ICIC Express Letters
Volume **12**, Number **10**, October **2018**

ICIC International ©2018 ISSN 1881-803X
pp. 1025–1031

A MARGINAL LINEARIZATION FUZZY INFERENCE MODELING METHOD BASED ON ENSEMBLE LEARNING

Degang Wang[1], Wenyan Song[2] and Shining Ding[1]

[1]School of Control Science and Engineering
Dalian University of Technology
No. 2, Linggong Road, Dalian 116024, P. R. China
wangdg@dlut.edu.cn

[2]School of Economics
Dongbei University of Finance and Economics
No. 217, Jianshan Street, Dalian 116025, P. R. China

Received March 2018; accepted June 2018

Abstract. *In this paper, a fuzzy modeling method combined with marginal linearization and ensemble learning is proposed. Firstly, fuzzy classifier method is applied to obtaining the partition of domains and to establishing fuzzy rules. Then fuzzy marginal linearization modeling (FMLM) method is utilized to deduce the mathematical representation of fuzzy system. By choosing different membership functions of input variables, several fuzzy models with different structures can be designed. Further, ensemble learning method is considered to build comprehensive model. Some numerical examples are provided to show the validity of the proposed method.*

Keywords: Fuzzy system, Fuzzy inference modeling, Marginal linearization method

1. **Introduction.** It is well known that fuzzy modeling technology can effectively handle human knowledge and data information, and can establish mathematical models for some complex systems. During the past few decades, fuzzy systems are proved to have universal approximation capabilities for nonlinear system. Accordingly, various fuzzy modeling methods have been proposed and widely applied in regression [1] and prediction [2].

The mathematical mechanism of fuzzy modeling is to obtain the relationship between the input variable and output variable. Different from some traditional machine learning methods, fuzzy modeling technology can handle not only historical numerical data but also priori information from experts. In [3], Mamdani fuzzy systems are designed to be universal approximators for nonlinear functions. In [4], the approximation capabilities of TS fuzzy systems are discussed. In some applications, people hope that fuzzy systems possess smooth approximation property. That is to say that fuzzy system can approximate both the original function and its derivate function. To solve this problem, in [5], fuzzy systems determined by some polynomial functions are provided.

Besides, in order to facilitate theoretical analysis, some linearization treatments are considered to improve fuzzy systems. In [6], fuzzy marginal linearization technology is proposed. The basic idea of this method is to let some fuzzy membership functions in fuzzy rules be rectangle functions and others be triangular functions. As thus, fuzzy system can transfer expert's knowledge and observed data into mathematical model with piecewise linear form models or piecewise nonlinear structure models which can be suitable for different model analyzing problem.

From existing results, we can find that fuzzy modeling technology is an effective tool for handling both qualitative knowledge and quantitative information. However, in practical application, when the objective system exhibits complexity and nonlinearity, a single fuzzy

DOI: 10.24507/icicel.12.10.1025

model may not reflect the whole characteristic of the system. Hence, how to synthesize several fuzzy models to express complicated system is an interesting problem. Motivated by aforementioned fact, in this paper, an ensemble learning scheme is applied in designing fuzzy system. By marginal linearization method, empirical and numerical information can be transferred into some kind of mathematical sub-models with different structures. Ensemble learning method is considered to establish an integrated model with several sub-models for solving modeling issue.

This paper is organized as follows. In Section 2, some preliminary knowledge of fuzzy modeling technology is introduced. In Section 3, a kind of fuzzy modeling method combining marginal linearization and ensemble learning is designed. In Section 4, some numerical examples are considered to demonstrate the validity of the proposed method. Finally, some conclusions are summarized in Section 5.

2. Preliminaries of Fuzzy Modeling Method.

In this section, we will introduce some basic notations and some preliminaries of FMLM method.

We take an n-input and single output system as an example to introduce FMLM method. The input universe is $X_1 \times \cdots \times X_n$ and the output universe is Y, where $(x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n$ is the input variable and $y \in Y$ is the output variable. Consider a group of input-output samples $(x_{1j_1}, \ldots, x_{nj_n}, y_{j_1\ldots j_n})$, where $j_k = 1, \ldots, p_k$, $k = 1, \ldots, n$.

In order to simplify the numbers of fuzzy rules, we utilize fuzzy C-means (FCM) method to divide these observed samples into several classifiers. Corresponding to clustering prototypes, we can get a series sub-regions of the input domain, which can be denoted by $(i_1, \ldots, i_n) = \left[x_{1i_1}, x_{1(i_1+1)}\right) \times \cdots \times \left[x_{ni_n}, x_{n(i_n+1)}\right)$, where $i_k = 1, \ldots, q_k$, $q_k \le p_k$ and $k = 1, \ldots, n$. Based on these partition points of sub-regions and the clustering centers, a group of fuzzy rules can be established as below,

$$\text{If } x_1 \text{ is } A_{1i_1} \text{ and } x_2 \text{ is } A_{2i_2} \text{ and } \cdots \text{ and } x_n \text{ is } A_{ni_n} \text{ then } y \text{ is } B_{i_1\ldots i_n}, \tag{1}$$

where A_{ki_k} is the fuzzy set of universe X_k and $B_{i_1\ldots i_n}$ is the fuzzy set of universe Y.

Then, we will use FMLM and fuzzy rules (1) to deduce mathematical representation of the fuzzy system.

First, in each phase of modeling, we only let one premise fuzzy set of fuzzy rules be triangular membership function and let the other input fuzzy sets be rectangle-shaped membership function. Accordingly, we can use fuzzy inference modeling method to obtain the output of this model, which can be denoted by f_k,

$$f_k(x_1, \ldots, x_n) = \sum_{i_1=1}^{q_1} \cdots \sum_{i_n=1}^{q_n} \left(A_{ki_k}(x_k) \cdot y_{i_1\ldots i_n} + A_{k(i_k+1)}(x_k) \cdot y_{(i_1+1)\ldots(i_n+1)}\right).$$

Then, we take the average sum of f_k $(k = 1, \ldots, n)$. Hence the fuzzy system with piecewise linearization structure can be represented by

$$f(x_1, \ldots, x_n) = \frac{1}{n} \cdot \sum_{i_1=1}^{q_1} \cdots \sum_{i_n=1}^{q_n} \left(\sum_{k=1}^{n} \left(A_{ki_k}(x_k) \cdot y_{i_1\ldots i_n} + A_{k(i_k+1)}(x_k) \cdot y_{(i_1+1)\ldots(i_n+1)}\right)\right). \tag{2}$$

Similarly, if two fuzzy sets of input dimensions in each process are denoted by triangular membership functions and the other fuzzy sets are rectangle-shaped membership functions, then we can obtain a fuzzy system with bilinear structure.

Particularly, if all the membership functions of input variables are chosen as triangular membership functions, then the fuzzy system can be represented by the so-called Mamdani fuzzy system, i.e.,

$$f(x_1,\dots,x_n)=\sum_{i_1=1}^{q_1}\cdots\sum_{i_n=1}^{q_n}\left(\prod_{k=1}^{n}A_{ki_k}(x_k)\cdot y_{i_1\dots i_n}\right).$$

As described above, we utilize FMLM method to generate a series of fuzzy systems with different structures. And the famous Mamdani fuzzy system can be seen as a particular case of FMLM method.

3. **Fuzzy Modeling Method Combined with Marginal Linearization and Ensemble Learning.** In this section, we will investigate the usage of FMLM and ensemble learning technique to solve time series forecasting problem.

The basic structure of the proposed fuzzy modeling method is shown in Figure 1. The designed model includes two stages. The first stage is to use FMLM to obtain several sub-models with different structures. The second stage is to ensemble outputs of every sub-model and to get the final result of the entire model.

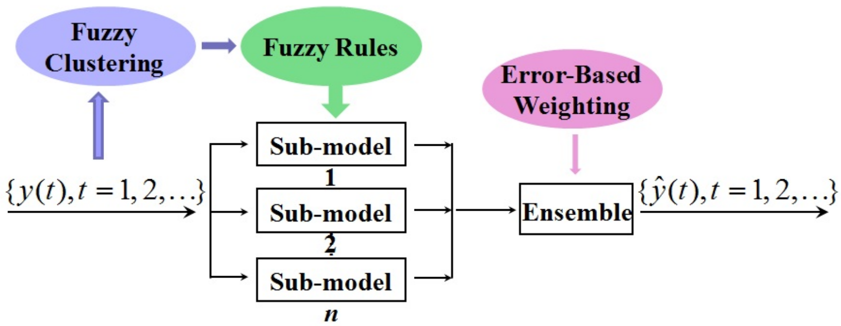

FIGURE 1. The basic structure of ensemble fuzzy modeling method

In the following, we will present the detailed process of this method.

Consider a time series: $\{y(t),t=1,2,\dots\}$. Without loss of generality, we take a dth-order time series as example to introduce the proposed procedure. In fact, a group of observed samples $\{(y(t-d+1),\dots,y(t),y(t+1)),\ t=d,d+1,\dots\}$ can be extracted from the original series, where $(y(t-d+1),\dots,y(t))$ is seen as the input variable, and $y(t+1)$ is the output variable hereinafter.

Firstly, FCM method is considered to divide the observed samples into K classifiers.

For each classifier, we take the cluster center as the peak point of fuzzy set, and take the minimum value and maximum value of this classifier as the left point and right point of corresponding fuzzy set respectively. Accordingly, we can establish a group of fuzzy rules as:

$$\text{If } y(t-d+1) \text{ is } A_{1i} \text{ and } \cdots \text{ and } y(t) \text{ is } A_{di} \text{ then } y(t+1) \text{ is } B_i, \tag{3}$$

where $i=1,\dots,K$.

As depicted in Equation (2), by FMLM we can obtain a fuzzy system with piecewise linear structure as follows,

$$\begin{aligned}y^{(1)}(t+1)&=f_1(y(t-d+1),\dots,y(t))\\&=\sum_{i=1}^{K}\frac{1}{d}\left(\sum_{j=1}^{d}\left(A_{ji}(y(t-d+j))\cdot y_i+A_{j(i+1)}(y(t-d+j))\cdot y_{i+1}\right)\right),\end{aligned}$$

where y_i is the peak point of fuzzy set B_i.

Similarly, based on fuzzy rules (3) with different membership functions, we can also determine other forms of sub-models denoted by f_l $(l=2,\dots,n)$.

Then, we will synthesize outputs of all the sub-models together. In this paper, two kinds of error based weights are considered in the ensemble learning stage [7,8]. They are defined as follows.

(1) Constant weights:

$$w_l = \frac{1/Err_l}{\sum_{l=1}^{n} 1/Err_l},$$

where Err_l is a kind of performance evaluation index of sub-model l for the training sample set.

For instance, Err_l can be settled as,

$$Err_l = \frac{1}{N}\sum_{t=1}^{N}\left(y(t) - y^{(l)}(t)\right)^2,$$

where N is the number of training data, and $y(t)$ and $y^{(l)}(t)$ are the desired output and the calculated output from the lth sub-model at time t. This Err_l is derived from the common used mean squared errors (MSE) accuracy index, and other types of Err_l can be similarly defined.

(2) Variable weights:

$$w_l(t+1) = \frac{1/Err_l(t)}{\sum_{l=1}^{n} 1/Err_l(t)},$$

where $Err_l(t)$ is the error of time t for sub-model l. In application, it can be designed as $Err_l(t) = \left(y(t) - y^{(l)}(t)\right)^2$ or other forms.

Obviously, for whether static weight or dynamic weight, a sub-model with more error receives less weight and vice versa. Hence, we can respectively compute the final output $\hat{y}(t+1)$ by the constant weights or variable weights:

$$\hat{y}(t+1) = \sum_{l=1}^{n} w_l \cdot y^{(l)}(t+1) = \sum_{l=1}^{n} w_l \cdot f_l(y(t-d+1), \ldots, y(t)),$$

$$\hat{y}(t+1) = \sum_{l=1}^{n} w_l(t+1) \cdot y^{(l)}(t+1) = \sum_{l=1}^{n} w_l(t+1) \cdot f_l(y(t-d+1), \ldots, y(t)).$$

In this way, we establish a fuzzy model by FMLM method and ensemble learning technique. The primary implementation steps of the proposed fuzzy modeling method are summarized as listed following.

Step 1. Use acquired knowledge and data information to classify samples.

Step 2. Determine partition points of input domains and output domains based on clustering results.

Step 3. Establish fuzzy membership functions of input variables and output variables and obtain fuzzy rule bases.

Step 4. With different shape of membership functions, use fuzzy marginal linearization method to deduce expressions of every sub-model.

Step 5. Take advantage of ensemble learning algorithm to integrate outputs of all the sub-models.

4. **Empirical Results.**

In this section, two datasets are provided to demonstrate the forecasting performance of the proposed method for time series.

Example 4.1. *Prediction of nonlinear dynamic system [9].*

This example is defined by the following nonlinear difference equation:

$$y(t+1) = \frac{y(t)y(t-1)[y(t)+2.5]}{1+y^2(t)+y^2(t-1)} + u(t),$$

where $y(0) = 0$, $y(1) = 0$ and $u(t) = \sin(2\pi t/25)$. It can been seen as a three-input single-output model, i.e., $y(t+1) = f(y(t-1), y(t), u(t))$. In order to compare with some literature, firstly 200 data sets are chosen as training data, while the other 200 data are chosen as testing data.

For the proposed method, the number of clusters is set to be 2, and 2 sub-models are applied in this simulation. One of the sub-models has piecewise linear form, and the other one has the piecewise quadratic bilinear structure. Hence, there are 4 rules and 30 parameters in the ensemble model in total. The root mean squared error (RMSE) comparisons for training set and testing set are listed in Table 1 and Table 2 respectively.

Table 1. Comparison between the proposed method and other algorithms for training set

Method	No. of rules	No. of parameters	RMSE
OLS [10]	65	326	0.0288
RBFAFS [11]	35	280	0.1384
DFNN [12]	6	48	0.0283
GDFNN [13]	6	48	0.0241
Proposed method	4	30	0.0014

Table 2. Comparison between the proposed method and other algorithms for testing set

Method	No. of rules	No. of parameters	RMSE
Farag's model [14]	75	48	0.201
SOFNN [15]	5	46	0.0151
SOFNN-GA [9]	4	34	0.0146
SOFNN-GAPSO [9]	4	34	0.0141
Proposed method	4	30	0.0014

In fact, in the process of training two sub-models based on marginal linearization method already have good results. RMSE of simple arithmetic average between aforementioned linear sub-model and the other quadratic sub-models is 0.0125. With error based ensemble strategy, we can get the improved performance. If the number of clusters is set to be 3, then the RMSE of testing set can be dramatically reduced to 7.8811e-12, where the number of estimated parameters is only 44.

We have also compared prediction accuracy of the proposed method to other references. While 500 data sets are chosen as training data, and the other 500 data are chosen as the testing data, then the results of the MSE are presented in Table 3.

Table 3. Comparison between proposed method and other algorithms (MSE)

Method	No. of rules	No. of parameters	MSE	
			Training data	Testing data
Bagis [16]	4	104	0.0341	0.0378
Li et al. [17]	4	40	0.0149	0.0115
Li et al. [18]	4	40	0.0102	0.0128
Proposed method	4	30	0.0002	0.0002

Example 4.2. *Prediction of river flow time series.*

This time series contains the monthly flow of the clear water river at Kamiah, Idaho, USA from 1911 to 1965. And the first 500 data are used as training set while the last 100 data are chosen as the testing set. The raw series is available at http://DataMarket.com.

In the simulation, 3 clusters are chosen to set up a third order lag model for prediction. The forms of sub-models are the same as in the previous example. If static weights are adopted, which are determined by the overall error of the training set for each sub-model, the mean absolute error (MAE) and the MSE of testing set are shown in Table 4. When dynamic weights are applied, the MAE and the MSE can be improved further, which means that for each time, its ensemble weights of sub-models are calculated according to the error of the first order time delay output in the ensemble algorithm.

Table 4. Comparison between the proposed method and other algorithms for river flow testing set

Error measures	Individual models		Combination methods	
	FANN [21]	EANN [20]	Avg. of FANN&EANN	[19]
MAE	0.66	1.036	0.751	0.638
MSE	1.217	2.19	1.158	0.978
	Individual models		Combination methods	
	Sub-model 1	Sub-model 2	Static Weights	Dynamic Weights
MAE	0.1466	0.1404	0.1398	0.1387
MSE	0.0348	0.0327	0.0325	0.0318

For this river flow time series, the proposed method can get more satisfactory results than other methods.

5. **Conclusions.** In this paper, a kind of fuzzy modeling method is proposed for forecasting time series. By marginal linearization scheme, fuzzy models with different structures can be established. Further, constant weight synthetic and variable weight synthetic are applied to calculating the forecasting value of the fuzzy model respectively. In the future research, it could be beneficial to discuss the parameters identification of fuzzy models for improving the forecasting accuracy of time series.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (61773088, 71571035).

REFERENCES

[1] Y. Zhang, Interval-valued variable weighted synthesis inference method for fuzzy reasoning and fuzzy systems, *ICIC Express Letters, Part B: Applications*, vol.4, no.3, pp.511-518, 2013.
[2] R. J. Dong and W. Pedrycz, A granular time series approach to long-term forecasting and trend forecasting, *Physica A*, vol.387, pp.3253-3270, 2008.
[3] X. J. Zeng and M. G. Singh, Approximation theory of fuzzy systems-MIMO case, *IEEE Trans. Fuzzy Systems*, vol.3, no.2, pp.219-235, 1995.
[4] L. Wu, X. Su, P. Shi and J. Qiu, Model approximation for discrete-time state-delay systems in the T-S fuzzy framework, *IEEE Trans. Fuzzy Systems*, vol.19, no.2, pp.366-378, 2011.
[5] H. X. Li, X. H. Yuan and J. Y. Wang, The normal numbers of the fuzzy systems and their classes, *Science in China (Series F)*, vol.53, no.11, pp.2215-2229, 2010.
[6] H. X. Li, J. Y. Wang and Z. H. Miao, Marginal linearization method in modeling on fuzzy control systems, *Progress in National Sciences*, vol.13, no.7, pp.489-496, 2003.
[7] C. Lemke and B. Gabrys, Meta-learning for time series forecasting and forecast combination, *Neurocomputing*, vol.73, pp.2006-2016, 2010.

[8] R. Adhikari and R. K. Agrawal, Performance evaluation of weight selection schemes for linear combination of multiple forecasts, *Artificial Intelligence Review*, pp.1-20, 2012.

[9] O. Khayat, M. M. Ebadzadeh, H. R. Shahdoosti, R. Rajaei and I. Khajehnasiri, A novel hybrid algorithm for creating self-organizing fuzzy neural networks, *Neurocomputing*, vol.73, pp.517-524, 2009.

[10] S. Chen, C. F. N. Cowan and P. M. Grant, Orthogonal least squares learning algorithm for radial basis function network, *IEEE Trans. Neural Networks*, vol.2, pp.302-309, 1991.

[11] K. B. Cho and B. H. Wang, Radial basis function based adaptive fuzzy systems and their applications to identification and prediction, *Fuzzy Sets and Systems*, vol.83, pp.325-339, 1996.

[12] S. Wu and M. J. Er, Dynamic fuzzy neural networks – A novel approach to function approximation, *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*, vol.30, no.2, pp.358-364, 2000.

[13] S. Wu, M. J. Er and Y. Gao, A fast approach for automatic generation of fuzzy rules by generalized dynamic fuzzy neural networks, *IEEE Trans. Fuzzy Systems*, vol.9, pp.578-594, 2001.

[14] W. A. Farag, V. H. Quintana and G. Lambert-Torres, A genetic-based neuro-fuzzy approach for modeling and control of dynamical systems, *IEEE Trans. Neural Networks*, vol.9, no.5, pp.756-767, 1998.

[15] C. T. Lin and C. S. G. Lee, *Neural Fuzzy Systems: A Neural Fuzzy Synergism to Intelligent Systems*, Prentice Hall, 1996.

[16] A. Bagis, Fuzzy rule base design using tabu search algorithm for nonlinear system modeling, *ISA Transactions*, vol.47, no.1, pp.32-44, 2008.

[17] C. Li, J. Zhou, X. Xiang, Q. Li and X. An, T-S fuzzy model identification based on a novel fuzzy c-regression model clustering algorithm, *Engineering Applications of Artificial Intelligence*, vol.22, pp.646-653, 2009.

[18] C. Li, J. Zhou, B. Fu, P. Kou and J. Xiao, T-S fuzzy model identification with a gravitational search-based hyper plane clustering algorithm, *IEEE Trans. Fuzzy Systems*, vol.20, no.2, pp.305-317, 2012.

[19] R. Adhikari, A neural network based linear ensemble framework for time series forecasting, *Neurocomputing*, vol.157, pp.231-242, 2015.

[20] C. P. Lim and W. Y. Goh, The application of an ensemble of boosted Elman networks to time series prediction: A benchmark study, *International Journal of Computational Intelligence*, vol.3, no.2, pp.119-126, 2005.

[21] G. Zhang, B. E. Patuwo and M. Y. Hu, Forecasting with artificial neural networks: The state of the art, *International Journal of Forecasting*, vol.14, no.1, pp.35-62, 1998.