A Software Tools Catalogue to Support the Statistical Process Control on the Software Context

Aluízio Ramos Pereira Neto and Sandro Ronaldo Bezerra Oliveira

Graduate Program in Computer Science, Institute of Exact and Natural Sciences, Federal University of Pará, Belém, Pará, Brazil

Keywords: Statistical Process Control, Software Tools, Empirical Software Engineering.

Abstract: Statistical Process Control (SPC) is applied to the software context in process analysis and improvement in high level maturity organizations. There are some studies that talk about the SPC in the context of software, however, these do not yet describe in depth the approaches related to it. The main goal of this study is to present the results of a Systematic Review of Literature, aiming to identify the SPC-related approaches (in this work, approaches are understood as techniques, frameworks, methods and tools to support the implementation of a process), which were put together in the form of a catalog. In this study, only the tools that implement the SPC will be presented, describing its characteristics, example of use, availability and ownership. With this study researchers will obtain valuable information for the possible future application of these tools in their development contexts.

1 INTRODUCTION

Software Engineering is a research area that has several study subareas. Among the most complex are Statistical Process Control (SPC), which depends on information generated by the execution of the Measurement process, which demands operational effort and cost for the collection, analysis and decision making. The SPC is executed in organizations with High Maturity (which implement levels A and B of MR-MPS-SW – Reference Model of Brazilian Software Process Improvement for Software, and 4 and 5 of CMMI-DEV – Capability Maturity Model Integration for Development), whose focus is on the ability of the process to be measured and optimized, which depends on a implementation of more institutionalized management, engineering and support processes among the different sectors (Rocha et al., 2012).

The SPC came into the industry with the use of descriptive statistics and began to be applied in the software context in the mid 1980's (Santos and Silva, 2013). In general, it investigates the capability and performance of a software process, adding the measurement process, descriptive statistics of data and control charts to analyze the aspects and assets related to a software process, to create mechanisms of improvement and quality of software process, establishing performance and capacity models (Rocha et al., 2012).

Many studies, such as (Santos and Silva, 2013; Zhang and Hou, 2010; Pettersen, 2011), on Statistical Process Control in the software engineering area, focus on the research methodology used, with few or no examples of implementation of SPC-related approaches identified.

The correct implementation of the SPC in an organization requires the knowledge and application of its related approaches (techniques, methods, frameworks, models and software tools). The lack of a knowledge repository about these approaches is a strong motivation for this work. Greater understanding of these approaches can enhance their performance and competitiveness. In addition, the use of SPC in the analysis of software processes in projects with high levels of maturity fosters the competitive advantages of the market (Rocha et al., 2012).

Thus, a study was proposed based on a Systematic Literature Review (SLR) with the objective of identifying SPC-related approaches (techniques, methods, frameworks, models and software tools), whose methodology and previous results were published in (Neto et al., 2017) and gathered in a catalogue of approaches.

In this paper, a summary of the results obtained in the review that gave rise to the catalogue of
approaches and the cataloging of the identified software tools will be presented, describing their characteristics, example of use, availability, ownership and tendency of adoption of the software tool. In addition, the work described in this document provides information about the catalogue evaluation process.

The SLR was chosen as a research methodology for the following reasons: it classifies primary studies and identifies approaches, has a specific focus (in this work to identify approaches to SPC) and provides a detailed analysis regarding the description and forms of use of the approaches (Travassos e Biolchini, 2007); is a non-biased, comprehensive, fair research that uses a reliable and rigorous methodology (Kitchenham et al., 2015); about the purpose, it is conducted according to a planning described in a review protocol, previously evaluated by an expert (Mafra and Travassos, 2006).

In addition to this introductory section, this paper has the following sections: Section 2 describes the theoretical foundation on SPC. Section 3 describes related works. Section 4 presents a brief summary of the SLR performed. Section 5 describes the catalogue of SPC-related software tools. Section 6 describes the catalogue evaluation process. Section 7 briefly describes the application of the catalogue and, finally, Section 8 presents the final considerations of this work.

2 THE STATISTICAL PROCESS CONTROL (SPC)

The SPC is able to identify variations in the behavior of the processes and to allow the analysis of their stability and the establishment of their capacity. For this, it uses graphs that associate methods of statistical control and graphical representation (control graphs) in order to detect signs of variation (unacceptable, special causes that need to be analyzed) in the behavior of the processes and differentiate them from the noises (acceptable, are the common causes) (Barcellos, 2009).

The stability of the process indicates that it is described within the established control limits. However, being a stable process does not mean being a good performing process. To verify that a stable process performs well, it is necessary to check its capacity. A process is considered capable when it: meets or exceeds organizational and customer expectations; and meets the strategic objectives of the organization and business. The performance of the process is indicated by the performance baseline (process voice) that the process has compared to the performance expected by the process (customer voice) (Rocha et al., 2012).

In order to identify and analyze the capacity of a process it is necessary to use the frequency histogram. It represents the values collected of process during its stability period. In the histogram, the control limits of the process baseline (process voice) and range specification limits (customer voice) are represented (Rocha et al., 2012). The performance baseline describes the selection of data collected in processes of several projects with the same characteristics and represented in a control chart. These graphs present the limits of performance control, which are the expected values for the process (Wheeler and Chambers, 2010).

The first step in using control charts is to select the type of measure to be analyzed. The data are collected by the measurement process and plotted on the charts where the control limits are defined, based on the organization's strategic objectives (business goals) (Rocha et al., 2012). The control charts have an Upper Control Limit (UCL) and Lower (LCL), which are at a distance of three standard deviations (σ) from the central line (Control Limit - CL). The centerline and boundaries can not be arbitrary, since they reflect the current behavior of the process. Their values are obtained by applying the expressions and constants defined by the type of control chart applied.

The central tendency of a control chart is the central indication of the measures. The key measures are the process inputs that have been “chosen” for analysis in the control charts. This analysis takes into account the process control charts, guided by the strategic objectives, which point out the upper and lower limits of an acceptable process. These objectives are formed by the customer's voice (market view) and the voice of the process (organization's process capability). They must be established and ensured by the organization in order for the process to achieve the expected stability (Rocha et al., 2012).

Control charts are the main tools used to analyze the data collected in the processes. Each chart type has its characteristics and specifications related to one or more contexts of use. This choice takes into account its purpose, context and type of data distribution (Barcellos, 2009).

3 RELATED WORKS

In order to identify studies on Systematic Review in the Statistical Process Control context, a literature
search about the work related to the topic was carried out. An ad hoc search was performed on the Google Scholar search engine, due to its favorable index resulting from other searches and to enable the identification of publications originating in the IEEE and ACM repositories. Some terms used in the search were: "Systematic Literature Review", "Statistical Process Control", "Techniques", "Approaches" and "Software Tools".

In the search it was identified that the book (Kitchenham et al., 2015) describes the methodology to perform an SLR in the software context. The works (Pettersen, 2011; Baldassarre et al., 2007) use the SLR as a research methodology and additionally inform some considerations about Statistical Process Control, presenting its theory in general. These three studies focus more on the SLR methodology than on its results.

The study (Garcés and Pino, 2014) is the closest to the objective of this paper. This paper presents a systematic review on Statistical Process Control with the objective of identifying the main approaches related to the SPC.

The study proposal reported in this paper has the differentials: (i) it is based on the accomplishment of an SLR, formalized in a protocol previously evaluated by an expert, (ii) execution of the SLR methodology, with planning (search strings, source, exclusion and quality criteria); execution (collection, selection of primary and secondary studies, quality evaluation), extraction and presentation of results; presents the results of SLR in a descriptive and bibliographical way (by means of graphics); the recommendations proposed in the catalogue detail present assets that were proposed or applied only in the context of software projects; describes in detail the approaches related to the SPC and presents examples of their use.

4 THE SYSTEMATIC LITERATURE REVIEW

The main objective of a systematic review is to evaluate and interpret the available research data. These are related to research questions from a thematic area or from a phenomenon of interest. SLR is a non-biased, comprehensive, fair research that uses a reliable and rigorous methodology (Kitchenham et al., 2015).

From the systematic literature review we can cite other characteristics: it is formulated from systematic and explicit research methods; evaluates and selects the most relevant searches for the research objective; is a transparent search (allows selection and quality criteria), comprehensive (selects the most relevant studies the research question) and non-biased (has a review protocol, free of charge and financial interest); and replicable (Mafra and Travassos, 2006).

The methodology of systematic review applied to carry out this study obeyed the norms defined by Kitchenham et al. (2015). Initially, revision planning was carried out, where search sources, search strings, quality criteria and other information were defined. In the next phase, the data were collected in the defined research sources and the selection and quality criteria were applied. Next, the selection of the primary studies was carried out and, with the selected works, the data extraction was performed.

This systematic review aimed to identify the approaches that support Statistical Process Control (SPC), in the context of software projects, regarding processes, methods, models, frameworks, methodologies, techniques, software tools and other ones. Based on the research objective of this review the following research question was defined, which guided this Systematic Literature Review:

- (QS1): What standards exist to support Statistical Process Control activities?
- (QS2): Are there software tools to support the Statistical Process Control?
- (QS3): If there are support software tools, what is the license to use?

A bibliometric analysis of the data generated by the systematic review was performed. We defined the scope (work in English, available by the UFPA domain, free search engine and others), constraints (duplicate studies, works that do not present key words and others) and research resources (4 researchers involved).

Manual search was performed on events and journals, and automatic search from the search string. The sources selected for the research were: ACM, Elsevier, IEEE Xplore Digital Library and Scopus. For each search source a search string was defined. These data can be seen in (Neto et al., 2017).

Next, the selection of the primary studies was carried out, which consists in evaluating the studies by means of Inclusion criteria (available for download, duplicate articles, repeated and others) and Exclusion of Primary Studies (SPC-related works). More detailed information about the planning and results obtained with the systematic literature review can be obtained from the work already published (Neto et al., 2017).
5 THE SOFTWARE TOOLS CATALOGUE

The Catalogue of Approaches related to Statistical Process Control (whose access is available at the following link: https://drive.google.com/file/d/0B_PDf6-qXCFcMHRveXJWbUdCNkk/view?usp=sharing) was developed based on the results extracted of the systematic literature review, described in Section 4, whose main objective was to identify the SPC-related approaches.

The catalogue was organized in two parts: the first presents techniques, methods, frameworks and models related to the SPC; and the second, which focuses the objective of this paper, presents the SPC support software tools cataloged in this study.

Each software tool presents: its type, which can be 1 - Statistics, which refer to techniques derived from descriptive statistics, 2 - Quality, which are techniques from software evaluation processes, and 3 - Software Measurement, measure and evaluate the software processes of an organization; an example of using the software tool described in the literature; and information on their availability.

5.1 Minitab

The Minitab proprietary software tool, cited in the work (Zhang and Hou, 2010; Rahman et al., 2008), is a computer program applied in statistical studies, developed in 1972. Its interface is similar to Microsoft Excel or Calc of OpenOffice, used in universities and in companies, has specific functions focused on process management and analysis of the Six Sigma suite. Minitab offers Quality Control tools, Experiment Planning (DOE), Reliability Analysis and General Statistics.

According to its official website (http://www.minitab.com/pr-BR/), Minitab has the following control charts available: X-bar-R chart, Xbar-S chart, I-MR-R / S, Zone Chart, P Cards, EWMA Chart, and more.

The Minitab, characterized as a type of 2 - Quality approach, was used in an experiment of an organization A, under the software context, presented in the work (Rahman et al., 2008). The organization produces high quality components for automotive engines. The company used several basic quality tools, such as: Pareto Diagram, Control Charts, Check Sheets, Ishikawa diagrams, among others, to analyze its collected data. According to the organization's managers, the team involved had a good understanding of Statistical Process Control, but it took a long time to implement this knowledge in the company. The company wanted to increase its knowledge and skills regarding the applicability of the SPC in its processes. For this, they participated in an experiment using Minitab version 13.2, where they used the tool and reported the use, advantages and disadvantages of the tool. The SPC was fully implemented in the organization and the use of the Minitab tool generated the following advantages: the tool enables a complete and accessible use of the SPC and provides the company with a quality process; the SPC with support from Minitab is able to detect abnormalities, critical parameters, variations and increase the stability of the process. Main drawback: Minitab version 13.2 did not bring the ease needed to build bar or pie control charts. This was overcome by using version 14.0 of the tool.

5.2 Clearquest

According to its official website (http://www-03.ibm.com/software/products/pt/clearquest), ClearQuest is a proprietary tool for error monitoring and change control that creates multiple forms with VB or PERL functions, which allow the development team to manage changes and errors in the project. The tool optimizes the software life cycle with an application development system for workflow management.

IBM Rational ClearQuest is a fully customizable system for the production and development of a database workflow application. It provides flexible tracking of changes and defects, customizable processes, near-real-time reporting, life cycle traceability to improve visibility and control of the software development lifecycle. IBM Rational ClearQuest provides scalable and cross-platform support for businesses of all sizes, allowing it to continue customizing processes as development needs evolve over time. The tool can be considered as 1 - Statistical approach; 2 - Quality; 3 - Measurement.

The work (Corrales et al., 2013) cites an example of using the ClearQuest tool, which investigates the validation unit of information technology of a financial organization (development and support). ClearQuest was used to archive the data used (referring to identified defects) in the construction of control charts. The organization used the measurement and data analysis (MA) process, implementing defect management and establishing quality and performance limits. In addition, it used process control and forecasting statistics.
5.3 7QC (PME)

The 7QC refers to the 7 major quality tools (Control Diagram, Flowchart, Check Sheet, Pareto Diagram, Fishbone Diagram, Histogram and Dispersion Diagram) described in (Rahman et al., 2008), which combined to SME, which are a set of small businesses that provide business management services to larger organizations, assists in implementing Statistical Process Control in organizations, reducing cost, increasing flexibility, and generating profit.

The work (Pettersen, 2011) does not detail how to use this proprietary tool. However, the authors apply the 7 quality tools and the business management process, aiming to implement statistical process control in organizations. The tool can be considered as 1 - Statistical approach.

5.4 Spc XL2000

The study (Rahman et al., 2008) presents the description of the SPC XL2000 proprietary tool, which implements techniques related to Statistical Process Control. This study analyzes the use of the SPC XL 2000 in the management of the production of disposable medical assets, focusing on the software development that automates this type of business. This software is used in the organization to collect, analyze and report on the processes of the organization. The tool approach is 1 - Statistical.

The work (Rahman et al., 2008) describes an experiment using the SPC XL2000 tool in an organization in data analysis and monitoring of key processes. The authors do not describe in detail the use of the tool, however it provides its main conclusions, where the SPC was implemented efficiently in the organization, establishing the capacity and stability of the processes, even when working with large amounts of data. The tool's website is http://www.winspc.com/.

5.5 SME-SPC

The construction of the SME-SPC proprietary tool is described in the study (Zain et al., 2009), which was proposed with the purpose of helping small and medium organizations to implement the SPC. Its development was based on a previous study on the SPC and the data collection of organizations that already used it. The second phase used the collaboration of an automotive company, which carried out the test phase of the tool, which contributed several suggestions for improvement. The tool approach is 1 - Statistical.

The work (Zain et al., 2009) describes the structure of the SME-SPC system and its main functionalities, such as: user system, configuration of data collection and analysis, data entry and subgroup characteristics, data status, specification of control chart data, chart types, chart plotting, comment and feedback space, frequency histogram definition and performance curves, summary and change of data. For access to tool feature screens access the full catalogue of approaches at: https://drive.google.com/file/d/0B_PDf6-qXCFcMHRvXJWbUdCNkk/view?usp=sharing.

5.6 WebAPSEE

WebAPSEE is a Process-Centered Software Engineering Environment (PSEE) developed as free software by LABES-UFPA. The main purpose of the environment is to meet organizational requirements to assist in the coordination of activities related to software development (Gonçalves et al., 2012). Its main functionalities are: visual modeling of processes and flexible execution of processes. WebAPSEE is strongly based on the measurement process and incorporates functionalities for SPC deployment such as: planning, identifying critical processes, plotting control charts, establishing baselines and identifying improvements. The tool's website is http://www.labes.ufpa.br. The tool presents the approaches of 1 - Statistics, 2 - Quality and 3 - Measurement.

The work (Gonçalves et al., 2012) describes an example of use of the WebAPSEE tool, where it analyzes the process of "productivity in the requirements activity ". Initially, the data for the control charts were selected and plotted on an XmR chart. The incidence of a variation point was investigated and subsequently solved. The tool has fields, where it can describe causes, point of deviations, problem status and action plans to be executed. To get access to the tool's feature screens, just go to the complete catalogue at: https://drive.google.com/file/d/0B_PDf6-qXCFcMHRvXJWbUdCNkk/view?usp=sharing.

5.7 MSChart

The work (Zhang and Hou, 2010) presents a study of the construction of the MSChart proprietary tool. This tool was proposed to meet the need of software development companies in the treatment of cases of statistical tolerance (ST), Statistical Process Control (SPC) and in the design of control charts. The tool's website is http://www.microsoft.com and is classified
as 1 - Statistical, 2 - Quality and 3 - Measurement approaches.

MSChart consists of modules: presentation (introduction), login, data collection, storage, analysis and management, ST and SPC parameters, design of control charts; data acquisition, data processing, control chart design (choice of control charts) and specification of statistical tolerance (definition of control limits, quality indexes, statistical tolerance and design of control charts). The work (Zhang and Hou, 2010) does not describe an example of tool use.

5.8 PAS System (Process Analysis System)

The Process Analysis System (PAS) was developed to meet the needs of small and medium enterprises of the software manufacturing industry in the use of quality control and the SPC. The PAS system encourages the use of SPC, aims to reduce costs, provides several types of control charts and accepts most types of data (Chang and Lee, 2013). Its main functionalities are: web service, user types, doubts page, scenario choice (control chart type), type of data storage and system procedures (history and data configuration). It is considered as 1 - Statistical, 2 - Quality and 3 - Measurement approaches.

The work (Chang and Lee, 2013) describes an example of using the PAS system, where the user initially defines the characteristics of the process, data type and control chart to be analyzed. It then configures the data to be inserted and can be viewed and filtered into tables. The user must define the number of data of each subgroup and the way the system will calculate the generated data. Subsequently, the charts are generated for analysis. Data can be entered and changed at any time. If there are more general doubts, the user can access the "doubts" field, or in cases of doubts related to control charts, the user must access "Control Guide". For access to tool feature screens access the full catalogue of approaches at: https://drive.google.com/file/d/0B_PDf6-qXCFcMHRvcXJWbUdCNkk/view?usp=sharing.

6 THE EVALUATION OF SOFTWARE TOOLS CATALOGUE

The catalogue of approaches was evaluated by means of a peer review method. In this method one or more experts of the researched area evaluate the study, observing its relevance, correctness and contributing with considerations for its improvement. Its choice is justified by the need for an expert, who understands the information provided in the catalogue, who has experience in the area in question, can contribute to the refinement of the catalogue and in the future put it under evaluation in the industry.

This catalogue was evaluated by an expert in the Software Engineering, Master in Computer Science at Software Process Improvement (SPI), MR-MPS-SW consultant and evaluator and with experience in implementation, consulting and evaluation in Software Process Improvement of more than 10 years, according to the models of CMMI and MPS.BR – Brazilian Software Process Improvement, and with more than 5 years experience with the use of Statistical Process Control.

For the evaluation of the catalogue an evaluation questionnaire was created, composed of 16 objective questions, divided into 2 groups: the first one concerns the Profile of the Reviewer of the catalogue, in which the questions aim to discover the level of knowledge of the reviewer regarding of the methodology, process improvement of the implementation the SPC, implementation of models for process improvement, methods of evaluation in the models and time of experience in evaluation of SPC processes.

The second group deals with the Proposal Presentation, whose purpose is to verify the evaluator's understanding of the work under evaluation, having as a matter of fact the degree of correctness and completeness of the catalogue and if it can be used as a reference in aiding the implementation of the SPC.

As an annex to the questionnaire, a subjective evaluation was requested to review the submitted material, based on (Neto et al., 2017), in which it was allowed to record comments by a table filled out by the evaluator, containing the identification of the comment, its category (HT - High Technician, indicating that a problem has been found in an item that, if not changed, will compromise the considerations; LT - Low Technician, indicating that a problem has been found in an item that it would be appropriate to change; Q - Questioning, indicating that there were doubts as to the content of the considerations; G - General, indicating that the comment is general in relation to the considerations, item a which corresponds (which can be related to a phase, task or in general of the catalog), the text of the comment itself, and a
suggestion with the proposal of the reviewer to circumvent the problem. The defined evaluation material and catalogue were sent to the selected reviewer via e-mail contact.

In general, the feedback received after the evaluation was very beneficial and favorable for the improvement of the catalogue. The evaluator considered the proposal of the catalogue complete and profound. There were some editorial considerations, referential and the inclusion of example of use of some approaches. After receiving the evaluation of the catalogue, the researchers made the corrections and the necessary observations, fulfilling all the mentioned considerations. The evaluation form can be accessed at: https://drive.google.com/file/d/1Y6wYRWaNTMep5X RyPusy31H5j3sLHpO/view?usp=sharing.

7 THE APPLICATION OF SOFTWARE TOOLS CATALOGUE

The application of this catalogue can be in the academic or industrial context in organizations that aim to implement the approaches related to Statistical Process Control and need clarification as to its definition and examples of practical use. In the catalogue we also mentioned some examples of software tools that implement the SPC and its methodology, with examples of use. The catalogue aims to contribute to the process of building knowledge of organizations regarding the SPC, clarify more general doubts and offer possibilities for improvement, in order to enhance its use and practical application in future projects.

8 CONCLUSIONS

The work carried out focused on the investigation of SPC-related approaches by a SLR. The software tools included in this catalogue, its description, use example, information about its availability and information on the ownership of the tool have been presented in this paper. The theoretical basis applied in this study was presented, related works, methodology and main results obtained by the systematic literature review, which gave rise to the catalogue of approaches, which, in this work, brought information from the cataloged software tools.

On the software tools presented in the catalogue it was possible to get a glimpse of how the SPC and its related approaches have been implemented. The catalogue can serve as a benchmark for future SPC research in software organizations.

So, the main scientific contribution of this work was to describe a set of software tools that can support during the implementation of Statistical Process Control in the software context from studies already published in conferences and journals, using as a means the application of a systematic review of the literature.

This work is a contributing part of a Master's Dissertation and other scientific productions (Neto et al., 2017; Neto et al., 2018). As evolution for this study some future work is defined: (i) reaplication of the systematic review, expanding the collection period and including new sources of research; (ii) the application of this catalogue in the industry, with its use in a software development project, and may be of great value for the maturation of this, because opportunities for improvement in relation to its structure, presentation form or even content may be found, in addition to its effectiveness as to what is proposed can be evaluated; (iii) proposing a survey with research teams in industry and academia to collect information on how to use the approaches identified in the catalogue, trying to describe the differences between theory and practical application of approaches, advantages, difficulties, social aspects and others; (iv) expanding the catalogue including new techniques, more examples of using approaches and software tools.

REFERENCES


A Software Tools Catalogue to Support the Statistical Process Control on the Software Context

España y Portugal, Sistema de Información Científica. Vol. 12, Pg: 55 -76, Portugal.


517