COCOMO 2.0
Post-Architecture Calibration

Bradford K. Clark

10/10/96
Presentation Outline:

⇒ Motivation

- Data collection activity
- COCOMO 2.0 Post-Architecture Model
- COCOMO calibration model
- Results to date
- Conclusions and future work
COCOMO Model Motivation:

- Need to address future software practices
- Update existing algorithmic cost model
- Model based on software engineering knowledge and observations
- Model values based on collected data from Industrial Affiliates to the Center for Software Engineering
Data Collection:

- Define the data needed.
- Collect data with a paper form or a computer software tool
- Affiliate Organizations providing majority of data.
  - Historical - whole project
- Site visits or phone interviews to record data
- Enter in data into the repository
  - Data is labeled with generic id
  - Stored in locked room
  - Limited access by researchers
Post-Architecture Model:

- Non-linear model:
  \[ PM = A \cdot (\text{Size})^B \cdot \prod_{i=1}^{17} EM_i \]

- B consists of 5 scale factors:
  \[ B = 1.01 + 0.01 \cdot \sum_{j=1}^{5} SF_j \]
COCOMO Calibration Model:

- Need linear model for regression:

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

- COCOMO 2.0 Post-Architecture is non-linear

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

- What should we do?
  - Expand COCOMO model
  - Transform products with logarithms to produce sums
Expanded COCOMO:

- Distribute the Scale Factors
- Results in 23 factors

\[ PM_{est} = A \cdot (Size)^{1.01} \cdot (Size)^{SF_1} \cdot (Size)^{SF_2} \cdots EM_1 \cdots EM_{17} \]

Log Transformed COCOMO:

\[ \ln(PM_{est}) = \ln(A) + 1.01\ln(Size) + SF_1\ln(Size) + \cdots + \ln(EM_{17}) \]

- Regression analysis will derive the coefficients, \( B_i \), for each factor
- \( \ln(A) \) is dropped
Presentation Outline:

- Motivation
- Data collection activity
- COCOMO 2.0 Post-Architecture Model
- COCOMO calibration model

⇒ Results to date

- Conclusions and future work
Results:

- 65 Observations from different Industrial categories:
  
  Commercial: 2
  
  Aerospace: 4
  
  FFRDC: 2
  
- Results improved with stratification of data by organization

- Forecast accuracy measured with proportional error:

\[
PE = \begin{cases} 
\left[ PM_{est} \div PM_{act} \right] - 1, & (PM_{est} - PM_{act}) \geq 0 \\
- \left[ PM_{act} \div PM_{est} \right] + 1, & (PM_{est} - PM_{act}) < 0
\end{cases}
\]
PE Before Regression

Std. CW = .81
man = .22
N = 65.00
Regression
(without stratification):

Adjusted R Square  .94570
Standard Error  .38872

Analysis of Variance

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>23</td>
<td>171.88624</td>
<td>7.47331</td>
</tr>
<tr>
<td>Residual</td>
<td>41</td>
<td>6.19527</td>
<td>.15110</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td>49.45804</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>LN_ACAP</td>
<td>0.364606</td>
<td>0.695296</td>
<td>0.524</td>
</tr>
<tr>
<td>LN_AEXP</td>
<td>0.231304</td>
<td>0.602883</td>
<td>0.384</td>
</tr>
<tr>
<td>LN_CPLX</td>
<td>1.497551</td>
<td>0.608187</td>
<td>2.462</td>
</tr>
<tr>
<td>LN_DATA</td>
<td>2.321239</td>
<td>0.999881</td>
<td>2.322</td>
</tr>
<tr>
<td>LN_DOCU</td>
<td>-0.153751</td>
<td>0.929772</td>
<td>-0.165</td>
</tr>
<tr>
<td>LN_LTEX</td>
<td>0.124795</td>
<td>0.848733</td>
<td>0.147</td>
</tr>
<tr>
<td>LN_PCAP</td>
<td>1.236443</td>
<td>0.844417</td>
<td>1.464</td>
</tr>
<tr>
<td>LN_PCON</td>
<td>0.831002</td>
<td>1.073967</td>
<td>0.774</td>
</tr>
<tr>
<td>LN_PEXP</td>
<td>0.402171</td>
<td>0.595131</td>
<td>0.676</td>
</tr>
<tr>
<td>LN_PVOL</td>
<td>-0.045234</td>
<td>0.676307</td>
<td>-0.067</td>
</tr>
<tr>
<td>LN_RELY</td>
<td>0.584032</td>
<td>0.574250</td>
<td>1.017</td>
</tr>
<tr>
<td>LN_RUSE</td>
<td>-0.799948</td>
<td>0.609753</td>
<td>-1.312</td>
</tr>
<tr>
<td>LN_SCED</td>
<td>2.698342</td>
<td>1.243080</td>
<td>2.171</td>
</tr>
<tr>
<td>LN_SITE</td>
<td>-0.947197</td>
<td>1.157316</td>
<td>-0.818</td>
</tr>
<tr>
<td>LN_STOR</td>
<td>2.003848</td>
<td>0.849830</td>
<td>2.358</td>
</tr>
<tr>
<td>LN_TIME</td>
<td>1.209235</td>
<td>0.770581</td>
<td>1.569</td>
</tr>
<tr>
<td>LN_TOOL</td>
<td>3.058214</td>
<td>1.065924</td>
<td>2.869</td>
</tr>
<tr>
<td>LNSIZE101</td>
<td>1.107444</td>
<td>0.123345</td>
<td>8.978</td>
</tr>
<tr>
<td>LNS_FLEX</td>
<td>0.696658</td>
<td>1.275185</td>
<td>0.546</td>
</tr>
<tr>
<td>LNS_PMAT</td>
<td>1.084491</td>
<td>1.788497</td>
<td>0.606</td>
</tr>
<tr>
<td>LNS_PREC</td>
<td>2.413984</td>
<td>0.960125</td>
<td>2.514</td>
</tr>
<tr>
<td>LNS_RESL</td>
<td>-2.269417</td>
<td>1.956745</td>
<td>-1.160</td>
</tr>
<tr>
<td>LNS_TEAM</td>
<td>3.715327</td>
<td>1.995218</td>
<td>1.862</td>
</tr>
<tr>
<td>(Constant)</td>
<td>-0.560723</td>
<td>0.425930</td>
<td>-1.316</td>
</tr>
</tbody>
</table>
Regression PE:

COCOMO Forum

Std. Dev = .40
Mean = .00
N = 65.00
Regression
(with stratification)

Adjusted R Square .95331
Standard Error .36043

Analysis of Variance

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>26</td>
<td>173.14497</td>
<td>6.65942</td>
</tr>
<tr>
<td>Residual</td>
<td>38</td>
<td>4.93654</td>
<td>.12991</td>
</tr>
<tr>
<td>$F = 51.26224$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Variable                B  SE B  T
---------------------------------
LN_ACAP                 0.219985 0.664825 0.331
LN_AEXP                 0.222100 0.561130 0.396
LN_CPLX                 1.843231 0.599430 3.075
LN_DATA                 2.852544 0.952041 2.996
LN_DOCU                 0.649755 1.011005 0.643
LN_LTEX                 0.423312 0.850311 0.498
LN_PCAP                 1.659577 0.841160 1.973
LN_PCON                 0.470698 1.017761 0.462
LN_PEXP                 0.586376 0.559265 1.048
LN_PVOL                 1.225317 0.778112 1.575
LN_RELY                 0.804534 0.538962 1.493
LN_RUSE                 -0.289814 0.613725 -.472
LN_SCED                 2.237578 1.237291 1.808
LN_SITE                 -1.030253 1.098791 -.938
LN_STOR                 0.739371 0.887684 .833
LN_TIME                 1.277919 0.729614 1.752
LN_TOOL                 2.366555 1.030659 2.296
LNSIZ101                1.024621 0.154880 6.616
LNS_FLEX                0.964983 1.321752 .730
LNS_PMAT                4.139000 2.761260 1.499
LNS_PREC                1.767164 1.141516 1.548
LNS_RESL                -1.807623 1.963190 -.921
LNSTEAM                 1.959708 2.128285 .921
ORG093                  -1.038590 0.464163 -2.238
ORG587                  -0.669031 0.366404 -1.826
ORG586                  -0.167731 0.274594 -.611
(Constant)              -0.047728 0.539532 -.088

11th COCOMO Forum
Regression PE:

\[ \text{Std. Dev} = 0.37 \]
\[ \text{Mean} = 0.00 \]
\[ N = 65.00 \]
Process Maturity Investigation with COCOMO 2.0:

- Assess effect of Process Maturity\(^1\) on Software Development Effort within context of other influencing factors.
- Data collected on either CMM level of KPA Goals.
- These results from stratified analysis show a generally positive influence.
- Data needs to be inspected to determine cause of variation.
- More data points are needed.

---

1. As defined by SEI's Capability Maturity Model.
Conclusions:

- Regression technique can be used to calibrate COCOMO locally.
- COCOMO calibrated to local organization is more accurate.
- Qualify your data - inspect it, decide before-hand what an outlier looks like.

Future Work:

- Negative coefficients do not make sense in the model (check correlation’s of parameter inputs).
- Schedule equation needs to be calibrated.
- Calibration of COCOMO Early Design model.