PREDICTING DELAY PERCENTAGE OF WATER AND SEWAGE PROJECTS IN EGYPT: REGRESSION ANALYSIS VERSUS STATISTICAL-FUZZY APPROACH

التنبؤ بنسبة التأخير في مشروعات المياه والصرف الصحي بمصر: مقارنة بين تحليل الإنحدار والطريقة الإحصائية للمنطق الضبابي

A. M. EL-KHOLY[1]

الملخص العربي

هذا البحث يقدم نموذجين للتنبؤ بنسبة التأخير في مشروعات المياه والصرف الصحي. النموذج الأول يعتمد على تحليل الانحدار. تم تجميع ٧٤ سبب من الأسباب التي تؤدى لتأخير تلك المشروعات وذلك من الدراسات السابقة. تم إجراء استبيان تم توزيعه على مقاولي التشييد لمشروعات المياه والصرف الصحي بمصر وذلك لتقييم الأهمية النسبية لتلك الأسباب. وقد تم معرفة أكثر تلك الأسباب تأثير على نسبة التأخير وهم ١٤ سبب وهي تمثل المتغيرات المستقلة في النموذج المقترح. تم تجميع بيانات تخص حدوث أو عدم حدوث الأسباب التي تؤدي لتأخير مشروعات المياه والصرف الصحي ونسبة التأخير المناظرة التي حدثت (المتغير التابع) لتلك المشروعات وذلك لعدد ٢٠ مشروع. تم تقسيم تلك المشروعات لجزئين: الجزء الأول يشمل ١٢ مشروع لبناء النموذج وقد أظهرت النتائج وجود علاقة خطيه قويه بين المتغير التابع و ٩ أسباب من ١٤ سبب السابقة. الجزء الثاني يشمل ٨ مشروعات وذلك لاختبار هذا النموذج ومقارنته بالنموذج الثاني. النموذج الثاني يعتمد على استخدام الطريقه الإحصائية لنظرية المنطق الضبابي وهذه الطريقة تتكون من استخدام نظرية المنطق الضبابي وتحليل الانحدار. حيث تم عمل معادلة انحدار لكل سبب منفرد مع المتغير التابع باستخدام (مشروعات الجزء الأول). وتم تحديد الوزن النسبي لكل سبب بواسطة معامل التحديد (R^2). وتم تمثيل تأثير السبب في صورة المنطق الضبابي بواسطة تقييم درجة الشده المناظره لتأثير السبب وذلك من نتائج الأستبيان. أيضا تم تمثيل نسبة التأخيرات بصفة عامة لمشروعات المياه والصرف الصحي في صورة (Trapezoidal membership function) باستخدام نظرية المنطق الضبابي اعتمادا على نسبة التأخير في ١٨ مشروع من ٢٠ مشروع السابقين و قد تم استبعاد مشروعين من تمثيل (Trapezoidal membership function) و ذلك لتباعد قيمهم عن باقي قيم المشروعات. من ذلك تم التنبؤ بنسبة التأخير المتوقعه باستخدام قوانين المنطق الضبابى. تم اختبار كلا من النموذجين باستخدام الجزء الثاني من المشروعات ، وقد أظهرت النتائج أن النموذج المعتمد على الأنحدار يعطى نتائج أدق في التنبؤ عن النموذج المعتمد على الطريقه الإحصائية لنظرية المنطق الضبابي.

ABSTRACT

This paper presents two models for predicting the delay percentage in water and sewage projects in Egypt. The first model based on regression analysis. 74 causes that lead to delay in water and sewage projects gathered from literature. A questionnaire survey was made on construction contractors of water and sewage projects in Egypt to evaluate the relative importance of these causes. 14 causes were obtained as the most significant causes that affect the delay percentage (DP) and these are the independent variables of the proposed model. Data for the occurrence of the previous causes on a yes/no basis and the corresponding DP (dependent variable) for 20 water and sewage projects was collected. The data was divided into two sets, the first set contains 12 projects for the purpose of model building. The results revealed that there was a strong linear relationship between DP and 9 causes from 14 causes that significantly affect DP of projects. The second set contains 8 projects for the validation purposes and comparison with the second model. The second model is a statistical fuzzy approach which is a hybrid approach from fuzzy logic and regression analysis. A regression equation between each cause and DP using projects of first set was extracted. The relative weight of each cause is determined by its coefficient of determination (R^2) value. The degree of severity each cause had received from questionnaire analysis was used to fuzzify this cause. A trapezoidal membership function

[1] Assistant Professor, Civil Eng. Dept., Delta Higher Institute of Eng. and Tech., Mansoura, Egypt

was used to represent the delay percentages in water and sewage projects in general depending on 18 out of the previous 20 projects. Two projects were excluded from this function due to their divergence values from other projects. Thus, the expected delay percentage of a project is then determined using fuzzy rules. Validation of the two models using projects of the second set revealed that regression model has prediction capabilities higher than that of statistical fuzzy model.

Keywords: Regression Analysis; Questionnaire Survey; Statistical Fuzzy Model
           Water and Sewage Projects.

1. INTRODUCTION

Although nowadays there are great efforts for accomplishing projects on time by people in charge, projects are accomplished with delays and costs that are higher than estimated (Ghoddosi, et al.) [1]. Delay is the time between pre-planned time and actual time of project's activities (Arditi and Robinson) [2]. Cost and time overruns are sometimes higher than the value of contract (more than 100%) (Peter and Hough , Morris, Flyvbjerg et al.) [3-5]. Delay in the completion of projects is often unavoidable circumstance. Considerable studies were performed about the delays in projects and the causes, it could be classified into two parts (Ghoddosi, et al. 2008)[1]. In the first part, delays and methods represented to analyse them. In this part, methods for analyzing delays included methods of comparing planned time schedule and actual time, increase in base time schedule and analysing time intervals delay (Cher; Michael; Terry) [6-8]. On the other hand, project delays were divided according to three criteria: (1) the party caused delay: owner, contractor, or third party; (2) occurrence time state: independent delays, serial delays, and concurrent delays; (3) compensability: impermissible delays and permissible (Kartam; Stumpf) [9,10]. The second part presents the precedent studies about

causes of delay. In these studies, time and cost overrun were classified into 2 categories: strategic decisions, and operational causes. Strategic decisions are made by high rank manager of organization (e.g., selecting the project delivery system) before concluding the contract. Operational causes are produced during the execution of the project (e.g., lack of material) (Peter and Hough) [3]. In addition to the above studies, other studies were carried out about the causes of project delays, some of which is as follows: Causes of delays in large building construction projects (Sadi et al.) [11], delay in public utility projects in Saudi Arabia (Al- khalil et al.)[12], construction delay: a quantitative analysis (Al-Moumani) [13], causes of construction delay: traditional contracts (Abdallah and Battaineh)[14], construction delays in Florida: an empirical study (Ahmed et al.)[15], significant factors causing delay and cost overruns in construction of ground-water projects in Ghana (Frimpong et al.)[16], factors affecting construction speed of industrialized building systems in Malaysia (Alaghbari) [17], fuzzy assessment of causes of time overrun (delays) in Iran's dam construction projects (Ghoddosi, et al.) [1], and causes of delay in construction industry in Libya (Tumi et al.)[18].

Construction projects, especially public water and sewage projects face high uncertainty. This makes problems that could obstruct the progress and cause delay in this type of the project (Ghafly)[19]. Ghafly [19], performed a preliminary survey of 76 public water and sewage projects to identify the frequency and extent of delay. These projects selected from different cities in the eastern province in Saudi Arabia completed by different contractors, and had varying contract values. He was found that about 60% (45 out of 76) were delayed. Also, he found that in a sample of other 20 projects selected randomly, the average extent of delay in these projects was about 110% of the original project duration. Al Sultan [20], surveyed actual time performance of different types of public projects. He concluded that approximately 70% of the public projects (101 out of 145) experienced time overrun.

Accordingly, it seems that delay is experienced in public water and sewage projects. Both contractors and public authorities, who suffer from delays in these projects, need to organize the causes of delay, their frequency, and severity, to take necessary precautions to control those causes. Thus, in this paper, two models are developed to predict the percentage delay to the planned project duration in public water and sewage projects in Egypt. The first model is based on regression analysis, while the other is based on fuzzy statistical approach. The independent variables, which are causes of delay were gathered from literature. These causes are candidate according to their significance through a questionnaire survey.

The paper is organized as follows: first, fuzzy logic literature in construction is highlighted. Causes of delay in the construction of public water and sewage projects based on literature are identified, and a questionnaire survey is then prepared. The survey response is then analysed and discussed. A regression model is then established. Statistical fuzzy approach is explained. Statistical fuzzy model is then developed. To illustrate statistical fuzzy model calculations, a case study is presented. Finally, validation of the two models is presented and discussed.

2. FUZZY LOGIC LITERATURE IN CONSTRUCTION

A fuzzy set approach pioneered by Zadeh [21] is useful for uncertainty analysis where a probabilistic data base is not available and or when interval values of input variable are uncertain. Zadah [21-25] outlined the theory of fuzzy sets for incorporating vague and imprecise data into analyses. Ayyub and Haldar [26] pioneered the use of fuzzy set theory to evaluate the impact of qualitative variables such as site conditions, weather conditions on activity cost and duration. The fuzzy set approach has been widely applied to represent the uncertainties of real life situations (Bogardi and Badossy; Anandalingam and westfall)[27-28]. Loterapong and Moselhi [29] calculated activity duration using traditional fuzzy set operations as well as newly developed fuzzy network scheduling (FNET). The results generated by FNET were reasonable but the computations were not as simple as for the program evaluation and review technique. Hanna and Lotfallah [30] used fuzzy logic to select the best crane type for a construction project. In their approach, they used the

max-min extension principal to transform linguistic information about the suitability of each crane type with respect to each factor of the project into overall efficiency of each crane type.

Hanna et al. [31] presented a hybrid approach to quantify the impact of change orders on construction projects using statistical regression and fuzzy logic. In their approach, fuzzy logic and regression analysis were integrated, where regression analysis was used to determine the membership functions of the input linguistic values. 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 determined from the output fuzzy membership function. This approach will be applied and adopted for developing the second model to predict the percentage delay in public water and sewage projects. Details of this approach will be given later.

3. CAUSES OF DELAY IN PUBLIC WATER AND SEWAGE PROJECTS

In this study, 74 reasons identified as causes of time overrun in public water and sewage projects, which were gathered from literature [15,16,18,19] and shown in Table 1. These causes will serve as the independent variables in the predictive model of delay percentage.

4. QUESTIONNAIRE SURVEY

A questionnaire was developed to collect data about the significance of causes of delay affect water and sewage projects compiled in Table 1. It was organized in the form of two priority scaling, one for occurrence frequency, while the other for severity scaling. The priority scaling for occurrence frequency was as follows: 5=Always, 4=often, 3=usually, 2= sometimes, and 1=scarcely, while the severity scaling was: 5=very severe, 4=severe, 3=somewhat severe, 2=little effect, 1=very little effect. The participants were asked to assign a number from 1 to 5 to each cause for both occurrence frequency and severity according to its significance. Besides, the questionnaire included collection of data for past water and sewage projects in a structured format. The data included occurrence of previous causes of delay presented in Table 1 on a yes/no basis.

The survey gathered data from contracting companies specialized in water and sewage projects. Some of them were involved in other construction works such as: administrative, commercial and residential buildings. Thirty-three contracting companies participated in the survey. Some of the questionnaires were sent via mail after contacting the participants through telephones, whereas, the other part was sent by some persons. Most of the participants were at the level of general manager or project manager.

5. SURVEY RESULTS AND ANALYSIS

As a result of mailing and follow up a total of twenty-three usable questionnaires were completed and returned. All the questionnaires were combined for the analysis. The respondents included general managers, technical office managers, and construction managers. All the participants were involved in public water and sewage projects, whereas 48% of them were involved in administrative &

commercial and residential buildings, in addition to public water and sewage projects. The author believes that the variations in positions besides the variations in the specialization for the participants will enrich this field study to a great extent. To give additional credibility for the findings of this survey, the participants were asked about their length of experience. Also, they were asked about their companies' length of experience and annual volume of work. 95% of respondents have an experience more than 10 years, whereas, 74 % have an experience more than 20 years. All of the companies have an experience more than 10 years, whereas 78% have an experience equal to or greater than 25 years. 91% of companies have an annual volume of work more than LE 10

millions, whereas 74% have an annual volume of work equal to or greater than LE 250 millions. The author believes that obtaining the needed information from such experienced and active contractors is one of the strengths of this survey.

An importance index (II) was established to assess the degree of significance for each cause in each returned questionnaire as given in Eq. 1. Then, the importance indices were used to measure the relative weight for each cause. The relative importance weight (RIW) was computed using Eq.2. Table 1 shows the causes rearranged in descending order according to their corresponding RIW. Then, the causes are ranked according to their RIW such that the cause received the highest RIW is assigned rank equal to 1.

$$\text{Importance Index (II)} = \text{Occurrence frequency} * \text{degree of severity} \tag{1}$$

$$\text{Relative Importance Weight (RIW)} = \frac{\Sigma \;\; \Pi * \text{corresponding no. of respondents}}{\text{Total no. of respodents}} \tag{2}$$

Table 1 shows that, changes in material prices comes out as the most important factor attributing to the causes of delay in water and sewage projects, it was ranked the first. Fluctuation in materials prices would make contractors spend some time waiting to obtain materials with cheaper price. Delay in progress payments by the owner received the second rank, due to failure of government departments to provide adequate funding resources to contractors. It seems that this cause affect contractors' cash flow, which lead to delay in purchasing material required to the project. The third significant cause identified was difficulties in obtaining work permits from the authorities concerned (see Table 1), due to

bureaucracy in government departments. The fourth item in Table 1, which was given by respondents as a cause of delay for water and sewage projects is original contract duration too short. Delay in issuance of change orders by the owner was ranked fifth. Change orders can frequently cause significant disruptions to a construction project, which may decreases the labor productivity of the contractor and extend the project duration. Change orders coupled with late issuance have tremendous negative impacts on project performance in terms of time. Inflation was ranked 6, since the trend of inflation is probably due to demand exceeding supply, which creates scarcity of goods and hence the prices of

materials increase. The factor, difficulties in financing the project by the contractor was ranked 7. It seems that, if the contractor suffers from negative cash flow for most or all other projects, he fails to finance the project under consideration. Effect of subsurface conditions, was ranked eighth, this factor is unknown and cannot be controlled. Changes in the scope of the project, economic conditions, delay in the settlement of contractor's claims by the owner, excessive bureaucracy in the owner administration were ranked, 9, 10, 11, 12, respectively. Delay in materials procurement was ranked 13, it demonstrates a complete absence of effective material procurement management on the part of contractors. The established bureaucratic structures in material procurement processes in developing countries seem to hinder the easy flow of construction materials, especially in public sector. The respondents declared that delivery of materials to the job site in a timely manner is essential to keep things going and maintain high productive level, hence, delay of project is minimized. Inefficient coordination by the owner in the early planning &design stages received a rank equal to 14.Causes received RIW less than 10 will not be considered in the predictive model to reduce the number of variables to a manageable number except (Inefficient coordination by the owner in the early planning &design stages) with a RIW (9.95), which is close to 10. Table 2 lists the final 14 causes (independent variables) used to develop the regression model.

Table 1: Causes of delay affecting public water and sewage projects

No	Cause Identification	RIW	Rank
1	Changes in material prices	15.09	1
2	Delay in progress payments by the owner	14.5	2
3	Difficulties in obtaining work permits from the authorities concerned	13.6	3
4	Original contract duration too short	13.04	4
5	Delay in issuance of change orders by the owner	12.9	5
6	Inflation	11.9	6
7	Difficulties in financing the project by the contractor	11.5	7
8	Effect of subsurface conditions	10.6	8
9	Changes in the scope of the project	10.6	9
10	Economic conditions	10.6	10
11	Delay in the settlement of contractor's claims by the owner	10.4	11
12	Excessive bureaucracy in the owner administration	10.2	12
13	Delay in materials procurement	10.0	13
14	Inefficient coordination by the owner in the early planning &design stages	9.95	14
15	Poor communications by the owner with construction parties, and government authorities concerned during construction	9.26	15
16	Cash flow problems faced by the contractor	9.20	16
17	Changes in drawings	9.14	17
18	Delay in making decision by the owner within a reasonable time	9.13	18
19	Design errors	8.95	19
20	Traffic control and restrictions at the job site	8.90	20
21	Ambiguities, defaults, and inconsistency of specifications, and drawings	8.60	21

Table 1: Causes of delay affecting public water and sewage projects (continued)

No	Cause Identification	RIW	Rank
22	Delay in furnishing and delivering the site to the contractor by the owner	8.55	22
23	Subcontractor problems with the contractor	7.86	23
24	Interference by the owner in the construction operation phases	7.80	24
25	Poor coordination by the contractor with the parties involved in the project	7.75	25
26	Imported materials.	7.65	26
27	Delay in the approval of contractor submittals by the engineer.	7.59	27
28	Poor coordination by the consultant engineer.	7.55	28
29	Design changes.	7.50	29
30	Uncooperative owner with contractor	7.45	30
31	Shortage of materials required	7.40	31
32	Changes in Specifications	7.38	32
33	Low performance of the lowest bidder contractor	7.30	33
34	Spending some time in finding a suitable subcontractor	7.18	34
35	Incomplete documents	6.90	35
36	Delay in mobilization to start the project	6.86	36
37	Negotiation during construction	6.76	37
38	Ineffective planning and scheduling of the project by the contractor	6.67	38
39	Dispute (variation order)	6.65	39
40	Shortage of manpower	6.53	40
41	Failure of equipment	6.45	41
42	Low skills of manpower	6.36	42
43	Delay in performing inspection and testing by the consultant engineer	6.24	43
44	Transportation delays	6.21	44
45	Suspension of work by the owner	6.16	45
46	Improper technical study of the project in the bidding stage by the contractor	6.09	46
47	Changes in materials specifications	6.00	47
48	Poor qualification of supervision staff of the consultant engineer	5.95	48
49	Ineffective control of project progress by the contractor	5.95	49
50	Poor subcontractor performance	5.86	50
51	Mistakes during construction	5.80	51
52	Interference with other contractor's work	5.76	52
53	Shortage of technical professionals in the contractor's organization	5.75	53
54	Lack of coordination on-site	5.68	54
55	Differing site conditions	5.55	55
56	Changes in government regulations and laws	5.53	56
57	Shortage of equipment required	5.32	57
58	Delay in the preparation of contractor submittals	5.15	58
59	Mismanagement by the contractor (financial, supplier support)	4.95	59
60	Political Influence	4.80	60
61	Underestimation of Productivity	4.56	61
62	Ineffective contractor head office involvement in the construction of the project	4.55	62
63	Inefficient quality control by the contractor	4.55	63
64	Poor qualifications of the contractor's technical staff	4.53	64
65	Severe weather conditions on the job site	4.40	65

Table 1: Causes of delay affecting public water and sewage projects (continued)

No	Cause Identification	RIW	Rank
66	Defective work	4.28	66
67	Labor injuries	3.85	67
68	Site accident	3.81	68
69	Improper construction method implemented by the contractor	3.72	69
70	Delay in the field survey done by the contractor	3.60	70
71	Loose safety rules and regulations within the contractor's organization	3.60	71
72	Social Influence	3.58	72
73	Shortage of contractor's administrative personnel	3.35	73
74	Labor Dispute and Strike	2.61	74

Table 2: Candidate independent variable final list

No.	Variable	(RIW)
1	Changes in material prices	15.09
2	Delay in progress payments by the owner	14.50
3	Difficulties in obtaining work permits from the authorities concerned	13.60
4	Original contract duration too short	13.04
5	Delay in issuance of change orders by the owner	12.90
6	Inflation	11.90
7	Difficulties in financing the project by the contractor	11.50
8	Effect of subsurface conditions	10.60
9	Changes in the scope of the project	10.60
10	Economic conditions	10.60
11	Delay in the settlement of contractor's claims by the owner	10.40
12	Excessive bureaucracy in the owner administration	10.20
13	Delay in materials procurement	10.00
14	Inefficient coordination by the owner in the early planning &design stages	9.95

6. REGRESSION BASED MODEL

Data for 20 water and sewage projects was collected. These data include the occurrence of causes of delays presented in Table 1 on a yes/ no basis, and the corresponding actual percentage delay. The data was divided into two sets. The first set contains 12 projects for the purpose of model building. The second set contains 8 projects for validation p rposes. Initial experimentation with a regression model that includes all 14 variables resulted in a model with less performance using SPSS 8 software. Forward-stepping and backward-stepping were used. Forward stepping begins with entering the most significant variables at the first step, and continues adding and deleting variables until none can significantly improve the fit. Backward stepping, on the other hand begins with all candidate variables, then removes the least significant variable at the first step and continues until no insignificant variable remains. Forward stepping or backward-stepping technique gave the same model for predicting the delay percentage (DP) for water and sewage projects depending on 9 variables out of 14 (see Table 3) with a squared multiple R=0.978. This indicates that the model is able to explain 97.8 % of the variability in

the data, which is an excellent indicator of the model's expected performance.

The variables excluded were: changes in material prices, delay in progress payments by the owner, delay in issuance of change orders by the owner, delay in the settlement of contractor's claims by the owner, and delay in materials procurement. The underlying formula of model is DP = -0.2 + 0.3 (Difficulties in obtaining work permits from authorities concerned) -0.3 (Original contract duration too short) + 0.15(Inflation) –0.1 (Difficulties in financing the project by the contractor)+0.65(Effect of subsurface conditions) + 0.4 (Changes in the scope of the project) -0.25 (Economic conditions) +0.05 (Excessive bureaucracy in the owner administration) -0.45 (Inefficient coordination by the owner in the early planning &design stages).Where each of the 9 variables can have a 0 (unused) or 1 (used) value.

To show how the model predicts the delay percentage, it can be used in predicting delay percentage for a project with the following characteristics: difficulties in obtaining work permits were occurred (1); original contract duration was too short (1); there was an inflation (1); there were no difficulties in financing the project by the contractor (0); subsurface conditions was affect the project (1); changes in the scope of the project occurred (1);economic conditions were affected the project (1); there was an excessive bureaucracy in the owner administration (1); there was inefficient coordination by the owner in the early planning & design (1). The predicted delay percentage will be obtained as follows: %   Delay   =-0.2+0.3*1-0.3*1+0.15*1-0.1*0+0.65*1+0.4*1–0.25*1+0.05*1- 0.45*1=0.35=35%, this

result means that the expected delay percentage is 35%. Validation of this model will be given later using the second part of projects (including this example).

7. STATISTICAL FUZZY APPROACH

Hanna et al. [31] used statistical data to choose the fuzzy membership functions and to form the fuzzy if-then rules as illustrated in the following example.

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, if the linguistic value P is fixed for positive, correspondingly, N and Z for negative and zero, respectively, then, the fuzzy membership functions are given by:

$$P(y) = \begin{cases} y & for \quad 0 \leq y \leq 1 \\ 0 & otherwise \end{cases}$$

$$N(y) = \begin{cases} -y & for \quad -1 \leq y \leq 0 \\ 0 & otherwise; \text{ and} \end{cases}$$

$$Z(y) = \begin{cases} 1-y & for \quad 0 \leq y \leq 1 \\ 1+y & \quad -1 \leq y < 0 \end{cases}$$

Also, if some statistical data represented by the points (x_i, y_i), for I= 1, ….,n, from which the regression Eq. 3 is extracted, where the values of a and b are picked to minimize the sum of the squares of the errors (Eq. 4).

In traditional regression analysis x and y are correlated and a≠ 0. If Eq. 3 rewritten in the form given in Eq. (5) and assuming a>0, then, if then rules can be sought as follows:

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.

$$y = f(x) = ax + b \tag{3}$$

$$E = \sum_{i=1}^{n} (yi - axi - b)^2 \tag{4}$$

$$y = f(x) = a(x - c) \tag{5}$$

Hanna et al [31], declared that according to the above rules, the belief strength that some particular value of x is high can be identified with the belief strength that the corresponding predicted value of y is positive. Thus, the fuzzy membership function H could be defined as follows:

$H(x) = P[f(x)]$ and similarly

$M(x) = Z[f(x)]$ and $L(x) = N[f(x)]$

Accordingly, in this example

$$H(x) = \begin{array}{ll} a(x-c) & \text{for } x \geq c \\ 0 & \text{otherwise} \end{array}$$

$$L(x) = \begin{array}{ll} a(c-x) & \text{for } x \leq c \\ 0 & \text{otherwise; and} \end{array}$$

$$M(x) = \begin{array}{ll} 1 - a(c-x) & \text{for } x \geq c \\ 1 + a(c-x) & \text{for } x \leq c \end{array}$$

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. Where, the values of H, M, and L are determined in a similar way (Hanna et. al)[31]. Hanna et al. [31] added that if the function $f(x)$ is used 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. Thus, if u is N, then y is N; if u is Z, then y is Z; and if u is P, then y is P. Then, the fuzzy

functional can be generalized to the: if u is A, then y is A, where A is any fuzzy set. Fuzzy rules can be simplified to a general rule as follows: if $f(x)$ is A, then y is A. In their work, Hanna et al. [31], stated that if K input variables x_1, x_2, ., x_K affecting an output variable y, the method declared above is used to get a separate regression function $ui = fi$ (xi) for each input variable xi with the generalized fuzzy rule (Ri): if $fi(xi)$ is A, then y is A. Also, for each input variable xi, a fuzzy statement: $fi(xi)$ is Ai is held. They added that, since the belief in rule Ri can be measured by the coefficient of determination $(Ri)^2$ of the correlation between xi, and y, thus, the belief that "y is Ai" is also measured by R^2. Accordingly, the fuzzy membership Ai is measured by Eq. 6

$$A(y) = \frac{\sum_{i=1}^{K} R_i^2 A_i(y)}{\sum_{i=1}^{K} R_i^2} \tag{6}$$

Calling $w_i = \dfrac{R_i^2}{\sum_{i=1}^{K} R_i^2}$

Eq. (6) can be written in the form given in Eq. (7)

$$A(y) = \sum_{i=1}^{K} w_i A_i(y) \tag{7}$$

where A= weighted average of the A_{is}. Now to defuzzify the output y, the centroid y* could be found from Eq.8.

$$y^* = \frac{\int y.A(y)dy}{\int A(y)dy} \tag{8}$$

In practice, the integration is replaced by summation. The above approach is adopted to develop a model to predict the delay percentage in water and sewage projects. It must be noted that $Ai(y)$ in current research represents the independent variables which are the previous 9 causes affect delay of water

and sewage projects. Also, y* represents the predicted value of DP.

8. STATISTICAL FUZZY BASED MODEL

The following steps will be adopted for developing the proposed model:

1. A regression Eq. is extracted between each cause and DP.
2. Each cause is fuzzified according to it's degree of severity received from questionnaire responses.
3. The weight (wi) and the corresponding Ai* (Severity effect) for each cause is then determined.
4. A trapezoidal membership function for representing delay in water and sewage projects is adopted to determine the discrete universe for for DP (y) in general.
5. Predicted value of DP is then determined using Eq.8.

The above steps will be applied next to a case study to illustrate how to predict DP.

9. CASE STUDY

Data, separate regression Eq.(s) for each cause, and corresponding weight for this case study (previous example used in regression model) are given in Table 4.

Causes were fuzzified according to the degree of severity each cause was received from questionnaires' responses. For example the cause; difficulties in obtaining work permits from the authorities concerned (x_1) received the ranks: 5(100%), 4(80%), and 3(60%) for 8, 6, and 9 respondents respectively. Accordingly, the weighted degrees of severity effect are: $(1*8/23) = 0.348$, $(0.8*6/23)=0.209$, and $(0.6*9/23) = 0.235$ for 1, 0.8, 0.6 memberships respectively. Hanna et al.[31], declared that for maintaining convexity for a fuzzy set, one can added an entry between two values by taking the average. To deduce weighted degree of severity for 0.5, 0.7, and 0.9 memberships, the author adds the entries $(0+0.235)/2 = 0.118$, $0.235+0.209)/2 = 0.222$, and $(0.209+0.348)/2= 0.279$, respectively for cause (x_1). Now, the discrete universe of severity effect for this cause is

$x_1= \{0, 0, 0, 0, 0.118, 0.235, 0.222, 0.209, 0.279, 0.348\}$ for memberships (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1), respectively.

The discrete universes for the other variables were calculated as presented for cause (x_1) and were as follows:

$x_2= \{0, \quad 0, 0.038, 0.076, 0.109, 0.143, 0.205, 0.267, 0.253, 0.238\}$
$x_3= \{0, 0.02, 0.030, 0.040, 0.120, 0.200, 0.167, 0.133, 0.206, 0.278\}$
$x_4= \{0, \quad 0, 0.020, 0.040, 0.080, 0.120, 0.280, 0.440, 0.295, 0.150\}$
$x_5= \{0, 0.01, 0.045, 0.080, 0.100, 0.120, 0.200, 0.280, 0.240, 0.200\}$
$x_6= \{0, \quad 0, 0.045, 0.091, 0.196, 0.300, 0.355, 0.409, 0.205, \quad 0\}$
$x_7= \{0, 0.033, 0.050, 0.067, 0.117, 0.167, 0.128, 0.089, 0.184, 0.278\}$
$x_8= \{0, \quad 0, 0.037, 0.730, 0.200, 0.327, 0.236, 0.145, 0.118, 0.090\}$
$x_9= \{0, \quad 0, 0.026, 0.052, 0.169, 0.287, 0.265, 0.243, 0.165, 0.087\}$

Table 5 shows (Ai)* (Severity effect) of causes for the example project. Applying Eq.7, calculation of A(y) for this project is given in Table 6, (values of wi are obtained from Table 4). A trapezoidal membership function was proposed to represent the discrete universe for the most common delay percentages in water

and sewage projects as in Fig. 1. Data for 18 projects of the previous 20 projects and fuzzy evaluations: the min. value (A), the max. value (B), less than average (α), and larger than average (β) are given in Table 7. Fig. 2 shows fuzzy values, A(5%), B(80%), α (30%), and β (35%). The time overrun corresponding to membership function 0.4 (for example) can be deduced from Fig. 2 as follows:

$[((0.3-0.05)*0.4)+0.05] = 0.15$ or $[0.35+((0.8-0.35)*0.6)] = 0.62$.

Accordingly, the discrete universe: {0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.44, 0.53, 0.62, 0.71} is corresponding to membership functions {0.2, 0.4, 0.6, 0.8, 1, 1, 0.8, 0.6, 0.4, 0.2}> It must be noted that membership value equal to 1 was repeated because there is an interval for delay percentage (30 to 35%), i.e. two values 30% and 35% correspond to a membership 1.
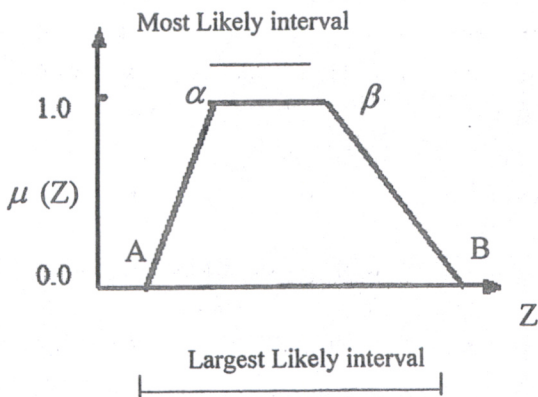

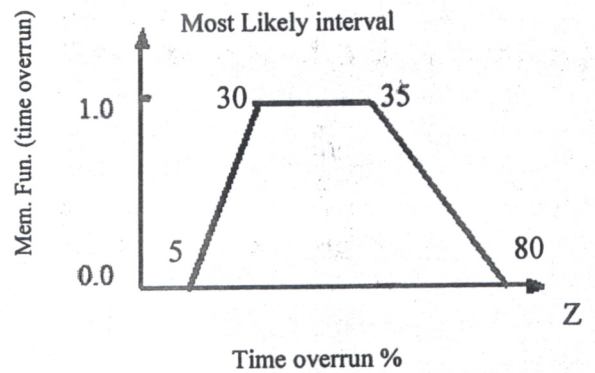
Fig 1. Membership function of delay percentage

Fig 2. Rate of time overrun in 18 projects

Table 4: Data and regression output for the example project

Cause	Occurr.	Regression outputs $ui=Ai=fi(xi)$		Weight (wi)
Difficulties in obtaining work permits from authorities concerned	1	$A_1=(0.21+0.175)$	$=0.385$	0.108
Original contract duration too short	1	$A_2=(-0.30+0.60)$	$=0.3$	0.222
Inflation	1	$A_3=(0.267+0.15)$	$=0.417$	0.236
Difficulties in financing the project by the contractor	0	$A_4=(0+0.375)$	$=0.375$	0.006
Effect of subsurface conditions	1	$A_5=(0.27+0.125)$	$=0.395$	0.180
Changes in the scope of the project	1	$A_6=(0.0545+0.3)$	$=0.3545$	0.004
Economic conditions	1	$A_7=(0.0889+0.283)$	$=0.3719$	0.026
Excessive bureaucracy in the owner administration	1	$A_8=(0.164+0.2)$	$=0.364$	0.036
Inefficient coordination by the owner in the early planning &design stages	1	$A_9=(-0.206+0.47)$	$=0.264$	0.182

Thus, the predicted value of delay percentage (y*) for the example problem is determined by applying Eq. 8. The values of discrete universe for delay percentage given previously are multiplied by the values of [A(y)] given in Table 6 to get the demonstrator of Eq. 8, while the denominator is produced from the summation of [A(y)] values.

$\int y.A(y)dy=0.1*0+0.15*0.002+0.2*0.026$
$+0.25*0.058+0.3*0.091+0.35*0.128+0.4$
$4*0.159+0.53*0.188+0.62*0.188+0.71*0.$
182
$=$
$0.5053 \int A(y)dy=0+0.002+0.026+0.058+$

$0.091+ 0.128 + 0.159 + 0.188 + 0.188 + 0.182 = 1.022$

$$y^* = \frac{\int y.A(y)dy}{\int A(y)dy} = \frac{0.5053}{1.022} = 0.494$$

y^*= 49.4%, thus, the expected percentage delay is 49.4%.

Table 5: Ai^* (Severity effect) of causes for the example project

Ai^* {Severity effect for causes}	
0.385	{0, 0, 0, 0, 0.118 , 0.235, 0.222, 0.209, 0.279, 0.348}
0.3	{0, 0, 0.038, 0.076, 0.109, 0.143, 0.205,0.267, 0.253,0.238}
0.417	{0, 0.020, 0.030, 0.040, 0.120, 0.200, 0.167, 0.133, 0.206, 0.278}
0.375	{0, 0, 0.020, 0.040, 0.080, 0.120, 0.280, 0.440, 0.295, 0.150}
0.395	{0, 0.010, 0.045, 0.080, 0.100, 0.120, 0.200 ,0.280, 0.240, 0.200}
0.3545	{0, 0, 0.045, 0.091, 0.196, 0.300 ,0.355, 0.409 ,0.205, 0}
0.3719	{0, 0.033, 0.050, 0.067 , 0.117 , 0.167, 0.128, 0.089,0.184, 0.278}
0.364	{0, 0, 0.037, 0.730, 0.200, 0.327, 0.236, 0.145,0.118, 0.090}
0.264	{0, 0, 0.026, 0.052, 0.169, 0.287, 0.265, 0.243,0.165, 0.087}

Table 6: Calculation of $A(y)$ for the example project

cause	$A(y)= wi*Ai(y)*$severity effect for causes									
1	0	0	0	0	0.005	0.009	0.009	0.008	0.012	0.014
2	0	0	0.019	0.038	0.055	0.072	0.103	0.134	0.127	0.119
3	0	0.002	0.003	0.004	0.012	0.019	0.016	0.013	0.02	0.027
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0.003	0.005	0.007	0.008	0.014	0.019	0.017	0.014
6	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0.001	0.002	0.001	0	0.002	0.003
8	0	0	0	0.009	0.003	0.004	0.003	0.002	0.002	0.001
9	0	0	0.001	0.002	0.008	0.014	0.0130	0.012	0.008	0.004
$A(y)$	0	0.002	0.026	0.058	0.091	0.128	0.159	0.188	0.188	0.182

Table 7: The rate of time overrun in 18 projects and fuzzy evaluation

Project No.	1	2	3	4	5	6	7	8	9
Time overrun (%)	5	10	15	15	20	20	25	25	30
Project No.	10	11	12	13	14	15	16	17	18
Time overrun (%)	30	35	35	40	50	50	50	65	80

Fuzzy evaluation	Min (A)	Max (B)	Average	Less than average(α)	Larger than average(β)
Time Overrun (%)	5	80	33.3	30	35

10. MODELS VALIDATION

A comparison between the regression model estimate and the statistical fuzzy model estimate is shown in Table 8. It provides the actual % delay, predicted % delay, and the analysis of the average

percent error for eight projects including the project of case study (these are projects of the second set) The regression model shows prediction accuracy better than (average % error = 30.27) statistical fuzzy model (average % error = 38.45).

11. SUMMARY AND CONCLUSIONS
This paper investigated the effect of causes of delay affecting water and sewage projects in Egypt through a questionnaire survey. These causes were established from literature. The questionnaire survey was also used a structured format to obtain information related to the occurrence of the previous causes in actual projects on a yes/no basis.

Based on the results of the questionnaires, a relative importance weight was established for each cause to quantify its effect on project delay. It was intended that causes received a relative importance weight equal to or higher than 10 are significant and will be incorporated into the model as independent variables. Accordingly, 14 significant causes were identified. The dependent variable was the delay percentage.

Two models were developed to predict the delay percentage in projects. The first model based on regression analysis. Data of 12 projects was used for model building, while the data of remaining 8 projects was used for validation purposes. The best regression model was found accurate in predicting the delay percentage contains 9 variables. The second model based on statistical fuzzy approach, where statistical results were used in forming fuzzy if then rules as well as choosing membership functions of linguistic values. Validation of the two models revealed that the regression model has prediction capabilities higher than that of statistical fuzzy model in predicting delay percentage in water and sewage projects in Egypt.

This research is relevant to both industry practitioners and researchers. It provides a systematic approach for practitioners to predict delay percentage for water and sewage projects. On the other hand, it provides researchers with a methodology to build regression and statistical fuzzy models suitable for delay percentage prediction. Computer implementation for statistical fuzzy model is suggested for future research for easy implementation.

Table 8: Validation of models for % delay prediction

Case project	Percent delay output		
	Project actual (X_{actual})	Regression model $(X_{estimated})$	Statistical fuzzy model $(X_{estimated})$
Example project	40	35	49.4
2	25	35	47.93
3	20	55	48.6
4	50	50	49.6
5	30	90	31.61
6	65	90	31.61
7	80	65	48.31
8	50	55	47.8
Average % error	0.00	30.27	38.45
Average % error = $|X_{actual} - X_{estimated}|/X_{estimated} \times 100$			

12. REFERENCES

[1] Ghoddosi, P., Hosseinalipour, M., and Jala, M.P., "fuzzy Assessment of Causes of Time Overrun (Delays) in Iran's Dam Construction Preojects."J. of Applied sciences 8 (19); 3423-3430, 2008.

[2] Arditi, D. and Robinson, M. A., Concurrent Delays in Construction Litigation" j. of Cost Eng., 37:20-28, 1995.

[3] Peter, W.G.M., and Hough, G.H., "The Anatomy of Major Projects,' 1st Edn. Wiley, New York, ISBN-10: 0471915513, 1987.

[4] Morris, S., "Cost and Time Overruns in Public Sector Projects." Econ. Political Weekly, 25: M154-M168.

[5] Flyvbjerg, B., Holm, M.K.S., and Buhl, S. L. "How common and How Large are Cost Overruns in Transport Infrastructure Projects. Transport Rev., 23:71-88, 1995.

[6] Cher, S., "Quantifying and Apportioning Delay on Construction Project." J. of Cost Eng., 37:11-11, 1995.

[7] Michael, R.F., "Window Analyses of Compensable Delays." J. of Construction Eng. Manage., 125:96-100, 1999.

[8] Teery, W., "Assessing Extension of Time Delays on Major Projects." Int. J. of Project Manage., 21:19-26, 2003.

[9] Kartam, S., "Generic Methodology for Analyzing Delay Claims." J. of Construct. Eng. and Manage., 125: 409-419.

[10] Stumpf, "Schedule Delay Analysis." J. of Cost Eng., 42: 32-43.

[11] Sadi, A.A., Al-Khlil, M., and Al-Hazmi, "Causes of Delay in Large Building Construction Projects." J. of Manage. In Eng., 11:45-50.

[12] Al- khalil, M.I., and Al- Ghafly, "Delay in Public Utility Projects in Saudi Arabia" Int. J. of Project Manage., 17: 101-106, 1999.

[13] Al-Moumani, H.A., "Construction Delay: A quantitative Analysis." Int. J of Project Manage., 18: 51-59.

[14] Abdalla, M.O. and Battaineh "Causes of Construction delay: Traditional Contracts." Int. J. of Project Manage., 20: 67-73, 2002.

[15] Ahmed, S.M., Azhar, S., Castillo, M., and Kappagntula, P., "Construction delays in Florida: An Empirical Study." Florida International University, Miami, March, 2010.

[16] Frimpong, Y., Oluwoye, J., and Grawford, L., "Causes of Delay and Cost Overruns in Construction of Groundwater Projects in a Developing Countries; Ghana as a case Study." Int. J. of Project Manage., 21: 321-326, 2003.

[17] Alaghbari, W. A. M., "Factors Affecting Construction Speed of Industrialized building Systems in Malaysia, M.sc. Thesis, University Putra Malaysia, Serdang, 2005.

[18] Tumi, S. A., Omran, A., and Pakir, A. H., "Causes of Delays in Construction Industry in Libya." P. of the Int. Con. On Economics and Administ., Univ. of Bucharest, Romania, 14-15th November, 2009.

[19] Al- Ghafly, M.A., "Delay in the Construction of Public Utility Projects in Saudi Arabia." M.Sc thesis, faculty of the Collage of graduate Studies, King Fahd Univ.

of Petroleum & Minerals, Dhahran, Saudi Arabia, 1995.

[20] Al Sultan, A.S., "Determination of Construction Contract Duration for Public Projects in Saudi Arabia." M.Sc thesis, King Fahd Univ. of Petroleum& Minerals, Dhahran, Saudi Arabia, 1989.

[21] Zadah, L.A., "Fuzzy Sets." Inf. Control., 8, 338-353, 1965.

[22] Zadah, L.A., "Concepts of a Linguistic Variable and its Application to Approximate Reasoning, I." Inf. Sci., 8, 199-249, (1975a).

[23] Zadah, L.A., "Concepts of a Linguistic Variable and its Application to Approximate Reasoning, Π ." Inf. Sci., 8, 301-357, (1975b).

[24] Zadah, L.A., "Concepts of a Linguistic Variable and its Application to Approximate Reasoning, III ." Inf. Sci., 9, 43-80, (1975c).

[25] Zadah, L.A., "A Theory of Approximate Reasoning." Machine Intelligence, Vol. 9, J. Hayes, D. Michie, and mikulich, L.I., eds., 149-194,1979.

[26] Ayyub, B.M., and Haldar, A., "Project Scheduling using Fuzzy Set Concept." J; of constr. Eng. And manage., 110 (2), 189-204, 1984.

[27] Bogardi, I., and Bardossy, A., "Regional Management of an Aquifer for Mining under Fuzzy Environmental Objectives." Water Resource Res., 19: 1394-1402, 1983.

[28] Anandalingam, G. and Westfall, M., "Selection of Hazardous Waste Disposal Alternative using Multi-Attribute Utility Theory and Fuzzy Set Analysis."J. of Environ. Syst., 18: 69-85, 1988.

[29] Loterapong, P., and Moselhi, O., "Project Network Analysis using Fuzzy Sets Theory.", J. of Constr. Eng. And Manage., 122(4), 308-318, 1996.

[30] Hanna, A. S., and Lotfallah, W.B., "A Fuzzy Logic Approach to the Selection of Cranes." Autom. Constr., 8, 597-608, 1999.

[31] Hanna, A. S., Lotfallah, W.B., and Lee, M.J., "Statistical- Fuzzy Approach to Quantify Cumulative Impact of Change Orders." J. of Computing in Civil Eng., Vol. 16, No.4 PP.252-258, October 1, 2002.