Building Tomorrow's Society
Bâtir la Société de Demain

Fredericton, Canada
June 13 – June 16, 2018/ *Juin 13 – Juin 16, 2018*

DATA ENVELOPMENT ANALYSIS (DEA) FOR SAFETY-BASED EFFICIENCY EVALUATION OF CONSTRUCTION SITES

Nahangi, Mohammad[1,3], Chen, Yuting[1] and McCabe, Brenda[2]
[1] Post-Doctoral Fellow, Department of Civil and Environmental Engineering, University of Toronto, Canada
[2] Professor, Department of Civil and Environmental Engineering, University of Toronto, Canada
[3] m.nahangi@utoronto.ca

Abstract: Identifying efficient construction sites in terms of safety is a challenge that must be addressed while assessing and evaluating performance criteria. This paper presents a method for comparative and relative analysis of construction sites in terms of their safety performance. The proposed method employs Data Envelopment Analysis (DEA) for identifying efficiency of construction sites, also known as decision making units (DMU's). Safety climate factor and number of incidents are respectively used as inputs and output of the DEA system. Four various scenarios are performed and analyzed, and results are compared to evaluate efficient construction sites. Results show that the number of incidents incurred at construction sites, which is the output of the DEA framework, is the dominant factor correlating with the efficiency of construction sites. Moreover, specific safety climate factors are investigated to examine their impact on the efficiency calculated. Some other important observations that all comply with practical expectations are also reported.

Keywords: construction safety, safety climate, incidents, data envelopment analysis (DEA), safety performance efficiency.

1. Introduction and background

Safety is critical for the construction industry. As an important factor affecting safety performance, many safety climate research studies have been conducted. Safety climate is people's perceptions toward safety and their management's commitment to safety (Zohar 1980). Most of the studies on construction safety climate were based on more than one site or operational unit and the results were investigated as aggregated results of multi-sites. For example, the work by (Chen et al. 2017b) were based on 837 construction safety surveys from 112 sites and the presented results were the aggregated results of the 112 sites. However, measuring the performance of individual sites, in particular, identifying weaknesses and strengths of individual sites regarding their safety climate and safety performance is significant to continuously improve safety. To fill in this gap, this paper applied Data Envelope Analysis (DEA) technique to assess individual performance of 112 surveyed Ontario construction sites in terms of their safety climate and incident occurrences.

DEA was first introduced as a tool to measure the productivity and efficiency performance of decision making units (DMUs) (Charnes et al. 1978). It has been widely applied to address various decision analysis in different sectors, e.g. bank performance (Yeh 1996), university performance (Avkiran 2001), and supplier election in manufacturing (Liu et al. 2000), etc. For safety research, the focus has been on road safety, e.g. (Hermans et al. 2009) applied DEA to assess road safety performance of 21 countries. However, for construction safety, very few research has been conducted, one example of which is that (El-Mashaleh et al. 2010) applied DEA to benchmark safety performance of 45 construction contractors. Thus, the application of DEA to construction sites in this paper expands the use of DEA.

DEA measures the efficiency of DMUs based on the input to the DMU which is considered as resources used, as well as the output of the DMU which is considered as the product of the system. In this paper, input is safety climate measured by six factors and output is safety performance measured by physical injuries, job stress, and unsafe events (Chen et al. 2017b). A DEA-based efficiency analysis framework based on the safety performance was built, and the factors that have the greatest impact on safety performance were identified.

2. Methodology

An overview of the proposed framework for safety-based efficiency analysis is illustrated in Figure 1. The input to the proposed system is the result of a safety-related survey performed on various construction sites. More specific information about the conducted survey used in this paper is provided in the next section. As seen in Figure 1, the proposed framework has three major modules: (1) Preprocessing: that is performed for applying the filtering rules to refine the input data (Section 3.1); (2) Efficiency analysis: that employs data envelopment analysis (DEA) to quantify efficiencies of the constructions sites relatively (Section 3.2); and (3) Identify influential factors: calculating efficiencies and identifying efficient and inefficient sites will lead to the investigation of the influential factors. The safety climate factors are then ranked and the final results are used to prioritize the safety climate improvement (Section 3.3). Detailed explanations about each module and the required metrics and parameters are provided in the following sections.

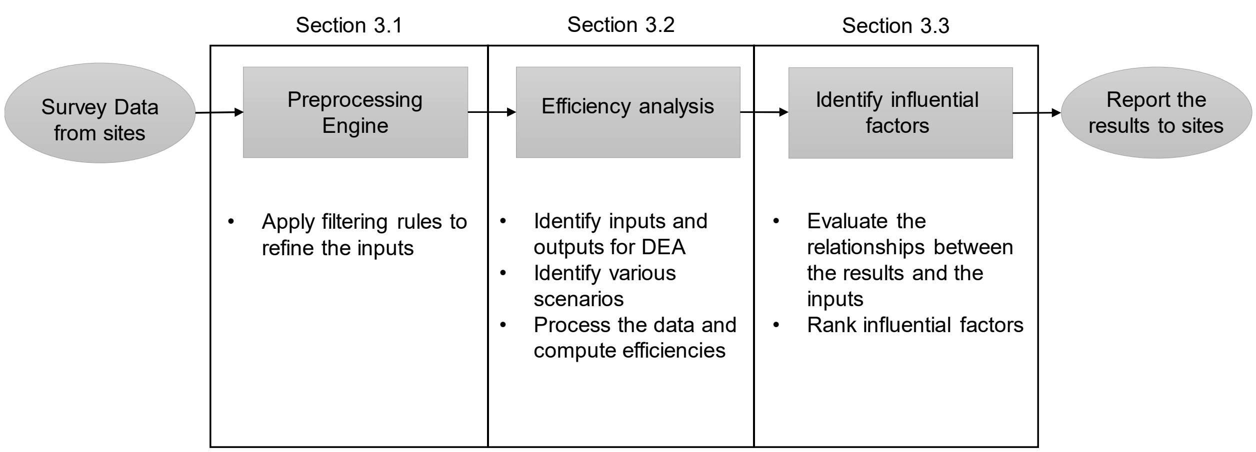

Figure 1: Overview of the proposed methodology and the flow of information between different components

2.1. Preprocessing

Collected data form a safety-related survey is the input to the proposed systems. Some filtering rules are performed for preprocessing and refining the input to the efficiency analysis module. Such rules and filters include:

- Grouping the collected data from the same construction sites: surveys are sent out to various construction firms who have participated. Collected data are sent back to the server at different time frames. Therefore, the first step to identify the efficiency at the site-level is to group the individuals who are from the same construction firm. This has been done by labelling and filtering the participants by their employers.

- Filtering the insufficient number of participants: in order to assure that the collected data are reliable for further analyses, a threshold value for number of participants have been considered. Construction firms not satisfying this rule are filtered out from the analyses.
- Filtering out the "Not Available" entries from the spreadsheets: non-entries and not-acceptable entries are removed from the calculations by applying the filtering rule in the programming engine.

2.2. Efficiency Analysis: Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) is employed to identify the relative efficiency of the construction firms participated in the safety survey. As mentioned earlier, DEA was first introduced as a tool to measure the productivity and efficiency performance of decision making units (DMU's) (Charnes et al. 1978). DEA measures the efficiency of DMUs based on the input(s) to the DMU which is considered as the resources used, as well as the output(s) of the DMU which is considered as the quantified product or service of the system. A brief introduction to the mathematical notation and formulation of the DEA method is provided in the following section.

2.2.1. Mathematical Formulation of DEA

Let X_i be the vector of inputs to the DMU i, and Y_i be the corresponding vector of the outputs. Consequently, assume X_0 and Y_0 are the input and output vectors of the DMU_0 for which the efficiency is being investigated. The efficiency of the DMU_0 can then be identified by solving the following linear program:

$$\begin{aligned} \text{Minimize: } & \theta, \text{ s.t.} \\ & \sum \lambda_i X_i \leq \theta X_0 \\ & \sum \lambda_i Y_i \leq Y_0 \\ & \lambda \geq 0 \end{aligned} \qquad \text{Eq.1.}$$

where λ_i is the weight given to DMU i to dominate DMU_0 and θ is the efficiency of the DMU being investigated DMU_0. Solving the linear program formulated above (Eq. 1) will lead to the identification of the efficiencies of the DMUs. It should be noted that the maximum value of the efficiency is $\theta = 1$. The DMU with $\theta = 1$ is therefore the most efficient unit. There might be multiple efficient units according to the inputs and outputs to the DEA linear problem.

In the case of safety environment evaluation of different construction sites and their relative efficiency, safety climate factors (6 factors) and number of incidents (3 types of incidents) are considered as the inputs and outputs of the system, respectively. Some manipulations are applied on the originally collected data in order to comply with the theoretical requirements of the DEA. Extensive description about such manipulations and modifications to the model for this study is provided in Section 4.

2.2.2. Influential Factors Identification

To identify and rank the most influential inputs (safety climate factors) on the efficiency of the sites being investigated in this paper, a sensitivity analysis is performed. The key reason for performing the sensitivity analysis is to investigate the impact of the inputs on the efficiency results. Because the variability of the input values is uncontrolled, the real impact of the safety climate factors on efficiency results cannot be measured directly. Therefore, each safety climate factor is artificially changed, while other factors are kept unchanged to reliably quantify its impact on the efficiency values. The same procedure is performed on all of the safety climate factors, and results are used to identify the most and the least influential factors on the efficiencies calculated. Results will help the decision makers in case an improvement in safety performance is desired in a unit. Most influential factors are the factors that will have the most impact on the safety performance, if they are improved. Figure 2 illustrated the procedure for the sensitivity analysis algorithm performed for ranking and identifying the influential factors.

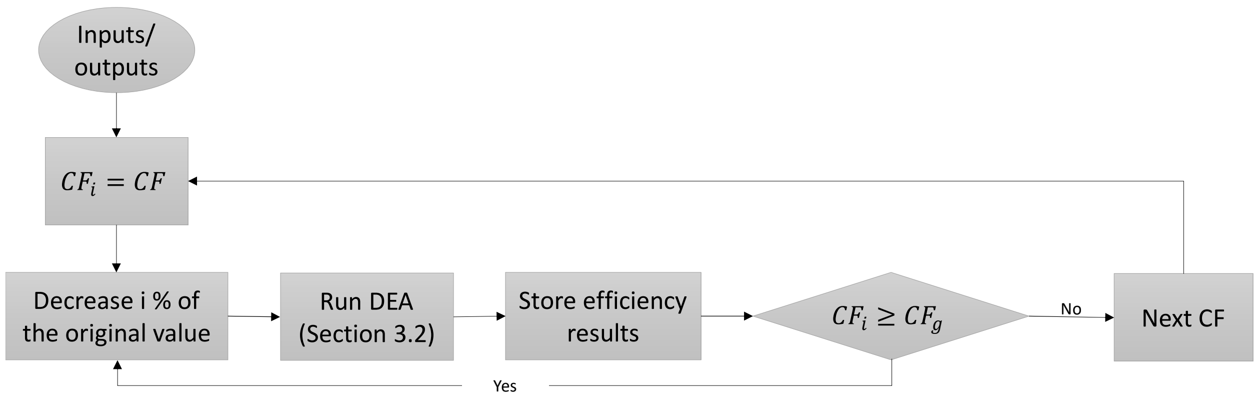

Figure 2: Flowchart for performing sensitivity analysis to rank and identify influential safety climate factors (CF). $CF_g = 0.5\ CF_i$ is the target climate factor for sensitivity analysis.

3. Case study and results

A case study in the form of a survey conducted at various construction firms and sites is used to verify and validate the proposed framework for safety-based efficiency identification using DEA. Information about the survey instrument and the data collection method is provided in the following sections.

3.1. Survey instrument and data collection

The survey used the self-administered questionnaire adapted from previous research (Chen et al. 2017a; Chen et al. 2017b; McCabe et al. 2008). The survey is designed in three key sections as follows:

- Demographic information that includes questions such as age, years with the company, and years in the construction industry, etc.
- Attitude statements where the participants and respondents identify the degree to which they agree with the statement. The answers are scaled between 1 for “strongly disagree” and 5 for “strongly agree”. The attitude statements are categorized as six different safety climate factors. This is going to be extensively discussed in the following sections. As discussed earlier, the attitude statements are imported to the DEA framework as the inputs to the system with a slight modification. For the purpose of consistency and compliancy with the theoretical formulation of the DEA, attitude statements are reversed in order to consider the positivity of higher values. In other words, higher numbers representing the attitude statements imply that the work environment has a better climate with potentially higher costs or resources associated with that number. Reversing the attitude numbers make the inputs consistent with the assumption made to formulate the DEA.
- Incident reports in which the respondents are asked to report their incidents in the past 3 months in three categories: (1) physical injuries, (2): unsafe events, and (3) psychological stress symptoms. As mentioned earlier, the incidents are considered as a metric to measure the safety of the work environment and are imported to the DEA framework as the outputs of the systems. Because the higher number of incidents imply a more unsafe work environment, incident numbers are also reversed, for the compliancy with the DEA assumptions. Safety metrics are generated by subtracting the incidents from the maximum number of incidents reported for that category. Therefore, “zeros” will represent the most unsafe situations.

3.2. Results

After calculating the efficiency values using the assumptions explained previously, the correlation between the efficiency of the construction firms and the DEA parameters is evaluated. First, the correlation between the inputs (safety climate factors) and the efficiency values is investigated (Section 4.2.1). To measure the influence of each safety climate factor a sensitivity analysis with controlled variability of the changing parameters is conducted (Section 4.2.2). Then, the correlation between the outputs of the DEA framework (incidents reported) is investigated to identify whether or not there is a linear relationship between them (Section 4.2.2). Finally, the correlation between the efficiency values and the demographic information (i.e. age, years of experience in construction, and years with the company) is investigated (4.2.4). Correlation coefficient and sensitivity analysis results are reported in the following sections.

3.2.1. Inputs and efficiency values

Efficiency values are calculated for various construction sites investigated in this study. The efficiency values versus each of the six input parameters (safety climate factors) is plotted to identify whether there is a significant linear relationship between them. As seen in Figure 3, there is not a significant linear relationship between any of the inputs and the efficiency values. The coefficients of correlation are also reported in Table 1, that signify the correlation is not strong between the investigated parameters and the efficiency values.

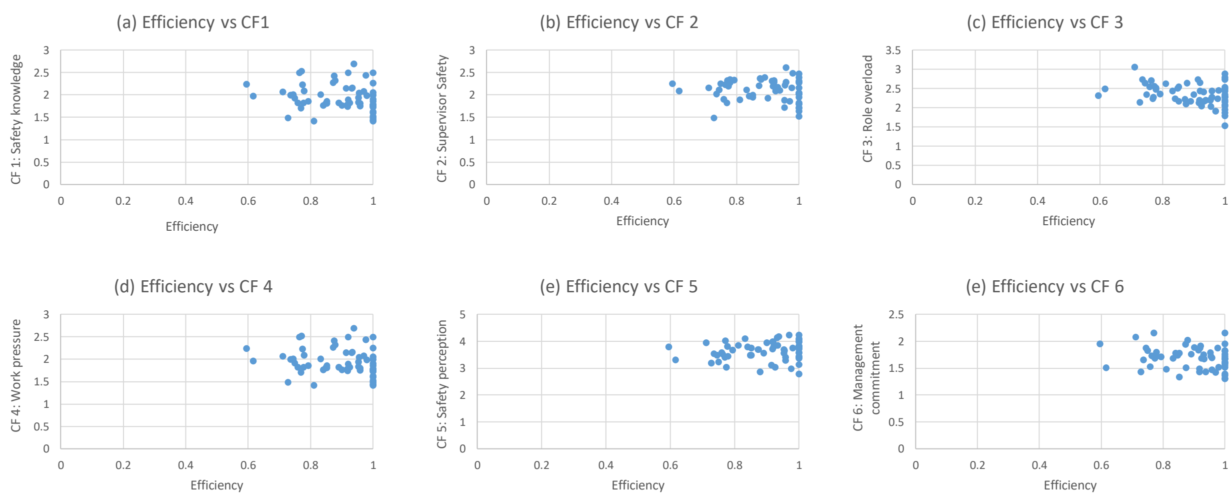

Figure 3: Relationship between the six safety climate factors (CFs) as inputs and the efficiency results. CF 1: Safety knowledge, CF 2: Supervisor safety, CF 3: Role overload, CF 4: Work pressure, CF 5: Coworker safety perception, CF 6: Management commitment.

Table 1: Coefficient of correlation between the safety climate factors (inputs) and the efficiency results

Safety Climate Factor	Correlation coefficient with Efficiency (ρ)
Safety knowledge	-0.185
Supervisor safety	-0.089
Role overload	-0.349
Work pressure	0.108
Co-worker safety perception	0.121
Management commitment	-0.210

Coefficients of correlation, reported in Table 1, signify that the correlation is not very strong ($\rho < 0.4$). Moreover, there is not even a direct or indirect significant relationship resulting from the fact that the calculated coefficients are not consistently positive or negative for all of the parameters.

3.2.2. Sensitivity analysis for identifying and ranking the most influential factors

The reason that the correlation between the inputs and the efficiency values is not very strong might be due the uncontrolled variability of the inputs for various construction firms. In other words, the influence of one climate factor on the efficiency values might be either exacerbated or attenuated by other climate factors significantly varying. This led to conduct a sensitivity analysis where the studied climate factor is changing, while other factors are kept unchanged. Figure 2 illustrates more detail on the method for the conducted sensitivity analysis to rank and identify the influential factors on the efficiency of the construction firms. Sensitivity analysis results using the procedure (explained in Section 3.2) is shown in Figure 4. The relative efficiency on the vertical axis is calculated as the summation of the efficiency values of all construction sites divided by the total number of construction sites remained after preprocessing (61 sites). In other words, relative efficiency is a metric to measure how efficient the construction sites are performing in terms of safety and with respect to the ideal situation, that the efficiency is 1 for all of the firms.

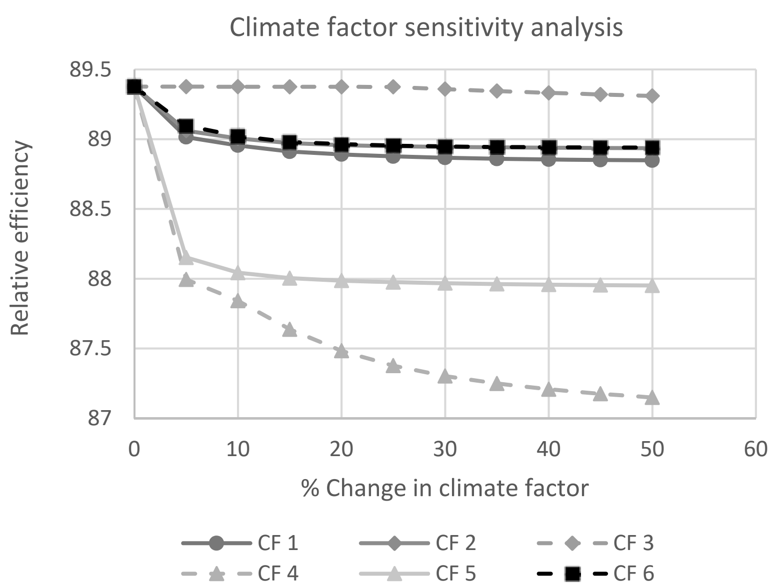

Figure 4: Sensitivity analysis of the inputs (i.e. safety climate factors) on the total relative efficiency. CF 1: Safety knowledge, CF 2: Supervisor safety, CF 3: Role overload, CF 4: Work pressure, CF 5: Coworker safety perception, CF 6: Management commitment.

The safety climate factors are then ranked based on their impact (i.e. the change) on the relative efficiency. As seen in Figure 4, the safety climate factors are ranked from the most to the least influential factors as the following:

- Rank 1: Work pressure (CF4),
- Rank 2: Supervisor safety (CF2),
- Rank 3: Safety knowledge, coworker safety perception, and management commitment (CF1, CF5, CF6),

- Rank 4: Role overload (CF3).

This ranking will be the key prioritizing the safety climate factors for safety-based efficiency improvement of the construction sites. In other words, if safety-based performance is to be improved in a construction firm, the highest priority is given to the “work pressure” as a safety climate factor. Similarly, role overload will have the lowest priority and will be the last safety climate factor to be improved considering the same cost for all of the safety climate factors.

3.2.3. Outputs and efficiency values

The correlation between the efficiency and the outputs of the DEA system is then evaluated to identify whether or not there is a linear relationship between the data. The coefficients of correlation are also calculated and reported in Table 2. Results show that there is a strong and direct correlation between the efficiency of the sites and the safety metric measured by reversing the number of incidents incurred, explained previously ($\rho > 0.6$).

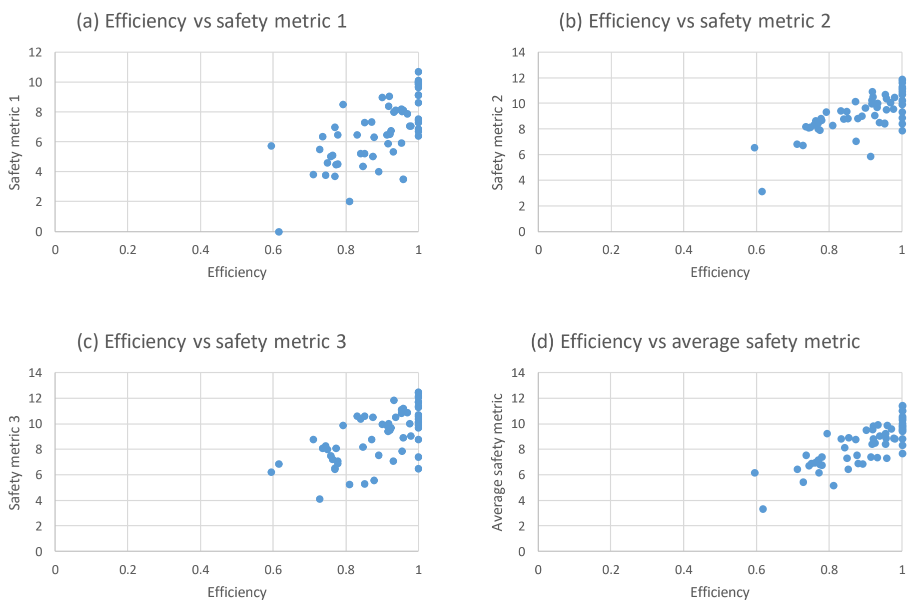

Figure 5: Relationship between the three safety metrics (a, b, and c) as well as the average safety metric (d) and the efficiency results. Safety metric 1: injuries, safety metric 2: unsafe events, safety metric 3: stress.

Table 2: Coefficient of correlation between the types of incidents (outputs) and the efficiency results

Type of incident	Correlation coefficient with Efficiency (ρ)
Injuries	0.727
Unsafe events	0.634
Stress	0.800

The strong correlation between the efficiency of the sites and the safety metric imply that reducing the number of incidents will directly improve the safety-based efficiency on construction sites, as expected.

This was also expected for the safety climate factors; however, because of the uncontrolled variability of the safety climate factors for various sites, the correlation was weak and inconsistent. For the outputs however, the impact of variability is dominated by their weight in the mathematical model for the DEA.

3.2.4. Demographics and efficiency values

The correlation between the demographics of the efficiency of the sites is also investigated, in order to identify whether or not there is a linear relationship between them. As shown in Figure 6 and Table 3, there is not a strong and consistently direct or indirect correlation between the data. The correlation test also implies that the efficiency results of construction sites do not significantly depend on the demographics parameters such as the average age of the respondents, average experience in the related industry or in the same company.

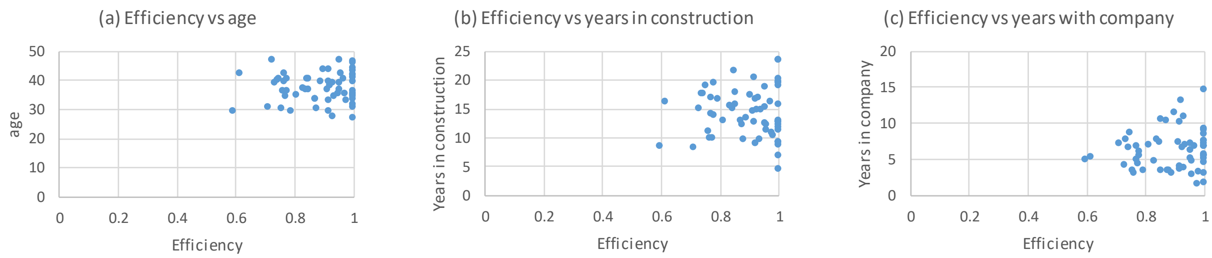

Figure 6: Analysis of the efficiency results using DEA vs age (a), years of experience in construction (b), and years with company (c).

Table 3: Coefficient of correlation between demographic factors and the efficiency results

Parameter	Correlation coefficient with Efficiency
Age	0.050
Years of experience in construction	-0.016
Years with company	0.107

4. Conclusions

A DEA-based framework was developed to evaluate the safety-based efficiency of construction sites. The research was conducted using a survey conducted at different construction sites in Ontario, Canada. A sensitivity analysis was also performed to rank and identify the most influential safety climate factors on the efficiency results. The key findings of this research are listed as follows:

- There was no direct and consistent correlation between the original safety climate factors and the efficiency of the construction sites. This might be due to the uncontrolled variability of the input data that dominates the weights calculated in the DEA framework.

- The sensitivity analysis showed that the “work pressure” and “role overload” are the most and the least influential safety climate factors on the efficiency results, respectively. In other words, to increase the efficiency of construction sites the “work pressure” parameter is the key factor to improve and must be given the highest priority; while the “role overload” improvement will have the least impact on the efficiency values.

- There was a direct and strong correlation between the safety metric measured by reversing the number of incidents and the efficiency results. Such strong and direct correlation implies that the efficiency values increase, if the safety metric on a construction site increases (i.e. or the number of incidents on the site decreases).

- The correlation between the demographics of the construction sites and the efficiency values was not strong and consistent either. This implies that the demographics parameters such as the average years of experience in the related industry or the average age of the respondents in site do not have a significant impact on the efficiency values calculated.

One limitation of this research is that it only considered safety climate factors as the inputs to the DEA system. Considering other scenarios for the DEA system such as the average of the input versus the average of the output will be a potential avenue for further exploration. Currently, the DEA framework and the efficiency values are implemented and measured relatively to the input dataset. Benchmarking the parameters of the DEA framework and measuring their impact on the relative efficiency will lead to a global metric for safety performance of construction sites. This is currently being investigated by the same group of authors.

References

Avkiran, N. K. (2001). "Investigating Technical and Scale Efficiencies of Australian Universities through Data Envelopment Analysis." *Socioecon. Plann. Sci.,* 35(1), 57-80.

Charnes, A., Cooper, W. W., Rhodes, E. (1978). "Measuring the Efficiency of Decision Making Units." *European Journal of Operational Research,* 2(6), 429-444.

Chen, Y., McCabe, B., Hyatt, D. (2017a). "Relationship between Individual Resilience, Interpersonal Conflicts at Work, and Safety Outcomes of Construction Workers." *J. Constr. Eng. Manage.,* 143(8), 04017042.

Chen, Y., McCabe, B., Hyatt, D. (2017b). "Impact of Individual Resilience and Safety Climate on Safety Performance and Psychological Stress of Construction Workers: A Case Study of the Ontario Construction Industry." *Journal of Safety Research,* 61, 167-176.

El-Mashaleh, M. S., Rababeh, S. M., Hyari, K. H. (2010). "Utilizing Data Envelopment Analysis to Benchmark Safety Performance of Construction Contractors." *Int. J. Project Manage.,* 28(1), 61-67.

Hermans, E., Brijs, T., Wets, G., Vanhoof, K. (2009). "Benchmarking Road Safety: Lessons to Learn from a Data Envelopment Analysis." *Accident Analysis & Prevention,* 41(1), 174-182.

Liu, J., Ding, F., Lall, V. (2000). "Using Data Envelopment Analysis to Compare Suppliers for Supplier Selection and Performance Improvement." *Supply Chain Management: An International Journal,* 5(3), 143-150.

McCabe, B., Loughlin, C., Munteanu, R., Tucker, S., Lam, A. (2008). "Individual Safety and Health Outcomes in the Construction Industry." *Canadian Journal of Civil Engineering,* 35(12), 1455-1467.

Yeh, Q. (1996). "The Application of Data Envelopment Analysis in Conjunction with Financial Ratios for Bank Performance Evaluation." *J. Oper. Res. Soc.,* 47(8), 980-988.

Zohar, D. (1980). "Safety Climate in Industrial Organizations: Theoretical and Applied Implications." *J. Appl. Psychol.,* 65(1), 96.