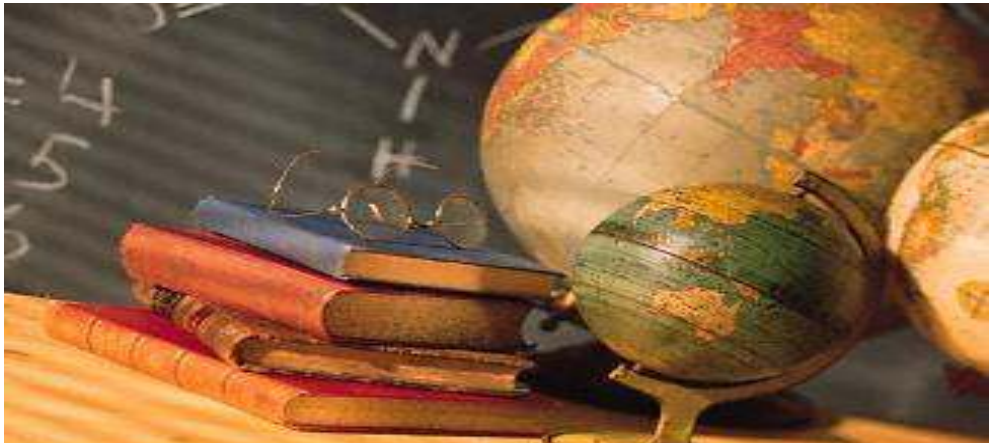
Estimation Models based on Functional Profiles



Alain Abran – Andra Panteliuc

Agenda

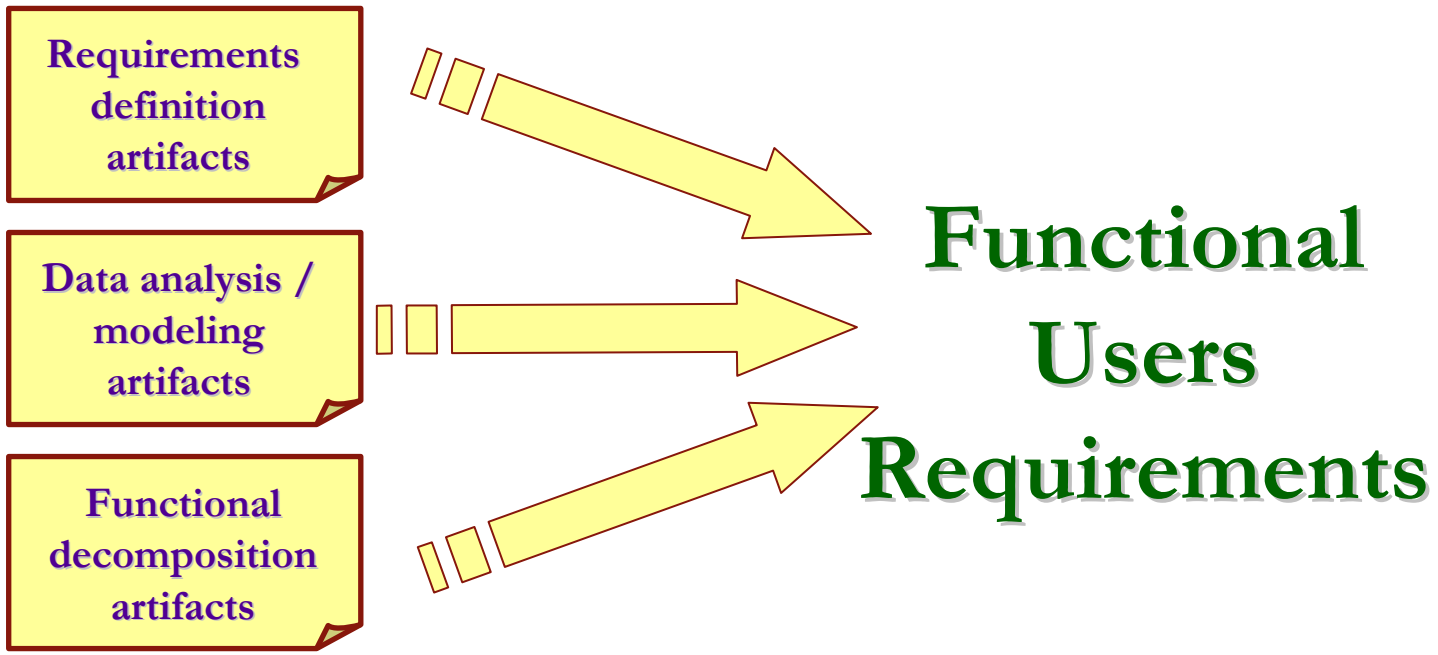
- Problem Statement
- Related Work
- Data Preparation
- Functional Profiles and Estimation Results

Problem Statement

- *Estimation models most often built using total Functional Size*
- *Can we produce better estimation models using the funcitonal profile of a software?*
- *Size measured with ISO 19761: COSMIC*

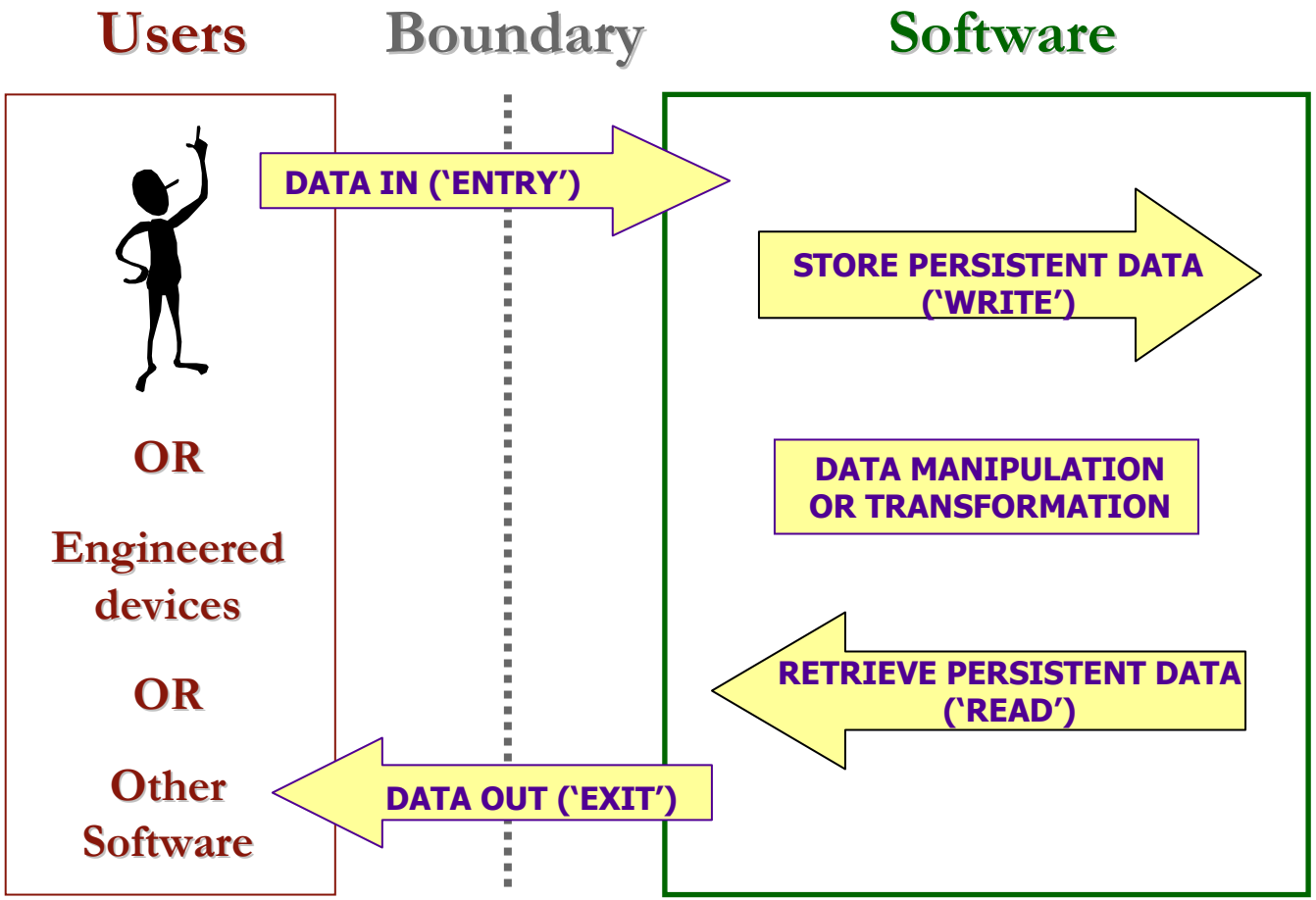
Functional Users Requirements

- Functional User Requirements (FUR) can be extracted from software engineering artifacts BEFORE the software exists… (using UML for instance)



User view of software functional requirements

COSMIC FFP Overview



Data Preparation

- *Repository 2006: International Software Benchmarking Standards Group – ISBSG*
- *Initial: 96 projects*
  - *5 with not enough quality*
  - *20 with not enough details by function type*
  - *Elimination of outliers on either size or effort*
- *Groups:*
  - *Development – 1 layer:* *15 projects*
  - *Development – Many layers:* *17 projects*
  - *Enhancement – All layer(s):* *18 projects*

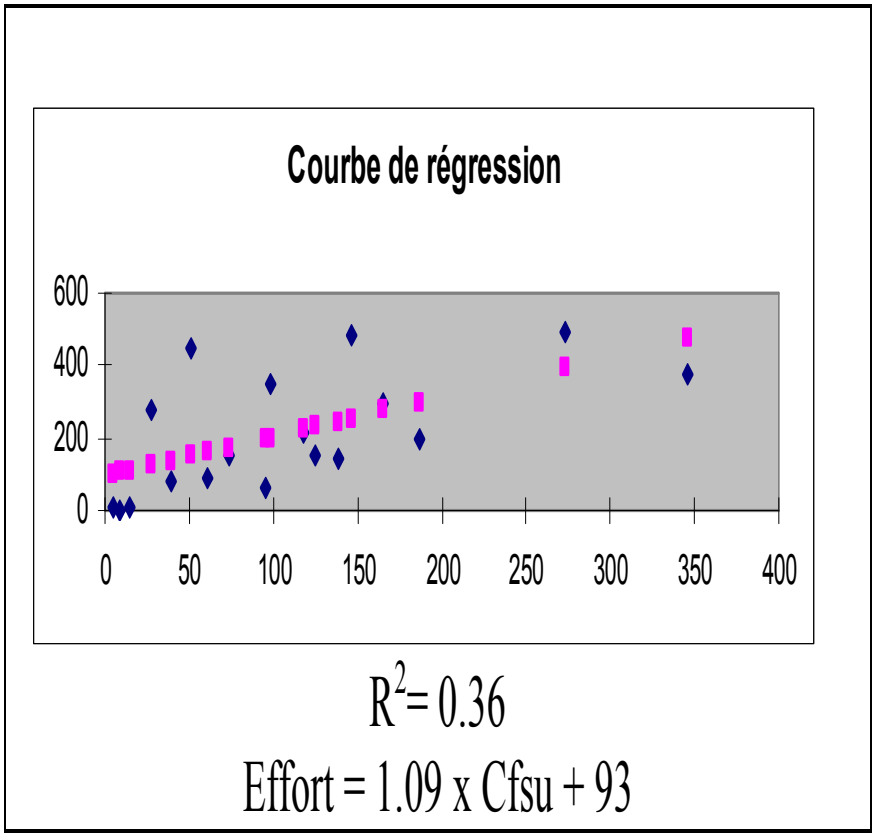


Figure 3: Enhancement projects (N=18)

Sample			
	No. of projects	R^2	Effort
Development Projects– 1 layer	15	0.60	15.95 Cfsu – 547
Development Projects – multiple layers	17	0.28	0.62 Cfsu + 135
Enhancement Projects	18	0.36	1.09 Cfsu + 93

Table 4: Enhancement projects	
Dominant function type	Reads
Percentage of projects within a range of ± 40% of the average of the dominant Read function type	77%
Number of projects within the ± 40% range (eg. Functional outliers)	14
Number of projects outside the range	4
Model for projects within the range	R^2=0.60

Table 2: Development projects – 1 layer	
Dominant function type	Exit
Percentage of projects within a range of ± 75% of the average of the dominant Exit function type	73%
Number of projects within the ± 75% range	11
Number of projects outside the ± 75% range (eg. Functional outliers)	4
Model for projects within the range	R^2 =0.83

Table 3: Development projects – multiple layers		
Dominant function type	Entries	Exits
Percentage of projects within the dominant selection and range	88%	76%
Range across the average of the dominant function type	± 50%	± 60%
Number of projects within the range	15	13
Number of projects outside the range (functional outliers)	2	4
Model- projects within range -R^2	0.28	0.22

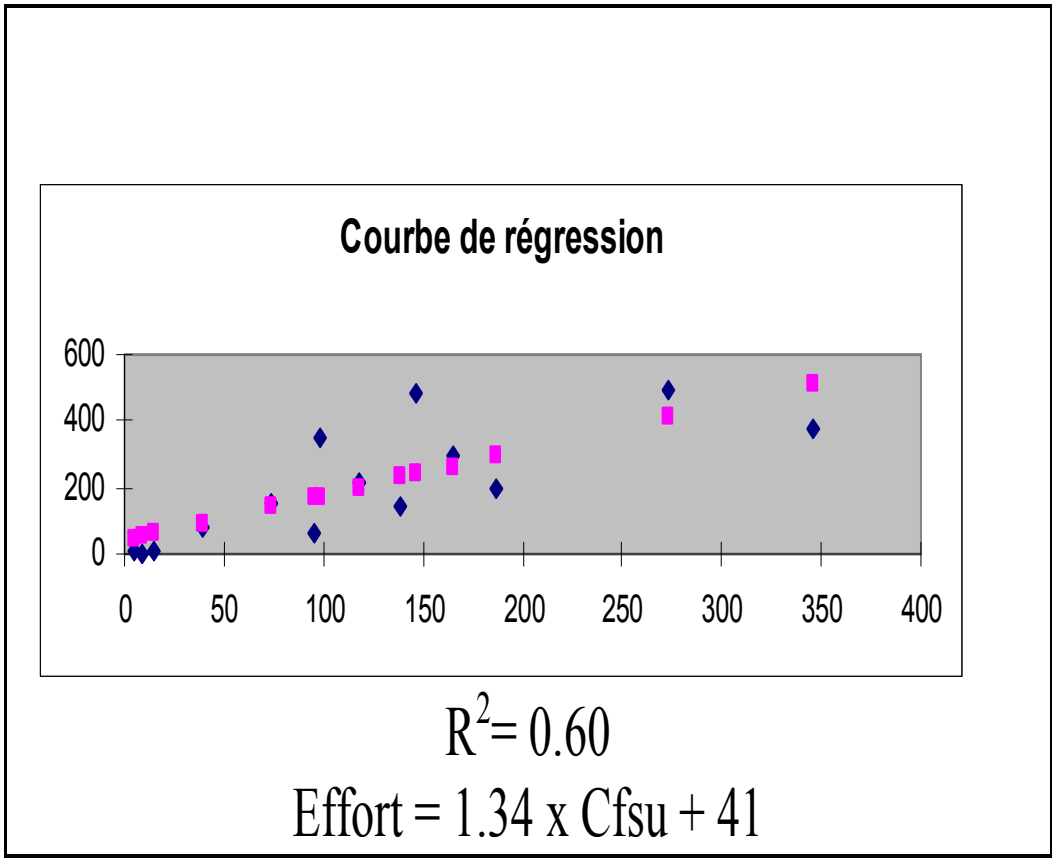
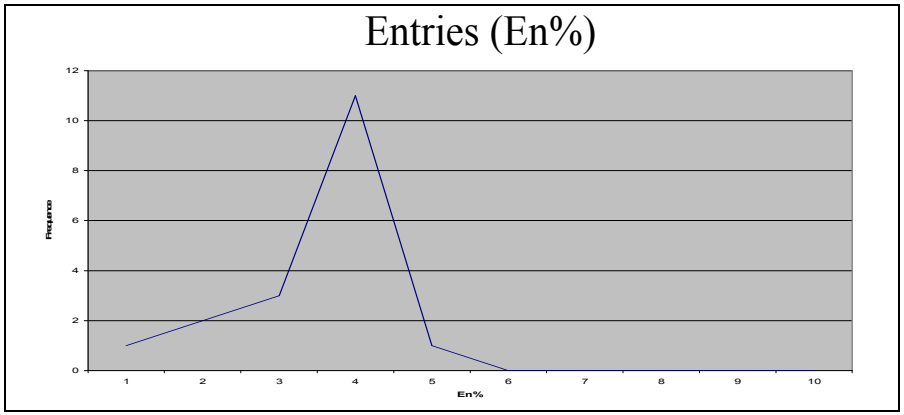
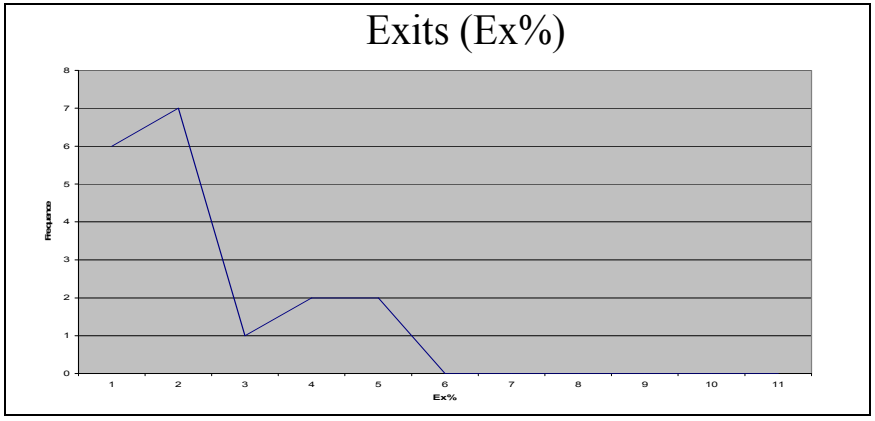
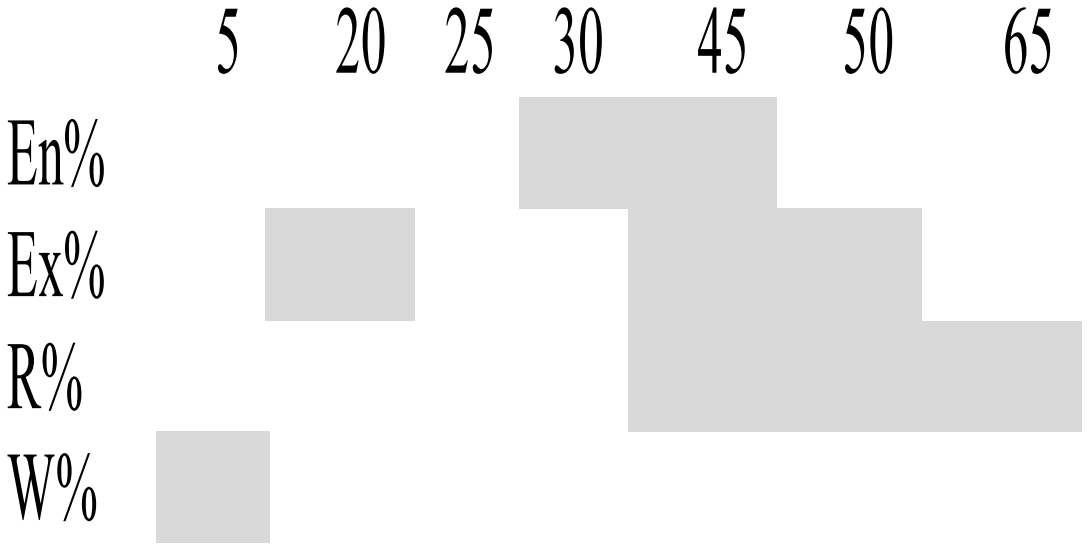


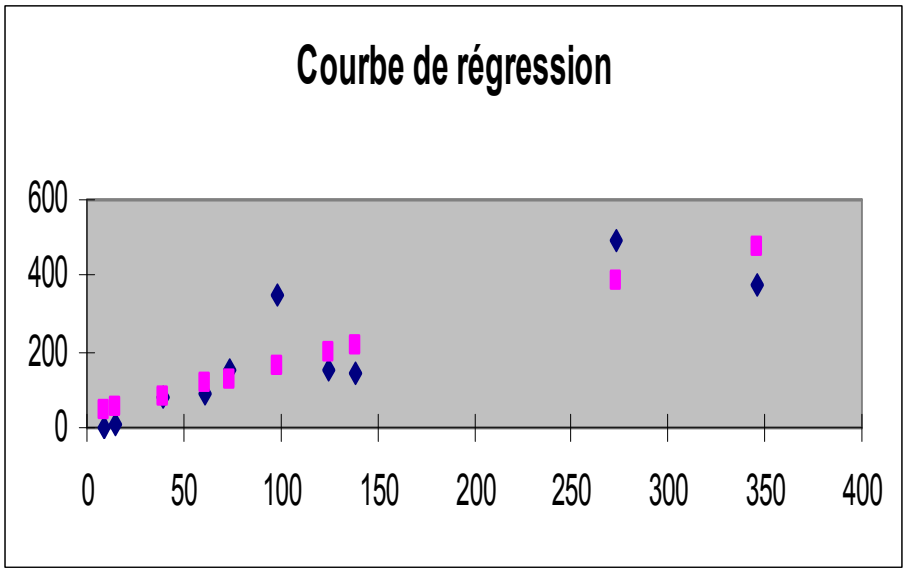
Figure 6: **Enhancement projects (N=14)**


Entries (En%)
Frequence
En%
0
2
4
6
8
10
12
1
2
3
4
5
6
7
8
9
10


Exits (Ex%)
Frequence
Ex%
0
1
2
3
4
5
6
7
8
1
2
3
4
5
6
7
8
9
10
11

: Most common frequency ranges by function type – Enhancements (N=18)





$R^2 = 0.72$

Effort = 1.27 x Cfsu + 35.11

Figure 12: Regression model - within the most common frequency ranges – enhancements (N=10)

Regressions models – Projects within the most common frequency ranges

Sample	# projects	R^2	Effort
Development projects – 1 layer	6	0.44	14.83 Cfsu – 182
Development projects – multiple layers	6	0.61	0.47 Cfsu + 66
Enhancement projects	10	0.72	1.27 Cfsu + 35

Conclusion

- *Functional profile are known early on*
- *Dissimilarity-similarity to average functional profiles is known early on*
- *This early knowledge can help build and select better estimation models*
  - *Numerous techniques for similarity to functional profiles*

Thank you !