A QUANTITATIVE ANALYSIS OF DATA COLLECTED FROM THE LAST PLANNER SYSTEM IN BRAZIL

Rodrigo C. Bortolazza[1] and Carlos T. Formoso[2]

ABSTRACT

In Brazil, a large number of companies have implemented the Last Planner system in order to improve the performance of their production systems. However, most research studies developed so far have emphasized the analysis of qualitative data, based on a small number of case studies. Therefore, there seems to be a good opportunity to learn more about production control by analyzing data from a larger sample of projects, based on performance measures such as the percentage of plans completed (PPC) and the causes for the non completion of work packages.

This paper presents further developments of a research project that was first presented at the IGLC13 Conference. It is based on data from 133 projects. Some of them have been investigated in previous research projects at the Building Innovation Research Unit (NORIE) of the Federal University of Rio Grande do Sul (UFRGS). The remaining data have been provided by companies that have been using the Last Planner system for several years.

At first, the average PPC and the reasons for the non completion of tasks were analysed using descriptive statistics. In a sample of 96 projects multivariate regression analysis was used for explaining the variance of weekly PPC, using as independent variables the causes for non completion of work packages, number of assignments completed and timing of the projects.

The results indicated that a major problem in most projects is the lack of effective implementation of look-ahead planning. Moreover, for multivariate regression analysis, some causes for the non completion of work packages, such as work force and planning have a strong contribution for explaining the variance of weekly PPC.

KEY WORDS

Last Planner, performance measurement, multivariate regression analysis.

INTRODUCTION

Nowadays, construction companies have been looking for new ways of improving the performance of their production systems in order to reduce time and cost in projects, besides increasing quality and productivity. In a highly uncertain environment production control has a key role in creating stability and predictability. The Last Planner system has been successfully used for achieving such aims, by using a mechanism known as shielding production (Ballard and Howell 1997). According to these authors, it has been claimed that it can potentially increase productivity and reduce variability in construction projects.

An evidence of its effectiveness is the fact that the Last Planner system has been implemented in the construction industry in a large number of companies from several countries since 1992 (Ballard and Howell 2003). In Brazil, a large number of construction companies have implemented the Last Planner since the mid Nineties, with the aim of improving the performance of their production systems. Initially, most initiatives were supported by research teams and consultants, but more recently many companies have started implementing by themselves. Despite its dissemination across industry, most research studies developed so far have mostly analysed qualitative data, based on single or on a small number of case studies (see, for instance, Oliveira 1999, Alves 2000, Marchesan and Formoso 2001, Bernardes and Formoso 2002, Soares et al. 2002).

Some recent studies have emphasised the importance of quantitative analysis in order to evaluate the effectiveness of the implementation of the Last Planner system as well as its impact. A few papers from Chile (Alarcón et al. 2005), Colombia (Botero and Alvarez 2005) and Brazil (Bortolazza et al. 2005) have been published on that matter. Ballard (2000) pointed out that it is necessary to quantify and understand the benefits of greater plan reliability for safety, quality, time and cost. There seems to be a good opportunity for investigating those benefits by comparing measures of plan reliability to projects goals. The main data produced through the implementation of the Last Planner system are the percentage of plan completed (PPC) and the causes for the non completion of work packages.

This paper presents an analysis of Last Planner data, which aims to discuss the plan reliability that was achieved in a sample of projects, as well as identify possible causes for implementation barriers.

Previous results of this research project were first presented at the IGLC13 Conference (Bortolazza et al. 2005). This study is part of the SISIND-NET Project, which has been developed by NORIE/UFRGS with the partnership of the Association of Contractors of the state of Rio Grande do Sul (SINDUSCON/RS), and the support of the National Council for Scientific and Technological Development (CNPQ). This project involves the design and implementation of a performance measurement system for benchmarking for construction companies. One of the aims is to devise a web site for collecting and disseminating data.

RESEARCH METHOD

A database was built at the Building Innovation Research Unit (NORIE) of the Federal University of Rio Grande do Sul (UFRGS), using data from 133 projects, carried out between 1996 and 2005. Some of those projects (57.9%) have been studied in previous research projects (M.Sc. dissertations and Ph.D. theses). The remaining data (42.1% of the projects) have been provided by companies that have been using the Last Planner system successfully for several years.

Initially, such data was prepared using some Knowledge Discovery in Databases (KDD) criteria and techniques, in order to improve the quality and the accuracy for further analysis. Initially, a sample of 133 projects was described by univariate analysis of PPC. The sample of projects was divided into four market sectors: industrial building, residential and commercial building, low-income housing and public sector. The causes for the non completion of work packages were investigated in 105 projects of the sample, in which such data was available. Also, a set of independent variables were used to explain the variance of weekly percentage of plans completed (PPC) in 96 projects.

Data mining tools such as decision trees and neural networks were initially used (Bortolazza et al. 2005). Afterwards, the decision was made to replace them with multivariate regression analysis, since those data mining techniques had increased complexity of the analysis without enough information about which features were more important to explain the dependent variable.

Database Creation

Data were initially stored in spreadsheets. In each of them, a project was described by the number of tasks planned, number of assignments completed and the causes for the non completion of work packages. The non completed packages were divided in nine groups, according to the main cause: work force, materials, equipments, design, planning, clients, weather conditions, suppliers and unknown (when the main reason was not described). Moreover, additional information about the origin of data (previous research studies or companies), start date, market sector, company, etc. were also stored in the database.

The average PPC and the total number of causes for the non completion of work packages were summarised in one spreadsheet. Several analyses were carried out: PPC histograms, descriptive statistics, analysis of variance (ANOVA) and the percentages of each category of problems in the project sample.

Basic descriptive statistics were used to characterize the projects in terms of data dispersion and central tendency. According to Han and Kamber (2001), the distribution is best represented by the following measures: minimum, Q1, median, Q3 and maximum. Besides, the standard deviation and the coefficient of variation (rate between standard deviation and the average) were also used as measures of dispersion. Regarding central tendency, "…the median best represents the real situation of the industry sub sector and has the effect of filtering out out-of-range data, which are included in calculation of the mean" (Ramírez 2004).

The analysis of variance (ANOVA) were used to assess whether the variances of project average PPC could be explained by some project variables. For instance, the impact of constraint analysis of the average PPC was analysed.

In another spreadsheet, the weekly number of assignments planned, number of assignments completed, PPC, number of causes for the non completion of work packages (classified into nine categories), timing and the market sector for 96 projects were stored. These data were exported to the statistical software SPSS[3], in which the multivariate regression analysis was undertaken.

[3] In this paper, data were analysed with Version 13.0 of the Statistical Package for Social Science – www.spss.com

Data Preparation

The stage of data preparation is very important for the quality of the results and can be described independently of an application (Weiss and Indurkhya 1998). For that reason, data preparation practices and techniques from Knowledge Discovery in Databases (KDD) (Cabena et al. 1997, Han and Kamber 2001, Pyle 1999, Weiss and Indurkhya 1998) were adopted to improve the data base for multivariate regression analysis. Pyle (1999) and Cabena et al. (1997) suggest that data preparation may take about 60% of the time needed for the entire process.

According to Cabena et al. (1997) data preparation can be divided into three phases: selection, preprocessing and data transformation. The third stage was not carried out in this research work, since it is more appropriate for techniques such as neural networks and nearest-neighbour methods (Weiss and Indurkhya 1998).

The goal of data selection is to identify the available data sources and extract interesting values for the preliminary analysis (Cabena et al. 1997). The design and organization of data base, including the setting of goals and the composition of features, is done according to the necessity of the researcher (Weiss and Indurkhya 1998).

Data preprocessing is the most difficult phase because real world data tend to be incomplete, noisy and inconsistent (Cabena et al. 1997, Han and Kamber 2001). Noisy data occur when one or more variables have values that are significantly diverging from what is expected for those ones. These observations are called outliers. However, this expression does not mean an undesirable value. They may reveal, for instance, some characteristics that would not be discovered in normal analysis. Therefore, the problematic cases are no more than invalid data. In these cases, outliers must be treated with some techniques for missing data (Cabena et al. 1997, Hair et al. 1998, Pyle 1999).

In this research work, the stage of data preparation started with outlier detection. Hair et al. (1998) have suggested that one identify the cases with measures larger than 3 or 4 standardized residuals (rate between error and the standard deviation) in samples larger than 80 observations. In this sample, 1.06% of total cases were identified with more than 3 standardized residuals. Most data between 3 and 4 standardized errors corresponded to low weekly PPC (lower than 50%). According to Hair et al. (1998), in this case, such measures should be kept, unless there are concrete evidences that they do not represent any possible case in the population.

Multivariate Regression Analysis

At this stage, special care is necessary for checking the sample size. Hair et al. (1998) suggest the proportion from 15 to 20 cases for each independent variable in order to make statistical generalization possible. Linear relationships among independent variables and the dependent one were also evaluated in this sample. In the variable "number of assignments completed" a log transformation was necessary to make this variable linearly related to the weekly PPC.

The multivariate regression analysis produced several statistic measures. In this sense, for Leech et al. (2005) adjusted R^2 contribute to comparisons among the regression models looking for the equation that explains more the dependent variable with the independent ones. For predictors, standardized beta coefficients highlight the independent variables that are more relevant. Moreover, the value t and significance indicates whether the variable is significantly contributing to the equation for predicting the dependent variable from the whole set of predictors (Leech et al. 2005).

RESULTS

Average PPC and Causes for the non-completion of Work Packages for the Projects

Initially, the whole sample of projects was divided according to the market segment. Table 1 presents average PPC, standard deviation, coefficient of variation, and the sample size. Figure 1 presents other measures of central tendency (median) and data dispersion (minimum, Q1,Q3 and maximum).

Table 1: Average PPC, standard deviation, coefficient of variation and sample size

	All Projects	Industrial building	Residential and commercial building	Low-income housing	Public Sector
Average	70.40%	72.24%	68.04%	68.90%	69.62%
Standard deviation	11.99%	10.30%	12.71%	13.42%	16.69%
Coefficient of variation	17.03%	14.26%	18.67%	19.48%	23.97%
Sample	133	67	36	21	9

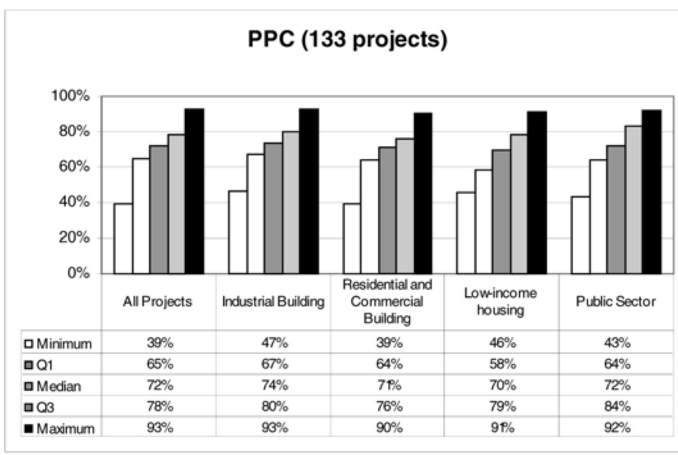

Figure 1: Data dispersion in the sample of 133 projects

Figure 1 indicates that PPC tends to higher in the industrial building sector. It is also the market segment in which the dispersion is the lowest. In contrast, low-income housing has the lowest measures for minimum, Q1 and median. Moreover, the coefficient of variation in this sample is one of the highest. In the public sector, the coefficient of variation is higher than low-income housing, but the sample is too small. The PPC histograms from industrial building and low-income housing are presented in figures 2 and 3, respectively.

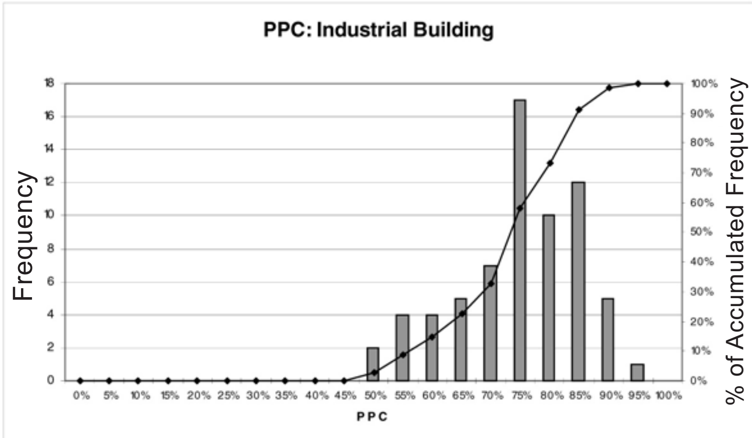

Figure 2: PPC histogram for industrial building market sector

In figure 2, the histogram confirms a better distribution compared to low-income housing (figure 3). In this market sector, the main reason for this relatively low PPC was that look-ahead planning and constraints analysis was not properly implemented in those projects. In figure 3, for instance, about 30% of values have a PPC lower than 60%.

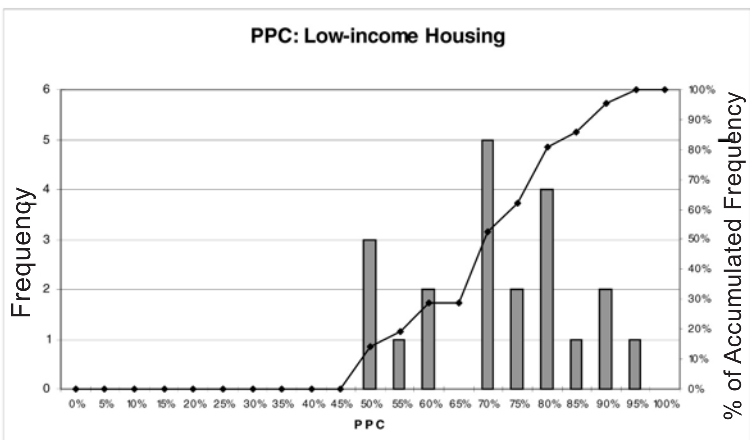

Figure 3: PPC histogram for low-income housing market sector

However, an analysis of variance (ANOVA), comparing average PPC in these four different market sectors resulted in a present probability (or p-value) of 0.34. This result exceeds the significance level considered of 0.1. Therefore, this evaluation was not statistically significant.

The industrial building market sector was also compared to the other projects. In this case, this analysis was statistically significant (p-value=0.07) and the average PPC of 72.24% for industrial building was higher when compared to another projects (68.53%). New analyses were done in this market sector, trying to explain this difference.

The industrial building project sample is mostly made up by projects from two different construction companies. Company "A" carried out 44.78% of the projects, while Company "B" for 38.60%. The remaining projects (22.39%) were carried out by other companies, but they all had the involvement of researchers in the implementation of the control system. Table 2 presents the average PPC and the standard deviation for those three groups.

Most Company "A" projects have high complexity, due to short lead time, interference from the clients in the production process, and high product flexibility (Soares et al. 2002). That construction company has been implemented Last Planner system since 1999 and production control is considered by their top managers as a critical process for the company competitive advantage.

In company B, the Last Planner system has been implemented since 2000 and most projects are university buildings. These are not so complex and uncertain when compared to Company A industrial projects. The ANOVA for comparing average PPC of those three subgroups indicate that their differences are statistically significant (p-value<0.01).

Table 2: Descriptive statistics in industrial building sector

Projects	N	Average	Standard Deviation	Minimum	Maximum
Company A	30	67.93%	10.14%	50.00%	89.00%
Company B	22	80.31%	6.26%	64.00%	93.00%
Others	15	69.03%	8.69%	47.00%	78.00%
TOTAL	67	72.24%	10.30%	47.00%	93.00%

Based on these analyses, some factors that might affect the average project PPC were identified. These are: constraints analysis at the look-ahead planning level, ISO9001 certified quality management system or research team intervention. The impact of the complexity in the projects, which seems to be very important, has not yet been evaluated. An ANOVA test resulted in statistically significant for those three factors (p-value<0,1). This comparison is summarised in table 3.

Table 3: Factors that affect the project average PPC in the projects

	Constraints Analysis		**Quality Management System**		**Researcher Intervention**	
	Yes	No	Yes	No	Yes	No
Average PPC	67.50%	71.69%	73.67%	66.18%	67.54%	72.18%

In table 3, projects in which constraints analysis was systematically carried out are supposed to have a higher average PPC. However, these results indicate that this process has not been implemented completely. Regarding quality management systems, the standardized process control in these projects also has contributed for plan relability. Finally, in those projects that had interventions from research teams the Last Planner system was in the early stages of implementation. For that reason, the average PPC tends to be lower compared to projects carried out by companies that had used the Last Planner for several years.

In 105 projects, the causes for non-completion of work packages were also investigated. Table 4 presents the main causes for the non completion of work packages.

Table 4: Causes for the non-completion of work packages in 105 projects

	All Projects	Industrial building	Residential and commercial building	Low-income housing	Public Sector
Work force	34.77%	30.24%	40.74%	42.45%	32.71%
Materials	5.82%	5.76%	7.85%	2.99%	11.21%
Equipments	3.39%	3.60%	1.69%	4.83%	2.80%
Design	2.83%	3.81%	2.02%	0.80%	0.93%
Planning	30.29%	30.64%	33.10%	26.66%	21.96%
Clients	4.08%	6.76%	0.21%	0.00%	4.67%
Weather conditions	14.17%	14.67%	9.46%	17.44%	21.50%
Suppliers	4.66%	4.52%	4.92%	4.83%	4.21%

When there is no systematic production control, external causes such as weather conditions are often pointed out as major sources of delays. However, considering the five groups of problems that are predominantly concerned with internal causes (work force, materials, equipments, design and planning) they correspond to approximately 70% of all problems. Weather conditions are more relevant in horizontal building sites, such as low-income housing and public sector (public schools, roads and electrical substations). However, even in those cases this category of problems corresponded to less than 25% of the all problems.

Multivariate Regression Analyses

A multivariate regression analysis was conducted to determine the best linear combination of the independent variables. The following predictors were tested: (a) eight categories of causes for the non-completion of work packages, (b) the variable "unknown" used for tasks that were not completed and did not have their problems identified, (c) number of assignments completed, (d) number of weekly plan work packages; and (e) timing of the project. As the "number of assignments completed" and the "number of weekly plan work packages" are significantly intercorrelated ($R=0.91$ with $p<0.01$), the variable with lower intercorrelation with PPC were not used for generating the model (in this case, the number of weekly plan work packages). Furthermore, the number of assignments completed was replaced by the logarithm of this value, in order to obtain a linear relation with PPC.

In the first model generated, the variable "timing" does not significantly contribute to the prediction and it was excluded ($p>0.1$). In the new model, the combination of the remaining variables significantly predicted PPC, all of them significantly contributing to the prediction.

The beta weights and t values, presented in table 5, suggest that the logarithm of the number of assignments completed contribute most to predicting PPC and that "unknown" problems, weather condition, work force and planning also contribute significantly for this prediction. The adjusted R squared value was 0.63%. This indicates that 63% of the variance in PPC was explained by the model.

However, about 37% of variance in PPC was not explained. For this reason, new models were evaluated in each market sector. The results, with their respective beta weights and t values for the statistically significant variables ($p<0.1$) are also in table 5.

Table 5: Multiple regression analysis explaining variance in PPC

	All Projects		Industrial building		Residential and commercial building		Low-income housing		Public Sector	
	β	t	β	t	β	t	β	t	β	t
CONSTANT		118,43		88,703		43,405		49,722		33,922
Work force	-,271	-17,08	-,246	-11,64	-,324	-11,34	-,356	-10,22	-,494	-8,900
Materials	-,048	-3,107	0	0	-,119	-4,301	-,179	-5,52	-,230	-4,225
Equipments	-,101	-6,841	-,113	-5,790	-,052	-1,898	-,126	-3,90	-,221	-3,773
Design	-,110	-7,325	-,150	-7,633	-,066	-2,417	0	0	0	0
Planning	-,250	-15,56	-,259	-11,96	-,424	-15,32	-,190	-5,75	0	0
Clients	-,125	-8,374	-,166	-8,452	0	0	0	0	-,136	-2,443
Weather conditions	-,286	-19,42	-,305	-15,80	-,133	-4,663	-,306	-9,44	-,369	-6,695
Suppliers	-,126	-8,516	-,116	-5,987	-,182	-6,808	-,160	-4,975	0	0
Unknown	-,300	-20,27	-,374	-19,17	-,363	-13,66	-,122	-3,369	-,258	-4,727
Log (Assignments Completed)	,632	41,344	,596	29,570	,593	22,102	,641	16,444	,659	11,184
Timing	0	0	0	0	0	0	,105	3,000	0	0
N	1720		992		296		359		73	
Adjusted R^2	62,90%		63,40%		79,6%		63,40%		79,40%	

According to table 5, the variables that contribute more for predicting PPC in the market sectors are: the logarithm of the number of assignments completed and work force. The variable “unknown” is more relevant in industrial building and residential and commercial building projects. It indicates that the larger the number of non completed work packages that do not have their causes identified the lower tends to be the PPC, since there is less feedback for production control. In low-income housing and in the public sector, weather conditions should be considered as an important factor, which confirms previous analyses with descriptive statistics. In the public sector, however, the category planning was excluded from the model because it was not statistically significant. In table 5, the highest t values in all the models are for the constant value. It indicates that there is a problem, because in the whole set of predictors the selected variables were not so important for predicting PPC when compared to a constant used to the adjustments of the model.

Regarding the adjusted R^2, in the models generated for industrial building and low-income housing projects, about 37% of variance in PPC was not explained. In industrial building maybe some feature related to the degree of complexity could improve the adjusted R^2. In low-income housing, the high importance of weather conditions could inform us of the relevance of the coverage to these projects. Therefore, when these houses do not have their roof completed the PPC values tend to be lower.

CONCLUSIONS

The only data available from the implementation of the Last Planner system in Brazil for a large number of projects is the weekly PPC and the causes for the non-completion of work packages. Therefore, data analysis so far has been limited to the implementation of the Last Planner system. In the future, as more data related to project goals (such as cost, duration, and customer satisfaction) are collected, it will be possible to analyse the impact of plan reliability in project performance.

In this paper, the multivariate regression analyses have suggested that about 37% of the variance in PPC can be explained by predictors that were not collected. For example, a variable representing the degree of project complexity might contribute improving the explanation of PPC in industrial building projects. Regarding the existing variables, only in low-income housing projects is the category "unknown" problems not significantly important. This means that when the weekly plan is poorly assessed, plan reliability is affected significantly.

Moreover, there should also be an effort in terms of improving the completeness and consistency of the short term planning information, in order to enable these data to be used for statistical analysis. That includes training the professionals in charge of coordinating weekly planning meetings, and developing a more appropriate classification of the causes for the non-completion of work packages. For instance, the two main categories of problems (planning and work force) seem to be too superficial for identifying the root cause.

ACKNOWLEDGEMENTS

This research project was made possible due to the financial support of CAPES (Coordenação para o Aperfeiçoamento de Pessoal de Ensino Superior) and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

REFERENCES

Alarcón, L.F., Diethelm, S., Rojo, O., Calderon, R. (2005). "Assessing the Impacts of Implementing Lean Construction". *Proceedings of the 13th annual conference of the International Group for Lean Construction*, Sydney.

Alves, T.C.L. (2000). "Diretrizes para a Gestão dos Fluxos Físicos em Canteiros de Obras: proposta baseada em estudo de caso". Porto Alegre: Post-graduate program of Civil Engineering, Federal University of Rio Grande do Sul. M.Sc.Dissertation.

Ballard, G. (2000). "The Last Planner System of Production Control". Birmingham: School of Civil Engineering, Faculty of Engineering, University of Birmingham. Ph.D. Thesis.

Ballard, G., Howell, G.A. (2003). "An Update on Last Planner". *Proceedings of the 11th annual conference of the International Group for Lean Construction*, Virginia.

Ballard, G., Howell, G. (1997). "Shielding Production: An Essential Step in Production Control". *Technical Report No.97-1*, Construction and Engineering Management Program, University of California, p.102-109.

Bortolazza, R.C., Costa, D.B., Formoso, C.T. (2005). "A Quantitative Analysis of the Implementation of the Last Planner System in Brazil". *Proceedings of the 13th annual conference of the International Group for Lean Construction*, Sydney.

Botero, L.F., Alvarez, M.E. (2005). "Last Planner: an advance in planning and controlling construction projects. Case study in the city of Medellin". *Proceedings of the 4th Brazilian Symposium on Construction Management and Economics*, Porto Alegre.

Bernardes, M.M.S. and Formoso, C.T. (2002). "Contributions to the Evaluation of the Production Planning and Control Systems in Buildings Companies". *Proceedings of the 10th annual conference of the International Group for Lean Construction*, Gramado, Brazil.

Cabena, P., Hadjinian, P., Stadler, R., Verhees, J. and Zanasi, A. (1998). "Discovering Data Mining: from concept to implementation". Prentice-Hall PTR: Upper Saddle River (NJ).

Hair, J.F., Anderson, R.E., Tatham, R., Black, W. (1998). "Multivariate Data Analysis". Upper Saddle River: Prentice-Hall.

Han, J. and Kamber, M. (2001). "Data Mining: concepts and techniques". Morgan Kaufmann Publishers.

Leech, N.L., Barrett, K.C., Morgan, G.A. (2005). "SPSS for Intermediate Statistics: use and interpretation". Lawrence Erlbaum Associates Publishers.

Marchesan, P.R.C. and Formoso, C.T. (2001). "Cost Management and Production Control for Construction Companies". *Proceedings of the 9th annual conference of the International Group for Lean Construction*, Singapore.

Oliveira, K.A.Z. (1999). "Desenvolvimento e Implementação de um Sistema de Indicadores no processo de Planejamento e Controle da Produção: proposta baseada em um estudo de caso". Porto Alegre: Post-graduate program of Civil Engineering, Federal University of Rio Grande do Sul. M.Sc.Dissertation.

Pyle, D. (1999). "Data Preparation for Data Mining". Morgan Kaufmann Publishers.

Ramírez, R.R., Alarcón, L.F.C., Knights, P. "Benchmarking Systems for Evaluating Management Practices in the Construction Industry". *Journal of Management in Engineering.* July, 2004. p.110-117.

Soares, A.C., Bernardes, M.M.S., Formoso, C.T. (2002). "Improving the Production Planning and Control System in a Building Company: Contributions After Stabilization". *Proceedings of the 10th annual conference of the International Group for Lean Construction*, Gramado, Brazil.

Weiss, S.M. and Indurkhya, N. (1998). "Predictive Data Mining". Morgan Kaufmann Publishers.