VALIDATION OF A COST MODEL FOR THE SUPERSTRUCTURE SERVICE IN JUVENILE PRISONS IN BRAZIL BY MEANS OF THE MONTE CARLO SIMULATION

Camila Isaton [a], Antônio Edésio Jungles [b], Jamil José Salim Neto[c]

[(a),(b),(c)] Federal University of Santa Catarina - Department of Civil Engineering, Laboratory Construction Management, Florianópolis, Brazil

[(a)] camila.isaton@pucpr.br, [(b)] ajungles@gmail.com, [(c)] jamil@unifap.br

ABSTRACT
The typology of edification studied in this research is limited to Juvenile Prisons. The parametric estimation via Multiple Linear Regression is the method used to create the cost model of the Superstructure service. The model presents response speed for the cost forecast, since the equation generated can be applied through simple information available in the phases Initials of the enterprise. The objective of the article is to present the construction of the cost model for the Superstructure service and its validation is given through the Monte Carlo Simulation. The model presented a power of explanation of costs in the order of 71%. In the Monte Carlo simulation validation it was possible to detect that for scenarios with too small or too big total areas the model presented non-applicable results, that is, validation was performed for the extreme conditions.

Keywords: Estimativa Paramétrica, Prisões Juvenis, Modelo de Custo, Simulação de Monte Carlo.

1. INTRODUCTION

Brazil occupies the 4th place among the countries with the largest prison population (International Center for Prison Studies, 2014). About 7.4% of the Brazilian prison population is composed of young people under the age of 18, and these are distributed in 350 prison units in the country (Secretary of Human Rights of the Presidency of the Republic of Brazil, 2012).

The research approaches the cost estimate for Juvenile Prisons in Brazil. The result of the research is a parametric cost equation that allows the fast calculation of the Superstructure servisse costs during the cost estimation.

The need to estimate costs for juvenile detention in Brazil is due to the lack of visibility of this topic at the academic level in this country, presenting itself in this way as a relevant analysis of social and budgetary importance in the conception of new enterprises.

The parameters that feed the cost model are like geometric characteristics of the buildings. Parisotto (2004), presents that the cost estimates with the analysis of geometric characteristics demand less complex information, that is, as accessible in the initial phases of the project.

The motivation for conducting the research also counts on the need for new Juvenile Prisons in Brazil. The Carta Capital newspaper (2014), through the Access to Information Law, published that among the 148 Socio-Educational Detention Units that make up the socio-educational system of the state of São Paulo (the largest state in population and wealth in the country), 54% of these are overcrowded.

The judicial website JusBrasil (2012) revealed that another Brazilian state, called Minas Gerais, only young perpetrators of serious crimes are interned in the 22 units of the state, young people suspected of homicide, armed robberies and rapes count on the incapacity of the system, which cannot receive them for the fulfillment of the sentences, the Socio-Educational Detention Units of this state show overcrowding in the order of 48.7%.

Brazil coexists with crime increasing and the need to create vacancies for young people to comply with the penalties in a closed regime, starting from this motivation, originates the research. The article is part of a master's degree, where, from a study involving 39 projects, a cost model was developed for the Superstructure Service. This service, chosen for analysis, was chosen because of its representativeness in the ABC curve of Services, for this type of building, representing 18,31% of the total cost.

The generated model is classified as statistical, the methodology used for the construction of the model is the parametric estimation, the method of construction of the model is developed with the help of the software "*Statistica* in the *Ultimate Academic* version",. The validation of the model is performed through Monte Carlo Simulation, with the support of Excel software.

It is expected that with the present research, to assist the public administration (who finances and operates the enterprises of that origin in Brazil) in the speed of response during the process of cost analysis and generation of cost estimate for Juvenile Prisons. Also, the objective is to draw the attention of the academic community to criminal investigations of public security function, which are currently little explored in research, but have a great social relevance, due to its function in meeting the demands and needs of the society.

2. MODEL OF COST FOR THE SUPTRUCTION SERVICE OF JUVENILE PRISONS IN BRAZIL

2.1.1. Objective

Development of a statistical model for an estimation of costs of the Superstructure Service with validation of this model via Monte Carlo Simulation, a service related to the scope of buildings for use of Juvenile Prison in Brazil.

2.1.2 Search Limitations

The cost information considered in the database refers to the contracted values of the project. The price readjustment are defined due to the technological changes that occurred during the execution and were incorporated into the raw data base, thus, the results seek to be compatible with the actual values spent in the execution of works.

The cost model of this research is not applied to construction systems different from the conventional Brazilian system (pillars, beams, slabs and reinforced concrete).

3. REVIEW OF LITERATURE

3.1. Parametric Estimation

The Department of Defense US (2011), describes that parametric cost estimation is an intermediate level classified estimative modality executed when projects are complete in the rage of 10% to 35%.
The parametric estimation can be applied in finalized projects where the number of information is greater. In this manner, in modeling for databases of historical works and project.
Watson and Kwak (2004), report the origin of the parametric estimation, that takes place in the American Air Force, during the World War II, the methodology was used to estimate the costs of airplanes, taking into account the speed, and the reach by the aircraft.
The parametric estimation is the modeling that uses parameters to predict costs in construction, through data from previous projects (Sonmez, 2008; Hyun JI; Park; Soo Lee, 2010).
Mascaró (1985), the precursor of this analysis for buildings in Brazil, cites that from the point of view of geometric analysis, the building behaves like a plane, made up of sets of planes, these being horizontal in intersection with vertical sets, thus forming spaces projected.
Cerea and Premoli (2010), reinforce that the methodology is widely used in complex projects, and is effective in the preliminary identification phase, where the main cost factors are likely to be detected. Cerea and Premoli (2010), present that parametric estimation or parametric modeling is expressed through an analytical function inserted in a set of variables.
Watson and Kwak (2004) state that the characteristics of the project may allow the application of an algorithm that determines the cost approximation, where the characteristics can be: physical attributes or performance specifications.
Martins, Jungles and de Oliveira (2010) contextualize that it is imminent the need of the developers of the products to use models and tools that provide information and options to control the variables that interfere in product development, a practice common to other industry sectors that have focus as on the cost of the product in its initial phase.
Keller, Collopy and Componation (2013), point out that the estimation through parametric models presents statistical limitations, the authors analyze the application of the methodology in the American aerospace industry and criticize the production of mass generalist models that tend to capture all the characteristics of the product, emphasize the prior recognition of cause and effect relations.

3.2. Juvenile Prisons

Juvenile prisons are enterprises funded by public administration in Brazil, and each state of the federation is responsible for building, maintaining and operating them. These undertakings are intended for young people under the age of 21, and therefore have a distinct function of jails, prisons, temporary detention centers and penitentiaries, these last buildings, housing the male and female population over 21 years of age in a closed system.

3.3. Monte Carlo Simulation

The simulation aims to represent a real system, considerations can be made without the need for modifications in the system under analysis (Oliveira, 2008).

The Monte Carlo Method is convenient and has its increasing use for problems involving common simulations, as well as simulations of economic specificity (Di Bernardi, 2012). Gavira (2003), reports that the method allows the resolution of non-probabilistic problems with the use of the simulation through the stochastic process.

4. RESEARCH METHOD

The steps of the research development, according to the flowchart, began with the research question, where the Service of Superstructure shows great impact of the costs of the projects. In this way, it is sought through the parametric estimation to produce a model for forecasting the amount of resources necessary for the execution of the service under analysis. So in a feasibility study, where complete designs are not ready, it is possible to define the cost of the superstructure quickly.

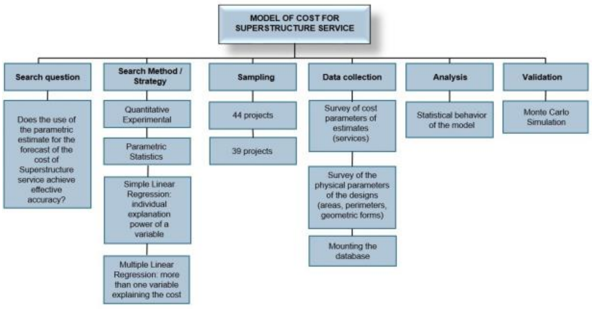

Figure 1: Flowchart of model execution and validation

The construction of the model started with the use of Simple Linear Regression with the measurement of forces between the independent variables and the dependent variable. Thus, a dependent variable such as the cost of the Superstructure was compared to an independent variable, such as Total Area, Total Perimeter etc. It was verified by means of this analysis that the simple linear regression models had no power of cost explanation for the service under study.

Then the model was constructed using multiple linear regression, in this phase, it was considered the hypothesis that more than one variable would influence the cost of the Superstructure.

The initial sampling was of 44 projects as shown in figure 1, suffering reduction for 39 projects, the elimination of 5 projects occurred due to the lack of architectural designs and standardized budgets, characterizing in this way as outliers.

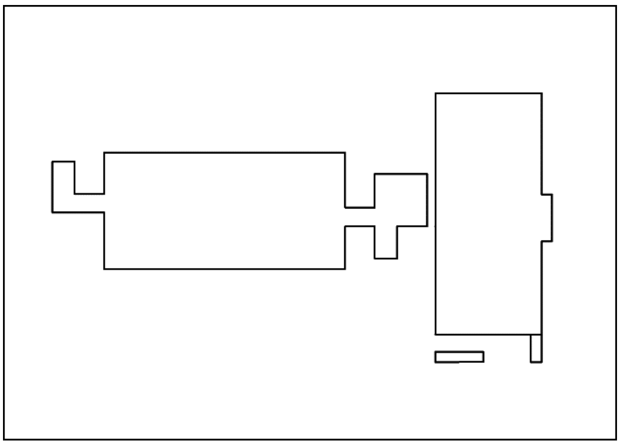

Figure 02: Example of a floor plan of a Brazilian Youth Prison - Form A

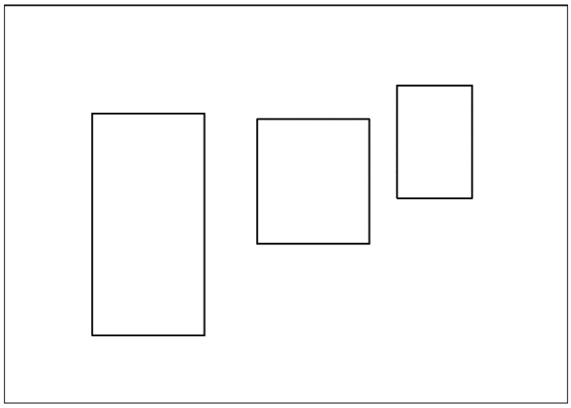

Figure 03: Example of a floor plan of a Brazilian Youth Prison - Form B

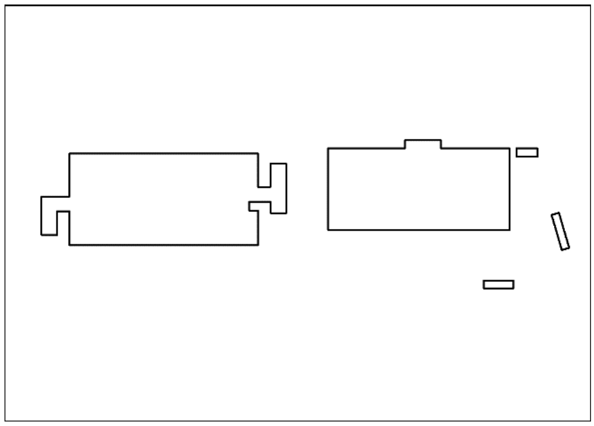

Figure 04: Example of a floor plan of a Brazilian Youth Prison - Form C

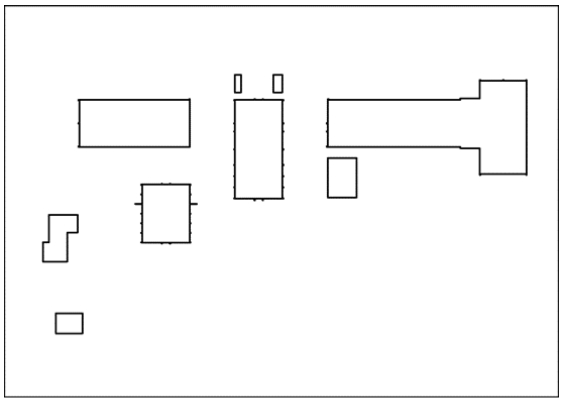

Figure 05: Example of a floor plan of a Brazilian Youth Prison - Form D

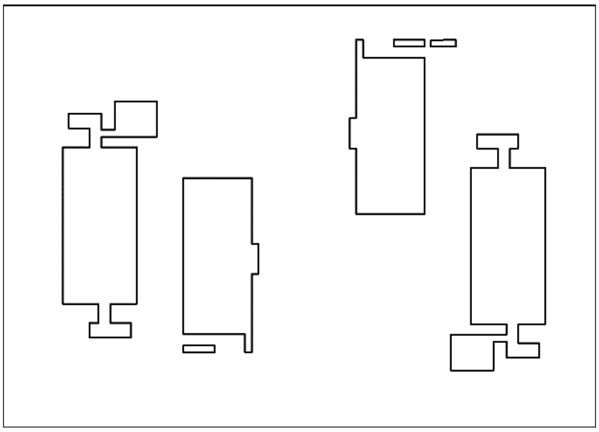

Figure 06: Example of a floor plan of a Brazilian Youth Prison - Form E

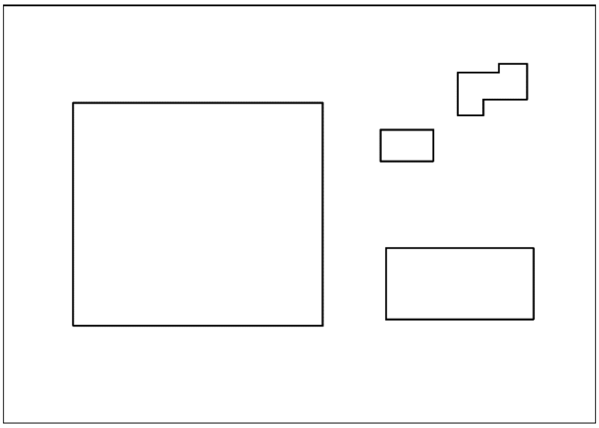

Figure 07: Example of a floor plan of a Brazilian Youth Prison - Form F

The data collection mentioned in figure 1 refers to the costs of the works removed from the budgets and the geometric parameters taken from the architectural designs.

The figures 02, 03, 04, 05, 06 and 07 are the low levels of architectural design, that is, some of the samples that make up the 39 projects used in the experiment.

The mathematical analysis of the model was done through the analysis of the Test F of Significance, performed the test and ANOVA analysis, also verified if there was behavior of multicollinearity.

The validation of the model was done through the Monte Carlo Simulation, where 1000 scenarios of architectural designs were generated, so that the model for a larger universe could be validated than the sample of the experiment.

4.1. Construction of the cost model for the Juvenile Prisons Superstructure service

The research tool is Parametric Statistics and Monte Carlo Simulation, for the feasibility of the research was used the software Microsoft Excel and Statistica, the latter being in the Ultimate Academic version. The latter software, according to Ogliari (2004), is an integrated program with capacity to manage data analysis and databases, characterizing itself as broad in the selection of the analytical process, from the basic levels to the advanced levels.

Montgomery (2009), emphasizes that the use of Multiple Linear Regression is recommended when the study presents situations that present more than one regressor.

NASA (2015), recommends as acceptable for cost estimates the coefficient R^2 on the order of 0.80 or higher. Care is needed in order to increase the R^2 value in the search for the best fit of the model, a model with higher R^2 with this addition may not perform better than the previous model (Montgomery, 2009).

The data of the 39 Youth Prisons were elaborated, tendered, contracted and executed from the year 2006 to the year 2015, and refers to a sample of one of the states of Brazil. We chose this time cut in the collection due to the standardization of budgets, descriptive memorials and other documentation of the works, from these, was taken from the geometric parameters that become the research data for the execution of the parametric estimation.

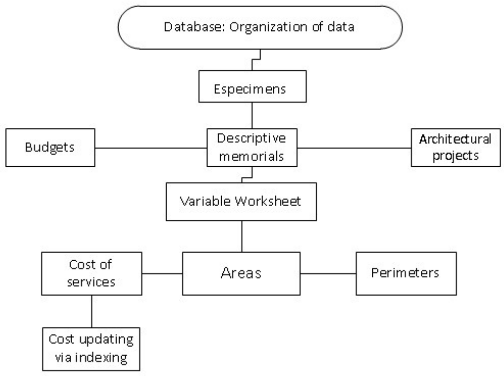

Figure 08: Assembly of the database for parametric estimation

The first step in parametric estimation is to select the samples (projects called Youth Prisons). The sample is composed of detailed budgets, descriptive memorials, architectural designs. These parameters, taken from the samples, form the variable worksheet, where the dependent and independent variables of the estimation will be chosen for later application of the Multiple Linear Regression.

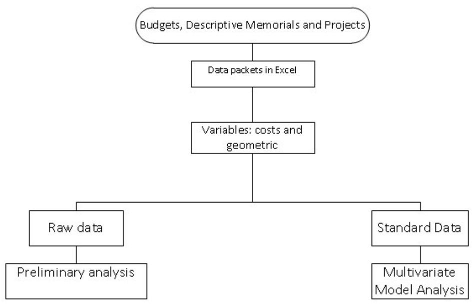

Figure 09: Standardization of data taken from specimens

The preference of the Superstructure service was due to its representativeness of 18.31% of the total amount of the costs of the projects of this nature, as shown by the ABC curve, presented in Table 01.

Description	% Partial	% Accumulated	
Security Frames	25,15	25,15	A
Superstructure	18,31	43,46	A
Roof	9,25	52,71	A
Sanitary facilities	7,52	60,23	A
Hidraulic instalations	7,27	67,50	A
Electrical and telephone installations	5,82	73,31	B
Masonry	5,52	78,83	B
Foundations	4,92	83,75	B
Painting	4,07	87,82	B
Various	3,23	91,05	B
Waterproofing, protection and gasket	2,15	93,20	C
Preliminary Services	2,13	95,34	C
Coatings	1,80	97,13	C
Earthmoving	1,50	98,64	C
Sidewalks, Guides and Gutters	0,49	99,13	C
Landscaping	0,48	99,61	C
Final Cleaning of the Work	0,34	99,94	C
Topography	0,06	100,00	C
TOTAL	**100,00**	**100,00**	

Table 1: ABC Curve of Services of Juvenile Prisons

The Suprastructure Service is composed of the following elements as indicated in figure 10: form fitting, forms, steel mounting, concrete pour, concrete cure, shoring of forms and false work.

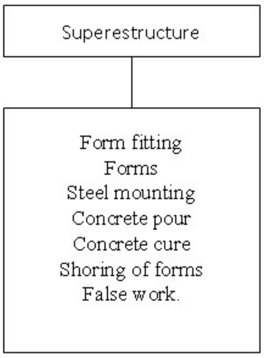

Figure 10: Superstructure Service

Starting from the analysis of the ABC Curve of Services, it is verified that the Suprastructure service is an eminent driver of costs, justifying its analysis in this research.

For the generation of the Superstructure model recommendations from NASA (2015), were observed, it shows that for cost estimates we have two types of uncertainties that may occur, in order to impair accuracy, one of them is related to problems in the method Resulting from the omission of cost variables, poor specification of coefficients and poor mathematical relationships, as well as the lack or inconsistency of historical data used.

Hamaker (1995) apud Watson and Kwak (2004), sustain that most of the estimates are linear in nature, having a single independent variable associated with a cost, citing that inflection points are rare, and reports that changes in costs can be associated with a learning curve that is calculated separately.

Based on the concept above, the model of this study would be mathematically representative through the use of Simple Linear Regression. In this way it can be combined: the total area with the costs of the service; the total perimeter versus service costs; or the regression between the density of walls and the costs of the Superstructure service.

However, individual models were generated and observed, however, they showed with low explanatory power of costs, leading to the attempt to search for more than one explanatory variable, so as to improve the representativeness of the proposed statistical model.

The correlation was generated with level of significance in the order of $p < 0.05$ measurement presented by "r" (Pearson's correlation coefficient). Correlation provided important information, which tested the parameters that had the best relationship with the dependent variable under study.

Based on the hypothesis, that more than one variable determines the cost of the Superstructure service, an explanatory equation is generated through the Multiple Linear Regression method, using more than one independent variable for the dependent variable.

For the specimens of 44 enterprises, 6 outliers were found, these were excluded from the sample, the coefficient used was 1.5, as recommended by the bibliography.

The parametric relationships can be defined: by the characteristics or properties of the product, according to Valle (2006). Or by the relation of costs with costs, as presented by the Department of Defense of the United States of America (2011). In summary, the quality of the parametric relationships depends on the validity, quality, size of the database and the purpose of the estimation.

In this research the fixed dependent variable was the costs of the Superstructure service, and the independent variables proposed were the parameters of areas and perimeters.

This research can also be performed by relating the cost parameters of other services. However, this requires a consistent database, a well-defined methodology for service definition, and preliminary knowledge of the mathematical correlation between variables (the variables should be related when both are observed on the x and y axis).

The model generated, for the sample of 39 Juvenile Prison units, had R^2 of determination presented in Figure 11 and Table 02.

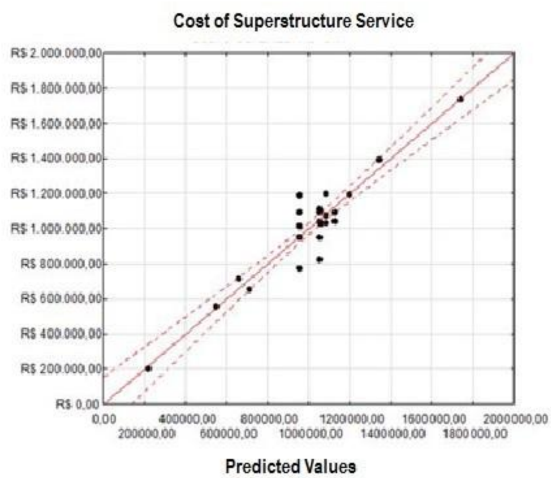

Figure 11: Result of Multiple Linear Regression for the Superstructure Service

Description	Value
R	0,86
R^2	0,74
R^2 adjusted	0,71
Standard error	110.682,00
Value p (F)	0,00

Table 2: Result of Multiple Linear Regression

The mathematical correlation relates the Cost of Service of Superstructure (CESUP) to the variables: External Perimeter (PE), Internal Perimeter (PI) and Total Area (TA) that are tested as cost drivers and which resulted in the following Equation 1.

CSUP= - 464.150,49 + 2.926,62 * PI + 94, 06 * AT + 177, 43 * PE

The model with the total cost of the Superstructure service as a dependent variable presented the adjusted coefficient of determination in the order of 0.71, that is, the model has explanatory power of 71%, when the parameters of the area and perimeters are available in the Design phase of the building.
The equation can be applied in cost estimates of Juvenile Prisons that have rectangular geometric shapes, built through the conventional constructional system: pillar system, beam and slab with concrete block walls. For the didactic effect of the application of the model that describes the cost of the Superstructure service, some parameters of the architectural design were selected, being: Total Area: 2020.00 m², External Perimeter: 150.73 m and Internal Perimeter: 434.82 m. Applying the Equation 01 produced in the research has the result:

CSUP= - 464.150,49 + 2.926,62 * PI + 94, 06 * AT + 177, 43 * PE
CSUP= - 464.150,49 + 2.926,62 * 150, 73 + 94, 06 * 2020, 00 + 177, 43 * 484, 82
CSUP= 253.001,76

That is, for a building containing these geometric parameters, the cost of the Superstructure service is obtained quickly for the estimation without the need of the structural design.

4.2. Validation of the model through Monte Carlo Simulation

In order to validate the model for projects that are outside the universe of samples, Monte Carlo Simulation was applied.
Monte Carlo Simulation is performed to include uncertainty and variability in the estimate to be performed with the cost model constructed through Multiple Linear Regression.
The Simulation is constructed from the generation of random numbers of probability and from the frequency distributions of each of the input parameters of the model: Total Area, External Perimeter and Internal Perimeter of the buildings.
Simulation of Monte Carlo for this situation and the benefits for diagnosing the effectiveness of the statistical cost model.
The number of 1000 scenarios was explained by Bruni, Fama and Siqueira (1998), where the authors clarify that the stabilization of the mean and standard deviation occurs around the 200th simulation, when performing 1000 simulations.
For Abreu and Stephan (1982, p. 152) apud Di Bernardi (2002), there is no need to repeat the simulation process more than 1,000 times, "since after a certain time the frequencies tend to stabilize and the modifications That occur only marginally affect the probability distribution."

Superstructure

CSUP= - 464.150,49 + 2.926,62 * PI + 94,06 * AT + 177,43 * PE	
Minimum Cost	R$ 80.438,67
Maximum Cost	R$ 2.606.788,61
Average Cost	R$ 1.084.317,22
Standard deviation	R$ 125.275,13

Table 03: Results of the application of the cost model for the scenarios generated through the Monte Carlo Simulation

The Superstructure service had an average cost of R $ 1,084,317.22, with a standard deviation of R$ 125,275.13. It was verified that when the model was applied in scenarios that represent Juvenile Prisons with low Total Area or high Total Area value, it presented values not feasible for the actual execution of an expedited estimate for the enterprise under analysis, when tested with extreme conditions parameters the model does not represent the cost of the Superstructure service effectively.

4.3. Analysis of the results

The cost drivers: Total Area, External Perimeter and Internal Perimeter were shown as explanatory independent variables for the modeling of the Superstructure service, that is, they are cost drivers.
The accuracy of the cost model for the Superstructure Service is compatible with the availability of information in the feasibility phase, in this initial stage of the enterprise, it is possible to estimate the costs with greater accuracy, when compared to simply dividing the overall value of the Works by the total area of the work, where the services are added having all the same weight in the estimate of costs.
In Monte Carlo Simulation, it was verified that when the model was applied in scenarios that represented juvenile prisons with low Total Area and high Total Area values (extreme conditions) it was not applicable in the execution of an expedited estimate of costs for the Superstructure service.

4.4. Final remarks

The model generated for the Superstructure service presented the adjusted coefficient of determination - R^2, in the order of 0.71, that is, the equation has the explanatory power of 71% of the costs for the service under analysis. It is considered the model of confidence, with moderate power of explanation, since the Department of Defense (2011), recommends indexes with values above 80% of representativeness. However, it is effective for the expedited estimation of Juvenile Prisons when compared to the classical methods of estimation where all services receive the same weight of costs in the project.

The error of the expedited estimate is acceptable, considering the information available in the initial phases of the projects.

The standard deviation generated by Monte Carlo Simulation for the cost model of the Superstructure Service was considered low, that is, the data generated through the random numbers showed homogeneity.

The application of the Monte Carlo Simulation to the model produced was the creation of scenarios (increasing the variability to show the uncertainty behavior of the model) for the application of the model for the Juvenile Prisons, with the generation of random numbers of the values of Total Area, Internal Perimeter, External Perimeter for a universe of different projects of the sample (considering that the data referring to descriptive statistics remained the same: mean and standard deviation, but the generated output values were different).

4.7 References

BRASIL. Sistema Nacional de Atendimento Socioeducativo - SINASE/ Secretaria Especial dos Direitos Humanos – Brasília-DF, 100p. Brasília, 2012.

BRUNI, A FAMA, R. SIQUEIRA, J. Análise do risco na avaliação de projetos de investimento: uma aplicação do método de Monte Carlo. Caderno de Pesquisas em Administração. v.1, n 6. São Paulo, 1998.

CARTA CAPITAL. Disponível em: <http://www.cartacapital.com.br/sociedade/um-em-cada-tres-unidades-da-fundacao-casa-tem-superlotacao-acima-do-permitido-pela-justica-2637.html>. Acesso em 15 set. 2015.

CEREA, A. P; PREMOLI, C. Stima parametrica del costo di costruzione. Individuazione di un metodo di stima, in fase di progettazione preliminare, del costo di costruzione tramite l'uso delle regressioni lineari. Tese (Doutorado em Engenharia Civil). Politecnico di Milano - Facoltà di Ingegneria Edile/Architettura, Milano (Itália), 2010.

DEPARTMENT OF DEFENSE. United States of America. Parametric Cost Estimating Handbook. 116p. Sine loco, Sine nomine, 2011.

DI BERNARDI, P.B. Análise de risco em investimentos imobiliários por simulação. 117p. Dissertação (Mestrado em Engenharia Civil). Universidade Federal de Santa Catarina, Florianópolis (SC), 2002.

GAVIRA. M de O. Simulação computacional como uma ferramenta de aquisição de conhecimento. 163p. Dissertação (mestrado em Engenharia de Produção). Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos (SP), 2003.

Internacional Centre for Prison Studies. Word Prision Brief. Available from: http://www.prisonstudies.org/ [accessed 15 July 2016].

JUSBRASIL. Disponível em: <http://amp-mg.jusbrasil.com.br/noticias/3050086/numero-de-infratores-em-centros-socioeducativos-supera-em-48-7-a-quantidade-de-vagas>. Acesso em: 15 set. 2015.

KELLER, S; COLLOPY, P; COMPONATION, P. What is wrong with space system cost models? A survey and assessment of cost estimating approaches. Elsevier – Acta Astronautica, [S. l.], 7p. 2013.

MARTINS, D. das N; JUNGLES, A. E; OLIVEIRA R. de. Avaliação da qualidade geométrica de projetos habitacionais e seu impacto no custo do empreendimento. Ln: Encontro Nacional de Tecnologia do Ambiente Construído, XIII ENTAC, Canela, 2010. Anais... Canela: ENTAC, 2010.

MASCARÓ, Juan Luís. O Custo das Decisões Arquitetônicas. 100p. São Paulo, 1985.

MONTGOMERY, D.D. Estatística aplicada e probabilidade para engenheiros. 493p. 4ª edição. Rio de Janeiro (RJ), 2009.

NASA. Cost Estimating Handbook. Version 4.0 Disponível em http://www1.jsc.nasa.gov/bu2/PCEHHTML/pceh.htm>. Acesso em 10 jul. 2015.

OGLIARI, P.J. Análise Estatística usando o Statistica 6.0. Departamento de Informática e Estatística.130p. Florianópolis (SC), 2004.

OLIVEIRA, Miriam. Caracterização de prédios habitacionais de Porto Alegre através de variáveis geométricas – uma proposta a partir das técnicas de estimativas preliminares de custo. 125p. (Dissertação em Engenharia Civil) Universidade Federal do Rio Grande do Sul. Porto Alegre (RGS), 1990.

PARISOTTO, J. A.; AMARAL, T. G. do; HEINECK, L. F. M. Análise de estimativas paramétricas para formular um modelo de quantificação de serviços, consumo de mão-de-obra e custos de edificações residenciais estudo de caso para uma empresa construtora. Ln: Encontro Nacional de Tecnologia do Ambiente Construído – X ENTAC, São Paulo, 2004. Anais...São Paulo: ANTAC, 2004.

SONMEZ, R. Parametric Range Estimating of Building Costs Using Regression Models and Bootstrap. Journal of Construction Engineering and Management - ASCE [S. l.], v. 134, n. 12, 6p. Dez, 2008.

VALLE, E. F. Análise de custos paramétricos de edificações não comerciais do oeste de Santa Catarina. Dissertação (Mestrado Profissionalizante em Desempenho de Sistemas Produtivos). Universidade Federal de Santa Catarina, Florianópolis, 2006.

WATSON, R.; KWAK, Y.H. Parametric Estimating in the Knowledge age: Capitalizing on Technological Advances. 2004, IAMOT Internacional Conference on Management of Tecnology. Washington, DC, Apr. 3-7, 2004.