COSMIC Light vs COSMIC Classic Manual: Case Studies in Functional Size Measurement

Tuna Hacaloglu[1-2], Huseyin Unlu[3], Onur Demirors[3], Alain Abran[4]

[1] Middle East Technical University, Ankara, 06531, Turkey
[2] Atilim University, Ankara, 06830, Turkey
tuna.hacaloglu@atilim.edu.tr
[3] Izmir Institute of Technology, Izmir, 35430, Turkey
[huseyinunlu,onurdemirors]@iyte.edu.tr
[4] École de Technologie Supérieure, Montréal, Canada
Alain.Abran@etsmtl.ca

Abstract. Functional size has been used in software engineering for more than 40 years. When measured early in the software development life cycle, it can serve as direct input for effort estimation. The COSMIC Functional Size Measurement (FSM) method developed by the Common Software Measurement Consortium (COSMIC) is the latest ISO-compliant functional sizing method. A streamlined manual titled "Software Development Velocity with COSMIC Function Points" summarizes the measurement process and shortens the learning time. The aim of this study is to compare the classic COSMIC FSM manual and this new "light" manual in terms of accuracy of the resulting FSM applied to case studies. The findings show that use of the light manual results in accurate measurement. In addition, there were no significant time differences between the two. With respect to the variations in COSMIC Function Points (CFP) values in the two case studies, they three causes were identified: the Object of Interest (OOI) concept and corresponding data groups, details regarding Functional Process Independence, and Error/ Confirmation messages related to the scope of the information included in the manuals.

Keywords: COSMIC, Function Points, Software Measurement, Size, ISO 19761

1   Introduction

Software size measurement is a significant activity in software project management, as it is the main input for effort and schedule estimation, which provides a key business advantage [1, 2]. Successful effort estimation may reduce potential risks such as schedule and budget overruns. Reliable software size measurement is therefore of critical importance.

Functional size measurement (FSM) has been one of the widely used size measures for more than 40 years [3]. Being based on standardized methods, FSM provides an objective size measurement method for establishing estimation models, internal process improvement, and benchmarking [4–6]. Functional size can also be utilized for keeping track of the project's scope change, establishing an agreement between the supplier and acquirer, and improving the organization's processes by normalizing performance and quality measures [7].

COSMIC [8] is one of the most commonly used [9] FSM methods accepted by ISO [10]. The measurement process of COSMIC is based on calculating data movements, such as Entry (E), Exit (X), Read (R) and Write (W), in functional processes that are triggered by functional users.

There are various measurement guidelines having varying levels of detail for measuring the size of software in different domains such as business application software [11], real-time software [12], and service-oriented software [13]; some of these are also supported with case studies.

All of these guidelines are based on the "COSMIC Measurement Manual" [8], which provides the rules and definitions in "The COSMIC Implementation Guide for ISO/IEC 19761: 2017" [10] as well as further guidance and examples to help measurers fully understand how to apply the method. However, because it describes the measurement process in detail, this manual, at 115 pages, can be overwhelming for software developers, especially if they are just at the learning stage or have limited time available. Presenting the method in a concise manner could speed up the learning process and encourage its use early on in the estimation process.

Accordingly, Abran [14] has recently published "Software Development Velocity with COSMIC Function Points", which presents a "light" version of COSMIC Function Points (FPs) and can be used to size the software functionality that can be delivered within a given time period. This new manual summarizes the COSMIC Measurement Manual in only 13 pages. Its aim is to facilitate the learning process for software developers being introduced to COSMIC as well as for those already using it. However, regardless which manual is chosen for learning COSMIC, the resulting size measurements must be consistent.

In the literature, only a handful of studies [15, 16] compare the effect of measurers' understanding and assumptions on measurement accuracy. There is no study comparing measurement accuracy resulting from the use of different manuals of the same FSM method.

In light of this situation, the aim of the present study is to explore how using either the official "COSMIC Measurement Manual" (Manual 1) or "Software Development Velocity with COSMIC Function Points" (Manual 2) affects measurement results. It reports on the results of two business application case studies where two projects were measured by two measurers using Manual 1 and Manual 2. We compare the COSMIC Function Points (CFP) values calculated against the effort actually spent, and we discuss the causes of the difference in measurement results. Our findings reveal the strengths and difficulties of each of the two manuals and also suggest possible improvements to both.

The remainder of this paper is structured as follows: Section 2 presents the research methodology and the cases. In Section 3, the results obtained are presented in terms of the calculated CFP values and effort spent, and the cause of the difference in

measurement results is discussed. Section 4 presents the threats to validity, and finally Section 5 presents the conclusion and possible future work.

2   Research Methodology

The aim of this study is to investigate how using either Manual 1 ("COSMIC Measurement Manual") or Manual 2 ("Software Development Velocity with COSMIC Function Points") affects measurement results. In other words, we wanted to evaluate the effect of two different representations of the measurement procedure on measurement results considering the accuracy of CFP values, the causes for the differences, and the effort spent on the measurement process. To achieve this aim, the case study research method was adopted. The following subsections present the case study details.

2.1   Case Study Design

To explore the effect of using either Manual 1 or Manual 2 on the measurement results, the following research questions were formulated:

RQ1. Is there a difference between CFP values when measuring using Manual 1 versus Manual 2?
RQ2. What are the causes of the difference in measurement results?
RQ3. How does the effort spent on measurement differ when using Manual 1 versus Manual 2?

Case Selection Criteria. In selecting the cases, the first criterion was to selected projects that measurers were not familiar with, in order to ensure the objectivity in the measurement process. We therefore selected cases from two different business application projects. Another criterion was the measurers' ability to understand the project domain; for this reason, we chose a regular business application type of software. The last criterion was the availability of functional requirements documentation at an understandable level of detail for both measurers.

Data Collection Procedure. The measurers collected and recorded the CFP values obtained as a result of the measurement, the time (in minutes) spent measuring the functional size of the projects, and the areas where they made assumptions or needed more information to make a judgment within the measurement process.

Measurement Planning. The size of each project was to be measured by two measurers: Measurer 1 (M1) has three years of experience with the COSMIC FSM method, and Measurer 2 (M2) has one year. The measurers have not taken the certification exam; however, both have taken a formal COSMIC course and together they have measured more than 20 projects with over 5,000 function points. The measurement process was planned in two stages:
- In the first stage, M1 was to measure the functional size of Case 1 using Manual 2, and M2 was to measure it using Manual 1.

- In the second stage, the measurers were to perform cross-measuring such that M1 measures Case 1 based on Manual 1, and M2 measures it based on Manual 2.

Next, we planned the same setup for Case 2, but reversed: M1 measures Case 2 based on Manual 2, and M2 measures it based on Manual 1.

On the one hand, each project was measured by each measurer and each of the manuals was used in the measurement process. On the other hand, although both measurers had experience with COSMIC measurement and had already read both manuals, while measuring they would keep to the guidelines of the manual (1 or 2).

Finally, for consistency, measurement of each case was scheduled to start at the same time by each measurer. The starting and end times of each measurement would also be recorded to calculate the amount of time spent on the measurement.

In addition, the results were discussed by the measurers immediately after the measurements, so all the differences and misunderstandings could be addressed and a consensus reached. The measurement plan is visualized in Figure 1.
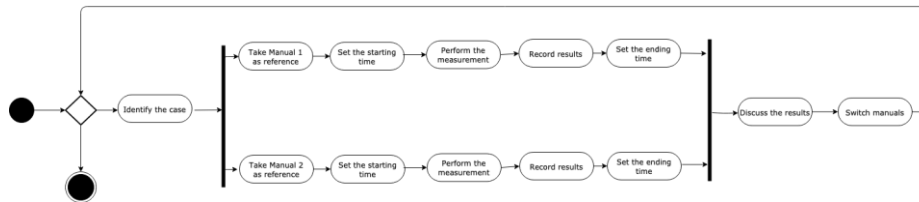

Fig. 1. Measurement plan with the two case studies

2.2 Description of Cases

The two case studies selected were from a domain familiar to both measurers. Unfortunately, we cannot share the case details due to privacy concerns.

- Case 1 is a business application software; written in a detailed way similar to a complete Software Requirements Specification (SRS) format, it includes 13 use cases. The documentation included use case descriptions, user interfaces, and an entity-relationship diagram (ERD) in the SRS document.
- Case 2 is also a business application. We selected the software analysis document from a single sprint. In contrast to Case 1, there were no use case descriptions or data models such as an entity-relationship diagram in this sprint documentation. Instead, the functional user requirements (FURs) were written in simple textual descriptions together with screen mock-ups.

2.3 Measuring the Case Studies

The measurement was performed using the functional user requirements and user interfaces included in the case documents as inputs. Following the measurement plan described in the previous section, the measurement for each case was performed according to the level of detail provided in the selected manual to identify the data movements (Entries, Reads, Writes, and Exits) and add them up to obtain the functional size of the software project. The effort spent was recorded by each measurer during each measurement process.

Upon completion of each measurement, the results were discussed in terms of effort, measured size, and guidance provided by the manual; this discussion also included an overall evaluation of the functionality of the two manuals. In addition, during these discussions, when an unclear functional user requirement for the measurement had been observed and when both measurers agreed, these were eliminated.

3 Results and Discussion

3.1 General Findings

The measurement results of each measurer are presented in Table 1 for the two case studies and the measurement manual used. The detailed COSMIC size measurement manual (Manual 1) is the correct size measure. The correctness of the measurement has been verified by an independent measurer. As shown in Table 1, there is no significant difference in terms of the amount of time spent between the two measurers when using different manuals for each of the two cases. Here it is worth repeating that Manual 1 has 5 chapters and 115 pages whereas Manual 2 has 4 chapters and 13 pages. However, even though different manuals—with very different levels of detail—were used, the time spent by the measurers is similar.

Table 1. Measurement effort and measured size for each case

Case	Measurement effort (minutes)		Measured size (CFP)	
	Manual 1 (Heavy)	Manual 2 (Light)	Manual 1 (Heavy)	Manual 2 (Light)
Case 1	88	93	71	85
Case 2	35	48	77	116

Interestingly, even though Manual 1 is much more detailed than Manual 2, the measurers who used Manual 1 as the reference completed the measurement in less time in both cases. However, when we look at the functional size values in terms of CFP, in both cases the measurers who used Manual 2 found higher CFP values.

These findings imply that measurers using Manual 2 (i.e., the light version of Manual 1) can:

- perform the measurement; and
- identify more functional size-related items, but it takes more time for this specific study.

One of the reasons for the change in CFP values observed in the measurement results of Case 2 is related to the EXIT data movement. For EXITs, both manuals specify the following rule: "*An Exit accounts for all data manipulation to create the data group attributes to be output and/or to enable the data group to be output (e.g. formatting and presentation manipulations) and to be routed to the intended functional user.*"

However, in Case 2, it can be seen that Measurer 1, using Manual 1, identified more EXIT data movements than Measurer 2, who used Manual 2. This is because in Case 2 there were requirements specifying that the reports displayed on the screen could be exported in Microsoft Excel format. The measurer who used Manual 1 considered this

extra type of display as another functional process and counted it as an EXIT data movement. Measurer 2, on the other hand, did not define these as new EXITs. Consequently, the authors suggest that this issue needs to be clarified in both manuals.

Apart from these comparisons of CFP values and measurement effort, upon completing the measurements the measurers met to evaluate the measurement process and identify the causes for any variations in CFP values for each requirement in the two case studies. From these discussions, the causes of these variations were classified into three groups with respect to detailed (Manual 1) versus summarized (Manual 2):

1) Object of Interest (OOI) concept and data groups
2) Functional Process independence, and
3) Error/Confirmation messages.

3.2   Variation in CFP Values Related to OOI and Data Groups

Object of Interest (OOI) is an important concept in the COSMIC FSM method, where a data movement can only be taken into account when it conveys a data group belonging to the same OOI. This concept is described in detail in Manual 1, which also refers to the Guideline for Sizing Business Application Software [11], which includes a special section encouraging measurers to use different data analysis methods such as Entity-Relationship Analysis, Relational Data Analysis, and UML Class Diagram to identify OOI. However, OOI is not mentioned in Manual 2. Moreover, identification of persistent and transient data groups used to define OOI are described in Manual 1 but are not mentioned in Manual 2. For this reason, when measuring Case 2, Measurer 1 paid special attention to the OOI concept and found more data movements while Measurer 2, who performed the measurement using Manual 2, did not consider this concept.

In addition:

- In Case 2, one of the functional user requirements refers to a search by a filtering feature: Measurer 2, using Manual 2, added an extra CFP for each filtering feature, although it is an attribute of the same OOI. This led to greater CFP values for the Case 2 measurement based on Manual 2.
- In Case 1, data analysis in the form of ERD in the requirements document helped Measurer 1 to identify the data groups without considering OOI.

3.3   Variation in CFP Values Related to Functional Process Independence

Another important component of COSMIC FSM is the Functional Process. Manual 1 gives a detailed definition of the Functional Process while Manual 2 offers only a short description. More specifically, in Manual 1, under the functional process identification section, "Independence of functional processes" is explained in detail whereas this detail is not mentioned in Manual 2.

Manual 1, states that "... *in the COSMIC method (as in all other FSM Methods) each functional process is defined, modeled and measured independently of, i.e. without reference to, any other functional process in the same software being measured…".* In addition, Manual 1 further specifies that "*when a statement of a FUR is implemented in software, any "functional commonality" may or may not be developed as reusable*

software. All implementation decisions including the extent of actual or potential software reuse must be ignored when measuring functional size.

The following observations apply to Case 2, where functional similarities had been observed:

- Complying with the guidance information given in Manual 1, Measurer 1 included the functional similarities in his CFP calculation.
- Since this detail is not mentioned in Manual 2, Measurer 2 eliminated the functional similarities in the functional size measurement.

3.4 Variations in CFP Values Related to Error/ Confirmation Messages

Manual 1 includes detailed guidance on error/ confirmation Messages, whereas in Manual 2 this is only briefly presented under the EXIT type data movement-related description.

More specifically, Manual 1 [8] states that "*If the FUR of the functional process does not require any type of error/confirmation message to be issued, do not identify any corresponding Exit*". Therefore, in Case 2, which did not have a FUR specifying the information related to system behavior regarding error or confirmation messages:

- Measurer 1, by referring to Manual 1, did not identify or count any related data movement.
- Measurer 2 made an assumption and included the error/ confirmation messages in the measurement process.

This lack of detail regarding error/ confirmation messages in Manual 2 causes inconsistencies in the measurement. For better measurement consistency, this information could also be added to Manual 2.

On the other hand, in Case 1, the system behavior regarding error or confirmation messages was specified in use case descriptions. Thus, there were no major inconsistencies in Case 1 regarding error or confirmation messages.

3.5 Other Comments Regarding Measurement

During the case studies, some problems already stated in the literature were also observed, such as the importance of the quality of requirements documents [17–19], difficulty identifying OOIs [14] due to the lack of data modeling [20] or errors in the data model (in Case 1), and challenges due to individual assumptions and understandings [15]. For example:

- Even though both measurers were consistent within their own measurements, their understanding of the Manuals may have differed: for instance, one of the measurers defined different EXITs for each of the confirmations and errors throughout the measurement whereas the other one counted them as pairs.
- The lack of data models and the paucity of information regarding requirements caused a challenge for Case 2: there was a combo box, but whether it retrieved its content from persistent storage or not was not specified in the requirements document. For this reason, the number of OOIs and the type of data movement involved were not clear. In addition, even though the Guideline for Sizing Business Application Software [11] specifies details related to combo boxes, neither Manual 1 or Manual 2 includes such detail.

4 Threats to Validity

One of the validity threats can be related to measurer experience. Measurer 1 had three years of COSMIC experience while Measurer 2 had only one. They have not taken the certification exam; however, both have taken a formal COSMIC course and together they have measured more than 20 projects with over 5,000 function points. To overcome the difference in measurer experience that could have changed the calculated CFP values and the measurement effort, cross-measurement was used: Each measurer measured one case based on Manual 1 and the other based on Manual 2. At the end of each case, the measurement results were discussed, and if there was a functional user requirement that was unclear to the measurers, it was eliminated. In addition, all measurements were verified by an established COSMIC measurer.

Another threat to validity is that both measurers had previous COSMIC knowledge and had previously measured different projects using Manual 1; this can affect the Manual 2 measurement results. To overcome the effect of Manual 1, measurers only applied the rules in Manual 2 for the measurements based on that manual: for instance, "object of interest" is not explained in Manual 2 and measurers did not consider this term in their measurements.

Our research methodology is clearly given in Section 3, and if another researcher replicates this study, then the same procedures can be applied.

5 Conclusion and Future Work

This study presented an experiment to explore the impact on measurement of using two distinct versions of the COSMIC FSM method [8, 14].

A major finding is that, when using Manual 2 as the reference, the measurers identified more data movements – e.g., higher CFP values. Therefore, it can be concluded that even though Manual 2 is the light version of Manual 1, it is still capable of yielding good measurement results.

Analysis of the detailed measurement results obtained while using these two different manuals identified that the measurement differences came from three measurable concepts--Object of Interest (OOI) concept and corresponding data groups, Functional process independence, and Error/Confirmation messages--in particular when the quality and completeness of the project documentation varied. The authors suggest the incorporation of details related to these concepts into Manual 2.

When the effort spent on measurement while using Manual 1 and Manual 2 was compared, no significant difference was noticeable. However, in both case studies, it took slightly more time to complete the measurement using Manual 2.

Manual 2, the light version of the COSMIC FSM method, is intended to facilitate and speed up the learning process for newcomers to the COSMIC method. Future work is needed to assess the usability and learnability of both manuals; for example, another experiment could be conducted with newcomers to the COSMIC FSM methods trained using either Manual 1 or Manual 2. In this way, both the learning time and the measurement results could be compared. In addition, effort estimation models could be developed and the effectiveness of each manual could be compared on the basis of the results they provide in terms of effort estimation. Future studies on questions such as

"Would it be useful/essential to go through the 115 pages of Manual 1 when the recommended minor additions to Manual 2 would also be valuable."

References

1. Salmanoğlu, M., Öztürk, K., Bağrıyanık, S., Ungan, E., Demirörs, O.: Benefits and challenges of measuring software size: early results in a large organization. Presented at the 25th International Workshop on Software Measurement and 10th International Conference on Software Process and Product Measurement, IWSM-Mensura (2015).
2. Abran, A.: Software project estimation: the fundamentals for providing high quality information to decision makers. John Wiley & Sons (2015).
3. Albrecht, A.J.: Measuring application development productivity. Presented at the Proc. Joint Share, Guide, and IBM Application Development Symposium, 1979 (1979).
4. Abran, A., Desharnais, J.-M., Zarour, M., Demirörs, O.: Productivity-based software estimation models and process improvement: an empirical study. Int J Adv Softw. 8, 103–114 (2015).
5. Commeyne, C., Abran, A., Djouab, R.: Effort estimation with story points and cosmic function points-an industry case study. Softw. Meas. News. 21, 25–36 (2016).
6. Trudel, S., Buglione, L.: Guideline for sizing Agile projects with COSMIC. Presented at the Proceedings of International Workshop on Software Measurement (2010).
7. Ozkan, B., Turetken, O., Demirors, O.: Software functional size: For cost estimation and more. In: European Conference on Software Process Improvement. pp. 59–69. Springer (2008).
8. COSMIC Measurement Manual Version 4.0.2. The Common Software Measurement International Consortium (2017).
9. Abran, A.: Automating Functional Size Measurement–a Survey. Presented at the UKSMA/COSMIC Conference 2011-22nd Annual conference on Metrics and Estimating: hosted in collaboration with COSMIC (2011).
10. ISO/IEC 19761:2017 - Software Engineering – COSMIC: A Functional Size Measurement Method. International Organization for Standardization (2017).
11. Guideline for Sizing Business Application Software. The Common Software Measurement International Consortium (2017).
12. Guideline for Sizing Real-time Software. The Common Software Measurement International Consortium (2015).
13. Guideline for Sizing Service-Oriented Architecture Software. The Common Software Measurement International Consortium (2015).
14. Abran, A.: Software Development Velocity with COSMIC Function Points. The Common Software Measurement International Consortium (2019).
15. Top, O.O., Demirors, O., Ozkan, B.: Reliability of COSMIC functional size measurement results: A multiple case study on industry cases. Presented at the 2009 35th Euromicro Conference on Software Engineering and Advanced Applications (2009).
16. Turetken, O., Top, O.O., Ozkan, B., Demirors, O.: The impact of individual assumptions on functional size measurement. In: Software Process and Product Measurement. pp. 155–169. Springer (2008).
17. Trudel, S., Abran, A.: Functional Size Measurement Quality Challenges for Inexperienced Measurers. In: Abran, A., Braungarten, R., Dumke, R.R., Cuadrado-Gallego, J.J., and Brunekreef, J. (eds.) Software Process and Product Measurement. pp. 157–169. Springer, Berlin, Heidelberg (2009). https://doi.org/10.1007/978-3-642-05415-0_12.
18. Desharnais, J., Buglione, L., Kocaturk, B.: Improving Agile Software Projects Planning Using the COSMIC Method. Presented at the workshop on Managing Client Value Creation Process in Agile Projects (Torre Cane, Italy (2011).

19. Desharnais, J.-M., Kocaturk, B., Abran, A.: Using the COSMIC Method to Evaluate the Quality of the Documentation of Agile User Stories. In: 2011 Joint Conference of the 21st International Workshop on Software Measurement and the 6th International Conference on Software Process and Product Measurement. pp. 269–272 (2011). https://doi.org/10.1109/IWSM-MENSURA.2011.45.
20. Hacaloglu, T., Demirors, O.: Measureability of Functional Size in Agile Software Projects: Multiple Case Studies with COSMIC FSM. In: 2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). pp. 204–211 (2019). https://doi.org/10.1109/SEAA.2019.00041.