SCIENCE CHINA
Information Sciences

• LETTER •

April 2020, Vol. 63 149201:1–149201:3
https://doi.org/10.1007/s11432-018-9545-8

Simplified outlier detection for improving the robustness of a fuzzy model

Yali JIN[1,2], Weihua CAO[1,2*], Min WU[1,2] & Yan YUAN[1,2]

[1] *School of Automation, China University of Geosciences, Wuhan 430074, China;*
[2] *Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, China University of Geosciences, Wuhan 430074, China*

Received 14 June 2018/Accepted 16 July 2018/Published online 27 August 2019

Citation Jin Y L, Cao W H, Wu M, et al. Simplified outlier detection for improving the robustness of a fuzzy model. Sci China Inf Sci, 2020, 63(4): 149201, https://doi.org/10.1007/s11432-018-9545-8

Dear editor,

Fuzzy models have been widely employed in various fields, such as complex industrial process modeling, classification, function approximation, and time series prediction. Extracting fuzzy rules is very important as they determine the capabilities of the fuzzy model. In recent years, data-based methods such as fuzzy clustering, genetic algorithm, and neural networks [1–3], have been proposed for extracting fuzzy rules. However, these methods require continual learning or possess complicated mechanisms that make them impractical for engineering applications.

The Wang-Mendel (WM) method [4], which is a classic method for fuzzy modeling, generates a fuzzy rule from a data point. This method is simple yet useful and rules generated from this method are easy to explain to nonspecialists. However, it lacks completeness and robustness so Wang [5] introduced an idea to generate a rule from multiple data points to improve the performance of the WM method. Unfortunately, despite this improvement, outliers still influence the accuracy of the fuzzy model derived with the WM method. For example, an outlier in A^1 (Figure 1(a)) has a great influence on the output subset because there are few data points in A^1. To further reduce the influence of the outliers, it was practical to use outlier detection to obtain the weights of the data points. In [6], the weights of the outliers were obtained by the fuzzy c-means (FCM) clustering algorithm, which increased the complexity of the WM method. Furthermore, other methods for outlier detection are always complicated and are difficult when they are used for a multi-input modeling problem.

To improve the robustness of the fuzzy model in a straightforward way, we simplified the outlier detection problem using the WM method to transform the outlier detection of the input and output data points to the outlier detection of the output points. In this way, the multidimensional outlier detection problem can be transformed into a one-dimensional outlier detection problem. The Gaussian distribution is then used to solve this one-dimensional outlier detection problem. To establish the fuzzy model, the WM method was first used to classify the fuzzy rules and the data points into different groups according to their inputs. Data points in the same group have inputs that are in the same space. If the output of a data point is far away from the other points in its group, it can be regarded as an outlier, as shown in Figure 1(a). Therefore, it is possible to just detect the outliers of the output points in a group without considering any of the inputs. Next, for every group of output points, we defined a Gaussian distribution and obtained the weight of every output point. To reduce the influence of outliers on the Gaussian distribution function, the median was used to replace the average as the parameter of this function. The weights of the outliers were small be-

* Corresponding author (email: weihuacao@cug.edu.cn)

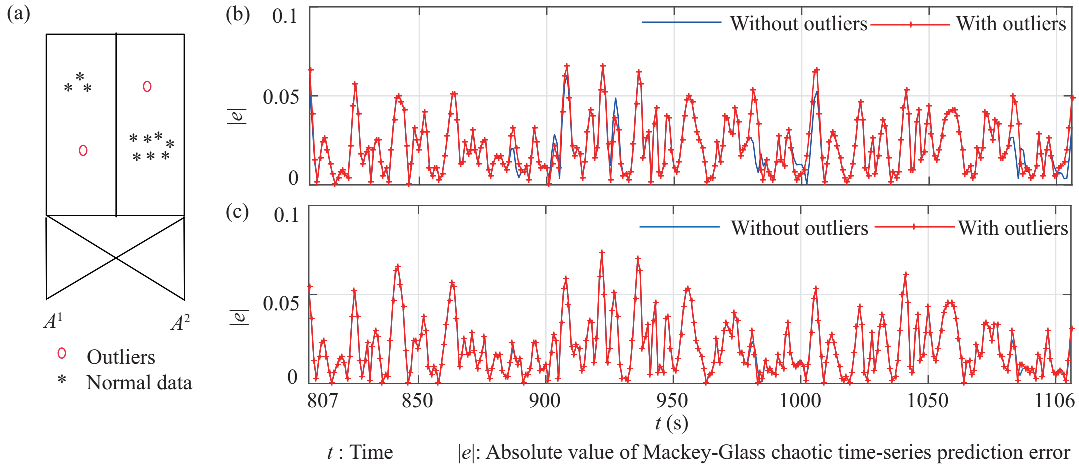

Figure 1 (Color online) (a) Fuzzy regions of the input-output space and sample data; (b) predicted results of the WM method; (c) predicted results of our proposed method.

cause the outliers were far away from the group median. Thus the influence of the outliers was reduced and the robustness of the fuzzy model was improved. Finally, fuzzy rules were all extracted from the groups, thus establishing the fuzzy model. Results of the chaotic time-series prediction experiment proved that our proposed method improved the robustness of the fuzzy model in a simple way.

Approach overview. In order to clarify the basic idea of our new approach, we take the simple two-input and one-output case as an example. Suppose we are given a set of input-output data points $(x_1^{(p)}, x_2^{(p)}, y^{(p)})$, $p = 1, 2, \dots, N$, where $x_1^{(p)}$ and $x_2^{(p)}$ are the p-th elements of the first and the second inputs, $y^{(p)}$ is the p-th element of the output, and N is the number of the data points.

To establish a fuzzy model, we first need to generate the fuzzy rules of every data point and classify the rules into different groups, so the data points are divided into different groups. Next, we can establish a Gaussian distribution for every group and obtain the weights of every output. At last, all the fuzzy rules are extracted from the groups and the fuzzy model can be established.

Generation and classification of fuzzy rules of every data point. We first define the fuzzy subsets of the two inputs as $A_1 = \{A_1^1, A_1^2, \dots, A_1^{c_1}\}$ and $A_2 = \{A_2^1, A_2^2, \dots, A_2^{c_2}\}$. Then, compute the membership values $\mu_{A_1^{l_1}}(x_1^{(p)})$ and $\mu_{A_2^{l_2}}(x_2^{(p)})$ of the p-th two inputs in $A_1^{l_1}$ and $A_2^{l_2}$, respectively. $l_1 = 1, 2, \dots, c_1$ is the index of subsets of the first input, and $l_2 = 1, 2, \dots, c_2$ is the index of subsets of the second input. Find $l_1^* \in (1, 2, \dots, c_1)$ and $l_2^* \in (1, 2, \dots, c_2)$ such that

$$\begin{cases} \mu_{A_1^{l_1^*}}(x_1^{(p)}) \geqslant \mu_{A_1^{l_1}}(x_1^{(p)}), \\ \mu_{A_2^{l_2^*}}(x_2^{(p)}) \geqslant \mu_{A_2^{l_2}}(x_2^{(p)}) \end{cases} \tag{1}$$

for all l_1 and l_2. From here, the rule can be determined that IF x_1 is $A_1^{l_1^*}$ and x_2 is $A_2^{l_2^*}$, THEN y is centered at $y^{(p)}$ for every data, which is called the data-generated rule needed by $(x_1^{(p)}, x_2^{(p)}, y^{(p)})$. We also define a weight for this rule as $w^{(p)} = \mu_{A_2^{l_2^*}}(x_1^{(p)})\mu_{A_2^{l_2^*}}(x_2^{(p)})$.

Thus, we can obtain N data-generated rules. These rules are divided into different groups, where the same groups share the same IF parts. Suppose there are M such groups. Let group m ($m = 1, 2, \dots, M$) have N_m rules such as: IF x_1 is $A_1^{l_1^m}$ and x_2 is $A_2^{l_2^m}$, THEN y is centered at $y^{(p_k^m)}$, where $k = 1, 2, \dots, N_m$, and p_k^m is the index for the data points in group m.

Outlier detection based on Gaussian distribution. The output points are also divided into M groups based on the groups of fuzzy rules. For every group, we can obtain a Gaussian distribution. The median is used to replace the average as the parameter of this function to reduce the influence of outliers on the Gaussian distribution function, The Gaussian distribution of the m-th group is defined as

$$\mathrm{aff}\left(y^{(p_k^m)}\right) = \frac{1}{\theta\sqrt{2\pi}} \mathrm{e}^{\frac{-\left(y^{(p_k^m)} - \bar{y}^{(m)}\right)^2}{2(\theta^{(m)})^2}}, \tag{2}$$

where $\bar{y}^{(m)}$ and $\theta^{(m)}$ are the median and the variance of the output points of the m-th group, respectively.

Fuzzy rules extraction and fuzzy modeling. We combine the N_m rules into a single rule as: IF x_1 is $A_1^{l_1^m}$ and x_2 is $A_2^{l_2^m}$, THEN y is $B^{(m)}$, where $B^{(m)}$ is a triangular fuzzy set [5] whose center $\mathrm{av}^{(m)}$ and variance $\sigma^{(m)}$ are computed as

$$\mathrm{av}^{(m)} = \frac{\sum_{k=1}^{N_m} w^{(p_k^m)} y^{(p_k^m)} \mathrm{aff}(y^{(p_k^m)})}{\sum_{k=1}^{N_m} w^{(p_k^m)} \mathrm{aff}(y^{(p_k^m)})}, \tag{3}$$

$$\sigma^{(m)}=\frac{\sum_{k=1}^{N_m} w^{(p_k^m)}|y^{(p_k^m)}-\mathrm{av}^{(m)}|\mathrm{aff}(y^{(p_k^m)})}{\sum_{k=1}^{N_m} w^{(p_k^m)}\mathrm{aff}(y^{(p_k^m)})}, \quad (4)$$

where $w^{(p_k^m)}$ is the rule weight. This combination is repeated for all groups, and thus the M rules are extracted.

Finally, the product inference engine, the singleton fuzzifier, and the center-average defuzzifier are used to construct the fuzzy model. The output is computed as

$$y=\frac{\sum_{m=1}^{M}\mathrm{av}^{(m)}\mu_{A_1^{l_1^{(m)}}}(x_1)\mu_{A_2^{l_2^{(m)}}}(x_2)}{\sum_{m=1}^{M}\mu_{A_1^{l_1^{(m)}}}(x_1)\mu_{A_2^{l_2^{(m)}}}(x_2)}. \quad (5)$$

Experiments. We conducted our experiment for prediction of Mackey-Glass chaotic time-series $\frac{\mathrm{d}x(t)}{\mathrm{d}t}=\frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)}-0.1x(t)$, where $x(0)=1.2$ and $\tau=17$. In this experiment, we used $x(t-6), x(t-5),\ldots,x(t)$ to predict $x(t+1)$. 800 points from $t=6$ to $t=805$ of the series are used as training data, and 300 points from $t=806$ to $t=1105$ are used as test data. $x(t-6),x(t-5),\ldots,x(t)$ are all divided into 7 fuzzy subsets with Gaussian membership functions. 6 outliers (Table S1) are given to test the robustness of WM method and our proposed method.

To expound the effect of the Gaussian distribution-based outlier detection to the fuzzy model, the membership functions are the same, regardless of whether the data base includes the outliers. The absolute value of the prediction error $|e|$ (Figure 1) and the output subset centers (Table S2) are given to evaluate the performances of the two methods. We also use the mean absolute percentage error $\mathrm{MAPE}=\frac{100}{N}\sum_{i=1}^{N}|e|$ as an evaluation criterion. With the WM method, MAPE changes from 2.43% to 2.60% after the outliers are added. With our proposed method, MAPE only changes from 2.32% to 2.33%. It is clear that the changes of the prediction results, output subset centers, and MAPE of the proposed method are smaller than those of the WM method. Therefore, we can affirm that the Gaussian distribution-based outlier detection can reduce the influence of the outliers, and improve the robustness of the WM method.

Conclusion. We developed a method to simplify the outlier detection with WM method in order to establish a fuzzy model with high robustness in a simple way. In this method, the data points are classified into different groups using the WM method. We can detect the outliers of the outputs in a group without considering any of the inputs because the inputs in the same group are in the same space. Thus, the multidimensional outlier detection problem is transformed into a one-dimensional outlier detection problem. The wights of outliers in the same group are small, so the influence of outliers is reduced. Results of the chaotic time-series prediction experiment prove that our proposed method improves the robustness of the fuzzy model.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61773354), Hubei Provincial Natural Science Foundation (Grant No. 2015CFA010), and Programme of Introducing Talents of Discipline to Universities (111 Project) (Grant No. B17040).

Supporting information Tables S1 and S2. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Setnes M. Supervised fuzzy clustering for rule extraction. IEEE Trans Fuzzy Syst, 2000, 8: 416–424
2 Palacios A M, Palacios J L, Sánchez L, et al. Genetic learning of the membership functions for mining fuzzy association rules from low quality data. Inf Sci, 2015, 295: 358–378
3 Hata R, Islam M M, Murase K. Quaternion neuro-fuzzy learning algorithm for generation of fuzzy rules. Neurocomputing, 2016, 216: 638–648
4 Wang L X, Mendel J M. Generating fuzzy rules by learning from examples. IEEE Trans Syst Man Cybern, 1992, 22: 1414–1427
5 Wang L X. The WM method completed: a flexible fuzzy system approach to data mining. IEEE Trans Fuzzy Syst, 2003, 11: 768–782
6 Gou J, Hou F, Chen W Y, et al. Improving Wang-Mendel method performance in fuzzy rules generation using the fuzzy C-means clustering algorithm. Neurocomputing, 2015, 151: 1293–1304