Applying Fuzzy-AHP for software effort estimation in data scarcity

Sumeet Kaur Sehra [#1], Yadwinder Singh Brar [*2], Navdeep Kaur [#3]

[#1] *Research Scholar, I.K.G. P.T.U., Kapurthala, Punjab, India.*
[*2] *Professor, I.K.G. P.T.U., Kapurthala, Punjab, India.*
[#3] *Associate Professor, S.G.G.S.W.U., Fatehgarh Sahib, Punjab, India.*

Abstract- *Project managers and estimators have considered software effort estimation as a most challenging task. Vast research has been conducted for finding the best effort estimation model but it has been proved that none of these models is completely suitable for all environments and datasets. Expert judgement is most prevalent method for estimation but requires documented data for estimating the effort. in case of data scarcity, Analytic Hierarchy Process (AHP), a multi-criteria decision making approach inspired by the intelligent behaviour of human beings can be used effectively. But AHP suffers from inconsistency and rank reversal so fuzziness of decision maker can be incorporated by using Fuzzy-AHP (FAHP). The motive of this paper is to propose FAHP for predicting the effort of project in data scarcity. The effort of the projects is estimated with minimum single known project effort. The proposed method is validated using IVR dataset of real projects and results obtained show better accuracy as compared to other existing effort estimation models.*

Keywords— *Effort Estimation, Multi-criteria Decision Making, Expert Judgement, Analytic Hierarchy Process.*

I. INTRODUCTION

Software effort estimation is, as its name suggests, the task of estimating the amount of effort required to develop new software. Knowing the estimated effort of particular software project early in the development cycle is a valuable asset. Industry and academia have always considered a reliable and accurate estimate a challenging task. However, a review of estimation surveys by [1] documents that still less progress has been made in the area of estimation performance. Thus, there is a high demand for more research on the topic of effort estimation in software development projects.

Estimators use different methods for estimation. They employ a single technique (formal model or expert judgement) or both the techniques for estimation. Passing and Shephard [2] advocated that expert judgement is leading estimation method adopted by organisations. But still it is unpredictable to define whether expert judgement is better or weaker than formal models. But practitioners needs to identify the situations, when to use expert estimation and when to use formal models. Jorgensen et al. [3] advocated that formal models should be developed as support to expert judgement.

The predominant obstacle in the effective estimation is the absence of reliable and systematic historic data. The factors contributing are collection of data from different sources [4], value reduction of data with time. So in the absence of local data, project manager has either to choose state-of-practice approach such as expert judgement or algorithmic models such as COCOMO-II, SEL, Halstead, Walston-Felix and Bailey-Basili models etc. The researchers have concluded that these models tend to deviate outside a given environment [5].

Analytic hierarchy Process (AHP) is a multi-criteria decision making approach inspired by the intelligent behaviour of human beings. It combines the historical as well expert judgement by quantifying subjective judgement. It qualifies for effort estimation as it enables expert to view problem in a more structured and systematic way. But imprecision and subjectivity are not dealt in traditional AHP. Further, any change in the relative values of the choices results in changed weights causing a problem called rank reversal [6].

Fuzzy-AHP overcomes the limitations of AHP by incorporating the fuzziness involved while considering the relative importance of one element to another. Instead of using single crisp value, FAHP uses a range of values to incorporate decision makers uncertainty. In this paper FAHP has been applied to estimate the effort of a project in case of data scarcity.

The remainder paper is structured as follows. In second section related work has been presented. Third section discusses in detail about expert judgement, formal estimation models and FAHP. Fourth section discusses the results obtained after applying FAHP. In last section, conclusion and future scope is presented.

II. RELATED WORK

Various models using different techniques have been proposed for predicting the effort more accurately by numerous researchers [7]–[10]. Cartwright and

Sheppard [11] have suggested that no single technique performs in ideal manner for all environments and datasets. Menzies [12] has suggested that a hybrid approach should be followed for estimating by combining expert judgement with formal models.

SEE problem can be viewed as a multi-criteria problem as suggested by [13], [14]. The MCDM approach preferably used for effort estimation out of many available MCDM approaches is Analytic hierarchy process [15]. But AHP suffers from the issues related to imprecision and subjectivity in the pairwise comparison process. These issues are effectively handled by using Fuzzy AHP [16], [17]. In FAHP, a range of values is used in place of single crisp value to incorporate decision maker's uncertainty [18], [19]. From this range, the value depicting the confidence level of decision maker can be selected. FAHP has been used in various domains of software engineering including risk analysis and planning [17], [20], [21], quality evaluation [16], [22], software project selection [23]–[25], and assessment of testing adequacy criteria [26], [27].

III.METHODOLOGY

A. *Expert judgement*

In expert judgement method, experts having similar domain knowledge are consulted to estimate the effort of software project. The estimate values obtained from this method are based on the intuition of experts [28], [29]. Jorgensen [30] has supported expert judgement and stated that formal models should support expert judgement for producing more reliable estimates [28], [31].

B. *Existing algorithmic models*

Algorithmic effort estimation techniques involve the application of mathematical formulas derived based upon historical data. The most popular algorithmic estimation models [32] include:

1) COCOMO-II Model: This model is an extension of COCOMO Intermediate model and effort is calculated as:

$$\text{Effort} = 2.9 * (\text{KLOC})^{1.10} \qquad \text{……} (1)$$

2) Software Engineering Laboratory (SEL) Model: This model has been suggested by Software Engineering Laboratory (SEL) of the Maryland University. According to SEL model , the effort estimate is evaluated as follows:

$$\text{Effort} = 1.4 * (\text{Size})^{0.93} \qquad \text{……} (2)$$

3) Bailey-Basili Model: Model developed by Bailey-Basili gives the relationship between delivered lines of source code and effort as below:

$$\text{Effort} = 5.5 * (\text{KLOC})^{1.16} \qquad \text{……} (3)$$

C. *Existing algorithmic models*

Fuzzy AHP approach solves uncertainty involved in human decision-making [37], [38]. In FAHP, different number patterns can be used including triangular and trapezoidal numbers. Triangular fuzzy numbers have been used extensively by various researchers [15], [33]–[35]. The extent analysis method as suggested by [36] has been used for evaluating the eigen ventor for FAHP. FAHP is a hybrid approach applicable for both qualitative and quantitative criteria comparisons [39], [40]. In FAHP, decision making of the experts is represented by using a range of values rather than using discrete values. The membership function used for creating the fuzzy set is given in equation 4, where x is the weight of relative importance of one criterion over other criterion.

$$\mu_A(x) = \left\{ \begin{array}{ll} 0 & for \quad x \leq l \\ \frac{x-l}{m-l} & for \quad l \leq x \leq m \\ \frac{u-x}{u-m} & for \quad m \leq x \leq u \\ 0 & for \quad x \geq n \end{array} \right\} \quad \text{……} (4)$$

Triangular fuzzy numbers (TFN) provide an opportunity in deciding the weight of one alternative over the other. TFN is represented by equation 5.

$$a_{ij} = (l_{ij}, m_{ij}, u_{ij}) \qquad \text{……} (5)$$

Where l,m,u are pessimistic, moderate and optimistic values respectively.

The modified Saaty scale using TFN is given in table I. In FAHP table 1 is used for construction comparison matrix A= (aij) nxn where i, j = 1, 2, 3n. The next step is to use extent analysis method to calculate the relative ranking of alternatives, the synthetic extent values are obtained by equation 6.

$$S_i = \sum_{j=1}^{m} N_{ci}^j \otimes \left[\sum_{i=1}^{n} \sum_{j=1}^{m} N_{ci}^j \right]^{-1} \qquad \text{……} (6)$$

The degree of possibility of M1 $\geqslant$ M2 is defined in equation 7.

$$V(N_1 \geq N_2) = 1 \text{ if } n_{11} \geq n_{21} \qquad \text{……} (7)$$

$$V(N_2 \geq N_1) = hgt\left(N_1 \cap N_2\right) = \mu_{N_1}(d) \qquad \text{……} (8)$$

In equation 8, d is representing ordinate of the highest intersection point between between μN1 and μN2.

TABLE I: Linguistic Scale for Fuzzy AHP

Linguistic scale for importance	Fuzzy numbers for FAHP	Membership function	Domain	Triangular fuzzy scale (l, m, u)
Just Equal	1			(1.0, 1.0, 1.0)
Equally important	1	$\mu M(x) = (3-x)/(3-1)$	$1 \leq x \leq 3$	(1.0, 1.0, 3.0)
Weak importance over each another	3	$\mu M(x) = (x-1)/(3-1)$	$1 \leq x \leq 3$	(1.0, 3.0, 5.0)
		$\mu M(x) = (5-x)/(5-3)$	$3 \leq x \leq 5$	
Essential importance over each other	5	$\mu M(x) = (x-3)/(5-3)$	$3 \leq x \leq 5$	(3.0, 5.0, 7.0)
		$\mu M(x) = (7-x)/(7-5)$	$5 \leq x \leq 7$	
Very strong importance over other	7	$\mu M(x) = (x-5)/(7-5)$	$5 \leq x \leq 7$	(5.0, 7.0, 9.0)
		$\mu M(x) = (9-x)/(9-7)$	$7 \leq x \leq 9$	
Extreme importance over other	9	$\mu M(x) = (x-7)/(9-7)$	$7 \leq x \leq 9$	(7.0, 9.0, 9.0)
The value of second element in comparison to first would be by reciprocal of TFN given as (1/u1,1/m1,1/l1)				

The degree of possibility for a convex fuzzy number, is defined by equation 9.

$$V (N \geqslant N1,N2,...,Nk) = (N \geqslant N1),...,(N \geqslant Nk) = \min V (N \geqslant Ni) \qquad \text{…… (9)}$$

In order to normalize the weight vector, equation 10 is used.

$$W_A = \frac{W^T}{\sum(W^T)} \qquad \text{…… (10)}$$

IV. ESTIMATING EFFORT

A project is taken as reference point and remaining projects are ranked relative to the reference point. Triangular fuzzy numbers corresponding to the relative importance are generated by using Table I. The weights are evaluated using methodology of FAHP. The normalized weights are used to estimate the effort of a project as depicted in equation 11 in which E_i and E_k are estimated effort and known effort respectively, whereas w_i and w_k are corresponding weights.

$$E_i = \left(\frac{w_i}{w_k}\right) E_k \qquad \text{…… (11)}$$

V. RESULTS

The empirical validation of proposed AHP model is performed using dataset of Interactive Voice Response (IVR) applications from software industry [43] as shown in table III. The dataset consists of data related to 48 IVR projects, out of which 20 projects have been used for this study. LOC and actual effort for the projects is presented in the dataset as depicted in Table III. The equations 1, 2 and3 have been used for evaluating the effort for

algorithmic models including COCOMO, SEL and Bailey-Basili respectively.

A. *Performance measures*

1) Mean magnitude of relative error (MMRE): MMRE measures the difference between actual and estimated effort relative to the actual effort as shown in equation 12

$$MMRE = \frac{1}{N}\sum_{i=1}^{N} \frac{Actual\ Effort - Predicted\ Effort}{Actual\ Effort} \quad (12)$$

2) Root mean square error (RMSE): RMSE is a measure of the imperfection of the fit of the estimator to the data and is evaluated as depicted in equation 13.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Actual\ Effort - Predicted\ Effort)^2}$$
$$(13)$$

The value of effort estimated by using different existing models and FAHP is presented in Table II. From Table III, it is evident that values of MMRE and RMSE are and respectively for FAHP which is less than the values for other existing models. Figures 1 and 2 depict the comparison of models.

TABLE III: Comparison of MMRE and RMSE for different models

	SEL	COCO MO	Bailey-Basili	Fuzzy-AHP
MMRE	0.7365	0.245	0.602	0.0685
RMSE	29.460 43	10.600 68	23.661 35	2.52607

TABLE II: Effort Estimation of IVR Dataset using Different Models

| Proj No.	LOC	Actual Effort	SEL	COCOMO	Bailey-Basili	Fuzzy-AHP
1	16.2	86.1	18.66	59.59	139.12	Ref. Proj.
2	5.34	24.02	6.65	18.58	38.4	26.32
3	7.6	36.05	9.23	26.92	57.82	35.43
4	4.7	20.74	5.9	16.25	33.11	22.39
5	3.1	12.85	4.01	10.5	20.43	14.21
6	5.2	23.3	6.49	18.07	37.23	25.62
7	6.8	31.72	8.32	23.95	50.82	34.28
8	6.4	29.59	7.87	22.47	47.37	33.20
9	7.2	33.88	8.78	25.43	54.31	36.32
10	5.4	24.34	6.72	18.8	38.9	26.89
11	8.5	41.01	10.24	30.27	65.84	43.35
12	7.8	37.15	9.46	27.66	59.59	40.18
13	12.5	63.9	14.66	45.38	102.99	61.11
14	10.4	51.71	12.36	37.41	83.2	52.23
15	9.5	46.6	11.36	34.02	74.91	48.5
16	3.4	14.29	4.37	11.57	22.74	16.23
17	6.8	31.73	8.32	23.95	50.82	35.22
18	5.8	26.42	7.18	20.27	42.26	29.12
19	7.4	34.96	9.01	26.17	56.06	38.45
20	7.2	33.88	8.78	25.43	54.31	37.32
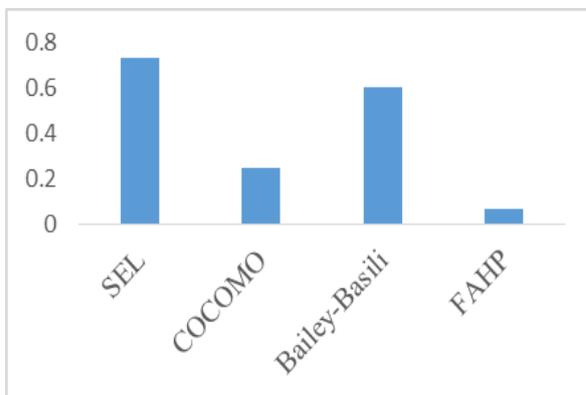
Figure 1: MMRE Comparison
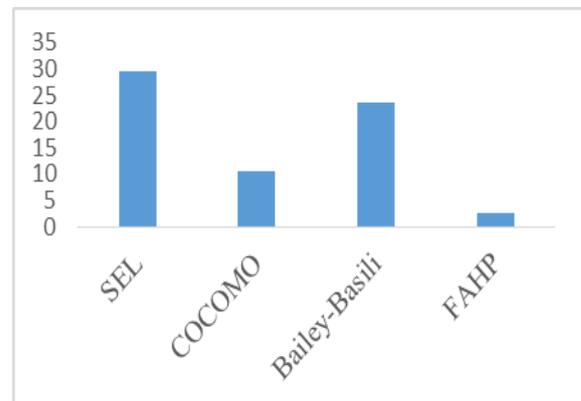
Figure 2: RMSE Comparison

VI.CONCLUSION

Software effort estimation has always been an area of concern to academia and industry. Although researchers have always focused on developing models for estimating the software effort more accurately, but it has been concluded that a single model does not conform to all datasets and environments. This limitation has created the scope of model development for data scarcity. FAHP has been effectively utilized to predict the effort of a project when only single point effort is known. By taking that project as reference point, effort of all other projects can be estimated using FAHP. FAHP generates the weights for the corresponding projects by ranking the projects relatively. The estimated value of effort using FAHP and other existing estimation models have been validated using IVR dataset. Comparison of MMRE and RMSE performance measures clearly depicts the dominance of FAHP over other models.

REFERENCES

[1] M. Jorgensen, "Selection of Effort Estimation Strategies," submitted to International Journal of Forecasting, 2008.

[2] U. Passing and M. Shepperd, "An experiment on software project size and effort estimation," in Proceedings. 2003 International Symposium on Empirical Software Engineering. IEEE, Sep. 2003, pp. 120–129.

[3] M. Jorgensen, B. Boehm, and S. Rifkin, "Software Development Effort Estimation: Formal Models or Expert Judgment?" IEEE Software, vol. 26, no. 2, pp. 14–19, Mar. 2009.

[4] G. Kadoda, M. Cartwright, L. Chen, and M. Shepperd, "Experiences using case-based reasoning to predict software project effort," in Proceedings of the EASE conference, Keele University, UK, 2000, pp.2–3.

[5] Y. Zhang, X. Zhang, X. Zhao, and T. Zhang, "Early effort estimation by AHP: A case study of project metrics in small organizations," in Computer Science and Automation Engineering (CSAE), 2012 IEEE International Conference on, vol. 1, May 2012, pp. 452–456.

[6] Y.-M. Wang and Y. Luo, "On rank reversal in decision analysis," Mathematical and Computer Modelling, vol. 49, no. 5-6, pp. 1221– 1229, Mar. 2009.

[7] R. de A. Araujo, A. L. I. de Oliveira, and S. C. B. Soares, "A´ morphological-rank-linear approach for software development cost estimation," in 21st IEEE International Conference on Tools with Artificial Intelligence, Newyork, USA, nov 2009, pp. 630–636.

[8] Amasaki and C. Lokan, "A replication study on the effects of weighted moving windows for software effort estimation," in 20th International Conference on Evaluation and Assessment in Software Engineering. Limerick, Ireland: ACM, jun 2016, p. 40.

[9] A. Sheta, D. Rine, and A. Ayesh, "Development of software effort and schedule estimation models using Soft Computing Techniques," in Evolutionary Computation, 2008. CEC 2008. (IEEE World Congress on Computational Intelligence). IEEE Congress on, 2008, pp. 1283–1289.

[10] B. Boehm and R. Valerdi, "Impact of software resource estimation research on practice: a preliminary report on achievements, synergies, and challenges," in 33rd International Conference on Software Engineering, Honolulu, USA, May 2011, pp. 1057–1065.

[11] M. Shepperd and M. Cartwright, "Predicting with Sparse Data," IEEE Transactions on Software Engineering, vol. 27, no. 11, pp. 987–998, Nov. 2001.

[12] T. Menzies, Z. Chen, J. Hihn, and K. Lum, "Selecting best practices for effort estimation," Software Engineering, IEEE Transactions on, vol. 32, no. 11, pp. 883–895, 2006.

[13] F. Ferrucci, C. Gravino, and F. Sarro, "How multi-objective genetic programming is effective for software development effort estimation?" in Search Based Software Engineering. Springer, 2011, pp. 274–275.

[14] S. Gu, R. Cheng, and Y. Jin, "Multi-objective ensemble generation," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 5, no. 5, pp. 234–245, 2015.

[15] M. B. Javanbarg, C. Scawthorn, J. Kiyono, and B. Shahbodaghkhan, "Fuzzy AHP-based multicriteria decision making systems using particle swarm optimization," Expert Systems with Applications, vol. 39, no. 1, pp. 960–966, Jan. 2012.

[16] J. S. Challa, A. Paul, Y. Dada, V. Nerella, P. R. Srivastava, and A. P. Singh, "Integrated Software Quality Evaluation: A Fuzzy Multi-Criteria Approach," Journal of Information Processing Systems, vol. 7, no. 3, pp. 473–518, Sep. 2011.

[17] M. Chang Lee, "Information Security Risk Analysis Methods and Research Trends: AHP and Fuzzy Comprehensive Method," International Journal of Computer Science and Information Technology, vol. 6, no. 1, pp. 29–45, Feb. 2014.

[18] Y.-M. Wang, Y. Luo, and Z. Hua, "On the extent analysis method for fuzzy AHP and its applications," European Journal of Operational Research, vol. 186, no. 2, pp. 735–747, Apr. 2008.

[19] A. Sarfaraz, P. Mukerjee, and K. Jenab, "Using fuzzy analytical hierarchy process (AHP) to evaluate web development platform," Management Science Letters, vol. 2, no. 1, pp. 253–262, Jan. 2012.

[20] F. Tysz and C. Kahraman, "Project risk evaluation using a fuzzy analytic hierarchy process: An application to information technology projects," International Journal of Intelligent Systems, vol. 21, no. 6, pp. 559–584, Jun. 2006.

[21] Wei, F. Dai, and j. Liu, "C2c E-commerce Risk Assessment Based on AHP and Fuzzy Comprehensive Evaluation," International Journal of Engineering and Manufacturing, vol. 1, no. 1, pp. 34–39, Feb. 2011.

[22] R. Kohli and S. K. Sehra, "Fuzzy Multi Criteria Approach for Selecting Software Quality Model," International Journal of Computer Applications, vol. 98, no. 11, pp. 11–15, 2014.

[23] T. Bakshi, B. Sarkar, and S. K. Sanyal, "A Novel Integrated AHP-QFD Model for Software Project Selection under Fuzziness," International Journal of Computer Applications (09758887), c, 2012.

[24] Y. Y. Jusoh, K. Chamili, N. C. Pa, and J. H. Yahaya, "Open source software selection using an analytical hierarchy process (AHP)," American Journal of Software Engineering and Applications, vol. 3, no. 6, pp. 83–89, 2014.

[25] M. Khan, A. Parveen, and M. Sadiq, "A method for the selection of software development life cycle models using analytic hierarchy process," in Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 International Conference on, Feb. 2014, pp. 534–540.

[26] P. R. Srivastava and M. P. Ray, "Multi-attribute Comparison of Automated Functional and Regression Testing Tools using Fuzzy AHP." in IICAI, 2009, pp. 1030–1043.

[27] V. Belton and T. Stewart, Multiple criteria decision analysis: an integrated approach. Springer Science & Business Media, 2002.

[28] M. Jrgensen, "A review of studies on expert estimation of software development effort," Journal of Systems and Software, vol. 70, no. 1, pp. 37–60, 2004.

[29] M. Jorgensen, "Forecasting of software development work effort: Evidence on expert judgement and formal models," International Journal of Forecasting, vol. 23, no. 3, p. 449462, Jul 2007.

[30] M. Jorgensen, B. Boehm, and S. Rifkin, "Software Development Effort Estimation: Formal Models or Expert Judgment?" Software, IEEE, vol. 26, no. 2, pp. 14–19, Mar. 2009.

[31] M. Jorgensen and T. Halkjelsvik, "The effects of request formats on judgment-based effort estimation," Journal of Systems and Software, vol. 83, no. 1, pp. 29 – 36, 2010, sI: Top Scholars.

[32] D. K. Srivastava, D. S. Chauhan, and R. Singh, "Vrs model: A model for estimation of efforts and time duration in development of ivr software system," International Journal of Software Engineering, vol. 5, no. 1, pp. 27–46, 1 2012.

[33] A. Sarfaraz, P. Mukerjee, and K. Jenab, "Using fuzzy analytical hierarchy process (AHP) to evaluate web development platform," Management Science Letters, vol. 2, no. 1, pp. 253–262, Jan. 2012.

[34] J. Agarwal, R. Nagpal, and R. Sehgal, "Reliability of Component based Software System using Fuzzy AHP and Fuzzy TOPSIS." International Journal of Computer Applications, vol. 106, 2014.

[35] S. K. Dubey, A. Mittal, and A. Rana, "Measurement of object-oriented software usability using fuzzy AHP," International Journal of Computer Science and Telecommunications, vol. 3, pp. 98–104, 2012.

[36] C.-H. Cheng and D.-L. Mon, "Evaluating weapon system by analytical hierarchy process based on fuzzy scales," Fuzzy Sets and Systems, vol. 63, no. 1, pp. 1 – 10, 1994.

[37] L. Mikhailov and P. Tsvetinov, "Evaluation of services using a fuzzy analytic hierarchy process," Applied Soft Computing, vol. 5, no. 1, pp. 23–33, 2004.

[38] L. Liu, H. Chen, and R. Zhang, "Comprehensive evaluation of examination quality based on fuzzy ahp,"

TELKOMNIKA Indonesian Journal of Electrical Engineering, vol. 11, no. 9, pp. 5384–5394, 2013.

[39] C.-N. Liao, "Fuzzy analytical hierarchy process and multi-segment goal programming applied to new product segmented under-price strategy," Computers & Industrial Engineering, vol. 61, no. 3, pp. 831–841, Oct. 2011.

[40] Y.-C. Tang, M. J. Beynon, and others, "Application and development of a fuzzy analytic hierarchy process within a capital investment study," Journal of Economics and Management, vol. 1, no. 2, pp. 207–230, 2005.

[41] S. K. Sehra, Y. S. Brar, and N. Kaur, "Multi Criteria Decision Making Approach for Selecting Effort Estimation Model," International Journal of Computer Applications, vol. 39, no. 1, pp. 10–17, 2012.

[42] Y.-M. Wang, Y. Luo, and Z. Hua, "On the extent analysis method for fuzzy AHP and its applications," European Journal of Operational Research, vol. 186, no. 2, pp. 735–747, Apr. 2008.