IPHIGÉNIE M.A. NDIAYE
A. Abran, G. Lévesque

Evaluation of SLIM Estimation Model Using ISBSG Repository

IWSM 2001,

Montreal, August 28-29, 2001

Agenda

- Introduction

- Project’s definition
- Project’s planning
- Operation
- Interpretation

- Conclusion

Introduction

- Duration and software cost estimation=> major elements in software projects planning
  - ❖ Accurate estimation of both of them is really important
  - ❖ => reliable estimation tool: Is SLIM reliable?

Project's definition

- 1. Motivation
  - ➢ Decision making is not an easy task
  - ➢ Lot of unknowns at the beginning of the project development
    [Abran et Robillard, 1993]

- Improve the quality of decision-making of the software projects managers.

Project's definition

2. Domain: MIS (Management information systems):
   - Software development
   - Software cost estimation models

3. Object: SLIM tool:
   - Based on Putnam's estimation model (1978)
   - Based on the Rayleigh curve

$$K = \left(LOC / \left(C * t^{4/3}\right)\right) * 3$$

Project's definition

- Sample: 789 projects from ISBSG database: *International Software Benchmarking Standard Group* (release 6 - 99), 1989-1998
  - Projects collected from 20 countries: 35% Australia, 34.4% North America, 29.2% Europe, 0.4 South America, 1% no identified.
- Projects mostly from domain of business application: 43% IS, 33% transaction processing applications, and 5% real-time related applications

Project's planning

- Directs criteria :
  - Error analysis (Conte et al, 1986)
    - Magnitude relative error $$MRE = |RE| = \left|\frac{Eact - Eest}{Eact}\right|$$
    - Mean magnitude of relative error MMRE = $$\overline{MRE} = \frac{1}{n}\sum_{i=1}^{n} MREi$$
    - Relative mean sqare error $$RMS = \left(\overline{SE}\right)^{1/2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Eact_i - Eest_i\right)^2}$$
    - Square root of mean relative error $$\overline{RMS} = \frac{RMS}{\frac{1}{n}\sum_{i=1}^{n} Eact_i}$$
    - Prediction level $$PRED(l) = \frac{k}{n}$$
  - Linear regression to measure the correlation between estimated effort and real effort

Project's planning

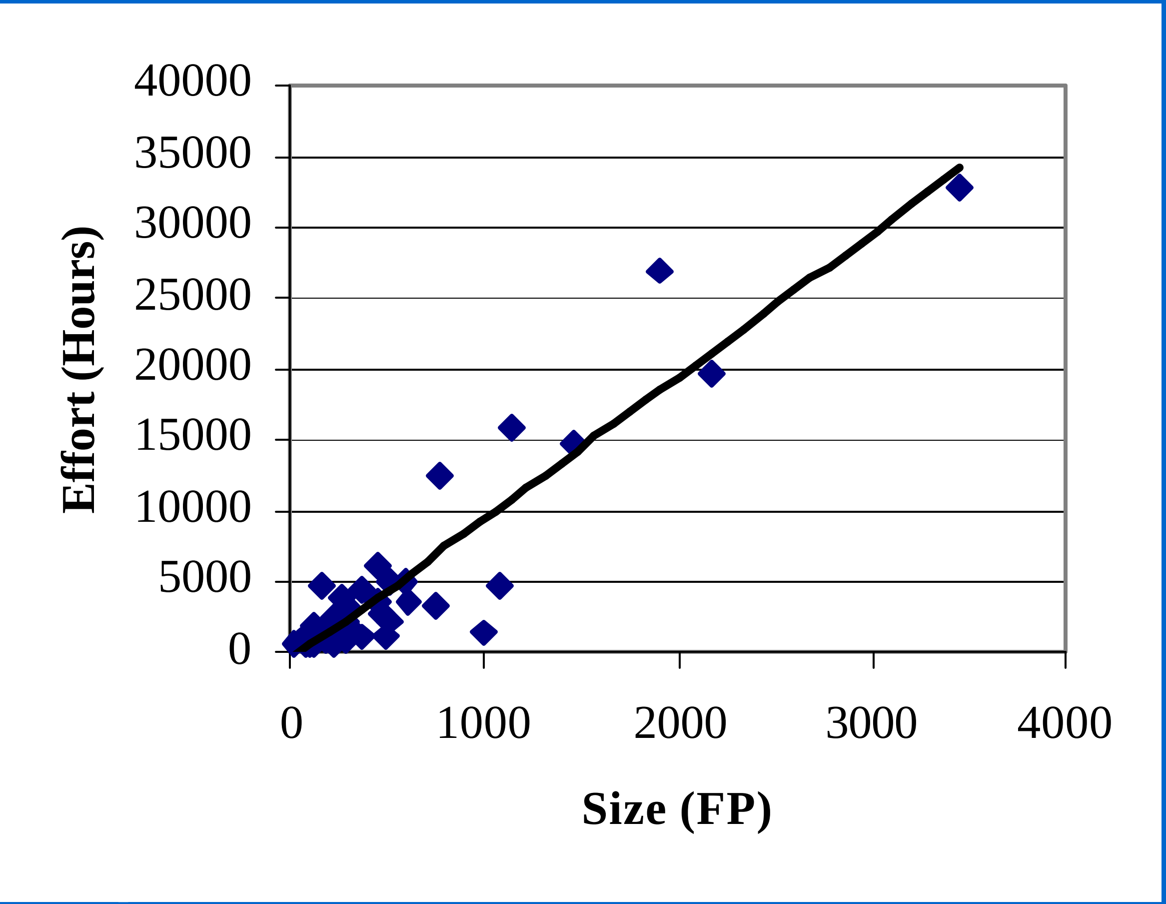
- Indirect criteria : Basis criteria of ISBSG-1999 (sample) :
  - No doubt about the quality of the data point :
    - each project has a quality tag assigned by ISBSG, based on whether or not the data received has fully met their data collection quality requirements, that is do the consider any specific data as fully credible
  - The project effort (in person-hours) is available and must be equal or greater than 400 p-h;
  - The project duration (n calendar month) is available
  - The programming language is available

Project's planning

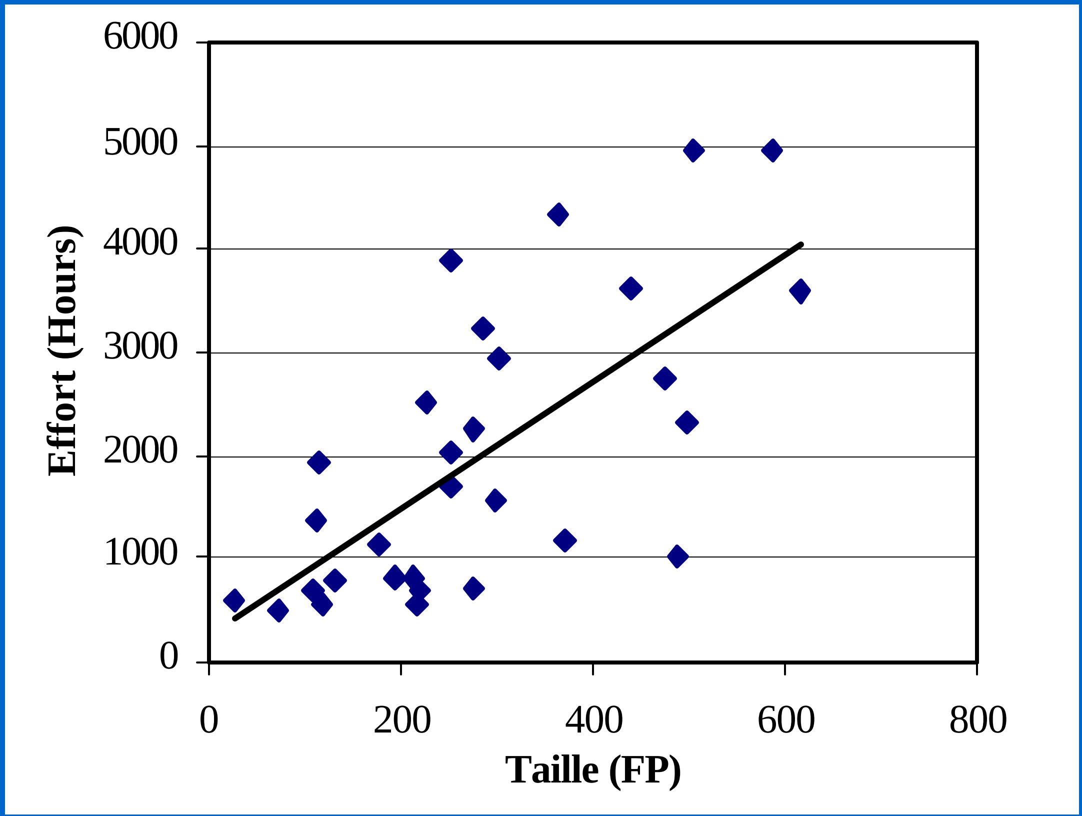
	Duration (month)	Effort (person-hour)
Number of projects	**497**	**497**
Maximum	**1**	**400**
Minimum	**84**	**138883**
Average	**10,5**	**6949**
Standard deviation	**9,2**	**13107**
Median	**8**	**2680**

Operation

- **Natural with all 41 projects:**
  - **Y = 10.05X – 648**
  - **R^2 = 0.85**

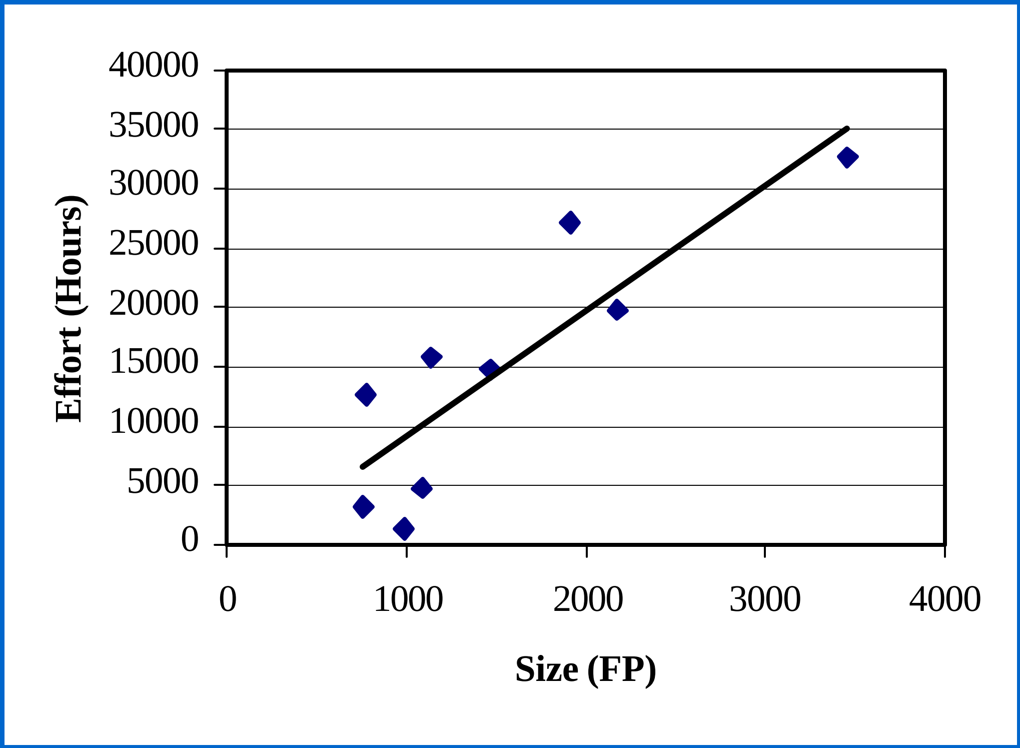


Operation



- Natural [20, 620]
- Y = 6.13X + 264
- $R^2 = 0.47$

Operation



- Natural [621, 3500]:
  - $Y = 10.53X - 1404$
  - $R^2 = 0,74$

Operation
Results and analysis

- Models without outliers
  - Real effort and SLIM
    - Natural [20, 620]: 60% under-estimated
    - Natural [621, 3500]: 56% over-estimated

    - MRE Natural [20, 620]: 248%
    - MRE Natural [621, 3500]: 850%

    - Is SLIM a good model?

Operation

Results and analysis

Based on the size, the duration and the language as parameters, SLIM 's estimations are very far from real effort

Operation

Results and analysis

- Real effort and SLIM:

Types of languages	Number of models	% under-estimated
3 GL	11	64
4 GL	17	88
APG	2	100

Operation
Results and analysis

- SLIM 's estimations are often less than real effort
- Those Results look like Kemerer 's (1987):
  - % Mean error of SLIM 's estimations = 772% with a minimum error of 21%
  - SLIM has been developed with the data of project of the department of american defense and claims to be now based on +7000 projects.

Operation
Results and analysis

- Real effort and ISBSG:
  - ❖Natural [20, 620]: 70% of 30 projects are overestimated
  - ❖Natural [621, 3500]: underestimated at 56%

Types of languages	Number of models	% underestimated
3 GL	11	55
4 GL	17	41
APG	2	50

Operation

Results and analysis

- SLIM et ISBSG:
  - ❖Natural [20, 620]: ISBSG > SLIM
  - ❖Natural [621, 3500]: ISBSG < SLIM

Types of language	Number of models	ISBSG > SLIM
3 GL	11	73
4 GL	17	88
APG	2	100

Mean relative error (MRE):

- Not many languages have a MRE of ISBSG> SLIM

- => ISBSG more reliable than SLIM

Operation
Results and analysis

- MRE:
  - 23/30 of sets of projects of the 3 types of languages have ISBSG's estimation more accurate than SLIM's
    - SLIM is the best only for others 4 GL[110, 950], Cobol II [181, 500], SQL
  - ISBSG is the best for others 3 GL, Access, Easytrieve, SQL Windows, APG, Telon
  - Aside those and others 4Gl [110, 950], neither SLIM, nor ISBSG is acceptable for projects 's effort estimation

Operation
Results and analysis

- Square root of mean relative error (RRMS) and prediction level (*PRED(l)*):

  - Models without outliers:
    - RRMS: SLIM > ISBSG
    - PRED (0,25): SLIM < ISBSG
    - RRMS high except for « other 3 GL »

Operation
Results and analysis

- Models without outliers
  - ISBSG's correlation coefficient always higher than SLIM's
- Link between real effort and the one estimated by ISBSG is more accentuated than the link between the same real effort and SLIM's estimation

Operation
Results and analysis

Models without outliers:

- $R^2 > 50\%$
  - SLIM: 25% (4GL), 20% (3GL), 0% (APG)
  - ISBSG: 50% (4GL), 56% (3GL), 50%(APG)

- $R^2 > 70\%$
  - SLIM: 6.25% (4GL), 11% (3GL), 0% (APG)
  - ISBSG: 25% (4GL), 22% (3GL), 50%(APG)

INTERPRETATION

- Interpretation context:
  - Only SLIM-estimate have been used
  - Use of analytic and statistical models
  - Goal reached
  - Field of posteriori productivity models evaluation

INTERPRETATION

- Extrapolation: Sample representativeness
  - Positive:
    - Large database (789 projects) =>457 projects
    - International projects
    - Completed projects
    - Various languages
  - Negative:
    - Not enough projects for some languages

INTERPRETATION

- Further Work:
  - Calibrate SLIM to the development environment of a specific project before estimating.
  - Adjust the gearing factor at the projects of the enterprise which is using SLIM, in order to adapt it at each enterprise context.
  - Using more than one tool (model) may be a possibility of estimations improvement.

CONCLUSION

« Adding man-power to a late software projects makes it later » [Brooks, 1975]. So, we must:

- ❖ Have a good cost and duration estimation
- ❖ Have a reliable estimation model

- But SLIM isn't eligible at this criteria of a good model in software engineering:
  - ➢ «a productivity model is considered good, if its MRE is between ± 25% for 75% of the observations».

References

[1] Heemstra, F.J., 1992, *How expensive is software? Estimation an control of software-development*, Kluwer.

[2] Boehm, B.W., 1981, *Software engineering economics*, Prenctice-Hall.

[3] Boehm, B.W., 1984, “ Software engineering economics ”, *IEEE Transactions on software engineering*, SE Vol. 10, p. 4-21.

[4] Herd, J.R.; Postak, J.; Russel, W.; Steward, K., 1977, “ Software cost estimation study - Study Results ”, Final technical report RADC-TR-77-220, Vol. 1, Doty Assoc., Rockville, Maryland.

References

[5] Putnam, L.H., 1978, “ A general empirical solution to the macro software sizing and estimating problem ”, *IEEE Transactions on software*, SE Vol.4, no 4, p. 345-361.

[6] Frieman, F.R.; Park, R.D., 1979, “ PRICE software model-version3 : An overview ”, *In proceedings IEEE PINY workshop on quantitative software models.*

[7] Rubin, H.A., 1983, “ Macro-estimation of software development parameters : The ESTIMACS system ”, *IEEE SOFTAIR Conference on software development tools, Techniques and Alternatives.*

[8] Albrecht, A.J.; Gaffney, J., 1983, ‘Software function, source lines of code, and development effort prediction ”; SE Vol.9, no 6, pp 639-648.

[9] http://www.qsma.com

Questions and comments

?