A Weighted Subsethood Mamdani Fuzzy Rules Based System Rule Extraction (MFRBS-WSBA) for Forecasting Electricity Load Demand – A Framework

Rosnalini Mansor[1], Maznah Mat Kasim[1], Mahmod Othman[2]
[1]School of Quantitative Sciences, UUM College of Arts and Sciences,
Universiti Utara Malaysia, 06010, Sintok, Kedah, Malaysia.
[2]Department of Fundamental and Applied Sciences, Faculty of Science & Information Technology,
Universiti Teknologi Petronas, 32610, Seri Iskandar, Perak, Malaysia.
rosnalini@uum.edu.my

***Abstract*—Fuzzy rules are very important elements that should be taken consideration seriously when applying any fuzzy system. This paper proposes the framework of Mamdani Fuzzy Rule-based System with Weighted Subsethood-based Algorithm (MFRBS-WSBA) for forecasting electricity load demand. Specifically, this paper proposed two frameworks: MFRBS-WSBA and WSBA framework where the WSBA is embedded in MFRBS-WSBA (fourth step in MFRBS-WSBA). The objective of this paper is to show the fourth step in the MFRBS-WSBA framework which applied the new electricity load forecasting rule extraction by WSBA method. We apply the proposed WSBA framework in Malaysia electricity load demand data as a numerical example in this paper. These preliminary results show that the WSBA framework can be one of alternative methods to extract fuzzy rules for forecast electricity load demand where the proposed method provide a simple to interpret the fuzzy rules and also offer a new direction to interpret the fuzzy rules compared to classical fuzzy rules.**

***Index Terms*—Fuzzy Rules; Forecasting; Electricity Load Demand.**

I. Introduction

A great number of literatures have studied the application of Fuzzy Rule-Based System (FRBS) in numerous real world problems, where the recent research is largely on the development of the specific techniques for data-driven machine learning in the FRBS. The extensive researches on such techniques have been motivated by the fact that rules generated by the techniques are able to provide the decision maker with the knowledge that embedded in the training samples. One should know that human expertise is important in the development of conventional fuzzy systems. However, extracting fuzzy rules automatically from data can be very meaningful to accommodate the decision making process when human experts are not available and may even provide new information which not known by the experts [1].

A number of approaches have been used to develop data-driven learning for FRBS in forecasting task. One of them is the use of a method that can automatically generate either membership functions or fuzzy rule structures or both from training data. The most popular choices for learning and reasoning from data among them is neural network [2], heuristic technique [3], Wang and Mandel method [4], qualitative description [5] and Song and Chissom method [6]. The proposed aforementioned methods are based on the Mamdani Fuzzy Rule Based System (MFRBS). In this research, the interest on the method of extracting fuzzy rules for forecasting problems is from Othman et al. (2013) This method uses fuzzy subsethood values in generating fuzzy rules.

The aim of this paper is to propose a framework model to forecast electricity load demand and to show the numerical example of fuzzy rules extraction using WSBA method. Furthermore, the numeral example is real data from Tenaga Nasional Berhad (TNB) even that is old data but this framework can implement to current data once the data is available for this study. This paper consists of four sections. The following sections discuss the propose framework. The last two sections are related to the explanation of numerical example results with a conclusion.

II. The Framework

Figure 1 shows the framework of Mamdani Fuzzy Rule-based System with Weighted Subsethood-based Algorithm (MFRBS-WSBA) for electricity load demand forecasting. The framework consists of six steps. Starting with step (1) Data Collection and Selection. The process includes identifying variables and identify time frame and time horizon based on the forecasting problem. Then, in step (2) Preprocessing Data, which are consist of data partition, handling any missing value and transformation if necessary. Step (3) Variables Selection; once the significant variables identified by using Rough set method, the framework's flow continue to step (4) Constructing Fuzzy Model and use the same variables to apply in other FRBS with difference fuzzy inference system (FIS) in step (5). Lastly, step (6) is Performance Evaluation.

The propose framework of Weighted Subsethood-based Algorithm (WSBA) fuzzy rule extraction method is embedded in step (4) MFRBS-WSBA framework. By referring step (4) in Figure 1, a fuzzy model has four components: fuzzification, fuzzy inference system (FIS), defuzzification and model output. Mamdani FIS in FRBS is the most widely applied because it can model the dynamic structure of systems where the robustness of this method has high power of approximation of nonlinear function [7].

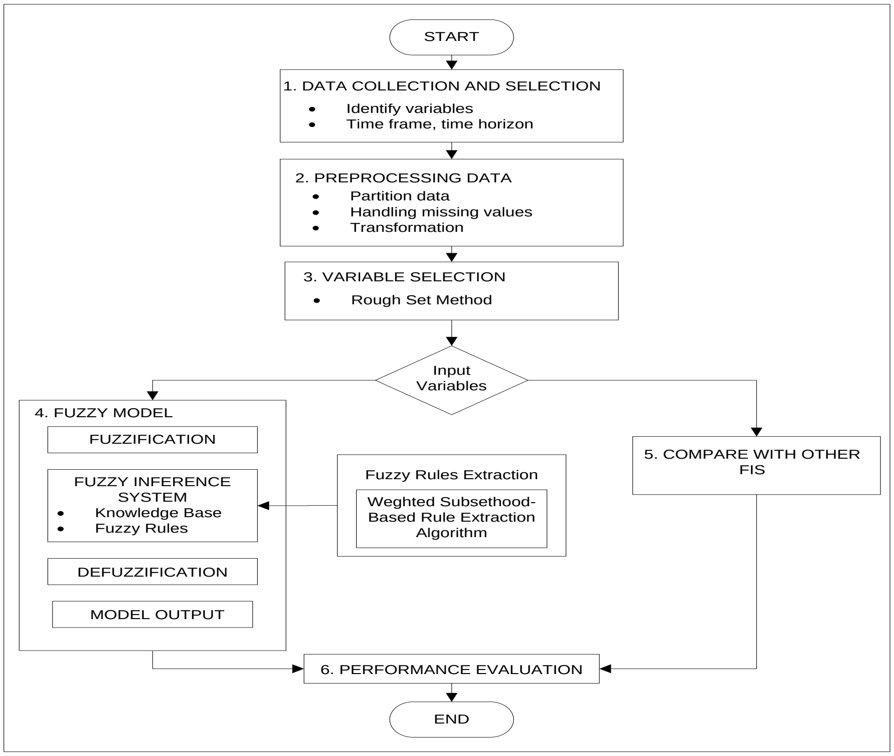

Figure 1: Proposed MFRBS-WSBA for electricity load demand forecasting

The WSBA fuzzy rules extraction is used as a fuzzy rule set extraction method as shown in Figure 1.

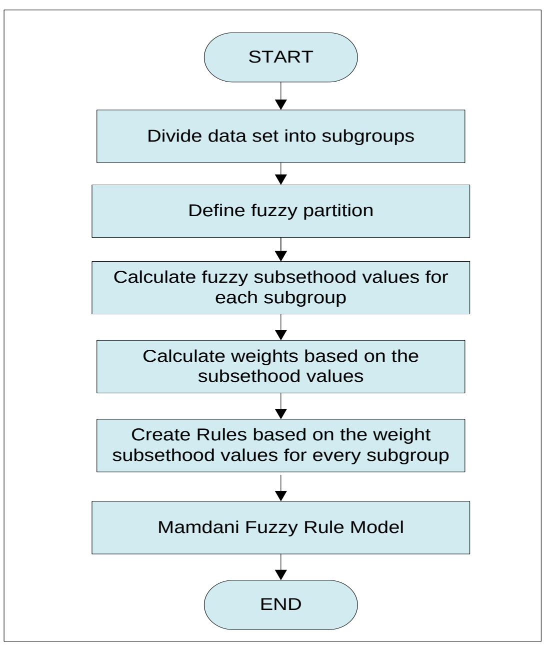

Figure 2: Proposed WSBS for fuzzy rules extraction

The WSBA framework in this study is shown in Figure 2 and the detail of six steps are as follow:

Step (1): Divide data set into subgroup
Step (2): Define fuzzy partition
Step (3): Calculate fuzzy subsethood values for each subgroup using the following definition [8].

Let *A* and *B* be two fuzzy set defined on the universe *U*. The fuzzy subsethood value of *A* with regard to *B*, *S(B,A)* represents the degree to which *A* is subsethood of *B*.

$$S(B,A) = \frac{M(B \cap A)}{M(B)} = \frac{\sum_{x \in U} \nabla(\mu_B(x),\ \mu_A(x))}{\sum_{x \in U} \mu_B(x)} \quad (1)$$

where $S(B,A) \in [0, 1]$ and ∇ is the *t*-norm operator.

Step (4): Calculate weights based on the subsethood values using the following definition [8].

The related weight for the linguistic terms Ai with regard to electricity load forecasting output *X* is calculated using Equation (2).

$$w(X, A_i) = \left(\frac{S(X, A_i)}{\max_{j=1,2,\ldots l} S(X, A_j)} \right) \quad (2)$$

Let $w(X,A_i) \in [0,1]$, $i = 1, 2,..., l$ and $A_i \in \{A_1, A_2, \ldots A_l\}$. The linguistic terms are attached with the weight generated which associates with the attributes. The Equation (3) is used to calculated weight conjunction of linguistic terms.

$$T(A) = \left(\frac{w_i}{w}(A_1) \nabla \ldots \nabla \frac{w_m}{w}(A_m) \right) \quad (3)$$

where A is the conditional attribute $T(A)$ is the compound weight and ∇ is the t-norm, A_i , i= 1,2,…,m are the linguistic terms of variable A which are conjunctively combined and w is the largest amongst the m associated weights. For the compound weight $T(B)$ of the weighted disjunction of linguistic terms associated with variables B is calculated as in Equation (4).

$$T(B) = \left(\frac{w_i}{w}(B_1) \Delta \ldots \Delta \frac{w_m}{w}(B_m) \right) \quad (4)$$

where Δ is the t- conorm and B_i, i=1,2,…, m are the linguistic terms B, which are disjunctively combined.

Step (5): Create fuzzy rules based on the weight subsethood values for every subgroup [9].

In this step, the weight Subsethood-Based Algorithm then used the weighted conjuction *T(A)* and weighted disjunction *T(B)* to generate fuzzy IF-THEN rules.

Step (6): Then, in this step, we apply the fuzzy rule sets for forecasting that are generated by using WSBA to forecast electricity load demand.

III. Numerical Example Of Fuzzy Rules Extraction (WSBA)

For this purpose, the study considers two input variables which are the previous day electricity load (L_{d-1}) and the similar day of last week electricity load (L_{d-7}). The next day forecast value (L) is the output variable of this study. This numerical example has adopted the data from Tenaga Nasional Berhad (TNB) from 27 October 2005 to 16 November 2005 as shown in Table 1.

The next sections show the step by step numerical calculation by applying the WSBA framework from this study.

A. *WSBA Step 1: Divide Data Set into Subgroup*

Based on the information in Table1, we created Table 2 where the data by case into subgroup with respect to the output variable linguistic term.

Table 1
TNB Data Set

Cases	Input Variables		Output variable	Output linguistic term
	L_{d-1}	L_{d-7}	L	
1	11,631	11,325	11,486	High
2	11,486	10,599	11,276	High
3	11,276	10,280	10,270	Medium
4	10,270	9,880	9,510	Medium
5	9,510	11,630	10,005	Medium
6	10,005	11,778	8,673	Low
7	8,673	11,631	8,342	Low
8	8,342	11,486	7,270	Low
9	7,270	11,276	7,536	Low
10	7,536	10,270	8,104	Low
11	8,104	9,510	8,381	Low
12	8,381	10,005	10,984	High
13	10,984	8,673	11,331	High
14	11,331	8,342	11,929	High
15	11,929	7,270	11,732	High
16	11,732	7,536	11,412	High
17	11,412	8,104	10,613	Medium
18	10,613	8,381	9,639	Medium
19	9,639	10,984	12,056	High
20	12,056	11,331	12,031	High
21	12,031	11,929	11,880	High

Table 2
Subgroup of data set with respect to the output

Subgroup	Cases	Next day forecast (Output)
1	6, 7, 8, 9, 10, 11	Low
2	3, 4, 5, 17, 18	Medium
3	1, 2, 12, 13, 14, 15, 16, 19, 20, 21	High

B. *WSBA Step 2: Define Fuzzy Partition*

Then, we define the fuzzy partition all linguistic term by transforming the data into membership set score by using the membership function definition μ, in Equation (5), (6) and (7) where a = 8051, b = 10099.5 and c = 11605.79. In order to construct load membership function, we apply k-means clustering to identify a, b and c value. The graphical of load membership functions as shown in Figure 3.

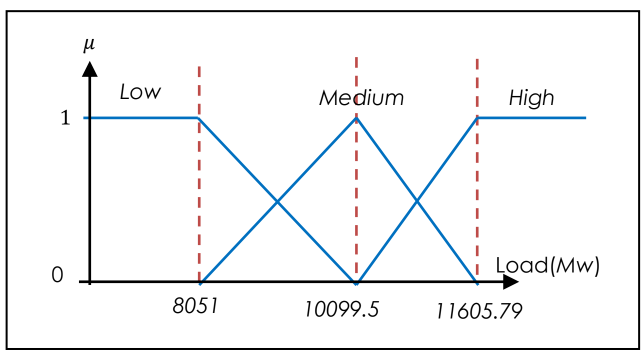

Figure 3: Graphical presentation of load membership function

$$\mu_{Low}(x) = \max\left\{0, \min\left[\frac{(b-x)}{(b-a)}, 1\right]\right\} \quad (5)$$

$$\mu_{Medium}(x) = \max\left\{0, \min\left[\frac{(x-a)}{(b-a)}, \frac{c-x}{c-b}\right]\right\} \quad (6)$$

$$\mu_{High}(x) = \max\left\{0, \min\left[\frac{(x-b)}{(c-b)}, 1\right]\right\} \tag{7}$$

We applied *k*-means clustering to assist us to construct the membership function for three linguistic term 'Low', 'Median' and 'High' [10]. Result from *k*-means clustering analysis showed that 8051, 10099.5 and 11605.79 is the final cluster canters for load data, then the graphical membership function as seen in Figure 3 constructed. Table 3 shows the membership function set score regarding to fuzzy partition in Figure 3.

Based on Subgroup of a data set with respect to the output in Table 2, we divide the membership function set score into three subgroups as presented in Table 4.

Table 3
Membership function sets score

	L_{d-1}			L_{d-7}			L		
Case	Low	Medium	High	Low	Medium	High	Low	Medium	High
1	0	0	1	0	0.1864	0.8136	0	0.0795	0.9205
2	0	0.0795	0.9205	0	0.6684	0.3316	0	0.2189	0.7811
3	0	0.2189	0.7811	0	0.8802	0.1198	0	0.8868	0.1132
4	0	0.9286	0.0714	0.1072	0.8928	0	0.2878	0.7122	0
5	0.2878	0.7122	0	0	0	1	0.0461	0.9539	0
6	0.0461	0.9539	0	0	0	1	0.6964	0.3036	0
7	0.6964	0.3036	0	0	0	1	0.8579	0.1421	0
8	0.8579	0.1421	0	0	0.4128	0.5872	1	0	0
9	1	0	0	0	0.2189	0.7811	1	0	0
10	1	0	0	0	0.8868	0.1132	0.9741	0.0259	0
11	0.9741	0.0259	0	0.2878	0.7122	0	0.8389	0.1611	0
12	0.8389	0.1611	0	0.0461	0.9539	0	0	0.4128	0.5872
13	0	0.4128	0.5872	0.6964	0.3036	0	0	0.1824	0.8176
14	0	0.1824	0.8176	0.8579	0.1421	0	0	0	1
15	0	0	1	1	0	0	0	0	1
16	0	0	1	1	0	0	0	0.1287	0.8713
17	0	0.1287	0.8713	0.9741	0.0259	0	0	0.6591	0.3409
18	0	0.6591	0.3409	0.8389	0.1611	0	0.2263	0.7737	0
19	0.2263	0.7737	0	0	0.4128	0.5872	0	0	1
20	0	0	1	0	0.1824	0.8176	0	0	1
21	0	0	1	0	0	1	0	0	1

Table 4
Three Subgroups According to the L (Nextday Forecast Output)

Linguistic variables		L_{d-1}			L_{d-7}			L		
Linguistic term Subgroup	Case	Low	Medium	High	Low	Medium	High	Low	Medium	High
Subgroup_1	6	0.0461	0.9539	0	0	0	1	0.6964	0.3036	0
	7	0.6964	0.3036	0	0	0	1	0.8579	0.1421	0
	8	0.8579	0.1421	0	0	0.4128	0.5872	1	0	0
	9	1	0	0	0	0.2189	0.7811	1	0	0
	10	1	0	0	0	0.8868	0.1132	0.9741	0.0259	0
	11	0.9741	0.0259	0	0.2878	0.7122	0	0.8389	0.1611	0
Subgroup_2	3	0	0.2189	0.7811	0	0.8802	0.1198	0	0.8868	0.1132
	4	0	0.9286	0.0714	0.1072	0.8928	0	0.2878	0.7122	0
	5	0.2878	0.7122	0	0	0	1	0.0461	0.9539	0
	17	0	0.1287	0.8713	0.9741	0.0259	0	0	0.6591	0.3409
	18	0	0.6591	0.3409	0.8389	0.1611	0	0.2263	0.7737	0
Subgroup_3	1	0	0	1	0	0.1864	0.8136	0	0.0795	0.9205
	2	0	0.0795	0.9205	0	0.6684	0.3316	0	0.2189	0.7811
	12	0.8389	0.1611	0	0.0461	0.9539	0	0	0.4128	0.5872
	13	0	0.4128	0.5872	0.6964	0.3036	0	0	0.1824	0.8176
	14	0	0.1824	0.8176	0.8579	0.1421	0	0	0	1
	15	0	0	1	1	0	0	0	0	1
	16	0	0	1	1	0	0	0	0.1287	0.8713
	19	0.2263	0.7737	0	0	0.4128	0.5872	0	0	1
	20	0	0	1	0	0.1824	0.8176	0	0	1
	21	0	0	1	0	0	1	0	0	1

C. WSBA Step 3: Calculate fuzzy subsethood value

In this step, in each subgroup, we calculated the fuzzy subsethood by using Equation (1). These values describe the relationship between L and every linguistic term. The fuzzy subsethood set value represents the larger the value, the closer the relationship between output linguistic term and the input linguistic term. Based on Table 2 and Table 4, we can see that there are three subgroups of cases. In each subgroup the next day forecast load term is fixed.

Take *Subgroup_1* as an example to show the fuzzy subsethood calculation. Based on Equation (1), "B" is 'Low' term in *L* of *Subgroup_1* and "A" is the 'Low' term in Ld-1. The meaning of fuzzy subsethood value is defined by using Equation (1), is measuring the degree in which, "B" is a subset of "A" defined by

$$S(B,A) = \frac{M(B \cap A)}{M(B)}$$

Therefore,

$$M(B) = \sum_{x \in U} \mu_B(x) = 0.6964 + 0.8579 + 1 + 1 + 0.9741 + 0.8389 = 5.3673\ 3$$

$$M(B \cap A) = \sum_{x \in U} \nabla\big(\mu_B(x), \mu_A(x)\big) = \min(0.6964,0.0461) + \min(0.8579,0.6964) + \min(1,0.8578) + \min(1,1) + \min(0.9741,1) + \min(0.8389,0.9741)$$

$$= 0.0461 + 0.6964 + 0.8579 + 1 + 0.9741 + 0.8389 = 4.4134$$

Hence, the value of *S(B,A)* is:

$$S(B,A) = \frac{M(B \cap A)}{M(B)} = \frac{4.4134}{5.3673} = 0.8223$$

Table 5 shows all fuzzy subsethood value from the similar calculation as presented in previous example regarding to the membership function linguistic term of L_{d-1}, L_{d-7} and L values in Table 4.

Table 5
Fuzzy subsethood value

Input linguistic variables	Linguistic Term	Output Linguistic variable, *L*: Low	Medium	High
	Low	0.8223	0.0823	0
L_{d-1}	Medium	0.7454	0.6100	0.4138
	High	0	1	0.7772
	Low	0.0536	0.5953	0
L_{d-7}	Medium	0.2956	0.4464	0.8741
	High	0	0.2493	0.3954

D. *WSBA Step 4: Calculate weight based on the fuzzy subsethood value*

By referring subsethood value in Table 5, we calculate the weighted subsethood value. Take L_{d-1} as an example to show the weighted fuzzy subsethood calculation. Based on Equation (2), in this example, "X" is 'Low' linguistic term of output variable *L* and "Ai" is 'Low' linguistic term of input variable L_{d-1}. The meaning of weighted fuzzy subsethood value is defined by using Equation (1), is measures the degree of 'importance' of linguistic term "Ai" to corresponding "X" defined by:

$$w(X, A_i) = \left(\frac{S(X, A_i)}{\max\limits_{j=1,2,3} S(X, A_j)} \right)$$

Therefore, $S(X,A_i) = 0.8223$

$$\max_{j=1,2,3} S(X, A_j) = max(0.8223\ , 0.7454\ , 0) = 0.8223$$

Hence,

$$w(X, A_i) = \left(\frac{S(X, A_i)}{\max\limits_{i=1,2,3} S(X, A_j)} \right) = \frac{0.8223}{0.8223} = 1$$

All the results from the similar calculation are shown in Table 6.

Table 6
Weighted subsethood value

Input linguistic variables	Linguistic Term	Output Linguistic variable, *L*: Low	Medium	High
	Low	1	0.0823	0
L_{d-1}	Medium	0.9065	0.6100	0.5324
	High	0	1	1
	Low	0.1814	1	0
L_{d-7}	Medium	1	0.7499	1
	High	0	0.4187	0.4524

E. *WSBA Step 5: Create fuzzy rules based on the weight fuzzy subsethood value for every subgroup*

We create the fuzzy rules based on information in Table 6. Figure 4 shows the three fuzzy rules obtain were consistent with results in Table 6. In the previous definition, 'OR' and 'AND' are fuzzy logical operators and are interpreted by *t*-conorm and t-norm respectively. Weights created from fuzzy subsethood values (WSBA value) work as multiplication factors for each linguistic term. During the extraction process, the ruleset is simplified as any linguistic terms that have a weight equal to 0 is automatically removed from the fuzzy rule.

Generally, *Rule 1* fuzzy rule in Figure 4 indicated the $L_{d-1=Low}$ and $L_{d-7=Medium}$ were most important linguistic term to L_{Low}. In other words, we can said that these two linguistic terms ($L_{d-1=Low}$ and $L_{d-7=Medium}$) were main factors contribution to L_{Low}. The similar interpretation from previous *Rule 1* for *Rule 2* and *Rule 3* based on the weighted values respectively. Then in the next step (step 6: Apply the fuzzy rule), the fuzzy rule sets can be used to apply for forecasting process. Once the fuzzy rules set is obtained, electricity load demand forecasting can be performed if data over the two variables are obtained.

Rule 1
IF L_{d-1} is (*Low* **OR** 0.9065*Medium*)
AND L_{d-7} is (0.1814 *Low* **OR** *Medium*)
THEN *L* is *Low*

Rule 2
IF L_{d-1} is (0.0823*Low* **OR** 0.*61Medium*)
AND L_{d-7} is (*Low* **OR** 0.7499*Medium* **OR** 0.4187*High*)
THEN *L* is *Medium*

Rule 3
IF L_{d-1} is (0.5324*Medium* **OR** *High*)
AND L_{d-7} is (*Medium* **OR** 0.4524*High*)
THEN *L* is *High*

Figure 4: Fuzzy rules based on WSBA for electricity load demand forecasting

IV. Discussion and Conclusion

This paper proposed two frameworks studies: MFRBS-WSBA framework and WSBA framework which the second framework (WSBA) is embedded in the MFRBS-WSBA. However, the aim of this paper is to show the fuzzy rules extraction process calculation in WSBA framework mainly. Step (6) in WSBA framework is cannot show in this paper due to the limited data set and this is the preliminary result only. This study will continue for further steps once the current data set available. Even though, the process of fuzzy rules extraction from this method can be one of alternative methods to extract fuzzy rules in other problems.

The problem to interpret the fuzzy rules among practitioner can be solving by SBA and WSBA methods. The interpretation between SBA and WSBA fuzzy rules is a totally different. SBA method provides the degree of input linguistic term is a subset of output linguistic term measurement. While WSBA provides the degree of 'importance' of the input linguistic term to corresponding output linguistic term. Moreover, WSBA value behaves as a multiplication factor for the input variables linguistic term. These fuzzy rules characteristic give a new dimension in fuzzy rules interpretation compared to classical fuzzy rules.

In conclusion, we have shown the example calculations for SBA and WSBA in the electricity load demand problem. The definition both of them also presented. The proposed fuzzy rules extraction method can give minimum number of rules regarding the number of output linguistics term. However, the fuzzy rules can added by adding the set of data to be many partitions of data or adding the linguistic term in the input and/or output variables. Furthermore, the interesting of WSBA fuzzy rules extraction method, it can express the degree of 'importance' input variables(s) linguistic term corresponding to output variable. In the meantime, this method able to show the hidden relationship between input variables and output variable. Moreover, this promising results provide a simple to interpret and able to improve the interpretability of derived fuzzy rules set compared to classical fuzzy rules. Future work will also focus on determining whether the number of fuzzy rules extraction from WSBA method influence the forecast error and FRBS accuracy.

References

[1] Kumar, S., Narula, P., and Ahmed, T. Knowledge Extraction from Numerical Data for the Mamdani Type Fuzzy Systems : A BBO Approach. *In Proceeding of the Innovative Practices in Management and Information Technology for Excellence*. India. 2010. 1–10.

[2] Abraham, A., and B. Nath., "A neuro-fuzzy approach for modelling electricity demand in Victoria," *Applied Soft Computing.* 1:127–138, 2001.

[3] Sachdeva, S., and Verma, C. M. Load Forecasting using Fuzzy Methods. *In Proceeding of the Power System Technology and IEEE Power India Conference*. 2008. 1–4.

[4] Cococcioni, M., Foschini, L., Lazzerini, B., and Marcelloni, F., "Complexity Reduction of Mamdani Fuzzy Systems through Multi-valued Logic Minimization," *Proceeding in IEEE International Conference on Systems, Man and Cybernetics*. 1782–1787, 2008.

[5] Rizwan, M., Kumar, D., and Kumar, R., "Fuzzy Logic Approach for Short Term Electrical Load Forecasting," *Electrical and Power Engineering Frontier*. 1(1): 8–12, 2012.

[6] Jarrett, J. E., and Plouffe, J. S., "The Fuzzy Logic Method for Simpler Forecasting," *International Journal of Engineering Business Management*. 3(3):25–52, 2011.

[7] Yazdania, A., Shariatib, S., and Yazdani-Chamzinic, A., "A risk assessment model based on fuzzy logic for electricity distribution system asset management," *Decision Science Letter*. 3: 343–352, 2014.

[8] Othman, M., Khalid, S., A., Abdullah, F., Hamzah, S. H. A., and Mahmud, K. R. K., "Hybrid Subjective Evaluation Method Using Weighted Subsethood-based (WSBA) Rule Generation Algorithm," *Journal of Arts, Science &Commerce.* IV (1(1)): 46–54, 2013.

[9] Othman, M., Saian, R., Nazri, M. N., Safiah, N., Kamal, A., and Liyana, N. Fuzzy Forecasting For Water Level of Flood Warning System in Perlis. *Proceeding in International Symposium on Mathematical Sciences and Computing Research*. 104–109, 2013.

[10] Tiwari, S., Singh, V., and Kumar, R., "An Intelligent System For Forecasting Electric Load And Price," *International Journal of Computer Sciences and Information Technologies.* 6(3): 2121–2126, 2015.