Proceedings of International Joint Conference on Neural Networks, Montreal, Canada, July 31 - August 4, 2005

# A Fuzzy Approach for Key Variables Identification of EMG E aluation System

Yanfeng Hou; Jacek M. Zurada
Department of Electrical and Computer Engineering
University of Louisville
Louisville, KY 40292
E-mail: y0hou002@louisville.edu

Waldemar Karwowski
Center for Industrial Ergonomics
Department of Industrial Engineering
University of Louisville
Louisville, KY 40292
E-mail: karwowski@louisville.edu

William S. Marras
Biodynamics Laboratory
Institute for Ergonomics
The Ohio State University
Columbus, OH 43210
E-mail: marras.1@osu.edu

***Abstract*— Identification of influence of input variables is very important for complex nonlinear systems with high dimensional input space. In this paper we propose a method using fuzzy average with fuzzy cluster distribution (FAFCD). To avoid the interference of different distributions of the sampling data, we deal with the distribution of fuzzy clusters in the sampling data, instead of the original data set. To discover the input-output relationship, we first use method of fuzzy rules and Fuzzy C-means to partition the original sampling data set into fuzzy clusters. We produce a new data set with the same distribution of the fuzzy clusters. Then the fuzzy average method is applied to the new data set. By doing this, the interference of distribution of the original sampling data is removed. This method is straight-forward and computationally easy. The performance is tested on both benchmark data and the electromyographic (EMG) signal Evaluation System.**

## I. Introduction

In modelling of complex nonlinear systems such as the EMG signal Evaluation System whose input-output relationship is not well understood, it is very helpful to find out the influence of each variable to the output of the system. If we can remove inputs that have little or no influence on the output and put emphasis on the important variables, we can build a more parsimonious and more effective model. Moreover, we will have a better understanding of those systems if we can discover how individual inputs affect the outputs.

The problem we are investigating can be stated as follows. For a nonlinear system with one output variable y and $n$ associated input variables $x_i$, we obtain $m$ sampling data points. From the sampling data of input-output pairs, we want to find out the relationship between each input variable and the output variable (the $x_i - y$ relationship, $i = 1, 2, ...n$).

This problem has a broad applicability and some methods have been proposed. In [1] and [2], different methods are employed to determine the importance of inputs. These methods can not determine in what way each input variable affects the outputs. Also, they either need to develop different neural networks or employ evolutionary computation, making them time consuming.

In [3] and [4], Lin et al. proposed their "fuzzy curves" method. The principal idea is to split the multiple-input-single-output (MISO) system into several single-input-single-output (SISO) systems, if the inputs are independent. For each input variable $x_i$, the $m$ data points are plotted in the $x_i - y$ space. A fuzzy rule is defined according to each sampling data point $(x_i^j, y^j)$ $(i = 1, 2, ..., n, j = 1, 2, ..., m)$ in the following form:

$R_i^j$ : IF $x_i$ is $\ _{ij}(x_i)$ THEN $y$ is $y^j$;

where $\ _{ij}(x_i)$ is a Gaussian membership function of $x_i^j$. A "fuzzy curve" can be produced using defuzzification method, which stands for the $x_i - y$ relationship. The importance of the input variables is ranked according to the ranges covered by the fuzzy curves.

This method is easy to understand and to calculate. The result obtained is straightforward. Lin et al. used this method in fuzzy-neural system modelling to determine model structure and set the initial weights in the model [3]. This method was also used in many other papers such as [4] - [12]. When trying to use this method in the evaluation model of electromyographic responses to manual lifting tasks, we found that it did not always work well. The distribution of the sampling data set will affect the result. Obviously this should not happen because for a system, the influence of each input variable is an inherent property of the system, regardless of the distribution of sampling data.

In section II, we will point out the limitation of the fuzzy curve method and improve it with a method we call FAFCD. In section III, the proposed method is tested on a benchmark data set and a real system.

## II. METHODS

As mentioned in the prior section, in the method of fuzzy curves, a fuzzy rule is defined according to each sampling data point $(x_i^j, y^j)$. From $m$ data points, $m$ fuzzy rules can be obtained. The fuzzy membership functions for input variable $x_i$ are Gaussian membership functions centered at $x_i^j$:

$$\ _{ij}(x_i) = \exp(-(\frac{x_i - \overline{x}_i^j}{\sigma})^2) \quad (1)$$

where $\overline{x}_i^j$ and $\sigma$ are the center and width of the membership function, respectively.

Then the fuzzy curve is produced from defuzzification:

$$C_i(x_i) = \frac{\sum_{j=1}^{m} y^j \mu_{ij}(x_i)}{\sum_{j=1}^{m} \mu_{ij}(x_i)} \quad (2)$$

The authors of [3] demonstrated and validated this method using a nonlinear system defined as:

$$y = (2 + x_1^{1.5} - 1.5\sin(3x_2))^2, 0 \le x_1, x_2 \le 3 \quad (3)$$

where $x_1$ and $x_2$ are two input variables. $y$ is the output variable.

Here we plot the defuzzified curves of $y$ in $x_1 - y$ and $x_2 - y$ space in Fig. 1 (A) and (B), respectively. The sample data were generated using formula (3) with $x_1$ and $x_2$ uniformly distributed in the interval [0,3] (Fig. 1 (C)). $R$ in the figure is the ratio of the range of $y$ covered by the curve to the whole range of $y$. We use it to stand for the influence of the corresponding input variable, instead of using the ranges covered by the fuzzy curves as in [3].

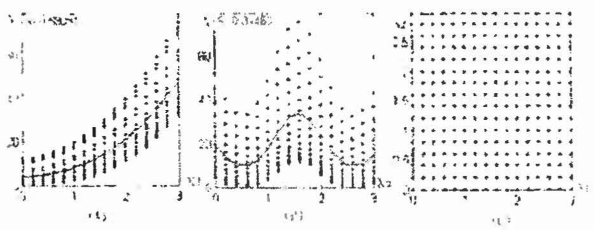

Fig. 1. The defuzzified curves of $y$ in $x_1 - y$ space (figure A) and in $x_2 - y$ space (figure B). The inputs $x_1$ and $x_2$ are uniformly distributed (figure C)

We can see that the curves can correctly reflect the $x_1 - y$ relationship and $x_2 - y$ relationship. However, can this still work if the distribution of the sample data changes? Fig. 2 (A) gives us a different result of $x_2 - y$ relationship when the data were generated with $x_1$ and $x_2$ shown in Fig. 2 (B). It seems that if the sampling data are not uniformly distributed, the curves will be distorted. This example shows that the fuzzy curve method puts restrictions on the distribution of sampling data. Below we will find out the reason and come up with a new approach.

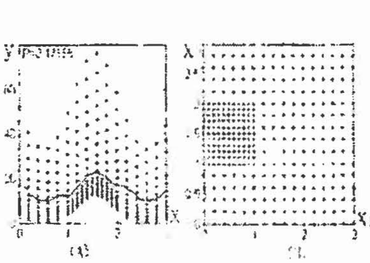

Fig. 2. The defuzzified curve of $y$ in $x_2 - y$ space (figure A) when the distribution of the input is uneven (figure B)

### 4 *The Method Depends on Distribution of Sampling Data*

As stated before, the importance of the input variables are ranked according to the ratio of the range of $y$ covered by the curve produced from defuzzification to the whole range of $y$. Let us define the ratio as Influence Rate $R$, then the range for variable $x_i$ can be written as

$$R_{x_i} = \frac{C_i(x_i^u) - C_i(x_i^l)}{a} \quad (4)$$

where $C_i(x_i^u)$ is the highest point on the curve and $C_i(x_i^l)$ is the lowest point on the curve. $a$ is the whole range of $y$.

$C_i(x_i^u)$ and $C_i(x_i^l)$ are calculated from (2). The membership of each value of $x_i$ to all the $m$ membership functions are calculated for (2). In (1), the width of the Gaussian membership function is often taken as about 20% of the length of the input interval of $x_i$. If the width $\sigma$ is very small, only those membership functions with a center (mean of the Gaussian function) close to current value of $x_i$ will have a big value, while the membership functions far from it will have a value close to zero. When $\sigma \to 0$, we obtain

$$R_{x_i}(\sigma \to 0) = \lim_{\sigma \to 0}\left(\frac{C_i(x_i^u) - C_i(x_i^l)}{a}\right) \quad (5)$$

Substituting (1) and (2) into (5),

$$R_{x_i}(\sigma \to 0) = \frac{1}{a} \times \lim_{\sigma \to 0}\left(\frac{\sum_{j=1}^{m} y^j \mu_{ij}(x_i^u)}{\sum_{j=1}^{m} \mu_{ij}(x_i^u)} - \frac{\sum_{j=1}^{m} y^j \mu_{ij}(x_i^l)}{\sum_{j=1}^{m} \mu_{ij}(x_i^l)}\right)$$

$$= \frac{1}{a} \times \lim_{\sigma \to 0}\left(\frac{\sum_{j=1}^{m} y^j \exp(-(\frac{x_i^u - \overline{x}_i^j}{\sigma})^2)}{\sum_{j=1}^{m} \exp(-(\frac{x_i^u - \overline{x}_i^j}{\sigma})^2)} - \frac{\sum_{j=1}^{m} y^j \exp(-(\frac{x_i^l - \overline{x}_i^j}{\sigma})^2)}{\sum_{j=1}^{m} \exp(-(\frac{x_i^l - \overline{x}_i^j}{\sigma})^2)}\right) \quad (6)$$

when $\sigma \to 0$, only those membership functions with their centers equal to $x_i^u$ and $x_i^l$ need to be taken into account. The memberships of $x_i^u$ and $x_i^l$ to the other membership functions are zero. Suppose there are $s$ membership functions with a center equal to $x_i^u$ and $h$ membership functions with a center equal to $x_i^l$, then (6) becomes:

$$R_{x_i}(\sigma \to 0) = \frac{1}{a} \times \left(\frac{\sum_{k=1}^{s} y_k^u}{s} - \frac{\sum_{k=1}^{h} y_k^l}{h}\right) \quad (7)$$

where $y_k^u(k = 1, 2, .., s)$ are the values of $y$ when $\mu_{ij}(x_i^u) = 1$; and $y_k^l(k = 1, 2, .., h)$ are the values of $y$ when $\mu_{ij}(x_i^l) = 1$.

Formula (7) indicates that the range of the curve (when $\sigma \to 0$) is the difference between the average value of $y$ at $x_i = x_i^u$ and the average value of $y$ at $x_i = x_i^l$.

If $\sigma$ is not approaching zero, the value of $C_{x_i}$ at $x_i = x_i^u$ takes those data points around $x_i = x_i^u$ into account. But it is still a weighted average. Since it has a meaning of average carried in a fuzzy sense, we call it Fuzzy Average.

Since the value of the fuzzy curve for $x_i$ at $x_i = x_i^u$ is a weighted average which takes the points around $x_i = x_i^u$ into account, each of the input variable should have the same distribution along the axis of $x_i$, respectively. Otherwise the curve can not reflect the $x_i - y$ relationship.

Fig. 1 is the 2-dimensional example with two input variables $x_1$ and $x_2$, in which $x_1$ has the same distribution along $x_2$ (Fig.

1 (C)). The fuzzy average of $y$ in the $x_2 - y$ space reflects the $x_2 - y$ relationship correctly. While in Fig. 2, $x_1$ does not have the same distribution along $x_2$ axis, so the fuzzy average of $y$ can not correctly reflects the $x_2 - y$ relationship.

For many practical applications, we can not assume each input variable (except $x_i$) of the same distribution along $x_i$ axis, respectively. This means for a certain system, if we obtain different sampling data sets, the fuzzy average may be different. Then the conclusion for the importance of input variables may be different.

This requires us to preprocess the sampling data set to make it a representative data set before using them to determine the influence of input variables.

### *B. Change the Distribution of Sampling Data Using Fuzzy Clustering*

Using fuzzy average method, the significance of variable $x_i$ can be correctly evaluated without the interference of other input variables only when all other input variables have the same distribution along $x_i$ axis. To transform the sampling data set into this form, we consider using fuzzy clustering to change the distribution of the data set. Again we use the example of the system with two input variables $x_1$ and $x_2$ shown in Fig. 1.

Suppose more data points fall into a small region than into other regions in the $x_1 - x_2$ space (as in Fig. 2 (B)). We consider using fuzzy clustering method to partition the data points into groups. The number of data points in each group (fuzzy cluster) will be different since the distribution of the data is uneven. If we use one data point (for instance, the fuzzy cluster center) to represent each partition, we can obtain a new data set with the distribution of fuzzy clusters. Since different number of sampling data in different partitions will be replaced by the same number of cluster center, we may obtain a new data set with better distribution.

Below we will first use Fuzzy C-means method to cluster the data. Then after discussing the drawback in this application, we will propose an improved method.

*1) Fuzzy C-means Method:* Fuzzy c-means allows one data point to belong to two or more clusters [15] [16]. It provides a method that group data points in multidimensional space into a specific number of clusters. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} {}_{ij}^{m} \| x_i - b^j \|^2 \qquad (8)$$

where $m$ is the number of clusters. $_{ij}$ is the degree of membership of $x_i$ in the cluster $j$, $x_i$ is the ith of n-dimensional measured data, $b_j$ is the n-dimension center of the cluster, and $\| \; \|$ is a norm expressing the distance between measured data and cluster center.

We first use Fuzzy c-means to cluster the data shown in Fig. 2 (B). The number of clusters need to be predefined. We assume it is 50. Then we use the produced 50 centers of the clusters to form a new data set. After this process, the distribution of the obtained new data set is shown in Fig. 3. The crosses are the original data points and the circles are the centers of the clusters.

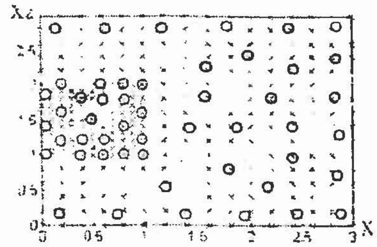


Fig. 3. Clusters generated using Fuzzy c-means

As we can see, the distribution of the data did not change. Why Fuzzy c-means clustering can not change the distribution? As we know that during the iteration process of Fuzzy c-means, the membership and the cluster centers are updated to minimize the total weighted distance between data points and the cluster centers of the fuzzy partition. Thus it is reasonable that in areas where more data points exist, there must be more cluster centers in order to make the total weighted distance between all data points and the cluster centers smaller. Therefore we can not change the data distribution when only using Fuzzy c-means.

*2) Generate Even Cluster Distribution :* To generate even cluster distribution, we partition the input space using fuzzy rules before applying Fuzzy c-means: we build a fuzzy rule base for the nonlinear system; those data points that can excite a particular fuzzy rule with high firing strength are grouped to the same partition. The fuzzy rule base is in the form of

IF $x_1$ is $A_{11}$ and ... and $x_n$ is $A_{n1}$ THEN $y$ is $y^1$
IF $x_1$ is $A_{12}$ and ... and $x_n$ is $A_{n2}$ THEN $y$ is $y^2$
...
IF $x_1$ is $A_{1m}$ and ... and $x_n$ is $A_{nm}$ THEN $y$ is $y^m$

where $A_{ij} (i = 1, 2, ..., n; j = 1, 2, ..., m)$ and $y^j$ are fuzzy sets of $x_i$ and $y$, respectively. If the width $\sigma^j$ of Gaussian membership functions are the same for all the fuzzy sets, the partition is an even partition.

The method is implemented as follows: we take the first sampling data point as center of a cluster and build a corresponding fuzzy rule. The center of the Gaussian membership function is $\overline{x}_i^j = x_i^j$; the width $\sigma^j$ is 1/30 of the normalized range of the input variable.

For every sampling data point, we calculate the firing strength (degree of fulfillment) of each existing rule:

$$_j = \prod_{i=1}^{n}( {}_{ij}(x_i)) = \prod_{i=1}^{n}\exp(-(\frac{x_i - \overline{x}_i^j}{\sigma^j})^2) \qquad (9)$$

AND operation is used in (9).

If the firing strength $_j$ , then the sampling data point is close to the data points in the partition. Thus it belongs to this partition. is a predefined threshold as the least acceptable degree and it determines the extent of the similarity to be classified into the partition. If the firing strength is less than the threshold , then a new fuzzy rule (a new partition) should be created.

After all the data are partitioned, we use Fuzzy c-means algorithm to cluster data points in each small partition. The same number of clusters are set for each small partition so that the distribution can be more even. Or, if the partition is small enough, we can set only one cluster for each partition and find its center by Fuzzy c-means. We would like to use the centers of the clusters to represent the clusters. But for real world systems, the corresponding output of the system to the centers are not available, if the centers are not coincident to the existing data points. So we use the sampling data point closest to the center of a cluster to represent the cluster. The closest data point is decided by its Euclidean distance to the center.

When only one cluster is set for each partition, the number of clusters is the same as the number of partitions. So if $\sigma'$ is too small, the distribution may not change and the number of clusters will be large; if it is too big, some clusters with less data points may be combined into other clusters and we lost its representation when only the cluster center is kept. We often take $\sigma'$ as 1/30 of the range of the normalized input variables. This can partition the input space into many small partitions which can represent the input space adequately, and at the same time remove redundant sampling data points inside the small partitions.

The procedure of FAFCD is listed as follows:

1. Normalize the original data set.
2. Partition the original data
3. Use Fuzzy c-means method to form a new data set.
4. Calculate the fuzzy average of $y$ in each input-output space on the new data set.
5. Identify key input variables according to their Influence Rate.

The data in Fig. 2 (B) becomes a even distribution after being processed by the above method (see Fig. 4 (B)). Hence the fuzzy average of $y$ in $x_2 - y$ space can reflect the $x_2 - y$ relationship correctly now (Fig. 4) (A). These results are very similar to the results generated in Fig. 1, which is a uniform distribution.

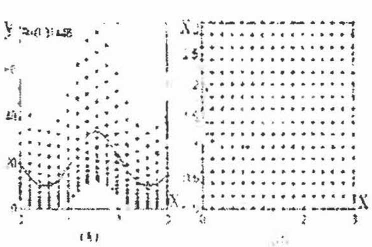

Fig. 4. Results using FAFCD. (A) The defuzzified curve of $y$ in $x_2 - y$ space (B) The input space $x_1 - x_2$

### C. Other Considerations

Some considerations and improvements should be implemented. First, before applying FAFCD, the data set should be normalized to avoid bias. Second, to reduce the number of membership functions to be calculated, only those membership functions have a center close to the current value of $x$, are taken into account. Third, to remove the outliers, a threshold is predefined and those clusters contain very few data points under the threshold are neglected.

## III. SIMULATIONS AND RESULTS

We will evaluate the performance of FAFCD using two nonlinear systems. In the first experiment, we apply our algorithm to the automobile fuel consumption prediction problem. In the second experiment, we apply it to the EMG signal evaluation system.

### A. Evaluate FAFCD on Automobile Fuel Consumption Prediction Problem

Data set used here is from the UCI Machine Learning Repository. It consists of 392 instances. Each instance is composed of six input variables (number of cylinders, displacement, horsepower, weight, acceleration, and model year) and one output variable (MPG). The input variables and output variable are denoted as $x_i$ $(i = 1, 2, ..., 6)$ and $y$, respectively. Fig. 5 shows the relationship between two input variables ($x_4$, $x_5$) and MPG.

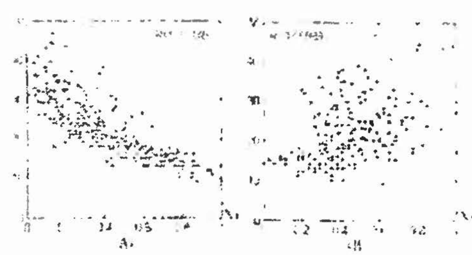

Fig. 5. Relationship between two input variables and the output obtained using FAFCD. (A) Weight ($x_4$) - MPG ($y$) relationship. (B) Acceleration ($x_5$) - MPG ($y$) relationship.

From Table I, we can get the conclusion that the variables weight and horsepower have most significant influence on MPG of a automobile.

TABLE I

INFLUENCE RATE OF ALL INPUT VARIABLES

| Input | Variable Name | Influence Rate |
|---|---|---|
| $x_1$ | Number of cylinders | 0.3852 |
| $x_2$ | Displacement | 0.4488 |
| $x_3$ | Horsepower | 0.5851 |
| $x_4$ | Weight | 0.5945 |
| $x_5$ | Acceleration | 0.5589 |
| $x_6$ | Model year | 0.4239 |

If the distribution (or density) of the original data set is changed by repeating some of the data points, the Influence Rate of acceleration obtained without clustering (the method used in [3]) becomes 0.6076 (Fig. 6 (A)), while the Influence

Rate of other variables remain almost unchanged. Then the conclusion becomes that acceleration is the most important variable, which is incorrect. If we use the FAFCD method, it can give an Influence Rate of acceleration of 0.5612 (Fig. 6 (B)). Thus the conclusion of importance of variables will not be affected.

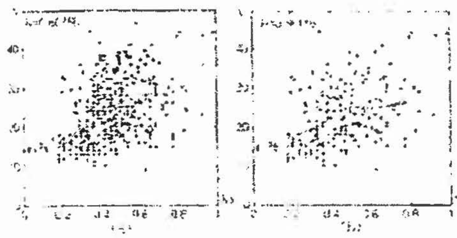

Fig. 6. Results on the modified MPG data using different methods. (A) Without clustering (method used in [3]) (B) Using FAFCD.

## B. Evaluate FAFCD on EMG Signal Evaluation System

EMG signal evaluation system is a model in which we use the kinematics parameters during a motion (kinematics variables) and the anthropometric characteristics of the subjects (subject variables) to evaluate the corresponding EMG signals in ten trunk muscles generated during the lifting motion of the subjects. Our objective is to find out those variables that have significant influence on the EMG signals and find out how they affect the EMG signals.

*1) Input variables:* Not knowing which variables really affect the EMG signals, all the associated kinematic variables and subject variables are recorded. The twelve kinematic variables are dynamic variables which change their values during the motion. While the fifteen subject variables are static variables which are the anthropometric characteristics of the subjects and they are the same during a motion for a particular subject. We want to find out how these variables affect the EMG signals. The sampling data set contains six trials of motions conducted by 249 subjects. Each trial has 20 sampling data points. Every subject conducted all the trials. Therefore the total number of data points N is 29880. Each data point consist of 27 input variables ($x_i, i = 1, 2, \ldots, 27$) and 10 output variables ($y_j, j = 1, 2, \ldots, 10$). The outputs are EMG signals of 10 trunk muscles.

*2) Results:* First, we use the distribution of the original data set without clustering to calculate the $x_i - y_j$ relationships. The results obtained are not satisfactory. For many of the input-output relationships, there is a drop in the middle of the range of variable $x_i$. An example is shown in Fig. 7 (A). It is uninterpretable from the ergonomics point of view. We think that it is probably because of the uneven distribution of the sampling data. Certain conditions may have appeared more frequently during the motion than other conditions and thus have distorted the fuzzy average curve.

Then we apply FAFCD to obtain the input-output relationships on the new data set produced using fuzzy clustering. A summary of cluster properties is shown in Table II. Those

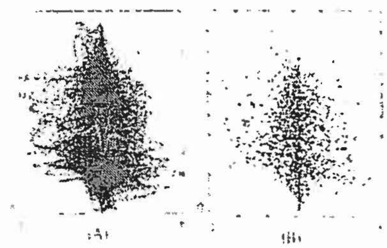

Fig. 7. The relationship between Axis Trunk Velocity ($x$ axis) and EMG signal ($y$ axis). (A) Without clustering (B) After clustering

clusters contain less than 10 data points (about 0.033% of the total) are considered as outliers and are removed from the data set.

TABLE II

SUMMARY ABOUT THE CLUSTERS

| | |
|---|---|
| Total number of data points | 29880 |
| Number of clusters | 4816 |
| Range of number of data points in the clusters | 1 - 1456 |
| Number of clusters with less than 10 data points | 66 |

Fig. 8. Relationship between two kinematic variables ($x$ axis) and EMG signal ($y$ axis) of muscle Right Latissimus Dorsi (RLD). (A) sagittal trunk moment (B) lateral trunk moment

Fig. 7 (B) shows the result of the same example as in Fig. 7 (A), using FAFCD. As we expected, the drop in Fig. 7 (A) disappeared and the result has a clear physical explanation now.

Fig. 8 shows the relationship between two kinematic variables and EMG signals of muscle Right Latissimus Dorsi. The relationships of inputs to the other muscles can be obtained similarly. With these relationships, we can have a better understanding to the muscle activities. At the same time, the importance of the input variables are indicated by their Influence Rate $R$. The importance of the kinematic variables and subject variables are ranked as shown in Fig. 9 and Fig. 10 respectively. It is clear that kinematic variables have more influence on the EMG signals than subject variables. Thus, these 12 kinematic variables should all be selected as inputs in modelling. As for subject variables, five variables (standing height, shoulder height, lower arm length, spine length, lower leg length) have bigger influence than the others. These variables should also be taken as inputs in modelling. While after examining the two variables "standing height" and "shoulder height", we found that the Influence Rate of these two variables are very similar, for every muscle. In other words, these two variables are correlated. Therefore we can remove one of them So at last 12 kinematic variables and four subject variables are kept. The input dimension is decreased from 27 to 16.

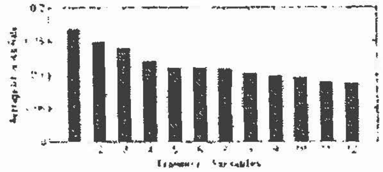

Fig 9. Rank kinematic variables by their Average Influence Rate (1: Sagittal trunk moment, 2: Lateral trunk moment 3: Axis trunk angle 4 Sagittal trunk velocity, 5: Axis trunk moment, 6: Sagittal trunk angle, 7: Axis trunk acceleration, 8: Sagittal trunk acceleration, 9: Lateral trunk velocity, 10. Axis trunk velocity, 11: Lateral trunk angle, 12: Lateral trunk acceleration)

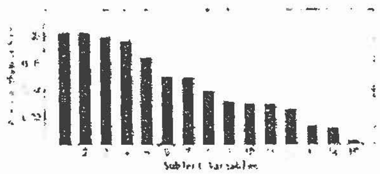

Fig. 10. Rank subject variables by their Average Influence Rate (1: Standing height, 2: Shoulder height, 3: Lower arm length, 4:Spine length, 5: Lower leg length, 6. Body weight, 7 Trunk breadth (xyphoid), 8: Trunk circumference, 9: Trunk depth (xyphoid), 10: Trunk breadth (pelvis), 11: Upper arm length, 12: Elbow height, 13: Upper leg length, 14: Trunk depth (pelvis), 15: Age)

Knowing the influence of each variable, we can use the selected variables as inputs, instead of using hypothetically selected variables. This will reduce the complexity of the model and time of modelling. In building a fuzzy logic model, we can adjust fuzzy rules according to this knowledge. In building a neural network model, less input variables means less free parameters and shorter training time.

## IV. Conclusions

FAFCD can find out the importance of specific input variables and how they influence the output, without the interference of the distribution of sampling data set. The method is straightforward and easy to implement. It was applied to the EMG signal evaluation system and obtained satisfactory results.


## Acknowledgment

This study was conducted under a research grant on the "Development of a Neuro-Fuzzy System to Predict Spinal Loading as a Function of Multiple Dimensions of Risk", sponsored by the National Institute for Occupational Safety and Health (DHHS).


## References


[1] E.B.Bartlett, "Self Determination of Input Variable Importance by Neural Networks", Neural, Parallel and Scientific Computations, vol. 2, no. 1, 1994 pp 103-114

[2] B. Yuan, G. Klir, "Data-driven identification of key variables", - pp 161 186

[3] Y. Lin, G.A. Cunningham, "A new approach to fuzzy-neural system modeling"; IEEE Transactions on Fuzzy Systems, vol 3, no. 2, May 1995, pp. 190-198.

[4] Y. Lin; G.A. Cunningham, S.V. Coggeshall, R.D. Jones, "Nonlinear system input structure identification, two stage fuzzy curves and surfaces", IEEE Transactions on Systems, Man and Cybernetics, Part A, vol. 28, no 5, Sept 1998 pp. 678 -684

[5] M.F. Azeem, ; M. Hanmandlu, N. Ahmad, "Structure identification of generalized adaptive neuro-fuzzy inference systems", IEEE Transactions on Fuzzy Systems, vol. 11, no 5, Oct. 2003, pp. 666 - 681

[6] F.C. Morabito, M. Versaci, "A fuzzy neural approach to localizing holes in conducting plates", IEEE Transactions on Magnetics, vol. 37, no 5, Sept. 2001 pp 3534 - 3537

[7] S.E. Papadakis, J.B. Theocharis, J. Kiartzis, A.G. Bakirtzis, "A novel approach to short term load forecasting using fuzzy neural networks", IEEE Transactions on Power Systems, vol. 13, no. 2, May 1998, pp. 480 - 492

[8] A.H Sung, "Ranking input importance in neural network modeling of engineering problems", Proceedings of IEEE World Congress on Computational Intelligence, vol 1, 4-9 May 1998, pp. 316 - 321 vol.1

[9] B. Bouqata, A. Bensaid, R. Palliam, A.F.G. Skarmeta, "Time series prediction using crisp and fuzzy neural networks: a comparative study", Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering, 26-28 March 2000, pp 170 - 173

[10] M.F. Azeem, M Hanmandlu, N. Ahmad, "A new criteria for input variable identification of dynamical systems"; Proceedings of IEEE Region 10 International Conference on Global Connectivity in Energy, Computer, Communication and Control, vol. 1, 17-19 Dec 1998, pp. 230-233

[11] F.C Morabito, M Versaci, "A fuzzy-neural approach to real time plasma boundary reconstruction in tokamak reactors", Proceedings of International Conference on Neural Networks, vol. 1, 9 12 June 1997, pp 43 -47

[12] S Papadakis, J Theocharis, "An efficient fuzzy neural modeling approach using the fuzzy curve concept", Proceedings of the Third IEEE International Conference on Electronics, Circuits, and Systems, vol 1, 13-16 Oct. 1996, pp. 279 - 282.

[13] J.C. Dunn, "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters"; Journal of Cybernetics vol 3 pp 32 57 1973

[14] J.C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algoritms", Plenum Press, New York 1981

[15] E.S ichrich, J. Ke, L.O. Hall, D.B Goldgof, "Fast accurate fuzzy clustering through data reduction"; IEEE Transaction on Fuzzy Systems, vol. 11, no. 2, April. 2003.

[16] S. Auephanwiriyakul, J.M. Keller, "Analysis and efficient implementation of a linguistic fuzzy c-means", IEEE Transactions on Fuzzy Systems, vol 10 , no. 5 , Oct 2002 pp. 563 - 582

[17] R.A. Fisher, "Multiple measurements in taxonomic problems"; Annals of Eugenics vol. 7 no. 2 pp. 179 - 188 1936

[18] J Zurada, W Karwowski, W.S. Marras, "A Neural Network-Based System for Classification of Industrial Jobs with Respect to the Risk of Low Back Disorders", Applied Ergonomics, vol. 28 no. 1, 1997 pp 49-58

[19] I.W. ee, W. Karwowski, W.S. Marras, D. Rodrick, "A Neuro-Fuzzy Modeling of Myoelectrical Activity of Trunk Muscles Due to Manual Lifting Tasks", Ergonomics, 2003, vol. 15 no. 1-3, pp. 285-309

[20] J.M Zurada, A. Malinowski, I. Cloete, "Sensitivity analysis for minimization of input data dimension for feedforward neural network",Procedings of 1994 IEEE International Symposium on Circuits and Systems, vol. 6, May 30, 1994, pp. 447-450

[21] Y. Hou, J.M. Zurada, W. Karwowski, W.S. Maras, "Identification of Key Variables using Fuzzy Average with Fuzzy Cluster Distribution", Submitted to IEEE Transactions on Fuzzy Systems.