Megaprojects are often associated with poor delivery performance and poor benefits realization. This article provides a method of identifying, in a quantitative and rigorous manner, the characteristics related to project management success in megaprojects. It provides an investigation of how stakeholders can use this knowledge to ensure more effective design and delivery for megaprojects. The research is grounded in 44 megaprojects and a systematic, empirically based methodology that employs the Fisher’s exact test and machine learning techniques to identify the correlation between megaprojects’ characteristics and performance, paving the way to an understanding of their causation.

KEYWORDS: megaprojects; case studies; statistical analysis; budget; schedule

INTRODUCTION

Megaprojects are temporary endeavors (i.e., projects) characterized by large investment commitment, vast complexity (especially in organizational terms), and long-lasting impact on the economy, the environment, and society (Brookes & Locatelli, 2015). Megaprojects include the development/creation of power plants, oil and gas extraction plants, airports and processing projects, railways, motorways, dams, and even cultural events such as the Olympic games or universal expositions (Van Wee, 2007). What megaprojects have in common is their requirements for the coordination and control of a vast and complex array of financial, social, and technical resources to turn them into reality (Hu, Chan, Le, & Jin, 2013; Locatelli, Mancini, & Romano, 2014). Megaprojects have significant implications for society, and they play a pivotal role in the implementation of both energy and transportation policies (Locatelli, Invernizzi, & Brookes, 2017; Locatelli, Mariani, Sainait, & Greco, 2017; Sovacool, Nugent, & Gilbert, 2014). Megaprojects represent the largest proportion of governmental and European Commission expenditure on infrastructure, and their successful design and delivery can have major implications for public finances (Flyvbjerg, Hon, & Fok, 2016). Despite their criticality, megaprojects are associated with extremely poor delivery performance and long-term benefits realization (Flyvbjerg, Bruzelius, & Rothengatter, 2003; Kardes, Ozturk, Cavusgil, & Cavusgil, 2013; van Marrewijk, Clegg, Pitsis, & Veenswijk, 2008).

The successful transfer of learning across projects and megaprojects has been a long-held desire by those involved in their design and delivery. The difficulties in learning are created by the very nature of projects themselves, that is, their separation from a “permanent” organization and their uniqueness (Jacobsson, Lundin, & Soderholm, 2015; Kujala, Artto, Aaltonen, & Turkulainen, 2010; Wikstrom, Artto, Kujala, & Soderland, 2010). Moreover, the size and complexity of megaprojects can make it very difficult to discern which actors and elements of their myriad configurations actually influence their success and delivery (Chang, Hatcher, & Kim, 2013; Chapman, 2016; van Marrewijk et al., 2008).

Over the past two decades, project management literature has vastly investigated the “success factors” that impact the success of the projects, measured through the so-called “success indicators” (or criteria). A “success factor” might be detailed front-end engineering and design (FEED) (Merrow, 2011) or the early engagement of external and internal stakeholders (Brookes & Locatelli, 2015). Project “success indicators” are the measures by which the successful outcome of a project is assessed, whereas “success factors”
are the elements of a project that can be influenced to increase the likelihood of success (Müller & Turner, 2007). Traditional “success indicators” in project management refers to the so-called iron triangle: cost, time, and quality. However this short-term, contract-based view has been challenged by researchers who consider multiple perspectives of different stakeholders in different time frames (Davis, 2014; Dimitriou, Ward, & Wright, 2013; Turner & Zolin, 2012). Williams (2016) emphasized that it is becoming increasingly recognized by the project management community (particularly by academics) that the nature of project success is multidimensional, with different criteria, only some of which are clearly measurable. Thus, there is still limited understanding of the causal chains through which success emerges. Zavadskas, Vilutien, Turskis, and Saparauskas (2013) also analyze common construction performance, comparing what they call project management “problems” against the “success factors,” illustrating how to assess projects’ efficiency using aggregated indicators. Gunathilaka, Tuuli, and Dainty (2013), reviewing papers about the relationship between project success factors and project success indicators, highlight the scarce empirical evidence supporting the actual correlation between them. Bassam (2013) does not limit his research to the construction field and employs statistical analysis to examine the correlations between the risk factors that are common to success indicators, concluding that there are some factors in the initiation phase that could lead to the occurrence of additional risk factors during the implementation and evaluation phases. A detailed example of this latter case related to EXPO 2015 is presented in Locatelli and Mancini (2010).

This article aims to provide a method to identify in a quantitative and rigorous manner the megaproject characteristics (i.e., the aforementioned “success factors”) correlated to project management success indicators. It also aims to provide a model for megaproject cost and time performance prediction.

As clarified in the literature review section, a key novelty of the research presented in this article is the transparent leveraging of project characteristics and case studies rather than, for instance, the use of a survey or proprietary database. Specifically, the article addresses 43 project characteristics (e.g., “there is planned long-term stability in usage and value” or “the project receives financial support from the European Union [EU]”; see the full list in the Appendix and Tables 4–10) and their correlation with three project management success indicators (“project had a cost overrun,” “delayed in the planning phase,” and “delayed in the construction phase”).

**Literature Review**

The majority of the existing literature about project management success indicators can be clustered into three groups:

1. Statistical analysis of large databases;
2. Surveys with project managers and stakeholders; and
3. Case study analysis.

**Statistical analysis of large databases.** Prominent research has been undertaken in the statistical analysis of large databases of megaprojects (Ansar, Flyvbjerg, Budzier, & Lunn, 2014; Cantarella, Flyvbjerg, & Buhl, 2012; Flyvbjerg, 2006; Merrow, 2011). For example, by analyzing a large database, Flyvbjerg (2006) investigates why projects are late or over budget and, once delivered, why they provide less benefit than planned. Optimism bias and strategic misrepresentation are significant contributory factors in the overestimation of benefits and underestimation of the costs of megaproject design and delivery. According to Merrow (2011), the main reasons for project failure are poor front-end loading (FEL), FEED, and misaligned incentives. Such statistical analyses provide invaluable insights into megaproject design and delivery, but the lack of availability of the data used to create these findings is inimical to further investigations.

A particular subset of statistical researchers has focused on a particular type of megaproject. Koch (2014), for example, investigated seven Danish and Swedish offshore wind farms; Ansar et al. (2014) analyzed a large database of dams; and Sovacool, Gilbert, and Nugent (2014a, 2014b) conducted empirical studies relying on an extensive database composed of 401 electricity infrastructure projects. Despite their large statistical significance, the sectoral construction of these investigations’ datasets makes it difficult to extend the findings to other megaproject sectors.

**Surveys of project participants.** Several researchers follow this path, directly surveying project managers about the success factors in their projects (e.g., Kog & Loh, 2012; Pinto & Mantel, 1990; Pinto & Slevin, 1987) or the “factors that were regarded as critical to that project’s outcome” (White & Fortune, 2002, p. 1) or the “project success factors for design and build projects and the relative importance of these factors on project outcome” (Chan, Ho, & Tam, 2001, p. 93). These papers provide a very interesting contribution, but the methodological choice they follow means that their results can only really be considered normative studies of what scholars and practitioners involved in the particular surveys think. Often, the responses are extremely constrained by the survey instruments utilized in the investigations.

**Case study analysis.** Case study analysis is a research methodology extensively used to describe and understand the behavior of a project (Yin, 2013). It can be considered a very effective methodology for theory building (Eisenhardt, 1989). Case studies do provide a useful approach to investigating megaprojects (Brookes, Hickey, Littau, Locatelli, & Oliomogbe, 2015). For example, Locatelli and Mancini (2012) analyzed the case studies of the nuclear
reactors Olkiluoto 3 and Flamanville 3. Greiman (2013) starts with a deep analysis of a single megaproject (the Central Artery/Tunnel Project in Boston, known unofficially as the Big Dig) to generalize a set of lessons and guidelines. The main limitation of case studies like these is in the emphasis on theory building rather than theory testing.

Despite the substantial amount of work that has already been undertaken to understand megaproject behavior, few studies have attempted to quantitatively express the relationships between project characteristics and success indicators. It is particularly unclear how to use this understanding to build a performance prediction model for megaprojects that would be of particular use to megaproject design and delivery professionals.

Method

Key Challenges

Given the limitations of the existing work in this area, the authors sought further approaches that relate megaproject characteristics to their success indicators. The main challenges in the process of identifying such relationships originate from the complexity of megaprojects and the size, availability, and representation of the data describing them.

The complexity of megaprojects. Because of their investment size, long duration, technological complexity, and political and social environment (including a large network of internal and external stakeholders), megaprojects are an extremely complex phenomenon. In order to use statistics to analyze megaprojects, it is necessary to cope with this complexity, specifically in terms of the number of characteristics that are being included in the study. Reductionism is dangerous, as it can move investigators away from identifying complex holistic phenomena, but it is necessary if statistical significance is sought with a preference for model simplicity (Easterby-Smith, Thorne, & Jackson, 2012).

The nature of megaproject data. Compared to the other types of projects across different sectors, the number of megaprojects is very limited. Moreover, information sensitivity issues can also strongly affect the availability and quality of specific megaproject data.

Data representation suitable for statistical analysis. When converting the real-life complexity of megaprojects to a dataset amenable to statistical analysis, it is necessary to identify the way of measuring and describing project characteristics (independent variables). Frequently, the conversion process adopted by researchers relays on integer Likert-type scales to rank qualitative variables. The application of Likert-type scales could potentially be interpreted differently by researchers and a wider practitioner audience (ranking the complexity of a megaproject on a scale of 1–7, for example, seems to be a subjective exercise). In addition, given the small size of the available megaprojects dataset, the high dimensionality of project characteristics could lead to inadequate statistical models, unable to capture the relationship between project characteristics and project performance ("curse of dimensionality") (Indyk & Motwani, 1998). Considering the previously defined challenges, when identifying significant relationships between project characteristics and performance, the authors applied a data-driven approach using the database presented in Brookes (2013) and a list of project characteristics presented in the Appendix. The method is based on two macro-phases:

1. Data collection and preparation, and
2. Data analysis using the Fisher’s exact test (FET) and machine learning (ML).

Data Collection and Preparation

Data collection and preparation consist of case collection and brainstorming, systematization, and definition of possible project characteristics. The authors identified 43 independent variables (i.e., project characteristics) for 44 megaproject cases.

Step 1: Case collection

Each case study is a megaproject delivered in the EU. The authors collected information about the specific case study and gained preliminary qualitative knowledge of the factors influencing successful project delivery.1

The final sample consists of 44 cases, clustered as follows:

• 30 transport: 6 motorways, 15 rail projects, 5 urban transport projects (4 metro lines and 1 tram), 2 bridges (road bridges), 1 tunnel (for road and rail traffic), and 1 airport
• 12 energy: 5 nuclear, 3 thermal, 2 windfarms, 1 solar, and 1 liquefied natural gas (LNG) extracting platform
• 2 hydrotechnical megaprojects: Mose in Venice, Italy, and the Raciborz reservoirs in Poland

The qualitative data describing these case studies are available in Brookes (2013) and University College London (2015).

Step 2: Identification of project characteristics as possible determinants of project management success

After the collection of cases, the authors identified a large range of megaproject characteristics that might be correlated with the success indicators. The list of project characteristics is, therefore, based on the knowledge acquired during the elaboration of these case studies, the researchers’ previous knowledge, and the literature summarized in the Appendix.

Step 3: Systematization

Following the identification process, the authors gathered to systematize the data. This “cleaning up” led to the final

1The authors acknowledge the contributions of scholars and practitioners involved in the “Megaproject Cost Action.” A full list of the people involved and the portfolio of projects analyzed are available at http://www.mega-project.eu/
Operationalization

The project was judged to have a cost overrun if the final cost of the project was greater than the 110% of the original estimate (adjusted for inflation). The estimated cost was taken to be a publicly available figure obtained either through direct interview with the project client or through public review at a time as close as possible to the point at which the first formal activity was entered into (for instance, the first stage in the acquisition of any land rights required for the project). The final cost was taken to be a publicly available figure obtained either through direct interview with the project client or through public review at the point at which the project entered operation. The final cost and initial estimate were assumed to have been made on the same basis.

The project was judged to be delayed in the planning phase if the actual commencement of physical construction was more than 12 months later than the planned date for the commencement of construction. The planned date for the commencement of construction was taken to be a publicly available figure obtained either through direct interview with the project client or through public review at the time as close as possible to the point at which the first formal activity was entered into (such as the first stage in the acquisition of any land rights required for the project). The actual date for the commencement of construction was taken at the point at which any physical construction activity related directly to key functionality of the project was undertaken as reported through direct interview with the project client or through public review.

The project was judged to be delayed in the construction phase if it exceeded the planned date for entry into service by 12 months (compared to the date set at the point of entry into construction). The planned date for the entry into service was taken to be a publicly available figure obtained either through direct interview with the project client or through public review at the time as close as possible to the commencement of construction work. The actual date for the entry into service was taken at the point at which the output from the project was first provided to its intended beneficiaries as reported through direct interview with the project client or through public review.

Table 1: Project management performance definitions (adapted from Brookes and Locatelli, 2015).

<table>
<thead>
<tr>
<th>Dependent Variable Construct</th>
<th>Operationalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>The project had a cost overrun</td>
<td>The project was judged to have a cost overrun if the final cost of the project was greater than the 110% of the original estimate (adjusted for inflation). The estimated cost was taken to be a publicly available figure obtained either through direct interview with the project client or through public review at a time as close as possible to the point at which the first formal activity was entered into (for instance, the first stage in the acquisition of any land rights required for the project). The final cost was taken to be a publicly available figure obtained either through direct interview with the project client or through public review at the point at which the project entered operation. The final cost and initial estimate were assumed to have been made on the same basis.</td>
</tr>
<tr>
<td>The project was delayed in the planning phase</td>
<td>The project was judged to be delayed in the planning if the actual commencement of physical construction was more than 12 months later than the planned date for the commencement of construction. The planned date for the commencement of construction was taken to be a publicly available figure obtained either through direct interview with the project client or through public review at the time as close as possible to the point at which the first formal activity was entered into (such as the first stage in the acquisition of any land rights required for the project). The actual date for the commencement of construction was taken at the point at which any physical construction activity related directly to key functionality of the project was undertaken as reported through direct interview with the project client or through public review.</td>
</tr>
<tr>
<td>The project was delayed in the construction phase</td>
<td>The project was judged to be delayed in the construction phase if it exceeded the planned date for entry into service by 12 months (compared to the date set at the point of entry into construction). The planned date for the entry into service was taken to be a publicly available figure obtained either through direct interview with the project client or through public review at the time as close as possible to the commencement of construction work. The actual date for the entry into service was taken at the point at which the output from the project was first provided to its intended beneficiaries as reported through direct interview with the project client or through public review.</td>
</tr>
</tbody>
</table>
Le, Caldas, Gibson, & Thole, 2009; Son, Kim, & Kim, 2012; Wang & Gibson, 2010; Williams & Gong, 2014).

ML techniques enable rigorous “pattern spotting” analysis of the existing (and relatively small) dataset that did not allow the application of multivariate statistical analysis. After conducting the FET, three different ML classifiers and two feature-selection techniques were applied. These techniques were adopted for small datasets to build models for predicting megaproject management success.

**Fisher’s Exact Test**
The purpose of the FET is to ascertain whether or not an independent variable is correlated with the presence (or absence) of a dependent variable (Leach, 1979). With respect to this research, the FET has two key features (Sheskin, 2011). First, the FET makes no assumption about distributions. It is a non-parametrical statistical significance test. It is not necessary to make a priori assumptions on the data distribution; therefore, this type of test can have a wide application. Second, the FET uses categorical data in the form of a contingency table and is used for categorical binary data. The probability of a relationship existing between the variables can be calculated exactly rather than estimated as in other statistical techniques. Further information about the FET and the steps to apply it in this kind of research are detailed in Brookes and Locatelli (2015).

**Machine Learning**
ML belongs to a continuum of data analysis techniques that use underlying data to describe structural patterns, explain trends, and make predictions (Kohavi & Provost, 1998; Witten, Frank, & Hall, 2011). In this research, the authors formulated the problem of project management success assessing correlations between project characteristics and performance where a set of classified examples (megaproject cases) is given as the input to an ML technique. Based on this classification, ML techniques map the relationships between project characteristics and success indicator classes. Examples are represented as binary vectors (i.e., the 43 project characteristics; see the Appendix) and a class label (success indicators; see Table 1), according to which the examples were classified. The output may also include an actual description of a structure that can be used to classify unknown examples. Descriptions can become fairly complex and are typically expressed as sets of rules. Because there are three success indicators adopted for investigation in this research (see Table 1), the authors built a separate classification model for each of them. The framework for building and evaluating the proposed models is described in the following sections.

**Testing Protocol**
In order to assess the quality of the analysis on the available megaproject data, which ends with the selection of the most informative characteristics for the specific success indicator, the authors proposed different classification models and estimated their performance based on the selected characteristics, with the reasonable assumption that more important project characteristics lead to better prediction accuracy. Models were evaluated on test data using the leave one out procedure that minimizes the negative effects of small sample size (Reich & Barai, 1999). The procedure iteratively divides the dataset with n examples (megaproject cases) into training (n-1 examples) and test part (one example). In each iteration, the test part rotates throughout the dataset, enabling the model to be both trained and tested on all available data (Reich & Barai, 1999). An aggregated confusion matrix, which accumulates classification results from each iteration, is maintained and further used to calculate standard model performance measures, in other words, precision, recall (for each class and overall), accuracy, and F-measure (Williams & Gong, 2014; Witten et al., 2011). The precision of the classification model for “class C” represents the percentage of examples classified as C, which actually belongs to C. The recall for C is the percentage of all examples from C in the test set that is predicted as C.

When comparing different models, it is convenient to use a single measure, such as accuracy (percentage of accurately classified examples in the test set) or weighted F-measure. The F-measure has been widely used in the field of data mining and information retrieval (He, Shu, Yang, & Wang, 2012; Japkowicz & Shah, 2011). It integrates recall and precision in a single indicator. Weighted F-measure is the weighted sum of harmonic means between class precision and recall. There is no unified standard for a threshold to be used for judging whether the results of a project management success prediction are acceptable. In past studies, both accuracy and F-measure equal or above 0.7 are often considered reasonably good (He et al., 2012; Thung, Lo, & Jiang, 2012; Zhang, Gong, & Versteeg, 2013).

**Selection of the Most Informative Project Characteristics**
Correlation-based feature subset selection (CFS) and selection of project characteristic based on information gain were applied to find the most informative subsets of project characteristic for predicting each of the three class labels (i.e., the project management success indicators).

CFS is based on the hypothesis that good feature sets contain features that are highly correlated with the class, yet uncorrelated with one another (Hall, 1999). The algorithm couples the evaluation formula with an appropriate correlation measure and a heuristic search strategy in order to find the best subset of available features. Opposite to CFS, information gain calculates the score of each feature independently of other features and selects the top n features from the sorted score list. The score is calculated based on information...
gain, which is the expected reduction in entropy caused by partitioning the examples according to a given project characteristic (Quinlan, 1986). The entropy characterizes the (im)purity of an arbitrary collection of examples with respect to their class. Although, in the case of CFS, the number of selected project characteristics is determined by the most informative produced subset, in this research, the top 10 most informative project characteristics by information gain were selected.

**Applied Learning Methods**

Three different learning methods—decision tree, Naive Bayes, and logistic regression—were trained on the available project characteristics. Decision tree is a non-parametric learning method used for classification and regression (Mitchel, 1997). The goal is to build a tree from the available data where, in each node, an example is tested against the value of the project characteristic associated with the node. Depending on the test result, the example is forwarded down the tree until it reaches the leaf node with the appropriate class label. The selection of the characteristics associated with the nodes is conducted using information gain by placing the most informative characteristics at the root node. The procedure is recursively repeated until all nodes are inserted into the tree. When built, human experts can interpret the decision tree because it provides rules for classification in the form of a sequence of if/then clauses (each branch from root to a leaf node is a conjunction of characteristics tests).

Naive Bayes is a probabilistic classifier that assumes the mutual independence between the characteristics given a class of an example (Yun & Caldas, 2009). It selects the class with the maximum posterior probability given an example represented with its characteristics by using the Bayes theorem. The classifier is easy to implement, but it is dependent on prior class probabilities—it tends to classify toward classes that occur more frequently in training data.

Logistic regression is a type of probabilistic statistical classification model used to predict the class based on one or more characteristics that are usually but not necessarily continuous. It measures the relationship between the class and the independent characteristics by using probability scores of the predicted values of the class (Hair, Black, Babin, & Anderson, 2009).

**Results**

Table 3 presents the overall results.

For ML, the quality of the prediction models is presented through “accuracy” and “F-measure” for three success indicators. For each success indicator, the authors used three learning techniques (decision tree, Naive Bayes, logistic regression) and three characteristic selection techniques (all characteristics, CFS, and information gain). The results obtained from ML tool application are shown in Table 2. As previously stated, accuracy and F-measure values equal to or above 0.7 are considered to be reasonably good.

The best performing models are:

- For cost overrun—Logistic regression with information gain (accuracy 0.718; F-measure 0.720; slightly higher F-measure than logistic regression with CFS);
- For delay in construction—Logistic regression with CFS (accuracy 0.732; F-measure 0.730); and
- For delay in planning—Decision tree with CFS (accuracy 0.718; F-measure 0.720).

Results show that the best-performing models are logistic regression and decision tree, which allow identification of the most informative project characteristics because of their expressive power. The prediction performance of the majority of models improves when the techniques for the selection of characteristics are applied, indicating that many of the megaproject characteristics taken into account with the available data are not sufficiently informative. Table 3 shows the most informative subsets of project characteristics correlated to cost overrun, delays in construction, and delays in the planning phases of megaprojects within the existing dataset.
The Successful Delivery of Megaprojects: A Novel Research Method

The most informative project characteristics for ML have an "X" indication (irrespective of having a positive or negative influence on megaproject outcome). Regarding the FET results, Table 3 presents the p-value and the type of correlation. A plus (+) indicates project characteristics that, if existing, are supportive, in other words, positively influence the project outcome. A minus (−) indicates project characteristics that, if they exist, are antagonist, in other words, negatively influence the project outcome. The results, therefore, represent the correlation between the individual project characteristics and the success indicators.

Regarding ML, for each success indicator there is a small subset of characteristics identified as the most informative for the prediction of megaprojects' success. For cost overrun, these mainly come from the categories project environment (legal and socioeconomic) and technological aspects of the project. Half of the most informative characteristics for both delay in construction and planning also fall into the category technological aspects. These are related to the complexity of megaprojects, because they are often the first of a kind in a country (characteristic T3) and they are challenging because of their sector-specific requirements, such as nuclear projects (characteristic T6), or their location, such as offshore projects (characteristic T7).
offshore projects (characteristic T7). For instance, according to ML, modularity when designing and building megaprojects (characteristic T1) helps prevent delays in the planning phase. However, the construction of "stand-alone" plants built with the interconnected modules is correlated with delays in construction (characteristic T2).

In the results of FET, project characteristics from the "stakeholders" category have the lowest p-value, in other words, stronger correlation with success indicators. "Litigation between client and EPC" and "the presence of an SPE" (special purpose entity) are the project characteristics correlated with all three success indicators.

In summary, out of 43 project characteristics, only 10 have been recognized as correlated with megaproject management success by both ML and FET. Regarding cost overrun, there are three project characteristics identified both by ML and FET: "environmental groups have been engaged ex ante, not ex-post;" "the project has a strong regulation system;" and "the project is a nuclear reactor." These prove that the role of external stakeholders is extremely relevant; indeed, actions from environmental groups, the regulation system (and agencies), and the national government are all strongly correlated with cost overrun performance. "Physical characteristics" are not correlated, unless the project is a nuclear reactor: All nuclear reactors under construction in Europe are over budget and late.

Regarding the delay in construction, the most correlated characteristics (identified by both ML and FET) are: "the project has national public acceptability;" “there is planned a long-term stability in usage and value;" "there was a formal litigation procedure during the contract between client and EPC;" "offshore project;" and “the project has an SPE." They are, again, mostly related to the project stakeholders: public acceptability, the contractual relationship between a client and a contractor, and the existence of an SPE. Interestingly, “the project has an SPE” is among the most relevant variables both for delay in construction and delay in planning. SPES are fenced organizations with limited, predefined purposes and a legal existence (Sainati, Locatelli, & Brookes, 2017). They require a long due diligence process (which often delays the beginning of the project) but after that, they help keep the project on schedule.

Regarding the delay in planning, three project characteristics are identified as correlated by both MLT and FET: "the project has a strong regulation system," "FOAK weak—country level," and "the project has an SPE." The key results show that the regulatory system and regulations have the strongest correlation with delay in the planning phase of a megaproject. Also, the FOAK in a technological sense and the usage of an SPE is correlated with this outcome. Only one out of seven project characteristics from the project management category (see Table 9 in the Appendix) has been identified as important by either ML or FET ("there was a formal litigation procedure during the contract between client and EPC"). However, the small sample of data about these characteristics may hide existing correlations.

**Limitations and Challenges in Practical Application**

In this research, the following limitations of both FET and ML methods were identified:

1. FET and ML models require a representative data sample that is hard to collect in the context of megaprojects. In this research, special attention has been made to the processes of collecting relevant project cases and their preparation (see the four steps in the section on Data Collection and Preparation); and

2. Project characteristics were represented in the form of binary (yes/no) attributes, leading to the loss of information. The proposed representation was needed because of the nature of the FET. Concerning the applied ML techniques and the effects of the curse of dimensionality, the characteristics could be modeled with multivalued attributes once the number of project cases in the available database increases.

Specific limitations regarding FET are:

1. Because of the availability of only a small sample size, possible correlations between project success indicators and characteristics could be considered not significant because the p-value is not lower than a certain threshold. Therefore, these correlations are disregarded, leading to a type II error. A type II error is committed when we fail to believe a truth (Leach, 1979); and

2. The test only considers the correlations between a single project characteristic and a project management success (i.e., characteristic C correlated with performance P). This does not allow correlations resulting from multiple project characteristics to be unveiled (i.e., simultaneous occurrence of C1 and C2 correlated with P).

Opposite to FET, ML can be used to assess if a group of project characteristics is correlated with the success criteria (project characteristics used to train more performant prediction models are likely to be more correlated with project management success). However, ML cannot be generalized well from small datasets. Therefore, the applied protocols for model building and validation were adapted to minimize the effects of the low number of available data.

When comparing the possible application of ML to FET in an EPC company, ML requires substantially more expert effort and knowledge; the results of both methods require expert interpretation and validation. However, our findings show that the best-performing ML methods (logistic regression and decision tree) are interpretable by human inspection,
The Successful Delivery of Megaprojects: A Novel Research Method

as opposed to other ML black-box methods (such as neural networks).

The FET can be implemented in a regular Excel spreadsheet or even executed from several free websites. The execution of ML is more challenging. The proposed ML experiments were conducted using open-source Weka package issued under the GNU General Public License (MLGATUOW, 2017). Weka is a collection of ML algorithms for data-mining tasks. In order to conduct the training and testing protocols suitable for small datasets, we needed to adapt Weka accordingly.

Conclusions

Megaprojects are large, unique, and complex projects. Their uniqueness and complexity are the results of their physical elements and the dynamic network of the stakeholders involved. Consequently, it is challenging to set up a “lessons learned system” for them. Nevertheless, there are certain project characteristics in megaprojects (e.g., types of contracts, financing schemes, technological choices) that are quite standard. By investigating these characteristics, it is possible to discover common patterns behind successful and unsuccessful projects.

This article provides a method to identify in a quantitative and rigorous manner how megaproject characteristics relate to success indicators. First, it provides an initial understanding of how stakeholders in megaprojects can use this knowledge to ensure the more effective design and delivery of megaprojects. Second, the analysis of the empirical data using statistical techniques such as the FET and ML investigates the correlations between project characteristics and success indicators.

The results show that stakeholder characteristics are strongly correlated with success indicators (respecting time and cost overrun). This finding supports the existing understanding in the project management research community and provides invaluable reinforcement for further research into these factors.

The project environment, especially legal and socioeconomic characteristics, have also been identified as having an important relationship with megaproject success. The influence of SPE in megaproject management success, which has previously received scant attention, is of particular importance.

The investigation outlined in this article indicates that, if the successful delivery of megaprojects is to be secured, projects need to:

- Engage better with external stakeholders of the megaproject (and especially environmental groups), the affected population, and regulators; and
- Understand how to make the best use of SPEs in the governance of megaprojects.

What makes the contribution provided by this analysis really relevant for practitioners and policymakers is that they will have insight into the project characteristics correlated with project outcomes even before starting a project. By being aware of the characteristics listed in Table 3, the stakeholders involved in megaproject design and delivery can use the characteristics of their megaprojects to identify potential problems and make their projects more resilient.

This investigation provides a starting point for future research. The success indicators (for schedule and cost overrun performance) only provide a partial understanding of megaproject management success. A first logical extension is to add other indicators to assess a “quality” dimension. Similarly, further project characteristics can be added. If more cases are collected, it will be possible to use a chi-squared test, the key advantage of which would be the possibility of using more complex contingency tables, allowing other, more elaborate, hypotheses to be tested.

In addition, the initial effort presented in this article in proposing a megaproject success indicators prediction model could be continued when data from more cases become available.

In general, the methods and variables presented in this article can be applied to specific types of megaprojects and sectors. For example, an EPC company might want to apply the method to its portfolio of oil and gas projects with the specific form of contracts as project characteristics and average production in the first two years as project success indicators.

Acknowledgments

This research has been partially financed by the EU with the “Cost Action” scheme and the grant TU1003 MEGAPROJECT: The Effective Design and Delivery of Megaprojects in the European Union, funded by the European Science Foundation. Further information can be found at http://www.mega-project.eu/.

References


Meacham, T. (2012). *Renewable energy: Community benefit and...*


**The Successful Delivery of Megaprojects: A Novel Research Method**


**Dr. Giorgio Locatelli** is a lecturer of infrastructure procurement and management at the University of Leeds. He has bachelor’s and master of science degrees in mechanical engineering (2006) and a PhD in industrial engineering, economics, and management from the Politecnico di Milano (Italy) (2010). His research includes project management in infrastructure/megaprojects; a focus on cost-benefit analysis, risk management, ethics and corruption, governance and temporary organizations, financing, and modularization; energy systems sustainability; large energy infrastructure; system engineering management and a focus on infrastructures and complex systems. He is the author of more than 100 international publications—the majority of them focused on large engineering projects. He is a member of the CEO Council on Transformational Megaprojects—World Economic Forum, the editorial board of the *International Journal of Project Management and Progress in Nuclear Energy*, and the core group of IPMA—Special Interest Group—Megaprojects. He can be contacted at g.locatelli@leeds.ac.uk

**Assistant Professor Miljan Mikic** graduated from the Faculty of Civil Engineering, University of Belgrade in 2007; in 2015 he obtained his PhD in the field of Large Infrastructure Project Risk and Sustainability Management from the Department of Construction Project Management, Faculty of Civil Engineering, University of Belgrade; as the Assistant Professor in the same department, he teaches Construction Project Management and Construction Project Risk Management. He was engaged in several international research and consultancy projects related to infrastructure development. He was an MC member of the COST Action TU1003 “Megaprojects: The Effective Design and Delivery of Megaprojects in the European Union” and in 2016 he was awarded with the Coimbra Group Scholarship for Young Researchers from European Neighborhood. His areas of interest include construction project management, project risk and sustainability management, and building information modeling (BIM). He can be contacted at mmikic@grf.bg.ac.rs

**Professor Milos Kovacevic** graduated in 1995 from the Faculty of Electrical Engineering, University of Belgrade, Serbia. In 2007 he obtained his PhD in computer science from the Faculty of Electrical Engineering, University of Belgrade. He is a professor in the Department of Construction Project Management, Faculty of Civil Engineering, University of Belgrade, where he is responsible for the group of courses related to ICT application in construction and civil engineering. His research interests cover web and text mining, application of machine learning in construction and geo sciences, and distributed programming. He can be contacted at milos@grf.bg.ac.rs

**Naomi Brookes** is the CEO of Projektlernen Ltd. and a visiting professor in the School of Civil Engineering in the University of Leeds. She has researched extensively in the field of megaproject management and has worked with a range of organizations to improve their project management practice—from large manufacturing blue chips to investment bankers. She was the chair of the MEGAPROJECT COST Action, a research program that brought together the efforts of more than 80 researchers from 24 countries to enable learning across megaprojects. She is an invited participant for organizations such as the OECD, the World Economic Forum, and the Chinese Academy of Engineering. She can be contacted at n.j.brooks@leeds.ac.uk

**Professor Nenad Ivanisevic** graduated from the Faculty of Civil Engineering, University of Belgrade in 1986; he graduated from the Faculty of Law in 1995; and, in 2007 he completed his PhD in the areas of project management and contract law and economics, fields in which he now lectures at the Faculty of Civil Engineering in Belgrade. He has vast experience in planning and administering various large construction projects nationally and internationally, especially concerning construction of highways (construction of remaining sections of highways on Corridor X through Serbia, a highway between Podgorica and Matecevo in Montenegro). He was a member of the Management Board of “Koridor 10” LLC with executive power and a member of the State Committee for Expert Control of Projects of Importance for Republic of Serbia. He was an MC member of the COST Action TU1003: “Megaprojects: The Effective Design and Delivery of Megaprojects in the European Union.” He has also served as a special advisor to the Minister of Regional Development and Local Self Government and between 2011 and 2012 he served as the advisor for the Serbian Deputy Prime Minister for economy and regional development. He can be contacted at nesa@grf.bg.ac.rs
### Appendix

#### Table 4: Project stakeholders—Internal (SI).

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Operationalization</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI1</td>
<td><strong>NO (0)</strong></td>
<td><strong>YES (1)</strong></td>
</tr>
<tr>
<td></td>
<td>The EPC has its main headquarter in the county hosting the project</td>
<td>The EPC has its main headquarter in a foreign country</td>
</tr>
<tr>
<td>SI2</td>
<td>The EPC is delivering the infrastructure for a certain customer</td>
<td>The EPC will own the infrastructure</td>
</tr>
<tr>
<td></td>
<td>In some projects (e.g., Flamanville 3), the EPC will also be the owner of the infrastructure (Locatelli &amp; Mancini, 2012)</td>
<td></td>
</tr>
<tr>
<td>SI3</td>
<td>There aren’t any documents to back up this characteristic</td>
<td>There are documents to back up this characteristic</td>
</tr>
<tr>
<td></td>
<td>It is a key factor (Pinto &amp; Slevin, 1987; Pinto &amp; Mantel, 1990)</td>
<td></td>
</tr>
<tr>
<td>SI4</td>
<td>Client and EPC have different nationalities (main headquarters in different countries)</td>
<td>Client and EPC have the same nationality (main headquarters in the same country)</td>
</tr>
<tr>
<td></td>
<td>The impact of multiculturalism in project is stressed in the literature as a key aspect of project governance (Ofori &amp; Toor, 2009; Rees-Caldwell &amp; Pinnington, 2013; Ruuska et al., 2011; Swart &amp; Harvey, 2011)</td>
<td></td>
</tr>
<tr>
<td>SI5</td>
<td>Client, EPC, and all the important first-tier contractors have different nationalities (main headquarters in different countries)</td>
<td>Client and EPC and all the important first-tier contractors have different nationalities (main headquarters in the same country)</td>
</tr>
<tr>
<td>SI6</td>
<td>The national state owns directly or indirectly less than 50% of the share in the project</td>
<td>The national state owns directly or indirectly more than 50% of the share in the project</td>
</tr>
<tr>
<td></td>
<td>When the customer is the government, the project is managed differently and the risk pattern changes (Aritua et al., 2011)</td>
<td></td>
</tr>
</tbody>
</table>

#### Table 5: Project stakeholders—External (SE).

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Operationalization</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE1</td>
<td>No evidence of action from environmental groups</td>
<td>The project has been openly censured by international environmental groups such as Greenpeace</td>
</tr>
<tr>
<td></td>
<td>Concerns from environmental groups can trigger scopes change or even stop the project (Ross &amp; Staw, 1993). The real effectiveness is assessed with this variable</td>
<td></td>
</tr>
<tr>
<td>SE2</td>
<td>There are relevant protests or referendums against the project at a national level</td>
<td>The population living in that nation was supportive (or not objected) about the project</td>
</tr>
<tr>
<td></td>
<td>Public acceptability is often advocated as a precondition for project success (Brunsting et al., 2013; Kaldellis et al., 2013)</td>
<td></td>
</tr>
<tr>
<td>SE3</td>
<td>There are relevant protests or referendums against the project at a local level</td>
<td>The local population was supportive (or did not object) about the project</td>
</tr>
<tr>
<td></td>
<td>Public participation is a key fact and support toward a certain infrastructure can evolve over time (Drazkiewicz et al., 2015)</td>
<td></td>
</tr>
<tr>
<td>SE4</td>
<td>External stakeholders have been involved after construction started</td>
<td>External stakeholders have been involved before construction started, particularly in the planning process</td>
</tr>
<tr>
<td></td>
<td>In large construction projects, the early involvement of external stakeholders such as “environmental groups” has been suggested as a best practice to avoid issues as the NIMBY syndrome (Alexander &amp; Robertson, 2004)</td>
<td></td>
</tr>
<tr>
<td>SE5</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
</tr>
<tr>
<td></td>
<td>Public participation is a key fact and support toward a certain infrastructure can evolve over time (Drazkiewicz et al., 2015)</td>
<td></td>
</tr>
</tbody>
</table>
Independent Variable | Operationalization | Justification
--- | --- | ---
The project has a strong regulation system as evidenced by:

EL1 | (a) The safety authority stopping the project or very similar projects in the same country | The definition does not apply to the project | A strong regulatory system, in case of not compliance, can foster the EPC and its contractor to expensive scope changes (Locatelli et al., 2011; Ross & Staw, 1993)

EL2 | (b) The authority giving a fine to the EPC or one of the internal stakeholders in the project | The definition applies to the project |

EL3 | (c) Action from the authority postponing the final completion of the project |

EL4 | The project fits in the long-term plan of the country's government | There is no evidence to support how the project fits in the long-term plan of the country's government | Long-term view is often advocated as a key aspect of project delivery (Ahola et al., 2008; Park, 2009)

| Independent Variable | Operationalization | Justification
--- | --- | ---
ES1 | There is planned a long-term stability in usage and value | There is no evidence of long-term value/stability planned | Long-term view is often advocated as a key aspect of project delivery. (Ahola et al., 2008; Park, 2009)

ES2 | Financial support from the European Union (EU) | The definition does not apply to the project | Infrastructural projects partially financed by the European Union are supposed to go through an independent cost-benefit analysis and third-party appraisal (CBA Guide Team, 2008; Kelly et al., 2015)

ES3 | Financial support from the national government |

ES4 | Unemployment in the area is above the national average | Unemployment in the area is above the national average | The deployment of megaprojects in areas with high unemployment creates job positions useful to reduce the NIMBY problem (Invernizzi et al., 2017; Martinát et al., 2014)

ES5 | The majority of the national population trusts the national authority | There are documents (e.g., pools) showing the trust of the national population toward the national authority | The trust on the national authority is linked to public acceptability (He et al., 2013). However, a “trustful national authority” might impose very restricting measures to the project, increasing the risks

ES6 | The compensation to the local community is above 0.1 of the total budget | The definition does not apply to the project | The compensation to the local community is a way to increase the local public acceptability of the project (NEI, 2003; Meacham, 2012)

ES7 | The density of the population of the province is below the national average | The definition does not apply to the project | Some projects, particularly controversial ones, might be delivered in areas scarcely populated to reduce the risk of local protest (Barrett & Lawlor, 1997; Lindén et al., 2015)
### Table 8: Project environment—Political (EP).

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>NO (0)</th>
<th>YES (1)</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP1 Support of the national government (not local)</td>
<td>The national government has not supported the plan through direct financial subsidies, loan guarantee, and tax exception</td>
<td>The national government has supported the plan; this includes direct financial subsidies, loan guarantee, and tax exception</td>
<td>The government is a key player in the megaproject. It can have several roles and influences the performances directly and indirectly. For instance, several megaprojects are delivered as public-private partnerships (PPP) (Evers &amp; de Vries, 2013; Greco et al., 2017; Liu et al., 2016; Locatelli &amp; Mancini, 2014)</td>
</tr>
<tr>
<td>EP2 Support of the local government (not national)</td>
<td>There are no official documents or incentives or subsidies from the local government to support the project</td>
<td>There are official documents or incentives or subsidies from the local government to support the project</td>
<td></td>
</tr>
<tr>
<td>EP3 Support of both national and local government</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td></td>
</tr>
<tr>
<td>EP4 Not supported by either national or local government</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 9: Project management (PM).

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>NO (0)</th>
<th>YES (1)</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM1 The project uses planning by milestones</td>
<td>There is no evidence that the project manager used a “planning by milestone” approach</td>
<td>There is evidence that the project manager used a “planning by milestone” approach</td>
<td>These three variables test the impact of well-known project management tools and techniques. (Golini et al., 2015; Mir &amp; Pinnington, 2014)</td>
</tr>
<tr>
<td>PM2 The project uses formal project management tools and techniques</td>
<td>There is no evidence that the project manager heavily used formal project management tools and techniques— at least: Gantt chart, PERT (or simulation), risk analysis, earned value, cost schedule control system</td>
<td>There is evidence that the project manager heavily used formal project management tools and techniques—at least: Gantt chart, PERT (or simulation), risk analysis, earned value, cost schedule control system</td>
<td></td>
</tr>
<tr>
<td>PM3 Usage of performance metrics</td>
<td>There is no evidence that the project manager used performance metrics</td>
<td>There is evidence that the project manager used performance metrics</td>
<td></td>
</tr>
<tr>
<td>PM4 Turnkey contract between client and EPC/main contractor</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td>The type of contract influences project management success (Suprapto et al., 2016) and turnkeys are blamed for poor risk allocation and, therefore, performance (Aussa et al., 2009)</td>
</tr>
<tr>
<td>PM5 There was a formal litigation procedure (e.g., international chamber of commerce) during the contract between client and EPC</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td>The alignment of goals between the stakeholders is key for the project delivery. Litigation is an indicator of misalignment among stakeholders</td>
</tr>
<tr>
<td>PM6 Project has a well-developed FEED (front-end engineering design)</td>
<td>Frequent design amendments and elaborations</td>
<td>There is no change of the FEED during construction and the FEED was finished before construction started</td>
<td>A well-developed FEED is often considered a key success factor for the delivery of the project (Merrow, 2011)</td>
</tr>
<tr>
<td>PM7 An experienced project director is present</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td>Key factors suggested in Pinto &amp; Slevin (1987)</td>
</tr>
<tr>
<td>Independent Variable</td>
<td>NO (0)</td>
<td>YES (1)</td>
<td>Justification</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>T1 The megaproject is composed of more than one identical independent unit</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td>Modularization is often advocated as a strategy to make the project more manageable and delivery on time and on budget (Locatelli, Bingham, et al., 2014). Modularization can be intended in two ways: (1) the construction of several small units with a total capacity comparable to a large plant or (2) as the decomposition of a large structure in dependent prefabricated modules (1), or (2) as FOAK projects (in particular megaproject) have several unknown unknowns (Ramasesh &amp; Browning, 2014) jeopardizing the planning and delivery. Often, FOAK projects are late and have a cost overrun (Merrow, 2011)</td>
</tr>
<tr>
<td>T2 The megaproject is composed by modular—dependent modules</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td></td>
</tr>
<tr>
<td>T3 FOAK weak—country level</td>
<td>At least a similar project was delivered somewhere in the country</td>
<td>The plant is absolutely the first in the country or the design has radical modification respect to existing ones</td>
<td></td>
</tr>
<tr>
<td>T4 FOAK strong—global level</td>
<td>At least a similar project was delivered somewhere in the world</td>
<td>The plant is the absolutely the first in the world or the design has radical modification with respect to existing ones</td>
<td></td>
</tr>
<tr>
<td>T5 Industrial sector (energy, transport, miscellaneous)</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td>This variable assesses the correlation of sector with the performance and support the machine learning (ML) algorithm</td>
</tr>
<tr>
<td>T6 The project is a nuclear reactor</td>
<td>The project is not a nuclear reactor</td>
<td>The project is the construction or major refurbishment of a nuclear reactor</td>
<td></td>
</tr>
<tr>
<td>T7 Offshore project</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td>Merrow (2011) reports offshore projects as particularly problematic and affected by poor performance in the delivery</td>
</tr>
<tr>
<td>T8 The project physically connects two countries</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td>Projects connecting two countries (such as the well-known Channel Tunnel) represent a challenge from several perspectives, including technology, governance, and stakeholder management (Anguera, 2006; Genus, 1997)</td>
</tr>
<tr>
<td>O1 Previous similar project was on time and budget (N/A for FOAK)</td>
<td>The definition does not apply to the project</td>
<td>The definition applies to the project</td>
<td>The deployment of similar facilities might benefit from the industrial learning effect, leading to better cost estimation and project delivery performances (Choi et al., 2009; David &amp; Rothwell, 1996; Locatelli, Bingham et al., 2014)</td>
</tr>
<tr>
<td>O2 The project has a special purpose entity (SPE)</td>
<td>No SPE is involved in the delivery of the project</td>
<td>One or more SPE are involved in the delivery of the project as client and/or EPC</td>
<td>Special purpose entities are temporary organizations often involved in project planning and delivery. They might reconcile the interests of several stakeholders toward the common goals of the project (Sainati et al., 2017)</td>
</tr>
</tbody>
</table>

Table 10: Technological aspects (T) and other (O).