Towards a Decision Support System for Project Planning: Multi-Criteria Evaluation of Past Projects Success

Miha Hafner
Elea iC d.o.o.
Department for Tunnels and Geotechnics, and
Jožef Stefan International Postgraduate School
Ljubljana, Slovenia
miha.hafner@elea.si

Marko Bohanec
Jožef Stefan Institute
Department of Knowledge Technologies
Ljubljana, Slovenia
marko.bohanec@ijs.si

Abstract

Project planning typically refers to the project management step in which project assets, timelines, budgets, milestones, subcontractors, etc., are determined before the new project starts. In this paper, we address infrastructure design projects in the context of a specific company (Elea iC) and explore the idea of using data about past-finished projects to help project managers and project leaders in project planning. A crucial requirement in this context is the ability to evaluate/assess the success of finished/new projects. This paper proposes a solution using a multi-criteria model to evaluate finished projects. This way, we add project success information to the finished projects database, which we shall use in the decision support system being designed to extract knowledge for the new project plan.

Keywords

Project success evaluation, multi-criteria model, decision support systems, data analysis, data mining, project management, project leading tools.

1 Introduction

Infrastructure, such as tunnels, bridges, schools, houses, sewage systems, roads, etc., and its design discipline play a vital role in society. Thus, infrastructure design must have properly and thoroughly defined requirements, objectives, scope and constraints concerning many expert fields such as civil engineering, architecture, geology, geotechnics, environmental engineering, urban planning, and other expert fields [1], [2], [3]. The term design is connected to the process that ends with technical documentation, technical approvals, models, and other deliverables prepared at the end of the design process. Each such process is referred to as the project [4]. The projects are expected to have clearly defined:

- *Goals* defining the project's desired result, e.g., a building permit for a bridge, static analysis of a retaining wall, architectural design for a subway station, geotechnical exploration for a tunnel, etc. [4].
- *Objectives* that support project goals include concrete and measurable project characteristics such as deliverables, milestones, and other steps and strategies to achieve the goals [7].
- *Scope and requirements* concerning project boundaries, e.g., the need for experts, potential subcontractors, technical equipment and other requirements to finish the project.
- *Constraints and limitations* concerning project deadlines, costs, etc. [8].

Besides that, each project should finish with the client's and stakeholders' satisfaction [8].

To achieve the above for the new project, project planning is vital at the beginning of each new project [6], [8]. It is the project management and project leaders' task to recognize and include all these in the project plan so that the work and processes lead to successful project completion.

This study aims to support this process in the context of Elea iC company, an interdisciplinary provider of engineering services and projects in Slovenia [5]. We wanted to include the knowledge obtained from past–finished projects in the project planning process for the new projects. The company collected this data from 2001. The assumptions are as follows:

1. The finished projects in the database offer valuable information for the new project planning phase.
2. The project workflows established in the company and requirements remained similar over the years.

The main challenge related to this question is the new project success assessment and its consideration in light of the available finished project data [7]. Unfortunately, the actual finished projects database does not contain much information about the finished projects' success. To bridge this, we had to construct a project success evaluation model, evaluate finished projects in the database and add this information to the database. The expected result of those steps is a database suitable for applying data-analysis and knowledge extraction methods, such as hierarchical clustering and machine learning [20].

This paper describes the finished project success evaluation component (hereafter called FPSE), which is part of the future decision support system (DSS, [12], [13], [14]) for project planning (hereafter called E-DSS). First, we present the general architecture of the E-DSS, explaining the role and integration of FPSE in its context. In section 3, we present the database of finished projects and its preparation for supporting the configuration of new projects. The evaluation model used and the experimental evaluation of FPSE are presented in sections 4 and 5, respectively. Section 6 concludes the paper.

Information Society 2024, 7–11 October 2024, Ljubljana, Slovenia

https://doi.org/10.70314/is.2024.scai.8463

2 E-DSS Architecture

E-DSS is a DSS under construction to support the project management and project leaders in the Elea iC company (hereafter called "the user") in configuring the new project plan parameters when a new project starts.

The user is expected to define the E-DSS input as shown in Figure 1: the new project objectives, requirements, desires, and expectations. Practically, this means that the user collects all the available new project data by:

- Extracting the new project data from the new project assignment and contracts containing relevant information for the project planning.
- Checking the company and potential subcontractors' state of the resources and assets needed to complete the new project.

Examples of those data include projected monetary value, project scope and goals, project start and finish date, the expert fields needed for project completion, etc.

The E-DSS output (Figure 1) consists of the new project plan configuration together with the corresponding success scores (+S). The configuration comprises the data such as the number of employees involved, the number of subcontractors, work distribution, work duration, the number of pauses, etc. Project success scores are assessed assuming this configuration settings.

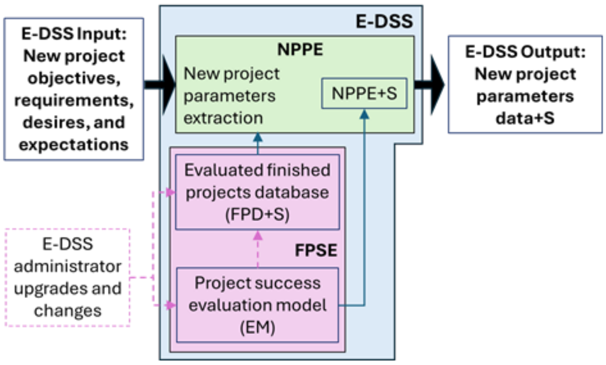

Figure 1: E-DSS architecture

Accordingly, E-DSS is composed of the following components (Figure 1):

- **NPPE** (New Project Parameters Extraction) is the component that extracts the potential new project configuration parameters and corresponding data to support the decision-making. NPPE is currently under construction and is aimed to operate interactively with the user and support: searching for similar projects in FPD+S according to a predefined range, searching the projects by desired success score, project segmentation, and project group identification—unsupervised descriptive analytics and parameter prediction by supervised machine learning methods. The component NPPE+S inside NPPE evaluates the success of the potential new project's configuration parameters obtained. The evaluation is made by EM, which is part of FPSE.
- **FPSE** (Finished Project Success Evaluation) consists of:
  - **EM** (Project success Evaluation Model) for evaluating the new project configuration (described in section 4).
  - **FPD+S** is the database of finished projects with project success evaluations (section 3).

Figure 1 also shows the element **E-DSS administrator** used to upgrade FPSE periodically by upgrading the database of the finished projects or making changes in EM according to the system's operational requirements and expected results.

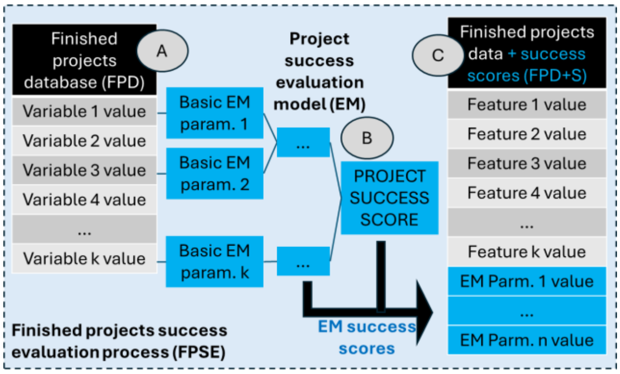

Figure 2: FPD+S development workflow

This paper focuses on the development of FPSE. The workflow is shown in Figure 2, consisting of the following steps:

Step A. Finished projects database preparation (FPD), as described in section 3.

Step B. Project success evaluation model (EM), as described in section 4.

Step C. Finished projects database with EM success scores (FPD+S): The result of the FPSE is the upgraded database of the finished projects with the finished projects' success scores (FPD+S).

3 Data Description

E-DSS is a data-driven system that operates on data from past-finished projects. This data was collected in Elea iC company from the year 2001 to 2023. At the beginning of the data collection, the number of the observed variables was relatively small, but it has grown substantially over the years. At the time of this study, the database contained data on 4704 finished projects, described by 39 numeric variables; 6 of them were date/time/year variables, and 2 categorical variables.

Data preparation (Step A, Figure 2) was carried out as follows:

1. Data cleaning: replacing "Nan" and deleting erroneous data;
2. Outlier's removal using the Interquartile range approach [18];
3. Data imputation: replacing the missing values using a descriptive statistic (e.g. mean, median, or most frequent) along each column or using a constant value [19]. We employed the mean strategy.
4. Sensitive data and information removal. For this reason, all numeric data was scaled to a range between 0 and 1.

We ended up with the database FPD containing data on 3132 finished projects described by 27 numeric variables. The variables describe the main project management characteristics, such as project financial results, workload distribution, number of employees, subcontractors, etc. Table 1 shows the list of all variables together with their basic statistics.

This way, the finished project database (FPD) was prepared for the FPSE component. FPD is the main resource for Exploratory Data Analysis for observing the data and its properties, such as variable correlations, variable information gain, etc. These operations are invoked interactively by the user in the context of NPPE and are not discussed further in this paper.

Table 1: Basic statistics of the variables after data cleaning, outliers' removal, and data scaling

	min	max	median	mean
Project_Value	0.0	1.0	0.4215	0.2476
Number_of_Hours	0.0	1.0	0.0017	0.0123
Number_of_Employees_on_Project	0.0	1.0	0.0452	0.0420
Emploees_Workhours_Load	0.0	1.0	0.0119	0.0114
Costs_of_Subcontractors	0.0	1.0	0.0000	0.0069
Number_of_Subcontractors	0.0	1.0	0.0000	0.0150
Travelling_km	0.0	1.0	0.0000	0.0044
Project_Income	0.0	1.0	0.0018	0.0368
Project_Duration	0.0	1.0	0.2024	0.2414
Work_Period_Duration	0.0	1.0	0.0271	0.0770
Pauses_Duration	0.0	1.0	0.0000	0.0558
Number_of_Pauses	0.0	1.0	0.0000	0.0473
Proj_Work_Concentration	0.0	1.0	0.4985	0.5780
Number_of_Work_Months	0.0	1.0	0.0282	0.0740
Work_dist_Mean	0.0	1.0	0.0084	0.0190
Work_dist_Std	0.0	1.0	0.0283	0.0364
Work_dist_Min	0.0	1.0	0.0019	0.0076
Work_dist_25%	0.0	1.0	0.0041	0.0110
Work_dist_50%	0.0	1.0	0.0067	0.0161
Work_dist_75%	0.0	1.0	0.0096	0.0227
Work_dist_Max	0.0	1.0	0.0089	0.0253
Work_dist_Kurtosis	0.0	1.0	0.2051	0.2047
Work_dist_Skewness	0.0	1.0	0.4807	0.4799
Year	0.0	1.0	0.4167	0.4852
Pauses_Time_Share	0.0	1.0	0.0000	0.2096
Hour_Income	0.0	1.0	0.4547	0.2923
Time_Reserve	0.0	1.0	0.8128	0.6891

4 Evaluation of Projects' Success

The project success evaluation model (EM), developed in Step B (Figure 2), is aimed at:

- The evaluation of the projects in FPD resulting in the FPD+S (Figure 2).
- The evaluation of potential new projects suggested through interaction between the user and NPPE+S (Figure 1).

Project success evaluation involves multiple criteria that have to be aggregated into a single evaluation score. Different criteria might be of different importance and affect the score differently, i.e., with different weights. For this purpose, we chose MAUT (Multi-Attribute Utility Theory) [11], a multi-criteria modelling approach that facilitates both hierarchical structuring of criteria and using weights for the aggregation of scores.

Considering the above requirements, available FPD data and other multi-criteria approaches to project evaluation ([15], [16], [17]), we developed the EM as presented in Figure 3.

EM consists of three components [10]: input parameters, evaluation parameters and aggregation functions.

Input parameters are variables in the leaf nodes of the model:

- *Project Work Concentration*: explains the distribution of the work on the project. If the value is closer to 0 or 1, the majority of the work is done at the beginning or at the end of the project, respectively.
- *Time Reserve*: explains if the project work ended earlier than defined in the contract.
- *Number of Pauses*: the number of times the work on the project stopped.
- *Pauses Time Share*: the ratio between the months the employees did not work and the total number of months.
- *Hour Income*: the ratio between project value and the number of work hours necessary to finish the project.

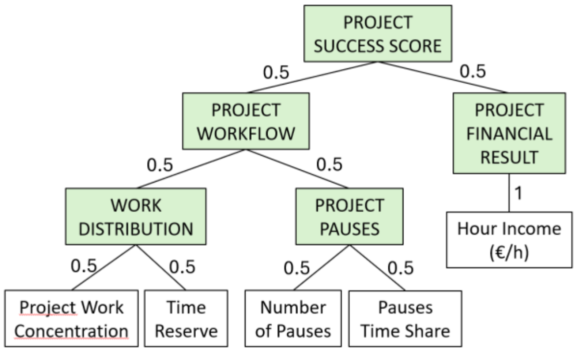

Figure 3: Multi-criteria model for the projects' success evaluation

Evaluation parameters represent outputs of the model:

- *WORK DISTRIBUTION*: assesses the characteristics of the work distribution in the project duration.
- *PROJECT PAUSES*: assesses the work pauses.
- *PROJECT WORKFLOW*: combines evaluation parameters WORK DISTRIBUTION and PROJECT PAUSES
- PROJECT FINANCIAL RESULT: assesses the project's success from the financial point of view.
- *PROJECT SUCCESS SCORE*: overall success score, determined by aggregation of all subordinate parameters.

Aggregation functions map subordinate EM parameters to the corresponding parent parameters. Employed is the weighted average function, using weights shown in Figure 3. Currently, weights are chosen to make all parameters equally important.

5 Experimental Evaluation of FPSE

Figure 4 shows an example of evaluating a project from FPD. Input parameters' values (terminal nodes) were obtained from the data base, while evaluation parameters' values (green nodes) were calculated by EM. The example project shows good workflow score (0.75), but has a poor financial score (0.29), both leading to an average success score (0.52) of the project. Several other projects of different types were evaluated in this way, confirming the appropriateness of EM structure and conformance with requirements of potential users. In this way, the quality of EM was assessed on a sample of past projects. Further assessment is planned in the next stages while configuring new projects, where EM's results can be confronted with opinions of project leaders actively involved in the process.

EM already enables evaluation of multiple finished projects. In Step C (Figure 2), FPD was extended by adding five variables corresponding the five Evaluation parameters of EM. All projects in FPD were evaluated by EM, resulting in FPD+S.

Basic statistics of FPD+S is presented by the distribution of the variables in Figure 5. The variables marked with red colour on the x-axis are E-DSS input parameters, the green uppercase variables are those corresponding to success scores, and the blue variables are potential new project parameters. The distribution of the final project evaluation, PROJECT_SUCCESS_SCORE, (average = 0.52, min = 0.15, max = 0.94) indicates that it well covers the range of possible outcomes and enables the discrimination and sorting of projects.

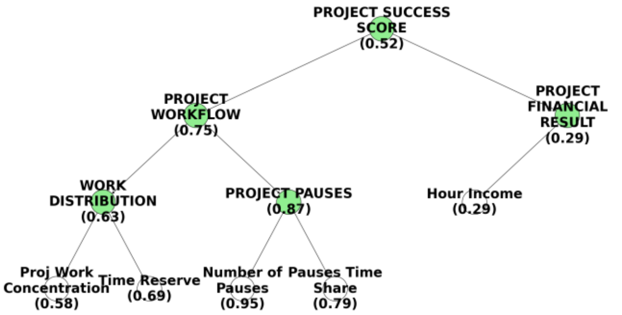

Figure 4: Example of evaluating a project using EM

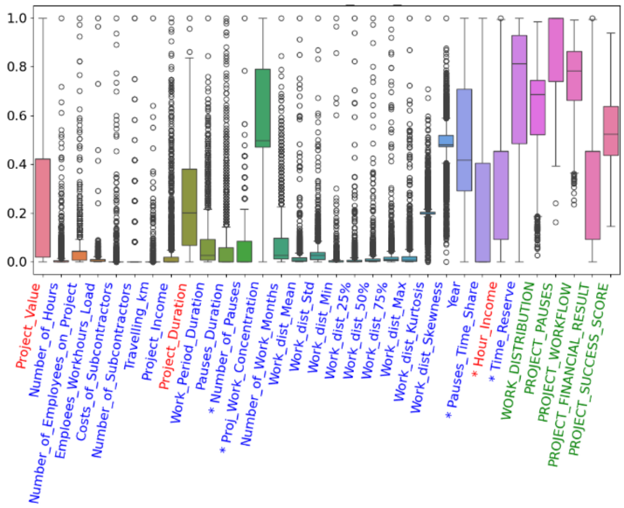

Figure 5: Distribution of the FPD+S features, including EM project success assessments

6 Conclusions

E-DSS is a DSS under construction, aimed at supporting the project management and project leaders' process in the new infrastructure project planning phase. We presented the design and development of the FPSE component, consisting of a multi-criteria project success evaluation model EM and a data base of projects, extended with success evaluation scores FPD+S.

EM has been developed using the MAUT approach and has turned out to be “fit-for-purpose”. It employs the data that is available in the projects’ database. It meaningfully describes aspects of the project's success and offers practical and functional model for the evaluation of multiple projects in the database.

FPSE is a key decision-support resource for E-DSS. E-DSS will allow the user to interactively search for similar past projects, to filter them according to the success score and simulate the effects of alternative project configurations, ultimately proposing the best one. Approaches based on unsupervised descriptive analytics (clustering) and supervised machine learning methods for prediction of E-DSS output parameters are foreseen for this purpose. Actually, we have already tested hierarchical clustering and decision tree classification methods on FPD+S, and first results are encouraging. We obtained meaningful clusters of past projects and created decision trees for prediction of individual output parameters that may lead to high new project success scores.

Future work will primarily continue by further data analysis and data mining of FPD+S, attempting to design effective algorithms for interactive exploration of past projects and suggesting as good as possible configurations of new projects. On this basis, we shall make a detailed functional specification of the NPPE+S component and design/implement the E-DSS.

Despite that E-DSS considered here is tailor-made for the specific business environment and is bound to the specific data base, the approach seems general enough to be applied to similar environments, projects and processes [9]. This work is a showcase of substantial efforts needed to prepare a corporate database for decision-support, which is often neglected in the literature. The main contribution is a combination of data processing with MAUT-based multi-criteria decision modelling.

References

[1] Fransje L. Hooimeijer, Jeremy D. Bricker, Adam J. Pel, A. D. Brand, Frans H.M. Van de Ven, Amin Askarinejad. 2022. Multi-and interdisciplinary design of urban infrastructure development. In Proceedings of the Institution of Civil Engineers: Urban Design and Planning. Vol.175. TU Delft. 153-168.

[2] Simon Christian Becker, Philip Sander. 2023. Development of a Project Objective and Requirement System (PORS) for major infrastructure projects to align the interests of all the stakeholders. In Expanding Underground - Knowledge and Passion to Make a Positive Impact on the World. CRC Press, London, UK, 3369-3376. DOI:10.1201/9781003348030-408.

[3] Michel-Alexandre Cardin, Ana Mijic, Jennifer Whyte. 2023. Data-driven infrastructure systems design for uncertainty, sustainability, and resilience. In D. M. Fabio Biondini, Life-Cycle of Structures and Infrastructure Systems. CRC Press, London, UK, 2565 – 2572. DOI: 10.1201/9781003323020-312.

[4] Saša Žagar. 2016. Organizacijski model v projektivnem podjetju Elea iC d.o.o.. Maribor, B.Sc. Thesis, Retrieved July 12, 2024 from https://dk.um.si/IzpisGradiva.php?id=58799&lang=eng.

[5] Elea iC webpage. https://www.elea.si/en/.

[6] Jürg Kuster, Eugen Huber, Robert Lippmann, Alphons Schmid, Emil Schneider, Urs Witschi, Roger Wüst. 2015. Project Management Handbook. Springer-Verlag, Berlin Heidelberg, Germany.

[7] Anton Hauc. 2007. Projektni management. (2nd. ed.). GV Založba, Ljubljana, Slovenija.

[8] Harvey A. Levine. 2002. Practical Project Management: Tips, Tactics, and Tools. John Wiley & Sons, Inc., New York, NY.

[9] Nadja Damij, Talib Damij. 2014. Process management. Springer-Verlag, Berlin Heidelberg, Germany.

[10] Marko Bohanec. 2012. Odločanje in modeli. DMFA – založništvo, Ljubljana, Slovenija.

[11] Salvatore Greco, Matthias Ehrgott, José Rui Figueira. 2016. Multiple Criteria Decision Analysis, State of the Art Surveys. Springer, Portsmouth, UK. DOI 10.1007/978-1-4939-3094-4

[12] Maria Rashidi, Maryam Ghodrat, Bijan Samali and Masoud Mohammadi. 2018. Decision Support Systems. In Management of Information Systems. IntechOpen, London, UK, 19-38. DOI: 10.5772/intechopen.79390.

[13] Daniel Joseph Power. 2013. Decision Support, Analytics, and Business Intelligence. Business Expert Press, New York, NY. DOI 10.4128/9781606496190.

[14] Sofiat Abioye, Lukumon Oyedele, Lukman Akanbi, Anuoluwapo Ajayi, Juan Manuel Davila Delgado, Muhammad Bilal, Olugbenga Akinade, Ashraf Ahmed. 2021. Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. Journal of Building Engineering 44, Elsevier. https://doi.org/10.1016/j.jobe.2021.103299.

[15] Erwin Berghuis. 2018. Measuring Systems Engineering and Project Success, Master’s Thesis. University of Twente. https://purl.utwente.nl/essays/75088

[16] Ali Beiki Ashkezari, Mahsa Zokaee, Amir Aghsami, Fariborz Jolai, Maziar Yazdani. 2022. Selecting an Appropriate Configuration in a Construction Project Using a Hybrid Multiple Attribute Decision Making and Failure Analysis Methods. Buildings, MDPI, Volume 12, 643. DOI: https://doi.org/10.3390/buildings9050112.

[17] Urban Pinter, Igor Pšunder. 2013. Evaluating construction project success with use of the M-TOPSIS method. Journal of civil engineering and management, Volume 19(1), 16–23. doi:10.3846/13923730.2012.734849

[18] Interquartile range. Retrieved May 15, 2024 from https://en.wikipedia.org/wiki/Interquartile_range

[19] SimpleImputer. Retrieved May 15, 2024 from https://scikit-learn.org/stable/modules/generated/sklearn.impute.SimpleImputer.html

[20] Aggarwal C. Aggarwal. 2015. Data Mining: The Textbook. Springer, New York, USA.