Calibration Approach and Results of the COCOMO II Post-Architecture Model

Sunita Devnani-Chulani, Brad Clark, Barry Boehm
({sdevnani, bkclark, boehm }@sunset.usc.edu)
Bert Steece
(berts@almaak.usc.edu)

1 INTRODUCTION

COCOMO II is an effort to update software cost estimation models, such as the 1981 **CO**nstructive **CO**st **MO**del and its 1987 Ada COCOMO update. Both these and other 1980's cost models have experienced difficulties in estimating software projects of the 90s due to new practices such as non-sequential and rapid-development process models; reuse-driven approaches involving commercial-off-the-shelf (COTS) packages, reengineering, applications composition, and application generation capabilities; object-oriented approaches supported by distributed middleware; software process maturity effects and process-driven quality estimation. The COCOMO II research effort has developed new functional forms reflecting these practices, and is concentrated on developing a model well-suited for the 1990s and then annually updating it for the forthcoming years of the 21st Century.

The initial definition of the COCOMO II model and its rationale are described in [Boehm95]. The model has three sub-models, Applications Composition; Early Design and Post-Architecture, which can be combined in various ways to deal with the current and likely future software practices marketplace.

The Post-Architecture Model, as the name suggests, is typically used after the software architecture is well defined and established. It estimates for the entire development life cycle of the software product and is a detailed extension of the Early-Design model. This model is the closest in structure and formulation to the Intermediate COCOMO '81 and Ada COCOMO models. It uses Source Lines of Code and/ or Function Points for the sizing parameter, adjusted for reuse and breakage; a set of 17 effort multipliers and a set of 5 scale factors, that determine the economies/diseconomies of scale of the software under development. The 5 scale factors replace the development modes in the COCOMO '81 model and refine the exponent in the Ada COCOMO model.

The Post-Architecture Model has been calibrated to a database of 83 projects collected from Commercial, Aerospace, Government and FFRDC organizations. This paper describes the calibration process incorporated and summarizes the results obtained. The Early Design Model calibration is obtained by aggregating the calibrated Effort Multipliers of the Post-Architecture Model as described in [USC-CSE1]. The Scale Factor calibration is the same in both the models. The Applications Composition Model has not yet been calibrated due to limitations in the availability of project data.

This paper describes our experience and results of the first calibration of the Post-Architecture model. The model determination process began with an expert Delphi process to determine apriori values for the Post-Architecture model parameters. A dataset of 83 projects was used in the multiple regression analysis. Projects with missing data or unexplainable anomalies were dropped. Model parameters that exhibited high correlation were consolidated. Multiple regression analysis was used to produce coefficients. These coefficients were used to adjust the previously assigned expert-determined model values. Stratification was used to improve model accuracy.

The resulting model produced estimates within 30% of the actuals 52% of the time for effort. Stratification by organization resulted in a model that produced estimates within 30% of the actuals 64% of the time for effort. It is therefore recommended that organizations using the model calibrate it using their own data. This increases model accuracy and produces a local optimum estimate for similar type projects.

Section 2 of this paper describes the data used for calibration. Section 3 describes the calibration procedures, results, and future calibration strategy.

2 DATA BASE USED FOR CALIBRATION

Data was recorded on a data collection form that asked between 33 and 59 questions depending on the degree of source code reuse [USC-CSE2]. The data collected was historical, i.e. observations were completed projects. As a baseline for the calibration database, some of the COCOMO 1981 projects and Ada COCOMO projects were converted to COCOMO II data inputs. The total observations used in the calibration was 83, coming from 18 different organizations.

This dataset formed the basis for an initial calibration.

Overall, the 83 data points ranged in size from 2 to 1,300 KSLOC, in effort from 6 to 11,400 person months, and in schedule from 4 to 180 months.

Figure 1. Data Distribution for Size

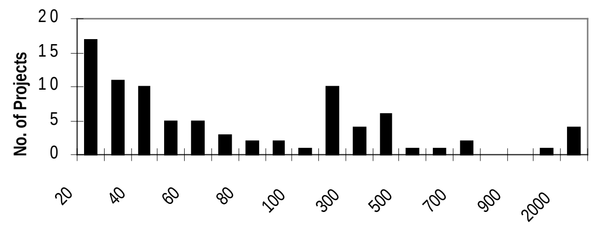


Figure 2. Data Distribution for Person Months

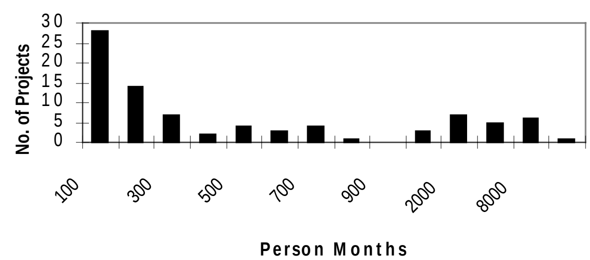
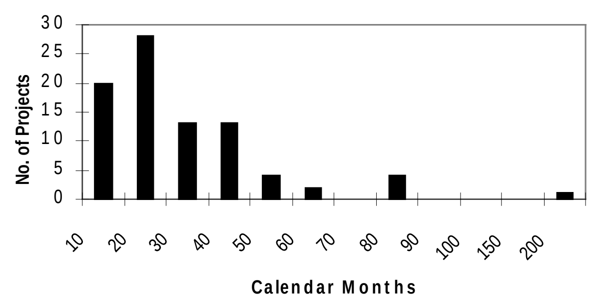


Figure 3. Data Distribution for Schedule



We have found that the different definitions of product and development in the 1990's make data sources less comparable than in the 1970's. As a result, even after data normalization the accuracy of COCOMO II.1997 was less accurate on 83 projects than COCOMO 1981 was on 63 projects.

3. MODEL CALIBRATION

The Post-Architecture Model calibration process [Chulani97] began by assigning expert-determined values to the 22 (17 Effort Multipliers + 5 Scale Factors) parameters. These assigned values were iterated with USC-CSE's affiliates to obtain the initial version of the model, now known as the "A-Priori Post-Architecture Model". Once the model was well-defined; an extensive data collection activity was initiated. This activity resulted in a MS Access database of 83 projects with complete ready-to-analyze data. Statistical analyses exhibited high correlation among four of the 17 Effort Multipliers which led to consolidating them into 15 Effort Multipliers [discussed further in Chulani97]. The data was then divided into two subsets; one subset of 59 datapoints that was used for the regression analysis and the other subset of 24 projects that was used for cross validation. Multiple Regression Analyses on the first subset of 59 datapoints determined the coefficients for the 20 (now 15 Effort Multipliers + 5 Scale Factors) parameters. A 10% weighted average (using 10% of the data-driven values and 90% of the a-priori values) approach was used to adjust the A-Priori expert-determined model parameters to obtain the calibrated "A-Posteriori Post Architecture Model" that reflected the characteristics of the actual 83 projects data. The 10% weighted average technique was preferred over a pure least squares approach due to uncertainties that were apparent in the contributed data.

Once, the A-Priori Post Architecture Model was calibrated; it was cross-validated using the second subset of 24 datapoints. Prediction Accuracy was computed in terms of Proportional Error (PE) which had a Normal Distribution and the following accuracies (see next page) for Effort Prediction were observed (for example, PRED(.30) = 64% means that 64% of the estimates were within 30% of the actuals).

Prediction	Before Stratification by Organization	After Stratification by Organization
PRED(.20)	46%	49%
PRED(.25)	49%	55%
PRED(.30)	52%	64%

In the above table, the column "Before Stratification by Organization" represents the Prediction Accuracy obtained by using the "A-Posteriori Post Architecture Model" on the 83 datapoints before any data stratification. On the other hand, the column "After Stratification by Organization" represents the Prediction Accuracy obtained by using the "A-Posteriori Post Architecture Model" on the 83 datapoints after stratifying the data by organization* and computing a new multiplicative constant for each of the organizations. It is clear from the above table that simple local calibration helps in improving the prediction accuracy of the model.

* Stratification "by Organization" does not mean "by Application Type". In the COCOMO II database, we have actual project data from 18 relatively homogeneous sources and the data was stratified into different sets based on the source of the data.

The 1997 calibration results indicated that the following changes from COCOMO ’81 to COCOMO II were successfully explaining sources of variation in the project data :

- Replacing the COCOMO ’81 Development Modes by the 5 exponent drivers Precedentedness, Development Flexibility,
- Architecture/Risk Resolution, Team Cohesiveness, and CMM-based Process Maturity.
- Adding multiplicative cost drivers for Amount of Documentation and Multisite Development.

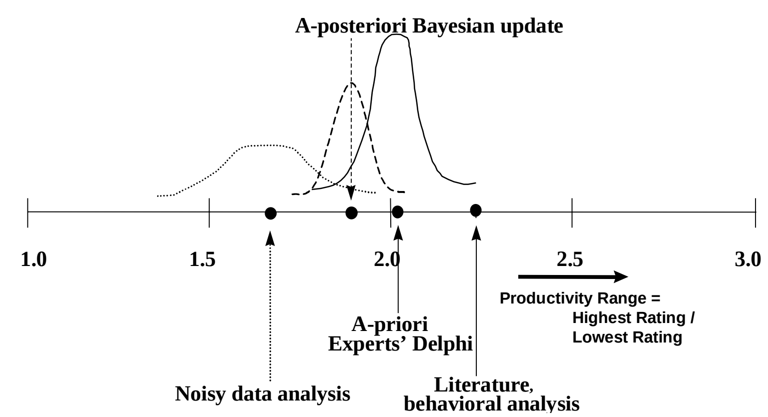
The evidence from the data was negative for another new multiplicative cost driver, Required Reusability, most likely due to misinterpretations in the rating scale during data collection.

4. STRATEGY AND FUTURE CALIBRATION

Based on our regression analysis results, we used a 10% weighted average approach versus a pure least squares regression. The 10% weighting factor was selected after comparison runs using 10%, 25% and 40% weighting factors were found to produce less accurate results than the 10% factors. This moves the model parameters in the direction suggested by the regression coefficients but retains the rationale contained within the apriori values. As more data is used to calibrate the model, a greater percentage of the weight will be given to the regression determined values. Thus the strategy is to release annual updates to the calibrated parameters with each succeeding update producing more data driven parameter values. Hence the COCOMO II model name is appended by a year date to denote the calibration database and approach on which the model is based, e.g. COCOMO II.1997.

With more data and a better understanding of the calibration process we have started doing Bayesian analysis to derive parameter coefficients. Each of the 22 parameters will have separate weights based on their relative variance and significance. For example, the parameter Analyst Capability (ACAP) is a well-understood parameter and the data collected has very little variance. In this case, we give more weight to the data and less weight to the a-priori value of the coefficient. In contrast, RUSE is a new parameter and data reported on this variable is incomplete and very often misinterpreted. Hence, we give a low weight to the data-determined values. See figure 4 for an overview of the suggested Bayesian approach.

Figure 4: Bayesian Approach



REFERENCES

[Boehm95] Boehm, B., B. Clark, E. Horowitz, C. Westland, R. Madachy, R. Selby, “Cost Models for Future Software Life Cycle Processes: COCOMO 2.0,” Annals of Software Engineering Special Volume on Software Process and Product Measurement, J.D. Arthur and S.M. Henry (Eds.), J.C. Baltzer AG, Science Publishers, Amsterdam, The Netherlands, Vol 1, 1995, pp. 45 - 60.

[Chulani97] Chulani S., Clark B., Boehm B. “Calibrating the COCOMO II Post Architecture Model”, USC-CSE-97-510.

[USC-CSE1] Center for Software Engineering , “COCOMO II Model Definition Manual,” Computer Science Department, University of Southern California, Los Angeles, Ca. 90089, http://sunset.usc.edu/Cocomo.html, 1997.

[USC-CSE2] Center for Software Engineering, “COCOMO II Cost Estimation Questionnaire,” Computer Science Department, University of Southern California, Los Angeles, Ca. 90089, http://sunset.usc.edu/Cocomo.html, 1997.

ACKNOWLEDGEMENTS

This work is being supported both financially and technically under the AFRL Contract F30602-96-C-0274, "KBSA Life Cycle Evaluation" and by the COCOMO II Program Affiliates: Aerospace, Air Force Cost Analysis Agency, Allied Signal, AT&T, Bellcore, EDS, Raytheon E-Systems, GDE Systems, Hughes, IDA, JPL, Litton, Lockheed Martin, Loral, MCC, MDAC, Motorola, Northrop Grumman, Rational, Rockwell, SAIC, SEI, SPC, Sun, TI, TRW, USAF Rome Lab, US Army Research Labs, Xerox.

ABSTRACT

The COCOMO II model was created to meet the need for a cost model that accounted for future software development practices. This paper describes some of the experiences learned in calibrating the COCOMO II Post-Architecture model from the 1997 dataset of 83 observations. The results of the multiple regression analysis, their implications, and the future calibration strategy of Bayesian analysis are discussed.

TITLE OF PAPER:

Calibration Approach and Results of the COCOMO II Post-Architecture Model

CONTACT PERSON:

Name: Sunita Devnani-Chulani

Address: Center for Software Engineering, University of Southern California, Henri Salvatori Room 330

L.A. CA 90089-0781

Email: sdevnani@sunset.usc.edu

Telephone: 213-740-6470

Fax: 213-740-4927

Biography:

Sunita Devnani-Chulani is a research assistant at the Center for Software Engineering at the University of Southern California. She is an active participant of the COCOMO research team and is primarily responsible for its calibration. She is working on a Cost/Quality Model which will be an extension to the existing COCOMO II model. Her main interests include software process improvement, software reliability modeling and software metrics and cost modeling.

Sunita holds a BE in Computer Engineering from Bombay University and a MS in Computer Science from University of Southern California. She is currently a Ph.D. candidate at the Center for Software Engineering at the University of Southern California.

Biographies of Co-authors

Bradford Clark is a Research Associate at the Center for Software Engineering at the University of Southern California and he is a consultant in the field of Software Metrics. He studied the effects of Software Process Maturity on Software Development effort. He has worked on the COCOMO II model definition, data collection, and model calibration since the COCOMO II project began in July of 1994. Previously he was engaged in promoting and assisting Software Process Improvement at the Naval Air Warfare Center - Weapons Division, in China Lake, California. He has also been involved in demonstrating the suitability of the Ada Programming Language for small real-time systems. Prior to this Brad was an A-6 pilot in the U.S. Navy for nine years. He received his B.S. in Computer and Information Science from the University of Florida (1978), an M.S. degree in Software Engineering from USC (1995), and a Ph.D. in Computer Science from USC (1997).

Barry Boehm is the TRW Professor of Software Engineering and Director of the Center for Software Engineering at the University of Southern California. His current research involves the WinWin groupware system for software requirements negotiation, architecture-based models of software quality attributes, and the COCOMO II cost-estimation model. Boehm received a BA in mathematics from Harvard University and an MS and PhD in mathematics from the University of California at Los Angeles. He is a fellow of the IEEE and the AIAA.

Bert Steece is Deputy Dean of Faculty and Professor of Statistics in the Marshall School of Business, University of Southern California. Prior to joining USC, he spent four years on the faculty of the University of Oregon and two years at the Jet Propulsion Laboratory.

Professor Steece, the 1990 Program Chair of the Business and Economic Section of the American Statistical Association Annual meeting, is past-president of the Decision Sciences Institute (Western Region). He also serves on two editorial boards as a consultant to several national firms.

Professor Steece is a specialist in statistical modeling, time series analysis (forecasting), and econometrics. He has published research papers in a number of journals including Management Science, Accounting Review, and Communications in Statistics. His papers in the Accounting Review received two outstanding tax manuscript awards from the American Taxation Association, and his paper on model adequacy received an award from the Decision Sciences Institute. Presently, Dr. Steece's research focuses on issues in Bayesian modeling, regression diagnostics, and statistical graphics. His areas of applications include economic forecasting, air pollution, hospital delivery systems, and the economic impact of taxes and regulations.