

SOFTWARE COST ESTIMATION WITH FUZZY INPUTS: FUZZY MODELLING AND AGGREGATION OF COST DRIVERS

MIGUEL-ÁNGEL SICILIA, JUAN-J. CUADRADO-GALLEGO, JAVIER CRESPO
AND ELENA GARCÍA-BARIOCANAL

Parametric software cost estimation models are well-known and widely used estimation tools, and several fuzzy extensions have been proposed to introduce a explicit handling of imprecision and uncertainty as part of them. Nonetheless, such extensions do not consider two basic facts that affect the inputs of software cost parametric models: cost drivers are often expressed through vague linguistic categories, and in many cases cost drivers are better expressed in terms of aggregations of second-level drivers. In this paper, fuzzy set elicitation techniques are used as a tool to model vague categories expressing cost driver quantities, focusing on two well-known COCOMO cost drivers. The results clearly indicate that such fuzzy set modelling approach affects significantly the estimation outcomes. In addition, the empirical adjustment of the DOCU cost driver as an aggregation of second-level documentation artifact measures is used to illustrate the modelling of flexible aggregation in the context of parametric estimation. Fuzzy set elicitation and aggregation operator modelling combined provide a novel approach to extending fuzzy parametric models for software estimation, which can be used as a complement to existing approaches.

Keywords: software cost estimation, fuzzy set elicitation, aggregation operator design

AMS Subject Classification: 68U35, 68T37, 03B52, 47S40, 28E10

1. INTRODUCTION

Parametric models are one of the principal software cost estimation techniques [3], used both in commercial tools like SLIM¹ and also in open, published models like the different versions of COCOMO [2]. These models essentially use mathematical expressions – obtained usually from conventional curve regression techniques – to derive effort of development estimates from a number of input variables that are often called *cost drivers*. Parametric models are dependent to some extent on the domain of development and vary with the technology used to support the Software Engineering process, so that it is currently accepted that no single model might ever fit every estimation situation [9]. In consequence, the problem of estimation

¹<http://www.qsm.com/>

4. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

Cost drivers in parametric software cost estimation are often expressed through linguistic assessments and they usually represent high-level concepts for which a single, precise measurement scale is not available. This motivates the use of fuzzy techniques to model estimation inputs and their assessment procedures. The use of common membership function elicitation techniques for RELY and AEXP has been described as an approach to analyze the human linguistic categories used to express cost drivers, resulting in evidence pointing out that existing crisp rating values are not adequate as a model for such expressions. In addition, the empirical adjustment of the aggregation process of second-level cost indicators into first-level cost driver inputs has been described using the DOCU cost driver. An empirically adjusted OWA operator has been demonstrated to be a better model than rating level selection rules and than the arithmetic mean.

Future work should address a more systematic and exhaustive experimental account of cost drivers, and the introduction of models of interaction between second-level cost drivers to obtain more detailed models of cost drivers. It should be noted that the modelling approaches described are not only useful for effort estimation but for general software assessment and benchmarking.

(Received May 15, 2004.)

REFERENCES

- [1] T. Bilgiç and T. Türksen: Measurement of Membership Functions: Theoretical and Empirical Work. In: Handbook of Fuzzy Sets and Systems (D. Dubois and H. Prade, eds.), Vol. 1, Chapter 3, Fundamentals of Fuzzy Sets, Kluwer 1999, pp. 195–232.
- [2] B. W. Boehm: Software Engineering Economics. Prentice-Hall, Englewood Cliffs, NJ 1981.
- [3] B. Boehm, C. Abts, and S. Chulani: Software Development Cost Estimation Approaches – A Survey. Technical Report USC-CSE-2000-505, Center for Software Engineering, University of California 2000.
- [4] S. Chulani, B. Clark, B. Boehm, and B. Steece: Calibration approach and results of the COCOMO II post – architecture model. In: Proc. 20th Annual Conference of the International Society of Parametric Analysts (ISPA) and 8th Annual Conference of the Society of Cost Estimating and Analysis (SCEA), 1998.
- [5] J. J. Crespo, M. A. Sicilia, and J. J. Cuadrado: On fuzzy regression in software cost estimation models. In: Proc. 2003 ACM-IEEE Internat. Symposium on Empirical Software Engineering (ISESE'03).
- [6] J. J. Crespo, M. A. Sicilia, and J. J. Cuadrado: On the use of fuzzy regression in parametric software estimation models: Integrating imprecision in COCOMO cost drivers. WSEAS Trans. on Systems 1 (2004), 3, 96–101.
- [7] J. J. Crespo, M. A. Sicilia, and J. J. Cuadrado: On aggregating second-level software estimation cost drivers: A usability cost estimation case study. In: Proc. Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU04).
- [8] J. J. Cuadrado-Gallego, O. Marbán, A. Amescua, L. García, and M. Sánchez: The importance of rating level selection method to obtain accurate estimations in parametric mathematical models. J. Cost Analysis and Management, Summer (2004), 14–20.
- [9] J. J. Dolado: On the problem of the software cost function. Information & Software Technology 43 (2001), 1, 61–72.

- [10] European Space Agency – Board for Software Standardisation and Control. 1991. ESA PSS-05-0 Software Engineering Standards Issue 2.
- [11] B. Ghahramani: Software reliability analysis: a systems development model. *Computers & Industrial Engrg.* 45 (2003), (2), 295–305.
- [12] A. Idri, A. Abran, and T.M. Khoshgoftaar: Fuzzy analogy: A new approach for software cost estimation. In: *Current Trends in Software Measurement* (Dumke and Abran, eds.), Shaker Publ., Aachen 2001, pp. 127–142.
- [13] B. Izyumov, E. Kalinina, and M. Wagenknecht: Software tools for regression analysis of fuzzy data. In: *Proc. 9th Zittau Fuzzy Colloquium*, Zittau 2001, pp. 221–229.
- [14] E. Kalinina and M. Wagenknecht: Fuzzy regression analysis and application to a crisp model. In: *Proc. 8th Zittau Fuzzy Colloquium*, Zittau 2000, pp. 9–18.
- [15] J. A. McCall, P. K. Richards, and G. F. Walters: *Factors in Software Quality*, Vol. 1–3. AD/A 049-014/015/055. Springfield 1977, VA: National Technical Information Service.
- [16] P. Musilek, W. Pedrycz, G. Succi, and M. Reformat: Software cost estimation with fuzzy models. *Appl. Comput. Rev.* 8 (2000), 2, 24–29.
- [17] NASA Software Measurement Guide Book. Revision 1. 1995. Software Engineering Laboratory Series, NASA-GB-001-94.
- [18] D. Palomar and M. A. Sicilia: Web page usability analysis based on vague perceptual concepts. In: *Proc. IADIS WWW Conference 2004*, to appear.
- [19] E. Rosch: *Principles of Categorization*. Readings in Cognitive Science, Erlbaum 1988, pp. 312–322.
- [20] M. A. Sicilia, E. García, and T. Calvo: An enquiry-based method for Choquet integral-based aggregation of interface usability parameters. *Kybernetika* 39 (2003), 601–614.
- [21] O. Souza Lima Jr., P. P. M. Farias, and A. D. Belchior: Fuzzy function point analysis. In: *Proc. 4th European Conference on Software Measurement and ICT Control*, Heidelberg 2001, pp. 161–172.
- [22] M. Van Welie, G. C. van der Veer, and A. Eliëns: Breaking down usability. In: *Proc. of Interact'99*, pp. 613–620.
- [23] Z. Xu, T. M. Taghi, and M. Khoshgoftaar: Identification of fuzzy models of software cost estimation. *Fuzzy Sets and Systems* 145 (2004), 1, 141–163.

*Miguel-Ángel Sicilia and Elena García-Barriocanal, Department of Computer Science, University of Alcalá, Ctra. Barcelona km.33.6 – 28871 Alcalá de Henares. Madrid. Spain.
e-mail: msicilia@uah.es, elena.garciab@uah.es*

*Juan-J. Cuadrado-Gallego, Department of Computer Science, University of Valladolid, Plaza de Santa Eulalia, 9 – 40005 Segovia. Spain.
e-mail: jjcg@infor.uva.es*

*Javier Crespo, Department of Computer Science, Complutense University, C/. Juan del Rosal, 8 – 28040 Madrid. Spain.
e-mail: jcrespoy@ucmail.ucm.es*