Statistical-Fuzzy Approach to Quantify Cumulative Impact of Change Orders

Awad S. Hanna, P.E., M.ASCE1; Wafik B. Lotfallah2; and Min-Jae Lee, S.M.ASCE3

Abstract: This paper presents a hybrid approach to quantify the impact of change orders on construction projects using statistical regression and fuzzy logic. There are many qualitative variables affecting the impact of change orders on labor productivity; statistical analysis falls short of addressing the fuzziness of those variables. Because of their complementary nature, fuzzy logic and regression analysis can be integrated; regression analysis is used to determine the membership functions of the input linguistic values. In this paper, each input variable is statistically treated before entering a general rule relating its space to the space of loss in labor productivity. The relative weight of each input variable is determined by its coefficient of determination ($R^2$) value. The expected loss of labor productivity and its standard deviation are then determined from the output fuzzy membership function. The proposed methodology is general and can be applied in areas of system analysis and decision making when a complex input-output function is to be predicted in the presence of some fuzzy knowledge and a large number of real input-output data.

DOI: 10.1061/(ASCE)0887-3801(2002)16:4(252)

CE Database keywords: Productivity; Fuzzy sets; Change order; Construction.

Introduction

A fact of life for a construction project is change. Changes result from the necessity to modify aspects of the construction project in reaction to circumstances that develop during the construction process. The changes may be small, well managed, and have little effect on the whole construction project. On the other hand, changes may be large, poorly managed, and have tremendous negative impacts on the construction project performance in terms of time and cost.

Problem

Change orders can frequently cause significant disruptions to a construction project, which may decrease the labor productivity of the contractor and extend the project duration. An earlier study showed that the United States construction industry spends $50 billion annually on new construction change orders (Ibbs and Allen 1995).

A recent audit report of state projects built in the state of Washington reviewed a total of 865 projects and found that 87% of the projects were completed with a combined total of 6,413 change orders of various sizes with an estimated value of $94 million. The audit report stated that one-third of the total number of change orders, or $35.4 million, could have been avoided. Inadequate field investigation, unclear specifications, plan error, and design change or mistakes by the consulting engineer were cited as causes for these changes (Cambridge Systematics, Inc. 1998).

Few studies have attempted to quantify the impact of change orders on project cost and schedule as well as labor productivity. Estimating the cumulative impact of change orders on project performance in terms of cost and schedule faces two major difficulties.

1. There are many input parameters that affect the loss of labor productivity; and
2. Many of these parameters are qualitative in nature and are hard to quantify—e.g., the quality of bid documents, the quality of contractors’ preplanning efforts, the contractors’ project management experience, and the quality of the engineering designs.

A literature review revealed that no research study has dealt with a statistical-fuzzy analysis of the qualitative variables affecting labor productivity in construction projects.

Limitations of Traditional Methods

A literature review revealed three academic methods that are generally used to predict the outcomes of engineering systems—namely, regression analysis, artificial neural networks, and fuzzy logic. However, when handling a problem like quantifying the impact of change orders on productivity, each of these methods has its own limitations.

1. Both regression analysis and neural networks have limitations in dealing with a qualitative input variable; and
2. As the system complexity increases, fuzzy logic faces the difficulty of determining the right set of rules and membership functions.
Hybrid Statistical-fuzzy Method

This paper presents a hybrid system that encapsulates both statistical analysis and fuzzy logic to study the effect of the quantitative and qualitative input variables on enhancing or reducing the impact of change orders on labor productivity. Regression analysis is used here to determine the membership functions of the input linguistic values as well as the forms of the if-then rules. Equivalently, each input variable is statistically treated before entering a general rule relating its space to the space of loss in labor productivity. Also, the relative weight of each input variable is determined by its coefficient of determination ($R^2$) value. When linguistic input values are determined, the general rules are separately fired and the resulting fuzzy values are aggregated with their relative weights. The expected loss of labor productivity as well as its standard deviation are then determined from the output fuzzy membership function.

Scope

This paper investigates the effects of change orders on the electrical and mechanical sectors of the construction industry. The reason for selecting electrical and mechanical contractors is that they represent a labor-intensive segment of the construction industry. Also, as subcontractors, these disciplines are typically the "last in line" and must carry the delays caused by previous trades. In addition, electrical and mechanical construction projects are complex, and small changes in plans and specifications can have large ripple effects on the rest of the project.

Background

After a contractor is awarded a construction project, an owner frequently finds it necessary to order changes on the project. Contract documents are an imperfect expression of the design professional’s and owner’s intent for a project. Circumstances develop during the construction process that make revisions of the drawings and specifications necessary. The design might prove to be inadequate. Materials specified might be unavailable or scarce. The owner’s budget or schedule might change and force a reduction of scope. Unforeseen natural events might occur. All of these reasons for change frequently cause disruption in the planned work schedule, and result in increased costs through rework and decreased efficiency of the base contract work. Some examples of causes of inefficiency due to change orders are (Fig. 1)

- Increased frequency of planning and replanning,
- Loss of efficiency due to interruption, interference, and lack of availability of tools, labor, and materials to meet the requirements of the changes,
- Increased project management and supervision involvement,
- Loss of efficiency due to a ripple impact that is a direct result of change orders, such as stacking of trades, schedule compression, and out-of-sequence work, and
- Difficulty in determining equitable adjustment compensation for the parties involved.

Labor Efficiency, Productivity, and Loss of Efficiency

Efficiency is the ratio of actual performance to the theoretical maximum performance, and therefore is dimensionless (Hanna et al. 1999a). Productivity can be defined either as the input (resources) divided by the output (completed work), or as the output divided by the input.

The variable that will be used to determine the loss of efficiency has been labeled “delta" (symbolically, $\Delta$). Delta is defined as the difference between the base project labor hours (actual total project hours less the estimate of change order hours) and the original estimate of labor hours at the contract award. Fig. 1 shows a graphical distribution of delta.

Delta can take a positive value or a negative value. Positive values of delta indicate that the actual productivity is less than the planned or estimated productivity. On the other hand, negative values of delta are an indication of higher efficiency than originally anticipated or estimated. Positive deltas can be attributed to a variety of factors such as a contractor’s low estimate, a contractor’s inefficiency, and the impact of productivity-related factors.
such as change orders, weather conditions, work interruptions, and rework, among others. To study the impact of change orders, the research team acquired data for projects where change orders were the main reason for loss of efficiency. Projects impacted by change orders were selected based on observing the deviation of actual manpower loading. Manpower loading is a graphical relationship between percent of time and work hours consumed per week. As a result, certain qualitative and quantitative criteria were developed to define projects that were impacted mainly by change orders (Hanna 1999a,b). Projects that were impacted by other factors or impacted because of the contractors’ low estimates were eliminated from the database.

**Change Orders Quantification: Literature Review**

Four studies are discussed below to provide an understanding of the current literature regarding the effects of change orders on project performance in terms of cost and time. In addition, the benefits and limitations of these studies are discussed.

Leonard et al. (1991) put forth a significant effort to quantify the effect of change orders on labor efficiency. The study used 90 cases that had resulted in disputes between owners and contractors. The data were collected for the following three different categories:

1. Electrical/mechanical contracts on building projects;
2. Electrical/mechanical contracts on industrial projects; and
3. Civil/architectural contracts on building and industrial projects.

The change order impacts were divided into three types—(1) minor; (2) medium; and (3) high. Graphs were presented in each impact category that related the loss of efficiency to the percentage of changes. There were, however, several deficiencies in this study:

- Limited number of variables. The study considered the amount of change as the only factor that was related to the loss of efficiency. The study did not consider other factors such as the timing of changes, the size of each change, or project-specific characteristics such as project size, labor type, and project delivery system, among others.
- Data adjustment. The study modified the loss of efficiency due to learning. This adjustment was not justified, because loss of productivity, as defined by Leonard, represents the cumulative impact that includes loss of efficiency due to learning.
- Combination of data. The study combined the data for electrical and mechanical trades without any evidence that the loss of efficiency between the two trades may be different.
- Biased sample. Data were collected from projects that reached a disagreement and dispute between parties. There were no opportunities to compare impacted and unimpacted projects.
- Classification of impact. The classification of impact as minor, medium, and high was subjective and based on the judgment of the researcher.

The Construction Industry Institute (CII) commissioned a study titled “Quantitative impacts of project change” (Ilbs and Allen 1995). In this study, a total of 89 projects were obtained from CII member companies. The data were gathered in order to examine the following three hypotheses:

1. Changes that occur late in a project are implemented less efficiently than those early on;
2. The greater the project change, the greater its negative impact on labor productivity; and
3. The hidden or unforeseeable costs of change (the cumulative change effect) increase with more project change.

This study showed several limitations, as follows:

1. In an attempt to relate the loss of productivity associated with change orders to several independent variables such as project percent complete and engineering percent complete, the study reported low values for the coefficient of determination $R^2$. Low $R^2$ values limit the ability of the owners and/or contractors to explain the variation present within the data.
2. The study assumed that the ratio between the installed material cost and the installed total cost is an indication of efficiency when late changes are implemented into a project. This can lead to difficulties when reductions of scope or changes occur that do not consume materials.
3. The study failed to support the concept that changes implemented late in a project are implemented less efficiently than changes that occur early in a project. Changes implemented late in a project cause a greater loss of labor efficiency because the peak of labor occurs 50–80% of the time.

There were two studies completed at the University of Wisconsin-Madison that used statistical methods to quantify the impact of change orders on labor productivity for mechanical and for electrical contracting (Hanna et al. 1999a, b). Both studies used questionnaires that were distributed to mechanical and electrical contractors, respectively. The data from the questionnaires were used in a regression analysis to determine a model to predict the delta as a percent of the total hours spent on the project.

The two studies identified four factors that impact the loss of labor efficiency. These were change order hours as a percent of original estimate hours, number of change orders, timing of change orders, and contractors’ project managers’ experience.

The mechanical and electrical studies at the University of Wisconsin have some shortcomings. Both studies looked at only a limited number of qualitative variables. The present research proposes to investigate over 70 possible factors that impact how change orders might affect productivity.

**Fuzzy Logic Literature in Construction**

Zadeh (1965, 1975a,b,c, 1979) outlined the theory of fuzzy sets for incorporating vague and imprecise data into analyses. Ayub and Haldar (1984) pioneered the use of fuzzy set theory to evaluate the impact of qualitative variables such as site conditions, weather conditions, and labor experience on activity cost and duration. They emphasized primarily the assessment of the mean, variance, and covariance of the activity duration.

Wu and Hadipriano (1994) introduced the fuzzy modus ponens deduction technique for construction scheduling to assess the impacts of qualitative factors on activity duration. They quantified linguistic values into numerical measures using angular fuzzy set theory. These numerical values are used to modify the activity duration affected by the cumulative impact of different site, climatic, resource, and management factors. Others have used the theory to provide start and finish times along with fuzzy project durations (Dubois and Prade 1980).

Lorterapong and Moselhi (1996) calculated activity duration using traditional fuzzy set operations as well as newly developed fuzzy network scheduling (FNET). The results generated by FNET are reasonable, but the computations are not as simple as for the program evaluation and review technique.

Recently, Hanna and Lotfallah (1999) used fuzzy logic to select the best crane type for a construction project. In their approach, they used the max-min extension principle to transform
linguistic information about the suitability of each crane type with respect to each factor of the project into an overall efficiency of each crane type.

**Fuzzy Set Concepts (Statistical-fuzzy If-then Rules)**

The novel approach in this paper is the use of statistical data to choose the fuzzy membership functions and to form the fuzzy if-then rules. The following example illustrates this method.

Suppose that a fuzzy if-then rule is to be formed between the universe of an input variable \( x \), and an output variable \( y \). Assume further that \( x \) lives in the unit interval \([0, 1]\) and \( y \) lives in the interval \([-1, 1]\). For the output variable \( y \), we fix the linguistic values \( P \) (for positive), \( N \) (for negative), and \( Z \) (for zero), whose fuzzy membership functions are given by

\[
P(y)=\begin{cases} \frac{y}{2} & \text{for } 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}
\]

\[
N(y)=\begin{cases} \frac{-y}{2} & \text{for } -1 \leq y \leq 0 \\ 0 & \text{otherwise} \end{cases}
\]

\[
Z(y)=\begin{cases} \frac{1-y}{2} & \text{for } 0 \leq y \leq 1 \\ \frac{1+y}{2} & \text{for } -1 \leq y < 0 \end{cases}
\]

We also have some statistical data represented by the points \((x_i, y_i)\), for \( i = 1, ..., n \), from which we get the regression line

\[
y = f(x) = ax + b
\]

where the values of \( a \) and \( b \) are picked to minimize the sum of the squares of the errors

\[
E = \sum_{i=1}^{n} (y_i - ax_i - b)^2
\]

In traditional regression analysis \( x \) and \( y \) are correlated and \( a \neq 0 \). We then prefer to write Eq. (1) in the form

\[
y = f(x) = a(x - c)
\]

Let us assume for now that \( a > 0 \). We then seek some if-then rules of the form

If \( x \) is \( H \), then \( y \) is \( P \); If \( x \) is \( M \), then \( y \) is \( Z \);

If \( x \) is \( L \), then \( y \) is \( N \)

where the values \( H \) (for high), \( M \) (for medium), and \( L \) (for low) are yet to be determined.

According to the above rules, it is plausible to say that our belief-strength that some particular value of \( x \) is high can be identified with our belief-strength that the corresponding predicted value of \( y \) is positive. This suggests that we can define the fuzzy membership function \( H \) by

\[
H(x) = P(f(x))
\]

and similarly

\[
M(x) = Z(f(x)) \quad \text{and} \quad L(x) = N(f(x))
\]

Thus, in our example we have

\[
H(x) = \begin{cases} a(x-c) & \text{for } x \geq c \\ 0 & \text{otherwise} \end{cases}
\]

\[
L(x) = \begin{cases} a(c-x) & \text{for } x \leq c \\ 0 & \text{otherwise} \end{cases}\]

The case when \( a < 0 \) is similarly treated, where the if-then rules will take the form

If \( x \) is \( H \), then \( y \) is \( N \); If \( x \) is \( M \), then \( y \) is \( Z \);

If \( x \) is \( L \), then \( y \) is \( P \)

The values of \( H, M, \) and \( L \) are determined in a similar way.

Our method can be applied even if the regression function \( y = f(x) \) is not a linear function. Another way to view this method is that we used the function \( f(x) \) to map the input variable \( x \) to an intermediate variable \( u = f(x) \) such that the relationship between \( u \) and the output variable \( y \) can be represented by an identity fuzzy functional; i.e., we have

If \( u \) is \( N \), then \( y \) is \( N \); If \( u \) is \( Z \), then \( y \) is \( Z \);

If \( u \) is \( P \), then \( y \) is \( P \)

The above functional can be generalized to the following statement:

If \( u \) is \( A \), then \( y \) is \( A \)

where \( A = \) any fuzzy set.

Thus, by statistically treating the input \( x \), we managed to simplify the fuzzy rules used to the general rule

If \( f(x) \) is \( A \), then \( y \) is \( A \)

**Fuzzy Set Concepts (Aggregating Rules)**

Starting with \( k \) input variables \( x_1, x_2, ..., x_k \) affecting an output variable \( y \), we use the method of the previous section to get a separate regression function \( u_i = f_i(x_i) \) for each input variable \( x_i \) with the generalized fuzzy rule

\[ R_i \quad \text{If } f_i(x_i) \text{ is } A_i \]

Now suppose that for each input variable \( x_i \) we have the fuzzy statement

\[ f_i(x_i) \text{ is } A_i \]

Since our belief in rule \( R_i \), can be measured by the coefficient of determination \( R_i^2 \) of the correlation between \( x_i \) and \( y \), we can deduce that our belief that “\( y \) is \( A_i \)” is also measured by \( R_i^2 \). We can then aggregate the fuzzy membership \( A_i \) by the formula

\[
A(y) = \frac{\sum_{i=1}^{k} R_i^2 \cdot A_i(y)}{\sum_{i=1}^{k} R_i^2}
\]

Calling

\[
w_i = \frac{R_i^2}{\sum_{i=1}^{k} R_i^2}
\]

Eq. (2) can be written in the form

\[
A(y) = \sum_{i=1}^{k} w_i A_i(y)
\]

where \( A \) = weighted average of the \( A_i \)s. Now to defuzzify the output \( y \), we find the centroid \( y^* \) from

\[
y^* = \frac{\int_{A(y)} y \cdot A(y) dy}{\int_{A(y)} A(y) dy}
\]
representing the expected value of the output $y$, and the standard deviation $\sigma$ with 

$$\sigma^2 = \frac{\int (y - \mu)^2 A(y) \, dy}{\int A(y) \, dy}$$

representing the accuracy of our estimation. In practice we may want to use discrete spaces for the input and the output variables, and replace integration by summation.

**Application: Impact of Change Orders on Productivity**

Using regression analysis, Hanna et al. (1999a, b) developed a linear model that quantifies the impact of change orders on labor productivity. In their model, the output dependent variable was the percent loss of labor productivity (Delta). They pointed out some significant independent variables affecting Delta. Table 1 gives the definition and the range of possible values for each independent variable. The regression model found was

$$\% \text{Delta} = 0.3495 + 0.139 \text{ Industrial} - 0.0984 \text{ EA}_P - 0.0368 \% \text{OwnInitiatedCO} - 0.190 \% \text{OwnInitiatedCO} \times \text{Industrial} - 0.100 \% \text{COHrsApproved} + 0.0627 \% \text{Additions} - 0.0593 \text{ Productivity-Track} + 0.0544 \text{ Absenteeism}$$

(3)

From their statistical findings, we get a separate regression linear function for Delta against each input variable. Table 2 shows those functions together with their $R^2$ value. Note that since the $R^2$ value of $\% \text{Additions}$ is 0.000, this variable ($x_5$) will not appear in our statistical-fuzzy model.

**Case Study**

To illustrate our method, we show the calculations for the following case study, where each input variable takes either a given crisp value or a fuzzy value determined by the available information:

1. Not an industrial project ($x_1 = 0$).
2. Estimated/actual peak work hours is 0.71 ($x_2 = 0.71$).
3. Change orders initiated by the owner is 90% ($x_3 = 0.9$).
4. Change orders hours approved by the owner is 81.3% ($x_4 = 0.813$).
5. Additions or deletions of change order hours is 75% ($x_5 = 0.75$).
6. The contractor’s productivity tracking system is not sophisticated, but adequate. In this case, we assume that $x_6$ has the following fuzzy value: $x_6 = [0.0, 0.2, 0.4, 0.6, 0.8, 1.0, 0.8, 0.6, 0.4, 0.2, 0.0]$.

Next we determine the fuzzy sets $A_i$ representing $f_i(x_i)$. As Delta ranges between $+50$ and $-50\%$, we take its discrete universe to be 

$$D = \{-0.5, -0.4, -0.3, -0.2, -0.1, 0.0, 0.1, 0.2, 0.3, 0.4, 0.5\}.$$

Since the input variables $x_1, \ldots, x_5$ take crisp values, their regression outputs $f_1(x_1), \ldots, f_5(x_5)$ take the following crisp values:

$$A_1 = f_1(0) = 0.077 \approx 0.1$$
$$A_2 = f_2(0.71) = 0.115 \approx 0.1$$
$$A_3 = f_3(0.9) = 0.035 \approx 0.0$$
$$A_4 = f_4(0.813) = 0.080 \approx 0.1$$
$$A_5 = f_5(0.75) = 0.0912 \approx 0.1$$

**Table 1. Significant Factors Affecting Delta**

<table>
<thead>
<tr>
<th>Input variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial</td>
<td>1 = Industrial projects; 0 = all other projects (commercial, institutional, etc.)</td>
</tr>
<tr>
<td>EA_P</td>
<td>Estimated/actual peak human-hours (0.15–3.25, from database)</td>
</tr>
<tr>
<td>%OwnInitiatedCO</td>
<td>Percent of change orders initiated by owner (0.00–1.00)</td>
</tr>
<tr>
<td>%COHrsApproved</td>
<td>Percent of change order hours approved by owner (0.00–1.00)</td>
</tr>
<tr>
<td>%Additions</td>
<td>Percent of change order hours that were additions or deletions (0.00–1.00)</td>
</tr>
<tr>
<td>Productivity-Track</td>
<td>Did the contractor track productivity for the project? (0 = No; 1 = Yes)</td>
</tr>
<tr>
<td>Absenteeism</td>
<td>Absenteeism ratio (1 = 0–5%; 2 = 6–10%; 3 = 11–20%; 4 = greater than 20%)</td>
</tr>
</tbody>
</table>

**Table 2. Regression Models with $x_i =$ Input Variable and $u_i =$ Delta Estimator**

<table>
<thead>
<tr>
<th>Input variable $x_i$</th>
<th>Regression function $u_i = f_i(x_i)$</th>
<th>$R^2$ value</th>
<th>Weight $w_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial ($x_1$)</td>
<td>$u_1 = 0.0313 \cdot x_1 + 0.077$</td>
<td>0.005</td>
<td>0.011</td>
</tr>
<tr>
<td>EA_P ($x_2$)</td>
<td>$u_2 = -0.122 \cdot x_2 + 0.202$</td>
<td>0.128</td>
<td>0.268</td>
</tr>
<tr>
<td>%OwnInitiatedCO ($x_3$)</td>
<td>$u_3 = -0.199 \cdot x_3 + 0.214$</td>
<td>0.097</td>
<td>0.203</td>
</tr>
<tr>
<td>%COHrsApproved ($x_4$)</td>
<td>$u_4 = -0.301 \cdot x_4 + 0.325$</td>
<td>0.073</td>
<td>0.153</td>
</tr>
<tr>
<td>%Additions ($x_5$)</td>
<td>$u_5 = -0.000105 \cdot x_5 + 0.0913$</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Productivity-Track ($x_6$)</td>
<td>$u_6 = -0.0765 \cdot x_6 + 0.132$</td>
<td>0.033</td>
<td>0.069</td>
</tr>
<tr>
<td>Absenteeism ($x_7$)</td>
<td>$u_7 = 0.121 \cdot x_7 - 0.082$</td>
<td>0.141</td>
<td>0.296</td>
</tr>
</tbody>
</table>
Table 3. Validation of Models for % Delta Prediction

<table>
<thead>
<tr>
<th>Case project</th>
<th>Industrial (x_1)</th>
<th>EA_P Initiated CO (x_2)</th>
<th>%Own Approved (x_3)</th>
<th>%COHrs Project (x_4)</th>
<th>%Additions (x_5)</th>
<th>Productivity-Track (x_6)</th>
<th>Absenteeism (x_7)</th>
<th>Project actual</th>
<th>Statistical model</th>
<th>Statistical fuzzy model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>No (0)</td>
<td>0.71</td>
<td>0.9</td>
<td>0.813</td>
<td>0.75</td>
<td>Not sophisticated, but adequate (1)</td>
<td>0–5% with strength 70%; 6–10% with strength 30%</td>
<td>10.80</td>
<td>23.70</td>
<td>7.20</td>
</tr>
<tr>
<td>E131</td>
<td>No (0)</td>
<td>0.82</td>
<td>1</td>
<td>0.94</td>
<td>0.65</td>
<td>Yes (1)</td>
<td>0–5% with strength 70%; 6–10% with strength 30%</td>
<td>10.61</td>
<td>17.39</td>
<td>8.51</td>
</tr>
<tr>
<td>E262</td>
<td>Yes (1)</td>
<td>0.28</td>
<td>0.95</td>
<td>0.5</td>
<td>0.65</td>
<td>Yes (1)</td>
<td>Greater than 20% with strength 70%; 12–20% with strength 30%</td>
<td>10</td>
<td>39.45</td>
<td>25.31</td>
</tr>
</tbody>
</table>

Average %Error

| %Error | —— | —— | —— | —— | —— | —— | 0.00 | 129.82 | 43.70 |

Note: Average %Error = |X_{actual} - X_{estimated}|/X_{estimated} x 100.

Since x_6 takes a fuzzy value, the outputs f_6(x_6) take the following fuzzy value:

\[ A_6 = f_6(\{0.0 + 0.2/0.1 + 0.4/0.2 + 0.6/0.3 + 0.8/0.4 + 1.0/0.5 + 0.8/0.6 + 0.6/0.7 + 0.4/0.8 + 0.2/0.9 + 0/1.0\}) \]

To get that value, we applied f_6 of Table 2 on the elements of the universe (denominators), rounded up the values to two decimal digits, and added the membership values (numerator) corresponding to the same denominator.

If we further try to represent A_6 as a subset of D, we have to round up the denominators even more to one decimal digit. However, as the regression function f_6 has a very small slope, A_6 gets approximated by the crisp number

\[ A_6 = 0.1 \]

indicating that our discrete model cannot capture the fuzziness of the variable x_6. Also, x_7 takes a fuzzy value and the corresponding f_7(x_7) takes the fuzzy value

\[ A_7 = f_7(\{0.7/1 + 0.3/2 + 0.0/3 + 0.0/4\}) = [0.7/0.039 + 0.3/0.160 + 0.0/0.281 + 0.0/0.402] = [0.7/0.0 + 0.3/0.2 + 0.0/0.3 + 0.0/0.4] \]

Thus, as a subset of D

\[ A_7 = [0.0, 0.0, 0.0, 0.7, 0.5, 0.3, 0.0, 0.0] \]

where we added the entry 0.5 = (0.7 + 0.3)/2 to maintain the convexity of A_7.

Now we use \( A(y) = \sum_{i=1}^{k} w_i A_i(y) \) to find the output fuzzy set A representing the relative loss in labor productivity, as follows:

\[ A = 0.011[0.0, 0.0, 0.0, 0.0, 0.1, 0.0, 0.0, 0.0] + 0.268[0.0, 0.0, 0.0, 0.1, 0.0, 0.0, 0.0, 0.0] + 0.203[0.0, 0.0, 0.0, 0.1, 0.0, 0.0, 0.0, 0.0] + 0.153[0.0, 0.0, 0.0, 0.1, 0.0, 0.0, 0.0, 0.0] + 0.000[0.0, 0.0, 0.0, 0.1, 0.0, 0.0, 0.0, 0.0] + 0.069[0.0, 0.0, 0.0, 0.1, 0.0, 0.0, 0.0, 0.0] + 0.296[0.0, 0.0, 0.0, 0.7, 0.5, 0.3, 0.0, 0.0] \]

Thus, the expected loss in labor productivity (centroid) is

\[ y^* = \frac{\sum y \cdot A(y)}{\sum A(y)} = \frac{0.410 + 0.1 \cdot 0.649 + 0.2 \cdot 0.089}{0.410 + 0.649 + 0.089} = 0.0827 \]

\[ y^* = 0.072 \approx 7.2\% \]
with variance

\[ \sigma^2 = \frac{\sum(y - y^*)^2 \cdot A(y)}{\sum A(y)} = \frac{(0 - 0.072)^2 \cdot 0.410 + (0.1 - 0.072)^2 \cdot 0.649 + (0.2 - 0.072)^2 \cdot 0.089}{0.410 + 0.649 + 0.089} = \frac{0.0041}{1.148} = 0.0036 \]

Thus, \( \sigma = 0.060 \approx 6\% \).

A comparison between the statistical-fuzzy model estimate and the pure statistical model estimate is shown in Table 3. It provides the actual \%Delta, predicted \%Delta, and the analysis of the average percent error for seven case studies (including this example). The statistical-fuzzy approach shows significant improvement in the prediction accuracy of productivity loss.

**Conclusion**

This paper provides a new methodology using the statistical-fuzzy approach to quantify the cumulative impact of change orders. The proposed methodology uses statistical results in forming fuzzy if-then rules as well as choosing membership functions of linguistic values.

The new methodology provides substantial improvement compared to traditional statistical models. The new model improves the prediction accuracy and is capable of integrating fuzzy knowledge into quantitative data.

The proposed methodology can also be applied in areas of system analysis and decision making when a complex input-output function is to be predicted in the presence of fuzzy knowledge and fuzzy variables. Computer implementation of this methodology is suggested for future research, so stakeholders can easily implement this methodology for their purposes.

**References**


