Focus

Soft Computing 8 (2004) 681–688 © Springer-Verlag 2003

Formulation of linguistic regression model based on natural words

Y. Toyoura, J. Watada, M. Khalid, R. Yusof

Abstract When human experts express their ideas and thoughts, human words are basically employed in these expressions. That is, the experts with much professional experiences are capable of making assessment using their intuition and experiences. The measurements and interpretation of characteristics are taken with uncertainty, because most measured characteristics, analytical result, and field data can be interpreted only intuitively by experts. In such cases, judgments may be expressed using linguistic terms by experts. The difficulty in the direct measurement of certain characteristics makes the estimation of these characteristics imprecise. Such measurements may be dealt with the use of fuzzy set theory. As Professor L. A. Zadeh has placed the stress on the importance of the computation with words, fuzzy sets can take a central role in handling words [12, 13]. In this perspective fuzzy logic approach is offten thought as the main and only useful tool to deal with human words. In this paper we intend to present another approach to handle human words instead of fuzzy reasoning. That is, fuzzy regression analysis enables us treat the computation with words. In order to process linguistic variables, we define the vocabulary translation and vocabulary matching which convert linguistic expressions into membership functions on the interval [0–1] on the basis of a linguistic dictionary, and vice versa. We employ fuzzy regression analysis in order to deal with the assessment process of experts' from linguistic variables of features and characteristics of an objective into the linguistic expression of the total assessment. The presented process consists of four portions: (1) vocabulary translation, (2) estimation, (3) vocabulary matching and (4) dictionary. We employed fuzzy quantification theory type 2 for estimating the total assessment in terms of linguistic structural attributes which are obtained from an expert.

Keywords Linguistic regression model, Natural word, Fuzzy regression model

1 Introduction

As Professor L. A. Zadeh has placed the stress on the importance of the computation with words, fuzzy sets can take a central role in handling words [12, 13]. In this perspective fuzzy logic approach is offten thought as the main and only useful tool to deal with human words. In this paper we intend to present another approach to handle human words instead of fuzzy reasoning. That is, fuzzy regression analysis enables us treat the computation with words.

We intend to abstract the latent structure under the relations between words. Human words can be translated into fuzzy sets such as fuzzy numbers, which is employed in a fuzzy controller. If it is possible such as fuzzy control to formulate a dictionary between fuzzy number and a word, we can build the relations under the data in terms of fuzzy regression analysis.

Consequently, only experts with much professional experiences are capable of making assessment using their intuition and experiences. The measurements and interpretation of these characteristics are taken with uncertainty, because most measured characteristics, analytical results, and field data can be interpreted only intuitively by experts. In such cases, judgments may be expressed by experts with linguistic terms. The difficulty in the direct measurement of certain characteristics makes the estimation of these characteristics imprecise. Such measurements may be dealt with the use of fuzzy set theory [2, 9–11].

Watada, Fu and Yao [4, 5] proposed a model of damage assessment by using the information given by experts through fuzzy multivariate analysis.

In order to process linguistic variables, we define the vocabulary translation and vocabulary matching which convert linguistic expressions into membership functions on the interval [0–1] on the basis of a linguistic dictionary, and vice versa. We employ Fuzzy Regression Analysis [2, 6] in order to deal with the assessment process [7, 8] of experts' from linguistic variables of features and characteristics of an objective into the linguistic expression of the total assessment.

Published online: 23 September 2003

Y. Toyoura
Osaka Institute of Technology, School of Industrial Engineering

J. Watada (✉)
Graduate School of Information, Production and Systems, Waseda University, 2-2 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0135 Japan,
e-mail: junzow@osb.att.ne.jp

M. Khalid, R. Yusof
Center for Artificial Intelligence and Robotics, Universiti Teknologi Malaysia

This research was supported in part by Grant-in Aid for Scientific Research(C-2); Grant No.11680459 of Ministry of Education of Science, Sports and Culture.

Table 1. Linguistic data given by experts

Training sample	Linguistic variables					Linguistic objective
1	$\mathbf{L}_1(1)$	$\cdots$	$\mathbf{L}_i(1)$	$\cdots$	$\mathbf{L}_K(1)$	$\mathbf{L}_0(1)$
2	$\mathbf{L}_1(2)$	$\cdots$	$\mathbf{L}_i(2)$	$\cdots$	$\mathbf{L}_K(2)$	$\mathbf{L}_0(2)$
3	"good"	$\cdots$	"bad"	$\cdots$	"very bad"	"bad"
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
ω	$\mathbf{L}_1(\omega)$	$\cdots$	$\mathbf{L}_i(\omega)$	$\cdots$	$\mathbf{L}_K(\omega)$	$\mathbf{L}_0(\omega)$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
n	$\mathbf{L}_1(n)$	$\cdots$	$\mathbf{L}_i(n)$	$\cdots$	$\mathbf{L}_K(n)$	$\mathbf{L}_0(n)$

2
Linguistic variable and vocabulary matching

In making assessments, experts often (*a*) evaluate various features and characteristics, and (*b*) assess the objective in a linguistic form. For instance, though it is possible to measure the production volume, it is difficult to analytically interpret the numerical value in terms of its influence of this amount on the future decision making.

On the other hand, experts may be able to express (a) the effect of a given sales trend as "good" or "not good", and (b) the total assessment or state in a linguistic form such as "very good", "good" or "not good" as shown in Table 1.

Consider the simplest possible value such as production volume. Let **L** be a linguistic variable which can have values in the set **X**. As examples of such a variable, let us consider a production volume as well as the number of visiting customers.

$$\mathbf{L}_{\text{volume}}(\text{production})\text{is/extremely bad/}$$

$$\mathbf{L}_{\text{number}}(\text{visiting customers})\text{is/bad/}$$

where subscripts (number) in $\mathbf{L}_{(\text{visiting customers})}$ denote the state of a visiting customers.

These expressions "good", "bad", "extremely bad" can be defined with fuzzy grades on [0,1] such as $\mathbf{U}_{(\text{good})}$, $\mathbf{U}_{(\text{bad})}$, $\mathbf{U}_{(\text{extremely bad})}$. Denote π as a possibility distribution. We can identify the possibility of the state of a sales trend with the degree of its descriptive adjective on [0,1]. For example,

$$\pi_{(\text{number})}(\text{visiting customer}) \equiv \pi_{(\text{state})}(\text{the number}) \equiv \mathbf{U}_{(\text{extremelybad})}$$

Define the dictionary of descriptive adjectives in which a descriptive adjective corresponds with its fuzzy grade on [0, 1].

The objective of this study is to model the experts' assessment process through which experts might evaluate the possibility of sales trend **L** on the basis of states of its features and characteristics $\mathbf{L}_i(i = 1, 2, \ldots, K)$, where these values **L** and $\mathbf{L}_i$ are expressed in a linguistic form. In other words, we intend to determine the linguistic assessment process **F** of linguistic variable $\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_K$, which produces a linguistic value of an objective **Z**. It can be written in the form

$$\mathbf{Z} = \mathbf{F}(\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_K) \tag{1}$$

We define the descriptive adjectives "extreme", "very", and so on. Let us make a dictionary for corresponding linguistic expressions $\mathbf{L}_i$ and fuzzy grades $\mathbf{U}_{\mathbf{L}_i}$. If we obtain the assessment $\mathbf{L}_i(i = 1, 2, \ldots, K)$, $\mathbf{L}_i$ is understood in terms of $\mathbf{U}_{\mathbf{L}_i}$ through the dictionary build by experts. Let us divide this linguistic assessment process F into three portion:

(1) translation of attributes form linguistic values $\mathbf{L}_i$ into fuzzy grades $\mathbf{U}_{\mathbf{L}_i}$

$$\mathbf{U}_{\mathbf{L}_i} \equiv (u_i, c_i^l, c_i^r)$$

where u_i denotes the central value with grade 1, c_i^l left side fuzziness and c_i^r right side fuzziness, respectively.

(2) estimation of total damage by the fuzzy assessment function

$$\mathbf{V} = f(\mathbf{U}_{\mathbf{L}_1}, \mathbf{U}_{\mathbf{L}_2}, \ldots, \mathbf{U}_{\mathbf{L}_K}) \tag{2}$$

which produces a fuzzy grade **V** in terms of fuzzy grades of attributes $\mathbf{U}_{\mathbf{L}_i}$ where a suffix **L** is not attached to **V** because it is unknown, and the detail of this calculation is explained in Sect. 3.

(3) linguistic matching of the fuzzy grade of the objective with the dictionary wherein the linguistic value **Z** is decided for the objective.

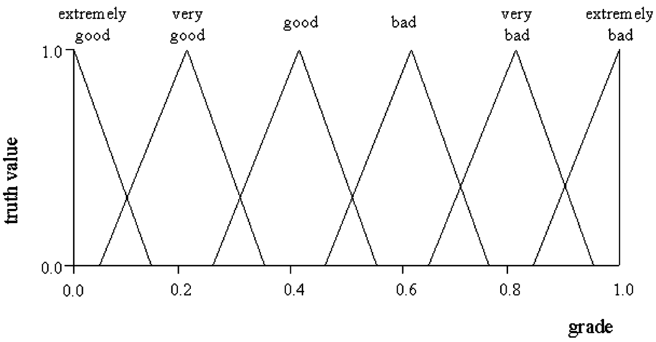
Let us define the vocabulary matching by using the following minimax calculation

$$\mathbf{Z}_0 \simeq \max_{\mathbf{W}_i \in \mathbf{D}} \left\{ \max_t \mu_{\mathbf{V}}(t) \bigwedge \mu_{W_i}(t) \right\} \tag{3}$$

where $\mathbf{Z}_0 \simeq \max_{\mathbf{W}_i \in \mathbf{D}} f(\mathbf{W}_i)$ denotes that $\mathbf{Z}_0$ is the word in **D** which realizes the maximum value of f, $\mu_{\mathbf{V}}$ denotes a membership function of **V**, and $\mu_{\mathbf{W}_i}$ denotes a membership function of a word $\mathbf{W}_i$ as included in the dictionary **D**. This procedure is illustrated in Fig. 1. We assign a word $\mathbf{L}_0$ "very bad" to the fuzzy grade of the total assessment **V** showed a dotted line in Fig. 2.

3
Determination of Fuzzy Regression Model

After the model of experts' assessment process is formulated, it is desirable to determine the fuzzy assessment function f of the total assessment as shown in Fig. 2. This is a fuzzy function with which K fuzzy grades of attributes, $\mathbf{U}_{\mathbf{L}_i}$, can be transformed to one fuzzy grade of its total assessment, **V**. We must determine this fuzzy regression model on the basis of training data $\omega(\omega = 1, 2, \ldots, n)$ as given by experts in order to mimic experts' assessment process. Let us employ the method proposed as Fuzzy Quantification Theory Type 1 by



$\mathbf{U}_{(extremely\ good)} = (0.00, 0.00, 0.15)$
$\mathbf{U}_{(very\ good)} = (0.20, 0.15, 0.15)$
$\mathbf{U}_{(good)} = (0.40, 0.15, 0.15)$
$\mathbf{U}_{(bad)} = (0.40, 0.15, 0.15)$
$\mathbf{U}_{(very\,bad)} = (0.80, 0.15, 0.15)$
$\mathbf{U}_{(extremely\ bad)} = (1.00, 0.15, 0.15)$

Fig. 1. Dictionary of descriptive adjectives

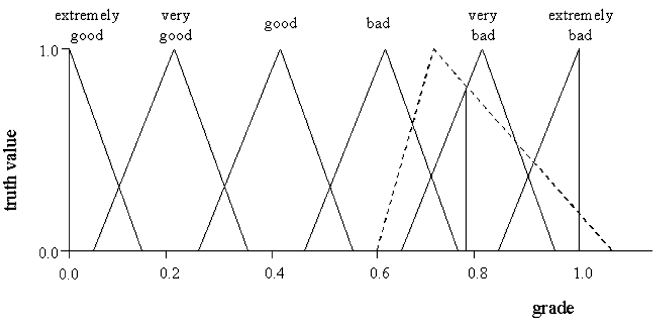


Fig. 2. Vocabulary matching

Watada, Tanaka and Asai [6] for determining this fuzzy regression model f.

Table 1 shows the training data given in linguistic form by experts. This data in Table 1 is translated into data of fuzzy grades in terms of the dictionary (see Table 2). In Table 2, the fuzzy grades of attributes i and of the total damage of sample structures ω are denoted by $\mathbf{U}_{\mathbf{L}_i}(\omega)$ and $\mathbf{V}_{\mathbf{L}_0}(\omega)$, respectively where $i = 1, 2, \ldots, K$ and $\omega = 1, 2, \ldots, n$. We should note that the model is estimated using given linguistic data $\mathbf{L}_0, \mathbf{L}_1, dots, \mathbf{L}_K$.

Assume that all fuzzy grades have triangular shapes which are normal and convex. We employ a fuzzy linear function for a fuzzy regression model f as

$$\mathbf{V}_{\mathbf{L}_0} = f(\mathbf{U}_{\mathbf{L}_1}, \mathbf{U}_{\mathbf{L}_2}, \ldots, \mathbf{U}_{\mathbf{L}_K}) = \sum_{i=1}^{K} \mathbf{A}_i \mathbf{U}_{\mathbf{L}_i}(\omega) \quad \omega = 1, 2, \ldots, n \tag{4}$$

It is noted that $\mathbf{V}_{\mathbf{L}_0}$ is given value instead of $\mathbf{V}$ is estimation.

Using n relation of training samples

$$\mathbf{V}_{\mathbf{L}_0}(\omega) = \sum_{i=1}^{K} \mathbf{A}_i \mathbf{U}_{\mathbf{L}_i}(\omega) \quad \omega = 1, 2, \ldots, n \tag{5}$$

We must specify the best fuzzy parameters $\mathbf{A}_i$ in terms of these relations. Two criteria are employed in order to define the goodness of the fuzzy linear function. One criterion is fitness of the fuzzy regression model, h, and the other is fuzziness included in the fuzzy regression model, $\mathbf{S}$.

(i) Fitness

Assumed that an estimated values $\mathbf{V}_{\mathbf{L}_0}(\omega)$ is obtained by the fuzzy linear function f, fitness $b(\omega)$ of $\mathbf{Y}_0(\omega)$ to a sample value $\mathbf{Y}(\omega)$ is $\mathbf{V}_{\mathbf{L}_0}(\omega)$ to a sample value $\mathbf{V}_{\mathbf{L}_0}(\omega)$ is defined by

$$h(\omega) = \bigvee_{y \in R} \left\{ \mu_{\mathbf{L}_0(\omega)}(y) \bigwedge \mu_{\mathbf{L}_0(\omega)}(y) \right\} \tag{6}$$

(ii) Fuzziness

The fuzziness $\mathbf{S}^{\alpha}$ included in the fuzzy function at α-level is defined by

$$\mathbf{S}^{\alpha} = \sum_{i=1}^{K} (\overline{a_i} - \underline{a_i}) \tag{7}$$

where $\overline{a_i}$ and $\underline{a_i}$ are numbers which specify an α-level set $\mathbf{A}_i^{\alpha}$, i.e.,

$$\mathbf{A}_i^{\alpha} = [\overline{a_i}, \underline{a_i}] \tag{8}$$

In this paper, we deal with triangular fuzzy numbers. Note that

$$\mathbf{A}_i^{\alpha} \equiv \left[a_i - (1-\alpha)c_i^l,\ a_i + (1-\alpha)c_i^r\right]$$

as $\mathbf{A}_i^{\alpha}$ is defined by

$$\mathbf{A}_i \equiv (a_i, c_i^l, c_i^r)$$

Note that these two indices of fuzziness and fitness are incompatible with each other. The higher fitness we seek, the resulting model will be.

Table 2. Fuzzy data translated

Training Sample	Fuzzy grade of variables					Fuzzy grade of the objective
1	$\mathbf{U}_{\mathbf{L}_1}(1)$	$\cdots$	$\mathbf{U}_{\mathbf{L}_i}(1)$	$\cdots$	$\mathbf{U}_{\mathbf{L}_K}(1)$	$\mathbf{V}_{\mathbf{L}_0}(1)$
2	$\mathbf{U}_{\mathbf{L}_1}(2)$	$\cdots$	$\mathbf{U}_{\mathbf{L}_i}(2)$	$\cdots$	$\mathbf{U}_{\mathbf{L}_K}(2)$	$\mathbf{V}_{\mathbf{L}_0}(2)$
3	(0.4,0.15,0.15)	$\cdots$	(0.6,0.15,0.15)	$\cdots$	(0.8,0.15,0.15)	(0.6,0.15,0.15)
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
ω	$\mathbf{U}_{\mathbf{L}_1}(\omega)$	$\cdots$	$\mathbf{U}_{\mathbf{L}_i}(\omega)$	$\cdots$	$\mathbf{U}_{\mathbf{L}_K}(\omega)$	$\mathbf{V}_{\mathbf{L}_0}(\omega)$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
n	$\mathbf{U}_{\mathbf{L}_1}(n)$	$\cdots$	$\mathbf{U}_{\mathbf{L}_i}(n)$	$\cdots$	$\mathbf{U}_{\mathbf{L}_K}(n)$	$\mathbf{V}_{\mathbf{L}_0}(n)$

3.1 Formulation of the problem

We formulate a fuzzy assessment function by minimizing its fuzziness **S** under the constraints that an estimated fuzzy grade of structural of total damage of each sample is fit to the fuzzy grade given by experts with the fitness greater than or equal to the given value h^0, called fitness standard.

Problem

If data are given such as that listed in Table 2, the problem is to determine a fuzzy linear function

$$\mathbf{V}_{\mathbf{L}_0}(\omega) = \sum_{i=1}^{K} \mathbf{A}_i \cdot \mathbf{U}_{\mathbf{L}_i}(\omega) \tag{9}$$

which minimizes the fuzziness

$$\mathbf{S} = \sum_{i=1}^{K} (\overline{a_i} - \underline{a_i}) \tag{10}$$

under the conditions that

$$h(\omega) = \bigvee_{y \in R} \left\{ \mu_{\mathbf{L}(\omega)}(y) \bigwedge \mu_{\mathbf{L}_0(\omega)}(y) \right\} \geq h^0 \quad \omega = 1, 2, \ldots, n \tag{11}$$

where $h(\omega)$ indicates the fitness of the estimated value with respect to a sample ω and h^0 denotes the fitness standard, and $\overline{a_i}$ and $\underline{a_i}$ are defined as

$$A_i^{h^0} = [\underline{a_i}, \overline{a_i}] \tag{12}$$

Note that we will employ triangular approximation for calculation of fuzzy number in this paper, although the precise calculation has been obtained in Tanaka, Watada and Asai [3].

That is, the membership function $\mu_{\mathbf{L}_0}(y)$ of the fuzzy grade of structural total damage $\mathbf{V}_0$ can be obtained through extension principle to

$$\mu_{\mathbf{V}_0}(y) = \bigvee_{(t_i,u_i) \left| \begin{array}{l} y = \sum_i t_i u_i \\ 0 \leq t_i \leq 1 \\ 0 \leq u_i \leq 1 \end{array} \right.} \bigwedge_{i=1}^{K} \left\{ \mu_{\mathbf{A}_i}(t_i) \bigwedge \mu_{\mathbf{L}_i}(u_i) \right\} \tag{13}$$

When the value $\mathbf{V}_{\mathbf{L}_0}(\omega)$ given Eq. (9) is obtained, Eq. (9) enables us to define its membership function using parameter t_i for $\mathbf{A}_i$ and parameter u_i for $\mathbf{L}_i$ of Eq. (9).

3.2 The fuzzy grade

In this section, we discuss the heuristic method to determine a fuzzy assessment function by using non-fuzzy grades of assessment attributes, i.e., $\mathbf{U}_i$ are fuzzy numbers in [0, 1].

According to the sign of $\mathbf{A}_i$, the production of fuzzy number $\mathbf{A}_i$ and $\mathbf{U}_{\mathbf{L}_i}$ is given in the following three cases:

(i) In this case where $\overline{a}_i \geq \underline{a}_i \geq 0$

$$(\mathbf{A}_i\mathbf{U}_{\mathbf{L}_i})^{h^0} = [\underline{a}_i\underline{u}_i,\ \overline{a}_i\overline{u}_i] \tag{14}$$

(ii) In this case where $\overline{a}_i \leq \underline{a}_i \leq 0$

$$(\mathbf{A}_i\mathbf{U}_{\mathbf{L}_i})^{h^0} = [\underline{a}_i\overline{u}_i,\ \overline{a}_i\underline{u}_i] \tag{15}$$

(iii) In this case where $\overline{a}_i \leq \underline{a}_i \leq 0$

$$(\mathbf{U}_{\mathbf{L}_i})^{h^0} = [\underline{a}_i\overline{u}_i,\ \overline{a}_i\underline{u}_i] \tag{16}$$

It is difficult to solve analytically this problem. Therefore, we employ the heuristic approach for solving the problem. The procedure is as follows.

An α-level set of the fuzzy degree of a structural attribute $\mathbf{U}_{\mathbf{L}_i}(i = 1, 2, \ldots, K)$ at h^0 is assumed to be denoted by Eq. (12).

Step 1

Let the trial count $r = 1$ and let

$$\left(\mathbf{A}_i^{(r)}\mathbf{U}_{\mathbf{L}_i}\right)^{h^0} = \left[\underline{a}_i^{(r)}u_i,\ \overline{a}_i^{(r)}u_i\right]$$

Determine $\overline{a}_i^{(r)}$, $\underline{a}_i^{(r)\sim}(i = 1, 2, \dot{s}, K)$ by the liner programing to minimize the fuzziness S defined by Eq. (10).

Step 2

(i) If $\underline{a}_i^{(r)}, {}^{\sim}\overline{a}_i^{(r)} \geq 0$, let

$$\left(\mathbf{A}_i^{(r+1)}\mathbf{U}_{\mathbf{L}_i}\right)^{h^0} = \left[\underline{a}_i^{(r+1)}\underline{u}_i,\ \overline{a}_i^{(r+1)}\overline{u}_i\right] \tag{17}$$

(ii) If $\underline{a}_i^{(r)}, {}^{\sim}\overline{a}_i^{(r)} \leq 0$, let

$$\left(\mathbf{A}_i^{(r+1)}\mathbf{U}_{\mathbf{L}_i}\right)^{h^0} = \left[\underline{a}_i^{(r+1)}\overline{u}_i,\ \overline{a}_i^{(r+1)}\underline{u}_i\right] \tag{18}$$

(iii) If $\underline{a}_i^{(r)} \leq 0 \leq \overline{a}_i^{(r)}$, let

$$\left(\mathbf{A}_i^{(r+1)}\mathbf{U}_{\mathbf{L}_i}\right)^{h^0} = \left[\underline{a}_i^{(r+1)}\overline{u}_i,\ \overline{a}_i^{(r+1)}\overline{u}_i\right] \tag{19}$$

Step 3

Determine $\overline{a}_i^{(r+1)}$, $\underline{a}_i^{(r+1)}(i = 1, 2, \ldots, K)$ by the liner programing to minimize the fuzziness S under the constraints (9) according to the judgement of $\left(\mathbf{A}_i^{(r+1)}\mathbf{U}_{\mathbf{L}_i}\right)^{h^0}$ in STEP2.

Step 4

If $\underline{a}_i^{(r)}\underline{a}_i^{(r+1)} \geq 0$ and $\overline{a}_i^{(r)}\overline{a}_i^{(r+1)} \geq 0 (i = 1, 2, \ldots, K)$ then go to STEP 6. Otherwise, let $r = r + 1$ and go to Step 5.

Step 5

If the trial count r is not beyond the maximum, then go to STEP 2. Otherwise, go to Step 6.

Step 6

Terminate the procedure.

4 An Illustrative Example

Let us illustrate our model by applying it to a example of expert's damage assessment of a structure. In real damage assessment of existing structures, we have to inspect many portions of the structure for the detection of various defects concerning with structural damage and non-structural damage. We shall also analyze collection of

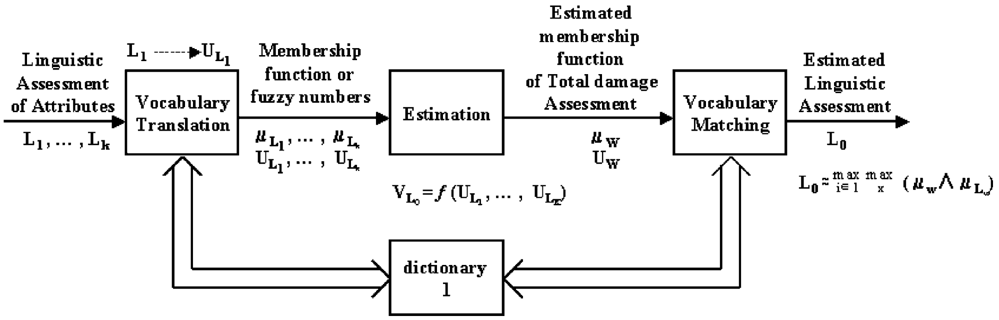


Fig. 3. Process of linguistic regression model

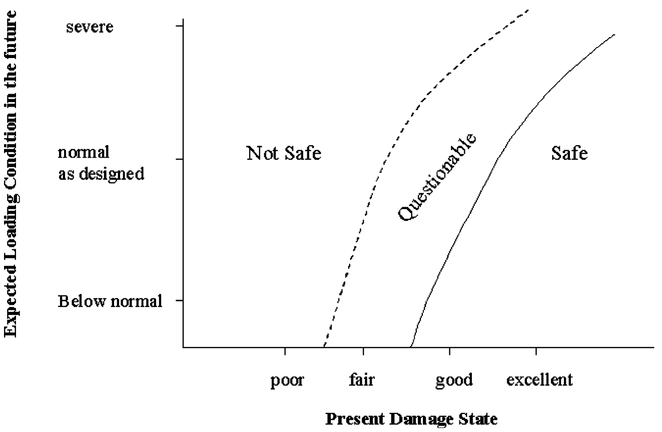


Fig. 4. Relation between the expected loading condition and the present damage state

instrumental records such as acceleration data. Sometimes historical records of the structure plays an important role in damage assessment. Moreover, its expected loading and environmental condition in future must be considered in forecasting possible failures of the structure. An overloaded structure is required to maintain greater strength than that required in normal usage. For instance, this relation between the present state of a structure and the expecting loading and environmental condition can be illustrated as in Fig. 4.

In this example, a simplified case is considered for the sake of the clarity in illustration. We will deal with only a small portion of a structure, say a beam. We are asked to model the expert's procedure of the total damage assessment of this portion of a structure. For this purpose, let us consider (1) cracking state, (2) corrosion state, and (3) expecting loading and environmental conditions to decide its total damage. Table 3 shows a dictionary of linguistic values and their fuzzy numbers employed in expressing the conditions of their attributes and the total damage, and Fig. 5 illustrates their fuzzy numbers.

Assume that sixteen training samples are given, which experts have evaluated their total damages and the conditions of their attributes in linguistic form as shown in Table 4. As illustrated in Table 5, these linguistic values given in Table 4 are translated into fuzzy numbers through the vocabulary translation unit in Figure 2 by using dictionary in Table 3. By applying the heuristic method to these fuzzy numbers, we can obtain the total assessment model on the basis of condition of attributes. Because it is essential to obtain high fitness of the expert model to the real cases, we have employed a value 0.8 for the fitness value. Table 5 illustrates the resulted model.

Let us consider sample 1 in Table 4. In the case of sample 1, the corrosion state is "extremely good" and, the loading & environment condition in the future is expected to be "below normal", but its cracking state is "extremely bad". On the basis of these information, experts assessed

Table 3. Dictionary

Name of attribute		Linguistic Value	Fuzzy Grade
Corrosion state	X_1	extremely good	(0.00,0.00,0.15)
		very good	(0.20,0.15,0.15)
		good	(0.40,0.15,0.15)
		bad	(0.60,0.15,0.15)
		very bad	(0.80,0.15,0.15)
		extremely bad	(1.00,0.15,0.15)
Cracking state	X_2	extremely good	(0.00,0.00,0.15)
		very good	(0.20,0.15,0.15)
		good	(0.40,0.15,0.15)
		bad	(0.60,0.15,0.15)
		very bad	(0.80,0.15,0.15)
		extremely bad	(1.00,0.15,0.15)
Expected loading environmental condition	X_3	below normal	(0.00,0.00,0.40)
		normal	(0.50,0.20,0.20)
		severe	(1.00,0.40,0.00)
Total assessment	Y	safe	(0.00,0.00,0.40)
		questionable	(0.50,0.20,0.20)
		not sage	(1.00,0.40,0.00)

its total damage state as "not safe" (see Table 4). Using the dictionary in our model as shown in Table 3, the words "extremely good" for corrosion state, "extremely bad" for cracking state and "below normal" for expected loading & environment condition are translated into $\mu_{(\text{corrosion})} = (0.00, 0.00, 0.15)$, $\mu_{(\text{cracking})} = (1.00, 0.15, 0.00)$ and $\mu_{(\text{loading})} = (0.00, 0.00, 0.40)$, respectively. By using the fuzzy assessment function

$$\begin{aligned}\mathbf{V} &= f(\mathbf{U}_{\mathbf{L}_1}, \mathbf{U}_{\mathbf{L}_2}, \mathbf{U}_{\mathbf{L}_3}) \\ &= (0.322, 0.000, 0.000)\mu_{(\text{corrosion})} \\ &\quad + (0.813, 0.000, 0.00)\mu_{(\text{cracking})} \\ &\quad + (0.373, 0.065, 0.65)\mu_{(\text{loading})}\end{aligned}$$

we can estimate its total damage as

$$\mathbf{V} = \mu_{(\text{total damage})} = (0.81, 0.12, 0.29).$$

Through the matching process in Fig. 3 we have

$$\max_t \{\mu_{(\text{total damage})} \bigwedge \mu_{(\text{safe})}(t)\} = 0.0,$$
$$\max_t \{\mu_{(\text{total damage})} \bigwedge \mu_{(\text{questionable})}(t)\} = 0.03, \text{ and}$$
$$\max_t \{\mu_{(\text{total damage})} \bigwedge \mu_{(\text{not safe})}(t)\} = 0.66.$$

Therefore, our model has assigned the same expression "not safe" to the total damage of this sample 1 as experts (see Table 8).

The estimated linguistic values can be obtained through the vocabulary matching unit with the dictionary as given in Table 7. Table 7 shows that our model can work well sufficiently. Let us check this model by new samples. Table 8 shows linguistic data of new samples and their estimated results of total assessment.

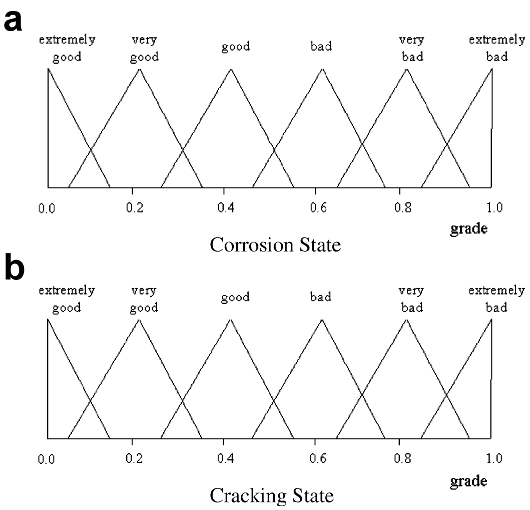


Fig. 5. Linguistic values in dictionary. **a** Corrosion state. **b** Cracking state

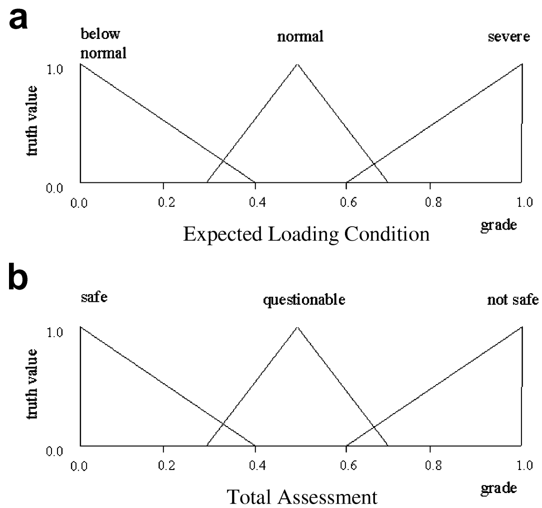


Fig. 6. Linguistic values in dictionary(continued). **a** Expected Loading Condition. **b** Total Assessment

Table 4. Linguistic data of training samples

Training Sample	Attribute			Total assessment ($\mathbf{Y}$)
	Corrosion state $(\mathbf{X})_1$	Cracking state $(\mathbf{X})_2$	Expected loading & environmental condition $(\mathbf{X})_3$	
1	extremely good	extremely bad	below normal	not safe
2	bad	very bad	below normal	not safe
3	very bad	very bad	below normal	not safe
4	very bad	very good	below normal	questionable
5	good	extremely good	below normal	safe
6	good	bad	normal	not safe
7	bad	very bad	normal	not safe
8	very bad	good	normal	not safe
9	bad	very good	normal	questionable
10	very bad	very good	normal	questionable
11	extremely good	bad	severe	not safe
12	very good	bad	severe	not safe
13	good	very bad	severe	not safe
14	very bad	good	severe	not safe
15	very good	good	severe	not safe
16	very good	very good	severe	questionable

Table 5. Fuzzy grades translated from linguistic data in Table 4

Training Sample	Attribute			Total assessment (**Y**)
	Corrosion state $(\mathbf{X})_1$	Cracking state $(\mathbf{X})_2$	Expected loading & environmental condition $(\mathbf{X})_3$	
1	(0.00,0.00,0.15)	(0.00,0.00,0.15)	(0.00,0.00,0.40)	(1.00,0.40,0.00)
2	(0.60,0.15,0.15)	(0.80,0.15,0.15)	(0.00,0.00,0.40)	(1.00,0.40,0.00)
3	(0.80,0.15,0.15)	(0.80,0.15,0.15)	(0.00,0.00,0.40)	(1.00,0.40,0.00)
4	(0.80,0.15,0.15)	(0.20,0.15,0.15)	(0.00,0.00,0.40)	(0.50,0.20,0.20)
5	(0.40,0.15,0.15)	(0.00,0.00,0.15)	(0.00,0.00,0.40)	(0.00,0.00,0.40)
6	(0.40,0.15,0.15)	(0.60,0.15,0.15)	(0.50,0.20,0.20)	(1.00,0.40,0.00)
7	(0.60,0.15,0.15)	(0.80,0.15,0.15)	(0.50,0.20,0.20)	(1.00,0.40,0.00)
8	(0.80,0.15,0.15)	(0.40,0.15,0.15)	(0.50,0.20,0.20)	(1.00,0.40,0.00)
9	(0.60,0.15,0.15)	(0.20,0.15,0.15)	(0.50,0.20,0.20)	(0.50,0.20,0.20)
10	(0.80,0.15,0.15)	(0.20,0.15,0.15)	(0.50,0.20,0.20)	(0.50,0.20,0.20)
11	(0.00,0.00,0.15)	(0.60,0.15,0.15)	(1.00,0.40,0.00)	(1.00,0.40,0.00)
12	(0.20,0.15,0.15)	(0.60,0.15,0.15)	(1.00,0.40,0.00)	(1.00,0.40,0.00)
13	(0.40,0.15,0.15)	(0.80,0.15,0.15)	(1.00,0.40,0.00)	(1.00,0.40,0.00)
14	(0.80,0.15,0.15)	(0.40,0.15,0.15)	(1.00,0.40,0.00)	(1.00,0.40,0.00)
15	(0.20,0.15,0.15)	(0.40,0.15,0.15)	(1.00,0.40,0.00)	(1.00,0.40,0.00)
16	(0.20,0.15,0.15)	(0.20,0.15,0.15)	(1.00,0.40,0.00)	(0.50,0.20,0.20)

Table 6. Coefficients in estimation unit

Fitting value = 0.70

Attribute		
Corrosion state $(\mathbf{X})_1$	Cracking state $(\mathbf{X})_2$	Expected loading & environmental condition $(\mathbf{X})_3$
(0.322,0.000,0.000)	(0.813,0.00,0.00)	(0.373,0.065,0.065)

Table 7. Linguistic data of training samples

Training sample	Value given by experts		Estimated value by the model	
	Linguistic value	Translated fuzzy grade	Estimated fuzzy grade	Matched word
1	not safe	(1.00,0.40,0.00)	(0.81,0.12,0.29)	not safe
2	not safe	(1.00,0.40,0.00)	(0.84,0.17,0.56)	not safe
3	not safe	(1.00,0.40,0.00)	(0.91,0.17,0.40)	not safe
4	questionable	(0.50,0.20,0.20)	(0.42,0.17,0.41)	questionable
5	safe	(0.00,0.00,0.40)	(0.13,0.05,0.41)	safe
6	not safe	(1.00,0.40,0.00)	(0.80,0.31,0.40)	not safe
7	not safe	(1.00,0.40,0.00)	(1.03,0.31,0.40)	not safe
8	not safe	(1.00,0.40,0.00)	(0.77,0.31,0.35)	not safe
9	questionable	(0.50,0.20,0.20)	(0.54,0.22,0.40)	questionable
10	questionable	(0.50,0.20,0.20)	(0.61,0.31,0.39)	questionable
11	not safe	(1.00,0.40,0.00)	(0.86,0.40,0.39)	not safe
12	not safe	(1.00,0.40,0.00)	(0.93,0.45,0.38)	not safe
13	not safe	(1.00,0.40,0.00)	(1.15,0.45,0.39)	not safe
14	not safe	(1.00,0.40,0.00)	(0.96,0.45,0.39)	not safe
15	not safe	(1.00,0.40,0.00)	(0.76,0.45,0.40)	not safe
16	questionable	(0.50,0.20,0.20)	(0.60,0.45,0.39)	questionable

Table 8. Linguistic data of new samples and their estimated total assessment

Training sample	Attribute			Estimated Total assessment (**Y**)
	Corrosion state $(\mathbf{X})_1$	Cracking state $(\mathbf{X})_2$	Expected Expected loading &.environmental condition $(\mathbf{X})_3$	
21	good	extremely bad	below normal	not safe
22	very good	very good	normal	not safe
23	bad	extremely good	below normal	questionable
24	very good	extremely good	below normal	safe
25	bad	very bad	normal	not safe

5 Concluding remarks

We have discussed the formulation of regression model based on natural words. In assessment, the role of experts is very important because existing structures are extremely complex and the accumulation of professional experience is often required. Experts frequently express their judgment in linguistic form rather than numerical form. Therefore, the linguistic treatment of assessment is central and essential for employing human inspective and subjective judgment into assessment process.

The presented process consists of four portions: (1) vocabulary translation, (2) estimation, (3) vocabulary matching and (4) dictionary. The emphasis of this paper should be placed on the use of natural words. The linguistic judgment of the total assessment is obtained through vocabulary matching on the basis of a dictionary as given by experts. We employed fuzzy quantification theory type 2 for estimating the total assessment in terms of linguistic structural attributes which are obtained from an expert.

References

1. **Dubois D, Prade H** (1980) Fuzzy Sets and systems: Theory and Applications. Academic Press, New York
2. **Tanaka H, Uejima S, Asai K** (1982) Linear Regression Analysis with Fuzzy Model. IEEE, Trans on System. Man Cybern. SMC-12 **6**: 903–907
3. **Tanaka H, Watada J, Asai K** (1983) Evaluation of Alternatives with Multiple Attribute based on Fuzzy Sets. Systems and Control The Japan Association of Automatic Control Engineers **27**(6): 403–409
4. **Watada J, Fu KS, Yao JTP** (1984) Damage Assessment Using Fuzzy Multivariant Analysis. Technical Report No. CE-STR-84-4, School of Civil Engineering, Purdue University, West Lafayette, IN, USA
5. **Watada J, Fu KS, Yao JTP** (1984) Fuzzy Classification Approach to Damage Assessment. First International Conference on Fuzzy Information Professing in Hawaii, USA
6. **Watada J, Tanaka H, Asai K** (1983) Fuzzy Quantification Theory Type I. The Japanese Journal of Behaviormetrics (The Behaviormetric Society of Japan) **11**(1): 66–73
7. **Watada J** (1994) Chapter 5 Applications in Business, 5.5: Multiattribute Decision-Making, Fuzzy System, by Terano T, Asai K, Sugeno M (Ed.) Applied AP Professional pp. 244–252
8. **Watada J** (2001) The Thought and Model of Linguistic Regression. In: Proceedings The 9th World Congress of International Fuzzy Systems Association, Canada
9. **Zadeh LA** (1965) Fuzzy Sets. Int. Control **8**: 338–353
10. **Zadeh LA** (1975) The concept of a Linguistic Variable and Its Applications to Approximate Reasoning. Part 1. Information Science **8**: 199–249
11. **Zadeh LA** (1975) The concept of a Linguistic Variable and Its Applications to Approximate Reasoning, Part 2. Information Science **8**: 301–357
12. **Zadeh LA** (1999) What is computing with words. Computing with Words in Information/Intelligent Systems. In: Physica-Verlag, Foundation, Zadeh LA, Janusz Kacpruzyk (Eds.) viii–ix
13. **Zadeh LA** (1999) Fuzzy Logic: = computing with words. Computing with Words in Information/Intelligent Systems. Zadeh LA, Janusz Kacpruzyk (Eds.) Physica-Verlag, Foundation, pp. 3–23