Calibrating the COCOMO II Post Architecture Model

Presented by:
Sunita Devnani-Chulani
sdevnani@sunset.usc.edu

On behalf of:
Brad Clark, Sunita Devnani-Chulani, Barry Boehm

ICSE 98
Outline

⇒ Brief History of COCOMO

◆ Modeling Methodology

◆ COCOMO II Calibration Approaches
  – COCOMO II.1997 Calibration
    ❖ Process
    ❖ Prediction Accuracies
  – COCOMO II.1998 Calibration
    ❖ Simple Cost Model Bayesian Prototype
    ❖ Status

◆ Updates and Plans
  – Plans for Improving Prediction Accuracies
COnstructive COst MOdel (COCOMO)

◆ COCOMO published since 1981

◆ Commercial implementations of COCOMO
  ❖ CoCoPro, CB COCOMO, COCOMOID, COSTMODL, GECOMO Plus, SECOMO, etc.

◆ Other models based on COCOMO
  ❖ REVIC, Gulezian

◆ COCOMO II
  ❖ Research effort started in 1994 to develop a 1990’s-2000’s software cost model
  ❖ Address new processes and practices
  ❖ COCOMO II.199Y/200Y
Outline

◆ Brief History of COCOMO

➡ Modeling Methodology

◆ COCOMO II Calibration Approaches
  – COCOMO II.1997 Calibration
    ❖ Process
    ❖ Prediction Accuracies
  – COCOMO II.1998 Calibration
    ❖ Simple Cost Model Bayesian Prototype
    ❖ Status

◆ Updates and Plans
  – Plans for Improving Prediction Accuracies
The Seven-Step Modeling Methodology

1. Analyze Existing literature
2. Perform Behavioral Analysis
3. Identify Relative Significance
4. Perform Expert-Judgment, Delphi Assessment
5. Gather Project Data
6. Determine Bayesian A-Posteriori Update
7. Gather more data; refine model

A-PRIORI MODEL + SAMPLING DATA = A-POSTERIORI MODEL
Literature, Behavioral Analysis (Steps 1-3)

Productivity Range = \( \frac{\text{Highest Rating}}{\text{Lowest Rating}} \)

1.54

Literature, behavioral analysis
Results of Delphi (Step 4)

A-priori
Experts’ Delphi

Productivity Range =
Highest Rating /
Lowest Rating

Literature,
behavioral analysis
Results of Sampling Data (Step 5)

Productivity Range = Highest Rating / Lowest Rating

A-priori
Experts’ Delphi

Noisy data analysis

Literature, behavioral analysis
Results of Bayesian Update: Using Prior and Sampling Information (Step 6)

- A-priori Bayesian update
- A-posteriori Bayesian update

Productivity Range = Highest Rating / Lowest Rating

Noisy data analysis
A-priori
Experts’ Delphi

1.28
1.42
1.54

Literature,
behavioral analysis
Outline

◆ Brief History of COCOMO
◆ Modeling Methodology
➔ COCOMO II Calibration Approaches
  – COCOMO II.1997 Calibration
    ❖ Process
    ❖ Prediction Accuracies
  – COCOMO II.1998 Calibration
    ❖ Simple Cost Model Bayesian Prototype
    ❖ Status
◆ Updates and Plans
  – Plans for Improving Prediction Accuracies
COCOMO II Calibration Process

- Began with expert-determined a-priori model parameters
  - Iterated with Affiliates (Result => Original Post Architecture Model)
- Collected Data
- Identified and consolidated highly correlated model parameters
- Statistically determined estimates of consolidated model parameters from data
  - Using logarithms to linearize regression
- Used data determined model parameters to adjust a-priori model parameters
  - Experimented with weighting factors
Consolidated Highly Correlated Parameters

<table>
<thead>
<tr>
<th>TIME</th>
<th>1.0000</th>
<th>0.6860</th>
<th>-0.2855</th>
<th>-0.2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>STOR</td>
<td>0.6860</td>
<td>1.0000</td>
<td>-0.0769</td>
<td>-0.0027</td>
</tr>
<tr>
<td>ACAP</td>
<td>-0.2855</td>
<td>-0.0769</td>
<td>1.0000</td>
<td>0.7339</td>
</tr>
<tr>
<td>PCAP</td>
<td>-0.2015</td>
<td>-0.0027</td>
<td>0.7339</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

TIME   STOR   ACAP   PCAP

• What do we do? ⇒ Combine:

TIME & STOR to give RCON (Resource Constraints)
ACAP & PCAP to give PERS (Personnel Factors)
Thus, 15 effort multipliers instead of 17 for calibration
Statistical Data Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Ratio (Max/Min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFFORT</td>
<td>6</td>
<td>11400</td>
<td>1900</td>
</tr>
<tr>
<td>SIZE</td>
<td>2.6</td>
<td>1292.8</td>
<td>497</td>
</tr>
</tbody>
</table>

Thus, we took log transforms to normalize the response variable. Also, we took log transforms to linearize the parametrized model.
COCOMO II Calibration Model

- Needed linear model for regression:

\[ Y = B_0 + B_1 X_1 + B_2 X_2 + \cdots + B_p X_p \]

- COCOMO II Post-Architecture is non-linear

\[ Y = B_0 X^{B_1} \]

- What did we do?
  - Expanded COCOMO II model
  - Transformed products with logarithms to produce sums
Expanded COCOMO II

• Distributed the Scale Factors
• Resulted in 21 predictor variables i.e. 15 Effort Multipliers + 5 Scale Factors + (Size)\(^{1.01}\)

\[ PM_{est} = A \cdot (Size)^{1.01} \cdot (Size)^{SF_1} \cdot (Size)^{SF_2} \cdots EM_1 \cdots EM_{15} \]

Log Transformed COCOMO:

\[ \ln(PM_{est}) - \ln(Size)^{1.01} = \ln(A) + SF_1 \ln(Size) + \cdots + \ln(EM_{15}) \]

• Regression analysis derived the coefficients, \( B_i \), for each factor
COCOMO II.1997 Calibration

- 83 projects
- Multiple Linear Regression
  - 10% weighted average between a-priori values and data-determined values

Develop for Reuse (RUSE)
## Accuracy Results

### Effort Prediction:

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Before Stratification By Organization</th>
<th>After Stratification By Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRED(.20)</td>
<td>46%</td>
<td>49%</td>
</tr>
<tr>
<td>PRED(.25)</td>
<td>49%</td>
<td>55%</td>
</tr>
<tr>
<td>PRED(.30)</td>
<td>52%</td>
<td>64%</td>
</tr>
</tbody>
</table>

### Schedule Prediction:

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Before Stratification By Organization</th>
<th>After Stratification By Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRED(.20)</td>
<td>48%</td>
<td>52%</td>
</tr>
<tr>
<td>PRED(.25)</td>
<td>54%</td>
<td>61%</td>
</tr>
<tr>
<td>PRED(.30)</td>
<td>61%</td>
<td>62%</td>
</tr>
</tbody>
</table>
COCOMO ‘81 & COCOMO II.1997

◆ COCOMO ‘81: Pred (.20) = 68%
◆ COCOMO II.1997: Pred (.20) = 46%

- Challenges faced in calibrating COCOMO II
  - GUI builders, COTS, 4GL’s, reuse, requirements breakage
    - Need to rethink size metrics
  - Distributed interactive applications
    - Web-based, object-oriented, event-based
    - Middleware effects
  - New process models (evolutionary, incremental, spiral)
    - Phases overlap
    - Where are cost measurement endpoints?
  - Lack of good data
    - not enough data (i.e. very little degrees of freedom)
    - lack of dispersion
    - heteroskedasticity
Multiple Linear Regression Well-Suited When...

- lot of data is available
- no data items are missing
- there are no outliers
- the predictor variables are not highly correlated
- the predictor variables have an easy interpretation when used in model

most are violated by current software engineering data
Bayesian Analysis: COCOMO II.1998

A Simple Cost Model

◆ Suppose we have a simple software cost model

\[ Effort = \alpha \cdot Size^{B_1} \partial \]

◆ Linearizing, we get

\[ \ln(Effort) = \ln(\alpha) + \beta_1 \cdot \ln(Size) + \ln(\partial) \]

\[ \ln(Effort) = \beta_0 + \beta_1 \cdot \ln(Size) + \varepsilon \]

◆ We need to determine the values of \( \beta_0 \) and \( \beta_1 \)
Bayesian Analysis on a Simple Cost Model

- Suppose we have a simple software cost model
  \[ \text{Effort} = \alpha \cdot \text{Size}^{\beta_1} \delta \]

- Linearizing, we get
  \[
  \ln(\text{Effort}) = \ln(\alpha) + \beta_1 \cdot \ln(\text{Size}) + \ln(\delta)
  \]
  \[
  \ln(\text{Effort}) = \beta_0 + \beta_1 \cdot \ln(\text{Size}) + \varepsilon
  \]

- We need to determine the values of \( \beta_0 \) and \( \beta_1 \)
## Modeling Under Complete Prior Uncertainty - Statistical Analysis

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>1.206</td>
<td>0.199</td>
<td>6.062</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>1.03</td>
<td>0.407</td>
<td>21.93</td>
</tr>
</tbody>
</table>

\[
\ln(\text{Effort}) = 1.206 + 1.03 \ln(\text{Size})
\]

\[
\text{Effort} = 3.3 \cdot \text{Size}^{1.03}
\]

where $3.3 = e^{1.206}$
Posterior Density Functions - Noninformative Prior

\[ f(\beta_0 \mid \ln(\text{Effort})) \]

\[ f(\beta_1 \mid \ln(\text{Effort})) \]

But, can B < 1? (i.e. economies of scale?)
Some Prior Knowledge on $\beta_1$

- Experience indicates that Software exhibits diseconomies of scale [Banker94, Gulledge93].
- Suppose we believe that $\beta_1 \geq 1$
- $f(\beta_1) = 1$ if $\beta_1 > 1.0$
- $f(\beta_1) = 0$ if $\beta_1 \leq 1.0$

![Diagram showing function $f(\beta_1)$]
Post Sample Density Functions - Inclusion of Prior Information

~ 50% of area under curve
Bayes Theorem

\[ g(\beta / y) = \frac{f(y / \beta) g(\beta)}{f(y)} \]

\[ g(\beta / y) \propto l(\beta / y) g(\beta) \]

posterior information \( \propto \) sample information \( \times \) prior information

A - Priori Information + Sampling Data = A - Posteriori Model
# Prediction Accuracies*

<table>
<thead>
<tr>
<th>Effort Prediction</th>
<th>Bayesian Approach 159 observations</th>
<th>COCOMO II.1997 83 observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRED(.20)</td>
<td>49%</td>
<td>46%</td>
</tr>
<tr>
<td>PRED(.25)</td>
<td>57%</td>
<td>49%</td>
</tr>
<tr>
<td>PRED(.30)</td>
<td>64%</td>
<td>52%</td>
</tr>
</tbody>
</table>

* Please note that the COCOMO II.1998 calibration is not yet complete. Hence these results may be different for the COCOMO II.1998 final calibration.
COCOMO II Calibration

Approaches

COCOMO II.1997
10% weighted-average approach

Bayesian approach - weight determined by data and prior significance

Productivity Range = Highest Rating / Lowest Rating
Outline

- Brief History of COCOMO
- Modeling Methodology
- COCOMO II Calibration Approaches
  - COCOMO II.1997 Calibration
    - Process
    - Prediction Accuracies
  - COCOMO II.1998 Calibration
    - Simple Cost Model Bayesian Prototype
    - Status
- Updates and Plans
  - Plans for Improving Prediction Accuracies
COCOMO II Research Group’s Aim

100% Expert Driven

100

50

10

100% Data Driven

~200 datapoints - Bayesian model

83 datapoints - COCOMO II.1997 version

Number of projects used in calibration