Using Regression Equations to Determine Cost Estimating Relationships for Software Development

Society of Cost Estimating and Analysis, Washington Metro Chapter
16 March 2011

William S. Barfield & Scott M. Allard

The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Abstract

The writing and testing of large-scale software is expensive and involves many substantial costs in addition to the development of the basic software itself.

A reliable method to estimate these costs is to employ Cost Estimating Relationships (CERs) applied to the various activities involved in the software development and delivery process.

CERs are regression equations typically based on normalized actual costs of prior analogous software development.

The Federal Aviation Administration (FAA) is responsible for management of our National Airspace System (NAS), which requires massive amounts of software development and maintenance.

We show the methodology, regression results and statistical confidence of new CERs available for the FAA to use in estimating costs for nine work breakdown structure elements of software development.

This is an excellent example of using applied statistics in business, industry, and government.
Using Regression Equations to Determine Cost Estimating Relationships for Software Development

- Background
- The Problem
- The Method
- The Results
- Questions, Comments
Background

• The FAA mission:
  • Provide the safest, most efficient aerospace system in the world

• Major FAA roles:
  • Develop and operate a system of air traffic control and navigation for both civil and military aircraft
  • Regulate civil aviation to promote safety
  • Research and develop the Next Generation Air Traffic Control System for the National Airspace System and for civil aeronautics
  • Ensure new, proposed, and existing NAS investments meet established business case and economic criteria

See also http://www.faa.gov/about/mission/
Two typical methodologies used to estimate investment costs for development, implementation, and maintenance of software:

- **Cost Factors** provide percentage multipliers against the historical cost of Development hardware (H/W) & S/W or Production H/W.
  - FAA factors were developed in 2002 from a survey of Department of Defense, Industry, and FAA sources
  - A problem with factors is they require a good understanding of the software and its environment to determine which factor to apply

- **Cost Estimating Relationships (CERs)** are equations derived by regression of normalized actual costs of prior analogous development, implementation, and maintenance
  - CERs may be confidently used and do not typically require detailed specifications or technical understanding
Using Regression Equations to Determine Cost Estimating Relationships for Software Development

• Background
• The Problem
• The Method
• The Results
• Questions, Comments
The Problem - Existing FAA Cost Factors

<table>
<thead>
<tr>
<th>WBS</th>
<th>DEVELOPMENT</th>
<th>Low Tendency</th>
<th>Low</th>
<th>ML</th>
<th>High</th>
<th>High Tendency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>System Engineering</td>
<td>Hardware Intensive</td>
<td>31%</td>
<td>60%</td>
<td>80%</td>
<td>Software Intensive</td>
</tr>
<tr>
<td>3.3.1.2</td>
<td>Hardware less NRE, AUC</td>
<td>All COTS</td>
<td>100%</td>
<td>150%</td>
<td>200%</td>
<td>New Development</td>
</tr>
<tr>
<td>3.3.3</td>
<td>HW/SW Integ., Ass'y, Test &amp; Chkout</td>
<td>Hardware Intensive</td>
<td>10%</td>
<td>16%</td>
<td>24%</td>
<td>Software Intensive</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Facility Planning &amp; Design</td>
<td>Software Intensive</td>
<td>2%</td>
<td>24%</td>
<td>47%</td>
<td>Hardware Intensive</td>
</tr>
<tr>
<td>3.5.1</td>
<td>System Dspt. Test &amp; Eval.</td>
<td>Minor Modification</td>
<td>5%</td>
<td>15%</td>
<td>27%</td>
<td>New Capability</td>
</tr>
<tr>
<td>3.6</td>
<td>Documentation</td>
<td>Minor Modification</td>
<td>1%</td>
<td>21%</td>
<td>27%</td>
<td>New Capability</td>
</tr>
<tr>
<td>3.7.3</td>
<td>Support &amp; Hdg Equip. Acq. (CSE)</td>
<td>Minor Modification</td>
<td>2%</td>
<td>8%</td>
<td>11%</td>
<td>New Capability</td>
</tr>
<tr>
<td>3.7.4</td>
<td>Support Fac. Const. / Conv. / Exp.</td>
<td>Software Intensive</td>
<td>10%</td>
<td>14%</td>
<td>20%</td>
<td>Hardware Intensive</td>
</tr>
<tr>
<td>3.7.5</td>
<td>Support Equip. Acq. / Mod. (PSE)</td>
<td>Minor Modification</td>
<td>1%</td>
<td>10%</td>
<td>34%</td>
<td>New Capability</td>
</tr>
<tr>
<td>3.7.7</td>
<td>Initial Spares &amp; Repair Parts Acq.</td>
<td>Software Intensive</td>
<td>1%</td>
<td>19%</td>
<td>39%</td>
<td>Hardware Intensive</td>
</tr>
<tr>
<td>3.7.8</td>
<td>Initial Training</td>
<td>Minor Modification</td>
<td>1%</td>
<td>10%</td>
<td>17%</td>
<td>New Capability</td>
</tr>
</tbody>
</table>

Factors applied to sum of (WBS 3.3.1 Hdw + WBS 3.3.2 SW) with exception of WBS 3.3.1.2
*Factor applied to WBS 3.3.5, Production, Average Unit Cost (AUC)

What is the FAA WBS?

Work breakdown structure.
A hierarchical decomposition of the work to be performed to accomplish an approved agency objective. It includes both internal and external work activities and each descending level represents an increasing definition of the work to be performed.

FAA Acquisition Management System, Appendix C.
The Problem – Buying a Lot of FAA Software

• The FAA is improving *Life Cycle Cost Estimates* in order to make informed investment decisions on the acquisition of NAS components

• Cost, the basis of an investment decision, is *difficult to estimate for software* because of the many elements, parameters and methodologies involved

• The FAA needed better method to estimate life cycle costs for WBS elements that are *significant cost drivers* of NAS software – and there is a lot of FAA software!!!

Just how much NAS software is there . . .
... a massive amount of software that controls . . .

(of course you really can’t read this!)
... a lot of North American airspace.

Snapshot of typical North American airspace at noon EST (5:10 PM Zulu) on June 10th.

Legend:

- **Yellow** – each dot is an aircraft above 1000 ft.
- **Green** – Convective weather
- **Red** – Lightning strikes
- **Ovals** – A/C in holding pattern
Using Regression Equations to Determine Cost Estimating Relationships for Software Development

- Background
- The Problem
- The Method
- The Results
- Questions, Comments
Overview of CER Development Process

• Survey FAA systems for program-specific cost data – every FAA program has a Resource Planning Document (RPD)
  – Has program baseline costs,
  – Which are time phased, and
  – Are in accordance with the FAA WBS

• Select those RPDs having (see Characterization of RPDs for details)
  – Achieved at least 70% of their life cycle (range from 3 to 15 years)
  – Non-zero Hardware and/or Software costs, and
  – Non-zero dependent variable costs (e.g. Program Management, System Engineering, etc.)

• Normalize the data to $BY09

• Develop CERs using regression analysis and statistical tests

• Validate each CER by comparing to the old FAA Factors

• Present results to FAA
# The Usable RPD Data Sets

## Potential RPD Data Points for CER Development

<table>
<thead>
<tr>
<th>WBS</th>
<th>Description</th>
<th>Non-Zero Value</th>
<th>And Non-Zero SW or HW</th>
<th>And Non-Zero SW and HW</th>
<th>Maximum Data Pts.</th>
<th>Minimum Data Pts.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Program Management</td>
<td>116</td>
<td>76</td>
<td>42</td>
<td>76</td>
<td>42</td>
</tr>
<tr>
<td>3.2</td>
<td>System Engineering</td>
<td>126</td>
<td>61</td>
<td>47</td>
<td>61</td>
<td>47</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Hardware Design and Dev</td>
<td>65</td>
<td>33</td>
<td>36</td>
<td>36</td>
<td>33</td>
</tr>
<tr>
<td>3.3.3</td>
<td>HW/SW Integration, Assembly, Test and Check</td>
<td>61</td>
<td>83</td>
<td>37</td>
<td>61</td>
<td>37</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Production Engineering</td>
<td>33</td>
<td>37</td>
<td>18</td>
<td>33</td>
<td>18</td>
</tr>
<tr>
<td>3.4</td>
<td>Phys / Airspace Infrastructure Dsn / Dev</td>
<td>37</td>
<td>29</td>
<td>17</td>
<td>29</td>
<td>17</td>
</tr>
<tr>
<td>3.5</td>
<td>Test and Evaluation</td>
<td>58</td>
<td>8</td>
<td>44</td>
<td>44</td>
<td>8</td>
</tr>
<tr>
<td>3.6</td>
<td>Data and Documentation</td>
<td>86</td>
<td>16</td>
<td>34</td>
<td>34</td>
<td>16</td>
</tr>
</tbody>
</table>

FAA WBS
Reminder – OLS Regression

• Ordinary **Least Squares** Regression - a statistical technique used to predict the behavior of a dependent variable

• A *linear regression equation* takes the form of $Y = a + bx + c$
  
  - $Y$ is the dependent variable being predicted
  - $x$ is the independent variable used to predict $Y$
  - $a$ is the $Y$-intercept of the line
  - $c$ is the regression residual

• “Best” fit is achieved when
  - the values of $a$ and $b$ are selected so that the square of the regression residuals is minimized,
  - a correlation coefficient is maximized, and
  - the sample t-test supports the null hypothesis that the IV is related to the DVs by at least $x\%$ correlation

• Assume the residuals are normally independent and identically distributed

• Caution - correlation does not prove causation!

The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Steps for Developing Each CER

1. Identify & Visualize the normalized Data Set
   - Identify independent and dependent variables
   - Graph relationship between dependent variable costs in increasing rank and the independent variables of Hardware and Software

2. Conduct statistical analysis and test linear and non-linear regressions to select a Best CER based on:
   - Sign of the Coefficients – negative and/or zero values not preferred
   - Fit statistics – significance of t-test, $R^2$, SE

3. Graphically compare Actual vs. Predicted cost
Step 1: Visualize the Data
(Example for WBS 3.5, T&E)
Step 2: Regression Results
(Example for WBS 3.5, T&E)

- **Linear, using both hardware and software as independent variables:**
  - \[ T&E = -0.01 \times SW + 0.39 \times HW, \quad R^2 = 94\%, \quad t_{SW} = -0.6, \quad t_{HW} = 15.7, \quad SE = 5.2 \]
  - Note negative SW coefficient, proximity of SW coefficient to zero, low t-statistic for SW

- **Logarithmic, using both hardware and software as independent variables:**
  - \[ T&E = SW^{0.18} \times HW^{0.48}, \quad R^2 = 81\%, \quad t_{SW} = 3.1, \quad t_{HW} = 6.9, \quad SE = 0.85 \]
  - Note that \( R^2 \) is lower than with Linear form, should be higher due to transformation reducing overall variability
  - Note that HW coefficient more powerful than SW coefficient, but data plot does not support this result

- **Linear, using software as independent variable:**
  - \[ T&E = 0.38 \times SW, \quad R^2 = 94\%, \quad t_{SW} = 20.7, \quad SE = 5.2 \]
  - Note removing HW made major improvement to t-Statistic
  - This CER judged to be the best and was selected

- **Logarithmic, using software as independent variable:**
  - \[ T&E = SW^{0.60}, \quad R^2 = 74\%, \quad t_{SW} = 9.2, \quad SE = 0.97 \]
  - Note that \( R^2 \) is lower than with Linear form, should be higher due to transformation reducing overall variability
  - \( t_{SW} \) lower than with linear form
Step 3: Actual vs. Predicted Plot  
(Example for WBS 3.5 T&E)

WBS 3.5, T&E, Actual vs. Predicted for 4 CER Forms

The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Step 3: Actual vs. Predicted Plot, detail
(Example of WBS 3.5 T&E)
Using Regression Equations to Determine Cost Estimating Relationships for Software Development

- Background
- The Problem
- The Method
- The Results
- Questions, Comments
CER Development: Summary of Results  
WBS 3 Solution Development

Cost Estimating Relationships derived for selected FAA WBS

<table>
<thead>
<tr>
<th>Ref</th>
<th>WBS</th>
<th>CER</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.1</td>
<td>Program Management</td>
<td>92%</td>
</tr>
<tr>
<td>2</td>
<td>3.2</td>
<td>System Engineering</td>
<td>94%</td>
</tr>
<tr>
<td>3</td>
<td>3.3.1</td>
<td>Hardware Design and Dev</td>
<td>67%</td>
</tr>
<tr>
<td>4</td>
<td>3.3.3</td>
<td>HW/SW Integ., Ass'y, Test and Checkout</td>
<td>74%</td>
</tr>
<tr>
<td>5</td>
<td>3.3.4</td>
<td>Production Engineering</td>
<td>78%</td>
</tr>
<tr>
<td>6</td>
<td>3.4</td>
<td>Phys / Airspace Infrastructure Dsn / Dev</td>
<td>81%</td>
</tr>
<tr>
<td>7</td>
<td>3.5</td>
<td>Test and Evaluation</td>
<td>94%</td>
</tr>
<tr>
<td>8</td>
<td>3.6</td>
<td>Data and Documentation</td>
<td>78%</td>
</tr>
<tr>
<td>9</td>
<td>3.7</td>
<td>Logistics Support</td>
<td>89%</td>
</tr>
</tbody>
</table>

Notes:

SW = Total cost from WBS 3.3.2, Software Design and Development  
HW = Total cost from WBS 3.3.5, Hardware Procurement / Production  
All CERs are based on baselined RPD costs, not on actual historical costs  
Hi/Lo dollar range and descriptive statistics for each CER is in the following slides

The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Actual vs. Predicted Chart

WBS 3.1 Program Management

CER

PM = 0.38 * SW + 0.21 * HW

$R^2$: 92%

Standard Error: +/- 29%

$t_{SW} = 7.8$

$t_{HW} = 12.2$

Data Set: 30 Programs

Total Cost Range (BY09$M):

SW: $0.10 - $369

HW: $0.11 - $268

PM: $2.1 - $172
Actual vs. Predicted Chart
WBS 3.2 System Engineering

CER

\[ SE = 0.81 \times SW + 0.06 \times HW \]
\[ R^2 = 94\% \]
Standard Error: +/- 18%
\[ t_{SW} = 20.0 \]
\[ t_{HW} = 1.7 \]

Data Set: 40 Programs
Total Cost Range (BY09$M):
- SW: $0.1 - $369
- HW: $0.1 - $268
- SE: $0.3 - $298

The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Actual vs. Predicted Chart
WBS 3.3.1 Hardware Design and Development

CER
HW Design & Dvlpmt = SW ^ 0.29 * HW ^ 0.37
R^2 = 67%
Standard Error: -54% / +68%
t_{SW} = 3.6
t_{HW} = 5.3
Data Set: 22 Programs
Total Cost Range (BY09$M):
SW: $0.0 - $11.6
HW: $0.0 - $258.5
HW Design & Dvlpmt: $0.2 - $11.8
The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Actual vs. Predicted Chart
WBS 3.3.4 Production Engineering

CER
Prod Eng = 0.45 * SW + 0.01 * HW
R² = 78%
Standard Error: 2.2
\( t_{SW} = 4.6 \)
\( t_{HW} = 2.1 \)
Data Set: 12 Programs
Total Cost Range (BY09$M):
SW: $1.0 - $16.1
HW: $0 - $268
PE: $0.1 - $9.5

The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Actual vs. Predicted Chart
WBS 3.4 Physical/Airspace Infrastructure Design & Development

CER
PAIDD = 1.14 * SW
$R^2 = 81\%$
Standard Error: 2.65
$t_{SW} = 7.2$

Data Set: 12 Programs
Total Cost Range (BY09$M):  
SW: $0.1 - $11.0  
PAIDD: $0.1 - $12.9

The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Actual vs. Predicted Chart
WBS 3.5 Test and Evaluation

CER

$TE = 0.38 \times SW$

$R^2$: 94%

Standard Error: 5.2

$t_{SW} = 20.7$

Data Set: 29 Programs

Total Cost Range (BY09$M)$:

SW: $0.1 - $237

TE: $0.3 - $98.1
Actual vs. Predicted Chart
WBS 3.6 Data and Documentation

CER
Data = 0.01 * SW + 0.01 * HW
R² = 78%
Standard Error: 0.96
\( t_{HW} = 3.8 \)
\( t_{SW} = 3.6 \)

Data Set: 21 Programs
Total Cost Range (BY09$M):
SW: $0.1 - $205
HW: $0.01 - $268
Data: $0.2 - $5.1
Actual vs. Predicted Chart
WBS 3.7 Logistics Support

CER
Log Spt = 0.07 * SW + 0.06 * HW
R^2 = 89%
Standard Error: 3.65
\(t_{SW} = 10.8\)
\(t_{HW} = 7.1\)

Data Set: 37 Programs
Total Cost Range (BY09$M):
SW: $0.01 - $369
HW: $0.1 - $368
Log. Spt: $0.04 - $30.1
Improvement of CERs over Factors

- CER results were compared to old FAA Factors:
  - Generally, results of CERs are within the range of the Factors, giving statistical confidence to using these new CERs
  - Exception: CER result for WBS 3.6 Data is at the low range of its Factor result

- CERs are derived solely from baselined FAA Programs
  - Factors used a mix of DoD, Industry, and FAA sources

- CERs are statistically derived estimators
  - Factors are simple multipliers

- CERs use separate unique inputs for Hardware and Software
  - Factors use a single combined cost input
Wrap-Up

Conclusion – Replace outdated Factors with the new CERs for estimating certain WBS 3 software development, implementation, and management costs

Result – FAA will have more accurate and defendable cost estimates to support software investment decisions

Future – Follow-up actions are identified

Lessons Learned
– Reporting of & Availability of actual & baselined spending data provides the basis for future robust cost estimates
– The development of these Cost Estimating Relationships is an excellent example of using applied statistics in business, industry, and government

Acknowledgements
Gia Grady FAA / IP&A
Al Clark Galorath Associates
Using Regression Equations to Determine Cost Estimating Relationships for Software Development

• Background
• The Problem
• The Method
• The Results
• Questions, Comments
Using Regression Equations to Determine Cost
Estimating Relationships for Software Development

BACKUP SLIDES
# FAA WBS Version 4.1
(Elements with New CERs in Yellow)

<table>
<thead>
<tr>
<th>FAA Work Breakdown Structure</th>
<th>3.5 Test and Evaluation</th>
<th>5.0 Watch Standing Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mission Analysis</td>
<td>3.5.1 System Development Test and Evaluation</td>
<td>5.0.7 Program Support</td>
</tr>
<tr>
<td>1.1 Identify Projected Demand for Services</td>
<td>3.5.2 System Operational Test and Evaluation</td>
<td>5.0.7.1 Program Planning, Authorization, Management and Control</td>
</tr>
<tr>
<td>1.2 Identify Technological Opportunities</td>
<td>3.5.3 System Independent Software Verification and Validation</td>
<td>5.0.7.2 Contract Management</td>
</tr>
<tr>
<td>1.3 Identify Projected Supply of Services</td>
<td>3.5.4 Independent Operational Test and Evaluation</td>
<td>5.0 Logistics</td>
</tr>
<tr>
<td>1.4 Mission Needs Analysis and Assessment</td>
<td>3.5.6 Data and Documentaion</td>
<td>5.0.8.1 Supply Support</td>
</tr>
<tr>
<td>1.5 Initial Requirements Definition</td>
<td>3.5.7 Logistics Support</td>
<td>5.0.8.2 Replenishment Spares</td>
</tr>
<tr>
<td>2 Investment Analysis</td>
<td>3.7.1 Logistics Support Planning</td>
<td>5.0.8.3 Repair</td>
</tr>
<tr>
<td>2.1 Initial Investment Decision</td>
<td>3.7.2 Test and Measurement Equipment Acquisition</td>
<td>5.0.8.4 Logistics Support Services</td>
</tr>
<tr>
<td>2.1.1 Planning</td>
<td>3.7.3 Support and Handling Equipment Acquisition</td>
<td>5.0.8.5 Support Equipment Maintenance</td>
</tr>
<tr>
<td>2.1.2 Analysis</td>
<td>3.7.4 Support Facilities Construction/Expansion</td>
<td>5.0.8.6 Technical Data</td>
</tr>
<tr>
<td>2.1.3 Documentation</td>
<td>3.7.5 Support Equipment Acquisition/Modification</td>
<td>5.0.8.7 Maintenance Support Facilities</td>
</tr>
<tr>
<td>2.2 Final Investment Decision</td>
<td>3.7.6 Support Facilities and Equipment Maintenance</td>
<td>5.0.8.8 Commercial Depot Logistics Service (CDLS) Contracts</td>
</tr>
<tr>
<td>2.2.1 Planning</td>
<td>3.7.7 Initial Spares and Repair Parts Acquisition</td>
<td>5.0.8.9 In-Service Training</td>
</tr>
<tr>
<td>2.2.2 Analysis</td>
<td>3.7.8 Initial Training</td>
<td>5.0.8.10 Airway Transportation System Specialists In-Service Training</td>
</tr>
<tr>
<td>2.2.3 Documentation</td>
<td>4 Implementation</td>
<td>5.0.8.11 Air Traffic Control Specialists In-Service Training</td>
</tr>
<tr>
<td>2.3 Rebaseline Decision</td>
<td>4.1 Program Management</td>
<td>5.10 Second Level Engineering</td>
</tr>
<tr>
<td>3 Solution Development</td>
<td>4.1.1 Program Planning, Authorization, Management and Control</td>
<td>5.10.1 Program Management and Infrastructure Support</td>
</tr>
<tr>
<td>3.1 Program Management</td>
<td>4.1.2 Contract Management</td>
<td>5.10.2 National Aerospace System (NAS) Field Support and Restoration</td>
</tr>
<tr>
<td>3.1.1 Program Planning, Authorization, Management and Control</td>
<td>4.1.3 Human Resources Planning and Staffing</td>
<td>5.10.3 Hardware and Software Engineering Support</td>
</tr>
<tr>
<td>3.1.2 Contract and Grant Management</td>
<td>4.2 Engineering</td>
<td>5.10.4 Configuration Management</td>
</tr>
<tr>
<td>3.2 System Engineering</td>
<td>4.3 Environmental and Occupational Safety and Health Compliance</td>
<td>5.10.5 Process Improvement</td>
</tr>
<tr>
<td>3.2.1 System Engineering Management</td>
<td>4.4 Site Selection and Acquisition</td>
<td>5.10.6 Quality Assurance</td>
</tr>
<tr>
<td>3.2.2 System Requirements and Definition</td>
<td>4.5 Construction</td>
<td>5.10.7 Information System Security</td>
</tr>
<tr>
<td>3.2.3 Analysis, design, and integration</td>
<td>4.6 Site Preparation, Installation, Test, and Checkout</td>
<td>5.10.8 Recurring NAS System Costs</td>
</tr>
<tr>
<td>3.2.4 Value Engineering</td>
<td>4.7 Joint Acceptance Inspection/Commission/Closure</td>
<td>5.10.9 Software Licenses</td>
</tr>
<tr>
<td>3.2.5 Supportability, Maintainability, and Reliability Engineering</td>
<td>4.7.1 NAS Charting and Aeronautical Information Management</td>
<td>5.11 Infrastructure Support</td>
</tr>
<tr>
<td>3.2.6 Quality Assurance Program</td>
<td>4.8 Telecommunications</td>
<td>5.11.1 Hazardous Materials Handling</td>
</tr>
<tr>
<td>3.2.7 Configuration Management</td>
<td>4.9 Implementation Training</td>
<td>5.11.2 Utilities, Building and Grounds Upkeep and Maintenance</td>
</tr>
<tr>
<td>3.2.8 Human Factors</td>
<td>5 In-Service Management</td>
<td>5.11.3 Telecommunications</td>
</tr>
<tr>
<td>3.2.9 Security</td>
<td>5.1 Preventive Maintenance/Certification</td>
<td>5.11.4 Building and Infrastructure Modernization and Improvements</td>
</tr>
<tr>
<td>3.2.10 System Safety Engineering and Management</td>
<td>5.1.1 Preventive Maintenance/Certification</td>
<td>5.11.5 Real Estate Management</td>
</tr>
<tr>
<td>3.2.11 Other System Engineering Specialties</td>
<td>5.1.2 System Management Office (SMO) Overhead</td>
<td>5.11.6 Physical Security</td>
</tr>
<tr>
<td>3.3 HW/SW Design, Development, Procurement, and Production</td>
<td>5.1.3 FAA Academy Maintenance</td>
<td>5.12 NAS Charting and Aeronautical Information Management</td>
</tr>
<tr>
<td>3.3.1 Hardware Design and Development</td>
<td>5.2 Corrective Maintenance</td>
<td>5.13 System Performance Assessment</td>
</tr>
<tr>
<td>3.3.2 Software Design and Development</td>
<td>5.2.1 Corrective Maintenance</td>
<td>5.14 System Operations</td>
</tr>
<tr>
<td>3.3.3 HW/SW Integration, Assembly, Test and Checkout</td>
<td>5.2.2 System Management Office (SMO) Overhead</td>
<td>5.15 Travel To And From Sites</td>
</tr>
<tr>
<td>3.3.4 Production Engineering</td>
<td>5.2.3 FAA Academy Maintenance</td>
<td>6 Disposition</td>
</tr>
<tr>
<td>3.3.5 Procurement/Production</td>
<td>5.3 Modifications</td>
<td>6.1 Program Management</td>
</tr>
<tr>
<td>3.4 Physical and Airspace Infrastructure Design and Development</td>
<td>5.4 Maintenance Control</td>
<td>6.2 Decommissioning</td>
</tr>
<tr>
<td>3.4.1 Facility Planning and Design</td>
<td>5.5 Technical Training</td>
<td>6.3 Engineering</td>
</tr>
<tr>
<td>3.4.2 Real Estate</td>
<td>5.5.1 Airway Transportation System Specialists Technical Training</td>
<td>6.4 Environmental Activities</td>
</tr>
<tr>
<td>3.4.3 Physical Infrastructure</td>
<td>5.5.2 Air Traffic Control Specialists Technical Training</td>
<td>6.5 Diamantine/Removal</td>
</tr>
<tr>
<td>3.4.4 Airspace Redesign</td>
<td>5.5.3 Other Staff Technical Training</td>
<td>6.6 Site Restoration/Closure</td>
</tr>
</tbody>
</table>

The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Characterization of Source RPD Data Sets

• By total number of usable RPDs:
  • Total number of RPDs obtained: 216 RPDs
  • With non-zero total cost: 208
  • With non-zero Hardware cost: 83
  • With non-zero Software cost: 76
  • With non-zero Software AND Hardware cost: 52

• Depending on CER hypothesis, maximum usable data sets range from 52 to 83 data points

• By number of RPDs with non-zero dependent variable values within applicable WBSs:
  • WBS 3.1 Program Management 116 RPDs
  • WBS 3.2 System Engineering 126
  • WBS 3.3.1 Hardware Design and Dev 65
  • WBS 3.3.3 HW/SW Integ, Asmbly, Test and Checkout 61
  • WBS 3.3.4 Production Engineering 33
  • WBS 3.4 Phys / Airspace Infrastructure Dsn / Dev 37
  • WBS 3.5 Test and Evaluation 83
  • WBS 3.6 Data and Documentation 58
  • WBS 3.7 Logistics Support 86

• Note that multiple RPDs contained non-zero dependent variable values but had zero values for one or both independent variables (i.e., Hardware and/or Software)
Approach for Development of Each CER

- Develop scatter plots of the data to observe outliers, relationships, and trends
- Determine availability of the two dependent variables
  - Software cost
  - Software schedule
- Determine candidate independent variables
  - Identify variables that can reasonably be size-estimated early in its Program life
  - Avoid using SLOC to estimate Program cost & schedule
  - Results will be correlated to 7 FAA domains, PM Staff Size, and Decision Duration (additional work pending)
- Develop hypotheses relating independent variables to dependent variables
  - H₀: IV is related to DVs by at least X% correlation
  - H₁: IV is not related to DV
- Transform selected data sets for development of non-linear CERs
- Calculate descriptive statistics (mean, standard deviation, coefficient of variation) to characterize the CER goodness-of-fit to the data
- Evaluate data set residuals and outliers
- Document the results
<table>
<thead>
<tr>
<th>Ref</th>
<th>Program</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>sheet76 ADS-B National Implementation Segment 1 and 2</td>
<td>Very low non-WBS 3.3, Hdw/SW Design costs</td>
</tr>
<tr>
<td>2</td>
<td>sheet143 ATO Strategy and Evaluation</td>
<td>Very low non-WBS 3.3, Hdw/SW Design costs</td>
</tr>
<tr>
<td>3</td>
<td>sheet162 Augmentation for GPS - Wide Area Augmentation System (WAAS) LPV Segment</td>
<td>Most costs mapped into WBS 3.3.1, Hardware Design and Dev, not a Candidate Independent Variable</td>
</tr>
<tr>
<td>4</td>
<td>sheet115 Continued General Support - Airspace Management Lab - ATDP</td>
<td>Portion of Ops costs only - not complete program</td>
</tr>
<tr>
<td>5</td>
<td>Sheet1 En Route Automation Program - En Route Auto Mod (eRAM)</td>
<td>Extremely large program, beyond range of most data - does not fit trends derived from smaller programs</td>
</tr>
<tr>
<td>6</td>
<td>sheet199 FLEX - Separation Management Approach Precision Approaches: Continued Development. Define the concepts, simulations, etc.)</td>
<td>Planning portion of program only</td>
</tr>
<tr>
<td>7</td>
<td>sheet93 FLEX - Separation Management Arrivals - Access and Environment RNAV/RNP with 3D and required time of arrival</td>
<td>Very low non-WBS 3.3, Hdw/SW Design costs</td>
</tr>
<tr>
<td>8</td>
<td>Sheet41 HAATS</td>
<td>Very low non-WBS 3.3, Hdw/SW Design costs</td>
</tr>
<tr>
<td>9</td>
<td>Sheet17 Instrument Flight Procedures Automation (IFPA)</td>
<td>Very low non-WBS 3.3, Hdw/SW Design costs</td>
</tr>
<tr>
<td>10</td>
<td>sheet189 Juneau Airport Wind System (JAWS)</td>
<td>Very low WBS 3.3, Hdw/SW Design costs relative to total WBS 3, Solution Development, costs</td>
</tr>
<tr>
<td>11</td>
<td>sheet77 NAS Voice Switch</td>
<td>Very low WBS 3.3.2, Software, and 3.3.5, Hardware costs relative to total WBS 3, Solution Development costs</td>
</tr>
<tr>
<td>12</td>
<td>sheet141 National Airspace System Interference Detection, Locating, and Mitigation (NAS IDLM)</td>
<td>Several missing values for Dependent WBS Elements (e.g., T&amp;E, Data, Logistics)</td>
</tr>
<tr>
<td>13</td>
<td>Sheet33 National Airspace System Recovery Communications (RCOM)</td>
<td>Very low non-WBS 3.3, Hdw/SW Design costs</td>
</tr>
<tr>
<td>14</td>
<td>sheet180 NEXRAD - Legacy, Icing, and Hail Algorithms</td>
<td>Algorithm development only - not complete program</td>
</tr>
<tr>
<td>15</td>
<td>sheet145 NextGen Trajectory-Based Operations (TBO)</td>
<td>Very low WBS 3.3, Hdw/SW Design costs relative to total WBS 3, Solution Development, costs</td>
</tr>
<tr>
<td>16</td>
<td>sheet181 Terminal Doppler Weather Radar (TDWR) - SLEP</td>
<td>SLEP - not complete development and implementation program</td>
</tr>
<tr>
<td>17</td>
<td>sheet182 Weather and Radar Processor (WARP) Sustain and Tech Refresh</td>
<td>Portion of Ops costs plut Tech Refresh - not complete program</td>
</tr>
</tbody>
</table>
SW and HW Costs
Used to Develop CER Comparisons to Factors

- Methodology:
  - Low = Mean of SW and HW range from PM CER (-1 Standard Deviation would result in negative values)
  - Medium = Average of SW and HW range from PM CER
  - High = +1 Standard Deviation of Medium

- Using above cost ranges yields 6 alternatives of cost range parings:

<table>
<thead>
<tr>
<th></th>
<th>Alt. 1</th>
<th>Alt. 2</th>
<th>Alt. 3</th>
<th>Alt. 4</th>
<th>Alt. 5</th>
<th>Alt. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>SW</td>
<td>$11.0</td>
<td>$11.0</td>
<td>$36.2</td>
<td>$36.2</td>
<td>$112.2</td>
<td>$112.2</td>
</tr>
<tr>
<td>HW</td>
<td>$29.7</td>
<td>$126.8</td>
<td>$29.7</td>
<td>$126.8</td>
<td>$29.7</td>
<td>$126.8</td>
</tr>
</tbody>
</table>

- Continue to next slide to see Results of CERs vs PEG comparisons -
Comparison between new CERs and Factors

The views expressed herein reflect the personal views of the authors and do not purport to be the views or positions of BAE Systems, the Federal Aviation Administration, or any other component of the Federal Government.
Follow-on Actions

• Determine name for new set of CERs
  • PEG 2.0?
  • FAACER (FAA CER)?
  • Other?

• Maintain and expand existing RPD data base using SPIRE
• Expand EVM data collection effort
• Refine existing statistical relationships with expanded RPD data and/or appropriate EVM data
• Develop CERs for Software and Hardware
• Develop additional CERs for WBS Elements in:
  • WBS 4 Implementation, and
  • WBS 5 In-Service Management
  • Utilize new CERs to develop Rough Order of Magnitude (ROM) cost estimates for elements within WBS 3 Solution Development

• Map these 9 WBS elements to the new FAA WBS version 5
CER Validation

- Limited EVM data available for CER validation
  - EVM data are not representative of completed development and fielding efforts
  - EVM data are inadequate to develop robust CERs
    - Insufficient quantity
    - Not representative of typical FAA investments
- Due to these shortcomings, the new CERs were run against only some EVM data to investigate accuracy in predicting and comparing against this limited set of actual cost data
Statistics 101

• **Non-linear least squares** is the form of least squares analysis which is used to fit a set of \( m \) observations with a model that is non-linear in \( n \) unknown parameters \((m > n)\). It is used in some forms of non-linear regression. The basis of the method is to approximate the model by a linear one and to refine the parameters by successive iterations. There are many similarities to linear least squares, but also some significant differences.

• The model function, \( f \), in **LLSQ** (linear least squares) is a linear combination of parameters of the form The model may represent a straight line, a parabola or any other linear combination of functions. In **NLLSQ** (non-linear least squares) the parameters appear as functions, such as \( \beta^2, e^{\beta x} \) and so forth. If the derivatives are either constant or depend only on the values of the independent variable, the model is linear in the parameters. Otherwise the model is non-linear.

• **A t-test** is any statistical hypothesis test in which the test statistic follows a Student's \( t \) distribution if the null hypothesis is supported. It is most commonly applied when the test statistic would follow a normal distribution if the value of a scaling term in the test statistic were known. When the scaling term is unknown and is replaced by an estimate based on the data, the test statistic (under certain conditions) follows a Student's \( t \) distribution.
### Change Log

Record of changes since release of 1st draft to FAA/IP&A on 16 August 2010

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Ver. in</th>
<th>Change</th>
<th>Ver. out</th>
<th>Date</th>
<th>By</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>1st draft is briefed to IP&amp;A (G. &amp; Grady, J Sullivan)</td>
<td>7</td>
<td>16 Aug 10</td>
<td>WB, SA</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>Added &quot;Draft&quot; to all slides</td>
<td>8</td>
<td>19 Aug 10</td>
<td>WB,SA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Added &quot;non-inclusion rational&quot; slide</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Removed references to REES.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Added &quot;next slide&quot; frame in &quot;...Step 3: Example...&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>Change title to better reflect content and results.</td>
<td>9</td>
<td>30 Aug 10</td>
<td>WB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Add &quot;Change Log&quot; slide.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct two coefficients in linear T&amp;E equation; fix $t_{SW} = 3.1$, modify linear comment. Correct logarithmic $R^2 = 74%$ from 84%.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Clarified definition of SW and HW in slide &quot;New CERs for FAA...&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Expanded the text and table to clarify the &quot;...comparisons to PEG&quot; slide</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>Version put on CD with supporting files given to Gia Grady / IP&amp;A</td>
<td>9</td>
<td>30 Aug 10</td>
<td>WB</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>Revise for MAA presentation; change title for math perspective; write new abstract and submit to MAA; coord. with Scott &amp; Katrina on 11/02.</td>
<td>10</td>
<td>15 Oct 10</td>
<td>WB</td>
</tr>
<tr>
<td>6</td>
<td>10.1</td>
<td>Added internal hyperlinks; consolidated stats with their charts; consolidate approach &amp; methodology slides to reduce number of slides;</td>
<td>11</td>
<td>1 Nov 10</td>
<td>WB</td>
</tr>
<tr>
<td>7</td>
<td>11</td>
<td>Further consolidated and simplified slides; added OLS slide; moved some slides to Backup.</td>
<td>12</td>
<td>5 Nov 10</td>
<td>WB</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>Clarified explanation of linear regression and fixed a broken internal hyperlink.</td>
<td>13</td>
<td>31 Dec 10</td>
<td>WB</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>Added on several slides hyperlinks to backup slides and returns; deleted duplicate Approach slide in backups.</td>
<td>14</td>
<td>13 Mar 11</td>
<td>wb</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>Adapted several last minute comments from MCR reviewers (Book, Covert); removed MCR logo;</td>
<td>15</td>
<td>16-25 Mar</td>
<td>wb</td>
</tr>
</tbody>
</table>