“Modeling Software Defect Introduction”

Sunita Devnani-Chulani
(sdevnani@sunset.usc.edu)
University of Southern California
Center for Software Engineering

California Software Symposium
November 7, 1997
Presentation Outline

► Motivation and Background Model

• Model Parameters and Modeling Methodology

• Initial Results and Ongoing Research
Motivation

- Insight on Determining Ship Time
- Assessment of Quality Investment Payoffs
- Understanding of Quality Strategy Interactions

“Quality” = number of non-trivial delivered defects/KSLOC
The Software Defect Introduction and Removal Model

Overall defect rate: 60/KDSI

Documentation Defects (15/KDSI)

Code Defects (15/KDSI)

Design Defects (25/KDSI)

Requirements Defects (5/KDSI)

Residual Software Defects

From Software Engineering Economics by Barry Boehm
Based on studies done by Thayer et al, Jones and Boehm in the 1970’s

chart 4
Presentation Outline

• Motivation and Background Model

  ➤ Model Parameters and Modeling Methodology

• Initial Results and Ongoing Research
Model Parameters and Modeling Methodology

Step 1 : Analyze literature for factors affecting Defect Introduction Rate (DIR)
Step 2 : Perform Behavioral Analyses of factor effects
Step 3 : Identify Relative Significance of DIR-drivers on DIR
Step 4 : Do Delphi, determine expert-based DIR-driver values and a-priori variances
Step 5 : Gather Data
Step 6 : Determine A-Posteriori Set of Model Parameters using Bayesian Regression
Step 7 : Refine Model
Initial Set of DIR Drivers (Step 1)

• 22 from COCOMO II
  – Project, Personnel, Product, Platform
  • Later dropped Development Flexibility, FLEX

• Disciplined Methods (Personal Software Process, Cleanroom Development, etc.)
  – Nominal  Defect prevention driven by RELY (Required Software Reliability) rating
  – High     Moderate/Extra emphasis given to defect prevention
  – Very High Strong emphasis given to defect prevention equivalent to CMM defect prevention KPA.
# Behavioral Analyses and Delphi Process (Steps 2-4)

## Required Software Reliability (RELY)

<table>
<thead>
<tr>
<th>RELY level</th>
<th>Requirements</th>
<th>Design</th>
<th>Code</th>
<th>Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VH</strong></td>
<td>Fewer Requirements Completeness, consistency defects due to detailed verification, QA, CM, standards, SSR, documentation, IV&amp;V interface, test plans, procedures</td>
<td>Fewer Design defects due to detailed verification, QA, CM, standards, PDR, documentation, IV&amp;V interface, design inspections, test plans, procedures</td>
<td>Fewer Coding defects due to detailed verification, QA, CM, standards, documentation, IV&amp;V interface, code inspections, test plans, procedures</td>
<td>Fewer Documentation defects due to requirements, design shortfalls</td>
</tr>
<tr>
<td>Nominal</td>
<td>Nominal level of defect introduction</td>
<td>0.70</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>VL</strong></td>
<td>More Requirements Completeness, consistency defects due to minimal verification, QA, CM, standards, PDR, documentation, IV&amp;V interface, test plans, procedures</td>
<td>More Design defects due to minimal verification, QA, CM, standards, PDR, documentation, IV&amp;V interface, design inspections, test plans, procedures</td>
<td>More Coding defects due to minimal verification, QA, CM, standards, PDR, documentation, IV&amp;V interface, code inspections, test plans, procedures</td>
<td>More Documentation defects due to requirements, design shortfalls</td>
</tr>
<tr>
<td>Initial Quality Range</td>
<td>1.43</td>
<td>1.45</td>
<td>1.45</td>
<td>1.22</td>
</tr>
<tr>
<td>Range - Round 1</td>
<td>2.24</td>
<td>2.24</td>
<td>2.24</td>
<td>1.77</td>
</tr>
<tr>
<td>Median - Round 1</td>
<td>1.2-2.61</td>
<td>1.2-2.61</td>
<td>1.2-2.61</td>
<td>1.2-1.75</td>
</tr>
<tr>
<td>Range - Round 2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>Final Quality Range (Median - Round 2)</td>
<td>1.8-2.24</td>
<td>1.9-2.3</td>
<td>2-2.24</td>
<td>1.5-2</td>
</tr>
</tbody>
</table>

| Initial Quality Range | 2.05 | 2.1 | 2.1 | 1.5 |
| Range - Round 2 | 1.8-2.24 | 1.9-2.3 | 2-2.24 | 1.5-2 |
| Final Quality Range (Median - Round 2) | 2.05 | 2.1 | 2.1 | 1.5 |
Analysis Summary

Quality Range of Required Software Reliability (RELY)

Literature, behavioral analysis
Analysis Summary

Quality Range of Required Software Reliability (RELY)

A-priori Experts’ Delphi
Analysis Summary

Quality Range of Required Software Reliability (RELY)

Noisy data analysis
Analysis Summary

Quality Range of Required Software Reliability (RELY)

A-posteriori Bayesian update

A-priori
Experts’ Delphi

Noisy data analysis

Literature, behavioral analysis

1.0  1.5  2.0  2.5  3.0
A-Priori Experts’ Reqts. Quality Range
Presentation Outline

• Motivation and Background Model

• Model Parameters and Modeling Methodology

 ➤ Initial Results and Ongoing Research
# Defect Data Reporting Scheme

<table>
<thead>
<tr>
<th>Activity →</th>
<th>Introduced + Unresolved / Resolved in Activity/Cost To Resolve by Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Artifact ↓</td>
<td>Reqts</td>
</tr>
<tr>
<td>Requirements</td>
<td>50/30</td>
</tr>
<tr>
<td>Design</td>
<td>55/25/1.0</td>
</tr>
<tr>
<td>Code</td>
<td></td>
</tr>
<tr>
<td>Documentation</td>
<td></td>
</tr>
</tbody>
</table>
# Results of initial-project data analysis

<table>
<thead>
<tr>
<th>Type of Artifact</th>
<th>1970’s Baseline DIRs</th>
<th>Quality Adjustment Factor</th>
<th>Predicted DIR</th>
<th>Actual DIR</th>
<th>Calibrated Constant (A)</th>
<th>1990’s Baseline DIRs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requirements</td>
<td>5</td>
<td>0.5</td>
<td>2.5</td>
<td>4.5</td>
<td>1.8</td>
<td>9</td>
</tr>
<tr>
<td>Design</td>
<td>25</td>
<td>0.44</td>
<td>11</td>
<td>8.4</td>
<td>0.77</td>
<td>19</td>
</tr>
<tr>
<td>Code</td>
<td>15</td>
<td>0.5</td>
<td>7.5</td>
<td>16.6</td>
<td>2.21</td>
<td>33</td>
</tr>
<tr>
<td>Documentation</td>
<td>15</td>
<td>0.57</td>
<td>8.55</td>
<td>3.9</td>
<td>0.45</td>
<td>7</td>
</tr>
</tbody>
</table>
Ongoing Research

• Develop Defect Removal Model
• Gather more data
• Statistically determine a-posteriori model
• Integrate to existing COCOMO II Post-Architecture Model
Integrated Cost/Quality Model

Software product size estimate
Software product, process, computer, and personnel attributes
Software reuse, maintenance, and increment parameters
Defect removal activity levels
Software organization’s project data

COCOMO
Quality Model

Software development, maintenance cost and schedule estimates
Cost, schedule distribution by phase, activity, increment
Software Quality Defects/KDSI or 10FPs
COCOMO recalibrated to organization’s data
Delphi Process

• Round 1
  1. Provided Participants with Round 1 Delphi Questionnaire with a proposed set of values for the Quality Ranges. This set was proposed based on experience.
  2. Received nine completed Round 1 Delphi Questionnaires.
  3. Ensured validity of responses by correspondence with the participants.
  4. Did simple analysis based on ranges and medians of the responses.

• Round 2
  1. Provided participants with Round 2 Delphi Questionnaire -- based on analysis of Round 1.
  2. Repeated steps 2, 3, 4 (above)
  3. Converged to Final Delphi Results which resulted in the definition of the initial model