MOTIVATION

- DOD desires more credible estimates
- Provide for more realistic Program Element budgets
- Permit better assessment of Contractor proposed costs
- Reduce bad press and the perception that it is wasteful of taxpayer dollars

To achieve this goal, DOD needs

- Benchmarks that can be used to assess reasonableness of budget submissions and Contractor proposals
- Consistent ways of addressing the many different size and cost parameters used in preparing estimates and cost-to-complete exercises
- Hard data that is statistically solid that justifies findings
- With this data, improvements can be made in processes used for planning, budgeting, and control

Data Quality Challenges

- DoD has many software cost data repositories
  - DCARC, NRO, AFCAA, ODASA-CE, NCQA, USAF, Aerospace, …
- Despite increased interest in data, it is surprising that so little effort has been committed reconciling inconsistencies within and across data repositories…
  1. No reporting of Equivalent Size Inputs – CM, DM, IM, SU, UNFM, Type
  2. No common SLOC reporting – logical, physical, etc.
  3. No standard definitions – Application Domain, Build, Increment, Spiral, …
  4. No common effort reporting – analysis, design, code, test, CM, QA, …
  5. Product size only reported in lines of code
  6. No reporting of quality measures – defect density, defect containment, etc.
- The data needs to be statistically sound and defensible
- A good opportunity to communicate and work towards standardization is the current MIL-STD 881 efforts

The key to standardization is encouraging broader participation among these groups.

Limited Research within DoD

1. Other contributors to Productivity besides effort and size, are being ignored by most analysts
   - Operating Environment, Application Domain, Product Complexity
   - Personnel Capability
   - Required Reliability
   - Quality – Defect Density, Defect Containment
   - Integrating code from previous deliveries – Builds, Spirals, Increments, etc.
   - Requirements Volatility

2. Converting to Equivalent SLOC
   - Categories like Modified, Reused, Adopted, Managed, and Used add no value unless they translate into single or unique narrow ranges of DM, CM, and IM parameter values. We have seen no empirical evidence that they do
   - Other categories like COTS, Converted, Generated and Rehosted are handled differently and there is no consistency when they’re used

Model Calibration Challenges

- Most program offices and support contractors rely heavily on software cost models for their estimates
- May have not been calibrated with most recent DoD data
- Calibration with recent data (2002-Present) will help increase program office estimating accuracy

SLIM-Estimate™

TruePlanning® by PRICE Systems

Consequence: Significant Cost Growth (%)

<table>
<thead>
<tr>
<th>Statistics</th>
<th>*Total System</th>
<th>**Software Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>-64%</td>
<td>-80%</td>
</tr>
<tr>
<td>Mean</td>
<td>45%</td>
<td>37%</td>
</tr>
<tr>
<td>Maximum</td>
<td>471%</td>
<td>623%</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>71%</td>
<td>107%</td>
</tr>
<tr>
<td>Milestone Phase</td>
<td>Development</td>
<td>Development</td>
</tr>
<tr>
<td>Sample Size</td>
<td>137</td>
<td>111</td>
</tr>
<tr>
<td>Year of Data</td>
<td>1993-2003</td>
<td>2002-2008</td>
</tr>
</tbody>
</table>

**Defense Automated Cost Information System (DACIMS)
**PROPOSED SOLUTION**

AFCAA in conjunction with USC and other DOD Cost Agencies, will publish a Manual to help analysts develop more credible estimates based on empirical data in a timely and consistent manner.

---

**Special Features**

- Augment NCCA/AFCAA Software Cost Handbook:
  - Empirical Productivity Benchmarks by Operating Environment, Application Domain, and Software Size
  - Default Equivalent SLOC Inputs (DM, CM, IM, SU, UNFM):
    - Requirements Volatility derived from empirical data
  - Empirical Code, Effort, and Schedule Growth Rates derived from SRDRs
  - Guidelines and Knowledge Base that capture best practice
- Empirical Cost Risk and Uncertainty Analysis Metrics
- Calibrated SLIM-Estimate™ using most recent DoD data
- Mapping between COCOMO, SEER, True S cost drivers
- Empirical Dataset for COCOMO, True S, and SEER Calibration
- Software Maintenance Cost Model

---

**Our Plan of Attack**

**Step 1: Data Definition Approach**

1. Review literature/past research results
2. Update USC’s cost model comparisons
3. Synthesize Overall Framework
4. Identify Candidate Application Domains
5. Define Counting Rules and Standards
6. Validate Framework Via Trial Use

**Operating Environments:**

- Avionics
- Business
- Unmanned Ground
- Manned Space
- Manned Ground
- Military Mobile
- Missile and Unmanned Airborne
- Shipboard
- Telecommunications
- Unmanned Space
- Web

Comparing results from an Avionics development with that of a business project makes no sense what-so-ever.

All comparison need to take software quality into consideration to be meaningful (threshold in terms of defects/FSLOC when delivered).
Define software size measures used as input to cost estimation models, and provide guidelines for counting and normalizing software size.

- Provide rules and guidelines to convert size inputs between models so projects can be represented in all models in a consistent manner.
- Logical source statements consisting of data declarations executables
- Rules for considering statement type, how produced, origin, build, etc.
- Providing automated code counting tools adhering to definition, including initial modified-code counting
- Providing conversion guidelines for physical statements

---

**Application Domains:**

<table>
<thead>
<tr>
<th>Application Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bus</td>
</tr>
<tr>
<td>2. Command &amp; Control</td>
</tr>
<tr>
<td>3. Communications</td>
</tr>
<tr>
<td>4. Controls &amp; Displays</td>
</tr>
<tr>
<td>5. Database</td>
</tr>
<tr>
<td>6. Executive</td>
</tr>
<tr>
<td>7. Information Assurance</td>
</tr>
<tr>
<td>8. Maintenance &amp; Diagnostics</td>
</tr>
<tr>
<td>9. Mission Management</td>
</tr>
<tr>
<td>10. Mission Planning</td>
</tr>
<tr>
<td>11. Payload</td>
</tr>
<tr>
<td>12. Platform</td>
</tr>
</tbody>
</table>

---

**Software Sizing: Counting Rules & Standards**

- For adapted software, apply the parameters:
  - DM: % of design modified
  - CM: % of code modified
  - IM: % of integration required compared to integrating new code
- Normal Adaptation Adjustment Factor AAF = 0.4*DM + 0.3*CM + 0.3*IM
- Reused software has DM = CM = 0.
- DM is not applied to the size of the reused software (e.g., 70M SLOC of Windows Vista) but to the size of the other software directly interacting with it (frequently estimated using a %)
- Modified software has CM > 0. Since data indicates that the AAF factor tends to underestimate modification effort due to added software understanding effects, two other factors are used:
  - Software Understandability (SU): How understandable is the software to be modified?
  - Unfamiliarity (UNFM): How unfamiliar with the software to be modified is the person modifying it?

---

**Equivalent SLOC: Stutzke’s Counting Rules**

- For adapted software, apply the parameters:
  - DM: % of design modified
  - CM: % of code modified
  - IM: % of integration required compared to integrating new code
- Normal Adaptation Adjustment Factor AAF = 0.4*DM + 0.3*CM + 0.3*IM
- Modified software has CM > 0. Since data indicates that the AAF factor tends to underestimate modification effort due to added software understanding effects, two other factors are used:
  - Software Understandability (SU): How understandable is the software to be modified?
  - Unfamiliarity (UNFM): How unfamiliar with the software to be modified is the person modifying it?

---

**Software Effort by Activity: Counting Rules & Standards**

- Guidelines to adjust effort inputs so projects can be represented in all models in a consistent manner

---

**Step 2: Identify Candidate Sources**

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Source</th>
<th>Format</th>
<th>Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space, Ground, Air</td>
<td>Defense Cost Analysis Resource Center</td>
<td>COCOMO 2020</td>
<td>460</td>
</tr>
<tr>
<td>Space</td>
<td>NIST, WILSON, SPS</td>
<td>NASA</td>
<td>75</td>
</tr>
<tr>
<td>Space</td>
<td>NIST</td>
<td>NASA</td>
<td>36</td>
</tr>
<tr>
<td>Space</td>
<td>WPROSS</td>
<td>NASA</td>
<td>97</td>
</tr>
<tr>
<td>Space, Air, Ground</td>
<td>NonStop Systems, Raytheon</td>
<td>COCOMO, SEER</td>
<td>41</td>
</tr>
<tr>
<td>Air, Ship, Ground</td>
<td>Naval Center for Cost Analysis</td>
<td>COCOMO 2020</td>
<td>68</td>
</tr>
<tr>
<td>Air</td>
<td>Lockheed Martin</td>
<td>COCOMO</td>
<td>70</td>
</tr>
<tr>
<td>Air</td>
<td>Army Cost and Economics Analysis Center</td>
<td>COCOMO 2018</td>
<td>18</td>
</tr>
<tr>
<td>Space</td>
<td>NRO CASI</td>
<td>SEER</td>
<td>68-69</td>
</tr>
<tr>
<td>Space</td>
<td>Aristope, Space &amp; Missile System Center</td>
<td>NASA</td>
<td>TBD</td>
</tr>
<tr>
<td>Space</td>
<td>NASA-JPL</td>
<td>NASA</td>
<td>TBD</td>
</tr>
<tr>
<td>Space, Air, Ground</td>
<td>USC AFFRATED</td>
<td>COCOMO</td>
<td>TBD</td>
</tr>
</tbody>
</table>

---

**Step 3: Collect Data**

- **USC** will interview program offices and developers to obtain additional information or resolve data anomalies.
  1. SLOC reporting – logical, physical, NCSS, etc.
  2. Requirements Volatility and Adaptation
     - Modified or Reused using DM, CM, IM, SU, UNFM as appropriate
  3. Size Type – Modified, Generated, New, Re-host, COTS, etc.
  4. Effort reporting – phase and activity
  5. Quality measures – defect density, defect containment, etc.
  6. Source – in-house, third party, Prior Build, Prior Spiral, etc.
  7. Requirements Volatility – % of ESLOC reworked or deleted due to requirements volatility
  8. Programming Languages
Generated Software created with automated source code generators using different Re-Host Rehosting software from one target environment to a similar environment.

COTS (Low) Modified (High)

Operating Environment
- Environment
- Technology driven
- Characterized differently using model cost drivers

USC Research Results
Productivity comparisons/benchmarks show best results achieved when similar application domains in similar operating environments are compared using actual data that is consistent and defendable.

Productivity Benchmarks (Example 1 – Not Endorsed)

<table>
<thead>
<tr>
<th>Operating Environment</th>
<th>Application Function</th>
<th>No. Projects</th>
<th>Project Range (ESLOC)</th>
<th>No. Person-Months</th>
<th>Mean Effort (Person-Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Command &amp; Control</td>
<td>3</td>
<td>4</td>
<td>120-1500</td>
<td>168</td>
<td>45</td>
</tr>
<tr>
<td>Control &amp; Display</td>
<td>4</td>
<td>5</td>
<td>150-2000</td>
<td>195</td>
<td>39</td>
</tr>
<tr>
<td>Executive</td>
<td>6</td>
<td>7</td>
<td>200-2500</td>
<td>225</td>
<td>32</td>
</tr>
<tr>
<td>Logistic</td>
<td>8</td>
<td>9</td>
<td>250-3000</td>
<td>285</td>
<td>31</td>
</tr>
<tr>
<td>Mission Planning</td>
<td>11</td>
<td>12</td>
<td>300-3500</td>
<td>330</td>
<td>28</td>
</tr>
<tr>
<td>Simulation &amp; Modeling</td>
<td>11</td>
<td>12</td>
<td>350-4000</td>
<td>385</td>
<td>32</td>
</tr>
<tr>
<td>Test &amp; Evaluation</td>
<td>12</td>
<td>13</td>
<td>400-4500</td>
<td>430</td>
<td>30</td>
</tr>
</tbody>
</table>

Source: Defense Automated Cost Information System (DACIMS), Proprietary Sources

USC Research Results
Productivity comparisons/benchmarks show best results achieved when similar application domains in similar operating environments are compared using actual data that is consistent and defendable.

Productivity Benchmarks (Example 2 – Not Endorsed)

<table>
<thead>
<tr>
<th>Operating Environment</th>
<th>Application Function</th>
<th>No. Projects</th>
<th>Project Range (ESLOC)</th>
<th>No. Person-Months</th>
<th>Mean Effort (Person-Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission Planning</td>
<td>4</td>
<td>5</td>
<td>100-150</td>
<td>130</td>
<td>26</td>
</tr>
<tr>
<td>Executive</td>
<td>6</td>
<td>7</td>
<td>150-200</td>
<td>180</td>
<td>30</td>
</tr>
<tr>
<td>Control &amp; Display</td>
<td>8</td>
<td>9</td>
<td>200-250</td>
<td>220</td>
<td>33</td>
</tr>
<tr>
<td>Logistic</td>
<td>10</td>
<td>11</td>
<td>250-300</td>
<td>280</td>
<td>35</td>
</tr>
<tr>
<td>Simulation &amp; Modeling</td>
<td>11</td>
<td>12</td>
<td>300-350</td>
<td>330</td>
<td>38</td>
</tr>
<tr>
<td>Test &amp; Evaluation</td>
<td>12</td>
<td>13</td>
<td>350-400</td>
<td>380</td>
<td>40</td>
</tr>
</tbody>
</table>

Source: Defense Automated Cost Information System (DACIMS), Proprietary Sources

Productivity Benchmarks (Example 3 – Not Endorsed)

<table>
<thead>
<tr>
<th>Operating Environment</th>
<th>Application Function</th>
<th>No. Projects</th>
<th>Project Range (ESLOC)</th>
<th>No. Person-Months</th>
<th>Mean Effort (Person-Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission Planning</td>
<td>4</td>
<td>5</td>
<td>100-150</td>
<td>130</td>
<td>26</td>
</tr>
<tr>
<td>Executive</td>
<td>6</td>
<td>7</td>
<td>150-200</td>
<td>180</td>
<td>30</td>
</tr>
<tr>
<td>Control &amp; Display</td>
<td>8</td>
<td>9</td>
<td>200-250</td>
<td>220</td>
<td>33</td>
</tr>
<tr>
<td>Logistic</td>
<td>10</td>
<td>11</td>
<td>250-300</td>
<td>280</td>
<td>35</td>
</tr>
<tr>
<td>Simulation &amp; Modeling</td>
<td>11</td>
<td>12</td>
<td>300-350</td>
<td>330</td>
<td>38</td>
</tr>
<tr>
<td>Test &amp; Evaluation</td>
<td>12</td>
<td>13</td>
<td>350-400</td>
<td>380</td>
<td>40</td>
</tr>
</tbody>
</table>

Source: Defense Automated Cost Information System (DACIMS), Proprietary Sources

- A similar table will be provided for the other 8 operating environments.
- Descriptive statistics and project descriptions will follow each table.

Raw Datasets (Example 5 – Not Endorsed)

<table>
<thead>
<tr>
<th>Year</th>
<th>Platform</th>
<th>Description</th>
<th>No. Projects</th>
<th>Project Range (ESLOC)</th>
<th>No. Person-Months</th>
<th>Mean Effort (Person-Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>Manned</td>
<td>Executive</td>
<td>1</td>
<td>100-150</td>
<td>120</td>
<td>24</td>
</tr>
<tr>
<td>2002</td>
<td>Manned</td>
<td>Control &amp; Display</td>
<td>2</td>
<td>150-200</td>
<td>180</td>
<td>36</td>
</tr>
<tr>
<td>2003</td>
<td>Manned</td>
<td>Mission Planning</td>
<td>3</td>
<td>200-250</td>
<td>240</td>
<td>48</td>
</tr>
<tr>
<td>2004</td>
<td>Manned</td>
<td>Simulation &amp; Modeling</td>
<td>4</td>
<td>250-300</td>
<td>280</td>
<td>56</td>
</tr>
<tr>
<td>2005</td>
<td>Manned</td>
<td>Test &amp; Evaluation</td>
<td>5</td>
<td>300-350</td>
<td>330</td>
<td>66</td>
</tr>
<tr>
<td>2006</td>
<td>Manned</td>
<td>Training</td>
<td>6</td>
<td>350-400</td>
<td>380</td>
<td>72</td>
</tr>
</tbody>
</table>

Not Real Data

UNCLASSIFIED

In case you want to challenge or refine the benchmarks and metrics provided in the manual, the raw datasets (Non Proprietary Version) will be included in the appendix.
**Current Status**

- Already collected a significant amount of data
  - 345 projects – Defense Cost Analysis Resource Center
  - 240 projects – Raytheon, Lockheed, Northrop Grumman, etc.
  - Expecting space software projects from National Reconnaissance Office, NASA, and Military Prime Contractors (>100 projects)
- Analyzing over 200 projects
- Common Data Definitions and Standards
  - Initial Review (May 2009)
  - Interim Review (International COCOMO Forum, Oct 2009)

Framework and definitions are done as is the initial data analysis – detailed data analysis is in process as is adding guidelines

- Manual Publication
  - Initial Release (Sep 2009)
  - Subsequent Releases (Sep 2010, Sep 2011, Sep 2012)

**Concluding Remarks**

- Goal is to publish a manual to help analysts develop quick software estimates using empirical metrics from recent programs
- Additional information is crucial for improving data quality across DoD
- We want your input on Productivity Domains and Data Definitions
- Looking for collaborators
- Looking for peer-reviewers
- Need more data
- Need even more data

**Questions or Comments?**

- Any questions
- Any pointers
- Any feedback

Contact:

Wilson Rosa, AFCAA
(703) 604-0395
Wilson.Rosa@pentagon.af.mil