International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 1 Issue 7, September - 2012

TAKAGI-SUGENO AND INTERVAL TYPE-2 FUZZY LOGIC FOR SOFTWARE EFFORT ESTIMATION

Author: Annepu BalaKrishna, RamaKrishna TK

Abstract

Software Effort Estimation carries inherent risk and this risk would lead to uncertainty and some of the uncertainty factors are project complexity, project size etc. In order to reduce the uncertainty, fuzzy logic is being used as one of the solutions. In this Chapter interval type-2 fuzzy logic is applied for software effort estimation. Two different methodologies have been discussed as two models, to estimate effort by using Takagi-Sugeno and Interval Type-2 fuzzy logic. The Formulas that were used to implement these models including Regression Analysis, Takagi-Sugeno membership functions, foot print of uncertainty intervals and de-fuzzification process through weighted average method were outlined along with analysis. The experimentation is done with NASA software data set on the proposed models, and the results are tabulated. The measured efforts of these proposed models are compared with available models from literature and finally the performance analysis is done based on parameters such as MARE, VARE and VAF.

1. Interval Type-2 Fuzzy Logic:

A type-2 fuzzy set, denoted as A, is characterized by a type-2 Membership Function (MF), $\mu_A(x, u)$, where $x \in X$ and $u \in J_x$, i.e.

$$\tilde{A} = \{((x,u),\ \mu_{\tilde{A}}(x,u)) \mid \forall x \in X,\ \forall x \in J_x \subseteq [0,1]\}$$

In which $0 \le \mu_A(x, u) \le 1$, if the universes of discourse X and the domain of secondary membership function J_x are continuous, A can be expressed as:

$$A = \int_{x \in x} \int_{x \in j_x} \mu_A(x,\mu) / (x,\mu),\ j_x \subseteq [0, 1]$$

Where ∫ denotes union over all admissible x and μ. If the universes of discourse X and J_x are both discrete, ∫ is replaced by Σ, A can also be expressed as:

$$A = \sum_{i=1}^{N} \sum_{j=1}^{M_k} \mu_A(x_i, \mu_j) / (x_i,\ \mu_j)$$

Where Σ denotes union over x and u. In the same way, if X is continuous and J_x is discrete or X is discrete and J_x is continuous, A can be expressed as:

$$A = \int_{x \in X} \sum_{j=1}^{M_u} \mu_A(x_i\ \ \mu_j) / (x_i,\ \mu_j) \text{ (or)}$$

$$A = \sum_{i=1}^{N} \int_{\mu \in j_x} \mu_A(x_i, \mu\ ) / (x_i,\ \mu)$$

The first restriction that is consistent with the T1 constraint $0 \le \mu_A(x) \le 1$. When uncertainties disappear a type-2 membership function is reduced to a type-1 membership function, in which case the variable u equals $\mu_A(x)$[11,12,13,14,15]. The second restriction that $0 \le \mu_A(x, u) \le 1$ is consistent with the fact that the amplitude of a membership function should lie between or equal to 0 and 1. When $\mu_A(x, u) \equiv 1$, A is an IT2 FS, it can still be expressed as a special case of general T2 FS as follow:

$$A = \int_{x\in X}\int_{\mu\in J_x} 1/(x,\mu) = \int_{x\in x}\left[\int_{\mu\in J_x} 1/\mu\right]/x, \quad \mu\in J_x \subseteq[0,1]$$

If universes of discourse X and J_x are both discrete, the above equation can be expressed as:

$$\tilde{A} = \sum_{x\in x}\left[\sum_{\mu\in J_x} 1/\mu\right]/x = \sum_{i=1}^{N}\left[\sum_{\mu\in J_{xi}}^{N} 1/\mu\right]/xi = \left[\sum_{x=1}^{m1} 1/\mu_{1k}\right]/x_1 + \dots + \left[\sum_{x=1}^{M_N} 1/\mu_{nk}\right]/x_N$$

In the above equation "+" denotes union.

Uncertainty in the primary memberships of an IT2 FS consists of a bounded region named as Footprint of Uncertainty (FOU). It is the union of all primary memberships, i.e.

$$FOU(A) = \underset{x\in x}{U}\, J_x$$

This is a vertical-slice representation of FOU, because each of primary membership is a vertical slice. The Upper Membership Function (UMF) and Lower Membership Function (LMF) of A are two T1 MFs that bound the FOU. The UMF is associated with the upper bound of FOU (A) and is denoted as $\bar{\mu}_A(x)$, $\forall\, x \in X$ and the LMF is associated with the lower bound of FOU(A) and is denoted as $\underline{\mu}_A(x)\ \forall\ x \in X$,i.e.

$$\bar{\mu}_A(x) = \overline{FOU(A)}\quad \forall\ x \in X$$
$$\underline{\mu}_A(x) = \underline{FOU(A)}\quad \forall\ x \in X$$

For an IT2FS, $J_x = [\underline{\mu}_A(x), \bar{\mu}_A(x)], \forall x \in x$ Therefore, the IT2FS A can be denoted as, Effort= $\sum_{i=1}^{n} [\underline{\mu}_A(x_j), \bar{\mu}_A(x_i)]/xi$ (*in discreate situation*) *or*

Effort = $\int_{x\in x}[\underline{\mu}_A(xi), \bar{\mu}_A(x_i)]/x$ (in continuous situation).

The following Figure 1 shows the components of Interval Type-2 Fuzzy Logic.

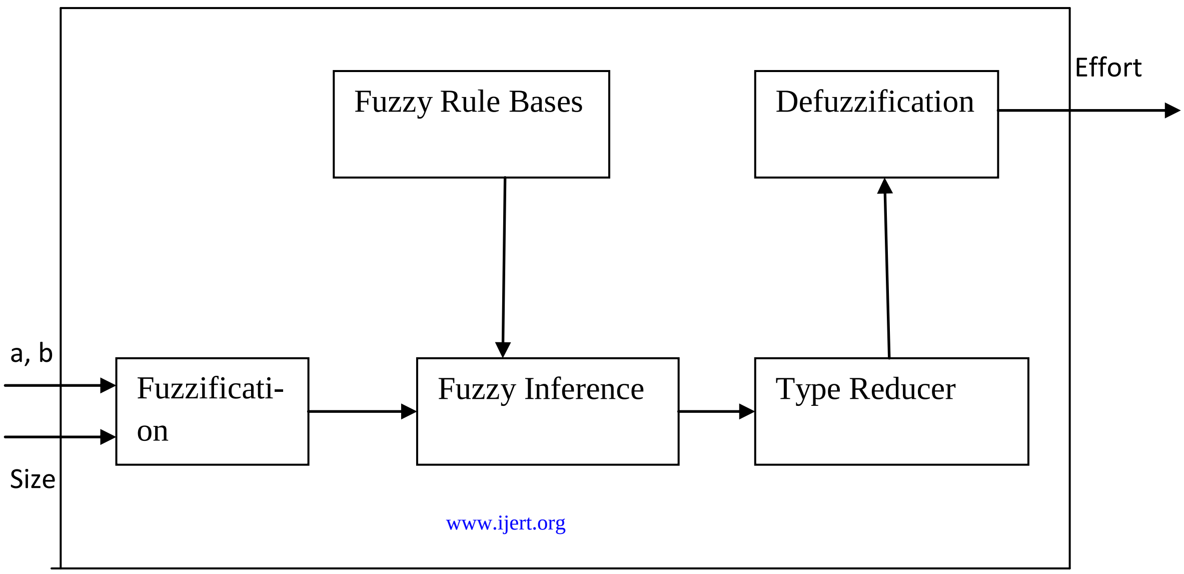


Figure 1: Structure of Interval Type-2 Fuzzy Logic

Fuzzification is the process which Translates inputs (real values) to fuzzy values. Inference System applies a fuzzy reasoning mechanism to obtain a fuzzy output. Knowledge Base contains a set of fuzzy rules, it is of the form R^i :if x_1 is F_1^i and …. x_n is F_n^i then Y is G^i ,i=1,2…m and a membership functions set known as the database. Type Reducer transforms a Fuzzy Set into a Type- 1 Fuzzy Set. The defuzzification traduces one output to precise values.
For an interval type-2 fuzzy system (ITF2S)

$$J_{f1} = \left[\frac{\mu_{PL1} + \mu_{PL2}}{2}, \frac{\mu_{PR1} + \mu_{PR2}}{2}\right] = \left[\underline{\mu}_P(x_i), \bar{\mu}_P(x_i)\right]$$

$$J_{f2} = \left[\frac{\mu_{NL1} + \mu_{NL2}}{2}, \frac{\mu_{NR1} + \mu_{NR2}}{2}\right] = \left[\underline{\mu}_N(x_i), \bar{\mu}_N(x_i)\right]$$

are the firing intervals for the membership functions positive and negative respectively, μ_{PL1}, μ_{PL2} are left hand side uncertainty region, μ_{PR1}, μ_{PR2} are right hand side uncertainty region.
The following section deals the two methodologies that have been used on the proposed models in order to estimate effort.

2 Model Description

2.1 Methodology for Model -1

In this model the mean of FOU `s as a firing interval in interval type-2, is considered to estimate the cost (effort) of the software.
Step1: This step estimate the parameters of a, b (the amplitude a and the exponent b) by regression analysis (power regression).
Step2: The variable "size" is then fuzzified by two input fuzzy sets named "Positive" and "Negative" respective. The mean of the sizes (L) is input for determining the fuzzy memberships. The representation shown in Figure3.2, the membership value $\mu_P(x_i)$ and $\mu_N(x_i)$ is either 0 or 1 when x_i is outside the interval [-L,L]. This process is known as Fuzzification.

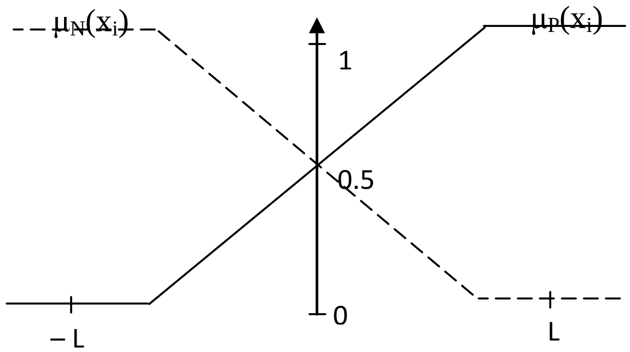


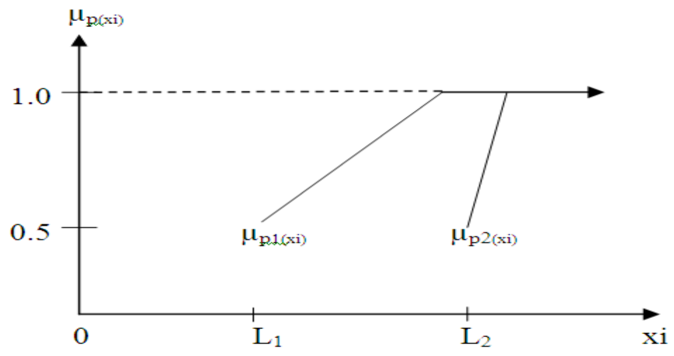
Figure 2: Universe of Discourse

Step 3: Then triangular membership function is applied to reduce the uncertainty (FOU-Foot Print of Uncertainty).The uncertainty at the left side (LMF) and right side (UMF) are then calculated by using following fuzzy rule: If x1 is μ(x1), x2 is μ(x2)…………xi is μ (xi) then μp1= [min (μ(x1), μ(x2)……, μ (xi))], [max (μ(x1), μ(x2)……,μ(xi))]. The same logic is used then for Negative membership function.
Step 4: After identifying the LMF and UMF for both of these positive and negative membership functions, the mean of the LMF and UMF FOU is employed as a firing interval for converting Type-2 into Type-1 fuzzy sets, this is known as Type Reducer.
Step 5: Finally in order to convert Fuzzy values into output (effort), weighted average defuzzification method is used.

2.2 Methodology for Model -2:
In this model Takagi-Sugeno Fuzzy Controller is considered, for determining the memberships, and Interval Type-2 logic and fuzzy operator for determining the firing intervals.
Step1: This step estimate the parameters of a, b (the amplitude a and the exponent b) by regression analysis (power regression).
Step 2: The variable "size" is fuzzified by two input fuzzy sets named "Positive1" and "Positive2" respectively. The mean and stddev (standard deviation) of size is used to determine the fuzzy memberships. The (mean+stddev) for the positive1 and (mean-stddev) for positive2 are the L values. The representation is shown in Figure3.3, the membership value $\mu_{p1}(x_i)$ and $\mu_{p2}(x_i)$ is either 0 or 1 when x_i is outside the interval [-L,L]. This process is taken as Fuzzification.
Step 3: Then triangular membership function is applied, to reduce the uncertainty (FOU). The uncertainty at the left side (LMF) and right side (UMF) are calculated by using following fuzzy rule: If x1 is μ(x1), x2 is μ(x2)…………xi is μ (xi) then μp1= [min (μ(x1), μ(x2)……, μ (xi))], [max (μ(x1), μ(x2)……,μ(xi))]. The same logic is used similarly for positive2 membership function.

Figure3.3: Membership Functions of Fuzzy sets in the Size's Space

Step 4: After identifying the LMF and UMF for both of these positive1 and positive2 membership functions, the mean of the LMF and UMF FOU, fuzzy operator max(OR), Left side Uncertainty interval, and Right side Uncertainty interval for Models are used as firing interval for converting Type-2 into Type-1 fuzzy sets.



Step 5: Finally in order to convert Fuzzy values into Output (Effort), weighted average defuzzification method is used.

3. MODEL ANALYSIS
The section analyses the proposed models.

Regression Analysis: By using power regression we calculate [www.xuru.com] a, b parameters $Y = ax^b$, Where x is the variable along the x-axis. The function is based on linear regression with both axis are scaled logarithmically.

Membership Functions for Model-1, Model-2

Membership Functions Used in Model-1: The mathematical definitions of the "Positive" and "Negative" fuzzy sets were identical for the input variable size are equation 1.1 and 1.2.

$$\mu_P(\Delta kloc) = \begin{cases} 0 & \Delta kloc < -L \\ \dfrac{\Delta kloc + L}{2L} & -L \le \Delta kloc \le L \\ 1 & \Delta kloc > L \end{cases} \tag{1.1}$$

and

$$\mu_N(\Delta kloc) = \begin{cases} 1 & \Delta kloc < -L \\ \dfrac{-\Delta kloc + L}{2L} & -L \le \Delta kloc \le L \\ 0 & \Delta kloc > L \end{cases} \tag{1.2}$$

The value of L affects the control performance and we take the mean of the input (size) as the L value.

Membership Functions Used in Model-2:

The mathematical definitions of the two "Positive" fuzzy sets were identical for the input variable size is equation 1.3 and 1.4.

$$\mu_{P1}(\Delta kloc) = \begin{cases} 0 & \Delta kloc < -L_1 \\ \dfrac{\Delta kloc + L_1}{2L_1} & -L_1 \le \Delta kloc \le L_1 \\ 1 & \Delta kloc > L_1 \end{cases} \tag{1.3}$$

and

$$\mu_{P2}(\Delta kloc) = \begin{cases} 0 & \Delta kloc < -L_2 \\ \dfrac{\Delta kloc + L_2}{2L_2} & -L_2 \le \Delta kloc \le L_2 \\ 1 & \Delta kloc > L_2 \end{cases} \tag{1.4}$$

The value of L affects the control performance and we taken (mean+stddev), (mean-stddev) of the input (size) as the L_1 and L_2 value.

Fuzzy Triangular Membership Function (TMF):

The triangular MF in specified by three parameters (α, m, β) as Figure :

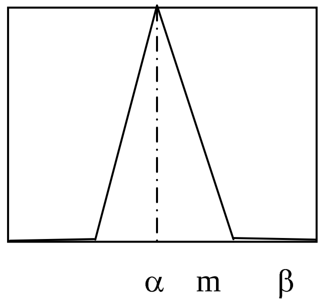


Figure 3: Triangular Member Function

The parameters (α, m, β) (with $\alpha < m < \beta$) determine the x – coordinates of the three corners of the underlying triangular MF.

Fuzziness:

Fuzziness of TFN (α, m, β) is defined as:

m = model value α = Left Boundary β = is right boundary

$$\text{Fuzziness of TFN (F)} = \frac{\beta - \alpha}{2m}, \qquad 0 < F < 1$$

The Higher the value of fuzziness, the more fuzzy is TFN. The value of fuzziness to be taken depends upon the confidence of the estimator. A confident estimator can take smaller values of F. Let (m, 0) divides internally, the base of the triangle in ration K : 1 where K in the real positive number.

$$\text{So that} \qquad m = \frac{\alpha + K\beta}{K+1}$$

As per the above definitions, $F = \dfrac{\beta - \alpha}{2m}$

$$\text{So} \qquad \alpha = \left(1 - \frac{2KF}{K+1}\right) * m \qquad \text{and} \qquad \beta = \left(1 + \frac{2F}{K+1}\right) * m$$

If we consider F = 0.1 and K = 1 then

$$\alpha = m * \left(1 - \frac{2 \times 0.1 \times 1}{2}\right) = 0.9m$$

$$\beta = m * \left(1 + \frac{2 \times 0.1 \times 1}{2}\right) = m(1 + 0.1) = 1.1m$$

<u>Footprint of Uncertainty in Universe Of Discourse:</u>

MODEL-1:

After applying fuzzification process on the size by using the positive and negative member functions, the Footprint of Uncertainty (FOU) in universe of discourse with uncertainty regions will be the one as shaded in the following figure 3.5.

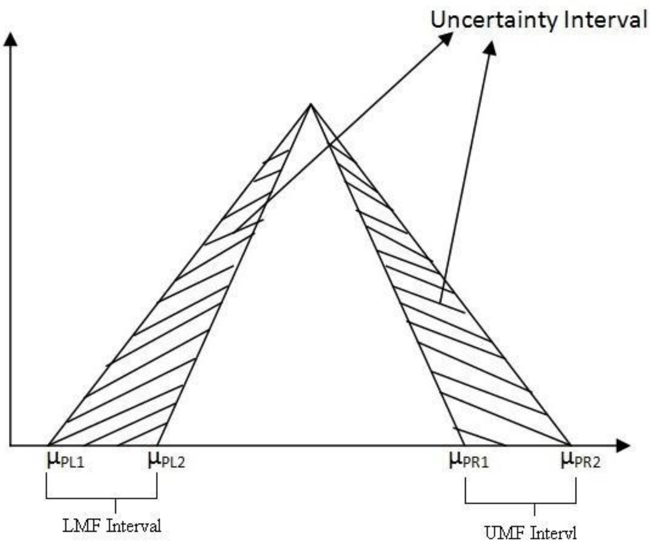


Figure 4: Footprint of Uncertainty and Prediction Intervals for Methodology-1

μ_{PL1}, μ_{PL2} are left hand side uncertainty region, μ_{PR1}, μ_{PR2} are right hand side uncertainty region.

MODEL-2:

After applying fuzzification process on the size by using two positive member functions, the Footprint of Uncertainty (FOU) in universe of discourse with uncertainty regions will be the one as shaded in the following figure 3.6.

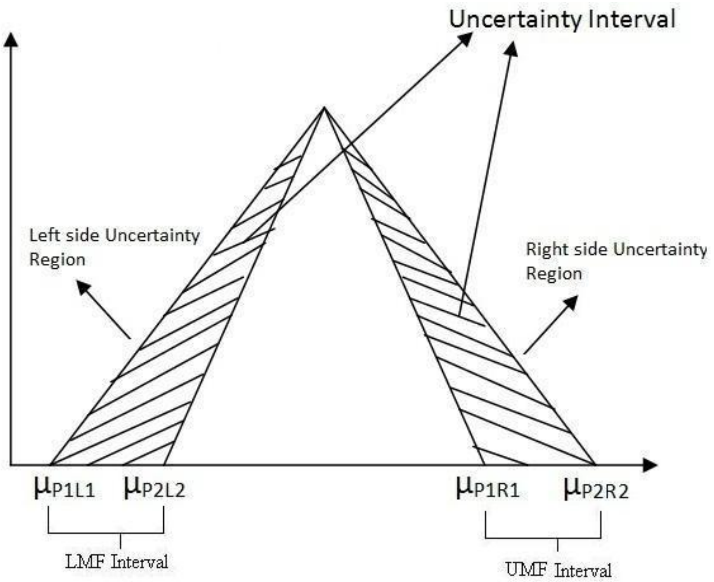


Figure 5: Footprint of Uncertainty and Prediction Intervals for Methodology-2

μ_{P1L1}, μ_{P2L2} are left hand side uncertainty region, μ_{P1R1}, μ_{P2R2} are right hand side uncertainty region.

Firing Intervals :

Firing Intervals for Methodology-1:

Here the means of FOU's are taken as firing strength.

$$J_{Px} = \left[\frac{\mu_{PL1} + \mu_{PL2}}{2}, \frac{\mu_{PR1} + \mu_{PR2}}{2} \right] = \left[\underline{\mu}_P(x_i), \bar{\mu}_P(x_i) \right]$$

$$J_{Nx} = \left[\frac{\mu_{NL1} + \mu_{NL2}}{2}, \frac{\mu_{NR1} + \mu_{NR2}}{2} \right] = \left[\underline{\mu}_N(x_i), \bar{\mu}_N(x_i) \right]$$

Firing Intervals for Methodology-2:

Case-I:

Here the means two positive member functions of FOU's are taken as firing strength.

$$J_{P_1x} = \left[\frac{\mu_{P_1L_1} + \mu_{P_1L_2}}{2}, \frac{\mu_{P_1R_1} + \mu_{P_1R_2}}{2} \right] = \left[\underline{\mu}_{P_1}(x_i), \bar{\mu}_{P_1}(x_i) \right]$$

$$J_{P_2 x} = \left[\frac{\mu_{P_2 L_1} + \mu_{P_2 L_2}}{2}, \frac{\mu_{P_2 R_1} + \mu_{P_2 R_2}}{2}\right] = \left[\underline{\mu}_{P_2}(x_i), \overline{\mu}_{P_2}(x_i)\right]$$

Case-II:

The uncertainty considered at the left, right hand side interval i.e. fuzzy operator OR (max) is used to determine the firing interval

$$J_{P\ x} = \left[\max\left[\underline{\mu}_{P_1}(x_i), \overline{\mu}_{P_1}(x_i)\right], \max\left[\underline{\mu}_{P_2}(x_i), \overline{\mu}_{P_2}(x_i)\right]\right]$$

Case-III:

The uncertainty considered only at the right hand side interval to determine the firing interval

$$J_{P\ x} = \left[\underline{\mu}_{P_2}(x_i), \overline{\mu}_{P_2}(x_i)\right]$$

Case-IV:

The uncertainty considered only at the left hand side interval to determine the firing interval

$$J_{P\ x} = \left[\underline{\mu}_{P_1}(x_i), \overline{\mu}_{P_1}(x_i)\right]$$

Defuzzification:

In these models weights average method, which is of the following form are considered.

$$C = \frac{\sum_{i=1}^{N} w_i\, \mu_i}{\sum_{i=1}^{N} w_i} \qquad (1.5)$$

where w_i is the weighting factor and μ_i is the membership obtained from triangular member function.

Performance Measures:

Three criterions were considered and they are outlined below

1) Variance Accounted For (VAF)

$$\%\ VAF = \left[1 - \frac{\text{var (Measured Effort} - \text{Estimated Effort)}}{\text{var (measured effort)}}\right] \text{x } 100$$

2) Mean Absolute Relative Error (MARE)

$$\%\ MARE = \text{mean}\left[\frac{\text{abs (Measured Effort} - \text{Estimated Effort)}}{\text{(measured effort)}}\right] \text{x } 100$$

3) Variance Absolute Relative Error (VARE)

$$\% \text{ VARE} = \text{var}\left[\frac{(\text{abs (Measured Effort} - \text{Estimated Effort)}}{(\text{measured effort})}\right] \text{x } 100$$

The following section describes the experimentation part of the work, and to conduct the study and in order to establish the effectively of the models dataset of 10 projects from NASA software project data [2] were used .

4. MODEL EXPERIMENTATION

Application on Model-1:

The membership function definitions and the memberships are shown here using equation 1.1 and 1.2; the L value is the mean of the input sizes i.e. 46. By applying power regression analysis (www.xuru.com) for the input sizes and effort the obtained values are : a=2.7 and b=0.8523. The membership functions are defined as follows using equation 3.1 and 3.2.

$$\mu_N(\Delta kloc) = \begin{cases} 1 & \Delta kloc < -46 \\ \dfrac{-\Delta kloc + L}{2L} & -46 \le \Delta kloc \le 46 \\ 0 & \Delta kloc > 46 \end{cases}$$

By applying Triangular membership function for the above membership functions the left and right boundaries obtained are shown below.

Foot print of uncertainty intervals for the μ_P is [0.4705 to 0.5751] for left hand side i.e. LMF and [0.9 to 1.1] for right hand side i.e. UMF. Foot print of uncertainty intervals for the μ_N is [0.3277 to 0.4005] for left hand side i.e. LMF and [0.4295 to 0.5249] for right hand side i.e. UMF. The means of FOU intervals is taken as firing strength.

$$J_{Px} = \left[\underline{\mu}_P(x_i), \overline{\mu}_P(x_i)\right] = [0.5228, 1]$$

$$J_{Nx} = \left[\underline{\mu}_N(x_i), \overline{\mu}_N(x_i)\right] = [0.3641, 0.4772]$$

The type reducer action by using the triangular membership function which is applied to the uncertainty region as a secondary member function and the results obtained are shown in Table 1and Table 2. Defuzzification process is done through weighted average method.

Table 1:Triangular Fuzzy Number of Adjusted Size and Effort Estimation for Positive Membership Function

S.No	Size(m)	$\alpha =$ 0.5228m	m	β =m	$a\alpha^b$	am^b	$a\beta^b$	E_p
1	2.1	1.09788	2.1	2.1	2.923672	5.081492	5.081492	4.978739
2	3.1	1.62068	3.1	3.1	4.074635	7.081925	7.081925	6.938721
3	4.2	2.19576	4.2	4.2	5.27833	9.174008	9.174008	8.9885
4	12.5	6.535	12.5	12.5	13.37204	23.24129	23.24129	22.77132
5	46.5	24.3102	46.5	46.5	40.97051	71.20884	71.20884	69.76892
6	54.5	28.4926	54.5	54.5	46.90638	81.52569	81.52569	79.87715
7	67.5	35.289	67.5	67.5	56.28813	97.83165	97.83165	95.85339
8	78.6	41.09208	78.6	78.6	64.08698	111.3865	111.3865	109.1341

9	90.2	47.15656	90.2	90.2	72.06489	125.2525	125.2525	122.7197
10	100.8	52.69824	100.8	100.8	79.22287	137.6934	137.6934	134.9091

Table 2: Triangular Fuzzy Number of Adjusted Size and Effort Estimation for Negative Membership Function

S.No	Size(m)	α = 0.3641m	M	β = 0.4772m	$a\alpha^b$	am^b	$a\beta^b$	E_N
1	2.1	0.76461	2.1	1.00212	2.147928	5.081492	2.704878	3.810078
2	3.1	1.12871	3.1	1.47932	2.993503	7.081925	3.769708	5.309992
3	4.2	1.52922	4.2	2.00424	3.877819	9.174008	4.883324	6.878626
4	12.5	4.55125	12.5	5.965	9.824005	23.24129	12.37133	17.4262
5	46.5	0	46.5	0	0	71.20884	0	71.20884
6	54.5	0	54.5	0	0	81.52569	0	81.52569
7	67.5	0	67.5	0	0	97.83165	0	97.83165
8	78.6	0	78.6	0	0	111.3865	0	111.3865
9	90.2	0	90.2	0	0	125.2525	0	125.2525
10	100.8	0	100.8	0	0	137.6934	0	137.6934

The following Table 3 shows the Measured Effort, Estimation Effort, Absolute Error and Relative Error.

Table3 : Error Calculations

S.No	Size	Measured Effort	Estimated Effort(E_p+E_n/2)	Absolute Error	Relative Error
1	2.1	5	4.394408	0.605592	0.121118
2	3.1	7	6.124357	0.875643	0.125092
3	4.2	9	7.933563	1.066437	0.118493
4	12.5	23.9	20.09876	3.801238	0.159048
5	46.5	79	70.48888	8.51112	0.107736
6	54.5	90.8	80.70142	10.09858	0.111218
7	67.5	98.4	96.84252	1.557483	0.015828
8	78.6	98.7	110.2603	11.5603	0.117126
9	90.2	115.8	123.9861	8.1861	0.070692
10	100.8	138.3	136.3013	1.998739	0.014452

Application on Model-2:

The membership function definitions and the memberships shown here are obtained using equation 3.3 and 3.4, the L value is the (mean + stddev) of the input sizes for Positive1 is (46+38.28) 84.28, and L value is the (mean – stddev) of the input sizes for Positive2 is (46-38.28) 7.72. By applying power regression (www.xuru.com) analysis for the input sizes and effort the obtained values of a, b are a=2.7 and b=0.8523

$$\mu_{P_1}(\Delta kloc) = \begin{cases} 0 & \Delta kloc < -84.28 \\ \dfrac{\Delta kloc + L_1}{2L_1} & -84.28 \le \Delta kloc \le 84.28 \\ 1 & \Delta kloc > 84.28 \end{cases}$$

and

$$\mu_{P_2}(\Delta kloc) = \begin{cases} 0 & \Delta kloc < -7.72 \\ \dfrac{\Delta kloc + L_2}{2L_2} & -7.72 \le \Delta kloc \le 7.72 \\ 1 & \Delta kloc > 7.72 \end{cases}$$

Case-1: The means of FOU intervals is taken as firing strength.

By applying Triangular membership function for the above membership functions the left and right boundaries obtained are shown in the following Table3.8 and Table3.9 [$\mu_{P1}(\alpha, m, \beta)$, $\mu_{P2}(\alpha, m, \beta)$]
Foot print of uncertainty intervals for the μ_{P1} is [0.5724 to 0.9] and Foot print of uncertainty intervals for the μ_{P2} is [0.6996 to 1.1]. The means of FOU intervals is taken as firing strength.

$$J_{Px} = \left[\underline{\mu}_P(x_i), \overline{\mu}_P(x_i)\right] = [0.7362, 0.8998]$$

The type reducer action by using the triangular membership function and associated results are shown in Table3.10. Defuzzification process is done through weighted average method.
Table 4: Triangular Fuzzy Number of Adjusted Size and Effort Estimation for Case-1
The following Table 5 shows the Measured Effort, Estimated Effort, Absolute Error and Relative Error.

S.No	Size (m)	α= 0.7362m	M	β =0.8998 m	$a\alpha^b$	am^b	$a\beta^b$	Effort
1	2.1	1.54602	2.1	1.88958	3.914099	5.081492	4.644189	4.817663
2	3.1	2.28222	3.1	2.78938	5.454964	7.081925	6.472469	6.714233
3	4.2	3.09204	4.2	3.77916	7.066424	9.174008	8.384511	8.697696
4	12.5	9.2025	12.5	11.2475	17.90196	23.24129	21.24119	22.03461
5	46.5	34.2333	46.5	41.8407	54.84972	71.20884	65.08075	67.5117
6	54.5	40.1229	54.5	49.0391	62.79644	81.52569	74.50975	77.2929
7	67.5	49.6935	67.5	60.7365	75.35636	97.83165	89.41245	92.75225
8	78.6	57.86532	78.6	70.72428	85.79716	111.3865	101.8008	105.6033
9	90.2	66.40524	90.2	81.16196	96.47767	125.2525	114.4735	118.7494
10	100.8	74.20896	100.8	90.69984	106.0605	137.6934	125.8438	130.5444

Table 5: Error Calculations for Case-1

S.No	Size	Measured Effort	Estimated Effort	Absolute Error	Relative Error
1	2.1	5	4.81766	0.182337	0.036467
2	3.1	7	6.71423	0.285767	0.040824
3	4.2	9	8.6977	0.302304	0.033589
4	12.5	23.9	22.0346	1.865395	0.07805
5	46.5	79	67.5117	11.4883	0.145422
6	54.5	90.8	77.2929	13.5071	0.148757
7	67.5	98.4	92.7523	5.647746	0.057396

8	78.6	98.7	105.603	6.903304	0.069942
9	90.2	115.8	118.749	2.949393	0.02547
10	100.8	138.3	130.544	7.75559	0.056078

Case 2: Using Fuzzy Operator for firing strength

This case deals with uncertainty at the left, right hand side interval i.e. fuzzy operator OR (max) is used here to determined the firing interval

$$J_{Px} = \left[\max\left[\underline{\mu}_{P_1}(x_i), \overline{\mu}_{P_1}(x_i)\right], \max\left[\underline{\mu}_{P_2}(x_i), \overline{\mu}_{P_2}(x_i)\right]\right]$$

Foot print of uncertainty intervals for the μ_{P1} is [0.5724 to 0.9], Foot print of uncertainty intervals for the μ_{P2} is [0.6996 to 1.1].The fuzzy operator max of FOU intervals is taken as firing strength.

$$J_{Px} = \left[\underline{\mu}_P(x_i), \overline{\mu}_P(x_i)\right] = [0.9, 1.1]$$

The Table 6 shows the Effort estimation using above firing intervals.

Table 6: Triangular Fuzzy Number of Adjusted Size and Effort Estimation for Case-2

S.No	Size(m)	α = 0.9m	m	β =1.1m	$a\alpha^b$	am^b	$a\beta^b$	Effort
1	2.1	1.89	2.1	2.31	4.645	5.081	5.511	5.265
2	3.1	2.79	3.1	3.41	6.473	7.081	7.681	7.337
3	4.2	3.78	4.2	4.62	8.386	9.174	9.95	9.506
4	12.5	11.25	12.5	13.75	21.245	23.241	25.208	24.082
5	46.5	41.85	46.5	51.15	65.093	71.208	77.234	73.786
6	54.5	49.05	54.5	59.95	74.523	81.525	88.424	84.476
7	67.5	60.75	67.5	74.25	89.429	97.831	106.11	101.373
8	78.6	70.74	78.6	86.46	101.82	111.386	120.812	115.419
9	90.2	81.18	90.2	99.22	114.495	125.252	135.851	129.786
10	100.8	90.72	100.8	110.88	125.867	137.693	149.345	142.678

The Table 7 shows the Measured Effort, Absolute Error, Estimated Effort and Relative Error.

Table 7: Error Calculations for Case-2

S.No	Size	Measured Effort	Estimated Effort	Absolute Error	Relative Error
1	2.1	5	5.265	0.265	0.0529
2	3.1	7	7.337	0.337	0.0481
3	4.2	9	9.506	0.506	0.0562
4	12.5	23.9	24.082	0.182	0.0076
5	46.5	79	73.786	5.214	0.066
6	54.5	90.8	84.476	6.324	0.0696
7	67.5	98.4	101.373	2.973	0.0302
8	78.6	98.7	115.419	16.719	0.1693
9	90.2	115.8	129.786	13.986	0.1207

10	100.8	138.3	142.678	4.378	0.0316

Case 3: Right-hand Side Uncertainty Interval as firing interval

In this case the uncertainty considered only at the right hand side interval i.e. Firing interval

$$J_{P\,x} = \left[\underline{\mu}_{P_2}(x_i), \overline{\mu}_{P_2}(x_i)\right]$$

Foot print of uncertainty intervals for the μ_{P1} is [0.5724 to 0.9] Foot print of uncertainty intervals for the μ_{P2} is [0.6996 to 1.1].The more uncertainty is on the right hand side.

$$J_{Px} = \left[\underline{\mu}_P(x_i), \overline{\mu}_P(x_i)\right] = [0.6996, 1.1]$$

The Table 8 shows the Effort estimation using above firing intervals.

Table 8: Triangular Fuzzy Number of Adjusted Size and Effort Estimation Case-3

S.No	Size(m)	α = 0.6996m	M	β =1.1m	$a\alpha^b$	am^b	$a\beta^b$	Effort
1	2.1	1.469	2.1	2.31	3.747	5.081	5.511	5.222
2	3.1	2.168	3.1	3.41	5.221	7.081	7.681	7.278
3	4.2	2.938	4.2	4.62	6.765	9.174	9.95	9.428
4	12.5	8.745	12.5	13.75	17.14	23.241	25.208	23.887
5	46.5	32.531	46.5	51.15	52.516	71.208	77.234	73.187
6	54.5	38.128	54.5	59.95	60.125	81.525	88.424	83.791
7	67.5	47.223	67.5	74.25	72.151	97.831	106.11	100.55
8	78.6	54.988	78.6	86.46	82.147	111.386	120.812	114.482
9	90.2	63.103	90.2	99.22	92.373	125.252	135.851	128.733
10	100.8	70.519	100.8	110.88	101.548	137.693	149.345	141.52

The Table 9 shows the Measured Effort, Absolute Error, Estimated Effort and Relative Error.

Table 9: Error Calculations for Case-3

S.No	Size	Measured Effort	Estimated Effort	Absolute Error	Relative Error
1	2.1	5	5.222	0.222	0.0444
2	3.1	7	7.278	0.278	0.0397
3	4.2	9	9.428	0.428	0.0475
4	12.5	23.9	23.887	0.013	0.0005
5	46.5	79	73.187	5.813	0.0735
6	54.5	90.8	83.791	7.009	0.0771
7	67.5	98.4	100.55	2.15	0.0218
8	78.6	98.7	114.482	15.782	0.1598
9	90.2	115.8	128.733	12.933	0.1116
10	100.8	138.3	141.52	3.22	0.0232

Case 4: Left-hand Side Uncertainty Interval

The uncertainty in this case is considered only at the left hand side interval i.e. firing interval

$$J_{P\ x} = \left[\underline{\mu}_{P_1}(x_i), \overline{\mu}_{P_1}(x_i)\right]$$

Foot print of uncertainty intervals for the μ_{P1} is [0.5724 to 0.9] Foot print of uncertainty intervals for the μ_{P2} is [0.6996 to 1.1].The more uncertainty is on the left hand side.

$$J_{Px} = \left[\underline{\mu}_P(x_i), \overline{\mu}_P(x_i)\right] = [0.5724, 0.9]$$

The Table 10 shows the Effort estimation using above firing intervals.

Table 10: Triangular Fuzzy Number of Adjusted Size and Effort Estimation for Case-4

S.No	Size(m)	α = 0.5724m	M	β =0.9m	$a\alpha^b$	am^b	$a\beta^b$	Effort
1	2.1	1.202	2.1	1.89	3.158	5.081	4.645	4.781
2	3.1	1.774	3.1	2.79	4.4	7.081	6.473	6.663
3	4.2	2.404	4.2	3.78	5.702	9.174	8.386	8.633
4	12.5	7.155	12.5	11.25	14.445	23.241	21.245	21.871
5	46.5	26.616	46.5	41.85	44.26	71.208	65.093	67.012
6	54.5	31.195	54.5	49.05	50.672	81.525	74.523	76.721
7	67.5	38.637	67.5	60.75	60.808	97.831	89.429	92.067
8	78.6	44.99	78.6	70.74	69.233	111.386	101.82	104.823
9	90.2	51.63	90.2	81.18	77.852	125.252	114.495	117.872
10	100.8	57.697	100.8	90.72	85.584	137.693	125.867	129.58

The Table 11 shows the Measured Effort, Absolute Error, Estimated Effort and Relative Error.

Table 11: Error Calculations for Case-4

S.No	Size	Measured Effort	Estimated Effort	Absolute Error	Relative Error
1	2.1	5	4.781	0.219	0.0438
2	3.1	7	6.663	0.337	0.0481
3	4.2	9	8.633	0.367	0.0407
4	12.5	23.9	21.871	2.029	0.0848
5	46.5	79	67.012	11.988	0.1517
6	54.5	90.8	76.721	14.079	0.155
7	67.5	98.4	92.067	6.333	0.0643
8	78.6	98.7	104.823	6.123	0.062
9	90.2	115.8	117.872	2.072	0.0178
10	100.8	138.3	129.58	8.72	0.063

5. RESULTS AND DISCUSSIONS

One of the objective of the present work is to employ Interval Type-2 fuzzy logic for tuning the effort parameters and test its suitability for software effort estimation. This methodology is then tested using NASA dataset provided by Boehm. The results are then compared with the models in the literature such as Baily-Basili, Alaa F. Sheta and Harish.

Comparison with other models:

The Table 12 compares effort estimation of TSFC- Interval Type-2 Models with other available models. The resulting data indicate that the approximation accuracy of the type-2 fuzzy systems methodology which is used in this chapter is comparable with the Bailey-Basili, AlaaF. Sheta, Harish models. The fuzzy systems approach to effort estimation has an advantage over the other models as the Interval Type-2 fuzzy systems architecture determines the firing intervals for inputs which reduces the factors of uncertainty, and the fuzzy rules be extracted from numerical data, which may easily be analyzed and the implementation is also relatively easy.

Table 12: Effort Efforts in Man-Months of Various Models with Interval Type-2 Models

S.No	Size	Measured effort	Bailey – Basili Estimate	Alaa F. Sheta G.E.model Estimate	Alaa F. ShetaModel 2 Estimate	Harish model1	Harish model2	Interval Type-2 Model-I	TSFC Model 2			
									Case-I	Case-II	Case-III	Case-IV
1	2.1	5	7.226	8.44	11.271	6.357	4.257	4.394	4.822	5.265	5.222	4.781
2	3.1	7	8.212	11.22	14.457	8.664	7.664	6.124	6.721	7.337	7.278	6.663
3	4.2	9	9.357	14.01	19.976	11.03	13.88	7.933	8.707	9.506	9.428	8.633
4	12.5	23.9	19.16	31.098	31.686	26.252	24.702	20.099	22.06	24.082	23.887	21.871
5	46.5	79	68.243	81.257	85.007	74.602	77.452	70.489	67.591	73.786	73.187	67.012
6	54.5	90.8	80.929	91.257	94.977	84.638	86.938	80.701	77.385	84.476	83.791	76.721
7	67.5	98.4	102.175	106.707	107.254	100.329	97.679	96.842	92.863	101.373	100.55	92.067
8	78.6	98.7	120.848	119.27	118.03	113.237	107.288	110.26	105.73	115.419	114.482	104.823
9	90.2	115.8	140.82	131.898	134.011	126.334	123.134	123.986	118.891	129.786	128.733	117.872
10	100.8	138.3	159.434	143.0604	144.448	138.001	132.601	136.301	130.7	142.678	141.52	129.58

Assessment through Graph Representation of Measured Effort Vs Estimated Effort:

The Figure 6 shows measured effort Vs estimated effort of interval type-2 models and one can notice that the estimated efforts are very close to the measured effort.

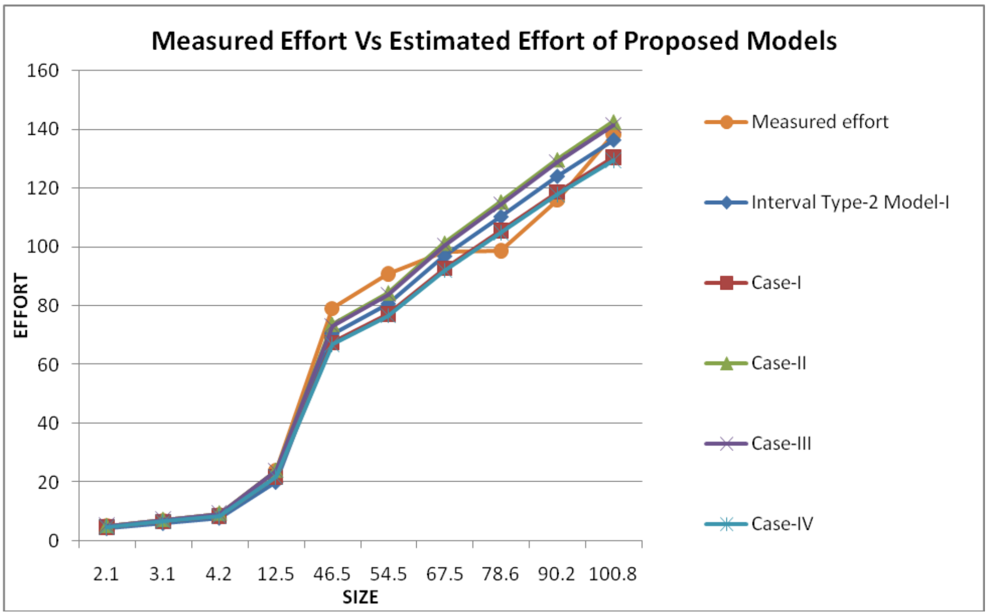


Figure 6: Measured Effort Vs Estimated Effort of TSFC models

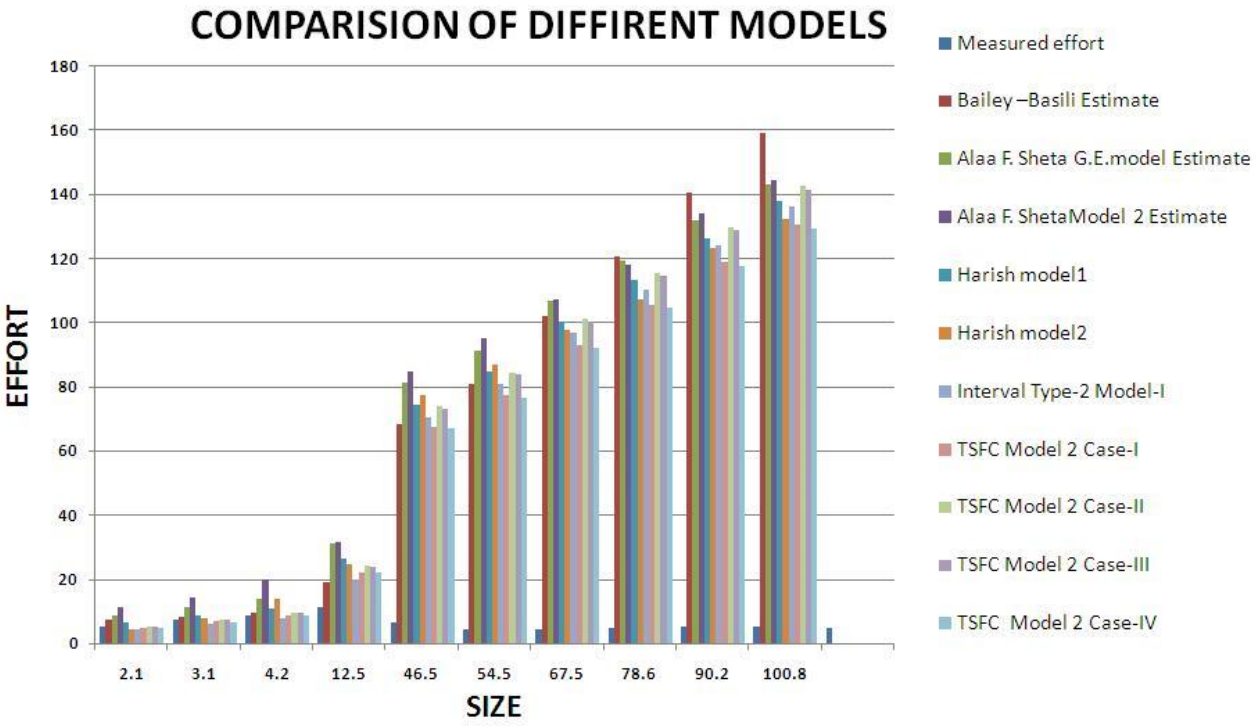


Figure 7: Effort Estimations of Various models Vs TSFC Models

PERFORMANCE ANALYSIS:

Parameters such as VAF, MARE, and VARE are employed to asses as well as to compare the performance of the estimation models. The integration of Takagi-Sugeno and Interval Type-2 fuzzy logic can be powerful tool when tackling the problem of effort estimation. It can be seen from the resulting data that the Fuzzy logic models for Effort estimation outperform the Baily-Basili, Alaa F. Sheta and Harish models. The computed MARE, VARE and VAF for all the models are indicated in Table 13.

Table 13 Summary Results of VAF, MARE and VARE

Model	Variance Accounted For (VAF%)	Mean Absolute Relative Error(MARE%)	Variance Absolute Relative Error(VARE%)
Bailey –Basili Estimate	93.147	17.325	1.21
Alaa F. Sheta G.E.Model Estimate	98.41	26.488	6.079
Alaa F. Sheta Model 2 Estimate	98.929	44.745	23.804
Harish model1	98.5	12.17	80.859
Harish model2	99.15	10.803	2.25
Interval Type-2 Model1	99.276	9.602	0.228
TSFC Model 2 Case-I	99.1	6.858	0.19
TSFC Model 2 Case-II	98.63	6.522	0.22
TSFC Model 2 Case-III	98.74	5.991	0.225
TSFC Model 2 Case-IV	98.98	7.312	0.21

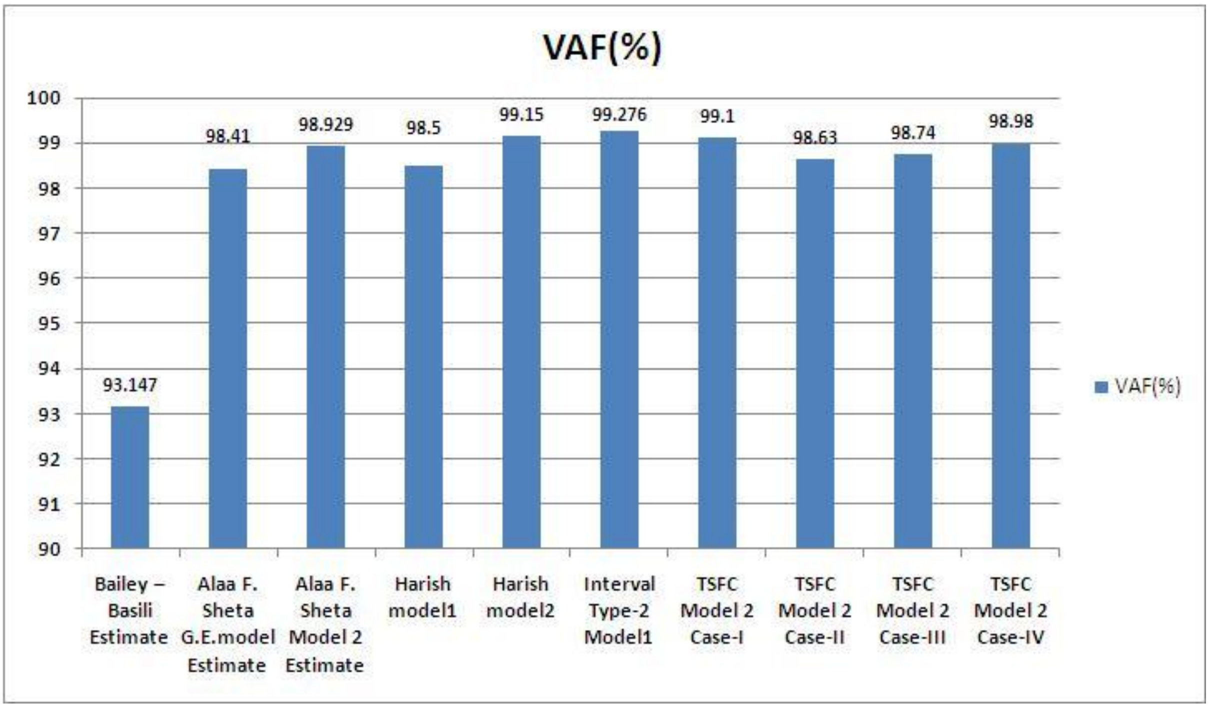


Figure 8: Variance Accounted For Of Various Models Vs TSFC Models

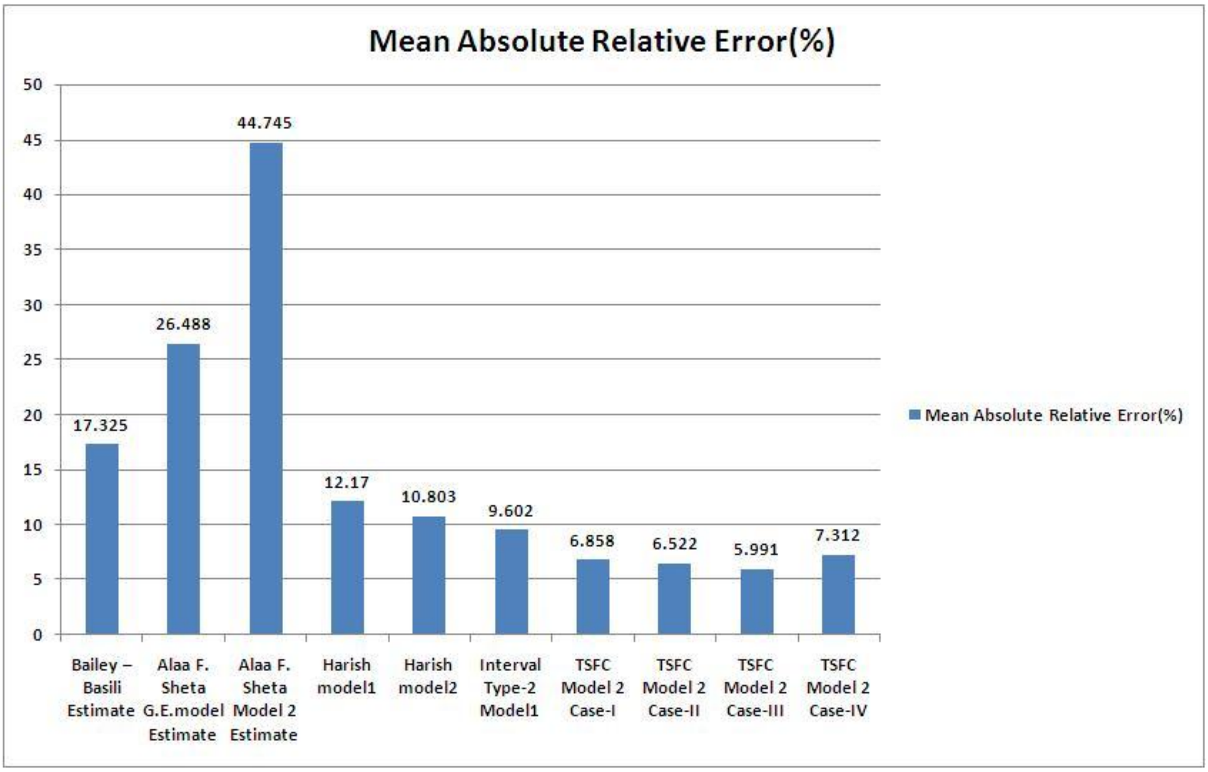


Figure 9: Mean Absolute Relative Error of Various models Vs TSFC Models

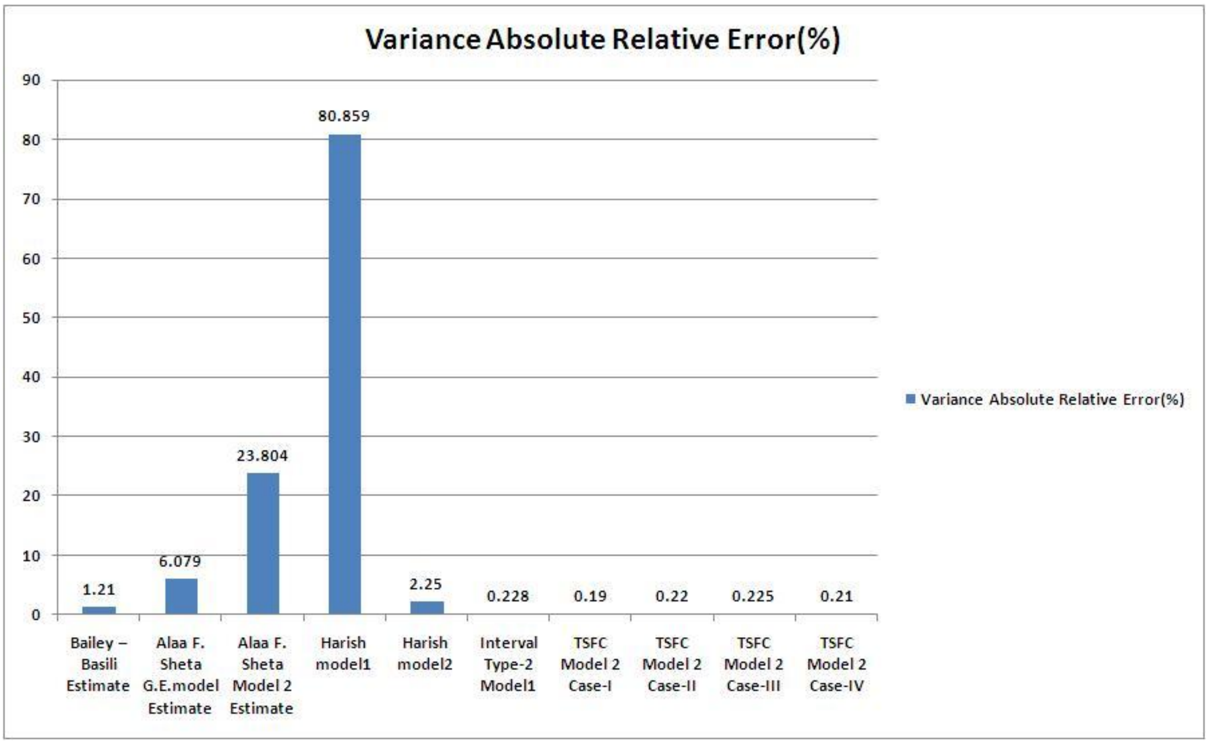


Figure 10: Variance Absolute Relative Error of various models Vs TSFC Models

6. CONCLUSION :

In this study we proposed new model structures to estimate the software cost (Effort) estimation. Interval Type-2 fuzzy sets is used for modeling uncertainty and impression to better the effort estimation. Rather than using a single number, the software size can be regarded as a fuzzy set yielding the cost estimate also in the form of a fuzzy set. These proposed models were able to provide good estimation capabilities as per the as per the experimental study taking parameters like VAF, MARE, and VARE. The work of Interval Type-2 fuzzy sets can be applied to other models of software cost estimation. However in these models only the size is used as input for estimating the effort. But there are so many "Cost Drivers" which have to be considered for measuring effort. In fact the main difficulty is to determine which cost driver really capture the reason for differences in estimated effort among the projects. Therefore for large projects of size>100 KDLOC the estimation process requires data to be more accurate, consistent with appropriate cost drivers. It is reasonable to assume that one should specify cost drivers for large projects as they are essential to calibrate the estimation model.

References

[1] Kim Johnson, Dept of Computer Science, University of Calgary, Alberta, Canada, "Software Cost Estimation: Metrics and Models" (pages 1 to 17).

[2] Hareton Leung,Zhang Fan Department of Computing The Hong Kong Polytechnic university," Software Cost Estimation".

[3] Robert W. Zmud. An Empirical Validation of Software Cost Estimation Models. Chris F.Kemerer, May 1987 ,Volume 30 Number 5.

[4] Krishna murthy and Douglas Fisha, "machine learning approaches to estimating software development effort", 1995.

[5] Lawrence H.putnam," A general empirical solution to the macro software sizing and estimating problem", 1978.

[6] Xishi Huang, Danny Ho, Jing Ren, Luiz F. Capretz,Improving the COCOMO model using a neuro-fuzzy approach. Applied Soft Computing 7 (2007) 29–40

[7] Petr Musflek ,Witold Pedrycz,Giancarlo Succi,Marek Reformat, "Software cost estimation with fuzzy, models",2000.

[8] Harish Mittal, Pradeep Bhatia"Optimization Criteria for Effort Estimation using Fuzzy Technique". CLEI ELECTRONIC JOURNAL, VOLUME 10, NUMBER 1, PAPER 2, JUNE 2007.

[9] Alaa F. Sheta," Estimation of the COCOMO Model Parameters Using Genetic Algorithms for NASA Software Projects" Journal of Computer Science 2 (2): 118-123, 2006, ISSN 1549-3636.

[10] Hao Ying, "The Takagi-Sugeno fuzzy controllers using the simplified linear control rules are nonlinear variable gain controllers", automatic, Vol:34,No.2,pp.157-167.

[11] Zehua Lv, Hai Jin, Pingpeng Yuan" The Theory of Triangle Type-2 Fuzzy Sets". IEEE Ninth International Conference on Computer and Information Technology.2009.

[12] Jerry M. Mendel, Jerry M. Mendel, Robert I. John]" Standard Background Material AboutInterval Type-2 Fuzzy Logic Systems That Can Be Used By All Authors".

[13] Shigeo Abe, Ming- Shong Lan,"Fuzzy rules extraction directly from numerical data for function approximation",IEEE transactions on systems,MAN,and cybernetics,Vol.25,no.1,1995.
[14] Jerry M. Mendel" Interval Type-2 Fuzzy Logic Systems Made Simple", 2006.
[15] Masaharu Mizumoto,Kokichi Tanaka"Fuzzy sets of type2 under algebraic product and algebraic sum", 2006
[16] Babak Rezaee]" A New Approach to Design of Interval Type-2 Fuzzy Logic Systems". Eighth International Conference on Hybrid Intelligent Systems..2008.
[17] Roberto Sepulveda, Patricia Melin, Antonio Rodriguez, Alejandra Mancilla, Oscar Montiel," Analyzing the effects of the Footprint of Uncertainty in Type-2 Fuzzy Logic Controllers",2006.
[18] Saikat Maity" Color Image Segmentation using Type-2 Fuzzy". International Journal of Computer and Electrical Engineering, Vol. 1, No. 3, August 2009.
[19] Juan R. Castro, Oscar Castillo, Luis G. Martínez" Interval Type-2 Fuzzy Logic Toolbox".2007.