

Prediction of Business Process Model Quality based on Structural Metrics

Laura Sánchez-González¹, Félix García¹, Jan Mendling², Francisco Ruiz¹,
Mario Piattini¹

¹Alarcos Research Group, TSI Department, University of Castilla La Mancha, Paseo de la
Universidad, n°4, 13071, Ciudad Real, España
{Laura.Sanchez | Felix.Garcia | Francisco.RuizG | Mario.Piattini}@uclm.es

² Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany,
jan.mendling@wiwi.hu-berlin.de

Abstract. The quality of business process models is an increasing concern as enterprise-wide modelling initiatives have to rely heavily on non-expert modelers. Quality in this context can be directly related to the actual usage of these process models, in particular to their understandability and modifiability. Since these attributes of a model can only be assessed a posteriori, it is of central importance for quality management to identify significant predictors for them. A variety of structural metrics have recently been proposed, which are tailored to approximate these usage characteristics. In this paper, we address a gap in terms of validation for metrics regarding understandability and modifiability. Our results demonstrate the predictive power of these metrics. These findings have strong implications for the design of modelling guidelines.

Keywords: Business process, measurement, correlation analysis, regression analysis, BPMN

1. Introduction

Business process models are increasingly used as an aid in various management initiatives, most notably in the documentation of business operations. Such initiatives have grown to an enterprise-wide scale, resulting in several thousand models and involving a significant number of non-expert modellers [1]. This setting creates considerable challenges for the maintenance of these process models, particularly in terms of adequate quality assurance. In this context, quality can be understood as “the totality of features and characteristics of a conceptual model that bear on its ability to satisfy stated or implied needs”[2]. It is well known that poor quality of conceptual models can increase development efforts or results in a software system that does not satisfy user needs [3]. It is therefore vitally important to understand the factors of process model quality and to identify guidelines and mechanisms to guarantee a high level of quality from the outset.

An important step towards improved quality assurance is a precise quantification of quality. Recent research into process model metrics pursues this line of argument by measuring the characteristics of process models. The significance of these metrics relies on a thorough empirical validation of their connection with quality attributes [4]. The most prominent of these attributes are understandability and modifiability, which both belong to the more general concepts of usability and maintainability, respectively [5]. While some research provides evidence for the validity of certain metrics as predictors of understandability, there is, to date, no insight available into the connection between structural process model metrics and modifiability. This observation is in line with a recent systematic literature review that identifies a validation gap in this research area [6].

In accordance with the previously identified issues, the purpose of this paper is to contribute to the maturity of measuring business process models. The aim of the empirical research presented herein is to discover the connections between an extensive set of metrics and the ease with which business process models can be understood (understandability) and modified (modifiability). This was achieved by adapting the measures defined in [7] to BPMN business process models [8]. The empirical data of six experiments which had been defined for previous works were used. A correlation analysis and a regression estimation were applied in order to test the connection between the metrics and both the understandability and modifiability of the models.

The remainder of the paper is as follows. In Section 2 we describe the theoretical background of our research and the set of metrics considered. Section 3 describes the series of experiments that were used. Sections 4 and 5 present the results. Finally, Section 6 draws conclusions and presents topics for future research.

2. Structural Metrics for Process Models

In this paper we consider a set of metrics defined in [6] for a series of experiments on process model understanding and modifiability. The hypothetical correlation with understandability and modifiability is annotated in brackets as (+) for positive correlation or (-) for negative correlation. The metrics include:

- Number of nodes (-): number of activities and routing elements in a model;
- Diameter (-): The length of the longest path from a start node to an end node;
- Density (-): ratio of the total number of arcs to the maximum number of arcs;
- The Coefficient of Connectivity (-): ratio of the total number of arcs in a process model to its total number of nodes;
- The Average Gateway Degree (-) expresses the average of the number of both incoming and outgoing arcs of the gateway nodes in the process model;
- The Maximum Gateway Degree (-) captures the maximum sum of incoming and outgoing arcs of these gateway nodes;
- Separability (+) is the ratio of the number of cut-vertices on the one hand to the total number of nodes in the process model on the other;
- Sequentiality (+): Degree to which the model is constructed out of pure sequences of tasks.

- Depth (-): maximum nesting of structured blocks in a process model;
- Gateway Mismatch (-) is the sum of gateway pairs that do not match with each other, e.g. when an AND-split is followed by an OR-join;
- Gateway Heterogeneity (-): different types of gateways are used in a model;
- Cyclicity (-) relates the number of nodes in a cycle to the sum of all nodes;
- Concurrency(-) captures the maximum number of paths in a process model that may be concurrently activate due to AND-splits and OR-splits.

3. Research Design

The empirical analysis performed is composed by six experiments: three to evaluate understandability and three to evaluate modifiability. The experimental material for the first three experiments consisted of 15 BPMN models with different structural complexity. Each model included a questionnaire related to its understandability. The experiments on modifiability included 12 BPMN models related to a particular modification task. A more detailed description of the family of experiments can be found in [9]. It was possible to collect the following objective data for each model and each task: time of understandability or modifiability for each subject, number of correct answers in understandability or modifiability, and efficiency defined as the number of correct answers divided by time.

Once the values had been obtained, the variability of the values was analyzed to ascertain whether the measures varied sufficiently to be considered in the study. Two measures were excluded, namely Cyclicity and Concurrency, because the results they offered had very little variability (80% of the models had the same value for both measures, the mean value was near to 0, as was their standard deviation). The remaining measures were included in the correlation analysis.

The experimental data was accordingly used to test the following null hypotheses for the current empirical analysis, which are:

- For the experiments on understandability,
H0,1: There is no correlation between structural metrics and understandability
- For the experiments on modifiability,
H0,2: there is no correlation between structural metrics and modifiability

The following sub-sections show the results obtained for the correlation and regression analysis of the empirical data.

4. Correlation analysis

Understandability: Understanding time is strongly correlated with number of nodes, diameter, density, average gateway degree, depth, gateway mismatch, and gateway heterogeneity in all three experiments. There is no significant correlation with the connectivity coefficient, and the separability ratio was only correlated in the first experiment. With regards to correct answers, size measures, number of nodes (-.704 with p-value of .003), diameter (-.699, .004), and gateway heterogeneity (.620, .014)

have a significant and strong correlation. With regard to efficiency, we obtained evidence of the correlation of all the measures with the exception of separability.

The correlation analysis results indicate that there is a significant relationship between structural metrics and the time and efficiency of understandability. The results for correct answers are not as conclusive, since there is only a correlation of 3 of the 11 analyzed measures. We have therefore found evidence to reject the null hypothesis **H0,1**. The alternative hypothesis suggests that these BPMN elements affect the level of understandability of conceptual models in the following way. It is more difficult to understand models if:

- There are more nodes.
- The path from a start node to the end is longer.
- There are more nodes connected to decision nodes.
- There is higher gateway heterogeneity.

Modifiability: We observed a strong correlation between structural metrics and time and efficiency. For correct answers there is no significant connection in general, while there are significant results for diameter, but these are not conclusive since there is a positive relation in one case and a negative correlation in another. For efficiency we find significant correlations with average (.745, .005) and maximum gateway degree (.763, .004), depth (-.751, .005), gateway mismatch (-.812, .001) and gateway heterogeneity (.853, .000). We have therefore found some evidence to reject the null hypothesis **H0,2**. The usage of decision nodes in conceptual models apparently implies a significant reduction in efficiency in modifiability tasks. In short, it is more difficult to modify model the model if:

- More nodes are connected to decision nodes.
- There is higher gateway heterogeneity.

5. Regression analysis

The previous correlation analysis suggests that it is necessary to investigate the quantitative impact of structural metrics on the respective time, accuracy and efficiency dependent variables of both understandability and modifiability. This goal was achieved through the statistical estimation of a linear regression. The regression equations were obtained by performing a regression analysis with 80% of the experimental data. The remaining 20% were used for the validation of the regression models. The first step is selected the prediction models with p-values below 0.05. Then, it is necessary to validate the selected models verifying the distribution and independence of residuals through Kolmogorov-Smirnov and Durbin-Watson tests. Both tests values are considered to be satisfactory.

The accuracy of the models was studied by using the Mean Magnitude Relative Error (MMRE) [10] and the prediction level Pred(25) and Pred(30) on the remaining 20% of the data, which were not used in the estimation of the regression equation. These levels indicate the percentage of model estimations that do not differ from the observed data by more than 25% and 30%. A model can therefore be considered to be accurate when it satisfies any of the following cases: a) $MMRE \leq 0,25$ or b) $Pred(0,25) \geq 0,75$ or c) $Pred(0,30) \geq 0,70$. Table 3 depicts the results.

Table 1. Prediction models of understandability

	Exp	Prediction model	MMRE	p(0,.25)	p(0,.30)
Understandability					
Time	E3	$47.04 + 2.46 \text{ n}^\circ\text{nodes}$.32	.51	.58
Correct answers	E2	$3.17 - 0.005 \text{ n}^\circ\text{nodes} - 0.38 \text{ coeff. of connectivity} + 0.17 \text{ depth} - 0.015 \text{ gateway mismatch}$.18	.79	.79
Efficiency	E3	$0.042 - 0.0005 \text{ n}^\circ\text{nodes} + 0.026 \text{ sequentiality}$	0.84	.22	.25
Modifiability					
Time	E4	$50.08 + 3.77 \text{ gateway mismatch} + 422.95 \text{ density}$.37	.31	.38
C.A.	E4	$1.85 - 3.569 \text{ density}$.23	.82	.83
Efficiency	E4	$0.006 + 0.008 \text{ sequentiality}$.62	.32	.42

Understandability: The best model for predicting the understandability time is obtained with the E3, which has the lowest MMRE value of all the models. The best models with which to predict correct understandability answers originate from the E2, and this also satisfies all the assumptions. For efficiency, no model was found that satisfied all the assumptions. The model with the lowest value of MMRE is obtained in the E3. In general, the results further support the rejection of the null hypothesis **H0,1**.

Modifiability: We did not obtain any models which satisfy all of the assumptions for the prediction of modifiability time, but we have highlighted the prediction model obtained in E4 since it has the best values. However, the model to predict the number of correct answers may be considered to be a precise model as it satisfies all the assumptions. The best results for predicting efficiency of modifiability are also provided by E4, with the lowest value of MMRE. In general, we find some further support for rejecting the null hypothesis **H0,2**. The best indicators for modifiability are gateway mismatch, density and sequentiality ratio. Two of these metrics are related to decision nodes. Decision nodes apparently have a negative effect on time and the number of correct answers in modifiability tasks.

6. Conclusions and Future Work

In this paper we have investigated structural metrics and their connection with the quality of business process models, namely understandability and modifiability.

The statistical analyses suggest rejecting the null hypotheses, since the structural metrics apparently seem to be closely connected with understandability and modifiability. For understandability these include Number of Nodes, Gateway Mismatch, Depth, Coefficient of Connectivity and Sequentiality. For modifiability Gateway Mismatch, Density and Sequentiality showed the best results. The regression analysis also provides us with some hints with regard to the interplay of different metrics. Some metrics are not therefore investigated in greater depth owing to their correlations with other metrics.

Our findings demonstrate the potential of these metrics to serve as validated predictors of process model quality. Some limitations in the experimental data are about the nature of subjects, which implies that results are particularly relevant to non-expert modellers. This research contributes to the area of process model measurement and its still limited degree of empirical validation. This work has implications both for research and practice. The strength of the correlation of structural metrics with different quality aspects (up to 0.85 for gateway heterogeneity with modifiability) clearly shows the potential of these metrics to accurately capture aspects that are closely connected with actual usage. From a practical perspective, these structural metrics can provide valuable guidance for the design of process models, in particular for selecting semantically equivalent alternatives that differ structurally. In future research we aim to contribute to the further validation and actual applicability of process model metrics

Acknowledgments. This work was partially funded by projects INGENIO (PAC 08-0154-9262); ALTAMIRA (PII2109-0106-2463), ESFINGE (TIN2006-15175-C05-05) and PEGASO/MAGO (TIN2009-13718-C02-01).

References

1. Rosemann, M., *Potential pitfalls of process modeling: part a*. Business process Management Journal, 2006. **12**(2): p. 249-254.
2. ISO/IEC, *ISO Standard 9000-2000: Quality Management Systems: Fundamentals and Vocabulary*. 2000.
3. Moody, D., *Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions*. Data and Knowledge Engineering, 2005. **55**: p. 243-276.
4. Zelkowitz, M. and D. Wallace, *Esperimental models for validating technology*. IEEE Computer, Computing practices, 1998.
5. ISO/IEC, *9126-1, Software engineering - product quality - Part 1: Quality Model*. 2001.
6. Sánchez, L., F. García, F. Ruiz, and M. Piattini, *Measurement in Business Processes: a Systematic Review*. Business process Management Journal, 2010. **16**(1): p. 114-134.
7. Mendling, J., *Metrics for Process Models: Empirical Foundations of Verification, Error Prediction, and Guidelines for Correctness*. 2008: Springer Publishing Company, Incorporated.
8. OMG. *Business Process Modeling Notation (BPMN), Final Adopted Specification*. 2006; Available from: <http://www.omg.org/bpm>.
9. ExperimentsURL, <http://alarcos.inf-cr.uclm.es/bpmnexperiments/>. 2009.
10. Foss, T., E. Stensrud, B. Kitchenham, and I. Myrtveit, *A Simulation Study of the Model Evaluation Criterion MMRE*. IEEE Transactions on Software Engineering, 2003. **29**: p. 985-995.