods), Exponential Smoothing (Holt-Winters), and ARIMA; and the SYSTEM B, which is composed by Artificial Neural Network, Exponential Smoothing (Holt-Winters), and ARIMA. This information is summarized in the table 1.

Table 1. Summary of forecast models

System A	System B
Forecast 1: Moving Average (5 periods)	Forecast 1:Neural Networks
Forecast 2: Exponential Smoothing (HW)	Forecast 2: Exponential Smoothing (HW)
Forecast 3: ARIMA	Forecast 3: ARIMA

Source: From the author

The forecasts models were chosen due to its relative high popularity, simple application, and relative low cost among all the statistic models available, as previously clarified in the literature review. The confrontation between moving average and neural network was set due to the huge technological gap between the methods. The moving average is the simple mean and do not consider any seasonality and trend but, as explained, might fit into some specific series patterns. On the other side, the neural network model is one of the most up-to-date among the others, which uses a modern soft-computing, and represents a disruption in the moving average technique that guides both the exponential smoothing and the auto regressive moving average. Consequently, the proposition is to set a comparison between the possible different impact on choosing these two “opposite” models.

All forecasts were generated using R Studio. The 5 period moving average forecast was calculated based on the normal average along the data. For all the other methods, that include parameters optimizations, they were achieved using rolling windows along the data. In other words, for each period the forecasts were generated using only past information and the parameters were redefined in every step. The methodology is better disclosed in the flowchart below representing the rolling window forecast algorithm that generates new parameters for all the forecasts generated.

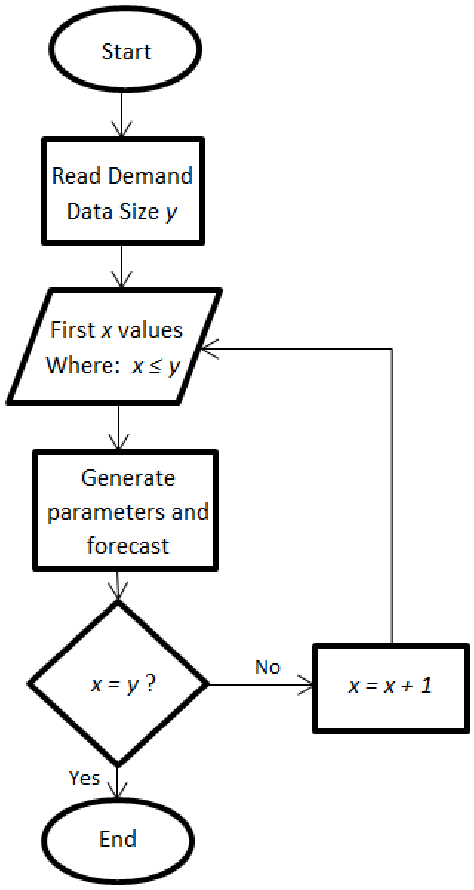

Figure 5. Rolling window forecast algorithm
Source: From the author

The Exponential smoothing was calculated using, firstly, the function *HoltWinters* to automatically fit the model parameters, and then the function *forecast.holtwinters* to execute the forecasts. The ARIMA forecast was calculated by using the function *auto.arima* to automatically fit the best model parameters, which includes seasonal ARIMA, and the forecast itself was calculated through the function *forecast.Arima.* Both functions *HoltWinters* and *auto.arima* use the algorithm proposed by Hyndman and Khandakar, (2008) to achieve the models' best parameters. The neural network model was calculated using the function *nnetar* to fit a neural network model and the function *forecast* to calculate the respective forecasts. All the functions cited above were found in the package *Forecast* (HYNDMAN; KHANDAKAR, 2008) or in its sub packages.

The APEs were calculated confronting the forecasts with the real data on the time t, using the formula below:

$$APE_t = \frac{|R_t - F_t|}{R_t} \tag{9}$$

Where $\mathrm{R_t}$ is the real data in the month t and $\mathrm{F_t}$ is the forecast for the month t.

The final outputs of the forecasting unit are APEs of each forecast model in the time t and forecasts for time $t+1$ and $t+3$, also for each forecast model

3.1.2. Combining Unit

The combining unit's inputs are the forecasting unit outputs. The combining unit consists in a rule-based fuzzy inference system that combines the forecasts into a final number, giving weights based on the respective APEs results. The fuzzy inference system is composed by five parts (JANG, 1993) as seen in the figure 5. The rule base block is a set of linguistic rules or conditional statements in the form of: "IF a set of conditions is satisfied, THEN a set of consequences are inferred". The database part contains the membership functions of the inference system. The fuzzification interface part transforms input numeric variables into degrees of match corresponding the linguistics values. The decision-making unit performs the inference operations based on the defined rules. Finally, the defuzzification interface block transforms fuzzy results into numeric results.

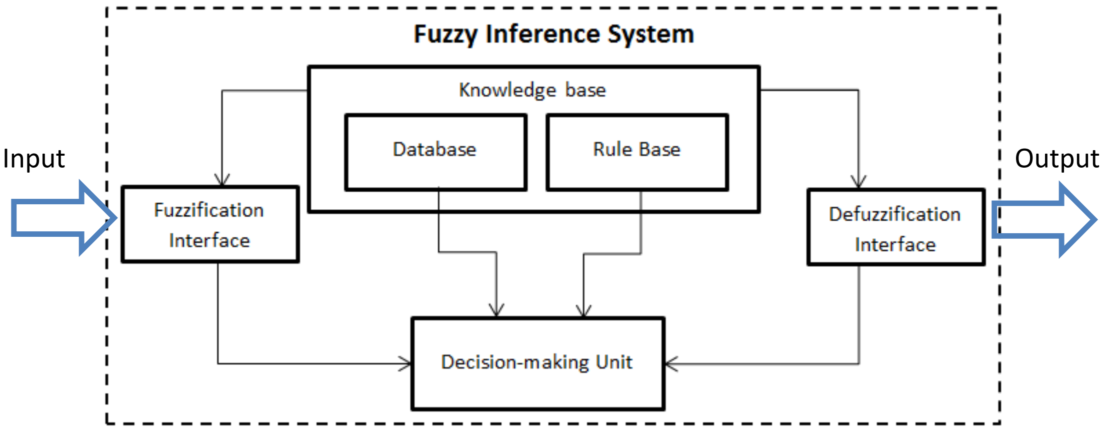

Figure 6. Fuzzy Inference System
Source: Adapted from Jang (1993)

Based on this structure, the inference system performs a step by step reasoning. Firstly, in the fuzzification process, crisp input variables are compared with the membership functions (database) and then transformed into fuzzy numbers. Secondly, the membership values are combined, where a final weight is given for each rule case. Thirdly, depending on the weight given, the final consequent is generated. Finally, in the defuzzification process the consequents are aggregated to produce a crisp value, which, in this case, represents the final forecasts. After defined all the input variables, which are the APEs and the forecasts ($t + 1$ and $t + 3$), and the output variable, which is the final forecast, it is necessary to define the system's rules and membership functions.

The fuzzy reasoning applied in this study is the Takagi Sugeno Kang (TSK) (SUGENO; KANG, 1988; TAKAGI; SUGENO, 1985). In this type of reasoning the output of each rule is a linear combination of the input variables, where the final output is the weighted average of each rule's output. Hence, it can be seen as a combination of linguistic and mathematical regression modeling in the sense that the antecedents describe fuzzy regions in the input space in which the consequent functions are valid (YESIL; KAYA; SIRADAG, 2012). The model's target is to give more weight in the final forecast for models that had lower APE's in the previous period. The antecedent proposition has three variables, namely APE values for each forecast method, and each APE is labeled by three linguistic terms: LOW, MEDIUM and HIGH.

The rules of the proposed system are listed in the table 2. The rules were based on the study proposed by Yesil, Kaya and Siradag (2012), however, the model presented in this study has one more membership function, demanding 19 more intermediary scenarios. The new weighting scenarios were achieved by rounding intermediary weights, driven by the scenario LOW. For instance LOW/LOW/HIGH is weighted 0,45/0,45/0,1 and LOW/HIGH/HIGH is weighted 0,8/0,1/0,1, respectively. The new scenario LOW/MEDIUM/HIGH is then set: firstly LOW is 0,6 by rounding the average between the previous ones. Consequently, the rest is 0,4, where MEDIUM takes the bigger part 0,3, so it gets more weight than the highest error, which will have 0,1 part.

Table 2. If-Then Rule Base

Rule #	if	APE 1	and	APE 2	and	APE 3	Then	TSK RULE (weights)
1	if	**LOW**	and	**LOW**	and	**LOW**	Then	0,333 F1 + 0,333 F2 + 0,333 F3
2	if	**LOW**	and	**LOW**	and	**MEDIUM**	Then	0,4 F1 + 0,4 F2 + 0,2 F3
3	if	**LOW**	and	**LOW**	and	**HIGH**	Then	0,45 F1 + 0,45 F2 + 0,1 F3
4	if	**LOW**	and	**MEDIUM**	and	**LOW**	Then	0,4 F1 + 0,2 F2 + 0,4 F3
5	if	**LOW**	and	**MEDIUM**	and	**MEDIUM**	Then	0,5 F1 + 0,25 F2 + 0,25 F3
6	if	**LOW**	and	**MEDIUM**	and	**HIGH**	then	0,6 F1 + 0,3 F2 + 0,1 F3
7	if	**LOW**	and	**HIGH**	and	**LOW**	then	0,45 F1 + 0,1 F2 + 0,45 F3

8	if	**LOW**	and	**HIGH**	and	**MEDIUM**	then	0,6 F1 + 0,1 F2 + 0,3 F3
9	if	**LOW**	and	**HIGH**	and	**HIGH**	then	0,8 F1 + 0,1 F2 + 0,1 F3
10	if	**MEDIUM**	and	**LOW**	and	**LOW**	then	0,2 F1 + 0,4 F2 + 0,4 F3
11	if	**MEDIUM**	and	**LOW**	and	**MEDIUM**	then	0,25 F1 + 0,5 F2 + 0,25 F3
12	if	**MEDIUM**	and	**LOW**	and	**HIGH**	then	0,3 F1 + 0,6 F2 + 0,1 F3
13	if	**MEDIUM**	and	**MEDIUM**	and	**LOW**	then	0,25 F1 + 0,25 F2 + 0,5 F3
14	if	**MEDIUM**	and	**MEDIUM**	and	**MEDIUM**	then	0,333 F1 + 0,333 F2 + 0,333 F3
15	if	**MEDIUM**	and	**MEDIUM**	and	**HIGH**	then	0,4 F1 + 0,4 F2 + 0,2 F3
16	if	**MEDIUM**	and	**HIGH**	and	**LOW**	then	0,3 F1 + 0,1 F2 + 0,6 F3
17	if	**MEDIUM**	and	**HIGH**	and	**MEDIUM**	then	0,4 F1 + 0,2 F2 + 0,4 F3
18	if	**MEDIUM**	and	**HIGH**	and	**HIGH**	then	0,5 F1 + 0,25 F2 + 0,25 F3
19	if	**HIGH**	and	**LOW**	and	**LOW**	then	0,1 F1 + 0,45 F2 + 0,45 F3
20	if	**HIGH**	and	**LOW**	and	**MEDIUM**	then	0,1 F1 + 0,6 F2 + 0,3 F3
21	if	**HIGH**	and	**LOW**	and	**HIGH**	then	0,1 F1 + 0,8 F2 + 0,1 F3
22	if	**HIGH**	and	**MEDIUM**	and	**LOW**	then	0,1 F1 + 0,3 F2 + 0,6 F3
23	if	**HIGH**	and	**MEDIUM**	and	**MEDIUM**	then	0,2 F1 + 0,4 F2 + 0,4 F3
24	if	**HIGH**	and	**MEDIUM**	and	**HIGH**	then	0,25 F1 + 0,5 F2 + 0,25 F3
25	if	**HIGH**	and	**HIGH**	and	**LOW**	then	0,1 F1 + 0,1 F2 + 0,8 F3
26	if	**HIGH**	and	**HIGH**	and	**MEDIUM**	then	0,25 F1 + 0,25 F2 + 0,5 F3
27	if	**HIGH**	and	**HIGH**	and	**HIGH**	then	0,333 F1 + 0,333 F2 + 0,333 F3

Source: From the author

To determine the membership functions' profiles of each linguistic term, this study takes two approaches: the Fuzzy System 1, which the parameters were defined through specialists opinions (PETROVIC; SWEENEY, 1994; ROSS, 2004; SEPTEM RIZA et al., 2014) and; Fuzzy System 2, which the parameters were defined by an optimization using the genetic algorithm method, where it is adjusted along time, as also proposed by the literature (HERRERA; LOZANO; VERDEGAY, 2005; KISSI et al., 2003; SEPTEM RIZA et al., 2014; SHIMOJIMA; FUKUDA; HASEGAWA, 1995).

3.1.2.1. Fuzzy System 1 – Specialist Based Static Membership Function

To take the specialists' point of view it was formulated a questionnaire, which is disclosed in the APPENDIX A. It was asked, mainly, what were a low and a high APE in the experts' opinions. For simplification it was assumed trapezoidal membership functions for both LOW and HIGH linguistics terms (YESIL; KAYA; SIRADAG, 2012) and triangular membership function to MEDIUM linguistic term. MEDIUM term was considered the mean between LOW and HIGH, and those last linguistic terms' membership functions do not overlap. 7 company's specialists involved with the forecasting decision answered the questionnaire and the final result is the average among all answers.

The one-step ahead results showed that an APE lower than 15% is considered totally low and an APE higher than 30% is considered totally high. Consequently, the medium APE function is centered on 22.5%. The graphic is in the figure 6.

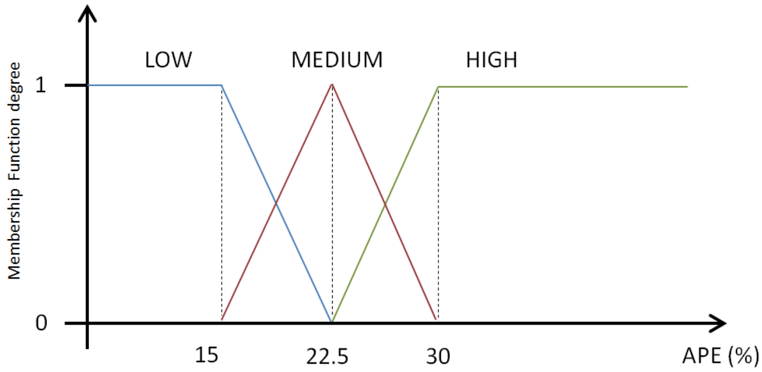

Figure 7. One-step ahead APE Membership Functions
Source: From the author

The three-step ahead results showed that an APE lower than 20% is considered totally low and an APE higher than 32.5% is considered totally high. Consequently, the medium APE function is centered on 26.25%. The graphic is in the figure 7.

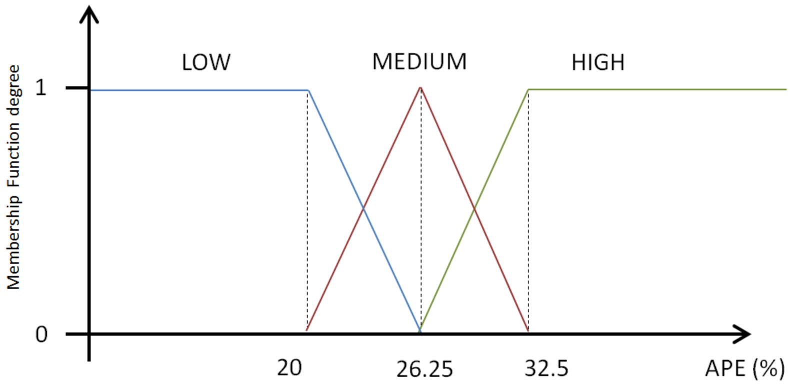

Figure 8. Three-step ahead APE Membership Functions
Source: From the author

3.1.2.2. Fuzzy System 2 – Genetic Algorithm Based Dynamic Membership Function

The other approach for defining the membership functions is data driven, through the genetic algorithm optimization. This strategy was investigated by many authors (AL-ADWAN et al., 2013; HERRERA; LOZANO; VERDEGAY, 2005; KARR, 1993; KARR; GENTRY, 1993; LIU et al., 2001; SHIMOJIMA; FUKUDA; HASEGAWA, 1995), and it is known that GA is effective, versatile, and straightforward enough to locate high-performance membership functions in complex control problems (KARR, 1991).

This approach consists proposing the functions parameters based on the Mean Absolute Percentage Error (MAPE) of three periods ($(t-2)$, $(t-1)$ and (t)). Then, the optimum parameters found these three periods are used in the fuzzy inference system parameters of the forecast of period $t+1$.

In this method, the membership functions types were kept the same: Trapezoidal for HIGH and LOW, and triangular for MEDIUM. To simplify the optimization process, 2 variables were used: x and y, where the x is the center of the MEDIUM function, and y is the distance from this point to the values purely HIGH or LOW, as seen in the figure 8.

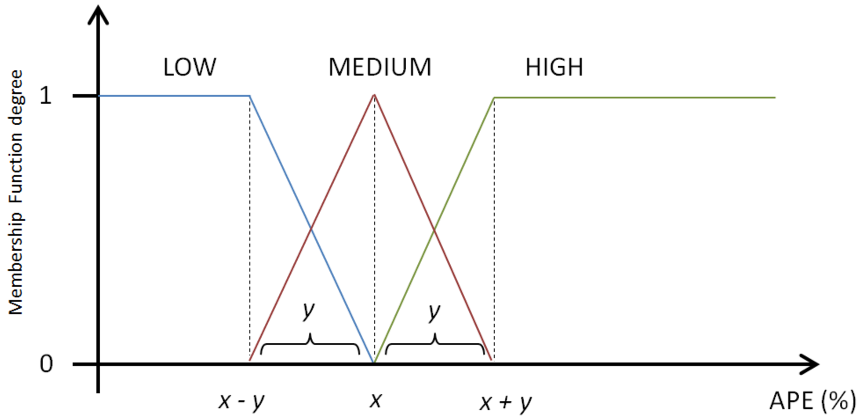

Figure 9. Genetic Algorithm Optimization Variables
Source: From the author

The function to be optimized, is the Mean Absolute Percentage Error (MAPE) of 3 periods before $t+1$ ($(t-2)$, $(t-1)$ and (t)), which the variables are x and y. The optimization process' flowchart is in the figure 9. Firstly, three types of classic forecasts are generated for the periods $t-2$, $t-1$ and t. Then a fuzzy model with variables x and y is set and optimized in function of the MAPE. The final results, x and y are applied in the fuzzy inference system to generate the final $t+1$ forecast.

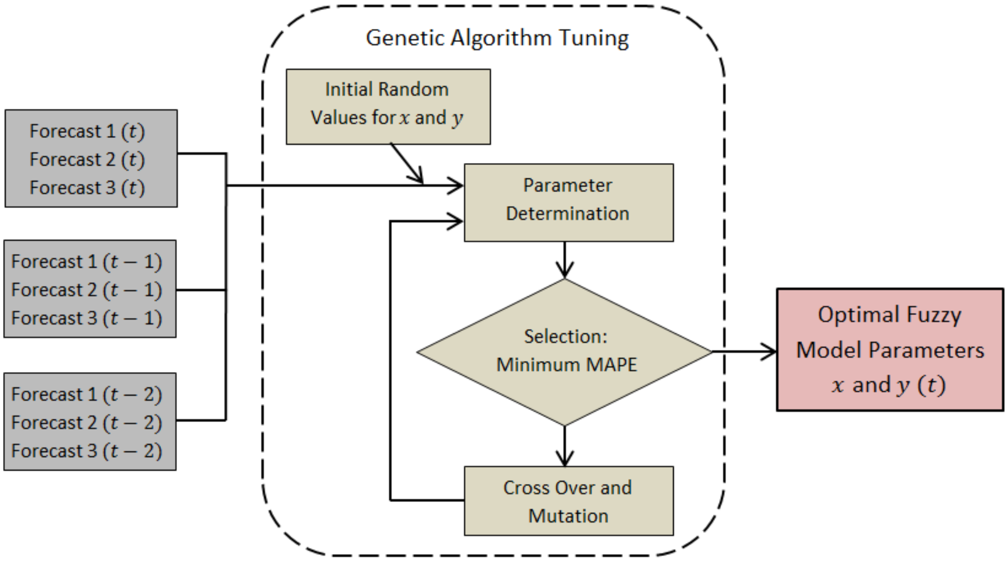

Figure 10. Membership Function Optimization Flowchart
Source: From the author

The GA optimization was performed through the function *ga* in the R package *GA* (SCRUCCA, 2013). The parameters were set as the function's default, which means that the probability of crossover is 0.8, the probability of mutation is 0.1 and the top 5% of individuals survive for the next generation

To summarize, four models were tested in this study: (1) Fuzzy System A-1, composed by Moving Average, Exponential Smoothing and ARIMA forecast models, with membership functions determined by experts; (2) Fuzzy System A-2, composed by Moving Average, Exponential Smoothing and ARIMA forecast models, with membership functions determined by genetic algorithm optimization. (3) Fuzzy System B-1, composed by Neural Network, Exponential Smoothing and ARIMA forecast models, with membership functions determined by experts; (4) Fuzzy System B-2, composed by Neural Network, Exponential Smoothing and ARIMA forecast models, with membership functions determined by genetic algorithm optimization. The previous description is summarized in the figure 10.

Fuzzy Inference System (FIS)	(System 1) Experts	(System 2) Genetic Algorithm
(System A) Moving Average, Exponential Smoothing, ARIMA	Fuzzy System A-1	Fuzzy System A-2
(System B) Neural Network, Exponential Smoothing, ARIMA	Fuzzy System B-1	Fuzzy System B-2

Figure 11. Summary of FIS models
Source: From the author

4. MODEL EVALUATION

4.1. Datasets

4.1.1. Input Dataset

As previously informed, the models' inputs are the historic sales data of nail polish of a multinational company in the beauty industry. The time series are composed by 66 observations each, from September 2011 to February 2017, on a monthly basis. The products demand data were submitted to a multiplicative factor, to avoid data exposure. The products names were also changed by numbers from 1 to 27 for the same purpose.

For illustration, the observed demand of the products 1 and 2 are shown in the figure 11 and 12 respectively.

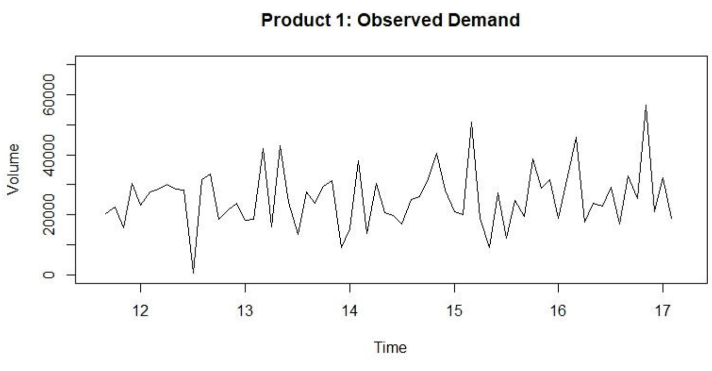

Figure 12. Product 1 demand
Source: Company

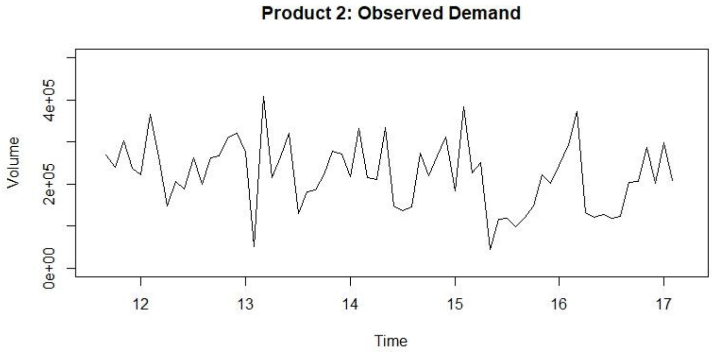

Figure 13. Product 2 demand
Source: Company

It is possible to see that the data has a high volatility, with huge variations along its history. Both products have similar profiles, despite having different volumes. Looking at the long-term point of view, in the product 1 it is possible to infer that there is a slight increase trend, while in the product 2 might be defined a slight decrease trend.

4.1.2. Benchmark Dataset

The forecasts were compared with the company's forecasts on the referred periods. The company's forecasts are the final forecast decision for the respective periods, and are not only a statistical forecast. Firstly, the statistic forecast is generated and then, this value represents an input for a monthly meeting composed by the company's experts, who are the representatives of the company's areas as Operations, Finance Management, Trade Marketing and Commercial Planning. In these meetings, other variables are considered, as the investment in marketing, inflation, industry entrants, among others, that the experts understand that will impact in the final decision. The final forecast decision is then taken, after the inter-organizational discussion.

For illustration, the company's forecasts of the products 1 and 2 are shown in the figure 13 and 14 respectively, in comparison to the observed data. The company's forecasts are in red and the observed data is in black.

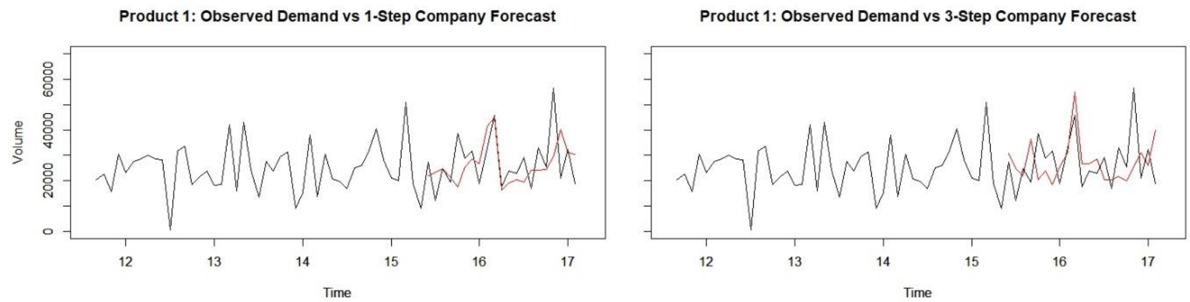

Figure 14. Product 1: 1-step and 3-step ahead company's forecast
Source: Company

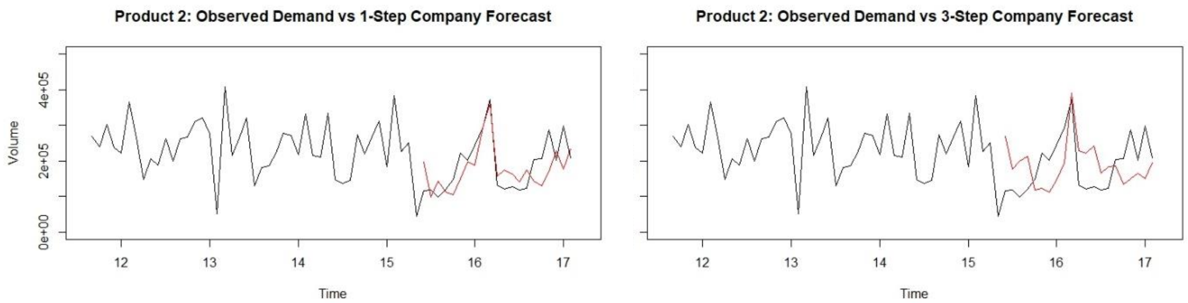

Figure 15. Product 2: 1-step and 3-step ahead company's forecast
Source: Company

4.2. Methodology of Evaluation

4.2.1. Hardware and Software used

The hardware used was an Intel Core i7-5600U processor at 2,60 GHz with 8GB of RAM installed. The operational system was a 64-bit Windows 10 Pro, and the application used to calculate all the forecasts and errors was the R Studio 1.0.153, based on R 3.4.4.

4.2.2. Functions Parameters

This section aims to clarify the main functions parameters designed to the forecasts and optimizations. All the models were used in accordance to the default proposed by the developers. To clarify, the table below discloses the main parameters

Table 3. Main packages informations

Package	Functionality	Main Default Information
GA	Genetic Algorithm Optimization	Mutation probability: 0,1 Elitism 2 Crossover probability 0,8 Population Size: 50 Iterations: 100
Stats	Exponential Smoothing Forecast	Seasonal: additive Parameters defined by minimizing squared prediction error
Forecast	Neural Network Forecast	Hidden Layer: (input nodes/2)+1 AIC to fit seasonality
Forecast	ARIMA Forecast	KPSS test for order of first differencing OCSB test for order of seasonal differencing

Source: From the author

4.2.3. Accuracy Measures

The results of the forecasts were compared to the benchmark time series using two different accuracy measures: mean absolute percentage error (MAPE) and percentage better (PB). The MAPE, is a widely used method for comparing forecasting tools, as can be seen in many studies in the field (HYNDMAN; KHANDAKAR, 2008; KAASTRA; BOYD, 1996; MAKRIDAKIS et al., 1993; VROMAN; HAPPIETTE; RABENASOLO, 1998). The formula, as previously cited, consists on the mean of all the absolute percentage errors observed, as can be seen below:

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\frac{|R_t - F_t|}{R_t} \qquad (10)$$

Where $\mathrm{R_t}$ is the real data in the month t, $\mathrm{F_t}$ is the forecast for the month t, and n is the number of observations.

The percentage better, present in the M3-Competition (MAKRIDAKIS; HIBON, 2000) is a relatively less used method than MAPE, but its importance is to measure how many times a given model has a smaller absolute percentage error than its benchmark, which in this study is the company's forecasts. Percentage better will be further identified as PB in this study.

4.2.4. Results

As a result, for each of the 27 products, it was generated 8 forecasts series with 21 periods length each, which consisted in: forecasting models Fuzzy System A-1, Fuzzy System A-2, Fuzzy System B-1 and Fuzzy System B-2 for one-step ahead forecast; and the same Fuzzy System A-1, Fuzzy System A-2, Fuzzy System B-1 and Fuzzy System B-2 models for three-step ahead forecast.

For illustration, the MAPE and PB results for the products 1 and 2 are disclosed in the table 9. All results for all products are disclosed in the Appendix B.

Table 4. Forecast Errors for Products 1 and 2

PRODUCT	FORECAST HORIZON	FORECAST TYPE	MAPE COMPANY (%)	MAPE FUZZY (%)	PERCENTAGE BETTER (%)
1	1-STEP	A1 - MA, EXP SMTH, ARIMA / Specialist	29,6	30,7	38,1
1	1-STEP	A2 - MA, EXP SMTH, ARIMA / GA	29,6	28,0	52,4
1	1-STEP	B1 - NN, EXP SMTH, ARIMA / Specialist	29,6	26,1	47,6
1	1-STEP	B2 - NN, EXP SMTH, ARIMA / GA	29,6	24,3	52,4
2	1-STEP	A1 - MA, EXP SMTH, ARIMA / Specialist	26,6	36,5	33,3
2	1-STEP	A2 - MA, EXP SMTH, ARIMA / GA	26,6	32,3	42,9
2	1-STEP	B1 - NN, EXP SMTH, ARIMA / Specialist	26,6	34,5	42,9
2	1-STEP	B2 - NN, EXP SMTH, ARIMA / GA	26,6	31,9	38,1
1	3-STEP	A1 - MA, EXP SMTH, ARIMA / Specialist	38,7	32,0	66,7
1	3-STEP	A2 - MA, EXP SMTH, ARIMA / GA	38,7	31,8	71,4
1	3-STEP	B1 - NN, EXP SMTH, ARIMA / Specialist	38,7	29,0	57,1
1	3-STEP	B2 - NN, EXP SMTH, ARIMA / GA	38,7	28,6	57,1
2	3-STEP	A1 - MA, EXP SMTH, ARIMA / Specialist	50,1	43,6	71,4
2	3-STEP	A2 - MA, EXP SMTH, ARIMA / GA	50,1	36,5	76,2
2	3-STEP	B1 - NN, EXP SMTH, ARIMA / Specialist	50,1	44,8	61,9
2	3-STEP	B2 - NN, EXP SMTH, ARIMA / GA	50,1	45,5	57,1

Source: From the Author

4.2.5. Regressions applied in the results analysis

To have a broader understanding on the relationship between the error results and other variables, two kinds of regressions were run, the Tobit and the Probit. The Tobit model is used to analyze relationships between a non-negative dependent variable and the independent ones. The model proposed fits this situation, since MAPEs and PBs do not accept negative values, being both left censored. Due to this fit, the regression model was set to this data analysis, enabling to reduce the regression bias.

The other model is the Probit. In this type of regression the dependent variable can take only two values in a binary profile, 0 or 1, or no/yes. The application in this study fits when analyzing which model (company vs proposed) is better than the other. For example, if the company's MAPE is better than the proposed one it is considered 1, consequently, the opposite situation is 0. This model is very useful to this study since it enables the investigation of one variable impacting specifically in the dependent variable, positively or negatively.

Both regressions were run in R Studio, where the Tobit model was found in the function *tobit* in the package *AER*, and the Probit model was run through the function *glm,* by changing the *family* character.

4.3. Results Discussion

MAPE and PB analysis might show similar results but are different approaches to the error investigation. While MAPE is a measure that involves each individual percentage error in the analysis, the percentage better approach does not focus on the error's size, but just computes which forecast was better, decreasing the impact of large errors in the measure. Firstly, it will be disclosed the investigation on the MAPE behavior, and then the PB's. For each analysis it was taken two approaches: one approach is based on the main statistics of the series and will be called series-based analysis. The other approach is based on the series of returns of the original demand, and will be called returns-based analysis. The idea of this second analysis is to further investigate the impact of the data volatility and the return behavior on the FIS accuracy.

4.3.1. MAPE Analysis

4.3.1.1. Series-based analysis

To understand the results' behaviors for each model it's necessary to investigate its distributions and dispersions in accordance to different variables.

By the table 13 in Appendix B, partially shown previously, it's possible to see that the MAPEs of the fuzzy distribution are, in general, higher than the benchmark in one-month (one-step ahead) forecast horizons and lower than the benchmark in three-month (three-step ahead) forecast horizons. Further in this analysis the data will be disclosed to understand the impact of each decision on the final forecast. This decision comprises if it was chosen between the neural network method or the moving average in the forecasting input or if it was chosen between genetic algorithm optimization or the specialist decision for the fuzzy inference system (FIS) parameters.

The graph below shows the boxplot of the distribution of MAPEs, dividing it in between the neural network and moving average, comparing them to the benchmark's (company) MAPE.

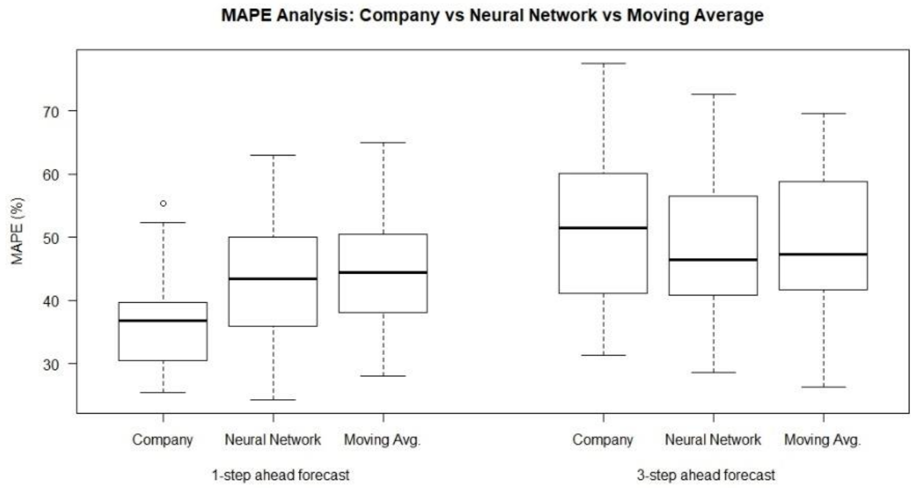

Figure 16. MAPE analysis: Company vs Neural Network vs Moving Average
Source: Company and the author

This graph shows that the median of the FIS that uses neural network is slightly lower than the moving average's median in both scenarios. The graph also illustrates that, in one-month horizon forecast, both distributions are considerably symmetric, with the data concentrated in the middle of the distribution. This is slightly different in the three-month forecast, where each method concentrates the data in different parts of the distribution. It is also important to notice that the benchmark forecast has a clear better performance in one-month forecast, but a slightly worse performance in the three-month forecast, and the dispersion of the data varies roughly in between the time horizons.

The graph in the figure 16 below shows the boxplot now dividing the information among the FIS that optimizes its parameters through genetic algorithm model or the FIS that is based on specialist parameters, compared to the same benchmark information. The graphs are like the previous boxplot, with few slight differences. The biggest difference is in the specialist-based FIS for a three-month period, that shows a larger dispersion than the average

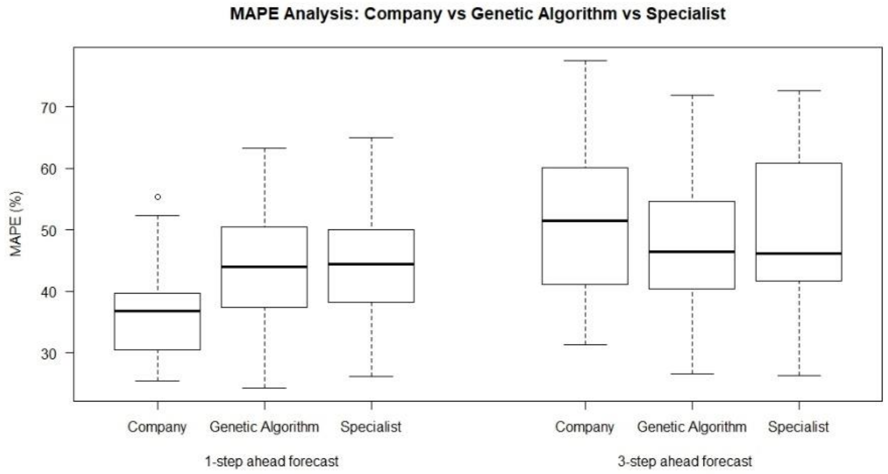

Figure 17. MAPE analysis: Company vs Neural Network vs Moving Average
Source: Company and the author

Another way of to analyze the data is looking at the probability density curves. In the graph below, the one-month forecasts are represented by the continuous lines, while the dashed lines represent the three-month ahead forecasts. The graph on the left compares the density profile of the benchmark and the neural network FIS. The graph on the right compares the remaining FIS models. It is visible that the FIS models have similar patterns and they differ basically when changing the time horizon. It is important to notice in that the MAPEs of the three-month period forecasts are more disperse than the one-step ahead forecast.

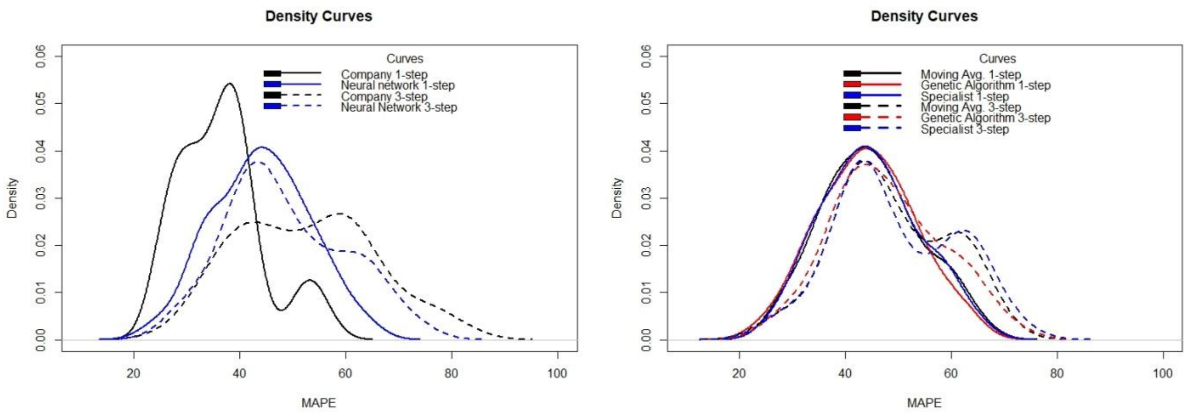

Figure 18. MAPE Density curves
Source: company and the author

Going further on the analysis, to better understand the variables that might impact on the MAPEs, it was set a comparison with the main characteristics of the series. The group of graphs below represents, on the left, the MAPEs of one-step ahead forecasts, and on the right, the MAPEs of the three-step ahead forecasts. The first line of graphs sets the relationship between the MAPEs and the coefficient of variation of the products' demand behavior. In the second line is found the comparison among MAPEs and the demand series' kurtosis, and, in the third line the comparison between MAPEs and the demand series' skewness.

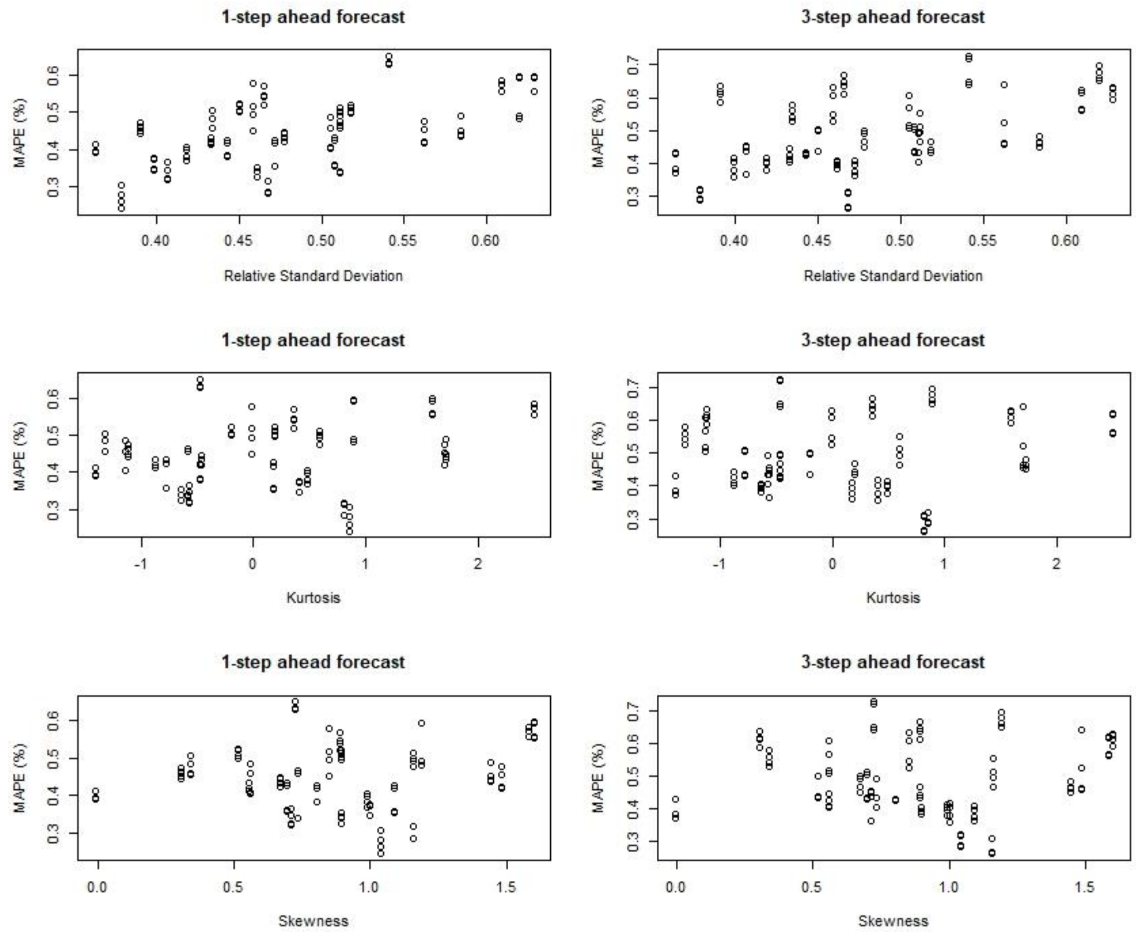

Figure 19. MAPE vs series' statistics
Source: Company and the author

The data is very dispersed, however, it is possible to see a slightly correlation between the MAPEs and the coefficient of variation (CV, or relative standard deviation), and also between the series kurtosis, what indicates that the error generated by the FIS might be correlated to those variables.

To better understand the impact of these variables in the MAPE of the forecasts it was tested a Tobit regression and a Probit regression. As previously informed, this kind of regression is fit in the situation since the MAPE can only be equal or greater than zero. In addition to the previous variables CV, skewness and kurtosis, it was added three dummy variables to test three things: (1) the impact of the decision between the specialist and the genetic algorithm for the parameters (2) the impact of the decision between the moving average and neural network for the input forecast and; (3) the impact of the forecast horizon choice.

The regression was run through the function *tobit* in the package *AER* in R, and it has the following profile, with MAPE as the dependent variable:

$$MAPE = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon \quad (11)$$

Where:

X_1 is the coefficient of variation of the observed demand data;

X_2 is the kurtosis of the observed demand data;

X_3 is the skewness of the observed demand data;

X_4 is the dummy variable of choosing specialists (1) instead the genetic algorithm (0) to define FIS parameters;

X_5 is the dummy variable of choosing moving average (1) instead of neural network (0);

X_6 is the dummy variable of running the three-month period (1) instead of the one-month period forecast (0);

β_n are the respective regressions' coefficients

ε is the error

The table 4 below shows the regression results

Table 5. MAPE Tobit regression

	Coefficient	Std. Error	Z-Value	Pr(>|z|)	
Intercept	0,050	0,045	1,113	0,000	***
Coefficient of Variation (CV)	1,304	0,094	13,931	0,000	***
Kurtosis	0,056	0,014	3,983	0,000	***
Skewness	-0,279	0,040	-7,010	0,000	***
Dummy Specialist/Genetic Alg.	0,011	0,010	1,131	0,258	
Dummy Moving Avg./Neural Net.	0,005	0,010	0,555	0,579	
Dummy 3-step/1-step	0,042	0,010	4,383	0,000	***
Log likelihood	265,3	8 D.F.			
Wald-statistic	225,1	6 D.F.	p-value:	0,000	
Signf. Level ***>99,9%	**>99%	*>95%	.>90%		

Source: From the author

By analyzing the coefficients it's possible to see that the decision between the specialist or the genetic algorithm and the decision between the moving average or the neural network do not have significant statistic impact on the MAPE. All the other variables have impact at more than 99,9% significant level. CV and kurtosis have positive impact on the dependent variable, on the other hand, skewness has a negative impact. It is important to notice that CV has a very large impact on the dependent variable, in relation to the other variables. In addition to that, it is also shown that a three-period horizon forecast impacts positively on the MAPE, what denotes that longer-term forecasts have a worse accuracy.

To a further analysis and to have more practical point of view it was also run a Probit regression. As previously clarified, in the Probit model the dependent variable can take only two values, in this case, 0 or 1. 1 in this study means when the MAPE is favorable for the company, and 0 means when the MAPE is favorable for the FIS. The Probit model was run through the function *glm* in the package *stats* in R. The right side of the Probit regression's formula is the same as the Tobit regression run previously, while the left side is a series where 1 is when the company's forecasts MAPE is better than the FIS and 0 is the opposite. The results of the regression are disclosed below in the table 5.

Table 6. MAPE Probit regression

	Coefficient	Std. Error	Z-Value	Pr(>|z|)	
Intercept	3,286	0,945	3,478	0,000	***
Coefficient of Variation (CV)	-1,843	1,933	-0,954	0,340	
Kurtosis	1,147	0,301	3,815	0,000	***
Skewness	-1,950	0,826	-2,362	0,018	*
Dummy Specialist/Genetic Alg.	0,174	0,196	0,888	0,375	
Dummy Moving Avg./Neural Net.	0,364	0,197	1,853	0,064	.
Dummy 3-step/1-step	-1,413	0,204	-6,936	0,000	***
Null Deviance	291,22	215 D.F.			
Residual Deviance	215,63	209 D.F.			
AIC	229,63				
Signf. Level ***>99,9%	**>99%	*>95%	.>90%		

Source: From the author

This regression exposes the variables that influence on which forecast method will be better using MAPE as reference: the company's forecasting system or the FIS. The results show that the CV, and, choosing between GA or specialist, do not have a statistic significant coefficient, meaning that it is not possible to tell if those variables impact in which forecast system wins. However, it is possible to conclude, with some statistical significance, that positive skewness influences on choosing FIS instead of the company's forecast, and, that choosing a moving average-based FIS disfavors the FIS forecast. The most relevant conclusions are that the kurtosis influences positively in choosing the company's forecast, meaning that series with long tails, consequently more extreme values, favors the company’s forecast. It is also important to notice that three-step ahead forecasts disfavor the company’s forecast system, where the FIS has an advantage.

The table 6 summarizes the variables that favors and disfavors the FIS forecast in relation to the company’s, based on the Probit regression.

Table 7. Summary of FIS favorable characteristics in MAPE/series dimension

Series' Characteristics		
Kurtosis	Disfavors	99,9% significance level
Skewness	favors	95% significance level
FIS Options		
3-step ahead	favors	99,9% significance level
Moving Avg.	disfavors	90% significance level

Source: From the author

4.3.1.2. Return-based analysis

It was identified that most of product demand time series have high volatility, with high variation of data. External information constantly influences the data behavior in a deep way, causing variation of more than 300% in the demand, for instance the policies around the IPI (Brazilian tax for industrialized products), or massive investments in marketing. In these situations, the company’s forecast can diminish the impact of these external influences, since this kind of information is discussed in the forecast meeting. However, the FIS statistical tool is penalized by these events. For illustration, the figure below shows the return behavior of the products 1 and 2.

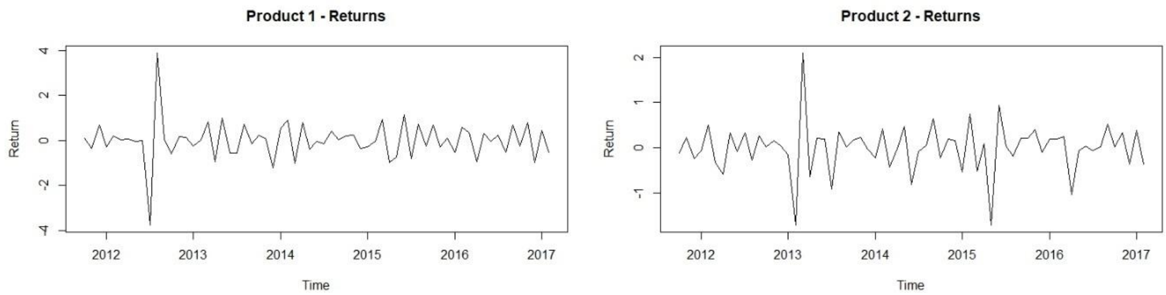

Figure 20. Returns of products 1 and 2
Source: the author

It is possible to see constant variations of more than 50% in the historic, and a considerable number of peaks of more than 100%. The Probit regression allow a deeper analysis on which returns characteristics impact directly in benefit in one of the models: the company’s or the FIS. The difference in this model is only the right side of the equation, where the series characteristics were changed by the returns characteristics

Table 8. MAPE and returns Probit regression

	Coefficient	Std. Error	Z-Value	Pr(>|z|)	
Intercept	0,4732	0,5517	0,8580	0,3911	
Return's Standard Deviation	-0,3541	0,8855	-0,4000	0,6892	
Return's Kurtosis	-0,5836	0,1801	-3,2400	0,0012	**
Return's Skewness	-1,3144	0,4059	-3,2390	0,0012	**
Dummy Specialist/Genetic Alg.	0,1596	0,1915	0,8330	0,4047	
Dummy Moving Avg./Neural Net.	0,3512	0,1923	1,8270	0,0677	.
Dummy 3-step/1-step	-1,3501	0,1968	6,8600	0,0000	***
Null Deviance	291,2200	215 D.F.			
Residual Deviance	227,43	209 D.F.			
AIC	241,43				
Signf. Level ***>99,9%	**>99%	*>95%	.>90%		

Source: From the author

Again, the company's forecast is penalized if it is forecasted a three-month period instead of the one-month forecast. This denotes that the FIS is better for the longer-term horizon. It is also possible to conclude with a considerable confidence level (>99%) that the returns' kurtosis has impact on which method will be better. In this case, it means that extreme values, what means the presence of very high or very low returns will actually increase the probability of the FIS model having a better result in the MAPE dimension. Negative skewness might also influence on the result, however, it is important to notice that it might be explained by the returns behavior, where positive and negative returns usually have different profiles.

Below is the summary of the conditions that impacts in the models' accuracy.

Table 9. Summary of FIS favorable characteristics in MAPE/series dimension

Returns' Characteristics		
Kurtosis	favors	99% significance level
Skewness	favors	99% significance level
FIS Options		
3-step ahead	favors	99,9% significance level
Moving Avg.	disfavors	90% significance level

Source: From the author

4.3.2. Percentage Better (PB) Analysis

4.3.2.1. Series-based analysis

The graph in the figure 19 below shows the impacts of each decision inside the FIS in the final results. First it is disclosed the results of the forecasting method's decision: moving average versus neural network, and then, choosing between the specialist or the GA to define FIS parameters. It is important to notice that while in MAPE the lowest value is pursued to a better result, in the PB comparison, the higher values show that the FIS was better than the benchmark a higher number of times.

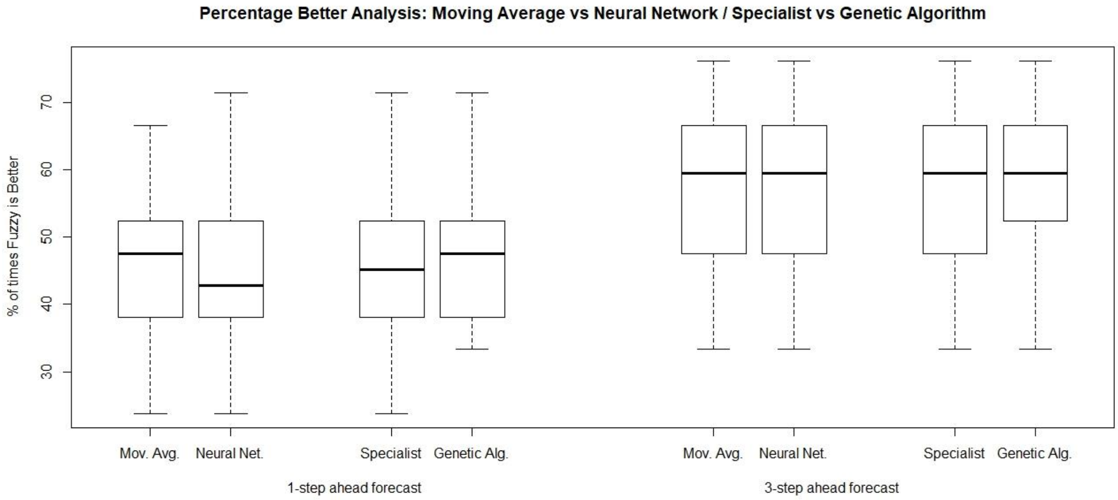

Figure 21. Percentage better analysis
Source: From the author

This graph shows on the right that for the three-step period forecasts, there is no expressive difference on which decision is made to configure the FIS, and, in all the situations, all the boxes are located above ~47%, showing that the data have a concentration in high PB positions, meaning that in three-month forecasts FIS is dominant. However, for the one-step forecasts, on the left, the medians have different positions according to which method is being used. It is important to notice that the GA method has the distribution boundaries much higher than the other methods. However, in all the cases the boxes are located in between ~38% and ~53%, meaning that the company forecast is dominant for this time horizon.

To better understand the variables that might impact on the PB numbers, as also done in the MAPE analysis, it was set a comparison with the main characteristics of the series. The group of graphs below represents, on the left, the PB of one-step ahead forecasts, and on the right, the PB of the three-step ahead forecasts. The first line of graphs sets the relationship between the PBs and the coefficient of variation of the products' demand behavior. In the second line of graphs is found the comparison among PBs and the demand series' kurtosis, and, in the third line the comparison between PBs and the demand series' skewness.

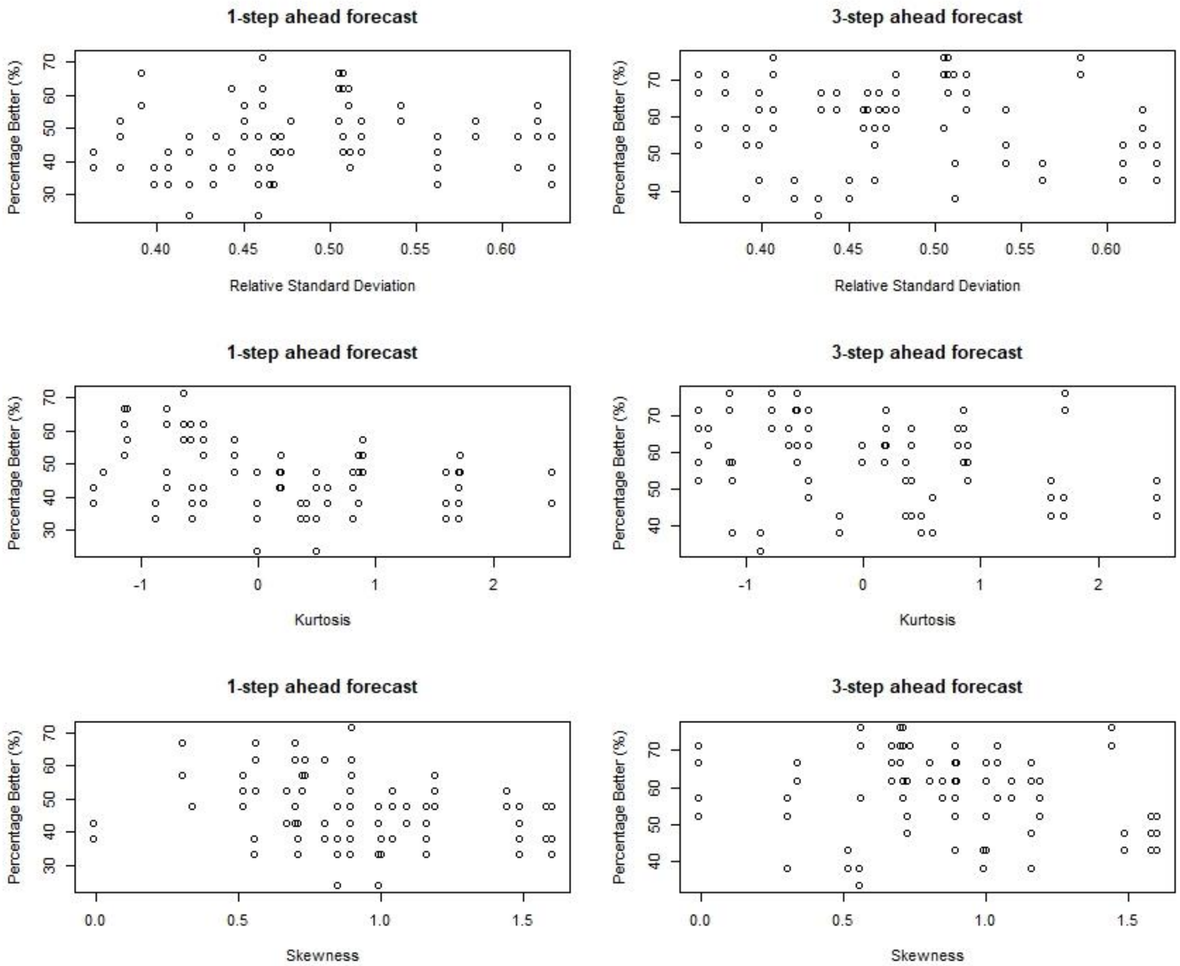

Figure 22. Percentage better vs series' statistics
Source: From the author

The graphs also show dispersed data, however they have a different profile than the MAPEs relationships. It is possible to infer a negative correlation between the series kurtosis and the PB, at least for the one-step ahead forecast. On the other side, it is hard to infer correlations with other variables, where they do not show strong correlation at first sight.

Looking at the probability density curves it is possible to see that the data for each forecasting period have similar profiles in between the time horizon groups. However, they differ a lot in when both groups are compared. The one-step ahead forecasts are the continuous lines, while the three-step ahead are represented by the dashed lines. As seen in the boxplots, the one-step ahead forecasts have a lower percentage better, meaning that the FIS has a worse performance in relation to the company in this time horizon. For the one-step ahead, the density curves profiles have an abrupt left side while it is more smooth transition in the right side. On the other hand, the three-period forecast has its probability more disperse, with the probabilities diluted in between 40% and 75%, with its peak around 65%, showing that the FIS model is dominant in this time horizon.

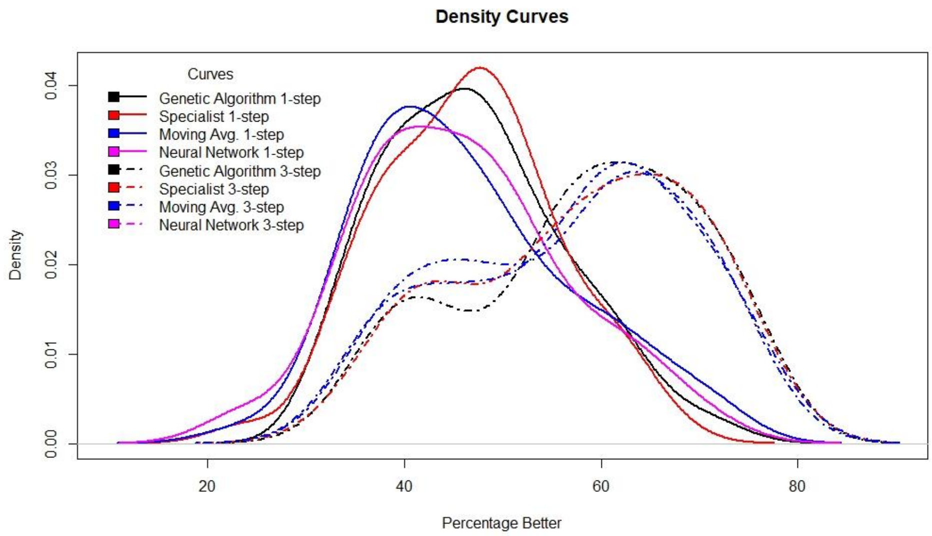

Figure 23. PB density curves
Source: From the author

To a further analysis on the percentage better behavior in comparison to the variables, it was used the same approach as the MAPEs analysis. The Tobit regression helps on understanding the relationships, according to the following function, which has the same profile to the one applied on MAPE:

$$PB = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon \quad (12)$$

Where:

X_1 is the coefficient of variation of the observed demand data;

X_2 is the kurtosis of the observed demand data;

X_3 is the skewness of the observed demand data;

X_4 is the dummy variable of choosing specialists (1) instead the genetic algorithm (0) to define FIS parameters;

X_5 is the dummy variable of choosing moving average (1) instead of neural network (0);

X_6 is the dummy variable of running the three-month period (1) instead of the one-month period forecast (0);

β_n are the respective regressions' coefficients;

ε is the error.

The table 7 below shows the regression results.

Table 10. PB Tobit regression

	Coefficient	Std. Error	Z-Value	Pr(>|z|)	
Intercept	0,1776	0,0624	2,8440	0,0045	**
CV	0,3996	0,0130	3,0670	0,0022	**
Kurtosis	-0,0920	0,0195	-4,7160	0,0000	***
Skewness	0,1209	0,0554	2,1820	0,0291	*
Dummy Specialist/Genetic Alg.	-0,0137	0,0134	-1,0180	0,3085	
Dummy Moving Avg./Neural Net.	0,0040	0,0134	0,2960	0,7675	
Dummy 3-step/1-step	0,1098	0,0134	8,1800	0,0000	***
Log likelihood	193,8	8 D.F.			
Wald-statistic	112,5	6 D.F.	p-value:	0,000	
Signf. Level ***>99,9%	**>99%	*>95%	.>90%		

Source: From the author

By analyzing the coefficients it's possible to see that the decision between the specialist or the genetic algorithm and the decision between the moving average or the neural network do not have significant statistic impact on the dependent variable. CV and skewness have positive impact on the PB result, with a considerable significance level. Kurtosis has a negative impact on the dependent variable, reinforcing the MAPE analysis. In addition to that, it is also shown that a three-period horizon forecast impacts positively on the PB result, giving better results for the FIS model.

The table 8 summarizes the variables that favors and disfavors the FIS forecast in relation to the company's, based on the Tobit regression. The Probit regression is not necessary since PB is already a comparative measure.

Table 11. Summary of FIS favorable characteristics in PB/series dimension

Series' Characteristics		
Kurtosis	Disfavors	99,9% significance level
CV	favors	99% significance level
Skewness	favors	95% significance level
FIS Options		
3-step ahead	favors	99,9% significance level

Source: From the author

4.3.2.2. Returns-based analysis

Below is disclosed the results of the regression based on the relationship of the PB and the returns characteristics. This time it was developed a Probit regression, as already seen in the MAPE versus returns analysis in the 4.3.1.2 item, instead of a Tobit regression with PB values as shown in the previous item. For the Probit regression, the right side of the equation is similar to the previous ones, but with the returns characteristics. However, the left side is composed by “0” or “1”. “0” was chosen to be the cases when the PB is below 50%, meaning that the company would have a better performance. “1” was chosen to be the cases when the PB is higher than 50%, meaning that FIS has the advantage. The table below shows the regression results.

Table 12. PB and returns Probit regression

	Coefficient	Std. Error	Z-Value	Pr(>|z|)	
Intercept	0,1608	0,5290	0,3040	0,7611	
Return’s Standard Deviation	-0,2082	0,8462	-0,2460	0,8057	
Return’s Kurtosis	-0,5251	0,1706	-3,0780	0,0021	**
Return’s Skewness	-1,1555	0,3848	-3,0030	0,0027	**
Dummy Specialist/Genetic Alg.	0,1664	0,1825	0,9120	0,3619	
Dummy Moving Avg./Neural Net.	-0,1148	0,1824	-0,6300	0,5290	
Dummy 3-step/1-step	-1,0694	0,1837	-5,8200	0,0000	***
Null Deviance	299,14	215 D.F.			
Residual Deviance	254,34	209 D.F.			
AIC	268,34				
Signf. Level ***>99,9%	**>99%	*>95%	.>90%		

Source: From the author

As also seen in the MAPE analysis, the FIS forecast is benefited if it is forecasted a three-month period instead of the one-month forecast. This denotes that the FIS is better for the longer-term horizon. It is also possible to conclude with a considerable confidence level (>99%) that the returns’ kurtosis has impact on which method will be better. In this case, it means that extreme values, what means the presence of very high or very low returns will actually increase the probability of the FIS model having a better result in the MAPE dimension. Negative skewness might also influence on the result, however, it is important to notice that it might be explained by the returns behavior, where positive and negative returns usually have different profiles. Below is the summary of the regression results, interpreted in the FIS point of view.

Table 13. Summary of FIS favorable characteristics in PB/returns dimension

Series' Characteristics		
Kurtosis	Favors	99% significance level

Skewness	Favors	99% significance level
FIS Options		
3-step ahead	favors	99,9% significance level

Source: From the author

5. CONCLUSION

As shown in the introduction, this study aims to revise the mechanism developed by Vroman, Happiette and Rabenasolo (1998), where a fuzzy inference system is used to give weights to determined variables, and apply to a forecast combiner, based in the model presented by Yesil, Kaya and Siradag (2012). The final objective is to explore the tools proposed with different forecast methods, including the classic statistic methods and the soft computing tools.

To compare all the models, the characteristics of the demand series and the series of returns (of the demands) were used as independent variables to verify possible impacts on the dependent variables, defined to be the MAPEs and PBs, in order to understand the behavior of the forecasts' accuracy.

First of all, the strongest result that is present in all regressions is that for the one-month forecast the benchmark is more likely to be better, and for three-month forecasts the FIS models are more like to exceed the benchmark's, what can configure a cost reduction opportunity for the company and less energy spent in the forecasting meetings. The one-step ahead forecast might be more fit for the qualitative forecast due to the influence of information that the FISs are not impacted. For example, participants of the team can control investments in marketing, and they have an idea of the impact of the investment in the final demand. The group of people is also aware of macro-economic variables, for instance governments incentives and when they might happen, so it is possible to improve their forecast to better meet the market needs. On the other side, for the longer-term, this qualitative influence might harm the accuracy, by introducing personal biases for the longer-term decisions.

Another important conclusion is that the four models shows very similar results among them. No significant difference was found among the models results, proving that choosing in between GA or specialists for the parameter definition is not determinant for the MAPE nor for the PB accuracy. On the other hand, NN-based models and MA-based models show slightly differences in some regressions. This might happen due to the fit of each model in each kind of data. It is important to investigate when the NN forecasts, that is a recognized superior model in most of times, are harmed due to overtraining or too much changing trends. These issues do not impact on MA-based forecasts, which are simply the average of the last 5 observations.

Looking at the analysis based on the series characteristics, the coefficient of variation (CV), the kurtosis and the skewness, they all impact on the MAPE. The Probit regression was run, to better understand if those characteristics favors one model over another, and it was possible to understand that CV does not have statistical proof on determining which model would have the better MAPE. However, with 99% of confidence level, it was shown that this variable impact positively on the PB. An explanation, is that this variable can impact differently in each observations' percentage error, but keeping overall similar MAPES, while the number of times one model is better than another has changed.

It was learned that series with high kurtosis and negative skewness disfavors the FIS forecasts in both MAPE and PB analysis, being more fit to the benchmark model. High kurtosis, meaning longer distribution tails, is explained by the existence of unusual very high of very low numbers. Considering that most of unusual and impactful events has results that can be predicted by the group of experts, for instance the end of a contract with a big retailer or the raise of minimum income, this analysis proof that external events like these can influence positively the company's qualitative forecast model, while the FISs are based only on the historic data.

According to the regressions, skewness also influences on the choice, however it can be naïve to jump into conclusions with this variable. The regressions showed that positive skewed series influence positively on choosing FIS over the company's forecast, however it can be just a consequence of the sales behavior in situations of extreme demand values, what means that unusual extremely low values are more extreme than the positive ones.

Understanding the different methodologies of forecasting and the huge impact of the qualitative information in some specific point of the series, it might be necessary to have a different approach to analyze the data and investigate the behavior of those extreme values. Again, marketing policies, tax incentives, or other odd events are almost impossible to predict with quantitative methods and it is understood, that they might impact negatively in the FIS results. A possible approach to analyze and measure these odd events can be the return analysis.

Looking at this return-based analysis, the outcomes were different than expected. For both MAPE and PB, the Probit regressions did not show any statistically significant correlation between the standard deviation of the returns, in other words, between its volatility and its different impact on the models. Against the common sense, the regressions showed that actually, for both PB and MAPE, the higher the kurtosis the better the FIS results in relation to the company's forecast.

The returns analysis might be harmed by possible correlation with the errors, but can be valuable tools to understand the impact of specific unusual volatilities, that are usually configured as being qualitatively predictable events in the overall forecast accuracy.

To summarize, these results suggest that the company's forecast system can be improved and optimized. Qualitative decisions are very important to interpret external inputs and adapt the forecast for specific situations of odd events. However, make use of it all the time can be not only costly and time consuming, but also drive to less accurate results than other options. Consequently, according to the results, the company could show better results by changing its forecast model for the FIS in the three-step ahead horizon, where the FIS overcomes the current model. Furthermore, for one-step ahead forecasts, the company should adopt a different hybrid system that demands human decision only when odd, and impactful, events are expected.

5.1. Limitations

This research is mainly limited by the number of products and the size of the forecasted period. More products and longer forecasted series are needed to reinforce the data analysis and have more robust conclusions. An interesting solution could be the Monte Carlo Simulation, however, unfortunately, the GA optimization currently takes long machine time to run (almost 20 minutes for each 21 periods forecast), what makes unfeasible to run the Monte Carlo simulations with the computer used, considering a 27 products range.

Another important limitation is to fit the unexpected behavior of the series in a specific and handy characteristic. This study proposed the returns to include these odd events in the study, however, opposite results might show an opportunity to define a more fit characteristic. This can lead to a better definition on when to use the company's forecast or a proposed FIS.

This study concentrates on characteristics of the series to determine the models' fit. More variables could be added to the regressions to better understand the situations to apply each model. For instance, economic factors or marketing investments, that could also be added to the FIS weighting engine. Furthermore, this study focuses only in one product line, which has a very similar demand profile among its products, and show fast-changing trends and high volatile data. Another product lines with different characteristics could enhance the data analysis, by probably showing different results patterns for each kind of profile.

5.2. Recommendations

It is important more researches in the forecasting field that focus to match different data behavior to the specific models. For instance, applying the models found in this study in another product lines with different behavior, from more steady profiles to even more volatile, would help to have better understanding on the impacts of choosing this type of FIS over other methods.

This study did not show a considerable difference among the fuzzy models' results. It is recommended to develop more variation types and test them into a robust data, to better investigate the FISs behaviors. For instance, a very important characteristic of the fuzzy model, the if-them rules, could have different values or even be optimized along time. In addition to that, the SARIMA models' sensitivity to its parameters can also drive to specific studies that go deeper on their impacts.

Furthermore, another opportunity is to develop models that are based on the return behavior, that could be more fit depending on the returns profiles. In addition to that it is also necessary more studies focusing on the return analysis. For example, analysis considering regressions with the returns of each period against the absolute percentage error for each period, to go deeper on the return investigation.

Aside to the return analysis, based on the similarity of the models on both dimensions analyzed in this study (MAPE and PB), a more specific quartile analysis can be developed in order to better understand the behavior of the data in the different parts of the distribution.

External variables are from extreme importance and need to be further investigated. More quantitative studies that test the main impacts on the product's demands would help to understand the data profile from a broader perspective than only the historic data.

The FISs proposed are mainly based on the historic data. Another suggestion is to develop more advanced models that can consist on interactive or hybrid FIS, consisting in a mix of expert and statistic forecast. An interactive could overcome a barrier between the qualitative and the quantitative analysis by default functioning as a forecasting FIS and, sporadically, demanding decisions from experts when necessary, in specific situations when its accuracy could be questioned.

APPENDIX A – SPECIALIST'S SURVEY

RESEARCH IN FORECAST ERRORS IN BEAUTY PRODUCTS SALES

This questionnaire's objective is to classify the forecast errors in the beauty product industry's universe.

Please insert your email: ___________________________________

1. **In your opinion, for the ONE-STEP AHEAD FORECAST, the Absolute Percentage Error (APE) is considered LOW when it is LOWER than:**

5%

10%

15%

20%

25%

Other value:______

2. **In your opinion, for the ONE-STEP AHEAD FORECAST, the Absolute Percentage Error (APE) is considered HIGH when it is HIGHER than:**

15%

20%

25%

30%

35%

Other value:_____

In your opinion, for the THREE-STEP AHEAD FORECAST, the same references are kept? In case not, specify the new references using LOWER THAN for LOW error and HIGHER THAN, for a HIGH error

Yes, the references are kept.

No, the other parameters are: ____

APPENDIX B – MAPE and PB results

Table 14. Error Results

Product	Forecast type	Forecast horizon	MAPE Company (%)	MAPE Fuzzy (%)	Fuzzy / Company Perc. Better (%)
1	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	29,6	30,7	38,1
1	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	29,6	28,0	52,4
1	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	29,6	26,1	47,6
1	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	29,6	24,3	52,4
2	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	26,6	36,5	33,3
2	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	26,6	32,3	42,9
2	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	26,6	34,5	42,9
2	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	26,6	31,9	38,1
3	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	39,9	34,1	61,9
3	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	39,9	35,3	57,1
3	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	39,9	32,5	71,4
3	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	39,9	34,0	71,4
4	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	31,4	44,6	42,9
4	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	31,4	44,4	42,9
4	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	31,4	42,2	52,4
4	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	31,4	43,4	42,9
5	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	52,1	59,4	47,6
5	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	52,1	59,2	52,4
5	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	52,1	48,2	52,4
5	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	52,1	49,0	57,1
6	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	39,4	58,3	38,1
6	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	39,4	55,6	47,6
6	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	39,4	57,2	38,1
6	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	39,4	57,2	38,1
7	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	36,8	51,2	38,1
7	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	36,8	49,2	42,9
7	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	36,8	50,0	38,1
7	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	36,8	47,6	42,9
8	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	40,1	46,5	57,1
8	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	40,1	45,8	61,9
8	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	40,1	33,8	61,9
8	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	40,1	34,0	61,9
9	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	42,2	48,5	52,4
9	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	42,2	45,8	52,4
9	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	42,2	40,6	66,7
9	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	42,2	40,5	61,9
10	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	37,9	56,9	33,3
10	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	37,9	51,9	38,1
10	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	37,9	54,5	33,3
10	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	37,9	54,1	38,1
11	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	39,2	59,7	47,6

11	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	39,2	59,3	38,1
11	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	39,2	55,5	33,3
11	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	39,2	55,6	38,1
12	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	32,2	38,4	61,9
12	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	32,2	38,1	61,9
12	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	32,2	41,9	42,9
12	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	32,2	42,5	38,1
13	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	26,8	38,2	23,8
13	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	26,8	36,9	33,3
13	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	26,8	40,6	42,9
13	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	26,8	39,8	47,6
14	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	28,2	35,6	47,6
14	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	28,2	35,4	47,6
14	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	28,2	41,7	42,9
14	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	28,2	42,7	42,9
15	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	31,9	48,5	47,6
15	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	31,9	50,5	47,6
15	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	31,9	45,6	47,6
15	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	31,9	45,7	47,6
16	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	33,0	39,3	38,1
16	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	33,0	39,0	42,9
16	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	33,0	41,3	38,1
16	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	33,0	39,5	42,9
17	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	37,2	52,1	57,1
17	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	37,2	52,3	52,4
17	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	37,2	50,0	47,6
17	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	37,2	50,5	47,6
18	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	34,8	41,9	33,3
18	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	34,8	41,3	38,1
18	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	34,8	43,4	38,1
18	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	34,8	41,9	38,1
19	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	55,3	65,1	52,4
19	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	55,3	63,3	52,4
19	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	55,3	62,8	52,4
19	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	55,3	63,0	57,1
20	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	36,3	49,4	47,6
20	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	36,3	45,1	38,1
20	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	36,3	57,7	23,8
20	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	36,3	51,7	33,3
21	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	38,9	44,1	52,4
21	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	38,9	43,6	47,6
21	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	38,9	45,1	47,6
21	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	38,9	49,0	47,6
22	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	38,5	43,4	42,9
22	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	38,5	42,5	47,6
22	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	38,5	35,6	61,9
22	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	38,5	35,9	66,7

23	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	25,5	37,6	38,1
23	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	25,5	37,3	33,3
23	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	25,5	34,8	38,1
23	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	25,5	34,6	33,3
24	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	52,4	49,6	42,9
24	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	52,4	50,2	52,4
24	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	52,4	51,2	52,4
24	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	52,4	52,0	47,6
25	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	41,1	45,0	66,7
25	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	41,1	44,4	57,1
25	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	41,1	47,4	57,1
25	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	41,1	46,2	57,1
26	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	26,4	28,5	47,6
26	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	26,4	28,5	47,6
26	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	26,4	31,7	33,3
26	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	26,4	31,6	42,9
27	A1 - MA, EXP SMTH, ARIMA / Specialist	1-step ahead	29,3	41,9	42,9
27	A2 - MA, EXP SMTH, ARIMA / GA	1-step ahead	29,3	42,1	47,6
27	B1 - NN, EXP SMTH, ARIMA / Specialist	1-step ahead	29,3	45,4	38,1
27	B2 - NN, EXP SMTH, ARIMA / GA	1-step ahead	29,3	47,6	33,3
Product	**Forecast type**	**Forecast horizon**	**MAPE Company (%)**	**MAPE Fuzzy (%)**	**Fuzzy / Company Perc. Better (%)**
1	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	38,7	32,0	66,7
1	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	38,7	31,8	71,4
1	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	38,7	29,0	57,1
1	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	38,7	28,6	57,1
2	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	50,1	43,6	71,4
2	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	50,1	36,5	76,2
2	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	50,1	44,8	61,9
2	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	50,1	45,5	57,1
3	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	42,9	40,6	61,9
3	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	42,9	38,3	61,9
3	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	42,9	39,4	66,7
3	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	42,9	40,3	61,9
4	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	59,6	49,1	71,4
4	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	59,6	49,8	71,4
4	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	59,6	46,7	61,9
4	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	59,6	44,8	66,7
5	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	77,5	69,7	57,1
5	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	77,5	67,8	57,1
5	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	77,5	66,1	52,4
5	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	77,5	65,0	61,9
6	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	46,7	61,6	42,9
6	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	46,7	62,2	42,9
6	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	46,7	56,1	47,6
6	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	46,7	56,5	52,4

7	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	46,6	55,3	38,1
7	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	46,6	51,2	38,1
7	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	46,6	49,4	47,6
7	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	46,6	46,6	47,6
8	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	63,4	49,3	71,4
8	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	63,4	49,1	71,4
8	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	63,4	40,5	71,4
8	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	63,4	43,3	71,4
9	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	70,4	60,8	57,1
9	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	70,4	50,7	76,2
9	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	70,4	56,8	71,4
9	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	70,4	51,7	76,2
10	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	60,2	64,7	52,4
10	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	60,2	63,5	52,4
10	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	60,2	66,8	42,9
10	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	60,2	61,3	57,1
11	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	54,7	62,9	47,6
11	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	54,7	59,2	52,4
11	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	54,7	62,6	42,9
11	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	54,7	61,1	47,6
12	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	51,5	43,0	66,7
12	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	51,5	43,0	66,7
12	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	51,5	42,4	61,9
12	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	51,5	42,7	66,7
13	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	34,1	40,3	38,1
13	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	34,1	37,9	42,9
13	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	34,1	41,4	38,1
13	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	34,1	39,7	38,1
14	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	39,5	37,5	61,9
14	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	39,5	39,6	57,1
14	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	39,5	40,8	61,9
14	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	39,5	36,2	57,1
15	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	60,0	57,9	61,9
15	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	60,0	56,0	66,7
15	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	60,0	52,7	66,7
15	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	60,0	54,2	61,9
16	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	48,7	42,9	52,4
16	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	48,7	43,1	57,1
16	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	48,7	37,1	71,4
16	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	48,7	38,4	66,7
17	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	40,7	43,7	42,9
17	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	40,7	43,5	42,9
17	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	40,7	49,9	38,1
17	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	40,7	50,2	38,1
18	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	38,6	44,5	33,3
18	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	38,6	40,3	38,1
18	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	38,6	42,6	38,1

18	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	38,6	41,0	33,3
19	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	76,6	64,1	61,9
19	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	76,6	64,9	61,9
19	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	76,6	72,7	47,6
19	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	76,6	71,9	52,4
20	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	57,8	60,8	57,1
20	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	57,8	52,7	61,9
20	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	57,8	63,2	57,1
20	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	57,8	54,7	57,1
21	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	59,6	45,1	76,2
21	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	59,6	48,2	71,4
21	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	59,6	46,5	71,4
21	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	59,6	46,2	71,4
22	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	61,8	51,1	66,7
22	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	61,8	50,5	71,4
22	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	61,8	43,1	76,2
22	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	61,8	43,5	76,2
23	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	38,8	41,7	42,9
23	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	38,8	40,4	52,4
23	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	38,8	37,9	61,9
23	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	38,8	35,8	66,7
24	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	63,0	43,4	66,7
24	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	63,0	43,4	66,7
24	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	63,0	44,2	71,4
24	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	63,0	46,7	61,9
25	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	59,3	61,8	52,4
25	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	59,3	58,8	57,1
25	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	59,3	63,5	38,1
25	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	59,3	61,0	38,1
26	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	31,3	26,2	61,9
26	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	31,3	26,6	66,7
26	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	31,3	30,7	61,9
26	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	31,3	31,1	61,9
27	A1 - MA, EXP SMTH, ARIMA / Specialist	3-step ahead	41,5	45,7	47,6
27	A2 - MA, EXP SMTH, ARIMA / GA	3-step ahead	41,5	46,2	42,9
27	B1 - NN, EXP SMTH, ARIMA / Specialist	3-step ahead	41,5	64,0	47,6
27	B2 - NN, EXP SMTH, ARIMA / GA	3-step ahead	41,5	52,3	42,9

Source: From the author

BIBLIOGRAPHY

ABIHPEC. **Caderno de Tendencias 2014/2015**ABIHPEC, , 2015.

ABIHPEC. **Panorama do Setor de Higiene Pessoal, Perfumaria e Cosméticos**ABIHPEC, , 2017.

ADYA, M.; COLLOPY, F. How effective are neural networks at forecasting and prediction? A review and evaluation. **Journal of Forecasting**, v. 17, n. 56, p. 481–495, 1998.

AL-ADWAN, I. et al. DESIGN OF AN ADAPTIVE FUZZY-BASED CONTROL SYSTEM USING GENETIC ALGORITHM OVER A pH TITRATION PROCESS. **International Journal of Recent Research and Applied Studies**, v. 17, n. November, p. 177–184, 2013.

ALCALÁ-FDEZ, J. et al. Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms. **Fuzzy Sets and Systems**, v. 160, n. 7, p. 905–921, 2009.

AMMERLAAN, J.; WRIGHT, D. Adaptive cooperative fuzzy logic controller. **ACSC '04 Proceedings of the 27th Australasian conference on Computer science - Volume 26**, p. 255–263, 2004.

BAKIRTZIS, A. G. et al. SHORT TERM LOAD FORECASTING USING FUZZY NEURAL NETWORKS. **IEEE Transactions on Power Systems**, v. 10, n. 3, p. 1518–1524, 1995.

BEASLEY, D.; BULL, D. R.; MARTIN, R. R. An overview of genetic algorithms : Part 1, fundamentals. **University Computing**, v. 2, n. 15, p. 56–69, 1993.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis**. Fourth Edi ed. Hoboken, New Jersey: JOHN WILEY & SONS, 2008.

CHEN, A. S.; LEUNG, M. T.; DAOUK, H. Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. **Computers and Operations Research**, v. 30, n. 6, p. 901–923, 2003.

CHEN, S.-M. Forecasting enrollments based on fuzzy time series. **Fuzzy Sets and Systems**, v. 81, n. 3, p. 311–319, 1996.

CHENG, C. H.; CHANG, J. R.; YEH, C. A. Entropy-based and trapezoid fuzzification-based fuzzy time series approaches for forecasting IT project cost. **Technological Forecasting and Social Change**, v. 73, n. 5, p. 524–542, 2006.

DARBELLAY, G. A.; SLAMA, M. Forecasting the short-term demand for electricity: Do neural networks stand a better chance? **International Journal of Forecasting**, v. 16, n. 1, p. 71–83, 2000.

DASE, R. K.; PAWAR, D. D. Application of Artificial Neural Network for stock market predictions: A review of literature. **International Journal of Machine Intelligence**, v. 2, n. 2, p. 14–17, 2010.

DE GOOIJER, J. G.; HYNDMAN, R. J. 25 Years of Time Series Forecasting. **International Journal of Forecasting**, v. 22, n. 3, p. 443–473, 2006.

DE JONG, K. A. **An Analysis of the Behavior of a Class of Genetic Adaptive Systems**. [s.l.] University of Michigan, 1975.

DE JONG, K. A. **Are Genetic Algorithms Functions Optimizers?** parallel Problems Solving from nature. **Anais**...North Holland, Elsevier, 1992

DICKEY, D. A.; FULLER, W. A. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. **Econometrica**, v. 49, n. 4, p. 1057–1072, 1981.

E. J. HANNAN; RISSANEN, J. Recursive Estimation of Mixed Autoregressive-Moving Average Order. **Biometrika**, v. 69, n. 1, p. 81–94, 1982.

FARAWAY, J.; CHATFIELD, C. Time series forecasting with neural networks: a comparative study using the air line data. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, v. 47, n. 2, p. 231–250, 1998.

FRANTTI, T.; MAHOMEN, P. Fuzzy logic-based forecasting model. **Engineering Applications of Artificial Intelligence**, v. 14, n. 2, p. 189–201, 2001.

GARDNER, E. S. Exponential smoothing: The state of the art-Part II. **International Journal of Forecasting**, v. 22, n. 4, p. 637–666, 2006.

GEERING, H. P. **Introduction to Fuzzy Control**. Zurich, Switzerland: [s.n.]. Disponível em: <http://scholar.google.co.uk/citations?view_op=view_citation&hl=en&user=vGEBO7E

AAAAJ&cstart=40&citation_for_view=vGEBO7EAAAAJ:TQgYirikUcIC>.

GEN, M.; TSUJIMURA, Y.; ZHENG, D. An application of fuzzy set theory to inventory control models. **Computers & Industrial Engineering**, v. 33, n. 3–4, p. 553–556, 1997.

GÓMEZ, V. Automatic Model Identification in the Presence of Missing Observations and Outliers. n. June 1994, p. 34, 1998.

HERRERA, F.; LOZANO, M.; VERDEGAY, J. L. Tuning fuzzy logic controllers by genetic algorithms. **International Journal of Approximate Reasoning**, v. 12, n. 3–4, p. 299–315, 2005.

HIPPERT, H. S.; PEDREIRA, C. E.; SOUZA, R. C. Neural networks for short-term load forecasting: a review and evaluation. **IEEE Transactions on Power Systems**, v. 16, n. 1, p. 44–55, 2001.

HO, K.; HSU, Y.; YANG, C.-C. SHORT TERM LOAD FORECASTING USING A MULTILAYER NEURAL NETWORK. **IEEE Transactions on Power Systems**, v. 7, n. 1, p. 141–149, 1992.

HOLMUKHE, R. M. et al. **Short Term Load Forecasting with Fuzzy Logic Systems for power system planning and reliability - A Review**. International Conference on Modeling, Optimization, and Computing. **Anais**...2010

HOLT, C. C. Forecasting seasonals and trends by exponentially weighted moving averages. **International Journal of Forecasting**, v. 20, n. 1, p. 5–10, 1957.

HSIEH, C. H. Optimization of Fuzzy Backorder Inventory Models. **Information Sciences**, v. 146, p. 29–40, 2002.

HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: The forecast package for R. **Journal Of Statistical Software**, v. 27, n. 3, p. C3–C3, 2008.

JACOBS, W.; ZANINI, R. R.; COSTA, M. Estudo comparativo de séries temporais para previsão de vendas de um produto. **Iberoamerican Journal of Industrial Engineering**, v. 6, n. 12, p. 112–133, 2015.

JACQUIN, A. P.; SHAMSELDIN, A. Y. Review of the application of fuzzy inference systems in river flow forecasting. **Journal of Hydroinformatics**, v. 11, n. 3–4, p. 202, 2009.

JANG, J.-S. R. **ANFIS: Adaptative Network Based Fuzzy Inference SIeee Transactions On Systems Man And Cybernetics**, 1993.

JARRETT, J. E.; PLOUFFE, J. S. The fuzzy logic method for simpler forecasting. **International Journal of Engineering Business Management**, v. 3, n. 3, p. 25–52, 2011.

KAASTRA, I.; BOYD, M. Designing a neural network for forecasting financial and economic time series. **Neurocomputing**, v. 10, n. 3, p. 215–236, 1996.

KARR, C. L. **Design of a cart-pole balancing fuzzy logic controller using a genetic algorithm**. SPIE 1468, Applications of Artificial Intelligence IX. **Anais**...The International Society for Optical Engineering, 1991

KARR, C. L. **Adaptive Process Control with Fuzzy Logic and Genetic Algorithms**. NASA, Washington, Technology 2002: The Third National Technology Transfer Conference and Exposition. **Anais**...Washington: NASA, 1993

KARR, C. L.; GENTRY, E. J. Fuzzy Control of pH Using Genetic Algorithms. **IEEE TRANSACTIONS ON FUZZY SYSTEMS**, v. I, n. I, p. 46–53, 1993.

KHASHEI, M.; BIJARI, M. An artificial neural network (p, d, q) model for timeseries forecasting. **Expert Systems with Applications**, v. 37, n. 1, p. 479–489, 2010.

KIM, K. et al. Implementation of Hybrid Short-term Load Forecasting System Using Artificial Neural Networks and Fuzzy Expert Systems. **IEEE Transactions on Power Systems**, v. 10, n. 3, p. 1534–1539, 1995.

KIMOTO, T. et al. Stock market prediction system with modular neural networks. **1990 IJCNN International Joint Conference on Neural Networks**, p. 1–6 vol.1, 1990.

KISSI, M. et al. Determination of fuzzy logic membership functions using genetic algorithms: application to olfaction. **Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (IEEE Cat. No.03EX795)**, v. 118, p. 6–9, 2003.

KO, M.; TIWARI, A.; MEHNEN, J. A review of soft computing applications in supply chain management. **Applied Soft Computing**, v. 10, n. 3, p. 661–674, 2010.

KWIATKOWSKI, D. et al. Testing the null hypothesis of stationarity against the alternative of a unit root. **Journal of Econometrics**, v. 54, n. 1–3, p. 159–178, 1992.

LACERDA, E. G. M.; CARVALHO, A. C. P. L. F. Introdução aos Algoritmos Genéticos. In: UFRGS (Ed.). . **Sistemas Inteligentes – Aplicações a Recursos Hídricos e Ciências Ambientais**. Porto Alegre: [s.n.]. p. 87–148.

LEE, M. H. et al. Fuzzy Time Series: An Application to Tourism Demand ForecastingNo Title. **American Journal of Applied Sciences**, v. 9, n. 1, p. 132–140, 2012.

LIU, B. et al. Design of Adaptive Fuzzy Logic Controller Based on Linguistic-Hedge Concepts and Genetic Algorithms. **IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS**, v. 31, n. 1, p. 32–53, 2001.

MAIERS, J.; SHERIF, Y. S. Applications of Fuzzy Set Theory. **IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS**, v. 15, n. 1, 1985.

MAKRIDAKIS, S. et al. The M2-competition: A real-time judgmentally based forecasting study. **International Journal of Forecasting**, v. 9, n. 1, p. 5–22, 1993.

MAKRIDAKIS, S.; HIBON, M. The M3-Competition: results, conclusions and implications. **International Journal of Forecasting**, v. 16, n. 4, p. 451–476, 2000.

MÉLARD, G.; PASTEELS, J.-M. Automatic ARIMA modeling including interventions, using time series expert software. **International Journal of Forecasting**, v. 16, n. 4, p. 497–508, 2000.

NOVÁK, V. Perception-based Logical Deduction as Alternative Approximate Reasoning Method. **Ieee International Conference On Fuzzy Systems**, p. 1032–1037, 2005.

ORD, J. K.; KOEHLER, A B.; SNYDER, R. D. Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models. **Journal of the American Statistical Association**, v. 92, n. 440, p. 1621–1629, 1997.

PEDRYCZ, W. Genetic algorithms for learning in fuzzy relational structures. **Fuzzy Sets and Systems**, v. 69, n. 1, p. 37–52, 1995.

PETROVIC, D.; SWEENEY, E. Fuzzy knowledge-based approach to treating uncertainty in inventory control. **Computer Integrated Manufacturing Systems**, v. 7, n. 3, p. 147–152, 1994.

POULSEN, J. R. **Fuzzy Time Series Forecasting - Developing a new forecasting model based on high order fuzzy time series**. [s.l.] Aalborg University Esbjerg, 2009.

ROSS, T. J. **Fuzzy Logic with Engieering Applications**. 2. ed. [s.l.] John Wiley & Sons, 2004.

SCRUCCA, L. GA : A Package for Genetic Algorithms in R. **Journal of Statistical Software**, v. 53, n. 4, 2013.

SEPTEM RIZA, L. et al. Learning from Data Using the R Package “ frbs ”. 2014.

SHIMOJIMA, K.; FUKUDA, T.; HASEGAWA, Y. Self-tuning fuzzy modeling with adaptive membership function, rules, and hierarchical structure based on genetic algorithm. **Fuzzy Sets and Systems**, v. 71, n. 3, p. 295–309, 1995.

SINGH, S. R. A robust method of forecasting based on fuzzy time series. **Applied Mathematics and Computation**, v. 188, n. 1, p. 472–484, 2007.

SINGHAL, S. et al. Application of Fuzzy Logic and Fuzzy Systems in Machining : A Literature Review. p. 1905–1911, 2016.

SLACK, N.; BRANDON-JONES, A.; JOHNSTON, R. **Operations Management**. 7th. ed. [s.l.] Pearson, 2013.

SONG, Q.; CHISSOM, B. S. Fuzzy time series and its models. **Fuzzy sets and systems**, v. 54, n. 88, p. 269–277, 1993a.

SONG, Q.; CHISSOM, B. S. Forecasting enrollments with fuzzy time series - Part I. **Fuzzy Sets and Systems**, v. 54, n. 1, p. 1–9, 1993b.

SRINIVASAN, D.; LEE, M. A. **Survey of Hybrid Fuzzy Neural Approaches to Electric Load Berkeley Initiative in Soft Computing**. 1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century. **Anais**...1995

SUGENO, M.; KANG, G. T. Structure Identification of Fuzzy Model. **Fuzzy Sets and Systems**, v. 28, p. 15–33, 1988.

SUHAIL, A.; KHAN, Z. A. Fuzzy control with limited control opportunities and response delay - A production-inventory control scenario. **International Journal of Approximate Reasoning**, v. 38, n. 1, p. 113–131, 2005.

TAKAGI, T.; SUGENO, M. Fuzzy Identification of Systems and Its Applications to Modeling and Control. **IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS**, v. 15, n. 1, p. 116–132, 1985.

TANG, K. S. et al. What are Genetic Algorithms ? Optimization Algorithms. **IEEE Signal Processing Magazine**, n. November, p. 22–37, 1996.

THRIFT, P. **Fuzzy Logic Synthesis with Genetic Algorithms**. International Conference on Genetic Algorithms. **Anais**...1991

TSENG, F. M.; TZENG, G. H. A fuzzy seasonal ARIMA model for forecasting. **Fuzzy Sets and Systems**, v. 126, n. 3, p. 367–376, 2002.

TÜRKŞEN, I. B. Review of fuzzy system models with an emphasis on fuzzy functions. **Transactions of the Institute of Measurement and Control**, v. 31, n. 1, p. 7–31, 2009.

UNION, W. A. Average-based fuzzy time series models for forecasting Shanghai compound. v. 4, n. 2, p. 104–111, 2008.

VROMAN, P.; HAPPIETTE, M.; RABENASOLO, B. Fuzzy Adaptation of the Holt–Winter Model for Textile Sales-forecasting. **Journal of the Textile Institute**, v. 89, n. 1, p. 78–89, 1998.

WANG, C.-C. A comparison study between fuzzy time series model and ARIMA model for forecasting Taiwan export. **Expert Systems with Applications**, v. 38, n. 8, p. 9296–9304, 2011.

WANG, J.-L. A supply chain application of fuzzy set theory to inventory control models – DRP system analysis. **Expert Systems with Applications**, v. 36, n. 5, p. 9229–9239, 2009.

WANKE, P. F. **Gerência de Operações: uma abortagem logística**. São Paulo:

Atlas, 2010.

YESIL, E.; KAYA, M.; SIRADAG, S. Fuzzy forecast combiner design for fast fashion demand forecasting. **INISTA 2012 - International Symposium on INnovations in Intelligent SysTems and Applications**, p. 0–4, 2012.

YULE, G. . U. On a Method of Investigating Periodicities in Disturbed Series , with Special Reference to Wolfer â€™ s Sunspot Numbers. **Philosophical Transactions of the Royal Society of London . Series A**, v. 226, n. 1927, p. 267–298, 1927.

ZADEH, L. A. Fuzzy sets. **Information and Control**, v. 8, n. 3, p. 338–353, 1965.

ZADEH, L. A. Is there a need for fuzzy logic? **Information Sciences**, v. 178, n. 13, p. 2751–2779, 2008.

ZHANG, G.; PATUWO, E.; HU, M. Y. Forecasting with artificial neural networks: The state of the art. **International Journal of Forecasting**, v. 14, p. 35–62, 1998.

ZHANG, P. G. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, v. 50, n. February, p. 159–175, 2003.

ZIMMERMANN, H. J. Fuzzy set theory. **Wiley Interdisciplinary Reviews: Computational Statistics**, v. 2, n. 3, p. 317–332, 2010.