ASSESSMENT OF IMPACTS OF PROJECT TECHNICAL COMPLEXITY ON BUILDING PRODUCTION USING CLUSTERING AND KNOWLEDGE-BASED SYSTEM

Arthur W T Leung, Senior Lecturer, bsawtl@cityu.eud.hk
Division of Building Science & Technology, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong
C M Tam, Professor, bctam@cityu.edu.hk
Department of Building & Construction, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong

ABSTRACT

Site production layout planning is highly influenced by the technical complexity of a building project. Building structures, building layouts, scales of project and external site conditions are the major components affecting allocation and positioning of site facilities and construction plant. The relationships between these attributes are well known by experienced project managers and in the planning and tendering process, project managers and planners would assess and decide the site production layout by applying their cognitive knowledge using intuition. They recognize the benefit of using quantitative models in decision making, which, however present much difficulty when modeling the intwined and complex relationships between these variables. This study proposes an assessment model to examine impacts of technical designs, building layout designs and site conditions on building production with respect to the site layout plan using a data-based platform, which can assist decision making in site planning.

The system consists of two components, the Building Production Impact Score (BPIS) and the Building Production Impact Database (BPIDB). The BPIDB adopts the natural clustering technique, the self-organizing map (SOM), to classify building project samples in terms of technical complexity and compute the BPIS for the sample projects. The sample projects and their indices are uploaded to the BPIDB forming the data records. In the assessment platform, planners can input the project information of a new project, and the system will return with a complexity score and three sample projects with similar scores. The objective of the proposed system is to generate both a quantitative complexity score derived by the clustering model and the cognitive knowledge through the selected projects to improve the quality of decisions. The conceptual framework of the system will be discussed and illustrated with examples.

Keywords: technical complexity, building production, clustering, database

1. INTRODUCTION

Technological advancement and movements toward the sustainable built environment have imposed multifunctional requirements for modern building developments. In meeting these requirements and the quality expectations of owners, building designs in Hong Kong become more complex and impose more constraints on building production. Construction managers advise that it has been very difficult to assess the impacts of building designs and site conditions on building production and they have relied on their cognitive knowledge and intuition. They also comment that there are difficulties in defining and evaluating project scopes, project scales and project characteristics as there are a lot of variables affecting building production. Generally, the contract sum, project duration, site area, floor area and building height could provide an overall picture of a project for assisting managerial decision making. On the other hand, researchers have attempted to identify project performance indicators, critical success factors and buildability to assist project managers in managing projects or to promote productivity. However, previous findings and models were largely derived from subjective judgments through opinion surveys. It is not denied that the validity of the findings can be assured with a large sampling size but this could be costly and induces problems in data collection for the maintenance of the models. Currently, there are few industry-wide objective tools to define and evaluate the relationship between these factors and building production in terms of buildability and the recent quantitative models were developed by the Building Construction Authority, Singapore (2000) and Lam et al, (2006). This study attempted to evaluate the technical complexity of building construction using an objective natural clustering technique.

Clustering or classification is commonly used in physical sciences, chemistry, natural science, geology, and sociology, etc. A reliable and scientific classification can always assist subsequent research and applications, for examples, the periodic table in chemistry, the Enneagram for personality in social science, the managerial grid in leadership style, etc. This project proposes to develop the Building Production Impact Database (BPIDB) using Factor Analysis and Self-organizing Map (SOM) unsupervised clustering technique to develop an objective and unbiased tool to differentiate the technical complexity for building projects. The trained model can be updated from time to time so as to cope with changes in practice and performance of the construction industry. The project clusters derived form the reference for formulating the BPIS, the technical complexity score, and to define scales and characteristics of building projects from site production perspective. The proposed BPIS score provides objective measurements, criteria, yardsticks and standards for comparing building projects and guiding managerial decisions in resource allocation, and more importantly to establish the fundamental academic framework for researchers in long-term.

2. TECHNICAL COMPLEXITY OF A PROJECT

The terms “complexity of a project” or “project complexity” have been used by researchers as one of the variables in research of project management (Pariff and Sanvido, 1993; Chan and Kumaraswamy, 1996). Baccarini (1996) reported that the concept of project complexity has received little attention and is not clearly defined. Also, “complexity” has seldom been defined in a quantitative term and it was often regarded as one of the variables rather than being studied as the main theme.

The design, nature and characteristics of building projects have been studied from two major perspectives. The first approach aims at facilitating construction by promoting buildability. The second group of research has established models to explain and classify building projects from various perspectives to assist and enhance managerial decision making. The terms “Buildability” and “Constructability” shares a similar meaning for the complexity of construction or project complexity. The study of buildability could be the earliest systematic research in defining the ease of construction since the release of the Banwell Report (1964) in the United Kingdom.

Previous studies have also explored project management from different perspectives such as elements for measuring project performance or successes for management and procurement of building services (Shoesmith, 1996); the key performance indicators (KPIs) and critical success factors (CSFs) for exploring possible ways of improving the effectiveness of project management (Chan et al., 2004). The findings provide practitioners with an overall picture about the characteristics and nature of building projects, such as forms of structures, the height of building, the headroom of the floors, the usage of the building etc., which are essential for making strategic decisions such as in tendering, project staffing, risk management, quality management and safety management. However, the studies have not addressed the impact of site conditions and site space on construction planning. The study on site layout planning focused more on optimization models regarding spacing planning and positioning of site facilities. Various quantitative researches had been conducted to derive mathematical models in selection of plant, optimization of plant positions and prediction of hoisting times (Furusaka & Gray 1984; Gray & Little 1985; Wijesundera & Harris 1989; Choi & Harris 1991; Zhang et al. 1996; Leung & Tam 1998, 1999a & 1999b; Leung 2001). Tommelein et al. (1991, 1992 & 1999) conducted a series of study on site layout planning. They considered space is one of the sixth overlooked resources and of secondary important in construction planning. Hence, with the inclusion of site condition variables, this would provide a more realistic view about the characteristics and technical complexity of building projects. The proposed BPIDB will examine the technical designs and technical-related factors for evaluating impacts of technical complexity on building production for building projects. In this context, building production refers to site layout planning, construction productivity and plant utilization.

3. DEVELOPMENT FRAMEWORK OF THE BPIDB

The development framework of the BPIDB consists of three stages; they are: collection of project data, development of the Building Production Impact Database (BPIDB) and development of the interactive assessment platform. The BPIDB system is described below and illustrated in Figure 1.

3.1 Collection of project data

Project data is to be collected through studying of construction drawings from public archives available in the Buildings Department, Hong Kong SAR. The use of the project data provided in government archives enables a good and balanced coverage of the building projects completed in Hong Kong and can provide up-to-date project information for long-term study. The project information includes the following five groups:

1) Structural frame – variables describing the form of building structures.
2) Structural elements – variables describing the characteristics of the structural elements such as slabs, beams and transfer structures.
3) External finishing – variables describing characteristics of external finishes, features or components.
4) Building layout – variables describing special arrangements for floor and building layout.
5) Site conditions - variables describing the site location, site space and adjacent site condition.

3.2 Development of the Building Production Impact Database (BPIDB)

The Self-organizing Map (SOM) unsupervised clustering technique, which is one of the prominent unsupervised clustering algorithms, will be used to classify project samples into clusters (Kohonen, 1982 and 2001). The SOM has often been used with neural networks to reveal the similarity of objects for forming meaningful clusters. The algorithms of the Kohonen model is described briefly in the following steps (SOMLib Digital Library Homepage) and shown in Figure 2.

i) An input layer consisting of n-dimensional observations $\boldsymbol{x}$, and an output layer (represented by a grid) consisting of 7 x 7 units, each of which is associated with a p-dimensional weight $\boldsymbol{m}$.
ii) At the start of the *t* learning process, an input pattern is assigned randomly to the output nodes and the weights are modified. The distance is calculated between the observations and the vector associated with each neuron.
iii) The neuron with the smallest distance, the winning neuron $\boldsymbol{m_c(t)}$, and the neighbourhood neurons around the "winner" are updated. The winner's weight vector $\boldsymbol{m}$ is brought closer to the input patterns $\boldsymbol{x}$, ie. $\boldsymbol{m_c(t+1))}$.
iv) During the learning process, the weight decreases and determines the movement of the vector.
v) As more input observations are assigned to the clusters and the adjustment repeats, the number of neighbourhoods decreases.
vi) The process is iterated until the weights to the clusters are stabilized.

Since the SOM is a natural clustering technique, the number of cluster groups formed is to be assigned by users, the project clusters formed are to be reviewed for determining the number of cluster groups for the model. When the number of clusters has been determined, the project variables will also be classified into the same number of clusters where appropriate for establishing the Building Production Impact Score (BPIS). If a 5-cluster model is selected, the scoring scales would then be gauged into a 5-point scale. In the feasibility study, it is observed that nine project clusters are appropriate for representing different characteristics of 30 project samples. The selection of the 9-cluster model is based on spread and coverage of the project samples across the project variables. Also, the cluster model was examined using the Hierarchical Cluster Analysis (HCA). In HAC, the number of appropriate clusters can be determined by examining the changes in the distance coefficients, which are measured by the Euclidean distance. If there is a significant change in the values of the coefficients between two clustering stage, the agglomeration shall be stopped and the number of clusters can then be formed from either one of the

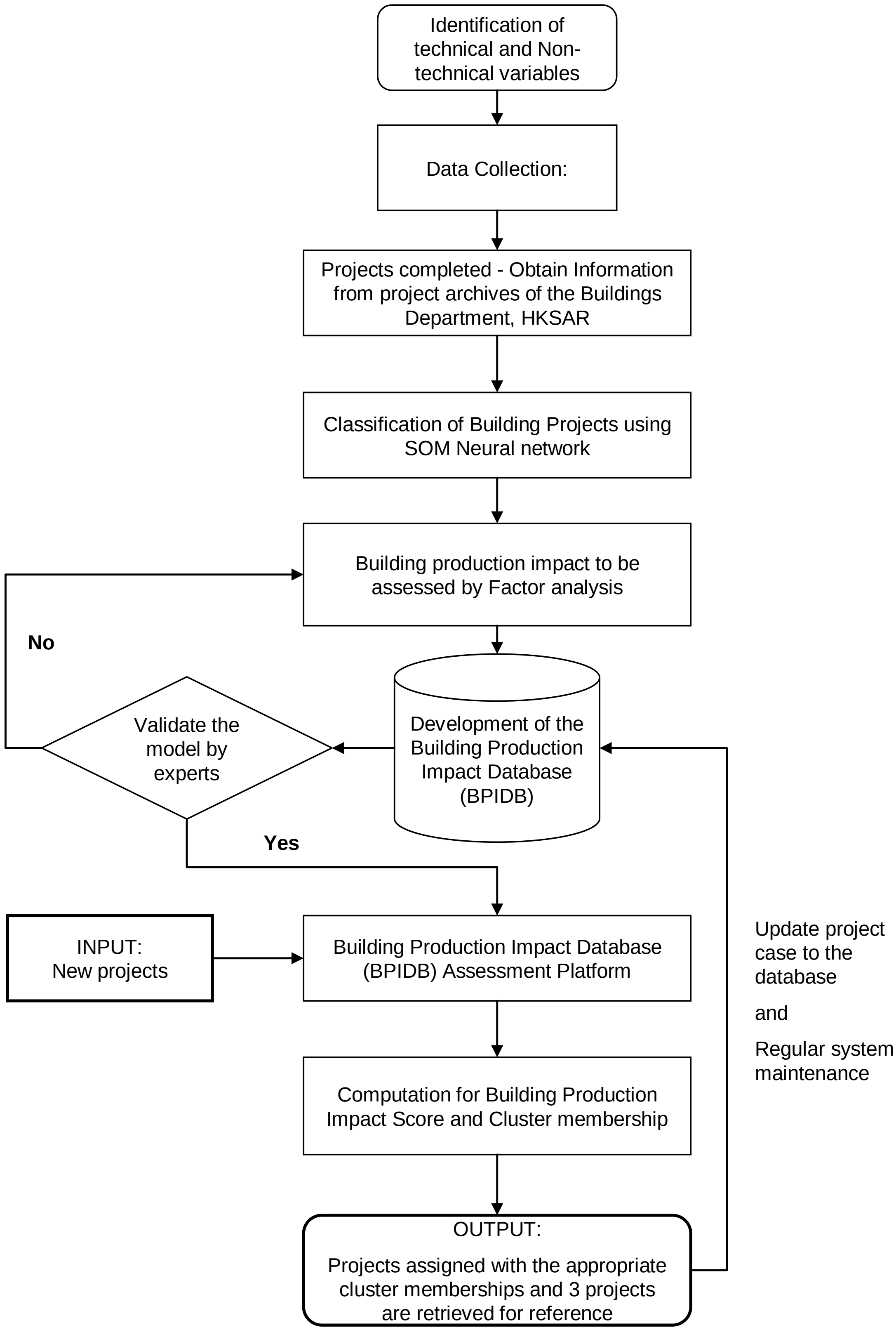

Figure 1 Development framework of the BPIDB

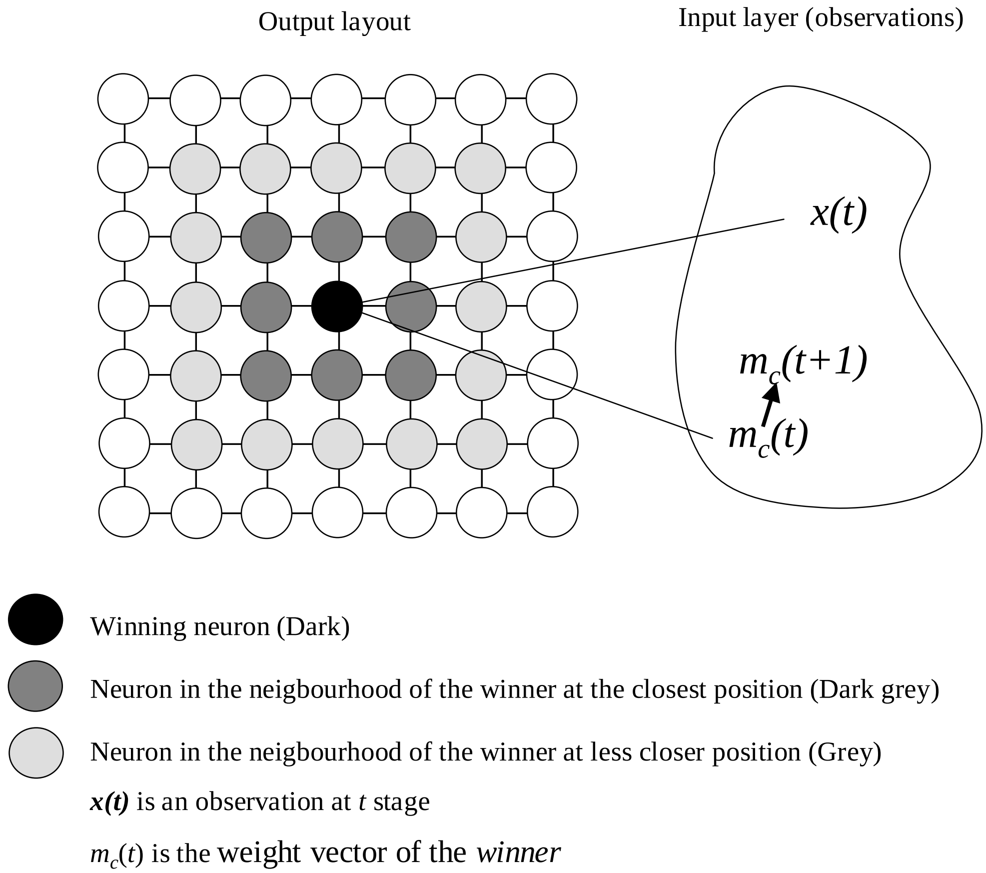

Figure 2: Architecture of a Kohonen SOM (adapted from SOMLib Digital Library Homepage)

stages. In this study, either a 8-cluster or 9-cluster model could be selected and a 9-cluster model was selected since it has a better fit to the SOM and HAC. Thus, the complexity scores for the variables are gauged to a 9-point scale based on the nine clusters formed by SOM as summarized in Table 1. The 9-cluster model has a good coverage of project scale, usage of building and site condition and is a good representation of the building projects in Hong Kong.

The BPIS scale is then derived using a 9-point scale. Table 2 shows an example for the BPIS scale for the variable "Typical floor area" classified by using the SOM. If a building project with a total typical floor area of 150,000 sq m, the floor area is within the range between 119,918 and 162,347 square meter. The BPIS for the project for the "Typical floor area" is 5. Further, the contributions of this attribute has to be determined by a weighted factor. The factor loadings derived by factor analysis are used to indicate the importance of the variables for computing the weighted BPIS. The use of factor analysis avoids the bias incurred in opinion surveys, which are generally adopted by previous researches. Thus, the contribution would be (5 x w_v), where w_v is the weighting for the corresponding variable. The total BPIS for a project is the summation of the weighted BPISs of the variables.

3.3 System Implementation

Visual Basic is used to build up the BPIDB. An Excel spreadsheet is used as the input and output interface. The BPIS scales generated by SOM is embedded in the system to compute project BPISs for assessing cluster memberships. The cluster memberships and BPIS are the important findings for helping practitioners to understand the characteristics of building projects with reference to the characteristics of the respective project clusters and the impacts of technical complexity on building production. The system input and output interfaces are illustrated in the example as shown in Table 3. The project database consists of 5 variable groups, namely Structural Frame, Structural Element, External Finishing, Building Layout and Site Condition. The project data, the variables, of a "new" project is input to a spreadsheet file which can be uploaded to the system. The project BPISs are calculated using the BPIS scales and the respective weighting. The system will examine the project data and will retrieve three projects with the highest similarity in the output report for user's review and comparison. With the use of the database system, users can compare the characteristics of the current project with the reference projects retrieved. The planning strategies and decisions could be reviewed, modified or adopted for the current planning exercises. The planning information of the "new" project would then be input to the system for future reference and thus the process keeps the system in line with the current planning practices. Apart from the use of the BPIS to gauge the characteristics of a project, the use of a data-base system enable the incorporation of other performance index into the system to enhance the reviewing and comparison processes.

Table 1 A 9-cluster Project Model

Project Cluster	**Project Description**
1	Small scale medium-rise, office buildings with good site conditions
2	Medium scale high-rise public housing buildings with very good site conditions
3	Medium scale high-rise private residential buildings with poor site conditions
4	Medium to large scale low-rise to very tall office and public buildings with poor site conditions
5	Large scale super high-rise private residential buildings with poor site conditions
6	Small scale high-rise public housing buildings with poor site conditions
7	Large scale high-rise public housing buildings with very good site conditions
8	Special large scale high-rise buildings with poor site conditions
9	Large scale low-rise private residential buildings with very good site conditions

Table 2 Building Production Impact Score scale derived by SOM for typical floor area

BPI Score	**Cluster Centre**	**Range**		
1	34,861	Below 22,381		
2	229,380	22,381	to	46,398
3	255,545	46,398	to	78,625
4	99,314	78,625	to	119,918
5	57,936	119,918	to	162,347
6	184,174	162,347	to	206,777
7	9,902	206,777	to	242,462
8	324,687	242,462	to	290,116
9	140,521	Above 290,116		

BPI Score =Building production impact score; Unit=square meter

Table 3 Interface for the BPIDB system

Project enquiry	***Residential development at ?***					**BPIS (0-100)**	**Project Performance Index**
Variable group	**Project variable**						Could be included in the system to assit decision making
Structural frame	V_{11}	V_{12}	V_{13}	…	V_{ij}		
	2	**4**	**3.2**	…	…	**17**	
Structural element	V_{21}	V_{22}	V_{23}	…	V_{ij}		
	…	…	…	…	…	**16**	
External finishing	V_{31}	V_{32}	V_{33}	…	V_{ij}		
	…	…	…	…	…	**6**	
Building Layout	V_{41}	V_{42}	V_{43}	…	V_{ij}		
	…	…	…	…	…	**13**	
Site condition	V_{51}	V_{52}	V_{53}	…	V_{ij}		
	…	…	…	…	…	**11**	
					Total BPIS	**63**	
Project Retrieved for review							
Project 28	…	…	…	…	…	**62**	
Project 45	…	…	…	…	…	**61**	
Project 55	…	…	…	…	…	**65**	

Project information and Project plans are available for review

Users can compare BPI scores between the projects retrieved and the new project

Possible extension for the BPIBD to integrate with project performance assessment

4. CONCLUSION

This proposal demonstrates the application of the BPIDB using factor analysis and SOM unsupervised clustering techniques with a database system to assess building production impacts. The project BPIS scores provide quantitative and objective assessment of technical impacts on production by building designs and site conditions. The BPIS computed for a 'new' project and the reference projects retrieved could be used to review the project characteristics and to assess the application of appropriate planning strategies, pricing strategies and project risks. The reliability of the system could be improved through regular maintenance of the database. However, the system reliability is also subject to the accuracy of factor analysis and project samples included. Therefore, the weightings need to be reviewed regularly with reference to professionals' opinions.

ACKNOWLEDGEMENTS

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administration Region, China (Project no.: CityU 117008).

REFERENCES

Baccarini, D. (1996), The concept of project complexity – a review, International Journal of Project Management, Vol. 4(4), pp.201-204.

Banwell, N. (1964), The placing and management of contracts for building and civil engineering work, HMSO, London. In: Griffith A. and Sidwell T. (1995), Constructability in Building and Engineering Projects, Macmillan Press Ltd., London.

Building Construction Authority (BCA), Singapore (2000), Code of Practice on Building Design, December 2000, Building Construction Authority, Singapore.

Chan, A.P.C, Scott, D. and Chan, A.P.L. (2004), Factors affecting the success of a construction project, Journal of Construction Engineering and Management, ASCE, Vol. 130, No. 1, February, pp. 153-155.

Chan, D.W.M. & Kumaraswamy, M.M. (1996), An Evaluation of construction time performance in the building industry, Building and Environment, Vol. 31(6), pp. 569-578.

Choi, C.W. & Harris, F.C. (1989), "A model for determining optimum crane position", Proceedings of the Institution of Civil Engineers, Vol. 90, June, pp627-634.

Furusaka, S. and Gray, C. (1984), "A model for the selection of the optimum crane for construction sites", Construction Management and Economic, Vol. 2, pp157-176.

Gray, C. and Little, J. (1985). A systematic approach to the selection of an appropriate crane for a construction site, Construction Management and Economic, Vol. 3, p121-144.

Kohonen, T. (1982), Self-organized formation of topologically correct feature maps, Biological Cybernetics, Vol. 43, pp. 59-69.

Kohonen, T. (2001), Self-Organizing Maps, 3rd ed. Springer Series in Information Sciences, New York.

Lam, P.T.I., Chan, P.C.A., Chan W.M.D., Chan H.W.E., Wong K.W.F. and Wong W.H.F. (2006), A model for buildability assessment in Hong Kong, Research Monograph, Hong Kong SAR China, The Hong Kong Polytechnic University.

Leung W.T.A. and Tam C.M. (1998), "Optimization of tower crane positions by using hoisting time prediction models", The Fifth International Conference on Tall Buildings, December, 1998, Hong Kong, pp. 1015-1020.

Leung,W.T.A. and Tam, C.M. (1999a), "Models for Assessing Hoisting Times of Tower Cranes", Journal of Construction Engineering and Management, ASCE, 125(6), 385-391.

Leung W.T.A. and Tam C.M. (1999b), "Prediction of hoisting time for tower cranes for public housing construction in Hong Kong", Construction Management and Economic, 17(3), May, pp. 305-314.

Leung W.T.A., Tam C M Tam and Tong Thomas KL (2001), "Estimation and Scheduling of Floor Construction Cycle for High-rise Building Construction", International Conference on Innovation in Architecture, Engineering and Construction (AEC), July 2001.

Parfitt, M.K. and Sanvido, V.E. (1993), Checklist of critical success factors for building projects, Journal of Management in Engineering, Vol. 9(3), pp. 243-249.

Shoesmith, D.R. (1996), A study of the management and procurement of building services work, Construction Management and Economics, Vol. 14, pp. 93-101.

SOMLib Digital Library Homepage, http://www.ifs.tuwien.ac.at/~andi/somlib/som.html.

Tommelein I.D. (1991), "Site Layout: "Where should it go?"", Construction Congress, Proceeding of the 1991 Conference, ASCE, pp.632-637.

Tommelein, I.D., Levitt, R.E., Hayes-Roth, B. (1992), "SightPlan Model for Site Layout". Journal of Construction Engineering and Management, ASCE, 118(1), January, pp.749-765.

Tommelein I.D. and Zouein P.P. (1999), "Interactive Dynamic Layout Planning", Journal of Construction Engineering and Management, ASCE Vol. 119(2), June, pp.266-287.

Wijesundera D.A. and Harris F.C. (1989). The selection of materials handling methods in construction by simulation, Construction Management and Economic, Vol. 7, p95-102.

Zhang, P., Harris, F.C. and Olonolaiye, P.O. (1996). A computer-based model for optimizing the location of a single tower crane, Building Research and Information, Vol. 24, No. 2, 1996, p113-123.