Predictive quality control in BPM: proposing a framework for predicting quality anomalies

Naef Saab\textsuperscript{a,b,*}, Remko Helms\textsuperscript{a}, Martijn Zoet\textsuperscript{b}

\textsuperscript{a}Open University, Heerlen, The Netherlands
\textsuperscript{b}Zuyd University of Applied Science, Heerlen, The Netherlands

Abstract

Business process management (BPM) literature suggests that more than 60\% of quality improvement projects fail due to factors associated with the lack of predictive quality control and the search for quality anomalies in quality performance continuously over time. Quality anomalies are indicators of extreme performance deviation from quality expectations and requirements. The findings suggest that quality performance control in BPM is the scientific method for producing quality anomaly knowledge and signalling opportunities for informed, systematic, and continuous quality improvement decisions. A framework based on the findings is proposed, which consists of a structural and procedural component.

Keywords: BPM; TQM; BPR; QFD; quality performance measurement; time series analysis; temporal data mining; predictive analytics; outlier and deviation mining; process mining; quality control; quality improvement; design science research; business process intelligence

1. Introduction

More than 60\% of quality improvement projects in Business Process Management (BPM) fail due to factors associated with predictive quality control, analytics and measurement. Popular reasons of these failures include the inability to detect and predict quality performance problems over time, the lack of long-term temporal performance focus from stakeholders’ perspectives, the lack of stakeholder involvement, unstructured performance measurement...
systems (PMS), and difficulties with prioritizing attention for solving quality performance problems continuously and proactively [1, 2]. Research has attempted to investigate and detect quality anomalies in quality performance control and found gaps related to predictive anomaly detection in performance measurement, such as the inability to predict temporal anomalies which are indexed over time, the focus on point anomalies only, and the limited use of available tools, methods and techniques to detect anomalies [3-5]. From a methodological perspective, Choong [6] notes that organizations need a continuous process for quality performance control in BPM, which defines stakeholders and quality standards, monitors and controls progress against quality standards, and improves processes towards and beyond quality standards sustainably and continuously over time [7]. In this regard, quality control in BPM addresses at least five dimensions of quality performance problems: time, compliance, nature of performance measurement, efficiency and effectiveness expectations, process design and product design [2]. To address these quality problems, this research therefore addresses the following research question and associated research objectives for generating a predictive quality performance control framework:

What framework could be proposed for predicting quality anomalies with quality performance control in BPM?

1. Understand the role of quality control and quality improvement in BPM;
2. Understand the analytical methods for implementing quality control and quality improvement;
3. Understand analytical methods for detecting and predicting anomalies in data indexed by time units.

2. Quality Anomalies and Quality Performance Control in BPM

Studies have shown that quality performance of enterprise capabilities has both positive and negative effects on financial performance and stakeholder satisfaction [8, 9]. Business or enterprise capability is the ability to integrate resources, knowledge, technology, processes and other capabilities in alignment with stakeholders’ objectives and satisfaction [10]. In line with this definition, this study defines quality performance as the degree of fulfilling critical-to-quality performance requirements and expectations (CTQPRE) of enterprise capabilities which satisfy stakeholders [11]. Stakeholders include but are not limited to customers, management, process owners, product owners and suppliers [12, 13]. Their roles include collaborators, recipients, influencers, and claimants [13]. From a stakeholder’s perspective, Golder, Mitra [9] propose furthermore that quality performance is a set of three different states consisting of perceived quality, produced quality and evaluated quality. Their proposed framework suggests that quality is perceived from a stakeholder’s perspective, is a process, and is subjected to quality performance control, evaluation, assessment, diagnostic, audit or process control [9, 13]. Therefore, quality performance is multidimensional, subjective and interpreted in a social context, consisting for example of continua between socially accepted value (good) and waste (bad) and between process and outcomes [14]. These dimensions are in line with ethical, economical, legal and philanthropic roles associated with stakeholder, agency and social responsibility theories [13]. Along these dimensions, extremely wide gaps between stakeholder CTQPRE and actual performance of enterprise capabilities reflect high variations in quality performance and indicate quality performance problems, quality weakness, quality performance gaps, quality deficiencies, or quality anomalies [15]. These quality anomalies are detected with evaluation, diagnosis, assessment, audit, or control of the variation between actual performance and CTQPRE, which are key activities of quality performance control [16].

From a performance measurement perspective in quality control, quality anomalies reflect quality performance problems as extreme variations in performance metrics relative to CTQPRE metrics [9]. Quality anomalies indicate whether CTQPRE are extremely far from being met (i.e. negative quality anomalies) or are being extremely exceeded (i.e. positive quality anomalies) [2]. From a conceptual perspective, positive quality anomalies in quality performance are problems which suggest opportunities for capitalizing on excess value or adjusting CTQPREs, while negative quality anomalies suggest opportunities for reducing excess waste or adjusting CTQPREs. Monitoring quality anomalies with quality performance control is furthermore an essential stakeholder management endeavor for perceiving, controlling, and improving the production process and outcomes of quality, with the aim of ensuring that performance aligns with stakeholders’ expectations, requirements, and satisfaction levels [9]. From an empirical perspective, quality anomalies are reflected and can be measured with data as well [17]. Literature suggests three types of quality anomalies which stakeholders should detect in quality performance control in BPM over time. As
pointed out by Ahmed, Mahmood [17], point anomalies refer to an individual instance in the data which exhibits extreme performance deviations from CTQPRE metrics. Collective anomalies is when multiple point anomalies in the data reflect extreme performance deviation from CTQPRE metrics simultaneously, such as multiple Enterprise Resource Planning (ERP) process instances experiencing quality anomalies simultaneously [11, 17-19]. Context anomaly is when data instances reflect extreme performance deviations from CTQPRE metrics in particular circumstances such as in particular seasons, months, years, days, experiments, improvements or implemented decisions [17, 20]. Point anomalies have received the most attention in business process literature [21].

As the occurrence of quality anomalies in quality performance transcends functional boundaries of business enterprises, all management disciplines are faced with the quality performance control challenge of detecting and managing quality anomalies over time. Business Process Management (BPM) is one such discipline which encounters quality anomalies, especially in ERP systems. BPM has received wide attention as a holistic, cross-functional discipline in business management with a focus on technology, processes, stakeholders, quality and improvement [22, 23]. Three streams of research have defined the BPM discipline from a quality performance control and improvement perspective. While Total Quality Management (TQM) takes a scientific, evidenced-based, and statistical approach for detecting and reducing quality anomalies through incremental quality performance improvement, Business Process Re-engineering (BPR) focusses on systematically reducing process anomalies through visualization and radical quality performance improvements [11]. Automation focuses on systematically reducing system anomalies through continuous quality performance improvement using information technology such as ERP systems, process mining, machine learning, business rules, workflow management, information systems, document management, and business intelligence [19]. Based on these streams of research, BPM has become a capability where resources, technology and processes are continuously combined, orchestrated, monitored for quality performance anomalies, and improved in alignment with stakeholders and their CTQPREs [10].

3. Research Methodology

Based on the works of Dresch, Lacerda [24], Wieringa [25], Sidorova and Isik [23], this study adopts a design science research approach in order to generate knowledge about the research objectives and to propose an actionable predictive quality performance control framework. This approach consists of thematic analyses for selecting relevant articles and theory grounding for addressing the research objectives and proposing an actionable predictive quality control framework for a visual evaluation. For each research objective, four quality criteria are employed.

Quality Criteria 1: Articles published between the years 1996 and 2017 in peer-reviewed English language journals are retrieved through EBSCOHost API integration for university libraries, using Boolean search terms derived from the research question and objectives. EBSCOHost API accesses a variety of popular scientific databases such as IEEE Xplore Digital Library, ScienceDirect, SpringerLink, Wiley Online Library, Academic Search, JSTOR, Sage Premier, and Google Scholar. Only articles published in journals such as computer science, performance, process, quality, business, analytics, statistics, BPM, TQM, and intelligence are considered.

Quality Criteria 2: Following the examples of Sidorova and Isik [23] for thematic analyses, this study applies Latent Semantic Analysis (LSA) of articles retrieved in quality criterion one. LSA in this regard analyses words in abstracts and generates latent themes using principal component analysis (PCA) in Python computer programming language. Latent themes are subsequently selected using elbows in the scree plots. K-means cluster analysis forms groups of article using distance measures based on the Euclidian distance. The number of groups are selected based on the maximum silhouette coefficients [26]. A word cloud per cluster subsequently displays the most commonly occurring themes in the abstracts of the articles. Only the clusters of articles displaying high frequencies of the themes related to quality, business, process, control, improvement, management, BPM, performance, framework, system, information, TQM, data, and modelling are selected for further analysis.

Quality Criteria 3: The titles of the articles which are excluded in quality criterion two are reviewed to ascertain if relevant topics where excluded. Subsequently, the selected articles are reviewed manually to ascertain their relevance in addressing the research questions and objectives. Articles exhibiting generalizable research methods, conceptual frameworks, and systematic literature reviews are subsequently selected for further analysis.

Quality Criteria 4: The final set of articles are grouped and compared per research objectives for theory grounding which consists of comparing methodologies and identifying underlying theory for generating a framework consisting of a structural for providing context and procedural component for providing guidance [24].
4. Results and Discussion

4.1 Quality Control and Quality Improvement in BPM

Several factors are hypothesized to create quality anomalies in BPM and drive the need for incremental, radical, and continuous quality performance improvement such as dynamic business environment, stakeholders performance expectations, competition, external customers, laws and regulation, management systems, fads and fashion, and importance of quality management [27]. These changes increase quality performance problems and create quality anomalies, often continuously [28]. BPM stakeholders encounter continuous weakness in quality performance, which could decrease stakeholder satisfaction and motivate stakeholders to take action for improving quality performance continuously [29]. In their literature review about continuous improvement (CI), Sanchez and Blanco [27] highlights CI as a key feature of quality improvement. They furthermore suggest that CI is a cycle consisting of activities which should be repeated over time for maintaining stakeholder satisfaction levels. Research on CI are classified into nine categories such as concept definition, implementation of CI, drivers/obstacles, methodologies, culture, control, management philosophies, technology, innovation, and human resources [27]. In this regard, the BPM literature highlights the importance of CI methods and control activities from different stakeholder perspectives, which explains the variety of CI methods organized under the BPM umbrella [19]. Siha and Saad [30] for example reviewed the critical success and failure factors of four popular quality improvement methods used in BPM: Six Sigma, Benchmarking, Business Process Reengineering, and Process mapping. Morais, Kazan [31] compare seven BPM improvement cycles against a professional BPM framework and found little alignment with strategy and stakeholders. These methods aim to reduce differences between CTQPRE and actual performance using applied research, evidence, data, statistics, modelling, observation, facts, reasoning and learning [32, 33]. Tort-Martorell, Grima [34] highlights this in their evidence-based management approach for reducing quality anomalies in business management, which relies on data for making informed decisions regarding performance improvement. Jeston and Nelis [19] and Association of Business Process Management Professionals (ABPMP) furthermore propose a BPM cycle for systematically integrating and improving BPM capabilities and aligning these with stakeholders’ expectations [31]. Literature suggests that these CI cycles are variants of Deming’s Plan-Do-Check-Act (PDCA) cycle for quality improvement as illustrated in Table 1 [29, 35].

<table>
<thead>
<tr>
<th>Scientific Method</th>
<th>Deming</th>
<th>Lean</th>
<th>Six Sigma</th>
<th>Evidenced-based Management</th>
<th>BPM</th>
<th>Process Performance Measurement</th>
<th>ABPMP</th>
<th>Design Science Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deductive</td>
<td>Do</td>
<td>Review</td>
<td>Measure</td>
<td>Questions, Data Analysis</td>
<td>Launch</td>
<td>Measure Process</td>
<td>Monitoring, Controlling</td>
<td>Evaluate</td>
</tr>
<tr>
<td>Reasoning</td>
<td>Check</td>
<td>Investigate, Verify, Control</td>
<td>Analyze, Control</td>
<td>Understand, Sustainable Performance</td>
<td>Assess Process Performance</td>
<td>Conduct Process Improvement</td>
<td>Refined, Apply</td>
<td></td>
</tr>
<tr>
<td>Reasoning</td>
<td>Act</td>
<td>Execute</td>
<td>Improve</td>
<td>Information</td>
<td>Innovate, People, Development, Implement, Realize</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The PDCA cycle is a scientific methodology for generating quality anomaly knowledge, solving quality performance problems, and improving quality performance [6, 36]. Scientific methodologies are preferred in business enterprises because they enable scientific knowledge production with measurement and validation; they are also more reliable and rigorous for making informed quality improvement decisions [34]. The desire to measure quality performance and achieve reliable quality improvement have demanded a more systematic approach to quality control, based on the scientific method [32]. The scientific method is a systematic approach for producing knowledge about phenomena with the purpose of filling knowledge gaps and/or solving practical problems [24, 32, 33]. The scientific method consists of inductive reasoning where limited observation and experiments are used to build expectations that are not always valid, to identify potential quality problems, or to identify potential solutions which are not always perfect. The scientific method also consists of deductive reasoning which relies on empirical evidence and data to validate assumptions, potential problems and solutions. Quality performance control also relies
on these reasoning methods, and these methods are commonly observed in auditing, process mining, process control, quality assurance, and performance measurement [37-39]. As such, quality control in BPM is the systematic generation of knowledge about quality anomalies in quality performance with the objective of improving quality performance of BPM capability. During the first inductive reasoning phase, quality control systematically produces tentative BPM CTQPRE knowledge using potential evidence from interviews, observations and experiments. During the deductive reasoning phase, quality control relies on evidence and performance measurement to produce quality performance knowledge and validate CTQPREs. During the third inductive phase, validated quality control knowledge is used for communicating significant quality performance weakness and driving potential ideas for quality improvement in BPM [19, 32].

4.2 Analytical Methods for Quality Control and Quality Improvement

Several tools, capabilities, and techniques are used for building quality control knowledge and conducting quality improvement such as Quality Function Deployment (QFD), Balanced Scorecard, process mining, statistical process control, brainstorming, interviews, questionnaires, cause-and-effect diagrams, flow charts, Pareto charts, scatter plots, stratification, machine learning, statistical modelling, run and control charts, and check-and-data collections sheets [11, 19, 40]. These tools and techniques are often combined and play an important role in the scientific method and measurement [32, 34]. They are essential for inductive reasoning in terms of generating CTQPRE knowledge about BPM performance capabilities; for deductive reasoning in terms of validating this knowledge with empirical evidence, data, and performance measurement; and for inductive reasoning in terms of suggesting areas and ideas for improvement in quality performance of business capabilities [32]. They are relevant for producing and applying quality performance control knowledge in BPM. Veit, Camargo [41] proposes two types of scientific knowledge production in BPM in this regard. Mode 1 Knowledge Production (KP) refers to producing knowledge about the scientific method from an academic perspective and Mode 2 KP refers to producing knowledge aimed at solving practical problems using the scientific method [41]. Mode 1 KP entails building knowledge about process modelling and improvement, process management, BPM methodologies, relationship and integration process, and process implementation and automation [6]. Mode 2 KP entails a variety of KP classes, including but not limited to performance indicators and measures, continuous improvement, supply chain, outsourcing, change management and innovation [6]. The scientific method also defines the nature of Mode 2 KP as can be seen with process performance measurement, PDCA, Lean, Process Mining, Six Sigma, evidenced-based management and design science research [24, 34]. Mode 2 KP is therefore a popular applied scientific method for quality performance control in BPM in terms of generating quality anomaly knowledge with the purpose of improving quality performance of BPM capabilities. Recently, business intelligence (BI) has gained momentum as combining Mode 1 for improving scientific methods and Mode 2 for applying scientific methods for quality control and improvement [24, 42]. BI consists of interconnected tools, people, processes, and technology, which are orchestrated in scientific discipline and systematically converted into actionable quality control knowledge for quality improvement. BI supports quality control and uses performance measurement methods systematically to identify quality anomalies, drive quality improvement, and validate quality improvement. In this regard, business process intelligence (BPI) is increasingly considered as a capability for generating and applying quality anomaly knowledge systematically in quality control using performance measurement frameworks, information communication technology (ICT), process mining, statistical process control and quality control in BPM [42]. Business process intelligence (BPI) enables Mode 1 & 2 KP in BPM using design science research for producing quality control knowledge and reporting artefacts [24].

4.3 Predicting Quality Performance and Detecting Quality Anomalies Over Time

One limitation of the current quality performance control methods in BPI is the reactive focus on historical and temporal data which are indexed over time, and the reactive focus on taking actions after quality anomalies have occurred [11]. Using time series or temporal data, proactive quality control predicts quality performance in the future and detects quality anomalies continuously [37]. Time series or temporal data are data points indexed by time units such as by hour, day, week, month or year [43]. In order to predict quality performance over time, the temporal data mining literature suggests different techniques for predicting temporal data. Temporal data mining, which is often referred to as time series analysis, is defined as the extraction of knowledge from large amounts of time series data using supervised, unsupervised, or reinforcement machine learning [44]. Liao, Chu [45] conducted a systematic review and summarized popular data mining techniques such as support vector machine, a priori algorithm, artificial
neural network, clustering, genetic algorithm, association rule, decision tree, neural network, classification, and feature selection. These techniques are useful for analyses which do not rely on statistical assumptions such as normal distribution, absence of autocorrelation and linearity in the data. However, auto regressive integrated moving average (ARIMA) has been the most popular statistical technique for modelling linearity and autocorrelation in time-series data [46]. Autocorrelation or serial correlation is an issue when data points correlate over time. Temporal data are both linear and non-linear, are often correlated over time, and are not always normally distributed. Therefore, combining statistical models for modelling linear components with machine learning for modelling non-linear components is becoming a popular ensemble machine learning method in data mining [47].

Machine learning and statistics play a key role in detecting quality anomalies in predictive quality control. Also known as outlier mining, deviation mining, novelty detection or exception mining; anomaly detection techniques are helpful for identifying quality anomalies in quality performance in time series or temporal data [48, 49]. Literature consists of several quality anomaly detection techniques for quality control. Agyemang, Barker [50] for example classify outlier mining techniques into numeric techniques, cluster-based techniques, and symbolic techniques. Numeric techniques are statistically based and are the most commonly used techniques used today in quality control and quality improvement, such as z-scores [11]. Cluster-based techniques rely on grouping of data and finding small sized groups with extremely large or small values. Symbolic based techniques convert unstructured data into numeric data for further analyses using numeric and cluster-based techniques [3, 17, 20, 51-54]. One common theme related to data mining from a machine learning and statistical perspective is the systematic process or scientific method which is deployed to produce knowledge from evidence and data using analytics. Known as the Cross Industry Standard Process for Data Mining, CRISP-DM consists of a method of understanding the business and its data, managing the data, developing the analysis model, evaluating the models for quality, and deploying the models, which is more in line with Mode 1 KP [44, 45, 55]. The Knowledge Discovery in Databases (KDD) model used for evidenced-based management extends the CRISP-DM cycle to practical applications including the BPM capabilities, which is in line with knowledge engineering, business intelligence capability, design science research and Mode 1 and 2 KP [42, 56]. The KDD model is proposed to play an essential role in BPI and design science research for building knowledge and data artefacts for solving quality performance problems in BPM.

5. Predictive Quality Performance Control Framework

This research proposes a predictive quality control framework with a structural component (1.0) for providing context, and a procedural component (2.0) for implementation. As shown in Figure 1, the structural component proposes three capabilities for supporting the procedure for predicting quality anomalies in quality control. Firstly, knowledge of BPM capability (1.1) is essential and should consist of different dimensions such as people, strategy, performance management, processes, governance, organization and ICT [10, 19]. Quality Function Deployment (QFD) is the second capability and consists of different Houses of Quality where stakeholder requirements and expectations are translated to activities and performance control via a quality chain. The QFD capability (1.2) starts with the defining the BPM stakeholders and their CTQPREs, subsequently converting the CTQPREs into products/services, processes, and activity design, and finally monitoring the performance of these components relative to the CTQPREs. The use of QFD methods and tools such as qualitative and quantitative research methods in BPM shifts the focus from performance management to quality performance management using scientific research. However, as the TQM literature suggests, the lack of technology and inability for continuous monitoring in quality performance management have also contributed to failure of quality improvement projects [11]. The growing amount of data and analytical methods has provided more opportunities for information technology in BPM and TQM to overcome the challenge of dealing with growing amount of data and evidence. Business Process Intelligence (BPI) in this regard is the third capability which combines different dimensions of the BPM capability and QFD for building and applying process knowledge systematically. The BPI capability (1.3) is further characterized by understanding business process needs from stakeholders’ CTQPREs, understanding process data, managing process data, producing knowledge and artefacts with models, evaluating models, and applying generated knowledge and artefacts for solving quality performance problems in BPM [42]. As such, the proposed framework assumes that BPM, QFD, and BPI are capabilities already in place and should be checked for before continuing with the procedural component. These capabilities align with the contextual capabilities in Wieringa [25] Design Science Research (DSR) framework: organization (BPM), social (QFD) and knowledge contexts (BPM, QFD and BPI).
Design science research is a systematic, applied, and actionable research process for building knowledge and artifacts with the aim for solving practical problems with rigor and relevance using inductive and deductive reasoning [24]. Artifacts are the predicted quality anomalies. Displayed in the center of Figure 1, the procedural component is reflected by a design science research process which aligns with Deming’s scientific PDCA cycle and Choong [6], Kueng [36] process performance measurement methods. During the Plan phase, practitioners start with determining, understanding, and designing BPM stakeholders and their tentative CTQPREs about BPM capabilities using inductive reasoning such as interviews, case studies, surveys, focus groups, and literature review. During the Do Phase, CTQPRE are translated into performance target measures and operationalized into metrics in partnership with BPM, QFD, and BPI stakeholders. Performance metrics are subsequently operationalized into data requirements and data retrieval in this phase with BPI stakeholders. Temporal or time series data are recorded and extracted and indexed by time units such as by hour, day, week, month or year. During the Study phase, quality performance predictions per time unit are generated and evaluated for future time units using statistical modelling [46] and machine learning [45] in BPI. Predictive quality performance is also determined at this phase as the difference between CTQPRE metrics and predicted performance metrics, which are indexed by time units. Using statistical z-score, practitioners can define the points at which anomalies occur [17, 20, 54]. Quality performance data points in time series which are located beyond a certain six sigma standard deviation from the mean are considered quality anomalies in quality performance or quality problems and are communicated as design science research artefacts in the Act Phase.

6. Conclusions and Avenues for Future Research

Quality performance control is the process of scientifically producing Mode 2 knowledge about extreme performance deviations from stakeholders’ requirements and expectations. Quality performance control is embedded in quality improvement and is a knowledge production process to guide quality improvement. Quality improvement methods in BPM are variants of scientific continuous improvement methods, which rely on continuous knowledge production of quality anomalies in quality performance. This research takes a design science approach for generating knowledge for addressing the research objectives and for proposing an actionable framework (i.e. artefact), which consists of a structural and procedural component. The structural component provides the context,
capabilities, and assumptions for enabling the procedural component. Future research should validate and extend the proposed framework with systematic literature review (SLR), case studies, and surveys in companies with and without the structural capabilities shown in Figure 1 and validate and improve the features in this framework. Other studies could conduct a case study and test the cycle in practice in situations where BPM, QFD, and BPI capabilities are already in place in order to further improve the framework. For example, process discovery in process mining can be used in BPM to retrieve the frequency counts of multiple process models per day over time in BPM. These help define or validate CTQPREs for practitioners with inductive reasoning in QFD. The predicted occurrences of these models could be converted into normalized z-scores and used to form three clusters of process model variants: one cluster with positive quality anomalies, one with negative quality anomalies and one with expected quality performance of process models. Communicating these collective quality anomalies with BPI reports to BPM stakeholders completes the design science research procedural component.

References


22. Dresch, D.P., Lacerda, and J.A. Antunes Ir, Design Science Research. 2015: Springer International Publishing Switzerland.
23. Wieringa, R.J., Design science methodology for information systems and software engineering. 2014: Springer.