# A REVIEW OF ESTIMATION OF SOFTWARE PRODUCTS DEVELOPMENT COSTS

**Osmanbegović, Edin; Suljić, Mirza; Agić, Haris**

*Source / Izvornik:* **Ekonomski vjesnik : Review of Contemporary Entrepreneurship, Business, and Economic Issues, 2017, 30, 195 - 207**

**Journal article, Published version**
**Rad u časopisu, Objavljena verzija rada (izdavačev PDF)**

*Permanent link / Trajna poveznica:* https://urn.nsk.hr/urn:nbn:hr:145:068645

*Rights / Prava:* Attribution-NonCommercial-NoDerivatives 4.0 International/Imenovanje-Nekomercijalno-Bez prerada 4.0 međunarodna

*Download date / Datum preuzimanja:* **2024-03-28**




*Repository / Repozitorij:*

EFOS REPOSITORY - Repository of the Faculty of Economics in Osijek






***Edin Osmanbegović***
*University of Tuzla*
*Faculty of Economics*
*Univerzitetska 8,*
*75000 Tuzla,*
*Bosnia and Herzegovina*
*edin.osmanbegovic@untz.ba*
*Phone: +38761101741*

***Mirza Suljić***
*University of Tuzla*
*Faculty of Economics*
*Univerzitetska 8,*
*75000 Tuzla,*
*Bosnia and Herzegovina*
*mirza.suljic@kreka.ba*
*Phone: +38761672907*

***Haris Agić***
*Pedagogical Institute of Tuzla Canton*
*Bosne Srebrene 119,*
*75000 Tuzla,*
*Bosnia and Herzegovina*
*agich59@hotmail.com*
*Phone: +38761109899*

***UDK: 004: 338.512***
***Review article***

*Received: March 10, 2016*
*Accepted for publishing: December 2, 2016*

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

# A REVIEW OF ESTIMATION OF SOFTWARE PRODUCTS DEVELOPMENT COSTS

**Abstract**

In the modern business and management of business processes, the standardization of procedures allows the creation of added value, increasing competitiveness and success in the business of an organization. Evaluation of the budget for software development is crucial to the success of an IT project, because the inability to make a realistic assessment leads to inadequate project plans, customer dissatisfaction, poor quality of software products, and reduced profits. In order to minimize such situations, making accurate and reliable software cost estimation should be carried out at all stages of the project life cycle. Although hundreds of research articles focusing on the application of different methods of budget estimates of the software product have been published so far, there is no comprehensive review of the current situation or review of research trends in the budget estimates of the software product. This paper aims to create a framework for estimation of costs of development of software products by providing an overview of the most influential researchers, the most influential articles published in the WoS database, the most used keywords for searching the articles, as well as a review of the estimation techniques used in budget estimates of the software product.

**Keywords:** Software cost estimation, software cost prediction, keyword analysis, estimation techniques

## *1. Introduction*

Regardless of decades of research, most IT projects still fail (Standing et al., 2006). At the same time, a high pace of technological change leads to a fact that modern organizations must be competitive in the harsh environment. Actually, they must respond quickly to problems and opportunities arising from the very dynamic conditions. As information technology becomes the main factor determining the survival of most organizations, it is in their interest to make their IT projects more successful. Software products must be delivered in a time frame within the predetermined budget, and meet the real needs of the client. Therefore, the assessment of the budget and its monitoring during the development of software products are key aspects for the success of an IT project. As the estimates are based on incomplete, imperfect knowledge and assumptions about the future, many estimates of software costs tend to be too low due to omissions of important product functions and project activities (Galorath, Evans, 2006).

Everyday practice shows that many IT organizations still propose unrealistic software costs, work within tight schedules, and finish their projects behind schedule and over budget. Therefore, the budget estimates of software development are the complex and challenging task which requires knowledge of the different parameters of the project for which the assessment is done. Parameters primarily include knowledge of the market, the effort the staff has to make, and characteristics of the product that is under development. The most important software product characteristics are size (Jiang et al., 2007), complexity and quality of the product, selected technology (Atkins et al., 2000) and applied development methodology. Therefore, planning the development of the software product needs to take into account all these factors, so that the outcome of the planning is as close as possible to the real processes. An approximate estimate is significant for the following reasons:

- Determining the priorities over the realization of planning activities within the organization,
- Determining required resources for creation of the project and the way they are going to be used,
- Matching estimated costs with the real price of the software product.

The effectiveness of budget estimates of software products is of crucial importance when it comes to an early indication of the project cost. If the estimate is lower than the real price, it may result in stepping out of the budget frame, or with incomplete functionality and the poor quality of the final product.

If the estimate is above the real cost it may result with too many initial resources involved in the project or it may result with the higher price of the contractual tender offer and eventual job loss. To avoid such situations as much as possible, it is especially important to make an accurate and reliable evaluation of software cost estimation in the early stadium of the project life cycle (Jørgensen, Halkjelsvik, 2010). Therefore, an accurate assessment of the budget in the process of the development of the software product is important and involves determining the effort, duration and cost of the project (in local currency). Although the cost, effort and duration of the project are closely related, they are not always related with a simple transactional function.

In the published literature on software development cost estimation, the research community has proposed various models. Within this context, (Jørgensen, Shepperd, 2007) a systematic review has been conducted, and 304 software cost estimation papers have been identified and the papers have been classified according to the research topic, estimation approach, research approach, study context and data set. The mentioned authors found that the most commonly used research topic was the introduction and evaluation of estimation methods, used in 61% of the surveyed papers. Also, some other researches have proposed the use of machine learning based methods, in order to achieve high prediction accuracy. In 2012, (Wen et al., 2012) a systematic review was conducted in which eight ML based methods were identified, and showed that CBR and ANN were the most widely used ML methods in the last two decades (1991–2010). Idri (Idri et al., 2015) has conducted a systematic map and review of analogy-based software estimation techniques and concluded that use of analogy-based models by practitioners is still limited. Researches also concluded that analogy-based models can be applied at an early stage of a software project and can mitigate problems with outliers. In recent years, in order to investigate models to estimate a software project's size, effort, duration and cost, many researchers have commonly used the ISBSG dataset. Moreover, (González-Ladrón-de-Guevara et al., 2016) a systematic mapping study has been conducted to determine which variables in the ISBSG dataset have been used in software engineering to build estimation models, and to what extent the variables were used. Considering the significant number of papers it can be noticed that the research area of software development cost estimation is an area with very active scientific work.

This paper reviews the literature on the area of assessment of the cost of software products published in the WoS database for the period from 1987 to 2015. The purpose of this paper is to meet the following two objectives. Firstly, to develop a methodological framework for conducting a comprehensive literature review based on our own experiences of conducting this scope of the study. Secondly, to use this framework to obtain an understanding of the current state of the software cost estimation (SCE) research area. Also, the intention is to be useful for academic and industrial communities, because it will provide an overview of the most influential researchers, journals, methods and used keywords.

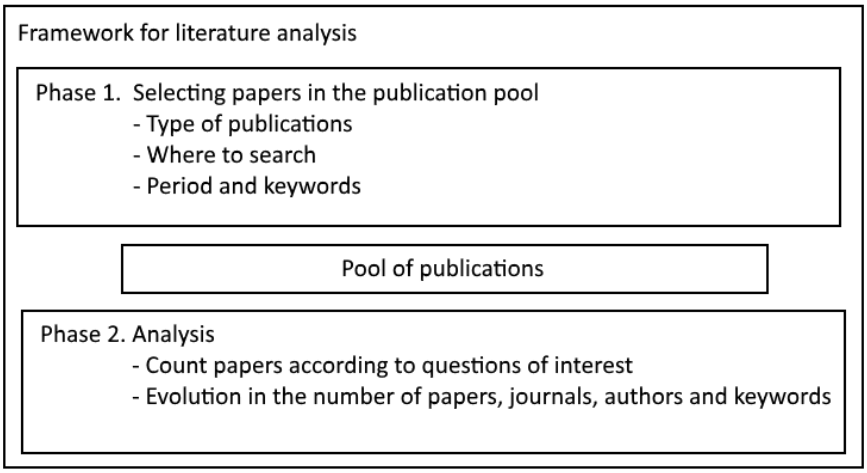
## 2. Methodology

This section describes the methodology for selection, collection and analysis of papers used in this report. It's pretty hard to limit a report to a specific area of software cost estimation because the relevant materials have been published in various journals. In order to identify works and papers that are related to the software cost estimation, the WoS databases have been searched. The search of the WoS database (see Table 1) was carried out for the entire period, limited to the title, abstract and keywords. In order to avoid constant repetition of the audit report, the November 1st 2015 has been selected as a final date. The following research questions were defined:

- Who are the most influential researchers in the research area of software cost estimation?
- Which scientific publications are relevant to the research area of software cost estimation?
- What are the methods most commonly used in the assessment of software cost estimation research area?
- What are the most commonly used keywords in the assessment of software cost estimation research area?

The methodology used for conducting a literature review, making it possible to answer the previous questions, will be presented below. To be able to get an overview of the research and answer the previous questions, an appropriate research methodology is required. The methodology is divided into two phases (see Figure 1): phase 1 – selection and accumulation of a journal publication of the WoS database in the publication pool, and phase 2 – identification of the publications by key researchers, journals, methods and keywords. In the following paragraphs each of the two phases will be discussed.

***Figure 1 Framework for literature analysis***




*Source: Created by the authors*

### *Phase 1: Choice and collection of a journal publication*

Regarding the literature review, it is particularly important to define clear boundaries to delimitate the research. The first decisions were adopted in order to identify the most relevant sources of information, taking into account research questions of interest. Additionally, it was necessary to decide in which period, which keywords and search terms were important, as well as their permutations/combinations to use in the search for relevant publications. Only journal publications were chosen for the review. There are three main reasons for focusing on journal papers:

(1) Journal papers contain more up-to-date data than books (Dale et al., 2001),

(2) A book is likely to duplicate material published previously in a paper (Dale et al., 2001) and

(2) Conference papers can be difficult to access.

***Table 1 Search query and additional options used while searching the WoS database***

| Search query | Additional search options |
|---|---|
| TS = (Software OR Project) AND TS = ("Effort Estimation" OR "Cost Estimation" OR "Effort Prediction" OR "Cost Prediction")<br>**Refined by: Databases:** (WOS OR CCC)<br>*Time span = All years*<br>*Search language = Auto* | Advanced search:<br>Release date: 1981 to January 2016<br>Search area: title; abstract; author keyword; keywords plus® |

*Source: Created by the authors*

For the search, journal publications in the Web of Science (WoS) databases were chosen. This gave access to a very high number of top journal publications. The search for related publications was carried out for the entire period and focused on the title, abstract and keywords. In Table 1 it is shown that a query with additional possibilities for the search has been applied within the WoS database.

This survey obtained 879 papers, of which 94.77% were categorized as journal publications, but not all of them were relevant for the research. The main criteria for inclusion/exclusion of a publication into a further analysis were the focus of the studies, which should be within the area of software cost estimation. The inclusion/exclusion criteria were applied case-by-case to the 1064 publications, by firstly reading the title and the abstract of each article. In order to answer the research questions, 210 papers have been finally selected for the literature analysis.

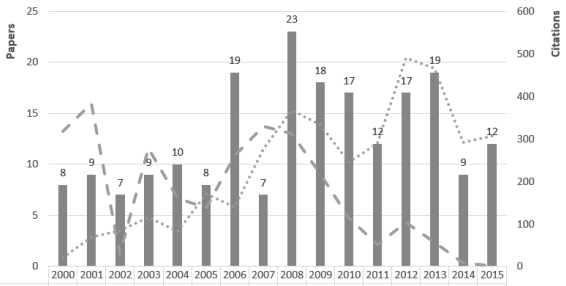
### *Phase 2: Analysis*

In this phase, as mentioned previously, the focus was on disclosing how the area had evolved during the period, which journals had published "Software Cost Estimation" papers, and finally which authors had contributed to the area. The questions were as follows:

- What is the distribution of publications across the time period?
- In which journals were Software Cost Estimation articles published?
- Which authors published Software Cost Estimation articles?
- Make a list of the keywords that were used in the published papers.

To be able to address the area of interest, according to the taxonomy, the descriptive dimensions were used to classify the papers. The basic pool of publication comprises of 229 papers. The distribution of the publications in the researched period (2000–2015) is shown in Figure 2 and total citations in journals for the same period.

***Figure 2 Distribution of publications per year across the period 2000-2015 for research area of software cost/effort product***




| | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Papers | 8 | 9 | 7 | 9 | 10 | 8 | 19 | 7 | 23 | 18 | 17 | 12 | 17 | 19 | 9 | 12 |
| Citations for papers published in the year | 318 | 383 | 29 | 277 | 161 | 138 | 260 | 330 | 312 | 217 | 114 | 52 | 103 | 56 | 8 | 2 |
| Citations in the year | 21 | 69 | 85 | 115 | 82 | 171 | 140 | 276 | 367 | 333 | 245 | 290 | 489 | 466 | 292 | 306 |

*Source: Authors' calculation from the WoS database*

While 1955 was the first year of publication where works were sought, the first published papers were from the year 1981. High numbers of publications were found in the period between 2006 and 2013, but it remained on a substantial level, which indicated a stable level of activity in the estimation of the cost/effort of the software product.

## 3. Results on the software project cost

### 3.1 Results on the influential researchers, papers, journals and keywords on software product cost estimation

#### Identification of the most influential researches

Based on the selected sample, 448 different authors have been identified who have published papers focusing on the assessment rates in the process of software product development. In order to identify the most influential researchers in the research area of software cost estimation, authors were classified as researchers with long-term interest in the research field (min. 5 years), with the number of papers in the mentioned period (min. 5 papers) and number of citations (min. 50). The analysis shows (see Table 2) that only 3 researchers met the previously mentioned criteria. Therefore, Lefteris Angelis, Witold Pedrycz and Ioannis Stamelos have been identified as the most influential researchers in the area of software cost estimation. Besides the previously mentioned researchers, Magne Jørgensen and Martin Shepperd have been identified as researchers with significant impact, although their primary focus was on software effort estimation (see Table 3), rather than software cost estimation.

***Table 2 The most influential software cost estimation researchers***

| | RESEARCHER | LONG-TERM INTEREST | | NUMBER OF PUBLICATIONS | | NUMBER OF CITATIONS |
|---|---|---|---|---|---|---|
| | | PERIOD | YEAR | WoS | DIFFERENT JOURNALS | WoS |
| Cost estimation | Angelis, Lefteris<br>Aristotle University of Thessaloniki | 2001-2015 | 14 | 14 | 4 | 155 |
| | Pedrycz, Witold<br>University of Alberta | 2002-2012 | 10 | 7 | 7 | 52 |
| | Stamelos, Ioannis<br>Aristotle University of Thessaloniki | 2001-2010 | 9 | 6 | 3 | 106 |
| | Mittas, Nikolaos<br>Technological Educational Institute of Kavala | 2008-2015 | 7 | 8 | 4 | 46 |
| | Menzies, Tim<br>North Carolina State University | 2010-2013 | 3 | 8 | 4 | 47 |

*Source: Created by the authors from the WoS database*

***Table 3 The most influential software effort estimation researchers***

| | RESEARCHER | LONG-TERM INTEREST | | NUMBER OF PUBLICATIONS | | NUMBER OF CITATIONS |
|---|---|---|---|---|---|---|
| | | PERIOD | YEAR | WoS | DIFFERENT JOURNALS | WoS |
| Effort estimation | Jørgensen, Magne<br>University of Oslo | 2003-2010 | 7 | 9 | 6 | 368 |
| | Shepperd, Martin<br>Brunel University | 2001-2011 | 10 | 8 | 7 | 326 |
| | Sicilia, Miguel-Angel<br>University of Alcalá | 2005-2008 | 3 | 7 | 6 | 35 |
| | Chiu, Nan-Hsing | 2006-2009 | 3 | 6 | 5 | 178 |
| | Kocaguneli, Ekrem<br>Microsoft | 2012-2013 | 2 | 6 | 4 | 36 |

*Source: Created by the authors from the WoS database*

In order to identify researchers who have had the greatest impact on publications published in the researched period a co-citation network of authors has been created.

***Figure 3 Researcher co-citation network***




*Source: Created by the authors using the VOSviewer*

### *Identification of the most influential journals*

In total, the papers have been published in 71 scientific journals in the area of software cost/effort estimation. In order to identify the most influential scientific journals within the researched area of software cost estimation, journals have been defined as journals with long-term interest in the area (min. 10 years), the number of papers in the mentioned period (min. 10 papers), and the number of citations (min. 100). The first four journals in Table 4 met the required criteria, with software cost estimation as the primary focus (90%) of scientific articles published in the magazine "Software Quality Journal".

***Table 4 The most influential software cost estimation journals***

| RANK | JOURNAL | LONG-TERM INTEREST (Year) | NUMBER OF PUBLICATIONS | NUMBER OF CITATIONS | IMPACT FACTOR |
|---|---|---|---|---|---|
| 1 | Information and software technology | 1987-2015 (28) | 38 | 778 | 1.328 |
| 2 | IEEE transactions on software engineering | 1992-2013 (21) | 29 | 1258 | 2.292 |
| 3 | Journal of systems and software | 1991-2015 (24) | 26 | 486 | 1.245 |
| 4 | Empirical software engineering | 2004-2014 (10) | 14 | 138 | 1.640 |
| 5 | Software quality journal | 2002-2011 (9) | 11 | 56 | 0.880 |

*Source: Created by the authors*

When the results of research with the systematic review of the literature (Jørgensen, Shepperd, 2007) in the research area of software cost estimation for the period to 2007 were compared, it can be concluded that the authors of this paper have come to a similar list of the most influential scientific journals.

### *Identification of the most used keywords*

As every scientific journal requires that the abstract is followed by several keywords related to the research, their selection requires great experience and can ensure wide availability of a publication. Keyword analysis shows that 777 unique words and phrases have been used for classification of the software cost estimation publications. The most commonly used keywords and phrases are: software cost estimation, cost estimation, cost estimation models, software development cost and cost estimation, with 145 appearances as well as software effort estimation, effort prediction and effort estimation with 85 appearances.

***Table 5 The most frequently used keywords***

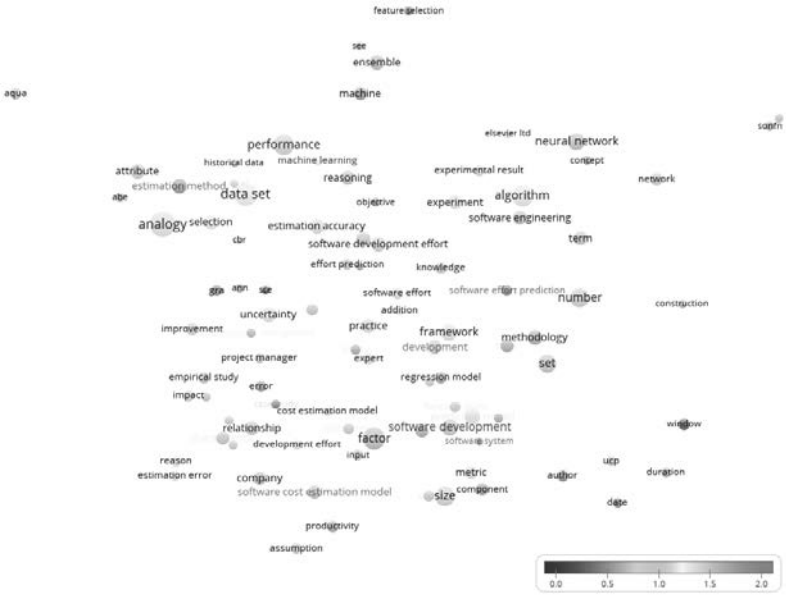
| | Keyword | Frequency |
|---|---|---|
| 1. | Models | 44 |
| 2. | Validation | 42 |
| 3. | Regression | 36 |
| 4. | Prediction | 32 |
| 5. | Software | 28 |
| 6. | Analogy | 22 |
| 7. | Simulation | 21 |
| 8. | Systems | 21 |
| 9. | Function points | 18 |
| 10. | Selection | 16 |
| 11. | Cocomo | 13 |
| 12. | Estimation models | 13 |
| 13. | Machine learning | 13 |
| 14. | Model | 13 |
| 15. | Accuracy | 12 |
| 16. | Estimation | 12 |
| 17. | Regression-models | 12 |
| 18. | Software engineering | 12 |
| 19. | Neural networks | 11 |
| 20. | Imputation | 10 |
| 21. | Management | 10 |
| 22. | Prediction models | 10 |
| 23. | Productivity | 10 |
| 24. | Size | 10 |

*Source: Created by the authors*

Twenty-four (24) keywords that have a frequency higher than 10 are given in Table 5. The results indicate that the focus of research is on different models (models, regression, analogy, function point, cocomo, machine of learning, and neural networks) of software cost/effort estimation.

To visualize a research area, a methodology of the term map was used based on the analysed sample of 229 publications. Using the term map or co-appearance of words has a long history of 30 years and it is a two-dimensional representation of the research area in which the strongly associated words are close to each other, and less associated words are distant from one another. Different areas on the map correspond to different research subarea or areas. Using the techniques of natural language processing, the words of the title and summary of the analysed publications were separated. This was used in order to obtain a list of words that appear in these publications. By applying the algorithm, there was a yield of 155 words or phrases that can be considered the most specific terms of the selected data sets within the research area. Only terms or phrases that co-appear at least ten times were taken into consideration. The location in two-dimensional space was determined by applying VoS mapping techniques in 2D space for each term. The terms which have strong ties tend to be located close to each other, while the terms that do not have strong ties are more distant from each other in two-dimensional space. Each circle in Figure 2 represents a term or phrase (terms with at least 10 co-appearances are shown), while the size of the bubble represents the intensity measured by the number of co-appearances.

***Figure 4 Term map of the software cost/effort estimation area***




*Source: Created by the authors using the VOSviewer*

In the map of terms (Figure 4) colour is used to indicate a difference in the practice of citation of various terms or phrases. The average citation impact of publications in which the term appears in the title or abstract is determined for each term or phrase. The colours range from blue (the average value of 0) to green (with an average value of 1) and red (with an average value of 2 or more). Thus, the blue colour indicates the fact that the publications have a low average citation impact, while the red colour indicates that the publication has a high average citation impact.

### *3.2 Results of the estimation techniques for software product estimation*

Many cost estimation techniques have been proposed over the last 40 years. The key question posed by managers is: „Which method is the best for my business?" The positive fact is that there are a lot of useful estimation methods. Still, the usefulness of the estimation methods primarily depends on the particular use situation and the context of the application. This chapter provides a general overview of software cost estimation methods including the recent advancement in the area. Review of the literature shows that the models for software cost estimation have been divided into two groups (Attarzadeh, Ow, 2014; Boehm, 1981; Khatibi, Jawawi, 2011a):

- Algorithmic and
- Non-algorithmic.

Algorithmic models have nonparametric form with fixed forms of formula, for which the parameters are calibrated based on historical data. The non-algorithmic models that are known for their efficiency in solving complex problems when conventional analytical methods are not profitable, are time consuming, and expensive. Examples of such methods are: fuzzy logic (Pedrycz et al., 1999), neural networks (Idri et al., 2002) and neural-fuzzy models (Banjanović-Mehmedović, 2011).

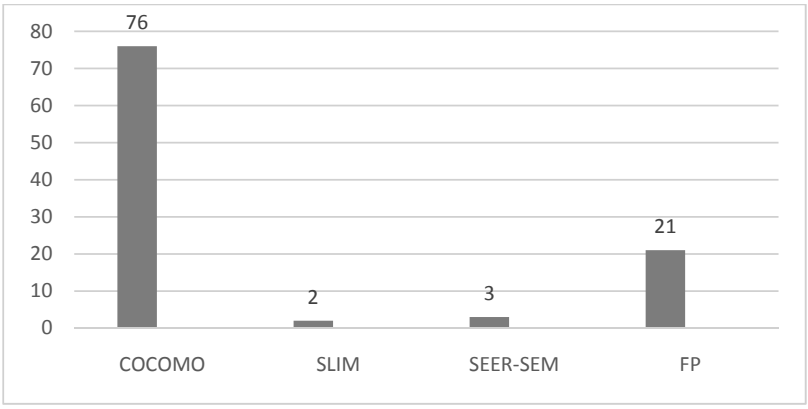
### *3.3 Algorithmic models*

These models are designed to provide a mathematical equation for software cost estimation. These equations are based on previous researches and historical data. They use different methods of measurement as the input, and they are based on the following: the number of lines of code, functional point etc. The valuation models use the general equation:

$$y = f(x_1, x_2, \dots x_n) \quad (1)$$

Where the vectors $(x_1, x_2, \dots x_n)$ represent cost factors of the software product. The main difference among the existing algorithmic methods refers to the selection of functions and the cost factors. All the cost factors which are used in those models can be divided into the following categories: product, hardware, staff and project (Khatibi, Jawawi, 2011b). Quantification of these factors is rather difficult to obtain, because some of them might have been ignored in certain software projects. According to Khatibi and Jawawi (Khatibi, Jawawi, 2011b), some of the most commonly used algorithmic models are: COCOMO (Boehm, 1981), SLIM (Putnam, Myers, 1992), SEER-SEM (Jensen, 1983) and Function Points (Albrecht, 1979). These techniques have several advantages, and the most prominent of which are objectivity, repeatability, the presence of supporting sensitivity analysis, and the ability to calibrate to previous experience. On the other hand, these models also have some disadvantages such as their lack of flexibility in adapting to the new development environment. In Figure 5 a distribution of used algorithmic models is shown. The most commonly used estimation method is the so-called COCOMO. It offers an off the-shelf fixed estimation model, in which all effort relationships are already predefined.

***Figure 5 Distribution of used algorithmic models***




*Source: Created by the authors*

Boehm (1981) developed the first COCOMO model using a multiple regression analysis. The most recent COCOMO II has been designed by calibrating the original model, which was conducted by using measurement data and expert judgment (Boehm et al., 2000). For this purpose, a hybrid approach to model parameters has been used, that has been learned from measurement data using statistical regression provided directly by human experts, and has been integrated by using Bayes' Theorem. Consequently, COCOMO II model requires very specific input data (Boehm et al., 2000).

### *3.4 Non-algorithmic methods*

Unlike the algorithmic methods, the methods from this group are based on an analytical comparison and interference. Using non-algorithmic methods requires information about previous projects similar to the project that has to be evaluated. Three general methods of estimation are given due to the fact that in recent years there have been published many papers on their application. According to Khatibi and Jawawi (Khatibi, Jawawi, 2011b), the most commonly used non-algorithmic models are: analogy (Li et al., 2009; Shepperd, Schofield, 1997; Sternberg, 1977), expert assessment (Khatibi, Jawawi, 2011a) and machine learning methods (Malhotra, 2015).

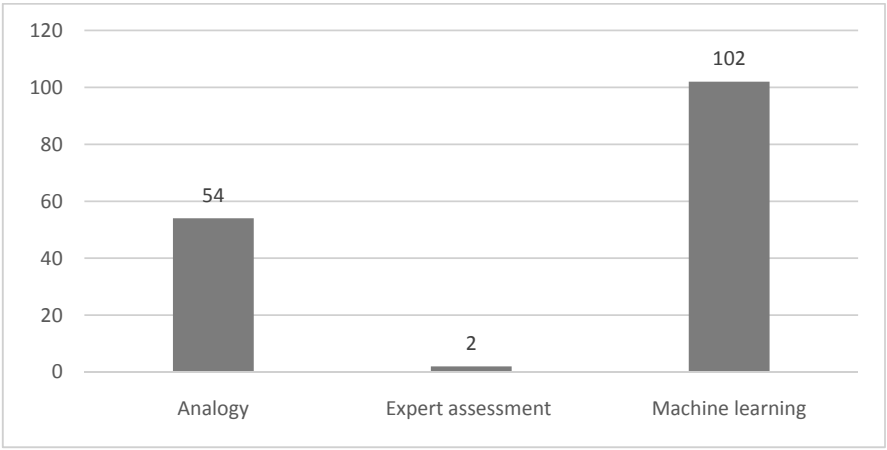
The analogy based software cost estimation, which is essentially a case-based reasoning approach, was first proposed by Sternberg (Sternberg, 1977). The analogy method is based on conclusion from cases, without the use of experts. In general, the basic idea of ABE is simple: when provided a new project for

estimation, compare it with historical projects to retrieve the most similar projects, which are then used to predict the cost of the new project. According to Li (Li et al., 2009) the ABE process system consists of four phases.

Estimation based on expert judgment is still the most popular estimation method in the software industry. This assessment is based on professional judgment, and it is made on the basis of advice received from the experts who have extensive experience on similar projects. This method is usually used when there are limitations in finding data and gathering requirements. The most commonly used methods based on this technique are: Guesstimation, Wideband-Delphi, Estimeeting, Planning Game, Analytic Hierarchy Process, and Stochastic Budget Simulation.

In recent years, machine learning methods have been applied in SCE domains. It is a consequence of the fact that managers had realized the potential of ML techniques and began applying them in the management of software projects. ML methods can be categorized into the following main groups: Fuzzy (Zadeh, 1965), ANN, ANFIS (Buragohain, Mahanta, 2008), GA (Banjanović-Mehmedović, 2011). In Figure 6 is shown a number of papers applied by software cost estimation of non-algorithmic models, in order to obtain a more detailed insight into the non-algorithmic models used in this paper. In Figure 6 is shown a distribution of used non-algorithmic models, which show that machine learning or data mining methods are the most frequently used techniques, followed by analogy.

***Figure 6 Distribution of the use of non-algorithmic models***




*Source: Created by the authors*

The most commonly used non-algorithmic estimation model is machine learning or data mining techniques. Machine learning (ML) methods, with several reported successful applications, have gained popularity in recent years. ML methods are focused on the effects of data pre-processing techniques in the context of software cost estimation (Huang et al., 2015). Papatheocharous and Andreou (Papatheocharous, Andreou, 2012) have addressed the issue of SCE and proposed the use of decision trees enhanced by fuzzy logic as a solution for establishing an accurate cost model. As software becomes more complex, the importance of research on developing machine learning methods for estimating software cost development has perpetually increased. Therefore, there is a need for increasing experimental research in the area of ML and SCE.

## *4. Discussion and conclusion*

In this paper, a comprehensive literature review has been conducted on software cost estimation. To be able to carry out such an assessment, a methodological framework has been developed, which consists of the two phases: phase one (selection and accumulation of articles) and phase two (analysis of the collected articles). Using this framework for conducting a review of the literature provides an

overview of the research area of software cost estimation, regardless of research disciplines, scientific field and research topics. The research shows that articles dealing with software cost estimation have been published in more than 70 journals, and that in the first five journals, regarding the number of published papers, about 35 percent of all papers have been published.

In the last decade, the subject of many researches has been directed towards finding the most important factors for failure of software product development. According to several studies mentioned in this paper, one of the common causes for failures of software projects are the inaccurate budget estimates in the early stages of the project. Thus, the introduction and emphasis on evaluation methods appears to be necessary to achieve accurate and reliable assessment of the software product. In this study the majority of evaluation techniques are presented systematically. Since the managers of software projects are used to selecting the best method of evaluation in accordance with the conditions and status of a project, it seems that the state and practices in the application of evaluation techniques may be useful in reducing the number of project failures.

There is no optimal method of assessment that can be presented as the best in all different situations i.e., every method of evaluation is appropriate for a certain project. In order to choose the best method of evaluation, it is necessary to understand the principles of each of the assessment methods. Each method of estimation depends on several parameters, such as the complexity of the project, its duration, staff expertise, methods of development, and so on. Improving the effectiveness of existing methods and the introduction of new methods for estimation, based on the current requirements of software product development, represents the future trend in this area.

Future research needs to focus on the creation of models for estimating the budget for the development of software products, as well as the analysis of the factors influencing the performance. As a methodological basis there are going to be used methods of artificial intelligence such as neural networks and neuro-fuzzy. Such analysis could lead to the creation of an intelligent system for decision support in the budget estimates for the development of software products, which would contribute to greater success of IT projects and quality of management of such projects as a whole.

## References


1. Albrecht, A. J. (1979), "Measuring application development productivity", Proceedings of the joint SHARE/GUIDE/IBM application development symposium, pp. 83-92.
2. Atkins, D. L., Mockus, A., Siy, H. P. (2000), "Measuring technology effects on software change cost", Bell Labs Technical Journal, Vol. 5, No. 2, pp. 7-18.
3. Attarzadeh, I., Ow, S. H. (2014), "Proposing an Effective Artificial Neural Network Architecture to Improve the Precision of Software Cost Estimation Model", International Journal of Software Engineering and Knowledge Engineering, Vol. 24, No. 6, pp. 935-953.
4. Banjanović-Mehmedović, L. (2011). Inteligentni sistemi. Tuzla: Harfograf Tuzla.
5. Boehm, B. W. (1981). Software engineering economics. Englewood Cliffs: Prentice-Hall.
6. Boehm, B. W., Madachy, R., Steece, B. (2000). Software cost estimation with Cocomo II with Cdrom. Prentice Hall PTR.
7. Buragohain, M., Mahanta, C. (2008), "A novel approach for ANFIS modelling based on full factorial design", Applied Soft Computing, Vol. 8, No. 1, pp. 609-625.
8. Dale, B., Elkjaer, M., Van der Wiele, A., Williams, A. (2001), "Fad, fashion and fit: An examination of quality circles, business process re-engineering and statistical process control", International Journal of Production Economics, Vol. 73, No. 2, pp. 137-152.
9. Galorath, D. D., Evans, M. W. (2006). Software sizing, estimation, and risk management: when performance is measured performance improves. Hoboken: CRC Press.
10. González-Ladrón-de-Guevara, F., Fernández-Diego, M., Lokan, C. (2016), "The usage of ISBSG data fields in software effort estimation: A systematic mapping study", Journal of Systems and Software, Vol. 113, March 2016, pp. 188-215.
11. Huang, J., Li, Y.-F., Xie, M. (2015), "An empirical analysis of data preprocessing for machine learning-based software cost estimation", Information and Software Technology, Vol. 67, November 2015, pp. 108-127.
12. Idri, A., Azzahra Amazal, F., Abran, A. (2015), "Analogy-based software development effort estimation: A systematic mapping and review", Information and Software Technology, Vol. 58, February 2015, pp. 206-230.
13. Idri, A., Khoshgoftaar, T. M., Abran, A. (2002), "Can neural networks be easily interpreted in software cost estimation?", FUZZ-IEEE '02: Proceedings of the 2002 IEEE International Conference on Fuzzy Systems, IEEE, Honolulu, pp. 1162-1167.
14. Jensen, R. (1983), "An improved macrolevel software development resource estimation model", 5th ISPA Conference, pp. 88-92.
15. Jiang, Z., Naudé, P., Jiang, B. (2007), "The effects of software size on development effort and software quality", International Journal of Computer and Information Science and Engineering, Vol. 1, No. 4, pp. 230-234.
16. Jørgensen, M., Halkjelsvik, T. (2010), "The effects of request formats on judgment-based effort estimation", Journal of Systems and Software, Vol. 83, No. 1, pp. 29-36.
17. Jørgensen, M., Shepperd, M. (2007), "A Systematic Review of Software Development Cost Estimation Studies", IEEE Transactions on Software Engineering, Vol. 33, No. 1, pp. 33-53.
18. Khatibi, V., Jawawi, D. N. (2011a), "Software Cost Estimation Methods: A Review", Journal of Emerging Trends in Computing and Information Sciences, Vol. 2, No. 1, pp. 21-29.
19. Khatibi, V., Jawawi, D. N. (2011b), "Software cost estimation methods: A review", Journal of Emerging Trends in Computing and Information Sciences, Vol. 2, No. 1, pp. 21-29.
20. Li, Y.-F., Xie, M., Goh, T. N. (2009), "A study of project selection and feature weighting for analogy based software cost estimation", Journal of Systems and Software, Vol. 82, No. 2, pp. 241-252.

21. Malhotra, R. (2015), "A systematic review of machine learning techniques for software fault prediction", Applied Soft Computing, Vol. 27, February 2015, pp. 504-518.
22. Papatheocharous, E., Andreou, A. S. (2012), "A Hybrid Software Cost Estimation Approach Utilizing Decision Trees and Fuzzy Logic", International Journal of Software Engineering and Knowledge Engineering, Vol. 22, No. 3, pp. 435-465.
23. Pedrycz, W., Peters, J. F., Ramanna, S. (1999), "A fuzzy set approach to cost estimation of software projects", in Meng M. (Ed.), 1999 IEEE Canadian Conference on Electrical and Computer Engineering, pp. 1068-1073.
24. Putnam, L., Myers, W. (1992). Measures for excellence. Yourdon Press Computing Series.
25. Shepperd, M., Schofield, C. (1997), "Estimating software project effort using analogies", IEEE Transactions on Software Engineering, Vol. 23, No. 11, pp. 736-743.
26. Standing, C., Guilfoyle, A., Lin, C., Love, P. E. (2006), "The attribution of success and failure in IT projects", Industrial Management & Data Systems, Vol. 106, No. 8, pp. 1148-1165.
27. Sternberg, R. J. (1977), "Component processes in analogical reasoning", Psychological Review, Vol. 84, No. 4, pp. 353-378.
28. Wen, J. F., Li, S. X., Lin, Z. Y., Hu, Y., Huang, C. Q. (2012), "Systematic literature review of machine learning based software development effort estimation models", Information and Software Technology, Vol. 54, No. 1, pp. 41-59.
29. Zadeh, L. A. (1965), "Fuzzy sets", Information and Control, Vol. 8, No. 3, pp. 338-353.

***Edin Osmanbegović***
***Mirza Suljić***
***Haris Agić***

# LITERALNI PREGLED PROCJENE TROŠKA RAZVOJA PROGRAMSKOG PROIZVODA

**Sažetak**

U suvremenom poslovanju i upravljanju poslovnim procesima standardizacija svih postupaka omogućava stvaranje dodatne vrijednosti, konkurentnosti i uspješnosti u poslovanju organizacije. Procjena budžeta za razvoja programskog proizvoda presudna je za uspjeh IT projekta, jer nemogućnost realne procjene dovodi do loših projektnih planova, nezadovoljstva klijenata, loše kvalitete programskog proizvoda i smanjenog profita. Kako bi takvih situacija bilo što je moguće manje, izradu točne i pouzdane estimacije troškova programskoga proizvoda potrebno je obavljati u svim fazama životnog ciklusa projekta. Iako je, do sada, objavljeno na stotine istraživačkih članka s naglaskom na primjenu različitih metoda procjene budžeta programskog proizvoda, ne postoji istraživanje i pregled trenutnog stanja, istraživačkih trendova u području procjene budžeta programskog proizvoda. Ovaj rad ima za cilj kreirati okvir za procjenu troškova razvoja softverskih proizvoda osiguravajući pregled najznačajnijih istraživača, najutjecajnijih članaka objavljenih u WoS bazi podataka, najčešće korištenih ključnih riječi u pretraživanju članaka, kao i pregled tehnika procjene koje se koriste u procjenama troška softverskog proizvoda.

**Ključne riječi:** procjena troška programskog proizvoda, predikcija troška programskoga proizvoda, analiza ključnih riječi, tehnike procjene