Estimating Effort and Cost in Software Projects

-

ISBSG A Multi-Organizational Project Data Repository for Project Estimation And Benchmarking

IEEE Computer Society
Buenos Aires, Argentina
9 and 10 May 2007

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

Ice Bag

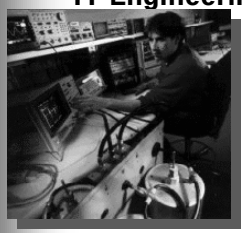
About ÉTS

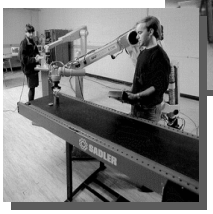
- One of Canada's leading schools of Engineering
- ÉTS motto is 'Engineering for Industry'.
- Over 4500 students, 125 professors, 25 internal senior lecturers and approximately 200 external lecturers.
- In 2005 only students completed about 2400 paid industrial internships in over 900 companies.
- A member of the Université du Québec network of establishments.
- Located in downtown Montreal, Canada

ÉTS Engineering Programs

Undergraduate Programs (7)

- **Construction Engineering**
- **Production Engineering**
- **Electrical Engineering**
- **Mechanical Engineering**
- **Logistics and Operations Engineering**
- **Software Engineering**
- **IT Engineering**





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

Software Engineering

- Sample of R & D topics:
  - Guide to the Software Engineering Body of Knowledge (SWEBOK)
  - Second generation of functional size methods: COSMIC
  - Estimation models
  - Quality engineering
  - Risk of measurement programs
  - Metrology
  - Software Maintenance Maturity Model
  - International Standards Development

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
- 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?
- Do you collect data that is never used?
- What is the reliability of this data?
- Do you track your estimates?
- 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?

What is software engineering ?

- IEEE 610.12:

  (1) The application of a systematic, disciplined, **quantifiable** approach to the development, operation, and maintenance of software; that is, the application of engineering to software.

  (2) The study of approaches as in (1).

Many Commercial Estimation Tools on the Market

- **Many well-known commercial tools:**
  - **Checkpoint**
  - **Project Workbench**
  - **PQM Plus**
  - **SLIM**
  - **…**
- See O. Mendes, A. Abran and P. Bourque. *Function Point Tool Market Survey*. Université du Québec à Montréal,1996 Available at *www.lrgl.uqam.ca/publications/pdf/204.pdf*

What is the result?

- Lots of numbers, graphs, nice slide presentation,
- Industry data often not verifiable,
- "Black box" approach,
- Bottom line: everybody sees what they want to see…
  - See 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.

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:

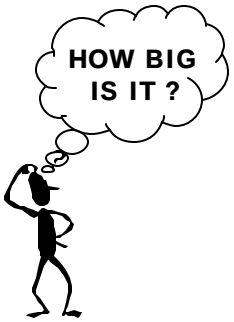
- To be **CREDIBLE**, an estimation process must inherently be **WHITE-BOX.**
- Software project estimation which has been plaguing the industry for years can only be solved through a COOPERATIVE DATA COLLECTION EFFORT.
- Much research still has to be done.

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
- Conclusion

Size of what …

Context...



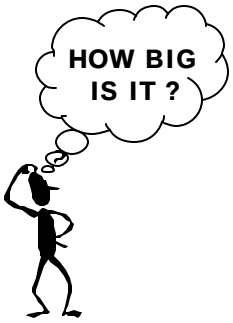
Project Size

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



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**

The 'Functional Size' of software

- **ISO/IEC/JTC1/SC7 Standard #14143 definition**:

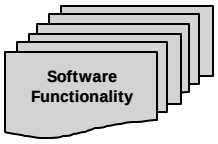
  " Functional Size : A size of software derived by quantifying the functional user requirements"

An analogy...

2000 sq. ft.

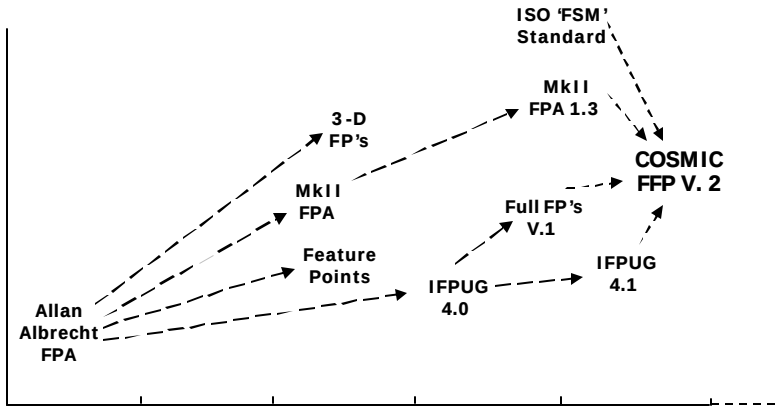
4000 sq. ft.

500 cfsu



1000 cfsu

Very Significant Amount of Work on Software Functional Size?"

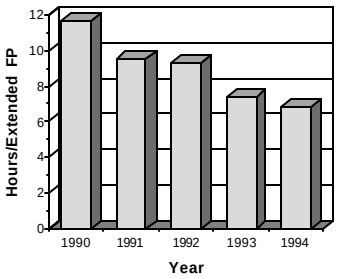


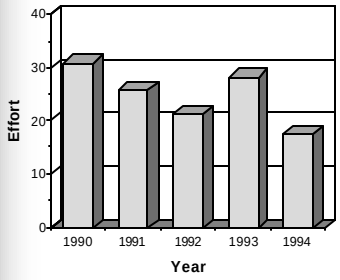
Usages of Software Sizing

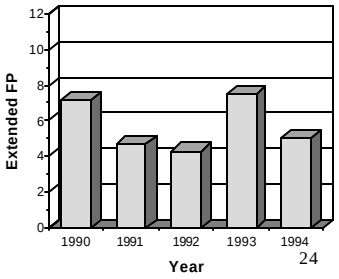
- Estimation
- Benchmarking
- Productivity Trend Analysis
- Contract Payment Mechanisms
  - Development
  - Corrective Maintenance and Support
- Quality Tracking

Productivity Trend Analysis

FP Unit Costs







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
- 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
- Allows DIRECT ACCESS TO SUBMITTED PROJECT DATA
- Broad representation of IT
  - technologies, organisation 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
- Conclusion

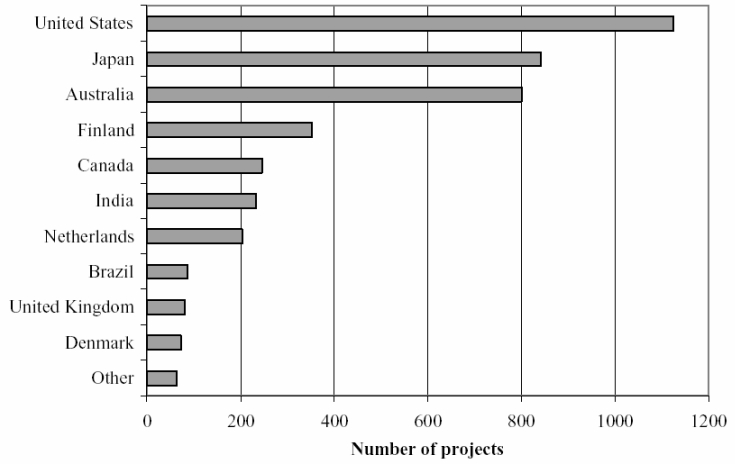
ISBSG Release 8

- Demonstration of ISBSG Release 8 Data Set

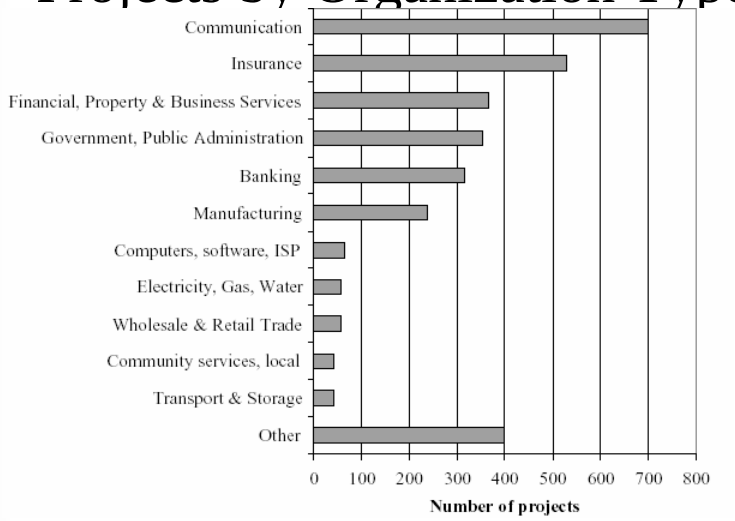
R10 Demographics

- Made available in January 2007
- Over 4000 projects
- 60 % or projects are less than 7 years old

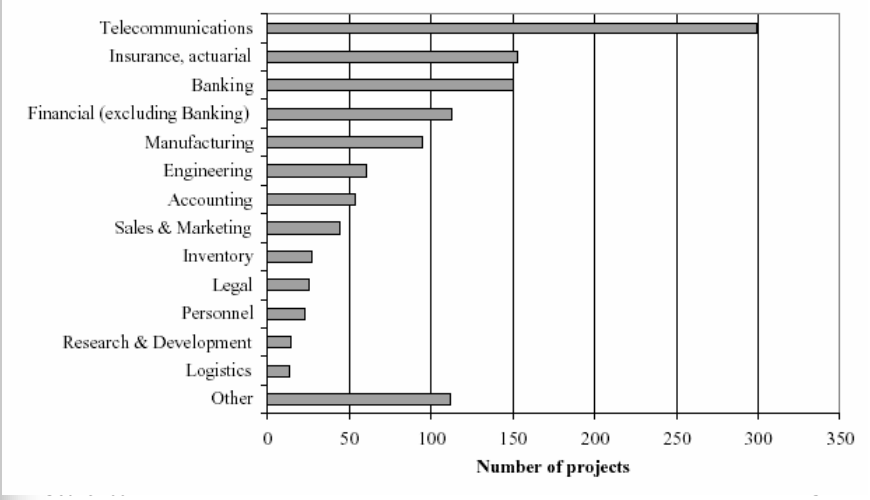
Projects by Country



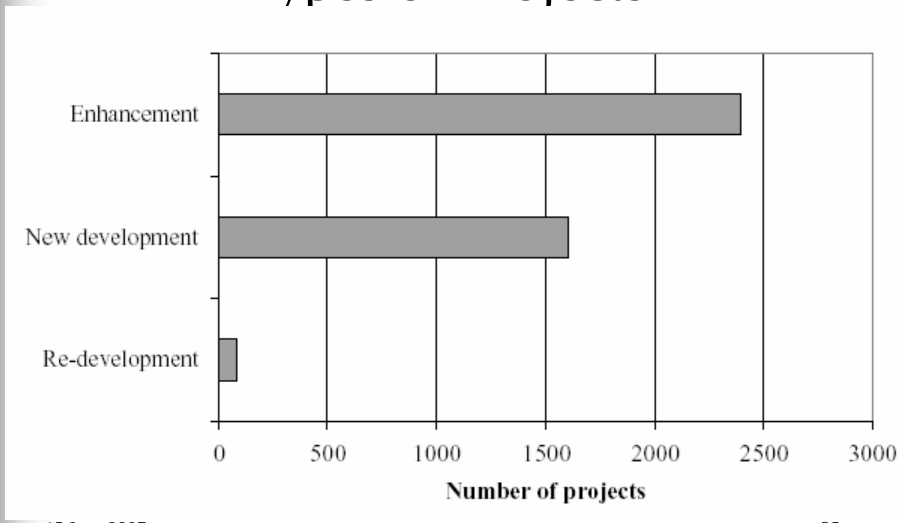
Projects by Organization Type



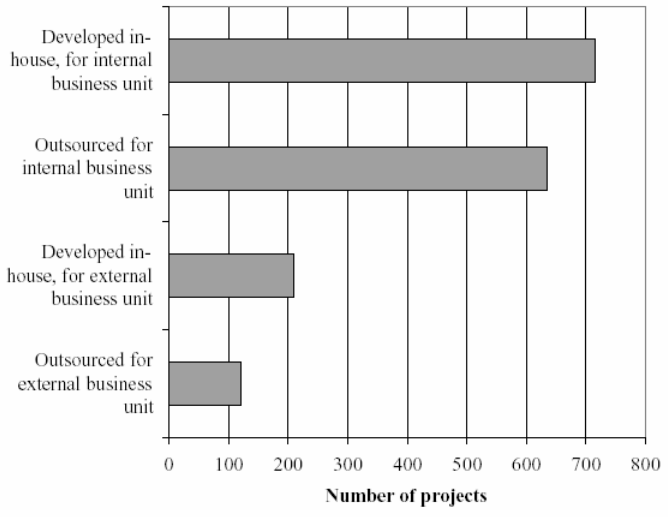
Projects by Business Area



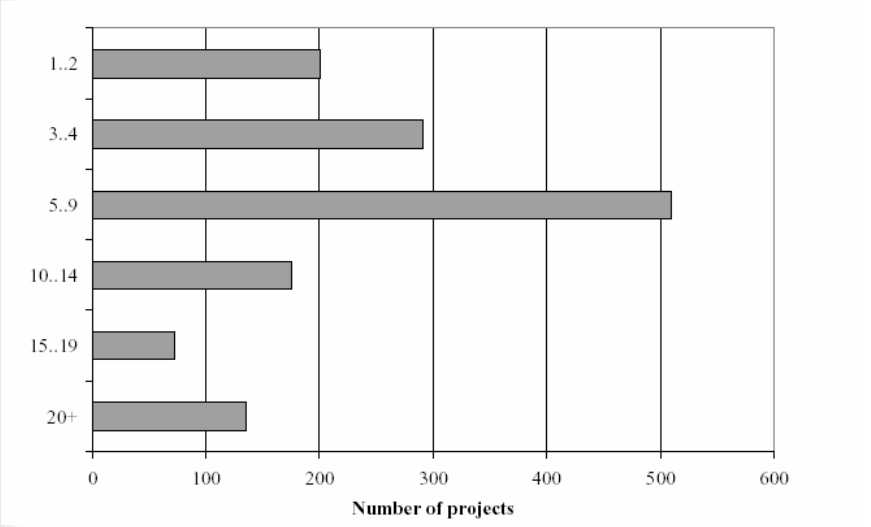
Types of Projects



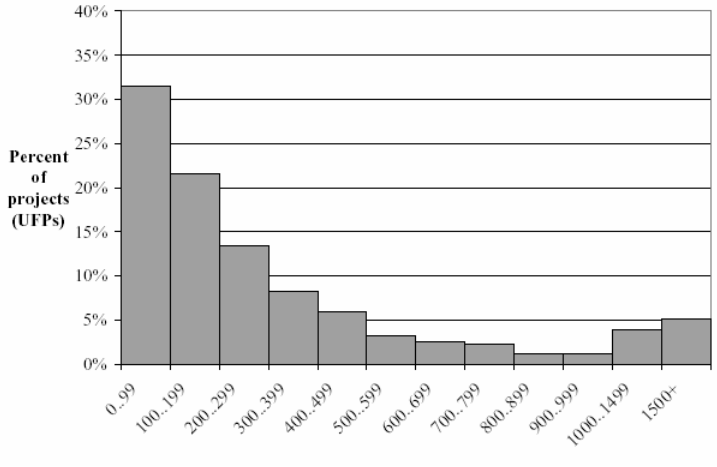
Intended Market



Maximum Team Size



Range of Sizes



The ISBSG Repository - Positioning

- Probably represents top 20% of industry
- Primarily MIS Applications (to be verified in recent releases)

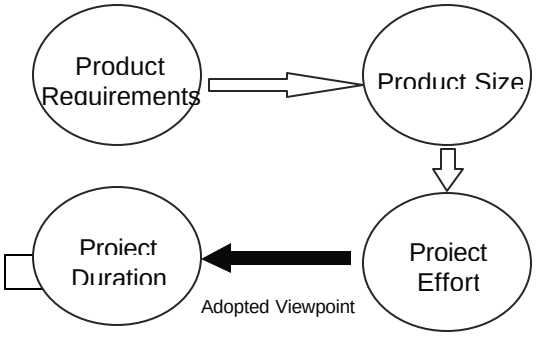
Agenda

- Principles of Credible Estimation
- Overview of Software Functional Sizing
- Overview of ISBSG
- Overview of the Repository
- An example of using ISBSG for Duration Estimation
- 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



2. Selecting a Data Sample

2.1 ISBSG release 4 (1997)

2.2 Basic selection criteria

2.3 Distribution analysis

- Effort
- Duration
- Summary

2.1 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

2.2 Basic Selection Criteria

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

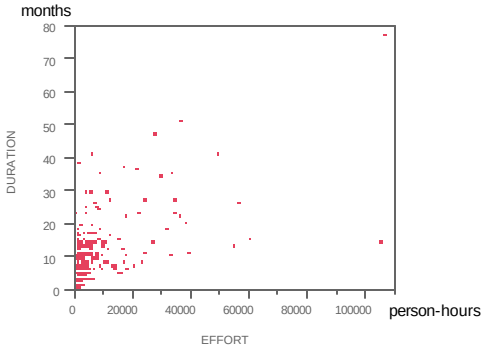
	Duration (D) in calendar months	Effort (E) in person-hours
Number of observations (n)	312	312
Minimum value	1.0	10
Maximum value	78,0	106 480
Mean value	10,5	5 933
Standard deviation	9,0	12 169
Median	8,0	2 228

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

↗ 312 projects satisfied all criteria

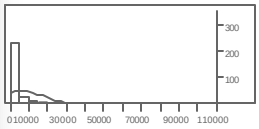
2.3 Basic Criteria

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



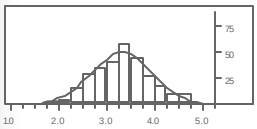
2.3 Distribution Analysis - Effort

No transform



Statistic	Value	Significance ($\alpha <= 0.05$)
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

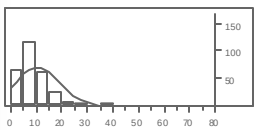
Log transformed



Statistic	Value	Significance ($\alpha <= 0.05$)
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

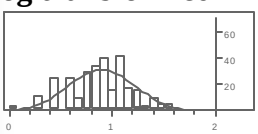
2.3 Distribution Analysis - Duration

No transform



Statistic	Value	Significance ($\alpha <= 0.05$)
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 ($\alpha <= 0.05$)
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

2.3 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.**

3. Deriving Models

3.1 Correlation analysis

3.2 Regression analysis:

Selected results

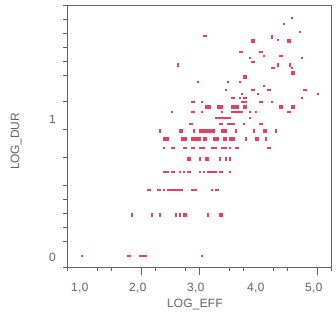
Residual analysis

The empirical models

Summary

3.1 Correlation Analysis (MF Platform)

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

3.2 Regression Analysis

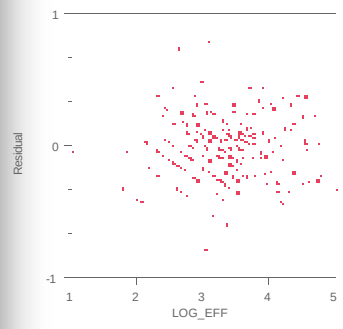
- **Regression hypotheses:**
  - linear relation judged adequate
  - residuals are randomly distributed
  - residuals independent from independent variable
  - variance of residuals is constant

3.2 Regression Analysis - Selected Results (MF Platform)

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

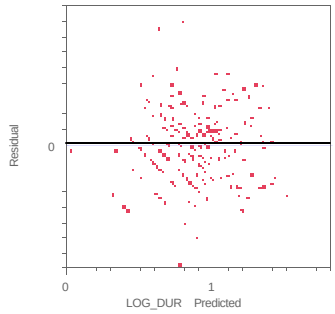
Selected results	Value
Sample size (n)	208
R^2	0.522
F(1.207)	224.865
Prob. > F	0.0001
Log(E) coefficient	0.366
Standard error of Log(E)	0.024
Constant	-0.339

3.2 Regression Analysis - Residual Analysis (MF Platform)



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

3.2 Regression Analysis - Residual Analysis (MF Platform)



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

3.2 Regression Analysis - The Empirical Model (MF Platform)

- **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)

3.2 Regression Analysis - The Empirical Models (MR and PC Platform)

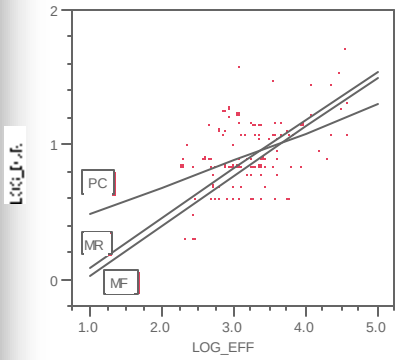
- **MR platform model (n = 65):**

  D = 0,548 * E ^{0,360} (E in person-hours)

- **PC platform model (n = 39):**

  D = 1,936 * E ^{0,201} (E in person-hours)

3.2 Regression Analysis - Summary



MF platform **: D = 0,458 * E ^{0,366}**

MR platform **: D = 0,548 * E ^{0,360}**

PC platform **: D = 1,936 * E ^{0,201}**

4. Assessing the Models

- **Value of Conte *et al.* (86) criteria for <u>untransformed</u> estimates:**

Criteria	ESCOM '97 ISBSG r. 3	MF model ISBSG r. 4	MR model ISBSG r. 4	PC model ISBSG r. 4
n	243	208	65	39
R^2	0,40	0,41	0,49	0,06
Rank	**3**	**2**	**1**	**4**
Avg. RE	- 0.18	- 0.15	- 0.14	- 0.16
Rank	**4**	**2**	**1**	**3**
Avg. MRE	0.48	0.45	0.47	0.48
Rank	**3**	**1**	**2**	**3**
Pred (0,25)	38 %	43 %	34 %	41 %
Rank	**3**	**1**	**4**	**2**
RMS	7,20	6,78	8,52	5,45
Rank	**3**	**2**	**4**	**1**
RMS bar	0,69	0,68	0,68	0,57
Rank	**3**	**2**	**2**	**1**

- Platform dependent models all show a performance equal or better than the ISBSG r.3 model (except for three values),
- Improvements are small though,
- In all models, magnitude of criteria underline the usefullness of model for "ballpark" estimates only. (ex.: Avg. MRE and Pred(0,25))

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
- Conclusion

Conclusion

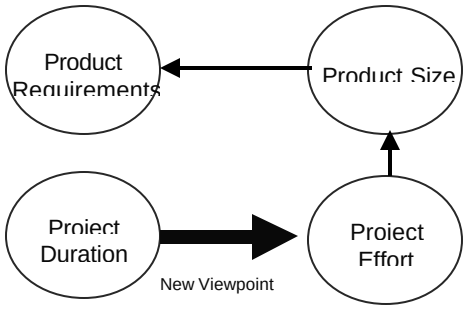
- Software sizing is different from estimation
- ISBSG data is available and can be analyzed by everyone.
- The steps taken to derive the example model and the assumptions behind it are known and the accuracy for this sample is published.
- Allows more intelligent tradeoffs and informed choices between 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.

6. Further Research Topics

↗ Turn the problem around



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 (To be published).
- S.D. Conte, H.E.Dunsmore, V.Y. Shen, Software engineering metrics and models. Menlo Park: The Benjamin/Cummings Publishing Company, Inc. 1986.
- O. Mendes, A. Abran and P. Bourque. Function Point Tool Market Survey. Université du Québec à Montréal,1996 Available at http://saturne.info.uqam.ca/Labo_Recherche/Lrgl/publi/treports/mo199701/mo199701.pdf.
- 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.