Software Qual J (2006) 14:339–357
DOI 10.1007/s11219-006-0039-0

Enhancing input value selection in parametric software cost estimation models through second level cost drivers

Juan J. Cuadrado-Gallego · Luis Fernández-Sanz · Miguel-Ángel Sicilia

Abstract Parametric cost estimation models are widely used effort prediction tools for software development projects. These models are based on mathematical models that use as inputs specific values for relevant cost drivers. The selection of these inputs is, in many cases, driven by public prescriptive rules that determine the selection of the values. Nonetheless, such selection may in some cases be restrictive and somewhat contradictory with empirical evidence, in other cases the selection procedure is somewhat subject to ambiguity. This paper presents an approach to improve the quality of the selection of adequate cost driver values in parametric models through a process of adjustment to bodies of empirical evidence. The approach has two essential elements. Firstly, it proceeds by analyzing the diverse factors potentially affecting the values a cost driver input might adopt for a given project. And secondly, an aggregation mechanism device for the selection of input variables based on existing data is explicitly devised. This paper describes the rationale for the overall approach and provides evidence of its appropriateness through a concrete empirical study that analyses the COCOMO II DOCU cost driver.

Keywords Parametric estimation models · Cost drivers · Software projects · COCOMO II · Empirical adjustment

J. J. Cuadrado-Gallego · M.-Á. Sicilia
E.U. Politécnica, Campus Universitario, Carretera Madrid-Barcelona, Km. 33,600, C.P. 28871, Alcalá de Henares (Madrid)
e-mail: jjcg@uah.es

M.-Á. Sicilia
e-mail: msicilia@uah.es

L. Fernández-Sanz ( )
Dep. Sistemas Informaticos, Universidad Europea de Madrid, C/Tajo s/n, 28670 Villaviciosa de Odon, Madrid, Spain
e-mail: luis.fernandez@uem.es

factor in the current project is selected using the table. The second step requires that the numerical values associated to the rating selected for each cost driver should be identified, (usually using a different table but, in same cases, it could be the same one, e.g. Table 6.)

As an example, we may consider that the quantity of documentation developed is excessive for our current project according to the software lifecycle needs. Tf we use Table 2, we obtain a rating level High and the first step is completed. Tn step two we consult Table 6 and obtain a value of 1.11.

Once numerical values for the cost drivers and scale factor have been determined, they are introduced in the Equation C.1 to obtain the value for E, the effort estimated for the project. Tn the example, the multiplier for DOCU is $x_4 = 1.11$ to be multiplied by the other 16 cost drivers, by A and the value of the size powered to the numerical value obtained after solving the scale factors expression.

A more extended explanation of how COCOMO TT Post-Architecture model works can be found in Boehm et al. (2000).

References

Baik, J., Boehm, B., Steece, B. 2000. The Effect of CASE Tools on Software Development Effort. Proceedings of the 15th Tnternational Forum on COCOMO and Software Cost Estimation Cost drivers Studies. Los Angeles, California (USA). 24–27/10.

Baik, J., Boehm, B., Steece, B. 2002. Disaggregating and calibrating the CASE tool variable in COCOMO TT. TEEE Trans. Software Eng. **28**(11) pp. 1009–1022.

Baylei, J., Basili. V. 1981. A Meta-Model for Software Development Resource Expenditures. Tn Proceedings of the Fifth Tnternational Conference on Software Engineering, pp. 107–116.

Boehm, B.W. 1981. Software Engineering Economics. Prentice Hall.

Boehm, B., Abts, C., Chulani, S. 2000. Software Development Cost Estimation Approaches-A Survey. Center for Software Engineering, University of California, Technical Report USC-CSE-2000–505.

Boehm, B., Clark, B., Horowitz, E., Madachy, R., Selby, R., Westland, C. 1995. Cost Model for Future Software Life Cycle Processes: COCOMO 2.0. Annals of Software Engineering. Special Volume on Software Process and Product Measurement. J. Arthur, S. Henry and J. Baltzer (eds.). Amsterdam: AG Science Publishers, (1): 45–60.

Boehm, B., Abts, C., Winsor B., A., Chulani, S., Clark, B., Horowitz, E., Madachy, R., Reifer, D., Steece, B. 2000. Software Cost Estimation with Cocomo TT. Prentice Hall.

Chulani, S., Clark, B., Boehm, B., Steece, B. 1998. Calibration Approach and Results of the COCOMO TT Post–Architecture Model. Tn Proceedings of the 20th Annual Conference of the Tnternational Society of Parametric Analysts (TSPA) and the 8th Annual Conference of the Society of Cost Estimating and Analysis (SCEA).

Chulani, S., Boehm, B., Steece, B. 1999. From Multiple Regression to Bayesian Analysis for COCOMO TT. Tn: Proceedings of the 21st Annual Conference of the Tnternational Society of Parametric Analysts (TSPA) and the 9th Annual Conference of the Society of Cost Estimating and Analysis (SCEA).

Chulani, S., Boehm, B., Steece, B. 1999. Bayesian analysis of empirical software engineering cost models. TEEE Transactions on Software Engineering, **25**(4):513–583.

Cuadrado-Gallego, J.J., Ernica E., Sánchez, M., Guzmán, J., Amescua, A. 2000. The cost estimator DOCU: An empirical and theoretical study. Proceedings of the 15th TnternationalForum on COCOMO and Software Cost Estimation Cost Drivers Studies. Los Angeles, California (USA), pp. 24–27/10.

Cuadrado-Gallego, J.J., Marbán, O., Sánchez, M., Garcín, L. (2005) The importance of rating level selection for input variables to determine accurate estimations in parametric mathematical models. Journal of Cost Analysis and Management, Summer 2004, pp. 12–24.

Crespo, J., Sicilia, M.A., Cuadrado, J.J. 2004. On Aggregating Second-Level Software Estimation Cost Drivers: A Usability Cost Estimation Case Study. Proceedings of TPMU, pp. 1255–1260.

Department of Defense (DOD), United States 1999. Parametric estimating handbook, 2nd Edition.

ESA, 1991. European Space Agency. Software Engineering Standards, ESA PSS-05-0 Tssue 2.

Farr, L., Zagorski, H. 1965. Quantitative Analysis of Programming Cost Factors: A Progress Report. Tn A. Frielink (ed.): Proceedings of the TCC Symposium on Economics of Automatic Data Processing. Amsterdam: North-Holland, Holland.
Ferens, D., Christensen, D. 1999. Calibrating software cost models to department of defense databases — a review of ten studies. Journal of Parametrics, **XIV**(1):33–52.
Fischman, L. 1997. Calibrating a Software Evaluation Model. Tn: Proceedings of the ARMS Conference.
Herd, J., Postak, J., Russell, W., Stewart, K. 1977. Software Cost Estimation Study – Study Results. Final Technical Report, RADC-TR-77-220. Doty Associates, Tnc.
Jensen, R. 1983. An TmprovedMacrolevel Software Development Resource Estimation Model. Tn: Proceedings of the 5th TSPA Conference, 88–92.
Linstone, H., Turoff, M., (eds.). 2002. The Delphi Method: Techniques and Applications. Web edition, available at http://www.is.njit.edu/pubs/delphibook/.
McCall, J., et al. 1977. Factors in software quality. Vol. 1,2,3. AD/A-049-014/015/055. Nat. Tech. Tnf. Service. Springfield.
Mertes, K., Ferens, D., Christensen. 1999. An empirical validation of the checkpoint software cost estimation model. Journal of Cost Analysis and Management, 35–44.
NASA, 1990. Manager's Handbook for Software Development. Revision 1. Software Engineering Laboratory Series. SEL-84-101.
NASA, 1995. Software Measurement Guide Book. Revision 1. Software Engineering Laboratory Series. NASA-GB-001–94.
NASA, 1996. Software Process Tmprovement Book. Revision 1. Software Engineering Laboratory Series. NASA-GB-001-95.
Nielsen, J. 1993. Usability Engineering. Morgan Kaufmann. Morgan Kaufmann Publishers, San Francisco, USA.
Parametric Estimating Tnitiative (PET) 1999. Parametric Estimating Handbook. 2nd edition.
Prather, P. 1995. Design and analysis of hierarchical software metrics. ACM Computing Surveys, **27**(1): 497–518.
Putnam, L. 1978. A general empirical solution to the macro software sizing and estimation problem. TEEE Transactions on Software Engineering, **4**:345–361.
Putnam, L., Mayers, W. 1992. Measures for Excellence. Reliable Software on Time, Within Budget. Englewood Cliffs, NJ: Yourdon Press.
Rubin, H. 1983. Macroestimation of Software Development Parameters: the Estimacs System. Tn: Proceedings of the SOFTFATR Conference Development Tools, Techniques and Alternatives. Arlington: TEEE Press.
Shrum, T. 1997. Calibration and Validation of the CHECKPOTNTModel to the Air Force Electronic Systems Center Software Database. Unpublished masters thesis. Dayton, OH, Air Force Tnstitute of Technology.
Sicilia, M., Cuadrado-Gallego, J.J., Crespo, J., García-Barriocanal E. 2005. Software cost estimation with fuzzy inputs: Fuzzy modelling and aggregation of cost drivers. Kybernetika, **41**(2):249–264.
Van Welie, M., van der Veer, G.C., Eliëns A. 1999. Breaking down usability: Tn: Proceedings of Tnteract 99, pp. 613–620.
Yager, R.R. 1988. Ordered weighted averaging aggregation operators in multi-criteria decision making. TEEE Trans. on Systems, Man and Cybernetics, **18**:183–190.

Juan J. Cuadrado-Gallego obtained his Ph. D. degree in Computer Science from the Carlos TTT University in 2001. From 2001 to 2004 he worked as a full-time lecturer at the Carlos TTT University, after which he joined the Universities of Valladolid and Alcalá. Also

he has taught lectures on Software Engineering at the foreign University of Rome Tre in Rome, Ttaly. His research interests are primarily in the areas of Software Engineering and Software Metrics. He is the Software Engineering leader of the Tnformation Engineering Research Unit at the University of Alcalá and collaborates as editorial board member in national and international journals related to Software Engineering and Software Metrics.

Luis Fernández-Sanz obtained degree in Computer Science from the Technical University of Madrid (Spain) in 1988 and a Ph. D. degree in Computer Science from the University of the Basque Country in 1997 (with a extraordinary award mention). From 1989 to 1996 he worked as a full-time lecturer at Technical University of Madrid, after which he joined Universidad Europea de Madrid where, since 2000, is director of the department of Computer Systems. His research interests are primarily related to the areas of Software Engineering and Quality as well as educational methods and TT labour market analysis. Since 2000 is the chairman of the Software Quality Group of ATT (the main Spanish association of TT professionals) in charge of the Spanish national conference on Software Quality and Tnnovation (JTCS). He is acted as editor, guest editor or editorial board member in different national and international journals (RETCTS, NOVATTCA, UPGRADE, etc.) in the area of Software Engineering.

Miguel-Ángel Sicilia obtained a Ms.C. degree in Computer Science from the Pontifical University of Salamanca, Madrid (Spain) in 1996 and a Ph.D. degree from the Carlos TTT University in 2003. He worked as a software architect in e-commerce consulting firms, being part of the development team of a Web personalization framework at Tntelligent Software Components (iSOCO). From 2002 to 2003 he worked as a full-time lecturer at the Carlos TTT University, after which he joined the University of Alcalá. His research interests are primarily in the areas of adaptive hypermedia and Web personalization, learning technology, and Software Engineering. He is the leader of the ATS STG on Reusable Learning Objects and collaborates as editorial board member in several international journals related to learning technology, metadata and the Semantic Web.