Neurocomputing 151 (2015) 1293–1304

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Short Communication

Improving Wang–Mendel method performance in fuzzy rules generation using the fuzzy C-means clustering algorithm

Jin Gou*, Feng Hou, Wenyu Chen, Cheng Wang, Wei Luo

College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China

ARTICLE INFO

Article history:
Received 26 April 2014
Received in revised form
24 August 2014
Accepted 27 October 2014
Communicated by M.-J. Er
Available online 11 November 2014

Keywords:
Fuzzy system
Fuzzy C-means clustering algorithm
Wang–Mendel method
Completeness
Robustness

ABSTRACT

The generation of fuzzy rules from samples for fuzzy modeling and control is significant. If samples contain noise and outliers, the Wang–Mendel (WM) method may lead to the extraction of invalid rules resulting in low confidence of the rules. The scale of the samples also affects the efficiency of the WM method. Interaction among input variables can help the WM method achieve high completeness and robustness. The fuzzy C-means clustering (FCM) algorithm can reduce the scale of samples and undo noisy data to some degree. This paper aims to develop an FCM-based improved WM method that adopts a modified FCM algorithm to preprocess the original samples and compute the interaction among the samples. Then, the optimized samples are used to generate fuzzy rules, thereby building a complete rule set through extrapolation. Experimental results from two nonlinear functions and short-term load forecasting case study show that the proposed method not only has high completeness and robustness, but also ensures better prediction accuracy of the fuzzy system.

1. Introduction

In recent decades various fuzzy systems based on fuzzy IF–THEN rules have been developed and applied in many different fields. These fuzzy rule systems are used for classification, control, and function approximation purposes. Fuzzy rules play a very important role in fuzzy systems, with their performance determining the capabilities of the fuzzy system. The key to building a fuzzy system is to extract fuzzy rules from samples. There are many methods for extracting fuzzy rules, such as heuristic methods, and methods based on genetic algorithms, fuzzy clustering, and neural networks, among others [1–5]. However, these methods either require continual repeated learning or are hard to implement owing to complex mechanisms. The Wang–Mendel (WM) method, proposed by Wang and Mendel, can generate fuzzy rules efficiently from sample data [6,7]. The method divides the input space into fuzzy regions, and then uses a lookup table to extract rules for each fuzzy subspace. The WM method is simple, yet useful, and does not continually need repeated learning. The WM method has become a classic method for fuzzy rule generation and has been widely applied to a variety of problems [8–12].

Fuzzy rule bases have been proven to be effective tools for modeling complex systems and approximating functions. A better fuzzy rule base generated from samples can improve the accuracy of the fuzzy system [13]. As a result, the dependence of the WM method on samples is relatively high. If the sample set is incorrect, so as is the fuzzy rule base extracted by the WM method [14]. Thus, the extracted fuzzy rule base does not have high completeness and robustness and the accuracy of the fuzzy system built using the fuzzy rule base will be poor [7,9]. All modeling approaches have drawbacks with respect to accuracy or interpretability. Galende-Hernandez et al. [15] established some works to show that a reduction in the complexity of a system can mean better interpretability of the fuzzy system. They proposed improving a fuzzy model by complexity reduction based on an accuracy-interpret-ability trade-off, implemented using a bi-objective (accuracy and interpretability) genetic approach via rule selection and a simple set of indices relating to accuracy-interpretability. Since the WM method generates rules from each sample, if the scale of samples increases, the complexity of the WM method also increases and the efficiency of the method declines sharply. The sample set therefore plays a decisive role in many of the problems associated with the WM method, such as completeness, robustness, and efficiency, among others. Research has shown that there are interactions among the samples. In general, the theoretical output variables become closer as the level of similarity among the input variables increases [16,17].

The fuzzy C-means clustering (FCM) algorithm involves unsupervised learning [18–21]. The FCM algorithm can learn the distribution of the original samples and form a new sample set with a smaller scale but good performance under the condition that it does not change the distribution of the original samples [22]. FCM can

* Corresponding author. Tel.: +86 13960396027.
E-mail addresses: goujin@gmail.com (J. Gou), houfeng0205@gmail.com (F. Hou), huiyu100@hqu.edn.cn (W. Chen), wangcheng@hqu.edu.cn (C. Wang), luowei.qinghua@gmail.com (W. Luo).

http://dx.doi.org/10.1016/j.neucom.2014.10.077
0925-2312/

also reduce the interference caused by noisy data in the original samples to a larger extent thereby improving the robustness of the fuzzy rule base [23].

Thus, an interesting goal would be to obtain a fuzzy model with adequate accuracy and a good level of explanation [24,25]. To ensure completeness and robustness of the fuzzy rule base, this paper presents a novel fuzzy rule generation method, that is, a combined modeling method based on a fuzzy system and clustering algorithm. The proposed method uses a modified FCM (AFCM) algorithm to preprocess the sample set and calculate the interaction among samples. Then, through an improved WM method, fuzzy rules are extracted systematically. Experimental results from two nonlinear functions and short-term load forecasting case study show that the proposed method not only has high completeness and robustness, but also ensures better prediction accuracy of the fuzzy system.

The rest of the paper is organized as follows. Section 2 describes the theory and details of the FCM and AFCM, the improved FCM algorithm. In Section 3, a FCM-based improved WM method is presented. This method adopts the AFCM algorithm to optimize the original sample set and calculate a newly defined interaction variable among the samples. Therefore, a complete and robust fuzzy rule base can be obtained through extraction. In Section 4, the prediction model based on a novel method is systematically described. Two nonlinear functions and a real case study are used to validate the proposed model in Section 5. Finally, Section 6 summarizes the research and presents the conclusions.

2. Fuzzy C-means clustering (FCM) algorithm

The fuzzy clustering algorithm, which effectively clusters the samples, can be used to identify inherent laws in a large number of samples. It is based on a foundation of fuzzy membership and differs from the hard-clustering algorithm. In contrast with other clustering algorithms, the FCM algorithm is better grounded on theory and has a more mature application. The purpose of fuzzy processing is to find the distribution of the samples. Through application of the FCM algorithm, the distribution of samples is clearly shown, while an optimized sample set is obtained at the same time.

2.1. Fundamentals of FCM clustering

Ruspini [18] first proposed the concept of a fuzzy partition in 1969. Based on this, researchers began using fuzzy set methods to solve clustering problems, which became known as fuzzy cluster analysis. Dunn [19] extended the concept of a fuzzy partition to a weighted within-groups sum of squared error function, and this was later improved to an infinite cluster of weighted within-groups sum of squared error by Bezdek [20]. Eventually this became the universal clustering criterion for the FCM clustering algorithm, which is a very effective fuzzy clustering algorithm. FCM clustering takes advantage of membership to express the relationships between data instances in samples and a cluster. Even if the variable is hard to classify, the FCM algorithm can realize a satisfactory relative clustering effect. Since FCM is simple in design, has a wide-ranging problem-solving ability, and is easily implemented on a computer, it has been applied in many fields.

Given a set to classify, $X=\{x_1,x_2,,x_N\}$, where each object x_k has n characteristic parameters, $x_k=\{x_{1k},x_{2k},\ldots,x_{nk}\}^T$. If set X can be divided into c classes, each class corresponds to a $c\times N$ Boolean matrix, $U=[u_{ij}]_{c\times N}$. The fuzzy partition space is denoted as

$$M_{fc}=\left\{U\subset R^{cN}|u_{ij}\in[0,1],\forall i,\forall j;\ \sum_{j=1}^{c}u_{ij}=1,\forall j;0<\sum_{i=1}^{N}u_{ij},\forall j\right\} \quad (1)$$

In addition, the basic idea of the FCM algorithm is to find a fuzzy partition matrix $U=[u_{ij}]_{c\times N}$ and c cluster centers $V=(v_1,v_2,\ldots,v_c)$, such that

$$J_m=\min\sum_{i=1}^{N}\sum_{j=1}^{c}(u_{ij})^m d^2(x_i,v_j) \quad (2)$$

where $\sum_{j=1}^{c}u_{ij}=1,\forall i;0<\sum_{i=1}^{N}u_{ij}\le N,0\le u_{ij},i=1,2,\ldots,N,j=1,2,\ldots,c$, $m\in[1,+\infty]$ is called a fuzzy weighted index, u_{ij} is the degree of membership of x_i in cluster j, v_j is the d-dimensional center of cluster j, and $d(x_i,v_j)$ is the distance between measured data x_i and the center v_j.

Fuzzy partitioning is carried out by iterative optimization of the objective function as shown above, with the update of membership u_{ij} and cluster centers c_j given by

$$v_i^{(b)}=\left[\sum_{i=1}^{N}(u_{ij}^{(b)})^m\cdot x_j\right]\Big/\left[\sum_{i=1}^{N}(u_{ij}^{(b)})^m\right],\quad i=1,2,\ldots,c \quad (3)$$

$$u_{ij}^{(b+1)}=\left[\sum_{k=1}^{c}(d_{ij}^{(b+1)}/d_{kj}^{(b+1)})^{2/(m-1)}\right]^{-1},\quad k=1,2,\ldots,c \quad (4)$$

where b is the number of iterations.

Details of the FCM algorithm are given in the following steps:

Step 1: Initialization. Set the number of cluster categories c, fuzzy weighted index m, iteration termination threshold value ε, and maximum number of iterations b_{max}, where $b=0,1,2,\ldots,b_{max}$. Set counter $b=0$ and initialize the membership matrix.

Step 2: Using Eq. (3), update the fuzzy cluster centers using the membership matrix $U^{(b)}$.

Step 3: Using Eq. (4), calculate or update the membership matrix $U^{(b+1)}$ according to the centers $V^{(b)}$.

Step 4: If $\|V^{(b+1)}-V^{(b)}\|\le\varepsilon$, the system is in a stable state. Terminate the iteration. At this time, obtain centers V and the fuzzy C-means partition matrix U. Otherwise, let $b=b+1$ and $U^{(b+1)}=U^{(b)}$, and return to step 2.

2.2. Modified FCM algorithm (AFCM)

The definition of membership degree in Eq. (4) may cause some serious problems in the presence of noisy observations. That is, FCM is a type of partitioning algorithm that divides the observations into c partitions regardless of whether they are noisy. However, it is more natural for noise points to have a very low membership degree in any cluster [26]. Most clustering methods, even those with feature weighting extensions, consider all samples to have an equal weight during the clustering process. However, it is not prudent to assume that every sample in a dataset has the same weight in cluster analysis. For instance, outliers should have less impact on the clustering results than other regular samples. Therefore, it would be very useful to exploit an applicable sample-weighted function in cluster analysis [27].

In practice, different features may have different contributions to the cluster analysis, and thus, certain important features of samples may not be found by the traditional FCM algorithm. Moreover, the contribution of noisy data will be enhanced. Because affinity is used to indicate the degree of affinity between a sample and the sample set, a sample weighted FCM algorithm with affinity, called the AFCM, is introduced to solve the above problems. Entropy is an important measure of the value of information. Higher levels of entropy can transmit more information. Taking advantage of the definition of entropy, AFCM applies E_i in Eq. (5) to measure the efficiency of u_{ij}:

$$E_i=-\sum_{j=1}^{c}u_{ij}\log_2 u_{ij} \quad (5)$$

where c is the number of classes, and u_{ij} is the degree of membership of x_i in cluster j.

Thus, the affinity of x_i, aff_i, can be computed as

$$aff_i = \frac{1-E_i}{N-\sum_{i=1}^{N} E_i} \tag{6}$$

where $\sum_{i=1}^{c} aff_i = 1$. Therefore, the objective function of AFCM is defined as

$$J_m = \min \sum_{i=1}^{N} \sum_{j=1}^{c} aff_i (u_{ij})^m d^2(x_i, v_j) \tag{7}$$

where u_{ij} is the degree of membership of x_i in cluster j, and $d(x_i, v_j)$ is the distance between measured data x_i and the center v_j.

Furthermore, the cluster centers c_j are updated by

$$v_i^{(b)} = \left[\sum_{i=1}^{N} aff_i (u_{ij}^{(b)})^m \cdot x_j \right] \Bigg/ \left[\sum_{i=1}^{N} aff_i (u_{ij}^{(b)})^m \right], \quad i = 1, 2, \ldots, c \tag{8}$$

where u_{ij} is the degree of membership of x_i in cluster j, and aff_i is the affinity of x_i.

Therefore, Eqs. (7) and (8) are applied to the AFCM algorithm. In addition, Eqs. (5) and (6) are used in step 2 of the AFCM algorithm to update the degree of affinity, which is considered to be equal in the initialization step.

A number of previous studies have investigated the convergence properties of FCM clustering [28–30]. The sample-weighted algorithm is not influenced by outliers, while its superiority and effectiveness have been demonstrated [27]. In this paper, a modified FCM (AFCM) algorithm is proposed. The AFCM algorithm is able not only to determine the sample weights automatically, but also to decrease the impact of the initialization on the clustering results during the clustering process.

3. Improved Wang–Mendel method based on FCM algorithm (FCAWM method)

Research on the WM method and other relevant methods has shown that if the samples are incomplete or contain noisy data, the fuzzy rule base generated by these methods may be incomplete or lack robustness [9]. There are interactions among the input variables [9,16,31,32], which can effectively be expressed by the degree of affinity. The degree of affinity among all the samples can reduce the interference caused by noisy data. Therefore, the proposed method takes advantage of the affinity to improve the completeness and robustness of the original WM method. Moreover, the accuracy of the fuzzy system generated by the WM method can be enhanced at the same time.

3.1. Problem description

There is a set of n input and single output data pairs:

$$T = \{(x^{(i)}, y^{(i)})\}, \quad i = 1, 2, \ldots, N \tag{9}$$

where $x^{(i)} = (x_1^{(i)}, \ldots, x_n^{(i)}) \in R_n$, $y^{(i)} \in R$, $x^{(i)}$ is the i-th element of the n-dimensional input, and $y^{(i)}$ is the output.

The basic problem is to extract rules describing how the output variable is influenced by the n input variables. The generated rules can take different forms. In this paper, the extracted fuzzy IF–THEN rules are expressed as follows:

$$\text{IF } x_{i1} \text{ is } A_{i1}^{(l)} \text{ and } \ldots \text{ and } x_{im} \text{ is } A_{im}^{(l)} \quad \text{THEN } y \text{ is } B^{(l)} \tag{10}$$

where $A_{ij}^{(l)}$ and $B^{(l)}$ are fuzzy sets defined on R, $l = 1, 2, \ldots, M$ is the index of a rule, and $(i1, \ldots, im)$ with $m \le n$ is a subsequence of $(1, \ldots, n)$.

Problem 1. Completeness: If the inputs fall into all the fuzzy subsets, the generated fuzzy rule base is complete.

Specifically, this paper defines M_i fuzzy sets for each input variable x_i $(i = 1, 2, \ldots, n)$ and K fuzzy sets for output variable y. Thus, $L = \prod_{i=1}^{n} (M_i)$, $A_1 = \{A_1^1, A_1^2, \ldots, A_1^{M_1}\}, \ldots, A_n = \{A_n^1, A_n^2, \ldots, A_n^{M_n}\}$ and $B = \{B^1, B^2, \ldots, B^K\}$. The complete fuzzy rule base corresponds to an $L \times (n+1)$ matrix, $V = [v_{ij}]_{L \times (n+1)}$:

$$M_{cb} = \left\{ V \,|\, v_{ij} \in A_j, \forall i, j \in [1, n]; v_{ij} \in B, \forall i, j = n+1; \bigcup_{i=1}^{L} \{v_{ij}\} = A_j, \forall j \right\} \tag{11}$$

Given a set of single input and single output data samples:

$$((x^{(1)}, y^{(1)}), \ldots, (x^{(4)}, y^{(4)})) \tag{12}$$

According to Fig. 1, input variable x is divided into five fuzzy subspaces, as is output variable y. Since the WM method can extract at most four rules from the sample set, there is no rule in the other fuzzy subsets.

Some conclusions can be drawn from the above example. Specifically, this paper defines M_i fuzzy sets for each input variable x_i $(i = 1, 2, \ldots, n)$. (1) If $N < \prod_{i=1}^{n} (M_i)$, the rule base is incomplete because of the incompleteness of the sample set, where N is the number of samples. (2) If $N \ge \prod_{i=1}^{n} (M_i)$ and the sample set is unevenly distributed, the rule base constructed by the WM method may be incomplete. Since some sample data are located

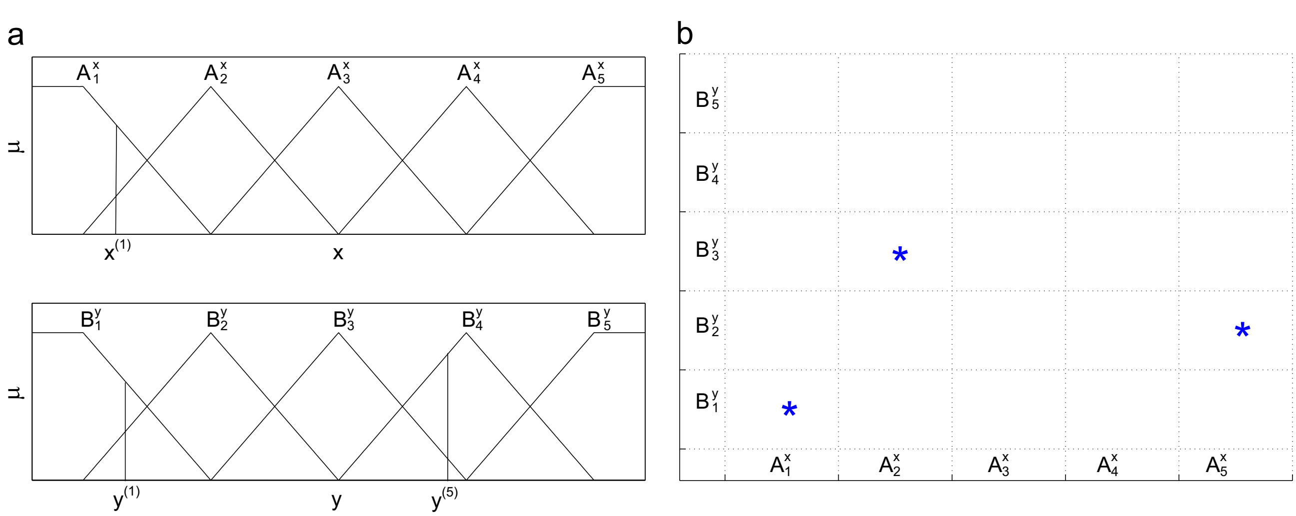

Fig. 1. Fuzzy regions of the input–output space and fuzzy rule base. (a) Fuzzy regions of input–output variables, (b) fuzzy rule base.

in the same fuzzy subspace, contradictory rules will be extracted, and the rule with the higher degree of confidence is used by the WM method. Therefore, when the sample set is evenly distributed and $N \geq \prod_{i=1}^{n}(M_i)$, the fuzzy rule base generated by the WM method has high completeness.

Problem 2. Robustness: If the sample set contains noisy data, the generated fuzzy system still has good approximation performance.

$$MAC_{\phi,\varphi} = \frac{(\phi^T \varphi)^2}{(\phi^T \phi)(\varphi^T \varphi)} \tag{13}$$

where ϕ is the generated fuzzy system, φ represents the actual model, ϕ^T and φ^T denote, respectively, the transpositions of ϕ and φ, $(\phi^T \varphi)$ represents the inner product of the two vectors, and $0 \leq MAC_{\phi,\varphi} \leq 1$. As the value approaches 1, the accuracy of the constructed fuzzy system improves.

Then, a noisy sample $(x^{(5)}, y^{(5)})$ is added into the above sample set, and assuming $x^{(5)} = x^{(1)}$, $y^{(5)} \neq y^{(1)}$. Therefore, the WM method can extract rules from the samples:

$$\text{Rule}^{(1)}: \text{IF } x \text{ is } A_1^x \quad \text{THEN } y \text{ is } B_1^y \tag{14}$$

$$\text{Rule}^{(5)}: \text{IF } x \text{ is } A_1^x \quad \text{THEN } y \text{ is } B_4^y \tag{15}$$

Since $x^{(5)} = x^{(1)}$, $y^{(5)} \neq y^{(1)}$, and the rules given above are ambiguous. If the WM method chooses Rule$^{(5)}$ with a higher degree of confidence, this will lead to the extraction of the wrong rule from noisy data. In conclusion, because the degree of support is not sufficient, it is preferable to use a single sample to generate a fuzzy rule.

Currently, the combined modeling method based on a fuzzy system and cluster algorithm is novel research attracting much attention. As a classic algorithm, FCM is remarkably beneficial for data preprocessing. In practical application, improving the performance of the WM method is an important issue. This paper attempts to address this.

3.2. Extrapolating fuzzy rules using the affinity among samples

The original WM method uses a single sample to extract rules, and thus, lacks high completeness and robustness. The sample data are not isolated, instead, interactions exist among samples [31,32]. Usually, the more similar the input variables, the closer are the corresponding theoretical output variables. Therefore, using the improved WM method, the output variables of the current rule can be calculated using all the samples. Sample data that are a better match with the input variables of the current rule have a greater impact on the process of extracting this rule [17,23].

Since the degree of affinity, *aff* in Eq. (6), can effectively express the contribution of the sample to the dataset, the FCAWM method uses *aff* to improve the completeness and robustness of the WM method. When the FCAWM method calculates the confidence of generating a fuzzy rule, it performs a weighted average operation for each sample datum with the help of the affinity among samples, *aff*. Then, in the processing step that computes the output variables, the FCAWM method selects *aff* again to perform a weighted average operation. Therefore, all the samples form the output variables of each rule. Even if the samples contain noisy data, by leveraging the effective shield of other correct data, the noisy data have relatively little influence on the output of the current rule, thereby improving robustness. Because all the fuzzy regions of the input variables are traversed, the generated fuzzy rule base can cover all the fuzzy subsets of the input variables. This kind of fuzzy rule base has high completeness.

The steps in the FCAWM are given as follows:

Step 1: Divide each input space and output space into fuzzy regions and compute the membership values. For each sample $(x^{(i)}; y^{(i)}), i = 1, 2, \ldots, N$ compute:

$$\omega^{(p)} = \prod_{j=1}^{m} \mu_{A_{ij}}(x_{ij}^{(p)}) \tag{16}$$

where the A_{ij} are the fuzzy sets in Eq. (10). If $\sum_{p=1}^{N} \omega^{(p)} = 0$, then no rule will be generated and the method terminates. Otherwise, $\omega^{(p)}$ is considered to be the weight of $y^{(p)}$.

Step 2: For each rule, compute

$$y_c^l = \frac{\sum_{i=1}^{N} y_c^{(i)} doc^{(i)} aff^{(i)}}{\sum_{i=1}^{N} doc^{(i)} aff^{(i)}} \tag{17}$$

where N is the number of samples, l is the current rule number, *doc* is the sample's degree of confidence, and *aff* is the degree of affinity among the samples. The fuzzy rules are generated as

$$\text{IF } x_{i1} \text{ is } A_{i1}^{(l)} \text{ and } \ldots \text{ and } x_{im} \text{ is } A_{im}^{(l)} \quad \text{THEN } y \text{ is } B^{(l)} \tag{18}$$

where $B^{(l)}$ is a fuzzy set $\mu_{B_{(l)}} = \mu(y; y_c^{(l)}, \sigma^{(l)})$, $y_c^{(l)}$ is given in Eq. (17) and $\sigma^{(l)}$ is computed as

$$\sigma^{(r)} = \frac{\sum_{k=1}^{n} |y^{(p_k^r)} - au^{(r)}| u^{(p_k^r)}}{\sum_{k=1}^{n} u^{(p_k^r)}} \tag{19}$$

$$u^{(r)} = \prod_{j=1}^{m} \mu_{A_{ij}}(x_{ij}^{(r)}) \tag{20}$$

$$au^{(r)} = \frac{\sum_{i=1}^{n} y^{(r)} u^{(r)} aff^{(r)}}{\sum_{i=1}^{n} u^{(r)} aff^{(r)}} \tag{21}$$

Step 3: Compute the degree of confidence of the extrapolation rule (18) as

$$doc^{(r)} = 1 - \frac{\sigma^{(r)}}{\max_{k,t=1}^{n} |y^{(p_k^r)} - y^{(p_t^r)}|} \tag{22}$$

Step 4: View the extrapolation rules in step 2 in the same way as the data-generated rules, and go to step 2 to repeat the process until all the extrapolation rules have been extracted.

4. Prediction model based on FCAWM method

The basic structure of a fuzzy system consists of a fuzzy inference engine, fuzzy rule base, fuzzifier, and defuzzifier [6,7]. The fuzzy rule base consists of a collection of fuzzy IF–THEN rules in the form of Eq. (10), extracted by the FCAWM method. The fuzzy system is a function $f(x)$ from R^m to R from an input–output point of view, and the function is used as a forecasting model with data and rule mining functions.

In this study, we used the continuous fuzzy prediction model to create a fuzzy system. We employed the center-averaged defuzzifier:

$$y = \frac{\sum_{l=1}^{M} y_c^l \cdot \omega^{(l)}}{\sum_{l=1}^{M} \omega^{(l)}}, \tag{23}$$

and obtained the following fuzzy system:

$$act(l) = \prod_{j=1}^{n} \mu_{A_{ij}^{(l)}}(x_{ij}), \tag{24}$$

$$f(x) = \frac{\sum_{l=1}^{M} y_c^l \cdot act(l)}{\sum_{l=1}^{M} act(l)}, \tag{25}$$

where y_c^l is the center of $B^{(l)}$, and $act(l)$ is the degree of activation of the l-th rule.

According to the theory of the WM method, fuzzy rules are generated from the samples. However, if the scale of the samples is too large, the number of rules is also large, making the computational model more complex. Moreover, noisy data may result in invalid rules [6]. The FCM algorithm can effectively learn the distribution of the samples, and can also decrease the interference from noisy data. At the same time, the algorithm can de-noise noisy data to some degree [33]. Using the optimized solution, updated samples are obtained. These optimized samples have higher confidence, less noise, and relatively higher completeness. Then, the improved WM method based on affinity is applied to extract fuzzy rules from the new samples. Because of the FCM algorithm, the novel method further reduces the influence of noisy data and the incompleteness of the original samples on the fuzzy rule base. In addition, the FCAWM effectively decreases the scale of the samples, thereby greatly reducing the complexity of the computational model.

In conclusion, the implementation steps for the proposed method are as follows:

Step 1: Input the training sample set. Then, learn and cluster the training sample set to obtain an optimized sample set and the degree of affinity using the AFCM algorithm.

Step 2: Input the new sample set created in step 1 to extract fuzzy rules using the improved WM method given in Section 3, and build the fuzzy rule base.

Step 3: Construct the fuzzy system based on the extracted fuzzy rule base.

A flowchart of the proposed method is shown in Fig. 2.

To make the fuzzy system more robust against noisy data in modeling tasks, the DM method proposed by Wang et al. was used to improve the WM algorithm [9]. The output values were replaced by membership function values to calculate the average output values for each rule, so that each fuzzy rule can be extracted from all the samples. Yang et al. presented a new combined modeling method based on fuzzy system and evolutionary algorithm (known as the PSO-WM method), which effectively enhances the accuracy of the WM method and demonstrates its competitiveness. The PSO-WM method adopts particle swarm optimization (PSO) to optimize the fuzzy rule centroid of the data thereby obtaining a complete fuzzy rule set through extrapolation [14]. In our previous work, an improved WM method, combining the FCM and COWM methods (known as the FCMcoWM method), was proposed to build a fuzzy system from the original samples and enhance the capabilities of the WM method [23]. On the other hand, the FCAWM method is applied to all the samples to generate each fuzzy rule so that missing rules over regions not covered by the samples can be extrapolated to ensure that the completeness of the FCAWM is high. In addition, the degree of affinity is introduced to express the relationship among samples. The FCAWM uses the modified FCM (AFCM) algorithm to optimize the training sample set, while the AFCM uses affinity to decrease the effect of noise. Thereafter, the affinity is used to improve the WM method, by which the interference of noise can be reduced further. Therefore, the robustness of the FCAWM method is enhanced. Since the contribution of each sample is different, the average output values for each rule are computed from the output values and their corresponding affinity in the proposed method, thereby enhancing the accuracy of the WM method.

5. Experiments and analysis

5.1. Single-input single-output model

A training sample set of 42 records was collected from the model in Eq. (26). Table 1 shows the training sample

$$f(x) = 0.5\cos(2\pi x) + 0.5\cos(3\pi x) \tag{26}$$

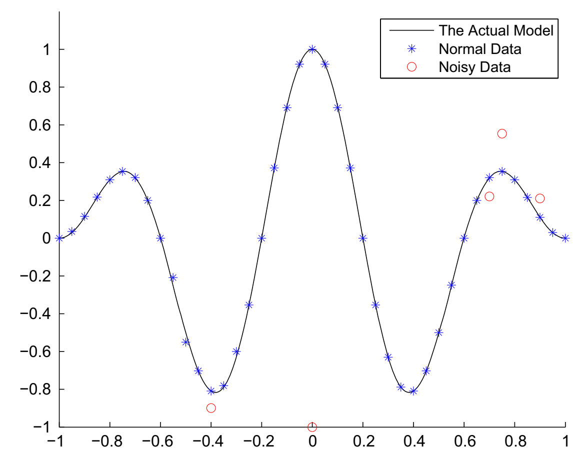

Fig. 3. Training sample and the actual model.

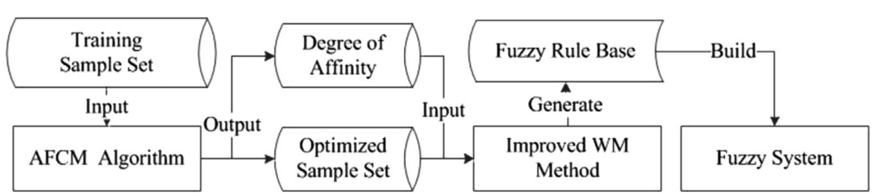

Fig. 2. Flowchart of the FCAWM.

Table 1
Training sample.

t	x	$f(x)$	t	x	$f(x)$	t	x	$f(x)$	t	x	$f(x)$
t_1	−1	0	t_{13}	−0.4	−0.809	t_{25}	0.2	0	t_{37}	0.8	0.309
t_2	−0.95	0.035	t_{14}	−0.35	−0.7807	t_{26}	0.25	−0.3536	t_{38}	0.85	0.2157
t_3	−0.9	0.1156	t_{15}	−0.3	−0.6	t_{27}	0.3	−0.63	t_{39}	0.9	0.1106
t_4	−0.85	0.2175	t_{16}	−0.25	−0.3536	t_{28}	0.35	−0.7877	t_{40}	0.95	0.03
t_5	−0.8	0.309	t_{17}	−0.2	0	t_{29}	0.4	−0.809	t_{41}	1	0
t_6	−0.75	0.3536	t_{18}	−0.15	0.3721	t_{30}	0.45	−0.7025	t_{42}	−0.4	−0.899
t_7	−0.7	0.321	t_{19}	−0.1	0.6904	t_{31}	0.5	−0.5	t_{43}	0.7	0.221
t_8	−0.65	0.2	t_{20}	−0.05	0.921	t_{32}	0.55	−0.2485	t_{44}	0.75	0.5536
t_9	−0.6	0	t_{21}	0	1	t_{33}	0.6	0	t_{45}	0.9	0.2106
t_{10}	−0.55	−0.2085	t_{22}	0.05	0.921	t_{34}	0.65	0.2	t_{46}	0	−1
t_{11}	−0.5	−0.55	t_{23}	0.1	0.6904	t_{35}	0.7	0.321			
t_{12}	−0.45	−0.7025	t_{24}	0.15	0.3721	t_{36}	0.75	0.3536			

In the actual model, x is the input variable of the fuzzy system with range $[-1,1]$. The range of input variable x is divided into nine fuzzy subsets, $A_1,\dots,A_9$, while the output variable y is also divided into nine fuzzy subsets, $B_1,\dots,B_9$. A Gaussian function is applied as the membership function for these variables. In the following sub-sections, this paper compares the results of the various models including the original WM method, COWM method, FCMcoWM method, DM method and FCAWM method, applied to the training sample given in Table 1. From Fig. 3, $t_{42}-t_{46}$ denote noisy data.

The methods were compared with respect to the following four cases: (1) the training sample set is complete and does not contain noisy data; (2) the training sample set is complete, but contains noisy data; (3) the training sample set is incomplete, but does not contain noisy data; (4) the training sample set is incomplete and contains noisy data.

The mean square error (MSE) is defined as follows:

$$MSE=\sqrt{\frac{1}{N}\left[\sum_{i=1}^{N}(f^E(x_i)-f(x_i))\right]^2} \tag{27}$$

In this definition, $f^E(x_i)$ is the theoretical value of the fuzzy system, while $f(x_i)$ is the reasoning value of the fuzzy system.

Table 2 presents a quantitative comparison of the performance of the five methods in the above four cases, where MAC is defined

Table 2
Comparison of fuzzy systems generated by WM, COWM, DM, FCMcoWM and FCAWM.

	Method	$x\Rightarrow$	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	MAC	MSE
Case 1	WM	$y\Rightarrow$	B_5	B_6	B_3	B_5	B_9	B_5	B_3	B_6	B_5	0.7495	0.2585
	COWM	$y\Rightarrow$	B_5	B_6	B_4	B_3	B_8	B_4	B_3	B_6	B_5	0.8687	0.2129
	DM	$y\Rightarrow$	B_4	B_5	B_2	B_2	B_7	B_2	B_2	B_5	B_4	0.8263	0.2095
	FCMcoWM	$y\Rightarrow$	B_5	B_6	B_4	B_5	B_7	B_4	B_2	B_6	B_5	0.7442	0.2856
	FCAWM	$y\Rightarrow$	B_5	B_6	B_3	B_3	B_9	B_4	B_3	B_6	B_5	0.8856	0.1949
Case 2	WM	$y\Rightarrow$	B_5	B_6	B_3	B_5	B_1	B_5	B_3	B_5	B_7	0.0205	0.6570
	COWM	$y\Rightarrow$	B_5	B_6	B_3	B_4	B_7	B_4	B_3	B_6	B_5	0.7955	0.2830
	DM	$y\Rightarrow$	B_4	B_5	B_2	B_2	B_7	B_2	B_4	B_6	B_5	0.8263	0.2095
	FCMcoWM	$y\Rightarrow$	B_5	B_6	B_5	B_2	B_7	B_2	B_4	B_6	B_5	0.6750	0.3093
	FCAWM	$y\Rightarrow$	B_5	B_6	B_3	B_4	B_8	B_4	B_3	B_6	B_5	0.8822	0.2088
Case 3	WM	$y\Rightarrow$	B_5	B_6	B_3	B_5	B_9	*	B_4	B_6	B_5	0.5545	0.3421
	COWM	$y\Rightarrow$	B_5	B_6	B_3	B_4	B_8	B_5	B_5	B_6	B_5	0.7429	0.2792
	DM	$y\Rightarrow$	B_4	B_5	B_2	B_2	B_7	B_7	B_4	B_5	B_4	0.2525	0.5200
	FCMcoWM	$y\Rightarrow$	B_5	B_6	B_3	B_5	B_9	B_5	B_5	B_6	B_5	0.5051	0.2965
	FCAWM	$y\Rightarrow$	B_5	B_6	B_3	B_4	B_9	B_5	B_4	B_6	B_5	0.7853	0.2524
Case 4	WM	$y\Rightarrow$	B_5	B_6	B_3	B_5	B_1	*	B_4	B_7	B_5	0.0332	0.7312
	COWM	$y\Rightarrow$	B_5	B_6	B_3	B_4	B_6	B_5	B_5	B_6	B_5	0.5183	0.3782
	DM	$y\Rightarrow$	B_4	B_5	B_2	B_2	B_7	B_1	B_4	B_5	B_4	0.6800	0.2879
	FCMcoWM	$y\Rightarrow$	B_5	B_6	B_5	B_2	B_8	B_5	B_4	B_6	B_5	0.6129	0.3192
	FCAWM	$y\Rightarrow$	B_5	B_6	B_3	B_4	B_9	B_5	B_4	B_6	B_5	0.7629	0.2537

Table 3
Paired-sample t-test (95% confidence interval of the difference).

	Method	Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	Lower	Upper			
Case 1	WM	−0.062	0.252	0.018	−0.098	−0.027	−3.506	199	0.001
	COWM	0.001	0.214	0.015	−0.029	0.030	0.043	199	0.966
	DM	0.125	0.169	0.012	0.101	0.149	10.447	199	0.000
	FCMcoWM	0.020	0.285	0.020	−0.020	0.060	0.987	199	0.325
	FCAWM	0.045	0.201	0.014	0.017	0.074	3.192	199	0.002
Case 2	WM	0.156	0.641	0.045	0.067	0.246	3.445	199	0.001
	COWM	0.053	0.279	0.020	0.014	0.092	2.705	199	0.007
	DM	0.125	0.169	0.012	0.101	0.149	10.447	199	0.000
	FCMcoWM	−0.003	0.285	0.020	−0.043	0.037	−0.147	199	0.883
	FCAWM	0.051	0.223	0.016	0.020	0.082	3.229	199	0.001
Case 3	WM	−0.141	0.314	0.022	−0.184	−0.097	−6.342	199	0.000
	COWM	−0.079	0.269	0.019	−0.116	−0.041	−4.125	199	0.000
	DM	−0.167	0.495	0.035	−0.236	−0.098	−4.760	199	0.000
	FCMcoWM	−0.142	0.280	0.020	−0.181	−0.103	−7.155	199	0.000
	FCAWM	−0.077	0.241	0.017	−0.110	−0.043	−4.498	199	0.000
Case 4	WM	0.203	0.706	0.050	0.105	0.302	4.068	199	0.000
	COWM	−0.017	0.340	0.024	−0.065	0.030	−0.719	199	0.473
	DM	0.083	0.277	0.020	0.045	0.122	4.255	199	0.000
	FCMcoWM	−0.120	0.271	0.019	−0.158	−0.082	−6.278	199	0.000
	FCAWM	−0.058	0.246	0.017	−0.093	−0.024	−3.356	199	0.001

in Problem 2. In addition, the results of paired t-test for this experiment are shown in Table 3.

5.1.1. Case 1: The training sample set is complete and does not contain noisy data

The training sample set S_1, that is, $T=\{t_1,t_2,\ldots,t_{40},t_{41}\}$ in Table 1, is generated by Eq. (26). S_1 is complete and does not contain noisy data. Moreover, its input variables cover all the input fuzzy subsets. In this case, the fuzzy rule bases generated by these four methods are similar, and therefore, the fuzzy systems are basically consistent. Since the FCAWM uses all the samples to generate each rule, in $[-0.05, 0.05]$ the wave crest of the FCAWM is relatively smaller than that of the WM method. The fuzzy rule bases generated by these methods are given in Table 2 while the results are shown in Fig. 4(a).

5.1.2. Case 2: The training sample set is complete, but contains noisy data

Training sample set S_2 is given as $T=\{t_1,t_2,\ldots,t_{45},t_{46}\}$ in Table 1. S_2 contains noisy data t_{42}, t_{43}, t_{44}, t_{45} and t_{46}. In this situation, the WM method is affected by the interference caused by the noisy data. The COWM method uses all the variables to calculate the weights for extracting rules. This can improve the robustness of the fuzzy system. The FCAWM uses the FCM algorithm to preprocess S_2, thereby reducing the impact of the noisy data and smoothing the influence of these data. Thereafter, fuzzy rules can be generated from the new sample set. The resulting fuzzy system performs better than those of the other methods. Fig. 4(b) shows the accuracy rates of these methods.

5.1.3. Case 3: The training sample set is incomplete, but does not contain noisy data

The training sample set S_3 is given as $T=\{t_1,t_2,\ldots,t_{21},t_{32},\ldots,t_{40},t_{41}\}$ in Table 1. Since S_3 does not cover all the input fuzzy subsets, it is incomplete. Using the WM method to extract fuzzy rules from S_3 causes the created fuzzy rule base to be incomplete as well. No rule is generated in position $\ast$ in Table 2 and thus, the WM method does not have high completeness. By allowing all the samples to vote, the COWM method has

Table 4
Experimental environment.

CPU	Intel(R) Core(TM) i5-2400 CPU at 3.10 GHz 3.10 GHz
RAM	4.00GB (2.91GB available)
System	Windows7 32bit
Software	Matlab 7.12.0(R2011a)

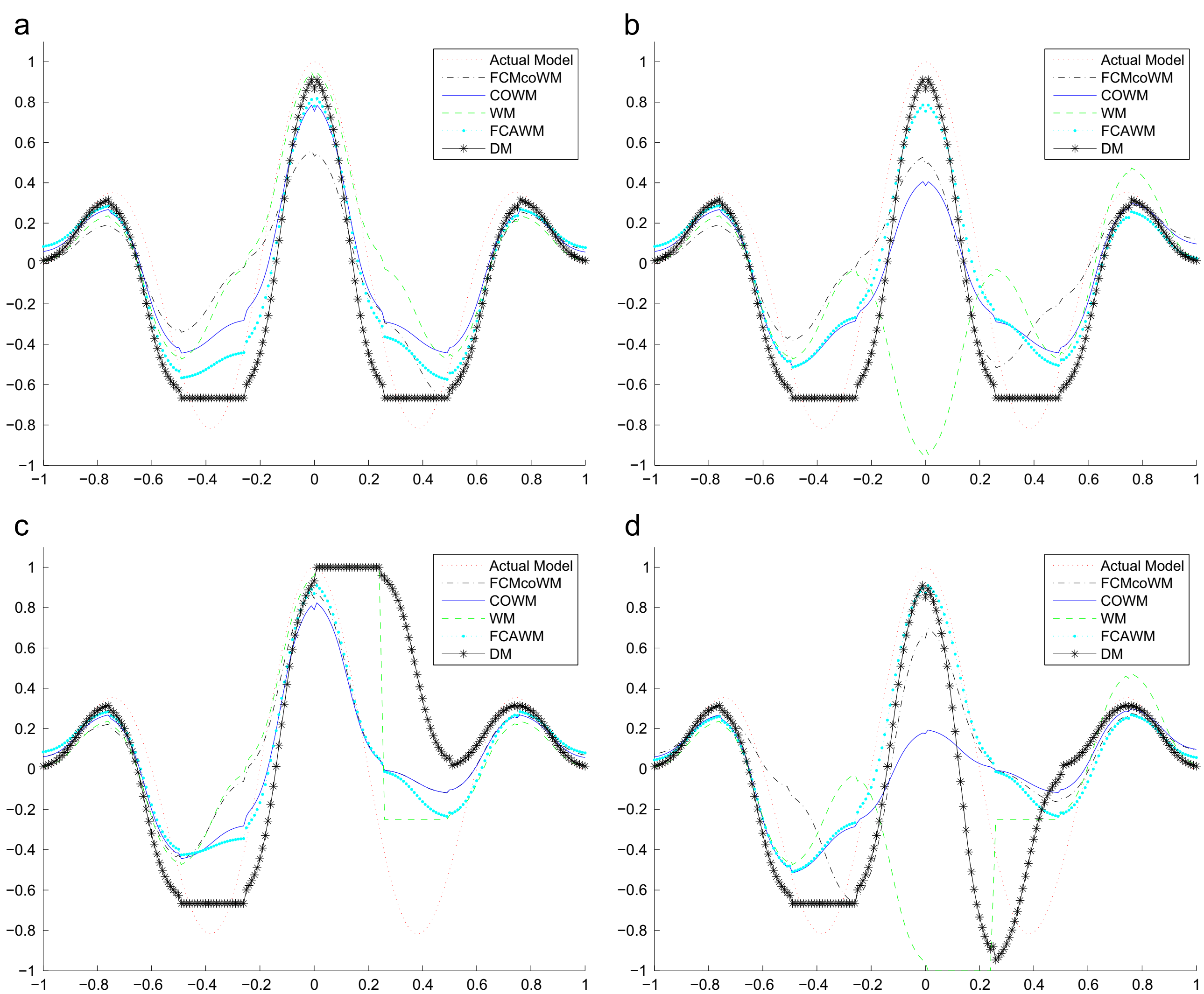

Fig. 4. Comparison of the experimental results of WM, COWM, DM, FCMcoWM, FCAWM and the actual model. (a) Case 1: The training sample set is complete and does not contain noisy data, (b) Case 2: The training sample set is complete, but contains noisy data, (c) Case 3: The training sample set is incomplete, but does not contain noisy data, (d) Case 4: The training sample set is incomplete and contains noisy data.

high completeness. The FCAWM method also has high completeness and reduces the scale of the samples. The system created by the FCAWM is similar to the fuzzy system generated by the COWM method. The results of this experiment are shown in Fig. 4(c).

5.1.4. Case 4: The training sample set is incomplete and contains noisy data

The training sample set S_4 is given as $T=\{t_1, t_2, \ldots, t_{21}, t_{32}, \ldots, t_{45}, t_{46}\}$ in Table 1, and so S_4 is incomplete. The WM method may extract invalid or incorrect rules if it selects noisy

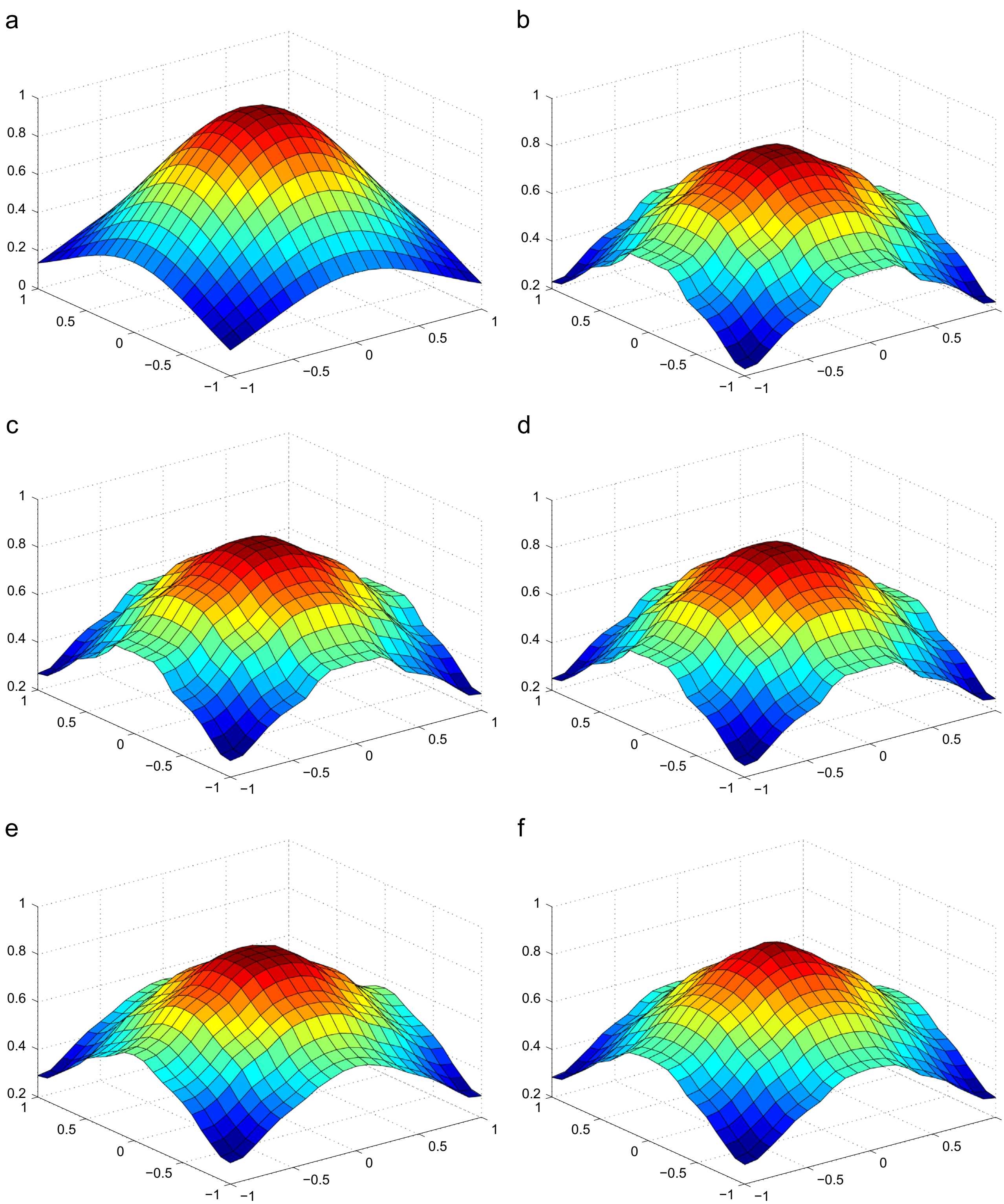

Fig. 5. Comparison of the fuzzy systems created by WM, COWM, DM, FCMcoWM, and FCAWM, and the actual model. (a) Actual model, (b) fuzzy system created by WM, (c) fuzzy system created by COWM, (d) fuzzy system created by DM, (e) fuzzy system created by FCMcoWM, (e) fuzzy system created by FCAWM.

Table 5
Experimental results of MISO model.

Method	*MSE*	Computation time (s)
WM	0.0909	8675
COWM	0.0919	15 476
DM	0.0890	12 703
FCMcoWM	0.1012	5891
FCAWM	0.0875	8542

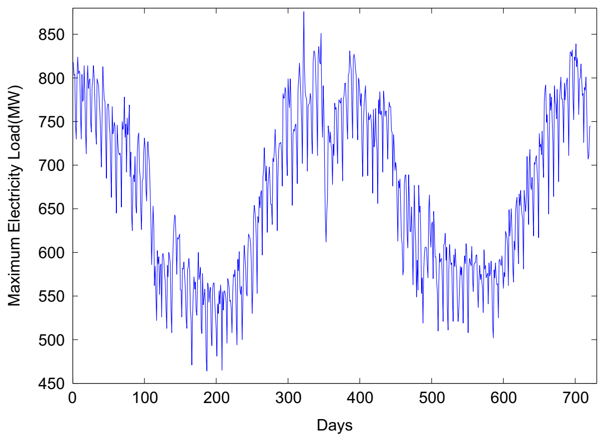

Fig. 6. Maximum daily loads from 1997 to 1998 (EUNITE).

Table 6
A subset of the training sample set.

L_{d-7}	L_{d-6}	L_{d-5}	L_{d-4}	L_{d-3}	L_{d-2}	L_{d-1}	W	H	D	L_d
726	685	765	770	770	746	737	6	1	1	709
685	765	770	770	746	737	709	7	0	2	688
765	770	770	746	737	709	688	1	0	3	799
770	770	746	737	709	688	799	2	0	4	778
770	746	737	709	688	799	778	3	0	5	766
746	737	709	688	799	778	766	4	0	6	799
737	709	688	799	778	766	799	5	0	7	781
709	688	799	778	766	799	781	6	0	8	719
688	799	778	766	799	781	719	7	0	9	654
799	778	766	799	781	719	654	1	0	10	735
778	766	799	781	719	654	735	2	0	11	741
766	799	781	719	654	735	741	3	0	12	739
799	781	719	654	735	741	739	4	0	13	746
781	719	654	735	741	739	746	5	0	14	747
719	654	735	741	739	746	747	6	0	15	714

data. In Table 2, there is still no rule extracted in position ⁎. According to Tables 2 and 3, COWM is shown to have good performance. However, the FCAWM uses the FCM algorithm to preprocess S_4 and calculate the affinity variances. Then, it applies the improved WM method based on the affinity among samples to extract fuzzy rules. The resulting fuzzy system has higher completeness and robustness than the other systems. Fig. 4(d) shows the approximation performance of the methods.

From Tables 2 and 3, if the training sample set is complete and does not contain noisy data, the fuzzy system generated by the FCAWM is similar to those generated by the other methods, but the *MSE* is smaller than in the other systems. The FCAWM performs relatively better than the other methods in Fig. 4. According to the statement of Problem 1, the fuzzy rule base built by the WM method is incomplete if the training sample set is incomplete, such as the position ⁎ in Table 2. When the training sample set contains noisy data, the *MAC* value of the WM method in Table 2 indicates that the robustness of WM method is poor. Moreover, the FCAWM has better prediction accuracy than the other methods, while also showing high completeness and robustness. In addition, it can effectively reduce the scale of samples thereby decreasing the complexity of the system. This helps reduce computation time.

5.2. Training set with large scale

In this experiment, this paper created fuzzy systems based on the generated fuzzy rule bases using the original WM method, COWM method, DM method, FCMcoWM method, and FCAWM method. A training sample set of ten thousands records was collected from the model in Eq. (28), where x_1 and x_2 are the input variables of the fuzzy system with range $[-1, 1]$. Then the results compared the approximation performance using the nonlinear model discussed below. Table 4 shows the environment for this experiment:

$$f(x) = e^{-(x_1^2 + x_2^2)} \tag{28}$$

The range of input variables x_1 and x_2 was set $[-1, 1]$, and both were divided into nine fuzzy subsets, as was output variable y. A Gaussian function was selected as the membership function for all of these. The actual model and fuzzy systems created by the five methods are shown in Fig. 5.

From Fig. 5 and Table 5, in the DM method, the output values are replaced by the membership function values to calculate the average output values for each rule, to ensure that fuzzy rules can

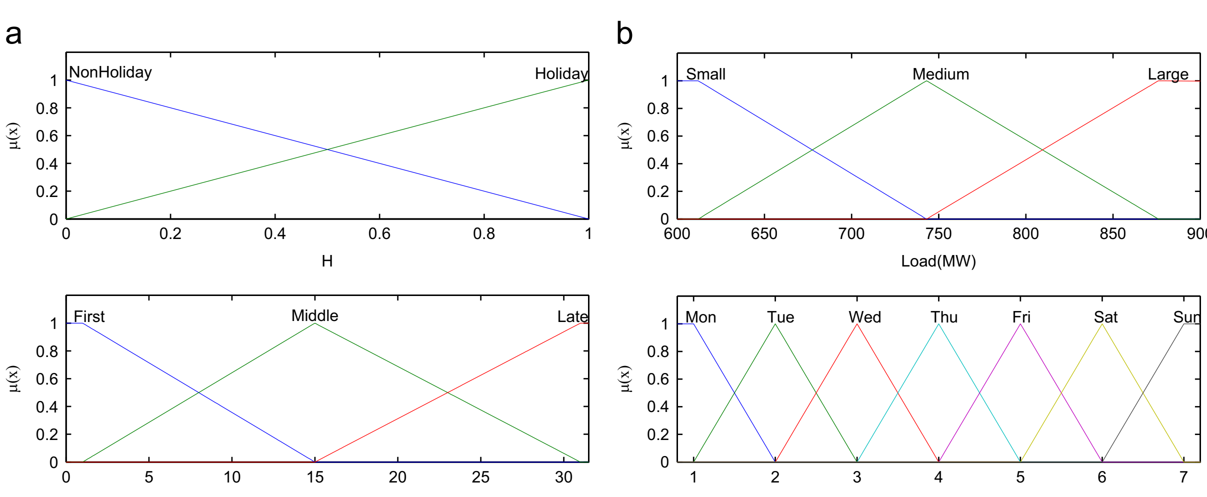

Fig. 7. Membership function for input variables. (a) Holiday attribute and date attribute, (b) daily maximum load and week attribute.

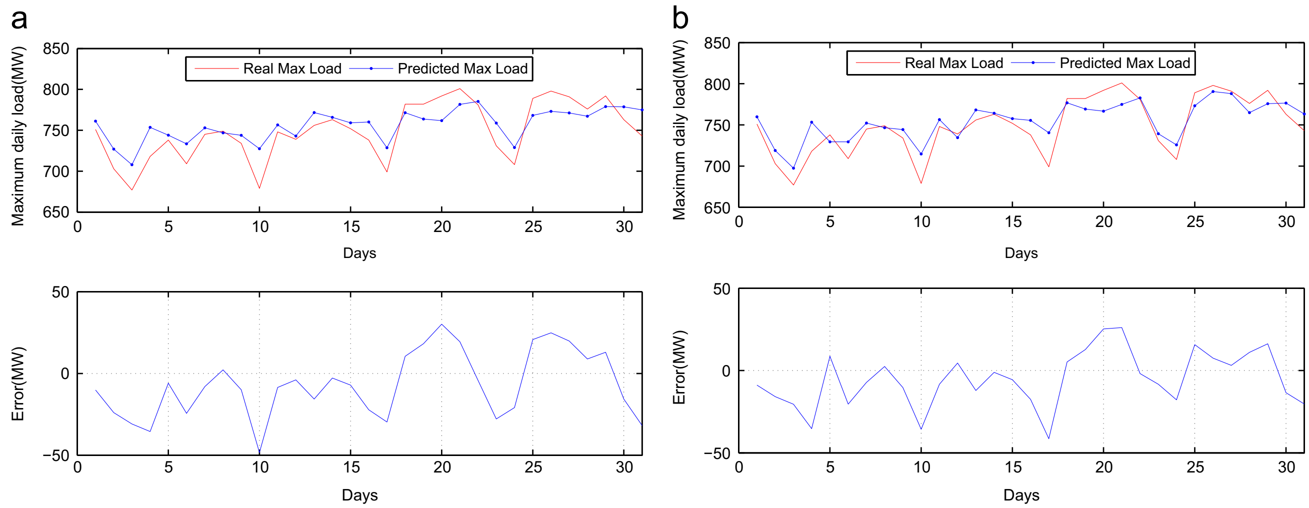

Fig. 8. Comparison of forecasting results obtained by WM and FCAWM methods. (a) Forecasting result of WM method, (b) forecasting result of FCAWM method.

be extracted from all the samples. The COWM performs the weighted average operation twice to improve completeness and robustness of the original WM method. Thus, the DM and COWM methods require more time to generate the fuzzy system than the WM method. However, the FCAWM is equipped with the capability of the FCM algorithm to optimize the original sample set and guarantees completeness and robustness of the fuzzy system. The scale of the sample set is greatly reduced, thereby decreasing the computation time for fuzzy system extraction.

5.3. Case study

In the multiple-input, single-output (MISO) model, practical data provided by the EUNITE network were adopted (http://neuron.tuke.sk/competition), including daily maximum loads from 1997 to 1998, and daily average temperatures from 1997 to 1998. Also included were the holidays for the same period. The actual goal was to predict the maximum daily values of electrical loads for January 1999 . The daily maximum loads for 1997 and 1998 are shown in Fig. 6.

The magnitudes of the E_{MAPE} and E_M errors were used to estimate the performance of the different solutions and the maximum error, respectively:

$$E_{MAPE} = 100 * \frac{\sum_{i=1}^{N} \left| \frac{L_{R_i} - L_{P_i}}{L_{R_i}} \right|}{N} \tag{29}$$

$$E_M = \max\left(\left|L_{R_i} - L_{P_i}\right|\right) \tag{30}$$

where L_{R_i} is the actual maximum daily electrical load on the “i” day of the year 1999, L_{P_i} is the predicted value of maximum daily electrical load on the “i” day of the year 1999, and N ($N=31$) is the number of days in January 1999.

By studying the maximum daily loads from 1997 to 1998, the load of the training data for one day represents the output variable L_d with the following attributes:

- Seven attributes for the maximum loads of the past 7 days.
- One numerical attribute indicating the day of the week.
- One binary attribute indicating whether this day is a holiday.
- One numerical attribute for the calendar date.

Therefore, the model for prediction this study is given as

$$x = [L_{d-7}, L_{d-6}, L_{d-5}, L_{d-4}, L_{d-3}, L_{d-2}, L_{d-1}, W, H_d, D] \tag{31}$$

$$y = L_d \tag{32}$$

where L_d denotes the forecast daily maximum load, $L_{d-7}, L_{d-6}, L_{d-5}, L_{d-4}, L_{d-3}, L_{d-2}, L_{d-1}$ are the daily maximum loads for the seven days prior to the forecast day, W expresses the week attribute of the forecast day (1 to 7 denotes Monday to Sunday, respectively), H_d is the holiday attribute (1 denotes a holiday, and 0 otherwise), and D is the date attribute.

This study used the daily maximum load, week attributes, holiday attributes, and date attributes for January, February, November, and December of 1997 and 1998 as training samples. For illustrative purposes, a subset comprising 15 samples is shown in Table 6.

For comparison with the PSO-WM method, the fuzzy sets are shown in Fig. 7.

The WM and FCAWM methods were separately applied to forecast the load in January 1999 for EUNITE, the results are illustrated in Fig. 8. The forecasting result of the WM method has $E_{MAPE}=2.4281$ and $E_M=48.3813$, while that for FCAWM method has $E_{MAPE}=1.9300$ and $E_M=41.4083$. Hence, this experiment indicates that the prediction model generated by the FCAWM effectively improves the accuracy of the WM method.

Compared with the methods used in other studies on the EUNITE network, the FCAWM method achieves a similar level of forecasting accuracy as the LIBSVM method [34] in terms of E_{MAPE} evaluation, and outperforms the LIBSVM in terms of E_M. Petter Otto applied as fuzzy time-series based method to predict EUNITE competition data and obtained $E_{MAPE}>11$ and $E_M \approx 200$ MW, while Ahamd Lotifi adopted as fuzzy reasoning system and obtained $E_{MAPE}>3.0$ and $E_M \approx 50$ MW. The result of the PSO-WM method is $E_{MAPE}=2.0372$ and $E_M=43.6498$ MW. These results show that the WM method achieves better accuracy than the two fuzzy methods, and that the FCAWM can enhance the performance of the WM method.

6. Conclusion

To improve the completeness and robustness of the original WM method, a novel method for fuzzy rule generation, called FCAWM, was proposed. The improved FCM algorithm was used to cluster and optimize the original sample set. The disturbing effect of noisy data was reduced, and the noisy data were smoothed. Since the WM method can extract rules from samples, the FCAWM also has this capability. In addition, it observably decreases the negative effects of incomplete sample sets and noisy data on fuzzy rule generation. Experimental results on a nonlinear function prove that the proposed method has high completeness and robustness with better accuracy than the original WM method. A case study on training set with large scale illustrates that the FCAWM method decreases the computation time for fuzzy system extraction. Moreover, the short-term load forecasting experiment confirms that the FCAWM method enhances prediction accuracy of the WM method.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grants 61103170 and 51305142, the Program for Prominent Young Talent in Fujian Province University under Grant JA12005, and the Promotion Program for Young and Middle-aged Teachers in Science and Technology Research at Huaqiao University under Grant ZQN-PY211.

References

[1] H. Ishibuchi, T. Yamamoto, Rule weight specification in fuzzy rule-based classification systems, IEEE Trans. Fuzzy Syst. 13 (4) (2005) 428–435.
[2] D. Chakraborty, N.R. Pal, A neuro-fuzzy scheme for simultaneous feature selection and fuzzy rule-based classification, J. Spacecr. Rockets 15 (1) (2004) 110–123.
[3] H. Ishibuchi, T. Nakashima, T. Murata, Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems, IEEE Trans. Syst. Man Cybern.—Part B: Cybern. 29 (5) (1999) 601–618.
[4] S. Guillaume, Designing fuzzy inference systems from data: an interpretability-oriented review, IEEE Trans. Fuzzy Syst. 9 (3) (2001) 426–443.
[5] W. Zhang, W. Liu, IFCM: fuzzy clustering for rule extraction of interval type-2 fuzzy logic system, in: 2007 46th IEEE Conference on Decision and Control, New Orleans, LA, 2007, pp. 5318–5322.
[6] L.X. Wang, J.M. Mendel, Generating fuzzy rules by learning from examples, IEEE Trans. Syst. Man Cybern. 22 (6) (1992) 1414–1427.
[7] L.X. Wang, The wm method completed: a flexible fuzzy system approach to data mining, IEEE Trans. Fuzzy Syst. 11 (6) (2003) 768–782.
[8] Y. Wang, T. Chai, Mining fuzzy rules from data and its system implementation, J. Syst. Eng. 20 (5) (2005) 497–503 (in Chinese).
[9] Y. Wang, D. Wang, T. Chai, Extract of fuzzy rules with completeness and robustness, Acta Autom. Sin. 36 (9) (2010) 1337–1342 (in Chinese).
[10] Y. Bai, T. Li, Robust fuzzy inference system for prediction of time series with outliers, in: 2012 International Conference on Fuzzy Theory and it's Applications (iFUZZY), Taichung, 2012, pp. 394–399.
[11] E. Cox, M. O'Hagan, The Fuzzy Systems Handbook, A Practitioner's Guide to Building, Using, and Maintaining Fuzzy Systems, second ed., Morgan Kaufmann, CA, USA, 1998.
[12] L.X. Wang, Adaptive Fuzzy Systems and Control: Design and Stability Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1994.
[13] E.R.R. Kato, O. Morandin Jr., M. Sgavioli, B.D. Muniz, Genetic tuning for improving Wang and Mendel's fuzzy database, in: IEEE International Conference on Systems, Man and Cybernetics, 2009, SMC 2009, San Antonio, TX, 2009, pp. 1015–1020.
[14] X. Yang, J. Yuan, J. Yuan, H. Mao, An improved wm method based on pso for electric load forecasting, Expert Syst. Appl. 37 (12) (2010) 8036–8041.
[15] M. Galende-Hernandez, G.I. Sainz-Palmero, M.J. Fuente-Aparicio, Complexity reduction and interpretability improvement for fuzzy rule systems based on simple interpretability measures and indices by bi-objective evolutionary rule selection, Soft Comput. 16 (3) (2012) 451–470.
[16] J. Casillas, O. Cordon, F. Herrera, Improving the Wang and Mendel's fuzzy rule learning method by inducing cooperation among rules, in: The 8th Information Processing and Management of Uncertainty in Knowledge-Based Systems Conference, Madrid, Spain, 2000, pp. 1682–1688.
[17] W. Chen, J. Gou, Improved Wang–Mendel scheme based on cooperation among input variables, Int. Rev. Comput. Softw. 7 (5) (2012) 2685–2689.
[18] J. Zhu, Some common key problems and their dealing methods in the application of fuzzy mathematical methods, Fuzzy Syst. Math. 6 (2) (1992) 57–63 (in Chinese).
[19] J.C. Dunn, A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters, J. Cybern. 3 (3) (1973) 32–57.
[20] J.C. Bezdek, R. Ehrlich, W. Full, FCM: the fuzzy c-means clustering algorithm, Comput. Geosci. 10 (2–3) (1984) 191–203.
[21] M.S. Yang, A survey of fuzzy clustering, Math. Comput. Modell. 18 (11) (1993) 1–16.
[22] S. Chen, Y. Chen, W. Zhang, W. Yan, A method of extracting fuzzy control rules of structural vibration using fcm algorithm, in: 2010 The 5th IEEE Conference on Industrial Electronics and Applications (ICIEA), Taichung, 2010, pp. 1550–1555.
[23] F. Hou, J. Gou, Fuzzy rule generation based on cowm and fcm algorithm, Int. J. Appl. Math. Stat. 45 (15) (2013) 20–27.
[24] J. Casillas, O. Cordon, L.M. FranciscoHerrera (Eds.), Accuracy Improvements in Linguistic Fuzzy Modeling, Springer, Berlin, 2003.
[25] J. Casillas, O. Cordon, L.M. FranciscoHerrera (Eds.), Interpretability Issues in Fuzzy Modeling, Springer, Berlin, 2003.
[26] J.V.D. Oliveira, W. Pedrycz, Advances in Fuzzy Clustering and its Applications, John Wiley & Sons, Chichester, UK, 2007.
[27] J. Yu, M.S. Yang, E.S. Lee, Sample-weighted clustering methods, Comput. Math. Appl. 62 (5) (2011) 2200–2208.
[28] J.C. Bezdek, R.J. Hathaway, M.J. Sabin, W.T. Tucker, Convergence theory for fuzzy c-means: counterexamples and repairs, IEEE Trans. Syst. Man Cybern. 17 (5) (1987) 873–877.
[29] W. Wei, J.M. Mendel, Optimality tests for the fuzzy c-means algorithms, Pattern Recognit. 27 (11) (1994) 1567–1573.
[30] J. Yu, M.S. Yang, Optimality test for generalized fcm and its application to parameter selection, IEEE Trans. Fuzzy Syst. 13 (1) (2005) 164–176.
[31] M. Tang, X. Chen, W. Hu, W. Yu, Generation of a probabilistic fuzzy rule base by learning from examples, Inf. Sci. 217 (25) (2012) 21–30.
[32] M. Dan, P. Zheng, Extracting linguistic rules from data sets using fuzzy logic and genetic algorithms, Neurocomputing 78 (1) (2012) 48–54.
[33] S.N. Sulaiman, N.A.M. Isa, Denoising-based clustering algorithms for segmentation of low level salt-and-pepper noise-corrupted images, IEEE Trans. Consumer Electron. 56 (4) (2010) 2702–2710.
[34] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines ⟨http://www.csie.ntu.edu.tw/~cjlin/libsvm/⟩ [Online], 2001.

Jin Gou (1978–) received the Ph.D. degree in computer science and technology from Zhejiang University, China, in 2006. Currently, he is an associate professor at Huaqiao University, China. His research interests include artificial intelligence and knowledge fusion.

Feng Hou (1990–) received the graduate degree in computer science and technology from Huaqiao University, China, in 2012. Currently, he is a graduate student at Huaqiao University, China. His research interests include artificial intelligence and data mining.

Wenyu Chen (1987–) received the graduate degree in computer science and technology from Huaqiao University, China, in 2010. Currently, he is a graduate student at Huaqiao University, China. His research interests include artificial intelligence.

Cheng Wang (1984–) received the Ph.D. degree in mechanics from Xi'an Jiaotong University, China, in 2012. Currently, he is a lecturer at Huaqiao University, China. His research interests include signal processing and artificial intelligence.

Wei Luo (1976–) received the M.S. degree in computer sciences from University of Wisconsin-Madison, USA, in 2004. Currently, he is a lecturer at Huaqiao University, China. His research interests include artificial intelligence.