International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 4, April 2012)

Modeling the Parametric Construction Project Cost Estimate using Fuzzy Logic

Nabil Ibrahim El Sawalhi
The Islamic University- Gaza
nsawalhi@iugaza.edu.ps

***Abstract*— The research aims to model to predict the parametric cost estimation in construction building projects in Gaza Strip using Fuzzy logic. The building factors which effect the parametric cost estimation were surveyed. A questionnaire survey and relative index ranking technique were used to indicate the most five important factors from the view points of contractors, clients and consultants. 106 Case studies from real executed construction project in Gaza Strip were collected for the most important five factors to build up Fuzzy logic Model. The results revealed the ability of Fuzzy Model to predict cost estimate to an acceptable degree of accuracy.**

***Keywords*—Cost estimating, fuzzy logic, modeling, parametric estimate.**

I. INTRODUCTION

It is a well-known fact that the earlier cost planning is started on; the more suitable results are obtained. Several cost estimation and calculation models can be used during construction process; beginning with the conception phase of the construction project and the schematic design, design development, construction documents phases respectively. However, building cost estimation is an important issue due to the incomplete nature of the project data in the schematic design phase [1].

Cost is one of the major criteria in decision making at the early stages of a building design process. In today's globally competitive world, diminishing profit margins and decreasing market shares. Cost control plays a major role for being competitive while maintaining high quality levels. The cost of a building is impacted significantly by decisions made at the design phase. While this influence decreases through all phases of the building project, the committed costs increase [2].

Due to the limited availability of information during the early stages of a project, construction managers typically leverage their knowledge, experience and standard estimators to estimate project costs. As such, intuition plays a significant role in decision making. Researchers have worked to develop cost estimators that maximize the practical value of limited information in order to improve the accuracy and reliability of cost estimation work and thus enhance the suitability of resultant designs and project execution work [3].

Most of the significant factors affecting project costs are qualitative such as client priority on construction time; contractor's planning capability, procurement methods and market conditions including the level of construction activity [4]. Traditionally cost estimates are deterministic i.e. point estimates for each cost element based on their most likely value [5]. Reliable predictions of cost and duration are amongst the highest determinants of success of construction projects. Construction practitioners are aware of uncertainty, incompleteness, unknown circumstances and complex relationships of factors affecting cost and duration of construction projects [6]. The primary function of cost estimation is to produce an accurate and reliable cost forecast of a construction project. However, which cost should be forecasted depends on the requirements of a client and also upon the information and data available to develop the model. For instance, a client or a contractor may need to know the lowest tender price at one stage and/or the final project cost at completion stage [6].

The factors influencing construction costs were formed into four groups/layers; project-specific factors, client-contractor related factors, competition and market conditions, and finally macroeconomic and political factors. The usefulness of each group of factors is in explaining regional cost increase disparities breaks down to whether the impact of these factors is confined to a specific project, region, or the whole country.

Project characteristic and client requirements such as size and quality could influence both the amount and the unit prices of the input resources needed to undertake a project and could increase the direct cost portion of the estimated construction costs. Import of materials and labor mobility may resolve shortages of resources. Causes of construction cost escalations can be numerous and any effort to ascertain them in order to explain regional disparities requires that all the major construction cost components affected by the increase must first be recognized. Imprecise concepts and categories as well as mix-up of what constitutes cost or price makes it difficult to systematically identify these cost constituents". The impact of project-specific factors on regional construction cost differences could mainly be linked to the indirect cost part of the construction costs where client-contractor related factors such as contractor/client type and the extent of relationship between contracting parties seem to

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 4, April 2012)

influence these costs. Besides, not all client-contractor related factors are helpful to explain regional cost increase differences. Client-contractor relationship is the only factor in this layer that presumably influences construction costs through indirect cost components where a long run and strong past relationship between the parties could reduce transaction costs and the incentive to price according to current demand [7].

One of the most important issues is updating unit cost of the composite element alternatives in the database due to the inflation. However, the unit cost of the building operational units and composite element alternatives can easily be updated, as the cost calculation system of the software is based on current prices of inputs (i.e. material, equipment and labor unit prices [1]

II. Fuzzy Logic

As software development has become an essential investment for many organizations, software estimation is gaining an ever-increasing importance in effective software project management [8]. In practice, software estimation includes cost estimation, quality estimation, risk analysis, etc. Accurate software estimation can provide powerful assistance when software management decisions are being made; for instance, accurate cost estimation can help an organization to better analyze the feasibility of a project and to effectively manage the software development process, therefore, greatly reducing the risk [9].

The aim of this research is use fuzzy logic in cost estimating for construction projects in Gaza Strip. The success of modern mathematics is largely due to the efforts of Aristotle and the philosophers. In their efforts to devise a concise theory of logic, and later mathematics, the so-called "Laws of Thought" were invented. One of these, the "Law of the Excluded Middle," stated that every proposition must either be True or False, A or not-A, either this or not this. For example, a typical rose is either red or not red. It cannot be red and not red. Every statement or sentence is true or false or has the truth value 1 or 0. Even then, there already were strong and immediate objections to this proposal [10].

In fact, almost two centuries earlier, Buddha saw the world filled with contradictions, that things could be a certain degree true and a certain degree false at the same time. However, it was Plato who laid the foundation for what would become fuzzy logic, indicating that there was a third region (beyond True and False) where these opposites "tumbled about[11].

In the early 1900s, Lukasiewicz came and proposed a systematic alternative to the bi-valued logic (*bivalence*) of Aristotle. He described a three-valued logic, which can best be translated as the term `possible', and assigned it a numeric value between True and False. Knuth, a former student of Lukasiewicz proposed a three-valued logic apparently missed by Lukasiewicz, whiched used an integral range [-1, 0 +1] rather than [0, 1, 2].

Nonetheless, this alternative failed to gain acceptance, and has passed into relative obscurity [11].

It was not until relatively recently that the theory of fuzzy logic was discovered. Lotfi A. Zadeh, a professor of UC Berkeley in California, soon to be known as the founder of fuzzy logic observed that conventional computer logic was incapable of manipulating data representing subjective or vague human ideas such as "an attractive person" or "pretty hot". Fuzzy logic hence was designed to allow computers to determine the distinctions among data with shades of gray, similar to the process of human reasoning. In 1965, Zadeh published his seminal work "Fuzzy Sets" which described the mathematics of fuzzy set theory, and by extension fuzzy logic. This theory proposed making the membership function (or the values False and True) operate over the range of real numbers [0.0, 1.0]. Fuzzy logic was now introduced to the world. [12].

Although, the technology was introduced in the United States, the scientist and researchers there ignored it mainly because of its unconventional name. They refused to take something which sounded so child-like seriously. Some mathematicians argued that fuzzy logic was merely probability in disguise. Only stubborn scientists or ones who worked in discrete continued researching it [13].

A. *Fuzzy Definition*

Zadeh (1994)[14] said Fuzzy Logic (FL) is a problem-solving control system methodology that lends itself to implementation in systems ranging from simple, small, embedded micro-controllers to large, networked, multi-channel PC or workstation-based data acquisition and control systems. It can be implemented in hardware, software, or a combination of both. FL provides a simple way to arrive at a definite conclusion based upon vague, ambiguous, imprecise, noisy, or missing input information. FL's approach to control problems mimics how a person would make decisions, only much faster.

B. *Fuzzy Basic Concepts*

Fuzzy Rules the notion central to fuzzy systems is that truth values (in fuzzy logic) or membership values (in fuzzy sets) are indicated by a value on the range [0.0, 1.0], with 0.0 representing absolute Falseness and 1.0 representing absolute Truth [13]. For example, let us take the statement: "Jane is old." If Jane's age was 75, we might assign the statement the truth value of 0.80. The statement could be translated into set terminology as follows: "Jane is a member of the set of old people."

This statement would be rendered symbolically with fuzzy sets as: mOLD (Jane) = 0.80.

Where m is the membership function, operating in this case on the fuzzy set of old people, which returns a value between 0.0 and 1.0.

IJETAE
EXPLORING RESEARCH AND INNOVATIONS

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 4, April 2012)

C. *Fuzzy Rules*

Fuzzy machines, which always tend to mimic the behavior of man, work the same way. However, the decision and the means of choosing that decision are replaced by fuzzy sets and the rules are replaced by fuzzy rules. Fuzzy rules also operate using a series of if-then statements. For instance, if X then A, if y then b, where A and B are all sets of X and Y. Fuzzy rules define fuzzy patches, which is the key idea in fuzzy logic.

A machine is made smarter using a concept designed by Bart Kosko called the Fuzzy Approximation Theorem (FAT). The FAT theorem generally states a finite number of patches can cover a curve as seen in the figure below. If the patches are large, then the rules are sloppy. If the patches are small then the rules are fine.

D. *Logic Operations*

Now that we know what a statement like "X is LOW" means in fuzzy logic, how do we interpret a statement like X is LOW and Y is HIGH or (not Z is MEDIUM), The standard definitions in fuzzy logic are:

truth (not x) = 1.0 - truth (x)
truth (x and y) = minimum (truth(x), truth(y))
truth (x or y) = maximum (truth(x), truth(y).

E. *Fuzzy Control*

The purpose of control is to influence the behavior of a system by changing an input or inputs to that system according to a rule or set of rules that model how the system operates. The system being controlled may be mechanical, electrical, chemical or any combination of these. Fuzzy control, which directly uses fuzzy rules, is the most important application in fuzzy theory. Using a procedure originated by Ebrahim Mamdani in the late 70s, three steps are taken to create a fuzzy controlled machine[15]:

1) Fuzzification (Using membership functions to graphically describe a situation)
2) Rule evaluation (Application of fuzzy rules)
3) Defuzzification (Obtaining the crisp or actual results)

FL offers several unique features that make it a particularly good choice for many control problems.

1) It is inherently robust since it does not require precise, noise-free inputs and can be programmed to fail safely if a feedback sensor quits or is destroyed. The output control is a smooth control function despite a wide range of input variations.

2) Since the FL controller processes user-defined rules governing the target control system, it can be modified and tweaked easily to improve or drastically alter system performance. New sensors can easily be incorporated into system simply by generating appropriate governing rules.

3) FL is not limited to a few feedback inputs and one or two control outputs, nor is it necessary to measure or compute rate-of-change parameters in order for it to be implemented. Any sensor data that provides some indication of a system's actions and reactions is sufficient. This allows the sensors to be inexpensive and imprecise thus keeping the overall system cost and complexity low.

4) Because of the rule-based operation, any reasonable number of inputs can be processed (1-8 or more) and numerous outputs (1-4 or more) generated, although defining the rule base quickly becomes complex if too many inputs and outputs are chosen for a single implementation since rules defining their interrelations must also be defined. It would be better to break the control system into smaller chunks and use several smaller FL controllers distributed on the system, each with more limited responsibilities.

5) FL can control nonlinear systems that would be difficult or impossible to model mathematically. This opens doors for control systems that would normally be deemed unfeasible for automation.

FL is used through the following steps [16]:

1) Define the control objectives and criteria: What am I trying to control? What do I have to do to control the system? What kind of response do I need? What are the possible (probable) system failure modes?
2) Determine the input and output relationships and choose a minimum number of variables for input to the FL engine (typically error and rate-of-change-of-error).
3) Using the rule-based structure of FL, break the control problem down into a series of IF X AND Y THEN Z rules that define the desired system output response for given system input conditions. The number and complexity of rules depends on the number of input parameters that are to be processed and the number fuzzy variables associated with each parameter. If possible, use at least one variable and its time derivative. Although it is possible to use a single, instantaneous error parameter without knowing its rate of change, this cripples the system's ability to minimize overshoot for a step inputs.
4) Create FL membership functions that define the meaning (values) of Input/output terms used in the rules.
5) Create the necessary pre- and post-processing FL routines if implementing in S/W, otherwise program the rules into the FL H/W engine.
6) Test the system, evaluate the results, tune the rules and membership functions, and retest until satisfactory results are obtained.

Determination methods break down broadly into the following categories:

1) Subjective evaluation and elicitation

As fuzzy sets are usually intended to model people's cognitive states, they can be determined from either simple or sophisticated elicitation procedures. At they very least, subjects simply draw or otherwise specify different membership curves appropriate to a given problem. These subjects are typcially experts in the problem area. Or they are given a more constrained set of possible curves from

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 4, April 2012)

which they choose. Under more complex methods, users can be tested using psychological methods.

2) Ad-hoc forms

While there is a vast (hugely infinite) array of possible membership function forms, most actual fuzzy control operations draw from a very small set of different curves, for example simple forms of fuzzy numbers. This simplifies the problem, for example to choosing just the central value and the slope on either side.

3) Converted frequencies or probabilities

Sometimes information taken in the form of frequency histograms or other probability curves are used as the basis to construct a membership function. There are a variety of possible conversion methods, each with its own mathematical and methodological strengths and weaknesses. However, it should always be remembered that membership functions are NOT (necessarily) probabilities.

4) Physical measurement

Many applications of fuzzy logic use physical measurement, but almost none measure the membership grade directly. Instead, a membership function is provided by another method, and then the individual membership grades of data are calculated from it.

5) Learning and adaptation.

III. METHODOLOGY

A questionnaire survey was conducted to identify the most important factors affecting the parametric project cost estimation. 5 factors were identified as the most important factors using relative importance index. The most important factors are: Area of typical floor, Number of floors, Number of elevators, Volume of HVAC, Type of external Plastering.

106 case studies from real implemented construction projects were collected. The collected cases were used to build up Fuzzy model to predict parametric cost estimation. Once the project cost most important factors have been defined, it is necessary to evaluate the consistency of the available data, in terms of measurability, reliability and completeness (i.e. real information content). In particular, with regard to the last point, data could result in being not suitable or not sufficient for the purpose "which inevitably leads to recycles on the previous phases or they could be redundant" which causes inefficiencies.

Data analysis has revealed the main input parameters to be used in the modeling and training of the network. These parameters were the predominant cost drivers of the case examples. They defined the buildings formal characteristics and the amount of material required for the structural and architectural construction of the building. The total area bears a strong relation to the total cost of the building data analysis and identification of the design variables.

A. Data

Proprietary data obtained from a lot of construction projects which implemented in Gaza Strip in all sectors and deferent scopes, durations and owners. This large collection of data will provide a rich source of information and previous experience to be used in the following stages as training. Each project must has a cost estimate and actual budget with time scheduling in defined specifications with final report of the project hand over, this detailed data is very useful in determined the factors. A full details data of 106 construction projects with all factors assigned as most critical in the final cost of the project was used to build up the model.

B. Data analysis

Once the project cost drivers have been defined, it is necessary to evaluate the consistency of the available data, in terms of measurability, reliability and completeness (i.e. real information content). In particular, with regard to the last point, data could result in being not suitable or not sufficient for the purpose "which inevitably leads to recycles on the previous phases or they could be redundant" which causes inefficiencies.

Measurability deals with the problem of assigning a range of values to each variable or driver. In the case of qualitative variables, this could imply the decomposition of the estimation problem and hence the generation of several parametric models.

Data analysis has revealed the main input parameters to be used in the modeling and training of the network. These parameters were the predominant cost drivers of the case examples. They defined the buildings formal characteristics and the amount of material required for the structural construction of the building. The total area bears a strong linear relation to the total cost of the building data analysis and identification of the design variables.

C. Research Strategy and work design

This research apply the fuzzy rules at cost estimating projects, to improve performance of cost estimating and avoid project performance fail.

The basic methodology which is considered to achieve the objectives of this research is literature review and past studies which give understanding to the origin of fuzzy, fuzzy principles and fuzzy tools and techniques. The research then discuss the applicable fuzzy at cost estimating, so a programming language where chosen (Matlab). Fuzzy is built in at Matlab. Figure 1 shows Fuzzy interface at Matlab.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 4, April 2012)

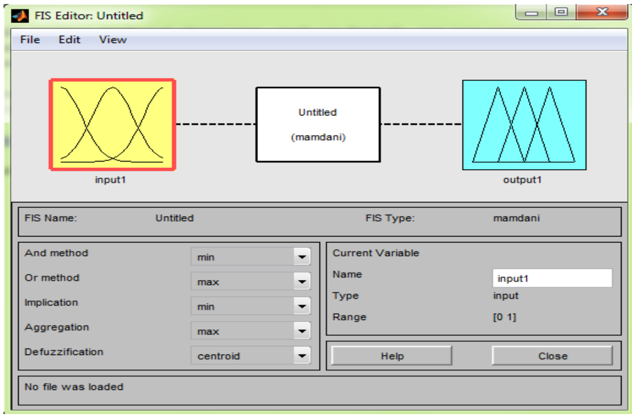

Figure 1 Fuzzy interface at Matlab

First step is fuzzification which includes write main factors affect the cost estimating (the input). Here there was five factors. Figure 2 shows the fuzzification process.

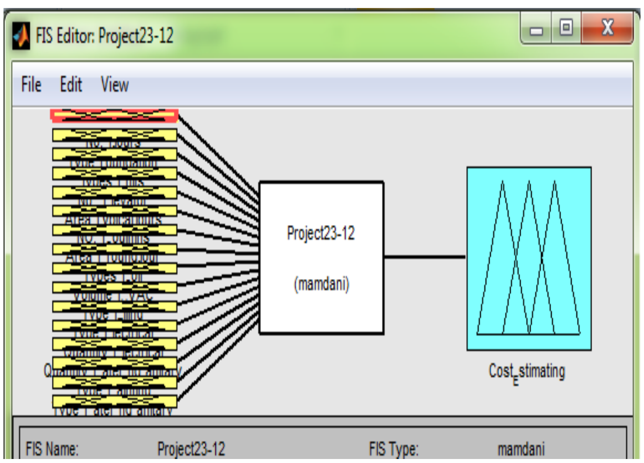

Figure 2 Fuzzy fuzzification

After determined of these factors their needs for the fuzzifier which mean transforms the input into linguistic terms using membership functions that represent how much a given numerical value of a particular variable fits the linguistic term being considered.

TABLE I

Variable Normalization Example

x	1	2	3	4	5	6
y	0	0.2	0.4	0.6	0.8	0

The chosen shape of the memberships was triangular. Figure shows the fuzzifier variables.

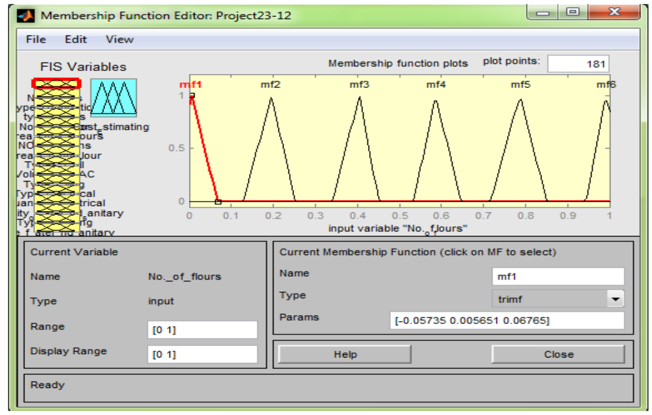

Figure 3 Fuzzy fuzzifier

The fuzzy inference engine performs the mapping between the input membership functions and the output membership functions using fuzzy rules that can be obtained from expert knowledge of the relationships being modeled. The greater the input membership degree, the stronger the rule fires, thus the stronger the pull towards the output membership function. Since several different output membership functions can be contained in the consequents of rules triggered, a defuzzifier carries out the defuzzification process to combine the output into a single label or numerical value as required. The required rules can be determined by (input)^membership.

C. Normalization Variable

According to the viable data the variable factors are discrete values with deferent ranges, the fuzzy set needs normalization the variable data to (0-1) values.

The normalization can make by the flowing equation

$$Y = \frac{X - \min .value}{Max.value - Minvalue}$$

Where: Y is the normalization value

X is the un- normalization value

Table 1 shows variable with six memberships and range from 1-6.

D. Validity test

Structure validity concerns the degree to which the variables, as measured by the research reflects the hypothesized construct [17]. After building the model and making all needs of rules, the model run and it give output for cost, many changed on input data done and the model give output, so the model is valid for cost estimating.

E. Reliability test

Test of reliability of model according to the model run. The reliability of an instrument is the degree of consistency which measures the attribute it is supposed to be measuring [17]. The results were compared with real projects cost, the outcome give a result near the real project cost. So the model is reliable for all projects.

IJETAE
EXPLORING RESEARCH AND INNOVATIONS

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 4, April 2012)

After making pilot study for the model the problem where determined at the rules number. Therefore if the model rules met with the project input the outcome will be near the real project cost.

A model with five important factors was built with six memberships and 27 rules. The tests made all the results near the real projects cost, so the cost estimating model is reliable if all rules added.

IV. RESULTS AND ANALYSIS

After finishing the model and end the (fuzzification) and start to inter variables data (rule evaluation), the outcome (Defuzzifecation) were suitable for projects. According to the error output, a model was built with five variables which had a high influence at projects cost, the results were logical and the cost could be predicted with a reasonable value of error. Figure 4 shows model output.

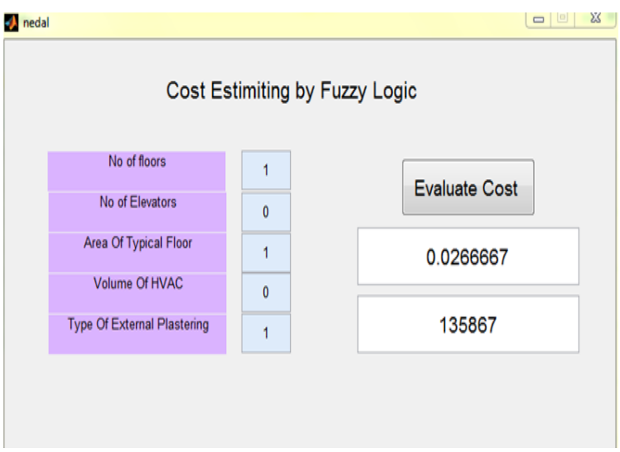

Figure 4 Model output

V. CONCLUSION

Construction industry is considered as an important sector in the industries. It is affected by clients, contractors, consultants, stakeholders, regulators, national economies and others. The main aim of this research is to model the parametric cost estimation using fuzzy logic approach. A triangular membership function in fuzzy technique was used for cost estimation and then it was validated with gathered data. Here, the advantages of fuzzy logic and good generalization are obtained. The main benefit of this model is its good interpretability by using the fuzzy rules and another great advantage of this research is that it can put together expert knowledge (fuzzy rules) project data into one general framework that may have a wide range of applicability in cost estimation. The results showed that Two-sided membership function is much better than other mentioned models. In order to achieve more accurate estimation, voting the estimated values of several techniques and combine their results maybe be useful.

References

[1] Hakan, Y., TAŞ, E., 2007. A building cost estimation model based on functional elements. Istanbul Technical University, Faculty of Architecture, Istanbul Turkey

[2] Gunaydın, M., and Dogan, Z., 2004. A neural network approach for early cost estimation of structural systems of buildings. International Journal of Project Management, Vol. 22, pp. 595–602.

[3] Cheng M., Tsai H., Sudjono E., 2010. Conceptual cost estimates using evolutionary fuzzy hybrid neural network for projects in construction industry. Expert Systems with Applications, Vol. 37, pp. 4224–4231.

[4] Nida Azhar, N., Farooqui, R., Ahmed, S., 2008. Cost overrun factors in construction industry of Pakistan. First International Conference on Construction In Developing Countries (ICCIDC–I), "Advancing and Integrating Construction Education, Research & Practice" , 4-5, Karachi, Pakistan.

[5] Baccarinibi, D., 1987. The maturing concept of estimating project cost contingency - A Review.

[6] Elhag and Boussabaine, 1998. An Artificial Neural System for Cost Estimation of Construction Projects, ARCOM 14th Annual Conference. Reading, UK.

[7] Warsame, A., 2006. Supplier structure and housing construction costs.

[8] Gray, A. and MacDonell S., 1997. Applications of Fuzzy Logic to Software Metric Models for Development Effort Estimation. The Information Science, Vol 97#10.

[9] Snell, D. 1997. Software Cost Estimation Website SCEW. http://www.ecfc.u-net.com/cost/compare.htm, Bournemouth University.

[10] Rush, C. and Roy, R. 2001. Workshop Expert judgement in cost estimating: Modelling the reasoning process. Concurrent Engineering, pp. 271-284.

[11] Lejewski, C. & Lukasiewicz,J. 1967 Encyclopedia of Philosophy, Vol. 5, MacMillan, NY, pp.104-107.

[12] Attarzadeh, I. and Hock Ow S. 2009. Software development effort estimation based on a new fuzzy logic model. International Journal of Computer Theory and Engineering, Vol. 1, No. 4, 1793-8201

[13] Idri, A. and Abran, A. 2001. Towards a fuzzy logic based measures for software projects similarity. 7th International Symposium on Software Metrics, pp. 85-96, England, UK.

[14] Zadeh, L. A. 1994. Fuzzy Logic. Neural Networks and Soft Computing. Communications of ACM, vol. 37, no. 3, pp.77-84.

[15] Zadeh, L. A., 1965. Fuzzy Set, Information and Control, vol. 8, pp. 338-353.

[16] Fuzzy Logic Toolbox User's Guide, 2010. MATLAB 6.0, Mathorks, Inc.,

[17] Fellows, R. and Liu, A. 2008. Research methods for construction. 3rd edition. United Kingdom, Blackwell Publishing Ltd.

Acknowledgement

Acknowledgement is given to my mater degree class students who participated in collecting and analyzing the research data.