Embedded Software Memory Size Estimation Using COSMIC: A Case Study

Sophie Stern[1], Cigdem Gencel[2],

[1] RENAULT, Embedded software group, Guyancourt Cedex (France)

[2] Blekinge Institute of Technology (BTH), Karlskrona (Sweden)

sophie.stern@renault.com, cigdem.gencel@bth.se

Abstract:

Early and accurate memory size estimation of Electronic Control Units (ECUs), which are dedicated embedded systems providing major software functionalities, is essential for automotive industry. Usually, experts using their expert judgments perform the estimations. Software functional size measures, which attempt at measuring the size by capturing the amount of functionality, have become popular in software industry particularly for effort estimation and project management purposes due to their availability early in the project lifecycle. This study investigates the potential usage of COSMIC Function Points as a predictor of memory size in comparison to expert opinion methods through conducting a case study in Renault.

Keywords

Memory size estimation, Embedded systems, Functional Size Measurement, COSMIC Function Points, Case Study.

1 Introduction

Optimal design of memory space is essential for developing cost effective embedded systems that execute on machines such as cars, airplanes, phones and television sets.

In automotive industry, most of the major functionalities are provided through dedicated ECUs (Electronic Control Units) connected to a communications network. ECU is a generic term for any embedded system that controls one or more of the electrical systems or subsystems; one example is the engine control module. Each ECU is composed of hardware and software. ECUs should be designed to have sufficient memory and processing capacity (ROM[1], RAM[2], E2P[3]) without inducing unnecessary costs [5].

[1] ROM: Read Only Memory
[2] RAM: Random Access Memory
[3] E2P: electrically erasable, programmable ROM

The hardware choice is made very early in the ECU's development process when actually choosing the ECU supplier. The hardware proposal is included in the supplier's commercial answer because it is technically and economically significant.

In case of over-estimation, the hardware would be over dimensioned and thus inducing unnecessary costs. On the other hand, if the estimations are too optimistic, the hardware will be under dimensioned at the beginning of the vehicle development project and thus bringing risks for delays. Therefore, early and accurate memory size estimation is essential for automotive industry.

The traditional way of memory size estimation is expert opinion. Recently, COSMIC (The Common Software Measurement International Consortium) Function Points (CFP) [7] has started to be used to predict the amount of memory needed in real time embedded systems (See the details in Section 2.2).

In this study, we further investigate the potential usage of CFP as a predictor of memory size in automotive industry. We conducted a series of empirical studies in Renault where COSMIC has been experimented for two years to measure the functional size of ECUs in order to construct productivity models and to predict software development costs.

This paper is organized as follows: Section 2 provides a short summary of the related work. Section 3 discusses the case study, the results and the validity threats. Section 4 presents the conclusions and the industrial perspectives of this work.

2 Background

2.1 Functional Size Measurement – COSMIC

Software functional size measurement (FSM) methods, which measure the software size by capturing the amount of functionality, have become popular in software industry particularly for effort estimation and project management purposes due to their availability early in the project lifecycle.

After the original description of Function Points (FP) by Albrecht and Gaffney [1][2], FSM evolved considerably [6][3]. Among all variations of the original method, the International Function Point Users Group (IFPUG) FPA [10][12], MarkII FPA [21], COSMIC [7][11], the Netherlands Software Metrics Association (NESMA) FSM [22][14] and the Finnish Software Metrics Association (FiSMA) FSM [24][15] methods have been accepted as international FSM standards by the International Organization for Standardization (ISO).

Among these methods we decided to use COSMIC in our empirical studies as COSMIC is a 2nd generation FSM method, which was designed to measure real-time and embedded systems as well as business information systems [11].

In COSMIC, the Functional User Requirements (FURs) are decomposed into their elementary components, called "Functional Processes". A Functional Process is defined as "an elementary component of a set of FUR comprising a unique, cohesive and independently executable set of data movements".

A Data Movement Type is defined as "a BFC, which moves one or more data attribute types belonging to a single data group type". Four kinds of BFC Types are recognized in COSMIC as:

- Entry: a data movement type that moves a data group from a functional user across the boundary into the functional process where it is required.
- Exit: a data movement type that moves a data group from a functional process across the boundary to the functional user that requires it.
- Read: a data movement type that moves a data group from persistent storage within reach of the functional process, which requires it.
- Write: a data movement type that moves a data group lying inside a functional process to persistent storage.

After identifying the BFC Types in the Functional Processes, the second step involves calculating the functional size of each BFC by applying a measurement function to the BFC Types and the related attributes. Then, the results are aggregated to compute the overall functional size of the software in CFP.

2.2 Functional Size and Memory Size Relationship

Toivonen [9] proposed a measure for software memory efficiency, defined as the amount of functionality (measured in CFP) packed per memory size in mobile terminals. He compared the memory efficiency of two different mobile phones using this measure.

Lind and Heldal [16][17][18] conducted a number of experiments in General Motors (GM) and identified very strong linear relationship (in all R^2 close to 0.99) between CFP and implemented code size in bytes. They developed a linear model to estimate the amount of memory required for ECUs in cars from CFP.

Gencel et al. [4] conducted another empirical study to compare the relationships between CFP and Source Lines of Code (SLOC) to between CFP and Bytes. Even though the strength of the relationship between SLOC and CFP found to be weak, the relationship for CFP and Bytes for the same components still showed significantly strong relationship. The authors stated that the reason might be that the compiler does a uniform job linking all the code used by the components and that measures the amount of functionality in bytes of code independent of how the functionality is implemented (as in COSMIC).

Later, Lind and Heldal [19] replicated their earlier experiments with software components of new types with the aim of identifying the factors affecting this relationship. They had two major findings. First, the nature of the linear relationship was different in each experiment even though the strengths of the relationships were similarly high in all experiments. Second, among different factors investigated, the functionality type, quality constraints, development methods and tools found to have significant impacts on this relationship. Therefore, they suggested that the real-time software components could be categorized to choose the suitable model for estimation.

3 Case Study

We performed a case study in Renault in order to investigate the potential usage of CFP as a predictor of memory size in automotive industry and to compare the results to the traditional expert opinion method that is currently being used in Renault and by the ECUs supplier.

Our research question (RQ) for this case study was the following:

RQ: ”Is COSMIC functional size a more reliable measure to predict the ECU memory size in real time embedded software than the traditional expert opinion methods?”

3.1 Data Collection

In order to answer the research question of this study, we collected data on 19 *functions* of the Body Control Module (BCM), which were designed by Renault and coded by one of its suppliers.

Table 1 shows the functional sizes of the *functions* measured by COSMIC method and the actual memory sizes for RAM, ROM and E2P for each *function*. Due to the confidentiality reasons, we cannot provide more information about the *functions* and instead use numbering in naming them.

The same team developed all the BCM *functions* using the same programming language, compiler and the development method. Therefore, we assume that we minimized the possible impact of these parameters to the results in this paper.

Two experienced measurers who have been using COSMIC for the last two years in Renault made the functional size measurements.

Table 1: *Functional Sizes and Actual Memory Sizes for the Case Functions*

Function	CFP	Actual Memory Size (kB)				
		ROM	RAM	E2P	RAM+E2P	Functional size/ ROM+RAM+E2P size (CFP/kB)
function 2	53	2.30	0.08		0.08	22.3
function 4	259	12.50	0.17	0.02	0.18	20.4
function 10	66	3.81	0.09		0.09	16.9
function 15	22	1.02	0.04		0.04	20.8
function16	73	3.81	0.09		0.09	18.7
function 17	748	32.96	0.58	0.16	0.73	22.2
function 18	90	31.08	0.63	0.13	0.75	2.8
function 21	7	0.58	0.07	0.03	0.10	10.2
function 22	8	0.85	0.04		0.04	9.0
function 23	34	1.79	0.04		0.04	18.7
function 24	39	2.14	0.05		0.05	17.8
function 25	45	4.96	0.53		0.53	8.2
function 26	54	2.51	0.07	0.02	0.08	20.8
function 27	57	1.97	0.08	0.02	0.11	27.5
function 28	91	4.72	0.14		0.14	18.7
function 29	110	4.01	0.07		0.07	27.0
function 30	136	5.71	0.12		0.12	23.3
function 31	155	10.50	0.16	0.02	0.18	14.5
function 32	157	5.14	0.10	0.02	0.11	29.9

Table 2 shows for each *function* the early size estimates made by Renault experts, by the supplier experts and the actual memory size measurements.

For *function 15* and *function 24*, only the supplier expert estimates and the actual size measurements were available. The estimates for E2P sizes were not available. Therefore, only the RAM and ROM size estimates are reported in this paper.

Table 2: *Comparison of the Memory Size Estimates and Residual Errors for the Case Functions*

Function	Actual memory size (kB)		COSMIC Based Memory Size Estimates (kB)		Renault Experts Memory Size Estimates (kB)		Sup. Exp. Mem Est. (kB)	% Residual (COSMIC)		% Residual (Renault Experts)		% Resid. (Suppl. Exp.)
	ROM	RAM +E2P	ROM	RAM +E2P	ROM	RAM	ROM	ROM	RAM +E2P	ROM	RAM	ROM
2	2.30	0.08	3.02	0.13	0.50	0.01	0.49	31.4	65.7	-78.1	-91.0	-78.7
4	12.50	0.18	11.96	0.30	25.78	0.37	25.20	-4.3	63.5	106.3	104.3	101.6
10	3.81	0.09	3.58	0.14	3.51	1.73	3.42	-6.0	44.6	-7.8	1729	-10.3
15	1.02	0.04	1.67	0.10			19.73	63.7	185.9			1831
16	3.81	0.09	3.89	0.14	5.99	0.11	5.86	1.9	52.3	57.0	20.8	53.7
17	32.96	0.73	33.18	0.70	27.08	0.20	26.46	0.7	-4.3	-17.8	-73.4	-19.7
21	0.58	0.10	1.02	0.09	0.17	0.03	0.20	74.9	-13.3	-70.6	-69.2	-66.6
22	0.85	0.04	1.06	0.09	0.51	0.03	0.49	25.9	127.6	-39.8	-30.0	-42.3
23	1.79	0.04	2.19	0.11	2.08	0.04	2.05	22.7	214.2	16.1	5.6	14.8
24	2.14	0.05	2.41	0.11			2.15	12.6	109.6			0.4
25	4.96	0.53	2.67	0.12	5.55	1.22	5.37	-46.1	-77.5	11.9	130.1	8.3
26	2.51	0.08	3.06	0.13	4.72	0.10	4.59	22.0	51.3	87.9	16.3	82.9
27	1.97	0.11	3.19	0.13	3.87	0.07	3.81	62.3	22.8	96.6	-35.2	93.6
28	4.72	0.14	4.67	0.16	7.39	0.12	7.23	-1.1	15.4	56.6	-10.0	53.2
29	4.01	0.07	5.49	0.17	4.72	0.10	4.59	37.1	140.1	17.7	35.1	14.6
30	5.71	0.12	6.62	0.20	5.72	0.11	5.57	16.0	61.1	0.2	-11.3	-2.4
31	10.50	0.18	7.44	0.21	10.99	0.16	10.74	-29.1	17.3	4.7	-12.0	2.3
32	5.14	0.11	7.53	0.21	5.59	0.07	5.47	46.4	87.5	8.6	-39.7	6.3

3.2 Data Analysis and Results

In order to answer our research question, we first investigated the strength of the relationship between functional size (CFP) and ROM size (kB) by looking at the scatter plot and calculating the correlation coefficient (R^2) (see Figure 1).

After making a correlation analysis, we obtained the correlation coefficient (R^2) as 0.56, which shows a medium strength relationship. On the other hand, the scatter plot demonstrated one outlier data point (*function 18*).

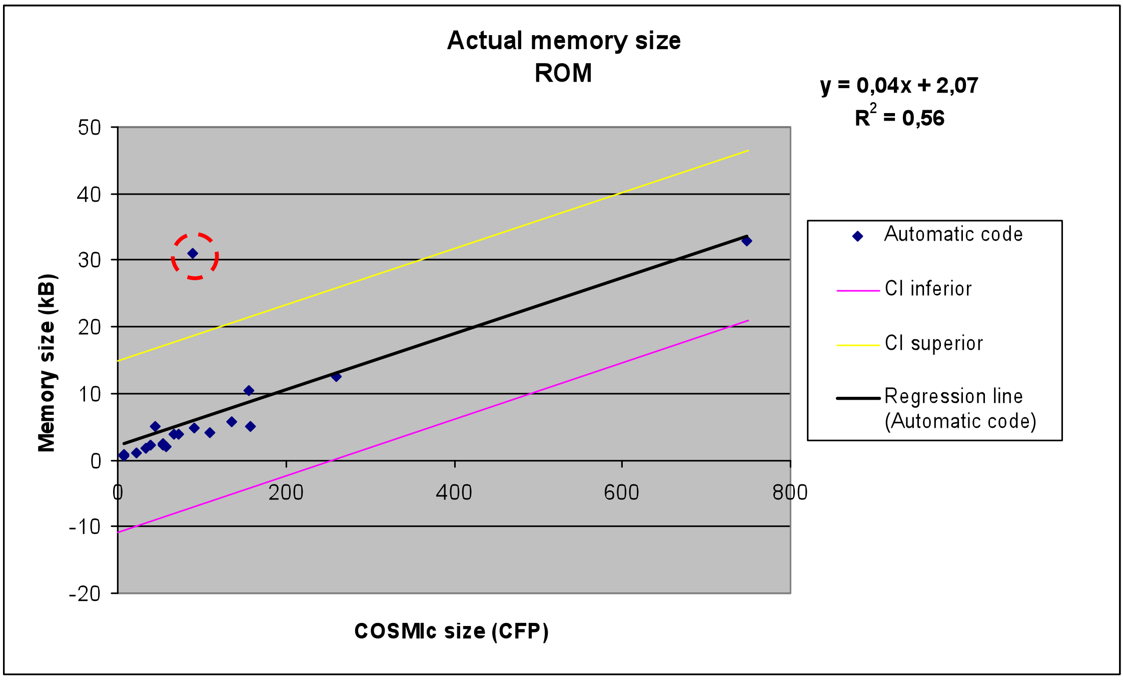


Figure 1: *CFP versus ROM size (kB)*

After investigating *function 18*, we identified that this is a particular case since the supplier had to do additional work on the Renault specification for the adaptation of this *function*. So, we excluded *function 18* from the study as an outlier and repeated the correlation analysis. The results showed a very strong relationship for the remaining functions (R^2=0.97) as shown in Figure 2.

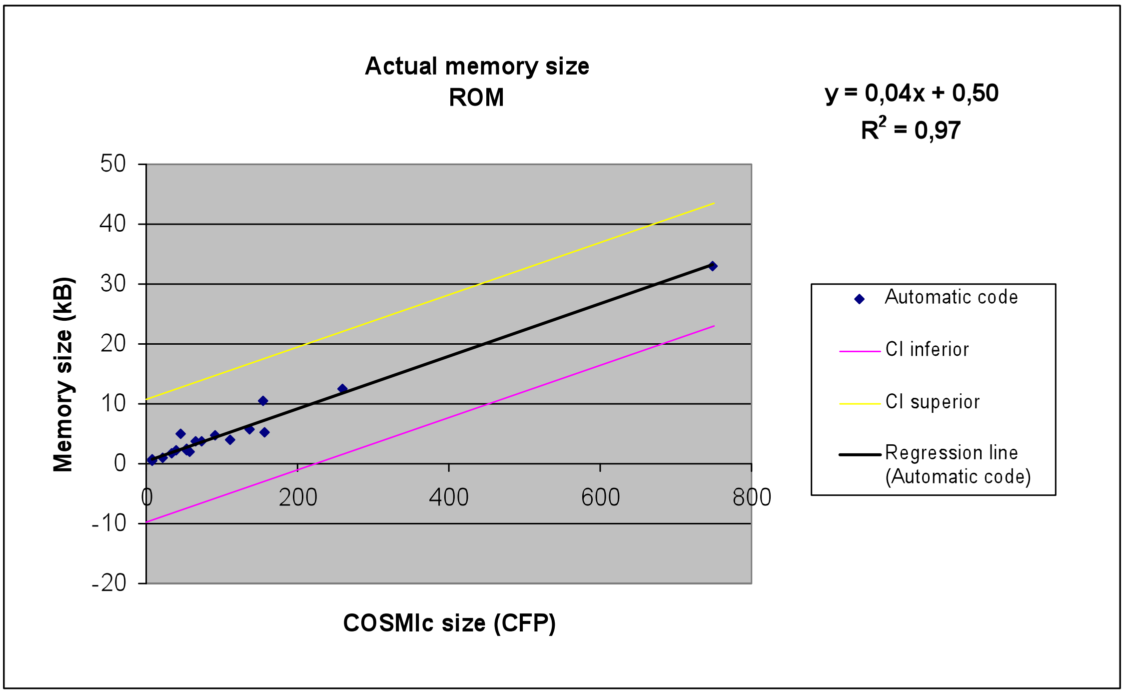


Figure 2: *CFP versus ROM size (kB) (excluding the outlier)*

Among other *functions* in the dataset, *function 17* has a very high functional size and memory size as compared to the others, which would have possibly made it an outlier as well. In order to avoid the possibility that such an outlier might distort the scale causing all other data points to look like a cloud with less importance and show false high correlation [20], we first removed this data point and checked whether R^2 was affected. Since there was no significant change, we decided not to remove this function as an outlier from the dataset.

These results are interesting in that the correlation coefficient is as high as what Lind and Rogardt [19] observed (R^2=0.99) for different types of distributable components developed in Saab and thus, confirming their findings.

Moreover, in this case study the *functions* have much higher variation in functional sizes (with a minimum 7 CFP and maximum 748 CFP) in comparison to the sizes of the distributable components (with a minimum 4 CFP to a maximum of 60 CFP) used for the experiments in [19].

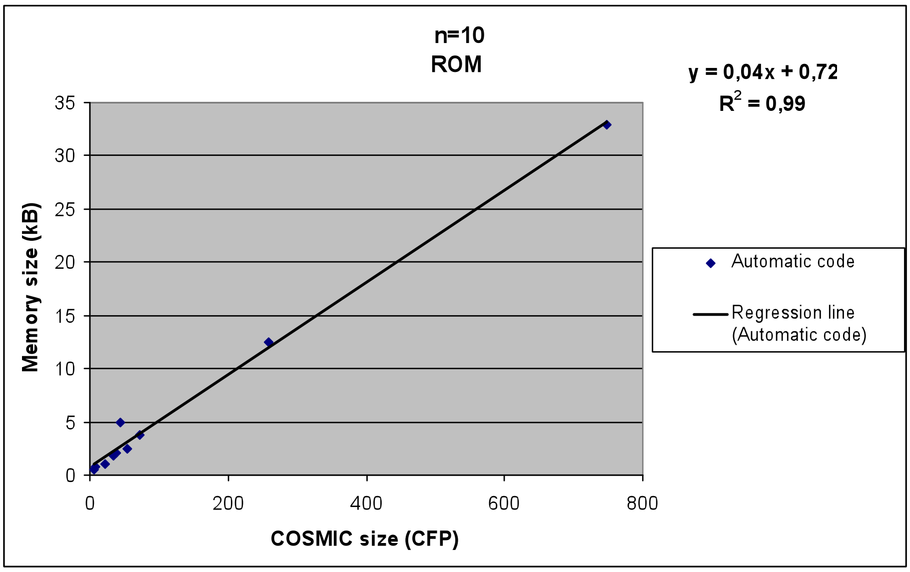
After identifying a strong relationship, we made further analysis to be able to compare the memory size estimates obtained by expert opinion methods to the estimation model we developed by building a linear model based on linear regression analysis.

First, we chose a sub-dataset of the functions (the highlighted functions in Table 2). We used *random sampling* function of Microsoft-Excel 'StatPlus for Mac' tool to randomly choose 10 *functions* out of 19 to build the model.

Then, we built a linear model performing linear regression analysis [20] on this sub-dataset. COSMIC size was found to be significant in explaining the variation in memory size and the built regression model was also significant (see Figure 3).

After, we estimated the memory sizes of all 19 *functions* by using the COSMIC based memory estimation model we developed from the sub-dataset. Finally, we compared the memory size estimates performed by three methods (COSMIC based estimation model, Renault experts, supplier experts) to the actual memory size values (see Table 2).

The comparison results showed some conflicting results. For some of the *functions*, the COSMIC based memory size estimates showed significantly better results than the expert estimates in Renault and at the supplier site and for the others, the contrary. Therefore, we further investigated the minimum and maximum % residuals (see Table 3).



Regression Statistics	
Multiple R	0.99
R Square	0.99
Adj. R Sq	0.99
Standard Error	0.92
Observations	10

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	892.61	892.61	1052.41	*8.89E-10*
Residual	8	6.79	0.85		
Total	9	899.40			

	Coeff.	*Std. Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0.72	0.34	2.12	0.07	-0.06	1.49
CFP	0.04	0.001	32.44	*8.88E-10*	0.04	0.05

Figure 3: *CFP versus ROM size (kB) (for the subset, n =10)*

Table 3: *Statistics on % Residuals for ROM, RAM and E2P Estimates*

	% Residual (COSMIC)		**% Residual (Renault Experts)**		**% Residual (Supplier Experts)**
Statistics	**ROM**	**RAM+ E2P**	**ROM**	**RAM**	**ROM**
Min	-46.10%	-77.50%	-78.10%	-91.00%	-78.70%
Max	74.90%	214.20%	106.30%	1728.90%	1831.20%

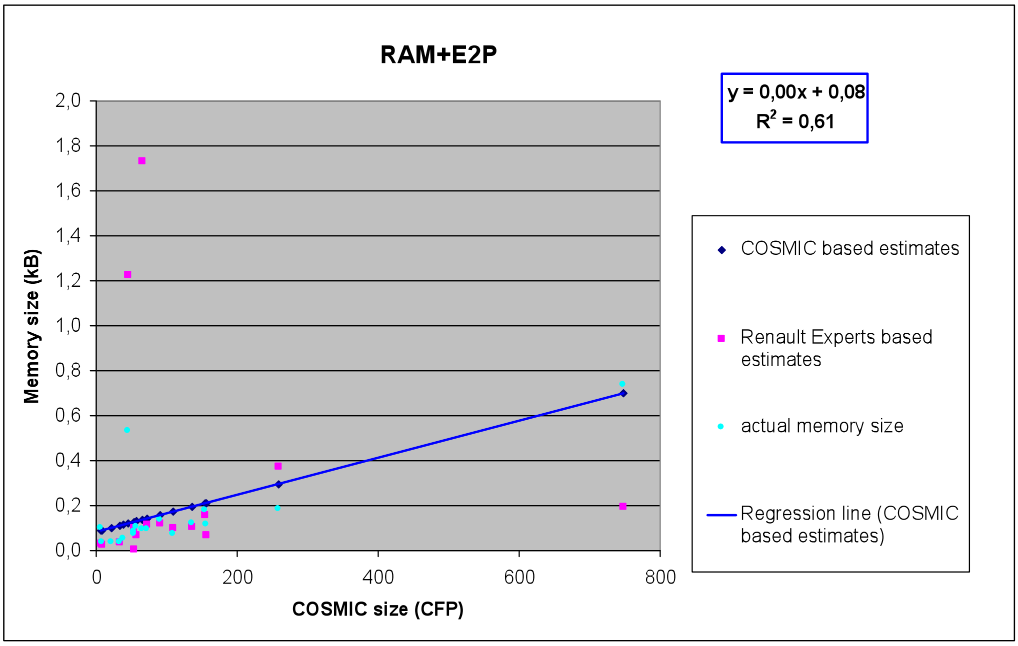
The COSMIC based memory size estimates (for ROM) showed underestimation for 5 of the *functions* (between -1.1% and -46.1%) out of 18 *functions*. The rest 13 *functions* were overestimated (between 0.7% to 74.90%). As for the RAM+E2P estimates, 3 out of 18 functions were underestimated (between -4.3% to -77.50%) and 15 functions were overestimated (between 15.4% to 214.20%).

Renault experts' estimates (for ROM) showed underestimation for 5 *functions* (between -7.8% to -78.1%) out of 16 *functions*. The remaining 11 *functions* were overestimated (between 0.2% to 106.30%). As for the RAM estimates, 9 out of 16 *functions* were underestimated (between -11.3% to -91.00%) and 7 *functions* were overestimated (between 5.6% to 1728.90%).

Supplier experts' estimates (for ROM) showed underestimation for 6 *functions* (between -2.4% to -78.7%) out of 18 *functions*. The remaining 12 *functions* were overestimated (between 0.4% to 1831.20%).

If the comparison of the accuracy of the estimates for ROM memory size is made globally, i.e., not for each *function*, the COSMIC model based estimates seems to be better than the expert opinion (see Figure 4).

Figure 4: *Comparison between memory size estimates (COSMIC model based, Renault experts, supplier experts) and actual memory sizes for ROM*



However, when looking at the estimates considering each *function*, we identified that the ROM size estimates based on experts' judgments are much better for a few *functions*. We attributed this to the fact that the experts in Renault have a lot of experience for such BCM *functions* and could make better estimates for those particular BCM *functions*. Interestingly, the Renault experts' estimates and supplier experts' estimates are in agreement for most of the cases even though the deviation from the actual values for some of the *functions* are much higher than the COSMIC based estimates.

As for the RAM+E2PROM memory size estimates, Renault experts' estimates showed great variation for some of the *functions* (see Figure 5). For the rest of the functions, sometimes the COSMIC based estimates were better and for some others; the Renault experts' estimates.

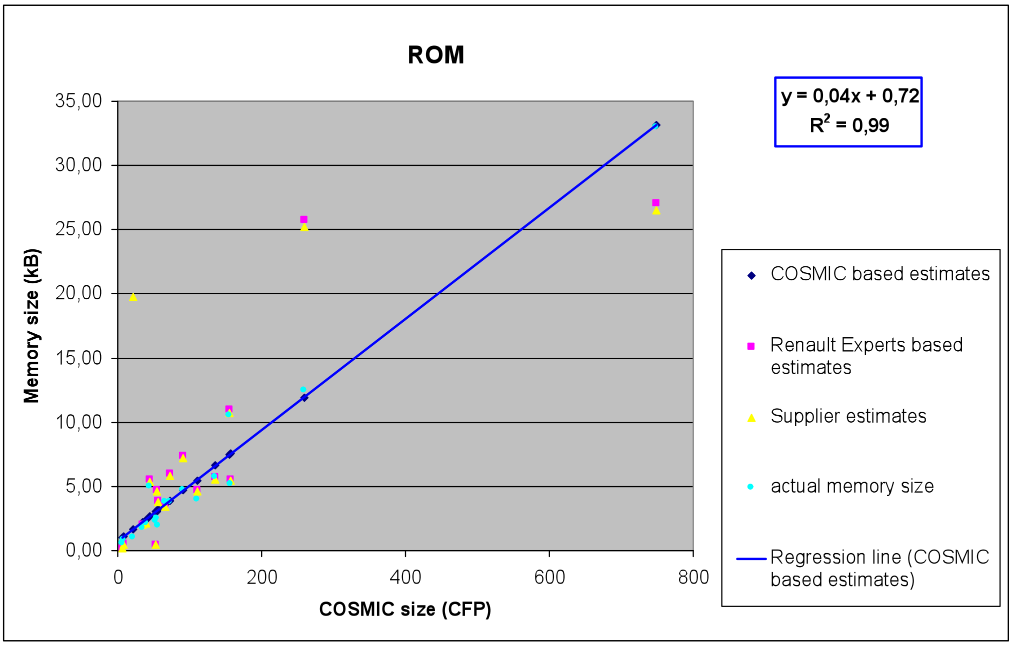


Figure 5: *Comparison between memory size estimates (COSMIC model based, Renault experts, supplier experts) and actual memory sizes, for RAM+E2*

3.3 Validity Threats

This section summarizes the validity threats of this study. Internal validity aims at ensuring that the collected data enables the researchers to draw valid conclusions [8].

In this study, we collected data on 19 *functions* belonging to BCM. The same team developed all of the *functions* by using the same programming language, compiler and the development method. Therefore, we made our analysis on a fairly homogenous dataset, which we believe that minimized the possible impact of these parameters to the estimations and thus to the conclusions of this paper. However, some other factors, which we could not fully control, might have had some impacts on our results.

Another internal validity threat might have occurred due to the inconsistencies among measurers who sometimes have to rely on their subjective interpretation of the functional requirements when measuring the functional size by COSMIC [23]. In order to minimize the effects of this threat, the same measurers who have the domain knowledge and experience in using COSMIC for the last two years in Renault performed the measurements.

External validity defines to what extent findings from the study can be generalized to and across populations of persons, settings, and time [8].

In this study, we collected data for similar *functions* of one particular ECU developed in Renault. Therefore, the results are valid only for this software organization and for one ECU type. However, since our results confirmed the previous results obtained by Lind and Heldal [18][17][16][19] that COSMIC FP can be used as a reliable measure to estimate memory size in Saab, we believe that our results can be generalisable to some extent for the automotive industry.

4 Conclusions

A few years ago, the automotive industry has taken a first bend with the introduction in mass in cars of electronics. Currently, the automotive industry is going into a second bend with the software increasing in cars to provide new functionalities such as multi-media, connectivity, advanced driver assistance systems, electrical vehicle, and so on.

Therefore, it is essential that the car production costs must be forecasted accurately to bring the best offer to demanding customers. In Renault, currently Renault experts and supplier experts have been performing the memory size estimates to meet this need.

This study investigated whether COSMIC functional size based estimations can be a good alternative method for memory size estimation. We found the results as promising. Therefore, we will continue to collect more data in order to develop a memory size estimation model based on COSMIC functional size to support experts' judgment in Renault.

We also think that repeating such case studies for other types of embedded systems' memory size estimations such as mobile phones, televisions, etc. is also promising.

5 Acknowledgements

We would like to thank Alice Marie, Mathematics Master student at UVSQ (University of Versailles Saint-Quentin en Yvelines) and trainee in Renault during last summer, for her contributions in performing the empirical studies.

6 References

[1] A.J. Albrecht, "Measuring Application Development Productivity", in Proc. of IBM Applications Development Symposium, Monterey, California, October 14-17, 1979.

[2] A.J. Albrecht, and J.E. Gaffney, "Software Function, Source Lines of Code, and Development Effort Prediction: A Software Science Validation", IEEE Transactions on Software Engineering, vol. SE-9, no. 6, November 1983.

[3] C. Gencel, and O. Demirors, "Functional Size Measurement Revisited", ACM TOSEM, Vol.17, No.3, 2008, 71-106.

[4] C. Gencel, R. Heldal, K. Lind, "On the Relationship between Different Size Measures in the Software Life Cycle", Proc. of the IEEE Asia-Pacific Software Engineering Conference (APSEC), Penang, Malaysia, 1-3 December 2009, pp.19-26.

[5] C. Srilatha and C.V. Guru Rao, "A Novel Approach for Estimation and Optimization of Memory In Low Power Embedded Systems", Intern. Journal of Computer Theory and Engineering, Vol. 1, No. 5, 2009, pp. 1793-8201.

[6] C. Symons, "Come Back Function Point Analysis (Modernized) – All is Forgiven!)", Proc. of the 4th European Conference on Software Measurement and ICT Control, FESMA-DASMA 2001, Germany, 2001, pp. 413-426.

[7] COSMIC: The Common Software Measurement International Consortium FFP, version 3.0, Measurement Manual, 2007.

[8] Creswell, J.W., Research design: qualitative, quantitative, and mixed method approaches. Second Edition, SAGE, ISBN: 0761924426, 2003.

[9] H. Toivonen, "Defining Measures for Memory Efficiency of the Software in Mobile Terminals", in Proc. of the 12th International Workshop on Software Measurement, October 7-9, Magdeburg, Germany, 2002.

[10] IFPUG, Function Point CPM, Release. 4.1, IFPUG, Westerville, OH, 1999.

[11] ISO/IEC 19761:2003: COSMIC Full Function Points Measurement Manual v. 2.2, 2003.

[12] ISO/IEC 20926, Software engineering - IFPUG 4.1 Unadjusted FSM Method - Counting Practices Manual, 2003.

[13] ISO/IEC 20968, Software engineering - Mk II Function Point Analysis - Counting Practices Manual, 2002

[14] ISO/IEC 24570:2005: Software engineering - NESMA FSM Method v.2.1 - Definitions and counting guidelines for the application of Function Point Analysis, 2005.

[15] ISO/IEC 29881:2008, Software Engineering -- FiSMA functional size measurement method version 1.1, International Organization for Standardization, 2008.

[16] K. Lind, and R. Heldal, "Estimation of Real-Time Software Code Size using COSMIC FSM", Proc. of the IEEE Intl. Symposium on Object/component/service-oriented Real-time distributed Computing (ISORC 2009), Tokyo-Japan, 17-20 March 2009.

[17] K. Lind, and R. Heldal, "Estimation of Real-Time Software Component Size", Nordic Journal of Computing (NJC), Vol. 14, Issue 4 (December 2008), pp. 282-300.

[18] K. Lind, and R. Heldal, "Estimation of Real-Time System Software Size using Function Points", Proc. of the Nordic Workshop on Model Driven Engineering (NW-MoDE), 2008.

[19] K. Lind, and R. Heldal, "Categorization of Real-time Software Components for Code Size Estimation", Proc. of the 4th Intern. Symposium on Empirical Software Engineering and Measurement (ESEM 2010), IEEE Press, 16-17 September 2010.

[20] K.D. Maxwell, Applied Statistics for Software Managers. Prentice Hall PTR, 2002.

[21] MkII FPA Counting Practices Manual Version 1.3.1, UKSMA: United Kingdom Software Metrics Association, 1998

[22] NESMA, Definitions and Counting Guidelines for the Application of Function Point Analysis, Version 2.0, NESMA, 1997.

[23] O. Top, O. Demirors, B. Ozkan, "Reliability of COSMIC Functional Size Measurement Results: A Multiple Case Study on Industry Cases", 35th Euromicro SEAA Conference, Greece, 27-29 August, 2009, pp. 327-334

[24] P. Forselius, Finnish Software Measurement Association (FiSMA), FSM Working Group: FiSMA Functional Size, 2004.