



Estimation of Maintenance Costs in Man-months using COCOMO 81 basic and Intermediate Model by Maintaining Annual Change Traffic and lines of code Values as Constant

C.V.S.R Syavasya
Master of Technology
Gitam University, India

Abstract: *Constructive Cost Model is one of the cost estimation models which contribute three different models namely basic, intermediate and detailed model. Maintenance of software is very much important so as to increase the functionality of software and decrease the cost incurred in extracting new system. The first level of COCOMO is Basic which is good for quick, early, rough order of magnitude estimates of software costs, but its accuracy is limited due to its lack of factors to account for difference in project attributes. In the Intermediate COCOMO takes these Cost Drivers into account and Detailed COCOMO additionally accounts for the influence of individual project phases. Here, in this paper, we mainly estimate maintenance cost of software using basic and intermediate model by maintaining Annual change traffic and thousand source lines of code as constant. We examine these values by varying on each case, thereby obtaining the final maintenance costs in man-months.*

Keywords: COCOMO, KLOC, ACT, MM_{DEV} , MM_{MAIN}

I. Introduction

Cocomo model [1]: The COCOMO cost estimation model is used by thousands of software project managers, and is based on a study of hundreds of software projects. Unlike other cost estimation models, it is an open model, so all of the details are published, including: The underlying cost estimation equations are obtained by software maintenance cost formulated. It is well defined, and because it doesn't rely upon proprietary estimation algorithms, Costar offers these advantages to its users: It is used to estimate more objective and repeatable than estimates made by methods relying on proprietary models. It can be calibrated to reflect your software development environment, and to produce more accurate estimates Costar is a faithful implementation of this model that is easy to use on small projects, and yet powerful enough to plan and control large projects.

Costar allows defining a software structure to meet your needs. Your initial estimate might be made on the basis of a system containing 3,000 lines of code. Your second estimate might be more refined so that you now understand that your system will consist of two subsystems. There are some sizing approaches for estimating the software maintenance efforts such as source lines of code (SLOC). COCOMO is used to compute the effort and calendar time based on the size and several characteristics of the software product.

A) *Review on background:*

COCOMO (*Constructive Cost Model*) is a model for estimating costs of software projects. It was created by Barry W. Boehm and published in 1981 using data collected from 63 projects. The quantity of projects studied and the care in its articulation have made the model popular and encouraged its use. The model offers empirical formulas for estimating software costs. We have taken the COCOMO as the starting point for our research, although we have borne in mind other data from the same author. After application of the first version of his model to a wide range of situations, Boehm concluded that coverage for only one mode of developing software was insufficient, so he provided for three modes: organic, semidetached and embedded. In this way, he recognized the influence of various characteristics such as size, communication needs, experience in similar projects, etc. Furthermore, Boehm provides the COCOMO in three versions: basic, intermediate, and detailed. The basic version is useful for quick estimations, although without fine precision. The intermediate version deals with 15 attributes of the project (reliability required, database size, memory restrictions, response time required, etc.), all of whose valuation acts as a multiplying factor in the model.

II. COCOMO Maintenance Cost Estimation [2]

The detailed version deals with estimates in each of the phases in the life cycle of the project. The basic version of the model offers the following formulas to calculate the cost of software development measured in MM (i.e., man-months):

The first version of COCOMO is COCOMO 81. This model, extensively described *The Software Engineering Economics* is still an interesting model to understand. The COCOMO 81 model is a static model with a hierarchy of three estimation models [3].

- 1) Basic Model: It provides a rough estimate of effort based on size and mode of software development.
- 2) Intermediate Model: It refines the basic model by use of 14 cost drivers-these are subjective attributes. The impact of cost drivers is considered at the project level, Effort is further adjusted to take into account any schedule constraint.
- 3) Detailed Model: This model further refines the estimation by considering the phase-wise impact of cost-drivers, and the computations in this are for various sub-systems/modules.

All the three COCOMO 81 models use size as the main input. Another factor used in all three models is the “mode” of project development- the way the project is going to be handled. The three modes considered are:

- 1) Organic Mode: Project developed by a small, experienced team working in a familiar environment. The team communicates easily. Projects done like these are small or medium sizes with minimal need for innovations. Environment, hardware and tools are stable.
- 2) Semi-detached Mode: This mode lies between organic and embedded mode.
- 3) Embedded Mode: Projects developed by large teams where communication is not achieved easily. Projects are medium to large, being developed under tight constraints, possibly with instable hardware and software environment and requiring complex, innovate processing.

A) Basic Model: [3]

The Basic Model is the simplest model and does not consider any cost drivers, using the size as the basis for estimation. It is approximately only for a rough and ready estimate.

The following is the calibration for size specified in KLOC. This is based on project data collected from multiple sources.

$$\text{Organic mode: } MM_{DEV} = 2.4 KS^{1.05} \quad (1)$$

$$\text{Semidetached mode: } MM_{DEV} = 3.0 KS^{1.12} \quad (2)$$

$$\text{Embedded mode: } MM_{DEV} = 3.6 KS^{1.20} \quad (3)$$

Where KS is an estimate of the delivered program size in thousands of instructions (or roughly, lines of source code).

To estimate the maintenance cost, another parameter is needed: the annual change traffic (ACT) which consists of the proportion of original instructions that undergo a change during a year by addition or modification.

$$ACT = (NNL + NML) / (NOL) \quad (4)$$

Where:

NNL = number of new lines,

NML = number of modified lines, and

NOL = number of original lines.

In the COCOMO, the cost in the maintenance stage is given by the product of the development cost and the annual change traffic:

$$MM_{MAIN} = ACT * MM_{DEV}$$

B. Intermediate Model [3]

The Intermediate Model refines the Basic Model by introducing “cost drivers”. It first uses a nominal cost equation for effort computation, and then applies a factor based on the cost drivers of it.

Here, cost driver is the product of a $*(size^b)$ implies:

$$MM_{dev} = a * (size^b) * Multiplier$$

Where the multiplier is the impact factor of the cost drivers. For the intermediate model, the calibration of a and b suggested in COCOMO 81 is shown below as follows:

$$\text{Organic Mode: } MM_{DEV} = 3.2 * KS^{1.05}$$

$$\text{Semidetached Mode: } MM_{DEV} = 3.0 * KS^{1.12}$$

$$\text{Embedded Mode: } MM_{DEV} = 2.8 * KS^{1.2}$$

$$MM_{MAIN} = ACT * MM_{DEV}$$

Where MM_{DEV} is the development costs in man-months, a and b are factors to calibrate using past data for each development mode.

There are totally 15 multipliers in Intermediate Model which are rated in different value starting from very low to extra high.

As COCOMO is a non-proprietary model, its details are available in the public domain [1], encouraging researchers and practitioners in the software engineering community to independently evaluate the model. There have been many extensions independently reported; use machine learning techniques to generate effort models from the original COCOMO model. Gulezian proposed a calibration method by transforming the model equation into a linear form and estimating the model parameters using standard linear regression techniques.

III. Source Lines of Code

The COCOMO calculations are based on your estimates of a project's size in Source Lines of Code (SLOC). SLOC is defined such that only Source lines that are DELIVERED as part of the product are included, test drivers and other support software is excluded, SOURCE lines are created by the project staff, code created by applications

generators is excluded. The original COCOMO 81 model was defined in terms of Delivered Source Instructions, which are very similar to SLOC. The major difference between DSI and SLOC is that a single Source Line of Code may be several physical lines. For example, an "if-then-else" statement would be counted as one SLOC, but might be counted as several DSI. Glance of cocomo model: The COCOMO cost estimation model is used by thousands of software project managers, and is based on a study of hundreds of software projects. Unlike other cost estimation models, it is an open model, so all of the details are published, including: The underlying cost estimation equations are obtained by software maintenance cost formulated. It is well defined, and because it doesn't rely upon proprietary estimation algorithms, Costar offers these advantages to its users: Its estimates are more objective and repeatable than estimates made by methods relying on proprietary models it can be calibrated to reflect your software development environment, and to produce more accurate estimates Costar is a faithful implementation of this model that is easy to use on small projects, and yet powerful enough to plan and control large projects. Typically, you'll start with only a rough description of the software system that you'll be developing, and you'll use Costar to give you early estimates about the proper schedule and staffing levels. As you refine your knowledge of the problem, and as you design more of the system, you can use Costar to produce more and more refined estimates. Costar allows defining a software structure to meet your needs. Your initial estimate might be made on the basis of a system containing 3,000 lines of code. Your second estimate might be more refined so that you now understand that your system will consist of two subsystems. There are some sizing approaches for estimating the software maintenance efforts such as source lines of code (SLOC). COCOMO is used to compute the effort and calendar time based on the size and several characteristics of the software product.

IV. Cost Drivers [4]

COCOMO II has 17 cost drivers to assess project, development environment, and team to set each cost driver. The cost drivers are multiplicative factors that determine the effort required to complete your software project. For example, if a project will develop software that controls an airplane's flight, you would set the Required Software Reliability cost driver to Very High. That rating corresponds to an effort multiplier of 1.26, meaning that your project will require 26% more effort than a typical software project. Constructive Cost Model is used for estimating costs of software projects. It was created by Barry W. Boehm and published in 1981 using data collected from 63 projects. The quantity of projects studied and the care in its articulation have made the model popular and encouraged its use. The model offers empirical formulas for estimating software costs. We have taken the COCOMO as the starting point for our research. After application of the first version of his model to a wide range of situations, Boehm concluded that coverage for only one mode of developing software was insufficient, so he provided for three modes: organic, semidetached and embedded. In this way, he recognized the influence of various characteristics such as size, communication needs, experience in similar projects, etc. Furthermore, Boehm provides the COCOMO in three versions: basic, intermediate, and detailed. The basic version is useful for quick estimations, although without fine precision. The intermediate version deals with 15 attributes of the project (reliability required, database size, memory restrictions, response time required, etc.), all of whose valuation acts as a multiplying factor in the model.

V. Discussion on Results Obtained:

Here, we first calculate Annual change traffic. To calculate ACT, we first assume the Total lines of code which is used in source code as 1000, number of lines added as 250 and number of lines modified as 50. Hence finally, the value of ACT is obtained as 0.3, which is maintained constant for three modes.

The value of KS is taken as 250 is assigned as constant for all the three modes namely organic, semidetached, embedded modes.

A) At Basic Model:

In Basic Model, by maintaining ACT and KS values as constant, we calculate MM_{DEV} for all the three modes as follows:

Organic Mode:

$$MM_{DEV} = 2.4 * (250)^{1.05} = 790.7$$

$$MM_{MAIN} = ACT * MM_{DEV} = 237.21$$

Semidetached Mode:

$$MM_{DEV} = 3.0 * (250)^{1.12} = 1454.8$$

$$MM_{MAIN} = 436.44$$

Embedded Mode:

$$MM_{DEV} = 3.6 * (250)^{1.20} = 2715.3$$

$$MM_{MAIN} = 814.59$$

B) At Intermediate Model:

In Intermediate Model, by maintaining ACT and KS values as constant, we calculate MM_{DEV} for all the three modes as follows:

Organic Mode:

$$MM_{DEV} = 3.2 * (250)^{1.05} = 1054.3$$

Multiplier Value:

The Multiplier values is calculate as,

Cost Driver (characteristic) * Cost Driver (characteristic) according to project's nature. This is termed as Multiplier Equation. According to the above, we assume the following:

Here, we assume DATA multiplier as Low and TOOL multiplier as High

It implies,

DATA (Low) = 0.94 (According to Cost driver Multipliers of COCOMO 81)

TOOL (High) = 0.91 (According to Cost driver Multipliers of COCOMO 81)

$$\text{Therefore, Multiplier Value} = \text{DATA (Low)} * \text{TOOL (High)} = 0.8554$$

Hence, Development costs of software are obtained by the product of Development cost value and Multiplier value obtained.

$$MM_{DEV} = a * (\text{size}^b) * \text{Multiplier that implies,}$$
$$= 1054.3 * 0.8554 = 901.8$$

Therefore

$$MM_{DEV} = 901.8$$

After obtaining Developments costs, now we calculate final maintenance cost in man-months:

$$MM_{MAIN} = ACT * MM_{DEV}$$
$$= 0.3 * 901.8 = 270.54$$

$$MM_{MAIN} = 270.54$$

Semidetached Mode:

$$MM_{DEV} = 3.0 * (250)^{1.12} = 1454.8$$

Multiplier Value:

The Multiplier values is calculate as,

Cost Driver (characteristic) * Cost Driver (characteristic) according to project's nature. This is termed as Multiplier Equation. According to the above, we assume the following:

Here, we assume Project's nature as Nominal because DATA multiplier is Nominal and MODP multiplier is Nominal.

It implies,

$$\text{Therefore, Multiplier Value} = 1.0$$

Hence, Development costs of software are obtained by the product of Development cost value and Multiplier value obtained.

$$MM_{DEV} = a * (\text{size}^b) * \text{Multiplier that implies,}$$
$$= 1454.8 * 1.0 = 1454.8$$

Therefore

$$MM_{DEV} = 1454.8$$

After obtaining Developments costs, now we calculate final maintenance cost in man-months:

$$MM_{MAIN} = ACT * MM_{DEV}$$
$$= 0.3 * 1454.8 = 436.44$$

$$MM_{MAIN} = 436.44$$

Embedded Mode:

$$MM_{DEV} = 2.8 * (250)^{1.20} = 2111.9$$

Multiplier Value:

The Multiplier values is calculate as,

Cost Driver (characteristic) * Cost Driver (characteristic) according to project's nature. This is termed as Multiplier Equation. According to the above, we assume the following:

Here, we assume TIME multiplier as Nominal and TOOL as Nominal multiplier as High

It implies,

DATA (Nominal) = 1.0 (According to Cost driver Multipliers of COCOMO 81)

TOOL (Nominal) = 1.0 (According to Cost driver Multipliers of COCOMO 81)

Therefore, Multiplier Value = 1.0

Hence, Development costs of software are obtained by the product of Development cost value and Multiplier value obtained.

$$MM_{DEV} = a * (size^b) * \text{Multiplier that implies,}$$

$$= 2111.9 * 1.0 = 2111.9$$

Therefore

$$MM_{DEV} = 2111.9$$

After obtaining Developments costs, now we calculate final maintenance cost in man-months:

$$MM_{MAIN} = ACT * MM_{DEV}$$

$$= 0.3 * 2111.9 = 633.57$$

$$MM_{MAIN} = 633.57$$

Finally, we consolidate the maintenance cost values obtained in Basic and intermediate model of all the modes in the form of table as follows:

TABLE I

Table showing the final maintenance cost values of Basic and Intermediate Models of all three modes:

Model/Mode	Organic Mode	Semi-Detached Mode	Embedded Mode
Basic Model	237.21	436.44	814.59
Intermediate Model	270.54	436.44	633.57

In the above table, we observe that as Multiplier value increases, Maintenance cost increases in Organic mode. In case of Semi-Detached, as the project's nature is Nominal in Intermediate model, the value obtained is same for both Basic and Intermediate model. In case of Embedded Mode, Maintenance cost decreases as the multiplier value remained Nominal.

VI. Graph:

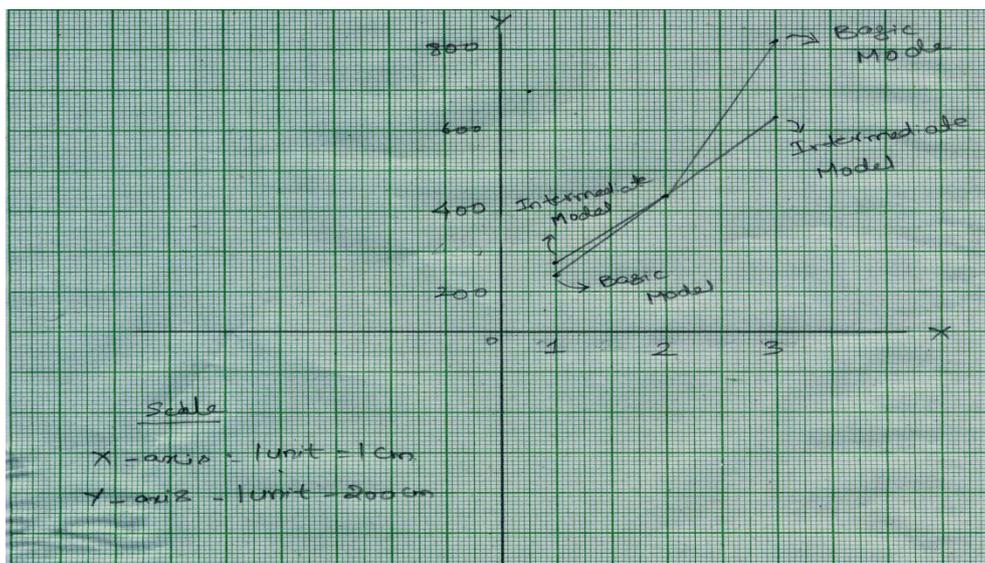


Fig. Graph showing the results obtained in basic and intermediate model of software maintenance cost.

A graph is drawn as shown in the above figure, it summarizes that as multipliers are added to intermediate model, the cost of maintenance is decreasing, where in basic model the graph increased.

VII. Conclusion:

Annual change traffic as well as Lines of code plays a crucial role in finding out the differences in new results obtained in basic and intermediate model. The main aim in taking up this paper is to compare and evolve the changes occurred when new set of resulted values are formed. This comparison helps the software engineering professionals to know the cost decreasing principles in software maintenance. Finally, we conclude this paper by obtaining the final maintenance costs and cost decreasing factors. As part of the future work, this project can be extended by assigning random multiplier values and check the difference in variations of software maintenance. The global aim here is to decrease maintenance costs at maximum extent so as to encourage the software developers to go for maintaining of software at the time of malfunctioning instead of going for new product, thereby benefiting the financial status of Information Technology sector in India.

References

- 1) "An Approach to Find Maintenance Costs Using Cost Drivers of Cocomo Intermediate Model" Vol 3 Issue. 1, January 2013, pg.154-158 in International Journal of Computational Engineering Research by cvs sr syavasya
- 2) <http://en.wikipedia.org/wiki/COCOMO>
- 3) "Software Requirements and Estimation" by Swapna Kishore and Rajesh Naik published by Tata McGraw Hill.
- 4) <http://www.softstarsystems.com/overview.htm>