A hybrid forecast model combining fuzzy time series, linear regression and a new smoothing technique

Fábio José Justo dos Santos[1,2] Heloisa de Arruda Camargo[1]

[1]Federal University of São Carlos, Brazil
[2] Federal Institute of Education, Science and Technology of São Paulo, Brazil

Abstract

In recent years, Fuzzy Time Series have been considered a promising tool to deal with forecasting problems due to the ease to model the problems, the satisfactory results obtained and also to the low computational cost required. However, the long experience with traditional methods coming from statistics, certainly brings a rich knowledge that can be used to enhance the computational methods employed to deal with Fuzzy Time Series. This paper introduces a forecast model where Fuzzy Time Series, linear regression and a new smoothing method are combined. Experiments were performed with the Taiwan Stock Exchange index and compared with eight others approaches found in the literature. The results confirm that the proposed model presents a good accuracy with relation to the other methods.

Keywords: fuzzy time series, linear regression, smoothing technique, forecasting.

1. Introduction

The good results obtained in forecast problems with Fuzzy Time Series (FTS) methods and the facility to build and to interpret its models, have been some of the reasons why the interest in this topic has increased recently. Since FTS was introduced in [1], several researches arose to deal with different scenarios of the real world. In [2] the authors introduced a methodology to work with a multilayer fuzzy inference system dedicated to estimate the power demand in a substation feeder. A model that combine concepts of support vector machines, evolutionary artificial neural networks and fuzzy rules was presented in [3], and validated using datasets of time series that contain the number of passengers in an airlines company, the number of car sales and the amount of milk production among others. The financial area is one of the most common areas of applications used in the experiments about forecast problems. In [4], a model that combines the concepts of particle swarm optimization, fuzzy time series and support vector machines to do forecasts in the Taiwan Stock Exchange (TAIEX) index datasets is presented. A methodology that combines concepts of Autoregressive Integrated Moving Average (ARIMA) model with the concepts of Artificial Neural Network (ANN) is presented in [5] and applied to the money exchange rate scenario.

It is possible to observe, from a literature review, that the integration of classic statistical models with the FTS concepts to address forecast problems has been little ex-plored. However, such integration could favor the development of several potential contributions. On one side, the advantages of traditional statistical forecasting methods are extensively pointed out. On the other side, some works have reported better accuracy by the FTS methods than traditional statistical methods, especially in time series with few samples or with irregular behavior [6, 7].

Towards this line of investigation, the authors of the present work have proposed recently an approach that combines FTS with simple exponential smoothing (SES) [8]. In this paper the method has been extended in several ways. The main contribution is the introduction of a new smoothing technique that is used with the Fuzzy Logical Relationships (FLR) extracted from fuzzy time series. The pre-processing method introduced in [9] is extended to include an additional step where the linguistic terms are defined over the edges of the time series domain. In addition, the introduced model aggregates to the forecasting the linear regression concepts that are applied when a suitable FLR to infer the forecasted value does not exist. The results obtained from experiments with TAIEX index were compared with eight others approaches found in the literature and confirm the good accuracy of proposed model.

The rest of this paper is organized as follows. Basic concepts of fuzzy time series and linear regression are presented in Section 2. The pre-processing model used and the new version of the forecasting model combining FTS with statistics concepts and including the new smoothing technique are introduced in Section 3. The experiments and conclusions are discussed, respectively, in Section 4 and Section 5.

2. Brief review of basic concepts

In this section, the basic concepts of FTS and linear regression used in the proposed model are presented.

2.1. Fuzzy time series

FTS were initially introduced in [1], as a forecasting model built from concepts of linguistic terms introduced by Zadeh's work [10]. According to [1], the basic definitions related to FTS are presented as following.

- Definition 1: Let $Y(t)$ $(t = 1, 2, 3, ...)$ be the universe of discourse, which is a subset of $\mathbb{R}$. Consider the fuzzy sets $f_i(t)$ $(i = 1, 2, 3, ...)$ defined in the universe of discourse $Y(t)$. If $F(t)$ is a collection of $f_i(t)$ $(i = 1, 2, 3, ...)$, then, $F(t)$ is called a fuzzy time series on $Y(t)$.

- Definition 2: Let $F(t)$ be caused by $F(t-1)$, $F(t-2), \ldots, F(t-n)$ this model is called nth-order fuzzy time series, and the fuzzy logical relationships between them is represented by $F(t-n)$, $\ldots, F(t-2), F(t-1) \rightarrow F(t)$.

The authors of the work in [1] also defined five basic steps that compose the forecast process by means of FTS. The first step is the definition of the universe of discourse. In the sequence, the universe of discourse is partitioned into several equal length intervals, on which the fuzzy sets will be defined. The third step is the fuzzification of the crisp observed values. In the sequence, the forecasting is computed in the form of a fuzzy set. Finally, the linguistic value is defuzzified.

2.2. Linear regression

The simple linear regression [11] represents the relation between one dependent variable Y and one independent variable X. This model has the purpose of defining an equation of a straight line that represents the smallest accumulated distance between the points plotted on the graphic and this straight line. In other words, the equation minimizes the sum of the deviations between y and $\hat{y}$, where y is the actual value and $\hat{y}$ is the value defined by equation of straight line. In our study Y and X are, respectively, the observed values in the time series and the time t. Thus, in this context, the linear regression provides a model that can be used to predict scores on variable Y from the relation identified between X and Y. Equation 1 represents this model.

$$\widehat{y_x} = \beta_0 + \beta_1 x \qquad (1)$$

where x is the time index, $\widehat{y_x}$ is the predicted value in time x, β_0 is the intercept, that is, the point where the straight line cuts the axis Y when x is equal to zero, and β_1 is the slope, that represents the straight line angular coefficient, i.e., the value increased or decreased in Y for each increase of one unit in X.

Least squares is the standard method used for estimating the parameters β_0 and β_1. Thus, β_0 is estimated by Equation 2.

$$\beta_0 = \bar{y} - \beta_1 \bar{x} \qquad (2)$$

where $\bar{y}$ and $\bar{x}$ are the arithmetic mean of the observed samples for each variable, and β_1 is estimated by Equation 3.

$$\beta_1 = \frac{\sum_{i=1}^{T} x_i\, y_i - \dfrac{\left(\sum_{i=1}^{T} x_i\right)\left(\sum_{i=1}^{T} y_i\right)}{T}}{\sum_{i=1}^{T} x_i^2 - \dfrac{\left(\sum_{i=1}^{T} x_i\right)^2}{T}} \qquad (3)$$

where T is the number of observed samples.

3. Forecasting proposed model

The forecasting model proposed in this article includes an extension to the pre-processing introduced in [9], a new smoothing technique and the use of linear regression. In the sequence, the pre-processing model and forecasting model are presented.

3.1. Pre-processing model

The purpose of the pre-processing phase is to identify and remove outliers, define the domain and the number of linguistic terms and, finally, define the support of each fuzzy set to fuzzify the time series. For this, an extension of pre-processing introduced in [9] is proposed to improve the process by which the number and the support of linguistics terms are defined.

Once the outliers have been identified and removed [9], the domain $[D_{min}, D_{max}]$ is calculated by Equation 4, where σ is the standard deviation, d_{min} is the smallest observed value and d_{max} is the highest observed value of the training samples.

$$D = [d_{min} - \sigma, d_{max} + \sigma] \qquad (4)$$

After the domain has been defined, the following six steps are used to calculate the number of linguistic terms according to [9].

Step 1. Sort the t numerical samples of the training dataset in ascending order.

Step 2. Calculate the threshold τ of stop condition as in Equation 5, where t is the index of the most recent sample and d_i is the sample at time i.

$$\tau = \frac{\sum_{i=1}^{t-1}(d_{i+1} - d_i)}{t - 1} \qquad (5)$$

Step 3. Consider each crisp value in the time series as one cluster $\{d_1\}, \{d_2\}, \ldots, \{d_k\}$, where the symbol "{ }" indicates a cluster, and k is the total number of clusters.

Step 4. Define the centroid $c_i, i = 1, 2, \ldots, k$. In clusters with more than one sample, the centroid c_i should be defined by the arithmetic mean of all samples in the cluster.

Step 5. Calculate the distance between centroids c_m and c_{m+1}, where $m = 1, 2, \ldots, k-1$.

Step 6. If the shortest distance identified in the previous step is smaller than the threshold τ, then merge the elements of the respective clusters into a cluster containing elements of both clusters, update the index of the centroids and go back to the Step 4. Otherwise, stop the algorithm.

The number k of linguistics terms found by these previous steps will be used as the input parameter in the Fuzzy C-Means (FCM) algorithm [12], providing the number of clusters to the algorithm. The centroids resulting from the FCM algorithm will be used to define the central parameters of fuzzy triangular sets, which will be used to fuzzify the time series.

The proposal to create the additional linguistic terms introduced in this paper, is composed by four steps and is presented as an extension to the procedure just described. The aim of this proposal is to add new linguistic terms in the intervals $[D_{min}, c_1]$ and $[c_k, D_{max}]$, to increase the accuracy of forecasts in problems where the samples in the test dataset are smaller than d_{min}, or

higher than d_{max}, respectively, smallest and highest observed values in the training dataset. The four steps to aggregate the new linguistics terms are introduced next.

Step 1. Calculate the arithmetic mean ave_{dc} of the differences between the centroids c_m and c_{m+1}, where $m = 1, 2, \ldots, k - 1$.

Step 2. Add triangular fuzzy sets to the interval $[D_{min}, c_1]$, while the central parameter of the triangular set being added is higher than or equal to D_{min}. Calculate the central parameter of the fuzzy sets by subtracting the average ave_{dc} from the centroid c_1 for the first fuzzy set and from the last fuzzy set added for the next ones. If the central parameter of the first calculation is smaller than D_{min}, add only one new triangular fuzzy set with central parameter defined by the arithmetic mean between D_{min} and c_1.

Step 3. Add triangular fuzzy sets, to the interval $[c_k, D_{max}]$, using a process similarly to Step 2, summing the average ave_{dc} to the central point of the last fuzzy set created, starting from c_k. If the sum of ave_{dc} with c_k the is higher than D_{max}, add only one new triangular fuzzy set with central parameter defined by the mean between D_{max} and c_k.

Step 4. Update the indexes of centroids.

Once the fuzzy sets have been defined, the crisp samples in the time series should be fuzzified and the knowledge base with fuzzy logic relations (FLR) should be created, before the forecasting method can be performed. To fuzzify the time series, the observed samples are converted to the linguistic terms with the higher membership degree.

3.2. Forecasting model

In this forecasting model a new smoothing technique is used for smoothing the consequents of the FLR as well as the impact of the FLR results. Linear regression is used to assist the calculation of predicted value when there is no FLR available in the knowledge base to do the forecasting.

After the fuzzification of the crisp samples, the second-order FLRs in the form shown in Equation (6) are extracted from the fuzzified time series,

$$L_i, L_{i+1} \rightarrow L_{i+2}, L_{i+3}, L_{i+4} \qquad (6)$$

where L_i is the linguistic term at time i. The set of all second-order FLR in the format of Equation 6 extracted from the time series represent the knowledge base of the system.

When a forecast is to been done, the last two linguistics terms L_{t-1} and L_t in the FTS indicate which FLR in the base will be used to compute the forecasted value. All FLRs in the base whose antecedents are $L_i = L_{t-1}$ and $L_{i+1} = L_t$ will be used to compute the predicted value.

To the forecasting proposed model, a new smoothing technique is introduced to be applied in the consequents of the FLR and in the values provide by them. The smoothing proposed is shown in Equation 7.

$$w_i = \frac{\alpha^i}{\sum_{j=0}^{n-1} \alpha^j} \qquad (7)$$

where w_i is the smoothing weight for term i, n is the number of terms to be smoothed and α is the smoothing factor. As it is possible to observe in Equation 7, to obtain the desired result, the index of the first smoothed term should be set to zero. For the smoothing factor, if $0 < \alpha < 1$, the weight of the last terms will be smaller than the weight of the first terms being smoothed. If $\alpha = 0$, only the first term will be consider. When $\alpha = 1$, all terms will have the same weight and, if $\alpha > 1$, the weights of the last terms will be higher than first terms.

To illustrate the behavior of the smoothing technique proposed, Figure 1 presents the smoothing weights when three terms are considered for α values smaller than 1.
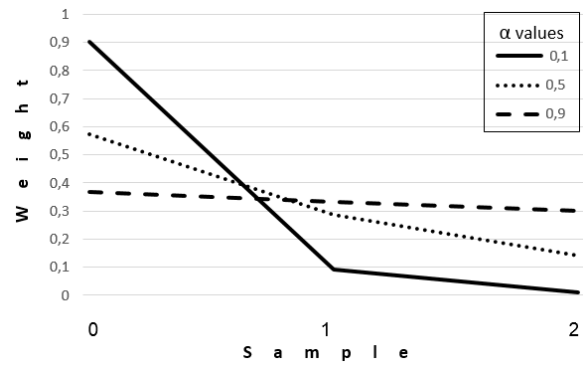

Fig. 1: Smoothing weights for α smaller than one.

In the graphic shown in Figure 1, it is possible to observe that the smoothing weight decreases drastically when α value is far from 1, as in the continuous line. If the α value is close to 1, as in dashed line, the smoothing becomes softer. Figure 2 shows the behaviors of the smoothing weight for α values higher than 1 considering three terms to be smoothed.
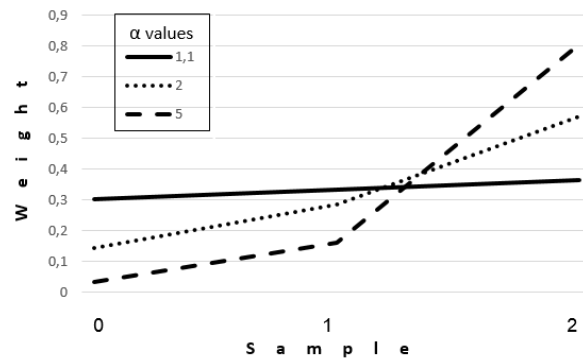

Fig. 2: Smoothing weights for α higher than one.

In Figure 2, it is possible to observe that the smoothing weight increases drastically when α values are far from 1. Otherwise, if the α value is close to 1, as in continuous line, the smoothing becomes softer.

Assume that, to perform a forecast from a FTS with L_{t-1} and L_t as the two most recent values, there are F FLRs in the knowledge base with these values as antecedents. The smoothing method introduced is applied to the consequents of these F FLRs, and then each FLR provides a smoothed value S_f. Equation 8 shows how each S value is obtained from the smoothing of the FLR.

$$S_f = \sum_{i=0}^{n-1} w_i \cdot l_i \qquad (8)$$

where S_f is the smoothed value inferred by FLR f, $f = 0, 1, \ldots, F-1$, n is the number of consequents in the FLR, l_i are the crisp values of the linguistic terms in the consequent of FLR, that is, the central parameter of triangular fuzzy set L_i and w_i is the smoothing weight obtained by Equation 7. The purpose of the method proposed here is to make the first term in the consequent of the FLR have a higher influence in the calculation of the inferred value, using a α value smaller than one. However, this α value can be adjusted as necessary to become suitable to different purposes.

These S_f values will compose a new series, where the value inferred by the oldest FLR will be the first sample of this new series. In an analogue way, the value inferred by Equation 8 using the most recent FLR with the antecedents L_{t-1} and L_t, will be the last sample in the new series. The smoothing is then applied again to this new generated series in order to obtain the forecasted value. Equation 9 shows how the forecasted value is obtained.

$$d_{t+1} = \sum_{f=0}^{F-1} w_f \cdot S_f \qquad (9)$$

In this equation d_{t+1} is the forecasted value, F is the total number of FLRs with the antecedents L_{t-1} and L_t, S_f is the smoothed crisp value obtained by Equation 8 through FLR f, and w_f is the smoothing weight obtained by Equation 7 to ponder the value S_f. In this work, the purpose of the model is to make the last samples in the new series, which represent the most recent behavior in the original time series, to have the higher influence in the calculation of the forecasted value. For this purpose, the α value should be set to a value higher than one. However, the smoothing factor can be adjusted to obtain the behavior as desired.

When the knowledge base does not have at least one equivalent FLR to be used in the process of forecasting, some forecast methods use the centroid of the last linguistic term in the FTS as the forecasted value [8, 13, 14]. Aiming at increasing the accuracy in forecasting in these cases, in the proposed model the use of simple linear regression combined with FTS is introduced. The aim of this procedure is to identify the current behavior of the time series and use it to calculate the forecasted values. Considering that the slope β_1, obtained by Equation 3 (see section 2.2), indicates how the time series increases or decreases at each new time unit, we aggregate β_1 to the crisp value l_t that represents the last linguistic term L_t in the time series. Thus, the prediction is made as shown in Equation 10.

$$d_{t+1} = l_t + \beta_1 \qquad (10)$$

where d_{t+1} is the forecasted value. Since the objective of this form of calculation is to identify the recent trend of the series, a reduced number of the most recent samples should be used in the calculation of the slope, especially in time series with irregular or seasonal behavior. The

use of linear regression in the forecasting model as described, allows a better accuracy of the forecasting model.

Regardless of the predicted value to be calculated by Equation 9 or by Equation 10, the last procedure in the forecasting process is to identify what is the linguistic term in which the forecasted value has the higher membership degree. The crisp value of this linguistic term will be the forecasting provided by the model. In the experiments using the crisp value l_i of linguistic term that represents d_{t+1} calculated by Equation 9 or Equation 10 as predicted value, it was possible to obtain a better accuracy than using the value d_{t+1} itself.

Finally, in order to avoid that the initial FLR base becomes obsolete, after the arrival of a new sample, a new FLR is included in the base. To illustrate, consider that the four most recent linguistics samples in a FTS are L_{t-3}, L_{t-2}, L_{t-1} and L_t. With the arrival of a new sample L_n, $n = t + 1$, the new FLR L_{t-3}, $L_{t-2} \rightarrow L_{t-1}, L_t, L_n$ will be added to the base. This feature allows the method to enhance the accuracy rate in scenarios where the test dataset presents a different behavior when compared with the training data.

4. Experiments and results

To demonstrate the effectiveness of the proposed model, the experiments were run using the TAIEX index between 2001 and 2004. In order to confirm the satisfactory accuracy of the proposed model, the results were compared with eight others approaches by means of root mean squared error (RMSE). The samples from January to October of each year were used as the training dataset, and the samples from November and December were used as the test dataset. Table 1 shows the number of training samples, the number of test samples, the domain and the number of linguistics terms (LT) defined by the pre-processing for each time series.

Table 1: Data of pre-processing in the TAIEX index.

TAIEX	Training Samples	Test Samples	Domain	LT
2001	201	43	[2732.30; 6818.20]	102
2002	205	43	[3190.59; 7121.75]	86
2003	206	43	[3580.74; 6666.89]	91
2004	205	45	[4861.44; 7489.53]	108

After the definition of the number of triangular fuzzy sets and their respective supports, the next step is to fuzzify the crisp samples and extract the FLR to compose the knowledge base. As an example, consider the first six samples 5600.05, 5526.32, 5638.53, 5834.89, 5810.08 and 5865.54 in the training dataset of the TAIEX 2002 index. Their respective linguistics terms, according to Table 2, are $L_{58}, L_{56}, L_{59}, L_{63}, L_{63}$ and L_{64}. Thus, the first two second-order FLRs to be extracted, according to Equation 6, to compose the knowledge base are, respectively, $L_{58}, L_{56} \rightarrow L_{59}, L_{63}, L_{63}$ and $L_{56}, L_{59} \rightarrow L_{63}, L_{63}, L_{64}$.

Table 2: Linguistics terms for TAIEX 2002 index and their respective representation values.

Linguistic Term	Value	Linguistic Term	Value
0	3209.20	43	4919.11
1	3254.98	44	4936.69
2	3300.76	45	4966.86
3	3346.54	46	5000.35
4	3392.32	47	5044.12
5	3438.10	48	5072.30
6	3483.88	49	5127.79
7	3529.66	50	5160.66
8	3575.44	51	5251.76
9	3621.22	52	5317.39
10	3666.99	53	5394.30
11	3712.77	54	5443.52
12	3758.55	55	5499.93
13	3804.33	56	5533.52
14	3850.11	57	5566.81
15	3917.56	58	5599.27
16	3956.42	59	5631.98
17	4071.91	60	5666.98
18	4131.56	61	5696.33
19	4167.11	62	5743.14
20	4188.86	63	5805.63
21	4208.86	64	5867.94
22	4222.81	65	5966.14
23	4280.93	66	6091.94
24	4287.06	67	6276.42
25	4328.56	68	6278.42
26	4386.80	69	6324.20
27	4430.67	70	6369.97
28	4463.98	71	6415.75
29	4531.21	72	6461.53
30	4560.86	73	6507.31
31	4579.39	74	6553.09
32	4598.21	75	6598.87
33	4635.28	76	6644.65
34	4646.03	77	6690.43
35	4660.64	78	6736.21
36	4668.14	79	6781.99
37	4700.41	80	6827.77
38	4720.96	81	6873.55
39	4766.15	82	6919.33
40	4801.85	83	6965.11
41	4854.20	84	7010.89
42	4886.03	85	7056.66

Once the FLR base has been defined, the forecasting process can be applied. The process starts by identifying the two most recent linguistics terms in the FTS to search for the equivalents FLR in the base. When one or more FLR are found, their consequents are smoothed by the process described in Section 3.2 and, each FLR infers a smoothed value S_f, $f = 0,1,...,F-1$ where F is the number of FLRs found. In the sequence, the values S_f are smoothed according to Equation 9 to produce a final inferred value. The crisp value of the linguistic term that better represents the number provide by Equation 9 will be the forecasted value defined by the model.

Assume that α_1 is the value used for smoothing the consequents of the FLR, therefore, used to define w_i in Equation 8 and, α_2 is the smoothing factor used to define w_f in Equation 9. Considering the different behaviors of the time series in the experiments, the smoothing factors are presented in Table 3.

Table 3: Smoothing factors used in the experiments.

TAIEX	α_1	α_2
2001	0.5	1.2
2002	0.4	0.2
2003	0.1	0.1
2004	0.4	7.0

To illustrate the forecasting process, if we consider the 34th prediction in the test dataset made for 2002/Dec/18, the two last linguistics terms of 2002/Dec/16 and 2002/Dec/17 are, respectively, L_{31} and L_{29}. Two FLR were found in the base, the first FLR $L_{31}, L_{29} \rightarrow L_{28}, L_{29}, L_{36}$ and the second FLR $L_{31}, L_{29} \rightarrow L_{31}, L_{30}, L_{38}$. Considering the smoothing factor $\alpha_1 = 0.4$, according to Equation 7 the smoothed weight for the consequents of the FLR are 0.64, 0.26 and 0.1, respectively for the first, the second and the third linguistics terms in the consequents. From these two FLR and considering Equation 8, the inferred values were $S_0 = 4502.15$ and $S_1 = 4589.15$. Given $\alpha_2 = 0.2$, the result obtained by Equation 9 is 4542.37. Considering the central parameter of the triangular fuzzy set with higher membership degree for this number, the forecasted value is 4531.21.

If there is not an FLR in the base to do the forecast, the prediction will be calculated with the aid of linear regression. An important parameter of the model is how many samples will be used in the regression. In the performed experiments, this number was equal to twenty percent of the total samples in the training dataset. Considering the training dataset of the TAIEX 2002 index, the slope was calculated using the forty-one most recent samples. For example, to do prediction on 2002/Nov/25, sixteen earlier samples were available in the test dataset. Thus, to complete the forty-one samples were necessary the twenty-five most recent samples of the training dataset, resulting in the predicted value 4720.95. The actual and the forecasted values for TAIEX indexes between 2001 and 2004 are shown in Figures 3, 4, 5 and 6, respectively.
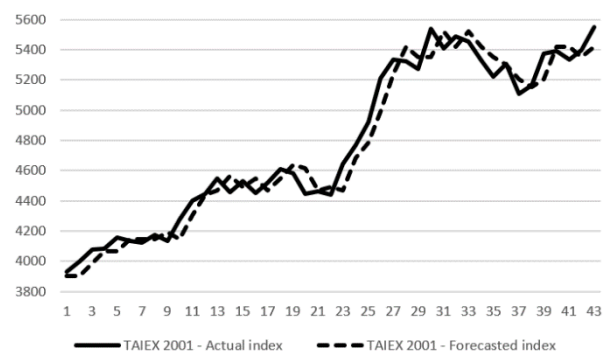

Fig. 3: Actual and forecasted indexes for November and December 2001.
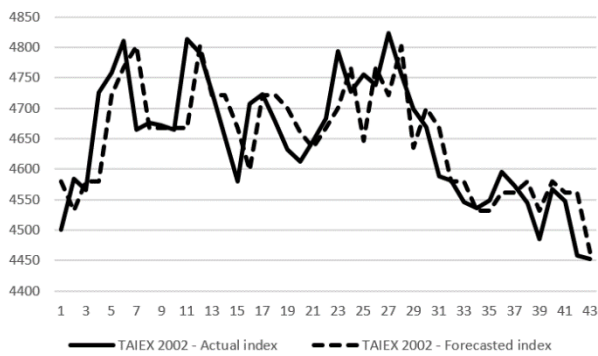
Fig. 4: Actual and forecasted indexes for November and December 2002.

Fig. 5: Actual and forecasted indexes for November and December 2003.
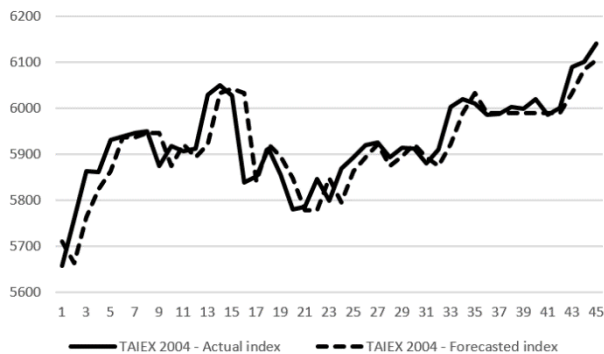

Fig. 6: Actual and forecasted indexes for November and December 2004.

Besides the satisfactory accuracy obtained by the method, the graphics also allow us to assert that the behaviors of the forecasted time series are very similar to those presented by the actual time series. Table 4 shows the root mean square error for the proposed method with linear regression and without linear regression compared with eight others approaches available in the literature.

The good performance of the proposed method indicates the effectiveness of the model compared with classic approaches as exponential smoothing [12], and with others recent approaches found in the literature [4, 13, 15, 16, 17, 18, 19]. An important feature observed in Table 4 is the essential role of linear regression in the model. When compared with the model without linear

regression, the accuracy increases, especially in the predictions where the FLRs were not found in the knowledge base to do the forecast.

Table 4: RMSE for forecasts of TAIEX indexes between 2001 and 2004.

Approach	2001	2002	2003	2004	Ave
Brown and Meyer [12]	123.1	66.0	52.6	56.9	74.65
Chen and Chen [15]	115.3	71.0	58.0	57.3	75.40
Chen, et. al.[16]	114.4	67.1	52.4	52.8	71.68
Santos and Camargo[13]	106.2	66.7	51.1	52.9	69.23
Chen and Kao [4]	114.5	76.8	54.2	58.1	75.90
Li and Chiang [17]	115.8	64.3	57.6	55.5	73.30
Chen and Chen [18]	114.6	64.7	53.6	52.9	71.45
Wang and Lee [19]	120.2	69.7	56.3	54.2	75.10
Proposed method [1]	107.5	64.8	49.2	**52.2**	68.43
Proposed method [2]	**97.2**	**63.2**	**48.9**	52.2	**65.38**

[1] without linear regression; [2] with linear regression

5. Conclusions

The forecasting model proposed in this article provides a significant contribution in a research area with growing interest and with a big potential for new developments. The combination of FTS with concepts and methods from statistics is certainly a promising field in the prediction of time series.

The first important feature of the proposed model is the improvement of the procedure to define the number of linguistics terms by the pre-processing method. This adjustment allowed increasing the accuracy in forecasting problems where the test dataset has values out of the domain presented by the samples in the training dataset. Other important contribution of this paper, the smoothing technique for time series, demonstrated to be coherent and effective through the results obtained by the model. Besides that, the smoothing method introduced here allows the adjustment of the smoothing factor as necessary.

Finally, the results showed the importance of the linear regression combined with the FTS and the smoothing process proposed. The combination of linear regression in the model provides a manner to improve the accuracy in situations where there are not equivalent FLRs in the base.

For future works, the model can be improved to do an automatic adjustment of the best number of samples to be used in the linear regression. Other possible improvement is the introduction of a dynamic domain, where new fuzzy sets can be aggregated up or down the limits of the initial domain, according to the current behavior of the test time series.

References

[1] Q. Song and B. S. Chissom: Fuzzy Time Series and its Models. *Fuzzy Sets and Systems* 54, pages 269–277, Elsevier, 1993.

[2] L. A. Moraes, R. A. Flauzino, M. A. Araújo and O. E. Batista: A Fuzzy Methodology to Improve Time

Series Forecast of Power Demand in Distribution Systems. *2013 IEEE Power and Energy Society General Meeting (PES)*, pages 1-5, IEEE, 2013.

[3] M. Stepnicka, P. Cortez, J. P. Donate and L. Stepnicková: Forecasting seasonal time series with computational intelligence: On recent methods and the potential of their combinations. *Expert Systems with Applications* 40, pages 1981–1992, Elsevier, 2013.

[4] S. M. Chen, P. Y. Kao: TAIEX forecasting based on fuzzy time series, particle swarm optimization techniques and support vector machines. *Information Sciences* 247, pages 62–71, Elsevier, 2013.

[5] G. P. Zhang: Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50, pages 159–175, Elsevier, 2003.

[6] M. N. Mehr, F. F. Samavati and M. Jeihoonian: Annual energy demand estimation of Iran industrial sector by Fuzzy regression and ARIMA. *Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pages 593-597, IEEE, 2011.

[7] C. C. Wang: A comparison study between fuzzy time series model and ARIMA model for forecasting Taiwan export. *Expert Systems with Applications 38,* pages 9296–9304, Elsevier, 2011.

[8] F. J. J. Santos and H. A. Camargo: Forecasting in Fuzzy Time Series by an Extension of Simple Exponential Smoothing. *Proceedings of 14th Ibero-American Conference on AI*, pages 257–268, Springer International Publishing, 2014.

[9] F. J. J. Santos and H. A. Camargo: Preprocessing in Fuzzy Time Series to Improve the Forecasting Accuracy. *12th International Conference on Machine Learning and Applications*, pages 170-173, IEEE, 2013.

[10] L. A. Zadeh: The Concept of a Linguistic Variable and its Application to Approximate Reasoning - Part 1. *Information Sciences* 8, pages 199–249, Elsevier, 1975.

[11] F. Galton: Natural Inheritance, 5th edition, New York, Macmillan and Company, page 226, 1894.

[12] J. C. Bezdek, E. C. Tsao and N. R. Pal: Fuzzy Kohonen Clustering Networks. *IEEE International Conference on Fuzzy Systems*, pages 1035-1043, 1992.

[13] R. G. Brown and R. F. Meyer: The fundamental theory of exponential smoothing. *Operations Research* 9, pages 673-685, 1961.

[14] C. Cheng, G. Cheng and J. Wang: Multi-attribute fuzzy time series method based on fuzzy clustering. *Expert Systems with Applications* 34, pages 1235–1242, Elsevier, 2008.

[15] S. M. Chen and C. D. Chen: TAIEX Forecasting Based on Fuzzy Time Series and Fuzzy Variation Groups. *IEEE Transactions on Fuzzy Systems* 19, pages 1-12, IEEE, 2011.

[16] K. Tanuwijaya and S. Chen: A new method to forecast enrollments using fuzzy time series and clustering techniques. *Proceedings of the Eighth International Conference on Machine Learning and Cybernetics*, pages 12-15, 2009.

[17] C. Li and T. W. Chiang: Complex Neurofuzzy ARIMA Forecasting: A New Approach Using Complex Fuzzy Sets. *IEEE Transactions on Fuzzy Systems* 21, pages 567-584, IEEE, 2013.

[18] S. M. Chen and S. W. Chen: Fuzzy Forecasting Based on Two-Factors Second-Order Fuzzy-Trend Logical Relationship Groups and the Probabilities of Trends of Fuzzy Logical Relationships. *IEEE Transactions on Cybernetics*, pages 1-5, IEEE, 2014.

[19] C. R. Wang and S.J. Lee: Temporal Prediction Using Self-Organizing Multilayer Perceptron. Proceedings of the 2014 International Conference on Machine Learning and Cybernetics, pages 585-591, IEEE, 2014.