Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Temperature Prediction Based on Fuzzy Clustering and Fuzzy Rules Interpolation Techniques

Yu-Chuan Chang[1] and Shyi-Ming Chen[1, 2]

[1]Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R. O. C.

[2]Department of Computer Science and Information Engineering, Jinwen University of Science and Technology, Taipei County, Taiwan, R. O. C.

***Abstract*—In this paper, we present a new method to deal with temperature prediction based on fuzzy clustering and fuzzy rules interpolation techniques. First, the proposed method constructs fuzzy rules from training samples based on the fuzzy C-Means clustering algorithm, where each fuzzy rule corresponds to a cluster and the linguistic terms appearing in the fuzzy rules are represented by triangular fuzzy sets. Then, it performs fuzzy inference based on the multiple fuzzy rules interpolation scheme, where it calculates the weight of each fuzzy rule with respect to the input observation based on the defuzzified values of triangular fuzzy sets. Finally, it uses the weight of each fuzzy rule to calculate the forecasted output. We also apply the proposed method to handle the temperature prediction problem. The experimental result shows that the proposed method gets higher average forecasting accuracy rates than Chen and Hwang's method [7].**

***Keywords*—fuzzy rules, temperature prediction, fuzzy clustering, fuzzy rules interpolation**

I. INTRODUCTION

Forecasting activities play an important role in our daily life, where there are many kinds of forecasting activities, such as stock market forecasting, earthquake forecasting, traffic flow forecasting, weather forecasting, economic growth rate forecasting, enrollments forecasting, etc. If we can make a forecast as precise as possible, we can prevent damages from the coming disasters, such as economic recession, company loss, traffic jam, storms, typhoons, etc. In [18], Zadeh presented the concepts of linguistic variables and fuzzy rules to approximate reasoning. In recent years, some forecasting methods have been presented based on fuzzy rules [2], [4], [5], [7], [10], [12], [13], [15].

However, fuzzy rule-based systems suffer from the problem of sparse fuzzy rule bases in which fuzzy rules incompletely cover the universe of discourse. Fuzzy rules in rule-based systems are usually limited to a few input variables, because a complete fuzzy rule base with K input variables and T fuzzy linguistic terms in each input variable needs T^K fuzzy rules, where the complexity of the rule base is exponentially increasing with the number of input variables. In order to increase the efficiency of fuzzy rule-based systems with multiple variables, it is necessary to reduce bigger fuzzy rule bases into smaller fuzzy rule bases while keeping the essential fuzzy rules in the rule bases. However, reducing fuzzy rule bases will cause sparse fuzzy rule bases which contain blank areas uncovered by fuzzy rules in the universe of discourse while conventional fuzzy inference methods only can handle complete fuzzy rule bases [14]. In recent years, some fuzzy rules interpolation methods [6], [8], [11], [16], [17] have been presented to handle inferences in sparse fuzzy rule bases for sparse fuzzy rule-based systems.

In this paper, we present a new method to deal with temperature prediction based on fuzzy clustering and fuzzy rules interpolation techniques. First, the proposed method constructs fuzzy rules from training samples based on the fuzzy C-Means clustering algorithm [1], where each fuzzy rule corresponds to a cluster and the linguistic terms appearing in the fuzzy rules are represented by triangular fuzzy sets. Then, it performs fuzzy inference based on the multiple fuzzy rules interpolation scheme [6], where it calculates the weight of each fuzzy rule with respect to the input observation based on the defuzzified values [9] of triangular fuzzy sets and uses the weight of each fuzzy rule to calculate the forecasted output. We also apply the proposed method to handle the temperature prediction problem. The experimental result shows that the proposed method gets higher average forecasting accuracy rates than Chen and Hwang's method [7].

II. PRELIMINARIES

A. *Fuzzy C-Means Clustering Algorithm* [1]

The fuzzy C-Means (FCM) clustering algorithm [1] is a widely used fuzzy clustering method in pattern recognition, which allows each data belonging to two or more clusters. The FCM clustering algorithm partitions data points X_j (j = 1, 2, …, n) into clusters C_i (i = 1, 2, …, c) based on the minimization of the following objective function [1]:

$$J_m = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{\ m}\left\| V_i - X_j \right\|^2, \tag{1}$$

where $\left\| V_i - X_j \right\|$ is the Euclidean distance between data point X_j and the cluster center V_i (i = 1, 2, …, c), u_{ij} is the membership grade of X_j belonging to cluster C_i, m is a fuzziness index [1], $m \geq 1$, n is the number of data points, and c is the number of clusters. The procedures of the FCM clustering algorithm are reviewed from [1] as follows:

Step 1: Randomize the membership grade u_{ij}, where $0 \leq u_{ij} \leq 1$, $\sum_{i=1}^{c} u_{ij} = 1$, $1 \leq i \leq c$ and $1 \leq j \leq n$.

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

Step 2: Calculate the cluster center V_i of cluster C_i,

$$V_i = \frac{\sum_{j=1}^{n} (u_{ij})^m \times X_j}{\sum_{j=1}^{n} (u_{ij})^m}, \tag{2}$$

where $1 \le i \le c$.

Step 3: Update the membership grade u_{ij} of X_j belonging to C_i, where

$$u_{ij} = \frac{1}{\sum_{d=1}^{c} \left(\frac{\| V_i - X_j \|}{\| V_d - X_j \|} \right)^{\frac{2}{m-1}}}, \tag{3}$$

$1 \le i \le c$ and $1 \le j \le n$.

Step 4: Repeat **Step 2** and **Step 3** until the value of J_m in Eq. (1) is no longer decreasing.

B. The Multiple Fuzzy Rules Interpolation Scheme [6]

A triangular fuzzy set A can be represented by three characteristic points (a, b, c), as shown in Fig. 1, where b is called "the center point" whose membership value in A is equal to 1, and a and c are called "the left point" and "the right point", respectively, whose membership values in A are equal to 0, respectively. The defuzzified value DEF(A) of the triangular fuzzy set A shown in Figure 1 is calculated as follows [9]:

$$\mathrm{DEF}(A) = \frac{a + 2 \times b + c}{4}. \tag{4}$$

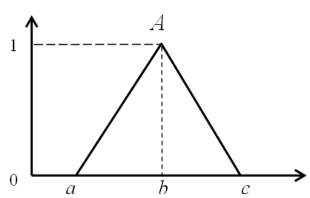

Figure 1. A triangular fuzzy set A.

Let us consider the multiple fuzzy rules interpolation scheme, which is shown as follows:

Rule 1: If $x_1 = A_{11}$ and $x_2 = A_{21}$ and … and $x_k = A_{h1}$ Then $y = B_1$
Rule 2: If $x_1 = A_{12}$ and $x_2 = A_{22}$ and … and $x_k = A_{h2}$ Then $y = B_2$
⋮
Rule p: If $x_1 = A_{1p}$ and $x_2 = A_{2p}$ and … and $x_k = A_{hp}$ Then $y = B_p$
Observations: $x_1 = A_1^*$ and $x_2 = A_2^*$ and … and $x_h = A_h^*$

Conclusion: $y = B^*$

where *Rule i* (i = 1, 2, .., p) is the ith fuzzy rule in the sparse fuzzy rule base, x_k denotes the kth antecedent variable (k = 1, 2, …, h), y denotes the consequence variable, A_{ki} denotes the kth antecedent fuzzy set of *Rule i*, B_i denotes the consequence fuzzy set of *Rule i*, A_k^* denotes the kth observation fuzzy set for the kth antecedent variable x_k, and B^* denotes the interpolated consequence fuzzy set. Figure 2 shows an example of the multiple fuzzy rules interpolation scheme with two antecedent variables using triangular fuzzy sets.

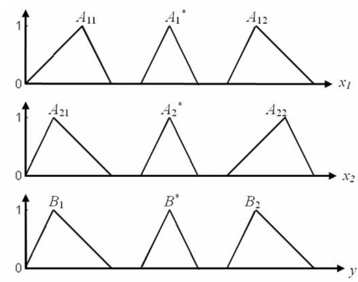

Figure 2. Multiple fuzzy rules interpolation scheme with two fuzzy rules using triangular fuzzy sets.

III. The Proposed Method for Handling Forecasting Problems Based on Fuzzy Clustering and Fuzzy Rules Interpolation Techniques

In this section, we present a new method to handle forecasting problems based on fuzzy clustering and fuzzy rules interpolation techniques. The flowchart of the proposed algorithm is shown in Figure 3.

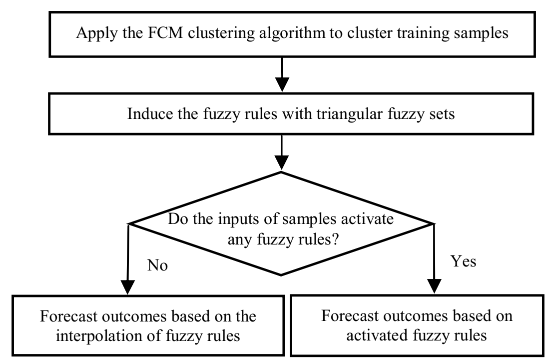

Figure 3. The flowchart of the proposed algorithm.

Assume that there is a forecasting dataset having n training samples X_1, X_2, …, and X_n, where the jth sample X_j is represented by ($I_j^{(1)}, I_j^{(2)}$, …, $I_j^{(h)}$, O_j), $I_j^{(k)}$ is the kth input of X_j, $1 \le k \le h$, and O_j is the desired output of X_j. Let c denote the number of clusters. The proposed algorithm to construct fuzzy rules from training samples is now presented as follows:

Step 1: Apply the FCM clustering algorithm [1] to update the cluster center V_i of cluster i and the membership grade u_{ij} of sample X_j belonging to cluster C_i based on Eq. (2) and Eq. (3), respectively, until the objective function J_m in Eq. (1) is no longer decreasing, where $1 \le i \le c$, $1 \le j \le n$, and the fuzziness index $m = 2$.

Step 2: Based on the clusters C_1, C_2, … and C_c obtained in **Step 1**, construct fuzzy rules *Rule* 1, *Rule* 2, …, and *Rule c* using triangular fuzzy sets, where *Rule i* corresponds to cluster C_i, shown as follows:

Rule i: If x_1 is A_{1i} and x_2 is A_{2i} and … and x_h is A_{hi} Then y is B_i,

where x_k is the kth antecedent variable (k = 1, 2, …, h), A_{ki} is the kth antecedent fuzzy set of *Rule i*, y is the consequence variable, B_i is the consequence fuzzy set of *Rule i*; the center point b_{ki}, the left point a_{ki} and the right point c_{ki} of the triangular fuzzy set A_{ki} of *Rule i* are calculated as follows:

$$b_{ki} = I_t^{(k)}, \quad \text{where } u_{it} = \max_{1 \le j \le n} u_{ij}, \tag{5}$$

$$a_{ki} = \frac{\sum_{j=1,2,\ldots,n \text{ and } I_j^{(k)} \le b_{ki}} u_{ij} \times I_j^{(k)}}{\sum_{j=1,2,\ldots,n \text{ and } I_j^{(k)} \le b_{ki}} u_{ij}}, \tag{6}$$

$$c_{ki} = \frac{\sum_{j=1,2,\ldots,n \text{ and } I_j^{(k)} \ge b_{ki}} u_{ij} \times I_j^{(k)}}{\sum_{j=1,2,\ldots,n \text{ and } I_j^{(k)} \ge b_{ki}} u_{ij}}, \tag{7}$$

where b_{ki} is the center point having the membership value of 1 in A_{ki}, a_{ki} and c_{ki} are the left point and the right point having the membership value of 0 in A_{ki}, $I_j^{(k)}$ is the kth input of X_j, $1 \le k \le h$ and $1 \le i \le c$; and similarly, the center point b_i, the left point a_i and the right point c_i of the triangular fuzzy set B_i of *Rule i* are calculated as follows:

$$b_i = O_t, \quad \text{where } u_{it} = \max_{1 \le j \le n} u_{ij}, \tag{8}$$

$$a_i = \frac{\sum_{j=1,2,\ldots,n \text{ and } O_j \le b_i} u_{ij} \times O_j}{\sum_{j=1,2,\ldots,n \text{ and } O_j \le b_i} u_{ij}}, \tag{9}$$

$$c_i = \frac{\sum_{j=1,2,\ldots,n \text{ and } O_j \ge b_i} u_{ij} \times O_j}{\sum_{j=1,2,\ldots,n \text{ and } O_j \ge b_i} u_{ij}}, \tag{10}$$

where O_j is the desired output of X_j and $1 \le i \le c$. Based on Eqs. (6)-(11), we can obtain the triangular fuzzy sets of the fuzzy rules *Rule* 1, *Rule* 2, … and *Rule c*, shown as follows:

Rule 1: If x_1 is A_{1i} and x_2 is A_{2i} and … and x_h is A_{hi} Then y is B_i,
Rule 2: If x_1 is A_{1i} and x_2 is A_{2i} and … and x_h is A_{hi} Then y is B_i,
⋮
Rule c: If x_1 is A_{1i} and x_2 is A_{2i} and … and x_h is A_{hi} Then y is B_i.

Step 3: If the inputs $I_j^{(1)}, I_j^{(2)}$, … and $I_j^{(h)}$ of the jth sample X_j activate some fuzzy rules, where $\min_{1\le k\le h} \mu_{A_{k,i}}(I_j^{(k)}) > 0$, $\mu_{A_{k,i}}(I_j^{(k)})$ is the membership value of the input $I_j^{(k)}$ belonging to triangular fuzzy set $A_{k,i}$, $1 \le i \le p$, and p denotes the number of activated fuzzy rules, then calculate the inferred output O_j^* as follows:

$$O_j^* = \frac{\sum_{i=1}^{p} \min_{1\le k\le h} \mu_{A_{k,i}}(I_j^{(k)}) \times \text{DEF}(B_i)}{\sum_{i=1}^{p} \min_{1\le k\le h} \mu_{A_{k,i}}(I_j^{(k)})}, \tag{11}$$

where DEF(B_i) denotes the defuzzified value of the consequence fuzzy set B_i of the activated fuzzy rule *Rule i*, and $1 \le i \le p$. Otherwise, go to **Step 4**.

Step 4: Based on the fuzzy rules *Rule* 1, *Rule* 2, … and *Rule c* obtained in **Step 2**, we have the following multiple fuzzy rules interpolation scheme [6]:

Rule 1: If $x_1 = A_{11}$ and $x_2 = A_{21}$ and …and $x_h = A_{h1}$ Then $y = B_1$
Rule 2: If $x_1 = A_{12}$ and $x_2 = A_{22}$ and …and $x_h = A_{h2}$ Then $y = B_2$
⋮
Rule c: If $x_1 = A_{1c}$ and $x_2 = A_{2c}$ and …and $x_h = A_{hc}$ Then $y = B_c$
Observations: $x_1 = I_j^{(1)}$ and $x_2 = I_j^{(2)}$ and … and $x_h = I_j^{(h)}$

Conclusion: $y = O_j^*$

where x_k denotes the kth antecedent variable (k = 1, 2, …, h), y denotes the consequence variable, $A_{k,i}$ denotes the kth antecedent fuzzy set of *Rule i* (i = 1, 2, …, c), B_i denotes the consequence fuzzy set of *Rule i*, $I_j^{(k)}$ denotes the kth input (k = 1, 2, …, h) of the jth sample X_j, and O_j^* denotes the inferred output with respect to the inputs $I_j^{(1)}$, $I_j^{(2)}$, …, and $I_j^{(h)}$ of X_j. Calculate the weight W_i of *Rule i* (i = 1, 2, …, c) with respect to the input observations $x_1 = I_j^{(1)}$ and $x_2 = I_j^{(2)}$ and … and $x_h = I_j^{(h)}$, where

$$W_i = \left(\sum_{d=1}^{c} \left(\frac{\left\| r^* - r_i \right\|}{\left\| r^* - r_d \right\|} \right)^2 \right)^{-1}, \tag{12}$$

r^* denotes the vector of the inputs ($I_j^{(1)}, I_j^{(2)}$, …, $I_j^{(h)}$), r_i denotes the vector of the defuzzified values of antecedent fuzzy sets of *Rule i* (DEF($A_{1,i}$), DEF($A_{2,i}$), ..., DEF($A_{h,i}$)) based on Eq. (4), $1 \le k \le h$, $0 \le W_i \le 1$ and $\sum_{i=1}^{c} W_i = 1$.

Step 5: Calculate the inferred output O_j^*, where

$$O_j^* = \sum_{i=1}^{c} W_i \times \text{DEF}(B_i), \tag{13}$$

where DEF(B_i) is the defuzzified value of consequence fuzzy set B_i based on Eq. (4), $0 \le W_i \le 1$, and $\sum_{i=1}^{c} W_i = 1$.

IV. Experimental Results

In this section, we apply the proposed method for temperature prediction based on the data set of the daily average temperature and the data set of the daily average cloud density from June 1996 to September 1996 in Taipei [3], as shown in Table I and Table II. If we want to forecast the daily average temperature of day i, then we use the proposed method to get the forecasted variation (i.e., the inferred output) of day i, and the forecasted daily average temperature of day i is equal to the daily average temperature of day i – 1 plus the forecasted variation of day i. We partition each data set into four groups, i.e., June 1996, July 1996, August 1996 and September 1996, and apply the proposed method to each group by using the variations of the daily average temperature and the daily average cloud density between any two adjacent days. Table III shows the variations of the daily average temperature and the daily average cloud density in June 1996, respectively. Let us consider a widow basis w using the historical data of the past w days to predict the forecasted data of the day being considered, where w is a positive integer. That is, the historical variations

T_{i-w}, …, T_{i-2} and T_{i-1} of the daily average temperature and the historical variations D_{i-w}, …, D_{i-2}, D_{i-1} of the daily average cloud density are used to predict the variation T_i of the daily average temperature of day i, where T_{i-w}, …, T_{i-2}, T_{i-1}, D_{i-w}, …, D_{i-2}, D_{i-1} and T_i form a training sample (T_{i-w}, D_{i-w}, …, T_{i-2}, D_{i-2}, T_{i-1}, D_{i-1}, T_i). Table IV shows the 27 training samples of June 1996 based on the window basis w = 2. For example, the historical variations T_2 = 1.5 and T_3 = 1.4 of the daily average temperature and the historical variations D_2 = -13 and D_3 = 0 of the daily average cloud density are used to predict the variation T_4 = 1.5 of the daily average temperature of June 4, 1996, where T_2, T_3, D_2, D_3 and T_4 form the training sample X_1 = (1.5, -13, 1.4, 0, 1.5). In the following, we apply the proposed method to forecast the daily average temperature of June 1996, where the number of clusters is set to five.

TABLE I. THE HISTORICAL DATA OF THE DAILY AVERAGE TEMPERATURE FROM JUNE 1, 1996 TO SEPTEMBER 30, 1996 IN TAIPEI, TAIWAN (UNIT: °C) [3]

Day \ Month	June	July	August	September
1	26.1	29.9	27.1	27.5
2	27.6	28.4	28.9	26.8
3	29.0	29.2	28.9	26.4
4	30.5	29.4	29.3	27.5
5	30.0	29.9	28.8	26.6
6	29.5	29.6	28.7	28.2
7	29.7	30.1	29.0	29.2
8	29.4	29.3	28.2	29.0
9	28.8	28.1	27.0	30.3
10	29.4	28.9	28.3	29.9
11	29.3	28.4	28.9	29.9
12	28.5	29.6	28.1	30.5
13	28.7	27.8	29.9	30.2
14	27.5	29.1	27.6	30.3
15	29.5	27.7	26.8	29.5
16	28.8	28.1	27.6	28.3
17	29.0	28.7	27.9	28.6
18	30.3	29.9	29.0	28.1
19	30.2	30.8	29.2	28.4
20	30.9	31.6	29.8	28.3
21	30.8	31.4	29.6	26.4
22	28.7	31.3	29.3	25.7
23	27.8	31.3	28.0	25.0
24	27.4	31.3	28.3	27.0
25	27.7	28.9	28.6	25.8
26	27.1	28.0	28.7	26.4
27	28.4	28.6	29.0	25.6
28	27.8	28.0	27.7	24.2
29	29.0	29.3	26.2	23.3
30	30.2	27.9	26.0	23.5
31		26.9	27.7	

TABLE II. THE HISTORICAL DATA OF THE DAILY AVERAGE CLOUD DENSITY FROM JUNE 1, 1996 TO SEPTEMBER 30, 1996 IN TAIPEI, TAIWAN (UNIT: %) [3]

Day \ Month	June	July	August	September
1	36	15	100	29
2	23	31	78	53
3	23	26	68	66
4	10	34	44	50
5	13	24	56	53
6	30	28	89	63
7	45	50	71	36
8	35	34	28	76
9	26	15	70	55
10	21	8	44	31
11	43	36	48	31
12	40	13	76	25
13	30	26	50	14
14	29	44	84	45
15	30	25	69	38
16	46	24	78	24
17	55	26	39	19
18	19	25	20	39
19	15	21	24	14
20	56	35	25	3
21	60	29	19	38
22	96	48	46	70
23	63	53	41	71
24	28	44	34	70
25	14	100	29	40
26	25	100	31	30
27	29	91	41	34
28	55	84	14	59
29	29	38	28	83
30	19	46	33	38
31		95	26	

Given the initial randomly generated membership grades shown in Table V, the proposed method to forecast the daily average temperature for June, 1996 is presented as follows:

[Step 1] After applying the FCM clustering algorithm [1], we can get the five clusters, as shown in Table VI, where the membership grades of the training samples belonging to each cluster are shown in Table VII.

[Step 2] Based on Eqs. (5)-(10), we can get the following five fuzzy rules *Rule* 1, *Rule* 2, *Rule* 3, *Rule* 4 and *Rule* 5 from the training samples, where

Rule 1: If x_1 is A_{11} and x_2 is A_{21} and x_3 is A_{31} and x_4 is A_{41} Then y is B_1,
Rule 2: If x_1 is A_{12} and x_2 is A_{22} and x_3 is A_{32} and x_4 is A_{42} Then y is B_2,
Rule 3: If x_1 is A_{13} and x_2 is A_{23} and x_3 is A_{33} and x_4 is A_{43} Then y is B_3,
Rule 4: If x_1 is A_{14} and x_2 is A_{24} and x_3 is A_{34} and x_4 is A_{44} Then y is B_4,
Rule 5: If x_1 is A_{15} and x_2 is A_{25} and x_3 is A_{35} and x_4 is A_{45} Then y is B_5,

the antecedent fuzzy sets A_{11} = (-0.257, -0.1, 0.39), A_{21} = (-3.298, 4, 8.245), A_{31} = (-2.1, -2.1, -0.429) and A_{41} = (17.573, 36, 38.353), A_{12} = (-0.629, -0.3, 0.753), A_{22} = (-18.589, -10, -3.577), A_{32} = (-0.842, -0.6, 0.315), A_{42} = (-15.537, -9, -0.671), A_{13} = (-1.225, -0.6, -0.011), A_{23} = (9.253, 26, 31.664), A_{33} = (0.08, 1.2, 1.276), A_{43} = (-31.212, -26, -13.441), A_{14} = (-0.563, -0.1, 0.413), A_{24} = (8.094, 22, 28.972), A_{34} = (-0.841, -0.8, 0.222), A_{44} = (-11.465, -3, 5.737), A_{15} = (0.37, 2, 2), A_{25} = (-8.781, 1, 7.796), A_{35} = (-0.79, -0.7, -0.05) and A_{45} = (2.769, 16, 19.623) are calculated by Eqs. (5)-(7), respectively, and the consequence fuzzy sets B_1 = (-1.032, -0.9, -0.058), B_2 = (-0.297, 0.6, 1.237), B_3 = (0.03, 1.2, 1.22), B_4 = (-0.553, 0.2, 0.737) and B_5 = (-0.315, 0.2, 0.611) are calculated by Eqs. (8)-(10), respectively.

TABLE III. THE VARIATION OF THE DAILY AVERAGE TEMPERATUE AND THE DAILY CLOUDY DENSITY IN JUNE 1996

Day	Daily Average Temperature	Variation of the Daily Average Temperature	Daily Average Cloud Density	Variation of the Daily Average Cloud Density
1	26.1		36	
2	27.6	1.5	23	-13
3	29.0	1.4	23	0
4	30.5	1.5	10	-13
5	30.0	-0.5	13	3
6	29.5	-0.5	30	17
7	29.7	0.2	45	15
8	29.4	-0.3	35	-10
9	28.8	-0.6	26	-9
10	29.4	0.6	21	-5
11	29.3	-0.1	43	22
12	28.5	-0.8	40	-3
13	28.7	0.2	30	-10
14	27.5	-1.2	29	-1
15	29.5	2	30	1
16	28.8	-0.7	46	16
17	29.0	0.2	55	9
18	30.3	1.3	19	-36
19	30.2	-0.1	15	-4
20	30.9	0.7	56	41
21	30.8	-0.1	60	4
22	28.7	-2.1	96	36
23	27.8	-0.9	63	-33
24	27.4	-0.4	28	-35
25	27.7	0.3	14	-14
26	27.1	-0.6	25	11
27	28.4	1.3	29	4
28	27.8	-0.6	55	26
29	29.0	1.2	29	-26
30	30.2	1.2	19	-10

TABLE IV. THE TRAINING SAMPLES FOR JUNE 1996 BASED ON THE WINDOW BASIS $w = 2$

Training Samples	$(I_j^{(1)}, I_j^{(1)}, I_j^{(1)}, I_j^{(1)}, O_j)$	Training Samples	$(I_j^{(1)}, I_j^{(1)}, I_j^{(1)}, I_j^{(1)}, O_j)$
X_1	(1.5, -13, 1.4, 0, 1.5)	X_{15}	(-0.7, 16, 0.2, 9, 1.3)
X_2	(1.4, 0, 1.5, -13, -0.5)	X_{16}	(0.2, 9, 1.3, -36, -0.1)
X_3	(1.5, -13, -0.5, 3, -0.5)	X_{17}	(1.3, -36, -0.1, -4, 0.7)
X_4	(-0.5, 3, -0.5, 17, 0.2)	X_{18}	(-0.1, -4, 0.7, 41, -0.1)
X_5	(-0.5, 17, 0.2, 15, -0.3)	X_{19}	(0.7, 41, -0.1, 4, -2.1)
X_6	(0.2, 15, -0.3, -10, -0.6)	X_{20}	(-0.1, 4, -2.1, 36, -0.9)
X_7	(-0.3, -10, -0.6, 9, -0.6)	X_{21}	(-2.1, 36, -0.9, -33, -0.4)
X_8	(-0.6, -9, 0.6, -5, -0.1)	X_{22}	(-0.9, -33, -0.4, -35, 0.3)
X_9	(0.6, -5, -0.1, 22, -0.8)	X_{23}	(-0.4, -35, 0.3, -14, -0.6)
X_{10}	(-0.1, 22, -0.8, -3, 0.2)	X_{24}	(0.3, -14, -0.6, 11, 1.3)
X_{11}	(-0.8, -3, 0.2, -10, -1.2)	X_{25}	(-0.6, 11, 1.3, 4, -0.6)
X_{12}	(0.2, -10, -1.2, -1, 2)	X_{26}	(1.3, 4, -0.6, 26, 1.2)
X_{13}	(-1.2, -1, 2, 1, -0.7)	X_{27}	(-0.6, 26, 1.2, -26, 1.2)
X_{14}	(2, 1, -0.7, 16, 0.2)		

[Step 3] Let us consider to infer the output (i.e., the forecasted variation) with respect to the inputs of the training sample X_1 shown in Table IV. That is, we want to forecast the daily average temperature of June 4, 1996. we can see that no fuzzy rules can be activated by the inputs $I_1^{(1)}$, $I_1^{(2)}$, $I_1^{(3)}$ and $I_1^{(4)}$ of X_1 shown in Table IV, where X_1 = ($I_1^{(1)}$, $I_1^{(2)}$, $I_1^{(3)}$, $I_1^{(4)}$, O_1) = (1.5, -13, 1.4, 0, 1.5). Because $\min_{1\le k\le 4}\mu_{A_{k1}}(I_1^{(k)})=0$ for *Rule* 1, $\min_{1\le k\le 4}\mu_{A_{k2}}(I_1^{(k)})=0$ for *Rule* 2, $\min_{1\le k\le 4}\mu_{A_{k3}}(I_1^{(k)})=0$ for *Rule* 3,

$\min_{1 \le k \le 4} \mu_{A_{k4}}(I_1^{(k)}) = 0$ for *Rule* 4 and $\min_{1 \le k \le 4} \mu_{A_{k5}}(I_1^{(k)}) = 0$ for *Rule* 5, we go to **Step 4**.

TABLE V. INITIAL RANDOMLY GENERATED MEMBERSHIP GRADES OF THE TRAINING SAMPLES BELONGING TO EACH CLUSTER

Training Samples \ Clusters	C_1	C_2	C_3	C_4	C_5
X_1	0.125	0.303	0.215	0.016	0.341
X_2	0.264	0.251	0.217	0.027	0.241
X_3	0.147	0.071	0.031	0.725	0.026
X_4	0.133	0.103	0.031	0.374	0.360
X_5	0.008	0.321	0.327	0.185	0.159
X_6	0.042	0.319	0.268	0.223	0.148
X_7	0.236	0.418	0.015	0.259	0.073
X_8	0.097	0.185	0.251	0.216	0.250
X_9	0.300	0.114	0.073	0.223	0.290
X_{10}	0.040	0.309	0.131	0.459	0.061
X_{11}	0.167	0.267	0.223	0.274	0.068
X_{12}	0.325	0.307	0.143	0.125	0.100
X_{13}	0.031	0.516	0.035	0.023	0.396
X_{14}	0.384	0.269	0.063	0.126	0.158
X_{15}	0.260	0.168	0.190	0.107	0.275
X_{16}	0.100	0.230	0.299	0.230	0.141
X_{17}	0.335	0.012	0.232	0.223	0.198
X_{18}	0.204	0.243	0.048	0.193	0.313
X_{19}	0.146	0.246	0.174	0.242	0.193
X_{20}	0.476	0.138	0.056	0.227	0.102
X_{21}	0.175	0.383	0.119	0.247	0.077
X_{22}	0.061	0.676	0.001	0.193	0.069
X_{23}	0.316	0.394	0.187	0.046	0.057
X_{24}	0.276	0.204	0.383	0.026	0.111
X_{25}	0.220	0.119	0.027	0.264	0.369
X_{26}	0.394	0.025	0.136	0.413	0.032
X_{27}	0.271	0.111	0.191	0.402	0.025

TABLE VI. CLUSTER CENTERS OF THE FIVE CLUSTERS

Clusters	Cluster Centers
C_1	(0.178, 0.789, -0.749, 33.746, -0.227)
C_2	(0.217, -13.063, 0.096, -5.622, 0.418)
C_3	(-0.890, 24.997, 0.470, -29.010, 0.331)
C_4	(-0.237, 16.446, 0.136, 0.841, -0.122)
C_5	(0.615, -0.728, -0.326, 14.242, 0.094)

TABLE VII. MEMBERSHIP GRADES OF THE TRAINING SAMPLES BELONGING TO EACH CLUSTER

Training Samples \ Clusters	C_1	C_2	C_3	C_4	C_5
X_1	0.023	0.844	0.013	0.035	0.085
X_2	0.048	0.463	0.119	0.227	0.142
X_3	0.046	0.683	0.021	0.060	0.189
X_4	0.068	0.025	0.007	0.044	0.855
X_5	0.148	0.068	0.045	0.452	0.287
X_6	0.035	0.093	0.161	0.622	0.089
X_7	0.010	0.921	0.012	0.025	0.032
X_8	0.010	0.918	0.010	0.024	0.038
X_9	0.276	0.057	0.014	0.052	0.601
X_{10}	0.022	0.032	0.057	0.841	0.048
X_{11}	0.039	0.614	0.066	0.153	0.128
X_{12}	0.022	0.833	0.014	0.041	0.090
X_{13}	0.060	0.333	0.041	0.211	0.354
X_{14}	0.024	0.012	0.003	0.016	0.945
X_{15}	0.058	0.046	0.032	0.707	0.157
X_{16}	0.039	0.135	0.620	0.134	0.073
X_{17}	0.104	0.544	0.066	0.104	0.183
X_{18}	0.838	0.029	0.011	0.032	0.089
X_{19}	0.110	0.091	0.204	0.446	0.149
X_{20}	0.943	0.008	0.004	0.012	0.033
X_{21}	0.021	0.037	0.833	0.077	0.033
X_{22}	0.094	0.437	0.162	0.148	0.159
X_{23}	0.088	0.564	0.082	0.109	0.158
X_{24}	0.117	0.311	0.028	0.084	0.460
X_{25}	0.032	0.047	0.025	0.768	0.129
X_{26}	0.616	0.035	0.013	0.057	0.279
X_{27}	0.003	0.006	0.973	0.014	0.005

[Step 4] Based on the five generated fuzzy rules obtained in **Step 2** and the inputs of the training sample X_1, we have the following multiple fuzzy rules interpolation scheme:

Rule 1: **If** x_1 is A_{11} and x_2 is A_{21} and x_3 is A_{31} and x_4 is A_{41} **Then** y is B_1,
Rule 2: **If** x_1 is A_{12} and x_2 is A_{22} and x_3 is A_{32} and x_4 is A_{42} **Then** y is B_2,
Rule 3: **If** x_1 is A_{13} and x_2 is A_{23} and x_3 is A_{33} and x_4 is A_{43} **Then** y is B_3,
Rule 4: **If** x_1 is A_{14} and x_2 is A_{24} and x_3 is A_{34} and x_4 is A_{44} **Then** y is B_4,
Rule 5: **If** x_1 is A_{15} and x_2 is A_{25} and x_3 is A_{35} and x_4 is A_{45} **Then** y is B_5,
Observations: $x_1 = I_1^{(1)}$ and $x_2 = I_1^{(2)}$ and $x_3 = I_1^{(3)}$ and $x_4 = I_1^{(4)}$

Conclusion: $y = O_1^*$

where A_{11} = (-0.257, -0.1, 0.39), A_{21} = (-3.298, 4, 8.245), A_{31} = (-2.1, -2.1, -0.429) and A_{41} = (17.573, 36, 38.353), B_1 = (-1.032, -0.9, -0.058), A_{12} = (-0.629, -0.3, 0.753), A_{22} = (-18.589, -10, -3.577), A_{32} = (-0.842, -0.6, 0.315), A_{42} = (-15.537, -9, -0.671), B_2 = (-0.297, 0.6, 1.237), A_{13} = (-1.225, -0.6, -0.011), A_{23} = (9.253, 26, 31.664), A_{33} = (0.08, 1.2, 1.276), A_{43} = (-31.212, -26, -13.441), B_3 = (0.03, 1.2, 1.22), A_{14} = (-0.563, -0.1, 0.413), A_{24} = (8.094, 22, 28.972), A_{34} = (-0.841, -0.8, 0.222), A_{44} = (-11.465, -3, 5.737), B_4 = (-0.553, 0.2, 0.737), A_{15} = (0.37, 2, 2), A_{25} = (-8.781, 1, 7.796), A_{35} = (-0.79, -0.7, -0.05), A_{45} = (2.769, 16, 19.623) and B_5 = (-0.315, 0.2, 0.611), the inputs $I_1^{(1)}$, $I_1^{(2)}$, $I_1^{(3)}$ and $I_1^{(4)}$ of X_1 are input observations, X_1 = (1.5, -13, 1.4, 0, 1.5), $I_1^{(1)} = 1.5$, $I_1^{(2)} = -13$, $I_1^{(3)} = 1.4$, $I_1^{(4)} = 0$, and O_1^* denotes the inferred output. Based on Eq. (12), we can get the weights W_1, W_2, W_3, W_4 and W_5 of *Rule* 1, *Rule* 2, *Rule* 3, *Rule* 4 and *Rule* 5 with respect to the input observations $x_1 = 1.5$, $x_2 = -13$, $x_3 = 1.4$ and $x_4 = 0$, respectively, where $W_1 = 0.0462$, $W_2 = 0.7042$, $W_3 = 0.0315$, $W_4 = 0.0535$ and $W_5 = 0.1646$.

[Step 5] Based on Eq. (13), we can calculate the inferred output O_1^* with respect to the training sample X_1, where O_1^* = 0.409.

Therefore, the forecasted variation of June 4, 1996 is 0.409. Because the forecasted daily average temperature of June 4, 1996 is equal to the daily average temperature of June 3, 1996 (i.e., 29) plus the forecasted variation of June 4, 1996 (i.e., 0.409), the forecasted daily average temperature of June 4, 1996 is equal to 29 + 0.409 = 29.409. Table VIII shows the forecasted variations and the forecasted daily average temperature from June 4, 1996 to June 30, 1996.

In this paper, we use the average error rate (AER) to evaluate the forecasting result for temperature prediction, where

$$\text{AER} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{T_{\text{Forecasted}}(i) - T_{\text{Actual}}(i)}{T_{\text{Actual}}(i)}\right| \times 100\%, \tag{14}$$

$T_{\text{Forecasted}}(i)$ and $T_{\text{Actual}}(i)$ denotes the forecasted temperature and the actual temperature of day i, respectively. The average error rate of the forecasting result of the proposed method shown in Table VIII is 2.68%. Because the FCM clustering algorithm [1] might produce different fuzzy clustering results depending on the initial randomly generated membership grades of the training samples belonging to each cluster, in this paper, we execute the proposed method 30 times and take the average of the average error rates of the forecasting results as the average error rate.

Table IX shows the average error rates of the forecasting results from June 1996 to September 1996 for different window bases based on Chen and Hwang's method [7]. Table X, Table XI and Table XII show the average error rates of the forecasting results from June 1996 to September 1996 for different window basis based on the proposed method with five, ten and fifteen generated fuzzy rules, respectively. From the forecasting results shown in Table IX, Table X, Table XI and Table XII, respectively, we can see that the proposed method gets smaller average error rates than Chen and Hwang's method [7]. That is, the proposed method gets higher average forecasting accuracy rates than Chen and Hwang's method [7].

TABLE VIII. FORECASTED VARIATION AND FORECASTED DAILY AVERAGE TEMPERATURE IN JUNE 1996

Day	Actual Temperature	Forecasted Variation	Forecasted Daily Average Temperature
1	26.1		
2	27.6		
3	29.0		
4	30.5	0.409	29. 409
5	30.0	0.431	30.931
6	29.5	0.336	30.336
7	29.7	0.110	29.610
8	29.4	0.053	29.753
9	28.8	0.146	29.546
10	29.4	0.535	29.335
11	29.3	0.507	29.907
12	28.5	-0.083	29.217
13	28.7	0.146	28.646
14	27.5	0.464	29.164
15	29.5	0.434	27.934
16	28.8	0.278	29.778
17	29.0	0.174	28.974
18	30.3	0.139	29.139
19	30.2	0.606	30.906
20	30.9	0.307	30.507
21	30.8	-0.476	30.424
22	28.7	0.249	31.049
23	27.8	-0.671	28.029
24	27.4	0.683	28.483
25	27.7	0.370	27.770
26	27.1	0.351	28.051
27	28.4	0.170	27.270
28	27.8	0.181	28.581
29	29.0	-0.495	27.315
30	30.2	0.913	29.913

TABLE IX. AVERAGE ERROR RATES FROM JUNE 1996 TO SEPTEMBER 1996 IN TAIPEI FOR DIFFERENT WINDOW BASIS BASED ON [7]

Window Basis / Month / AER	$w = 2$	$w = 3$	$w = 4$	$w = 5$	$w = 6$	$w = 7$	$w = 8$
June	2.88%	3.16%	3.24%	3.33%	3.39%	3.53%	3.53%
July	3.04%	3.76%	4.08%	4.17%	4.35%	4.38%	4.38%
August	2.75%	2.77%	3.30%	3.40%	3.18%	3.15%	3.15%
September	3.29%	3.10%	3.19%	3.22%	3.39%	3.38%	3.38%

TABLE X. AVERAGE ERROR RATES FROM JUNE 1996 TO SEPTEMBER 1996 IN TAIPEI FOR DIFFERENT WINDOW BASIS BASED ON THE PROPOSED METHOD WITH FIVE GENERATED FUZZY RULES (AFTER EXECUTING THE PROPOSED METHOD 30 TIMES)

Window Basis / Month / AER	$w = 2$	$w = 3$	$w = 4$	$w = 5$	$w = 6$	$w = 7$	$w = 8$
June	3.25%	2.79%	2.43%	2.64%	2.57%	2.45%	2.45%
July	2.74%	2.40%	2.69%	2.74%	2.77%	2.88%	3.75%
August	2.64%	2.32%	2.21%	2.62%	2.62%	2.81%	3.18%
September	2.86%	2.68%	2.81%	2.38%	2.47%	2.37%	2.53%

TABLE XI. AVERAGE ERROR RATES FROM JUNE 1996 TO SEPTEMBER 1996 IN TAIPEI FOR DIFFERENT WINDOW BASIS BASED ON THE PROPOSED METHOD WITH TEN GENERATED FUZZY RULES (AFTER EXECUTING THE PROPOSED METHOD 30 TIMES)

Window Basis / Month / AER	$w = 2$	$w = 3$	$w = 4$	$w = 5$	$w = 6$	$w = 7$	$w = 8$
June	2.25%	2.03%	1.95%	1.89%	1.85%	1.83%	1.73%
July	1.97%	2.23%	2.22%	2.25%	2.28%	2.40%	3.23%
August	2.07%	1.89%	1.88%	2.02%	2.12%	2.69%	3.05%
September	2.21%	2.22%	2.08%	1.85%	1.97%	2.48%	2.42%

TABLE XII. AVERAGE ERROR RATES FROM JUNE 1996 TO SEPTEMBER 1996 IN TAIPEI FOR DIFFERENT WINDOW BASIS BASED ON THE PROPOSED METHOD WITH FIFTEEN GENERATED FUZZY RULES (AFTER EXECUTING THE PROPOSED METHOD 30 TIMES)

Window Basis / Month / AER	$w = 2$	$w = 3$	$w = 4$	$w = 5$	$w = 6$	$w = 7$	$w = 8$
June	1.70%	1.50%	1.38%	1.37%	1.28%	1.13%	0.97%
July	1.62%	1.77%	1.74%	1.68%	1.77%	1.72%	3.04%
August	1.60%	1.48%	1.24%	1.30%	1.28%	2.41%	2.97%
September	1.44%	1.51%	1.35%	1.20%	2.02%	2.49%	2.42%

V. CONCLUSIONS

In this paper, we have presented a method to deal with temperature forecasting based on fuzzy clustering and fuzzy rules interpolation techniques. First, the proposed method constructs fuzzy rules from training samples based on the fuzzy C-Means clustering algorithm [1], where each fuzzy rule corresponds to a fuzzy cluster and the linguistic terms appearing in the fuzzy rules are represented by triangular fuzzy sets. Then, it performs fuzzy inference based on the multiple fuzzy rules interpolation scheme. It calculates the weight of each fuzzy rule with respect to the input observation and uses the weight of each fuzzy rule to get the inferred output. We also have applied the proposed method to handle the temperature prediction problem. From the experimental result, we can see that the proposed method gets higher average forecasting accuracy rates than Chen and Hwang's method [7].

ACKNOWLEDGMENT

This work is supported in part by the National Science Council, Republic of China, under Grant NSC 97-2221-E-011-107-MY3.

REFERENCES

[1] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. NY: Plenum Press, 1981.

[2] G. Cardoso and F. Gomide, "Newspaper demand prediction and replacement model based on fuzzy clustering and fuzzy rules," *Inform. Sci*., vol. 177, no. 21, pp. 4799-4809, 2007.

[3] Central Weather Bureau, *The Historical Data of the Daily Average Temperature and Daily Cloud Density* (from January 1995 to September 1996), Taipei, Taiwan, R. O. C., 1996.

[4] A. Celikyilmaz and I. B. Turksen, "Enhanced fuzzy system models with improved fuzzy clustering algorithms," *IEEE Trans. Fuzzy Syst*., vol. 16, no. 3, pp. 779-794, 2008.

[5] P. C. Chang and C. H. Liu, "A TSK type fuzzy rule based systems for stock price prediction," *Expert Systems with Applications*, vol. 34, no. 1, pp. 135-144, 2008.

[6] Y. C. Chang, S. M. Chen, and C. J. Liau, "Fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on the areas of fuzzy sets," *IEEE Trans. Fuzzy Syst*., vol. 16, no. 5, pp. 1285-1301, 2008.

[7] S. M. Chen and J. R. Hwang, "Temperature prediction using fuzzy time series," *IEEE Transactions on Syst., Man, and Cybern.-Part B: Cybernetics*, vol. 30, no. 2, pp. 263-275, 2000.

[8] S. M. Chen and Y. K. Ko, "Fuzzy interpolative reasoning for sparse fuzzy-rule-based systems based on cutting and transformation techniques," *IEEE Trans. Fuzzy Syst*., vol. 16, no. 6, pp. 1626-1648, 2008.

[9] S. M. Chen, "Evaluating weapon systems using fuzzy arithmetic operations," *Fuzzy Sets and Systems*, vol. 77, no. 3, pp. 265-276, 1996.

[10] C. H. Cheng, G. W. Cheng, and J. W. Wang, "Multi-attribute fuzzy time series method based on fuzzy clustering," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1235-1242, 2008.

[11] Z. H. Huang and Q. Shen, "Fuzzy interpolation and extrapolation: A practical approach," *IEEE Trans. Fuzzy Syst*., vol. 16, no. 1, pp. 13-28, 2008.

[12] C. K. Kwong, K. Y. Chan, and H. Wong, "Takagi-Sugeno neural fuzzy modeling approach to fluid dispensing for electronic packaging," *Expert Systems with Applications*, vol. 34, no. 3, pp. 2111-2119, 2008.

[13] H. C. W. Lau, E. N. M. Cheng, C. K.M. Lee, and G. T. S. Ho, "A fuzzy logic approach to forecast enery consuption change in a manufacturing system," *Expert Systems with Applications*, vol. 34, no. 3, pp. 1813-1824, 2008.

[14] M. Mizumoto and H. J. Zimmermann, "Comparison of fuzzy reasoning methods," *Fuzzy Sets and Systems*, vol. 8, no. 3, pp. 253-283, 1982.

[15] M. Setnes, "Supervised fuzzy clustering for rule extraction," *IEEE Tran. Fuzzy Syst*., vol. 8, no. 4, pp. 416-424, 2000.

[16] K. W. Wong, D. Tikk, T. D. Gedeon, and L. T. Koczy, "Fuzzy rule interpolation for multidimensional input spaces with applications: A case study," *IEEE Trans. Fuzzy Syst*., vol. 13, no. 6, pp. 809-819, 2005.

[17] Y. Yam, M. L. Wong, and P. Baranyi, "Interpolation with function space representation of membership functions," *IEEE Trans. Fuzzy Syst*., vol. 14, no. 3, pp. 398-411, 2006.

[18] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning - I," *Inform. Sci*., vol. 8, pp. 199-249, 1975.