When metrics mean business

Mauricio Aguiar

Abstract

Measurement specialists often wonder why so many Brazilian organizations use functional sizing as a basis for software development contracts. As a matter of fact organizations from several industries such as banks, airlines, telecommunication companies, and government agencies use function points as a basis for software development contracts. The Brazilian government has issued a directive known as "IN04" instructing all government agencies to use metrics – usually function points – in software development contracts to define productivities, prices, and quality criteria. In those organizations most software development services are priced and billed based on the functional size. Even support activities such as measurement and testing have been known to be priced and billed based on function points. Function point analysis is now an indispensable part of the business cycle and a daily activity in many Brazilian organizations. This paper discusses the business drivers that have made Brazil the number one function point user in the world as well as the business processes involved.

1. Introduction

Having been founded in 1986, The International Function Point Users Group (IFPUG) is probably the oldest software measurement association in the world. As of April 2012, Brazil was the number one country in IFPUG memberships, with 34% of all IFPUG members, followed by the U.S. (22%) and Italy (13%). IFPUG certifies individuals that pass the IFPUG exam. As of March 2012, Brazil had 42% of all IFPUG certified individuals, followed by Italy (20%) and the U.S. (13%) [1]. These numbers make Brazil the number one country in function point utilization. As a result several software measurement companies and independent consultants appeared in the Brazilian market in the last decade. Even though there is no publicly available data, a single Brazilian company claims to count 60,000 function points per month [2]. At the cost of US$ 1,000 per function point, that number could mean US$ 60,000,000 changing hands each month based on function point counts performed by just one Brazilian software measurement organization [3].

There are several possible explanations for the growth of software measurement in Brazil in the last several years. The following paragraphs intend to shed some light on this topic.

2. A Very Short History of Software Measurement in Brazil

The first Brazilian book on software measurement was Aguinaldo Aragon's "Gerência Efetiva de Software Através de Métricas" ("Effective Software Management through Metrics") published in 1995. Aragon's book included sections on function point analysis, COCOMO, and linear regression as effort estimation techniques. It also presented several applications of metrics to software management [4].

Function point analysis has been in use in Brazil since the eighties. However, it became more popular in the nineties, when UNISYS Eletronica became its main local sponsor. This was the first Brazilian function point movement, or the "First Wave". UNISYS Eletronica joined IFPUG in 1989 and started sending employees to IFPUG conferences in 1990. UNISYS Eletronica promoted function point user meetings in Brazil, called ENUPFs, from

1991 to 1994, featuring several international participants. UNISYS Eletronica sponsored a Certified Function Point Specialist (CFPS) exam in Brazil where a few of its employees became CFPS. Unfortunately, for internal reasons that company severely reduced its FPA sponsoring efforts around 1995 so there were no significant changes in the Brazilian measurement market until 1998. The first Brazilian function point analysis book was published in 1996 [5].

In 1998 a group of Brazilians became IFPUG members and founded the Brazilian Function Point Users Group (BFPUG) that had a significant role in promoting software measurement and function point analysis in Brazil. This was the second Brazilian function point movement, or the "Second Wave". BFPUG received active support from IFPUG starting 1999. As a result, a Brazilian was elected to the IFPUG Board of Directors in 2000. After a couple of years another Brazilian joined the IFPUG Board of Directors. A Brazilian – this author – was IFPUG President from 2005 to 2007. Also as a part of the Second Wave, the second Brazilian function point analysis book was published in 2003 [6].

The CFPS exam was brought to Brazil for the second time in 2001 and became a regular event, being held twice a year in three or four major Brazilian cities until its automation in 2008. A total of 1098 exam candidates sat for the exam in Brazil from 2001 to 2007. Many received the CFPS designation.

A possible reason why The Second Wave was more successful than the first is that the former came as a result of Brazilian government initiatives. Some parts of the Brazilian government were concerned with the cost of software development and were looking for ways of managing and possibly reducing that cost. Caixa Econômica Federal ("CAIXA"), a Brazilian government bank, was the leader of that movement. CAIXA decided to transition its software development projects to a "price per function point" model and launched a large function point based RFP in 1999. Other government agencies such as the Brazilian Post Office ("Correios") and the Brazilian IRS Data Processing Service ("SERPRO") soon followed. A government directive known as "IN04" was issued in 2008 (updated in 2010) and stated that IT services should not be billed based on effort (person-hours) or job positions (persons-month) [7, 8]. This confirmed function points as the measure of choice for Brazilian government software development contracts. Those facts got the attention of private organizations such as banks, telecom companies, airlines, and others that would also adopt variations of the method.

The success of function point analysis in Brazil led IFPUG to hold its 2010 annual conference in Sao Paulo, Brazil's largest city. ISMA Cinco (ISMA is the International Software Measurement & Analysis Conference. "Cinco" is Portuguese for "Five") was successful both in terms of international participation as well as local attendance. The conference featured international speakers such as Ricardo Valerdi (MIT), Carol Dekkers (Quality Plus Technologies), Joe Schofield (IFPUG Vice-President), Donald Beckett (Quantitative Software Management), Pam Morris (Total Metrics), Luigi Buglione (Engineering IT), Thomas Cagley (IFPUG Past President), Terry Vogt (Booz-Allen-Hamilton), Daniel Galorath (Galorath Corporation), and many others.

BFPUG has held an annual software measurement and analysis conference in Brazil since 2009, featuring international speakers from the software & systems measurement arena.

In November 2010 the Brazilian government, through its Ministry of Planning, published guidelines to the application of function point analysis to software development contracts [15].

In March 2012 Brazil had 42%, of all CFPS, making Brazil the number one country in number of IFPUG-certified specialists.

As more Brazilian organizations adopt the "price per function point" method, it is likely that more measurement-related methods and techniques will become popular. For instance, the COCOMO II estimation model and the Practical Software & Systems Measurement framework are already used in Brazil; The Netherlands Software Metrics Users Association (NESMA) methods for early size estimation and enhancement counts are also used; the COSMIC measurement method is increasing its popularity, as well as the recently published IFPUG Software Non-functional Assessment Method (SNAP).

3. Why Brazilian Organizations Use Functional Sizing in Software Development Contracts

A typical software development project includes a request for proposal issued by the acquirer. Following that request, one or more suppliers present their proposals. These are usually based on an estimated amount of effort – a number of person-hours or person-months that constitutes the main input to the pricing process. Because the estimated amount of effort is highly dependent on the supplier's resources and development process bids may not be easily evaluated by the acquirer. This pricing mechanism could be called "process-oriented pricing", where the price is based on the inputs – the resources needed to complete the project. An alternate pricing mechanism could be designated "result-oriented pricing" where the price would be determined by the output, i.e., the amount of software delivered.

Process-oriented pricing, the more traditional way of pricing services, puts control in the hands of the supplier. This happens because the acquirer does not have the expertise or information to assess the supplier's estimate. A simple analogy may illustrate this situation. Suppose you are on the road and your car breaks down. You manage to find a car repair shop nearby, where they agree to take a look at your car and tell you to come back after a couple of hours. When you come back the chief mechanic hands you a budget with several parts to be replaced and an estimated number of work hours. If you are not at least an amateur mechanic you will not be able to make a rational decision. In this situation the price is completely controlled by the supplier. Car repair shops will probably continue to control prices in their market, but there is a better way in the software development arena. Far from being new, these better methods have been in use for more than ten years. The government of the State of Victoria, Australia, is the creator of southernSCOPE and refers to it as the "$ per function point" method. According to the southernSCOPE Reference Manual, the method "results in the customer paying the software developer an agreed price for each 'unit' of functionality in the delivered software" [9]. The Finnish Software Metrics Association (FiSMA) built on the southernSCOPE concept and launched the northernSCOPE method, allegedly an improved version of southernSCOPE [10]. Other countries where the government has recommended or required the use of function points are Italy and Korea.

Because that kind of pricing mechanism is based on the outputs rather than the inputs, the "$ per function point" or "price per function point" method is a result-oriented pricing mechanism. In this type of method the price is controlled by the client. That is probably the main reason why Brazilian organizations have been consistently transitioning to that software acquisition method in the last several years.

It is important to note that the "price per function point method" could also be called the "price per functional size unit" method, since any functional size unit may be used in the method, not necessarily function points.

Suppliers initially tend to oppose the change to the new method, possibly for fear their profits will decrease. That may actually happen in the beginning, but as they learn to work

with the method they realize their risks tend to be much lower than with fixed-price contracts, so they tend to become supporters of the "price per function point' way of doing business.

Functional size measures can serve as the basis for a result-oriented pricing method because they:

- are result-oriented by definition, as they measure the output (software) and not the input (work hours)
- can be understood and verified by both the acquirer and the supplier, bringing transparency to the business relationship
- can be standardized (IFPUG, MkII, COSMIC, and NESMA function points have become ISO standards [11])
- can be benchmarked (for example, see the ISBSG [12])
- can be used to manage the project scope [9, 10]

All those reasons have led Brazilian organizations to use functional size measures in software development contracts for new developments as well as enhancements. Acquirers aim to get control of the pricing process and ultimately pay lower average prices. They also want to be able to manage project scope and schedule, make sure budgets are not exceeded, and pay only for software actually required and delivered. Suppliers want to make sure every change request is adequately measured so that they get correctly paid. They want to guarantee that if the system size increases so does their revenue. If wisely used, functional size measures will support all those contract objectives.

4. How Brazilian Organizations Use Functional Sizing in Software Development Contracts

Brazilian organizations use functional sizing in software development contracts according to several business models. Before getting to those particular models, it will be useful to address the main types of models used.

4.1. Types of Models

A few types of models are used with functional sizing. Each has specific characteristics and serves a different purpose, so it is very important to distinguish them. Here we will consider explanatory, predictive, and prescriptive models.

Explanatory models seek to understand phenomena, discover associations, identify causes and effects, and acquire knowledge. In the social sciences they are often used to understand behaviour, as well as to support and guide decisions. These models need information not available early in the cycle and are typically used after all the facts are known. One general example would be an explanatory model to find the causes of a problem in order to prevent it from happening again. An example from the software measurement context would be to model productivity as a function of several variables in order to guide productivity improvement initiatives.

Predictive models seek to predict the future behaviour of a variable, or of a set of variables. They are usually based on the past behaviour either of the target variable, or of associated variables. These models are used by estimators. Predictive models ideally leverage existing data as well as estimator expertise. Several estimation models are known to the software measurement community, some being implemented by commercial software tools. One example would be to model effort as a function of size and past productivity in order to estimate the effort of a project.

Prescriptive models are used to regulate relationships. A prescriptive model can be seen as a system of business rules designed to implement a business agreement. A predictive model

may be used to calibrate a prescriptive model. Once the prescriptive model is calibrated it should be self-contained. Prescriptive models are not to be confused with predictive models. An example would be a model to set prices per function point for different project types according to a set of rules.

The main use of function points in Brazil is in prescriptive models. Those are not estimation models.

4.2. Function Point Based Business Models

A simple definition of a business model is "a way of doing business". The most common function point-based business models used in the Brazilian community are estimation models and pricing models.

Estimation models are predictive models used to calibrate a prescriptive model. Before using the "price per function point method" an organization will typically perform a study to determine the ideal price per function point to be paid the supplier. That type of study includes the definition of project types, statistical analysis of historical and benchmark data, specific data collection, as well as the elicitation of business objectives. The outputs of such a study will be project types, productivity ranges, prices per function point, and/or possibly other measures of interest. Those numbers will be used in a prescriptive model that will ultimately be part of a contract. It should be noted that a contract for the development of a new software system will have different estimation requirements than an umbrella contract for an unknown number of enhancements to be performed over a period of years.

The most basic type of estimation model defines project effort as the product of size (in function points) and productivity (in hours per function point). Productivity is sometimes called delivery rate and usually varies with project type, being determined in a specific study as described above. Some more sophisticated estimation models – parametric models – define project effort as a function of size and several other factors. All estimation models use statistical techniques in their calculations. Some models like COCOMO II use more specialized techniques such as Delphi and Bayesian Statistics. Several models are implemented by tools such as SEER, SLiM, COSTAR, PRICE-S, KnowledgePLAN, Capers Jones's SRM, and others.

In addition to contracts, estimation models are used in budgets, RFPs, make or buy studies, etc.

Pricing models are prescriptive models that establish the pricing and billing rules to be used in a software development contract. Those models can become quite complicated in contracts that deal with several types of service. While there are several model variations, only a few basic pricing models are used in Brazil, typically the Productivity-based Model, the Price-per-FP Model, and the Baseline-based Model.

The Productivity-based Model is based on the simple estimation model that defines effort as the product of size (FP) and productivity (Hours/FP). There will be a distinct productivity for each project type. Each new development or enhancement is sized in function points. The FP size is multiplied by the corresponding productivity giving the number of effort hours. The number of hours is multiplied by the contract price per hour giving the amount to be paid the developer.

The Price-per-FP Model simply assigns a different price per function point for each project type. Each new development or enhancement is sized in function points. The FP size is multiplied by the corresponding price per FP giving the amount to be paid the developer.

The Baseline-based Model assigns a price per function point to an installed application baseline. A fixed monthly fee is charged for a service package, e.g. application maintenance and support. The price per function point is periodically updated as the application baseline

grows. A service level agreement (SLA) handles all the details, including schedule, quality, incentives, penalties, etc.

Several other models are in use in Brazil, of which the following deserve to be mentioned:

Defect-based Model – This is a quality control model. A penalty is associated with a defect threshold, typically based on a defect density measure (defects per function point).

Negotiation-based Model – Even though this is a prescriptive model, it is negotiation-intensive. The supplier receives a request for proposal and is required to respond with the functional size, the number of hours, and the price estimated to complete the project. The acquirer receives the proposal and enters the FP size into an estimation model not necessarily known by the supplier. If the number of hours or price returned by the estimation model is greater than the number provided by the supplier, the acquirer accepts the proposal. Otherwise the acquirer rejects the proposal and starts a negotiation process to bring the number of hours or price to the level indicated by the estimation model. Estimation models used in this type of business model are not always objective. Some Brazilian acquirers use a customized COCOMO-like model in this manner.

Phase-based Model – Some acquirers outsource different parts (phases) of a project to different suppliers. For example, one supplier may develop the software product and other supplier may test it. In this business model each project phase is assigned a percentage of the total lifecycle, and the price per FP is divided accordingly. Each supplier is paid according to the defined project phase percentages and the project phases they are assigned to. Phase percentages are typically defined using historical data.

5. Technical Considerations Related to the Use of Functional Sizing in Software Development Contracts

Even though there are many benefits associated with the use of functional sizing in software development contracts there are still many issues to be resolved. Those have been addressed by Brazilian acquirers and suppliers both from the government and private sectors.

5.1. Predictive versus Prescriptive Models

One issue has to do with the difference between predictive and prescriptive models. Because most of the software development market still uses conventional effort-based pricing specialists tend to focus on predictive models to estimate effort. However, in Brazil the most important use of functional sizing is in pricing and billing. Even though there are similarities between predictive and prescriptive models, there are also differences that must be considered before using an estimation technique as a basis for a pricing model. A few of them are listed on Table 1 below.

Table 1: Estimation vs. Pricing

Estimation (Predictive Model)	Pricing (Prescriptive Model)
1. Estimated value should be close to actual	1. Prescribed value should be close to actual
2. Method is expected to give approximate results	2. Method is expected to give exact results
3. Different estimators may produce different values (depending on their expertise and skill)	3. Different model operators must produce the same values
4. Input values do not need to be objective – may depend on estimator's opinion/assessment	4. Input values must be objective – must not depend on estimator's opinion/assessment

Both models share the goal of producing values that should be as close to the actual values as possible. Even though neither will be able to match the actual values exactly, both are expected to approximate them in the long run.

Estimating models are expected to give approximate results. A slight change in the inputs may not be reflected in the outputs. On the other hand, pricing models are expected to give exact results. A small change in the inputs should create a (hopefully small) change in the outputs.

Different estimators are expected to produce different results. A more skilled and more experienced estimator is expected to produce a better estimate than a novice, whereas an operator of a pricing model is expected to follow exact rules and produce exactly the same result as any other trained operator.

Estimating models often have subjective parameters. For example, the COCOMO II estimating model has parameters such as ACAP (Analyst Capability) and PCAP (Programmer Capability) with ratings 15%, 35%, 55%, 75%, and 90% [13]. An analyst or programmer team that falls in the 15% level is rated very low – at the estimator's discretion. Two distinct estimators could potentially disagree on those levels. Estimation models often have to be customized before they can be used for pricing purposes. Pricing models, on the other hand, have no room for ambiguity or subjectivity.

5.2. Items Not Covered by Functional Sizing

By definition, functional sizing does not take non-functional items into account. This leaves room for unaccounted effort variation in projects where effort is predicted or prescribed using functional measures only. A simple solution is to define several different project types and assign them different productivities, hoping they will account for all non-functional effects. However, there will always be variation even in a well specified and calibrated model. Another solution would be to use a parametric model such as COCOMO II, where all variation due to factors other than size would hopefully be accounted for by model parameters (there are 22 of them in COCOMO II). This works well for estimation, but suppliers and acquirers are not happy when variation is not accounted for. In Brazil some suppliers built tables that transform certain non-functional characteristics into an equivalent number of function points. That type of solution was used in government bids [14]. Other ways of dealing with so-called "non-measurable items" for several types of activities including documentation and testing are defined in [15]. Many of those methods add equivalent function points to the functional size, or multiply the functional size by a specified factor.

Even though adding "equivalent function points" to the functional size violates several measurement principles, the Brazilian industry has been using the workarounds above for lack of a better solution. The industry will typically adopt an imperfect solution as long as it is (or seems) better than the alternatives. The software measurement community still has a lot of work to do on non-functional assessment models such as IFPUG SNAP to fix this situation.

5.3. Sizing Enhancements

A significant part of the Brazilian software measurement community believes that the IFPUG method for sizing enhancement projects is not suitable for pricing models. They prefer to use the NESMA enhancement sizing method instead. That method assigns different weights for added, deleted, and changed function points [16]. According to Brazilian users the NESMA method provides more accurate results than the IFPUG method.

6. Benefits and Challenges of Using Functional Sizing in Software Development Contracts

The "price per function point" method potentially leads to better productivity and represents an improvement over previous effort-based methods. It brings transparency and objectivity to the negotiation process, being good for any application domain, development process, and technology.

Special care must be taken when determining the initial productivities in order to establish a balanced relationship between acquirer and supplier. While a good pricing model will reduce variation to an acceptable level, it is important to note that bad requirements do not favour accurate sizing. Poor requirements will increase the uncertainty in the sizing process. Most organizations will want to improve requirements before transitioning to the "price per function point" method. Any functional sizing method may be used, but it is highly recommended that measurement be performed or at least supervised by certified professionals. This will reduce differences in the interpretation of counting rules, especially between acquirer and supplier. Sizing may also be outsourced to a neutral third party organization in order to improve transparency and minimize conflict. Non-functional items will continue to be a challenge until a non-functional measurement solution is found and accepted by the measurement community. Most of all, one should keep a win-win attitude and be aware that when using functional sizing in pricing models there will be gains and losses, but at the end of the day things will balance and everybody will win.

7. Conclusion

This paper has presented a short description of the utilization of software measurement in Brazil. After providing a historical perspective, several relevant topics were addressed, such as the difference between predictive and prescriptive models, and why the latter is so important in Brazil; why and how Brazilian organizations in the government and private sectors use functional sizing for estimating and pricing; the main technical difficulties that have been encountered and how they have been addressed.

It is hoped that this report will contribute to raise interest in the use of functional sizing in pricing models, since so much work is still needed to improve those models.

8. References

[1] IFPUG, unpublished internal reports, IFPUG, Princeton Junction, 2012.

[2] Aguiar, Mauricio, "Domain-Oriented Modeling, Estimation and Improvement for Commercial and IT Systems Domains" (presentation), 26th International Forum on Systems, Software, and COCOMO Cost-Modeling", USC, Los Angeles, 2011, http://csse.usc.edu/csse/event/2011/COCOMO/presentations/Panel_Aguiar_2011-11-03.ppt

[3] Jones, Capers, "Applied Software Measurement – Global Analysis of Productivity and Quality", 3rd Edition, McGraw-Hill, New York, 2008, pg. 310.

[4] Fernandes, Aguinaldo Aragon, "Gerência Efetiva de Software Através de Métricas", Atlas, São Paulo, 1995.

[5] Braga, Antônio, "Análise de Pontos de Função", Infobook, Rio de Janeiro, 1996.

[6] Vazquez, Carlos E., Simões, Guilherme S. e Albert, Renato M., "Análise de Pontos de Função: Medição, Estimativas e Gerenciamento de Projetos de Software", Érica, São Paulo, 2003.

[7] Santos, Rogério S., "Instrução Normativa Nº 4", SLTI, Brasília, 2008.

[8] Santos, Maria da Glória G., "Instrução Normativa Nº 04", SLTI, Brasília, 2010, http://www.governoeletronico.gov.br/biblioteca/arquivos/instrucao-normativa-no-04-de-12-de-novembro-de-2010/download
[9] Government of Victoria, Australia, "southernSCOPE Reference Manual, Version 1",: Government of Victoria, Victoria, Australia, 2000, http://www.egov.vic.gov.au/pdfs/SSRefManual.PDF
[10] FiSMA, "northernSCOPE: customer-driven scope control for ICT projects" (brochure), FiSMA, http://www.fisma.fi/wp-content/uploads/2008/09/northernscope-brochure-v152.pdf
[11] The International Organization for Standardization (ISO), website, http://www.iso.org
[12] The International Software Benchmarking Standards Group, website, http://www.isbsg.org
[13] Boehm, Barry et al., "Software Cost Estimation with COCOMO II", Prentice-Hall, New York, 2000.
[14] Caixa Econômica Federal, "Concorrência Nº 001/2006, Anexo I: Projeto Básico – Item VIII", Caixa Econômica Federal, Brasília, 2006.
[15] Ramos, Carlos R. S. et al., "Roteiro de Métricas de Software do SISP – Versão 1.0", Ministério do Planejamento, Orçamento e Gestão, Brasília, 2010.
[16] Netherlands Software Metrics Users Association, "Function Point Analysis for Software Enhancement – Guidelines, Version 1.0", NESMA, 2001.