Generating of Fuzzy Rule Bases with Gaussian Parameters Optimized via Fuzzy C-Mean and Ordinary Least Square

Samingun Handoyo, Achmad Efendi

Abstract: *The fuzzy rule bases has a central role in the fuzzy inference system. Generating of rule bases based on input-output data pairs by using Fuzzy C-Mean clustering (FCM) requires parameters setting of the membership function (mf). In the Gaussian mf, the mean and spread parameters must be determined. The outputs of FCM clustering include the cluster center and the partition matrix of each object. The value of cluster center can be used as the center parameter, but the spread parameter is usually determined as a constant value. The research proposes a method to determine spread parameter of Gaussian mf by using Ordinary Least Square (OLS) approach with using the partition matrix as the membership degree of each object in the cluster. The magnitude of cluster centers (n) of 3, 5, and 7 are considered as FCM input to cluster the weekly price of soybeans in East Java, Indonesia on the period of January 2014 to December 2017. Based on each cluster center, the optimal spread value of a Gaussian mf is obtained via OLS. This research succeeded in getting a spread values that could approach almost perfectly each element of the partition matrix.*

Keywords: *Fuzzy C-Means Clustering, Fuzzy rule bases, Gaussian membership function parameters, Optimization*

I. INTRODUCTION

The main motivation of the Fuzzy Inference System (FIS) is to map an input space in the form of fuzzy numbers obtained from the fuzzification process to produce output in the form of crips number yielded by inference and defuzification process [1-3]. The FIS has been widely applied to solve problems in various fields including engineering [4-6], health [7], [8] and finance [9-11]. The inference process in the component inference engine uses Fuzzy Rule Bases (FRB) to produce an output. The main components of FIS are expressed as in Figure 1 as follows [1]:

Based on Figure 1, there is shown that the FIS is arranged of 4 main components including the process of fuzzification, inference, defuzzification and inference engine (FRB). In all components, it always involves the linguistic value and degree of membership in which these two characteristics are the main features of the fuzzy set. The degree of membership of an observation to a linguistic value can be computed when the parameters of the membership function of the fuzzy set have been determined.
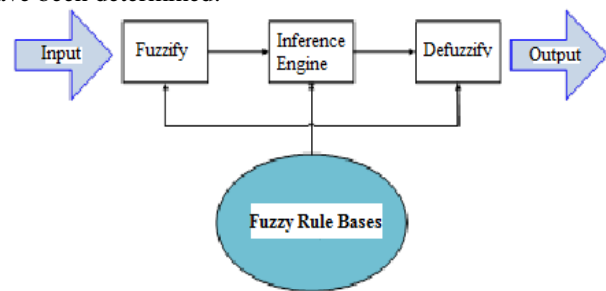

Figure 1. The Components of Fuzzy Inference System (FIS

An FRB which is a core component of FIS initially comes from experts or operators. The presence of many empirical data in the real world has led some researchers for developing methods in generating FRB which are the lookup table schemes by Wang and Mendel [12], heuristic method based on numerical data by Nozaki et al [13], and the association rule mining method by both of Kuang, and Li [14] and Arif, et al [15]. The methods of generating the FRB above produce an FRB with a large number of rules that result in a slow inference process which affects that the implementation of FIS will consume much time.

A lot of effort has been done by a group of researchers, among them by Yager, and Filev [16] who have generated an FRB with mountain clustering which produces many rules same as many clusters formed. The clustering method that is able to produce optimal clusters is needed in order to produce an FRB which is able to capture all the patterns of the data. The clustering method that groups objects fairly based on the greatest degree of membership of the object in a cluster is fuzzy c-mean [17-18]. The Fuzzy c-mean (FCM) can group objects quickly and robustly [19] and can also be used to group data in very large sizes [20]. Two main outputs of FCM are both of the cluster center and the membership degree of an object in each cluster. With regard to FRB, the cluster center formed as the center (mean parameter) of the Gaussian membership function (MF), but the spread parameter of the Gaussian MF is determined by a formula deterministically that result in the magnitude of spread contradiction with reality.

Samingun Handoyo*, Statistics Department, Brawijaya University, Malang, Indonesia. Email: samistat@ub.ac.id
Achmad Efendi, Statistics Department, Brawijaya University, Malang, Indonesia. Email: a_efendi@ub.ac.id

Estimation of spread parameters with optimization techniques has been widely discussed in literature such as by Rao and Bard [21], Pratama and Handoyo [22] estimated the Gaussian MF parameters with genetic algorithms,

Handoyo, et al [23] estimated the ARMA model parameters with particle swarm optimization (PSO), and estimation of normal distribution parameters and multiple linear regression model parameters can use the Ordinary Least Square (OLS) method as found in Heaney, and. Poitras [24], Lee [25], and also in Rosipal, and Krämer [26]. The OLS method has also been applied to the optimization of the consequent part of FRB generated with the lookup table scheme by Handoyo and Marji [11].

In this research proposed FRB generation method using FCM and OLS which has optimal Gaussian membership function parameters. The central parameter is obtained from the cluster center of the FCM output and the spread parameter is obtained through OLS optimization of the partition matrix of the FCM output which is the membership degree of the object in each cluster. In this study, it will also be shown that the more clusters that are formed, the greater of the spread parameters resulted through clustering data into 3, 5, and 7 clusters

II. RESEARCH METHOD

In this section is explained theoretically the methods used in the study including the discussion about fuzzy rule bases and Gaussian membership function, explained the concept of Fuzzy C-Means clustering, and the mathematical detail of the proposed method.

A. The Fuzzy Rules Base and Gaussian Membership Function

According to Wang [1], human knowledge can be represented in the form of the "IF-THEN" fuzzy rule. Propositions that follow the IF are called antecedent and the proposition that follows THEN is consequent. The general form of "IF-THEN" fuzzy rule is stated in the statement follows

The antecedent part is the proportion or composition of the proportions that follow the "IF" part, whereas the consequent part is the part following "THEN" in the form of a proposition or in the form of a linear equation or a constant. Consider a rule like the following:

IF <*fuzzy proposition*>, THEN <*fuzzy proposition*>   (1)

Or in more detail the form of the fuzzy rule of equation (1) is expressed as

IF $< x_1$ is A_1 and x_2 is A_j and x_3 is $A_3>$ THEN $< y$ is $A_j >$

The above rule can be interpreted as coming from a system consisting of 3 input variables and an output with each variable consisting of linguistic values of n. Any linguistic value will correspond to a membership function in order to the inference process on the inference engine can be performed.

According to Cannon, et al.[18], the membership function is a curve that shows the mapping of points of observation value into membership value which is often called the degree of membership which has interval 0 to 1. One way that can be used to get membership value is through a function approach one of which is the Gaussian membership function that has a bell-like curve which is determined by two parameters: central (mean) and spread. The Gaussian function is expressed in equation (2) below:

$$\mu_A(x) = exp\left(-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2\right) \tag{2}$$

where x is the crips value or the observed value of a variable, c is the center or the mean value and σ is the standard deviation (spread).

B. The Fuzzy C-Means (FCM) Clustering

Fuzzy C-Means Clustering (FCM) algorithm was introduced by Bezdek , et al [17]. The FCM has an objective function that is formulated with:

$$J(Y; U, C) = \sum_{h=1}^{q} \sum_{t=p^*+1}^{n} (\mu_{ht})^m \| y_t - c_h \|^2 \tag{3}$$

with tne constraints:

$$\sum_{h=1}^{q} \mu_{ht} = 1 \text{ and } t = p^* + 1, p^* + 2, \dots, n \tag{4}$$

Equations (3) and (4) can be combined into a Lagrange multiplier function as follows [1]:

$$L(Y; U, C, \lambda) =$$
$$\sum_{h=1}^{q} \sum_{t=p^*+1}^{n} (\mu_{ht})^m \| y_t - c_h \|^2 - \sum_{t=p^*+1}^{n} \lambda_t \left(\sum_{h=1}^{q} \mu_{ht} - 1 \right) \tag{5}$$

with $\| y_t - c_h \| = \sqrt{(y_t - c_h)'(y_t - c_h)}$,

where

Y	:	Observation data in a matrix structure
U	:	Partition matrix
C	:	Center value matrix that its columns are the cluster centers vector
λ	:	vector coefficient Lagrange
μ_{ht}	:	degree of tth observation membership on the h-cluster
m	:	exponent weights (fuzzy exponent), m> 1
y_t	:	observation vector (input-output) on t-observation with order $(p + 1) \times 1$
c_h	:	vector of center value (center cluster) for cluster h with order $(p + 1) \times 1$

q is the number of clusters assigned, so with a certain q value we will find μ_{ht} and c_h which minimize equation (3) with the constraint on equation (4). Through the Lagrange function in equation (5), μ_{ht} and c_h can be found by making the gradient $L(Y; U, C, \lambda)$ to μ_{ht} , c_h and λ_t equals to 0 as follows:

$$\frac{\partial L(Y;U,C,\lambda)}{\partial \mu_{ht}} = 0, \quad \frac{\partial L(Y;U,C,\lambda)}{\partial c_h} = 0, \text{ and } \frac{\partial L(Y;U,C,\lambda)}{\partial \lambda_t} = 0 \tag{6}$$

Through the mathematical decomposition process are obtained cluster center and degree of observation membership on each cluster as follows:

$$c_h = \frac{\sum_{t=p^*+1}^{n} (\mu_{ht})^m y_t}{\sum_{t=p^*+1}^{n} (\mu_{ht})^m} \tag{7}$$

$$\mu_{ht} = \frac{1}{\sum_{s=1}^{q} \left(\frac{\| y_t - c_h \|}{\| y_t - c_s \|}\right)^{\frac{2}{m-1}}} \tag{8}$$

The equations (7) and (8) are indirect solutions of Lagrange function optimization problems of the equation (5), so that c_h and μ_{ht} are solved by iteration using the FCM algorithm.

C. The Proposed Method

The parameters of Gaussian membership function based on (2) are central and spread. The output of FCM is the cluster center and degree of membership. The cluster center can be used as a central parameter, but the spread parameter is usually set to a constant value with the formula:

$$Spread_h = \frac{(y_{hmax} - y_{hmin})}{(n-1)} \tag{9}$$

Where y_{hmax} is the highest observation value at y_h and y_{hmin} is the lowest observation value at y_h, whereas n is the number of linguistic values or the number of fuzzy sets.

The spread parameter obtained by formula (9) will result in a degree of membership different from the degree of membership of the FCM output. Therefore, by writing the Gaussian membership function equal to the optimal partition value of the FCM output then it is obtained:

$$\mu_{ht} = exp\left(-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2\right) \qquad (10)$$

If there are as many as k attributes in each object, then the equation (1) can be expanded to become as follows:

$$\mu_{ht} = \prod_{i=1}^{k} exp\left(-\frac{1}{2}\left(\frac{x_i - c_i}{\sigma_i}\right)^2\right) \qquad (11)$$

The both sides of equation (11) are regarded as being obtained

$$\ln(\mu_{ht}) = \frac{1}{2}\ln\left(exp\left(-\left(\frac{x_1 - c_1}{\sigma_1}\right)^2\right)\right) +$$
$$\frac{1}{2}\ln\left(exp\left(-\left(\frac{x_2 - c_2}{\sigma_2}\right)^2\right)\right) + \cdots + \frac{1}{2}\ln\left(exp\left(-\left(\frac{x_k - c_k}{\sigma_k}\right)^2\right)\right)$$

$$\ln(\mu_{ht}) = -\frac{1}{2}\left(\left(\frac{x_1 - c_1}{\sigma_1}\right)^2 + \left(\frac{x_2 - c_2}{\sigma_2}\right)^2 + \cdots + \left(\frac{x_k - c_k}{\sigma_k}\right)^2\right)$$

$$2\ln(\mu_{ht}) = -\frac{1}{\sigma_1^2}(x_1 - c_1)^2 - \frac{1}{\sigma_2^2}(x_2 - c_2)^2 - \cdots - \frac{1}{\sigma_k^2}(x_k -$$

with substitution of $2\ln(\mu_{ht}) = y, -\frac{1}{\sigma_i^2} = b_i$ and $(x_i - c_i)^2 = x_i$, we get the equation (12) as follows

$$y = b_1 x_1 + b_2 x_2 + \cdots + b_k x_k \qquad (12)$$

If there are observed values of i = 1,2, ..., n, finally the equation (11) can be expressed in the matrix notation as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \cdots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$Y = X\theta + \varepsilon \text{ so that, } \varepsilon = Y - X\theta$$

By using OLS optimization which minimizes sum square of error through first derivative is elaborated as follows:

$$\frac{\partial \varepsilon^T \varepsilon}{\partial \theta} = 0,$$
$$\frac{\partial (Y - X\theta)^T (Y - X\theta)}{\partial \theta} = 0$$
$$\frac{\partial (Y^T Y - Y^T X\theta - \theta^T X^T Y + \theta^T X^T X\theta)}{\partial \theta} = 0$$
$$\frac{\partial (Y^T Y - 2X^T Y\theta^T + \theta^T X^T X\theta)}{\partial \theta} = 0$$
$$\frac{\partial (Y^T Y - 2X^T Y\theta^T + \theta^2 X^T X)}{\partial \theta} = 0$$
$$-2X^T Y + 2X^T X\theta = 0$$
$$\theta = (X^T X)^{-1} X^T Y$$

Refer to the equation (12), furthermore it is yeilded $\theta = [b_1, b_2, \dots, b_k]^T$, so that be obtained, $\sigma_i^2 = -\frac{1}{b_i}$. The optimal spread of Gaussian membership function is obtained as $spread = sqrt\left(abs\left(-\frac{1}{b_i}\right)\right)$

III. RESULTS AND ANALYSIS

In this section will be discussed the results obtained from this study include descriptions of data with the summary statistics, parameters setting deterministic of the Gaussian membership function, construct fuzzy-rule bases using FCM

compute optimal spread parameters by OLS, and construct fuzzy rule bases with the optimal of both mean and spread parameters .

A. The Data Description

Weekly data on soybean prices in East Java Province of Indonesia from March 2014 to December 2017 were used as the case of study. The data were obtained from the East Java BULOG Corporation, which publishes weekly soybean prices on the http://siskaperbapo.com.

Based on Figure 2. the weekly prices of soybeans in East Java have a negative trend with respect to time, it continues to decline gradually. This price reduction was caused by the government policy to maintain some small industries that use soybean as the raw materials. In this study, a data structure will be arranged in the form of an input-output data pairs.in the matrix data structure assuming that there are several lags influencing the soybean price at the current time that will be considered for the number of lag = 3, 5, and 7. Based on the input-output data pairs, it also will be clustered the data being the cluster number (cl) of 3, 5, and 7. Table 1 below is a statistical summary for the number of lags that are assumed influential, namely k = 3.
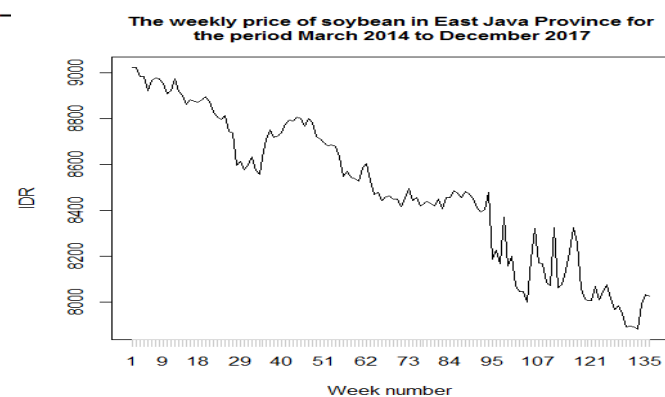

Figure 2. The weekly price of soybean commodity in East Java province of Indonesia

Based on Table 1. it is clear that the statistical values of each variable in the leftmost column are all almost the same because they come from time series data where each variable is only 3 different values.

Table 1. Summary of weekly soybean prices for lag of k = 3

	Min.	Max.	Mean	Std.
Yt-1	7885	8983	8479.81	310.74
Yt-2	7885	9021	8487.23	311.79
Yt-3	7885	9021	8494.97	312.18
Yt	7885	8983	8472.63	310.08

B. Parameters setting of the Gaussian membership function

The membership function has an important role in the process of inference in fuzzy systems because the membership function will determine the amount of fire strength of each rule on the rule bases. One of the most frequently used membership functions is the Gaussian [11] membership function.

Generating of Fuzzy Rule Bases with Gaussian Parameters Optimized via Fuzzy C-Mean and Ordinary Least Square

The bell-shaped Gaussian curve is characterized by two parameters, namely the center of the curve and spread (standard of deviation).

In some FIS implementations such as in Handoyo and Marji [9], Handoyo et al [10], Arief et al [15] and Wang [1], the two parameters are determined in a simple way. The spread parameter is a constant which amount is calculated that divides range by (n-1), where n is the number of fuzzy sets, while the curve centers are obtained through the minimum value plus k * the spread value, for k = 0,1,2,3, ..., (n-1), in this case, the minimum value automatically becomes the center of the lowest curve, and the maximum value automatically becomes the center of the highest one.

The values of both parameters (the mean and spread) of weekly soybean price data in East Java are presented in table 2, where the number of fuzzy sets are n = 3. 5, and 7. The corresponding Gaussian curves are given in Figure 3.

The spread and mean parameter values determined in this way are not affected by the number of lags considered as input from the system. All input and output variables will have spread parameters and the same mean. Visually the Gaussian curves for n = 3, 5, and 7 are presented in Figure 3 as follows.

Table 2. Mean and spread parameters where the cluster numbers of 3, 5, and 7

cluster	n=3		n=5		n=7	
	Std	mean	Std	Mean	std	mean
1	284	7885	142	7885	95	7885
2	284	8453	142	8169	95	8074
3	284	9021	142	8453	95	8264
4			142	8737	95	8453
5			142	9021	95	8642
6					95	8831
7					95	9021

Based on Figure 3, the resulting Gaussian curve is very proportional and seems regular. But actually, there is a contraction state that is by increasing fuzzy sets or increasing the Gaussian curve will be followed by a decrease in the observed value represented by each curve. The spread parameters are greatly influenced by the number of observations. Thus the value of the spread parameter should increase with the increasing number of fuzzy sets.
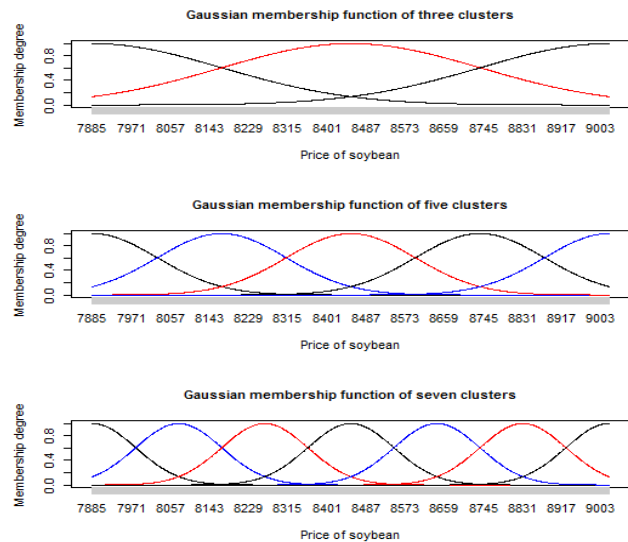
Figure 3.The Gaussian curves with the cluster numbers of 3, 5, and 7

C. Establishment of fuzzy rule bases using FCM

The fuzzy rule base generated by the lookup table scheme such as in Handoyo and Marji [9] tends to produce a large number of rules. Of course, the more basic rules on the basis of fuzzy rules will have the effect of slowing down the FIS to produce output because the fuzzification process and its inference will take a long time. One of ways to generate a FRB with a small number of rules is to use FCM. The generating FRB with FCM will produce a fewer number of rules whose numbers are always the same as the number of clusters specified. In addition, the FCM method also produces cluster centers that can be used as the mean parameter of the Gaussian membership function.

Table 3. The cluster centers with the lag numbers of 3 and 5 where the cluster numbers of 5 and 7

Cl	Numbers of lag=3					Numbers of lag=5						
	Yt-1	Yt-2	Yt-3	Yt	Memb.	Yt-1	Yt-2	Yt-3	Yt-4	Yt-5	Yt	Memb
1	8908	8917	8925	8900	21	8771	8780	8783	8784	8783	8758	17
2	8176	8197	8200	8154	22	8653	8666	8681	8697	8712	8638	13
3	8002	8002	8012	8004	19	8555	8561	8567	8575	8590	8551	13
4	8458	8463	8468	8455	38	8913	8921	8928	8937	8945	8908	18
5	8707	8715	8720	8697	33	8139	8164	8178	8186	8176	8121	23
6						8439	8443	8446	8450	8454	8437	32
7						7997	7996	7999	8011	8029	7996	15

Based on table 3, it is exposed that each input has a different of cluster center, so the Gaussian function also varied, but the spread parameter of each cluster is determined in the same way as in the previous session, namely bases the spread value of n=5 is 142 and n= 7 is 95. In addition, in Table 3 there is also information about the total of members in each cluster. If the number of clusters increases, it will decrease the total of members of each cluster. The previous information leads to a contradiction when a setting of smaller spread value will increase the number of clusters. Because the smaller of total of cluster members will cause the spread value to be even greater. The Figure 4 shows the Gaussian membership function of the variable Yt where the number of clusters are 3, 5, and 7.
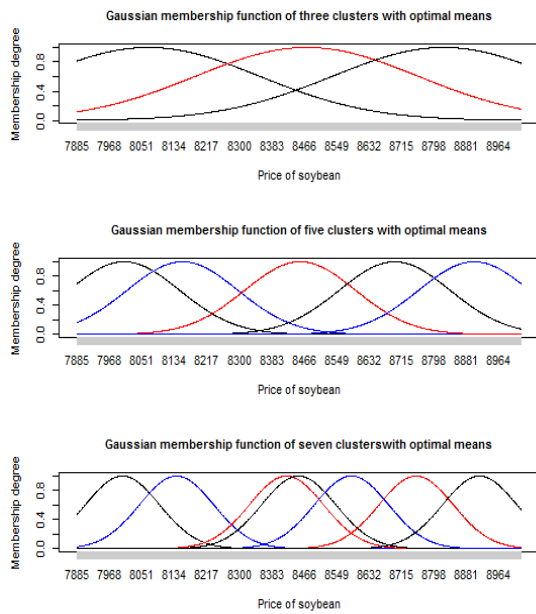
The main difference of the Gaussian curves in Figure 4 is compared in Figure 3 that the center of the first and the last curves are not the minimum and the maximum value. In addition, the center of the curves almost coincides such as in the bottom curve. More realistic than the Gaussian curves whose center of the curve is determined by a specific formula such as the previous Gaussian curves in Figure 3.

D. The optimized spread parameters use OLS

As explained in the previous section, the determination of the magnitude of the spread which is getting smaller when the increasing number of clusters leads to a contradiction. The correct logic is that the increasing of the cluster number will be followed by the increasing of the spread value. In this study, each combination has a lot of lag (nlag) determined and many clusters (ncl) are calculated according to the value of the corresponding spread. As an illustration for the combination of nlag = 3 with ncl = 5, and nlag = 5 with ncl = 7, the spread values in table 4 are obtained.
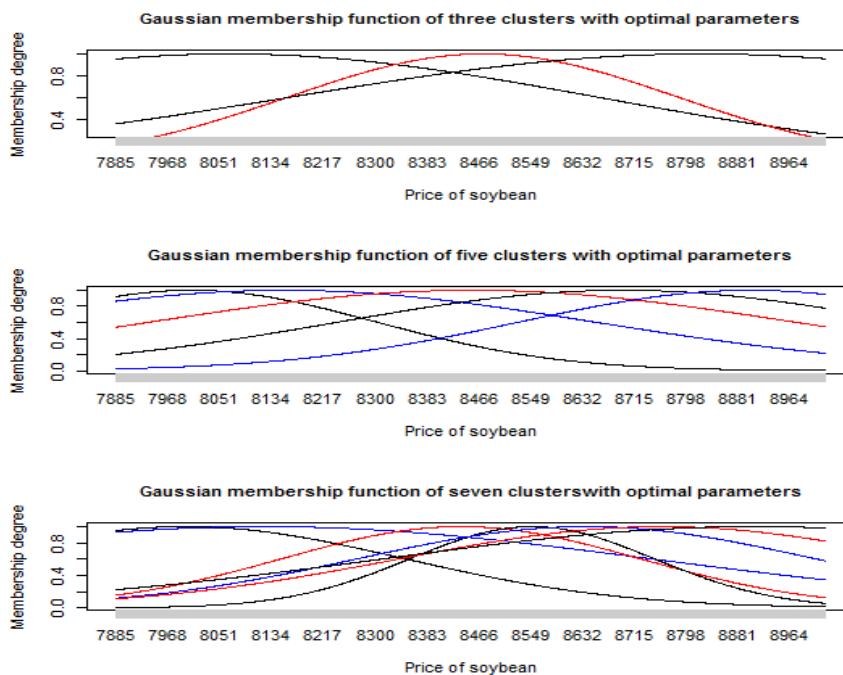
Figure 4. The Gaussian curves where the numbers of clusters are 3, 5, and 7 where the means are optimal

Table 4. Spread values on lag numbers of 3 and 5 with cluster numbers of 5 and 7

Cl	Number of lags=3					Number of lags=5						
	Yt-1	Yt-2	Yt-3	Yt	Memb.	Yt-1	Yt-2	Yt-3	Yt-4	Yt-5	Yt	Memb.
1	434	490	407	368	21	717	1004	943	1097	400	417	17
2	1184	513	503	495	22	2568	434	377	267	2167	365	13
3	568	459	247	295	19	422	1114	359	670	390	195	13
4	662	271	690	514	38	556	781	427	438	668	592	18
5	519	605	448	454	33	740	2446	931	1685	455	617	23
6						373	800	718	490	307	286	32
7						466	423	863	1486	268	350	15

Figure 5. The Gaussian curves with optimized mean and spread parameters on the cluster number of 3, 5, and 7

Based on table 4, it can be seen that there are 20 spread values in the nlag = 3 and ncl = 5, while in the nlag = 5 amd ncl = 7, the number of spread values are 42. The spread values are getting larger when the number of clusters increases. In the nlag = 3 and ncl = 5, the largest spread value is 1184 which is on Yt-1 variable of cluster 2. In the nlag = 5 and ncl = 7, the largest spread value is 2568 on Yt-1 variable of cluster 2. The spread values described in Table 4 show clearly that the setting of the spread values with the specific constant for each cluster on all lag variables is a wrong decision because there is a contradiction that the spread value gets smaller if the number of cluster increases. The following are the results of plotting Gaussian functions for ncl = 3.5, and 7 with optimal spread values.

Based on Figure 5, the Gaussian curves generated on the ncl= 7 overlap each other, so that cannot specific represent the observed values. At then ncl = 3 and ncl = 5 visually can still be recognized the Gaussian curves of each cluster. The result is supported by the FIS implementation carried out by Handoyo and Marji [9] that the best-performing FIS is on ncl = 3 or 5.

E. Fuzzy rule bases with optimal parameters

The formation of a fuzzy rule bases with FCM is done by sorting the cluster centers of each lag variable. The ranking number of a lag variable states the linguistic value of the lag variable. The following is an illustration of the formation of FRB in nlag = 5 with ncl = 5. Suppose that linguistic terms are Low (L), Normal (N), and High (H) are used. The cluster numbers of each cluster center in left side of Table 5 can be converted into linguistic values, i.e.

.

$$1 = very\ Low(L3), 2 = moderate\ Low(L2), 3 = Low(L1), 4 = Normal(N)$$
$$5 = High(H1), 6 = moderate\ High(H2), 7 = very\ High(H3)$$

Thus FRBs formed are

1. IF (Y_{t-1} is H2 and Y_{t-2} is H2 and Y_{t-3} is H2 and Y_{t-4} is H2 and Y_{t-5} is H2) THEN (Y_t is H2)
2. IF (Y_{t-1} is H1 and Y_{t-2} is H1 and Y_{t-3} is H1 and Y_{t-4} is H1 and Y_{t-5} is H1) THEN (Y_t is H1)
3. IF (Y_{t-1} is N and Y_{t-2} is N and Y_{t-3} is N and Y_{t-4} is N and Y_{t-5} is N) THEN (Y_t is N)
4. IF (Y_{t-1} is H3 and Y_{t-2} is H3 and Y_{t-3} is H3 and Y_{t-4} is H3 and Y_{t-5} is H3) THEN (Y_t is H3)
5. IF (Y_{t-1} is L2 and Y_{t-2} is L2 and Y_{t-3} is L2 and Y_{t-4} is L2 and Y_{t-5} is L2) THEN (Y_t is L2)
6. IF (Y_{t-1} is L1 and Y_{t-2} is L1 and Y_{t-3} is L1 and Y_{t-4} is L1 and Y_{t-5} is L1) THEN (Y_t is L1)
7. F (Y_{t-1} is L3 and Y_{t-2} is L3 and Y_{t-3} is L3 and Y_{t-4} is L3 and Y_{t-5} is L3) THEN (Y_t is L3)

Table 5. The cluster centers and the appropriate linguistic values of nlag = 5 and ncl =7

ncl	Cluster centers						Cluster numbers					
	Yt-1	Yt-2	Yt-3	Yt-4	Yt-5	Yt	Yt-1	Yt-2	Yt-3	Yt-4	Yt-5	Yt
1	8771	8780	8783	8784	8783	8758	6	6	6	6	6	6
2	8653	8666	8681	8697	8712	8638	5	5	5	5	5	5
3	8555	8561	8567	8575	8590	8551	4	4	4	4	4	4
4	8913	8921	8928	8937	8945	8908	7	7	7	7	7	7
5	8139	8164	8178	8186	8176	8121	2	2	2	2	2	2
6	8439	8443	8446	8450	8454	8437	3	3	3	3	3	3
7	7997	7996	7999	8011	8029	7996	1	1	1	1	1	1

The Gaussian membership function of each rule uses the mean parameters in Table 5 and the spread parameters in Table 4 for the cluster number and the appropriate lag number. Suppose the rule number seven if stated in full with the membership function is as follows:

$$IF \left(Y_{t-1}\ is\ \mu_{L3} = exp\left(-\frac{1}{2}\left(\frac{x-7997}{466}\right)^2\right)\ and\ Y_{t-2}\ is\ \mu_{L3} = exp\left(-\frac{1}{2}\left(\frac{x-7996}{423}\right)^2\right)\ and\ Y_{t-3}\ is\ \mu_{L3}\right.$$
$$= exp\left(-\frac{1}{2}\left(\frac{x-7999}{863}\right)^2\right)\ and\ Y_{t-4}\ is\ \mu_{L3} = exp\left(-\frac{1}{2}\left(\frac{x-8011}{1486}\right)^2\right)\ and\ Y_{t-5}\ is\ \mu_{L3}$$
$$\left. = exp\left(-\frac{1}{2}\left(\frac{x-8029}{268}\right)^2\right)\right)\ THEN\ \left(Y_t\ is\ \mu_{L3} = exp\left(-\frac{1}{2}\left(\frac{x-7996}{350}\right)^2\right)\right)$$

It is clear that although the linguistic value of rule number seven is all very Low (L3) but has a different membership function, which has both different mean and spread values. The same thing also happens to rule number one to number 6. Generating FRB in this way for nlag = 5 and ncl = 7 will be obtained as many as 42 unique membership functions. The main difference with the generation of FRB with FCM is that this method (FCM and OLS Optimization) employs 42 parameter spread values while at FCM only one spread value is employed for 42 Gaussian functions. Rule number 7 if using FCM is as follows:

$$IF\left(Y_{t-1} \text{ is } \mu_{L3} = exp\left(-\frac{1}{2}\left(\frac{x-7997}{95}\right)^2\right) \text{ and } Y_{t-2} \text{ is } \mu_{L3} = exp\left(-\frac{1}{2}\left(\frac{x-7996}{95}\right)^2\right) \text{ and } Y_{t-3} \text{ is } \mu_{L3}\right.$$

$$= exp\left(-\frac{1}{2}\left(\frac{x-7999}{95}\right)^2\right) \text{ and } Y_{t-4} \text{ is } \mu_{L3} = exp\left(-\frac{1}{2}\left(\frac{x-8011}{95}\right)^2\right) \text{ and } Y_{t-5} \text{ is } \mu_{L3}$$

$$\left.= exp\left(-\frac{1}{2}\left(\frac{x-8029}{95}\right)^2\right)\right) THEN \left(Y_t \text{ is } \mu_{L3} = exp\left(-\frac{1}{2}\left(\frac{x-7996}{95}\right)^2\right)\right)$$

More extreme is the use of the Gaussian membership functions found in Handoyo and Marji [9] where each linguistic values only has a deterministic parameters of the Gaussian membership function. In the case, the rule number seven above can be stated as follows:

$$IF\begin{pmatrix} Y_{t-1} \text{ is } L3 \text{ and } Y_{t-2} \text{ is } L3 \text{ and} \\ Y_{t-3} \text{ is} \\ L3 \text{ and } Y_{t-4} \text{ is} \\ L3 \text{ and } Y_{t-5} \text{ is } L3 \end{pmatrix} THEN (Y_t \text{ is } L3)$$

Where the membership function of very Low (L3) linguistic value is

$$\mu_{L3} = exp\left(-\frac{1}{2}\left(\frac{x-7885}{95}\right)^2\right)$$

If we observe the rules in equation (13), we consider more aspects of the pattern found in the data. Thus the basis of the rule generated by FCM and OLS optimized to obtain spread parameters will be able to capture data patterns more accurately so that it is very possible to obtain FIS that has better performance.

IV. CONCLUSION

The generating fuzzy rule bases using FCM and OLS generates unique spread parameter for each linguistic values, so a Gaussian membership function will be formed that is unique for both the mean and spread parameters. The determining parameters of spread with OLS can eliminate the contradiction in the generating FBR The resulting FBR is more realistic so it is hoped that be able to support FIS performing to be better.

REFERENCES

1. L.X. Wang. A Course Fuzzy Systems and Contro., New Jersey. Prentice Hall PTR. 1997.
2. G. Chen, and T. T. Pham. Introduction to fuzzy sets, fuzzy logic, and fuzzy control systems. CRC press, 2000.
3. S. Handoyo, A P S Prasojo. Applied Fuzzy System with R Software (Sistem Fuzzy terapan dengan Software R). First Edition. Malang. UBPress. 2017.
4. D.M. Atia, et al. "Modeling and control PV-wind hybrid system based on fuzzy logic control technique." Telkomnika 10.3 (2012): 431-441.
5. A. Azriyenni , M. M. Wazir, Z. Naila, Fuzzy Neural Network for Classification Fault In Protection System. Indonesian Journal of Electrical Engineering and Computer Science. 2014; 12(8): 5969-5975.
6. A.B. Khizer, , A. M. Soomro D. Yaping, and X. X. Yang. "Identification of Nonlinear System Based on Fuzzy Model with Enhanced Gradient Search." TELKOMNIKA Indonesian Journal of Electrical Engineering 12.7 (2014): 5261-5267.
7. I.G.M.N.A. Yasa, I.K.G.D. Putra, and N.M.I.M. Mandenni. "Hypertension expert system with c5. 0 algorithm and fuzzy logic." Telkomnika Indonesian Journal of Electrical Engineering 12.7 (2014): 5669-5677.
8. S. Handoyo, and H. Kusdarwati. "Implementation of Fuzzy Inference System for Classification of Dengue Fever on the villages in Malang". In IOP Conference Series: Materials Science and Engineering . 2019. (vol. 546, no. 5, p. 052038).
9. S. Handoyo, and Marji. The Fuzzy Inference System with Least Square Optimization for Time Series Forecasting. Indonesian Journal of Electrical Engineering and Computer Science. 2018; 11(3): 1015-1026.
10. S. Handoyo, Marji, I.N. Purwanto, and F. Jie. "The Fuzzy Inference System with Rule Bases Generated by using the Fuzzy C-Means to Predict Regional Minimum Wage in Indonesia". The International Journal of Operation and Quantitatve Management . 2018. 24(4): 101-116
11. A. Efendi, S. Handoyo, A.P.S. Prasojo, and Marji. "The Implementation Of The Optimal Rule Bases Generated By Hybrid Fuzzy C-Mean And Particle Swarm Optimization", " Journal of Theoretical & Applied Information Technology". 2019. 97(16): 4453-4453.
12. L.X. Wang, and J.M. Mendel. "Generating fuzzy rules by learning from examples." IEEE Transactions on systems, man, and cybernetics 22.6 (1992): 1414-1427.
13. K. Nozaki, H. Ishibuchi, and H. Tanaka. "A simple but powerful heuristic method for generating fuzzy rules from numerical data." Fuzzy sets and systems 86.3 (1997): 251-270.
14. G. Kuang, and Y. Li. "Using fuzzy association rules to design e-commerce personalized recommendation system." TELKOMNIKA Indonesian J. Elec. Engin 12.2 (2014): 1519-1527.
15. M. F. A,rif, B Anoraga, S. Handoyo, H. Natsir. Algorithm Apriori Association Rule in Determination of Fuzzy Rule Based on Comparison of Fuzzy Inference System (FIS) Mamdani Method and Sugeno Method. Business Management and Strategy. 2016; 7(1): 103-124.
16. R.R. Yager, and D. P. Filev. "Generation of fuzzy rules by mountain clustering." Journal of Intelligent & Fuzzy Systems 2.3 (1994): 209-219.
17. J.C. Bezdek, R. Ehrlich, and W. Full. "FCM: The fuzzy c-means clustering algorithm." Computers & Geosciences 10.2-3 (1984): 191-203.
18. R.L. Cannon, V.D. Jitendra, and J.C. Bezdek. "Efficient implementation of the fuzzy c-means clustering algorithms." IEEE transactions on pattern analysis and machine intelligence 2 (1986): 248-255.
19. W. Cai, S. Chen, and D. Zhang. "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation." Pattern recognition 40.3 (2007): 825-838.
20. T.C. Havens, et al. "Fuzzy c-means algorithms for very large data." IEEE Transactions on Fuzzy Systems 20.6 (2012): 1130-1146.
21. S.S. Rao, and J. Bard. Engineering optimization: theory and practice. New York, NY, USA: John Wiley & Sons, 1997.
22. A.W. Pratama, S. Handoyo. Optimasi Fungsi Keanggotaan Sistem Fuzzy Time Series Dengan Algoritma Genetika (Studi Kasus Harga Harian Saham Bank CIMB Niaga). Jurnal Mahasiswa Statistik. 2015; 3(3): 169-172.
23. S. Handoyo, A. Efendi, F. Jie, A. Widodo. Implementation of particle swarm optimization (PSO) algorithm for estimating parameter of arma model via maximum likelihood method. Far East Journal of Mathematical Sciences. 2017;102(7):1337-1363.
24. J. Heaney, and G. Poitras. "Estimation of the optimal hedge ratio, expected utility, and ordinary least squares regression." The Journal of Futures Markets (1986-1998) 11.5 (1991): 603-615.
25. T. Lee. "On algorithms for ordinary least squares regression spline fitting: a comparative study." Journal of statistical computation and simulation 72.8 (2002): 647-663.
26. R. Rosipal, and N. Krämer. "Overview and recent advances in partial least squares." International Statistical and Optimization Perspectives Workshop" Subspace, Latent Structure and Feature Selection". Springer, Berlin, Heidelberg, 2005.

Retrieval Number: D8561118419/2019©BEIESP
DOI:10.35940/ijrte.D85611.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

5793

Generating of Fuzzy Rule Bases with Gaussian Parameters Optimized via Fuzzy C-Mean and Ordinary Least Square

AUTHORS PROFILE

Samingun Handoyo is a lecturer and a researcher at Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya, Malang, Indonesia. He is the member both of the Indonesian Mathematical Society (IndoMS) and the International Association of Engineers (IAENG). He received the B.S. degree in Mathematics from Universitas Brawijaya, Malang, Indonesia in 1997 and the M.Cs degree in Computer Science from Universitas Gadjah Mada, Yogyakarta, Indonesia in 2010. His research interests include statistical computing, Neural Network, Fuzzy System, and Partial Least Square path modeling using R Software.

Achmad Efendi is a lecturer and a researcher at Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya, Malang, Indonesia. He is of the Indonesian Mathematical Society (IndoMS). He received the B.S. degree in Statistics from Sepuluh Nopember Institute of Technology (ITS), Surabaya, Indonesia in 2004. And then, he received Master of Science (M.Sc) and Doctoral (Ph.D) in Biostatistics from CenStat, Universiteit Hasselt and L-Biostat, Katholieke Universiteit Leuven, Belgium respectively, in 2009 and 2013. His research interests include computational statistics, mixed modeling, biostatistics, econometrics, and statistical modeling with R.