

IMPLEMENTAÇÃO DE MÉTODOS ESTATÍSTICOS PARA PREDIÇÃO E CONTROLE DE CUSTOS EM PROCESSOS DA CONSTRUÇÃO CIVIL

Laion L. Boaventura¹, Paulo H. Ferreira¹, Rosemeire L. Fiaccone¹, and
Jorge E. R. Cajazeira²

¹Departamento de Estatística, IME, Universidade Federal da Bahia, Avenida Milton Santos s/n,
Campus de Ondina, CEP: 40.170-110, Salvador, BA, Brasil; *englimaboaventura@gmail.com*,
paulohenri@ufba.br, *fiaccone@ufba.br*.

²Departamento de Ciências Sociais Aplicadas, Universidade Estadual de Feira de Santana, Avenida
Transnordestina s/n, Novo Horizonte, CEP: 44.036-900, Feira de Santana, BA, Brasi;
jercajazeira@uefs.br

RESUMO

Segundo o Guia do Conhecimento em Gerenciamento de Projetos (PMBOK, sigla do inglês Project Management Body of Knowledge), são quatro os elementos necessários para o gerenciamento de custos de um projeto: o plano de gerenciamento de custos, a estimativa de custos, a determinação de orçamentos e o controle de custos. Com a crescente importância da elaboração de planejamento para empreendimentos, a utilização de antigas ferramentas para estimação do custo vem se tornando obsoleta. No Brasil, é comum o uso de modelos determinísticos para previsão de custos e prazos de obras e serviços. Indicadores de consumo retirados de tabelas de composição de preços, como os fornecidos pelo Sistema Nacional de Pesquisa de Custos e Índices da Construção Civil (SINAPI), vêm sendo utilizados há décadas como fonte para se efetuar orçamentos analíticos e estabelecer preços, custos e índices da construção civil. Porém, a dificuldade de adaptação sempre foi a principal crítica ao modelo usado pelo SINAPI, que ainda não dispõe de flexibilidade das definições e não considera as diferentes realidades de canteiro. Neste contexto, a utilização dos dados históricos da empresa para construção dos próprios modelos de previsão de custos pode ser a chave para otimização dos resultados esperados. Desta forma, o emprego de modelos estatísticos vêm desempenhando um papel cada vez mais importante no setor de Construção Civil. Seja para planejar ou melhorar projetos existentes, ou desenvolver controle do processo de produção, a utilização correta de modelos estatísticos diminui a probabilidade de cometer erros grosseiros na estimação de custos e prazos. Sendo assim, o objetivo deste trabalho é propor a implementação de um sistema de monitoramento e estimação de resultados dos processos dentro da Construção Civil (como, por exemplo, custos de serviços, controle de estoque, quantificação de materiais etc.), que faça uso de técnicas de Controle Estatístico de Processos (CEP) em combinação com modelos estatísticos com ênfase nos conceitos de Modelos Lineares Generalizados (MLGs) e suas ferramentas de regularização. Para isto, neste trabalho é feita uma aplicação prática, com base em um conjunto de dados reais, de processos sob controle, do Serviço de Alvenaria de Vedação com Blocos de Concreto, a fim de realizar um comparativo qualitativo e quantitativo entre a técnica tradicional, baseada em modelos determinísticos, e o sistema aqui proposto. Como resultado final, é possível destacar que, apesar de não haver grandes diferenças comportamentais ao analisar os gráficos de controle entre os modelos utilizados, houve sim, um ganho substancial na capacidade preditiva dos custos nos Modelos Estatísticos quando comparados ao Modelo Determinístico.

Palabras claves: construção civil; gerenciamento de custos; gráfico de controle; modelos

1 Introdução

A partir da década de 1950, por iniciativa dos profissionais atuantes na área de gestão, em diversos seguimentos da indústria, as práticas da administração de projetos, no nível operacional, foram compiladas e organizadas e transformaram-se em um conceito, ou seja, uma estrutura organizada de conhecimentos. Este grupo compõe o denominado Instituto de Gerenciamento de Projetos, mais conhecido como PMI (sigla do inglês *Project Management Institute*).

Assim, liderados pelo PMI, os avanços na arte da administração de projetos cresceram ao ponto que, em 1996, surgiu o Guia do Conhecimento em Gerenciamento de Projetos, ou simplesmente, PMBOK (sigla do inglês *Guide to the Project Management Body Of Knowledge*). Assim, o PMBOK é uma publicação nuclear sobre gestão de projetos, que nasceu com o propósito de padronizar e difundir as práticas mais eficientes, testadas e comprovadas por gerentes de projetos do mundo inteiro. Sua publicação é regularmente revisada e disponibilizada pelo PMI, aos seus filiados, no site do grupo: <https://pmi.org.br>. Neste trabalho, utilizou-se como referência a 6a, e última, versão do PMBOK, publicada pelo PMI em 2017.

De acordo com o PMBOK, administrar um projeto é um macroprocesso, composto por cinco grupos de processos gerenciais ligados entre si: iniciação, planejamento, execução, monitoramento e controle, e encerramento. Esses grupos gerenciais dividem-se em 49 processos menores, e estão diretamente integrados a oito áreas de concentração dos conhecimentos de um projeto (ou variáveis da administração): a integração, o escopo, a administração do tempo e do prazo, a administração dos recursos humanos, a administração das comunicações, a administração dos riscos, a administração de suprimentos, a administração da qualidade e a administração de custos.

Por sua vez, o processo pode ser entendido como a combinação dos seguintes fatores: fornecedores, produtores, pessoas, equipamentos, materiais de entrada, métodos e meio ambiente que trabalham juntos para produzir o resultado (produto), sendo que os clientes correspondem aos elementos que utilizam o resultado (ver [1]). Os processos gerenciais sugerem as melhores práticas de como

gerenciar o projeto. Já as áreas do conhecimento são o conteúdo do projeto, ou seja, o que deve ser, de fato, administrado. Dado o enfoque nas áreas de administração (gestão ou gerenciamento) da qualidade e dos custos dentro de um processo, neste trabalho aborda-se a necessidade de saber como medir e obter informações sobre o desempenho do processo ante essas duas áreas.

O PMBOK aborda qualidade sob dois aspectos: a qualidade do produto e a qualidade do processo. Essas duas vertentes do gerenciamento da qualidade são complementares, pois dada a qualidade encontrada em um produto, é muito importante considerar que o processo de produção daquele produto é uma boa fonte de partida para a produção de novos produtos semelhantes a esse. Por outro lado, se existe um processo de produção o qual possui qualidade, presume-se que o que está sendo produzido tem grandes chances de também ter qualidade. Quando a qualidade no processo é atendida, ela pode facilitar, em muito, a qualidade final do produto.

Segundo a Associação Brasileira de Normas Técnicas (ABNT) NBR ISO 9000 (ver [2]), “uma organização focada em qualidade promove uma cultura que resulta em comportamentos, atitudes, atividades e processos que agregam valor através da satisfação das necessidades e expectativas dos clientes e de outras partes interessadas pertinentes. A qualidade dos produtos e serviços de uma organização é determinada pela capacidade de satisfazer os clientes e pelo impacto pretendido e não pretendido nas partes interessadas pertinentes. A qualidade dos produtos e serviços inclui não apenas sua função e desempenho pretendidos, mas também seu valor percebido e o benefício para o cliente.”

Dentro desse conceito, o PMBOK trata ainda algumas diretrizes muito importantes para a formação, garantia e melhoria contínua da qualidade, que vêm sendo estabelecidas há décadas. Essas diretrizes estão expostas na Figura 1, extraída da 6a edição do PMBOK.

Quando o assunto é gerenciamento da qualidade, um poderoso conjunto de ferramentas utilizado na obtenção da estabilidade do processo e na melhoria de sua capacidade através da redução da variabilidade, é sempre lembrado. Esse conjunto de ferramentas é conhecido como Controle Estatístico de Processos (CEP). Para informações mais detal-



Figure 1: Diretrizes para gerenciamento da qualidade.

hadas, ver, por exemplo, Louzada et al. e Montgomery [3, 4].

Atuante também na área de administração de custos, o CEP auxilia no monitoramento do custo agregando aos processos que envolvem planejamento, estimativa, orçamento e controle de gastos, uma maior precisão para tomadas de decisão e previsão orçamentária. O que, segundo o PMBOK, é a chave para gerir os custos de qualquer projeto.

As tarefas de estimar custos e controlá-los são as que demandam maior esforço, já que as medições são complexas de serem feitas e tornam-se uma área à parte de estudos para que um bom gerenciamento de custos possa ser feito. Para a execução dos processos referentes ao gerenciamento de custos, três itens são importantes: as entradas e saídas, as ferramentas e técnicas, e o modelo de estimacão.

As entradas e saídas referem-se aos mecanismos utilizados em cada processo, os quais podem oferecer informações ou dados referentes ao projeto, oriundos de fatores ambientais da empresa (determinações já estabelecidas e que devem ser observadas para o trabalho), ou de fatores externos (como calendário dos recursos disponíveis), ou ainda, gerados a partir de outros processos de gerenciamento do projeto (como a *baseline* do escopo do projeto, plano de riscos, dentre outros).

Já as ferramentas e técnicas utilizadas podem ser um padrão (utilizadas em todos os projetos da empresa) ou ainda estarem sendo utilizadas pela

primeira vez no projeto em questão. Elas podem ser desde estimativas de três-pontos e análise de reservas, passando por custos relacionados à qualidade, até uma ferramenta de software de gerenciamento de projetos. Ou seja, é o modo operacional do projeto. Por exemplo, é estabelecer como serão processadas as contratações, como serão feitas as aquisições dos materiais e das ferramentas, entre outros aspectos operacionais.

Por fim, o modelo de estimacão de custos é o procedimento mais importante no gerenciamento dos custos. É ele o responsável direto pelo orçamento e previsão de lucro do projeto. Afinal, quando bem feito, ele permite a visão de como o projeto irá caminhar durante a sua execução. O modelo de estimacão tem por objetivo principal quantificar, a partir do escopo de todo o projeto, a previsão de qualquer custo envolvido em cada um dos processos previstos, refletindo nos prazos e no planejamento estratégico de toda a execução.

Visto isso, pode-se admitir que a previsão orçamentária de um projeto envolve diversos elementos, com inúmeros cálculos, e exige uma atenção especial para que os erros sejam minimizados ao máximo. Ainda assim, em diversos setores da indústria, processos que envolvem estimacão paramétrica de custos, muitas vezes, assumem, de maneira determinística, valores empíricos, desprezando os dados e as técnicas estatísticas existentes.

Desta forma, para a construção dos modelos de previsão de custos, a maior parte das empresas ainda utilizase de tabelas determinísticas de com-

posição de custos dos serviços. Para entender como essas tabelas funcionam, primeiramente é necessário entender o que são custos fixos e custos variáveis.

- **Custos Fixos:** é a parcela do custo que não varia de acordo com o número de unidades produzidas ou em relação a um parâmetro do projeto. Por exemplo, o valor nominal de um litro do combustível utilizado por uma máquina é o mesmo, independentemente do número de horas que ela trabalhe. Ou o valor pago por hora-aula de um professor continua sendo o mesmo, ainda que sua carga horária aumente
- **Custos Variáveis:** é a soma dos fatores variáveis de produção. Ou seja, é a parcela do custo que varia de acordo com o número de unidades produzidas ou em relação a um parâmetro do projeto. Por exemplo, horas gastas para uma máquina executar um serviço, ou o salário de um professor que recebe seus dividendos em horas trabalhadas.

Sendo assim, a Tabela de Composição de Custos é o detalhamento dos insumos destinados aos custos variáveis de um determinado serviço. Nela, estão expressas a descrição, quantidades, produtividades e custos unitários dos materiais, mão-de-obra e equipamentos necessários à execução de uma unidade de medida desse serviço.

Dentro do setor de Construção Civil, o sistema de gerenciamento e controle de custos também é baseado nessas Tabelas de Composição de Preço Unitário ou de Custos. Neste caso, as tabelas apresentam uma lista de todos os insumos necessários à execução de cada serviço, com suas respectivas quantidades, e seus custos unitários e totais. No Brasil, existem duas fontes de composições de custos mais utilizadas por órgãos públicos e privados: o SINAPI e a TCPO. Ambas trazem composições organizadas de acordo com a classificação própria e estão disponíveis tanto na forma impressa como na forma eletrônica.

Criada em 1955 pela editora PINI, a TCPO conta hoje com mais de 8.500 registros de tipos de serviços e preços de referências direcionados ao orçamento da obra [5]. Todos os serviços nela listados estão modelados a partir de equações lineares e tiveram seus índices estimados a partir de três indicadores gerais e quatro indicadores próprios,

sendo eles:

Indicadores Gerais:

- **Custo Unitário Básico (CUB)**, o qual definido pelo Sinduscon (Sindicato da Indústria da Construção Civil), e é o mais popular dos índices do setor, precifica o metro quadrado da edificação em uma determinada região. Além da localidade, outro fator que altera o CUB unitário é o tempo. Esse índice mantém uma relação direta com o custo e, de acordo com a PINI, por isso é considerado na TCPO
- **Índice Setorial da Construção Civil**, que também é calculado pelo Sinduscon de São Paulo, e entra na TCPO porque toma como referências variações do custo global, da mão-de-obra e dos materiais a cada mês
- **Fundação Instituto de Pesquisas Econômicas (FIPE)**, que também analisa o mercado da construção civil e estipula preços para as edificações.

Indicadores Próprios (gerados em suas pesquisas nas obras das regiões Sul e Sudeste do Brasil):

- **CUPE** (Custos Unitários PINI de Edificações)
- **IPCE** (Índices PINI de Custos de Edificações)
- **IPCI** (Índice PINI de Custos da Construção Industrializada)
- **Índices PINI de Custos de Obras de Infraestrutura.**

Quanto ao SINAPI, o mesmo foi implantado em 1969 pelo Banco Nacional da Habitação (BNH), e hoje é indicado como fonte oficial de referência de preços de insumos e de custos de composições de serviços pelo Decreto 7.983/2013, que estabelece regras e critérios para elaboração do orçamento de referência de obras e serviços de engenharia, contratados e executados com recursos dos orçamentos da União, e pela Lei 13.303/2016, que dispõe sobre o estatuto jurídico da empresa pública, da sociedade de economia mista e de suas subsidiárias. Assim como a TCPO, todos os serviços listados pelo SINAPI estão modelados a partir de equações lineares e tiveram seus índices estimados a partir de indicadores gerais e próprios, fornecidos pelo Instituto Brasileiro de Geografia e Estatística (IBGE), com o apoio da Fundação para o Desenvolvimento Tecnológico da

Engenharia (FDTE), instituição sem fins lucrativos formada por docentes e pesquisadores da Escola Politécnica da Universidade de São Paulo (Poli-USP).

Estudos comparativos entre as composições fornecidas pelo SINAPI e TCPO indicam uma proximidade de resultados nos coeficientes estimados. Um estudo realizado no ano de 2016, por Wilson Menezes de Melo Filho, como parte da Especialização em Produção e Gestão do Ambiente Construído, da Universidade Federal de Minas Gerais, intitulado como “Estudo Comparativo de Composições de Preço Unitário dos Sistemas SINAPI-Caixa Econômica Federal e TCPO-PINI”, verificou que o SINAPI é um pouco mais conservador, já que, de maneira geral, o orçamento com base na TCPO resultou em um valor total 3,03% menor.

Talvez essa proximidade de resultados entre os dois principais referenciais de composição de preços na construção civil no Brasil seja explicada por Marchesan [6], que ressalta que os sistemas de gestão de custos tradicionalmente empregados pelo setor da construção civil, de um modo geral, enfatizam informações voltadas às exigências legais, fiscais e acionárias, sem um foco propriamente gerencial. Ainda segundo o autor, os modelos são simplesmente ajustados preocupados em cumprir o objetivo de estimar o controle financeiro global de um produto ou serviço, e ainda assim as estimativas (ou informações) geradas são tidas como pouco relevantes aos tomadores de decisão.

Os modelos de previsão de custos dos processos dentro do setor de Construção Civil são elaborados em duas etapas:

1. Na primeira etapa, estimam-se, a partir do escopo inicial, os quantitativos de cada um dos insumos pertencentes ao processo. Por exemplo, é sabido que no processo de assentamento cerâmico existem 4 insumos relacionados ao custo do processo. São eles: o número de horas necessárias para um pedreiro assentar $1 m^2$ de cerâmica, o número de horas necessárias para um servente (ajudante) contribuir com o pedreiro no assentamento desse $1 m^2$, bem como a quantidade de pisos e de quilos de argamassa utilizados para assentar o mesmo $1 m^2$ de cerâmica. Tais estimativas são feitas através de modelos lineares simples,

em que cada um dos insumos se relaciona com a quantidade de cerâmica determinada no escopo do projeto, por meio de um coeficiente β_i , estimado deterministicamente.

Assim, dada a quantidade total, em m^2 , necessária para executar todo o projeto, consegue-se estimar as quantidades totais consumidas de cada um dos insumos e, conseqüentemente, seus custos.

2. Feito isto, a segunda etapa do processo contempla a construção do modelo de previsão de custos do processo. Este modelo relaciona linearmente os quantitativos estimados a um custo total do processo (C), e pode ser escrito da seguinte forma:

$$C = \sum_i^p \hat{x}_i c_i, \quad (1)$$

em que x_i é a quantidade estimada e c_i é o custo fixo estipulado, associados ao i -ésimo insumo do processo.

De acordo com Vrijhoef e Koskela [7], essa forma de prever custos pode ser considerada como deficiente, principalmente pelo fato de não levar em conta a natureza do produto e do processo de produção, uma vez que generaliza a estimação dos quantitativos para toda e qualquer situação ocorrente. Para Black [8], a falta de estudos e pesquisas em gestão de custos a partir de uma perspectiva teórica consistente é um dos principais fatores impeditivos para o avanço e melhorias nas práticas de gestão e controle de custos na construção civil.

1.1 Objetivo

Com o avançar dos anos, a competitividade de mercado no setor de Construção Civil exige resultados cada vez mais fiéis ao que se é planejado. Dentro dessa exigência por melhoria, existe o confronto entre dois conceitos de estimação de custos: a modelagem determinística e a modelagem estatística. Ambas travam um duelo de imposição sobre qual das abordagens se faz mais eficaz em determinados contextos. A verdade é que, seja no contexto determinístico ou estatístico, simular sistemas e processos significa modelar de tal forma que o modelo escolhido imite as respostas do sistema real numa sucessão de eventos que ocorrem ao longo do tempo.

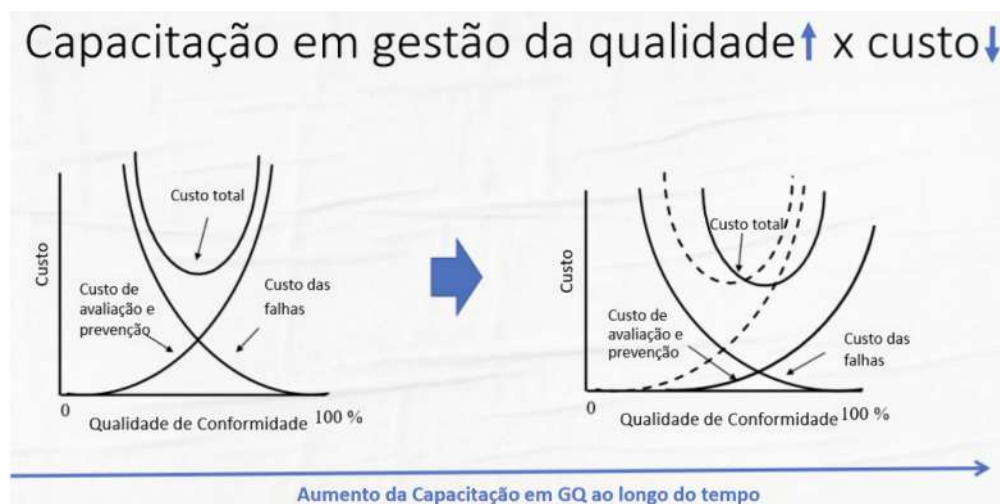


Figure 2: Qualidade × Custo.

De acordo com o PMBOK, a eficiência de um projeto está diretamente ligada ao controle do desempenho dos custos e da qualidade. Para otimizar a qualidade do processo, é necessário capacitar o poder de estimação do custo, o que é conhecido na engenharia como custo da conformidade. Segundo Carpinetti [9], o dinheiro gasto para evitar falhas durante um projeto é, na maioria das vezes, inferior ao dinheiro investido para corrigir essas falhas. Ainda segundo o autor, o aumento da capacidade de gestão da qualidade na estimação do custo otimiza a qualidade de conformidade e reduz o custo total do processo numa relação de causa e efeito, conforme mostrado na Figura 2 (Introduzido por Juran [21], o conceito de custos da qualidade pode ser expresso como a soma do custo da boa qualidade e da má qualidade. O custo da boa qualidade é o resultado da soma dos custos de avaliação e prevenção, enquanto que o custo da má qualidade consiste na soma dos custos decorrentes de falhas internas e externas.)

Posto isto, o presente trabalho tem como objetivo propor a implementação de métodos estatísticos para previsão de custos e tomada de decisão a respeito de um processo do setor de Construção Civil. Para isso, se faz necessário aprofundar-se em conceitos de modelagem estatística e CEP, explorando suas especificidades, pressupostos e aplicações.

A fim de contrapor a metodologia tradicional utilizada na estimação de custos, uma aplicação prática do modelo foi realizada, com dados reais,

de maneira a apontar aspectos positivos e negativos das abordagens determinística e estatísticas propostas. Desta forma, neste trabalho, dois tipos de avaliação comparativa são apresentados:

- (i) **Avaliação quantitativa do custo.** Realizada com base em modelos lineares como: modelos de regressão linear simples, modelos lineares generalizados estimados a partir de estimadores de máxima verossimilhança, e modelo de regressão Ridge, esta avaliação tem por objetivo comparar a capacidade preditiva do método determinístico tradicional com ferramentas de modelagem estatística, porém de forma a garantir a explicabilidade dos insumos do processo e, conseqüentemente, salvaguardar uma comparação mais direta com o modelo determinístico proposto pelo SINAPI
- (ii) **Avaliação qualitativa do processo.** A partir do proposto por Pedrini e Caten [10], no qual os autores mesclam as propostas dos Gráficos de Controle de Regressão Múltipla, de Haworth [11], com a de monitoramento em duas fases, introduzida por Woodall e Montgomery [12] e Vining [13], esta avaliação tem por objetivo verificar se existe similaridade quanto à tomada de decisão do processo, entre as saídas gráficas construídas a partir dos modelos determinístico e estatísticos.

1.2 Organização do Trabalho

O restante deste artigo está dividido em três seções, dispostas da seguinte forma. Na Seção 2

é abordado o desenvolvimento dos modelos aqui utilizados, tanto o determinístico quanto os estatísticos. As aplicações computacionais, inclusive as simulações, realizadas neste trabalho, foram elaboradas com o auxílio do software R [14]. A Seção 3 apresenta os resultados da avaliação do custo, por meio de uma tabela comparativa, e da qualidade, através dos gráficos de controle construídos com base nos conceitos abordados na Seção 2. Por fim, a Seção 4 encerra o artigo apresentando algumas considerações finais e perspectivas de trabalhos futuros.

2 Material e Métodos

2.1 Modelo de Previsão de Custo na Construção Civil

No setor de Construção Civil, os projetos têm o quantitativo de cada processo predeterminado ainda na fase de definição do escopo; para maiores detalhes, ver PMBOK, 6a Edição, 2018. Ou seja, por exemplo, no projeto de construção de uma casa, o tamanho das paredes, a quantidade de área coberta, o número de banheiros e o tamanho dos ambientes, são alguns dos processos previamente conhecidos às fases de estimativas de planejamentos de custos, prazos e qualidade.

Entretanto, os insumos relacionados a cada um desses processos, apesar de conhecidos, não têm seu quantitativo mensurado especificadamente na fase de escopo. Desta forma, por exemplo, a estimativa de custos de cada um dos processos é definida a partir do quantitativo previamente conhecido.

Por exemplo, suponha o processo de execução do telhado de uma casa. Neste processo, ainda na fase de escopo, sabe-se a quantidade total de área da casa que deve ser coberta. Contudo, apesar de já ser entendido que: quantidade de telhas, quantidade de madeira, quantidade de pregos, horas trabalhadas pelo ajudante e horas trabalhadas pelo carpinteiro, são insumos relacionados ao processo, seus respectivos quantitativos não são conhecidos e devem, portanto, ser mensurados a partir da quantidade total de área da casa a ser coberta.

Assim, no setor de Construção Civil, a partir de relações lineares simples, estimase os quantitativos dos insumos com base no quantitativo conhecido do processo. Matematicamente, essa relação pode

ser representada da seguinte forma:

$$X_i = Y\beta_i, \text{ para } i = 1, 2, \dots, p, \quad (2)$$

em que X_i representa o i -ésimo insumo do processo, e β_i é o parâmetro que explica a relação desse insumo com a quantidade Y conhecida desse processo.

Após estimar as quantidades relacionadas a cada insumo, é dado o momento de estimar o custo do processo. No setor de Construção Civil, conforme mencionado na Seção 1, para a estimação dos custos considera-se uma relação linear, desta vez múltipla, entre os custos relacionados a cada insumo, e o custo total do processo.

Desta forma, a partir de custos unitários fixados, e dos quantitativos estimados de cada insumo, pode-se prever o custo do processo. Esse modelo de previsão pode ser matematicamente representado por:

$$C_j = \hat{X}_{1j}c_1 + \hat{X}_{2j}c_2 + \dots + \hat{X}_{ij}c_i + \hat{X}_{pj}c_p, \quad (3)$$

em que \hat{X}_{ij} é a quantidade estimada do i -ésimo insumo na j -ésima observação; e c_i é o custo fixo conhecido do i -ésimo insumo.

Esse conceito de modelagem é muito comum na área de Engenharia e Economia, em que geralmente recebe o nome de Modelagem de Comportamento. Nesse tipo de modelagem, as variáveis são caracterizadas como endógenas ou exógenas, podendo, a depender do contexto, uma mesma variável ser endógena para um modelo e exógena para outro. Entende-se por variável exógena aquela que fora determinada por forças externas ao modelo, mas que influencia no comportamento das variáveis endógenas, enquanto que as variáveis endógenas são definidas dentro do modelo, podendo ser tanto a variável resposta quanto as variáveis explicativas ao processo.

No contexto da Modelagem para Previsão de Custo na Construção Civil, a variável Y pode ser vista como exógena, uma vez que seus valores são determinados por circunstâncias não descritas pelo modelo. Já a variável C pode ser considerada como uma variável endógena, dado que seus valores são observados como conjunto solução de um modelo. Em contrapartida, as variáveis X_i ($i = 1, 2, \dots, p$) assumem diferentes papéis no processo: quando utilizadas conforme (2.1), atuam como variáveis endógenas na relação direta

com a variável exógena Y ; já quando em (2.2), X_i é considerada dada naquele modelo, tendo assim características de uma variável exógena. Por isso, ainda que admitase a existência de um erro associado na geração de X_i como conjunto solução de Y , esse erro é desconsiderado quando em (2.1) x_i é considerada exógena ao modelo.

2.1.1 Alvenaria de Vedação com Blocos de Concreto

Segundo o Guia de Melhores Práticas para Alvenaria de Vedação com Blocos de Concreto, publicado pela Associação Brasileira de Cimento Portland em 2012, entendese por alvenaria “o subsistema do edifício constituído por elementos que compartimentam e definem os ambientes internos, controlando a ação de agentes indesejáveis, entre os quais: intrusos, animais, vento, poeira e ruído”. Em outras palavras, alvenaria é uma parede que tem por objetivo fechar as partes externas de um edifício e estabelecer suas divisões internas.

O Serviço de Alvenaria pode ser dividido em dois seguimentos: a Alvenaria de Vedação e a Alvenaria Estrutural. A alvenaria de vedação é projetada para resistir basicamente ao seu próprio peso. Enquanto que, a alvenaria estrutural, além de resistir ao seu próprio peso, deve suportar outras cargas e ações, como a de um telhado, por exemplo. Os dois sistemas, no entanto, podem utilizar blocos de concreto ou cerâmicos.

Considerando que cada material tem as suas propriedades específicas, escolher entre blocos de concreto ou cerâmicos pode ser uma tarefa relacionada simplesmente à disponibilidade na região do projeto ou preferência estética. Desta forma, neste trabalho, dado a disponibilidade dos dados reais, abordase o processo de alvenaria de vedação com blocos de concreto.

O Processo de Alvenaria de Vedação com Blocos de Concreto conta com seis componentes, ou insumos, relacionados à quantidade de alvenaria, geralmente expressa em metros quadrados, definida no escopo do projeto. São eles: horas trabalhadas pelo pedreiro, horas trabalhadas pelo servente (ajudante), número de blocos consumidos, quantidade de argamassa utilizada, metragem da tela de aço para amarração entre as paredes e pinos de aço para fixação dessas telas.

Dessa forma, o processo de estimação do custo de

um Processo de Alvenaria de Vedação com Blocos de Concreto é feito em duas etapas:

1. Estimar as quantidades referentes a cada um dos seis insumos;
2. A partir de valores fixos para os custos de cada um dos insumos, usar as quantidades estimadas para prever o custo total do processo.

Neste trabalho, esse custo é estimado por meio de métodos determinísticos, já consolidados no setor de Construção Civil, bem como através de técnicas estatísticas, ambos expressos nas subseções posteriores. Para os custos fixos de cada insumo, os valores ilustrados na Tabela 1 serviram de base para a construção de todos os modelos utilizados, em que $X_1 - X_6$ são descritas na Subseção 2.3.

Tabela 1: Valores unitários das variáveis.

Variável	Valor unitário com encargos
X_1	R\$ 16,67
X_2	R\$ 10,54
X_3	R\$ 1,56
X_4	R\$ 186,50
X_5	R\$ 14,18
X_6	R\$ 29,80

Fonte: Cidade de Salvador - Bahia, 2017.

2.2 Modelo Estatístico \times Modelo Determinístico

Um modelo é dito determinístico quando se tem um conjunto de entradas conhecido e deste resultará um único conjunto de saídas. Em geral, um sistema determinístico é modelado analiticamente; isto só não ocorre quando o modelo se torna muito complexo, envolvendo um grande número de variáveis ou de relações. Neste caso, recorre-se à simulação como alternativa de solução. Dentre as principais aplicações da simulação determinística, podese citar o planejamento financeiro dentro das engenharias e a simulação de sistemas macroeconômicos.

Em contrapartida, um modelo é dito estatístico quando parte-se de um conjunto de entradas conhecido, porém essas entradas conduzem a saídas aleatórias. As saídas aleatórias, por sua vez, são consideradas estimativas das verdadeiras características do sistema, e como qualquer estimativa, carrega um componente aleatório denomi-

nado “erro”. Este erro aleatório é relativo ao distanciamento entre os valores reais observados e os valores estimados pelo modelo. Assim, denomina-se como Análise de Regressão, a técnica estatística utilizada para investigar e modelar a relação entre essas variáveis. Por exemplo, em uma modelagem estatística, a equação da reta ou do plano (se $p > 1$) recebe um vetor de componentes aleatórios dos erros em sua representação:

$$\begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{pmatrix} = \begin{pmatrix} 1 & W_{11} & W_{12} & \dots & W_{1p} \\ 1 & W_{21} & W_{22} & \dots & W_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & W_{n1} & W_{n2} & \dots & W_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_p \end{pmatrix}, \quad (4)$$

ou, na forma matricial $\mathbf{Z} = \mathbf{W}\beta + \epsilon$

Esses erros aleatórios são estimados a partir dos chamados resíduos. Os resíduos são calculados pela diferença analítica entre o vetor de variáveis respostas \mathbf{Z} e o vetor de respostas estimadas $\hat{\mathbf{Z}}$. Os resíduos são componentes aleatórios que, quanto mais próximos a zero, indicam boas estimativas para os parâmetros β do modelo. Para informações mais detalhadas, ver [15].

Desta forma, entende-se que, ao contrário da situação determinística, a modelagem estatística está mais próxima da realidade. Uma vez que se utiliza de experimentos amostrais das variáveis aleatórias, a fim de reduzir os erros estimados, reproduzindo, conseqüentemente, a realidade da forma mais fiel possível.

2.3 Modelo Determinístico

Conforme abordado na Seção 1, no setor de Construção Civil, a estimação dos quantitativos dos processos ocorre, tradicionalmente, de maneira determinística. E os parâmetros utilizados para relacionar os insumos à quantidade prevista sobre o processo estão dispostos através de tabelas de composição de custos. Dentre as tabelas apresentadas na Seção 1, neste trabalho optase por utilizar o SINAPI. Decisão esta embasada no fato de que, além de ser fonte oficial de referência de obras públicas, conforme estabelecido pelo Decreto 7.983/2013 e pela Lei 13.303/2016, existem trabalhos que comprovam a informação de que o SINAPI é, quando comparado às demais referências determinísticas concorrentes, a tabela de índices mais conservadores para estimação dos quantitativos dos processos. Neste

caso, utilizar coeficientes determinísticos mais conservadores implica em uma menor generalização do modelo construído, quando comparado aos modelos resultantes das demais tabelas.

Sendo assim, para a construção do modelo determinístico estudado neste trabalho, utilizou-se o Caderno Técnico de Composições para Alvenaria de Vedação - Lote 1 do SINAPI, cuja vigência é a partir de junho de 2014, com última revisão de atualização em julho de 2017. Este caderno foi escolhido devido à semelhante faixa temporal de referência do conjunto de dados reais utilizado.

Para informações mais detalhadas sobre o Serviço de Alvenaria, no caderno técnico de serviços do SINAPI, o Serviço de Alvenaria de Blocos de Concreto está codificado pelo número 87449 e descrito como “ALVENARIA DE VEDAÇÃO DE BLOCOS VAZADOS DE CONCRETO DE 14 19 39 CM (ESPESSURA 14 CM) DE PAREDES COM ÁREA LÍQUIDA MENOR QUE 6 M2 SEM VÃOS E ARGAMASSA DE ASSENTAMENTO COM PREPARO EM BETONEIRA”.

Com covariáveis e seus respectivos coeficientes dispostos de acordo com a Figura 3, extraída do próprio caderno técnico.

A partir da análise sobre a Figura 3, pôde-se construir, com respeito à Equação (2.1), seis modelos lineares simples para estimação dos quantitativos de cada um dos insumos do Serviço de Alvenaria. Esses modelos podem ser representados da seguinte forma:

- Para estimação da quantidade de horas trabalhadas pelo pedreiro: $X_{1j} = 0,99Y_j$
- Para estimação da quantidade de horas trabalhadas pelo servente: $X_{2j} = 0,495Y_j$
- Para estimação da quantidade de blocos consumidos no processo: $X_{3j} = 13,35Y_j$
- Para estimação do volume, em m^3 , de argamassa consumida: $X_{4j} = 0,0103Y_j$
- Para estimação do comprimento, em m , de tela de aço utilizada: $X_{5j} = 0,785Y_j$
- Para estimação do cento de pinos de aço utilizados: $X_{6j} = 0,0189Y_j$

em que Y_j é a quantidade de alvenaria, em m^2 , estabelecida no j -ésimo projeto.

Neste trabalho, a partir de uma amostra real

Classe: PARE				
Tipo: 0065 - Alvenaria de Blocos de Concreto				
COMPOSIÇÃO				
Item	Código	Descrição	Unidade	Coefficiente
C	88309	PEDREIRO COM ENCARGOS COMPLEMENTARES	H	0,9900
C	88316	SERVENTE COM ENCARGOS COMPLEMENTARES	H	0,4950
I	651	BLOCO VEDAÇÃO CONCRETO 14 X 19 X 39 CM (CLASSE D - NBR 6136)	UN	13,3500
C	87292	ARGAMASSA TRAÇO 1:2:8 (CIMENTO, CAL E AREIA MÉDIA) PARA EMBOÇO/MASSA ÚNICA/ASSENTAMENTO DE ALVENARIA DE VEDAÇÃO, PREPARO MECÂNICO COM BETONEIRA 400 L. AF 06/2014	M3	0,0103
I	34547	TELA DE AÇO SOLDADA GALVANIZADA/ZINCADA PARA ALVENARIA, FIO D = *1,20 A 1,70* MM, MALHA 15 X 15 MM, (C X L) *50 X 12* CM	M	0,7850
I	37395	PINO DE AÇO COM FURO, HASTE = 27 MM (AÇO DIRETA)	CENTO	0,0189



Figure 3: Qualidade × Custo.

com 1.000 observações da quantidade de alvenaria produzida, onde cada uma dessas observações representa um projeto diferente e independente do outro, dado os modelos listados acima, um banco de dados composto por $j = 1.000$ vetores $(\hat{X}_{1j}, \hat{X}_{2j}, \hat{X}_{3j}, \hat{X}_{4j}, \hat{X}_{5j}, \hat{X}_{6j})$ estimados deterministicamente de Y_j , foi simulado com o objetivo de fornecer as quantidades necessárias para estimação dos custos dos processos. Este banco de dados pode ser representado matricialmente da seguinte maneira:

$$\begin{bmatrix} Y_1 & \hat{X}_{1,1} & \hat{X}_{2,1} & \hat{X}_{3,1} & \hat{X}_{4,1} & \hat{X}_{5,1} & \hat{X}_{6,1} \\ Y_2 & \hat{X}_{1,2} & \hat{X}_{2,2} & \hat{X}_{3,2} & \hat{X}_{4,2} & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \\ Y_{1.000} & \hat{X}_{1,1.000} & \hat{X}_{2,1.000} & \hat{X}_{3,1.000} & \hat{X}_{4,1.000} & \hat{X}_{5,1.000} & \hat{X}_{6,1.000} \end{bmatrix} \quad (5)$$

2.3.1 Simulação do Modelo

Desta forma, a partir do banco de dados representado acima, o interesse foi gerar um vetor simulado dos j -ésimos custos dos processos, a fim de, posteriormente neste trabalho, avaliar o modelo determinístico: (i) qualitativamente, através da técnica de CEP apresentada na Subseção 2.5; e (ii) quantitativamente, por meio da análise sobre a precisão do modelo quanto à estimação do custo global dos projetos. Assim, com base na Tabela 1 e na Equação (2.2), calculouse os custos estimados \hat{C}_j , para $j = 1, 2, \dots, 1.000$, através da seguinte expressão:

$$\hat{C}_j = 16,67\hat{X}_{1j} + 10,54\hat{X}_{2j} + 1.56\hat{X}_{3j} + 186,5\hat{X}_{4j} + 14,18\hat{X}_{5j} + 29,8\hat{X}_{6j}, \quad (6)$$

Após análise descritiva sobre a distribuição dos custos estimados (os resultados não são mostrados aqui por limitação de espaço), atrelada às informações a respeito da coleta das observações amostrais dos dados reais, pôde-se concluir que o pres- suposto estocástico (normalidade), necessário para aplicação do método de CEP proposto por Pedrini e Caten [10] (ver Subseção 2.5), foi atendido.

2.4 Modelos Estatísticos: Regressões Lineares

A partir do conjunto de dados reais do Serviço de Alvenaria de Vedação com Blocos de Concreto (descrito em detalhes na Subseção 2.6), foram desenvolvidos modelos estatísticos para a estimação das quantidades de cada insumo relacionado ao processo, bem como para a sua previsão de custo.

Como descrito ao longo deste artigo, no setor de Construção Civil, os quantitativos dos insumos que compõem um processo são estimados a partir de regressões lineares simples. É a partir dessas estimativas que, conforme apresentado na Equação (2.2), os custos dos processos e, conseqüentemente, do projeto, são calculados. Desta forma, neste trabalho foram construídos modelos estatísticos para as estimativas dos insumos do Serviço de Alvenaria de Vedação com Blocos de Concreto a partir de regressões lineares simples. E, com essas estimativas, utilizouse os conceitos de Modelos Lineares Generalizados (MLGs; Nelder e Wedderburn [16]), e posteriormente, de regressão Ridge (Hoerl e Kennard [17]), para construir o modelo de estimação do custo

2.5 Gráficos de Controle de Regressão

Neste trabalho, é utilizado o método proposto por Pedrini e Caten [10], o qual mescla o gráfico de controle de regressão múltipla, proposto por Haworth [11], e a proposta de monitoramento em duas fases, apresentada por Woodall e Montgomery [12] e Vining [13], permitindo, desta forma, o monitoramento direto da variável resposta do processo ao invés do monitoramento dos resíduos padronizados do modelo.

Posto isto, na aplicação da Fase I (também chamada de análise retrospectiva, em que os limites gráficos são estimados com base em amostras

provenientes do processo a ser monitorado), após a validação do modelo de regressão, são encontrados os j limites das linhas de controle para cada j -ésima observação da variável resposta, através das seguintes equações:

$$LIC_j = \hat{Y}_j - 3\sqrt{MQR}, LC_j = \hat{Y}_j, LSC_j = \hat{Y}_j + 3\sqrt{MQR} \quad (7)$$

em que LIC denota o limite inferior de controle, LC a linha central, LSC o limite superior de controle, \hat{Y} é o j -ésimo valor estimado da variável resposta através do modelo de regressão utilizado, e MQR é o quadrado médio dos resíduos, dado pela seguinte equação:

$$MQR = \frac{\sum_{j=1}^n (Y_j - \hat{Y}_j)^2}{n - p} \quad (8)$$

sendo n igual ao número de observações e p igual ao número de parâmetros estimados no processo. Assim, em um processo controlado, espera-se verificar que nenhum dos Y 's coletados extrapole a região formada pelos \hat{Y} 's estimados a partir do modelo de regressão utilizado na Fase I, evidenciando, assim, que o modelo utilizado para a construção dos limites de controle é, de fato, viável para a tomada de decisão acerca do processo. Perspectiva esta ilustrada a partir da Figura 4.

Contudo, a Fase I só finaliza quando verifica-se, através do Gráfico de Controle de Extrapolação, se os valores das variáveis de controle indicam estar sob a ação de causas especiais. O Gráfico de Controle de Extrapolação foi desenvolvido pelos próprios autores e consiste em monitorar os pontos de alavanca (*leverage values*):

$$h_{jj} = \mathbf{X}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}_j \quad (9)$$

com base em um $LSC = \max\{h_{jj}\}$. Segundo Pedrini e Caten, caso ocorram muitos pontos de extrapolação da região original dos dados, ou seja, muitos $h_{jj} > \max\{h_{jj}\}$, recomenda-se retornar ao início da Fase I e alterar as definições acerca do modelo. Caso contrário, desenvolve-se o gráfico de controle de regressão da Fase II (também chamada de monitoramento do processo, que é quando as características estimadas do processo são monitoradas dentro dos limites estimados na Fase I).

Na Fase II, antes de monitorar novas amostras, são corrigidos os limites de controle da Fase I, adotando o desvio-padrão para a previsão de uma

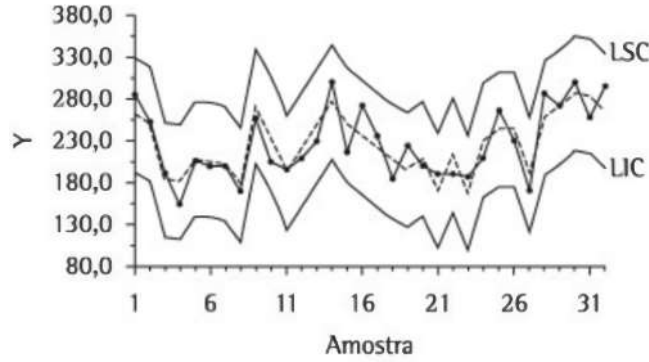


Figure 4: Representação da proposta de Pedrini e Caten (2011).

nova observação, conforme descrito por Montgomery et al. [18]. Assim, os limites de controle e a linha central do gráfico de controle de regressão são definidos como segue:

$$\begin{aligned} LIC_j &= \hat{Y}_j - 3\sqrt{\text{MQR}(1 + h_{ij})} \\ LC_j &= \hat{Y}_j, \\ LSC_j &= \hat{Y}_j + 3\sqrt{\text{MQR}(1 + h_{jj})} \end{aligned} \quad (10)$$

Além disso, Pedrini e Caten indicam que, ao monitorar novas amostras oriundas de processos de status desconhecidos, deve-se primeiro monitorar as variáveis h_{jj} das novas observações da Fase II, tendo o valor máximo dos h_{jj} da Fase I como LSC. Isto porque, segundo Pedrini e Caten (2011), a escolha desse limite de controle deriva do fato de que o valor máximo dos h_{jj} da Fase I delimita a região composta esperada para todos os vetores X_j , com $j = 1, 2, \dots, n$.

2.6 Descrição e Exploração dos Dados

O conjunto de dados utilizado neste trabalho foi fornecido por uma empresa do setor de Construção Civil atuante na Região Metropolitana da cidade de Salvador - Bahia, Brasil. Segundo informações da empresa fornecedora, os dados foram coletados de maneira independente, em que cada uma das 1.000 observações representa uma obra (ou projeto) realizada pela empresa no período entre janeiro de 2016 e agosto de 2017. Cada uma das 1.000 observações coletadas traz informações a respeito da quantidade de cada um dos insumos relacionados ao Serviço de Alvenaria de Vedação, descritos na Subseção 2.1.1, bem como as quan-

tidades, em m^2 , do serviço realizado. Denotando cada uma dessas informações como:

- Y_j para a quantidade de alvenaria executada, em m^2 , na j -ésima observação, sendo esta a variável resposta
- X_1 para a quantidade de horas trabalhadas pelo pedreiro para a realização dos serviços
- X_2 para a quantidade de horas trabalhadas pelo servente para a realização dos serviços
- X_3 para a quantidade de blocos utilizados para a realização dos serviços
- X_4 para a quantidade de argamassa, em m^3 , destinada para a realização dos serviços
- X_5 para a quantidade de tela de aço, em m , utilizada para a realização dos serviços
- X_6 para a quantidade de pinos de aço utilizados para a realização dos serviços

calculouse algumas medidas descritivas a fim de entender melhor a disposição das observações dentro de cada variável. Essas medidas estão ilustradas na Tabela 2. Ao analisar as informações da Tabela 2, pode-se concluir que a análise da relação linear entre as variáveis não poderia ser feita por meio da covariância, afinal, as diferentes unidades de medida poderiam resultar num peso maior para determinada variável. Sendo assim, essa análise foi feita por meio da matriz de correlação entre as variáveis.

Para a construção do gráfico da matriz de correlação, utilizou-se a função `corrplot.mixed()` do pacote `corrplot` (Wei e Simko [19]) do software R. Com essa função, pôde-se verificar que existe uma alta correlação entre as variáveis, o que pode

Tabela 2: Análise descritiva do banco de dados.

Variável	Mínimo	Media	Mediana	Máximo	Desvio-Padrão
Y	12,00	179,30	146,00	393,00	85,79
X ₁	14,40	211,10	173,10	459,80	99,80
X ₂	7,04	105,84	86,14	231,87	50,26
X ₃	165,00	2.565,00	2.054,00	5.562,00	1.464,00
X ₄	0,11	1,63	1,33	3,58	0,77
X ₅	4,50	67,84	55,44	149,30	32,50
X ₆	0,40	5,35	4,35	11,79	2,55

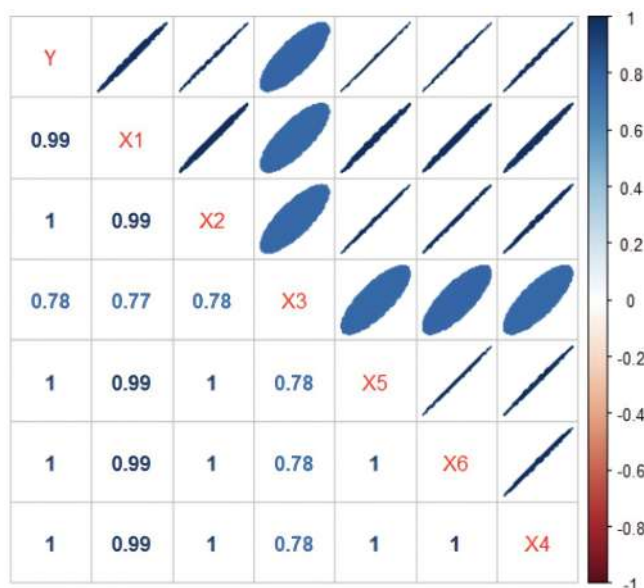


Figure 5: Matriz de correlação.

acarretar em problemas no ajuste do modelo de regressão para a previsão de custo, assim como na veracidade dos coeficientes encontrados. A Figura 5 traz a saída gráfica do pacote, na qual as elipses em azul são resultados de uma correlação linear positiva entre as variáveis, e conforme mais afilada a elipse, maior o nível de correlação linear entre as variáveis.

Ainda a partir da análise da matriz de correlação construída, é possível destacar também que, aparentemente, é possível pensar em relações lineares simples para explicar a relação entre Y e cada um dos insumos X_i ($i = 1, 2, 3, 4, 5, 6$). Entretanto, dado o interesse na construção dos modelos para a estimação desses insumos, faz-se necessário analisar a presença de possíveis *outliers* na disposição das observações. Assim, por meio da função `boxplot(.)` do R, o gráfico `boxplot` de cada um dos insumos contidos no banco de dados foi

gerado, apresentando os seguintes resultados: (i) todos os insumos apresentaram assimetrias positivas; (ii) apresentaram também possíveis pontos tidos como *outliers*. Já em relação a Y, além do `boxplot`, a fim de compreender melhor a distribuição dos seus dados amostrais, o histograma foi traçado sob a densidade de uma distribuição normal com parâmetros equivalentes à média e à variância amostrais de Y. Assim como os insumos, Y apresentou assimetria positiva e alguns possíveis pontos de *outliers*. Os valores excessivos de dados entre 100 e 150 m^2 , aparentemente distanciam a distribuição de Y de uma distribuição normal. Todos esses resultados acerca de Y, assim como os apresentados sobre os insumos, estão ilustrados nas Figuras 6 e 7.

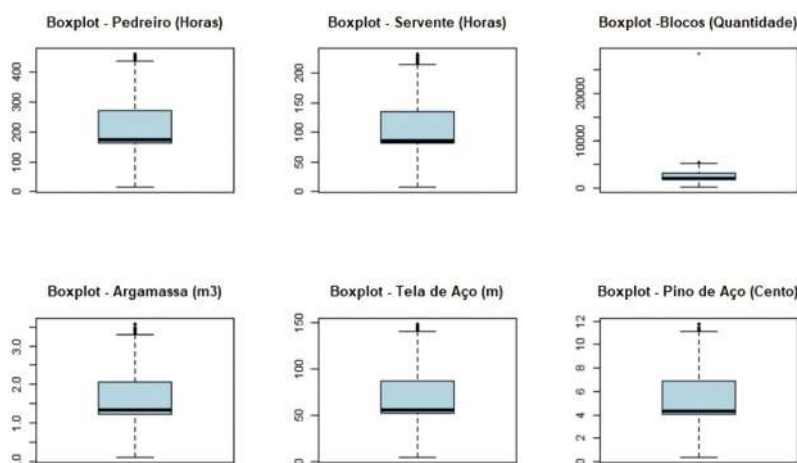


Figure 6: Matriz de correlação.

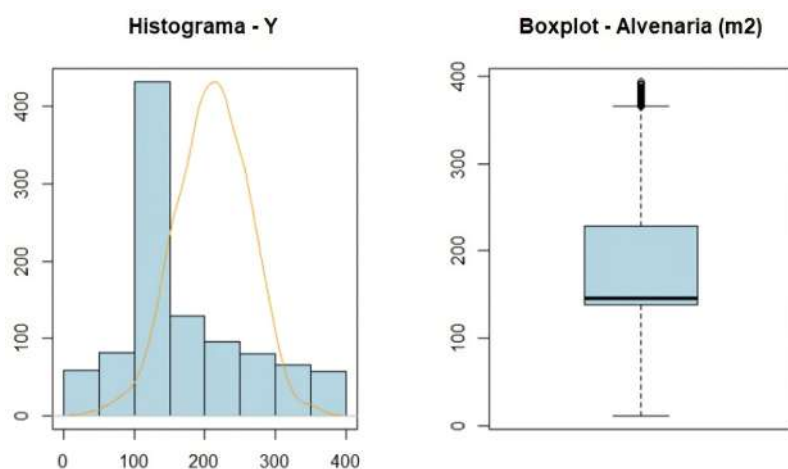


Figure 7: Matriz de correlação.

3 Resultados

3.1 Modelagem Estatística

Após processo de seleção (os resultados em detalhes não são apresentados aqui por restrição de espaço), o modelo escolhido foi o MLG de família Gaussiana com função de ligação canônica e intercepto. Entretanto, ao verificar os pressupostos para a aplicação do modelo selecionado, foi evidenciada alta correlação linear entre as variáveis explicativas, o que, consequentemente,

resultou em singularidade aproximada da matriz de custos dos insumos, e, dado a isso, em problemas de multicolineariedade. Com o problema de multicolineariedade evidenciado, apesar dos estimadores de máxima verossimilhança (EMVs) ap-

resentarem propriedades muito importantes, como a de serem, assintoticamente, estimadores não viesados e de terem variância mínima, dentre os estimadores não viesados (Bolfarine e Sandoval [19]), foi visto que seu uso pode não ser o mais eficaz para modelagem de regressão. Sendo assim, foi proposta como alternativa uma técnica de regularização que introduz um pequeno viés nos EMVs dos MLGs, com o intuito de diminuir suas variâncias e minimizar o efeito da colinearidade. Essa técnica foi a regressão Ridge (Hoerl e Kennard [17]). Com o modelo de regularização utilizado, os resultados de validação obtidos foram satisfatórios. Assim, nas próximas subseções, os três modelos (o modelo determinístico, o MLG com EMVs, e o modelo Ridge) são analisados quanto à sua capacidade preditiva para o custo

Tabela 3: Análise gráfica sobre Y

Modelo	V (R\$)	\hat{V} (R\$)	Δ (R\$)	P(%)	status
Determinístico	8.907.253	10.071.881	1.164.628	86,90	Superestimado
MLG (EMVs)	8.907.253	8.003.361	-903.892	89,85	Subestimado
Ridge	8.907.253	8.696.789	-210.464	97,64	Satisfatório

do processo (Subseção 3.2), e quanto à sua competência em monitorar qualitativamente o processo (Subseção 3.3).

3.2 Análise de Controle do Custo

Para a avaliação quantitativa, verificouse qual dos modelos estudados apresentou custo estimado total mais próximo do custo real gasto com o serviço de alvenaria. Para isso, primeiramente, a partir do conjunto de dados apresentado na Subseção 2.6, incluindo as observações tidas como outliers, e utilizando-se dos preços unitários apresentados anteriormente (ver Tabela 2), encontrou-se um valor global gasto igual a R\$ 8.907.253,00.

Esse valor serviu de referência para a averiguação da capacidade preditiva de cada um dos modelos construídos. Assim, denotando esse valor global real por V , o valor global estimado pelos modelos como \hat{V} , a diferença absoluta entre o valor global real e o valor global estimado por Δ , e a precisão de predição, em porcentagem, de cada modelo como P , podese observar, como ilustrado na Tabela 3, adotando como status satisfatório os valores de precisão maiores que 95%, que o rendimento quantitativo da estimação feita pela regressão Ridge, é bastante superior quando comparado aos demais modelos verificados.

3.3 Análise de Controle da Qualidade

Para a avaliação qualitativa dos modelos, a proposta foi verificar se existia diferença na tomada de decisão sobre a conformidade do processo de execução do serviço de alvenaria entre os três modelos.

Essa tomada de decisão seria observada mediante análise dos gráficos de controle implementados com base em uma adaptação da proposta de Pedrini e Caten, descrita na Subseção 2.5. Dado o objetivo didático em construir uma comparação qualitativa entre os modelos determinístico, MLG

e Ridge, a fim de analisar a validade deles para o monitoramento do processo, aliado ao fato de termos somente as 1.000 amostras já utilizadas para a construção dos modelos, implementamos neste trabalho somente a Fase I da proposta de Pedrini e Caten.

Inicialmente, visto que os dados são oriundos de um processo sabidamente controlado, ainda que os valores da variável a ser monitorada não estejam padronizados, em um modelo válido para monitoramento, o esperado é que a linha que representa os valores reais observados, acompanhem aleatoriamente, dentro dos limites de controle, o desenvolvimento do gráfico. Ou seja, o desejável é que o custo real do processo monitorado apresente comportamento próximo ao comportamento médio dos custos estimados do processo, em cada uma das observações.

Assim sendo, começamos a implantação da Fase I com o monitoramento do custo real de cada uma das j -ésimas observações, para $j = 1, 2, 3, \dots, 1.000$, do conjunto de dados real apresentado (excluindo as observações tidas como *outliers*).

Para isso, utilizando os modelos construídos neste trabalho, através das equações:

$$LI_j = \hat{C}_j - 3\sqrt{MQR} \text{ e } LS_j = \hat{C}_j + 3\sqrt{MQR} \quad (11)$$

calculou-se os limites de controle de cada observação para cada um dos modelos, em que \hat{C}_j é o custo estimado na j -ésima observação e MQR é a média dos quadrados dos resíduos. Estes limites são válidos nos três modelos estudados. Isto porque o MLG selecionado é o modelo Gaussiano, o que indica resíduos com distribuição normal; o modelo Ridge apresentou resíduos padronizados bastante acomodados aos quantis de uma distribuição normal padrão (ver Figura 10 do Apêndice); e o modelo determinístico, ainda que não assuma a ocorrência de erros aleatórios, é implicitamente construído sob uma regressão linear múltipla, o que evidencia pressuposição de normalidade dos resíduos.

Posto isto, ao analisar a Figura 8, pode-se observar que o MLG, apesar de apresentar, em média, maior proximidade de resultados na maior parte do gráfico, a partir da observação 800 apresenta comportamento completamente diferente do real. Este fato pode ser atribuído à multicolinearidade encontrada na construção do modelo. Assim, entende-se que o problema encontrado no pressuposto do modelo afeta a tomada de decisão acerca do processo. Desta forma, o modelo para estimação dos custos construído a partir do MLG não seria ideal para uma tomada de decisão sobre a conformidade do processo de alvenaria.

De maneira análoga, é possível verificar que o gráfico construído a partir do modelo Ridge acompanha de maneira muito próxima o custo médio esperado do processo em cada uma das j observações. Sendo assim, pode-se entender o modelo Ridge como apto a monitorar o processo e, assim, ser implementado na Fase II. Situação esta similar à observada no gráfico construído a partir do modelo determinístico. Apesar de não apresentar, em média, valores tão próximos aos esperados, quando comparados aos encontrados no modelo Ridge, todas as observações monitoradas encontram-se dispostas dentro da região de conformidade, acompanhando o comportamento esperado dos custos do processo. Apresentando, assim, os pré-requisitos necessários para ser graficamente implementado na Fase II.

Contudo, segundo Pedrini e Caten, os resultados obtidos até então não são suficientes para validar a utilização dos modelos na Fase II do CEP. Isto porque é preciso analisar a ocorrência de pontos de extrapolação ao valor máximo dos h_{jj} calculados a partir dos dados utilizados para a construção dos limites de controle.

Assim, tomando os valores estimados de X_{ij} , com respeito à Subseção 2.3, para o modelo determinístico, e à Tabela 4 (ver Apêndice), para o modelo Ridge, calculou-se o LSC do Gráfico de Extrapolação, para que, após isso, fossem monitorados todos os X_{ij} estimados em ambos os casos, incluindo as observações apontadas como *outliers*.

O LSC do Gráfico de Extrapolação do modelo determinístico foi aproximadamente igual a 0,0042. Resultado este que indicou a não conformidade do processo em 112 das 1.000 observações. Já o LSC do Gráfico de Extrapolação do modelo Ridge apresentou valor aproximadamente igual a 0,007, o

que levou à tomada de decisão que 23 das 1.000 observações estavam fora de controle. Tais resultados estão ilustrados na Figura 9.

Com o auxílio da função *identify(.)* do R, observou-se que, 17 dos 23 pontos destacados como amostras não conformes no Gráfico de Extrapolação do modelo Ridge, foram considerados pontos de outliers no momento de construção dos modelos de regressão linear simples para a estimação dos X_{ij} utilizados. Desta forma, entende-se que apenas 6 das amostras tidas como não conformes podem ser consideradas como erros em tomadas de decisão sobre o processo. Uma vez que as demais apontadas já haviam sido evidenciadas como desconformes no processo de estimação.

Em contrapartida, o grande número de observações h_{jj} extrapoladas do intervalo de investigação das variáveis de controle utilizadas na construção dos limites de controle do modelo determinístico, evidencia a necessidade de ajuste no intervalo de variação das variáveis de controle, o que, neste caso, significa a inviabilidade do modelo determinístico em monitorar o processo.

Em suma, ao analisar os gráficos ilustrados, percebe-se que existe diferença na tomada de decisão do gráfico gerado a partir do modelo Ridge e do gráfico implementado com base no modelo determinístico. Enquanto que o gráfico criado a partir do modelo determinístico indica aparente desconformidade do processo, o gráfico oriundo das técnicas de modelagem estatística apresentadas, de uma forma geral, evidencia a conformidade do processo, ainda que indique possíveis amostras não conformes.

4 Conclusões

No desenvolvimento deste trabalho, com o objetivo de propor a implementação de métodos estatísticos para a tomada de decisão sobre um processo do setor de Construção Civil, realizou-se uma aplicação prática, acerca do Custo do Serviço de Alvenaria com Blocos de Concreto. Para isso, além da análise criteriosa a respeito de um banco de dados real, sabidamente controlado, deste serviço, foi necessário aprofundar-se em conhecimentos sobre variáveis endógenas e exógenas, construção de MLGs e seus modelos de regularização, bem como sobre gráficos de CEP e modelagem determinística.

Assim, além do estudo acerca dos MLGs convencionais, foi abordada também a regressão Ridge, aplicada com o objetivo de corrigir problemas de multicolinearidade já esperados no Processo de Alvenaria, e detectados no banco de dados real. Com relação aos procedimentos de CEP, o método utilizado foi o gráfico de controle proposto por Pedrini e Caten (2011). Este método mescla a proposta dos Gráficos de Controle de Regressão Múltipla, de Haworth (1996), com a proposta de monitoramento em duas fases, introduzida por Woodall e Montgomery (1999) e Vining (2009). Contudo, dado o caráter didático deste trabalho, aliado ao fato de possuímos apenas as amostras utilizadas na construção dos modelos, a proposta de Pedrini e Caten (2011) foi implementada apenas em sua Fase I. Também foram realizados estudos de simulação com o intuito de gerar um banco de dados para a produção de alvenaria a partir do modelo determinístico estudado.

Por fim, os resultados acerca dos gráficos de controle mostraram que o modelo determinístico, apesar de acomodar bem os custos preditos, mediante análise do Gráfico de Extrapolação, trouxe a falsa impressão de desequilíbrio do processo. Isto porque, ao monitorar os pesos atribuídos aos valores reais de consumo de cada uma das covariáveis pertencentes ao modelo, baseando-se nos limites de controle construídos a partir dos valores estimados pelo modelo determinístico, incluindo possíveis situações atípicas (outliers), notou-se que muitas observações estariam fora de controle. Este parecer traz evidências de que, de maneira global, o modelo determinístico pode, de fato, ser utilizado para a estimação do custo do processo. Entretanto, quando se busca um controle sob os custos direcionados a cada um dos insumos relacionados ao processo, este não se mostrou eficaz, uma vez que acusou não conformidade em um processo originalmente controlado. Em contrapartida, o gráfico construído com base nos modelos estatísticos, constatou conformidade do processo. Pode-se dizer, assim, que o método estatístico utilizado para a definição do modelo de controle do processo, foi eficaz tanto na previsão do custo do serviço de alvenaria, quanto na estimação da quantidade destinada a cada insumo relacionado ao processo.

De maneira análoga, quantitativamente, os resultados apresentados pelos modelos estatísticos, até mesmo o MLG (ainda que não tenha atendido

os pressupostos de independência linear entre as covariáveis), foram bastante satisfatórios quando comparados com o modelo determinístico. Por exemplo, ao analisar a Tabela 3, percebe-se uma diferença de precisão dos custos de aproximadamente 10,5% entre o modelo de melhor rendimento (regressão Ridge) e o modelo determinístico.

A conclusão que fica deste trabalho é que a implementação do controle do Serviço de Alvenaria em Blocos de Concreto foi mais eficaz quando feita através de técnicas estatísticas. E que é possível sim utilizar modelagem estatística para a previsão de custos, sem menosprezar a ideia da explicabilidade dos insumos do processo.

Fica também, como reflexão a partir dos resultados, que, em uma economia competitiva na qual o preço de venda dos serviços é determinado pelo mercado, onde as empresas precisam trabalhar bem a sua gestão, os seus processos e o seu gerenciamento de custos para maximizar a lucratividade, utilizar técnicas estatísticas significa sobrevivência.

Sendo assim, entende-se que a situação ideal para controle de custos dentro de processos da construção civil, não é desprezar os dados e continuar balisando-se por metodologias determinísticas. Mas sim, implementar, dentro de cada organização, uma estrutura de análise de decisão a partir dos dados. Para que, assim, a utilização, por exemplo, de CEP e modelos estatísticos, venham a ajudar na prevenção contra eventos indesejados, evitando prejuízos nos processos.

Em suma, espera-se, então, que os resultados obtidos neste trabalho, a partir de um conjunto de dados reais, possam, de fato, auxiliar as empresas do setor de Construção Civil a quebrar o paradigma do uso de técnicas estatísticas para controle de processos, tanto em âmbito qualitativo, quanto em âmbito quantitativo.

Como trabalho futuro, pretende-se realizar pesquisas mais aprofundadas na literatura a fim de propor gráficos de controle de regressão mais generalistas, que minimizem a necessidade de informação sobre a relação entre as variáveis explicativas e a variável resposta. Deseja-se, ainda, expandir os estudos de simulação para dados massivos, de maior dimensão, e também estabelecer a comparação entre os gráficos já existentes, a fim de verificar o ganho qualitativo e quantitativo em

sua utilização.

Por fim, intenta-se aplicar a metodologia proposta a outros conjuntos de dados dentro da área de

Construção Civil e demais áreas, como indústria, saúde, etc.

Apêndice

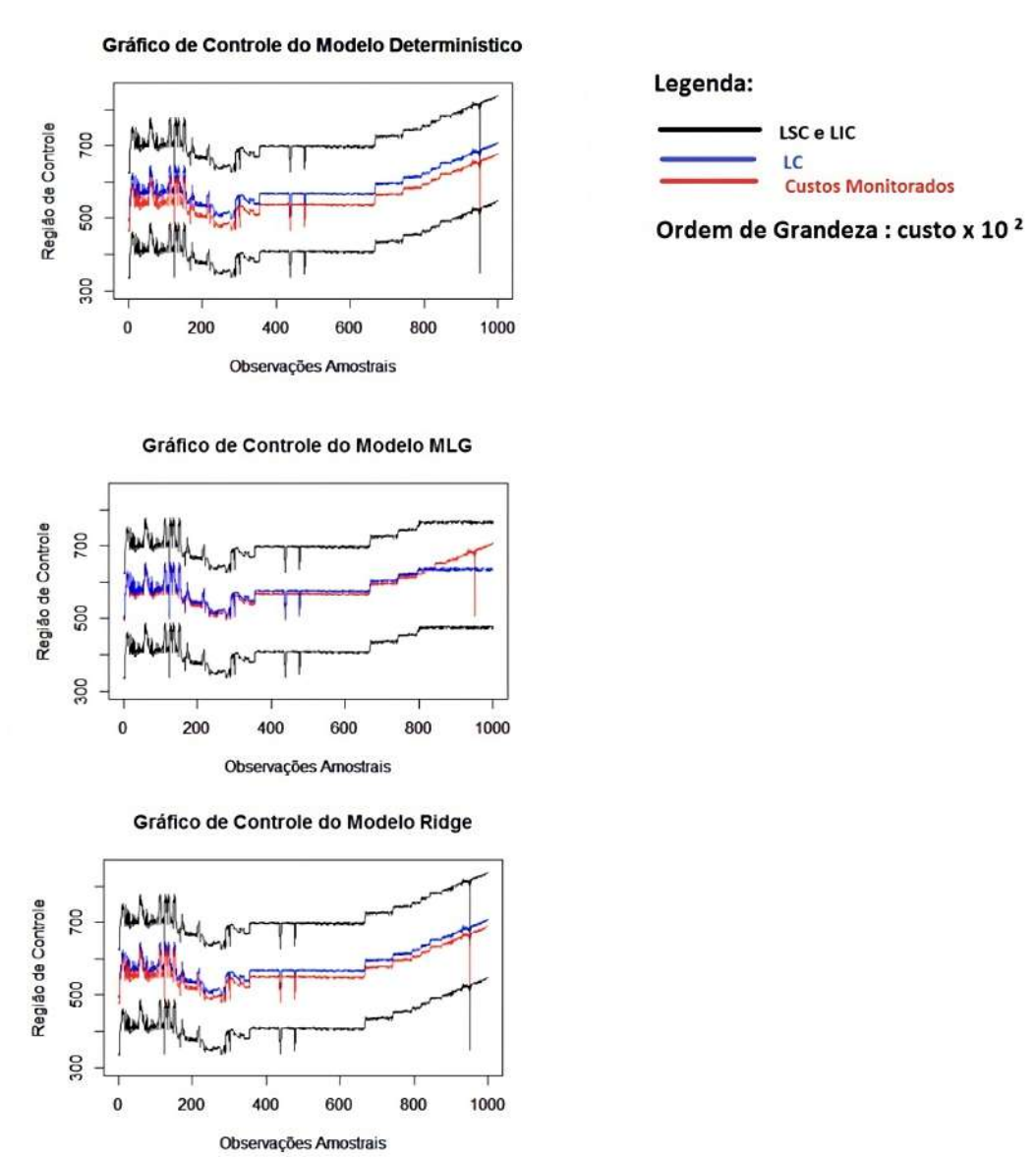


Figure 8: Fase I - Etapa de construção dos gráficos de controle.

Tabela 4: Modelo final de estimação para cada um dos insumos.

Modelo	$\hat{\beta}_0$	$\hat{\beta}_1$
$X_1 \sim Y$	1,79	1,16
$X_2 \sim Y$	0,83	0,59
$X_3 \sim Y$	*	14,16
$X_4 \sim Y$	0,0136	0,00897
$X_5 \sim Y$	*	0,378
$X_6 \sim Y$	*	0,029

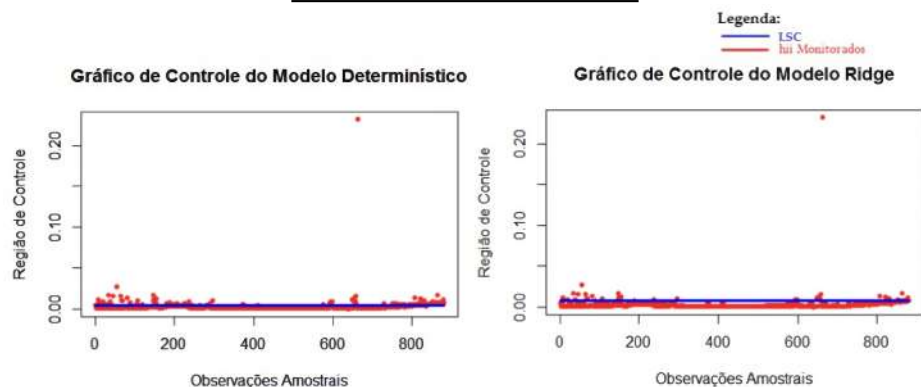


Figure 9: Fase I - Etapa de construção do Gráfico de Extrapolação.

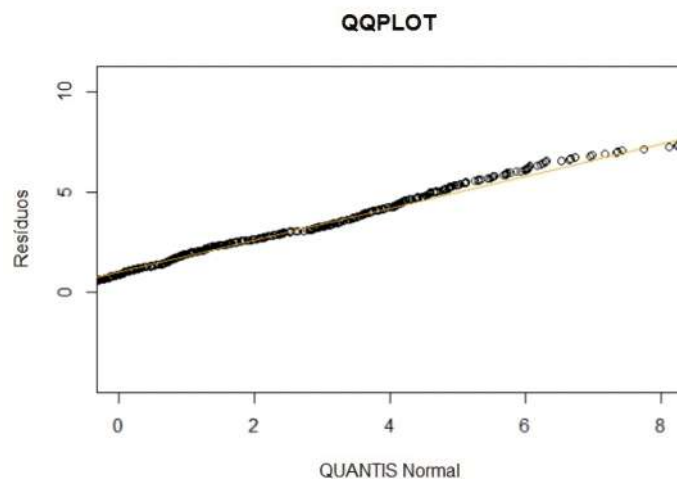


Figure 10: Gráfico quantil-quantil (ou Q-Q plot) dos resíduos estudentizados do modelo de regressão Ridge.

Referencias

- [1] Hautrive, T.P. (2021). *Ciência e Tecnologia de Alimentos*. Editora Insular.
- [2] Associação Brasileira de Normas Técnicas NBR ISO 9000 (2015). *Sistemas de Gestão da Qualidade - Fundamentos e Vocabulário*.
- [3] Louzada, F., Diniz, C., & Ferreira, P. (2000).
- [4] Montgomery, D. C. (2009). *Introduction to Statistical Quality Control*, Sixth Edition. John Willey and Sons. New Jersey. Hal, 303-304.

Controle estatístico de processos: uma abordagem prática para cursos de engenharia e administração. Grupo Gen-LTC.

- [5] PINI (2022). TCPO web. Disponível em: <http://tcpoweb.pini.com.br/home/home.aspx>. Acesso em: 22 out. 2022.
- [6] Marchesan, P.R.C. (2001). *Modelo Integrado de Gestão de Custos e Controle da Produção para Obras Cíveis*. Dissertação de Mestrado (mestrado em Engenharia Civil) – Programa de Pós-Graduação em Engenharia Civil. Universidade Federal do Rio Grande do Sul, Porto Alegre.
- [7] Vrijhoef, R., & Koskela, L. (2000). The four roles of supply chain management in construction. *European journal of purchasing & supply management*, 6(3-4), 169-178.
- [8] Black, J.H. (1984). Cost engineering planning techniques for management. M. Dekker.
- [9] Carpinetti, L.C.R. (2016). *Gestão da Qualidade: Conceitos e Técnicas*, 3ed., São Paulo, Atlas.
- [10] Pedrini, D. C., & Caten, C. S. T. (2011). Method for applying regression control charts to process monitoring. *Production*, 21, 106-117.
- [11] Haworth, D.A. (1996). Regression control charts to manage software maintenance. *Journal of Software Maintenance: Research and Practice*, 8(1), 35-48.
- [12] Woodall, W. H., & Montgomery, D. C. (1999). Research issues and ideas in statistical process control. *Journal of Quality Technology*, 31(4), 376-386.
- [13] Vining, G. (2009). Technical Advice: Phase I and phase II control charts. *Quality Engineering*, 21(4), 478-479.
- [14] R Core Team (2019). R: A Language and Environment for Statistical Computing (Version 3.6.1). R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- [15] Paula, G. A. (2004). *Modelos de regressão: com apoio computacional*. São Paulo: IME-USP.
- [16] Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 135(3), 370-384.
- [17] Hoerl, A.E.; Kennard, R.W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.
- [18] Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley & Sons.
- [19] Wei, T.; Simko, V. (2017). R package “corrplot”: Visualization of a Correlation Matrix (Version 0.84). Available from <https://github.com/taiyun/corrplot>.
- [20] Bolfarine, H., & Sandoval, M. C. (2001). Introdução à inferência estatística (Vol. 2). SBM.
- [21] Juran, J. M., Seder, L. A., & Gryna, F. M. (1951). *Quality control handbook (Vol. 1)*. McGraw-Hill.