Adaptive Fuzzy C-Regression Modeling for Time Series Forecasting

Leandro Maciel[1] **André Lemos**[2] **Rosangela Ballini**[3] **Fernando Gomide**[1]

[1]Department of Computer Engineering and Automation, School of Electrical and Computer Engineering
University of Campinas, Brazil
{maciel,gomide}@dca.fee.unicamp.br
[2]Department of Electronics Engineering, Federal University of Minas Gerais, Brazil
andrepl@cpdee.ufmg.br
[3]Department of Economy Theory, Institute of Economics, University of Campinas, Brazil
ballini@eco.unicamp.br

Abstract

The aim of the 2015 IFSA-EUSFLAT International Time Series Competition, *Computational Intelligence in Forecasting* (CIF), is to evaluate the performance of computational intelligence-based approaches to forecast time series of different nature. The participants must propose a unique consistent methodology for all time series. This paper suggests an adaptive fuzzy c-regression modeling approach (aFCR) for time series forecasting. The aFCR is a fuzzy clustering with affine prototypes modeling approach to develop fuzzy functional rule-based models. The approach uses participatory learning to adapt the model structure as it processes data as a stream of time series values. Computational experiments show that the aFCR forecaster is an effective tool to forecast time series.

Keywords: Adaptive clustering, fuzzy modeling, time series forecasting, stream data.

1. Introduction

Time series forecasting (TSF) models aim at predicting the behavior of a given phenomenon based solely on the past data of the same phenomenon. Predictive analytics and forecasting are important tools to support operation, planning, and strategic decision making. The issue of predicting time series has widespread applications in business, economy, finance, engineering, biology, agriculture, and medicine.

Statistical TSF methods such the ARIMA family and its variants, still are dominant in practice [1], but recently computational intelligence (CI) techniques have received a great deal of attention [2]. Relevant examples of CI approaches applicable to TSF include artificial neural networks [3], recurrent neural networks [4], evolutionary computation [5], support vector machines [6], immune systems [7], fuzzy techniques [8], fuzzy rule-based modeling [9], and their hybridizations [10, 11]. Fuzzy and neural fuzzy-based methods have shown to be particularly effective to handle series with imprecise data.

The 2015 IFSA-EUSFLAT competition *Computational Intelligence in Forecasting* (CIF) gives a new opportunity to evaluate the predictive accuracy of forecasting approaches derived from computational intelligence methodologies. The data set of the competition consists of time series with yearly, quarterly, and monthly frequencies. The aim of the CIF is to develop forecasts for a number of distinct time series and time horizons using a unique consistent methodology. The data set has a total of 91 times series with different number of samples each. Forecasts must be developed for different time horizons, that is, the number of values to be forecasted varies. We assume that long term, multiperiod forecasts must be based on previous forecasted values.

The forecasting approach proposed in this paper is based on an adaptive/evolving fuzzy modeling approach [12, 13] called adaptive fuzzy c-regression (aFCR). aFCR is a method based on recursive clustering with affine prototypes. The forecasting model is a fuzzy functional rule-based model of the Takagi-Sugeno (TS) type. The rule base emerge from the cluster structure because each cluster translates in a functional fuzzy rule. The model structure and functionality evolves as time series data is input. Time series data are viewed as a stream of data, and each datum of the series is processed sequentially. This is particularly helpful in online and real-time domains. The TS model is useful to reflect the ideia of local modeling in fuzzy regions. TS rule antecedents encapsulate fuzzy regions of the input space, and rule consequents are affine functions of the input variables. Nonlinearities are captured by combining the local affine models that assemble the multi-model structure. This enables Takagi-Sugeno models to approximate virtually any continuous function in compact domains [14].

Construction of TS rule-based models requires the identification of its structure and estimation of the parameters of the local affine models [15]. In a data stream processing framework, structure identification can be done using recursive fuzzy clustering techniques. The parameters of the rule consequent can be estimated using a recursive least squares algorithm [16]. Recursive modeling simultaneously learns the model structure and the corresponding model parameters using incremental learning procedures. The main difference among the various recursive fuzzy modeling approaches is how the model structure is derived, usually how the clustering step is performed. Examples of different types of recursive, incremental fuzzy rule-based and fuzzy neural modeling approaches include the following.

After the pioneering evolving Takagi-Sugeno (eTS) modeling [17] approach and its extensions (e.g. Simpl_eTS [18], eXtended eTS (xTS) [19]) an autonomous modeling scheme called eTS+ [12] was developed. Here, autonomy means that no user-defined control parameters are needed by eTS+. Instead, eTS+ uses criteria such as age, utility, local density, and zone of influence to update the model structure. Later, ePL+ was developed in the realm of participatory learning clustering [20]. ePL+ extends the ePL approach [21] and uses similar updating mechanisms as eTS+.

An alternative method for evolving TS modeling is given in [22] based on a recursive form of the fuzzy c-means (rFCM) algorithm to learn the model structure. Later, the rFCM modeling was translated in a recursive Gustafson-Kessel (rGK) algorithm. The aim of the rGK is to capture different cluster shapes [23] that appear in a cluster structure. Combination of the rGK algorithm and evolving mechanisms such as adding, removing, splitting, merging clusters, and recursive least squares became a powerful evolving fuzzy modeling approach called eFuMo [24].

An evolving fuzzy modeling approach using tree structures, namely, evolving fuzzy trees (eFT) was introduced in [25]. The eFT model is a fuzzy linear regression tree whose topology can be continuously updated employing a statistical model selection test. A fuzzy linear regression tree is a fuzzy tree with a linear model in each leaf. Experiments in the realm of time series forecasting have shown that the fuzzy evolving regression tree is a promising approach for adaptive system modeling.

A distinct, but conceptually similar approach for TS modeling, is the dynamic evolving neural fuzzy inference system model (DENFIS) [10]. DENFIS uses distance-based recursive clustering to adapt the rule base structure. The weighted recursive least squares with forgetting factor algorithm updates the parameters of the rule consequent.

An on-line sequential extreme learning (OS-ELM) algorithm for single hidden layer feedforward networks with either additive or radial basis function hidden nodes in a unified framework was developed in [26]. Computational experiments using stream data of benchmark problems drawn from the regression, classification, and time series forecasting has shown that OS-ELM has high generalization capability and is computationally fast.

A recursive clustering algorithm derived from a modification of the vector quantization technique called evolving vector quantization is another way to construct flexible fuzzy inference systems (FLEXFIS) [27]. More recent examples of evolving modeling include the neural fuzzy type-2 self-organizing fuzzy modeling and modified least-squares network (SOFMLS) [28], and the sequential adaptive fuzzy inference system (SAFIS) [29]. SOFMLS employs an evolving nearest neighbor clustering algorithm. SAFIS uses a distance criterion in conjunction with an influence measure of the new rules created to develop and update the rule base. Further instances of evolving methods include [30, 31, 32].

The structure of the evolving/adaptive model suggested in this paper is based on the fuzzy c-regression method (FCRM) introduced in [33]. The idea is to find affine prototypes instead of cluster centers, and use the affine prototypes to construct the adaptive fuzzy c-regression approach (aFCR). The main advantage is that both, model structure and model parameters, are derived simultaneously by the same algorithm. Affine prototypes constitute local models whereas fuzzy clusters contain information on the local models domains in terms of their membership functions. Recently [34] developed a recursive version of the fuzzy c-regression method together with an evolving algorithm to add new local models whenever current data of a data stream can not be explained by the current cluster structure.

The fuzzy c-regression modeling framework developed in this paper uses participatory learning to adapt the cluster structure, and an utility measure to evaluate the quality of the current cluster structure. The utility measure allows the rule base to shrink by removing rules with low utility (the data pattern shifted away from the domain of the rule) and provides a simpler and more relevant rule base by considering the rules with higher summarization power.

After this introduction, this paper proceeds as follows. Section 2 introduces the adaptive fuzzy c-regression modeling framework. The forecasting methodology, the forecaster design, and forecasting performance analysis using the data made available are addressed in Section 3. Finally, conclusions and issues for further development are summarized in Section 4.

2. Adaptive fuzzy c-regression modeling

2.1. Fuzzy model structure

The adaptive fuzzy c-regression modeling approach relies on local first order Takagi-Sugeno (TS) fuzzy models. TS fuzzy models consist of a set of fuzzy functional rules of the form:

$$\mathcal{R}_i: \text{ IF } \mathbf{x} \text{ is } \mathcal{A}_i \text{ THEN } y_i = f(\theta_i, \mathbf{x}), \qquad (1)$$

where $\mathcal{R}_i$ is the i-th fuzzy rule, $i = 1, 2, \ldots, c$, c is the number of fuzzy rules, the q-element vector $\mathbf{x} = [x_1, \ldots, x_q]^T \in \Re^q$ denotes the input, $y_i \in \Re$ is the output of the i-th local model, and $f(\theta_i, \mathbf{x}) = \theta_{i0} + \theta_{i1}x_1 + \ldots + \theta_{iq}x_q$ is an affine function of the input. $\Re$ is the set of real numbers. $\mathcal{A}_i$, $i = 1, \ldots, c$, is a multidimensional fuzzy set whose membership function is $\mathcal{A}_i(\mathbf{x}) : \Re^q \rightarrow [0, 1]$.

Fuzzy inference using TS rules (1) is done as follows:

$$y = \sum_{i=1}^{c} \left(\frac{\mathcal{A}_i(\mathbf{x})y_i}{\sum_{j=1}^{c} \mathcal{A}_j(\mathbf{x})} \right). \tag{2}$$

Expression (2) can be rewritten using normalized degree of activation λ_i:

$$y = \sum_{i=1}^{c} \lambda_i y_i, \tag{3}$$

where

$$\lambda_i = \frac{\mathcal{A}_i(\mathbf{x})}{\sum_{j=1}^{c} \mathcal{A}_j(\mathbf{x})}, \tag{4}$$

is the normalized firing level of the i-th rule.

TS modeling methodology requires two steps: i) learn the rule antecedent using a fuzzy clustering algorithm, and ii) estimate the parameters of the affine rule consequents [17]. The adaptive fuzzy c-regression modeling approach aFCR performs fuzzy clustering using affine prototypes whose parameters are estimated using the weighted recursive least squares algorithm (wRLS). The participatory learning principle is employed to adapt the cluster structure, that is, to create, exclude or merge fuzzy rules depending on input data. These steps are detailed in the next sections.

2.2. Fuzzy clustering using affine prototypes

The main issue in clustering-based fuzzy modeling is how to partition the input-output space. Clustering techniques use prototypes as proxies of the cluster structure. Often, prototypes are selected points of the input-output space called cluster centers or centroids. In this paper cluster prototypes are affine functions. This means that clusters are represented by hyperplanes in the data space [34].

In the evolving/adaptive fuzzy c-regression approach, the fuzzy c-regression clustering method (FCRM) [33] is used to partition the input-output space viewing data of a data set as a source of a data stream. In FCRM the distance between the current input data and the prototypes is the estimation error:

$$d_{ik}^2 = \left(y_k - \mathbf{x}_{ek}^T\theta_i\right)^2, \; i = 1, \ldots, c, \tag{5}$$

where $\mathbf{x}_{ek}^T = [1 \; \mathbf{x}_k^T]$ is the expanded input vector at k, y_k the output at k, and θ_i the parameters of the consequent of the i-th fuzzy rule.

Distance (5) is used similarly as in the fuzzy c-means clustering algorithm [35]. The minimization of the objective function subject to constraints on membership degrees, as in the fuzzy c-means, gives membership degrees μ_{ik}:

$$\mu_{ik} = \left(\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{\eta-1}} \right)^{-1}, \tag{6}$$

where η is a real number that governs the influence of membership degrees. Usually $\eta = 2$.

The affine prototypes rely on the estimation of the consequent parameters as follows.

2.3. Consequent parameters estimation

The value of a local model output is a weighted combination of the inputs:

$$y_i = \mathbf{x}_e^T\theta_i, \; i = 1, \ldots, c. \tag{7}$$

Using (7), expression (3), which is the sum of c local affine models, can be rewritten as:

$$y = \sum_{i=1}^{c} \lambda_i \mathbf{x}_e^T \theta_i, \tag{8}$$

where $\theta_i = [\theta_{i0}, \theta_{i1}, \ldots, \theta_{iq}]^T$ is the vector of parameters.

If the data set has n samples, each augmented regression input is $\lambda_{ik}\mathbf{x}_{ek}^T$, $k = 1, \ldots, n$, $i = 1, \ldots, c$. We can assemble the n augmented regression inputs in a $(q + 1) \times n$ data matrix as follows:

$$\Psi_i = \left[\lambda_{i1}\mathbf{x}_{e1}^T, \ldots, \lambda_{in}\mathbf{x}_{en}^T\right]. \tag{9}$$

Thus, the $n \times 1$ output data vector can be written as $\mathbf{y}_i = [y_{i1}, \ldots, y_{in}]^T$. Hence

$$\mathbf{y}_i = \Psi_i^T\theta_i. \tag{10}$$

When dealing with data streams, the estimation of consequent parameters can be done using the weighted recursive least squares algorithm (wRLS) [36] as in [12]. The wRLS can be summarized as follows. First, compute the weighted estimation error e_{ik} at k as follows [34]:

$$e_{ik} = y_{ik} - \lambda_{ik}\mathbf{x}_{ek}^T\theta_{ik}, \; i = 1, \ldots, c, \tag{11}$$

Next, compute the gain vector $\mathbf{K}_{ik}$:

$$\mathbf{K}_{ik} = \mathbf{P}_{i,k-1}\lambda_{ik}\mathbf{x}_{ek}\left(\gamma + \lambda_{ik}\mathbf{x}_{ek}^T\mathbf{P}_{i,k-1}\lambda_{ik}\mathbf{x}_{ek}\right)^{-1}, \tag{12}$$

and the covariance matrix $\mathbf{P}_{ik}$ of the estimation error:

$$\mathbf{P}_{ik} = \frac{1}{\gamma}\left(\mathbf{I} - \mathbf{K}_{ik}\lambda_{ik}\mathbf{x}_{ek}^T\right), \tag{13}$$

Here γ is the forgetting factor, $\mathbf{P}_{i0} = 10^3 I$, and I is an identity matrix of order $q+1$. Finally, update the parameters θ_{ik} of the current local model as follows:

$$\theta_{ik} = \theta_{i,k-1} + \mathbf{K}_{ik} e_{ik},\ \theta_{i0} = 0,\ i = 1, \dots, c. \quad (14)$$

Next section addresses the adaptation of the cluster structure using participatory learning.

2.4. Participatory learning

Participatory learning (PL) is a learning paradigm which assumes that the learning process depends on what the system has already learned from the data. The current knowledge is part of the learning process itself and influences the way in which new data are used for self-organization. An essential property of participatory learning is that the impact of new data in causing self-organization and model revision depends on its compatibility with the current knowledge, or equivalently in the context of this paper, on the compatibility of current input data with the existing cluster structure [21].

In participatory learning, a cluster structure is updated using a compatibility measure, $\rho_{ik} \in [0,1]$ and an arousal index, $a_{ik} \in [0,1]$. While ρ_{ik} measures how much a data point is compatible with the current cluster structure, the arousal index a_{ik} acts as a critic to sinalize when current cluster structure should be revised in front of new information brought by new data.

Due to its unsupervised, self-organizing nature, the PL clustering procedure may create a new cluster, or delete, merge, or modify the existing ones at each learning step. If the arousal index is greater than a threshold value $\tau \in [0,1]$, then a new cluster is created. The arousal index a_{ik} is updated as follows:

$$a_{ik} = a_{i,k-1} + \beta(1 - \rho_{ik} - a_{i,k-1}),\ a_{i0} = 0, \quad (15)$$

where $\beta \in [0,1]$ controls the rate of change of arousal, the closer β is to one, the faster the system is to sense compatibility variations, and

$$\rho_{ik} = 1 - |y_k - \mathbf{x}_{ek}^T \theta_i|, \quad (16)$$

stands for the compatibility index, $\rho_{ik} \in [0,1], \forall\, i, k$.

The arousal index can be interpreted as the complement of the confidence we have in the truth of the current belief, the rule base structure in this paper. The arousal mechanism monitors the performance of the system by observing the compatibility of the current model as new data are input.

In this paper, the quality of the cluster structure is monitored at each step considering the utility measure suggested in [12]. The utility measure is an indicator of the accumulated relative firing level of a rule:

$$\mathcal{U}_{ik} = \frac{\sum_{l=1}^{k} \lambda_{il}}{k - I^{i*}} \quad (17)$$

where I^{i*} is the step (time tag) that indicates when fuzzy rule $i*$ was created. Once a rule is created, the utility indicates how often the rule is being used. This quality measure aims to remove unused clusters from the structure. Clusters with low quality are deleted. In other words:

$$\text{IF } \mathcal{U}_{ik} < \epsilon \text{ THEN delete cluster } i:\ c \leftarrow c - 1, \quad (18)$$

where $\epsilon \in [0.03; 0.1]$ is a threshold that controls the utility of each cluster [12]. This principle guarantees higher relevance cluster structures and corresponding fuzzy local models. Alternative quality measures such as age, support, zone of influence and local density may be adopted.

2.5. Adaptive FCR algorithm

This section summarizes the adaptive fuzzy c-regression (aFCR) modeling approach. The steps proceed sequentially as the data set is scanned. The algorithm updates the existing model whenever new input data brings new information. Its recursive nature means that, as far as data storage is concerned, it is memory efficient.

```
Adaptive Fuzzy C-Regression Modeling
Set $c = 1$, $\theta_1 = 0$, $a_1 = 0$, $\mathbf{P}_1 = 10^3 I$, and $\eta = 2$.
Choose control parameters $\tau$, $\beta$, and $\epsilon$.
for $k = 1, 2, \dots, n$ do
    for $i = 1, 2, \dots c$ do
        compute $\rho_{ik} = 1 - |y_k - \mathbf{x}_{ek}^T \theta_i|$
        compute $a_{ik} = a_{i,k-1} + \beta(1 - \rho_{ik} - a_{i,k-1})$
          if $a_{ik} \geq \tau$ then create a new cluster
            $c \leftarrow c + 1$
          end if
        compute $d_{ik}^2 = \left(y_k - \mathbf{x}_{ek}^T \theta_i\right)^2$
        compute $\mu_{ik} = \left(\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{\eta-1}}\right)^{-1}$
    end for
    if $\mathcal{U}_{ik} < \epsilon$ then delete cluster $i$
        $c \leftarrow c - 1$
    end if
    update $\theta$ using the wRLS
end for
```

3. Computational experiments

This section discusses how aFCR forecasters have been developed. It addresses input selection and choice of the control parameters. Forecasting results for selected data sets are shown to illustrate the performance of the adaptive fuzzy c-regression approach. The results of aFCR are compared with ARIMA, seasonal ARIMA (SARIMA), and with state of the art evolving fuzzy and neuro-fuzzy modeling approaches, DENFIS [10], OS-ELM [26], eTS [17], xTS [37], eTS+ [12], ePL [21], and eFT [25].

The 2015 IFSA-EUSFLAT competition *Computational Intelligence in Forecasting* (CIF) provided

91 time series with different time frequencies, that include yearly, quarterly and monthly data. The competition task is to develop forecasts for all time series applying a unique consistent approach, the adaptive fuzzy c-regression modeling aFCR in this paper. Each time series has a different number of samples and forecasting horizon (number of values to be forecasted).

Model input selection for the evolving, ARIMA, and SARIMA approaches were based on the analysis of the partial autocorrelation function at 5% significance level. The input vector is assembled using the current and previous lags of series values as dictated by the autocorrelation function. Control parameters of the evolving techniques were chosen using simulations to select the parameters that provide highest accuracy average.

Performance evaluation of aFCR considers the criterion recommended by the competition organizers, the Symmetric Mean Absolute Percentage Error (SMAPE):

$$\mathrm{SMAPE} = \frac{1}{T}\sum_{t=1}^{T} \frac{|\hat{y}_t - y_t|}{(|y_t| + |\hat{y}_t|)/2}, \qquad (19)$$

where $\hat{y}_t$ is the t-th forecasted value, y_t the t-th actual value, and T is the forecast horizon.

To verify the performance of the evolving fuzzy c-regression modeling approach in time series forecasting using the competition data, six time series with distinct dynamics were chosen. Table 1 gives the details of each time series, including the number of inputs considered for all the algorithms. Label ID identifies the time series according to competition organizers notation, and T is the forecast horizon. Recall that the inputs are current and lagged time series values, as indicated by the partial autocorrelation functions.

Series ID	T	# samples	# inputs
6	90	266	5
11	180	909	3
37	12	36	1
40	12	51	2
74	4	14	2
87	12	206	5

Table 1: Time series description.

The computational experiments remove the last T samples from each time series data. Next, the forecasting models are run to predict the remaining T values. SMAPE is computed accordingly.

After 200 simulations for each of the series, control parameters of aFCR were set: $\tau = 0.21$, $\beta = 0.19$, and $\epsilon = 0.08$. Table 2 shows the SMAPE values for each time series achieved by ARIMA, SARIMA, DENFIS, OS-ELM, eTS, xTS, eTS+, ePL, eFT and aFCR. Overall, the aFCR modeling approach provided good results for all time series considered. For the time series whose IDs are 6, 11 and 40, the best result was achieved by aFCR. The ARIMA and SARIMA models developed the worst result for all series. The performance of the adaptive/evolving approaches are similar. The dynamic evolving neuro fuzzy inference system (DENFIS) gives the most accurate result for the time series ID 37, whereas the OS-ELM and the evolving fuzzy tree (eFT) achieve the best performance for time series IDs 74 and 87, respectively.

ID	6	11	37
ARIMA	0.326	1.162	0.315
SARIMA	0.268	1.069	0.291
DENFIS	0.179	0.795	**0.163**
OS-ELM	0.1.87	0.811	0.174
eTS	0.198	0.991	0.285
xTS	0.193	0.915	0.240
eTS+	0.185	0.843	0.266
ePL	0.213	0.917	0.260
eFT	0.299	0.804	0.178
aFCR	**0.159**	**0.714**	0.196
ID	40	74	87
ARIMA	0.123	0.096	0.065
SARIMA	0.104	0.074	0.052
DENFIS	0.078	0.032	0.022
OS-ELM	0.102	**0.030**	0.041
eTS	0.087	0.047	0.032
xTS	0.091	0.042	0.037
eTS+	0.079	0.042	0.031
ePL	0.082	0.054	0.026
eFT	0.138	0.042	**0.014**
aFCR	**0.066**	0.039	0.016

Table 2: SMAPE values.

Figures 1-6 show actual time series values, including the one used to develop the forecasting models, and the forecasted values given by the adaptive fuzzy c-regression approach in two cases. The first case concerns one step ahead forecast, denoted by aFCR-1 (red line), i.e., the forecasting models use actual values of the lagged time series values. The second case concerns forecasts from the current up to T steps ahead, denoted by aFCR-T (green line). As it can be seen, the aFCR approach gives accurate results for one step ahead forecasts using actual past time series values, but accuracy deteriorates when forecasting T steps ahead using forecasted values of the lagged values.

Another experiment was conducted to verify how the aFCR performs when the forecast horizon increases. For this purpose, time series ID 6 and ID 11 were selected once they show the highest forecast horizons, $T = 90$ and $T = 180$, respectively. The SMAPE values achieved by aFCR are shown in Table 3. In this case, aFCR attains higher accuracy in short term forecasts, roughly for forecast horizons from 1 up to 30 steps ahead. This demonstrates the potential of aFCR to deal with time series forecasting tasks. Similarly as in the previous experiments, past forecasts were used as values of the corresponding lagged inputs.

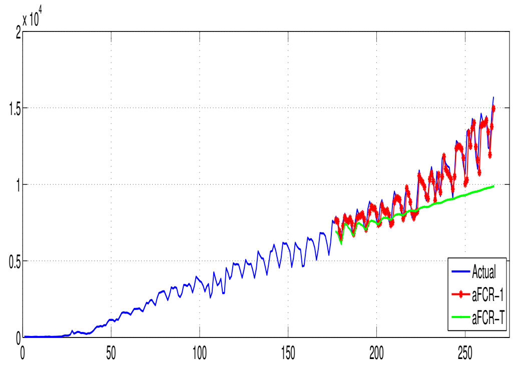

Figure 1: Time Series ID 6 - Actual and forecasted values for one step ahead (aFCR-1) and for T steps ahead (aFCR-T).

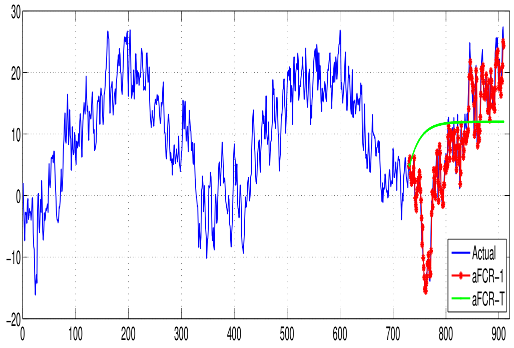

Figure 2: Time Series ID 11 - Actual and forecasted values for one step ahead (aFCR-1) and for T steps ahead (aFCR-T).

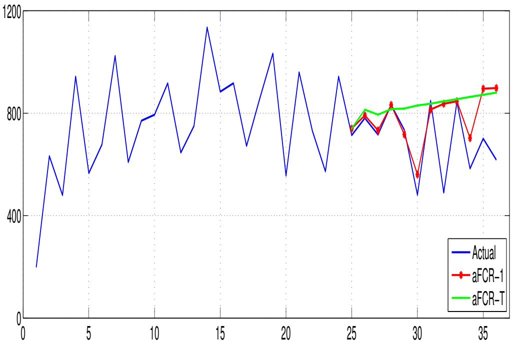

Figure 3: Time Series ID 37 - Actual and forecasted values for one step ahead (aFCR-1) and for T steps ahead (aFCR-T).

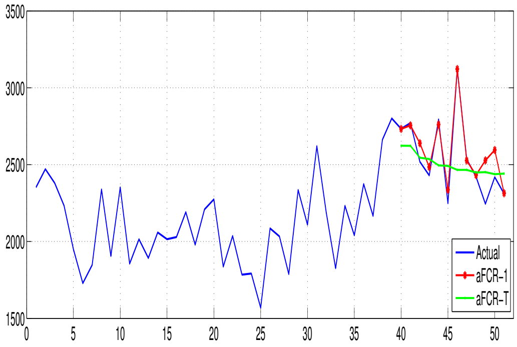

Figure 4: Time Series ID 40 - Actual and forecasted values for one step ahead (aFCR-1) and for T steps ahead (aFCR-T).

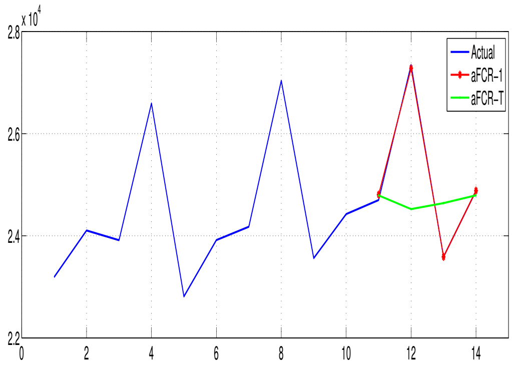

Figure 5: Time Series ID 74 - Actual and forecasted values for one step ahead (aFCR-1) and for T steps ahead (aFCR-T).

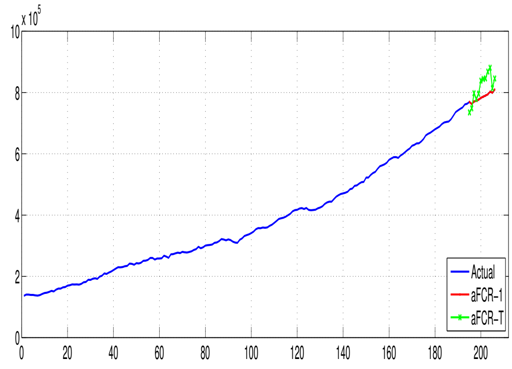

Figure 6: Time Series ID 87 - Actual and forecasted values for one step ahead (aFCR-1) and for T steps ahead (aFCR-T).

4. Conclusion

This paper has suggested an adaptive fuzzy c-regression modeling approach for time series modeling and forecasting. The approach is a clustering-based fuzzy modeling that develops fuzzy functional TS models. Clustering uses affine functions as prototypes, and function parameters are computed using the weighted least squares algorithm. Computations are done recursively and process samples of data sets sequentially. The idea is to adapt the model structure and parameters whenever required by new input data. Adaptation is based on participatory learning. The usefulness of the approach has been shown using selected data sets offered by the 2015 IFSA-EUSFLAT International Time Series Competition organizers. The results are encouraging when compared against ARIMA, SARIMA

T (steps ahead)	ID 6	ID 11
1	0.011	0.029
10	0.054	0.098
30	0.093	0.187
50	0.119	0.229
70	0.126	0.513
90	0.159	0.639
130	-	0.681
190	-	0.714

Table 3: SMAPE values of aFCR for different forecast horizons.

and alternative evolving fuzzy models such as DENFIS, OS-ELM, eTS, xTS, eTS+, ePL and eFT. Further work shall consider the generalization of the approach to handle clusters of distinct shapes, as well as the use of more sophisticated input selection methods instead of autocorrelation based criteria for time series forecasting.

Acknowledgement

The authors thank the anonymous referees for the comments and suggestions that help to improve the paper. This work was supported by the Brazilian Ministry of Education (CAPES), the Research Foundation of the State of São Paulo (FAPESP), and the Brazilian National Research Council (CNPq) grant 305906/2014-3.

References

[1] S. Makridakis, S. Wheelwright, and R. Hyndman. *Forecasting methods and applications.* John Wiley & Sons, 2008.

[2] A. K. Palit and D. Popovic. *Computational intelligence in time series forecasting: Theory and engineering applications.* Springer, Berlin, 2005.

[3] S. F. Crone, M. Hibon, and K. Nikolopoulos. Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction. *International Journal of Forecasting*, 27(3):635–660, 2011.

[4] L. Maciel and R. Ballini. Neural networks applied to stock market forecasting: An empirical analysis. *Learning and Nonlinear Models*, 8:3–22, 2010.

[5] P. Cortez, M. Rocha, and J. Neves. Evolving time series forecasting ARMA models. *Journal of Heuristics*, 10(4):415–429, 2004.

[6] K. Miiler, A. Smola, G. Ratsch, B. Scholkopf, J. Kohlmorgen, and V. Vapnik. Predicting time series with support vector machines. In *Proceedings of the 7th international conference on artificial neural networks*, pages 999–1004, 1997.

[7] I. Nunn and T. White. The application of antigenic search techniques to time series forecasting. In *Proceesings of The Genetic and Evolutionary Computation Conference (GECCO)*, pages 353–360, 2005.

[8] J. Aznarte, J. Benítez, and J. Castro. Smooth transition autoregressive models and fuzzy rule-based systems: Functional equivalence and consequences. *Fuzzy Sets and Systems*, 158:2734–2745, 2007.

[9] I. Luna and R. Ballini. Top-down strategies based on adaptive fuzzy rule-based systems for daily time series forecasting. *International Journal of Forecasting*, 27(3):708–724, 2011.

[10] N. K. Kasabov and Q. Song. DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *IEEE Transactions on Fuzzy Systems*, 10(2):144–154, 2002.

[11] J. Peralta, L. Xiaodong, G. Gutierrez, and A. Sanchis. Time series forecasting by evolving artificial neural networks using genetic algorithms and differential evolution. In *Proceedings of the International Joint Conference on Neural Networks*, 2010.

[12] P. Angelov. Evolving Takagi-Sugeno fuzzy systems from data streams (eTS+). In P. Angelov, D. Filev, and N. Kasabov, editors, *Evolving intelligent systems: Methodology and applications*, pages 21–50. Wiley & IEEE Press, Hoboken, NJ, USA, 2010.

[13] E. D. Lughofer. *Evolving fuzzy systems – Methodologies, advanced concepts and applications.* Springer-Verlag, Berlin Heidelberg, 2011.

[14] V. Kreinovich, G. C. Mouzouris, and H. T. Nguyen. Fuzzy rule based modeling as a universal approximation. In H. T. Nguyen and M. Sugenno, editors, *Fuzzy Systems: Modeling and Control*, pages 135–196. Kluwer Academic Publishers, 1998.

[15] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-15(1):116–132, 1985.

[16] S. L. Chiu. Fuzzy model identification based on cluster estimation. *Journal of Intelligent Fuzzy Systems*, 2:267–278, 1994.

[17] P. Angelov and D. Filev. An approach to online identification of Takagi-Sugeno fuzzy models. *IEEE Transactions on Systems Man and Cybernetics – Part B*, 34(1):484–498, 2004.

[18] P. Angelov and D. Filev. Simpl_eTS: A simplified method for learning evolving Takagi-Sugeno fuzzy models. *IEEE International Conference on Fuzzy Systems, Reno, Nevada, USA*, pages 1068–1073, 2005.

[19] P. Angelov and X. Zhou. Evolving fuzzy systems from data streams in real-time. *International Symposium on Evolving Fuzzy Systems, Ambleside, Lake District, United Kingdom*, pages 29–35, 2006.

[20] L. Maciel, F. Gomide, and R. Ballini. En-

hanced evolving participatory learning fuzzy modeling: An application for asset returns volatility forecasting. *Evolving Systems*, 5(1):75–88, 2014.

[21] E. Lima, H. Hell, R. Ballini, and F. Gomide. *Evolving Fuzzy Modeling Using Participatory Learning*. John Wiley & Sons, Ltd, 2010.

[22] D. Dovžan and I. Škrjanc. Recursive fuzzy c-means clustering for recursive fuzzy identification of time-varying processes. *ISA Transactions*, 50(2):159–169, 2011.

[23] D. Dovžan and I. Škrjanc. Recursive clustering based on a Gustafson-Kessel algorithm. *Evolving Systems*, 2(1):15–24, 2011.

[24] D. Dovzan, V. Loga, and I. Skrjanc. Solving the sales prediction with fuzzy evolving models. In *WCCI 2012 IEEE World Congress on Computational Intelligence*, pages 10–15, June, Brisbane, Australia, 2012.

[25] A. P. Lemos, W. Caminhas, and F. Gomide. Fuzzy evolving linear regression trees. *Evolving Systems*, 2:1–14, 2011.

[26] N. Liang, G. Huang, P. Saratchandran, and N. Sundararajan. A fast and accurate online sequential learning algorithm for feedforward networks. *IEEE Transactions on Neural Networks*, 17(6):1411–1423, 2006.

[27] E. D. Lughofer. FLEXFIS: A robust incremental learning approach for evolving takagi-sugeno fuzzy models. *IEEE Transactions on Fuzzy Systems*, 16(6):1393–1410, 2008.

[28] J. D. J. Rubio. SOFMLS: Online self-organizing fuzzy modified least-squares network. *IEEE Transactions on Fuzzy Systems*, 17(6):1296–1309, 2009.

[29] H. Rong, N. Sundarajan, G. Huang, and P. Saratchandran. Sequential adaptive fuzzy inference system (SAFIS) for nonlinear system identification and prediction. *Fuzzy Sets and Systems*, 157(9):1260–1275, 2006.

[30] K. Subraminan, A. K. Das, S. Sundaram, and S. Ramasamy. A meta-cognitive interval type-2 fuzzy inference system and its projection based learning algorithm. *Evolving Systems*, pages 1–12, 2013.

[31] K. Subraminan and S. Suresh. A meta-cognitive sequential learning algorithm for neuro-fuzzy inference system. *Applied Soft Computing*, 12(11):3603–3614, 2012.

[32] H. J. Rong, N. Sundarajan, G. Huang, and G. Zhao. Extended sequential adaptive fuzzy inference system for classification problems. *Evolving Systems*, 2:71–82, 2011.

[33] R. Hathaway and J. C. Bezdek. Switching regression models and fuzzy clustering. *IEEE Transactions on Fuzzy Systems*, 1(3):195–204, 1993.

[34] I. Škrjanc, D. Dovžan, and F. Gomide. Evolving fuzzy-model-based on c-regression clustering. *IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, pages 1–7, 2014.

[35] J. C. Bezdek. *Pattern recognition with fuzzy objective function algorithm*. Plenum, New York, 1981.

[36] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ, 1988.

[37] P. Angelov and X. Zhou. Evolving fuzzy-rule-based classifiers from data streams. *IEEE Transactions on Fuzzy Systems*, 16(6):1462–1475, 2008.