Measuring Size for the Development of A Cost Model: A Comparison of Results Based on COSMIC FFP and SLIM Back-Firing Function Points

Ghislain Lévesque and Valéry Bevo
Department of Computer Science
Université du Québec à Montréal

tel.: 514-987-3000 ext. 3864
fax: 514-987-8477
e-mail: levesque.ghislain@uqam.ca

Abstract

Research has been initiated to develop a cost model for an outsourcing company which is oriented toward the development and maintenance of management information systems for a large telecommunications company. A measurement process has been implemented to collect post-project data and to develop a cost model for estimation purposes. As an initial step, these data were analyzed using the SLIM metrics tool and the COSMIC FFP method in parallel for the measurement of size. In this paper, the cost models derived from the two methods are presented, compared and discussed for their validity. Results of a short-cut estimation method based on the COSMIC FFP measures are also included. Finally, research avenues emerging from this study are identified.

1. Introduction

In this article, we present the first part of research initiated one year ago to develop a cost model with the COSMIC Full Function Point method for an outsourcing company which is oriented toward the development and maintenance of management information systems for a large telecommunications company. At the same time, the company set up a program to measure the size of its software projects using the SLIM metrics tool for determining the size of its projects in function points.

The data used for measurement came from closed projects for which a SLIM form had to be completed by project managers in order to obtain the breakdown of efforts and cost by project phase and the number of logical lines of code developed or modified by projects. At the same time, the requirement specification documents were collected for the size measurement using COSMIC FFP. There was no verification or audit to control the quality of data, and only one person was in contact with the project manager to collect all the information needed.

During the early months of the project, a measurement manual was prepared to set out the guidelines for implementing the measurement process within the company and to ensure the repeatability of the process.

Here, we present the results of the first cost models based on COSMIC FFP, and then we compare the results derived using the COSMIC FFP method with those produced by the SLIM tool. The comparison leads to some unanticipated findings which remains to be fully validated. Finally, an analysis of by-process COSMIC FFP measures is carried out to determine whether or not there is a possibility of finding a short-cut size estimation method with the same sample of projects.

2. COSMIC FFP Measurement Results

The results presented in this article were obtained by applying the COSMIC FFP (1999) method, version 2.0, and the process described in the Measurement Manual, version 1.0, prepared for the company. In Table 1, all the measured projects are listed. Three of the projects measured are not included in this sample for various reasons related to cost and timing. Only the names of the projects have been changed to preserve confidentiality. The relation between the cost of each project and its corresponding functional size is shown in Figure 1, where the correlation coefficient is 0.62. The average COST/COSMIC FFP ratio is $6.95 K per FFP and the average COSMIC FFP /man-month is 4.09.

The regression line for the cost model indicates that the fixed cost to start a project is $36,351 and that the variable cost is $3,400 per function point. All the projects used to calculate these values have a cost which is higher than this fixed cost. In terms of man-months (Figure 2), the fixed cost is 4.34 man-months per project and the variable cost is 0.26 man-months per function point. This fixed cost is higher than the amount observed for three of the projects (fewer than 4 man-months). These figures are valid for projects in the 0 to 350 COSMIC functional size units range.

Table 1: Results for the COSMIC FFP method

Projects	FFP	Costs $K	man-months	Cost/FFP	FFP/MM
A	30	49.0	3.8	1.63	7.89
B-1	7	51.0	5.0	7.29	1.41
C	170	254.0	21.7	1.49	7.83
B-2	17	41.0	3.4	2.41	5.00
B-3	61	59.0	7.8	0.97	7.82
D	167	137.0	12.6	0.82	13.25
E	8	40.0	3.2	5.00	2.50
F	13	81.0	6.5	6.23	2.02
G	158	574.0	53.1	3.63	2.98
H	349	1617.0	121.2	4.62	2.88
I	18	713.0	55.7	39.61	0.32
J	60	259.0	19.4	4.32	3.09
K	15	91.0	7.4	6.07	2.03
L	8	97.5	9.3	12.19	0.86
M	35	278.0	25.1	7.94	1.39
Mean	74.40	289.43	23.67	6.95	4.09
Average deviation	72.85	271.43	21.39	5.23	2.85
Standard deviation	96.79	419.16	31.79	9.55	3.60

Figure 1: Relation between cost and size, COSMIC FFP

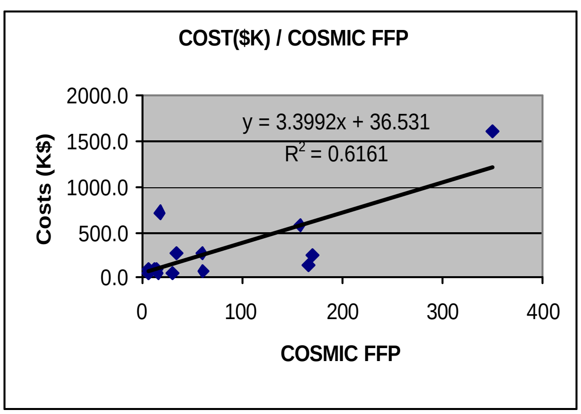
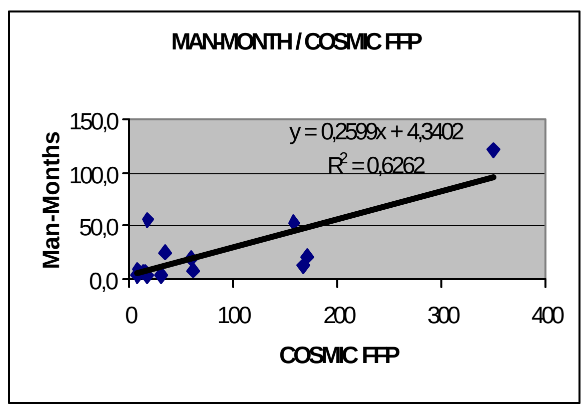


Figure 2: Relation between size and effort, COSMIC FFP



The quality of these data is quite good, as can be seen from Figures 3 and 4. In Figure 3, for the mean cost per function point, only one point is out of the 1-sigma range, and the deviation is more than 3-sigma for this point. This example is taken from the “I“ (for infrastructure) project. In the case of Figure 4, for the number of function points per man-month, four points are out of the 1-sigma range, one of them being out of the 2-sigma range, which confirms the earlier observation.

Figure 3: COST ($K) / COSMIC FFP mean and standard deviations

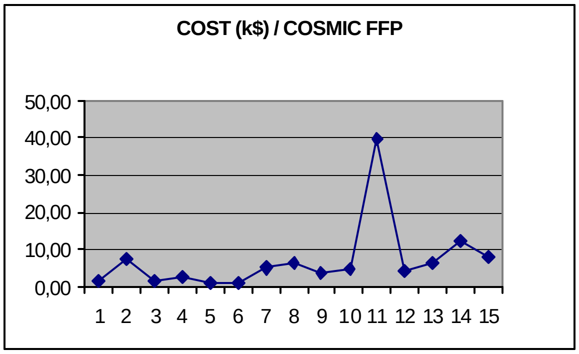
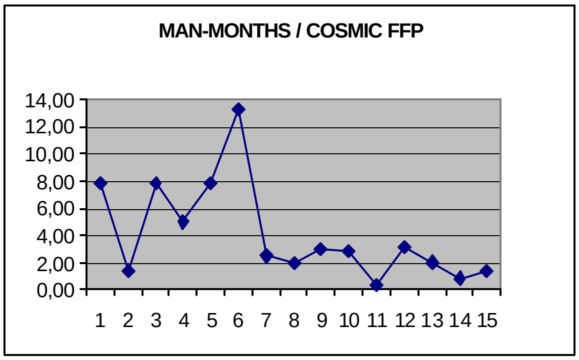


Figure 4: MAN-MONTH / COSMIC FFP mean and standard deviations



When divided into subgroups according to project type, two subgroups have shown promising results; however, the size of the sample is too small to provide more information on these.

3. Back-Firing Function Points

For the same group of projects analyzed in section 2, the results based on Back-Firing Function Points (BF-FP) are included in Table 2. One project is kept out of the sample for the same reason as before, the cost of the project being out of range. The relation between the cost of each project and its functional size is shown in Figure 5, where the correlation coefficient is 0.79. The regression line for the cost model shows that the fixed cost for starting a project is $52,838 and the variable cost is $2,542 per function point. The fixed cost determination is thus excluding four of the projects included in the sample for estimation purposes (the cost of four of the projects is less than $51 K). In terms of man-months (Figure 6), the fixed cost is 6.34 man-months per project and the variable cost is 0.19 man-months per function point. In this case, the same four projects have a cost that is less than 6.44 man-months, the fixed time to charge to a project. So, this model would not be a valid estimator for projects costing less than $53,000 and 6.44 man-months. The problem identified is associated with very small projects sized in BF-FP.

Table 2: Results for all projects Using SLIM Metrics (bf-fp)

PROJECTS	Back-Firing FP	Costs $K	Man-Months	Cost ($K)/ BF-FP	BF-FP/ MM
A	0.37	49.0	3.8	132.43	0.10
B-1	0.9	51.0	5.0	56.67	0.18
C	94.2	254.0	21.7	2.70	4.34
B-2	39.25	41.0	3.4	1.04	11.54
B-3	0.56	109.4	8.9	105.36	0.06
D	54.48	137.0	12.6	2.51	4.32
E	0.16	40.0	3.2	250.00	0.05
F	89.98	81.0	6.5	0.90	13.95
G	36.56	574.0	53.1	15.70	0.69
H	558.1	1617.0	121.2	2.90	4.60
I	161.7	713.0	55.7	4.41	2.90
J	251.4	259.0	19.4	1.03	12.96
K	2.89	91.0	7.4	31.49	0.39
L	49.53	97.5	9.3	1.97	5.33
M	56.04	278.0	25.1	4.96	2.23
Mean	93.07	289.43	23.67	40.94	4.24
Average deviation	92.41	271.43	21.39	50.76	3.65
Standard deviation	146.21	419.16	31.79	70.97	4.83

Figure 5: Relation between cost and size

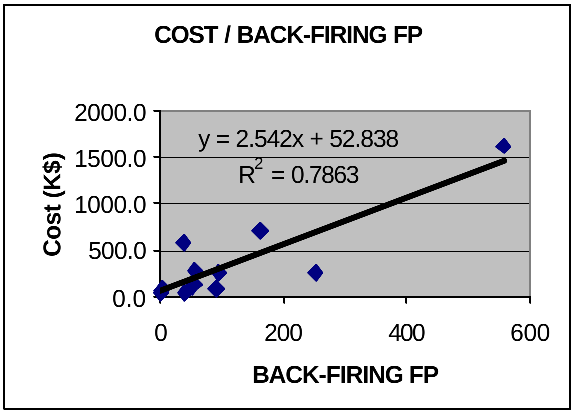
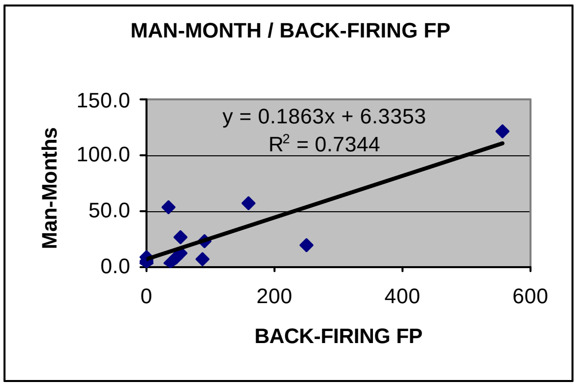


Figure 6: Relation between size and effort



The quality of the data can be assessed from Figures 7 and 8. Contrary to what was observed with COSMIC FFP in section 2, there are more data points in the 2-sigma and 3-sigma range here. The standard deviation to the mean then indicates that there are larger variations with the measures obtained by BF-FP. For example, the mean cost is $6,950 per COSMIC FFP unit with a variation of $9,550 for 1-sigma, while for BF-FP the mean cost is $46,940 per function point with a variation of $80,000 for 1-sigma, corresponding to 137% and 170% respectively. Moreover, it should be noted that there is no project in the interval from 250 to 558 BF-FP (see Table 2 or Figure 5). Consequently, even though the correlation factor seems good, we would not recommend using these regression curves for estimation purposes.

Figure 7: COST ($K) / BF-FP mean and standard deviations

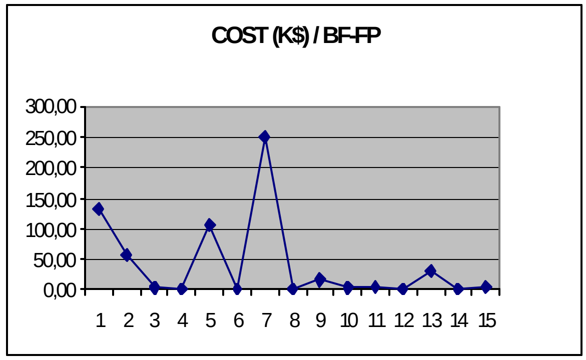
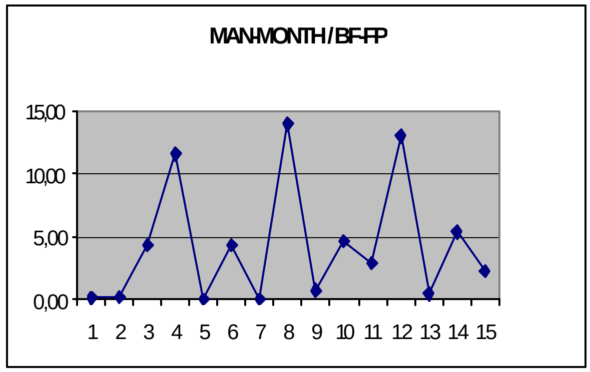


Figure 8: MAN-MONTHS / BF-FP mean and standard deviations



When taking the “H“ project out of the previous sample, which has 558 BF-FP, the fourteen remaining projects in the 0 to 250 BF-FP range leads to regression curves where the correlation coefficients drop to 0.26 and 0.19 (Figures 9 and 10), which confirms the previous observations and becomes completely non-significant.

Figure 9: Relation between cost and size

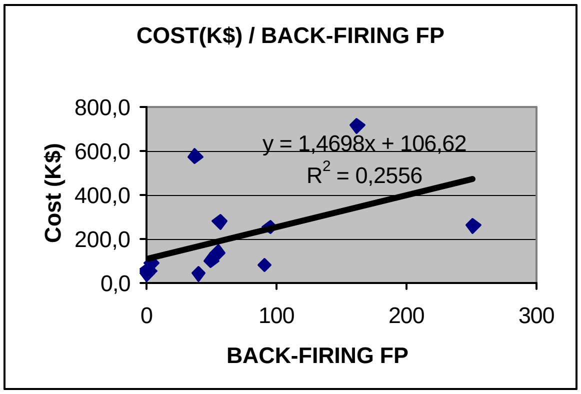
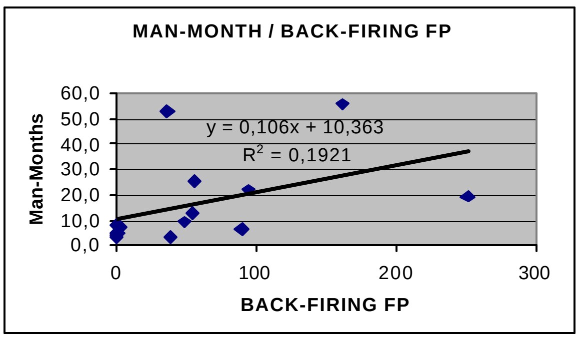


Figure 10: Relation between size and effort



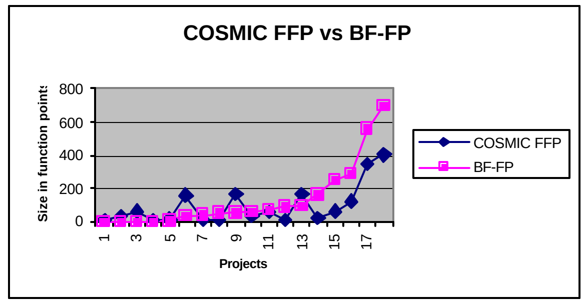
4. Comparison of COSMIC FFP and SLIM BF-FP results

Table 3 contains the sample used for this comparison (note that the ordering is based on the ascending Back-Firing Function Points). Figure 11 is drawn to illustrate the relation between the two methods more clearly.

Table 3: Project size measures with both methods

Project acronym	COSMIC FFP	BF-FP
E	8	0.16
A	30	0.37
B-3	61	0.56
B-1	7	0.9
K	15	2.89
G	158	36.56
B-2	17	39.25
L	8	49.53
D	167	54.48
M	35	56.04
O	59	68.5
F	13	89.98
C	170	94.2
I	18	161.7
J	60	251.4
P	125	290.5
H	349	558.1
N	401	694

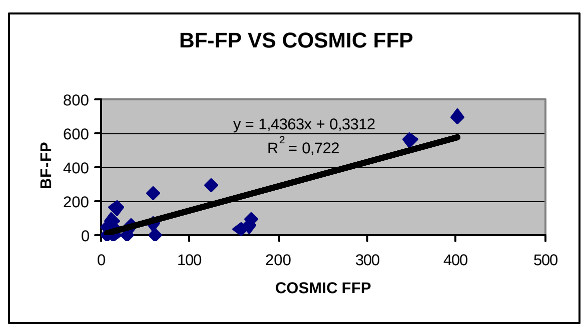
Figure 11: Comparison of COSMIC FFP and BF-FP Results



Essentially, COSMIC FFP is a sizing method which measures process activities in terms of data movements. There is no weighting for the complexity of data elements manipulated, as is the case for IFPUG 4.1. The SLIM estimates in function points are obtained by multiplying the number of logical lines of code by a conversion factor which is a constant determined for each language used. This is why there are probably more similarities between COSMIC FFP and SLIM BF-FP measures than between SLIM BF-FP and IFPUG 4.1 measures.

The linear regression on this sample of data has a correlation factor of 0.72 (Figure 12).

Figure 12 : Regression and Correlation on Size Measures



However, it can be observed from Figure 11 that projects larger that 100 BF-FP are marked with an inflection point. If this sample of data is broken down into two subsets, one for BF-PF < 100 FP and one for

BF-PF > 100 FP, we obtain the correlation results shown in Figures 13 and 14.

Figure 13: Regression and Correlation, Subset < 100 BF-FP

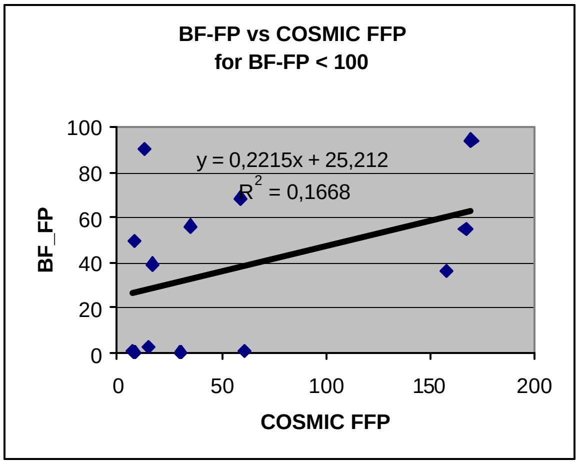
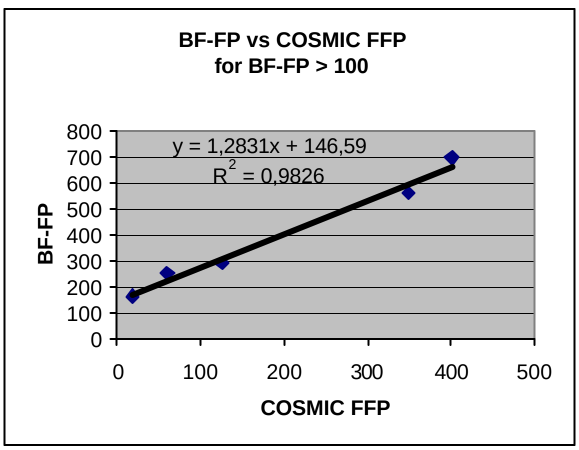


Figure 14 : Regression and Correlation Subset > 100 BF-FP,



It can be concluded from these graphs that:

1. *there is* <u>*no correlation at all*</u> *between the measures in COSMIC FFP and SLIM BF-FP for all projects under 100 function points according to the BF-FP measures (Figure 33c).*
2. *there is a* <u>*high correlation (0.98)*</u> *between the measures in COSMIC FFP for all projects over 100 functions points according to the BF-FP measures (Figure 33d).*

These conclusions are limited only by the size of the samples, 12 in the first case and 5 in the second case.

This means that the measures obtained using the BF-FP method for projects under 100 function points are highly questionable in terms of their validity as software size measures, if we allow that the COSMIC FFP method is a sound method for measuring the size of software. It also means that according to COSMIC FFP the measures could be valid for projects larger than 100 function points obtained from the SLIM Back-Firing method. The only limitation to this conclusion is the size of the sample, which is restricted to five data points for the time being. The range of this subset is in the interval 0 to 400 COSMIC Cfsu and 100 to 700 BF-FP.

5. By-Process COSMIC FFP measures

Another type of analysis has been carried out with the idea of determining a short-cut estimation method based on the data accumulated during size measurement using COSMIC FFP. As the system specifications identify all the processes to be modified or developed to satisfy the requirements of a project, we have tried to establish a relation between the size as measured in COSMIC FFP and the number of processes or reports impacted by a project. The sample of data used is presented in Table 4.

Table 4: By-Process COSMIC FFP Measures

PROJECTS	User interface req.			Downstream int. mod.			JCL & D. W. script			Report production		
	Process	Cfsu	mean	Process	Cfsu	mean	Process	Cfsu	mean	Report	Cfsu	mean
A				9	18	2.00	6	12	2.00			
B-1	1	4	4.00				1	3	3.00			
C	31	151	4.87	6	19	3.17						
B-2	2	17	8.50									
B-3							6	61	10.17			
D	4	16	4.00	11	39	3.55	19	112	5.89			
E				2	8	4.00						
F				2	13	6.50						
G				10	49	4.90	1	5	5.00	17	104	6.12
H	17	141	8.29	14	65	4.64	7	19	2.71	31	124	4.00

PROJECTS	User interface req.			Downstream int. mod.			JCL & D. W. script			Report production		
	Process	Cfsu	mean	Process	Cfsu	mean	Process	Cfsu	mean	Report	Cfsu	mean
I				3	18	6.00						
J	2	15	7.50	8	45	5.63						
K	3	15	5.00									
L	2	8	4.00									
M	6	27	4.50							2	8	4.00
N	19	101	5.32	12	143	11.92	18	61	3.39	26	96	3.69
TOTAL	87	495		77	417		58	273		76	332	
Occurrences			10.00			10			7			4
Mean			5.6			5.2			4.6			4.5

Measures have been subdivided into 4 categories:

- *User interface requirements* : these measures address the changes related to the presentation of inputs and outputs to the user via a terminal, Web-based or otherwise;
- *Downstream interface modules*: these measures concern all the intermediate systems impacted by a modification to produce the functionalities required;
- *JCL and Data Warehouse modules*: these measures represent the lower layers of applications to be changed by requirement modification.
- *Report production*: these measures represent the number of reports to be modified or developed to satisfy the requirements.

The figures under the columns *Process* represent the number of processes impacted by a project according to each of the previous categories. The Cfsu (COSMIC functional size unit) figure represents the sum of COSMIC full function points corresponding to these processes. The corresponding mean is determined by process and for each category as a whole. The average number of COSMIC function points by process is 5 Cfsu / process with a variation from 4.5 to 5.6, depending on the category.

To check the validity of these averages, a comparison of the actual measures in COSMIC FFP has been carried out with weighted means by type of process to obtain the weighted estimates (third column). The result is shown in Table 5 and its validity is determined by the correlation factor (Figure 15).

Table 11: Comparison of actual COSMIC FFP measures with a weighted estimate based on average by category

ALL PROJECTS	Actual Measures Cfsu	Weighted Estimates Cfsu	Average Estimate 5 Cfsu/Proc.	Average Estimate 5.1 Cfsu/Proc.
A	30	75	75	77
B-1	7	10	10	10
C	170	205	185	189
B-2	17	11	10	10
B-3	61	28	30	31
D	167	167	170	173
E	8	10	10	10
F	13	10	15	15
G	158	133	140	143
H	349	339	345	352
I	18	16	15	15
J	60	53	50	51
K	15	17	15	15
L	8	11	10	10
M	35	42	30	31
N	401	368	375	383
TOTAL	1517	1495	1485	1515

The results indicate a correlation factor of 0.97 (weighted estimates vs actual measures), which is excellent and confirms that the weighted average based on the process could be a good predictor of size in COSMIC FFP. To test whether or not a further simplification was possible, we repeated the test, this time with a simple average of 5.0 Cfsu by process (fourth column, Table 5). The correlation factor for this test is presented in Figure 16 (average estimate 5 Cfsu / proc. vs actual measures). The result is a correlation factor of 0.98, which is slightly better, with the added benefit that the estimation process has been simplified.

When comparing the totals of columns two, three and four in Table 5, we observe a slight difference between the totals for estimations and the total for the actual measures. If the average estimate is determined with a factor of 5.1 Cfsu / process, which has no impact on the correlation factor, the resulting total becomes 1515 Cfsu (column six in Table 5).

Figure 15: Comparison of Weighted estimate versus COSMIC FFP

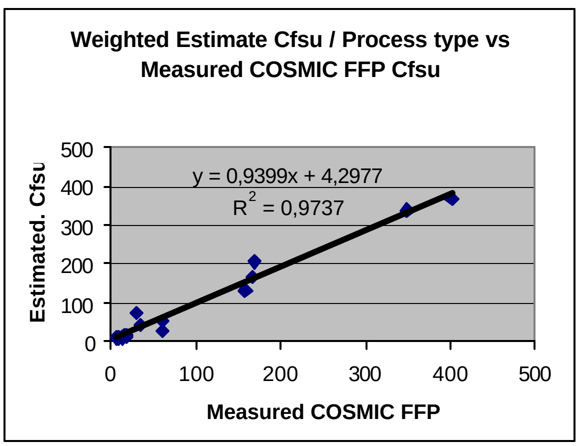
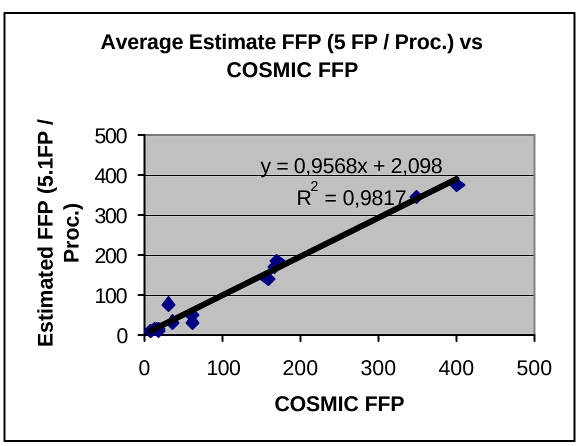


Figure 16: Comparison of average estimate versus COSMIC FFP



We can conclude from this that there is a possibility of applying a simple short-cut estimation process based on the number of processes and reports to be modified to determine an approximate size of a project in COSMIC FFP, in this company for a size range in the interval 0 to 400 Cfsu. This size measure can then be used with the regression line of a cost model to provide an estimate of the cost for a project.

7. Conclusion

In this paper, we have presented a model for costing development projects based on COSMIC FFP, and compared the results with other measures carried out using the SLIM metrics tool. The data sample included projects in the size range 0 to 400 COSMIC FFP.

The cost model obtained with COSMIC FFP seems to be valid, although the correlation coefficient is low, in the order of 0.62. We should recall that there has been no verification of the effort or cost data delivered by the project managers. This would be a necessary step for continuing the research and obtaining a stable model for costing. The development of subcategories of projects would probably also lead to a higher correlation coefficient.

The same sample of projects estimated for their size with SLIM metrics did not provide a reliable cost model.

The results of the comparison of COSMIC FFP measures with those of SLIM metrics were unanticipated. The fact that the results are highly correlated for projects over 100 Back-Fired Function Points can be explained by the fact COSMIC FFP only measures data movements - the processes - as is probably the case for logical lines of code for large projects. These results should not, however, be taken for granted, since the sample is actually limited to 5 projects. But there is a possibility of promising results if more projects are added to this sample.

Finally, the by-process analysis for the same group of projects opens up the possibility of using a short-cut COSMIC FFP estimation method in a company where the projects have been measured. Again, the sample is limited to 16 projects and it would be safer to go further with this test. Nevertheless, these results are also really promising.

References

Abran, A., Desharnais, J.-M., Oligny, S., St-Pierre, D. and Symons, C., 1999, COSMIC-FFP, Measurement Manual, version 2.0, Common Software Measurement International Consortium, October.

Boehm B., Clark. B., Horowitz E. and Westland C., 1997, Cost Models for future software life cycle processes: COCOMO 2.0, in Software Engineering Management, 2nd ed., Edited by Richard H. Thayer, N.Y., IEEE Computer Society, p. 270-307.

Gaffney J.E. and Cruickshank, R. D., 1997, *How to Estimate Software System Size*, in Software Engineering Management, 2nd ed., Edited by Richard H. Thayer, N.Y., IEEE Computer Society, p. 246-266.

Meli R., 1997, Early Function Points: a new estimation method for software projects, ESCOM 97, Maggio, Berlin.

Oligny S., Abran, A., Desharnais, J.M. and Morris, P., Functional Size of Real Time Software: Overview of the Field Tests, 1998, http: /www.lrgl.uqam.ca.

Santillo L. and Meli R., 1998, Early Function Points: some practical experiences of use, ENCOM-ENCRESS 98, Rome, Italy, May 27-29.

Stutzke R.D., 1997, *Software Estimating Technology: A Survey*, in Software Engineering Management, 2nd ed., Edited by Richard H. Thayer, N.Y., IEEE Computer Society, p.218-229.