

A Mamdani Fuzzy Logic Inference System to Estimate Project Cost



Daniel Helder Maia and Arjab Singh Khuman

Abstract The precision and reliability of estimations of project costs are essential, especially in significant cooperation. The level of uncertainty when estimating projects can cause issues down the line during a project. For generations, humans are more often than always in a predicament where estimation for a project size or cost appears to be complicated. The methodology adopted in this research included using the literature to review the topic of project estimation and explore the use of fuzzy logic in order to define an initial fuzzy system. The development of a system to estimate project costs is based on findings from the literature. This work seeks to demonstrate the benefits of using fuzzy logic in estimating the cost for business. Analysis of the results attained during testing and research shows that the system could be beneficial for estimating the cost of projects. The results show that the system can produce an appropriate result when estimating project cost. The study concludes that there is still room for improvement and that further development and testing could lead to improvements; however, the current system gives a foundation for further development such that the system can be put to use in a real-world situation. Whether it is for business or personal circumstances where any or most cases, cost estimation is required.

Keywords Cost estimation · Time estimation · Size estimation · Effort estimation · Expenses estimation

1 Introduction

How can businesses use a fuzzy inference system (FIS) to estimate how much a particular project should cost based on variables such as, time, size, expenses and travel time? Over the years, project managers, either underestimate or overestimate their projects. As humans, expecting to accurately estimate 100% the price of a particular project is based on how long it will take to complete by date/time, travel

D. H. Maia · A. S. Khuman (✉)
De Montfort University, Leicester, UK
e-mail: arjab.khuman@dmu.ac.uk

time, etc. This approach is fundamentally impossible. There have been many research papers, and researchers who have tried to calculate ways for specific tasks, but yet it has become almost virtually impossible to estimate a particular outcome accurately.

Due to unexpected changes of circumstances to a project, it can be difficult to estimate what should be completed when by a particular date, however, we still cannot be sure this will happen however the prediction is made. In essence, as an example, if we have project A with time, size, travel time, expenses and effort, as humans we can measure and predict that the project can/should be completed within say 6 months and cost X, yet the project completion may well end up taking 12 months costing the company more than the profit from the project. This principle applies to any other type of estimation, including the expenses of a project and predicting travel time spent. This is an ongoing issue that project managers face on a day-to-day basis, in fact, everyone faces it in our personal lives and personal projects.

With this context in mind, how can we use a Mamdani Fuzzy Logic Inference System, which first appeared in 1965 [1] to assist project managers to better estimate projects based on variables? We will incorporate this idea and build a FIS system which will allow and assist project managers to carefully estimate a project cost to get an even closer accurate estimation of projects.

2 Literature Review and Motivation

To understand the role and implementation of fuzzy logic in estimating projects, firstly we need to look at one of the most widely used models such as the Constructive Cost Estimation Model (COCOMO) [2]. This model enables software developers to estimate, time, effort and cost based on three categories. These are organic, semi-detached and embedded. The organic category is considered a group of experienced individuals in a small team with a small level of complexity, whereas, semi-detached aimed at experienced and new individuals in small to medium-size teams with a comparatively larger complexity than organic. Finally, embedded is aimed at a large experienced team with high complexity.

There is evidence in research demonstrating that fuzzy logic gives better software project estimates than the traditional COCOMO model [2]. Likewise, Reddy and Raju [3] agree that COCOMO is less efficient in its estimation compared to using fuzzy logic. Reddy and Raju [3] and Attarzadeh and Hock Ow [2] all agree that it yields uncertainty in the output, which leads to errors. It appears that fuzzy logic-based estimation models are more appropriate when dealing with ambiguous and unreliable information.

There is a general agreement from the research conducted and collected: it appears that the FIS is more accurate compared to the COCOMO model. However, there appears to be no further research on change of circumstances during the project process. Even though compared to other models, further research in depth to determine for which are suitable for further careful estimation as the design of the FIS.

Building on the idea that a FIS can be used to calculate estimation carefully, we need to look at the benefits of fuzzy logic in business and whether businesses would benefit. As we have seen thus far, it can be beneficial to implement fuzzy logic into different aspects of businesses. Specifically, in estimating projects, however, what about other aspects of business such as purchasing items?

Bezdek [4] collected information by approaching sales assistants and other employees from a Baumax store in Zlin. After asking 30+ questions, the outcome was that they overwhelmingly found that the staff were in favour of using fuzzy logic in their business. Furthermore, using the idea implemented in the paper, he/she found in three simple steps the best lawnmower identified for the customer and claimed the use of fuzzy logic therefore accelerates customer service. In accordance with the evidence outlined above and the results from the questionnaire, it appears fuzzy logic can be used and implemented in many areas of business. However, what has not been considered is whether or not fuzzy logic can alternatively cause a complication using this approach.

This next section provides a general discussion of how businesses encounter failed projects due to sparse estimation or in uncommon cases overestimated. Looking at how or why projects fail will allow us to be in a better position to design a fuzzy logic system to prepare for project estimation.

Doloi [5] discusses the shortfalls of traditional practices of estimation in which it plays a part in the failure of projects. Although he/she explains that estimation has an impact on the failure of projects, the author also stated that political, economic, technical and behavioural perspectives play a significant role in business case development. Although useful, there was not enough research found to get a concise cause of issues related to project failures. It would be interesting to know at which point projects fail and why precisely.

We will continue to look at some of the main points which are vital in the estimation process, which includes but is not limited to size, cost and time. Research was conducted on each of the variables and how they can be optimised with the implementation of fuzzy logic.

Several reports have shown that to estimate project quality or any aspect of the project, the critical inputs are time, cost and size. Studies conducted by Kharola and Singh [6] used time and cost as inputs to determine the quality of a project, by merely placing time and cost into a FIS and using the quality as the output. Similarly, Marandi and Khan [7] conducted a similar study, with minor differences in the inputs. For instance, Marandi and Khan [7] took consideration of the cost of failure into account, along with the general time and cost. Also, Reddy and Raju [3] both concentrated on the size of the project in conjunction with the COCOMO methods.

These approaches show that it is viable to use the variables of time, cost and size to create a custom FIS for project estimation ultimately. However, not all research conducted shows any debate on the use of different membership functions. Marandi and Khan [7] conducted the tests using only the trapezoidal and triangular membership functions. Therefore, further research is necessary for investigating additional membership functions and applying them to collect further samples to analyse further.

3 The System

3.1 System Design Overview

The proposition outlined below for this system is to allow businesses the ability to have their projects' cost estimated based on input variable values passed to the system. As we all know in project management and estimating projects, in general, cost is amongst one of the most significant issues all businesses encounter. We will look into a FIS designed using the same principles already researched in the literature review in Sect. 2, along with the understanding of the problem.

The view of the problem, as mentioned, is that businesses are unable to estimate project costs. Businesses tend to underestimate or overestimate. Perception of the solution is to use Fuzzy Logic to minimise estimation from a crisp binary result of either overestimated or underestimated and allowing fuzzy logic to give a vague estimation on the project cost.

This system has been split into two systems, starting with essential parts of estimation for any project which includes the time and size of the project. The outcome of those inputs will provide us with the effort required for the project. It is followed by using the output of that system as the input for the second inference system. Since the FIS not only has an output but also is used as an input for the second FIS, we add the expenses and travel time as the additional inputs to the second FIS. We now have the two inputs being expenses, travel time and the output from the first FIS known as effort. These inputs will make up the second FIS which will then give the output of the project cost.

A full visual diagram of the system can be seen in Fig. 1, how each of the components is modelled, to produce the full concept of the system. The system is broken into two parts to get a better understanding of how each of the parts plays in their system.

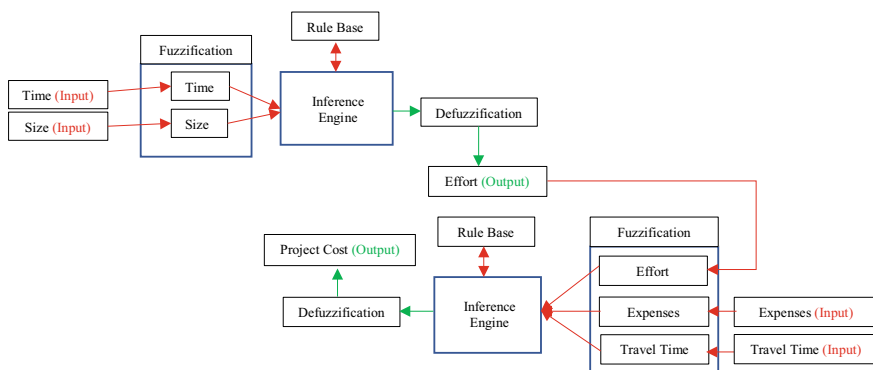


Fig. 1 Full fuzzy inference system

3.2 *Project Effort Fuzzy Inference System*

Overview

With the first inference system, we will consider the project effort. In this section, we will have two inputs and one output, and each of the inputs will allow us to calculate the output for the required effort for a particular project. As seen in Fig. 1, the system consists of two inputs, time and size.

Justification

Each of the variables was chosen based on research papers conducted in the literature review. How long it takes to complete the project truly depends on the size and the time it takes to complete. With this information, we will calculate the effort to pass through to the second inference system.

With regards to time, the decision to use a range of 0–100% was based on research conducted in the literature review. However, we can consider that a small percentage represents a meagre amount of time for the project to be completed, whereas the higher the percentage, the more time is available for the project. Kharola and Singh [6] used the same range values of 0–100% within their research. The interval values for this work were picked according to those used by Kharola and Singh [6].

The size variable will also have a range of 0–100%. The same principle applies as the time variable. The smaller the percentage, the smaller the size of the project and vice versa, the larger the percentage, the bigger the project. Even though the specified range was correctly picked, other researchers such as Sharma and Verma [8] used the same values in their research.

The range for the effort was chosen based on a paper by Sharma and Verma [8], where they specified the range of 0–1000 to represent the size. It would be more appropriate to have a range of 0–100% for not only simplicity but also the ease of measurement in general. The intervals, on the other hand, were chosen based on simplicity since it will not make much of a difference or effect to the system from using either 1000 or 100.

Fuzzification

The three types of membership functions used vary from the Triangular to the Trapezoidal and Gaussian membership functions.

The input for time utilises five membership functions which all five use the Triangular membership function. The decision to use the membership functions was taken based on Kharola and Singh [6].

The size input has been developed with six Gaussian membership functions. These membership functions were used based on research conducted by Sharma and Verma [8], however, since a new interval was added another membership function was also added based on the rest of the membership functions used for simplicity and consistency. The output utilises six Triangular membership functions.

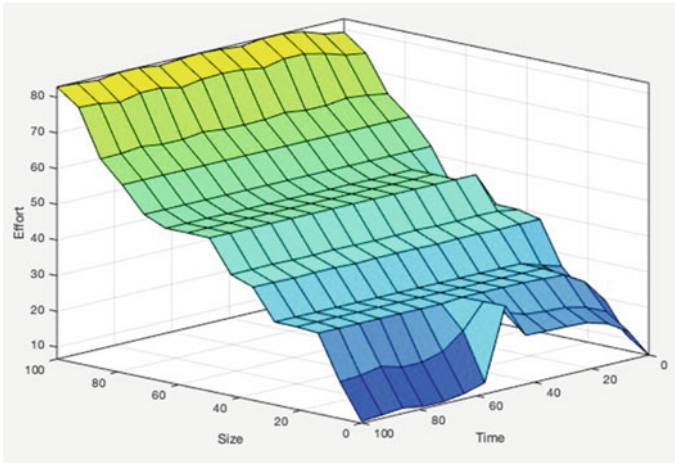


Fig. 2 Surface view of the project effort fuzzy inference system

Fuzzy Rule Base

Forty-one rules were created using the two inputs of time and size, over time during testing it will be amended to identify the rules to keep, add or remove. The fuzzy rule base infers the output based on the two input variables. The rule base has utilised the modus ponens as such

Antecedent – IF x is A THEN
Consequent – y is B

Defuzzification

The defuzzification is part of the process in which it produces a quantifiable result in crisp logic format. The first part of the FIS, defuzzification has been kept using the centroid to calculate the centre area of the method, which is commonly called the centroid. Then the x coordinate of the centroid is the defuzzied value (Fig. 2).

3.3 Project Cost Fuzzy Inference System

Overview

In the second fuzzy inference system, we will look at how the system will be completed. The output of the first inference system, Effort, will then be used with the other two inputs, expenses and travel time to complete the full FIS inputs; the final output will produce the project cost. The second inference system has been designed,

which utilises the output from the first project effort inference system to continue through the system to calculate the project cost. The project cost inference system has been designed to utilise the output from the project effort inference system to continue through the system to calculate the project cost.

Justification

As previously mentioned in the first inference system, why the decision to use the measurements of 0–100% and the values of the interval will be the same used for this new inference system.

The project cost is going to be measured in the same way as the other inputs of 0–100%. The choice of these measurements is due to maintaining consistency with the other inputs and based on the research conducted by Kharola and Singh [6]. In contrast, Attarzadeh and Hock Ow [2] used the measurement of 0–1.25, however, the preference of measurement chosen here for simplicity and consistency remains at 0–100%.

In general, time spent travelling for business will always differ based not only on a project but also by location. An initial idea for this is to use a range of 48 h. Nowadays, people can travel to the other side of the world within 24 h. The other 24 h are provided for leeway for travel disruptions and flight connections. However, since some business individuals travel for a living without a home base, it would be difficult to find an accurate representation in terms of exact hours for such cases.

In which case, a range of 0–100% is again the most appropriate range to use on the basis that we can say that anything above say 70% is viewed as a very long time spent travelling, likewise on the other side of the spectrum where anything less than 20% can be seen as short travel time or 0% seen as no travel time. This way we can adequately represent all cases.

Since measuring expenses is similar to that when dealing with cost, the judgement of using the same range and intervals as the project cost appears to be the most appropriate to use.

Fuzzification

Part of this second FIS has made use of two different membership functions, and this includes the Triangular and the Trapezoidal ones. There are three inputs and one output, and we will look into which membership functions are used for those variables. The first input we will look into is the output from the first inference system, which is project effort. The result of the first inference system will feed into this second inference system. In this case, the same membership functions have been used, which is the Triangular membership function. The second input is expenses, based on research conducted earlier in the report, having used the same membership function as the output (project cost) since expenses are calculated the same way very much. The third input is travel time. A mixture of Triangular and Trapezoidal membership functions was used. The intervals short, medium and long have been used for the Triangular functions whereas concise and very lengthy were used for the trapezoidal function. Finally, the project cost, which is the output, makes use

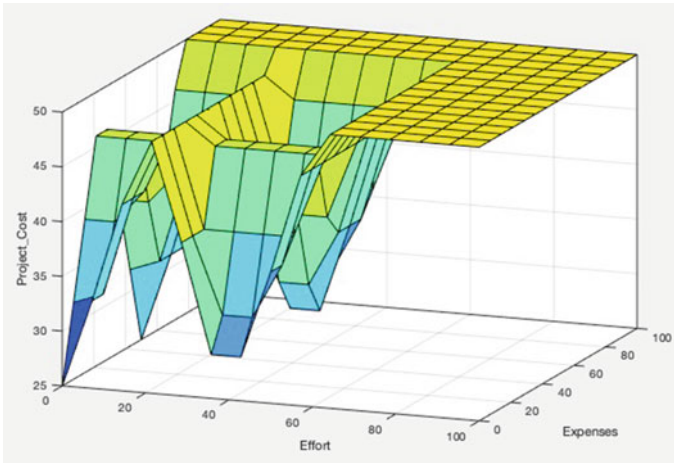


Fig. 3 Surface view of the project cost fuzzy inference system

of the triangular membership function. This decision was made again on research conducted earlier in the paper Kharola and Singh [6].

Fuzzy Rule Base

Thirty-six rules were created for the initial development of the system. The output from the first inference system is used as an input for this inference system, with the addition of two other inputs of travel time and expenses for the project cost calculation. For example, one of the rules is if the effort is very low and expenses are very low and travel time is very short, then the project cost will be very low.

Defuzzification

The Largest of Maximum (LOM) will be used initially to get the most significant outcome from the output. It can be seen later that tests were conducted with other defuzzification methods to identify which method best suited the system. After creating the surface viewer, further testing was required. See Fig. 3 of the first fuzzy inference system, the surface viewer.

4 Experimental Design and Evaluation

The tests were conducted using MATLAB version R2019b. The three most common membership functions were used; these are the Triangular, Trapezoidal and the Gaussian functions. The process of identifying the best suitable membership function for each of the inputs and outputs was tricky. However, those that have been used seem the most suitable at this time and following research conducted during the literature review.

Test One: Rules Adjustments

Changes

Firstly, we will look at going over a small set of rules to ensure the most appropriate defuzzification method to use, which will give us a steppingstone onto further tests, attempting to ensure that the rules for very low time and very large size, in particular, provide an appropriate result before proceeding with additional rules and tests.

Expected Outcome

Expecting only small differences when testing across all the defuzzification methods and believing there will be a suitable method upon completion of this test.

Results of the test

The changes did not produce accurate results across the two operators and all the defuzzification methods. There were only two output results that stood out, and it was felt that the AND operator was not the most suitable one to use for this particular system since the results were inconsistent (Table 1).

Defuzzification

The five defuzzification methods were used to get a crisp output for this inference system. Each of the defuzzification methods has produced results against all the sample data. The produced results varied, however, in some instances; centroid (rather than LOM) produced the closest prediction.

Analysis

During the tests, it was found that all the defuzzification methods produced some relevant results; however, many were not relevant. The decision was taken to alter the naming of the fuzzy sets for both size and time to make it easier to identifying and implementing the rules. Changing the membership functions for the fuzzy set size to use the Triangular membership function may help. Minor changes were made to the membership functions for time. The fuzzy sets were changed because, in the instance of size, the original naming convention did not appear to be most appropriate. In reality, when we measure size, we think small, medium, large work well just as we would when shopping for clothes which is why the decision to rename it was made. The time fuzzy set names were kept the same with one minor change and removing the Extremely High (EH) since it was no longer required and would have no effect since the rules would not make use of it.

Test Two: Re-test Test 1 Changes

Changes

Membership functions and their values have been changed for both inputs and the effort output. The fuzzy sets have also been modified, as mentioned in the test one analysis. We added further rules to this test, however, not expecting the outputs to change too much.

Table 1 Represents AND and OR operators for test one

AND operator (%)						
Time	Size	Centroid	Bisector	MOM	SOM	LOM
90	5	7.16091954	7	4	0	8
5	90	92.8609338	93	96.5	93	100
1	90	92.8609338	93	96.5	93	100
90	1	7.16091954	7	4	0	8
99	1	6.357552581	6	0	0	0
1	99	93.64244742	94	100	100	100
10	90	92.83908046	93	96	92	100
90	10	7.16091954	7	4	0	8
50	25	50	50	50	50	50
25	50	50	50	50	50	50
OR operator (%)						
Time	Size	Centroid	Bisector	MOM	SOM	LOM
90	5	6.351945602	6	0	0	0
5	90	93.4	94	98	96	100
1	90	93.64244742	94	100	100	100
90	1	6.351945602	6	0	0	0
99	1	6.351945602	6	0	0	0
1	99	93.64245073	94	100	100	100
10	90	92.8609338	93	96.5	93	100
90	10	6.904077562	6	3	0	6
50	25	9.5	9	9.5	0	19
25	50	9.909172282	10	9.5	0	19

Expected Outcome

There was an expectation of improved results due to the alteration to the membership functions replacing the old fuzzy set names with something more appropriate. The tests used the same sample data as before to remain consistent across all the tests.

Results of the test

There were drastic changes when using the OR operator, along with all the defuzzification methods compared to the last test. They seem far from appropriate compared to the previous results. However, the results for using the AND operator produces the exact results expected when using the default defuzzification method of Centroid (Table 2).

Table 2 Represents AND and OR operators for re-testing test one

AND operator (%)						
Time	Size	Centroid	Bisector	MOM	SOM	LOM
90	5	5.3030303	5	3	0	6
5	90	93.9230769	94	95	90	100
1	90	93.9230769	94	95	90	100
90	1	5.3030303	5	3	0	6
99	1	4.70588235	4	0.5	0	1
1	99	95.2941176	96	99.5	99	100
10	90	93.9230769	94	95	90	100
90	10	6.07692308	6	5	0	10
50	25	50	50	50	50	50
25	50	50	50	50	50	50
OR operator (%)						
Time	Size	Centroid	Bisector	MOM	SOM	LOM
90	5	50	50	50	0	100
5	90	50	50	50	0	100
1	90	50	50	50	0	100
90	1	50	50	50	0	100
99	1	49.9774368	50	0	0	0
1	99	50	50	50	0	100
10	90	50	50	50	0	100
90	10	46.9514563	47	3	0	6
50	25	37.5	38	37.5	25	50
25	50	62.5	63	62.5	50	75

Defuzzification

All the same five defuzzification methods were used to generate and get a crisp output.

Analysis

Across all the defuzzification methods and operators, the most appropriate from a simple test is when expecting the time to be around the 10% region, therefore very low, and the size expected to be around 90%; the expected effort would be around 90%.

The same goes when the time is very high at around 90% and the size at around 10%. The expectation is to have the effort to be lower. With Centroid, those exact results are produced making this the best defuzzification method for this system along with using the AND operator.

Test Three: Weight Value Changes

Changes

The weight will be amended to the rules to see if the weight changes will have any effect on the output. The weight values will be amended to 1, 0.75, 0.5, 0.25 and 0.

Expected outcome

Not expecting the results to change much from the most recent results produced. However, there will be minor changes to the results but not much to warrant any significant further changes to the system.

Results of the test

As expected, there was not much of a change in the results from the outcome of the tests. Much of the results were comparable to the previous tests.

Defuzzification

Only the centroid defuzzification method was the most important for this test, however, it created a table of comparisons against the other defuzzification methods. The operator has not changed and will remain using AND since the OR was weak in the previous tests.

Analysis

As expected from the results, the weight did not make much of a change to the results in which case the weight will remain at 1 for the remainder of the tests. The result was the same across all the defuzzification methods.

Test Four: Prod Implication and Sum Aggregation Methods

Changes

During this test, the implication and aggregation will be altered for the inference system, to see if this has any effect on the result. The test will include all the defuzzification methods across all the weight changes which have been tested as before.

Expected outcome

With the expectation to have significant changes to the output; more specifically, the result will have a considerable significant difference.

Results of the test

The prod implication method test was again used against the same sample data and modified rules as the previous test. The sum aggregation method, as investigated, found the results of using the prod implication method for this inference system along with a comparison to all the defuzzification methods to be beneficial (Table 3).

Table 3 Prod Implication and Sum Aggregation methods for AND and OR operators for test four

AND operator (%)						
Time	Size	Centroid (weight: 1)	0.75	0.5	0.25	0
90	5	4.666666667	4.66666667	4.6666667	4.66666667	50
5	90	95.33333333	95.3333333	95.333333	95.3333333	50
1	90	95.33333333	95.3333333	95.333333	95.3333333	50
90	1	4.666666667	4.66666667	4.6666667	4.66666667	50
99	1	4.666666667	4.66666667	4.6666667	4.66666667	50
1	99	95.33333333	95.3333333	95.333333	95.3333333	50
10	90	95.33333333	95.3333333	95.333333	95.3333333	50
90	10	4.666666667	4.66666667	4.6666667	4.66666667	50
50	25	50	50	50	50	50
25	50	50	50	50	50	50
OR operator (%)						
Time	Size	Centroid (weight: 1)	0.75	0.5	0.25	0
90	5	50	50	50	50	50
5	90	50	50	50	50	50
1	90	50	50	50	50	50
90	1	50	50	50	50	50
99	1	49.82972334	49.8297233	49.829723	49.8297233	50
1	99	50	50	50	50	50
10	90	50	50	50	50	50
90	10	45.64941129	45.6494113	45.649411	45.6494113	50
50	25	37.5	37.5	37.5	37.5	50
25	50	62.5	62.5	62.5	62.5	50

Defuzzification

All five defuzzification methods were tested as before, including centroid, bisector, MOM, LOM and SOM. This was also including the testing against the two operators AND and OR.

Analysis

In this instance, prod did not show much improvement to the system across all of the defuzzification methods, except the Centroid and Bisector where there was some minuscule improvement; with regards to the other defuzzification methods, they made the results much worse. Similarly, the use of sum made slight changes, again to the Centroid and Bisector defuzzification methods—there were small changes in results and like prod, the other defuzzification methods did not show an improvement in the results.

4.1 Project Cost Fuzzy System Tests

Test One: Rules Adjustments

Changes

As we did in the first inference system, we will first look at the current rule base and reduce the number of rules. Currently, there are 36 rule entries. The goal of this test is to reduce the number of rules and to ensure the inputs and outputs have the most appropriate membership functions associated with their fuzzy set, and these rule bases were altered. It is tested across all the defuzzification methods with both the AND and OR operators to determine which will be the most appropriate.

Expected outcome

Expected to have a much-reduced rule base list and notice that the membership function used may need altering as it did in the first inference system. This will provide a more accurate result on the test data.

Results of the test

After reducing the number of rules from 36 to 15, it made a significant difference in the system outcome, and it is still providing unexpected results. Not only is it suitable from a testing perspective since the system will be quicker to run, but also the rule base is sufficient to carry on with the tests further. After looking at the membership functions, alteration is required, especially for the output and the expenses input (Table 4).

Defuzzification

The five defuzzification methods were tested across the test data with the minimised rule base along with the two operators to get a crisp output for this inference system.

Table 4 Results from the data against the rule base

Project effort	Expenses	Travel time
5.303030303	100	100
93.92307692	25	50
93.92307692	50	25
5.303030303	10	0
4.705882353	75	25
95.29411765	20	75
93.92307692	5	5
6.076923077	30	50
50	50	50
50	35	0
50	0	0

Analysis

During this test, as expected, the membership functions need to be altered for both the project cost output and the expenses input. The travel time input appears to be just as good for the time being, however, it cannot rule out future changes at this moment in time. With regards to the results, it is currently inconclusive since the membership functions will need altering.

As seen in the result table, this gives inconclusive results, for example, when using the AND operator in particular for Centroid, the project effort returns 95% from the first inference system, expenses 5%, and travel time at 5% I would then expect the result to be around 80% or 85%, however, in this case, the result was 90%.

Test Two: Prod Implication and Sum Aggregation Methods

Changes

The changes made will cater to both the implication method and aggregation. When using the implication method, it will use the prod and the aggregation will use the sum. These two methods will test across all defuzzification methods and weights to decide on the most appropriate one to use.

Expected outcome

Expecting to get similar results as before for the first inference system, whereas most of the minor changes will be for the OR operator across all the defuzzification methods.

Results of the test

As expected, the results appeared to make the most difference using the OR operator. However, the implication method across all the weights based on the chosen defuzzification method caused some changes. Using the prod implication method based on the chosen defuzzification method, the values stayed consistent, meaning that no matter what the weight is the results will remain consistent. On the other hand, while using the sum aggregation method, their minimal differences.

Defuzzification

All five defuzzification methods will be tested against as before including centroid, bisector, MOM, LOM and SOM so that it will allow us to collect an exciting result set to compare against to make a better decision to proceed.

Analysis

In the previous tests, when the weight is set at 0, the results remain at 50 regardless of the defuzzification, implication and aggregation methods. Since using weight 0 makes no sense since project cost cannot be the same value across all the scenarios based on the effort, expenses and travel time, it is ignored as it was in the previous tests. However, in the results from the remaining weights across all defuzzification methods, the differences were very subtle.

When noticing, except for weight 0, the lower the weight starting from 1 using the OR operator with the sum aggregation method, the results tend to decrease slightly compared to finding the decrease of values when the weights are lowered. In contrast, while using the centroid and bisector, they intend to increase whereas the values across the weights using LOM defuzzification fluctuate. When compared to the prod implication method, the results are consistent across all the weights. The initial result is set based on the defuzzification method; then when the different weights are applied, the results are the same with no change.

Test Three: Weights against the default Max Aggregation and Min Implication

Changes

In this test, we will take the current system and its results by testing different values for the weight values of 1, 0.75, 0.5, 0.25 and 0 for all the rules across all the defuzzification methods to see if there will be any significant changes to the system and whether or not any will suit the system and provide an appropriate output.

Expected Outcome

Not expecting to get significant changes with the current results, especially for the AND operator, but for the OR operator we expect there will be significant changes. It may help deciding whether or not the change of weight will be beneficial to the system or if other changes need to be made to the system as a whole, in particular, changing the functions.

Results of the test

Firstly, the results from using the AND operator across all the weights do not appear to show much change. However, during testing using the OR operator, the results of changing the weight appear to be more significant than the AND operator. The most consistent result was when the weight was set to 0 across all defuzzification methods using both operators; there was no change. In every instance, and the return value was 50%.

Defuzzification

In this test, all five defuzzification methods have been used to test against all the weight changes with the two operators AND and OR. The weights used were 1, 0.75, 0.5, 0.25 and 0.

Analysis

Throughout the changes to the weight for the AND operator across the defuzzification methods, there was little change of any significance to the output based on the expected outputs. However, in some cases, the results from the OR operator across the defuzzification methods appear to show more noticeable results compared to those expected for the output.

Looking at Centroid, these results take the effort of 93%, expenses at 100% and travel time at 100%. I would then expect to have an output of 70–90% because the

more effort is required for the project then, the more the project is going to cost, however, in this instance it is giving an output of 50% which is slightly lower when using the OR operator. So, in this case, the Centroid defuzzification method would not be the best suited across all different weights.

The Bisector defuzzification method results in a similar output to the centroid with the expectation too when the project effort is at 5%, expenses at 10% and travel time at 0% giving a result of 19%. Similarly, when the effort is at 93%, expenses at 5% and travel time at 5% resulting in an output of 91% again, this was the expectation to the output result, however, across the other inputs, the values are not consistent and provide the expected result. Again, when using the OR operator, the outputs across all weights are inconsistent since the outputs using the OR operator are the opposite of that of the AND operator.

The Middle of Maximum calculates the most likely result. There was an output which resulted in expected output. While using the OR operator where the effort is at 5%, expenses and travel time at 100% with a weight at 0.25 returns an output of 63.9%, which is an output which is expected. However, for the rest of the outputs it was inconsistent, in particular, when the effort is at 95%, expenses at 20% and travel time at 75%, the output was 20% when expecting a moderately higher output. Similarly, it is the same in the case when using the AND operator.

The Smallest of Maximum as the name suggests returns the smallest value from the output. In this instance, while using the AND operator, there were two instances where the output made no sense. There were when the effort was at 95%, travel time at 75% and expenses at 20%, expect a higher output than 50%. Whereas when the effort is it at 93%, expenses at 5% and travel time at 5%, in expecting the output to be lower than the previous output of 95% effort. However, this is not the case, where the output returned a much larger output, in which it is expected to get when the effort was at 95%, expenses at 20% and travel time at 75%.

The Largest of Maximum is the opposite of the SOM, where it returns the maximum value from the output. While using the AND operator, the effort is at either 5% or 93% across all weights are equal. In this instance, the expectation of having different output values. Since more effort is required, the more the project is going to cost. Similarly, when using the OR operator, if the effort is at 5%, expenses and travel time are at 100% and the output is at 100% which again is not the output I am expecting for the system.

All in all, the system could still benefit from further development, especially the second inference system. It was manageable in narrowing down of which operator to use, in this case, the OR operator, however, further minor changes need to be made to the rule base, but most importantly to the membership function and its values. Creating an additional custom defuzzification method should be the next step to see if this will alter the results to become closer to those expected.

5 Discussion

Firstly, research was conducted on the topic to get an overview of what has already been found and if it can assist with testing. Secondly, a prototype of the system needs to be constructed on the basis of information gathered from the research work. Lastly, tests were conducted to see the robustness of the prototype. As expected, there were flaws with some positives and negatives. The tests were thorough and assisted as to which areas are needed to be improved and changed in order to improve the quality of the system outputs.

During the early stages of testing, the rule base was minimal. An increase in the number of rules improved the visibility of the scope of the system, but it hindered at the same time because it was much harder to achieve a match of what was expected compared with the actual results. So, the rule base was later reduced in the tests.

The system has full potential in most if not all environments related to cost estimation however. The limitation does not take into consideration the use of further variables, which would be another way to further develop the system. The system can be developed to something far more significant than it currently is; it has much potential with further testing and fine-tuning of the current system. For a future iteration, it would be useful to reconsider the variables used by comparing them to variables used commonly by businesses, for example, to estimate cost. The variables used currently are based on research and from our findings on how cost estimation is conducted.

The action plan is to conduct further background research, along with that already employed. It is not limited to only research papers but also talking to people within the field which can provide up-to-date insight of what they do on a day-to-day basis which can provide valuable information to produce additional variables for the system.

6 Conclusion

The system has evolved bit by bit throughout the work, and there is still much more testing required, specifically for the second FIS. However, based on the research already conducted where other researchers have already had a similar idea, it is close to being able to apply this system in a real-world scenario, whether that is at a personal project or in business planning. The system still requires further testing, as mentioned in the discussion section, but the system has been put through its paces during the testing phases.

The scepticism during the development of the project is the use of the fuzzy logic toolbox. In the essence of that, when prototyping a system, there are two issues which were encountered. The first being that we are dealing with two inference systems combined with the output from the first inference system, which is then fed into the second system as an input. You can only create one system at a time and cannot combine the two systems. The other is that when you have a system which has been

prototyped, then the need to convert the code from a .fis file to a .m file in itself can be a slight inconvenience but it just means more time needs to be spent. However, on the contrary, the option to save that time we could potentially write the code directly in the .m file, thus bypassing the toolbox completely but ultimately is down to preference. In this case, we prefer the visual representation to have a clear vision of the direction of the system.

Despite this, the opportunity to develop a system is to realise the idea of a system that handles business planning and estimating project costs. Applying the concept of fuzzy logic does not only give me a better understanding of how project cost estimations are conducted, but how effectively fuzzy logic can assist in getting better results in cost estimations; this is based on the research conducted at the beginning of the paper within the literature review and from the tests already conducted. This proves there is great potential for the use of fuzzy logic in cost estimations in almost all situations where cost estimations are required, and we believe it is not limited to business use; it could also be applied to other situations.

References

1. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.
2. Attarzadeh, I., & Hock Ow, S. (2010). Improving the accuracy of software cost estimation model based on a new fuzzy logic model. *World Applied Sciences Journal*, 8(2), 177–184.
3. Reddy, C. S., & Raju, K. (2009). Improving the accuracy of effort estimation through fuzzy set representation of size. *Journal of Computer Science*, 5(6), 451–455.
4. Bezděk, V. (2014). Using fuzzy logic in business. *Procedia - Social and Behavioral Sciences*, 124, 371–380.
5. Doloi, H. K. (2011). Understanding stakeholders' perspective of cost estimation in project management. *International Journal of Project Management*, 29(5), 622–636.
6. Kharola, A., & Singh, S. (2014). Development of fuzzy logic model for project management success (PMS). *PM World Journal*, 3(1), 1–11.
7. Marandi, A., & Khan, D. (2017). Software quality improvement and cost estimation using fuzzy logic technique. *International Journal of Applied Engineering Research*, 12(16), 5433–5440.
8. Sharma, V., & Verma, H. (2010). Optimized fuzzy logic-based framework for effort estimation in software development. *International Journal of Computer Science*, 7(2), 30–38.