Estimating Effort and Cost in Software Projects - ISBSG A Multi-Organizational Project Data Repository for Project Estimation And Benchmarking

IEEE Computer Society
Long Island, NY
29 September 2009

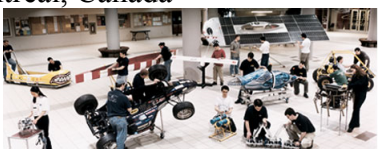
Pierre Bourque, Ph.D.
École de technologie supérieure
Montreal, Canada

Ice Bag

ISBSG

About ÉTS

- One of Canada's leading schools of Engineering
- ÉTS motto is 'Engineering for Industry'.
- Over 4500 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



ETS Software Engineering Lab.

- Recognized world-wide in software engineering for:
  - Building consensus in software engineering
  - Leadership of world-wide initiatives
  - Strong applied research focus
- Over 800 publications and contributions over roughly the past 15 years
- www.gelog.etsmtl.ca

Agenda

- Principles of Credible Estimation
- Overview of Software Functional Sizing
- 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 ».

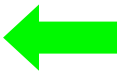
Underlying Principles of CREDIBLE Estimation-3

« 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 
- 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

Context...

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)

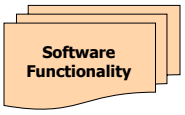
The 'Functional Size' of software

- **" Functional Size : A size of software derived by quantifying the functional user requirements"**

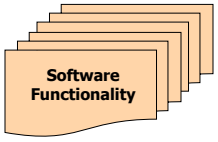
An analogy...

2000 sq. ft.

4000 sq. ft.



500 cfsu

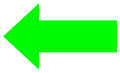


1000 cfsu

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
- 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
- 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

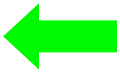
- Many organization types
- Many business areas
- Enhancements and New developments
- Many programming languages
- …
- See www.isbsg .org for full details

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

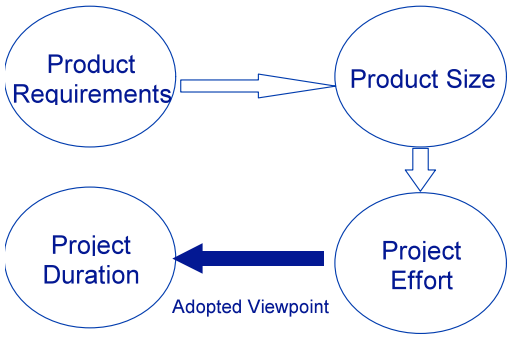
- Principles of Credible Estimation
- Overview of Software Functional Sizing
- Overview of ISBSG
- Overview of the Repository
- An example of using ISBSG for Duration Estimation
- Example of a reality-check using ISBSG
- Southern Scope
- Conclusion



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 developped 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

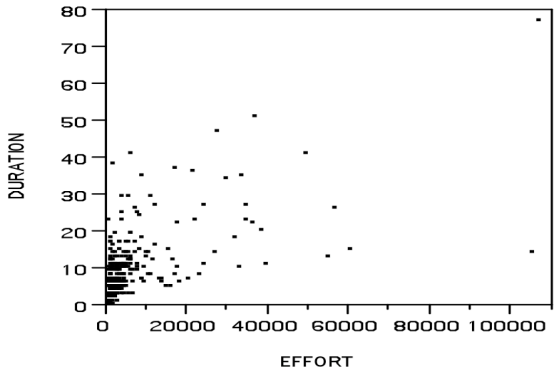
	Duration (D) in calendar months	Effort (E) in person-hours
Number of observations (n)	312	312
Minimum value	1,0	10
Maximum value	78,0	106 480
Mean value	10,5	5 933
Standard deviation	9,0	12 169
Median	8,0	2 228

D range: 1 to 78 months
E range: 0,1 to 761 person-months

↗ 312 projects satisfied all criteria

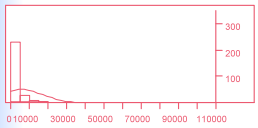
Basic Criteria

Scatter plot of effort vs. duration (n=312)



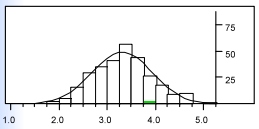
Distribution Analysis - Effort

No transform



Statistic	Value	Significance (α <= 0,05)
Skewness ($\sqrt{b_1}$)	4,87	Hypothesis of normality rejected
Kurtosis (b_2)	33,69	Hypothesis of normality rejected
Combined (K^2)	344,25	Hypothesis of normality rejected

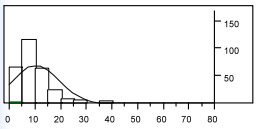
Log transformed



Statistic	Value	Significance (α <= 0,05)
Skewness ($\sqrt{b_1}$)	0,05	Hypothesis of normality NOT rejected
Kurtosis (b_2)	3,26	Hypothesis of normality NOT rejected
Combined (K^2)	1,28	Hypothesis of normality NOT rejected

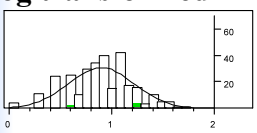
Distribution Analysis - Duration

No transform



Statistic	Value	Significance (α <= 0,05)
Skewness ($\sqrt{b_1}$)	2,88	Hypothesis of normality rejected
Kurtosis (b_2)	15,78	Hypothesis of normality rejected
Combined (K^2)	222,49	Hypothesis of normality rejected

Log transformed



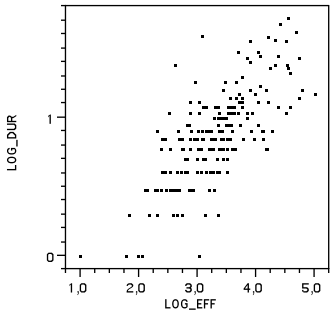
Statistic	Value	Significance (α <= 0,05)
Skewness ($\sqrt{b_1}$)	0,07	Hypothesis of normality NOT rejected
Kurtosis (b_2)	3,37	Hypothesis of normality NOT rejected
Combined (K^2)	2,17	Hypothesis of normality NOT rejected

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

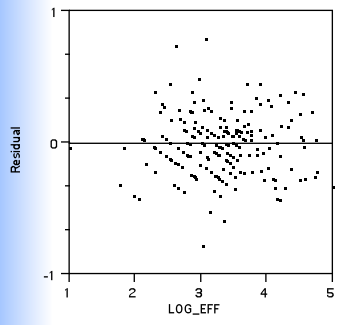
- **Regression analysis hypotheses:**
  - linear relation judged adequate
  - residuals are randomly distributed
  - residuals independent from independent variable
  - variance of residuals is constant

Regression Analysis - Selected Results (MF Platform)

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

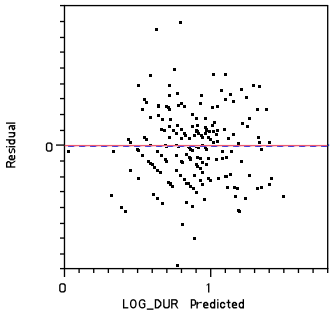
Selected results	Value
Sample size (n)	208
R^2	0,522
F(1,207)	224,865
Prob. > F	0,0001
Log(E) coefficient	0,366
Standard error of Log(E)	0,024
Constant	-0,339

Regression Analysis - Residual Analysis (Mainframe Platform)



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

Regression Analysis - Residual Analysis (Mainframe Platform Only)



Variance of residuals is constant over the range of the dependent variable Log(D)

3.2 Regression Analysis - The Empirical Model (Mainframe Platform Only)

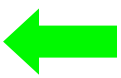
- **Directly from regression results:**

  Log(D) = (0,366 * Log(E)) - 0,339 (E in person-hours)

- **Converted to the usual format:**

  D = 0,458 * E ^{0,366} (E in person-hours)

Agenda

- Principles of Credible Estimation
- Overview of Software Functional Sizing
- 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

No	ID	Description du processus	Événement déclencheur	Description des sous- processus	Groupe de données	Type du mouvement de données	CFP	ΣCFP
2.2 2.2.1 2.2.2	1	Ajouter client	L'utilisateur sélectionne : nouveau client	Le system génère un NIP pour le nouveau client	Données client	X	1	
				L'utilisateur entre les infos du client	Données client	E	1	
				Le system valide les données et vérifie si le client existe	Données client	R	1	
				Le système crée le nouveau client	Données client	W	1	
				Affichage message d'erreur	Messages	X	1	
								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** Recommandation:
  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

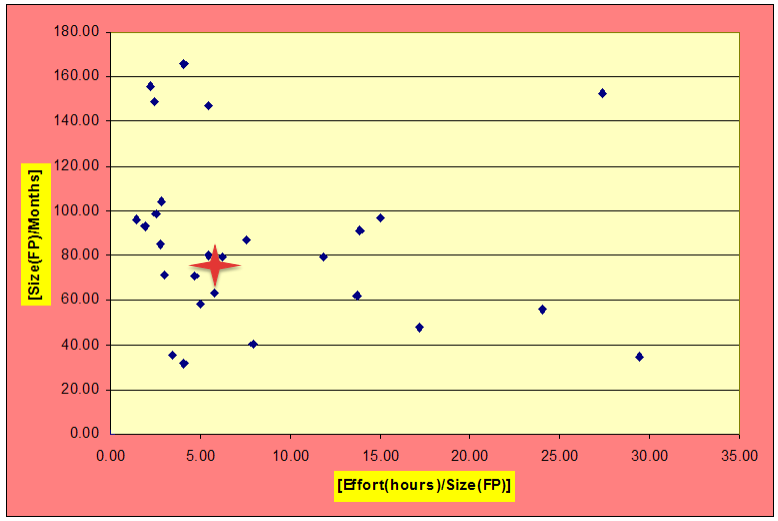
- Data Quality = A

and

- Functional Size Range: 800-1200

Total : 28 projects

2 : Functional Size: 800-1200



Planned project

2 : Functional Size : 800-1200

28 projects	Min	Average	Max	St. Dev.	1st quartile	Median	3rd quartile
Team Size	2.0	8.2	25.0	7.4	3.3	5.0	7.0
Months	5.0	13.7	30.0	6.9	10.5	12.0	17.3
PDR	1.5	8.5	29.4	7.8	3.0	5.5	12.4
Size (FP)	805.0	962.8	1181.0	113.2	871.0	955.5	1045.8
Effort	1634.0	8155.5	28855.0	7534.8	2895.8	4949.5	11130.8
Plan	35.0	2922.1	7617.0	3099.5	809.0	1253.0	4966.0
Specify	168.0	1049.7	3238.0	907.6	305.0	915.0	1430.0
Build	1345.0	5094.8	21238.0	5635.7	1981.0	3079.0	5483.3
Test	100.0	1513.8	5787.0	1589.1	567.0	1279.0	1808.5
Implement	50.0	957.3	2946.0	1269.0	190.0	328.0	1573.0
Design	5.0	1774.8	35677.0	5355.7	183.5	600.0	1232.5
Size/Months	31.9	86.0	165.6	38.1	61.0	80.1	97.0

3: Year of implementation

- Data Quality = A

and

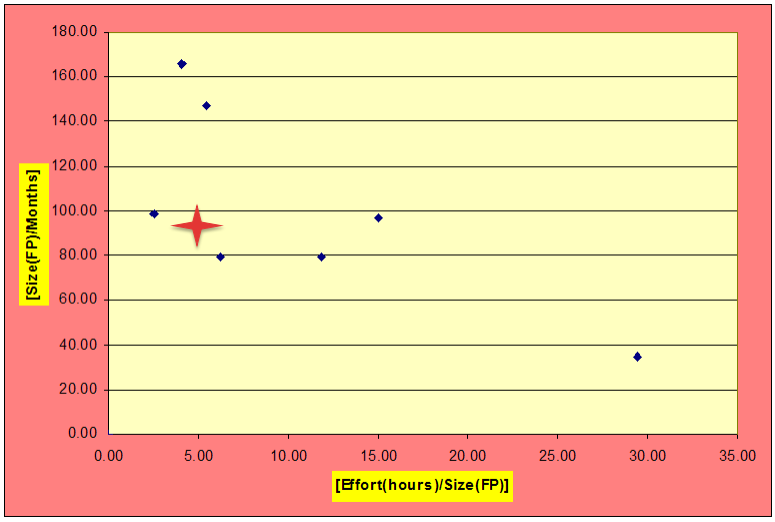
- Functional Size: 800-1200

et

❑Year of implementation >2000

- Total : 7 projects

3: Year of Implementation > 2000



3: Year of implementation > 2000

7 projects	Min	Average	Max	St. Dev.	1st quartile	Median	3rd quartile
Team Size	3.3	17.1	25.0	8.7	16.0	17.0	24.0
Months	5.0	12.1	28.0	7.6	8.5	11.0	12.0
PDR	2.5	10.7	29.4	9.4	4.8	6.3	13.5
Size (FP)	828.0	967.0	1181.0	122.2	880.0	957.0	1021.5
Effort	3006.0	10412.3	28855.0	9417.3	4091.5	5482.0	13680.0
Plan	1253.0	4435.0	7617.0	4500.0	2844.0	4435.0	6026.0
Specify	168.0	1080.4	3238.0	1253.7	246.0	777.0	973.0
Build	1960.0	6267.2	21238.0	7490.5	2282.3	3086.0	5435.5
Test	755.0	1333.6	1997.0	496.2	976.0	1320.0	1620.0
Implement	144.0	1396.0	2648.0	1770.6	865.5	1396.0	2022.0
Design	5.0	1774.8	35677.0	5355.7	183.5	600.0	1232.5
Size/Month	35.0	100.3	165.6	44.0	79.7	96.5	122.9

4 : New Development Projects Only

- Data Quality : A

and

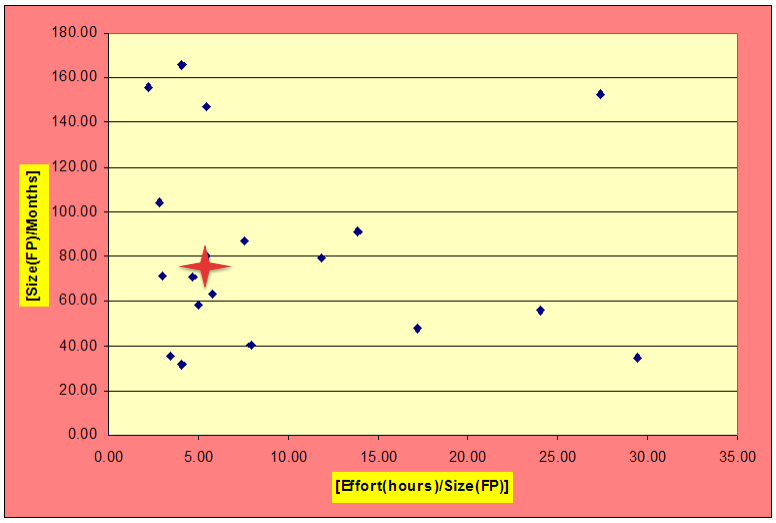
- Functional Size : 800-1200

and

❑ New Development projects only

TOTAL : 19 projects

4: New Development Projects Only



4: New Development Projects Only

19 projects	Min	Average	Max	St. Dev.	1st quartile	Median	3rd quartile
Team Size	2.0	6.8	25.0	6.5	3.0	5.0	6.0
Months	5.0	15.1	30.0	7.9	9.5	12.0	20.0
PDR	2.2	9.8	29.4	8.6	4.1	5.5	12.9
Size (FP)	828.0	959.8	1164.0	100.5	884.0	957.0	1000.0
Effort	2078.0	9438.0	28855.0	8307.7	3516.0	5800.0	13276.5
Plan	375.0	3833.4	7617.0	3254.5	1243.0	3120.0	6812.0
Specify	168.0	1253.3	3238.0	978.0	659.8	1117.5	1595.0
Build	1345.0	6033.3	21238.0	6295.9	2080.0	3283.0	6153.0
Test	100.0	1696.5	5787.0	1814.6	661.0	1127.5	2072.5
Implement	50.0	793.2	2946.0	1215.7	328.0	328.0	498.0
Design	5.0	1774.8	35677.0	5355.7	183.5	600.0	1232.5
Size/Months	31.9	82.9	165.6	43.4	52.1	71.3	97.7

Summary of Analysis of Product Delivery Rates (Hours/CFP)

Selected Subset	Sample Size	Average	Std Dev.	1st Quartile	Median	3rd Quartile
800-1200	28	8.5	7.8	3.0	5.5	12.4
Year of implem. > 2000	7	10.7	9.4	4.8	6.3	13.5
Only new devel. projects	19	9.8	8.6	4.1	5.5	12.9

Planned PDR of Project: 6.4 hours/CFP

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

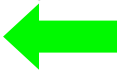
Selected Subset	Sample Size	Average	Std Dev.	1st Quartile	Median	3rd Quartile
800-1200	28	86	38.1	61.0	80.1	97.0
Year of implem. > 2000	7	100.3	44.0	79.7	96.5	122.9
Only new devel. projects	19	82.9	43.4	52.1	71.3	97.7

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
- 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

- http://www.egov.vic.gov.au/index.php?env=-innews/detail:m1816-1-1-8-s-0:n-832-1-0--

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 ».

These principles are taken from: R. E.Park, W. B. Goethert and J.T. Webb. Software Cost and Schedule Estimating: A Process Improvement Initiative. Pittsburgh, PA Software Engineering Institute,1994.

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
- 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.

Refererences

- Abran, A., Ndiaye, I., and Bourque, P. Evaluation of a Black-Box Estimation Tool: A Case Study, *Software Process: Improvement and Practice*, 2007, 12(2): 199-218.
- P. Bourque, S. Oligny, A. Abran, Developing Project Duration Models in Software Engineering, Journal of Computer Science and Technology, Springer, Vol 22, No 3, 2007.
- S.D. Conte, H.E.Dunsmore, V.Y. Shen, Software engineering metrics and models. Menlo Park: The Benjamin/Cummings Publishing Company, Inc. 1986.
- R. E.Park, W. B. Goethert and J.T. Webb. Software Cost and Schedule Estimating: A Process Improvement Initiative. Pittsburgh, PA Software Engineering Institute,1994.