Article:

Invernizzi, DC orcid.org/0000-0001-8178-9557, Locatelli, G orcid.org/0000-0001-9986-2249 and Brookes, N (2018) A methodology based on benchmarking to learn across megaprojects: the case of nuclear decommissioning. International Journal of Managing Projects in Business, 11 (1). pp. 104-121. ISSN 1753-8378

https://doi.org/10.1108/IJMPB-05-2017-0054

A methodology based on benchmarking to learn across megaprojects:

the case of nuclear decommissioning

DC Invernizzi, G. Locatelli, N. Brookes 2017 *"A methodology based on benchmarking to learn across megaprojects: the case of nuclear decommissioning"* - International Journal of Managing Projects in Business

Please go to the journal website to find the exact Volume, Issue, Page and doi.

Abstract

The literature lacks a single and universally accepted definition of major and megaprojects: usually, these projects are described as projects with a budget above $1 billion and a high level of innovation, complexity & uniqueness both in terms of physical infrastructure and stakeholder network. Moreover, they often provide fewer benefits than what were originally expected and are affected by delays and cost overruns. Despite this techno-economic magnitude, it is still extremely hard to gather lessons learned from these projects in a systematic way.

This paper presents an innovative methodology based on benchmarking to investigate good and bad practices and learn from a portfolio of unique megaprojects. The methodology combines quantitative & qualitative cross-comparison of case studies and statistical analysis into an iterative process. Indeed, benchmarking offers significant potential to identify good and bad practices and improve the performance of project selection, planning and delivery.

The methodology is exemplified in this paper using the case of Nuclear Decommissioning Projects and Programmes (NDPs). Indeed, due to their characteristics, NDPs can be addressed as megaprojects, and are a relevant example for the application of the methodology presented here that collects and investigates the characteristics that mostly impact the performance of (mega)projects, through a continuous learning process.

Keywords: Megaprojects, Methodology, Benchmarking, Cross-Case Study, Nuclear Decommissioning.

1 Introduction

Major and megaprojects are often defined as projects with a budget above $1 billion with an high level of innovation and complexity (Flyvbjerg et al. 2003; Van Wee 2007; Merrow 2011; Locatelli, Mariani, et al. 2017). However, already in the mid 80ies, Warrack (1985) argued that $1 billion is not a constraint in defining megaprojects, since sometimes a relative approach is needed. In fact, in some contexts, a much smaller project (such as one with a $100 million budget), could constitute a megaproject. Similarly, Hu et al. (2013) claim that a deterministic cost threshold is not appropriate for all countries, and a relative threshold such as the GDP should be used instead.

Even without defining a single threshold, megaprojects share the characteristics of not only being extremely expensive and long, but also politically sensitive, since they are often commissioned at least partially by the Government and involve a large number of external and internal stakeholders. Moreover, these projects are both influenced by the context in which they are delivered and they are able to influence the context themselves (Merrow 2011). Additionally, due to the size and complexity of both their physical infrastructure and stakeholder network, it is still extremely hard to gather and investigate lessons learned from these projects in a systematic way.

Due to this techno-economic, political and social magnitude, megaprojects have risen significant interest not only among practitioners, but also among academics. Nevertheless, due to their uniqueness, it is still extremely hard to gather good and bad practices and develop empirically-based guidelines in a systematic way.

This paper addresses this challenge, presenting a methodology to improve learning across projects and uncover good and bad practices, and ultimately investigates the project characteristics (i.e. the independent variables) that impact most on the project performance (i.e. the dependent variable).

This methodology is based on benchmarking. Indeed, “benchmarking” refers to the process of comparing projects and, as explained in section 2, it offers significant potential to investigate the characteristics that impact most on the project performance. This methodology is applied to Nuclear Decommissioning Projects and Programmes (NDPs), as NDPs are extremely complex, long and expensive, with a budget that often exceed $1 billion; they are politically sensitive and involve a large number of external and internal stakeholders (LaGuardia & Murphy 2012; D.C. Invernizzi, G. Locatelli, et al. 2017). Therefore NDPs can be addressed as megaprojects.

Nevertheless, this methodology can be adapted on all major and megaprojects where the uniqueness of projects and the low number of cases available hinder the use of analysis based on big numbers.

The rest of the paper is organised as follows. Section 2 critically reviews recent researches on benchmarking and compares benchmarking studies applied on the construction industry. Section 3 stems from the literature and proposes a methodology to adapt the benchmarking approach to the situation where the number of cases is low and the information available are scattered. Then, this methodology is exemplified using the case of Nuclear Decommissioning Projects and Programmes (NDPs) in section 4. Nevertheless, this methodology is suitable to be applied on major and megaprojects where the alleged uniqueness of projects and the low number of cases available hinder the use of analysis based on big numbers. Section 5 is dedicated to discussion and conclusions.

2 Benchmarking analysis in the literature

The meaning of the term "benchmarking" has been widely discussed in the last decades and, as shown in Table 5 in the appendix, there are different definitions of "benchmarking" and of the benchmarking "steps and/or phases" in the literature. Already in 1992, benchmarking had already been described through 49 definitions (Anand & Kodali 2008, quoting (Spendolini, 1992)) and through a different number of steps and phases. More recently, Anand & Kodali (2008), reviewed 35 published models and highlighted that there are only 13 common steps of the benchmarking analysis, out of 71 investigated. Therefore, before performing a "benchmarking analysis", it is critical to agree on its definition. In this research, the authors follow the PMBOK definition (2013, p.116), where benchmarking involves "comparing actual or planned practices, such as processes and operations, to those of comparable organizations *to identify best practices, generate ideas for improvement"* and it provides *"a basis for* measuring performance". Additionally, Garnett & Pickrell (2000, p.57) assert that benchmarking is "a continuous process of establishing critical areas of improvement within an *organization [...]"*, that it offers "the means to *identify why `best practice' organizations are* high achievers, and how others can learn from best practice processes to improve their own approach". Ramirez et al (2004) also state that it is necessary to complement a quantitative benchmarking system with a qualitative based one, in order to establish causal relationships. This demonstrates the need to adapt benchmarking case-by-case. Within the construction industry, benchmarking has already been used to compare projects in order to identify successful projects and the reasons for their success, and the interest in benchmarking is significantly increasing because, through finding examples of superior performance, firms can adjust their policies and practices to improve their own performance (El-Mashaleh et al. 2007; Costa et al. 2006; Ramirez et al. 2004). Table 6 in the appendix compares benchmarking analysis applied to the construction industry, highlighting, per each study, (1) the aim of the research, (2) the method or model described or adopted, (3) the steps of the analysis and highlights, and (4) the data collection and the number of case studies investigated.

Concerning benchmarking applied to the construction industry (Table 6), the following conclusions can be drawn:

- The benchmarking analysis is suitable to determine the performance of a company, using input metrics (i.e. safety expenses and management expenses) and output metrics (i.e. schedule adherence, cost performance, customer satisfaction, safety performance, and profit) (El-Mashaleh et al. 2007).
- *"L*essons learned from other companies can be used to establish improvement targets and to promote changes in the organization" (Costa et al. 2006, p.158), but there is a need to upgrade existing benchmarking initiatives and devise new ones.
- Qualitative benchmarking can enable the comparison of management practices, discover relationships between performance data, and determine industry trends. Also, being based on the perception of key personnel, this approach can be applied as part of a continuous improvement programme (Ramirez et al. 2004).
- The benchmarking process is as important as the benchmarks themselves (Garnett & Pickrell 2000), therefore the selection of cases is pivotal.

In conclusion, the benchmarking analysis is recognized to be a valuable tool to improve the performance of projects delivered in different industrial sectors and in different countries. However, the aforementioned analyses are not directly applicable when the number of projects is low and/or the information available scattered (e.g. construction megaprojects and NDPs), and where a single and globally-recognized benchmark is missing. Therefore, a new framework needs to be developed to deal with the complexity and low number of major and mega projects. Table 1 compares a few techniques for benchmarking and highlight those that are suitable for megaprojects. Section 3 explains this framework, which is exemplified in section 4 using the case of NDPs. Other statistical analysis, such a qualitative comparative analysis (Schneider & Wagemann 2012), will be considered at a later stage of the research.

Ref, aim of the paper and data collection	**Method, model or analysis implemented**	**Applicable for benchmarking megaprojects?**
"Benchmarking System for Evaluating Management Practices in the *Construction Industry"* (Ramirez et al. 2004) This paper presents the results from the application of the benchmarking system through different methods, i.e. qualitative benchmarking, correlation analysis, factor analysis, multivariate linear regression and sectors trends. Thirteen companies participated to the initial application of the benchmarking system.	(1)**Qualitative benchmarking** with the class median, used to enable each company to evaluate its position compared to the worse and best case scenario and the median. This comparison is highlighted using the Radar graph.	Yes, qualitative benchmarking is suitable between 2 or 3 megaprojects. However, it is not suitable to calculate the median, due to the low number of projects.
	(2)**Correlation analysis,** used to investigate the intensity of the linear relationship between two variables, Xi and Xj. To measure this intensity of the correlation, the Pearson's coefficient is used. The Pearson's correlation is a measure of the strength and direction of the linear relationships that exists between two variables measured on an interval scale.	No, as to use the Pearson's correlation, variables should be approximately normally distributed and there should be no significant outliers (Laerd Statistics 2016). Moreover, the cases should represent a random sample from the population. These assumptions are not met by megaprojects.
	(3)**Factor analysis**, that uses the principal components to determine the underlying structure among the different management dimensions and identify relationships not previously established.	No, as the principal component analysis requires assumptions (e.g. linearity (Shlens 2005)), that are not met by megaprojects.
	(4)**Multivariate linear regression**, that was implemented but discarded due to the weak correlation coefficient caused by the low number of data quantity of data.	No, as assumptions for the multivariate linear regression (e.g. linearity, homoscedasticity, etc.) are not met by megaprojects.
	(5)**Sector trends** by management dimensions, by job categories, and by subsectors are used to categorize and analyse survey results.	Yes, as trends highlighted during the descriptive analysis of the collected data can yield interesting conclusions.
"Management of Construction Firm Performance Using Benchmarking" (El-Mashaleh et al. 2007) This research presents a benchmarking model that uses input metrics to determine the company performance. Data were collected from 74 construction firms through a survey,	**Data Envelopment Analysis** (DEA). DEA is concerned with evaluation of the activities of organizations such as business firms, hospital and government agencies. The organization responsible for converting inputs into outputs is called Decision Making Unit (DMU). DEA uses mathematical linear programming to determine which of the DMU forms an envelopment surface, i.e. an efficient frontier.	No, as the number of megaprojects and the information available is too low to implement the DEA.
"Power plants as megaprojects: Using empirics to shape policy, planning, and *construction management"* (Brookes & Locatelli 2015) This paper investigates the correlation between characteristics of power plant megaprojects and their costs and schedule cost performance.	This research implements the **Fisher Exact Test** to a dataset of o a dataset of 12 case studies from several industries, e.g. the nuclear, coal, and renewable resources. The Fisher Exact Test investigates the correlation of single independent variables vs dependent ones and is able to identify correlations within small data sets.	Yes, as the Fisher Exact Test is able to identify correlations within small data sets (< 30 cases), as it investigates each project characteristics independently.
*"Empirical researc*h on infrastructural megaprojects: what really matters for their successful delivery*"* (Locatelli, Invernizzi, et al. 2017) This research investigates the relationship between project characteristics and performance using a pool of 44 case studies.	This paper implements **the Fisher Exact Test** and **Machine Learning** techniques. Machine Learning enable rigorous "pattern spotting" analysis of the existing, relatively small dataset, which did not allow the application of multivariate statistical analysis. Three different learning methods are implemented, i.e.: Decision tree, Naïve Bayes and Logistic Regression	Yes, both the Fisher Exact Test and Machine Learning are applicable to megaprojects. Being the Logistic Regression a type of probabilistic model used to predict the class based on one or more attributes (not necessarily continuous), it can be applied to the case of megaprojects.

Table 1. Techniques for benchmarking (Invernizzi, Locatelli, et al. 2017)

3 The methodology to learn across megaprojects

The methodology presented here is a development of the seminal work by Kathleen Eisenhardt (1989), who recommends data collection using multiple methods, introduces the concept of "theoretical saturation", and promotes the deep analysis both of a single case and across case-studies to develop theories. In particular, the cross-case comparison is an iterative process, where the first step refines the initial hypothesis, the second step verifies the relationships among hypothesis and empirical evidence, and the third step critically compares new theories with existing ones. The case-method is described by several authors (e.g. (Yin 2009)) and is of significant interest for the current research, even if it is sometimes criticized due to its limited rigour (Easterby-Smith et al. 2012).

The methodology proposed here is largely based on empirical evidence, and employs an "inductive" method (rather than a "deductive" one) where "induction" is defined as follows (Gill & Johnson 2002; Brookes et al. 2015, p.6): *"the induction of particular inferences from* particular instances or the development of a theory from the ob*servation of empirical reality."* Figure 1 shows the research framework that has been developed by the authors to ultimately collect good and bad practices, and investigate what drives the project performance.

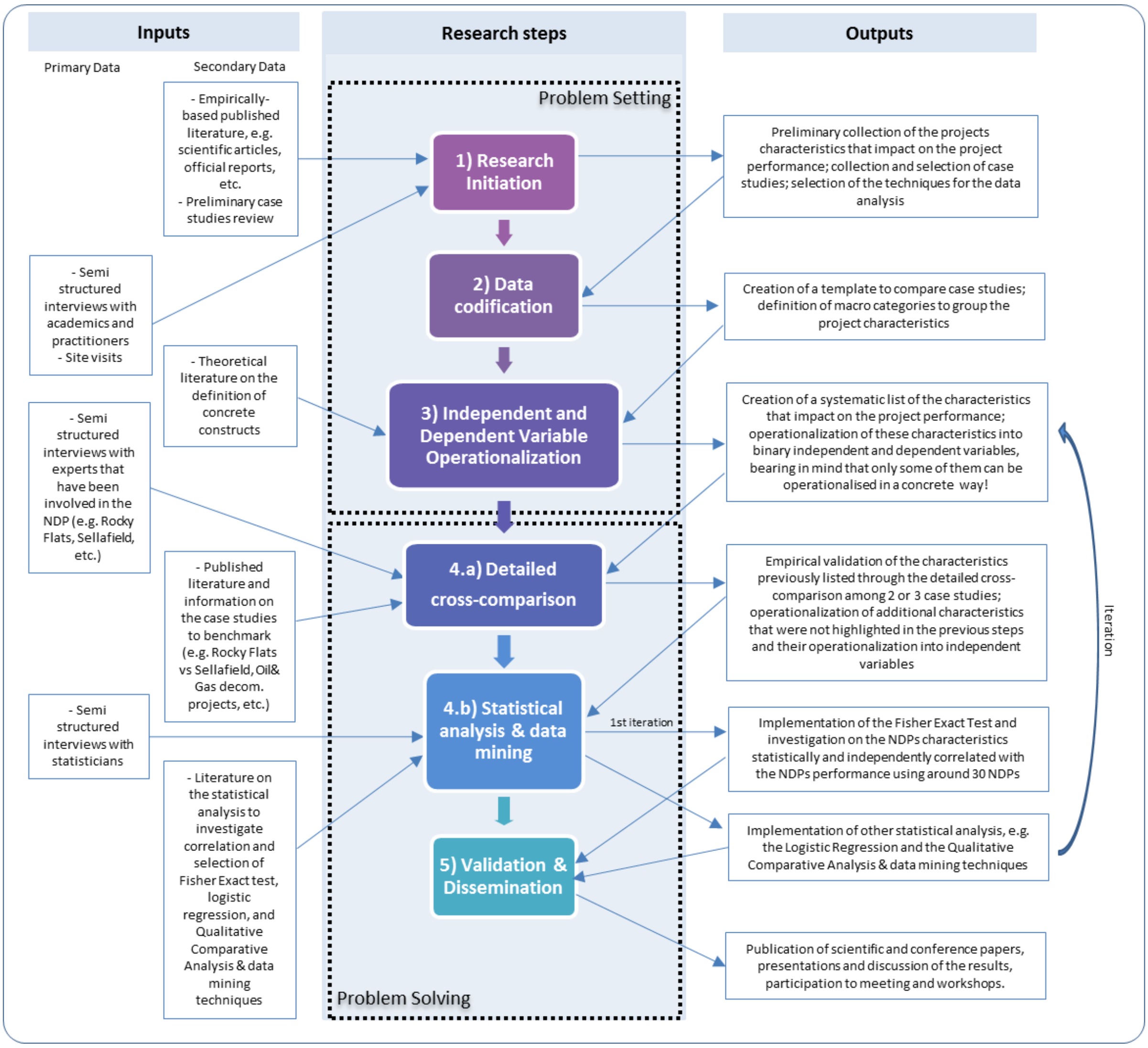

Figure 1. The five-step methodology

The first step embraces a preliminary literature review and the collection of case studies. This is complemented by semi-structured interviews and site visits. Case studies are selected according to their relevance and their completeness, and the availability of information. The date when these projects have been delivered is also significant, so the rule “the newer the better” applies. The output of the first step is (I) the preliminary collection of the projects’ characteristics that impact on the projects’ performance, (II) the selection of case studies and (III) of the techniques for the data analysis.

The second step consists of the data codification. Indeed, the selected case studies need to be recorded in a standard template. This template contains several pieces of information grouped into macro-categories, such as:

- an overview of the project, its physical characteristics and its final end-state;
- governance, funding and contacting schemes;
- stakeholders & stakeholder engagement.

The output of the second step is the development and population of a standard template to allow an easier comparison of projects. From this template, lessons learned can been listed and analysed.

The third step consists of the operationalisation of the independent and dependent variables, i.e. respectively the project characteristics and their performance. To do this it is necessary to firstly differentiate between "concepts" and "constructs", where a construct is a more formalised definition of a concept, a concept being a "general idea in our heads about a variable which has a part to play in our theories" but that still cannot be observed directly (Lee & Lings, 2008). The measurement of a construct is "the process of moving our theoretical constructs into the real world" […], therefore "once we work out exactly how we can represent our constructs in the real world, we have what can been called an operational definition" […]. So, the operational definition outlines exactly "what in the real word we say represents our theoretical constructs" (Lee & Lings 2008, p. 161) and implicitly means that operational definitions and constructs are not the same thing, as shown in Figure 2. Constructs can describe the world, which is qualitative, quantitative, complex and dynamic. However, these constructs are not directly observable, therefore observable measures have to be used instead.

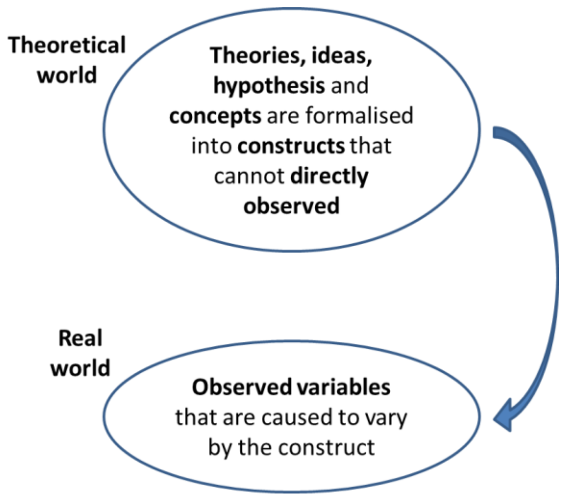

Figure 2. Theoretical word and real world, adapted from (Lee & Lings 2008)

Rossiter (2002) adopts the definition of Edwards & Bagozzi (2000) that describes constructs as phenomena "of theoretical interest" and suggests describing them in terms of (1) the object, including its constituents or components, (2) the attribute, including its components, and (3) the rater entity, where:

- The object part of the construct can be singular, collective or have multiple components, and can be concrete or abstract;
- The attribute in the construct is the dimension on which the object is being judged, and can be concrete singular, abstract formed, and abstract eliciting;
- The rater can be individual, expert or a group.

The output of the third step is the creation of a systematic list of the characteristics that impact on the project performance and their operationalization into binary independent and dependent variables, bearing in mind that, due to their nature, only some of them can be operationalised in a concrete way. The fourth step consists of the actual data analysis and it is split into two stages, i.e. the qualitative & quantitative cross-comparison and the statistical analysis and data mining, respectively 4.a. and 4.b in Figure 1.

The qualitative & quantitative cross-comparison of step 4.a highlights the good practices that empirically resulted to be relevant for the successful performance of a project. The correlation of these practices, together with “lessons learned” gathered from published literature (e.g. journal articles, official reports, case studies), semi structured interviews with experts and site visits is then investigated in step 4.b.

The statistical analysis employed needs to address: 1) the low number of cases and 2) their complexity, in other words, their (alleged) uniqueness. This is why the Fisher Exact Test is implemented first. Indeed, the Fisher Exact Test is able to identify correlations within small data sets (Leach 1979), e.g. 20-30 projects and to evaluate whether or not a single independent variable (e.g. a project characteristic) is associated with the presence (or absence) of a dependent variable (e.g. the project performance), using categorical data in the form of a contingency table as input. The output of the test is a p-value, which represents how likely it is that the result detected by the implementation of this statistical analysis could have resulted from chance rather than due to a real relationship between the variables in question. In this respect, the smaller the “p-value” is, the better. Key features, limitations and the implementation of the Fisher Exact Test can be found in (Brookes & Locatelli 2015; Locatelli, Invernizzi, et al. 2017).

Regarding the value of the p-value, the authors suggest to adopt a higher significance level than the one traditionally used, such as a $p\text{-value} < 0.15$ rather than a more typical value of $p\text{-value} < 0.05$. This means that statistically significant findings must be dealt in a circumspect fashion and the actual causation between project characteristics and their performance requires further investigation and validation, e.g. through pilot projects and interviews with experts.

4 Example of results from the cross-comparison and the statistical analysis

NDPs are complex and affected by high uncertainties, can be characterized by activities that reach national multibillion projects, involve large numbers of partners and stakeholders, and are often (at least partially) commissioned by governments. Therefore, this paper addresses NDPs as megaprojects and uses NDPs to exemplify the methodology described in section 3. Section 4.1 and 4.2 exemplify some results from the cross-comparison, while 4.3 exemplifies preliminary findings regarding the statistical analysis.

4.1 Cross-comparison between two "similar but different" NDPs

The cross-comparison of NDPs assists the collection of relevant good (and bad) practices, and therefore is envisaged to be performed both within the UK and against other countries (Table 2).

	Nuclear	**Non-nuclear**
UK	(1) Benchmarking NDPs across the UK	(3) Benchmarking non-nuclear decommissioning projects across the UK
Non-UK	(2) Benchmarking NDPs across several countries	(4) Benchmarking projects across countries and in different industrial sectors

Table 2. Benchmarking across decommissioning projects

Some of the "lessons learned" from the comparison of two similar-but-different NDPs are highlighted below, where the cross-comparison between Rocky Flats (US) and Sellafield (UK) is summarized in Invernizzi et al. (2017). Lessons learned from ten Oil & Gas decommissioning projects are also collected and summarized in section 4.2.

Rocky Flats (US) and Sellafield (UK) represented in Figure 3 are compared because these two NDPs:

- Are recent NDPs;

- Share a reasonably similar history (e.g. both facilities were opened for military purposes in the 1940s/1950s and have been affected by major nuclear accidents);
- Have a comparable size;
- Had a decommissioning budget in the order of tens of billions of $.

Moreover, there is publically available information in English regarding both these NDPs.

Rocky Flats was a military nuclear weapons facility in Colorado that produced plutonium and enriched uranium from 1953, and stopped operations in 1989. It was owned by the US Department of Energy (DOE) and was managed by a series of weapons contractors. When Rocky Flats was shut down in 1989, due to the significant radioactivity on site, the US DOE estimated it would have taken 70 years and $36 billion to decommission it. The project was however completed by a joint venture in less than 10 years and $ 3.5 billion (DOE 2013; Cameron & Lavine 2006; Bodey 2006). Sellafield is a 6 square kilometres UK nuclear site that contains 99% of the UK radioactivity. The UK NDA estimates it would take more than 100 years and £ 53 billion to decommission it (NDA 2016; Sellafield Ltd 2016).

Nevertheless, these NDPs have also very different aspects. Rocky Flats stopped operations 1989. Conversely, Sellafield is still an operating nuclear site that handles radioactive material shipped both from other countries and other nuclear sites in the UK. Moreover, during its decommissioning, its waste was shipped to other states in the US. Conversely, Sellafield is still an operating nuclear site that handles radioactive material shipped both from other countries and other nuclear sites in the UK (Invernizzi et al. 2017).

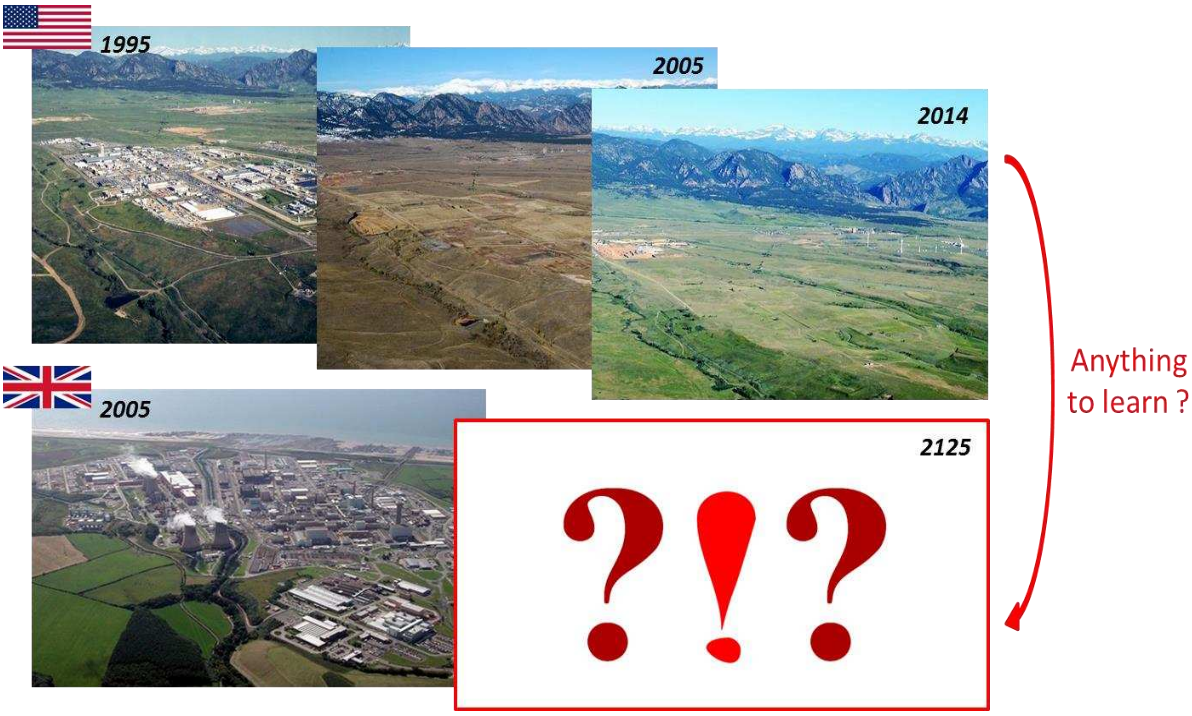

Figure 3. Rocky Flats NDP vs Sellafield NDP
(Source of the single pictures combined by the authors: Wikipedia)

Despite these differences, early results and "lessons-learned" from this cross comparison are remarkable. Within the others:

- **Funding arrangements and contracting schemes, especially if tailored on single employees.** Indeed, Rocky Flats adopted the so-called "abundance approach", where the aim was to fill the gap between forecasted (successful) performance and "spectacular" performance, i.e. to achieve positive deviance by closing the abundance gap (Cameron & Lavine 2006). This, together with incentives singularly allocated to employees to promote feasible ideas can improve the performance of the NDP.
- **The size of the free space available within the perimeter of the nuclear site to manage radioactive waste**. In fact, even if the size of Rocky Flats is comparable to Sellafield, Rocky Flats had a "buffer zone" which surrounded the site that proved to be helpful for the management of radioactive material (Cameron & Lavine 2006). Sellafield, on the

contrary is “packed with buildings” (informal talks with Sellafield employees), which hinders the construction of new facilities to treat and confine the radioactive material.

- **Early and timely engagement of stakeholders.** Indeed, effective communication and the involvement of stakeholders in collaborative action support the smooth delivery of the project (Invernizzi, Locatelli, et al. 2017).

These empirically-based lessons learned, together with good and bad practices gathered from the literature are tested with the statistical analysis of step 4.b.

4.2 Cross comparison among Oil & Gas decommissioning projects

The decommissioning of North Sea Oil & Gas facilities in the North Sea is projected to cost £40.6 billion over the next 25 years, £16.9 billion of which will be over the next decade. This is an increase of £2.3 billion on the 2014 report’s ten-year forecast of £14.6 billion and is primarily due to 47 new projects entering this year’s survey (Oil&Gas UK 2015; Offshore Energy 2014). As NDPs, not only Oil & Gas decommissioning projects:

- have a multi-million budget;
- are partly funded by the government;
- are affected by a highly regulated environment;
- are affected by cost overrun and schedule slippage ;
- have a potentially high environmental impact, as Natural Occurring Radioactive Material might build up, which might cause unexpected radiological issues
- are less uncertain than NDPs, but still are affected by high uncertainties.

Therefore, the lessons learned from these ten Oil & Gas projects (summarized in Table 3) are also considered to populate the list of project characteristics whose correlation with the project performance is tested through the statistical analysis applied on NDPs. Indeed, most of the cost and schedule drivers highlighted in Table 3 are shared with the nuclear decommissioning industry and are therefore tested through the statistical analysis of step 4.b.

Case Study	Within budget ?	Within schedule ?	Cost & Schedule drivers highlighted in the close out reports
Frigg Field (Total E&P Norge AS 2011) (Total E&P Norge AS 2003)	No	Yes	More complex operation than originally foreseen; change in removal method
North West Hutton (NWH) (Jee 2014) (British Petroleum 2006)	No	No	General lack of track record; lack of available benchmarking; pipeline cutting and removal taking longer than expected; additional vessel mobilisations were necessary; trenching activities took longer than scheduled due to the soil type encountered; delayed due to the intention to combine NWH decommissioning scope with other works to enable technical synergies and cost efficiencies
Indefatigable (Shell E&P UK 2014) (Shell U.K. Limited 2007)	No	No	Costs figures to estimates cost were too old and not corrected with inflation
Linnhe Field (Shell E&P UK 2014) (Shell U.K. Limited 2007)	No	No	Inclement weather; greater than estimated duration of the work; need of a guard vessel
Manifold and Compression Platform (MCP) – 01 (Total E&P UK 2013) (Total E&P UK 2007)	No	Yes	Additional engineering required to ensure a safe and stable removal activity; additional man-hours required to execute the significantly larger work scope; presence of hazardous materials not previously recorded on register gave increased activity both offshore and onshore; additional time at site required additional flotel attendance; more visits by heavy lift vessel required than had been estimated; the decision by the contractor to use the "piece-small" removal process on a large scale; difficulty to contract enough experienced and skilled labour; knowledge management; pre-qualification of "new" techniques should be conducted; control on the availability and reliability of cranes and tools
Kittiwake SAL Export System (Centrica Energy 2012) (Venture Production plc 2009)	No	No	The over-spend related to the cost of preservation and onshore storage made necessary by early recovery (£0.4m) and the weather delay during load in (£0.3m); availability of a suitable vessel; potential synergies with other projects
Shelley (PremierOil 2015) (PremierOil 2010)	Yes	Yes	The impossibility of utilising water jetting methods; the re-use of end fittings to be re-used to make new jumpers; the complexity of the recovery of the "Polyoil" resin-based cable clamps, due to complete disintegration; the discharge of oil-based mud residue during the wellhead cut
Tristan NW (bridge energy 2010) (Silverstone Energy Limited 2010)	No	No	Operational and extensive weather delays
Fife, Fergus, Flora and Angus (FFFA) (HESS 2014)(HESS 2012)	No	Yes[1]	Additional scope of work
Camelot (Helix Energy Solutions 2013) (Energy Resource Technology Ltd 2012)	Yes	Yes	Impact of processing naturally occurring radioactive material
Total Positive	**2/10**	**5/10**	Only two of the Oil & Gas decommissioning projects were completed within the estimated budget. Five over ten were completed within the schedule

Table 3. Summary 10 selected Oil & Gas decommissioning case studies (DECC 2016)

[1] Not explicit

4.3 Example of results from the first iteration of the statistical analysis: the Fisher Exact Test

The output of the statistical analysis is to highlight the correlation between the project characteristics and their performance. Table 4 lists four country-specific independent variables, two of which resulted in being correlated with the project performance according to the first statistical test implemented (i.e. the Fisher Exact Test) to a pool of 30 NDPs. This is a preliminary result, applied on a pool of NDPs. This research aims to increase the number of NDPs to improve the reliability of the results of the statistical analysis.

Independent variables, i.e. the NDP characteristics	**Correlation of the independent variables with the dependent variable "50% cost overrun"**
The country scores a corruption perception index > 60[2]	The fact that the corruption perception index in a country is less than 60 is correlated with **the presence** of 50% of cost overrun. The p-value is lower than 10%, showing a correlation.
The legal timeframe for review of decommissioning plans is less 2 years	The fact that the legal timeframe for review of decommissioning plans is less 2 years is strongly correlated to **the absence** of 50% of cost overrun. The p-value is lower than 10%, showing a correlation.
There are other nuclear facilities still operating in the country	The fact that there are other nuclear facilities operating in the country **is not correlated** to the absence of 50% of cost overrun. The p-value is >>15%, showing no correlation.
The NDP is state owned	The fact that the NDP is state owned **is not correlated** with the absence of 50% of cost overrun. The p-value is >>15%, showing no correlation.

Table 4. Example of independent variables statistically correlated to 50% cost overrun

[2] From Transparency International, as in (Locatelli, Mariani, et al. 2017)

5 Discussion and Conclusion

Due to their techno-economic, political and social magnitude, megaprojects have risen significant interest not only among practitioners, but also among academics. However, it is still extremely hard to gather good practices and develop empirically-based guidelines in a systematic way.

This paper presents an innovative methodology based on benchmarking, that combines qualitative and quantitative analysis to collect, select and investigate good and bad practices and learn from a portfolio of projects. This methodology is exemplified using the case of decommissioning projects (and nuclear ones in particular), which are still remarkably under investigated, although extremely significant in terms of complexity and budget.

The methodology proposed in this paper starts with the selection of representative case studies (megaprojects) and the listing of the project characteristics that impact on the project performance, according to the literature. This is supported by semi structured interviews, and the cross-comparison and qualitative analysis of the information collected. Then, this methodology suggests to apply statistical analysis to assess the correlation and causation between project characteristics and their performance and to validate the correlation through pilot projects.

In particular, the Fisher Exact Test is envisaged to be applied as the first statistical method, as it is able to identify correlations within small data sets and to evaluate whether or not a single independent variable (e.g. a project characteristic) is associate with the presence (or absence) of a dependent variable (e.g. the project performance). The output of the test is a p-value, which represents how likely the result detected by the implementation of this statistical analysis could have resulted from chance rather than due to a real relationship between the variables in question. Other statistical analysis & data mining techniques can be applied at later stages of the research.

References

Anand, G. & Kodali, R., 2008. Benchmarking the benchmarking models. Benchmarking: An International Journal, 15(3), pp.257–291.

Bodey, E., 2006. Making the Impossible Possible: Closing Rocky Flats: ahead of schedule and under budget. Radwaste Solutions, (October 2005), pp.39–45.

bridge energy, 2010. Tristan NW Field Decommissioning Programmes - Close Out Report, Available at: https://whitehall-admin.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/43398/5022-tristan-nw-close-out-report.pdf.

British Petroleum, 2006. North West Hutton - Decommissioning Programme, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/43406/nw-hutton-dp.pdf.

Brookes, N., Locatelli, G. & Mikic, M., 2015. Learning Across Megaprojects, Available at: http://www.mega-project.eu/assets/exp/docs/Learning_Across_Megaprojects.pdf.

Brookes, N.J. & Locatelli, G., 2015. Power plants as megaprojects: Using empirics to shape policy, planning, and construction management. Utilities Policy, 36, pp.57–66.

Büyüközkan, G. & Maire, J.-L., 1998. Benchmarking process formalization and a case study. Benchmarking: An International Journal, 5(2), pp.101–125.

Cameron, K. & Lavine, M., 2006. Making the Impossible Possible: Leading Extraordinary Performance, The Rocky Flats Story, Berrett-Koehler.

Camp, R., 1989. Benchmarking - The Search for Industry Best Practices That Lead to Superior Performance, Milwaukee, WI: ASQC Quality Press.

Centrica Energy, 2012. Kittiwake SAL Export System Decommissioning Programme - Close Out Report, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/43402/6213-kittiwake-sal-close-out-report.pdf.

Costa, D. et al., 2006. Benchmarking initiatives in the construction industry: lessons learned and improvement opportunities. Journal of Management in Engineering, 22(4), pp.158–168.

DECC, 2016. Oil and gas: decommissioning of offshore installations and pipelines - Detailed guidance. UK GOvernment official website. Available at: https://www.gov.uk/guidance/oil-and-gas-decommissioning-of-offshore-installations-and-pipelines [Accessed February 24, 2016].

DOE, 2013. Rocky Flats Closure Legacy report - Office of Legacy Management. Office of Legacy Management official website. Available at: http://www.lm.doe.gov/Rocky_Flats_Closure.pdf#TOC [Accessed June 16, 2016].

Easterby-Smith, M., Thorpe, R. & Jackson, P.R., 2012. Management Research, SAGE Publications.

Edwards, J.R. & Bagozzi, R.P., 2000. On the Nature of Direction of Relationships between Constructs and Measures. Psychological methods, 5(2), pp.`155–174.

Eisenhardt, K.M., 1989. Building Theories from Case Study Research. The Academy of Mangement Review, 14(4), pp.532–550.

El-Mashaleh, M., Minchin, R.E. & O'Brien, W.J., 2007. Management of construction firm performance using benchmarking. Journal of Management in Engineering, 23 (1)(January), pp.10–17.

Energy Resource Technology Ltd, 2012. Camelot CA Platform, Camelot CA Pipelines, Camelot CB Pipelines Decommissioning Programmes, Aberdeen (UK). Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/255075/Camelot_dp.pdf.

Fernandez, P., McCarthy, I.P. & Rakotobe-Joel, T., 2001. An evolutionary approach to benchmarking. Benchmarking: An International Journal, 8(4), pp.281 – 305.

Flyvbjerg, B., Bruzelius, N. & Rothengatter, W., 2003. Megaprojects and Risk: An Anatomy of Ambition, Cambridge University Press.

Garnett, N. & Pickrell, S., 2000. Benchmarking for construction: theory and practice. Construction Management and Economics, 18(1), pp.55–63.

Gill, J. & Johnson, P., 2002. Research Method for Managers Fourth Edi., SAGE Publications.

Helix Energy Solutions, 2013. Camelot CA Platform, Camelot CA Pipelines, Camelot CB Pipelines Decommissioning Programmes - Close Out Report, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/239536/D13_957976__Helix_ERT_Close_out_report_FINAL_dated_14th_Aug_13._DT.PDF.

HESS, 2012. Fife, Fergus, Flora and Angus Field - Decommissioning Programmes, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/43395/4884-fife-flora-fergus-angus-decomm-prog.pdf.

HESS, 2014. Fife, Fergus, Flora and Angus Fields Decommissioning Programmes - Close Out Report, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/478970/FFFA_Close-Out_Report.pdf.

Invernizzi, D.C. et al., 2017. Similar but different: a top-down benchmarking approach to investigate Nuclear Decommissioning Projects - paper accepted for the International Conference on Nuclear Engineering 2017. In ICONE25. Shangai, China.

Invernizzi, D.C., Locatelli, G. & Brookes, N.J., 2017. How benchmarking can support the selection, planning and delivery of nuclear decommissioning projects - accepted for publication. Progress in Nuclear Energy.

Invernizzi, D.C., Locatelli, G. & Brookes, N.J., 2017. Managing social challenges in the nuclear decommissioning industry: a responsible approach towards better performance - article in press. International Journal of Project Management. Available at: http://dx.doi.org/10.1016/j.ijproman.2016.12.002.

Jee, 2014. North West Hutton Decommissioning Programme - Close Out Report, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/371545/NWH_Decommissioning_Programme_Close_Out.pdf.

Kyro, P., 2004. Benchmarking as an action research process. Benchmarking: An International Journal, 11(1), pp.52–73.

Laerd Statistics, 2016. Pearson's Product-Moment Correlation using SPSS Statistics. Laerd Statistics Official Website.

LaGuardia, T.S. & Murphy, K.C., 2012. Financing and economics of nuclear facility decommissioning. In Nuclear Decommissioning: Planning, Execution and International Experience. pp. 49–86. Available at: http://linkinghub.elsevier.com/retrieve/pii/B978085709115450004X.

Leach, C., 1979. Introduction to Statistics: a nonparametric approachfor the social science John Wiley., New York.

Lee, N. & Lings, I., 2008. Doing Business Research: A Guide to Theory and Practice, SAGE Publications.

Locatelli, G., Mariani, G., et al., 2017. Corruption in public projects and megaprojects: There is an elephant in the room! International Journal of Project Management, 35(3), pp.252–268.

Locatelli, G., Invernizzi, D.C. & Brookes, N.J., 2017. Project characteristics and performance in Europe: an empirical analysis for large transport infrastructure projects. Transportation Research Part A: Policy and Practice, 98, pp.108–122.

Longottom, D., 2000. Benchmarking in the UK: an empirical study of practitioners and academics. Benchmarking: An International Journal, 7(2), pp.98–1717.

Merrow, E.W., 2011. Industrial Megaprojects: Concepts, Strategies and Practices for Success 1st ed. John Wiley & sons, ed., Cambridge University Press.

NDA, 2016. Strategy - effective from April 2016, Available at: https://www.gov.uk/government/consultations/nuclear-decommissioning-authority-draft-strategy.

Offshore Energy, 2014. Decommissioning North Sea platforms to cost $66.3 B |. Offshore Energy Today official website. Available at: http://www.offshoreenergytoday.com/decommissioning-north-sea-platforms-to-cost-66-3-b/ [Accessed February 26, 2016].

Oil&Gas UK, 2015. Decommissioning Insight 2015, Available at: http://oilandgasuk.co.uk/wp-content/uploads/2015/11/Decommissioning-Insight-2015-updated.pdf.

PMBOK, 2013. A Guide to the Project Management Body of Knowledge - Fifth Edition,

PremierOil, 2010. Shelley Field Decommissioning Programmes, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/43399/shelley-dp.pdf.

PremierOil, 2015. Shelley Field Decommissioning Programmes - Close Out Report, Available at: https://whitehall-admin.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/43400/5021-shelley-close-out-report.pdf.

Ramirez, R.R., Alarcón, L.F.C. & Knights, P., 2004. Benchmarking System for Evaluating Management Practices in the Construction Industry. Journal of Management in Engineering, 20(3), pp.110–117.

Rossiter, J.R., 2002. The C-OAR-SE procedure for scale development in marketing. International Journal of Research in Marketing, 19(4), p.30.

Schneider, C.Q. & Wagemann, C., 2012. Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis, Cambridge University Press.

Sellafield Ltd, 2016. Sellafield Ltd. Sellafield Ltd official website. Available at: http://www.sellafieldsites.com/ [Accessed October 30, 2016].

Shell E&P UK, 2014. Indefatigable Field Platforms and Pipelines Decommissioning Programme - Close Out Report, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/362848/Inde_Close_Out_Report.pdf.

Shell U.K. Limited, 2007. Indefatigable Field Platforms and Pipelines - Decommissioning Programme. , (May). Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/43411/inde-dp__1_.pdf.

Shlens, J., 2005. A Tutorial on Principal Component Analysis. University of Pennsylvania Law Review, 154(3), p.477. Available at: citeulike-article-id:80546\nhttp://www.snl.salk.edu/~shlens/pub/notes/pca.pdf.

Silverstone Energy Limited, 2010. Tristan NW Field Decommissioning programmes - revision 3,

Stapenhurst, T., 2009. The Benchmarking Book: a how to guide to best practice for managers and practitioner,

Total E&P Norge AS, 2003. Frigg Field Cessation Plan, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/43408/frigg-dp.pdf.

Total E&P Norge AS, 2011. Frigg Field Cessation Plan - Close Out Report, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/43409/4204-frigg-close-out-report.pdf.

Total E&P UK, 2013. Decommissioning, dismanling and disposal of the MCP - 01 installation - Close Out Report, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/204864/MCP-A-RP-0009_-_DECOMISSIONING__DISMANTLING_and_DISPOSAL_of_the_MCP-01_I....pdf.

Total E&P UK, 2007. MCP - 01 - Decommissioning Programme, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/43405/mcp01-dp.pdf.

Venture Production plc, 2009. Kittiwake SAL Export System Decommissioning Programme, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/4

3401/kittiwake-sal-export-dp.pdf.

Warrack, A.A., 1985. Resource Megaproject Analysis and Decision Making, Institute for Research and Public Policy, Western Resource Programme, Victoria, BC.

Van Wee, B., 2007. Large infrastructure projects: a review of the quality of demand forecasts and cost estimations. Environment and Planning B: Planning and Design, 34(4).

Yin, R.K., 2009. Case Study Research: Design and Methods 4th ed.,

Appendix I – Benchmarking in the literature

Reference	Definition of benchmarking	Highlights relevant for the current research	Steps and/or phases of the benchmarking process
"The benchmarking book" (Stapenhurst 2009)	Stapenhurst (2009) provides a list of definitions, finally summarizing benchmarking in the "method of measuring and improving our organizational performance by comparing *ourselves with the best"*	Stapenhurst (2009) provides a comprehensive investigation into the benchmarking analysis, focusing on the benchmarking process and the managerial and organizational aspects of benchmarking.	Stapenhurst (2009) summarizes the benchmarking process into three main phases, i.e.: 1. Planning: develop a project proposal 2. Benchmarking performance: recruit and work with participants, collect and compare data 3. Improvement: improve the organization
"Benchmarking the benchmarking models" (Anand & Kodali 2008)	According to (Anand & Kodali 2008), the definition of benchmarking vary. Key themes include measurement, comparison, identification of best practices, implementation and improvement.	Anand & Kodali (2008) highlight that there are many kinds of classification schemes for benchmarking and provides an overview of different classification schemes and types of benchmarking. In total, Anand & Kodali (2008) have investigated 35 published models. The author provides a **taxonomy for benchmarking models**, dividing them into consultant/expert-based models, academic/research-based models and organization-based models, and finally states that benchmarking should at least be classified into **internal** and **external benchmarking.** Anand & Kodali (2008) reveal the presence of **71 benchmarking steps** in the 35 different publications analysed. In those, around 13 steps have been addressed by many researchers (>40% of them) and are therefore called "common steps" of benchmarking. Excluding the "common steps", Anand & Kodali (2008) additionally lists 18 practices, that had an occurrence between 14% and 45%. Lastly, Anand & Kodali (2008) propose a 12-phases that include 54-detailed-step benchmarking model.	Anand & Kodali (2008) highlight **the 13 common steps** of the benchmarking process: 1. Identify the benchmarking subject 2. Identify benchmarking partners 3. Perform benchmarking study 4. Determine current competitive gap 5. Establish functional goals 6. Develop action plans 7. Implement of action plans to bridge the gap 8. Recalibrate the benchmark 9. Understand the current situation by collecting and analysing the existing information on the subject to be benchmarked 10. Monitor results of the implemented actions 11. Identify the critical success factors or indicators of the subject to be benchmarked 12. Measure the existing state of the subject to be benchmarked with respect to the critical success factors/indicator 13. Form a benchmarking team and identify a leader of the team to carry benchmarking study
"Benchmarking as an action research process" (Kyro 2004)	"Even though definitions vary between scholars, the aspects of evaluation and improvement by learning from others are embedded in the definitions regardless of the definer" (Kyro 2004)	Adopting an **interpretative, comparative concept analysis**, Kyro (2004) suggests that benchmarking can be regarded as a special kind of action research. If so, the author affirms that, more attention should be put on "preliminary planning, observation, reflection and the use of theoretical frames". Kyro (2004) also provides a detailed comparison of benchmarking and action research, focusing on: (1) the similarity of their purpose, i.e. to improve practices, (2) the researcher's role, (3) how the processes take place and how they can be either adapted or created, (4) the phase of action research vs the phases of benchmarking.	Kyro (2004) provides a review of benchmarking models and a classifications of benchmarking phases compared to action research. Kyro (2004) approaches the benchmarking process as a "two-cycle spiral, where the actual data collection phase is regarded as an *action of the first cycle [...] and at the same time is generic enough to* be adopted for different forms of benchmarking"
"An evolutionary approach to Benchmarking"	"Benchmarking is the process that facilitates learning and understanding of the organization and it processes. It	Fernandez et al. (2001) propose an evolutionary classification method called **cladistics,** that distinguishes between different organisational types according to how they evolve and develop new ways of working. The authors firstly list the benchmarking styles (classified into internal, competitive, functional and generic	Fernandez et al. (2001) provide a structured literature review of 6 earlier publications on benchmarking models, highlighting the five generic steps of benchmarking models: 1. Planning

(Fernandez et al. 2001)	enables organizations to identify the key processes that need improvement, and to search for applicable solutions for the best in class" (Fernandez et al. 2001)	and strategic), highlighting their advantages and disadvantages. Fernandez et al. (2001) also affirm that the benchmarking process is not "a universal yardstick, as it is impossible to establish an absolute measurement in the benchmarking process". The author also state that benchmarking provides a "situational *analysis"*, but not necessarily a "strategic roadmap". Of interest is the consideration that benchmarking models can be used to benchmark "both single functions and an entire organization" and that the **reductionist approach** (in opposition to the systemic approach) seeks to understand systems by reducing them into manageable individual parts	2. Analysis and data collection 3. Comparison and results 4. Change 5. Verification and maturity
"Benchmarking in the UK: an empirical study of practitioners and academics" (Longottom 2000)	Longbottom (2000) highlights respectively Camp's definition of benchmarking, as the "search for industry best practices that will lead to superior performance" (Camp 1989)	Longbottom (2000) investigates the status of benchmarking **within the UK** through the analysis of answers to questionnaires (1,020 questionnaires were issued over a period of nine months, achieving a total response of 560). This study revealed that benchmarking is not so well-established as suggested by the literature, highlighting four major areas that require further discussion, i.e.: (1) the link between benchmarking and the strategic planning process; (2) the development of customer benchmarking methods; (3) the critical factors for transferring best practices across organizations, (4) the adaptation to post-modern attitudes to benchmarking	Longbottom (2000) identifies four major stages of the benchmarking process: 1. Planning 2. Analysis 3. Implementation 4. Review
"Benchmarking process formalization and a case study" (Büyüközkan & Maire 1998)	"Benchmarking is the continuous process of evaluation of products, services and practices, with respect to those of the strongest competitors or of the enterprises recognized as *leaders"* (Büyüközkan & Maire 1998, quoting Camp (1995)) "In a direct way, benchmarking is a process of evaluation and improvement of performance" (Büyüközkan & Maire 1998)	Büyüközkan & Maire (1998) state that benchmarking is one of the most efficient and effective management tools to help an enterprise to improve its performance. The author also points out some of the obstruction against the benchmarking approach, e.g. industrialists that think that their business processes are **very company specific** and that it is **not ethical** to look at other companies' technology and manufacturing methodology, and the **lack of formal benchmarking methodology.** Büyüközkan & Maire (1998) define a general benchmarking process to cover the different types of benchmarking (i.e. internal, external, industry, competitive, and generic benchmarking). This process is divided into the following 5 phase, divided into 15 steps and is a **cyclical, "never-ending and learning" process**. Büyüközkan & Maire (1998) also state the serious difficulties of implementing a continuous improvement activity is that "there are no standard performance metrics to be utilized in such studies". The author then illustrates the methods and tools for the first 5 steps of the benchmarking process through a case study.	The 5-phase, 15-steps benchmarking process: 1. self-analysis: a. Define the activities and customers of enterprise b. Determine performance criteria and performance measures c. Revise and improve current performance 2. pre-benchmarking a. Establish priorities and choose benchmarking subject b. Choose benchmarking partners c. Determine methods and tools of data, information and knowledge collection 3. benchmarking a. Collect and organize information b. Determine and analyse the performance gap c. Evaluate future enterprise realization 4. post-benchmarking a. Communicate benchmark findings and establish functional goals b. Set objectives and develop action plans c. Implement specific actions 5. observation and adjustment a. Review the benchmarking integration and learn the results b. Estimate success of the project and recalibrate benchmarks c. Adjust the objectives and return to step 1

Table 5. "Benchmarking" in the literature

Appendix II– Benchmarking applied to the construction industry

Ref	(El-Mashaleh et al. 2007)	(Costa et al. 2006)	(Ramirez et al. 2004)	(Garnett & Pickrell 2000)
Aim of the research	The aim of this research is to present a comprehensive benchmarking model that uses input metrics (i.e. (1) safety expenses and (2) management expenses) and output metrics (i.e. (1) schedule adherence, (2) cost performance, (3) customer satisfaction, (4) safety performance, and (5) profit) to determine the company performance.	The final aim of this paper is the comparison of four benchmarking approaches is to use the lessons learned and upgrade the existing benchmarking initiatives to devise new ones. All four initiatives use an interactive online tool for the collection and evaluation of the benchmarking measures.	This research presents the results from the application of different benchmarking system through different methods: (1)the qualitative benchmarking with the class median, (2)the correlation analysis, (3)the factor analysis, (4)the multivariate linear regression, and (5)sector trends.	This paper discusses the development and testing of a benchmarking model to be implemented in the construction industry. Benchmarking models which ranged from 5 step to 11 step processes are reviewed and the key features of each stage are highlighted.
Method or model described or adopted	Firstly, this paper lists and criticizes three models that provide insight into overall firm performance and support trade-off analyses among multiple key performance metrics. These three models are criticized for (1) being project specific; (2) not supporting the understanding of the trade-offs among the different variables that affect the performance; (3) providing no insight into the relationship between how resources are expended and the relative success of out- comes; (4) not allowing the measurement of the impact of certain technological and managerial factors on overall firm performance. Secondly, this paper proposes and describes a comprehensive model that uses (1) schedule adherence, (2) cost performance, (3) customer satisfaction, (4) safety performance, and (5) profit to assess the performance in the construction industry, since these appear to be the most critical to overall success. This comprehensive model also incorporate Data Envelopment Analysis (DEA) where the organization under study is called Decision Making Unit (DMU). DEA uses mathematical linear programming to determine which of the DMU under study form an envelopment surface, i.e. an efficient frontier.	Firstly, this research revises the role of performance measurement in the benchmarking process and highlights the importance of abstraction of best practices before they are transmitted and applied in the company. Secondly, four benchmarking initiatives were analysed: (1) KPI from the UK; (2) the National Benchmarking System for the Chilean Construction Industry (NBS-Chile); (3) the Construction Industry Institute Benchmarking and Metrics (CII BM&M) from the United States; and (4) the Performance Measurement for Benchmarking in the Brazilian Construction Industry (SISIND-NET Project). This paper highlights that the most common performance measures involved are cost & cost predictability, time & time predictability and safety.	(1)**Qualitative benchmarking** with the class median is used to enable each company to evaluate its position compared to the worse & best case scenario and the median. This comparison is highlighted using the Radar graph. The main steps of this analysis are: send questionnaire, collect data, calculate performance indicator per each company, compare with other companies performance. (2)**Correlation analysis** is used to investigate the intensity with which two variables, Xi and Xj, are linearly related. To measure the intensity of the correlation, the Pearson's coefficient is used. The correlations investigated are (2.a)between indicators and management dimensions, (2.b)for the central office, and (2.c)for the construction site. The correlation analysis found that safety performance is strongly related to companies having superior planning and control, quality management, cost control, and subcontractor management policies. (3)**factor analysis** is used (3.a)for central office, and (3.b)for construction site. The factor analysis uses the method of principal components to determine the underlying structure among the different management dimensions and identify relationships not previously established. The commonality table is	This paper focuses on three main questions: (1)what is benchmarking? (2)how can it be used?(3)when can it be used? The authors also highlight the p**roblems in benchmarking**, i.e. (1) insufficient client resources, time, money, staff, etc.; (b) internal resistance; (c) previous bad experiences; (d) difficulty in identifying and obtaining partners; (e) difficulty in obtaining data. Also the uniqueness of projects, their various location, the inability of identifying best practices, and the low number of good benchmarks hinders the benchmarking analysis of the construction industry. This paper compares different benchmarking models to subsequently define a 7–step model and case study analysis through action research. The authors highlight that, to be successful, the benchmarking process is as important as the benchmarks themselves. Another key strength of the methodology is that the theoretical basis is aligned with that of benchmarking, i.e. social constructivism instead of positivism. The difference in approach between benchmarks and benchmarking reflects the theoretical schools of positivism and social constructivism; the former being based on fact finding, the latter an interactive activity whose benefits are as much in the social interaction as the measurement process.

Steps of the analysis and highlights	The three benchmarking models are: Fisher et al. (1995) that collected data from 17 companies on 567 projects; Hudson and the Construction Industry Institute (1997-2000) that collected data of 901 projects from 37 owner and 30 contractors; and the Construction Benchmarking Programme or key performance indicator (KPI) model (1998).	The four methods are described. The barriers of the implementation of performance-measured systems in the construction industry arise because (1) the construction industry is a project-oriented industry with “unique” projects; (2) establishing a project performance measurement system and incorporating it into the company routine requires an intense effort; (3) the responsibilities for data collection, processing, and analysis, in general, are not well defined at the beginning of the project; (4) each project usually has a different managerial team and the use of measures will depend on the capabilities and motivation of each manager.	one of the outputs of this analysis, and it represents the proportion of the variance explained by the component or factor. The factor analysis found that central office priorities focus on strategic management policies having longer-term competitive impact, while site management emphasizes tactical management dimensions having short-term impact. (4)**multivariate linear regression** was used but discarded due to the weak correlation coefficient caused by the reduced quantity of data. (5)**sector trends** are used to categorize and analyse survey results according to construction industry subsectors. The main steps for this analysis are: prepare histograms of the responses to each questions, analyse responses firstly according to job category, secondly according to construction industry subsectors, draw conclusions. In the analysis of industry trends, construction companies working in the civil works and low-rise housing subsectors were generally found to lag management performance levels registered for the high-rise and heavy construction subsectors.	The steps presented in this paper are: definition of the benchmarking model, collection of data for the case study analysis, description of the implementation of the selected case studies, presentation of the results obtained, that suggest that benchmarking could be a powerful tool in investigating and managing change on construction projects.
Data collection and number of cases	The data were collected through a survey questionnaire that was divided into (1) collection of general information regarding the interviewee, (2) firm general information, (3) firm overall performance. Data were collected from 74 construction firms, including general contractors, construction management companies, design/build firms, and subcontractors, involved in residential, commercial, industrial, and heavy/ highway construction. The minimum number of DMU in any model should be three times the number of variables, and the model’s discriminatory power increases with more DMU and fewer variables.	The information about the KPI was obtained from its website, published papers, and reports. For the CII BM&M, the information was obtained from its website, reports, and interviews with a project researcher and the associate director. The information about the NBS-Chile was obtained from its websites, from published papers and reports, and from its existing database. Also, semi-structured interviews were conducted with managers of seven construction companies involved in the program and the current coordinator of the initiative. Finally, two of the authors of the SISIND-NET Project have been directly involved in this project. Different numbers of data per method are analysed.	This paper advances the use of a structured questionnaire to evaluate aspects related to the organizational culture and management of construction companies. The results of the questionnaire are then correlated against the quantitative performance indices obtained from the CDT’s national benchmarking study to establish causal relationships. 13 companies initially participated to the assessment, 42 questionnaires were completed by central office personnel and 87 by construction site representatives.	2 case studies are analysed: construction retail client and research organization. The authors stressed the fact that, “often data, loosely termed benchmarks, are determined by comparative analysis, experience and gut feeling rather than through focused analysis”.

Table 6. “Benchmarking” in the construction industry