# Development of Safety Level Assessment Models for Building Projects using Multiple Linear Regression and Bayesian Network

**Gholamreza Heravi [1*], Amirbahador Ketabi [2]**

[1] Associate Professor, School of Civil Engineering, College of Engineering, University of Tehran, Tehran, Iran.

[2] Former Post-graduate Student, School of Civil Engineering, College of Engineering, University of Tehran, Tehran, Iran.

## ABSTRACT

The construction industry accounts for a large share of the industry in various countries. The rate of accidents and deaths in this industry is much higher than other industries. Safety is a major concern in the construction industry. One of the most dangerous parts of the industry is the construction of high-rise buildings. In this research, an attempt has been made to evaluate the level of factors affecting the safety of the project, to be able to predict the level of safety of future projects based on the records and information of previous projects. Because with a correct understanding of the impact of factors on safety, it is possible to take the necessary policies to improve safety and reduce damage to this part of the industry and prevent the waste of resources. To achieve this purpose, first, the factors affecting safety have been identified by reviewing the literature and reviewing the unofficial statistics of work-related accidents, and then the relationship between these factors and the level of project safety for 95 projects in Tehran and Kish Island is developed. It was conducted with the help of a questionnaire, with the help of a Bayesian network and multiple linear regression. In the proposed model, the effect of each of the groups affecting the safety of the project is determined. “Monitoring and supervising the safety" and "safe process of doing work" are among the most influential factors on the safety level of high-rise building projects; regression coefficients were 0.338 and 0.264, respectively.

## KEYWORDS

**Construction industry, Building projects, Safety hazards, Safety risks, Project safety level.**

[*] Corresponding Author: Email: heravi@ut.ac.ir

## 1-**Introduction**

The construction industry is recognized as one of the most unsafe sectors among all industries [1]. Workers perform construction activities through various procedures using different types of machinery, tools, and materials. Potential hazards for construction workers often cause serious safety risks. The rates of accidents, injuries, and deaths reported in the construction industry are significantly higher compared to other industries [2]. Iran's construction industry is no exception. The studies report that while only 29% of the total workforce are involved in the construction industry, roughly 40% of the reported incidents have occurred in the construction sector [3]. An unofficial report from the Bureau of Technical Execution System (2016) identifies the most common accidents in Iran's construction industry during 2012 - 2015 as follows: (1) Falls from heights, 56%; (2) Struck-by equipment and materials, 21%; (3) Caught in/between equipment or material, 15%; and (4) Electrocution, 3.5%.

Many types of research have been done to identify the safety influential factors (SIF) of projects. For example: a research reviewed 90 studies on the safety and identified 13 main factors (consisting of numerous sub-factors) influencing safety performance in construction projects [4]. This research aims to show how a model could be developed for the preliminary estimation of project safety level (PSL) by SIFs. In past literature, many different types of predictive models have been used for predicting PSL – for instance: fuzzy predictive tool, Bayesian belief network (BBN), linear models such as: Multiple Linear Regression (MLR) and etc.

## 2-**Methodology**

To achieve the objectives goals, first, accidents caused severe damage to workers have been identified by studying the unofficial statistics of accidents in the country. Because the presence or probability of occurrence of these risks has been used as a criterion for determining the level of safety in the present study, the high number of occurrences of these accidents or the high probability of their occurrence in a project indicates a low level of project safety and vice versa. Next, six factors that affect the level of safety of projects are identified and selected by reviewing the literature and unofficial statistics of accidents in the country, as follow:

- Safety investigation, hazard evaluation, and safety risk assessment processes ($SIF_1$) [5];
- Safety of scaffolds, access equipment, and protective structures ($SIF_2$) [6];
- Safe work procedure ($SIF_3$) [7];
- Safety training ($SIF_4$) [8];
- Maintenance, repair, and proper use of tools and machines ($SIF_5$) [9]; and
- Supervisory System ($SIF_6$) [10].

In the next step, the data of 95 high-rise building projects [based on Cochran's formula] were collected using a questionnaire; then, this information was entered into Bayesian and Regression models and processed to find a relationship between the factors affecting safety and safety level. The general research process can be summarized according to the flowchart of Figure 1.

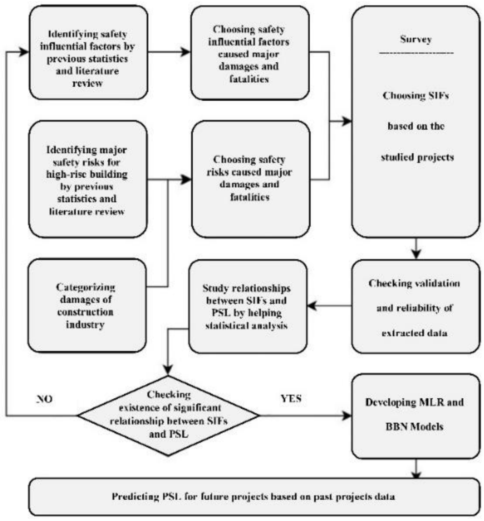


**Figure 1. Flowchart of Research Methodology**

Assessing the SIFs levels and PSL is very complex and ambiguous and be accomplished under many uncertain conditions; As a result of these complex and ambiguous conditions, the use of qualitative and linguistic terms is inevitable [11]. Therefore, a six-point Likert scale has been used to assess the levels of SIFs and PSL.

## 3-**Results and Discussion**

After a pilot survey, data from 95 steel structural building projects (18 - 26 floors) were collected through structured interviews and questionnaires. About 90% of the questionnaires were used to build the models, and the other 10% was applied to measure the accuracy of the models. Table 1 summarizes the SIFs and PSL statistics collected from the conducted interviews and questionnaires.

**Table 1. Descriptive statistics of selected safety data**

| SIF | Mean | Standard Deviation | Coefficient of Variation | Lower Level | Upper Level |
|---|---|---|---|---|---|
| **$SIF_1$** | 4.27 | 0.895 | 0.210 | 2 | 6 |
| **$SIF_2$** | 3.98 | 1.024 | 0.257 | 2 | 6 |
| **$SIF_3$** | 4.20 | 1.166 | 0.278 | 1 | 6 |
| **$SIF_4$** | 4.27 | 0.707 | 0.166 | 3 | 6 |
| **$SIF_5$** | 4.15 | 1.014 | 0.244 | 2 | 6 |
| **$SIF_6$** | 4.04 | 0.941 | 0.233 | 1 | 6 |
| **PSL** | 4.21 | 0.832 | 0.198 | 2 | 6 |

As the next step, MLR and Bayesian models were developed using the first category of extracted data to predict PLS based on previous projects' information. "SPSS" software was used to build the MLR model. The best and final regression model for predicting the PSL level can be predicted by MLR using four factors: $SIF_6$, $SIF_3$, $SIF_4$, and $SIF_2$, is represented by Eq. 1. To ensure that each independent variable is linearly related to the dependent variable, the assumptions of multiple linear regression should be properly verified before developing models: (1) Lack of multicollinearity, (2) Multivariate normality, and (3) Homoscedasticity. Violations of any of the following assumptions lead to the rejection of the MLR model [12].

$$PSL = (0.216 \times SIF_2) + (0.264 \times SIF_3) + (0.219 \times SIF_4) + (0.338 \times SIF_6) - 0.055 \quad (1)$$

According to the research results, Supervisory System "$SIF_6$" and safe work procedure "$SIF_3$" are the most influential factors on the project safety level of high-rise building projects with regression coefficients of 0.338 and 0.264.

In continue, "GeNIe V.2/00 Academic" software was utilized to build the Bayesian model. The main formula applied in the Bayesian Network is Bayes' probability formula, as presented in Eq. 2. The value of P(Hi|E) in this equation indicates how new values of probabilities update P(Hi). This update is a crucial concept of learning in Bayesian Networks using past experiences [7]. In this way, after the network is built, the Bayesian network variables are regularly updated with the entry of new data.

$$P(H_i|E) = \frac{P(H_i) \times P(E|H_i)}{\sum_{i=1}^{m} P(H_i) \times P(E|H_i)} \quad (2)$$

The Bayesian network also proved the direct relationship of all six factors to the target parameter, which is the PSL.

The MAPE calculated for the created models is 8.14%, which indicates a very accurate accuracy for the model. Also, the MAPE calculated equal to 11.78% for the predicted PSL by using the second category of data, which indicates the accurate prediction ability of models. Therefore, created models could provide a good forecast for similar new projects.

## 4- Conclusions

The primary purpose of this research is to show how to create a framework that can provide a step-by-step preliminary estimation to determine the PSL level of current projects based on information from similar projects. Therefore, two methods that have relatively high functionality among previous researches, including linear regression and Bayesian network, were studied, and the present study confirmed the application of both methods for the studied projects. An advantage of the Bayesian network over the linear regression model is the ability to predict PSL when some factors data is not available in a project. In this case, the Bayesian model possibly has high potential in predicting the future project safety level; but in this situation, the regression model needs sufficient information from all factors to predict the future.

## References

[1] D. Fang, H. Wu, Development of a Safety Culture Interaction (SCI) model for construction projects, Safety science, 57 (2013) 138-149.
[2] S. Gilani-nia, Optimization of safety processes in the construction industry by helping a decision support system, Journal of Business Management, 31 (2008) 159 - 178.
[3] A. Saberi, key project management skills, Faraa, 2006.
[4] A. Mohammadi, M. Tavakolan, Y. Khosravi, Factors influencing safety performance on construction projects: A review, Safety science, 109 (2018) 382-397.
[5] N.J. Bahr, System safety engineering and risk assessment: a practical approach, CRC Press, 2014.
[6] R.M. Choudhry, D. Fang, Why operatives engage in unsafe work behavior: Investigating factors on construction sites, Safety science, 46(4) (2008) 566-584.
[7] K. Imriyas, An expert system for strategic control of accidents and insurers' risks in building construction projects, Expert Systems with Applications, 36(2) (2009) 4021-4034.
[8] C. Tam, S. Zeng, Z. Deng, Identifying elements of poor construction safety management in China, Safety Science, 42(7) (2004) 569-586.
[9] J. Hinze, X. Huang, L. Terry, The nature of struck-by accidents, Journal of construction engineering and management, 131(2) (2005) 262-268.
[10] B.H. Guo, T.W. Yiu, Developing leading indicators to monitor the safety conditions of construction projects, Journal of Management in Engineering, 32(1) (2015) 04015016.
[11] S.Q. Wang, M.F. Dulaimi, M.Y. Aguria, Risk management framework for construction projects in developing countries, Construction Management and Economics, 22(3) (2004) 237-252.
[12] E. Pereira, S. Han, S. AbouRizk, U. Hermann, Empirical Testing for Use of Safety Related Measures at the Organizational Level to Assess and Control the On-Site Risk Level, Journal of Construction Engineering and Management, 143(6) (2017) 05017004.