Using Function Points in Agile Projects: A Comparative Analysis Between Existing Approaches

Eduardo Garcia Wanderley[1(✉)], Alexandre Vasconcelos[2], and Bruno Tenório Avila[2]

[1] IFPE – Instituto Federal de Educação, Ciência e Tecnologia, Garanhhuns, Pernambuco, Brazil
eduardo.wanderley@garanhuns.ifpe.edu.br

[2] UFPE – Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil
amlv@cin.ufpe.br, brunotavila@gmail.com

Abstract. Agile Software Development has become increasingly common in the software development environment, but effort estimates in software projects using Agile methodologies are made differently from those made in traditional way projects. This paper presents a comparative analysis of the different approaches of applying Function Point Analysis (FPA) in software projects that make use of some existing agile methodologies. Through an experimental, empirical and controlled research, the existing proposals in the literature in order to test your application and analyze its results were evaluated. The results showed that in the context studied, the approach Agile Estimation Using Functional Metrics was best suited up.

Keywords: Effort estimation · Cost estimation · Size estimation · Agile Scrum

1 Introduction

Agile methods are approaches for software development based on iterative and incremental development, direct involvement of customer, early delivery of higher business values and rapid responses to changes [1]. According to Soares [2], agile methods do not reject process and tools, documentation, contract negotiation or planning, but simply define them as secondary importance.

As reported by Mens and Demeyer [3], the companies' demand for products and services in the information technology field has been consistently increased, however the budgets and schedules are getting inflexible, due, for instance, to the need of cost-cutting policy and shorter deadlines for on time delivery performance. Therefore, the increase in quality and productivity becomes essential to software development. For this purpose, software metrics are used to: (i) serve as source of information to the monitoring of software process current situation; (ii) determine the evolution of software development; (iii) identify delays and deviations during project implementation; (v) prioritize accurate data over subjectivity and intuition.

Researcher's reviews have illustrated that in the last few decades, too little progress was made in the software estimation field [4]. This is a major concern in industry, because the low performance in estimation process very often leads to exceeded budget,

V. A. Santos et al. (Eds.): WBMA 2017, CCIS 802, pp. 47–59, 2018.
https://doi.org/10.1007/978-3-319-73673-0_4

delays, sub-optimal allocated resources, contract losses or very low quality software [5]. Due to these problems, there is a high demand for enhancements in estimation process.

The development time of a feature in software can be measured by its size. However, there are several forms of measuring the size of software [6]. One of the best forms is the Automated Function Point (AFP) specification [7] which was standardized by the ISO/IEC 20.926 regulation [8] as a functional size measurement, added to the ISO/IEC 14.143 regulation [9]. A research from the Quality Assurance Institute [10] pointed that AFP is the proper metric to measure quality and productivity in software projects. In 1993, AFP became the most applied and studied metric in software engineering [11].

The advantages of using AFP are: (i) it is an objective method to find the score of an element; (ii) it might serve as a comparative basis between two different projects or even two different teams [12]; (iii) to equate communication and expectative about the system size; (iv) system size is defined by the customer insight [13]. Due to these circumstances, it is noticed that the use of AFP has become consolidated [14]. About 371 out of 893 of the licensed AFP professionals in the world are from Brazil [15].

In Brazil, many private and public institutions started using AFP on account of the several benefits this metric presents, and yet, by recommendation of regulatory agencies, it has been utilized in the contracts of development and maintenance of software in government organizations, which made AFP unanimous in the national metric industry [16]. In Brazilian public administration domain, normative instruction no. 4 [17], article 10 clearly outlines that the estimated financial cost must be measured. Article 14 affirms that the procedures planning and the measuring of the offered services – including metrics, indicators and values must be done.

Consequently, software development companies that already employ agile methodology must now have planning and measuring at top of agenda. Nonetheless, during agile development process, the software is incrementally developed, made in small iterations and the customers' feedback is an important asset to the following iterations. This implies that planning and estimation must be accomplished progressively [18]. Furthermore, taking into account the high number of iterations in agile methods, planning and estimation are done differently from classic software development [19]. From the perspective of estimation, All metrics and conventional life cycle models can be used in agile methods, but they require adaptation [20]. Therefore, specific techniques for agile methods were proposed and are being utilized in projects [1].

In this article a comparative analysis has been made between three different approaches (Extending Function Point Analysis [21], Function Point Analysis [22], and Functional Metrics [23]) to assess which was the ideal form of measuring effort in agile projects. The choices were mainly based on a systematic literature mapping. From this mapping, the three most relevant metrics were selected.

As a result of this work, it is noticed that two approaches can be satisfactorily used, and that a third one had been assessed truly insufficient and, therefore, may be discarded from all types of projects, not only the ones mentioned in this study. It is expected that the outcomes of this article may also support the using of effort metrics in agile methodology projects, benefiting mainstream software development community and, possessing a more accurate effort estimation, may bargain longer deadlines and better costs, assisting both customer and developer and, also, simplifying the adoption of AFP techniques by the companies that already make use of agile methods in their projects, leading the academic community to find new paths of research.

Due to the page break, some aspects of the experiment were suppressed. To obtain a more complete version, please contact us. This article is based on the master degree dissertation, presented by the main author, in 2015 [24].

2 Method

In software engineering, an experiment is a procedure carried out to combine facts with assumptions, speculations and beliefs so plentiful in software development [25]. It presents a systematic, disciplined and computable mode of assessment of human activity. Experiments are proper processes to achieve the validation of theories, confirm common knowledge, explore the relationships, assess model predictions or validate the measures. The possibilities of repetition, the full control over the process and the variables compose the major asset of the experiment.

The goals of an experiment study are the description, assessment, prediction, control and enhancement of products, processes, resources, models, theories etc. [26].

2.1 Experiment Planning

2.1.1 Goals

The GQM (Goal Question Metric) [26] approach was used to organize the objectives of this work, whose structure is presented as follows:

To assess the process of development effort estimation.
Intending to understand the utilization of different estimation processes.
In terms of accuracy and ease of application.
From the developers **point of view.**
In the context of the system analysts.

2.1.2 Metric and Question

What is the most accurate approach? Extending Function Point Analysis [21], Function Point Analysis and Cost Estimation in An Agile Development Environment [22] or Agile Estimation Using Functional Metrics [23]? The choices were given through a systematic literature mapping. From it, the three most relevant metrics were selected. The article about the non-selected approaches that were considered relevant, did not present suiTable detailing about their right use, and for this reason, were disregarded. The criteria used to choose the best estimate was to elect the closest to the real value. For instance, if the real value was 50, the best estimate was the one that presented the closest value to 50. In order to appraise the accuracy of the experiment, the estimated value (measured in weekly function points) will be considered.

2.2 Experiment Description

2.2.1 Hypothesis

The main hypothesis of this study is the null hypothesis that says there are no significant statistic differences between the three approaches. Therefore, the work attempted to refute this hypothesis. The possible hypotheses are listed below:

- **Null hypothesis:** The estimated value (EV) is not significantly different when utilized the three distinct approaches.
- **Alternative hypothesis 1:** The estimated value is more accurate utilizing *Extending Function Point Analysis* approach.
- **Alternative hypothesis 2:** The estimated value is more accurate utilizing *Function Point Analysis and Cost Estimation in an Agile Development Environment* approach.
- **Alternative hypothesis 3:** The estimated value is more accurate utilizing *Agile Estimation Using Functional Metrics* approach.

2.2.2 Unity and Experimental Subject

The objects on which the experiment is based are called experimental units or experimental objects. Two studies from a municipal company were selected.

The person who applies the methods and techniques in the experimental unity is called experimental subject. This study employed eight system analysts who have never had contact with function points and development effort estimation.

2.2.3 Parameters and Factors

They are the input variables for the controlled experiment that defines the parameters (variables whose values will be kept constant throughout the experiment) and factors (Variables representing the procedures to be received by the experimental objects). They have as independent variables the measurement approaches and as dependent variables the estimated value. Since the analysts having different professional experience in estimation and the projects being distinct influence the experiment outcomes, non-experienced analysts were selected. To avoid interference in the experiment, the eight analysts were randomly arranged.

2.3 Execution Planning

2.3.1 Context

The experiment execution is part of the curriculum of the postgraduate student of Federal University of Pernambuco Informatics Center. The analyst team was preferred among the industry of Pernambuco state. All members are, at least, five years experienced system analysts.

2.3.2 Training and Participant Profile

Intending to standardize and ensure a minimum level of knowledge before the experiment performance, a few trainings were planned and performed earlier than the launch of the experiment pilot experiment scheme. The first training group was related to the function point technique. Firstly, they were introduced to the concepts and how to do the counting. After the conceptual starting point, exercises were practiced to assure that there was a proper understanding of the techniques use. Already knowing about function points, the participants had training in the experiment approaches. A total of 19 h were spent at the training sections, where 10 h were employed for function points whilst the approaches had been left with 3 h each.

Aiming to form cohesive and uniform teams, in which the effect related to the knowledge and experience could be mitigated, it was created and applied a characterization survey, where the participants could answer and be profiled.

2.4 Experiment Design

It is known that the proposed design for experimentation critically influences the formation of the teams. The Latin square was chosen as experimental model, due to its ideal suitability to the experiment. As stated in Juristo and Moreno [27], the Latin square presents as characteristic that each factor occurs once in each line and once in each column, it means that, it may take place only once to each possible combination of two blocking variables. For one to use the Latin square, it is necessary to set three different teams, where each of them will execute all approaches in each of the projects.

2.4.1 Round Execution and Pilot Scheme

With the purpose of systematizing data collection, some data (approaches/study, projects and analyst group) were divided into three groups, where two groups had three analysts each and one group had two of them. All groups executed the proposed activities and collected the data containing in the data collection form, to each one of the studies and with dissimilar tools.

A pilot stage was launched in order to validate the elements of instrumentalization and the data processing. At this stage, a fake project and different participants from the real experiment were used. At this point it was possible to diagnose the main issues and doubts the participants could have and, therefore, have them answered and not compromise the experiment performance.

2.4.2 Data Analysis Processing

In order to not interfere with the data analysis, approaches and analysts' info were omitted from data collection before data analysis.

2.5 Methods for Data Analysis

The analysis of the study aimed to compare the data collected from the experiment performance and check if the null hypothesis could be rejected. This analysis assessed the approaches utilized to estimate the software development effort in agile projects.

The performed experiment presented only one factor, and the distribution of the projects and approaches between the subjects were randomly made, by lot. With the purpose of rejecting null hypothesis, the ANOVA (analysis of variance) statistical model was utilized [29]. To assess the magnitude of these differences, it was chosen the Turkey's range test which is a single-step multiple comparison procedure and statistical test [30].

The ANOVA [29] statistical model says that in an experiment, each Y_{ij} observation may be decomposed as the following model equation:

$$Y_{ij} = \mu + \tau i + \in i_{ij} i = 1, \ldots, I e j = 1, \ldots, J$$

Where μ is the constant effect (general mean), τ_i is the effect of the *ith* procedure, ϵ_{ij} is the error related to the *ith* procedure in the *jth* experimental unit.

ANOVA is based on the partition of the total variance of response variable into components attribuTable to the procedures (variation among) and experimental error (variation within). This variation can be measured by the sum of squares defined to each of the components as presented in the following equations:

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \boldsymbol{SST}(\boldsymbol{Sum\ of\ Squared\ Total})$$

$$\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \boldsymbol{SSR}(\boldsymbol{Sum\ of\ Squared\ Regression})$$

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \boldsymbol{SSE}(\boldsymbol{Sum\ of\ Squared\ Error})$$

Thereby, the equation is written: SST = SSR + SSE, in which the sum of squared totals is decomposed into the sum of squared regression and sum of squared error. This sum of squares is organized in a Table, named variance analysis Table (Table 2).

To test the H0 hypothesis (null), the F-test was used, as presented at Variance Analysis Table (Table 1).

Table 1. Variance analysis table.

Causes of variation	Degrees of freedom	Sum of squares	Median squares	F calculated
Procedures	I−1	SSR	MSR = SSR/(I−1)	MSR/MSE
Residual	I(J−1)	SSE	MSE = SSE/(I(J−1))	
Total	IJ−1	SST		

So whether F-calculated is higher than F-Tabled, then the null hypothesis (H0) must be rejected, it means that there are evidences of significant difference among at least one pair of treatment means, at α chosen level of significance. Otherwise, null hypothesis (H0) must not be rejected; it means that there are no significant differences between the treatments, at α chosen level of significance.

Consequently, the collected data will be assessed in order to try to reject the null hypothesis revealing that the mean of techniques are not equal, taking into account a certain significance level (α). At this experiment, regarding the small quantity of participants involved, there will be considered a 0,05 level of significance (α).

In other words, there will be adopted a 95% confidence interval on the outcomes of this study.

After concluding that there is significant differences among treatments by F-test, it was assessed the magnitude of these differences by utilizing a single-step multiple comparison procedure.

Tukey's range test allow one to test any contrast, always between two treatment means, it implies that it does not permit to compare groups with each other. The test is based on the Honest Significant Difference (HSD). The test statistic is given by the equation below:

$$TSD = q_{\alpha}(k, N - k)\sqrt{MSE/n}$$

Where q is the total studentized amplitude, Tabled, MSR is the mean square of residuals and n is the number of repetitions. The value of q depends on the number of treatments and residual degree of freedom. It was utilized a 5% level of significance. In case that the contrast is higher than Δ, the measures differ at α level of significance.

Since Tukey's test is somewhat independent from F-test, it is possible that, even with the value of F-calculated being significant, one may not find significant differences among the mean of contrasts [28].

3 Outcomes

The project/studies utilized at the experiment were chosen from a public company of Recife/PE, whose real values were selected without participants knowing. Thereafter, it were compared the estimations made by each one of them with real values, calculating the estimation error. The time necessary for colleting was not considered, because it was not an experiment concern. This fact gave participants freedom to spend the time they need.

3.1 Description of the Study

Previously than any outcome analysis it is necessary to present the characteristics of each study. These characteristics are:

- E1 study: New system with several features, developed in 3 iterations.
- E2 study: New feature for an ongoing system, with 2 iterations.

3.2 Collected Data

Listed below are the data collected from the participants. The Tables 2 and 3 illustrate the data divided into two studies. The first lines of each Table represent the estimation error value, presented in number of weeks. The value was obtained by subtracting the time estimated by each participant by the real time spent at the study development. The line 'means' reveals the mean of error of the approach. The line called 'general mean' illustrates the mean error value among the approaches. The line named 'effect' demonstrates the effect of every approach on response variable. The effect is calculated by subtracting the error mean of each approach by the general mean of the experiment. It is worth noting that the participant 8 did not deliver the data related to E1 study. Due to this fact, the E1 study remained with seven valid experiments.

Table 2. Experiment data with E1 study.

E1 study	CA approach	FM approach	EX approach
Participant 1	2,5875	0,1500	0,0309
Participant 2	2,5875	0,1125	0,2441
Participant 3	2,5875	0,6750	0,8333
Participant 4	1,2750	0,1875	0,3226
Participant 5	2,5875	0,4875	0,3844
Participant 6	1,2750	0,1125	0,0084
Participant 7	1,2750	0,6000	0,5025
Mean	2,0250	0,3321	0,3325
General Mean		0,8965	
Effect	1,1285	−0,5643	−0,5642

Table 3. Experiment data with E2 study.

E2 study	CA approach	FM approach	EX approach
Participant 1	1,5000	0,0750	0,3630
Participant 2	1,5000	0,0750	0,3630
Participant 3	1,5000	0,7875	0,7875
Participant 4	1,5000	0,2250	0,5020
Participant 5	1,5000	0,0750	0,4328
Participant 6	1,5000	0,4125	0,5896
Participant 7	1,5000	0,4125	0,5896
Participant 8	1,5000	0,0375	0,3806
Mean	1,5000	0,2625	0,5010
General Mean		0,7545	
Effect	0,7455	−0,4920	−0,2535

Tables 4 and 5 represent the errors (residuals) of each observation in relation to the error mean of each approach. These values illustrate the difference between the values obtained from each participant and the mean of each approach.

Table 4. Residuals related to each observation at E1 study.

E1 study	CA approach	FM approach	EX approach
Participant 1	0,5625	−0,1821	−0,3014
Participant 2	0,5625	−0,2196	−0,0882
Participant 3	0,5625	0,34286	0,5010
Participant 4	−0,7500	−0,1446	−0,0097
Participant 5	0,5625	0,1554	0,0521
Participant 6	−0,7500	−0,2196	−0,3239
Participant 7	−0,7500	0,2678	0,1702

Table 5. Residuals related to each observation at E2 study.

E2 study	CA approach	FM approach	EX approach
Participant 1	0,0000	−0,1875	−0,1380
Participant 2	0,0000	−0,1875	−0,1380
Participant 3	0,0000	0,5250	0,2865
Participant 4	0,0000	−0,0375	0,0010
Participant 5	0,0000	−0,1875	−0,0682
Participant 6	0,0000	0,1500	0,0886
Participant 7	0,0000	0,1500	0,0886
Participant 8	0,0000	−0,2250	−0,1204

3.3 Null Hypothesis

Considering that each study presents their own characteristics, it was not possible to group their outcomes in a single experiment, thus each of them was assessed as a separated experiment and due to this fact, the null hypothesis was rejected in each project.

The ANOVA statistical method was utilized as above-mentioned, and it states that one must compare the value of F-calculated with the value of F-Tabled, and whether F-calculated > F-Tabled then the null hypothesis must be rejected. The P1 project presents treatment and residual degrees of freedom (2 and 21 respectively). Consulting T-Student Table [30] with the values of F-Tabled F(2,21) will reveal the value of 3,467. At P2 project, the value of F (2,18) was used and the number found was 3,555. At Tables 6 and 7, the calculated value of F was always higher the Tabled value, and consequently the null hypothesis must be reject in all performed experiments.

Table 6. Analysis of variance data related to E1 study.

Causes of variance	Degrees of freedom	Sum of squares	Median squares	F-calculated
Treatments	2	13,3722	6,6861	
Residuals	18	3,8035	0,2113	
Total	20			31,6423

Table 7. Analysis of variance data related to E2 study.

Causes of variance	Degrees of freedom	Sum of squares	Median squares	F-calculated
Treatments	2	6,8967	3,4484	
Residuals	21	0,6331	0,0301	
Total	23			114,37

3.4 Tukey's Test

It was calculated an honest significant difference (HSD) to each project. Conforming to Tukey [30], if the difference between the means is higher than HSD one must reject the hypothesis of equality among the mean levels. According to Table 8, the HSD value of 0.63 in E1 study is higher than the differences among the means of each approach, and for this reason, the hypothesis of equality among the mean levels must not be rejected. In E2 study, presented at Table 9, it may be noticed that the HSD for this experiment was 0.22 and the means of the Extending Function Point Analysis (EX), Function Point Analysis and Cost Estimation in An Agile Development Environment (CA) and Agile Estimation Using Functional Metrics (FM) approaches were 1.5000, 0.2625 and 0.5010 respectively. Since the differences between all approaches were higher than HSD, one must reject the hypothesis of equality among the mean levels with a 95% level of significance. Seeing that the means represent the errors in estimates, one concludes that in E2 study the Function Point Analysis and Cost Estimation in An Agile Development Environment approach were found to have the ideal performance.

Table 8. The means values and the differences among them at E1 study.

E1 study	Mean	Dif. with CA	Dif. with FM	Dif. with EX
HSD	0,6300	N/A	N/A	N/A
CA approach	2,0250	N/A	1,6929	1,6925
FM approach	0,3321	−1,6929	N/A	−0,0004
Ex approach	0,3325	−1,6925	0,0004	N/A

Table 9. The means values and the differences among them at E2 study.

E2 study	Mean	Dif. with CA	Dif. with FM	Dif. with EX
HSD	0,2200	N/A	N/A	N/A
CA approach	1,5000	N/A	1,2375	0,9990
FM approach	0,2625	−1,2375	N/A	−0,2385
Ex approach	0,5010	−0,9990	0,2385	N/A

3.5 Outcome Assessment

As illustrated in Sect. 3.3, the null hypothesis was effortlessly rejected, what reveals that the approaches are different from each other. In every studies, the Function Point

Analysis and Cost Estimation in An Agile Development Environment – CA [26] approach presented very divergent numbers from other approaches and a high level of errors.

At E1 study, the Extending Function Point Analysis – EX and Agile Estimation Using Functional Metrics – FM approaches handed out very similar outcomes and a reduced error mean, and for that it may be concluded that in similar projects, both approaches can be considered acceptable.

At E2 study, the Agile Estimation Using Functional Metrics – FM approach was, considerably the ideal approach to measure projects with characteristics related to the study. This was largely due to a problem with the Extending Function Point Analysis – EX approach. It presents a statement that in new stories developed in iteration, only 53.52% of the activities can be developed in the same iteration, in other words, it becomes possible to conclude only up to 53.53% of the development of a new story. It leads the iterations that have only one story to low performance, since they become limited to the above-mentioned percentage, which is the case of E2 Study.

4 Concluding Remarks

From the controlled experiment results, it was noticed that the Function Point approach Analysis and Cost Estimation in an Agile Development Environment - CA is a very specific to the environment which the approach was described, since it determines fixed values of hours for each ALI found independent ALI found independent of the features that will be developed and because of this did not go well when it was used in an environment other than proposed by the author. It was also noticed that, despite the Extending Function Point Analysis - EX and Agile Estimation Using Functional Metrics - FM have similarities and values, the FM approach presented a lower mean error, especially with the E2 study and therefore was considered the most adequate for the study.

4.1 Threats to Validity

4.1.1 Internal Validity

At this experiment a few system analysts were utilized as subject and they may have suffered influence throughout the project and it may lead to some changes at the outcome. In relation to the subjects' commitment, they could have become discouraged during working time, although to bypass the problem it was accorded to their managements that they could earn a time off for their help in the project. The fatigue of estimating could have been another factor, and for this the delivery time was increased in 40% so they could relieve the pressure on the estimates delivery.

4.1.2 Conclusion Validity

It is about to the correctness of applying the statistical tests on the outcomes obtained during the experiment, and for this, a statistic professional monitored the project. At this study, the Tukey's Test was utilized to compare the outcome data.

4.1.3 Construct Validity

The training applied to the developers on the estimate metrics may not be fully understood by them, thus affecting the outcomes of the experiment.

4.1.4 External Validity

At this study, metrics to estimate software development effort using agile method were tested. In order to guarantee such validity, one may repeat the experiment in different groups with other characteristics in consideration of ensure the outcomes can be generalized.

4.1.5 Empirical Reliability

The procedures for carrying out the research were documented in detail, seeking to serve as a source of information and so, to enable that other researchers can repeat it in the future, contributing to a greater empirical reliability.

4.2 Future Works

The following works are recommended:

- Develop case studies about the theme involving Brazilian software development companies;
- Develop action researches for testing if the outcomes of this experiment may be confirmed by using them with other companies.

References

1. Sommerville, I.: Engenharia de Software. Editora Addison-Wesley (2011)
2. dos Santos Soares, M.: Comparação entre Metodologias Ágeis e Tradicionais para o Desenvolvimento de Software. Unipac-Universidade Presidente Antônio Carlos (2010)
3. Mens, T., Demeyer, S.: Future trends in software evolution metrics. In: Proceedings of the 4th International Workshop on Principles of Software Evolution, IWPSE 2001, pp. 83–86 (2001)
4. França, L.P.A., et al.: Medição de Software para Pequenas Empresas: uma Solução Baseada na Web. PUC-RJ, Rio de Janeiro (1998)
5. Usman, M., et al.: Effort estimation in agile software development: A systematic literature review. In: Proceedings of the 10th International Conference on Predictive Models in Software Engineering, pp. 82–91. ACM (2014)
6. Kan, S.: Metrics and Models in Software Quality Engineering. Addison-Wesley, Boston (2002)
7. Albrecht, A.J.: Measuring application development productivity. In: Proceedings of the IBM Applications Development Symposium, p. 83. GUIDE, IBM Corp., Monterey (1979)
8. SO/IEC 20926: Disponível em (2002). www.iso.org/iso/cataloguedetail.htm
9. Dekkers, C.: Measuring the logical or function a Size of Software Projects and Software Application. Spotlight Software, ISO Bulletin, May 2003
10. Perry, W.E.: The best measures for measuring data processing quality and productivity. Quality Assurance Institute Technical Report (1986)

11. Jones, C.: Function points. Computer **27**(8), 66–67 (1994)
12. Santana, C., Gusmão, C.: Uso de Análise de Pontos de Funções em Ambientes Ágeis. In: Engenharia de Software Magazine, pp. 33–40, 20 December 2009
13. Oest, C.: Quando a Análise de Pontos de Função se Torna um Método Ágil? In: 2nd Conferência Brasileira de Medição e Analise de Software, São Paulo, Brasil 2011
14. BFPUG: Brazilian Function Point Users Group, Número de CFPS por País (2008)
15. SISP: Roteiro de Métricas de Software do SISP: V. 2.0. Ministério do Planejamento, Orçamento e Gestão: Secretaria de Logistica e Tecnologia da Informação, Brasília (2012)
16. Governo Federal: Disponível em (2008). http://www.governoeletronico.gov.br/anexos/instrucao-normativa-n-04
17. Usman, M., et al.: Effort estimation in Agile Software Development: A systematic literature review. In: Proceedings of the 10th International Conference on Predictive Models in Software Engineering, pp. 82–91. ACM (2014)
18. Cohn, M.: Agile Estimation and Planning. Addison-Wesley, Upper Saddle River (2005)
19. Schmietendorf, A., et al.: Effort estimation for agile software development projects. In: 5th Software Measurement European Forum (2008)
20. Fuqua, A.M.: Using function points in XP - considerations. In: Marchesi, M., Succi, G. (eds.) XP 2003. LNCS, vol. 2675, pp. 340–342. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-44870-5_46
21. Banerjee, A.U., et al.: Estimating agile iterations by extending function point analysis. In: WORLDCOMP 2012 (2012)
22. Alexander, A.J.: Case Study: Function Point Analysis and Cost Estimation in An Agile Development Environment (2011)
23. Cagley, T.: Agile Estimation Using Functional Metrics. The IFPUG Guide to IT and Software Measurement IFPUG. CRC Press (2009)
24. Wanderley, E.G.: Aplicação de Pontos por Função em Projetos que Usam Métodos. Dissertação de Mestrado UFPE (2015)
25. Kitchenham, B., et al.: Towards a framework for software measurement validation. IEEE Trans. Softw. Eng. **21**(12), 929–943 (1995)
26. Sousa, K., De, D., et al.: Uso do GQM para avaliar implantação de processo de manutenção de software. Universidade Católica de Brasília (2005)
27. Juristo, N., Moreno, A.M.: Basics of Software Engineering Experimentation. Springer, Heidelberg (2010). https://doi.org/10.1007/978-1-4757-3304-4
28. Travassos, G.H., et al.: Introdução à Engenharia de Software Experimental (2002)
29. Fisher, R.A.: Statistical Methods for Research Workers. Oliver & Boyd, Edinburgh (1925)
30. Tukey, J.W.: Exploratory Data Analysis. Addison-Wesley, Reading (1977)