# Application of Regression Analysis and Artificial Nueral Network in Construction Project

**Mrs. Anjana T**
Department of Civil Engineering
CCET, Valanchery, Kerala, India

*Abstract*— Project can be defined as a sequence of unique, complex, and attached activities having one goal or purpose and that must be completed by a specific time, within budget, and according to specification. Success of construction projects depends mainly on success of performance of a project. Project performance is predicted mainly based on the four performance metrics i.e., schedule and quality performance. This study is carried out to identify the factors affecting the performance of construction project and their significance on the schedule, and quality performance. The next objective of this research is to develop regression models and neural network models to predict schedule performance and quality performance.

***Keywords:*** Regression, Artificial Nueral Network, Correlation, SPSS Software, MATLAB Software

## I. INTRODUCTION

Construction trade is very large and multifaceted industry which plays a vital role in the progress of any nation. Every construction projects requires diverse teams to plan, design, construct and uphold the project. It is commonly agreed that for a construction project to be fruitful it has to be completed on time within budget and according to the specification. Forecasting project performance is one of the most demanding tasks in predicting whether the project will be successful. The effective performance of construction project cannot be achieved without challenges and obstacles. To meet these challenges and beat these obstacles, an organization must have a clear awareness of its performance.

Based on the description on the background above, the thought arises to examine the factors affecting construction project performance and to develop models for predicting construction project performance in terms of schedule and quality level. In most of the construction projects in Kerala, the estimation of cost and schedule is based on a guess work or rough estimate rather than proper calculated effort. Quality performance and satisfaction are often neglected in construction industry. These results in cost and time overrun in construction projects.

## II. RESEARCH METHODOLOGY

There are a tremendous number of qualitative and quantitative factors that affect the construction project performance. Creating a project performance models using multiple linear regression and artificial neural network is the prime objective of this research. A relevant literature review and meetings with the industry personnel helped to identify the significant factors which affect the performance of construction projects. After the identification of factors, structured questionnaires were distributed among the industry experts. Important factors are identified and ranked using RII method. The prediction models created using MLR and ANN is expected to accurately predict the actual performance of a construction project. The designed models are validated using collected data from the field. A comparison of the two model's results has been performed to select the appropriate model that provides closest results to the real world practice.

## III. SOFTWARES USED

The abbreviation of SPSS stands for Statistical Package for Social Sciences. SPSS is a comprehensive and flexible statistical analysis and data management solution. SPSS can take data from almost any type of file and use them to generate tabulated reports, charts, and plots of distributions and trends, descriptive statistics, and conduct complex statistical analyses. In addition to statistical analysis, data management (case selection, file reshaping, creating derived data) and data documentation (a metadata dictionary was stored in the data file) are features of the base software. SPSS is a Windows based program that can be used to carry out data entry and analysis and to create tables and graphs.

Microsoft Excel is a spread sheet application created by Microsoft. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for applications. It has been a very widely applied spread sheet for these platforms, especially since version 5 in 1993, and it has replaced Lotus 1-2-3 as the industry standard for spread sheets. Excel forms part of Microsoft Office. MATLAB is the high-level language and interactive environment used by engineers and scientists all over the world. It lets you explore and visualize ideas and collaborate across disciplines including signal and image processing, communications, control systems, and computational finance.

MATLAB can be used in projects such as modelling energy consumption to build smart power grids, developing control algorithms for hypersonic vehicles, analysing weather data to visualize the track and intensity of hurricanes, and running millions of simulations to pinpoint optimal dosing for antibiotics. In this study tool application is used to develop the ANN models. The data is imported in to the software. The input data set and the target data set are selected and the number of hidden nodes is also selected. The number of hidden nodes will be around the value of number of inputs. A range is selected between which all values of hidden nodes are given and corresponding network is obtained after training.

## IV. DATA ANALYSIS

The primary data was collected through self-administered questionnaire and analysed. A sample set of 30 was taken. The survey was conducted among builders and contractors of construction projects. The factors were rated using a Likert scale system ranging from 1 "very low" to 5 "veryhigh".

After the analysis of primary data, secondary data was collected directly from records of builders and through

 

structured questionnaire. The data includes schedule details and quality. The data were collected from different builders. Data analysis was further done with the help of SPSS 16 and MATLAB – 2015b to achieve the best fit project performance prediction models

$$RII = \sum a*100/A*n$$

## V. DATA ANALYSIS METHOD

Spearman's correlation analysis is done to identify the predictor variables which have a high degree of association with four performance measures (schedule growth and quality performance). It measures how closely a change in one variable is tied to the change in another variable, and vice versa. Random variables are treated symmetrically, where the correlation between X1 and X2 is the same as the correlation between X2 and X1. Spearman's correlation assesses monotonic relationships and appropriate for both continuous and discrete variables. The factors which have a high degree of association with the performance measures are selected and those factors are the input for ANN model. SPSS 16 is used for correlation analysis.

To validate that these variables are indeed the key determinants that affect project performance, ANN methodology was employed, to check whether these variables can reasonably predict project performance. ANN modelling was chosen because it has a robust learning capability, and produces fairly accurate predictions, even if information furnished is incomplete. Each ANN model is an expert system that can estimate one of the project performance metrics in Table 2, by learning from numerical examples and representing complex nonlinear relationships of these examples. ANN models have the self-learning ability, by adjusting their parameters to reduce the error of estimation.

The structure of the neural network model includes an input layer that receive input from the outside world, hidden layers that serve the purpose of creating an internal representation of the problem, and an output layer, or the solution of the problem. Before solving a problem, neural networks must be ''trained''.

Networks are trained as they examine a smaller portion of the dataset just as they would a normalsized dataset. Through this training, a network learns the relationships between the variables and establishes the weights between the nodes. Once this learning occurs, a new case can be entered into the network resulting in solutions that offer more accurate prediction or classification of the case. Neural network models are generally developed through the following six basic steps: Identify the problem, decide what information to be used and what will the network do; come to a decision of how to gather the information and symbolize it; define the network, select network inputs and identify the expected outputs; structure the network; train the network; and analyze the trained network. This engages addressing novel inputs to the network and evaluates the network's results with the authentic life results.

## VI. RESULTS

It was found that the neural network model predicted the actual schedule growth and quality in construction projects. This will help the practitioners to estimate the schedule performance and quality performance even before the construction stage.

Data sets of 30 projects were received from different construction firms. Correlation analysis was undertaken to identify variables/factors) that are significantly correlated to each performance metric. The final list of variables that affect each project performance metric is presented in Table 3.

Schedule performance of a construction project is measured in terms of schedule growth. It shows the delay beyond the date for completion specified by the contract. To predict the schedule growth an ANN model is developed with 15 input factors. MAPE value is 6.67% and the model is Robust.

Quality performance is measured qualitatively in a scale of 1-5. To predict the quality performance an ANN model is developed with 16 input factors. ANN model is acceptable because it has a low MAPE value of 8.67%.

## VII. CONCLUSION

The aim of this research was to identify the most important factors affecting construction project performance. The factors were identified through literature survey and a questionnaire was prepared to identify the most important factors among them. Once the survey response was checked for reliability the response was ranked based on their RII value. After checking the significance through one sample T test 16 significant factors were selected for second round data collection.

Artificial Neural Network is a machine learning technique suitable for complex cases requiring large number of parameters to be considered in parallel. Models were developed with the help of MATLAB. The models were trained and tested using the data collected from different projects. It was found that the neural network model predicted the actual schedule growth and quality performance in construction projects. This will help the practitioners to estimate schedule performance and quality performance even before the construction stage.

## REFERENCES

[1] Florence Yean Yng Ling and Min Liu, "Using neural network to predict performance of design-build projects in Singapore" Building and Environment 39, 2004;1263 – 1274 Rubio "Design Guidelines towards Compact Litho-Friendly Regular cells" SPIE Photomask Technology 2012
[2] Adnan Enshassi, Sherif Mohamed and Saleh Abushaban, "Factors affecting the performance of construction projects in the Gaza strip", 2009;JCEM, Vol.15, No.3, 20pp.:269-280.
[3] Kumar Neeraj Jha a & CT Chockalingam , "Prediction of schedule performance of Indian construction projects using an artificial neural network" Construction Management and Economics ISSN 0144-6193 ,2011;print/ISSN 1466- 433X
[4] Baccarini D, "The logical framework method for defning project Success". Project Management Journal 1999;30(4):25–32.
[5] Florence Yean Yng Ling, Sui Pheng Low, ShouQing Wang, and Temitope Egbelakin , "Models for Predicting

 

Project Performance in China Using Project Management Practices Adopted by Foreign AEC Firms" Journal of Construction Engineering and Management, 2008;Vol. 134, No. 12

[6] Sai On Cheung , Peter Shek Pui Wong , Ada S.Y. Fung , and W.V. Coffey "Predicting project performance through neural networks" International Journal of Project Management 24, 2006;24 207–215

[7] Dissanayaka SM, Kumaraswamy MM. "Evaluation of factors affecting time and cost performance in Hong Kong building projects", Engineering, Construction and Architectural Management 1999;6(3):287–98.

 