Estimating the Effort of Independent Verification and Validation in the Context of Mission-Critical Software Systems – A Case Study

Haruka Nakao[a], Adam Trendowicz[b], Jürgen Münch[b]
[a] *Japan Manned Space Systems Corporation, Ibaraki, Japan*
[b]*Fraunhofer Institute for Experimental Software Engineering, Kaiserslautern, Germany*
haruka@jamss.co.jp, {Adam.Trendowicz, Juergen.Muench}@iese.fraunhofer.de

Abstract

The ability to generate sufficiently accurate effort estimates can be seen as a key success factor for multi-organizational projects focusing on the development of large and critical software systems. This is caused, for instance, by the need for synchronizing multiple development and verification and validation processes. Paradoxically, effort predictions in the critical software systems domain are still relying on human judgment. This requires much overhead and its reliability depends largely on the expertise and individual preferences of the involved experts. In particular, verification and validation by independent entity (IV&V) needs an estimation method that supports negotiating and managing IV&V costs in the context of sparse measurement data and low availability of domain experts. In order to address these problems, in this paper we propose applying a hybrid effort estimation method called CoBRA® for estimating effort for IV&V of mission-critical software systems. When applied in an industrial context, CoBRA® improved estimation accuracy and precision by about 40%, on average, compared to experts estimates and OLS regression.

1. Introduction

The average company spends about 4 to 5 percent of its revenue on information technology, with those that are highly IT-dependent - such as financial and telecommunications companies - spending more than 10 percent on it [5]. Now, a great part of those investments is wasted because software organizations are still proposing unrealistic software costs, work within tight schedules, and, in consequence, finish their projects behind schedule and budget (about 50% of projects), or do not complete them at all (more than 25% of projects). Moreover, even though projects are completed within a target plan, the functionality and quality of products delivered are usually cut to fit this plan [11]. This indicates that software project planning is a critical success factor of a software project.

Project planning in the safety-critical domain is particularly important and difficult at the same time. Large functional constraint, high quality requirements, and involvement of several independent parties make it much more challenging to plan mission-critical projects than to plan ordinary, non-critical software development projects. In that context, effective synchronizing activities of all involved parties are a key factor for project success. One example activity requiring such synchronization is verification and validation done by an independent organization, or independent verification and validation (IV&V). A mismatch between the IV&V plan and the overall project plan may lead to significant delays or (in extreme cases) to the skipping of certain IV&V activities. In consequence, the whole mission might be exposed to the high risk of a significant loss of money, and, in the worst case, injury or death of people. Yet, comprehensive support for planning and managing IV&V is missing.

Effort estimation approaches proposed by the research community have traditionally focused on planning and tracking classical, in-house software development. Effort estimation methods that grew upon those objectives focus on providing exact estimates. They do not, however, support an easily understandable, systematic and reliable analysis of the most relevant causal effort dependencies. Even though an accurate prediction is provided, software practitioners have hardly any support to prevent potential project overruns. In the short-term perspective, this would mean a lack of a solid basis for effectively mitigating project risks, and in the long-term perspective, a limited ability to identify process improvement areas and to learn. Moreover, estimation methods promoted by the research community require large data sets, whereas methods commonly employed by industry extensively involve domain experts.

All those aspects significantly reduce the applicability of existing estimation methods in the IV&V context, where reliable and comprehensive project management has to be provided despite the minimal availability of quantitative data and human expertise.

In this paper, we propose applying the Cost Estimation, Benchmarking, and Risk Analysis method (CoBRA®) [12][17] to estimate the effort of the IV&V of mission-critical software systems. CoBRA® is a hybrid method that combines analytical and expert-based estimation. It provides a systematic way to transform various sources of organizational knowledge (minimal set of measurement data and expert judgment) into a transparent and reusable effort model that supports achievement of a variety of project management objectives, such as risk management or negotiating of project costs.

The remainder of the paper is organized as follows: Section 2 briefly characterizes the IV&V context. Section 3 gives an overview of existing software effort estimation methods, followed by a more detailed description of the CoBRA® method in Section 4. Section 5 presents the empirical results of an industrial application of CoBRA® for planning the effort of IV&V, followed by lessons learned (Section 6) in the study. The paper ends up with a brief summary and conclusions given in Section 7.

2. Independent Verification and Validation

2.1. Characteristics of IV&V

Independent verification and validation (IV&V) can be defined as a process where software work products generated by a development team are verified and validated by a completely independent organizational entity. Independence is considered here [7] in terms of technical, managerial, and financial independence. IV&V is typically applied in the context of safety- or mission-critical software systems, such as space and nuclear plant systems.

The typical constraint of IV&V, as compared to classical in-house V&V, is limited information on processed artifacts. On the one hand, there is limited knowledge about the software development environment; on the other hand, IV&V has to handle various types of mission-critical systems. This variety does not allow for collecting many historical project data. Moreover, involvement of three sites (customer-, development-, and IV&V-entity) in the software development process contributes to frequent and unpredictable requirements change.

In that context, managing an IV&V project's resources is critical and difficult at the same time.

2.2. Objectives of Effort Estimation

Besides traditional estimation objectives such as precise planning and tracking software resources, project needs decision-making support. In particular, explicit identification of factors having the greatest impact on IV&V cost and their quantitative impact should be supported. Identification of customer-specific factors (e.g., level of customer support) may be used to justify and negotiate IV&V costs that cannot be influenced by the IV&V supplier. On the other hand, identification of the IV&V supplier's characteristics (process and human capabilities) that have the greatest impact on increased IV&V costs will allow targeting improvement actions to specific process areas and improving the efficiency of IV&V. In consequence, project risks can be mitigated timely and critical organizational processes can be improved.

2.3. Current Effort Estimation Practices

A survey about current estimation practices at Japan Manned Space Systems Corporation (JAMSS) revealed that measurement data from around 10 already completed projects have been collected within the past 10 years. However, the data suffered from significant incompleteness (around 20% of missing data) and large variability – due to the high uniqueness of the considered projects. Since hardly any data-driven estimation method that would meet the estimation objectives (Section 2.2) can be applied reasonably, estimates are typically based on the judgment of one or more domain experts. Although sparse project data are available, experts based their estimates solely on personal experiences. One of the reasons is that the available simple size measures, such as pages of software requirements document, are believed not to reflect the amount of IV&V effort reliably. Yet, expert-based estimation did not provide satisfactory support for project management. First, the reliability of the estimates depends largely on individual expertise and preferences of involved domain expert. In consequence effort estimates are not accurate and vary widely across projects (see Table 4 and Table 5 in Section 5.5). Moreover, estimation costs much effort each time it is performed, and since it does not provide any explicit effort model, it hardly supports decision making in a project.

3. Related Work

Numerous types of estimation methods have been developed over the last decades. In this section we provide a brief overview of existing estimation methods from the viewpoint of their applicability in the context of IV&V. For a comprehensive review and comparative evaluation of existing methods, please refer to [16].

Existing effort estimation methods differ basically with respect to the type of inputs they require and the form of the estimation model they do provide. With respect to input data, we differentiate between three major groups: data-intensive, expert-based, and hybrid methods (combining available data and expert knowledge in order to come up with estimates). Among the data-intensive methods, some require past project data for building customized models (*define-your-own-model* approaches), others provide an already defined model, where factors and their relationships are fixed based on a set of multi-organizational project data (*fixed-model* approaches). The major advantage of fixed-model approaches is that they, theoretically, do not require any historical data to be applied. Those methods might be especially attractive in the IV&V context, where very sparse (if any) data are typically available. Yet, in practice, fixed models, such as COCOMO [2][1], are developed for a specific context (typically different from IV&V) and are, by definition, only suited for estimating the types of projects for which the fixed model was built. The applicability of such models for the IV&V context is, in practice, very limited. In order to improve their performance, a significant amount of organization-specific project data would be required for calibrating the generic model. In that case, the little usefulness of the fixed-model approaches for IV&V effort estimation would not differ much from the define-your-own-model approaches, which require a significant amount of context-specific data to build customized effort models [15]. Application of the define-your-own-model methods in the context of IV&V is further limited by the additional requirements of specific methods. Parametric approaches, such as regression [14], for instance, make several assumptions about underlying project data (completeness, normal distribution, etc.) that are rarely met in the software domain. Non-parametric methods originating from the machine learning domain, such as artificial neural networks (ANN) [3] or Decision Trees/rules [15], make practically no assumptions about the data but are quite sensitive to their parameter configuration and there is usually little universal guidance regarding how to set those parameters. Thus, finding appropriate parameter values requires some preliminary experimentation.

In contrast to data-intensive methods, *expert-based estimation* does not require any project measurement data because estimates are based on the judgment of one or more human experts [2]. Expert estimation is, in fact, commonly used in the software industry (including IV&V). It does, however, have several significant limitations. First, much effort is required each time estimation is performed, and the reliability of the outputs it provides largely depend on the expertise and individual preferences of the human experts involved. Moreover, since the rationale underlying final estimates is not modeled explicitly, there is hardly any support for effective decision making in a project (risk management, process improvement, project scope negotiations, etc.). Recently, a few hybrid methods have been proposed to cope with deficits of data-intensive and expert-based estimation. They combine a reduced amount of both measurement data and human expertise to provide more reliable estimates with limited estimation overhead. Empirical applications [17][10] report on their higher estimation accuracy and stability when compared to data- or expert-based methods. Moreover, methods that employ explicit causal effort modeling (e.g., CoBRA® [17]) have proven to greatly contribute to the achievement of a variety of organizational objectives, such as risk management or process/productivity improvement.

4. The CoBRA® Method

CoBRA® [12][15] is a hybrid method combining data- and expert-based effort estimation approaches. CoBRA® the method is based on the idea that project effort consists of two basic components: nominal project effort and an effort overhead portion as presented below.

Nominal effort is the effort spent only on developing a software product of a certain size in the context of a hypothetical "ideal" project that runs under optimal conditions; i.e., all project characteristics are the best possible ones ("perfect") at the start of the project. *Effort overhead* is the additional effort spent on overcoming the imperfections of a real project environment, such as insufficient skills of the project team. In this case, a certain effort is required to compensate for such a situation, e.g., team training has to be conducted. In CoBRA®, effort overhead is modeled by a causal effort model that consists of factors affecting project effort within a certain context. The causal model is obtained through <u>expert knowledge acquisition</u> (e.g., involving experienced project managers).

$$\text{Effort} = \underbrace{\text{Nominal Productivity} \cdot \text{Size}}_{\text{Nominal Effort}} + \text{Effort Overhead} \quad (1)$$

$$\text{Effort Overhead} = \sum_i \text{Multiplier}_i(\text{Effort Factor}_i) + \sum_i \sum_j \mathit{Multiplier}_{ij}(\text{Effort Factor}_i, \text{Indirect Effort Factor}_j) \quad (2)$$

An example is presented in Figure 1. The solid and dashed arrows indicate direct and indirect relationships, respectively. For instance, *Requirements volatility* has a direct impact on development effort. The strength of this negative influence on effort may, however, be modified (compensated) by *Disciplined requirement management* (indirect influence). The effort overhead portion resulting from indirect influences is represented by the second component of the sum shown in (2).

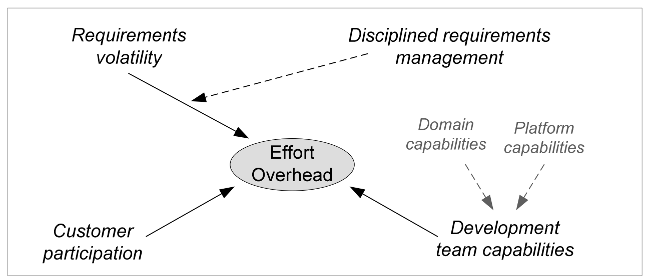

Figure 1: Example of a Causal Effort Model

The influence on effort and between different factors is quantified for each factor using experts' evaluation. The influence is measured by means of effort overhead, i.e., a relative percentage increase of the effort above the nominal project. In order to capture the uncertainty of evaluations, experts are asked to give three values: the maximal, minimal, and most likely cost overhead for each factor (triangular distribution).

The second component of CoBRA®, the nominal project effort, is based on data from past projects that are similar with respect to certain characteristics (e.g., development type, life cycle type) that are not part of the causal model. These characteristics define the context of the project. Past project data is used to determine the relationship between cost overhead and costs (see equation #1). Since it is a simple bivariate dependency, it does not require much measurement data. In principle, merely project size and effort are required, whereby both can be measures using any valid metric representing project size and effort.

Based on the quantified causal model, past project data, and current project characteristics, an effort overhead model is generated using a simulation algorithm (e.g., Monte Carlo). The probability distribution obtained could be used further to support various project management activities, such as effort estimation, evaluation of effort-related project risks, or benchmarking. More details regarding the CoBRA® method can be found in [12][15].

5. Case Study

5.1. Context of the Study

The study was performed in the context of JAMSS, a company that performs IV&V of space software systems (embedded software domain). JAMSS has been, for instance, supporting IV&V for critical space software systems created by the Japan Aerospace Exploration Agency (JAXA) for more than 10 years.

In this study, we focused on IV&V of the software requirements specification documents using the *document review* technique. The document review process starts with a risk analysis to identify a software system's operational risks. Software requirements are then reviewed in more detail based on their operational risks with respect to one or more review objectives. In principle, there were six review objectives [8] (Table 2): (O1) risk analysis, (O2) state transition completeness and consistency, (O3) design completeness for exceptional behavior, (O4) timing correctness and consistency, (O5) interface correctness and consistency, and (O6) traceability.

There were three ***domain experts*** involved in the study (Table 1) who provided their knowledge to build the effort overhead model. The main fields of expertise covered by involved experts included: software product quality & safety assurance (SPQSA), software safety reviews (SR), and safety assurance in operation (SAO).

Table 1. Involved domain experts

Expert	Expertise	Domain experience [#years]	Estimation experience [#projects]
1	SR	7	8
2	SPQSA	8	9
3	SAO	4	6

As project measurement data, the number of document pages was selected as the size of a software requirement because even if the complexity of a requirement complexity is related to the effort for reviewing a document, the document itself has to be read by the IV&V team in order to find out what this complexity is.

IV&V effort data from five projects were collected for each project. In practice, because some IV&V objectives were not addressed, effort data were not collected for each IV&V objective except for one project. Therefore, weekly working statuses of IV&V were used to abstract the effort for each IV&V objective. ***Measurement data*** available for the estimation included size and effort. Size was measured in pages of software requirements for objectives O1 to O5 and system specification (software and hardware) for objective O6 additionally. The effort was measured in person-days (PD).

Table 2. Review objectives considered in the study

Id	Objective	#projects
O1	Risk analysis	5
O2	State transition completeness/consistency	5
O3	Design completeness	5
O4	Interface completeness/consistency	4
O5	Timing consistency/correctness	3
O6	Traceability with correctness	5

5.2. Study Objectives

The objective of the study was to validate accuracy and precision of CoBRA in the context of JAMSS IV&V (compared to expert judgment and Ordinary Least Squares method) and its contribution to the achievement of defined organizational objectives (Section 2.2).

5.3. Study Design

5.3.1. Effort Estimation Procedure

Motivated by its numerous benefits, the CoBRA® Method was proposed as best fitting the effort estimation capabilities and objectives of JAMSS. First of all, CoBRA® proposes a systematic way to build an explicit and reusable effort model based on both implicit knowledge of domain experts and sparse measurement data. Moreover, it provides on the output a transparent and intuitive model of causal effort dependencies specific for the context where it was applied. The first step of the effort estimation procedure included development of the CoBRA® model using the knowledge of the involved domain experts and measurement data (size and effort) from already completed (historical) projects. For each of the six IV&V objectives specified in the study (Table 2), a separate CoBRA® model was developed. After the CoBRA® models had been created, each was validated on the historical data in a leave-one-out cross-validation experiment.

5.3.2. Study Hypotheses

In order to effectively support achievement of the estimation objectives, the outputs of CoBRA® need to be reliable. In our study, we evaluate reliability byvalidating the predictive performance of the estimation outputs, measured in terms of predictive accuracy and precision. We expect that CoBRA® will outperform the currently employed expert-based estimation as well as the Ordinary Least Squares method (OLS), one of a few data-driven methods that are applicable in the study context (due to very sparse measurement data). This leads us to two study hypotheses:

H1. *CoBRA® provides more accurate and more precise estimates than estimation based on expert judgment.*

H2. *CoBRA® provides more accurate and more precise estimates than estimation based on OLS.*

5.3.3. Evaluation of Estimation Performance

The effort models created in the study effort models were evaluated with respect to their predictive performance. We define *predictive performance* as the ability of the effort model to provide accurate and precise estimates. *Estimation accuracy* refers to the nearness of an estimate ($\hat{E}$) to the true value (E). In order to remain comparable to other estimation studies, we use common estimation error measures and accuracy measures [4], such as *relative error* (*RE* in equation #3) and *mean magnitude of relative error* (*MMRE*).

$$RE_i = \left(\hat{E}_i - E_i\right) / E_i \quad (3)$$

The Conte's RE and MRE measures are the subject of common criticism in the software research community [6]. One of the alternative measures of estimation error proposed is the so-called *z* measure (equation #4) [6]. It quantifies the ratio of the estimate to the actual value

$$z_i = E_i / \hat{E}_i \quad (4)$$

Estimation precision refers to the degree to which several estimates are very close to each other (i.e., the scatter in the data). For the purpose of comparability to other studies, we adopt the *Pred.m* measure. The *Pred.m* measures the percentage of estimates that are within *m%* of MRE [4]. In our study, we use *m = 25%* as typically employed in software estimation studies. Moreover, we adopt *relative standard deviation* (*RSD*) (5) proposed by [6] for software effort estimation as uncorrelated with size (S_i) (which is a weakness of classical standard deviation measures).

$$RSD = \sqrt{\frac{\sum_{i=1}^{n}\left[\left(E_i - \hat{E}_i\right)/S_i\right]^2}{n-1}} \quad (5)$$

5.4. Study Execution

During the study execution, six CoBRA® models were created for each of the IV&V objectives.For each model, domain experts identified several factors (Table 3) that are responsible for the variance of IV&V efficiency for a certain objective (Figure 2).

5.5. Results and Interpretation

This section presents the results of the empirical study. Table 4 and Table 5 present the aggregated measures of the predictive performance. In order to test the significance of the observed effects, appropriate statistical test were performed [13] (at α = 0.05). The results of a *Shapiro-Wilk W* test indicated that the MRE and z results come from normal population; in that case, a parametric *Paired T-test* for homogeneity of means was used. Since the RSD data violated the normality assumption, a non-parametric *Wilcoxon Matched-Pairs Signed Rank* test was used. Note that expert estimates were available for a subset of the past projects considered (indicated as *n* in Table 4 and Table 5). Finally, as we were afraid that for such a small data sample, statistical tests would not have enough power (β-1 ≥ 80%), we performed a power analysis.

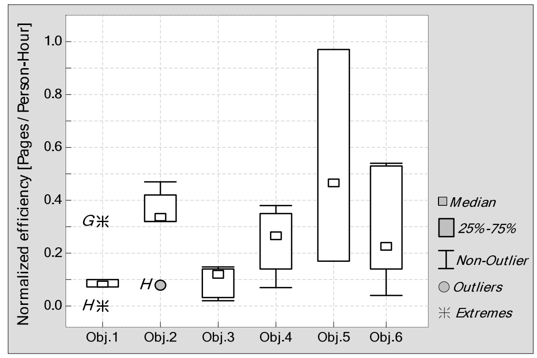

Figure 2. Overall efficiency of IV&V

Table 3. Effort factors considered in the study

Factors influencing IV&V efficiency	Objectives
Domain experience of the IV&V team	O1-O3
Requirements volatility allowed within an initial contract	O2-O4, O6
Novelty of applied IV&V technique	O3
Number of system's interfaces to other (sub)systems	O4, O6
Time pressure in the last IV&V phase	O5
Level of risk assessment done by a supplier or customer	O1
Fault Tree Analysis done by IV&V company	O1
Timing consistency objective included in IV&V	O5
Field Programmable Gate Array (FPGA) review performed	O5
New (inexperienced) personnel involved in IV&V	O1

5.5.1. Hypothesis H1

The results of the empirical investigation (Table 4) suggest that, in principle, CoBRA® provides noticeably more accurate estimates than either expert judgment or OLS for all considered IV&V review objectives (a few exceptions are marked in gray). For example, it improves MMRE by 70% and 40%, on average, compared to OLS and domain experts, respectively. Only few of the obtained results are statistically significant at the chosen α level (marked in bold). As expected, in most of the cases, statistical significance testing did not provide meaningful results (β-1 << 80%). The only powerful results are marked in italics. Summarizing, we conclude that hypothesis ***H1 is valid***.

Table 4. Effort estimation accuracy

	CoBRA		OLS		Expert		
Obj.	***MMRE***	***Mean z***	***MMRE***	***Mean z***	***MMRE***	***Mean z***	***n***
O1	18.2%	97.9%	60.0%	44.1%	42.7%	57.3%	2
O2	25.4%	96.6%	34.0%	66.0%	37.1%	116.5%	4
O3	22.4%	101.%	32.1%	73.0%	36.0%	64.0%	3
O4	***24.1%***	98.6%	33.2%	66.8%	***44.4%***	55.6%	3
O5	39.6%	105.8%	46.3%	63.0%	72.2%	94.4%	3
O6	24.5%	93.1%	44.5%	55.5%	13.8%	88.8%	2

5.5.2. Hypothesis H2

The analysis of estimation precision (Table 5) suggests that, in principal, CoBRA® noticeably outperforms both expert judgment and OLS for all considered IV&V review objectives (a few exceptions are marked in gray). For example, it reduces RSD by 54% and 40%, on average compared to OLS and domain experts, respectively. Similar to accuracy, in most of the cases, statistical significance testing did not provide meaningful results (β-1 << 80%). Summarizing, we conclude that hypothesis ***H2 is valid***.

Table 5. Effort estimation precision

	CoBRA		OLS		Expert		
Obj.	***Pred.25***	***RSD***	***Pred.25***	***RSD***	***Pred.25***	***RSD***	***n***
O1	80.0%	**26.3%**	20.0%	**74.0%**	00.0%	71.3%	2
O2	60.0%	10.7%	60.0%	17.6%	25.5%	12.6%	4
O3	60.0%	15.1%	60.0%	42.1%	33.3%	22.6%	3
O4	***50.0%***	09.9%	50.0%	20.6%	00.0%	***17.3%***	3
O5	00.0%	07.7%	33.3%	12.5%	33.3%	15.0%	3
O6	80.0%	***07.8%***	40.0%	***28.1%***	100.0%	01.9%	2

5.6. Threats to Validity

Several threats to the validity of the presented case study were identified. First, the very sparse project measurement data available prevented us from achieving sufficient power of performed statistical tests. Moreover, expert estimates used to compare CoBRA®'s performance were available only for some of the past projects considered in the study. Finally, the conclusions drawn in the study are limited to the specific context of IV&V reviews at JAMSS. Generalization of the study findings requires further replications.

6. Lessons Learned

The following practical lessons were learned while applying the CoBRA® method for estimation effort of IV&V:

(**LL1**) *Effort estimation scope*: Since IV&V activities differ depending on the objective of IV&V, the scope of effort estimation (the context for which an effort model is built) should be limited to a single IV&V objective. Total effort is the sum of effort over all objectives.

(**LL2**) *Size and complexity of review*: The complexity of a document under review should be considered as an effort driver beyond simple size measures, such as number of document pages.

(**LL3**) *Effort drivers*: Considering effort drivers other than size is a very important aspect of effort modeling. We experienced that a single factor may multiply effort by as much as 10 times (e.g., a complete lack of risk assessment already done by a software supplier may increase the effort of independent risk analysis by up to 20 times). Such an effect is impossible to investigate based only on historical size and effort data.

7. Summary

In this paper, we proposed adapting the CoBRA® software estimation method to predict the effort of independent verification and validation (IV&V). The method provides a potential solution to estimation problems in the context of IV&V. By integrating data- and expert-based estimation, CoBRA® requires minimal amounts of project measurement data and reduces the involvement of domain experts. As a result, it provides a reusable model that supports strategic project/process objectives, such as risk management for effort overrun for each IV&V objectives. At the same time, as reported by several empirical studies, it provides accurate and precise estimates.

When applied in the context of an example IV&V organization, CoBRA® proved to provide more reliable estimates than both the expert-based estimation currently applied and ordinary regression (OLS) - one of few data-intensive methods applicable in the IV&V context. It improved the accuracy and precision of estimates by 40%, on average. At the same time the method provided a transparent, context-specific effort model that supported IV&V practitioners in achieving project and process management objectives (e.g., negotiating project scope or improving the effectiveness of IV&V activities).

References

[1] B.W. Boehm, C. Abts, A.W. Brown, S. Chulani, B.K. Clark, E. Horowitz, R. Madachy, D. Refer, and Steece B. *Software Cost Estimation with COCOMO II*, Prentice Hall, 2000.

[2] B.W. Boehm, *Software Engineering Economics*, Prentice-Hall, 1981.

[3] G. Boetticher, “An Assessment of Metric Contribution in the Construction of a Neural Network-Based Effort Estimator”, *Proc. Int'l Workshop Soft Computing Applied to Soft. Eng*., 2001, pp. 59-65.

[4] S.D. Conte, H.E. Dunsmore, V.Y. Shen, *Software Engineering Metrics and Models*. The Benjamin-Cummings Publishing Company, Inc. 1986.

[5] R.N. Charette, ”Why Software Fails [Software Failure],” *IEEE Spectrum*, vol. 32, no. 9, Sept. 2005, pp. 42-49.

[6] T. Foss, E. Stensrud, B. Kitchenham, I. Myrtveit, ”*A simulation study of the model evaluation criterion MMRE*,” IEEE Trans. Soft. Eng., vol. 29, no. 11, Nov. 2003, pp. 985-995.

[7] *IEEE Std. 1012-2004, IEEE Standard for Software Verification and Validation*, IEEE computer Society, June, 2005.

[8] N. Kohtake, A. Katoh, N. Ishihama, Y. Miyamoto, T. Kawasaki, M Katahira, “Software Independent Verification and Validation for Spacecraft” *Proc. IEEE JAXA Aerospace Conference*, 2008.

[9] M. Lother, R. Dumke, “Point Metrics. Comparison and Analysis,” in *Current Trends in Software Measurement* (ed. R. Dumke, A. Abran), Shaker Publ., 2001, pp. 228-267.

[10] E. Mendes, ”A Comparison of Techniques for Web Effort Estimation,” *Proc. Int'l Symp. Emp. Soft. Eng. & Meas*., Madrid, Spain, 20-21 Sept., 2007, pp. 334-343.

[11] T. Menzies, J. Hihn, ”Evidence-based Cost Estimation for Better Quality Software,” *IEEE Software*, vol. 23, no. 4, 2006, pp. 4-6.

[12] M. Ruhe, R. Jeffery, I. Wieczorek, “Cost Estimation for Web Applications,” *Proc. Int'l Conf. Soft. Eng*., 2003, pp. 285-294.

[13] D.J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedure*, (3rd ed.), Chapman & Hall/CRC; 2003.

[14] P. Sentas, L. Angelis, I. Stamelos, G.L. Bleris,. ”Software Productivity and Effort Prediction with Ordinal Regression,” J. Inf. & Soft. Tech., vol. 47, no. 1, Jan. 2005, pp. 17-29.

[15] Q. Song, M. Shepperd, M. Cartwright, C. Mair, ”Software Defect Association Mining and Defect Correction Effort Prediction,” *IEEE Trans. Soft. Eng*., vol. 32, no. 2, 2006, pp. 69-82.

[16] A. Trendowicz, “Software Effort Estimation – Overview of Current Industrial Practices and Existing Methods,” Tech. Rep. 06.08/E, Fraunhofer IESE, Kaiserslautern, Germany, 2008.

[17] A. Trendowicz, J. Heidrich, J. Münch, Y. Ishigai, K. Yokoyama, N. Kikuchi, “Development of a Hybrid Cost Estimation Model in an Iterative Manner,” *Proc. Int'l Conf. Soft. Eng.*, 2003.