Estimating Effort and Cost in Software Projects

ISBSG A Multi-Oorganizational Project Data Repository for Project Estimation and Benchmarking

IEEE Computer Society
Western Canada Speaking Tour
October 2009

Pierre Bourque, ing., Ph.D.
École de technologie supérieure
Montreal, Canada

Ice Bag
Large sculpture recently exhibited at the Guggenheim Museum in Manhattan

artsearch.nga.gov.au/Detail-LRG.cfm?IRN=37808&View=LRG

About ÉTS

- One of Canada’s leading schools of Engineering
- ÉTS motto is ‘Engineering for Industry’.
- Roughly 5000 students, 125 professors, 25 internal senior lecturers and approximately 200 external lecturers.
- In 2005 only students completed about 2400 paid industrial internships in over 900 companies.
- A member of the Université du Québec network of establishments.
- Located in downtown Montreal, Canada
Agenda

- Principles of Credible Estimation
- Overview of Software Functional Sizing and COSMIC
- Overview of ISBSG
- Overview of the Repository
- An example of using ISBSG for Duration Estimation
- Example of a reality-check using ISBSG
- Southern Scope
- Conclusion

How do you build your estimates?

- How do you build estimates in your organization?
- What are the inputs to your estimates?
- Is your estimation process documented?
- Do you collect and use historical data?
- What is the reliability of this data?
- How do you size the amount of work to be done?
- Are the quality of your estimates very much based on the competency of a few key people in your organization?
- Do you monitor actuals versus estimates? How do you re-estimate? Do you resize?
- Is this acceptable in other walks of life, in other forms of engineering?
Mixed PUBLISHED Results for Software Cost Models...

- Panoply of software cost models
- Several studies have been conducted to assess the accuracy of these models on various databases
- However, no study has proven the superiority of any models excepted for limited applications
- Often small data samples

Underlying Principles of CREDIBLE Estimation - 1

As defined in Park et al. (94):

« Estimates are made by people, not by models. They require reasoned judgments and commitments to organizational goals that CANNOT be delegated to any AUTOMATED process ».
Underlying Principles of CREDIBLE Estimation-2

« All estimates are based on COMPARISONS. When people estimate, they evaluate how something is like, and how something is unlike, things that they or others have seen before ».

Underlying Principles of CREDIBLE Estimation-3

« Before people can estimate, they must acquire knowledge. They must collect and quantify information from other projects, so that they can place their comparative evaluations on DEMONSTRABLY SOUND FOOTINGS ».
These Principles Imply:

- To be **CREDIBLE**, an estimation process must inherently be **WHITE-BOX**.
- Software project estimation which has been plaguing the industry for years can only be solved through a **COOPERATIVE DATA COLLECTION EFFORT**.
- Much research still has to be done.

Agenda

- Principles of Credible Estimation
- Overview of Software Functional Sizing and COSMIC
- Overview of ISBSG
- Overview of the Repository
- An example of using ISBSG for Duration Estimation
- Example of a reality-check using ISBSG
- Southern Scope
- Conclusion
Acknowledgements

Many thanks to my colleague Alain Abran, Ph.D., Full Professor at École de technologie supérieure, for kindly providing the slides in this section.

Size of what ...

Project Size
The total effort, estimated or actual in work-hours or staff-months

Software size
the size of the requirements (functions) or of the deliverables (modules, lines of code)
**Software Size Measurement**

**Context...**

**TECHNICAL**

Mmm... so many programs, so many lines of code...

- Meaningful to the technical staff,
- Meaningless to management,
- Poor portability,
- Only known precisely when too late to use

**FUNCTIONAL**

Mmm... so much functionality delivered to the users...

- Meaningful to management,
- Meaningful to technical staff,
- Portable,
- Can be measured early on,
- Must be independent from effort, method or technology

---

**An analogy...**

2000 sq. ft. vs. 4000 sq. ft.

500 fsu vs. 1000 fsu

Software Functionality
The ‘Functional Size’ of software

“Functional Size: A size of software derived by quantifying the functional user requirements”

Functional Users Requirements FUR

Includes:
- data transfer (ex: Input customer data; Send control signal)
- data transformation (ex: Calculate bank interest; Derive average temperature)
- data storage (ex: Store customer order; Record temperature over time)
- data retrieval (ex: List current employees; Retrieve latest aircraft position)

Excludes:
- quality constraints (ex: usability, reliability, efficiency and portability)
- organizational constraints (ex: locations for operation, target hardware & compliance to standards)
- environmental constraints (ex: interoperability, security, privacy and safety)
- implementation constraints (ex: development language, delivery schedule)
Functional Users Requirements

- Functional User Requirements (FUR) can be identified **BEFORE** the software exists…
  (using UML diagrams, for instance)

Context...

Functional Users Requirements

- Functional User Requirements (FUR) can also be identified **AFTER** the software has been built

Context...

Functional Users Requirements
ISO standards on Functional Size

Meta-standards on functional size: ISO 14143
- Part 1: Definitions and Concepts
- Part 2: Conformity Assessment
- Part 3: Verification Guide
- Part 4: References Functional User Requirements
- Part 5: Software Domains

- Specific standards:
  - 1st generation measurement methods:
    - MKII: ISO 20698
    - IFPUG: ISO 20926
    - NESMA: ISO 24570
  - 2nd generation: COSMIC – ISO 19761

User View of Software Functional Requirements

COSMIC Overview

Users
- OR Engineered devices
- OR Other Software

Boundary
- DATA IN (ENTRY)
- STORE PERSISTENT DATA (WRITE)
- DATA MANIPULATION OR TRANSFORMATION
- RETRIEVE PERSISTENT DATA (READ)
- DATA OUT (EXIT)
COSMIC Overview

Model of Software

Functional User Requirements

Functionality = Data movements and Data manipulations

COSMIC Model of Software

F.U.R.

... only, as a reasonable approximation each data movement is assumed to have an associated constant average amount of data manipulation

Functionality = Data movements + some processing
**Definition: Data Movements**

- A base functional component which moves a single data group (type).

**Notes:**
- **1:** 4 types of data movements:
  - Entry (E),
  - Exit (X),
  - Read (R) and
  - Write (W).

- **2:** For measurement purposes, each data movement type is considered to include certain associated data manipulations.
- **3:** It is an occurrence of a data movement, not the type, that actually moves the data group occurrences (not types).

---

**Unit of measure**

- **Unit of measure:** COSMIC Function Point = CFP

- **Yardstick (by convention):**
  1 CFP = 1 Data Movement,

- **Data Movement types = same size unit = 1 CFP**
  - Entry (E),
  - Exit (X),
  - Read (R) and
  - Write (W)
Aggregation function

- COSMIC size = addition of the size units assigned to all functional processes.
- No upper limit to the functional size of a functional process.
- A functional size can thus be obtained for each functional construct (piece of software, layer, software application...).

Usages of Functional Sizing

- Estimation
- Benchmarking
- Productivity Trend Analysis
- Contract Payment Mechanisms
  - Development
  - Corrective Maintenance and Support
- Quality Tracking
Agenda

- Principles of Credible Estimation
- Overview of Software Functional Sizing and COSMIC
- Overview of ISBSG
- Overview of the Repository
- An example of using ISBSG for Duration Estimation
- Example of a reality-check using ISBSG
- Southern Scope
- Conclusion

ISBSG Mission

- “To help improve the management of IT resources, by both business and government, through the provision and exploitation of a public repositories of software engineering knowledge which are standardized, verified, recent and representative of current technologies.”
International Membership

Current membership:
• Australia, China, Finland, Germany
• India, Italy, Japan, Korea, Netherlands,
• Spain, Switzerland,
• United Kingdom, USA

ISBSG Strengths

• Not profit motivated
• Cooperative industry initiative
  – Strongly encouraged to not only use the submitted data but also to submit your OWN data
• Broad representation of IT
  – technologies, organization types, geography
• See www.isbsg.org
Agenda

- Principles of Credible Estimation
- Overview of Software Functional Sizing and COSMIC
- Overview of ISBSG
- Overview of the Repository
- An example of using ISBSG for Duration Estimation
- Example of a reality-check using ISBSG
- Southern Scope
- Conclusion

R11 Demographics

- Made available in June 2009
- Over 5000 projects
- 70% of the projects being less than 9 years old
- Projects have been submitted from 24 countries
- Projects were “built” in 29 different countries
R11 Demographics

- Many organization types
- Many business areas
- Enhancements and New developments
- Many programming languages
- …
- See www.isbsg.org for full details

The ISBSG Repository - Positioning

- Probably represents top 20% of industry?
- Primarily Business-type Applications
- Less than 200 entries with Application Type stated as being “Process-Control, Real-Time Applications”
Agenda

• Principles of Credible Estimation
• Overview of Software Functional Sizing and COSMIC
• Overview of ISBSG
• Overview of the Repository
• An example of using ISBSG for Duration Estimation
• Example of a reality-check using ISBSG
• Southern Scope
• Conclusion

Strategic Importance of Time-to-Market

• Project manager’s dream:
  – Complete and stable product requirements
  – High quality
  – Low costs
  – Short time-to-market
• Time to market or project duration is often the hardest one to pin down
• A variation of the other three have a determining effect on it
Adopted Viewpoint in this Research

ISBSG Release 4

- Release 4 (1997) contains 396 completed projects
  - Contribution from 13 countries
  - 2/3 new development, 1/3 enhancements & re-development
  - 34% Txn proc., 38% MIS, 14% office information
  - 3/4 developed in-house for internal use
  - 67% Mainframe
  - 46% 3GL, 38% 4GL
Basic Selection Criteria

- No reasonable doubts on data validity according to ISBSG screening
- Known effort, known duration and known platform

<table>
<thead>
<tr>
<th>Number of observations (n)</th>
<th>D range: 1 to 78 months</th>
<th>E range: 0.1 to 761 person-months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum value</td>
<td>1.0</td>
<td>10</td>
</tr>
<tr>
<td>Maximum value</td>
<td>78.0</td>
<td>106490</td>
</tr>
<tr>
<td>Mean value</td>
<td>10.5</td>
<td>5933</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>9.0</td>
<td>12169</td>
</tr>
<tr>
<td>Median</td>
<td>8.0</td>
<td>2228</td>
</tr>
</tbody>
</table>

312 projects satisfied all criteria

Basic Criteria

Scatter plot of effort vs. duration (n=312)
Distribution Analysis - Effort

No transform

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>Significance (α &lt;= 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness (Vb1)</td>
<td>4.87</td>
<td>Hypothesis of normality rejected</td>
</tr>
<tr>
<td>Kurtosis (b2)</td>
<td>33.69</td>
<td>Hypothesis of normality rejected</td>
</tr>
<tr>
<td>Combined (K^2)</td>
<td>344.25</td>
<td>Hypothesis of normality rejected</td>
</tr>
</tbody>
</table>

Log transformed

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>Significance (α &lt;= 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness (Vb1)</td>
<td>0.05</td>
<td>Hypothesis of normality NOT rejected</td>
</tr>
<tr>
<td>Kurtosis (b2)</td>
<td>3.26</td>
<td>Hypothesis of normality NOT rejected</td>
</tr>
<tr>
<td>Combined (K^2)</td>
<td>1.28</td>
<td>Hypothesis of normality NOT rejected</td>
</tr>
</tbody>
</table>

Distribution Analysis - Duration

No transform

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>Significance (α &lt;= 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness (Vb1)</td>
<td>2.88</td>
<td>Hypothesis of normality rejected</td>
</tr>
<tr>
<td>Kurtosis (b2)</td>
<td>15.78</td>
<td>Hypothesis of normality rejected</td>
</tr>
<tr>
<td>Combined (K^2)</td>
<td>222.49</td>
<td>Hypothesis of normality rejected</td>
</tr>
</tbody>
</table>

Log transformed

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>Significance (α &lt;= 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness (Vb1)</td>
<td>0.07</td>
<td>Hypothesis of normality NOT rejected</td>
</tr>
<tr>
<td>Kurtosis (b2)</td>
<td>3.37</td>
<td>Hypothesis of normality NOT rejected</td>
</tr>
<tr>
<td>Combined (K^2)</td>
<td>2.17</td>
<td>Hypothesis of normality NOT rejected</td>
</tr>
</tbody>
</table>
Distribution Analysis Summary

- Skewness due to the natural distribution of projects
- Normal distribution cannot be assumed without log transformation
- Log transformed data selected for modeling purposes.

Correlation Analysis (Mainframe Platform Only)

Scatter plot of Log(effort) vs. Log(duration), n=208

- Pearson correlation coef. (r): 0.72
- Significant at 0.05 confidence level
- Linear model preferred
Regression Analysis

- Regression analysis hypotheses:
  - linear relation judged adequate
  - residuals are randomly distributed
  - residuals independent from independent variable
  - variance of residuals is constant

Regression Analysis -
Selected Results
(MF Platform)

- Independent variable: Log(Effort)
- Dependent variable: Log(Duration)
- Linear regression model

<table>
<thead>
<tr>
<th>Selected results</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size (n)</td>
<td>208</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.522</td>
</tr>
<tr>
<td>$F(1,207)$</td>
<td>224,865</td>
</tr>
<tr>
<td>Prob. &gt; F</td>
<td>0.0001</td>
</tr>
<tr>
<td>Log(E) coefficient</td>
<td>0.366</td>
</tr>
<tr>
<td>Standard error of Log(E)</td>
<td>0.024</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.339</td>
</tr>
</tbody>
</table>
Regression Analysis - Residual Analysis (Mainframe Platform)

- Residuals are randomly distributed
- Residuals are independent of Log(E)

Variance of residuals is constant over the range of the dependent variable Log(D)
3.2 Regression Analysis -
The Empirical Model
(Mainframe Platform Only)

- Directly from regression results:

  \[ \log(D) = (0.366 \times \log(E)) - 0.339 \]  
  (E in person-hours)

- Converted to the usual format:

  \[ D = 0.458 \times E^{0.366} \]  
  (E in person-hours)

---

Agenda

- Principles of Credible Estimation
- Overview of Software Functional Sizing and COSMIC
- Overview of ISBSG
- Overview of the Repository
- An example of using ISBSG for Duration Estimation
- Example of a reality-check using ISBSG
- Southern Scope
- Conclusion
Actual 2008 « Reality Check »
Example*

• March 2008: Can we develop and implement this software by the end of January 2009? Is this realistic?

« REALITY CHECK » WAS PERFORMED BY PIERRE BOURQUE AND VASILE STROIAN. COSMIC FFP COUNTING WAS PERFORMED BY LEILA CHEIKHI UNDER THE SUPERVISION OF PIERRE BOURQUE. NAME OF ACTUAL COMPANY INVOLVED IN THIS EXAMPLE WILL REMAIN UNDISCLOSED.

How can we answer this question?

– Count (Estimate) the functional size of the software to be developed
– Analyze actual actual speed of delivery and productivity rates of similar projects in the ISBSG database
– Compare with required speed of delivery and productivity rate to complete the development by the planned deadline
Results of Detailed COSMIC Counting

- Counting was performed on the « actual previous software » taking into consideration a number planned changes.
- Roughly 60 hours of work for the counting
- A detailed counting report was delivered to the customer.
  - Limits of counting: ex: No documentation available
  - Assumptions when counting
  - Identification of Objects of Interest and Data Groups: 29
  - Identification of Functional Processes: 135
  - Identification of all data movements
- Count of functional size: 784 CFP or 929 CFP depending on the set of assumptions
- Estimate of functional size: 1000 CFP

Sample Data Movements in Counting

<table>
<thead>
<tr>
<th>No</th>
<th>ID</th>
<th>Description du processus</th>
<th>Événement déclencheur</th>
<th>Description des sous-processus</th>
<th>Groupe de données</th>
<th>Type du mouvement de données</th>
<th>CFP</th>
<th>LCDFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Ajouter client</td>
<td>L’utilisateur sélectionne : nouveau client</td>
<td>Le système génère un NIF pour le nouveau client</td>
<td>Données client</td>
<td>X</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.1</td>
<td>Ajouter client</td>
<td>L’utilisateur entre les infos du client</td>
<td>Données client</td>
<td>E</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.2</td>
<td>Ajouter client</td>
<td>Le système valide les données et vérifie si le client existe</td>
<td>Données client</td>
<td>R</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.3</td>
<td>Ajouter client</td>
<td>Le système crée le nouveau client</td>
<td>Données client</td>
<td>W</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.4</td>
<td>Ajouter client</td>
<td>Affichage message d'erreur</td>
<td>Messages</td>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5
Planned Effort and Duration

- Estimated Size: 1000 CFP
- Planned Duration: 13 months
- Planned Speed of Delivery: 77 CFP/Month
- Planned Effort: Not disclosed
- Planned Effort for purposes of illustration: 6400 hours (≈ 4 person years)
- Planned PDR: 6.4 hours/CFP

Initial Remarks on ISBSG

- ISBSG version 9 – 2005
- Number of projects: 3024 projects
1 : Initial Selection of Sample

- **ISBSG** Recommendation:
  Only use projects with Data Quality = A or B

- Analyze projects with Data Quality = B and Functional Size Range: 800-1200
  - Eliminate projects with Data Quality=B after analyzing outliers

2 : Functional Size: 800-1200

- Data Quality = A
  - Functional Size Range: 800-1200

Total : 28 projects
2 : Functional Size: 800-1200

2 : Functional Size: 800-1200

<table>
<thead>
<tr>
<th>28 projects</th>
<th>Min</th>
<th>Average</th>
<th>Max</th>
<th>St. Dev.</th>
<th>1st quartile</th>
<th>Median</th>
<th>3rd quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team Size</td>
<td>2.0</td>
<td>8.2</td>
<td>25.0</td>
<td>7.4</td>
<td>3.3</td>
<td>5.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Months</td>
<td>5.0</td>
<td>12.7</td>
<td>30.0</td>
<td>6.5</td>
<td>10.5</td>
<td>12.0</td>
<td>17.3</td>
</tr>
<tr>
<td>PCR</td>
<td>1.5</td>
<td>8.5</td>
<td>29.4</td>
<td>7.8</td>
<td>3.0</td>
<td>5.5</td>
<td>12.4</td>
</tr>
<tr>
<td>Size (FP)</td>
<td>896.0</td>
<td>962.8</td>
<td>1181.0</td>
<td>113.2</td>
<td>871.0</td>
<td>955.5</td>
<td>1045.0</td>
</tr>
<tr>
<td>Effort</td>
<td>1634.0</td>
<td>8155.5</td>
<td>28855.0</td>
<td>7334.8</td>
<td>2895.8</td>
<td>4945.5</td>
<td>11130.8</td>
</tr>
<tr>
<td>Plan</td>
<td>35.0</td>
<td>2922.1</td>
<td>7617.0</td>
<td>3989.5</td>
<td>803.0</td>
<td>1261.0</td>
<td>4196.0</td>
</tr>
<tr>
<td>Specify</td>
<td>168.0</td>
<td>1044.7</td>
<td>3238.0</td>
<td>997.6</td>
<td>305.0</td>
<td>915.0</td>
<td>1430.0</td>
</tr>
<tr>
<td>Build</td>
<td>1345.0</td>
<td>5094.0</td>
<td>21238.0</td>
<td>5735.7</td>
<td>1981.0</td>
<td>3079.0</td>
<td>5483.3</td>
</tr>
<tr>
<td>Test</td>
<td>100.0</td>
<td>1593.0</td>
<td>5787.0</td>
<td>1385.1</td>
<td>567.0</td>
<td>1273.0</td>
<td>1800.5</td>
</tr>
<tr>
<td>Implement</td>
<td>50.0</td>
<td>957.3</td>
<td>2946.0</td>
<td>1399.0</td>
<td>190.0</td>
<td>328.0</td>
<td>1573.0</td>
</tr>
<tr>
<td>Design</td>
<td>5.0</td>
<td>1714.8</td>
<td>25677.0</td>
<td>5355.7</td>
<td>183.5</td>
<td>600.0</td>
<td>1732.5</td>
</tr>
<tr>
<td>Size/Months</td>
<td>36.9</td>
<td>86.0</td>
<td>165.6</td>
<td>38.1</td>
<td>61.0</td>
<td>89.1</td>
<td>97.0</td>
</tr>
</tbody>
</table>
3: Year of implementation

- Data Quality = A
  and
- Functional Size: 800-1200
  et
- Year of implementation > 2000

- Total : 7 projects
3: Year of implementation > 2000

<table>
<thead>
<tr>
<th>Projects</th>
<th>Min</th>
<th>Average</th>
<th>Max</th>
<th>St Dev.</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team Size</td>
<td>3.3</td>
<td>17.1</td>
<td>25.0</td>
<td>8.7</td>
<td>16.0</td>
<td>17.0</td>
<td>24.0</td>
</tr>
<tr>
<td>Months</td>
<td>5.0</td>
<td>12.1</td>
<td>28.0</td>
<td>2.6</td>
<td>8.5</td>
<td>11.1</td>
<td>12.0</td>
</tr>
<tr>
<td>POR</td>
<td>2.5</td>
<td>10.7</td>
<td>29.4</td>
<td>9.4</td>
<td>4.8</td>
<td>6.3</td>
<td>13.5</td>
</tr>
<tr>
<td>Size (FP)</td>
<td>828.0</td>
<td>967.0</td>
<td>1181.0</td>
<td>122.2</td>
<td>808.0</td>
<td>957.0</td>
<td>1021.5</td>
</tr>
<tr>
<td>Effort</td>
<td>206.0</td>
<td>1041.2</td>
<td>2355.0</td>
<td>947.3</td>
<td>4091.5</td>
<td>5482.0</td>
<td>13680.0</td>
</tr>
</tbody>
</table>

Plan     | 1725.0 | 4426.0 | 7617.0| 4950.0   | 2044.0       | 4455.0 | 6826.4       |
Specify  | 168.0  | 1009.4 | 3238.0| 1253.7   | 246.0        | 777.0  | 973.0        |
Build    | 1960.0 | 6267.2 | 21280.0| 7490.5   | 2202.3       | 3086.0 | 5435.6       |
Test     | 755.0  | 1333.6 | 1987.0| 496.2    | 975.0        | 1320.0 | 1620.0       |
Implement| 114.0  | 1396.0 | 2548.0| 1770.6   | 865.5        | 1356.0 | 2022.0       |
Design   | 5.0    | 1774.8 | 35677.0| 5355.7   | 183.5        | 600.0  | 1232.5       |
Size/Months| 35.0 | 100.3  | 165.6 | 44.0     | 70.7         | 96.5  | 122.9        |

4: New Development Projects Only

- Data Quality: A
- Functional Size: 800-1200
- New Development projects only

TOTAL: 19 projects
### 4: New Development Projects Only

#### Data Table

<table>
<thead>
<tr>
<th>Project Type</th>
<th>Min</th>
<th>Average</th>
<th>Max</th>
<th>St. Dev.</th>
<th>1st quantile</th>
<th>Median</th>
<th>3rd quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team Size</td>
<td>2.0</td>
<td>6.8</td>
<td>25.0</td>
<td>6.5</td>
<td>3.0</td>
<td>5.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Months</td>
<td>5.0</td>
<td>15.1</td>
<td>30.0</td>
<td>7.9</td>
<td>9.5</td>
<td>12.0</td>
<td>20.0</td>
</tr>
<tr>
<td>PDR</td>
<td>2.2</td>
<td><strong>9.8</strong></td>
<td>29.4</td>
<td><strong>8.6</strong></td>
<td><strong>4.1</strong></td>
<td><strong>6.5</strong></td>
<td><strong>12.9</strong></td>
</tr>
<tr>
<td>Size (FP)</td>
<td>828.0</td>
<td>928.0</td>
<td>1164.0</td>
<td>100.5</td>
<td>864.0</td>
<td>957.0</td>
<td>1000.0</td>
</tr>
<tr>
<td>Effort</td>
<td>2078.0</td>
<td>9438.0</td>
<td>28855.0</td>
<td>8307.7</td>
<td>3516.0</td>
<td>5800.0</td>
<td>13276.5</td>
</tr>
<tr>
<td>Plan</td>
<td>375.0</td>
<td>3833.4</td>
<td>7617.0</td>
<td>3254.5</td>
<td>1243.0</td>
<td>3120.0</td>
<td>6812.0</td>
</tr>
<tr>
<td>Specify</td>
<td>168.0</td>
<td>1253.3</td>
<td>3230.0</td>
<td>978.0</td>
<td>659.8</td>
<td>1117.5</td>
<td>1595.0</td>
</tr>
<tr>
<td>Build</td>
<td>1345.0</td>
<td>6033.3</td>
<td>21238.0</td>
<td>6295.9</td>
<td>2080.0</td>
<td>3283.0</td>
<td>6153.0</td>
</tr>
<tr>
<td>Test</td>
<td>100.0</td>
<td>1696.5</td>
<td>5787.0</td>
<td>1814.6</td>
<td>661.0</td>
<td>1127.5</td>
<td>2072.5</td>
</tr>
<tr>
<td>Implement</td>
<td>90.0</td>
<td>733.2</td>
<td>2546.0</td>
<td>1213.7</td>
<td>326.0</td>
<td>328.0</td>
<td>498.0</td>
</tr>
<tr>
<td>Design</td>
<td>5.0</td>
<td>1774.8</td>
<td>35677.0</td>
<td>5355.7</td>
<td>183.5</td>
<td>600.0</td>
<td>1232.5</td>
</tr>
<tr>
<td>Size/Months</td>
<td>31.9</td>
<td><strong>82.9</strong></td>
<td>165.6</td>
<td><strong>43.4</strong></td>
<td><strong>52.1</strong></td>
<td><strong>71.3</strong></td>
<td><strong>97.7</strong></td>
</tr>
</tbody>
</table>
### Summary of Analysis of Product Delivery Rates (Hours/CFP)

<table>
<thead>
<tr>
<th>Selected Subset</th>
<th>Sample Size</th>
<th>Average</th>
<th>Std Dev.</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>800–1200</td>
<td>28</td>
<td>8.5</td>
<td>7.8</td>
<td>3.0</td>
<td>5.5</td>
<td>12.4</td>
</tr>
<tr>
<td>Year of implem. &gt; 2000</td>
<td>7</td>
<td>10.7</td>
<td>9.4</td>
<td>4.8</td>
<td>6.3</td>
<td>13.5</td>
</tr>
<tr>
<td>Only new devel. projects</td>
<td>19</td>
<td>9.8</td>
<td>8.6</td>
<td>4.1</td>
<td>5.5</td>
<td>12.9</td>
</tr>
</tbody>
</table>

Planned PDR of Project: 6.4 hours/CFP

### Summary of Analysis of Speed of Delivery (CFP/month)

<table>
<thead>
<tr>
<th>Selected Subset</th>
<th>Sample Size</th>
<th>Average</th>
<th>Std Dev.</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>800–1200</td>
<td>28</td>
<td>86</td>
<td>38.1</td>
<td>61.0</td>
<td>80.1</td>
<td>97.0</td>
</tr>
<tr>
<td>Year of implem. &gt; 2000</td>
<td>7</td>
<td>100.3</td>
<td>44.0</td>
<td>79.7</td>
<td>96.5</td>
<td>122.9</td>
</tr>
<tr>
<td>Only new devel. projects</td>
<td>19</td>
<td>82.9</td>
<td>43.4</td>
<td>52.1</td>
<td>71.3</td>
<td>97.7</td>
</tr>
</tbody>
</table>

Planned Speed of Project: 77 CFP/month
Conclusion « Reality Check »

- Planned duration and effort can be viewed as realistic but in reality are quite aggressive in terms of the overall industry
  - ISBSG data most probably represents the upper portion of the industry
- High standard deviation indicates significant variability in productivity and in speed of delivery
- Need to aggressively manage scope of project
- Need to aggressively manage delays in project
- Need to apply software engineering best practices

Agenda

- Principles of Credible Estimation
- Overview of Software Functional Sizing and COSMIC
- Overview of ISBSG
- Overview of the Repository
- An example of using ISBSG for Duration Estimation
- Example of a reality-check using ISBSG
- Southern Scope
- Conclusion
Discussion and Review of Southern Scope


Underlying Principles of CREDIBLE Estimation - 1

As defined in Park et al. (94):

« Estimates are made by people, not by models. They require reasoned judgments and commitments to organizational goals that CANNOT be delegated to any AUTOMATED process ».

Underlying Principles of CREDIBLE Estimation-2

« All estimates are based on COMPARISONS. When people estimate, they evaluate how something is like, and how something is unlike, things that they or others have seen before ».

Underlying Principles of CREDIBLE Estimation-3

« Before people can estimate, they must acquire knowledge. They must collect and quantify information from other projects, so that they can place their comparative evaluations on DEMONSTRABLY SOUND FOOTINGS». 
Conclusion

• Software sizing is different from estimation
• Software functional sizing is rigorously defined and used extensively around the world
• ISBSG data is available and can be analyzed by everyone.
• The steps taken to derive the example model and the assumptions behind it are known and the accuracy for this sample is published.
• Allows more intelligent tradeoffs and informed choices between various scenarios.

Conclusion

• Development of demonstrably sound quantitative models is a difficult and key problem in this industry.
• Can only be solved with an inherently white-box approach.
• Credibility of results depends entirely on the transparency of the method, data, definitions and assumptions that were used to derive this estimate.
References