MÉTODOS QUANTITATIVOS

Pesquisa Operacional

2

VOLUME

Editora Poisson

Editora Poisson

Métodos Quantitativos – Pesquisa Operacional Volume 2

1ª Edição

Belo Horizonte
Poisson
2018

Editor Chefe: Dr. Darly Fernando Andrade

Conselho Editorial

Dr. Antônio Artur de Souza – Universidade Federal de Minas Gerais
Dra. Cacilda Nacur Lorentz – Universidade do Estado de Minas Gerais
Dr. José Eduardo Ferreira Lopes – Universidade Federal de Uberlândia
Dr. Otaviano Francisco Neves – Pontifícia Universidade Católica de Minas Gerais
Dr. Luiz Cláudio de Lima – Universidade FUMEC
Dr. Nelson Ferreira Filho – Faculdades Kennedy

Dados Internacionais de Catalogação na Publicação (CIP)

M593

**Métodos Quantitativos- Pesquisa Operacional
Volume 2/ Organização: Editora Poisson -
Belo Horizonte - MG : Poisson, 2018
256p**

Formato: PDF

ISBN: 978-85-7042-007-7

DOI: 10.5935/978-85-7042-007-7.2018B001

Modo de acesso: World Wide Web

Inclui bibliografia

**1. Métodos Quantitativos 2. Pesquisa
Operacional.3. Estatística I. Título**

CDD-658

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos seus respectivos autores.

www.poisson.com.br

contato@poisson.com.br

Sumário

Capítulo 1: Simulação, Teoria das Restrições (TOC) e Gestão da Capacidade: Casos de Aplicação na área de Saúde	7
Francisco Santos Sabbadini, Antonio Augusto Gonçalves , Mário Jorge Ferreira de Oliveira , Antonio Henriquez de Araujo Jr. , Rosinei Ribeiro Batista	
Capítulo 2: Mineração de dados com detecção de outliers em tarefas de predição de séries temporais.....	20
Lidio Mauro Lima de Campos · Gustavo Augusto Lima de Campos	
Capítulo 3: O AHP como um modelo matemático: uma análise de sensibilidade simples	31
Mylene Cristina Rezende Pacheco, Fernando Luiz Goldman	
Capítulo 4: A teoria de filas como ferramenta de apoio na análise de um serviço de atendimento	45
Byanca Porto de Lima, Artur da Silva Siqueira de Novais, Caroline Cristina de Carvalho, Larissa Melo Negrão	
Capítulo 5: Confiabilidade dos Questionários: Uma análise da Consistência Interna da Escala F baseada no Coeficiente Alfa de Cronbach.....	54
Carolina Riul de Freitas, Carla Regina de Moraes, Francisco Severo de Lima Junior, Maria Cristina Vendrameto, Rafael Lustosa da Costa	
Capítulo 6: Uso do método Analytic Hierarchy Process (AHP) para escolha de fornecedor de farinha de trigo: um estudo de caso..	68
Pâmela Braga Bertahone, Nilson Brandalise	
Capítulo 7: Otimização via Simulação Monte Carlo na Melhoria de Processos Produtivos Aeronáuticos	80
Igor Makiyama de Souza Barbosa, Murilo Vieira Strongoli, Aneirson Francisco da Silva, Fernando Augusto Silva Marins, Erica Ximenes Dias	

Sumário

Capítulo 8: Integração de método AMD com técnicas GIS para classificar ZEIS: Um caso de estudo na capital do Ceará da região Nordeste do Brasil.....	94
Luiz Otavio Ribeiro Afonso Ferreira, Luís Miguel Alçada Tomás de Almeida, Carlos Jorge Lima de Freitas, Mario Fracalossi Jr.	
Capítulo 9: Jogos Cooperativos na alocação de custos de estoques de peças compartilhados.....	104
Natália Nogueira Ferreira Souza, Bernardo Santos Aflalo, Takashi Yoneyama	
Capítulo 10: Matheurística baseada em PSO para o problema de custo de disponibilidade de recursos com múltiplos modos.....	115
Lettiery D'Lamare Portela Procópio, Lucídio dos Anjos Formiga Cabral, Marco Cesar Goldbarg	
Capítulo 11: Recomendação de grupos de Facebook a partir de textos e imagens de usuários	127
Gustavo Ramos Lima, Patrick Marques Ciarelli	
Capítulo 12: Modelo multiobjetivo de análise envoltória de dados combinado com o desenvolvimento de funções empíricas e otimização via Simulação Monte Carlo.....	136
Aneirson Francisco da Silva, Fernando Augusto Silva Marins, Erica Ximenes Dias, Marcelo Vilela Figueiredo, Rafael de Carvalho Miranda	
Capítulo 13: Horário universitário personalizado por meta-heurística: um estudo de caso.....	153
Rosalvo Ferreira de Oliveira Neto, Caique César Leite Mariano, Max Santana Relemberg Farias	
Capítulo 14: Detecção de minas terrestres via análise multivariada de imagem.....	164
Leandro Vitral Andraos, Paulo Henrique Coelho Maranhão	

Sumário

Capítulo 15: Estudo da propagação dos erros em processos produtivos multivariados em estágios	176
--	-----

Josiane da Silva Jesus, Paulo Henrique Coelho Maranhão

Capítulo 16: Algoritmo de busca local baseado em indicador aplicado ao problema de Roteamento de Veículos Multiobjetivo com Coleta Opcional	188
---	-----

Ramon Rocha Leite, Luciana Pereira de Assis , Alessandro Vivas Andrade , Marcelo Ferreira Rego, Cristiano Grijó Pitangui

Capítulo 17: Mensuração da eficiência em serviços especializados terceirizados a partir das dimensões de qualidade via análise envoltória de dados: uma aplicação numa organização do setor automotivo da região Sul-fluminense	200
--	-----

Daniel Maia de Souza, Pauli Adriano de Almada Garcia, Pftias Teodoro

Capítulo 18: Análise numérica em um problema de programação linear com a utilização do método dos multiplicadores de Lagrange	210
--	-----

Rafaela Boeira Cechin, Leandro Luís Corso

Capítulo 19: Aplicação do controle estatístico nos processos decolagens de voos regulares dos aeroportos brasileiros	217
---	-----

Luiz Antonio Figueiredo, Jose Elmo Menezes, Ricardo Luiz Machado

Capítulo 20: Modelagem matemática no sequenciamento de projetos com restrição de recursos: um estudo de caso.....	234
--	-----

Clarisse da Silva Vieira, Natanael Salgado Coutinho, Gustavo José de Aguiar Gomes, Costa, Vinícius de Souza Bandeira, Vinícius Rossi Oliveira

Autores:	244
-----------------------	-----

Capítulo 1

SIMULAÇÃO, TEORIA DAS RESTRIÇÕES (TOC) E GESTÃO DA CAPACIDADE: CASOS DE APLICAÇÃO NA ÁREA DE SAÚDE

Francisco Santos Sabbadini

Antonio Augusto Gonçalves

Mário Jorge Ferreira de Oliveira

Antonio Henriquez de Araujo Jr.

Rosinei Ribeiro Batista

Resumo: A Teoria das Restrições encontra extensa documentação na manufatura industrial. Verificam-se, entretanto, poucos trabalhos relacionados à sua utilização na área de saúde. Em relação a simulação de eventos discretos, sua aplicação na área de saúde está bem documentada. Entretanto a utilização integrada de ambas está pouco referenciada na literatura. Este artigo apresenta um conjunto de casos aplicados, resultado de 10 anos de estudo que consolidam e mostram a aplicação de forma integrada da simulação e da TOC na área de saúde, a saber: em uma unidade de pronto atendimento, num centro de tratamento de câncer do Instituto Nacional do Câncer (INCA) e numa unidade de emergência hospitalar no Hospital Municipal de Emergência Henrique Sérgio Gregori. Para atender aos objetivos pretendidos são referenciados estudos e apresentados aspectos conceituais da Teoria das Restrições e da simulação na área de saúde. Neste artigo é descrita a utilização integrada da TOC e da DES para identificar e gerenciar gargalos em processos de atendimento de cada uma das instituições e são apresentados os resultados obtidos.

Palavras Chave: Simulação - Teoria das Restrições - Gestão da capacidade - Serviços de Saúde - Pesquisa Operacional

*Artigo originalmente publicado nos anais do Simpósio de Excelência em Gestão e Tecnologia 2013

1. INTRODUÇÃO

Existem poucos trabalhos sobre a utilização integrada da Teoria das restrições e da simulação, orientadas para gestão da capacidade de atendimento na área de saúde. Trabalhos nesse sentido foram desenvolvidos por Gonçalves (2004), Sabbadini (2005) e Sabbadini et. Al. (2008). Rébula et. al. (2009) utilizaram a teoria das restrições integradas a outras técnicas de produção como mapeamento de processos e gestão da capacidade. Este artigo apresenta um conjunto de casos aplicados, resultado de alguns anos de estudo que consolidam e mostram a aplicação de forma integrada da simulação e da TOC na área de saúde, a saber: em uma unidade de pronto atendimento, num centro de tratamento de câncer do Instituto Nacional do Câncer (INCA) e numa unidade de emergência hospitalar no Hospital Municipal de Emergência Henrique Sérgio Gregori. Assim sendo, está estruturado além desta introdução em três partes: i) fundamentação teórica sobre teoria das restrições; ii) simulação; iii) gestão da capacidade; iv) estudos de caso; e por fim a v) conclusão.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. TEORIA DAS RESTRIÇÕES

A teoria das restrições foi originalmente desenvolvida por Eliyahu Goldratt, um físico israelense que publicou os fundamentos no seu livro A Meta. Goldratt introduziu o termo teoria das restrições ("Theory of Constraints" – TOC) em 1987 englobando vários conceitos como o sistema de gerenciamento de desempenho, os cinco passos para melhoria contínua, refinamentos de conceitos de como gerenciar a produção e outras áreas de uma empresa.

Existem duas premissas em sua abordagem. A primeira é considerar a organização como um sistema cujo sucesso ou fracasso depende da forma como diferentes processos interagem entre si. A segunda é a de que uma restrição é qualquer coisa que limite o sistema de atingir sua meta. Sendo assim, todo sistema deve ter pelo menos uma restrição ou fator limitante. Uma restrição não é boa nem má. Ela existe e deve ser gerenciada (GOLDRATT, 1990; 1994). De fato, o reconhecimento da existência de um fator limitante é uma excelente oportunidade para a

melhoria, pois permite o foco na identificação e gerenciamento das restrições.

Apesar da teoria das restrições ter nascido no setor industrial, sua filosofia pode ser aplicada em diferentes tipos de organização, já existindo aplicações da TOC na área de saúde. Segundo Burton T. (2001), consultora do instituto Goldratt, as metas de uma organização de saúde de oferecer tratamento com qualidade e a de gerar lucro são complementares e essenciais. Na Inglaterra esta teoria foi aplicada com sucesso na redução das longas listas de espera que são administradas pelo sistema nacional de saúde conforme descrito por Phipps (1999). Na área do tratamento de câncer, Kershaw (2000) apresenta um estudo interessante realizado no setor de quimioterapia de uma clínica oncológica. Um dos princípios fundamentais da teoria das restrições baseia-se no processo de focalização buscando a melhoria contínua. Os cinco passos para se aplicar a teoria das restrições na área de saúde podem ser descritos da seguinte forma:

Identificar as restrições do sistema

Restrições físicas devem ser imediatamente identificadas como, por exemplo, o número de salas de exames, equipamentos, médicos, enfermeiras e técnicos. Em muitas instituições de saúde, a restrição está no número de auxiliares administrativos. Quando existe a necessidade de redução de custos, grande parte das organizações decidem por reduzir o pessoal administrativo preservando o corpo clínico. Uma ação desta natureza, sem uma análise detalhada do processo atual pode resultar em um crescimento de atividades administrativas, sendo desempenhadas por médicos e enfermeiras e a diminuição do tempo destinado ao atendimento do paciente. A velocidade do fluxo de pacientes atendidos cai, o faturamento cai e o custo operacional por paciente aumenta. Restrições políticas como decisões e estratégias ineficientes, apesar de muitas vezes mais desastrosas do que problemas operacionais são mais difíceis de identificar e gerenciar.

Decidir como explorar as restrições do sistema

Se a restrição é física, o objetivo é gerenciar os fatores limitantes tornando o processo o mais eficiente possível. O elo mais fraco corresponde à força e resistência de toda a corrente. A restrição dita a velocidade do fluxo de atendimento dos pacientes. O tempo

perdido com um fator limitante é um tempo perdido em todo o processo.

Sendo assim, uma vez identificada a restrição ela deve ser imediatamente explorada e gerenciada. Isto pode ser realizado de diferentes maneiras, dependendo do tipo de restrição. Se a limitação está no número de equipamentos de diagnóstico a decisão está em como programar sua utilização para que o maior número de pacientes possa utilizá-lo sem que haja o comprometimento da qualidade do exame e do diagnóstico.

Todos os esforços devem ser direcionados para aumentar a taxa de utilização dos equipamentos considerados recursos gargalos. Ações de curto prazo, que aumentem a eficiência do agendamento de pacientes e reduzam o tempo do exame, devem ser exploradas e executadas.

Subordinar/Sincronizar todos os recursos restantes às decisões tomadas anteriormente.

Subordinar e sincronizar todo o restante para as decisões tomadas nos dois primeiros passos é a etapa mais difícil e geralmente a mais importante. Sem esta subordinação, os planos de explorar a restrição podem não ocorrer, sendo colocados de lado em função dos problemas e incêndios do dia-a-dia.

Elevar a capacidade das restrições do sistema

Em contraste com o passo 2 em que são realizadas ações para aumentar o fluxo de atendimento aos pacientes sem que hajam gastos significativos, este passo requer um investimento nos recursos considerados gargalo. Por exemplo, pode ser necessário o aumento da equipe médica ou o crescimento da capacidade instalada com a compra de um novo equipamento.

Se a restrição se deslocar, retornar ao passo 1 e não permitir que a inércia se instale.

Não se pode permitir que a inércia se instale e se transforme numa restrição do sistema. O ambiente está em constante mudança. Uma solução tende a se deteriorar ao longo do tempo. A teoria das restrições é um processo de aprimoramento contínuo.

Quando se analisa o fluxo de tratamento de pacientes em um hospital ou clínica, pode-se observar que o mesmo é composto de uma sequência de atividades tais como, registro, consultas ambulatoriais, exames de diagnóstico e condutas terapêuticas, que configuram uma sequência linear de eventos

ou uma cadeia de processos interdependentes. Cada elo dessa cadeia possui a habilidade de executar suas respectivas atividades em diferentes taxas médias de atendimento. É importante ressaltar que esta cadeia é tão forte quanto o seu elo mais fraco que corresponde ao processo com a menor taxa média de pacientes atendidos.

2.2. SIMULAÇÃO

A crescente elevação da complexidade da gestão dos serviços de saúde, associada à redução da oferta de recursos nesta área revelou-se um campo potencial para a aplicação das técnicas de Pesquisa Operacional (PO). Diversos trabalhos foram desenvolvidos neste campo por De Oliveira e Filho (2003), De Oliveira e Toscano (2003), Chagas (2003), Gonçalves (2004), Sabbadini (2005) e Magalhães (2006), que apresentam ferramentas e iniciativas que proporcionaram respostas significativas em termos de resultados. O uso da simulação está amplamente documentado em Brailsford (2009).

Neste contexto seja por falta de informação, de recursos ou de organização dos processos administrativos e operacionais, ficam perceptíveis as dificuldades e os desafios com os quais se defrontam os administradores hospitalares. Nesse sentido a melhoria na eficiência destes serviços é de interesse particular para a maioria das comunidades, em qualquer lugar do mundo, e segundo De Oliveira e Toscano (2003) um componente crítico do sistema é a resposta resultante do gerenciamento dos serviços em termos de atendimento da demanda. Em função disso os gestores tem procurado por instrumentos confiáveis para compreensão, análise do sistema de serviço e implementação de mudanças apropriadas. A simulação computacional é uma técnica que preenche essa lacuna.

Para Banks e Carson (1984), a simulação é a imitação da operação de processos reais ou sistemas num dado período de tempo. Essa técnica envolve a geração de uma história artificial de um sistema, e a observação desta para fazer inferências relativas às características de operação do sistema real. De acordo com Gonçalves (2004) o custo relativamente baixo, pouco risco em relação à experimentação diretamente no sistema real e a capacidade de suportar de maneira confiável as flutuações estatísticas são

algumas das características que recomendam a utilização desta técnica.

2.3 GESTÃO DA CAPACIDADE

Capacidade é o nível máximo que um processo pode atingir ao operar, sob condições normais, em determinado período de tempo (SLACK, 2002). O intuito de gerenciar essa capacidade é atender a demanda de maneira eficiente.

Outra questão que deve ser considerada, tanto em operações de manufatura quanto nas de serviços, é que algumas partes trabalham abaixo da sua capacidade e outras em seu limite máximo. A esse respeito Slack (2002, p.345), registra que "(...) dependendo da natureza da demanda, o uso de diferentes partes de uma operação pode atingir sua capacidade máxima e atuar como uma restrição para toda a operação".

Além das questões relativas aos gargalos, particularmente em operações de serviço, há o desafio de se estabelecerem medidas de capacidade. Segundo Slack (2000), "o principal problema com a medição da capacidade é a complexidade da maior parte dos processos produtivos. Somente quando a produção é altamente padronizada e repetitiva é fácil definir a capacidade". As principais decisões sobre capacidade devem considerar os seguintes aspectos: Estimar as capacidades das instalações atuais, prever as necessidades futuras de capacidade para todos os produtos e serviços, identificar e analisar as diferentes fontes e formas de alterar a capacidade no curto, médio e longo prazo, identificar diferentes formas de influenciar a demanda, avaliar o impacto da decisão a respeito de capacidade sobre a qualidade do serviço prestado e, por fim, escolher fontes alternativas para incrementar a capacidade do sistema.

As decisões sobre a capacidade requerem a identificação das áreas do sistema em que os benefícios da alteração vão repercutir no aumento da capacidade global. Uma vez efetuada esta etapa, as políticas de alteração de capacidade devem considerar modificações de curto, de médio e longo prazo. As estratégias de gestão da capacidade e da demanda são instrumentos de que os gestores dispõem face ao desafio de equilibrar a oferta à demanda.

Neste contexto, faz-se necessária uma abordagem sistemática, estruturada e

consistente que trate dessas questões e ofereça alternativas para o desenvolvimento de uma estratégia adequada de gestão da capacidade em cada caso, como será apresentado nos casos a seguir.

3. ESTUDOS DE CASO

Nesta seção serão apresentadas aplicação práticas na área de saúde com a implementação do processo de focalização da TOC integrado a simulação. O primeiro estudo de caso descreve o processo numa unidade de pronto atendimento. Em seguida em um centro de tratamento de câncer numa unidade assistencial do Instituto Nacional do Câncer (INCA), na cidade do Rio de Janeiro, RJ. O segundo estudo de caso foi realizado na unidade de emergência hospitalar do Hospital Municipal de Emergência Henrique Sérgio Gregori (HMEHSG), em Resende, RJ.

3.1 HOSPITAL GERAL DE NOVA IGUAÇU

Neste caso especificamente utilizou-se a simulação de eventos discretos para identificar a restrição e posteriormente aplicou-se o processo de focalização. A metodologia de simulação utilizado está fundamenta em Bank e Carlson (1984) e em Rossetti et al (1999). O Hospital Geral de Nova Iguaçu (Hospital da Posse) foi inaugurado no segundo trimestre de 1982. A unidade possui 15 mil metros quadrados de área construída e é dividida em cinco blocos: Bloco A (ambulatório e laboratório), Bloco B (direção, raios-X e CPD), Bloco C (emergência), Bloco D (refeitório, cozinha, UTI, centro cirúrgico, RPA e esterilização) e Bloco E (enfermagem e prédio da subestação elétrica). Em área anexa, está o Centro de Referência Parteira Mariana Bulhões (CRMB). O Serviço de pronto atendimento apresentava uma média mensal de 16.279 atendimentos, com uma média diária de 543 pacientes e uma taxa horária média de 54 pacientes, o que têm refletido no congestionamento e longas filas de espera que causam desconforto aos pacientes e pressão sobre a equipe médica.

3.1.1- SERVIÇO DE PRONTO ATENDIMENTO (SPA)

O Serviço de Pronto-Atendimento (SPA) oferece as especialidades de pediatria e clínica médica 24 horas por dia, além do setor de ortopedia e de pequenas cirurgias que funciona de 7h às 19h. Um aspecto que chamou a atenção no SPA é que este funciona como uma das portas de entrada para a emergência do hospital, com a qual possui um corredor de acesso. Além disso, neste setor efetua-se uma classificação de risco dos pacientes nas filas, de modo a priorizar o atendimento dos casos mais graves. Este procedimento é realizado por dois enfermeiros que efetuam uma breve entrevista com os pacientes de modo a estabelecer uma avaliação inicial. Os recursos disponíveis neste segmento são: 02 recepcionistas, 02 enfermeiras, 02 médicos no Pronto-Atendimento, 02 médicos na área de observação, 02 consultórios no pronto-atendimento e 01 sala de observação.

Outra questão identificada *in loco* durante a visita ao setor é a diversidade de casos em termos de tipo de ocorrência, que vão desde pacientes buscando atestado médico ou atenção, até casos de emergência. Uma emergência é um estado ou situação inesperada, de quebra da ordem estabelecida, podendo ser esta de vários tipos, dependendo do sistema no qual se apresenta. Segundo Sabbadini (2005) esta classificação pode ser agrupada em três categorias:

Emergenciais: Casos que apresentam um estado que indica perigo de morte e necessidade de atenção imediata;

Urgentes: Aqueles que apesar de não apresentarem perigo de vida devem receber atenção médica especializada, no decorrer de poucas horas, sob o risco de tornarem-se uma emergência.

Ambulatoriais: Casos crônicos que não exigem atenção imediata, nem oferecem riscos imediatos ao estado de saúde do paciente.

3.1.2 SIMULAÇÃO COMPUTACIONAL

O SPA vem enfrentando uma elevação de demanda por seus serviços e os gestores identificaram problemas de sobrecarga no setor. Em face disso desenvolveu-se um estudo que resultou num modelo de

simulação a eventos discretos. A metodologia utilizada baseou-se em Rossetti et al (1999) e considera os seguintes passos:

3.1.2.1 DEFINIÇÃO DO PROBLEMA E DOS OBJETIVOS

A partir da análise do fluxo dos pacientes e da elaboração do modelo conceitual pode-se identificar a existência de um gargalo na unidade de observação de pacientes com casos urgentes que se relaciona diretamente com problemas de congestionamento e também resultam em elevado tempo de permanência do paciente no setor. A partir desta avaliação inicial definiu-se como objetivo identificar a configuração de recursos humanos e materiais que minimizassem os efeitos da restrição existente. A hipótese inicial foi a de que a disponibilidade de leitos representava o recurso restritivo.

3.1.2.2 DESENVOLVIMENTO DO MODELO CONCEITUAL

Definido o problema e determinados os objetivos da simulação, formulou-se a partir da análise do fluxo de pacientes o modelo conceitual do SPA, sendo que este estudo teve como foco o processo de atendimento de pacientes com casos urgentes naquela unidade de serviço, o qual é descrito a seguir:

O paciente chega ao SPA e aguarda numa fila. Dois enfermeiros fazem entrevistas com estes para efetuar uma classificação de risco, quanto à severidade do quadro clínico de cada paciente. Estes podem ser classificados com um caso de emergência (risco de vida), um caso urgente ou um caso ambulatorial (crônico), tendo a respectiva prioridade no atendimento.

A partir da avaliação dos enfermeiros cada paciente segue um fluxo distinto dentro das instalações do SPA. Sendo um caso de emergência é encaminhado diretamente para o pronto-socorro do hospital para os procedimentos necessários e se for o caso internação, saindo posteriormente do sistema SPA e entrando noutro processo, pertinente a área de emergência. Sendo classificado como um caso urgente, o paciente é encaminhado a uma sala de observação, onde receberá os primeiros cuidados e será monitorado por um dos dois médicos alocados naquela área. Havendo agravamento do quadro, será encaminhado para a área de emergência.

Estando reabilitado, o paciente sai do sistema.

Nos casos relacionados aos pacientes ambulatoriais (crônicos) verifica-se a disponibilidade da especialidade clínica no SPA sendo o paciente encaminhado para atendimento em uma das duas salas disponíveis. Não havendo a especialidade clínica no setor o paciente é encaminhado para atendimento na rede primária municipal de Nova Iguaçu e sai do sistema. O modelo conceitual descrito é serviu de base para a elaboração do modelo computacional.

3.1.2.3 COLETA DE DADOS

Na primeira etapa do estudo as informações necessárias à elaboração do modelo conceitual foram obtidas *in loco*. Este trabalho foi realizado com o acompanhamento de uma enfermeira do SPA. Os dados relativos aos volumes de procedimentos e demais dados estatísticos foram obtidos nos registros disponíveis na página eletrônica e de outras fontes do HGNI. As taxas de chegada e os tempos de atendimento foram estimados com base na experiência dos profissionais de saúde envolvidos no projeto. Cabe ressaltar que foram identificadas distribuições de probabilidade que se ajustassem aos dados coletados. Os resultados obtidos indicaram que a chegada de pacientes na área de observação ocorre segundo uma exponencial de média 11. Para o tempo de atendimento pelos médicos utilizou-se uma distribuição triangular com valor mínimo de 10 minutos, com o valor mais provável de 20 minutos e o valor máximo de 30 minutos.

3.1.2.4 CODIFICAÇÃO DO MODELO COMPUTACIONAL

A partir do fluxo de tratamento a pacientes e do modelo conceitual foi desenvolvido o modelo computacional, com a utilização do software Arena, versão de treinamento 10.0, cujos módulos pré-definidos facilitaram a sua construção.

A simulação é terminal, considerando um período equivalente a duas horas de aquecimento. Foram efetuadas cinquenta replicações, cada uma equivalendo a um dia de operação da unidade de pronto atendimento, com a finalidade de obter mais

acuracidade e eliminar distorções decorrentes de flutuações estatísticas. O tempo definido no relógio de simulação inicia-se às 8 horas e termina às 16 horas.

3.1.2.5 VALIDAÇÃO E VERIFICAÇÃO DO MODELO

Nesta fase realizou-se uma apresentação estruturada do modelo desenvolvido para os gestores do hospital e a equipe envolvida no projeto, na qual foi descrita a forma como este se relaciona com o sistema real. A verificação é o processo de assegurar que o modelo de simulação foi construído corretamente e é completada quando o programa funciona da maneira como o analista de simulação pretendia. Ela pode ser efetuada rodando-se a simulação e monitorando-se sua operação. Se os parâmetros de entrada e a estrutura lógica do modelo estiverem corretamente representados no código, a verificação é finalizada. Para garantir a correta representação do modelo foi realizada a verificação lógica dos módulos com o fluxo de pacientes. O experimento foi realizado diversas vezes até que os resultados obtidos replicassem o modelo real.

3.1.2.6 EXPERIMENTAÇÃO, EXECUÇÃO DA SIMULAÇÃO E ANÁLISE DOS RESULTADOS

Considerando a configuração atual da área de observação de casos de pacientes urgentes foram realizados diversos experimentos com várias rodadas de simulação variando a quantidade de leitos disponíveis. Os experimentos realizados compreenderam um período de 8 horas iniciando-se às 8 horas e terminando às 16 horas. Esse corte de tempo é devido ao maior fluxo de pacientes estar concentrado neste período. Para cada experimento foram realizadas 50 replicações com o intuito de eliminar distorções e garantir a consistência dos resultados. Cada replicação equivale a um período completo de oito horas de operação nas condições mencionadas. Dos resultados das simulações puderam-se extrair informações importantes para uma avaliação daquele setor. A figura 1 apresenta os resultados da primeira hipótese, onde se consideraram como fixos o número de médicos e de enfermeiras e variou-se o número de leitos disponíveis.

Figura 1: Disponibilidade de leitos e taxa de ocupação da equipe médica

A situação específica explorada nos experimentos partiu da hipótese de que a disponibilidade de leitos na área de observação era uma restrição à agilidade no atendimento e causava efeito sobre a taxa de ocupação da equipe médica. Pode-se verificar na figura 1 que à medida que se aumenta o número de leitos eleva-se a taxa

de ocupação dos médicos e das enfermeiras. No mesmo sentido pode-se verificar um aumento da produtividade da equipe pela redução no número de pacientes no sistema ao final das rodadas de simulação, o que conforme se pode verificar na figura 2, indica um maior número de pacientes atendidos e liberados.

Figura 2: Efeito da disponibilidade de leitos sobre pacientes no sistema

Uma segunda situação foi verificada após a realização dos primeiro experimentos, a partir da observação do comportamento do tempo médio de espera do paciente por leito.

Conforme se pode ver na figura 3, observa-se que a partir do quinto leito ocorre um ligeiro aumento no tempo de espera.

Figura 3: Relação entre número de leitos e tempo médio de espera do paciente

Partindo-se da hipótese que tal efeito é resultado do fato de o número de médicos e enfermeiras ter permanecido fixo, realizaram-se experimentos acrescentando-se um médico a partir do sexto leito para avaliar o efeito sobre o tempo médio de espera (em horas). Verificou-se que houve pouco efeito sobre o tempo de espera, e confirmou-se a disponibilidade de leitos como uma restrição física. Aplicou-se o processo de focalização da teoria das restrições, identificando-se a necessidade estratégica de implantar-se um sistema de classificação de risco acompanhado de agendamento para as consultas ambulatoriais.

3.2 INSTITUTO NACIONAL DO CÂNCER

No tratamento do câncer, um dos fatores críticos de sucesso é o diagnóstico precoce. A agilidade no encaminhamento do paciente a adequadas condutas terapêuticas, com a redução do tempo de espera é fundamental para o aumento da sobrevivência, melhoria da qualidade de vida e chances de cura. O fluxo de tratamento de pacientes é composto de três grandes processos: Processo de Triagem: Identifica através de exames clínicos se a suspeita de câncer procede ou não; Processos de Diagnóstico: Detectam através de uma série de exames o tipo de câncer, sua localização e o estágio de evolução clínica da doença. Somente após estes exames o tratamento é realmente iniciado; Processos Terapêuticos: Encaminhamento do paciente a adequadas condutas terapêuticas, que correspondem, em sua grande maioria, a uma cirurgia, aplicações de radioterapia e quimioterapia.

Estes processos são interdependentes. O desempenho de cada um depende do anterior. O desempenho do sistema depende da estrutura desta cadeia de processos. O que importa é o sistema como um todo, e não os processos isolados.

O problema em questão está relacionado à identificação de possíveis gargalos no processo de diagnóstico e à melhor alocação dos equipamentos hospitalares, componentes estratégicos da capacidade instalada de um hospital. Estes gargalos são considerados como restrições à meta de oferecer um tratamento ágil e com qualidade. Neste contexto, pode-se utilizar a TOC que é um conjunto de princípios e ferramentas para resolução e minimização de gargalos existentes, para melhoria do desempenho do sistema como um todo. A seguir será descrita a aplicação do processo de focalização:

Identificar as restrições do sistema

Para aplicar o primeiro passo do processo de focalização, que é a identificação da restrição, elaborou-se o fluxo de paciente, desde sua triagem até o início de um procedimento terapêutico. Quando um paciente é encaminhado ao hospital, a primeira etapa possui o objetivo de confirmar através de exames clínicos, o diagnóstico de câncer. Caso o diagnóstico não seja confirmado, o paciente é encaminhado para outros hospitais. Caso seja confirmado ou haja a necessidade de exames complementares, o paciente é matriculado para que se detecte a localização e o estágio de evolução clínica da doença. Após estes exames, o tratamento é realmente iniciado.

O fluxo do paciente foi analisado de forma detalhada e através do processo de focalização foi possível identificar que o departamento de radiologia desempenhava função estratégica, já que qualquer atraso na área afetava todo o processo. Os demais exames de patologia clínica e anatomia patológica apresentavam resultados com intervalos de tempo bem menores que os de imagem, não se constituindo em fatores limitantes no processo de diagnóstico.

Através da aplicação da TOC ficou clara a necessidade de aumentar a capacidade diária de execução de exames, com repercussão imediata na redução do tempo entre a marcação e sua realização. O gargalo estava localizado exatamente entre os processos de agendamento e realização dos exames de imagem, que levavam mais de 30 dias no caso da tomografia computadorizada. Neste contexto a redução do tempo era fundamental para agilizar o início do tratamento.

Decidir como explorar a restrição do sistema

Identificada a restrição, o passo seguinte foi decidir como explorá-la. Todos os esforços foram direcionados para aumentar a taxa de utilização dos equipamentos considerados recursos gargalos. De modo a atender a todos os pacientes e evitar a ociosidade dos tomógrafos, a gestão de curto prazo é fundamental, para através de uma análise de sensibilidade antecipada, encontrar uma solução de equilíbrio. As ações de curto prazo consideradas para aumentar eficiência do sistema no sentido de reduzir o tempo dos exames foram à atribuição de prioridade de atendimento de acordo com o paciente e o tipo de exame, efetuar o agendamento buscando minimizar a ociosidade dos equipamentos disponíveis e ajustar os níveis de capacidade no caso de sobrecargas.

Para avaliar essas alternativas o problema em questão foi o foco de um estudo de simulação (GONCALVES E DE OLIVERA, 2004). Durante a modelagem e a fase de levantamento de dados, pontos importantes foram debatidos com as áreas envolvidas. A discussão e o entendimento de cada atividade foi tratada de forma integrada com o corpo clínico da área de radiologia, para que aspectos relevantes do processo replicado fossem representados no modelo construído.

Os resultados da simulação permitiram uma análise apurada de indicadores estratégicos, tais como taxa de utilização dos tomógrafos,

quantidade de pacientes atendidos, tempo de permanência nas filas, fornecendo informações importantes para a tomada de decisão em cada cenário analisado.

Subordinar/Sincronizar todos os recursos restantes às decisões tomadas acima

Identificadas as alternativas mais adequadas o passo seguinte foi subordinar e sincronizar todos os recursos restantes às decisões tomadas acima. A utilização da simulação foi importante na identificação antecipada de eventuais problemas na agenda de um dia típico que foram solucionados através de ações gerenciais tais como: 1) a utilização de roteiros alternativos, que consistem na reprogramação do agendamento visando maximizar a utilização dos tomógrafos; 2) horas extras para atender a demanda no caso de estouro da capacidade planejada.

Para garantir a máxima utilização dos recursos gargalo foi desenvolvido e implementado o Sistema de Gestão de Atendimento (GONÇALVES E DE OLIVEIRA, 2003; 2004), que indica quando há necessidade de intervenção para manter o fluxo de atendimento planejado. Sua abordagem inovadora permite aos gestores acompanhar a situação das filas em cada uma das clínicas.

Elevar a capacidade da restrição do sistema

Ao se implementar os passos 2 e 3 obtém-se como resultado a elevação da capacidade da restrição do sistema pela redução da ociosidade no gargalo. Não houve a necessidade de elevar a restrição pela contratação de mais recursos humanos ou pela aquisição de equipamentos, uma vez que a solução encontrada atendeu os objetivos de um atendimento mais ágil.

Se a restrição se deslocar, retornar ao passo 1 e não permitir que a inércia se instale.

O simulador e o Sistema de Gestão de Atendimento permitem que o gestor tenha total domínio do dia-a-dia da clínica, aumentando sua eficiência nas tomadas de decisão relativas à programação da agenda. Desse modo é possível gerenciar a restrição e atuar de maneira efetiva no sentido de garantir a eficácia do fluxo desejado de serviço prestado ao paciente.

O resultado após a implementação das alternativas mostra que o intervalo de tempo médio entre a requisição do exame de tomografia computadorizada e sua realização

caiu de 30,71 para 22,71 dias, com uma redução de 25% do tempo de espera, comprovando a eficácia do processo.

3.3 HOSPITAL MUNICIPAL DE EMERGÊNCIA HENRIQUE SÉRGIO GRÉGORI

Nos hospitais, o departamento de emergência é uma área estratégica, onde situações críticas freqüentemente ocorrem. No atendimento de emergência o tempo é um fator crucial. A meta é a avaliação rápida, a estabilização, o tratamento e pronta admissão do paciente ao hospital. Quanto mais eficaz for esse processo, maiores serão as chances de restabelecimento, redução de seqüelas e cura do paciente.

O Hospital Municipal de Emergência Henrique Sérgio Gregori (HMEHSG) é um componente do sistema de saúde de Município de Resende, integrando o Sistema Único de Saúde (SUS), tendo como finalidade o atendimento integral a pacientes em estado de urgência ou emergência. Devido à eficácia na prestação de atendimento de casos de emergência, além da sua infra-estrutura e localização geográfica, o HMEHSG tornou-se ao longo dos anos uma referência neste tipo de serviço para o município de Resende e a micro-região vizinha. É reconhecido também como hospital de referência no atendimento de emergências decorrentes de acidentes ocorridos da Rodovia Presidente Dutra.

A unidade de emergência do hospital funciona 24 horas, todos os dias, atendendo diariamente uma média de 192 pacientes. A direção do HMEHSG vem observando mudanças no perfil da população de pacientes atendidos e ao aumento da procura por seus serviços. Como consequência, sua capacidade de atendimento é inferior à demanda, verificando-se com freqüência congestionamento na unidade de emergência, a formação de longas filas de espera e o aumento do tempo de permanência de pacientes no hospital, tendo como consequência pressão sobre os serviços e a equipe médica.

Outro fato constatado pela direção do hospital é que a maioria dos pacientes que procuram a unidade de emergência não necessita de atendimento imediato, prejudicando os paciente de urgência e os de emergência. Neste contexto a teoria das restrições é adequada para a caracterização e tratamento do problema em questão, uma vez que o

objetivo principal é agilizar o fluxo de tratamento de pacientes, identificando e gerenciando gargalos. O processo de focalização da TOC é descrito a seguir:

Identificar as restrições do sistema

Para esse fim, foi elaborado o fluxo de tratamento a pacientes no HMEHSG desde a triagem até o tratamento. Toda a elaboração foi realizada *in loco* com o acompanhamento do médico chefe da emergência e posteriormente validada por outros membros do corpo clínico, envolvendo médicos e enfermeiros da unidade. O fluxo foi estudado para cada perfil de paciente segundo critérios de emergência (risco de vida), urgência (caso grave) e ambulatorial. Um fluxo geral é apresentado na figura 5.

O paciente chega à unidade de emergência, sendo recebido por uma das recepcionistas, que efetua o preenchimento e o registro do boletim de atendimento médico. Em seguida solicita ao paciente que aguarde na recepção. O médico disponível recebe o paciente e inicia o atendimento. Faz a avaliação deste e emite o diagnóstico. Constatado ser um caso urgente é encaminhado para o pronto-socorro sendo atendido por um médico da emergência. Tratando-se de um paciente ambulatorial após a consulta este é liberado.

O fluxo de pacientes, que é composto pelos processos de triagem, diagnóstico e tratamento foi analisado detalhadamente. São operações interdependentes, em que o desempenho de uma depende do resultado do imediatamente anterior. Na análise do fluxo de tratamento constatou-se o descompasso estrutural entre a oferta de recursos e a demanda, que resultam na formação de filas de espera, aumento do tempo de permanência de pacientes para o início do tratamento. A menor taxa de atendimento está nas clínicas, conforme apresentado na figura 4.

Associada a essa restrição há pressão sobre os serviços e a equipe médica, que divide sua atenção entre a demanda de pacientes que chegam para o primeiro atendimento e a daqueles que retornam dos serviços de laboratório e radiologia, com os resultados dos exames.

Conjuntamente, os pacientes ambulatoriais, que não necessitam de atendimento imediato, concorrem com casos urgentes pela atenção dos médicos, o que leva ao aumento do

tempo de espera pelo primeiro atendimento, protelando o início do tratamento. Nesta situação o paciente com quadro urgente aguardava na fila de espera, sofrendo

agravamento no seu quadro clínico, o qual somente seria detectado efetivamente quando este fosse atendido pelo médico.

Figura 4: – Fluxo de atividades na unidade de emergência no HMEHSG

O processo se configura como uma restrição à meta da unidade de emergência de oferecer um tratamento ágil ao paciente urgente. A agilidade no encaminhamento ao tratamento é fundamental para a redução de seqüelas e melhoria da qualidade do serviço prestado a esse paciente.

Decidir como explorar a restrição do sistema

Identificada a restrição, no sentido de aproveitar ao máximo a capacidade existente no gargalo e manter o fluxo de atendimento desejado, optou-se por duas alternativas de reorganização dos processos. A primeira delas foi à implementação de uma triagem especializada para direcionar os pacientes com quadro urgente para áreas de tratamento sem restrição. A segunda alternativa consiste na alocação dos médicos nas clínicas em função da demanda por cada uma delas, principalmente nos horários de pico, que são os mais críticos.

Para avaliar a viabilidade dessas alternativas definiu-se um modelo de simulação cujo escopo compreendeu os setores de Pronto Atendimento e Pronto-Socorro da unidade de

emergência, considerando as instalações, os recursos humanos e materiais relativos a cada processo analisado. Com base no fluxo de tratamento a pacientes e do modelo conceitual foi desenvolvido o modelo computacional, com a utilização do *software* MedModel, versão 3.5, cuja abordagem utiliza módulos pré-definidos que facilitam a sua construção. Durante todo o desenvolvimento houve o envolvimento dos médicos responsáveis pela unidade. A animação do modelo computacional desenvolvido permitiu o seu completo entendimento pelos gestores e demais envolvidos no projeto, que o validaram como uma representação do sistema real.

O foco da aplicação do modelo computacional é na agilização do fluxo de pacientes urgentes pela implementação de uma triagem especializada com classificação de risco, antes do gargalo, de modo a priorizar os casos que demandam maior atenção naquele hospital, uma vez que os casos de emergência são prontamente atendidos, não havendo espera do paciente pelo atendimento médico.

Os resultados da simulação permitiram a análise de indicadores estratégicos como a taxa de ocupação dos médicos, quantidade de pacientes atendidos, ociosidade e sobrecarga nas clínicas e o tempo de espera por atendimento nas filas fornecendo informações fundamentais para a tomada de decisão pelos gestores.

A implementação deste passou implicou na alteração de procedimentos e rotinas na unidade de emergência, como por exemplo a disponibilização de enfermeiras profissionais de outros setores para atuar na triagem, no sistema de admissão da unidade de emergência. Do mesmo modo a alocação planejada para os médicos previu a disponibilidade destes em função da demanda por cada clínica, sendo modificada dinamicamente nos horários de pico.

Subordinar/Sincronizar todos os recursos restantes às decisões tomadas acima

A implementação deste passou implicou na alteração de procedimentos e rotinas na unidade de emergência, como por exemplo a disponibilização de enfermeiras profissionais de outros setores para atuar na triagem, no sistema de admissão da unidade de emergência. Do mesmo modo a alocação planejada para os médicos previu a disponibilidade destes em função da demanda por cada clínica, sendo modificada dinamicamente nos horários de pico.

Para evitar que a restrição seja quebrada no passo 4, e no sentido de aproveitar ao máximo os recursos no gargalo, um Aplicativo de Monitoramento da Demanda foi desenvolvido e implementado (SABBADINI, 2005). Esta ferramenta permite identificar visualmente os dias da semana e horários dentro de um dia com maior pico de procura por atendimento, assim como analisar o comportamento, hora a hora, dia a dia e semana a semana, dentro de cada mês ao longo do ano, melhorando a previsibilidade acerca da demanda e desse modo permitindo aos gestores atuar no sistema de maneira eficaz.

Elevar a capacidade da restrição do sistema

A programação do fluxo de atendimento foi executada melhorando a alocação dos médicos nas clínicas em função da demanda e possibilitou elevar a capacidade na restrição sem a necessidade de inicialmente efetuar investimentos na aquisição de

equipamentos, ampliação de instalações ou contratação de mais médicos.

Se a restrição se deslocar, retornar ao passo 1 e não permitir que a inércia se instale.

Dada a dinâmica dos sistemas, as soluções tende a se deteriorar ao longo do tempo. Nesse sentido a restrição pode mudar de lugar, deslocando-se para outro processo. A utilização do simulador e o Aplicativo de Monitoramento de Demanda possibilitam aos gestores do hospital o controle do dia-a-dia da unidade de emergência, contribuindo para o gerenciamento da restrição no sentido de não permitir que a inércia se instale e se transforme numa restrição do sistema.

Os resultados obtidos através do simulador e comparados com o processo real apresentaram uma diminuição no tempo médio de espera dos pacientes urgentes de 44,28 minutos para 28,86 minutos, com uma redução de tempo de 34,95%.

4 CONCLUSÃO

A análise do resultado dos três estudos de caso evidencia as possibilidades do uso da simulação, teoria das restrições e gestão da capacidade de forma integrada. Em todos os três casos estudados as técnicas apresentadas permitiram o gerenciamento estratégico de gargalos e a ação em termos de gestão da capacidade, que resultaram na melhoria no gerenciamento dos processos de tratamento a pacientes.

Deu-se especial ênfase a ao processo de gerenciamento da TOC, dado que tanto a simulação com os fundamentos e aplicações de gestão da capacidade são bastante consolidados. No primeiro caso, Hospital Municipal de Nova Iguaçu, em particular, partiu-se da abordagem de simulação para posteriormente através do processo de focalização identificar as ações de melhoria requeridas na gestão estratégica da capacidade. Nos demais casos, INCA e HMEHSG, partiu-se do processo de focalização da Teoria das Restrições para identificar o gargalo e posteriormente realizaram-se estudos de simulação para orientar a gestão da capacidade.

O gerenciamento baseado na teoria das restrições assume uma dimensão estratégica num ambiente complexo em que se encontra o desafio de gerenciar permanentemente recursos. Ao redirecionar as ações de

melhoria nos recursos gargalo, pode-se ajustar a capacidade à demanda e atuar nos pontos do sistema em que os benefícios da alteração vão repercutir no aumento da capacidade global. Em todos os casos os ganhos obtidos possuem implicações importantes como a satisfação do paciente no

processo de tratamento e sua percepção em relação à qualidade dos serviços prestados. Outro aspecto a ser considerado é a melhoria da assistência prestada a estes que passam a ter seu tratamento iniciado mais rapidamente, contribuindo para a elevar a sua qualidade de vida, chances de cura e restabelecimento.

REFERÊNCIAS

- [1]. BANKS J.; CARSON, J. S. Discrete-event system simulation. Prentice-Hall international series in industrial and system engineering. New Jersey: Prentice Hall, 1984.
- [2]. BRAILFORD, S. C., HARPER, P. R.; PATEL, B.; PITTI M.. 2009. "An Analysis of the Academic literature on Simulation and Modelling in Health Care." J of Simulation 3(3):130-140
- [3]. Burton, T, GOLDRATT INSTITUTE, 2001, "TOC Case Study: Healthcare – What if my organization's goal is not to make money", disponível em <http://www.goldratt.com/for-cause/partnerperspsep2001toct.htm>
- [4]. CHAGAS, C. D. Simulação em emergência médico-hospitalar: um estudo de caso no Hospital Municipal Miguel Couto. Dissertação (Mestrado em Engenharia de Produção) - COPPE/UFRJ, Rio de Janeiro, 2003.
- [5]. DE OLIVEIRA, MJF;TOSCANO, LNP, Emergency information support system for Brazilian public hospital, In: M.S. Rauner and K. Heidenberger (Eds.), Quantitative approaches in health care management, Peter Lang, (2003) 235-251.
- [6]. DE OLIVEIRA, M. J. F. e ACHÃO FILHO, N. A simulação como método de avaliação da qualidade de atendimento na emergência de um hospital municipal. In: Simpósio de Pesquisa Operacional da Marinha, 6., Simpósio de Logística da Marinha, 7., Anais. Rio de Janeiro, 2003.
- [7]. GOLDRATT, E. M. What is this thing called Theory of Constraints and how should it be implemented? Massachusetts: North River Press, 1990.
- [8]. GOLDRATT, E.M., 1994, A Meta , 17ª edição, ed. Educator, São Paulo.
- [9]. GONÇALVES, A. A. Gestão da capacidade de atendimento em hospitais de câncer. Dissertação (Doutorado em engenharia de produção). Rio de Janeiro : COPPE/UFRJ, 2004.
- [10]. GONÇALVES A.A., DE OLIVEIRA M.J.F., 2004, The use of simulation to improve cancer treatment access, Accessibility and Quality of Health Services, Peter Lang, Frankfurt Main/Berlin/Bern/New York/Paris/Viena, Germany, pp.149-160.
- [11]. KERSHAW R., 2000, Using TOC to cure Healthcare problems", Management Accounting Quaterly, Spring.
- [12]. MAGALHÃES, M. S. (2006). Simulação do Sistema de Admissão de Emergência do Hospital Universitário Antônio Pedro. Dissertação de M.Sc, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- [13]. PHIPPS B., 1999, "Hitting the Bottleneck' , Health Management, United Kingdon, disponível em Goldratt Institute www.goldratt.com/hlthmgmtmn.htm.
- [14]. ROSSETTI, M. D. et al. Emergency department simulation and determination of optimal attending physician staffing schedules. In 1999 Winter Simulation Conference. Proceeding , Arizona 1999.
- [15]. RÉBULA, U. et. al. Integrando técnicas e procedimentos de gestão de operações: uma aplicação em um banco comercial brasileiro de grande porte. Revista Produção. 2009.
- [16]. SABBADINI, F.S. Gerenciamento de restrições em hospital de emergência: um estudo de caso no Hospital Municipal Henrique Sérgio Gregori. Dissertação (Mestrado em Administração e Desenvolvimento Empresarial). Rio de Janeiro. UNESA, 2005.
- [17]. SABBADINI, F. S. ; De Oliveira, M. J. F. ; Gonçalves, A A . Administration of constraints in a Brazilian public emergency hospital service. In: Xiolan Xie; Françoise Lorca; Éric Marcon.. (Org.). Operations Research for Health Care Delivery Engineering.. Saint-Étienne: Saint- Étienne: Publications de l'Université de Saint-Étienne, 2008, v. 33, p. 367-378.
- [18]. SLACK, N. Administração da produção. 2. ed. São Paulo : Atlas, 2002.

Capítulo 2

MINERAÇÃO DE DADOS COM DETECÇÃO DE OUTLIERS EM TAREFAS DE PREDIÇÃO DE SÉRIES TEMPORAIS

Lidio Mauro Lima de Campos

Gustavo Augusto Lima de Campos

Resumo: Este artigo apresenta um Sistema Neural Híbrido aplicado ao processo de descoberta de conhecimento em base de dados (KDD). Utiliza-se uma Rede SOM que realiza a etapa de pré-processamento (eliminação de outliers) após a etapa de seleção dos dados brutos, os mesmos são filtrados eliminando aqueles que deveriam seguir um padrão esperado, mas não o fazem. A saída da Rede SOM classificará os dados entre bons e ruins (outliers) para a mineração de dados. Após essa etapa, realiza-se a transformação dos dados por meio da redução e normalização “z-score”. Na etapa de mineração de dados utilizam-se dois classificadores recorrentes não lineares recorrentes ARX e NARX apresentados nas secções 5.1 e 5.2 respectivamente, selecionando-se o melhor preditor. Após essa escolha, visualizam-se os dados e interpretam-se os resultados gerando o conhecimento a cerca da predição da série temporal estudada. Conclui-se que : as Redes NARX são superiores as Redes ARX . O Sistema Neural Híbrido constitui-se uma boa ferramenta para predição de séries temporais, que apresentam peculiaridades específicas, como é o caso do Índice de Preços ao Consumidor-IPC.

Palavras-Chave: Sistema Neural Híbrido, Rede de Kohonen, Rede Neural Recorrente, Descoberta de Conhecimento em Base de Dados.

1 INTRODUÇÃO

A extração de Conhecimento, geralmente referenciada na literatura como “*Knowledge Discovery in Database*” (KDD) é uma área multidisciplinar que incorpora técnicas utilizadas em diversas áreas como Banco de Dados, Inteligência Artificial, Estatística e visualização de dados, dentre outras. Fayyad, Piatetsky-Shapiro & Smyth [1] definem KDD como sendo: “Extração de Conhecimento de Base de Dados é o processo de identificação de padrões válidos, novos, potencialmente úteis e compreensíveis embutidos nos dados”. Segundo Goebel e Gruenwald [6], o termo KDD é usado para representar o processo de tornar dados de baixo nível em conhecimento de alto nível, enquanto mineração de dados pode ser definida como a extração de padrões ou modelos de dados observados.

Neste contexto, os estudos em KDD, despontam como uma área de grande interesse cuja demanda por resultados vem crescendo nos últimos anos. O Processo de Descoberta de Conhecimento tem na Mineração de Dados o cerne de sua execução, pois é nesta fase em que são buscadas informações implícitas ao conjunto de dados em análise que sejam potencialmente interessantes e utilizáveis, e previamente desconhecidas [1].

As Redes Neurais Artificiais (RNA) tem sido utilizadas de forma crescente, como classificadores no processo de Mineração de Dados, no que diz respeito predição de séries temporais em vários ramos de negócio, economico-financeiro [7][8], predição de séries caóticas [9], consumo de energia [10] dentre outras. As RNAs podem aprender a partir de exemplos, reconhecer um padrão escondido em observações históricas e usá-los para prever valores futuros. Além disso, são capazes de lidar com informação incompleta ou dados ruidosos e podem ser muito eficaz, especialmente em situações onde não é possível definir as regras ou etapas que levam à solução de um problema. As Redes de Kohonen Auto Organizáveis (SOM) são comumente e eficazmente utilizadas para fins de Mineração de Dados em tarefas de Clusterização [11]. Além disso, podem ser utilizadas para facilitar a remoção de “outliers” de base de dados [12].

O problema em si que se pretende resolver nessa pesquisa é a Predição do Índice de Preços ao Consumidor, “k” passos à frente,

adotando como “dataset” a série temporal do IPC no período de jan-1990 a dez-2010 fornecida pelo Sistema de Gerenciamento de séries do Banco Central do Brasil disponível em [14]. O conhecimento gerado será útil na definição de políticas públicas e planejamento público no diz respeito a contratações coletivas, negociações salariais, índices inflacionários e como facilitador na implantação de projetos empresariais.

O objetivo da presente pesquisa é inicialmente utilizar uma Rede de Kohonen para realizar a etapa de pré-processamento de dados: remoção de “outliers” de uma base de dados com dados representativos da série histórica do IPC - Índices de Preços ao Consumidor no período de jan-1990 a dez-2010. Após essa etapa, caminhando no processo de Extração de Conhecimento em Base de Dados, a série temporal modificada passará por normalização “z-score”. Dessa forma, após isso os dados tratados serão entrada para a etapa principal do processo de KDD, que é a Mineração de Dados.

As Redes Neurais Recorrentes (RNAR) ainda não foram totalmente exploradas, no processo de Mineração de Dados, principalmente no que diz respeito à predição de séries temporais, devido ao seu tedioso treinamento e por conta de suas estruturas complexas[13]. Sendo, portanto, um campo aberto para pesquisas o estudo de classificadores não lineares recorrentes, baseados em modelos Auto Regressivos com entrada exógena (ARX) e (NARX) [15]. Assim sendo, nessa pesquisa testam-se como classificadores as redes ARX e NARX, comparando-se o desempenho das duas e escolhendo-se a melhor.

Para que os objetivos da pesquisa sejam alcançados, na secção 2 discutem-se conceitos sobre mineração de dados, a secção 3 explica o que é o Índice de preços ao Consumidor e sua utilização pelo Governo Brasileiro. Na secção 4 apresenta-se a Rede Neural SOM, utilizada para clusterização, na secção 5 apresentam-se duas Arquiteturas de Redes Neurais Recorrentes utilizadas na predição do Índice de Preços ao Consumidor: a rede ARX (Auto-Regressiva com entrada externas) e a NARX (Auto Regressiva não linear com entrada externas). Na secção 6, discute-se a metodologia que contempla todas as etapas do processo de descoberta de conhecimento em Base de Dados. Na secção 7, apresentam-se a execução das etapas da metodologia e os resultados de

simulação obtidos e finalmente nas secções 8 e 9 as conclusões e as referências bibliográficas consultadas.

2. MINERAÇÃO DE DADOS

Mineração de dados, ou “data mining”, é termo utilizado para nomear o processo de análise de conjuntos de dados com o objetivo de encontrar padrões que representem informações úteis e não triviais. Para tanto, utiliza-se de métodos matemáticos, heurísticas e algoritmos. A mineração de dados é parte de um processo maior e mais abrangente, o de descoberta de conhecimento em bancos de dados, que tem por objetivo extrair conhecimento de alto nível a partir de dados de baixo nível no contexto de grandes conjuntos de dados [1].

Descoberta de conhecimento em bancos de dados, ou “Knowledge Discovery in Databases”, é o termo, criado em 1989, que se refere ao amplo processo de descobrir conhecimento em dados armazenados em bancos de dados. Tal processo objetiva extrair conhecimento implícito e previamente desconhecido, buscando informação potencialmente útil nos dados. O processo, descrito em [1], consiste em uma sequência de cinco etapas, partindo dos dados existentes e chegando à descoberta do conhecimento extraído dos mesmos.

Seleção dos dados: a primeira etapa consiste em escolher qual o conjunto de dados que

será submetido ao processo. Seleciona-se um conjunto de dados alvo, ou foca-se em um subconjunto de variáveis ou amostras de dados.

Pré-processamento: nesta etapa, os dados podem sofrer uma qualificação, a fim de corrigir erros e inconsistências que poderão existir. Incluem-se limpeza de dados, eliminação de dados ruidosos, falta de dados e normalização.

Transformação: aqui os dados são convertidos em um formato adequado para serem acessados pelos algoritmos de mineração. É nela que também se realiza uma possível redução no número de variáveis, resumindo os dados que serão submetidos à mineração.

Mineração: é a etapa mais importante do processo. É nela que o algoritmo escolhido é aplicado sobre os dados a fim de se descobrir padrões interessantes. É fundamental para que esta etapa obtenha resultados de qualidade a correta aplicação dos passos anteriores.

Interpretação dos dados e Visualização: nesta última etapa do processo, os resultados obtidos na mineração são interpretados.

As técnicas de mineração de dados podem ser aplicadas a tarefas como classificação, estimativa, associação, segmentação e sumarização. Essas tarefas são descritas na Tabela 1.

Tabela 1 – Tarefas em Mineração de Dados

Tarefa	Descrição	Exemplos
Classificação	Consiste em construir um modelo de algum tipo que possa ser aplicado a dados não classificados visando categorizá-los em classes. Um objeto é examinado e classificado de acordo com uma classe definida [1].	<ul style="list-style-type: none"> classificar solicitações de pedidos de crédito. -esclarecer fraudes na declaração do imposto de renda.
Regressão	Regressão é aprender uma função que mapeia um item de dado para uma variável de predição real estimada [1].	<ul style="list-style-type: none"> prever a demanda futura de um novo produto. -estimar expectativa de vida média dos brasileiros.
Associação	Identificação de grupos de dados que apresentem coocorrência entre si.	<ul style="list-style-type: none"> quais produtos são colocados juntos em carrinhos de supermercado.
Segmentação (ou Clustering)	Processo de partição de uma população heterogênea em vários subgrupos ou grupos mais homogêneos	<ul style="list-style-type: none"> agrupamento de clientes com comportamento de compras similar. comportamento de clientes em compras realizadas na web para uso futuro.
Detecção de desvios (outliers)	Identificação de dados que deveriam seguir um padrão esperado, mas não o fazem.	<ul style="list-style-type: none"> detecção de intrusão em redes de computadores.

3.ÍNDICE DE PREÇOS AO CONSUMIDOR – IPC

O IPC quantifica o custo de produtos em diferentes momentos, em outras palavras são medidas do nível de preços de bens e serviços adquiridos pelas famílias através do tempo, sendo útil para o cálculo da inflação. O IPC é um dos produtos estatísticos que é alvo de uma utilização mais intensa e variada. Nos planos econômicos, o IPC desempenha um papel insubstituível na formulação da política de preços e rendimentos. Nomeadamente, o indicador taxa de inflação média dos últimos doze meses desempenha um importantíssimo papel enquanto referencial habitualmente utilizado no âmbito das negociações salariais enquadradas por contratações coletivas.

O IPC é também muitas vezes indispensável na avaliação da eficácia das políticas económicas seguidas. Por exemplo, a avaliação da competitividade da produção nacional face aos bens de origem externa não pode ser feita sem o auxílio do IPC. No âmbito da política de Segurança Social, o IPC é muitas vezes utilizado na formulação de medidas destinadas a favorecer grupos sociais delimitados. Por outro lado, o IPC é frequentemente utilizado a nível jurídico, em contratos públicos e privados, como

indexante das obrigações assumidas pelas partes contratantes.

4.REDES MAPAS AUTO ORGANIZÁVEIS (SOM)

Um mapa de Kohonen é uma forma de rede neuronal investigada por Kohonen [2] nos anos 1980. Essas redes tem forte inspiração neurofisiológica, as mesmas são baseadas no mapa topológico presente no córtex cerebral, observa-se que os neurónios estão espacialmente ordenados dentro dessas áreas, os que estão próximos tendem a responder a padrões ou estímulos semelhantes, esse processo foi inspirado nas funcionalidades inerentes a subáreas do córtex: auditivo, visual, lobo frontal. A rede SOM é uma rede biologicamente plausível, ou seja, apresenta respaldo biológico.

O propósito do mapa de kohonen é agrupar dados de entrada em diversos grupos (clusters) não conhecidos previamente. Um mapa de kohonen tem duas camadas : uma de entrada e outra de agrupamento, que serve como camada de saída. Cada nó de entrada é conectado a todo nó da camada de agrupamento e, tipicamente, os nós na camada de agrupamento são organizados em forma de grade, Figura 2.

Figura 2 – Rede de Kohonen em arranjo bidimensional

O treinamento da rede de kohonen é o seguinte: iniciam-se todos os pesos com pequenos valores aleatórios, estabelece-se a taxa de aprendizado η com um pequeno valor positivo. Um vetor de entrada é apresentado à camada de entrada do mapa. Esta camada alimenta os dados de entrada para a camada de agrupamento. O neurônio na camada de agrupamento que melhor combine com os dados de entrada é declarado vencedor. Este

neurônio fornece a classificação de saída do mapa e também tem seus pesos atualizados. Para determinar o neurônio vencedor, seus pesos são tratados como um vetor, que será comparado ao vetor de entrada, o neurônio cujo vetor de pesos for o mais próximo do vetor de entrada será o vencedor. A função e ativação da Rede de Kohonen é baseada na medida de distância euclidiana equação 1.

$$\gamma_j = \sum_{i=1}^n \|x_i - w_{\mu}\| \quad (1)$$

Onde n é o número de neurônios na camada de entrada e consequentemente, o número de elementos do vetor de entrada. O neurônio para o qual a distância euclidiana for menor será o vencedor. Nesse processo, apenas o neurônio vencedor e seus vizinhos dentro de um certo raio ou área de vizinhança atualizam

seus pesos. Durante o treinamento, a taxa de aprendizado e o raio da vizinhança são continuamente decrementados. A equação 2 ilustra como são atualizados os pesos da unidade vencedora e daqueles situados na sua vizinhança.

$$w_{\mu}(t+1) = \begin{cases} w_{\mu}(t) + \eta(t)(x_i(t) - w_{\mu}(t)), & \text{se } j \in \Lambda(t) \\ w_{\mu}(t), & \text{caso contrário} \end{cases} \quad (2)$$

Onde $w_{ij}(t)$ é o peso da conexão entre o elemento de entrada $x_i(t)$ e o neurônio j , $\eta(t)$ é taxa de aprendizado e Λ é a vizinhança do neurônio vencedor, tudo isso no instante de tempo t .

O treinamento da Rede SOM ocorre em duas fases : fase de ordenação e fase de convergência. Durante a primeira, ocorre a ordenação topológica dos vetores de pesos que são inicialmente orientados de forma aleatória (com pesos iniciais aleatórios). Nessa fase , o treinamento busca agrupar os neurônios do mapa topológico em agrupamentos, de modo a refletir a distribuição dos padrões de entrada. Dessa forma, a rede descobre quantos grupos ela deve identificar e suas posições relativas no mapa. Na segunda fase é feito um ajuste fino do mapa, sendo utilizada uma taxa de aprendizado baixa da ordem de 0.01 ou menos, e o raio da vizinhança envolve um ou nenhum vizinho, essa fase sofisticada o mapeamento realizado no estágio anterior, aprimorando o agrupamento realizado [4] [3].

A seguir apresentam-se dois modelos de Redes Neurais Recorrentes que foram utilizadas em conjunto com a Rede SOM no processo de descoberta de conhecimento em base de dados (KDD).

5. CLASSIFICADORES NEURAIS RECORRENTES NÃO LINEARES

A predição de séries temporais vem sendo realizada com o uso dos tradicionais modelos Auto Regressivos e/ou médias móveis (AR e

ARMA) . Tratam-se de modelos paramétricos através dos quais a previsão das futuras observações é obtida a partir da combinação linear de valores passados e, quando for o caso, com as componentes de ruído da série. Os sistemas dinâmicos são não lineares, dessa forma nessas aplicações devem-se escolher modelos com não linearidade pelo fato de produzirem certos regimes dinâmicos que os lineares não conseguem representar [5].

As RNAs (Redes Neurais Artificiais) são um exemplo bem difundido de representação não linear que atualmente são cada vez mais usadas para predição de séries temporais, devido à sua capacidade de agregar conhecimento em sua estrutura, a partir de exemplos, um preditor baseado em rede neural tem condições de estimar o comportamento futuro de uma série temporal apenas a partir de suas amostras passadas. Um modelo de predição deste tipo é chamado de não paramétrico, uma vez que não há necessidade de se conhecer os parâmetros do processo que gera o sinal. O modelo do processo é estimado através de um algoritmo de aprendizagem onde os exemplos são apresentados à rede neural e seus pesos são atualizados de acordo com o erro de predição. A seguir apresentam-se as representações não lineares baseadas em RNAs utilizadas na presente pesquisa.

5.1. REDE RECORRENTE COM SAÍDA REALIMENTADA – REDE ARX

Na presente seção apresenta-se uma

arquitetura de Rede Neural Recorrente (RNAR) com a saída realimentada, Figura 3. Essa RNAR foi baseada no modelo ARX [5], que nada mais é do que uma rede MLP cuja entrada consiste da própria saída realimentada com atrasos. Esta rede neural é equivalente ao Modelo **ARX “Autoregressive with exogenous inputs”**, dado pela equação (5). Onde $x(n)$ é a entrada do sistema e $y(n)$ a

saída, em que a função $f(\cdot)$ é uma função não linear, geralmente desconhecida e $x(n)$ e $y(n)$ correspondem à entrada e saída no tempo n , enquanto $dy > 0$, é a ordem da memória de entrada. Quando esta função é aproximada por uma rede “perceptron” de múltiplas camadas, a topologia resultante é chamada rede recorrente **ARX**, sendo um caso particular da rede apresentada na secção 5.2.

$$y_n = f[a_1 y(n-1) + a_2 y(n-dy) + x_n] \quad (3)$$

Para as deduções dos modelos apresentados nessa secção e nas posteriores, considere que A seja o número de unidades da camada de entrada, conforme determinado pelo comprimento dos vetores de entrada de treinamento, C o número de unidades da camada de saída e B o número de unidades da camada oculta. As camadas de entrada e oculta têm cada uma, uma unidade extra usada como limite, portanto as unidades dessas camadas às vezes serão indexadas pelos intervalos $(0, \dots, A)$ e $(0, \dots, B)$. Denotam-se os níveis de ativação das unidades da camada de entrada por x_j da camada oculta por h_j e da camada de saída por o_j . Os pesos que conectam a camada de

entrada a camada oculta são denotados por w_{1ij} , onde i indexa as unidades de entrada e j , as unidades ocultas. Da mesma forma, os pesos que conectam a camada oculta à camada de saída são denotados por w_{2ij} com i indexando as unidades ocultas e j as unidades de saída.

As modificações realizadas no “backpropagation” para aproximação do modelo ARX pelo modelo neural recorrente foram as seguintes: as saídas na camada intermediária são agora dadas pela equação (4). Observa-se a inclusão das contribuições das recorrências dadas pelo somatório de $m=0$ a C para os termos $o_m(t-1)$ na mesma.

$$h_k(t) = \frac{1}{1 + e^{-\sum_{m=0}^C o_m(t-1) \cdot w_{3mk}(t) + \sum_{j=0}^A x_{nj}(t) \cdot w_{1jk}(t)}} \quad (4)$$

para $t > 0$, $h_0 = 1$, que é o valor do bias. O termo $o_m(t-1)$ refere-se a cada saída que é realimentada e C é o número de neurônios da camada de saída, para $t=0$ o seu valor é $o_m(t-1) = 0$. A atualização dos pesos das realimentações são dadas pelas equações 5

e 8. Após deduções matemática chegou-se a conclusão que as atualizações dos pesos da camada de entrada para a intermediária e da camada intermediária para a de saída são dadas pelas equações (6) e (7).

$$\Delta w_{3mk}(t+1) = \Delta w_{3mk}(t) - \eta \frac{\partial \varepsilon_n}{\partial w_{3mk}(t)} \quad (5)$$

Figura 3 – Rede ARX aproximação do Modelo ARX

$$\Delta w_{1,k}(t+1) = \Delta w_{1,k}(t) + \eta [(1-h_k)h_k \sum_{m=0}^{\infty} (y_{nm} - o_m)(1-o_m) \cdot o_m \cdot w_{2,km}] x_{nk} \quad (6)$$

$$\Delta w_{2,km}(t+1) = \Delta w_{2,km}(t) + \eta [(y_{nm} - o_m)(1-o_m) o_m] h_k \quad (7)$$

$$\Delta w_{3,mk}(t+1) = \Delta w_{3,mk}(t) + \eta (1-h_k(t)) h_k(t) \sum_{m=0}^{\infty} (y_{nm}(t) - o_m(t))(1-o_m(t)) \cdot o_m(t) \cdot o_m(t-1) w_{2,km}(t) \quad (8)$$

5.2. REDE RECORRENTE NARX

O modelo estudado nessa seção é o modelo NARX que nada mais é que uma rede MLP cuja entrada consiste na própria saída realimentada com atrasos de tempo e uma entrada exógena, também com atrasos. Nessa arquitetura, as saídas estimadas da

rede são introduzidas novamente às entradas, essa representação é equivalente ao modelo estatístico **NARX** “Nonlinear Autoregressive model with eXogenous input”, dado pela equação (9).

$$y_n = f[ay(n-1) + \dots + y(n-dy), u(n-1) + \dots + u(n-du+1)] \quad (9)$$

Em que a função $f(\cdot)$ é uma função não linear, geralmente desconhecida e $u(n)$ e $y(n)$ correspondem à entrada e saída no tempo n , enquanto $du > 0$ e $dy > 0$, $du \leq dy$ são as ordens da memória de entrada e memória de saída. Quando esta função é aproximada por uma rede “perceptron” de múltiplas camadas,

a topologia resultante é chamada rede recorrente NARX. A Figura 4 ilustra uma rede **NARX**, com uma camada oculta e um laço de realimentação global. Na presente pesquisa considerou-se a utilização da rede NARX no modo de identificação paralelo.

Figura 4 – rede NARX com du entradas e dy atrasos na saída

Fonte: Hecht-Nielsen (1,987)

As modificações realizadas no “backpropagation” original foram as seguintes: as saídas na camada intermediária são agora dadas pela equação (10). Observa-se a inclusão das contribuições das recorrências dadas pelo somatório de $m=0$ a

C para os termos $om(t-1)$ que correspondem a própria saída realimentada com atrasos e uma várias entradas exógenas com atrasos, representados pelos termos $X_{nj}(t)+X_{nj}(t-1)+\dots+X_{nj}(t-du)$ na equação (10).

$$hk(t) = \frac{1}{1 + e^{-\sum_{m=0}^C om(t-1) \cdot w3(t)_{mx} + \sum_{j=0}^A [X_{yj}(t) + X_{yj}(t-1) + \dots + X_{yj}(t-du)] w1(t)_j}} \quad (10)$$

Considerando-se a modificação da função $hk(t)$ para a rede NARX, após deduções matemáticas chegou-se a conclusão que a atualização dos pesos da camada intermediária para a de saída é dada pela

equação (7), a atualização dos pesos das realimentações são calculados pela equação (8) e finalmente a atualização dos pesos da camada de entrada para a intermediária é dada pela equação (11).

$$\Delta w1_{jx}(t+1) = \Delta w1_{jx}(t) + \eta [(1 - h_x) h_x \cdot \sum_{m=0}^C (y_{mx} - o_{mx}) (1 - o_{mx}) \cdot o_{mx} \cdot w2_{mx} \cdot \{x_{xj}(t) + x_{xj}(t-1) + \dots + x_{xj}(t-du)\}] \quad (11)$$

6. METODOLOGIA

O primeiro passo foi obter a série temporal do IPC do Sistema Gerenciador de Séries Temporais do Banco Central do Brasil [14], para o emprego de uma rede neural artificial como preditor, é necessário definir quais valores passados da série serão utilizados na definição da entrada da rede neural. Feito isso, o problema de síntese do preditor se transforma em um problema de treinamento supervisionado, onde o que se deseja é obter um mapeamento multidimensional não linear de entrada-saída.

O segundo passo foi o tratamento dos dados brutos (seleção de um subconjunto de dados alvos), pré-processamento (limpeza de dados, eliminação de ruídos) e transformação (redução dos dados, normalização), todas essas etapas influenciam na eficiência do Algoritmo de Mineração de Dados escolhido (Classificação) para a busca de padrões, que são necessários para interpretação e visualização dos dados. A seguir descrevem-se como as etapas de pré-processamento, transformação e normalização de dados foram realizadas.

Utilizou-se uma Rede SOM para realizar a etapa de pré-processamento (eliminação e outliers) após a etapa de seleção dos dados

brutos, para que os dados sejam filtrados eliminando dados que deveriam seguir um padrão esperado, mas não o fazem. A saída da Rede SOM classificará os dados entre bons e ruins (outliers) para a mineração de dados. Após essa etapa realiza-se a transformação dos dados por meio da redução e normalização “z-score”. Na etapa de mineração de dados utilizaram-se dois classificadores recorrentes não lineares recorrentes ARX e NARX apresentados nas seções 5.1 e 5.2 respectivamente, selecionando-se o melhor preditor. Após isso, visualizam-se os dados e interpretam-se os resultados gerando o conhecimento a cerca da predição da série temporal estudada. A execução dessas etapas é feito na seção 7.

7. RESULTADOS DE SIMULAÇÕES E DISCUSSÕES

O Gráfico 1 mostra o comportamento da série temporal do IPC utilizada nos experimentos, a mesma apresenta periodicidade mensal e se estende do período de jan-1990 a dez-2010. Percebe-se que a mesma apresenta comportamento bastante diferenciado nos períodos de jan-1990 a jun-1994 e de jul-1994 a dez 2010. No segundo período os dados apresentam-se mais homogêneos, visto que

em 1994 surgiu o plano Real e com isso veio estabilização da economia com índices inflacionários mais estáveis. A segunda etapa realizada foi selecionar um conjunto de dados alvo relevantes da série e pré-processar os mesmos visando eliminar “outliers”, para isso utilizou-se clusterização por meio de uma Rede Neural de Kohonen, o Gráfico 2 mostra os resultados desse processo, considerando os dados do período de jan-1990 a dez-2010. A Rede de Kohonen dividiu os dados brutos do IPC em duas classes, os valores de IPC maiores que 28.41 foram classificados como classe 2 e os menores que esse valor como classe 1.

Com base no mapa de Kohonen foram eliminados os “outliers”: valores de IPC

maiores que 28.41 que correspondiam a valores antes de jul-1994 e, portanto anteriores ao período de estabilização econômica. Após essa etapa, partiu-se para a normalização dos dados para que o conjunto inteiro fosse padronizado de modo a ter uma determinada propriedade, para isso utilizou-se a normalização “z-score” dada pela equação 1, onde “mean x” é a média dos valores do atributo x e “stand_dev” x o desvio padrão, a transformação dada pela equação cria uma nova variável xnew que tem uma média 0 e desvio padrão 1, com isso os valores foram escalonados com objetivo de minimizar problemas oriundos do uso de unidades de dispersão distintas entre variáveis.

$$x_{new} = \frac{x_{old} - mean_x}{stand_dev_x} \quad (1)$$

Gráfico 1 – Série temporal do IPC (jan. de 1990 a dez. de 2010)

Fonte: Banco Central do Brasil

Gráfico 2 – Clusterização com Rede Neural de Kohonen para Remoção de outliers.

A terceira etapa foi definir quais valores passados foram utilizados na predição, consideraram-se os dados de janeiro de 1990 a dezembro de 2010, sendo que se separou 2/3 para treino e 1/3 para testes, considerando que foram eliminados dados com valores de IPC acima de 28.41 (outliers). As etapas finais do processo de Descoberta de Conhecimento em Base de Dados (KDD): Mineração de Dados, Interpretação e Visualização dos Dados são mostradas a seguir.

Realizaram-se vários treinamentos para várias configurações das Redes ARX e NARX. Inicialmente considerou-se os seguintes parâmetros um (1) neurônio na camada de entrada, quatro (6) na intermediária e um (1) na de saída, taxa de aprendizagem 0.01 e 30000 épocas sendo as funções de ativação de todos os neurônios do tipo sigmoide. Obtiveram-se os resultados da Tabela 2, percebe-se que os valores não foram satisfatórios, pois a Rede ARX indicou MAPE 29.02% de e a NARX 22.81%.

Tabela 2 – Valores obtidos para o teste da Rede ARX e NARX

Desejado	Obtido ARX	Obtido NARX	EMQ ARX	EMQ NARX	Predição
0.03477	0.0275132	0.0403254	2.63303e-005	1.5431e-005	4 meses
0.03250	0.0305615	0.0403249	1.87894e-006	3.06144e-005	7 meses
0.03136	0.027618	0.0403244	7.00123e-006	4.01803e-005	1 ano
0.03288	0.0304439	0.0403239	2.96727e-006	2.7706e-005	3 anos

Os melhores resultados foram obtidos utilizando-se os seguintes parâmetros : um (1) neurônio na camada de entrada, quatro (9) na intermediária e um (1) na de saída, taxa de aprendizagem 0.01 e 30000 épocas, com funções de ativação sigmoide para todos os neurônios. O Gráfico 3 mostra os resultados de validação para o modelo ARX. A Tabela 3 apresenta os resultados de teste obtidos, para quatro, sete meses, um e três anos respectivamente, percebeu-se que pelos resultados o MAPE situou-se em torno de 8,52%.

Após o teste da Rede ARX realizou-se o treinamento usando-se uma Rede NARX. Os parâmetros utilizados foram um (1) neurônio na camada de entrada, oito (8) na intermediária e um (1) na de saída, taxa de aprendizagem 0.01 e 90000 épocas, as

funções de ativação de todos os neurônios foram sigmoide. O Gráfico 4 ilustra a validação para a Rede NARX. A Tabela 3 apresenta os valores de predição obtidos para os mesmos períodos utilizados para a Rede ARX o MAPE situou-se em torno de 3.20%. Realizaram-se estudos de Predição para períodos acima de quatro anos, entretanto os resultados apresentados para a Rede NARX e ARX não foram satisfatórios. Contudo as predições de até 4 anos a frente foram boas para ambas as Redes ARX e NARX, sendo que os resultados para a arquitetura NARX são superiores com valores de MAPE (Erro Percentual Médio Absoluto) de até 3.20%. A rede NARX apresentou, para quatro anos valor de predição de 0.0319129 para um valor desejado de 0.03288, com Erro médio Quadrático (EMQ) de 4.67642e-007 e MAPE de 2.94%.

Gráfico 3 - Validação do Modelo Rede ARX

Gráfico 4 - Validação do Modelo Rede NARX

Tabela 3 – Valores obtidos para o teste da Rede ARX e NARX

Desejado	Obtido ARX	Obtido NARX	EMQ ARX	EMQ NARX	Predição
0.03477	0.031836	0.0322401	4.30427e-006	3.20008e-006	4 meses
0.03250	0.0318389	0.0322396	2.18528e-007	3.39162e-008	7 meses
0.03288	0.0318418	0.032239	5.38902e-007	3.86287e-007	1 ano
0.02643	0.0318448	0.0322384	1.46598e-005	2.05845e-007	3 anos

8. CONCLUSÕES

De acordo com os resultados apresentados por simulação percebe-se que a Rede de Kohonen representaram um papel importante no que diz respeito à clusterização dos dados, a série temporal do Índice de Preços ao Consumidor, por ser uma série com dados econômicos, apresenta uma peculiaridade específica no que diz respeito à amplitude de valores para diferentes períodos antes e depois da estabilização econômica, ou seja, antes de meados de 1994 e depois. Com a eliminação dos “outliers” diminui-se a amplitude de valores usados no treinamento para valores mais condizentes com os padrões desejados.

Após os experimentos, chegou-se a conclusão que a Rede NARX, além de apresentar bom desempenho para o problema de predição do IPC, possuiu convergência mais rápida e capacidade de generalização melhor que a ARX. Isto ocorre porque o vetor de entrada dos modelos NARX

são construídos por meio de uma linha de atraso com derivação deslizada sobre o sinal de entrada, junto com uma linha de atraso com derivação formada pelas realimentações do sinal de saída da rede [12].

A rede ARX não mostrou uma boa capacidade de generalização nas simulações tendo sido descartada dos estudos de predições a passos futuros. O Sistema híbrido estudado aplicado a descoberta de conhecimento em base de dados (KDD) constitui-se uma ferramenta útil para a predição da série temporal do Índice de Preços ao Consumidor, para períodos de até quatro anos a frente, contribuindo assim para o planejamento governamental a curto e médio prazo. A metodologia utilizada, ainda encontra-se em fase de testes e pode ser utilizada em trabalhos futuros para predição de série mais complexas, como é o caso de séries caóticas.

REFERENCIAS

- [1] Fayyad, Usama; Piatetski-Shapiro, Gregory; Smyth, Padhraic (1996). The KDD Process for Extracting Useful Knowledge from Volumes of Data. In: Communications of the ACM, pp.27-34, Nov.1996
- [2] Kohonen, T. Self-Organization and associative memory. Springer-Verlag, Berlin, 3 edition, 1989.
- [3] Hecht-Nielsen, R. Counterpropagation networks. In Proceedings of the IEEE . International Conference on Neural Networks. IEEE Computer Press, 1987.
- [4] B. Antônio de Pádua, C. André Ponce de Leon F., Ludemir, Teresa Bernarda. Redes Neurais Artificiais. Teorias e Aplicações. 2 ED. Rio Janeiro : LTC, 2007.
- [5] Aguirre, L.A., Introdução à Identificação de Sistemas, Editora UFMG, terceira edição, 2007.
- [6] Goebel, M.; Gruenwald L. “A survey of data mining and knowledge discovery software tools”. In: ACM SIGKDD Explorations Newsletter. 1. ed. vol. 1 1999.
- [7] Lam M., Neural network techniques for financial performance prediction : integrating fundamental and technical analysis. Decision Support Systems 37, 2004; 567-581.
- [8] Yao J.T., Towards a better forecasting model for economic indices. In: Proceedings of the 6th Joint Conference on Information Science, 2002; 299-303.
- [9] José M. Menezes Jr. & Guilherme A. Barreto (2006), A New Look at Nonlinear Time Series Prediction with NARX Recurrent Neural Network, IX Brazilian Neural Networks Symposium (SBRN'2006), Ribeirão Preto-SP.
- [10] LIN, T. et al. Learning long-term dependencies in NARX recurrent neural networks. IEEE Transactions on Neural Networks, v. 7, n. 6, p. 1424_1438, 1996.
- [11] James Malone , Kenneth McGarry , Stefan Wermter, Chris Bowerman. Data mining using rule extraction from Kohonen self-organising Maps. Neural Comput & Applic (2005) 15: 9–17 DOI 10.1007/s00521-005-0002-1. Springer-Verlag London 2005.
- [12] Bartkowiak, Anna. Zdziarek, Joanna. Evelpidou, Niki. Pejaś, Jerzy. Piegat, Andrzej. Choosing representative data items: Kohonen, Neural Gas or Mixture Model Book Title: Enhanced Methods in Computer Security, Biometric and Artificial Intelligence Systems. Springer US. Isbn: 978-0-387-23484-7. Computer Science. Pag. 337-304. Doi: 10.1007/0-387-23484-5_33, 2005.
- [13] S. Crone . Stepwise Selection of Artificial Neural Network Models for Time Series Prediction, Journal of Intelligent Systems, Vol. 14, No. 2-3, 2005, pp. 99-122
- [14] Banco Central do Brasil. Sistema Gerenciador de Séries Temporais do Banco Central do Brasil. Disponível em: <<http://www4.bcb.gov.br/pec/series/port/aviso.asp>>. Acesso em: 05 dez. 2011.
- [15] H. Kantz, T. Schreiber, Nonlinear Time Series Analysis, 2nd ed., Cambridge University Press, 2006.

Capítulo 3

O AHP COMO UM MODELO MATEMÁTICO: UMA ANÁLISE DE SENSIBILIDADE SIMPLES

Mylene Cristina Rezende Pacheco

Fernando Luiz Goldman

Resumo: Somos todos fundamentalmente decisores e os decisores hoje se percebem imersos em contextos de incerteza knightiniana e racionalidade limitada não podendo ser maximizadores perfeitos. Muita ênfase vem sendo dada na literatura sobre empresas à decisão em si. Os métodos clássicos de Pesquisa Operacional são utilizados quando apenas uma variável-objetivo é relevante ou quando todos os objetivos podem ser representados pela mesma unidade de medida, o que nem sempre ocorre. Assim, os diferentes modelos da Análise de Decisão Multi Critério são importantes ferramentas da Pesquisa Operacional e entre eles se destaca o Processo Analítico Hierárquico, um modelo aditivo compensatório. No entanto, usualmente, pouca atenção é dada aos benefícios da análise de sensibilidade ao se utilizar o AHP, daí ser comum a referência a ele como um método e não como um modelo que realmente é. Este artigo tem como objetivo geral apresentar um exemplo da análise de sensibilidade na resolução de um problema real de decisão usando o modelo AHP, com a utilização de uma planilha eletrônica não dedicada. Este artigo descreve uma pesquisa de caráter teórico-conceitual, que se caracteriza como qualitativa e exploratória, trazendo em seu desenvolvimento a aplicação de um modelo matemático já existente e discussões pertinentes a sua utilização na análise de sensibilidade de uma tomada de decisão com auxílio de uma planilha eletrônica. Conclui-se pela viabilidade da utilização do AHP – associado a uma planilha eletrônica de fácil acesso, sem depender de nenhum software proprietário – não só como um método composto de passos bem definidos, que levam a uma tomada de decisão, mas principalmente como um modelo matemático, que propicia variações e análise de sensibilidade, fácil de aplicar e capaz de propiciar ao decisor, mais do que verificar a robustez de sua decisão, aprender e aperfeiçoar seu processo decisório.

Palavras chave: AHP; Multicritério; Processo Decisório; Sensibilidade; Consistência.

*Artigo originalmente publicado nos anais do Simpósio de Excelência em Gestão e Tecnologia 2016

1. INTRODUÇÃO

Como destaca Saaty (2008, p. 83), “somos todos fundamentalmente decisores”. Hoje, os decisores se percebem imersos em contextos de incerteza knightiniana e racionalidade limitada (GOLDMAN, 2013, p. 33), não podendo ser maximizadores perfeitos, pois “em função de sua experiência pregressa e de limitações cognitivas, se deixam guiar pelas rotinas construídas” (CORAZZA; FRACALANZA, 2004, p. 129-130). Para Bazerman e Moore (2010, p. 7) ainda há outros limitadores (tempo, custo ou a própria percepção e inteligência dos próprios decisores).

Muita ênfase vem sendo dada na literatura sobre empresas à decisão em si. Almeida (2011, p. 16) afirma que os “métodos clássicos de pesquisa operacional” são utilizados quando “apenas uma variável-objetivo é relevante” ou quando todos os objetivos podem ser representados pela mesma unidade de medida, o que nem sempre ocorre. Assim, os diferentes modelos da Análise de Decisão Multi Critério (MCDA) são importantes ferramentas da Pesquisa Operacional (PO) e entre eles vem se destacando o Processo Analítico Hierárquico (AHP – do inglês Analytic Hierarchy Process), um modelo aditivo compensatório (PACHECO, 2015).

O AHP foi desenvolvido por Thomas Saaty na década de 1970 nos EUA (SHIMIZU; 2006, p. 278). Foi pioneiro na Escola Americana no que diz respeito aos problemas de decisão que envolvem múltiplos critérios (REIS; LÖBLER, 2012, p.400), permitindo combinar dados qualitativos e quantitativos num só problema (ABREU; et al, 2000, p.) e por esses motivos tem tido aplicação em várias áreas.

No entanto, usualmente, pouca atenção é dada aos benefícios da análise de sensibilidade ao se utilizar o AHP, daí ser comum a referência a ele como um método (um passo-a-passo) e não como um modelo que realmente é.

Este capítulo tem como objetivo geral apresentar um exemplo da análise de sensibilidade na resolução de um problema real de decisão usando o modelo AHP, com a utilização de uma planilha eletrônica não dedicada e facilmente disponível.

O desenvolvimento do capítulo faz-se em cinco seções, incluída esta Introdução. A

segunda, **Aspectos metodológicos**, qualifica a pesquisa descrita quanto aos métodos e meios empregados. A terceira, **Exemplo de aplicação do AHP**, mostra a aplicação do modelo estudado em um problema real. A quarta, **Resultados e discussão**, apresenta e analisa os resultados obtidos. A quinta, **Considerações finais**, traz limitações da pesquisa, sugestões para pesquisas futuras e uma conclusão sucinta.

2. ASPECTOS METODOLÓGICOS

Este artigo descreve uma pesquisa de caráter teórico-conceitual, que se caracteriza como qualitativa e exploratória, trazendo em seu desenvolvimento a aplicação de um modelo matemático já existente e discussões pertinentes a sua utilização na Análise de Sensibilidade de uma tomada de decisão com auxílio de uma planilha eletrônica facilmente disponível.

Todos os cálculos foram realizados com o software MICROSOFT® EXCEL®, que é suficiente e de fácil utilização para a aplicação do modelo AHP. Para o exemplo de aplicação descrita neste artigo foi usada uma única planilha, com células contendo fórmulas simples, de acordo o passo-a-passo proposto por Saaty (1990), largamente conhecido, cuja descrição detalhada, além de ser bastante documentada, fugiria ao escopo desse artigo.

Como é bem sabido, o processo de normalização, usado neste artigo, de cálculo bem mais prático, possui resultados extremamente próximos aos obtidos com o cálculo algébrico do autovetor matricial, sendo perfeitamente adequado ao tipo de precisão requerido em problemas do tipo do discutido aqui.

3. EXEMPLO DE APLICAÇÃO DO AHP

Para análise de sensibilidade com a utilização de uma planilha eletrônica não dedicada será apresentada a resolução de um problema fictício de decisão usando o modelo AHP, como segue.

Uma empresa fictícia, denominada doravante pelo nome XPTO, utiliza o modelo AHP para seleção dos fornecedores de determinado material. Este artigo exemplificará a utilização do modelo AHP para escolha do material, considerando os seguintes critérios:

Qualidade, Preço, Condições de pagamento, Frete e Prazo de entrega.

A decomposição hierárquica proposta para o problema descrito acima pode ser visualizada

através de uma árvore de decisão da Figura 1.

Figura 1: Árvore de decisão do problema descrito

A partir de entrevistas, com o representante do setor de compras da XPTO, por exemplo, é possível coletar os dados pertinentes à percepção do decisor em relação aos critérios considerados na compra.

4. RESULTADOS E DISCUSSÃO

O número de julgamentos necessário à comparação de n critérios, como é bem sabido, é dado pela expressão $n(n-1)/2$.

Como no caso presente teremos $n=5$ critérios, isto faz com que o número de julgamentos para comparar par a par os critérios seja igual a 10.

Após a primeira percepção descrita pelo decisor, deve ser utilizada a Escala Fundamental de Saaty, no Quadro 1, para comparar, par a par, os critérios utilizados para esta seleção e o resultado de tais comparações pode ser visto na Tabela 1.

Quadro 1: Escala Fundamental de Saaty

Intensidade de Importância	Definição	Explicação
1	Mesma Importância	As duas atividades contribuem igualmente para o objetivo.
3	Importância pequena de uma sobre a outra	A experiência e o julgamento favorecem levemente uma atividade em relação à outra.
5	Importância grande ou essencial	A experiência e o julgamento favorecem fortemente uma atividade em relação à outra.
7	Importância muito grande ou demonstrada	Uma atividade é muito fortemente favorecida em relação à outra; sua dominação de importância é demonstrada na prática.
9	Importância absoluta	A evidência favorece uma atividade em relação à outra com o mais alto grau de certeza.
2,4,6,8	Valores intermediários entre valores adjacentes	Quando se procura uma condição de compromisso entre duas definições.

Fonte: Adaptado de Saaty (1990)

Tabela 1: Variações de julgamentos para Análise de Sensibilidade

Comparações	Valor original	Variações
Qualidade-Preço	1	1
Qualidade-Forma de PGTO	3	3
Qualidade-Frete	3	3
Qualidade-Prazo	5	5
Preço-Forma de PGTO	5	5
Preço-Frete	3	3
Preço-Prazo	5	5
Forma de PGTO-Frete	1	1
Forma de PGTO-Prazo	2	2
Frete-Prazo	1	1

A partir dos julgamentos iniciais do decisor é montada uma primeira Matriz de Comparações Pareadas dos critérios, como

mostrado na Tabela 2. Deve ser observado que os cálculos devem ser feitos a partir da coluna "Variações".

Tabela 2: Matriz de comparação pareada dos critérios

Crítérios	Qualidade	Preço	Forma de PGTO	Frete	Prazo
Qualidade	1	1	3	3	5
Preço	1	1	5	3	5
Forma de PGTO	1/3	1/5	1	1	2
Frete	1/3	1/3	1	1	1
Prazo	1/5	1/5	1/2	1	1
Soma	2,8667	2,7333	10,5000	9,0000	14,0000

Os pesos atribuídos para cada critério, obtidos através do processo de normalização estão mostrados na Tabela 3.

Tabela 3: Normalização e obtenção dos pesos

Crítérios	Qualidade	Preço	Forma de PGTO	Frete	Prazo	Média	Peso (%)
Qualidade	0,3488	0,3659	0,2857	0,3333	0,3571	0,3382	33,82%
Preço	0,3488	0,3659	0,4762	0,3333	0,3571	0,3763	37,63%
Forma de PGTO	0,1163	0,0732	0,0952	0,1111	0,1429	0,1077	10,77%
Frete	0,1163	0,1220	0,0952	0,1111	0,0714	0,1032	10,32%
Prazo	0,0698	0,0732	0,0476	0,1111	0,0714	0,0746	7,46%
Soma	1	1	1	1	1	1	100,00%

Nota-se que os critérios Preço e Qualidade foram avaliados como tendo grande e equivalente importância, enquanto os demais obtiveram pesos de menor importância.

A Tabela 4 mostra os elementos da análise de Consistência.

Tabela 4: Vetor Soma e Vetor de Consistência

Critérios	Qualidade	Preço	Forma de PGTO	Frete	Prazo	Peso (%)	Vetor Soma	V. Consistência
Qualidade	1	1	3	3	5	33,82%	1,7203	5,0871
Preço	1	1	5	3	5	37,63%	1,9358	5,1447
Forma de PGTO	1/3	1/5	1	1	2	10,77%	0,5482	5,0881
Frete	1/3	1/3	1	1	1	10,32%	0,5237	5,0745
Prazo	1/5	1/5	1/2	1	1	7,46%	0,3746	5,0198
Soma	2,8667	2,7333	10,5000	9,0000	14,0000	1,0000	5,1026	

Os valores obtidos para a média aritmética da coluna do Vetor de Consistência (λ máx), o Índice de Consistência (CI), o Índice

Randômico (RI) e a Relação de Consistência (CR), estão registrados na Tabela 5, finalizando, o teste de consistência:

Tabela 5: Teste de Consistência

λ máx	CI	RI	CR
5,0829	0,0207	1,1200	0,0185

Através do teste de consistência realizado pôde-se constatar que a comparação pareada do decisor obteve uma Relação de Consistência de 0,0185, que atende ao requisito de Saaty de que uma inconsistência aceitável deve ser menor ou igual a 0,1.

Ao comparar de forma pareada, cada alternativa de fornecedor em relação ao critério Qualidade, segundo o julgamento do hipotético decisor, foram obtidas as seguintes matrizes:

Tabela 6: Matriz de comparação pareada dos fornecedores: Critério Qualidade

Qualidade	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4
Alternativa 1	1	1/2	1	3
Alternativa 2	2	1	1	3
Alternativa 3	1	1	1	9
Alternativa 4	1/3	1/3	1/9	1
Soma	4,3333	2,8333	3,1111	16,0000

Tabela 7: Normalização e desempenho dos fornecedores: Critério Qualidade

Qualidade	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Média	Peso (%)
Alternativa 1	0,2308	0,1765	0,3214	0,1875	0,2290	22,90%
Alternativa 2	0,4615	0,3529	0,3214	0,1875	0,3309	33,09%
Alternativa 3	0,2308	0,3529	0,3214	0,5625	0,3669	36,69%
Alternativa 4	0,0769	0,1176	0,0357	0,0625	0,0732	7,32%
Soma	1	1	1	1	1	100,00%

O novo teste de consistência pode ser visualizado nas Tabelas 8 e 9.

Tabela 8: Vetor Soma e Vetor de Consistência: Critério Qualidade

Qualidade	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Peso (%)	Vetor Soma	V. Consistência
Alternativa 1	1	1/2	1	3	22,90%	0,9810	4,2829
Alternativa 2	2	1	1	3	33,09%	1,3754	4,1572
Alternativa 3	1	1	1	9	36,69%	1,5856	4,3214
Alternativa 4	1/3	1/3	1/9	1	7,32%	0,3006	4,1067
Soma	4,3333	2,8333	3,1111	16,0000	100,00%	4,2426	

Tabela 9: Teste de Consistência: Critério Qualidade

λ máx	CI	RI	CR
4,2171	0,0724	0,9000	0,0804

A CR se mostrou aceitável.

O desempenho dos fornecedores em relação ao Critério Preço, segundo o julgamento do

hipotético decisor, pode ser verificado na Tabela 10 e 11 e a respectiva Análise de Consistência nas Tabelas 12 e 13.

Tabela 10: Matriz de comparação pareada dos fornecedores: Critério Preço

Preço	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4
Alternativa 1	1	7	1/5	7
Alternativa 2	1/7	1	1/9	1
Alternativa 3	5	9	1	9
Alternativa 4	1/7	1	1/9	1
Soma	6,2857	18,0000	1,4222	18,0000

Tabela 11: Normalização e desempenho dos fornecedores: Critério Preço

Preço	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Média	Peso (%)
Alternativa 1	0,1591	0,3889	0,1406	0,3889	0,2694	26,94%
Alternativa 2	0,0227	0,0556	0,0781	0,0556	0,0530	5,30%
Alternativa 3	0,7955	0,5000	0,7031	0,5000	0,6246	62,46%
Alternativa 4	0,0227	0,0556	0,0781	0,0556	0,0530	5,30%
Soma	1,0000	1,0000	1	1	1	100,00%

Tabela 12: Vetor Soma e Vetor de Consistência: Critério Preço

Preço	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Peso (%)	Vetor Soma	V. Consistência
Alternativa 1	1	7	1/5	7	26,94%	1,1362	4,2178
Alternativa 2	1/7	1	1/9	1	5,30%	0,2139	4,0360
Alternativa 3	5	9	1	9	62,46%	2,9253	4,6832
Alternativa 4	1/7	1	1/9	1	5,30%	0,2139	4,0360
Soma	6,2857	18,0000	1,4222	18,0000	100,00%	4,4893	

Tabela 13: Teste de Consistência: Critério Preço

λ máx	CI	RI	CR
4,2432	0,0811	0,9000	0,0901

O teste de consistência para o desempenho dos quatro fornecedores perante o critério preço foi satisfatório obedecendo ao requisito: $0,0901 \leq 0,10$.

Para o critério Forma de PGTO, foram feitas as seguintes comparações:

Tabela 14: Matriz de comparação pareada dos fornecedores: Critério Forma de PGTO

Forma de PGTO	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4
Alternativa 1	1	5	3	8
Alternativa 2	1/5	1	1	1
Alternativa 3	1/3	1	1	1
Alternativa 4	1/8	1	1	1
Soma	1,6583	8,0000	6,0000	11,0000

Tabela 15: Normalização e desempenho dos fornecedores: Critério Forma de PGTO

Forma de PGTO	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Média	Peso (%)
Alternativa 1	0,6030	0,6250	0,5000	0,7273	0,6138	61,38%
Alternativa 2	0,1206	0,1250	0,1667	0,0909	0,1258	12,58%
Alternativa 3	0,2010	0,1250	0,1667	0,0909	0,1459	14,59%
Alternativa 4	0,0754	0,1250	0,1667	0,0909	0,1145	11,45%
Soma	1,0000	1,0000	1	1	1	100,00%

Tabela 16: Vetor Soma e Vetor de Consistência: Critério Forma de PGTO

Forma de PGTO	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Peso (%)	Vetor Soma	V. Consistência
Alternativa 1	1	5	3	8	61,38%	2,5964	4,2299
Alternativa 2	1/5	1	1	1	12,58%	0,5089	4,0458
Alternativa 3	1/3	1	1	1	14,59%	0,5908	4,0494
Alternativa 4	1/8	1	1	1	11,45%	0,4629	4,0433
Soma	1,6583	8,0000	6,0000	11,0000	100,00%	4,1590	

Tabela 17: Teste de Consistência: Critério Forma de PGTO

λ máx	CI	RI	CR
4,0921	0,0307	0,9000	0,0341

Para o critério Frete, foram feitas as seguintes comparações, segundo o julgamento do

hipotético decisor e as respectivas Análises de Consistência:

Tabela 18: Matriz de comparação pareada dos fornecedores: Critério Frete

Frete	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4
Alternativa 1	1	6	3	9
Alternativa 2	1/6	1	1/7	1
Alternativa 3	1/3	7	1	9
Alternativa 4	1/9	1	1/9	1
Soma	1,6111	15,0000	4,2540	20,0000

Tabela 19: Normalização e desempenho dos fornecedores: Critério Frete

Frete	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Média	Peso (%)
Alternativa 1	0,6207	0,4000	0,7052	0,4500	0,5440	54,40%
Alternativa 2	0,1034	0,0667	0,0336	0,0500	0,0634	6,34%
Alternativa 3	0,2069	0,4667	0,2351	0,4500	0,3397	33,97%
Alternativa 4	0,0690	0,0667	0,0261	0,0500	0,0529	5,29%
Soma	1,0000	1,0000	1	1	1	100,00%

Tabela 20: Vetor Soma e Vetor de Consistência: Critério Frete

Frete	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Peso (%)	Vetor Soma	V. Consistência
Alternativa 1	1	6	3	9	54,40%	2,4199	4,4486
Alternativa 2	1/6	1	1/7	1	6,34%	0,2555	4,0292
Alternativa 3	1/3	7	1	9	33,97%	1,4414	4,2437
Alternativa 4	1/9	1	1/9	1	5,29%	0,2145	4,0528
Soma	1,6111	15,0000	4,2540	20,0000	100,00%	4,3314	

Tabela 21: Teste de Consistência: Critério Frete

λ máx	CI	RI	CR
4,1935	0,0645	0,9000	0,0717

Para o critério Prazo de Entrega, foram feitas as seguintes comparações, segundo o julgamento do hipotético decisor:

Tabela 22: Matriz de comparação pareada dos fornecedores: Critério Prazo

Prazo	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4
Alternativa 1	1	7	1	7
Alternativa 2	1/7	1	1/5	2
Alternativa 3	1	5	1	5
Alternativa 4	1/7	1/2	1/5	1
Soma	2,2857	13,5000	2,4000	15,0000

Tabela 23: Normalização e desempenho dos fornecedores: Critério Prazo

Prazo	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Média	Peso (%)
Alternativa 1	0,4375	0,5185	0,4167	0,4667	0,4598	45,98%
Alternativa 2	0,0625	0,0741	0,0833	0,1333	0,0883	8,83%
Alternativa 3	0,4375	0,3704	0,4167	0,3333	0,3895	38,95%
Alternativa 4	0,0625	0,0370	0,0833	0,0667	0,0624	6,24%
Soma	1,0000	1,0000	1,0000	1,0000	1,0000	100,00%

Tabela 24: Vetor Soma e Vetor de Consistência: Critério Prazo

Prazo	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4	Peso (%)	Vetor Soma	V. Consistência
Alternativa 1	1	7	1	7	45,98%	1,9042	4,1410
Alternativa 2	1/7	1	1/5	2	8,83%	0,3567	4,0388
Alternativa 3	1	5	1	5	38,95%	1,6028	4,1153
Alternativa 4	1/7	1/2	1/5	1	6,24%	0,2501	4,0094
Soma	2,2857	13,5000	2,4000	15,0000	100,00%	4,1137	

Tabela 25: Teste de Consistência: Critério Prazo

λ máx	CI	RI	CR
4,0761	0,0254	0,9000	0,0282

A matriz de decisão, a seguir, mostra o peso atribuído a cada critério, o desempenho das alternativas (fornecedores) em relação a cada critério e o desempenho global (Resultado)

que corresponde ao somatório das multiplicações entre o peso de cada critério e o respectivo percentual de desempenho das alternativas naquele Critério:

Tabela 26: Desempenho final dos fornecedores (Matriz de Decisão)

	Qualidade	Preço	Forma de PGTO	Frete	Prazo	Resultado
Pesos	33,82%	37,63%	10,77%	10,32%	7,46%	
Alternativa 1	22,90%	26,94%	61,38%	54,40%	45,98%	33,54%
Alternativa 2	33,09%	5,30%	12,58%	6,34%	8,83%	15,85%
Alternativa 3	36,69%	62,46%	14,59%	33,97%	38,95%	43,89%
Alternativa 4	7,32%	5,30%	11,45%	5,29%	6,24%	6,71%
Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Dessa forma, ao aplicar o Modelo AHP para a seleção de fornecedores, o melhor desempenho foi da alternativa 3 com 43,89%; que, comparativamente com os demais desempenhos 33,54%, 15,85% e 6,71% das alternativas 1, 2 e 4 respectivamente, pode ser considerado um resultado robusto neste primeiro momento.

O tipo de Análise de Sensibilidade empregado neste trabalho é o de alterações

dos pesos relativos dos critérios, como descrito a seguir.

1ª Simulação: A comparação entre o critério Qualidade e os demais, foi alterada considerando o critério Qualidade um pouco mais importante em todos os casos. As demais comparações permanecem iguais.

A Tabela 27 mostra os julgamentos originais e as variações propostas para a Análise de Sensibilidade:

Tabela 27: Variações de avaliação para Análise de Sensibilidade

Comparações	Valor original	Variações
Qualidade-Preço	1	6
Qualidade-Forma de PGTO	3	5
Qualidade-Frete	3	5
Qualidade-Prazo	5	6
Preço-Forma de PGTO	5	5
Preço-Frete	3	3
Preço-Prazo	5	5
Forma de PGTO-Frete	1	1
Forma de PGTO-Prazo	2	2
Frete-Prazo	1	1

Vale destacar que com a montagem aqui proposta basta substituir os valores da coluna variações na Tabela 27 (que corresponde à Tabela 1) para obter de forma rápida e simples (automática) uma nova solução para o problema. As Tabelas 28 e 29 mostram a

nova Matriz de Comparações Pareadas obtida a partir dos novos julgamentos e o resultado do respectivo teste de consistência. Deve ser observado que os novos cálculos apresentados nas tabelas a seguir foram feitos na mesma planilha inicial.

Tabela 28: Matriz de comparação pareada dos Critérios

Critérios	Qualidade	Preço	Forma de PGTO	Frete	Prazo
Qualidade	1	6	5	5	6
Preço	1/6	1	5	3	5
Forma de PGTO	1/5	1/5	1	1	2
Frete	1/5	1/3	1	1	1
Prazo	1/6	1/5	1/2	1	1
Soma	1,7333	7,7333	12,5000	11,0000	15,0000

Tabela 29: Teste de Consistência

λ máx	CI	RI	CR
5,4012	0,1003	1,1200	0,0895

O resultado final pode ser visto na Tabela 30.

Tabela 30: Desempenho final dos fornecedores (Matriz de Decisão)

	Qualidade	Preço	Forma de PGTO	Frete	Prazo	Resultado
Pesos	52,15%	24,63%	8,91%	7,92%	6,39%	
Alternativa 1	22,90%	26,94%	61,38%	54,40%	45,98%	31,30%
Alternativa 2	33,09%	5,30%	12,58%	6,34%	8,83%	20,75%
Alternativa 3	36,69%	62,46%	14,59%	33,97%	38,95%	41,00%
Alternativa 4	7,32%	5,30%	11,45%	5,29%	6,24%	6,96%
Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

2ª Simulação: Nesta simulação o critério Preço teve maior importância em relação aos demais. Os novos julgamentos, a nova Matriz

de Comparações Pareadas e o respectivo teste de consistência, são vistos nas Tabelas 31, 32 e 33, respectivamente.

Tabela 31: Variações de julgamentos para Análise de Sensibilidade

Comparações	Valor original	Variações
Qualidade-Preço	1	1/5
Qualidade-Forma de PGTO	3	3
Qualidade-Frete	3	3
Qualidade-Prazo	5	5
Preço-Forma de PGTO	5	7
Preço-Frete	3	5
Preço-Prazo	5	7
Forma de PGTO-Frete	1	1
Forma de PGTO-Prazo	2	2
Frete-Prazo	1	1

Tabela 32: Matriz de comparação pareada dos Critérios

Critérios	Qualidade	Preço	Forma de PGTO	Frete	Prazo
Qualidade	1	1/5	3	3	5
Preço	5	1	7	5	7
Forma de PGTO	1/3	1/7	1	1	2
Frete	1/3	1/5	1	1	1
Prazo	1/5	1/7	1/2	1	1
Soma	6,8667	1,6857	12,5000	11,0000	16,0000

Tabela 33: Teste de Consistência

λ máx	CI	RI	CR
5,2028	0,0507	1,1200	0,0453

O novo peso de cada critério, bem como o desempenho final dos fornecedores pode ser visto na Tabela 34.

Tabela 34: Desempenho final dos fornecedores

	Qualidade	Preço	Forma de PGTO	Frete	Prazo	Resultado
Pesos	21,79%	55,47%	8,58%	8,01%	6,15%	
Alternativa 1	22,90%	26,94%	61,38%	54,40%	45,98%	32,39%
Alternativa 2	33,09%	5,30%	12,58%	6,34%	8,83%	12,28%
Alternativa 3	36,69%	62,46%	14,59%	33,97%	38,95%	49,01%
Alternativa 4	7,32%	5,30%	11,45%	5,29%	6,24%	6,32%
Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

3ª Simulação: A última simulação considerou o critério Forma de Pagamento com grande importância em relação aos outros critérios, como mostra a Tabela 35. A Matriz de

Comparações Pareadas dos Critérios e a respectiva análise de consistência também podem ser visualizadas nas Tabelas 36 e 37.

Tabela 35: Variações de julgamento para Análise de Sensibilidade

Comparações	Valor original	Variações
Qualidade-Preço	1	1
Qualidade-Forma de PGTO	3	1/2
Qualidade-Frete	3	3
Qualidade-Prazo	5	5
Preço-Forma de PGTO	5	1/5
Preço-Frete	3	3
Preço-Prazo	5	5
Forma de PGTO-Frete	1	5
Forma de PGTO-Prazo	2	7
Frete-Prazo	1	1

Tabela 36: Matriz de comparação pareada dos Critérios

Critérios	Qualidade	Preço	Forma de PGTO	Frete	Prazo
Qualidade	1	1	1/2	3	5
Preço	1	1	1/5	3	5
Forma de PGTO	2	5	1	5	7
Frete	1/3	1/3	1/5	1	1
Prazo	1/5	1/5	1/7	1	1
Soma	4,5333	7,5333	2,0429	13,0000	19,0000

Tabela 37: Teste de Consistência

λ máx	CI	RI	CR
5,1848	0,0462	1,1200	0,0413

Os pesos alterados para cada critério e o resultado do desempenho dos fornecedores a partir desta simulação são mostrados na Tabela 38.

Tabela 38: Desempenho final dos fornecedores (Matriz de Decisão)

	Qualidade	Preço	Forma de PGTO	Frete	Prazo	Resultado
Pesos	21,84%	18,90%	46,95%	6,90%	5,40%	
Alternativa 1	22,90%	26,94%	61,38%	54,40%	45,98%	45,15%
Alternativa 2	33,09%	5,30%	12,58%	6,34%	8,83%	15,05%
Alternativa 3	36,69%	62,46%	14,59%	33,97%	38,95%	31,12%
Alternativa 4	7,32%	5,30%	11,45%	5,29%	6,24%	8,68%
Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Ao finalizar as três simulações, pôde-se comparar, como mostra a Tabela 39, os pesos obtidos para cada critério em relação ao peso original. Na avaliação original, o critério Preço tinha maior importância, seguido dos critérios: Qualidade, Forma de Pagamento, Frete e Prazo. Na primeira

simulação e critério Qualidade teve maior destaque. Na segunda simulação, o critério Preço teve um peso bem maior que os demais, e por fim, na terceira simulação, o maior peso foi do critério Forma de Pagamento.

Tabela 39: Comparação dos pesos obtidos por simulação

	Peso original	1ª Simulação	2ª Simulação	3ª Simulação
Qualidade	33,82%	52,15%	21,79%	21,84%
Preço	37,63%	24,63%	55,47%	18,90%
Forma de PGTO	10,77%	8,91%	8,58%	46,95%
Frete	10,32%	7,92%	8,01%	6,90%
Prazo	7,46%	6,39%	6,15%	5,40%

Sendo assim, os desempenhos globais das alternativas (fornecedores), segundo a avaliação original e perante cada simulação, podem ser analisados na Tabela 40.

Tabela 40: Comparação do desempenho dos fornecedores

	Desempenho original	1ª Simulação	2ª Simulação	3ª Simulação
Alternativa 1	33,54%	31,30%	32,39%	45,15%
Alternativa 2	15,85%	20,75%	12,28%	15,05%
Alternativa 3	43,89%	41,00%	49,01%	31,12%
Alternativa 4	6,71%	6,96%	6,32%	8,68%

Para as duas primeiras simulações, a alternativa 3 se manteve com o melhor desempenho; entretanto, ao considerar a Forma de Pagamento com grande importância – situação não comum, mas que pode vir a acontecer em determinado momento a partir de uma série de fatores que influenciam interna e externamente em uma empresa – a alternativa 1 se torna a mais adequada.

Todos os tipos de Análise de Sensibilidade possíveis poderiam ser facilmente empreendidos para o problema em questão, desde que observadas as devidas consistências.

5. CONSIDERAÇÕES FINAIS

Naturalmente, o presente artigo sofre as limitações comuns à descrição de uma longa pesquisa conceitual que necessita ser sintetizada em poucas linhas. A partir de seu caráter exploratório, é possível desdobrar propostas para pesquisas futuras, entre as quais se destaca a realização de problemas

reais para um maior aprofundamento no uso do modelo estudado.

Ao adotar uma planilha eletrônica não dedicada e facilmente disponível em empresas de qualquer porte, reforça-se o caráter de modelo do AHP, viabilizando que com a construção de uma única planilha, como a usada aqui, sejam resolvidos vários diferentes problemas de tomada de decisão sem a necessidade de laboriosas contas (pelo menos conscientemente), nem de várias planilhas.

Conclui-se, assim, pela viabilidade da utilização do AHP – associado a uma planilha eletrônica de fácil acesso, sem depender de nenhum software proprietário – não só como um método composto de passos bem definidos, que levam a uma tomada de decisão, mas, principalmente, como um modelo matemático, que propicia variações e Análise de Sensibilidade, fácil de aplicar e capaz de propiciar ao decisor, mais do que verificar a robustez de sua decisão, aprender e aperfeiçoar seu processo decisório.

REFERÊNCIAS

- [1] Abreu, L. M. de; Granemann, S. R.; Gartner, I.; Bernardes, R. S. Escolha de um Programa de Controle da Qualidade da Água para Consumo Humano: Aplicação do Método AHP. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.4, n.2, p.257-262, 2000. Disponível em: <<http://www2.ic.uff.br/~emitacc/AMD/Artigo%204.pdf>>. Acesso em: 8 fev. 2015.
- [2] Almeida, A. T. O Conhecimento e o Uso de Métodos Multicritério de Apoio a Decisão. 2. ed. Recife: Ed. Universitária da UFPE, 2011. 234p.
- [3] Bazerman, M. H.; Moore, D. *Processo Decisório*. 7.ed. Rio de Janeiro: Elsevier, 2010. 319 p.
- [4] Corazza, R. I.; Fracalanza, P. S. Caminhos do pensamento neo-schumpeteriano: para além das analogias biológicas. *Nova Economia*, v. 14, n. 2, p. 127-155, Mai./Ago., 2004.
- [5] Goldman, F. L. A Dinâmica da Criação do Conhecimento Organizacional: um estudo sobre inovação no Sistema Eletrobrás. 2013. 243f. Tese (Doutorado) – Instituto de Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.
- [6] Pacheco, M. C. R. Modelos multicriteriais de apoio à decisão: o método AHP como auxílio à seleção de fornecedores em uma confecção. 2015. 80 f. Trabalho de Conclusão de Curso (Graduação em Engenharia de Produção) - Centro Universitário Serra dos Órgãos, Teresópolis. 2015.
- [7] Reis, E. dos; Löbler, M. L. O Processo Decisório Descrito pelo Indivíduo e Representado nos Sistemas de Apoio à Decisão. *Revista de Administração Contemporânea*, Rio de Janeiro, v. 16, n. 3, p. 397-417, Mai/Jun. 2012. Disponível em <<http://www.scielo.br/pdf/rac/v16n3/v16n3a05.pdf>> . Acesso em: 8 mar. 2015.
- [8] Saaty, T.L. How to make a decision : The Analytic Hierarchy Process. *European Journal of Operations Research*, v. 48, n.1, p. 9-26, 1990.
- [9] Saaty, T.L. Decision making with the analytic hierarchy process, *Int. J. Services Sciences*, v. 1, n. 1, p.83–98, 2008.
- [10] Shimizu, T. *Decisão nas Organizações*. 2.ed. São Paulo: Atlas, 2006. 419p.

Capítulo 4

A TEORIA DE FILAS COMO FERRAMENTA DE APOIO NA ANÁLISE DE UM SERVIÇO DE ATENDIMENTO

Bianca Porto de Lima

Artur da Silva Siqueira de Novais

Caroline Cristina de Carvalho

Larissa Melo Negrão

Resumo: A teoria das filas consiste em obter modelos adequados de circunstâncias que envolvem filas, prevendo assim o seu comportamento. Esse comportamento é expresso por diversas medidas de desempenho, designadamente, a taxa de chegada dos clientes e taxa de atendimento do sistema. A presente pesquisa objetiva analisar e evidenciar a importância da teoria de filas na avaliação da capacidade de atendimento da Loteria do Lazer, filial da Caixa Econômica Federal da cidade de Barra Mansa no estado do Rio de Janeiro. A insatisfação dos clientes pelo tempo esperado na fila é decorrente da grande procura pelos serviços da agência, que fica localizada em um local no qual há um grande fluxo de pessoas. Neste estudo de caso foi analisado apenas o caixa preferencial, onde foi possível identificar a capacidade de atendimento, simular e analisar o comportamento quanto ao tempo de espera na fila e no sistema e seus tamanhos, respectivamente, de forma a dimensionar a capacidade do serviço, buscando a satisfação dos clientes e a viabilidade econômica do projeto. Demonstrando que é possível aplicar e simular de forma simples e coerente o funcionamento da teoria de filas com a finalidade de propor soluções adequadas caso algum problema for detectado.

Palavras-chave: Teoria de filas, Taxa de chegada, Taxa de atendimento.

1. INTRODUÇÃO

Atualmente, a qualidade torna-se cada vez mais um fator determinante para a escolha de métodos de serviços. Dessa forma, é importante o estudo da teoria das filas, pois esta visa aperfeiçoar o desempenho de um sistema, reduzindo seus custos e gerando satisfação.

Neste mundo contemporâneo, em todos os serviços, a relação interpessoal direta ou indireta sempre envolve a obrigação de esperar. No entanto, oferecer um serviço de qualidade significa saber compreender a demanda de clientes, organizar sistemas que a controlem e ajustar a capacidade de atendimento em função da demanda, e como consequência, reduzir a espera e manter um serviço adequado para o cliente (GIANESE; CORRÊA, 1996).

A Teoria das Filas é utilizada, com frequência, para resolução de problemas que envolvem tempo de espera, ou seja, em um determinado sistema clientes chegam para serem atendidos, recebem o serviço e depois se retiram do sistema (ROMERO et al. 2010).

Conforme Abensur (2011), o motivo das filas está relacionado com a capacidade de

atendimento de um servidor e a demanda existente para ser atendida. Nesse sentido, é importante para o gerenciamento de serviços, já que é inviável conhecer toda uma demanda existente durante um período, logo a técnica se torna representativa para um dado intervalo de tempo, sendo esta um método analítico que aborda o assunto por meio de fórmulas matemáticas, no qual se estuda as relações entre as demandas e os atrasos sofridos pelo usuário do sistema, para avaliação das medidas de desempenho dessa relação em função da disposição deste sistema (ARENALES et al. 2007). Em Morabito; Lima (2000) são descritos dois tipos de sistema de atendimento. O primeiro referente à n atendentes para uma única fila, com disciplina de atendimento FIFO (First In, First Out), admitindo que λ clientes chegam em um determinado intervalo médio de tempo e que cada servidor possui a mesma capacidade atendimento μ , em um determinado tempo. O segundo modelo, também FIFO, admiti m atendentes que servem filas paralelas com chegadas de λ clientes por intervalo de tempo e cada servidor possuindo sua capacidade de atendimento μ . A Figura 1 ilustra estas duas situações decorrentes em sistemas de filas.

Figura 1: (a) Processo para uma única fila (M/M/m) (b) sistema para filas em paralelo (M/M/1).

O processo de chegada é quantificado através da taxa média de chegada em um determinado intervalo de tempo (λ) e que ocorre de maneira aleatória (ARENALES et al. 2007). Para caracterizar esta aleatoriedade corretamente, um processo de chegada deve dispor de uma distribuição de probabilidade, tal como uma distribuição Normal, de Poisson, exponencial negativa, Erlang, etc. Os mesmos autores ainda esclarecem que o processo de chegada não varia ao longo do tempo e que

não é afetado pelo número de usuários presentes no sistema, sendo um processo determinístico.

Já o processo de atendimento é quantificado através do ritmo médio de atendimento (μ) e do tempo, ou, duração média do serviço. Arenales (2007) admite que não mais de um usuário pode ser atendido por um servidor no mesmo instante, que o processo não varia ao longo do tempo e que não é afetado pelo número de usuários presentes no sistema. O

sistema pode possuir um ou mais servidores, sendo que esta quantidade é influenciada pela demanda existente.

A satisfação dos clientes está na qualidade dos serviços oferecidos e prestados. Quanto mais satisfeito o cliente está com o produto, maior é sua lealdade com este, já que ele atende a todos seus anseios e com o aumento da lealdade, clientes satisfeitos podem fornecer vantagens lucrativas, como assegurar receitas ao longo do tempo, a reduzir custos de transações futuras, a diminuir a elasticidade de preço e minimizar a probabilidade de os clientes trocarem de fornecedor caso haja uma queda no nível da qualidade (ANDERSON; FORNELL, 1999).

Para aplicação destes preceitos, foi selecionada a Loteria do Lazer localizada no município de Barra Mansa, visto que esta apresenta um fluxo constante de pessoas proporcionando alto nível de confiabilidade.

2. DESCRIÇÃO DO SISTEMA

O estudo foi executado na Loteria do Lazer, uma filial do sistema bancário da Caixa Econômica Federal localizada no município de Barra Mansa no interior do estado do Rio de Janeiro. Considerando que a casa lotérica está localizada no centro do município, foi possível observar um constante fluxo de pessoas. Neste estabelecimento é possível realizar o pagamento de contas, fazer apostas em jogos de loteria (como a Loto fácil, por exemplo) e pequenas operações bancárias, tais como receber o seguro desemprego e entre outras.

A loteria dispõe de cinco caixas de atendimento, sendo um deste preferencial

para idosos, gestantes e pessoas com necessidades especiais.

Verificando o alto fluxo de pessoas neste caixa e sua maneira diferente de atuar, já que é preferencial, foi escolhido apenas este para ser analisado e estudado, no qual os clientes chegam em uma fila única e aguardam por atendimento em ordem de chegada.

3. MODELAGEM DO SISTEMA

Para se estudar de forma adequada a casa lotérica escolhida, deve-se verificar como é a metodologia do sistema. Além de coletar os dados necessários como a taxa de chegada e saída dos clientes, o tempo de espera na fila, a taxa média de atendimento e fazer a correta modelagem. Este estudo de caso baseou-se em apenas um caixa, o modelo adotado para análise é, portanto, o MM1. Onde há uma única fila, na qual os clientes são atendidos de acordo com a ordem de chegada.

Por se tratar de um estudo de caso simples e com objetivo de demonstrar a funcionalidade da teoria de filas, apenas uma coleta de dados foi realizada.

3.1. TAXA DE CHEGADA DE CLIENTES

O levantamento de dados foi coletado observando quantos clientes chegam a cada minuto e aguardam atendimento na fila preferencial. Esta coleta teve início às 11 horas 56 minutos e terminou às 12 horas 56 minutos, pois foi observado grande fluxo neste período. Neste período de uma hora foram coletados os seguintes dados apresentados na Tabela 1.

Tabela 1: Dados coletados de entrada de clientes.

Minuto	Frequência Observada	Minuto	Frequência Observada	Minuto	Frequência Observada
1	3	21	1	41	3
2	0	22	3	42	2
3	2	23	1	43	7
4	0	24	4	44	2
5	1	25	2	45	2
6	1	26	8	46	1
7	0	27	2	47	1
8	2	28	4	48	5
9	5	29	4	49	4

Tabela 1: Dados coletados de entrada de clientes (continuação)

Minuto	Frequência Observada	Minuto	Frequência Observada	Minuto	Frequência Observada
10	2	30	3	50	1
11	3	31	2	51	3
12	0	32	1	52	0
13	2	33	2	53	3
14	4	34	5	54	1
15	1	35	1	55	1
16	1	36	4	56	1
17	1	37	4	57	3
18	2	38	2	58	0
19	5	39	2	59	2
20	2	40	1	60	3

Em primeira análise, pode-se perceber que durante 12:10 e 12:40 o número de chegadas aumentou consideravelmente. Tendo seu pico

entre 12:10 e 12:20. Conforme pode ser observado no gráfico abaixo:

Gráfico 1: Taxa de chegada na loteria por minuto durante uma hora.

3.2. TAXA DE CHEGADA DE CLIENTES

Também foi coletado no mesmo período e utilizando a mesma metodologia utilizada para

a coleta de dados de chegada de clientes, a quantidade de clientes que saíam da agência lotérica. Os dados desta coleta podem ser observados na tabela 2 abaixo:

Tabela 2: Dados coletados de saída de clientes.

Minuto	Frequência observada	Minuto	Frequência observada	Minuto	Frequência observada
1	1	21	1	41	3
2	2	22	1	42	4
3	0	23	8	43	3
4	2	24	2	44	2
5	4	25	1	45	4
6	1	26	5	46	2
7	2	27	4	47	0
8	1	28	5	48	2
9	1	29	1	49	4
10	1	30	4	50	0
11	3	31	5	51	2
12	1	32	2	52	3
13	4	33	2	53	3
14	4	34	2	54	1
15	8	35	4	55	3
16	4	36	3	56	3
17	1	37	3	57	0
18	2	38	1	58	3
19	2	39	4	59	1
20	4	40	1	60	4

Quanto a saída, pode-se analisar que há também vezes que mais clientes foram atendidos dentro de um minuto. Porém tende

a demonstrar um comportamento mais constante (com menos picos). Como pode-se verificar no gráfico abaixo:

Gráfico 2: Taxa de saída na loteria por minuto durante uma hora.

Com todos estes dados coletados, é possível verificar se esta estrutura de sistema se encaixa na teoria das filas. Ou seja, através do cálculo da frequência calculada para a realização dos testes de Aderência e do Qui-quadrado, isto é, testar se a variável em questão, número de chegadas por unidade de tempo, segue uma distribuição de Poisson

com taxa 3,014 clientes por minuto, ao 5% de significância.

4. TESTE DE ADERENCIA

Os dados coletados, foram organizados nas tabelas a seguir para facilitar o entendimento dos valores usados:

Tabela 3: Resumo das informações coletadas sobre a chegada de clientes.

Número de Chegadas/min (Xi)	Frequência Observada	P(nº obs = K)	Frequência Esperada
0	6	0,10025884	6,015531
1	16	0,23059534	13,83572
2	16	0,26518464	15,91108
3	9	0,20330823	12,19849
4	7	0,11690223	7,014134
5	4	0,05377503	3,226502
6	0	0,02061376	1,236826
7	1	0,00677309	0,406386
8	1	0,00194726	0,116836

Tabela 4: Resumo das informações coletadas sobre a saída de clientes.

Número de Saídas/min (Xi)	Frequência Observada	P(nº obs = K)	Frequência Esperada
0	4	0,07679109	4,607465375
1	15	0,19709713	11,8258278
2	13	0,252941317	15,176479
3	10	0,216405349	12,98432093
4	13	0,138860099	8,331605927
5	3	0,071281517	4,276891043
6	0	0,030492649	1,829558946
7	0	0,011180638	0,67083828
8	2	0,003587121	0,215227282

A partir dos dados coletados, realizou-se o teste Qui-quadrado com os valores de taxa de entrada e taxa de saída. É importante ressaltar, ainda, que há uma regra em que é necessário que a frequência esperada seja maior que do que 5. Portanto alguns grupos

tiveram que ser agrupados para se chegar neste valor de frequência esperada. Para utilização deste, foi necessário a montagem de um quadro com os resultados apontados da frequência esperada.

Tabela 5: Resumo das informações coletadas na entrada com X^2 calculado.

Frequência Observada Acumulada	Frequência Esperada Acumulada	X^2 calculado
6	6,01553062	0,000040096
16	13,8357204	0,33855165
16	15,9110785	0,00049695
9	12,1984935	0,83865772
13	12,0006823	0,08321493

Somando ainda o valor do X^2 calculado pode-se obter o valor de 1,26096135. Ainda seguindo o teste de aderência é necessário verificar o valor de V .

$$V = 5 - 1 - 1 = 3$$

Desta forma, foi aplicada a fórmula do teste Qui-quadrado, obtendo o valor abaixo.

$$X^2_{\text{calc}} = 1,26$$

De acordo com a tabela de Teste de Aderência, foi encontrado o X^2 tabelado.

$$X^2_{\text{tab}} = 7,82$$

Foi encontrado que X^2 tabelado é maior que X^2 calculado, logo, aceita-se a hipótese de que segue a distribuição Poisson.

Entretanto, ainda é necessário realizar o mesmo com os dados de saída de clientes. Logo, seguindo o mesmo procedimento, foi montada uma tabela para resumir dados coletados no local a respeito da saída de clientes.

Tabela 6: Resumo das informações coletadas na saída com X^2 calculado.

Frequência Observada Acumulada	Frequência Esperada Acumulada	X^2 calculado
19	16,43329317	0,400892498
13	15,176479	0,312131744
10	12,98432093	0,68591738
13	8,331605927	2,615810614
5	6,99251555	0,567766806

Com os dados em mãos e o X^2 calculado é possível continuar com a verificação da hipótese com o teste Qui-quadrado. Sabe-se que a soma do X^2 calculado dos dados de saída é de 4,582519042.

$$V = 5 - 1 - 1 = 3$$

Desta forma, foi aplicada a fórmula do teste qui-quadrado, obtendo o valor abaixo.

$$X^2_{\text{calc}} = 4,58$$

De acordo com a tabela de Teste de Aderência, foi encontrado o X^2 tabelado.

$$X^2_{\text{tab}} = 7,82$$

Foi encontrado que X^2 tabelado é maior que X^2 calculado, logo, aceita-se a hipótese de que segue a distribuição Poisson.

5. A TEORIA DE FILAS

O modelo adequado para esta situação é ajustado segundo os dados obtidos. Então com isso pode-se afirmar que é MM1 com $\lambda = 2,3$ clientes/min e $\mu = 2,567$ /min. Os parâmetros a serem calculados e que representam a operacionalidade do sistema atual estão descritos na tabela 7 abaixo:

Tabela 7: Resultado das variáveis de desempenho da Teoria de Filas.

Variáveis de Desempenho da Teoria de Filas	Resultados
Probabilidade de não existir nenhum cliente no sistema (P_0)	12%
Quantidade de clientes no sistema (L)	6,42
Quantidade de clientes na fila (L_q)	5,54
Taxa de entrada (λ)	2,3
Fator de utilização (ρ)	0,88
Tempo médio de espera no sistema (W)	2,84
Tempo médio de espera na fila (W_q)	2,45

6. CONCLUSÃO

O trabalho realizado a respeito da Loteria do Lazer apresentado neste estudo de caso, exemplifica e demonstra uma aplicação de teoria das filas em um sistema cotidiano, evidenciando que é possível aplicar de forma simples a pesquisa operacional no setor de serviços para coletar dados e comprovar hipóteses com o fim de otimizar espaços, balancear sistemas, melhorar atendimentos e aumentar a satisfação de clientes e funcionários.

Conforme descrito anteriormente, a loteria estudada apresenta cinco caixas de atendimento e um caixa preferencial, sendo este o escolhido para estudo e aplicação da teoria de filas. Assim sendo, este enquadra-se no modelo de fila do tipo M/M/1, que por sua vez caracteriza-se por apresentar uma única fila para um único posto de atendimento.

Para melhor análise do desempenho, deve-se realizar uma pesquisa de satisfação com os clientes em relação aos serviços prestados.

REFERÊNCIA

- [1]. ANDERSON, E.W. & FORNELL, C. (1994) - A customer satisfaction research prospectus. In: Service Quality: new directions in theory and practice, RUST, R.T. & OLIVER, R.L. (eds.) Thousand Oaks: Sage Publications.
- [2]. ANDRADE, Eduardo Leopoldino de. Introdução à Pesquisa Operacional: métodos e modelos para análise de decisões. 3ª edição. Rio de Janeiro: LTC, 2002.
- [3]. ANDRADE, E. L. Problemas de Congestionamento das Filas. In: ANDRADE, E. L. Introdução à Pesquisa Operacional: Métodos e modelos para análise de decisões. Ed. 4. Rio de Janeiro: LTC, 2009.
- [4]. ARENALES, M. ET AL. Pesquisa operacional: para cursos de engenharia. Rio de Janeiro: Elsevier, 2007.
- [5]. AURELIO, Marco. Teoria das filas e das simulações (apostila). Santo André: UNIA, 2004.
- [6]. DOS SANTOS, MAURÍCIO. Apostila de Pesquisa Operacional da Universidade Estadual do Rio de Janeiro (UERJ). R.J., Brasil, 2003.
- [7]. Estatística – Qui Quadrado. Disponível em: <<http://www.ufpa.br/dicas/biome/biopdf/bioqui.pdf>> Acesso em 03 de novembro de 2015.
- [8]. FIGUEIREDO, Danielle Durski; ROCHA, Silvana Heidemann. Aplicação da Teoria de Filas na Otimização do número de caixas: um estudo de caso. Disponível em: <<http://periodicos.unicesumar.edu.br/index.php/iccsumar/article/view/1300/1113>> Acesso em 03 de novembro de 2015.
- [9]. FILHO, Paulo José de Freitas. Introdução a modelagem e simulação com aplicações em arena. Disponível em: <<http://bookstore.com.br/shop/sumarios/01113.pdf>> Acesso em 03 de novembro de 2015.
- [10]. GIANESI, Irineu G. N.; CORRÊA, Henrique Luis. Administração estratégica de serviços. São Paulo. Atlas, 1996
- [11]. KRAJEWSKI, Lee, Larry; RITZMAN, Larry; MALHOTRA, Manoj. Administração de Produção e Operações. 8. Ed. São Paulo: Pearson, 2009.

[12]. TAHA, Hamdy A. Pesquisa Operacional. 8. Ed. São Paulo: Pearson, 2008 HILLIER, Frederick S. & LIEBERMAN, Gerald J. Introdução à Pesquisa Operacional. 8ª Edição. São Paulo: McGraw-Hill, 2006.

[13]. TORRES, Oswaldo Fadigas. Elementos da teoria de filas. Disponível em: <<http://bibliotecadigital.fgv.br/ojs/index.php/rae/article/viewFile/40705/39458>> Acesso em 03 de novembro de 2015.

Capítulo 5

CONFIABILIDADE DOS QUESTIONÁRIOS: UMA ANÁLISE DA CONSISTÊNCIA INTERNA DA ESCALA F BASEADA NO COEFICIENTE ALFA DE CRONBACH

Carolina Riul de Freitas

Carla Regina de Moraes

Francisco Severo de Lima Junior

Maria Cristina Vendrameto

Rafael Lustosa da Costa

Resumo: Esta pesquisa ocupou-se da Confiabilidade dos Surveys. Objetivou analisar o processo de estimação da Confiabilidade de um questionário a partir da análise de Consistência Interna, adotando o Alfa de Cronbach como estatística para esta verificação. Considerando o paradigma crítico ao qual a pesquisa se vincula, a decisão acerca dos métodos foi pautada nas necessidades e características do objeto pesquisado, de modo que a abordagem indutiva e procedimentos estatísticos mostraram-se adequados para seu empreendimento. Classificada como de natureza aplicada, quantitativa quanto à abordagem do problema, analítica em relação aos objetivos e bibliográfica, documental e de levantamento no que diz respeito aos procedimentos, confirmou a hipótese que a medida de Confiabilidade do caso observado, a saber, os resultados da aplicação da Escala de Fascismo nos alunos do curso de Tecnologia em Secretariado da Fatec Carapicuíba, poderia ser incrementado a partir da exclusão de determinadas questões. A aplicação da Escala F em sua versão de com 17 itens indicou um coeficiente Alfa de 0,705, que foi elevado para 0,732 após a exclusão de dois itens. Mesmo com resultados moderados, é possível afirmar que tal estratégia se apresenta como uma alternativa para o incremento da precisão dos instrumentos de coleta de dados.

Palavras Chave: Confiabilidade - Alfa de Cronbach - Escala F

*Artigo originalmente publicado nos anais do Simpósio de Excelência em Gestão e Tecnologia 2016

1. INTRODUÇÃO

As pesquisas do tipo *Survey* caracterizam-se pela interrogação dos sujeitos cujo o comportamento, opinião ou atitude deseja-se conhecer. Em síntese, escolhe-se um grupo, solicita aos partícipes deste grupo que emitam informações acerca de dado problema ou objeto em evidência e concede tratamento quantitativos aos dados, para que a partir das informações geradas se estabeleçam conjectura ou conclusões (GIL, 2008).

A noção de censo pode ser observada na história cultural da humanidade por meio dos vestígios egípcios que registram os indícios da origem das pesquisas do tipo *Surveys*. Já na pesquisa social, a investigação desenvolvida por Marx, em sua obra *A Worker's Inquiry* acerca das condições da classe operária francesa pode ser considerada pioneira na utilização deste procedimento (BABBEL, 2003).

A Confiabilidade destes instrumentos de coleta de dados é o tema deste estudo. Apesar do grande enfoque que o campo da Psicometria dá a este tema, sua grande marca é a transdisciplinaridade da sua aplicação, de modo que existe vasto campo para a aplicação dos conceitos aqui abordados. Vale-se da Estatística e da Matemática para trabalhar a questão da medida em questionários e testes. A Psicometria trata, essencialmente, das propriedades métricas da Confiabilidade e da Validade dos Testes nas Ciências Psicossociais. Como parte integrante das pesquisas de abordagem quantitativa ou integrada, caracteriza-se pela utilização de método de procedimento estatístico abrangendo vários níveis cognitivos de abordagem do objetivo de pesquisa, os testes psicológicos, utilizados em diversas áreas do conhecimento, não ficando restrito às Ciências Psicossociais. São aplicados por meio de instrumentos de pesquisa conhecidos como *Surveys* com o objetivo de extrair de uma determinada população seu posicionamento, em diversos níveis comportamentais, acerca de um dado objeto.

Esta pesquisa norteia-se pela interrogação sobre a medida de Confiabilidade da Escala de Fascismo, (Escala F), elaborada por Adorno *et al.* nos Estados Unidos, década de 1950 e adaptada no Brasil por Crochík (2009). Sua intenção é verificar se a exclusão de itens do teste apresenta-se como alternativa

factível para o incremento da medida em questão.

Mesmo sem o protagonismo de ser observado como objeto desse estudo, a reflexão acerca dos condicionantes e as relações de causa e consequência do Fascismo, abordado aqui não enquanto movimento político, mas enquanto ideologia, se justifica pelo seu próprio empreendimento. Se em dado momento da História, o horror fascista foi possível, a vigília para que as condições desta regressão à barbárie sejam extintas se fazem urgentes. “A barbárie continuará existindo enquanto persistirem no que têm de fundamental as condições que geram esta regressão [...] tanto a estrutura básica da sociedade como os seus membros, responsáveis por termos chegado onde estamos, não mudaram” (ADORNO, 1993).

Nesse ponto reside o problema da racionalidade burguesa: o Fascismo pode ser considerado um subproduto da Era das Máquinas. A tecnologia, materialidade desta Razão, pode ser o meio da intensificação ou da abolição do trabalho forçado, da escassez ou da abundância, da Liberdade ou do Autoritarismo. O nazi-Fascismo é um exemplo de como uma economia altamente racionalizada e maquinizada, portanto, produtivamente eficiente, pode funcionar a favor da dominação e da imposição da escassez (MARCUSE, 2001).

Mas, quais foram as bases da ascensão do Fascismo? A concepção marxista ortodoxa não consegue atender a demanda de explicar, por meio das condições materiais e objetivas da sociedade, contradição entre a deterioração das condições de vida dos trabalhadores e da economia política burguesa alemã e o surgimento de um movimento antirrevolucionário. Uma vez que, diferente de outros regimes reacionários, o Fascismo obtém apoio das massas, cabe à investigação das condições ideológicas destas e do mecanismo de introjeção da ideologia dominante, uma vez que essa é a manifestação da ideologia da classe dominante, nos sujeitos. Assume-se aqui, que a reprodução da condição econômica na consciência política não é um movimento automático de transferência, o que evidencia a ação da ideologia que, neste caso, é a ideologia fascista. O descrédito ou o desconhecimento da estrutura do caráter das massas na explicação de fenômenos desta natureza pode produzir questionamentos

estéreis e, por conseguinte, respostas estéreis (REICH, 1988).

A investigação sobre como a ideologia fascista se instala nos sujeitos, uma vez que o Fascismo se caracteriza como um movimento de massas, é imperativa. As contribuições da Escola de Frankfurt são elucidativas para a interpretação deste fenômeno. Tanto identificaram os traços latentes de indivíduos potencialmente fascistas, o que se chamou de Personalidade Autoritária, quanto criaram um instrumento para identificar e mensurar traços desta personalidade nos sujeitos, a Escala de Fascismo (Escala F).

O construto da Escala F é a noção de Personalidade Autoritária. Em síntese última, o paradigma da Personalidade Autoritária se assenta em características como: Adesão ao ideário burguês, acriticismo em relações de autoridade moral, agressividade autoritária, hostilidade e desprezo às humanidades, ênfase nas relações de dominação-submissão, superstição e estereotipia, oposição à subjetividade, projetividade dos temas externos para elementos exteriores a si e preocupação excessiva com a sexualidade e podem ser observadas tanto na vida pessoal quanto nas atitudes dos sujeitos (ADORNO et al., 1950). Esses traços foram investigados nos sujeitos deste estudo: alunos do curso de Tecnologia em Secretariado da Faculdade de Tecnologia de Carapicuíba, São Paulo.

Considerando o paradigma científico crítico ao qual a pesquisa se vincula, a decisão acerca dos métodos foi pautada nas necessidades e características do objeto da pesquisa, de modo que o método indutivo de abordagem e o estatístico de procedimento configuraram-se como os mais adequados para o desenvolvimento da pesquisa que foi classificada como sendo de natureza aplicada, quantitativa quanto à abordagem do problema, analítica em relação aos objetivos e bibliográfica, documental e de levantamento no que diz respeito aos procedimentos.

A hipótese que norteia esta pesquisa é de que a confiabilidade do instrumento de coleta é satisfatória, havendo oportunidade de incremento da medida a partir da exclusão de itens

2. REFERENCIAL TEÓRICO

2.1 O FASCISMO

Um erro tão crasso quanto comum é reduzir o Fascismo à dimensão política e localizá-lo estritamente nos territórios alemães e japoneses. Reich (1988 p.11) define o Fascismo como “a atitude emocional básica do homem oprimido da civilização autoritária da máquina, com sua maneira mística e mecanicista de encarar a vida”.

A impossibilidade de conciliar a expansão internacional do Capital e a preservação das identidades das nações, provoca o conflito entre essas forças. A resposta nacional ao transnacionalismo liberal tende à ruptura entre estes entes (MARCUSE, 2001).

O antagonismo entre a universalidade que se atribui ao Capital e ao interesse particular da nação é falso, uma vez que o Fascismo não ataca as bases de manutenção do modo de produção Capitalista. As burguesias nacionais se amparam nas benesses dos estados fascistas quando existe algum ruído na manutenção de seus interesses. A base social da massa fascista era constituída da elite que buscava diferenciação daqueles que eram apontados como a raiz do problema, os banqueiros judeus, e aqueles atingidos pela crise, a classe média, e parte da classe operária (ADORNO & HORKHEIMER, 1985).

A crítica fascista ao Capital é mentira manifesta, uma vez que um dos seus propósitos é de justamente naturalizar a posse como um direito do mais forte, ou seja, a raça superior. Mas essa naturalidade da superioridade natural é tão mentirosa quanto a universalidade do Capital. “A raça, hoje, é a autoafirmação do indivíduo burguês integrado à coletividade bárbara” (ADORNO & HORKHEIMER, 1985 p.158).

O Fascismo é a expressão de toda a irracionalidade do caráter do homem médio. Atribuir a teoria de raça fascista meramente ao preconceito é incorrer em erro; esta teoria é condicionante do Fascismo e não sua consequência. O Fascismo é a expressão racional, materializada em sua forma política, da irracionalidade do ódio racial. (REICH, 1981).

Não é possível neutralizar o Fascismo apenas pela esfera política. É necessário entender onde o Fascismo vive no indivíduo. Neste ponto, deve-se notar a afirmação, deveras corajosa, de Reich (1981, p.12) no sentido de que “não existe um único indivíduo que não

seja portador, na sua estrutura, de elementos do pensamento e do sentimento fascistas”.

2.2 A PERSONALIDADE AUTORITÁRIA E A ESCALA F

A personalidade superficial do homem médio expressa-se como atenciosa e responsável. É no nível da subconsciência que se alojam impulsos menos nobres como a crueldade e o sadismo. É a repressão de tudo que é antissocial. Desta forma, qualquer impulso advindo do cerne biológico tenha que atravessar essas pulsões de perversidade e ser distorcido por elas até que se materializem em ação (REICH, 1988). É nessa zona intermediária de pulsos que reside o arcabouço da irracionalidade do homem médio, onde se verificam os traços da Personalidade Autoritária. Esta personalidade relaciona-se intimamente com a ideologia fascista, e atitudes de sujeitos propensos ao Fascismo podem ser descritas como

traços de uma personalidade Autoritária (TEIXEIRA & POLO, 1975).

Esse é o contexto onde se inserem os questionamentos de Adorno et al. (1950 p.2): “Se esse indivíduo potencialmente fascista existe, como ele é, precisamente? [...] Como se constitui seu pensamento antidemocrático?” Essas dúvidas orientaram a produção de uma série de estudos acerca da natureza da constituição psíquica do indivíduo potencialmente fascista. Como não estudaram nessa obra sujeitos declaradamente fascistas, o estudo assim refere-se aos que carregam em si a base para a introjeção da propaganda fascista (ADORNO et al., 1950).

Valendo-se de um sofisticado arcabouço de técnicas de pesquisa, os autores demonstraram que existe correlação entre determinados traços de personalidade e o preconceito. O Quadro 1 apresenta os traços da Personalidade Autoritária:

Quadro 1: Traços da Personalidade Autoritária.

Traço (continua)	Descrição das Características (continua)
Convencionalismo	Adesão ao convencional. A hipótese é de que a inclinação para o Fascismo é característica de quem gravita em torno da classe média; por outro lado, os indivíduos não convencionais são mais inclinados a serem livres de preconceitos. O convencionalismo autoritário é aquele que resulta não de escolha espontânea, mas de pressões sociais externas.
Agressividade Autoritária	Facilidade em controlar, vigiar, repelir, condenar, punir quem violar as normas convencionais. É a carga de hostilidade contra algo para qual não se pode canalizar, é despejada em outros objetivos “razoavelmente justificáveis”. O grupo minoritário a que são dirigidos tais sentimentos “oferece boas razões” para ser alvo, guiado por um senso de moralidade e de patriotismo distorcidos.
Destruição e Cinismo	Hostilidade e desprezo por tudo que é humano. Percebeu-se neste pessimismo universal a necessidade de descobrir justificativas posteriores à própria hostilidade e aos outros impulsos interiores não-aceitos.
Poder e Rudeza	Ênfase nas dimensões “dominação-submissão”. Identificação com figuras de poder. Acentuação dos atributos convencionais do ego. Afirmações drásticas relativas aos conceitos de força e rudeza. Decorrente de possível falta de força interior, a que se procura suprir apoiando-se em estruturas poderosas. Assim, o fraco sente-se forte
Projeção	Capacidade de “transferir” problemas interiores para o mundo exterior (impulsos, tabus, fraquezas, temores, responsabilidades). Atinge por vezes formas mesmo ridículas de juízo na “maré montante” do mal, dos vícios e dos perigos que “mais do que nunca” ameaçam o mundo, no que interessa aos dois campos “privilegiados”: agressividade (crime) e impulsos reprimidos.
Traço (conclusão)	Descrição das Características (conclusão)
Superstição e Estereotipia	Crença nos destinos fatais do homem e indivíduo, a propensão para raciocínios esquematizados e categorizados. Fuga da responsabilidade para um álibi sobrenatural. Cria-se um halo místico em torno dos líderes. A estereotipia age como ferramenta de reducionismo da realidade. Materializado nos Slogans surreais utilizados pela propaganda.
Sexualidade	Atitude exageradamente preocupada do autoritário com relação aos atos da sexualidade constitui as características salientes no quadro dos sintomas da Personalidade Autoritária. A repressão e a ansiedade consequentes são as matrizes prováveis de tais excessos.
Traço (conclusão)	Descrição das Características (conclusão)

Fonte: Adaptado de Teixeira & Polo (1975)

Outro ponto a ressaltar é que o estudo conseguiu produzir um instrumento para medir estas características em estratos da população. A Escala de Fascismo pode ser descrita como uma evolução das Escalas de Etnocentrismo (Levinson, 1950a) e de Conservadorismo Econômico (Levinson, 1950b), sendo a Escala F o instrumento principal do estudo. Seus objetivos era avaliar, ainda que indiretamente, o preconceito etnocêntrico e as propensões dos sujeitos em ter uma visão fascista da vida (TEIXEIRA & POLO, 1975). A validação desta escala foi realizada via comparação de entrevistas com sujeitos da Alto e baixo Score.

A noção de preconceito não pode se construir dissociadamente das noções de sociedade, grupo, indivíduo, racismo e cultura. O preconceito engloba a ideologia autoritária à suas formas psíquicas, impondo atitudes sobre minorias, sejam elas religiosas, étnicas, políticas, econômicas ou de sexualidade. O preconceito individualizado ganha força quando encontra grupos, manifestando uma visão distorcida da realidade, ao emitir opinião formada antecedendo ao conhecimento real dos fatos ou assunto.

O sujeito preconceituoso é um personagem opressivo e intransigente. Autoritários tendem a ter posturas severas e tradicionais. Ao observar o mundo com esse tipo de comportamento, não confia no ser humano, receia e repudia todos os grupos da sociedade os quais não fazem parte, sendo assim o preconceito apresenta-se atrás da sua desconfiança e do receio (ADORNO, 1950).

O preconceito se instala nos sujeitos a partir da relação entre seus conflitos psíquicos e a defesa contra estes, o pensamento estereotipado, o que denota sua gênese aos elementos culturais. Por isso, nota-se que, de certo modo, o preconceito independe do objeto a qual se destina, mantendo relações com o estereótipo em que o indivíduo se baseia para mediar a relação com o objeto de seu preconceito (CROCHÍK, 1996).

2.3 A CONFIABILIDADE NA TEORIA CLÁSSICA DOS TESTES

Atualmente, duas correntes teóricas que amparam o pensamento em Psicometria: A Teoria de Resposta ao Item (TRI) e a Teoria Clássica dos Testes (TCT).

A TRI nasceu na década de 1950 e vale-se de um arcabouço estatístico sofisticado, onde a computação é quase imprescindível. Com o incremento da capacidade de processamento, sua utilização vem crescendo nos últimos quinze anos, sua aplicabilidade ainda encontra obstáculos nos horizontes da estimação e no horizonte da automação (MARQUES, 2000). A TRI “propõe modelos para os traços latentes, ou seja, características do indivíduo que não se podem ser observadas diretamente” (ANDRADE, TAVARES & VALLE, 2000 p.3)

Um ponto em comum entre estas abordagens é que os parâmetros de Validade e Confiabilidade do Teste ocupam posição de protagonismo. Validade refere-se ao procedimento estruturado a partir de várias fontes para justificar a adesão ou rejeição das interpretações propostas pelas pontuações dos instrumentos. É vista como (BECKHAM & COOK, 2006).

Na literatura brasileira não existe consenso sobre a nomenclatura desta medida: precisão de um teste: Lebrão (1978) refere-se a medida como “Fidedignidade”. Em Freitas, Noronha e Otatti (2003) o parâmetro é nomeado como “Confiabilidade”. Gouvea et al. (2014) o termo utilizado é “Consistência Interna”. Pasquali (2009) nota que a terminologia pode estar relacionada com a estatística utilizada para a análise dos dados. Neste estudo adotou-se o termo Confiabilidade.

Ao explorar a noção de Confiabilidade, percebe-se que o termo original vem da língua inglesa, a saber, Reliability. A American Psychology Association (2016) define Reliability como “The degree to which a test produces similar scores each time it is used; stability or consistency of the scores produced by an instrument”.

2.4 A CONFIABILIDADE NA TEORIA CLÁSSICA DOS TESTES

Atualmente, duas correntes teóricas que amparam o pensamento em Psicometria: A Teoria de Resposta ao Item (TRI) e a Teoria Clássica dos Testes (TCT).

A TRI nasceu na década de 1950 e vale-se de um arcabouço estatístico sofisticado, onde a computação é quase imprescindível. Com o incremento da capacidade de processamento, sua utilização vem crescendo nos últimos quinze anos, sua aplicabilidade

ainda encontra obstáculos nos horizontes da estimação e no horizonte da automação (MARQUES, 2000). A TRI “propõe modelos para os traços latentes, ou seja, características do indivíduo que não se podem ser observadas diretamente” (ANDRADE, TAVARES & VALLE, 2000 p.3)

Um ponto em comum entre estas abordagens é que os parâmetros de Validade e Confiabilidade do Teste ocupam posição de protagonismo. Validade refere-se ao procedimento estruturado a partir de várias fontes para justificar a adesão ou rejeição das interpretações propostas pelas pontuações dos instrumentos. É vista como (BECKHAM & COOK, 2006).

Na literatura brasileira não existe consenso sobre a nomenclatura desta medida precisão de um teste: Lebrão (1978) refere-se a medida como “Fidedignidade”. Em Freitas, Noronha e Otatti (2003) o parâmetro é nomeado como “Confiabilidade”. Gouvea *et al.* (2014) o termo utilizado é “Consistência Interna”. Pasquali (2009) nota que a terminologia pode estar relacionada com a estatística utilizada para a análise dos dados. Neste estudo adotou-se o termo Confiabilidade.

Ao explorar a noção de Confiabilidade, percebe-se que o termo original vem da língua inglesa, a saber, *Reliability*. A *American Psychology Association* (2016) define *Reliability* como “*The degree to which a test produces similar scores each time it is used; stability or consistency of the scores produced by an instrument*”.

A TCT é a teoria fundamental para o pensamento psicométrico. Teve nas contribuições de Spearman (1904) sua gênese. O argumento da constituição do Score observado (X_i) de uma a escala se dar a partir do somatório do Score real (T_i) e do Score do Erro (E_i), não necessariamente sendo erro estritamente atribuído às variáveis em tela (PASQUALI, 2003)

Primi (2012) comenta que a discrepância entre o X_i e T_i é a medida de erro E_i , medida essa que é não-observável e estocástica, podendo assim assumir valores negativos e positivos o que faz com que T_i possa vir a ser maior ou menor que X_i . Neste sentido pode-se estabelecer os seguintes postulados: em grandes conjuntos E_i será igual a 0; não existe correlação entre X_i e T_i e não há correlação entre os E_i de testes paralelos. O que serve de base para as deduções apresentadas no Quadro 2

Quadro 2: Postulados da Teoria Clássica dos Testes

Dedução	Representação Matemática
Em dada população, a média dos <i>Scores</i> observados é a média dos <i>Scores</i> verdadeiros.	Quando $\overline{E_i} = 0$; $X_i = T_i + E_i \Rightarrow \overline{X_i} = \overline{T_i} + \overline{E_i} \therefore \overline{X_i} = \overline{T_i}$
Se $\rho_{X_i E_i} = 0$ a variância (σ_x^2) de X_i se dá do seguinte modo	$\sigma_x^2 = \sigma_T^2 + \sigma_E^2 + 2\rho_{XE}\sigma_T\sigma_E = \sigma_x^2 = \sigma_T^2 + \sigma_E^2 + 2(0)\sigma_T\sigma_E \therefore$ $\sigma_x^2 = \sigma_T^2 + \sigma_E^2$
Como não existe correlação entre o <i>Score</i> de erro e o <i>Score</i> verdadeiro, a variância do <i>Score</i> observado não sofrerá contribuição da covariância entre o <i>Score</i> verdadeiro e de erro. Sobrando apenas a variância dos <i>Scores</i> observados.	$\sigma_{TX}^2 = \frac{\sum_{i=1}^N (T_i - \overline{T})(X_i - \overline{X})}{N} = \frac{\sum_{i=1}^N [(T_i - \overline{T})(T_i - E_i) - \overline{T}E_i]}{N}$ $\sigma_x^2 = \frac{\sum_{i=1}^N [(T_i^2 - T_i\overline{T} + T_iE_i - \overline{T}T_i + \overline{T}^2 - \overline{T}E_i)]}{N}$ $\sigma_{TX}^2 = \frac{\sum_{i=1}^N (T_i^2 - 2T_i\overline{T} + \overline{T}^2) + \sum_{i=1}^N T_iE_i - N\overline{T}\sum_{i=1}^N E_i}{N}$ $\sigma_{TX}^2 = \frac{\sum_{i=1}^N (T_i^2 - 2T_i\overline{T} + \overline{T}^2) + 0 - 0}{N}$ $\sigma_{TX}^2 = \frac{\sum_{i=1}^N (T_i - \overline{T})^2}{N} = \sigma_T^2 \therefore \sigma_{TX}^2 = \sigma_T^2$
Aplicando esta dedução no cálculo da correlação entre os <i>Scores</i> verdadeiro e observado nota-se que esta é a razão entre os desvios dos <i>Scores</i> do erro e observado	$\rho_{TX} = \frac{\sum_{i=1}^N (T_i - \overline{T})(X_i - \overline{X})}{N\sigma_T\sigma_X} = \frac{\sum_{i=1}^N (T_i - \overline{T})(X_i - \overline{X})}{N} \cdot \frac{1}{\sigma_T\sigma_X}$ $\sigma_T^2 \cdot \frac{1}{\sigma_T\sigma_X} = \frac{\sigma_T^2}{\sigma_T\sigma_X} = \frac{\sigma_T}{\sigma_X} \therefore \rho_{TX} = \frac{\sigma_T}{\sigma_X}$
A grandeza para a correlação da dedução anterior são os desvios de cada variável. Entretanto, a precisão expressa-se por variâncias.	$\rho_{TT} = (\rho_{TX})^2 = \left(\frac{\sigma_T}{\sigma_X}\right)^2 = \frac{\sigma_T^2}{\sigma_X^2} \therefore \rho_{TT} = \frac{\sigma_T^2}{\sigma_X^2}$

Fonte: Adaptado de Sijtsma (2009).

Neste sentido, é confortável afirmar que conhecendo as variâncias dos *Scores* verdadeiros e observados, pode-se estimar a precisão do teste. Sabe-se que o *Score* verdadeiro não é uma grandeza observável. Para estimar este índice, é necessário se amparar no conceito de testes paralelos,

assumidos que hajam dois testes em evidência, a e b, e que estes tenham mesma média, variância e correlação entre os as formas (GUILFORD, 1954). Primi (2012) nota que o cálculo da correlação entre os *Scores* observados segue a seguinte métrica:

$$\rho_{ab} = \frac{\sum_{i=1}^N (T_{ia} - \bar{T})(X_{ib} - \bar{X})}{N\sigma_a\sigma_b} = \frac{\sum_{i=1}^N ((T_{ia} + E_{ia}) - \bar{T})((T_{ib} + E_{ib}) - \bar{T})}{N\sigma_a\sigma_b} =$$

$$\rho_{ab} = \frac{\sum_{i=1}^N [T_i^2 + T_i E_{ib} - T_i \bar{T} + E_{ia} T_i + E_{ia} E_{ib} - E_{ia} \bar{T} - \bar{T} T_i - \bar{T} E_{ib} + \bar{T}^2]}{N\sigma_a\sigma_b}$$

Dados os critérios: $\overline{E_{ia}} = 0, \overline{E_{ib}} = 0, \rho_{X_{ia}E_{ia}} = 0, \rho_{X_{ib}E_{ib}} = 0$, a correlação se dá por:

$$\rho_{ab} = \frac{\sum_{i=1}^N [T_i^2 - 2T_i \bar{T} + \bar{T}^2]}{N\sigma_a\sigma_b} = \frac{\sum_{i=1}^N (T_i - \bar{T})^2}{N\sigma_a\sigma_b} = \frac{\sigma^2}{\sigma_a\sigma_b}$$

Como $\sigma_a = \sigma_b = \sigma_X$, observa-se que $\rho_{ab} = \rho_{TT} = \frac{\sigma_T^2}{\sigma_X^2}$

A literatura psicométrica brasileira não existe consenso sobre a nomenclatura desta medida precisão de um teste: Lebrão (1978) refere-se a medida como “Fidedignidade”. Em Freitas, Noronha e Otatti (2003) o parâmetro é nomeado como “Confiabilidade”. Gouvea et. al (2014) o termo utilizado é “Consistência Interna”. Pasquali (2009) nota que a terminologia pode estar relacionada com a estatística utilizada para a análise dos dados. Neste estudo adotou-se o termo Confiabilidade.

Ao explorar a noção de Confiabilidade, percebe-se que o termo original vem da língua inglesa, a saber Reliability. A American Psychology Association (2016) define Reliability como “The degree to which a test produces similar scores each time it is used; stability or consistency of the scores produced by an instrument”.

Gochyyev & Wilson (2013) definem Reliability como “an index of how consistently a test measures whatever it. [...] If the instrument is sufficiently reliable, then the measurer can assume that measurement errors are sufficiently small to justify using the observed score.”

Pasquali (2003, p.192) estabelece que Confiabilidade “diz respeito à característica que ele deve possuir, a saber, de medir sem erros [...] significa que o mesmo teste, medindo os mesmos sujeitos em ocasiões diferentes, ou testes equivalentes, medindo os mesmos sujeitos na mesma ocasião, produzam resultados idênticos.”. Beckham & Cook (2006) estabelecem os métodos para a obtenção da Confiabilidade de um teste demonstrados no Quadro 3

Quadro 3: Fontes, descrições, estatísticas e definições do parâmetro de Confiabilidade.

Fonte	Descrição	Estatística	Definição
Consistência Interna	Todos os itens de um instrumento medem o mesmo construto? (Se um instrumento medir mais de um construto, apenas um <i>Score</i> não vai medir corretamente cada construto.) *	Método das metadespartidas	Análise das correlações entre as metades do instrumento
		Coeficiente de Kuder-Richardson	Equivalente ao método das metadespartidas, para respostas dicotômicas
		Alfa de Cronbach	Uma generalização da fórmula de Kuder-Richardson
Estabilidade Temporal	O instrumento produz resultados similares quando aplicado uma segunda vez no sujeito?	Método do Teste-Reteste	Administrar o Instrumento duas vezes na mesma pessoa com dado intervalo de tempo.
Testes Paralelos	Diferentes versões do instrumento produzem resultados iguais?	Formas alternativas de Confiabilidade	Administrar versões diferentes do mesmo instrumento simultaneamente ou separado
Confiabilidade Inter Avaliador (Concordância)	Quando da utilização de Avaliadores, existe diferença entre os <i>Scores</i> dos Avaliadores?	Concordância	Percentual de respostas idênticas
		Phi	Correlação simples
		Kappa	Concordância corrigida para o acaso
		Tau de Kendall	Concordância para dados classificados
		Correlação Intraclass	Análise da variância para estimar as coincidências das diferentes Avaliações
Teoria da Generalização	Quanto o erro de medição contribui para cada fator?	Coeficiente de Generalização	Modelo Sofisticado, que permite a estimação de múltiplas fontes de erro
*Medida para instrumentos que medem um único construto. Instrumentos com múltiplos construtos, podem ter a Consistência verificada via metodologia de Instrumentos Paralelos.			

Fonte: Adaptado de Beckham (2006).

Para entender como se dá a utilização dos diferentes modos de extração da Confiabilidade dos testes, este estudo verificou a distribuição das estatísticas utilizadas em pesquisas publicadas no Repositório Scielo refinando resultados conforme o quadro 3 Áreas Temáticas:

Saúde, Ciências Humanas e Sociais; Período de Publicação: entre 2010 e 2016; Metodologia de extração da Confiabilidade: Método da Consistência Interna e Estabilidade Temporal. Os resultados estão relatados na Figura 1

Figura 1: Consistência Internas nas Áreas da Saúde, Ciências Humanas e Ciências Sociais publicadas no Repositório SCIELO, em porcentagem

Fonte: Pesquisa de Campo.

Os resultados desta observação apontam para uma tendência de hegemonia na

utilização do Alfa de Cronbach (α) quando o método escolhido para a verificação da

Confiabilidade do Teste é a verificação da Consistência Interna. Mesmo hegemônica, a utilização do Alfa de Cronbach encontra algumas críticas. Sijtsma (2009) argumenta que a Confiabilidade de um teste está além da análise da Consistência Interna.

Face a isto, o presente estudo se justifica no sentido de tornar clara as bases para aplicação de tal ferramenta estatística, uma vez que as contribuições de Warmbold (2001) revelam que apenas 7% das pesquisas

publicadas na edição 41 do *The Journal of Agricultural Education* fizeram o uso correto da estatística Alfa de Cronbach na verificação da Consistência Interna.

2.5 O ALFA DE CRONBACH

Kuder & Richardson (1937) desenvolveram uma estatística para o teste de Confiabilidade de testes com variáveis dicotômicas, a saber:

$$r_{TT} = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^k p_i q_i}{\sigma_x^2} \right); (i = 1, 2 \dots n)$$

Onde p_i é a proporção de itens que receberam resposta 1, e q_i é proporção de

itens que receberam resposta 0. Cronbach (1951) apresentou uma adequação da fórmula para itens que não sejam dicotômicos, conforme a fórmula abaixo:

$$r_{tt} = \frac{n}{n-1} \left(1 - \frac{\sum_i V_i}{V_t} \right)$$

Onde V_t é a variância do Score do teste e V_i é a variância do item em evidência. A adequação proposta por Cronbach se demonstra mais eficiente que em um comparativo com as fórmulas utilizadas nos métodos de metades-partidas, que sofreram severas críticas acerca do problema (CRONBACH, 1951). Uma vez que o Alfa não trabalha questões relativas à homogeneidade das variância e covariâncias inter-itens, permite determinar o limite inferior da consistência interna de um grupo de ou itens, que será a correlação esperada entre a escala usada e outras escalas paralelas (NUNES & PESTANA, 2008).

É prudente notar que o Alfa sofre influência da homogeneidade da amostra, da distribuição dos Scores da escala e do número de itens do teste. Nota-se que quanto maior a variabilidade das respostas dos sujeitos, maior será o Alfa, o que remete à necessidade de haver uma amostra

suficientemente heterogênea e corretamente dimensionada em um Teste (MAROCO & MARQUES, 2006).

3 PROCEDIMENTOS

O teste utilizou os, descritos aqui em ordem de apresentação, seguintes instrumentos: Questionário de dados gerais dos sujeitos: constava de questões relacionadas à idade, religião, curso, cor da pele e se a pessoa era portadora de necessidades especiais ou se tinha ou tinha contato com pessoas que apresentam deficiências Intelectuais. Não havia a necessidade de identificação. As informações foram autodeclaradas. Escala F: A versão da Escala F utilizada foi proposta por Adorno et al (1950). e adaptada por Crochik et. al na sua versão com 16 itens. Cada questão fora nomeada com um código. O Quadro 4 apresenta a codificação:

Quadro 4: Nomes das variáveis relativas aos itens da Escala F.

Nome	Itens da Escala
F06	Algum dia se provará talvez que a astrologia pode explicar muitas coisas
F08	A obediência e o respeito à autoridade são as principais virtudes que devemos ensinar a nossas crianças.
F10	Um indivíduo de más maneiras, maus costumes e má educação dificilmente pode fazer amizade com pessoas decentes.
F14	Este país necessita, primordialmente, antes de leis ou planos políticos, é de alguns líderes valentes, incansáveis e devotos em quem o povo possa depositar a sua fé.
F18	Não se concebe nada mais baixo do que uma pessoa que não sente profundo amor, gratidão e respeito por seus pais.
F23	Os homens podem ser divididos em duas classes definidas: os fracos e os fortes.
F29	Só por meio do sofrimento se aprendem as coisas verdadeiramente importantes.
F31	A ciência tem o seu lugar, mas há muitas coisas importantes que a mente humana jamais poderá compreender.
F34	Às vezes, os jovens têm ideias rebeldes que, com os anos, deverão superar para assentar os seus pensamentos.
F37	Se falássemos menos e trabalhássemos mais, todos estaríamos melhor.
F39	Todos devemos ter fé absoluta em um poder sobrenatural, cujas decisões devemos acatar.
F41	Os homossexuais são quase criminosos e deveriam receber um castigo severo.
F43	Nenhuma pessoa decente, normal e em seu são juízo pensaria em ofender um amigo ou parente próximo.
F46	Hoje em dia, as pessoas se intrometem cada vez mais em assuntos que deveriam ser estritamente pessoais e privados.
F53	Crimes sexuais tais como o estupro ou ataques a crianças merecem mais que prisão; quem comete esses crimes deveria ser açoitado publicamente ou receber um castigo pior.
F57	A maioria de nossos problemas sociais estaria resolvida se pudéssemos nos livrar das pessoas imorais, dos marginais e dos débeis mentais.
F60	Quando alguém tem problemas ou preocupações, é melhor não pensar neles e se ocupar de coisas mais agradáveis.

Fonte: Crochík (2009).

As respostas possíveis eram dispostas da seguinte forma: Concordância Plena, Concordância Moderada, Concordância Leve, Discordância Leve, Discordância Moderada, Discordância Plena. Nesta sequência, fora atribuído à cada pergunta um valor, a saber: 1,2,3,5,6 e 7. A ferramenta de automação estatística utilizada para a automação dos cálculos foi o IBM SPSS, programa especializado em análise quantitativa de dados. Sua escolha se deu por conta da interface amigável e pela facilidade de tratar grande volume de dados (MOOI & SARSTEDT, 2011).

4 RESULTADOS

De modo censitário, foram entrevistados 59 alunos concluintes do curso de secretariado, todos do sexo feminino com média de idade de 24,75 anos. Os aspectos socioculturais da

amostra revelam uma população majoritariamente não branca, cristã, não portadora de necessidades especiais e não portadoras de deficiência intelectual. O tempo médio de aplicação foi de 35 minutos. O teste de Alfa de Cronbach retornou o valor de 0,705, considerado “Aceitável” a partir do critério proposto por Nunnally (1978) onde resultados superiores a 0,7 para o estágio exploratório do problema são considerados aceitáveis, embora patamares mais favoráveis estejam em coeficientes a partir de 0,8. O tema da interpretação do Alfa será objeto de discussão em momento oportuno neste estudo. O SPSS possibilita a análise de cenários acerca da Confiabilidade apresentando o Alfa de Cronbach da escala caso o item seja excluído. A Tabela 1 apresenta os resultados obtidos:

Tabela 1: Média, Desvio Padrão (σ) e Alfa de Cronbach da Escala F em sua versão com 16 itens.

Item	Média	σ	Alfa se excluído	Item	Média	σ	Alfa se excluído
F06	4,5	2,1	0,721	F37	2,7	2,0	0,682
F08	2,9	2,1	0,701	F39	4,4	2,3	0,689
F10	4,2	2,1	0,696	F41	6,6	1,2	0,719
F14	4,2	2,3	0,686	F43	4,2	2,4	0,694
F18	3,5	2,4	0,676	F46	3,1	2,4	0,676
F23	4,6	2,1	0,685	F53	3,6	2,2	0,685
F29	4,1	2,1	0,689	F57	5,6	1,9	0,693
F31	2,6	1,9	0,702	F60	3,3	2,2	0,680
F34	3,4	1,9	0,689	Escala	3,9	2,0	-

Fonte: Pesquisa de campo.

A avaliação dos dados revela que a exclusão do item F06 contribui para o incremento da

Confiabilidade do instrumento. Portanto o item em tela será excluído do instrumento.

Tabela 2: Média, Desvio Padrão (σ) e Alfa de Cronbach da Escala F em sua versão com 15 itens.

Item	Média	σ	Alfa se excluído	Item	Média	σ	Alfa se excluído
F08	2,9	2,1	,719	F37	2,7	2,0	0,696
F10	4,2	2,1	,717	F39	4,4	2,3	0,703
F14	4,2	2,3	,702	F41	6,6	1,2	0,733
F18	3,5	2,4	,693	F43	4,2	2,4	0,709
F23	4,6	2,1	,704	F46	3,1	2,4	0,694
F29	4,1	2,1	,708	F53	3,6	2,2	0,707
F31	2,6	1,9	,720	F57	5,6	1,9	0,708
F34	3,4	1,9	,706	F60	3,3	2,2	0,699
				Escala	3,9	2	-

Fonte: Pesquisa de campo

No cenário descrito, a exclusão do item F41 eleva o Coeficiente Alfa, o que justifica a exclusão desse item. Novamente, fora

calculado o Alfa para os itens que constam na versão da Escala F com 15 itens. O resultado desta rodada está expresso na Tabela 3:

Tabela 3: Média, Desvio Padrão (σ) e Alfa de Cronbach da Escala F em sua versão com 14 itens.

Item	Média	σ	Alfa se excluído	Item	Média	σ	Alfa se excluído
F08	2,9	2,1	0,732	F37	2,7	2,0	0,712
F10	4,2	2,1	0,730	F39	4,4	2,3	0,717
F14	4,2	2,1	0,715	F43	4,2	1,2	0,723
F18	3,5	2,3	0,706	F46	3,1	2,4	0,705
F23	4,6	2,4	0,719	F53	3,6	2,2	0,719
F29	4,1	2,1	0,722	F57	5,6	1,9	0,723
F31	2,6	2,1	0,732	F60	3,3	2,2	0,713
F34	3,4	1,9	0,718	Escala	3,7	2,0	-

Fonte: Pesquisa de campo

Ao observar o cenário estabelecido no Quadro 3 percebe-se que não há mais alternativa de incremento do Alfa a partir da

exclusão de itens da escala. O Alfa, portanto, é de 0,733.

5 CONSIDERAÇÕES FINAIS

Não é possível discutir as questões acerca da Confiabilidade desta aplicação da Escala F sem antes repercutir, mesmo sem a pretensão de realizar avaliações mais aprofundadas sobre estes, seus resultados. É necessário reafirmar a ciência de que essa proposição resulta em um deslocamento, ainda que temporário, do objeto desta pesquisa para uma posição secundária em favor da discussão, ainda que sem a devida densidade que esta temática necessita, dos resultados da aplicação da Escala. A medida de Confiabilidade teria um fim em si, caso optasse pela negação dessa reflexão.

Isto posto é desejável relembrar a pontuação atribuída a cada resposta: Concordância Plena, 1 ponto, Concordância Moderada recebeu 2 pontos, Concordância Leve, 3 pontos. A Discordância Leve, recebeu 5 pontos, a Discordância Moderada 6 pontos e 7 pontos para Discordância Plena.

A aplicação da Escala F revelou uma tendência da amostra para um ponto entre o ponto neutro da escala e a Concordância Leve, com Score médio de 3,7. Portanto, é possível identificar traços da Personalidade Autoritária na amostra. Uma análise superficial acerca destes resultados pode considerar que os traços marcantes na amostra são a Submissão Acrítica e dada medida de Agressividade Autoritária. Nesse sentido, cabe recomendar que tais análises sejam objeto de estudos mais aprofundados acerca dos traços da Personalidade Autoritária nesta população e seus condicionantes além de possíveis relações com adesão à ideologia, atitudes preconceituosas e outras atitudes.

Nas questões acerca da Confiabilidade dos testes, vale notar que esta se refere à precisão do instrumento utilizado na pesquisa. A literatura internacional mostrou-se mais homogênea em relação a adesão em relação à terminologia Reliability para se referir a este parâmetro. Na literatura brasileira, encontraram-se alguns termos para o parâmetro.

Apesar dos avanços da Teoria de Resposta ao Item dos últimos 15 anos, a Teoria Clássica dos Testes ainda é hegemônica. Baseia sua argumentação na assunção de que o Score observado se compõe do somatório do Score real e no Score do erro. Assume que esse erro não pode ser mensurado, sendo determinado como uma variável estocástico-aleatória.

É da razão entre os desvios-padrão da variância dos Scores verdadeiros e observados que se produz a precisão, ou seja, a Confiabilidade. Para isso, estabeleceu-se um cenário de testes paralelos onde as medidas idênticas de uma mesma variável com mesma média, variância e correlação entre os pares possíveis, estando ambas sujeitas a erros. Portanto, é confortável afirmar que a estimação da Confiabilidade de um teste está relacionada com a correlação de medidas paralelas.

Neste cenário, notam-se diversas abordagens para esta estimação, dentre os quais fora escolhido para o estudo do caso em evidência aqui, a análise da Consistência Interna, onde o Alfa de Cronbach se revela a estatística mais utilizada nos periódicos das ciências Psicossociais e de Saúde no Brasil. O Alfa se estabelece como uma generalização do coeficiente KR-20 e sendo mais eficiente que o método de metades-partidas.

Para descrever a aplicação e as estratégias de aplicação desta estatística, propôs-se um experimento de aplicação da Escala F na população descrita. A análise da Confiabilidade do instrumento retornou um coeficiente minimamente aceitável, a saber, 0,705. Neste ponto cabem algumas reflexões acerca da avaliação qualitativa da ferramenta quantitativa, que sempre estará a cargo do pesquisador.

A interpretação do Alfa é não é consensual na literatura. Apesar de amplamente utilizado na literatura, a interpretação do critério de Nunnally é alvo de críticas, que se amparam no fato de muitas pesquisas não atentam para o fato de o autor ter sido claro quando relaciona determinados coeficientes alfa com determinados estágios da pesquisa em observação. Interpretações menos severas, como aceitabilidade a partir de coeficientes maiores que 0,6 devem ser avaliadas criteriosamente em decorrência da fase e dos objetivos do estudo. Outra ponderação que deve ser observada é que coeficientes muito próximos de 1 podem revelar um instrumento redundante. O Alfa não pode ter um fim em si, cabendo ao pesquisador uma avaliação criteriosa e contextualizada do coeficiente.

Mesmo sendo predominante, o Alfa tem determinada limitação. Conforme observado na literatura, o Alfa não se apresenta como uma alternativa interessante de estimação da precisão do teste quando este se propõe às

escalas multidimensionais, onde outras abordagens de estimação da Confiabilidade devem ser consideradas, como a Estabilidade Temporal.

Ainda no tocante às limitações do Alfa, verifica-se que existem variáveis que influenciam o resultado do teste com o alfa, como o tamanho e homogeneidade da amostra, pode incrementar artificialmente seu resultado, o que pode levar o tomador de decisão a interpretações erradas.

No intuito de incrementar a confiabilidade do teste, a estratégia utilizada, todos os aspectos e variáveis apresentadas, foi a exclusão de itens do questionário, que resultou em um Alfa de 0,733 após a exclusão de 2 itens. Considera-se que mesmo não apresentando

resultados de grande expressividade, a estratégia de exclusão de item deve ser considerada em cenários onde seja necessário aumentar o Alfa.

Neste sentido e considerando todas as exposições acerca da noção de confiabilidade, é confortável afirmar que a hipótese que amparou este estudo foi confirmada. A Escala F se coloca como um instrumento válido e confiável para medir as potencialidades dos sujeitos à em uma interpretação mais sofisticada, demonstrar atitudes preconceituosas, podendo ser aplicado por organizações que se atentam para tais variáveis.

REFERÊNCIAS

- [1] Adorno, T. Consignas. Buenos Aires. Amorrortu Editores. 1993.
- [2] Adorno, T. et al. The Authoritarian Personality. Berkley. W.W. Norton Company 1950.
- [3] Adorno, T.W; Horkheimer, M. Dialética do Esclarecimento. Tradução de Guido de Almeida. Rio de Janeiro. Zahar 1985. 224 p.
- [4] American Psychological Association. Glossary. Disponível em: <http://www.apa.org/research/action/glossary.aspx?tab=17>. Acesso em 30 de abril de 2016.
- [5] Babbal, E. Métodos de Pesquisa Survey. Belo Horizonte . Editora UFMG.2003.
- [6] Beckman, T. J; Cook, D.A. Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. The American Journal of Medicine, N 119, ed.16 p.166.e7-166. Fev.2006
- [7] Crochik, J. L. et al . Relações Entre Preconceito, Ideologia E Atitudes Frente À Educação Inclusiva. Estud. psicol. (Campinas), Campinas, v. 26, n. 2, p. 123-132,Junho de 2009. Disponível em:
- [8] <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-166X2009000200001&lng=en&nrm=iso>. Acesso em 20 de maio de 2016. <http://dx.doi.org/10.1590/S0103-166X2009000200001>.
- [9] Crochik, J. L. et al . Relações entre preconceito, ideologia e atitudes frente à educação inclusiva. Estud. psicol. (Campinas), Campinas , v. 26, n. 2, p. 123-132, Junho de 2009 . Disponível em:
- [10] <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-166X2009000200001&lng=en&nrm=iso>. Acesso em 16 de Maio de 2016. <http://dx.doi.org/10.1590/S0103-166X2009000200001>.
- [11] Crochik, J. L. Notas sobre a nação ÉTICA-Psicologia. Psicol. cienc. prof., Brasília , v. 12, n. 2, p. 10-15, 1992 . Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1414-98931992000200003&lng=en&nrm=iso>. Acesso em 20 de maio de 2016. <http://dx.doi.org/10.1590/S1414-98931992000200003>.
- [12] Crochik, J. L. Preconceito, indivíduo e sociedade. Temas psicol. Ribeirão Preto , v. 4, n. 3, p. 47-70, dez. 1996 . Disponível em <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S141389X1996000300004&lng=pt&nrm=iso>. Acesso em 27 de maio de 2016.
- [13] CRONBACH, J.L. Coefficient Alpha And The Internal Structure Of Tests. Madison. Psychometrika. Vol. 16, No. 3 Set., 1951.
- [14] Gerrig, R J.; Zimbardo, P.G. Psychology And Life.16° ed Boston. Pearson Education.2002.
- [15] Gil,A. C. Métodos e Técnicas da Pesquisa Social. 6° ed. São Paulo. Editora Atlas. 2008.
- [16] Gouveia, V V. et al. Escala de Procrastinação Ativa: evidências de validade fatorial e consistência interna. Psico-USF, Itatiba , v. 19, n. 2, p. 345-354, Ago. 2014 . Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-82712014000200016&lng=en&nrm=iso>. Acesso em 15 de maio de 2016. <http://dx.doi.org/10.1590/1413-82712014019002008>.
- [17] Guilford, J. P. Psychometric Methods. New York: McGraw-Hill.1954

- [18] Hubley, A. M.; Zumbo, B. D. A dialectic on validity: where we have been and where we are going. *Journal of General Psychology*. Nova York. v. 23, n.3, p.207-215, 1996.
- [19] Kuder, J. M Madison. The Theory Of The Estimation Of Test Reliability. Madison. *Psychometrika*. Vol. 2, No. 3 September, 1937.
- [20] Lebrao, M. L. Análise da fidedignidade dos dados estatísticos hospitalares disponíveis na Secretaria de Estado da Saúde de São Paulo em 1974. *Rev. Saúde Pública, São Paulo* , v. 12, n. 2, p. 234-249, jun. 1978 . Disponível em <http://www.scielo.org/scielo.php?script=sci_arttext&pid=S0034-89101978000200014&lng=pt&nrm=iso>. Acessos em 15 de mai. 2016. <http://dx.doi.org/10.1590/S0034-89101978000200014>.
- [21] Levinson, D. Politico-Economic Ideology And Group Memberships In Relation To Ethnocentrism. In: *The Authoritharian Personality*. Berkley. W.W. Norton Company 1950.
- [22] Levinson, D. The Study Of Ethnocentric Ideology. In: *The Authoritharian Personality*. 1993. W.W. Norton Company
- [23] Marcuse, H. Algunas implicaciones sociales de la tecnologia moderna. In: *Guerra, Tecnologia e Fascismo. Colección Otraparte*. Medellín.. Editorial Universidad de Antioquia. 2001. 352 p.
- [24] Maroco, J; Marques, T.G. Qual a Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas? *Laboratório de Psicologia*, Lisboa, v. 4, n. 1, p. 65-90, 2006.
- [25] MARQUES, R. M. Apresentação. In: *Teoria Da Resposta Ao Item: Conceitos E Aplicações*. SINAPE. 2000.
- [26] Mooi, E; Sarstedt, M. A Concise Guide to Market Research. The Process, Data, and Methods Using IBM SPSS Statistics. Nova York. Springer.2011.
- [27] NUNNALLY, J. C. *Psychometric theory*.2º ed. Nova York. McGraw-Hill. 1978.
- [28] Ottati, F.; Noronha, A. P. Parâmetros psicométricos de instrumentos de interesse profissional. *Estud. pesqui. psicol.*, Rio de Janeiro , v. 3, n. 2, p. 37-50, jul. 2003 . Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1808-42812003000200003&lng=pt&nrm=iso>. Acesso em 15 de mai. 2016.
- [30] Pasquali, L. *Psicometria*. Rev. esc. enferm. USP, São Paulo , v. 43, n. spe, p. 992-999, Dec. 2009 . Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S008062342009000500002&lng=en&nrm=iso>. Acesso em 8 de mai. 2016. <http://dx.doi.org/10.1590/S0080-62342009000500002>.
- [31] Pasquali, L. *Psicometria. Teoria dos Testes na Psicologia e na Educação*. 5º Ed. Petropolis. editora Vozes. 2003.
- [32] Pasquali, L. *Teoria e métodos de medida em ciências do comportamento*. Organizado por Luiz Pasquali. — Brasília: Laboratório de Pesquisa em Avaliação e Medida do Instituto de Psicologia. Brasília. Editora UnB: INEP, 1996.432p.
- [33] Pestana, M. H.; Gageiro, J. N. Análise de Dados para Ciências Sociais. A complementaridade do SPSS. 5ª edição revista e corrigida. Lisboa. Edições Sílabo. 2008.p. 527-528.
- [34] Primi, R. *Psicometria: fundamentos matemáticos da Teoria Clássica dos Testes*. Aval. psicol., Itatiba , v. 11, n. 2, p. 297-307, ago. 2012 . Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712012000200015&lng=pt&nrm=iso>. Acesso em 26 de mai. 2016.
- [37] Reich, W. *Psicologia De Massas Do Fascismo*.2º Ed. São Paulo Martins Fontes. 1988.
- [38] Sijsma, K. On The Use, The Misuse, And The Very Limited Usefulness Of Cronbach's Alpha. *Madison. Psychometrika*—Vol. 74, No. 1, 107–120. 2009 Doi: 10.1007/S11336-008-9101-0
- [39] Spearman, C. General Intelligence," Objectively Determined and Measured". *Champaign. The American Journal of Psychology*. Vol. 15, No. 2 Abr., 1904, pp. 201-292.
- [40] Teixeira, J. B. C.; Polo, A. A Personalidade Autoritária: Componentes E Gênese Psicológica. *Arquivos Brasileiros Psicologia Aplicada*, 1975. Disponível em :<<http://bibliotecadigital.fgv.br/ojs/index.php/abpa/article/viewFile/17538/16282>>. Acesso em 30 de abril de 2016.

Capítulo 6

USO DO MÉTODO ANALYTIC HIERARCHY PROCESS (AHP) PARA ESCOLHA DE FORNECEDOR DE FARINHA DE TRIGO: UM ESTUDO DE CASO

Pâmela Braga Bertahone

Nilson Brandalise

Resumo: Esse trabalho tem como objetivo auxiliar no processo de escolha entre três fornecedores, de farinha de trigo, para uma Padaria e Confeitaria, localizada na cidade de Barra do Piraí – RJ. Para isso, foi utilizado um método de auxílio à decisão sob múltiplos critérios, o Método da Análise Hierárquica (AHP). O estudo de caso foi realizado por pesquisa, com a participação de funcionários do nível operativo, que responderam questionário contendo 4 (quatro) fatores, que são: entrega, prazo, qualidade e preço, para 3 (três) alternativas, sendo analisado no software SuperDecision, que é uma ferramenta de suporte à decisão que implementa o AHP. Além disso, houve uma pesquisa bibliográfica sobre como funciona esse método. A principal contribuição do trabalho consiste nos resultados encontrados, que identificaram qual é o melhor fornecedor de farinha de trigo para a padaria estudada, contribuindo para o seu melhor desempenho.

Palavras-chave: Padaria, Método AHP, SuperDecision, Tomada de Decisão.

1. INTRODUÇÃO

Toda empresa busca um conjunto de características muito importantes ao realizar uma compra: entrega, prazo, qualidade e preço. Com isso, vem a necessidade de escolher um fornecedor que atenda aos seus pré-requisitos. Se o produto é entregue com segurança e rapidez, apresenta boas condições de pagamento, possui um preço justo e competitivo e é de boa qualidade, então o fornecedor se torna competitivo.

Uma boa relação do consumidor com o fornecedor é essencial em qualquer empresa. De acordo com o artigo 2º, da Lei Nº 8078 (Código de Defesa do Consumidor), de 11 de setembro de 1990, consumidor é toda pessoa física ou jurídica que adquire ou utiliza produto ou serviço como destinatário final. Já o artigo 3º da mesma lei, diz que o fornecedor é toda pessoa física ou jurídica, pública ou privada, nacional ou estrangeira, bem como os entes despersonalizados, que desenvolvem atividade de produção, montagem, criação, construção, transformação, importação, exportação, distribuição ou comercialização de produtos ou prestação de serviços.

Dessa forma, entende-se que o fornecedor é aquele que fornece produtos e presta serviços a terceiros. Segundo Bertaglia (2006), “o processo de seleção de fornecedor não é simples. A complexidade aumenta em função das características do item ou serviço a ser comprado, pois as exigências podem ser maiores ou menores”. Dessa forma, o processo de decisão torna-se muito difícil. Deve-se buscar escolher aquele que atenda melhor as suas expectativas, tenha a melhor avaliação ou desempenho. Para isso, é de grande importância que a escolha seja apoiada em algum método consolidado.

Com o objetivo de colocar essa questão em prática, esse trabalho utilizou um método de auxílio à decisão sob múltiplos critérios – o Método da Análise Hierárquica (*Analytic Hierarchy Process* - AHP). Ele foi empregado em um problema de escolha sobre o melhor fornecedor de farinha de trigo para uma Padaria e Confeitaria, localizada na cidade de Barra do Piraí – RJ.

O setor de panificação foi escolhido, pois além de ter uma relação direta com fornecedores, é uma área de grande importância no Brasil. De acordo com o Sistema FIRJAN, esse setor

vem ocupando posição de destaque no *ranking* dos maiores setores industriais do país, considerando-se o número de estabelecimentos e a geração de empregos.

A decisão de usar o método AHP está ligada à praticidade que ele oferece, por representar rapidez na escolha do fornecedor e por apresentar uma base teórica consolidada. Aliado a isso, também foi usado o *software SuperDecision*, como auxílio ao emprego e análise do AHP. Dessa forma, percebe-se que este trabalho é muito interessante, pois o método foi aplicado em um caso real, e foi possível observar os resultados na panificadora, posteriormente.

Em suma, este artigo está organizado da seguinte forma: a seção 2 explica o método AHP; a seção 3 descreve como a pesquisa foi realizada; a seção 4 apresenta o estudo de caso descrito anteriormente com as características dos fornecedores a serem analisados; a seção 5 apresenta a discussão e os resultados encontrados; e, finalmente, a seção 6 apresenta a conclusão.

2. FUNDAMENTAÇÃO TEÓRICA

O método AHP (*Analytic Hierarchy Process*) foi criado por Tomas L. Saaty, na década de 70, baseado na matemática e psicologia e é um método de multicritério amplamente utilizado e conhecido atualmente. O AHP foi desenvolvido com a finalidade de auxiliar nos processos de tomadas de decisão, planejamentos empresariais, resoluções de conflitos e diversas outras modalidades em que a multiplicidade de fatores está presente (SCHIMIDT, 1995).

A ideia principal do método é a comparação aos pares de diferentes situações. Segundo Saaty (1991), o método consiste na decomposição e síntese das relações entre os critérios até que se chegue a uma priorização dos seus indicadores, aproximando-se de uma melhor resposta de medição única de desempenho.

O *Decision Support Systems Glossary* (DSS, 2006) define AHP como: “uma aproximação para tomada de decisão que envolve estruturação de multicritérios de escolha numa hierarquia. O método avalia a importância relativa desses critérios, compara alternativas para cada critério, e determina um ranking total das alternativas”.

Além de possuir a finalidade de possibilitar uma avaliação subjetiva de um conjunto de alternativas selecionadas por meio de múltiplos critérios e organizada em uma estrutura hierárquica (SINUANY-STERN ET AL., 2000; AZEREDO, 2010), o método AHP, (SAATY, 1980) permite a atribuição de pesos relativos para múltiplas características, ou múltiplas alternativas para certa característica, enquanto permite a comparação pareada entre os atributos simultaneamente (AZEREDO, 2010). Essa possibilidade permite que, mesmo que as alternativas não tenham relações entre si, a alternativa mais importante seja reconhecida no processo de decisão.

Como todo método, o AHP apresenta suas vantagens e desvantagens. Segundo Saaty (1994), o benefício do método é que, como os valores dos julgamentos das comparações paritárias são baseados em experiência, intuição e também em dados físicos, o AHP pode lidar com aspectos qualitativos e quantitativos de um problema de decisão. Porém, uma das limitações do método é a sua aplicação inadequada, isto é, em ambientes desfavoráveis onde a aplicação é percebida como simplificação excessiva ou como desperdício de tempo (GRANDZOL, 2005).

Dessa forma, segundo Costa (2002, p. 16-17) este método baseia-se em três etapas de pensamento analítico:

(i) Construção de hierarquias: De acordo com Marins (2006), o problema deve ser estruturado em níveis hierárquicos descendentes formando uma hierarquia ou árvore de decisão no qual o objetivo geral deve ficar no topo da estrutura, sucedido dos critérios associados com o problema a ser solucionado. Tais critérios podem ser ainda diluídos em subcritérios, em um nível abaixo. Se houver necessidade, este processo poderá continuar até que todos os critérios sejam bem entendidos e especificados. Após os critérios, em um nível logo abaixo, são alocadas as alternativas.

De acordo com Bornia e Wernke (2001), a ordenação hierárquica possibilita ao decisor ter uma “visualização do sistema como um todo e seus componentes, bem como interações destes componentes e os impactos que os mesmos exercem sobre o sistema”. E a compreender de forma global, o problema e a relação de complexidade, ajudando na avaliação da dimensão e conteúdo dos critérios, através da comparação homogênea dos elementos.

A figura 1 apresenta a estrutura hierárquica básica do método AHP:

Figura 1 - Estrutura Hierárquica Básica

(ii) Definição de prioridades: Depois da estruturação da hierarquia, deve ser feita uma comparação, par a par, de cada elemento de um determinado nível da hierarquia, através de uma matriz de decisão quadrada. Essa matriz representa as preferências do decisor entre os elementos comparados sob o enfoque de um elemento proveniente de um nível imediatamente superior da hierarquia.

Este processo se estende até que todos os elementos de um mesmo nível forem comparados segundo os elementos de um nível acima.

De acordo com Gomes, Araya e Carignano (2004), o decisor quando comparar dois elementos deve buscar responder as seguintes perguntas: qual dos dois elementos contribui mais para ocorrer o outro? Quantas

vezes um elemento contribui mais que outro?
Desta forma, uma matriz será gerada

conforme figura abaixo:

Figura 2:Matriz AHP

$$A = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 1/a_{21} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 1/a_{n1} & 1/a_{n2} & \cdots & 1 \end{bmatrix}, \text{ onde:}$$

$a_{ij} > 0 \Rightarrow \text{positiva}$
 $a_{ij} = 1 \therefore a_{ji} = 1$
 $a_{ij} = 1/a_{ji} \Rightarrow \text{recíproca}$
 $a_{ik} = a_{ij} \cdot a_{jk} \Rightarrow \text{consistência}$

A quantidade de julgamentos necessários para a construção de uma matriz de julgamentos genérica A é $n(n-1)/2$, onde n é o número de elementos pertencentes a esta matriz.

Cada elemento a_{ij} da matriz representa a dominação do fator A_i sobre o fator A_j . A

diagonal principal da matriz é preenchida com o valor 1, que representa a não dominância de um fator a outro. A comparação das alternativas é feita, utilizando uma escala própria que varia de 1 a 9, sendo denominada escala fundamental de Saaty (Tabela 1):

Tabela 1 – Escala Fundamental de Saaty

Escala Numérica	Escala Verbal	Explicação
1	Ambos elementos são de igual importância.	Ambos elementos contribuem com a propriedade de igual forma.
3	Moderada importância de um elemento sobre o outro.	A experiência e a opinião favorecem um elemento sobre o outro.
5	Forte importância de um elemento sobre o outro.	Um elemento é fortemente favorecido.
7	Importância muito forte de um elemento sobre o outro.	Um elemento é muito fortemente favorecido sobre o outro.
9	Extrema importância de um elemento sobre o outro.	Um elemento é favorecido pelo menos com uma ordem de magnitude de diferença.
2, 4, 6, 8	Valores intermediários entre as opiniões adjacentes.	Usados como valores de consenso entre as opiniões.
Incremento 0.1	Valores intermediários na graduação mais fina de 0.1.	Usados para graduações mais finas das opiniões

Fonte: Roche (2004, p. 6)

Depois que a matriz de decisões estiver totalmente preenchida, é necessário fazer a normalização dos valores atribuídos aos critérios comparados. Para isso, devem-se somar os valores da entrada de cada coluna da matriz e dividir cada entrada da matriz pela soma dos valores de sua respectiva coluna. Com a matriz normalizada estima-se o peso (w_i) de cada critério, através da média aritmética de cada linha da matriz normalizada, formando assim o vetor peso. Estes pesos obtidos indicarão quais critérios são mais relevantes na comparação das

alternativas e auxiliarão no processo decisório. (SAATY, 1991, *in* BESTEIRO et al., 2009)

(iii) consistência lógica: o ser humano tem a habilidade de estabelecer relações entre objetos ou idéias de forma que elas sejam coerentes, tal que estas se relacionem bem entre si e suas relações apresentem consistência (SAATY, 2000). Assim o método A.H.P. se propõe a calcular a Razão de Consistência dos julgamentos, denotada por $RC = IC/IR$, onde IR é o Índice de Consistência Randômico obtido para uma

matriz recíproca de ordem n , com elementos não-negativos e gerada randomicamente.

O Índice de Consistência (IC) é dado por $IC = (\lambda_{\max} - n) / (n - 1)$. O fator $\lambda_{\max}$ é o autovalor da matriz de decisão. Segundo Gomes, Araya e Carignano (2004) $\lambda_{\max}$ pode ser obtido a partir da equação:

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^n v_i \frac{[Aw]_i}{w_i} \quad (1)$$

É desejável que o valor encontrado para o RC seja menor que 0,1 para que o grau de consistência seja satisfatório, caso contrário o método deve ser revisto. (SAATY, 1991, in BESTEIRO, et al. 2009).

No final, realiza-se uma síntese de tudo o que foi feito e elabora-se um ranking das alternativas avaliadas. Devem-se combinar as matrizes das alternativas com a matriz dos critérios. Cada matriz de critério será multiplicada pelo vetor de peso dos critérios conforme a fórmula abaixo:

$$\bar{v}_k(A_i) = \sum_{j=1}^n \bar{v}_j(A_j) / n \quad i=1, \dots, n \quad (2)$$

3. MÉTODO DE PESQUISA

A pesquisa foi realizada em uma Padaria e Confeitaria, localizada na cidade de Barra do Piraí – RJ, em maio de 2017. O método empregado foi o AHP, com a utilização de 4 (quatro) critérios, como auxílio na escolha do melhor entre 3(três) fornecedores de farinha de trigo.

Para aplicar o método, primeiramente é necessária a escolha de alguns critérios para que a estrutura hierárquica seja montada. Dessa forma, decidiu-se optar pelos seguintes critérios:

- - Entrega: é o ato de entregar. Nesse caso, foi analisado principalmente a rapidez na entrega. Mas também se ela é feita com segurança e se todos os produtos são entregues em boa condição e/ou se não há mercadoria faltando ou sobrando.
- - Prazo: é o tempo determinado para que alguma tarefa seja realizada. No caso deste trabalho, foi analisado o prazo de pagamento oferecido aos clientes. Se as condições de pagamento são boas ou não.
- - Qualidade: é o grau negativo ou positivo de excelência. É um conceito subjetivo, é a propriedade de qualificar os mais diversos serviços, objetos, indivíduos, entre outros. No caso dessa pesquisa, foi analisada a qualidade do produto em si, se ele está em conformidade com as exigências dos compradores, etc.
- - Preço: é o valor monetário que é atribuído a algo. Esse valor monetário é expresso em dinheiro e indica a quantia que o comprador ou o cliente deve disponibilizar para se apoderar de um produto ou para usufruir de um serviço. No caso da padaria, foi comparado o preço de venda da farinha de trigo dos três fornecedores.

Abaixo, segue um a tabela 2 com os critérios citados anteriormente, juntamente com a maior importância dada a cada um:

Tabela 2 –Importância dos critérios escolhidos

Critérios	Importância atribuída
Entrega	Rapidez
Prazo	Condições de pagamento
Qualidade	Condições do produto oferecido
Preço	Valor cobrado

Com os critérios definidos, foi elaborado um questionário (anexo 1) com o auxílio do *SuperDecisions*, que é um *software* de suporte à decisão que implementa o AHP e a

ANP. O questionário foi aplicado em uma amostra, por conveniência de 3 (três) funcionários da padaria estudada.

A razão desta amostra se deu, por lidarem diretamente com os fornecedores, ou então trabalharem diretamente com os produtos. Com isso, foram escolhidos duas auxiliares

administrativas e um padeiro. Na tabela abaixo, segue algumas informações sobre cada funcionário:

Tabela 3 – Informações sobre os funcionários entrevistados

Função	Formação	Tempo que trabalha na Padaria	Atividades realizadas relacionadas aos fornecedores
Auxiliar Administrativo	Bacharel em Matemática	5 anos	Recebimento de mercadorias, realização dos pedidos dos produtos, pagamento de contas.
Auxiliar Administrativo	Bacharel em Matemática	17 anos	Recebimento de mercadorias, realização dos pedidos dos produtos, pagamento de contas.
Padeiro	Curso de Padeiro	10 anos	Produção de pães e outros produtos.

4. ESTUDO DE CASO

Com já dito antes, a empresa do estudo de caso é uma padaria de médio porte, situada na cidade de Barra do Piraí, interior do estado do Rio de Janeiro. Nela, é produzido diversos tipos de pães e bolos todos os dias. Seus produtos têm um demanda considerável dentro da cidade. Para a produção dos produtos feitos pelo próprio estabelecimento, a matéria-prima é comprada de vários fornecedores.

Como a maior produção realizada dentro da padaria são os pães, a matéria-prima mais usada é a farinha de trigo, que é comprada de 3 fornecedores diferentes: Viviana, Bunge e Farisur. Abaixo, encontram-se algumas características de cada uma:

- Viviana: De acordo com seu site, a empresa tem sede no Rio Grande do Sul e surgiu em 1955. Conta com um parque industrial preparado, incluindo equipamentos modernos e profissionais capacitados.

- Bunge: De acordo com seu site, no Brasil, a empresa é uma das principais empresas de agronegócio e alimentos. Possui cerca de 17.000 colaboradores, é líder em origem de grãos e processamento de soja e trigo, na fabricação de produtos alimentícios e em serviços portuários. Desde 2006, atua também no segmento de açúcar e bioenergia. É uma das maiores exportadoras do país (a primeira em agronegócio) e contribui de maneira substancial para o saldo positivo da balança comercial e para as divisas para a economia nacional. Apresenta mais de 100 instalações no Brasil, entre fábricas, usinas, moinhos, portos, centros de

distribuição, silos e instalações portuárias. Foi fundada em 1818, com sede em White Plains, Nova York, EUA.

- Farisur: De acordo com seu site, a empresa está presente no Brasil desde 1996 e possui filiais no Rio de Janeiro, São Paulo e Espírito Santo. Ela importa, processa e comercializa farinhas de trigo e misturas para panificação produzidas na Argentina em diversas marcas.

A proprietária da padaria, conforme relato, optou por comprar de três fornecedores diferentes como medida de segurança, pois se um fornecedor atrasasse, já teria a farinha de outro e a padaria não ficaria com falta desse produto. Porém esses atrasos passaram a acontecer com mais frequência, juntamente com problemas na entrega, como falta de produtos. Dessa forma, viu-se a necessidade de uma reavaliação nessa forma de comprar e se é válido comprar de três fornecedores diferentes e mesmo assim passar por problemas, ou então se não é necessário.

5. DISCUSSÃO E RESULTADOS

O questionário (Anexo 1) foi aplicado a 3 (três) funcionários do estabelecimento de forma conjunta, onde as perguntas foram feitas para os 3 (três) ao mesmo tempo, e de forma consensual, foram escolhidas as respostas. Com isso, os resultados foram passados para o *software SuperDecisions*.

Primeiramente, foi elaborada a estrutura hierárquica, conforme Figura 3:

Figura 3: Estrutura Hierárquica Fornecedores

Nessa figura é possível visualizar melhor a relação entre o problema, os critérios e as alternativas. Dessa forma, para escolher o melhor fornecedor, serão avaliados os critérios entrega, prazo, qualidade e preço em relação aos fornecedores Viviana, Bunge e Farisur.

Com isso, foi possível colocar as respostas do questionário realizado com os funcionários.

Na figura 4 estão presentes as avaliações a respeito dos critérios. Pode-se perceber que essa parte teve um índice de inconsistência de 0.08859 e o critério melhor avaliado foi a qualidade, com um peso de 0.63204, seguida pelo preço, entrega e prazo, com os índices de 0.21384, 0.10246 e 0.05166, respectivamente.

Figura 4: Avaliação dos Critérios

1. Choose

Node Cluster

Choose Node

Fornecedor

Cluster: Escolha Fornecedor

Choose Cluster

Critéria

Restore

2. Node comparisons with respect to Fornecedor

Graphical Verbal Matrix Questionnaire Direct

Comparisons wrt "Fornecedor" node in "Criteria" cluster

Entrega is moderately more important than Prazo

1. Entrega	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Prazo
2. Entrega	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Preço
3. Entrega	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Qualidade
4. Prazo	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Preço
5. Prazo	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Qualidade
6. Preço	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Qualidade

3. Results

Normal Hybrid

Inconsistency: 0.08859

Entrega	0.10246
Prazo	0.05166
Preço	0.21384
Qualidade	0.63204

Completed Comparison

Copy to clipboard

Após isso, foram repassados os resultados dos questionários a respeito da comparação entre os fornecedores, de acordo com cada critério. Dessa forma, a figura 5 apresenta as avaliações em relação a entrega realizada.

Nesta parte, o índice de inconsistência foi de 0.03112 e a empresa Bunge teve o melhor resultado com 0.75825, seguida, respectivamente, pela Viviana com 0.15125 e Farisur com 0.09051.

Figura 5: Avaliação da Entrega

Comparisons for Super Decisions Main Window: Fornecedores de Farinha de Trigo.sdmod

1. Choose

Node: Cluster

Choose Node: Entrega

Cluster: Criteria

Choose Cluster: Alternatives

Restore

2. Node comparisons with respect to Entrega

Graphical Verbal Matrix Questionnaire Direct

Comparisons wrt "Entrega" node in "Alternatives" cluster

Bunge is very strongly more important than Farisur

1. Bunge >=9.5 9 8 7 6 5 4 3 2 2 3 4 5 6 7 8 9 >=9.5 No comp. Farisur

2. Bunge >=9.5 9 8 7 6 5 4 3 2 2 3 4 5 6 7 8 9 >=9.5 No comp. Viviana

3. Farisur >=9.5 9 8 7 6 5 4 3 2 2 3 4 5 6 7 8 9 >=9.5 No comp. Viviana

3. Results

Normal Hybrid

Inconsistency: 0.03112

Bunge	0.75825
Farisur	0.09051
Viviana	0.15125

Completed Comparison

Copy to clipboard

A figura 6 mostra o resultado a respeito do prazo. Essa parte teve um índice de inconsistência de 0.00885. A Bunge,

novamente, teve a melhor avaliação, 0.53961, seguida por Viviana (0.29696) e Farisur (0.16342).

Figura 6: Avaliação do Prazo

Comparisons for Super Decisions Main Window: Fornecedores de Farinha de Trigo.sdmod

1. Choose

Node: Cluster

Choose Node: Prazo

Cluster: Criteria

Choose Cluster: Alternatives

Restore

2. Node comparisons with respect to Prazo

Graphical Verbal Matrix Questionnaire Direct

Comparisons wrt "Prazo" node in "Alternatives" cluster

Bunge is moderately more important than Farisur

1. Bunge >=9.5 9 8 7 6 5 4 3 2 2 3 4 5 6 7 8 9 >=9.5 No comp. Farisur

2. Bunge >=9.5 9 8 7 6 5 4 3 2 2 3 4 5 6 7 8 9 >=9.5 No comp. Viviana

3. Farisur >=9.5 9 8 7 6 5 4 3 2 2 3 4 5 6 7 8 9 >=9.5 No comp. Viviana

3. Results

Normal Hybrid

Inconsistency: 0.00885

Bunge	0.53961
Farisur	0.16342
Viviana	0.29696

Completed Comparison

Copy to clipboard

Já na figura 7, encontra-se as avaliações em relação a qualidade, critério considerado o mais importante anteriormente. Aqui, o índice de inconsistência foi de 0.09609 e a melhor

colocada foi a Bunge com um peso de 0.75041, bem a frente da Farisur (0.17134) e Viviana (0.07825).

Figura 7: Avaliação da Qualidade

Comparisons for Super Decisions Main Window: Fornecedores de Farinha de Trigo.sdmod

1. Choose

Node Cluster

Choose Node

Qualidade

Cluster: Criteria

Choose Cluster

Alternatives

Restore

2. Node comparisons with respect to Qualidade

Graphical Verbal Matrix Questionnaire Direct

Comparisons wrt "Qualidade" node in "Alternatives" cluster

Bunge is strongly to very strongly more important than Farisur

1. Bunge	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Farisur
2. Bunge	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Viviana
3. Farisur	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Viviana

3. Results

Normal Hybrid

Inconsistency: 0.09609

Bunge	0.75041
Farisur	0.17134
Viviana	0.07825

Completed Comparison

Copy to clipboard

O último critério avaliado foi o preço. Na figura 8 está presente os resultados. Nessa parte, o índice de inconsistência foi de 0.00000. A

melhor colocada, como em todos outros, foi a Bunge com 0.50000, seguida com um empate entre Farisur e Viviana, com 0.25000 cada.

Figura 8: Avaliação do Preço

Comparisons for Super Decisions Main Window: Fornecedores de Farinha de Trigo.sdmod

1. Choose

Node Cluster

Choose Node

Preço

Cluster: Criteria

Choose Cluster

Alternatives

Restore

2. Node comparisons with respect to Preço

Graphical Verbal Matrix Questionnaire Direct

Comparisons wrt "Preço" node in "Alternatives" cluster

Bunge is equally to moderately more important than Farisur

1. Bunge	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Farisur
2. Bunge	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Viviana
3. Farisur	>=9.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	<=9.5	No comp.	Viviana

3. Results

Normal Hybrid

Inconsistency: 0.00000

Bunge	0.50000
Farisur	0.25000
Viviana	0.25000

Completed Comparison

Copy to clipboard

Com todas as avaliações feitas, o *SuperDecisions* fez uma junção para se chegar a um resultado final. Com isso, na figura 9, percebe-se os pesos atribuídos a cada critério e os valores dados aos fornecedores. Em relação aos critérios, a qualidade teve a maior prioridade, com um

peso de 0.63240. Depois dela, vieram o preço com 0.21384 e a entrega e prazo, com 0.10246 e 0.05166, respectivamente. Em relação aos fornecedores, a empresa Bunge ficou em primeiro lugar, com 0.68677, seguida pela Farisur (0.17947) e Viviana (0.13375).

Figura 9: Resultado Final

Icon	Name	Normalized by Cluster	Limiting
No Icon	Bunge	0.68677	0.343386
No Icon	Farisur	0.17947	0.089736
No Icon	Viviana	0.13375	0.066877
No Icon	Entrega	0.10246	0.051229
No Icon	Prazo	0.05166	0.025832
No Icon	Preço	0.21384	0.106921
No Icon	Qualidade	0.63204	0.316018
No Icon	Fornecedor	0.00000	0.000000

Com isso, percebe-se que o critério considerado o mais importante na escolha do fornecedor foi a qualidade. E o melhor fornecedor para a padaria estudada foi a empresa Bunge, que ficou bem a frente da Viviana e Farisur em todos os critérios, principalmente entrega e qualidade.

6. CONCLUSÃO

Após a realização dessa pesquisa pode-se concluir o quão é interessante e eficiente o método AHP. Em conjunto com o *software*

SuperDecision, foi possível realizar uma análise prática e rápida do questionário aplicado aos funcionários da padaria. Sua aplicação promoveu uma interação entre várias pessoas de áreas diferentes, o que torna o modelo mais justo e completo.

Desta forma, a principal contribuição do trabalho consiste nos resultados encontrados, que identificou qual é o melhor fornecedor de farinha de trigo para a padaria estudada, contribuindo, desse modo, para o seu melhor desempenho.

REFERÊNCIAS

- [1] Azeredo, Jeanderson da Silva; Junior, Geraldo Galdino de Paula; Gonçalves, Tiago José Menezes. A Escolha de um Sistema Integrado de Gestão Empresarial (ERP) através do Método de Análise Hierárquica (AHP). Universidade Federal de Santa Maria - Centro de Tecnologia, 2010. Disponível em: <<http://ojs.ingepro.com.br/index.php/ingepro/article/view/180/157>>. Acesso em: 10 abr. 2017.
- [2] Bertaglia, P. R. Logística e gerenciamento da cadeia de abastecimento. São Paulo, Saraiva, 2006.
- [3] Besteiro, A.M. et al. A Utilização do método AHP para traçar, como ferramenta para o auxílio a decisão de um candidato, a escolha de um curso de engenharia. UNESP - São Paulo, 2009.

- [4] Bornia, Antonio Cezar; WERNKE, Rodney. A contabilidade gerencial e os métodos multicriteriais. Revista Contabilidade & Finanças. FIPECAP - FEA - USP. v.14, n. 25, p. 60- 71, jan./abr. 2001.

- [5] Brasil, Lei Nº 8.078, de 11 de Setembro DE 1990. Disponível em: <http://www.planalto.gov.br/ccivil_03/LEIS/L8078.htm> Acesso em: 16 mai. 2017

- [6] Costa, Helder Gomes. Introdução ao método de análise hierárquica: Análise multicritério no auxílio à decisão. Niterói: H.G.C., 2002.

- [7] Firjan. O setor de panificação e o Sistema Firjan. Disponível em: <<http://www.firjan.com.br/o-sistema-firjan/setores-de-atuacao/panificacao.htm>> Acesso em: 16 mai. 2017

- [8] Food Ingredients Brasil. Panificação: Os ingredientes enriquecedores. Rev. n. 9. 2009. Disponível em: <<http://www.revista-fi.com/materias/114.pdf>> Acesso em: 16 mai. 2017
- [9] Gomes, L. F. A. M.; Araya, M. C. G.; Carignano, C. Tomada de Decisões em Cenários Complexos. São Paulo. Pioneira Thomson Learning, 2004. 168p.
- [10] Grandzol, J. R. Improving the faculty election process in higher education: a case for the analytic hierarchy process. IR Application, S.I., v.6, p.1-13, 2005.
- [11] Marins, C. S., Souza, Daniela De Oliveira, Barros, Magno da Silva, O uso do método de análise hierárquica (AHP) na tomada de decisões gerenciais – Um estudo de caso. Disponível em: <<http://www2.ic.uff.br/~emitacc/AMD/Artigo%204.pdf>> Acesso em: 16 mai. 2017
- [12] Marins, C. S.; Souza, Daniela De Oliveira, Freitas; André Luis Policani. A metodologia de multicritério como ferramenta para a tomada de decisões gerenciais: um estudo de caso. GEPROS, 2006.
- [13] Power, D. J. Decision Support Systems Glossary, DSSResources.COM, World Wide Web, Disponível em: <<http://DSSResources.COM/glossary/>>, 2014. Acesso em: 16 mai. 2017.
- [14] Saaty, T. L. Some Mathematical Concepts of the Analytic Hierarchy Process Behaviormetrika, v. 29, p. 1-9, 1991.
- [15] Saaty, T. L., The Analytic Hierarchy Process: Planning, Priority, Setting and Resource Allocation, McGraw-Hill, Inc., 1980.
- [16] Schmidt, A.M.A. Processo de apoio à tomada de decisão, abordagens: AHP e MACBETH. Dissertação (Mestrado), UFSC-Eng. Produção, 1995.
- [17] Sinuany-Stern, Z.; Mehrez, A.; Hadad, Y. An Ahp/Dea methodology for ranking decision making units. In: International Transactions in Operational Research, 2000.
- [18] Tomada de Decisão em Projetos – Método AHP. Disponível em: <<http://blog.mundopm.com.br/2012/05/02/tomada-de-decisao-em-projetos-%E2%80%93-metodo-ahp>>. Acesso em: 2 mai. 2017

ANEXO 1 - QUESTIONÁRIO SOBRE OS CRITÉRIOS E OS FORNECEDORES DE FARINHA DE TRIGO

De acordo com a escala abaixo, responda as questões:

Valor	Ordem de Importância
1	Mesma importância
3	Pouco mais importante
5	Muito mais importante
7	Bastante mais importante
9	Extremamente mais importante
2, 4, 6, 8	Valores intermediários

1) Na sua opinião, em relação aos critérios abaixo, quanto um é mais importante em relação ao outro?

Entrega	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Prazo
Entrega	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Preço
Entrega	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Qualidade
Prazo	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Preço
Prazo	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Qualidade
Preço	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Qualidade

2) Em relação aos critérios abaixo, quanto um fornecedor é melhor em relação ao outro?

a) Entrega

Bunge	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Farisur
Bunge	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Viviana
Farisur	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Viviana

b) Prazo

Bunge	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Farisur
Bunge	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Viviana
Farisur	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Viviana

c) Preço

Bunge	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Farisur
Bunge	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Viviana
Farisur	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Viviana

d) Qualidade

Bunge	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Farisur
Bunge	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Viviana
Farisur	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Viviana

Capítulo 7

OTIMIZAÇÃO VIA SIMULAÇÃO MONTE CARLO NA MELHORIA DE PROCESSOS PRODUTIVOS AERONÁUTICOS

Igor Makiyama de Souza Barbosa

Murilo Vieira Strongoli

Aneirson Francisco da Silva

Fernando Augusto Silva Marins

Erica Ximenes Dias

Resumo: A indústria aeronáutica tem características que a distinguem de outros setores industriais, atribuindo-lhe desafios dinâmicos e peculiares. Devido à grande quantidade de variáveis concernentes aos processos industriais, a tomada de decisão relacionada a melhoria de tais processos se mostra uma atividade complexa (RODRIGUES, 2016). Neste trabalho, buscou-se utilizar a Otimização via Simulação Monte Carlo (OvSMC) na previsão de parâmetros de usinagem de peças do setor aeronáutico, os quais deveriam ser utilizados para atender a um período de tempo total determinado. Para tanto, foram utilizados dados de um processo de usinagem de uma indústria do setor aeronáutico brasileiro. Com base nos resultados obtidos com o trabalho, concluiu-se que a OvSMC é uma ferramenta que contribui na melhora da determinação de parâmetros envolvidos em processos de manufatura de peças do setor aeronáutico.

Palavras Chave: Otimização - Simulação - Monte Carlo - Setor Aeronáutico –

1. INTRODUÇÃO

A indústria aeronáutica tem características que a distinguem de outros setores industriais, atribuindo-lhe desafios dinâmicos e peculiares. Por conta de sua complexidade, são necessários elevados níveis de investimento em pesquisa, conhecimento e recursos humanos. Além disso, por conta dos longos processos de desenvolvimento de produto, o setor gera retornos crescentes e economia de escala (MORAES et. al., 2016).

Em função desse contexto, a indústria aeroespacial brasileira se destaca no aspecto econômico por apresentar alto potencial de crescimento, conteúdo tecnológico relevante, alto valor agregado do produto e expressiva contribuição para a balança comercial (MONTORO & MIGON, 2009).

Nas Figuras 1 e 2 é possível visualizar a importância do setor aeronáutico nacional, por meio da contribuição socioeconômica do mesmo, como apresentado anteriormente.

Figura 1: Receita do setor aeronáutico em bilhões de dólares, de 2010 a 2015.

Fonte: Associação das Indústrias Aeroespaciais do Brasil, AIAB (2015).

Figura 2: Empregos gerados pelo setor aeronáutico em bilhões de dólares, de 2010 a 2015.

Fonte: Associação das Indústrias Aeroespaciais do Brasil, AIAB (2015).

Segundo Gomes (2015), o quadro de tendências que se delineia para os próximos anos aponta para o gradual aumento da

concorrência a partir de 2015 devido ao incentivo estatal de países como China,

Rússia e Japão ao desenvolvimento da indústria aeronáutica nacional.

Baseado nessas tendências, é possível inferir que nos próximos anos haverá um esforço maior relacionado à melhoria dos processos industriais, considerando que se deve manter o controle sobre os parâmetros de qualidade, visto que os recursos estão cada vez mais escassos. A determinação de arranjos ideais para atingir metas estipuladas representa uma grande parcela de tal esforço (BOYLAN et al., 2013).

Devido à grande quantidade de variáveis concernentes aos processos industriais, a tomada de decisão relacionada a melhoria de tais processos se mostra uma atividade complexa. Além do elevado número de variáveis, ainda é necessário entender o grau de influência entre as variáveis independentes (influenciadoras) e dependentes (influenciadas) (RODRIGUES, 2016). No auxílio a essa difícil atividade, é possível empregar métodos de otimização da Pesquisa Operacional (PO) que resultam em uma maior eficiência destes processos (PAIVA, FERREIRA & PAIVA, 2007).

Entre as diversas técnicas da PO existentes, Yang (2010) comenta que a Simulação Monte Carlo (SMC) tem sido aplicada a todas as áreas do conhecimento. Neste trabalho, a SMC foi combinada com um conjunto de técnicas de otimização gerando um procedimento que pode ser encontrado na literatura como Otimização via Simulação Monte Carlo (OvSMC). Tal procedimento é indicado para solucionar problemas complexos que apresentem vários ótimos locais, envolvendo um *mix* de variáveis contínuas e discretas (KROESE, TAIMRE & BOTEV, 2011).

Uma vez compreendido o contexto no qual o trabalho está inserido, é possível levantar o seguinte questionamento: há vantagens em

se utilizar a Otimização via Simulação Monte Carlo (OvSMC) na determinação de parâmetros, relacionados aos processos de manufatura de peças aeronáuticas, de forma a atender metas de tempo estipuladas pelos especialistas do setor?

Como proposta de resposta para a questão levantada, o objetivo geral deste trabalho foi utilizar a OvSMC na predição dos tempos de usinagem e ajustagem do processo de manufatura de peças aeronáuticas. Os objetivos específicos envolveram dois aspectos:

- Desenvolver funções que representem adequadamente o processo de manufatura de peças aeronáuticas;
- Propor uma ajustagem dos parâmetros vinculados ao processo de manufatura, a partir de um tempo de processo determinado.

Ehrgott et al. (2014) comentam que no processo de otimização ainda há lacunas entre teoria e prática, pois não é simples para o gestor aplicar os conceitos, modelos e técnicas desenvolvidas pelos pesquisadores diretamente nos seus problemas do dia a dia que possuem características especiais. Os dados observados na Figura 3 justificam a aplicação da OvSMC na indústria aeronáutica, sendo possível observar um crescente número de publicações e citações relacionadas com OvSMC, utilizando a palavra-chave "Optimization via Monte Carlo Simulation" na base de dados Web of Science, enquanto há somente uma publicação quando associamos as palavras-chave "Optimization via Monte Carlo Simulation" e "Aeronautical Industry" na mesma base de dados. Observa-se que, embora a área de estudo esteja em alta, a aplicação prática no ramo aeronáutico ainda pode ser mais explorada.

Figura 3: Progressão anual de publicações e citações sobre OvSMC e Indústria Aeronáutica.

Fonte: Web of Knowledge.

Este trabalho recebe, segundo Bertrand e Fransoo (2002), a classificação de pesquisa aplicada, sendo seus objetivos empíricos descritivos, uma vez que as relações causais existentes na realidade são descritas adequadamente pelo modelo desenvolvido. O método de pesquisa utilizado foi a modelagem e simulação, seguindo uma abordagem quantitativa.

Para a execução do trabalho, foi utilizado o *software Crystal Ball®* em conjunto com o Excel® para modelagem e simulação e posterior otimização do problema. O *software MiniTab17®* foi utilizado para testar a normalidade dos resíduos do modelo. A FEG – UNESP possui computadores adequados para a execução deste estudo. Dessa forma, não houve limitações físicas ou de *software* para que a pesquisa fosse desenvolvida.

Este trabalho está estruturado em mais três seções além desta seção introdutória. A Seção 2 contemplou a fundamentação teórica sobre os temas abordados na pesquisa. A Seção 3 abordou a solução do problema e análise dos resultados, e por fim, a Seção 4 contemplou as conclusões e recomendações para futuras pesquisas.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 SIMULAÇÃO E FUNÇÕES EMPÍRICAS

De maneira geral, a simulação é uma ferramenta usada tanto na concepção de sistemas, como na adequação de sistemas já

existentes, incluindo reconfigurações físicas, no controle e/ou nas regras que regem a operação de tais sistemas. Segundo Sakurada e Miyake (2009), a aplicabilidade da simulação como ferramenta de auxílio à tomada de decisão vem se expandindo em todas as áreas, fato que permite um conhecimento mais consistente a respeito das organizações e seus processos.

Realizar uma simulação significa realizar uma série de etapas, que podem ser resumidas, segundo Hillier e Lieberman (2010), da seguinte forma:

- Definir o estado atual do sistema.
- Identificar os possíveis estados do sistema.
- Identificar os eventos que podem alterar os estados do sistema.
- Estabelecer uma forma de controlar o tempo de simulação.
- Estabelecer um método para gerar eventos aleatórios.
- Formular um modelo que identifique os eventos e altere o estado do sistema.

Segundo Helene (2013), a análise de regressão pode ser usada para estimar uma função empírica, sendo aplicada com essa finalidade em muitas áreas do conhecimento. Por meio da construção de uma função empírica, é possível executar a fase de formulação do modelo, proposta por Hillier e Lieberman (2010).

A análise de regressão é uma ferramenta utilizada para avaliar a relação entre variáveis dependentes e independentes, por meio de uma equação linear, tal como expresso em (1):

sendo y_i variável dependente, $x_1, x_2, \dots, x_k$ as variáveis independentes, ε o fator de erro, e n o número de amostras observadas.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i \quad (1)$$

Existem diversas técnicas de análise de regressão, neste trabalho optou-se pelo algoritmo OLS - Ordinary Least Squares (HELENE, 2013, FERRARO & GIORDANI, 2012), que consiste em se desenvolver um modelo por meio de regressão linear ou não linear e refinar os parâmetros com base em sucessivas iterações (LIU *et al.*, 2016).

Estudando uma forma de solucionar problemas de blindagem em reatores nucleares, os pesquisadores Von Neumann e Ulam desenvolveram, em meados de 1940, a ferramenta conhecida como Simulação Monte Carlo (SMC) (CORRAR, 1993).

A função da SMC é satisfazer a 5ª etapa descrita por Hillier e Lieberman (2010), gerando números (eventos) aleatórios para alimentar o modelo matemático. Silva *et al.* (2014) comentam sobre a natureza estocástica de grande parte das variáveis em sistemas reais, entre as quais é possível citar o tempo das atividades de um projeto, objeto de estudo deste trabalho.

As amostras de números aleatórios são, em geral, obtidas computacionalmente (OLIVEIRA, BARROS e DOS REIS, 2007). Devem possuir parâmetros como média, desvio padrão e/ou seus limites máximo e mínimo, sendo modeladas seguindo distribuições de probabilidade como, por exemplo, Normal, Uniforme e Triangular, dentre outras. A função de densidade de

distribuição de probabilidade triangular foi utilizada na realização do trabalho, como será apresentado adiante.

Assis *et al.* (2013) ressaltam que a SMC se destaca por, entre outras características, tratar-se de um método quantitativo, conforme ocorre a estabilização da curva de distribuição de probabilidade das variáveis dependentes do modelo matemático. Contudo, algumas desvantagens merecem destaque:

1. Demanda por grande capacidade computacional para realizar o processamento de dados.

2. Necessidade de conhecer a fundo o sistema com que se está trabalhando, a ponto de ser capaz de determinar as distribuições de probabilidade das variáveis de entrada.

Segundo Shahinidis (2004), a otimização por meio da simulação estocástica é aplicável a diversas frentes do conhecimento. Sua aplicabilidade se justifica por possibilitar a análise de diferentes cenários, pois gera informações relevantes acerca do sistema estudado (ABDELAZIZ, AOUNI & EL FAYEDH, 2007; AOUNI, ABDELAZIZ & MARTEL, 2005; DEB, 2001).

Abdelaziz, Aouni e El Fayedh (2007) propõem uma formulação genérica para a otimização estocástica:

$$\text{Max} \quad \sum_{j=1}^n \tilde{C}_j x_j \quad (2)$$

S.A:

$$\sum_{j=1}^n \tilde{A}_{ij} x_j \leq \tilde{b}_i \quad i = 1, 2, \dots, m \quad (3)$$

$$x \in X, X \text{ é o conjunto de soluções viáveis} \quad (4)$$

Sendo $\tilde{C}_j, \tilde{A}_{ij}$ matrizes aleatórias e $\tilde{b}_i$ um recurso aleatório estocástico.

O software *Crystal Ball®*, que será utilizado na realização deste trabalho, foi criado com base na SMC, e uma de suas ferramentas para solução é o *Optquest* (ORACLE, 2015). O *Optquest* é um otimizador que trata o modelo de simulação como uma “caixa-preta”, ou seja, o usuário só toma conhecimento das entradas e saídas do modelo de simulação. Ainda, esta ferramenta utiliza de maneira combinada as metaheurísticas de Busca Tabu, Redes Neurais e Busca Dispersa em uma única heurística de busca (KLEIJNEN & WAN, 2007).

O uso desta combinação de metaheurísticas significa que o otimizador *Optquest* do software *Crystal Ball®* busca as melhores soluções candidatas, entre essas soluções, aquelas que não atenderem às restrições são descartadas. Tais restrições podem ser definidas pelo usuário, sendo possível também especificar os limites para as saídas da simulação aleatória. Ele requer a definição de valores mínimos, sugeridos e máximos para as variáveis a serem otimizadas. O ponto de partida da busca são os valores sugeridos, e assim sendo, esta escolha afeta a eficiência e eficácia da busca (RODRIGUES, 2016).

3. SOLUÇÃO DO PROBLEMA E ANÁLISE DOS RESULTADOS

Este trabalho foi desenvolvido numa indústria do setor aeronáutico. O foco do estudo

envolveu um processo de usinagem de componentes aeronáuticos, buscando otimizar os tempos de usinagem e ajustagem da máquina-ferramenta utilizada, dadas as dimensões da peça a ser usinada referentes a bitola, ao comprimento e a largura.

A modelagem foi realizada com base nos dados históricos do processo fornecidos pela empresa. Havia 541 linhas de dados, com as dimensões da peça a ser usinada: bitola (em polegadas), comprimento (em milímetros), largura (em milímetros) e, também, os tempos de usinagem e ajustagem referentes a cada peça (ambos em horas). A título de exemplo, os dados referentes a algumas linhas estão apresentados na Figura 4.

Com base no que foi apresentado pela empresa, definiu-se como variáveis de decisão:

- $X1$ – Bitola [pol]
- $X2$ – Largura [mm]
- $X3$ – Comprimento [mm]

As variáveis dependentes foram escolhidas, após consulta aos gestores da empresa, como sendo:

- $Y1$ – Tempo de usinagem [h]
- $Y2$ – Tempo de Ajustagem [h]
- $Y3 = Y1 + Y2$ - Tempo total do processo [h]

Figura 4: Histórico de dados.

	A	B	C	D	E
1	Bitola (pol)	Largura (mm)	Comprimento (mm)	Usinagem (Y1)	Ajustagem (Y2)
2	0,75	275	225	1,35	0,5
3	1	275	225	1,5	0,55
4	1,25	185	890	2,6	0,9
5	0,75	165	120	0,65	0,3
6	1,25	130	160	0,85	0,4
7	1,25	370	170	1,45	0,6
8	1,25	170	380	1,45	0,6
9	2,5	90	120	1,05	0,4
10	2,5	90	120	1,05	0,4
11	2,25	100	110	1,05	0,4
12	2,25	115	75	0,95	0,35
13	2,5	90	120	0,85	0,4
14	2,5	115	95	1,05	0,4
15	2,5	90	120	0,85	0,4
16	2,5	125	95	0,9	0,4
17	1,75	195	100	0,95	0,4
18	2,5	115	100	0,9	0,45
19	2,25	120	130	1,15	0,45
20	2,25	100	120	1,05	0,4
21	1,5	230	330	1,7	0,7
22	0,75	155	125	0,65	0,3
23	2,5	205	140	1,35	0,6
24	2,75	160	190	1,65	0,6
25	1,75	80	380	1,4	0,5
26	1,75	70	180	1,1	0,35

Fonte: Empresa estudada

A modelagem do problema foi realizada utilizando-se ferramentas de regressão linear

disponíveis no Excel® por meio do algoritmo dos mínimos quadrados ordinários (OLS). Os

resultados da regressão estão apresentados na Figura 5 (tempo de usinagem) e na Figura

6 (tempo de ajustagem).

Figura 5: Resultados da Regressão do tempo de usinagem.

1	RESUMO DOS RESULTADOS								
2									
3	<i>Estatística de regressão</i>								
4	R múltiplo	0,936704524							
5	R-Quadrado	0,877415364							
6	R-quadrado ajustado	0,876730534	>70%						
7	Erro padrão	0,158164601							
8	Observações	541							
9									
10	ANOVA								
11		<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>			
12	Regressão	3	96,15282774	32,05094258	1281,215623	3,0502E-244			
13	Resíduo	537	13,43361403	0,025016041					
14	Total	540	109,5864418						
15									
16		<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>% superior</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
17	Interseção	-0,286678377	0,026835829	-10,68267274	2,76304E-24	-0,339394449	-0,23396	-0,339394449	-0,233962305
18	Variável X 1	0,293673616	0,01040537	28,22327345	3,81E-108	0,273233395	0,314114	0,273233395	0,314113837
19	Variável X 2	0,002721168	7,83747E-05	34,71997506	2,297E-139	0,002567209	0,002875	0,002567209	0,002875126
20	Variável X 3	0,002746155	6,80774E-05	40,3387366	1,1706E-164	0,002612425	0,00288	0,002612425	0,002879886
21									

Fonte: Excel®

Figura 6: Resultados da Regressão do tempo de ajustagem.

1	RESUMO DOS RESULTADOS								
2									
3	<i>Estatística de regressão</i>								
4	R múltiplo	0,956382486							
5	R-Quadrado	0,91466746							
6	R-quadrado ajustado	0,914190741							
7	Erro padrão	0,048895988							
8	Observações	541							
9									
10	ANOVA								
11		<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>			
12	Regressão	3	13,76161151	4,587203837	1918,675742	1,7887E-286			
13	Resíduo	537	1,283869081	0,002390818					
14	Total	540	15,04548059						
15									
16		<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
17	Interseção	-0,090832329	0,008296195	-10,94867334	2,55182E-25	-0,107129303	-0,074535354	-0,107129303	-0,074535354
18	Variável X 1	0,114306693	0,003216781	35,5344972	3,8922E-143	0,107987676	0,120625709	0,107987676	0,120625709
19	Variável X 2	0,001037978	2,42292E-05	42,83986868	2,2885E-175	0,000990382	0,001085573	0,000990382	0,001085573
20	Variável X 3	0,001020693	2,10459E-05	48,49849354	2,3142E-198	0,00097935	0,001062035	0,00097935	0,001062035
21									

Fonte: Excel®

Como visto nas Figuras 5 e 6, os valores de R^2 -ajustados para as variáveis Y_1 e Y_2 foram maiores que 70% (88% e 91%, respectivamente), apresentaram p -value menor que 5% e resíduos normais (como se destacará adiante). Dessa forma, conclui-se que não há tendenciosidade nas funções empíricas (HELENE, 2013). Portanto o nível descritivo dos coeficientes encontrados é satisfatório, sendo possível construir modelos para as variáveis dependentes, como descrito na Equação 5 e na Equação 6:

$$\hat{Y}_1 = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + \tilde{\beta}_2 x_2 + \tilde{\beta}_3 x_3 \quad (5)$$

$$\hat{Y}_2 = \tilde{\beta}'_0 + \tilde{\beta}'_1 x_1 + \tilde{\beta}'_2 x_2 + \tilde{\beta}'_3 x_3 \quad (6)$$

Para verificar se os valores estimados pelo modelo matemático representavam a realidade, foi feita uma análise dos resíduos utilizando o *software MiniTab17®*. Como verificado nas Figuras 7 e 8, os valores dos resíduos seguiram uma distribuição normal (os valores de p -value foram maiores que 5%, aceitando-se a hipótese nula H_0), o que mostra que não há tendenciosidade no modelo, validando o mesmo.

Figura 7: Análise de resíduos do modelo Y1.

Fonte: MiniTab17®

Figura 8: Análise de resíduos do modelo Y2.

Fonte: MiniTab17®

Com o modelo validado, partiu-se para a otimização do mesmo por meio do *software Crystal Ball®* em conjunto com o otimizador *OptQuest*. A primeira etapa consistiu em informar as variáveis de decisão como discretas (ou seja, as variáveis assumiram valores limitados pelo banco de dados da empresa, evitando que houvesse uma otimização de valores impossíveis de serem obtidos).

A próxima etapa consistiu em atribuir valores aleatórios para cada um dos coeficientes (β e β') das expressões (5) e (6), para tal, foram atribuídos valores $x \sim T [LI, VP, LS]$, com distribuição de probabilidade triangular (KOKONENDJI; KIESSÉ, ZOCCHI, 2007). Os valores dos parâmetros estão dispostos na Figura 8.

Figura 8: Parâmetros da distribuição de probabilidade triangular.

Y1	Coeficientes	95% inferiores (LB)	95% superiores (UB)
Interseção	-0,28667838	-0,339394449	-0,233962305
Variável X 1	0,293673616	0,273233395	0,314113837
Variável X 2	0,002721168	0,002567209	0,002875126
Variável X 3	0,002746155	0,002612425	0,002879886
Y2	Coeficientes	95% inferiores (LB)	95% superiores (UB)
Interseção	-0,09083233	-0,107129303	-0,074535354
Variável X 1	0,114306693	0,107987676	0,120625709
Variável X 2	0,001037978	0,000990382	0,001085573
Variável X 3	0,001020693	0,00097935	0,001062035

Fonte: Excel®

A probabilidade triangular deve atender à Equação (7):

$$f(x) = \begin{cases} \frac{2(x-LI)}{(LS-LI)(VP-LI)} & LI \leq x < VP \\ \frac{2}{(LS-LI)} & x = VP \\ \frac{2(LS-x)}{(LS-LI)(LS-VP)} & VP < x \leq LS \\ 0 & \text{para qualquer outro valor} \end{cases} \quad (7)$$

Sendo x uma variável aleatória e $f(x)$ a função que descreve sua densidade de probabilidade triangular, LI o limite inferior, VP o valor mais provável e LS o limite superior.

Como exemplo, na Figura 9, é possível ver a distribuição triangular para o coeficiente $\tilde{\beta}_0$ da variável dependente Y1.

Figura 9: Distribuição triangular do coeficiente $\tilde{\beta}_0$.

Fonte: Crystal Ball®

A OvSMC foi utilizada para determinar os valores das variáveis de entrada (dimensões da bitola, comprimento e largura) que atendem a uma certa especificação de tempo

(valores-alvo). Para tal, foi utilizada a Equação (8) sujeita às restrições expressas nas Equações (9) e (10).

$$\hat{Y}_3 = (\hat{Y}_1 + \hat{Y}_2) \in Y \text{ (conjunto dos valores – alvo)} \quad (8)$$

Sujeito a:

$$\hat{Y}_1, \hat{Y}_2 \geq 0 \text{ (Tempo maior que zero)} \quad (9)$$

$$x_i \in X \text{ (Conjunto discreto de soluções viáveis)}$$

Foi determinado, em conjunto com os gestores da empresa, que seria pertinente definir como meta para Y_3 o valor de 2 horas. Os resultados da OvSMC são apresentados nas Figuras 10 e 11.

Figura 10: Resultado para a variável resposta “Tempo de Usinagem”

Fonte: Crystal Ball®

Figura 11: Resultado para a variável resposta “Tempo de Ajustagem”

Fonte: Crystal Ball®

Observa-se nas Figuras 10 e 11 os valores dos limites inferior e superior das variáveis dependentes (ver Tabela 1) para um intervalo de confiança de 95%. Os ajustes propostos

para as variáveis independentes (fatores ou variáveis x 's), que atendem à meta estipulada de 2 horas, estão especificados na Tabela 2.

Tabela 1: Intervalos $[L/, LS]$ obtidos na OvSMC para variáveis de resposta.

Valor final das funções objetivo	$L/$	LS
$\hat{Y}_1$	1.37	1.5
$\hat{Y}_2$	0.55	0.58

Tabela 2: Ajuste proposto para as variáveis independentes

Variável de Entrada	Valor
x_1 [pol]	1,5
x_2 [mm]	245
x_3 [mm]	225

A fim de tornar os resultados do modelo mais próximos da realidade, realizou-se ainda um teste de correlação entre as variáveis independentes por meio do *software MiniTab17®*. Para tal foi utilizado o Teste de

Correlação não paramétrico de Spearman, pois as variáveis em questão não seguem distribuição normal de probabilidades (Figura 12).

Figura 12: Teste de correlação entre as variáveis independentes.

	X1	X2	X3
X2	-0,095 0,027		
X3	-0,150 0,000	-0,042 0,330	

Cell Contents: Spearman rho
P-Value

Fonte: MiniTab17®

Pelo Teste de Correlação de Spearman foi possível observar que há relação entre as variáveis comprimento ($X3$) e bitola ($X1$); largura ($X2$) e bitola ($X1$), pois em ambos os casos o *p-value* foi menor do que 5%, rejeitando-se H_0 , sendo tal hipótese referente à “não há correlação”.

Mesmo a correlação entre as variáveis sendo fraca (menor que 0,3), para maior acuracidade nos resultados da OvSMC, os valores encontrados foram adicionados à matriz de correlação do *Crystal Ball®*. Os resultados da nova simulação estão ilustrados nas Figuras 13 e 14.

Figura 13: Resultado para a variável resposta “Tempo de Usinagem” com correlação. Fonte: *Crystal Ball®*

Figura 14: Resultado para a variável resposta “Tempo de Ajustagem” com correlação. Fonte: *Crystal Ball®*

Na Tabela 3 são apresentados os limites superior e inferior, com 95% de confiança, para as variáveis dependentes. Na Tabela 4

são apresentados os ajustes propostos para as variáveis independentes.

Tabela 3: Intervalos $[L/I, LS]$ obtidos na OvSMC para variáveis de resposta com correlação.

Valor final das funções objetivo	L/I	LS
$\hat{Y}_1$	1.38	1.5
$\hat{Y}_2$	0.55	0.58

Tabela 4: Ajuste proposto para as variáveis independentes com correlação

Variável de Entrada	Valor
X_1 [pol]	1,5
X_2 [mm]	290
X_3 [mm]	180

Entre os gestores consultados, todos pertencentes à empresa que foi objeto do estudo deste trabalho, foi consensual que os ajustes propostos eram possíveis e atenderam à meta de tempo estipulada, sendo os resultados da OvSMC considerando o efeito da correlação mais aderentes a realidade da empresa.

4. CONCLUSÕES E RECOMENDAÇÕES PARA FUTURAS PESQUISAS

Todos os objetivos gerais e específicos foram plenamente atendidos, e os resultados gerados foram promissores e aplicáveis na prática, respondendo à questão da pesquisa.

Com base nos resultados obtidos, é possível concluir que a OvSMC é uma ferramenta que contribui na melhora da determinação de parâmetros envolvidos em processos de manufatura de peças do setor aeronáutico. Tal conclusão se justifica à medida que os ajustes propostos foram validados junto aos gestores da empresa e atenderam ao tempo

de processo estipulado pelos mesmos gestores, dentro de um intervalo de confiança de 95%.

Os objetivos específicos foram satisfeitos ao passo que a função empírica desenvolvida foi adequada para representar o processo real de manufatura da empresa, bem como forneceu ajustes apropriados para bitola, comprimento e largura, predizendo os tempos de usinagem e ajustagem por meio da OvSMC.

Vale ressaltar que, na elaboração de um modelo que represente de forma aderente à realidade, a análise de correlação foi importante e influenciou no ajuste das variáveis de decisão, pois, no caso apresentado, foram geradas variações na solução proposta, mesmo a correlação tendo sido fraca (menor que 0,3).

Para futuros trabalhos se sugere a aplicação de modelos de OvSMC em indústrias de diferentes setores ou em problemas de natureza diferente como, por exemplo, no mercado financeiro.

REFERÊNCIAS

- [1] Abdelaziz, F. B.; Aouni, B. & El Fayedh, R. Multi-objective stochastic programming for portfolio selection. *European Journal of Operational Research*, v. 177, 1811-1823, 2007.
- [2] Aouni, B., Ben Abdelaziz, F. & Martel, J. M. Decision-makers preferences modeling in the stochastic goal programming. *European Journal of Operational Research*, v. 162, p. 610-618, 2005.
- [3] Assis, A. Francis, F. Rabechini JR, R. Gestão de Riscos: Fundamentos e Práticas – Análise Qualitativa e Quantitativa de Riscos. 100 p. Goiânia, 2013.
- [4] Associação das Indústrias Aeroespaciais do Brasil – AIAB. Números da Associação das Indústrias Aeroespaciais do Brasil. 2017 Acesso em: 11 de abril de 2017. Disponível em: < <http://www.aiab.org.br/numeros-da-aiab.asp>>.
- [5] Bertrand, J. W. M, Fransoo, J. C. Operations management research methodologies using quantitative modeling. *International Journal of Operations and Production Management*, v. 22, p. 241 – 264, 2002.
- [6] Boylan, G. L.; Goethals, P. L.; CHO, B. R. Robust parameter design in resource-constrained environments: an investigation of trade-offs between costs and precision within variable processes. *Applied Mathematical Modelling*, v. 37, p. 2394-2416, 2013.
- [7] Corrar, L. J. O modelo econômico da empresa em condições de incerteza - aplicação do método de simulação de Monte Carlo. *Caderno de Estudos*, São Paulo: FIPECAFI/FEA/USP, n. 8, p. 1-13, 1993.
- [8] Deb, K. Multi-Objective optimization using evolutionary algorithms. Chichester: John Wiley & Sons, 2001.
- [9] Ehrgott, M.; Ide, J.; Schöbel, A. Minmax robustness for multi-objective optimization problems. *European Journal of Operational Research*, v. 239, p. 17- 31, 2014.
- [10] Ferraro, M. B.; Giordani, P. A. Multiple linear regression model for imprecise information. *Metrika*, v. 75, p.1049-1068, 2012.
- [11] Gomes, S. B. V. A indústria aeronáutica no Brasil: evolução recente e perspectivas. Rio de Janeiro: BNDES, 2015.
- [12] Helene, W. . Rethinking the linear regression model for spatial ecological data. *Ecology*. 94, p. 2381 -2391, 2013.
- [13] Hillier, F.S., Lieberman, G.J. Introduction to Operations Research, ninth ed. McGraw-Hill, New York, 2010.
- [14] Kleijnen, J. P. C.; WAN, J. Optimization of simulated systems: Optquest and alternatives. *Simulation Modelling Practice and Theory*, v. 15, p. 354-362, 2007.
- [15] Kokonendji, C. C., Kiessé, T. S.; Zocchi, S. S. Discrete triangular distributions and non-parametric estimation for probability mass function.

Journal of Nonparametric Statistics, v. 19, p. 241–257, 2007.

[16] Liu, H.; Wei, G.; XU, Z.; Liu, P.; LI, Y. Quantitative analysis of Fe and Co in Co-substituted magnetite using XPS: The application of non-linear least squares fitting (NLLSF). Applied Surface Science, v. 389, p. 438–446, 2016.

[17] Montoro, G. C. F.; Migon, M. N. Cadeia Produtiva Aeronáutica Brasileira: oportunidades e desafios. Rio de Janeiro: BNDES, 2009. Acesso em: 11 de abril de 2017. Disponível em: <<https://web.bndes.gov.br/bib/jspui/handle/1408/1911>>

[18] Moraes, M. B.; Oliveira, E. A. A. Q.; Lobosco, A.; Souza, V. B. R. Regional Innovation System of Brazilian Aeronautical Industry. Latin American Journal of Business Management, v. 7, n. 1, p. 152-166, 2016.

[19] Oliveira, P.H.D. Barros, N.R. dos Reis, S.G. Aplicabilidade do Método de Simulação Monte Carlo na Previsão dos Custos de Produção de Companhias Industriais: O Caso da Companhia Vale do Rio Doce. Revista de Contabilidade e Organizações – USP. São Paulo, 2010. Acesso em: 14 de abril de 2017. Disponível em: <<http://www.revistas.usp.br/rco/article/view/34781>>

[20] Oracle. Oracle Crystal Ball – Visão Geral. Disponível em: <<http://www.oracle.com/br/products/applications/crystalball/overview/index.html>>. Acesso em: 22 de abril de 2017.

[21] Paiva, E. J.; Ferreira, J. R.; Paiva, A. P. Otimização de processos de manufatura com múltiplas respostas baseada em índices de capacidade Cpm. ENEGEP, Foz do Iguaçu, Paraná, Brasil, 2007. Acesso em: 22 de abril de 2017. Disponível em: <http://www.abepro.org.br/biblioteca/enegep2007_T580439_0302.pdf>

[22] Rodrigues, D. M. Otimização Multiobjetivo em Problemas de Delineamento de Experimentos. Guaratinguetá: FEG – UNESP; 2016. Acesso em: 22 de abril de 2017. Disponível em: <<https://repositorio.unesp.br/handle/11449/144446>>.

[23] Sakurada, N. Miyake, D.I. Aplicação de Simuladores de Eventos Discretos no Processo de Modelagem de Sistemas de Operações de Serviços. Gestão da Produção, v.16, n.1, p. 25-43. São Carlos, 2009.

[24] Shahinidis, N. V. Optimization under uncertainty: state of the art and opportunities. Computers and Chemical Engineering, v. 28, p. 971–983, 2004.

[25] Silva, A.F., Silva, M.F.F., Marins, F.A.S. Otimização Estocástica com Múltiplos Objetivos e Simulação Monte Carlo no Desenvolvimento de Estratégias de Vendas. PODes – Revista Eletrônica Pesquisa Operacional para o Desenvolvimento, v.6, n.1, p.35-53. Rio de Janeiro, 2014.

Capítulo 8

INTEGRAÇÃO DE MÉTODO AMD COM TÉCNICAS GIS PARA CLASSIFICAR ZEIS: UM CASO DE ESTUDO NA CAPITAL DO CEARÁ DA REGIÃO NORDESTE DO BRASIL

Luiz Otavio Ribeiro Afonso Ferreira

Luís Miguel Alçada Tomás de Almeida

Carlos Jorge Lima de Freitas

Mario Fracalossi Jr.

Resumo: Obter o título de propriedade residencial é desejo de muitos brasileiros das regiões carentes da sociedade. Percebendo esta necessidade, o governo preocupa-se na regularização fundiária das zonas especiais de interesse social – ZEIS, através da definição dos parâmetros urbanísticos para cada ZEIS, considerando aspectos quantitativos, qualitativos e julgamentos subjetivos. Este estudo concentra-se na capital do estado do Ceará, sendo à 5ª do Brasil em população, totalizando 2.891.188 habitantes. Utilizou-se o apoio multicritério à decisão – AMD, como metodologia para auxiliar na tomada de decisão com relação à sequência de intervenções nas ZEIS, integrado ao Sistemas de Informação Geográfica – SIG/GIS. Aplicou-se o método multicritério ELECTRE TRI por influência da integração com o ambiente GIS. Conforme os dados obtidos pelo tomador de decisão, foi possível apresentar de forma incontestável a categoria enquadrada nas ZEIS em duas versões: uma mais e outra menos rigorosa (pessimista, otimista), conforme os critérios e níveis de referência.

Palavras chave. Apoio Multicritério à Decisão, Sistemas de Informação Geográfica, Zonas Especiais de Interesse Social.

*Artigo originalmente publicado nos anais do Simpósio Brasileiro de Pesquisa Operacional 2016

1. INTRODUÇÃO

Na década de 80, os brasileiros das mais diversas regiões carentes do país, realizaram uma mobilização social para demonstrar o desejo de receber o título de propriedade residencial e logrou-se êxito com o encaminhamento ao congresso constituinte de uma emenda popular pela reforma urbana. Com a inserção do capítulo da política urbana na constituição de 1988 e a vinculação constitucional entre direito de propriedade e sua função social, cujo entendimento da lei é o de que haja a prevalência do interesse comum sobre direito individual de propriedade, isso implica o uso socialmente justo e ambientalmente equilibrado do espaço urbano.

A partir deste momento o plano diretor tornou-se obrigatório para os municípios com mais de 20.000 habitantes, convertendo-se em um valioso instrumento, estabelecendo em quais condições a propriedade urbana deve cumprir a sua função social. Posteriormente, a lei federal de nº 10.257/2001, conhecida como estatuto da cidade, estabeleceu as diretrizes gerais da política urbana a serem seguidas por todos os municípios brasileiros instituindo, entre outras ferramentas, o instrumento jurídico-político das ZEIS. Pretendeu-se, desta forma, facilitar a regularização fundiária e a urbanização de assentamentos da população de baixa renda, mediante normas urbanísticas especiais capazes de integrar os territórios da cidade.

Assim, o zoneamento urbano pode ser um impeditivo para implantação da política de regularização fundiária, motivo pelo qual a legislação municipal deve prever as ZEIS em áreas da cidade que demandem um tratamento especial e prioritário.

A prefeitura municipal de Fortaleza atende à nova ordem legal urbana, através do plano diretor participativo – lei complementar 062 de 02 de fevereiro de 2009, para promover o direito às cidades sustentáveis, segundo [Bezerra et al. 2015].

Este artigo tem por objetivo classificar e realizar uma ordenação das ZEIS de uma forma racional, baseando-se em modelos matemáticos para realizar esta avaliação e integrar com os Sistemas de Informação Geográficas – SIG, apresentando uma sistemática robusta para se trabalhar toda área urbanística com as ZEIS mapeadas. O Apoio Multicritério à Decisão – AMD, é um segmento dentro da área de pesquisa

operacional utilizada no meio acadêmico na indústria em geral e nos órgãos governamentais entre outros. Ao lidar com critérios quantitativos e qualitativos, este método permite que seja levado em consideração a subjetividade inerente ao processo de urbanização, trabalhando a questão de ordenação e classificação das ZEIS.

Para um cidadão adquirir crédito junto aos órgãos de fomento, entre outros, faz-se necessário o registro do imóvel, entretanto, os moradores das Zonas Especiais de Interesse Social não conseguem o título de propriedade por não existirem parâmetros de urbanização definidos junto às prefeituras. Sendo o bairro uma unidade já parametrizada, tenta-se utilizar os mesmos procedimentos para as ZEIS localizadas na região, porém, não é possível ou viável o enquadramento dos mesmos parâmetros. Para tanto, o governo municipal preocupado em resolver esta não conformidade dos parâmetros urbanísticos e atender à lei federal de nº 10.257/2001, resolve através do IPLANFOR – Instituto de Planejamento de Fortaleza – elaborar uma agenda de regulamentação das ZEIS. Esta ação pretende acelerar e viabilizar a emissão dos títulos de propriedade de cada cidadão, logo, o sucesso deste processo depende diretamente do nível de modernização/evolução de cada ZEIS, sendo os principais indicadores do início do processo a causar demora na liberação do título.

O AMD auxiliará a efetuar as comparações necessárias de forma mais lógica, reduzindo os erros inerentes num processo decisório, causados muitas vezes pela falta de tempo hábil para uma análise mais profunda, pelos prazos reduzidos dos projetos, por opiniões e preferências da alta gestão, tornando-se a ferramenta mais adequada de classificação para as ZEIS.

É importante para a organização que as respostas às necessidades aos mais diversos problemas tenham embasamento técnico e forneçam informações para argumentação. A confiabilidade destas informações é que ajudarão os tomadores de decisão a direcionar melhor os recursos.

A utilização da ferramenta de GIS integrada a AMD proporcionará ao processo de decisão uma avaliação sistêmica e flexível, visando reduzir o tempo e a complexidade na avaliação das ZEIS, proporcionando

benefícios como a classificação e ordenação das ZEIS.

Na introdução foram apresentados os objetivos da pesquisa. Este trabalho é organizado da seguinte maneira: Revisão bibliográfica, para se obter conhecimento a respeito do AMD para justificar a escolha do método; O método ELECTRE TRI, descrição e integração com o GIS; Aplicação prática do ELECTRE TRI integrado com QGIS. Para o estudo do método, escolheu-se avaliar as ZEIS de Fortaleza, utilizando dados espaciais coletados pelo PLHISFOR, utilizar ferramenta computacional para integrar GIS com AMD na resolução do problema e a conclusão baseada nos resultados obtidos, bem como propostas futuras.

2.METODOLOGIA

Para elaboração deste trabalho, primeiramente, foi realizada uma revisão bibliográfica. Esta fase consiste na pesquisa, estudo e compreensão dos métodos AMD e integração com ferramentas GIS disponíveis principalmente em livros, artigos de periódicos e materiais disponibilizados na internet. Tem por objetivo principal gerar conhecimento e embasamento teórico para aplicação prática dirigida ao caso de estudo.

Em seguida, na fase do caso de estudo, seguiremos os procedimentos da figura 1, na elaboração do problema de decisão:

Figura 1: adaptado de [Silva et al. 2014]

Os métodos de classificação comparam as alternativas com valores pré-determinados pelo tomador de decisão, o que normalmente expressa o seu nível de aspiração com relação as alternativas. O método ELECTRE TRI [Yu 1992], é empregado para classificar alternativas ou ações potenciais com níveis de referência estabelecidos previamente para qualificá-las, por exemplo, em muito viáveis, viáveis, pouco viáveis e não viáveis. O GIS está sendo integrado com diversas áreas, inclusive com AMD como ressalta [Silva et al. 2014]. São ferramentas com poderosos recursos computacionais e de fácil acesso, tanto as ferramentas comerciais quanto as de software livre.

Esta parte se dedica à busca de conhecimento que sustentará a elaboração deste artigo. É feita uma revisão sobre AMD: breve histórico, classificação, integração com GIS e justificativa para utilização da solução de integração de AMD e GIS.

3. APOIO MULTICRITÉRIO À DECISÃO

A necessidade de enfrentar situações específicas, nas quais um tomador de

decisão deveria resolver um problema onde haviam muitos objetivos a serem alcançados de forma simultânea, deu-se a origem dos primeiros métodos de apoio multicritério à decisão – AMD.

Os componentes básicos são um conjunto finito ou infinito de ações, no mínimo dois critérios e um tomador de decisão. Pode ser caracterizado como tomador de decisão, um indivíduo ou um grupo de indivíduos, que tomam a decisão de acordo com suas preferências ou desempenho das alternativas ou ambos os casos; decisão é um processo que permite a escolha de uma ou mais alternativas dentre todas apresentadas e a alternativa é uma das possibilidades que compõem o objeto da decisão, qualificada como potencial quando existe algum interesse no âmbito do processo decisório ou quando esta pode ser implementada e pode ser identificada no início ou no decorrer do processo.

Conforme [Gomes et al. 2012], os elementos básicos de um problema de decisão multicritério são: Tomador de decisão, Analista/Agente de decisão, Especialistas,

Pessoas afetadas pela decisão, Modelo, Alternativas, Critérios.

Em muitos problemas, as funções das pessoas envolvidas numa questão de decisão se confundem, pois o agente de decisão pode ser um especialista e este, além do conhecimento específico numa área pode ser também um tomador de decisão.

Segundo [Bana et al. 1995], em um sistema aberto, composto pelos atores, seus valores e seus objetivos, a tomada de decisão deve buscar a opção que apresente o melhor resultado e a as suas disponibilidades em adotá-la, considerando a relação entre elementos objetivos e subjetivos.

De acordo com [Gomes et al. 2012], existem quatro tipos de problemática de referência na qual pode-se identificar o tipo de problema no contexto do AMD: Escolha, Classificação (*sorting*), Ordenação (*ranking*) e Descrição. São quatro as situações fundamentais e mutuamente excludentes das preferências do tomador de decisão: Indiferença (I), Preferência Estrita (P), Preferência fraca (Q) e Incomparabilidade (R ou RC), o autor ainda complementa com relação a estrutura de preferência, tem-se: Ordem completa, Pré-ordem completa, Quase-ordem e Ordem de intervalo, Pré-ordem parcial, Pseudo-ordem.

Alguns autores demonstram suas classificações para os métodos de AMD (ver tabela 1).

Tabela 1: classificações para os métodos de AMD

Nº	[Almeida 2013]	[Pardalos et al. 1995]	[Gomes et al. 2012]	[Vincke 1992]
1	Método de Critério Único;	Programação Matemática Multiobjetivo;	Primeiro Grupo a) Escola Americana; b) Escola Francesa; c) Métodos Iterativos; d) Híbrida; e) Outras Escolas.	Teoria de Utilidade Multiatributo;
2	Método de Sobreclassificação (Outranking, Surclassment); superação, prevalência ou subordinação;	Teoria da Utilidade Multiatributo (MAUT);	Segundo grupo;	Os Métodos de Subordinação;
3	Métodos Iterativos.	Método de Sobreclassificação (outranking); superação, prevalência ou subordinação;	Terceiro Grupo.	Os Métodos Iterativos.
4		Abordagem de Desagregação de Preferências.		

Os principais métodos de AMD são: MAUT [Keeney and Raiffa 1993], SMART [Edwards and Barron 1994], AHP (Saaty), MACBETH [Bana et al. 1995], PROMÉTHÉE [Brans and Vincke 1985], TODIM [Gomes and Others 2009], UTADIS [Devaut et al. 1980] e métodos ELECTRE [Bouyssou and Roy 1993].

Para completar o assunto, temos os tipos de critérios utilizados no AMD: Critérios verdadeiros, semi-critério, critério de um intervalo e pseudo-critério. Os métodos de superação também podem ser chamados de sobreclassificação, prevalência ou subordinação em síntese [Gomes et al. 2012].

4. MÉTODO ELECTRE TRI

O Método ELECTRE TRI caracteriza-se por tratar de problemas específicos de classificação ordenada. Ou seja: dado um conjunto $\underline{A}$ de alternativas, associá-las a um conjunto de classes ordenadas $\underline{C} = [C_1, C_2, \dots, C_n]$.

Isso é feito considerando o desempenho de A à luz de um conjunto de critérios $\underline{F} = [g_1, g_2, \dots, g_m]$. Um conjunto formado por $(p+1)$ classes, delimitadas por p limites de classes. Uma classe genérica C_h é delimitada por um limite inferior b_h e um limite superior b_{h-1} .

Segundo [Yu 1992] e [Mousseau et al. 2000], esse método integra funções que dão suporte ao tomador de decisão no processo de

preferência e reduzem o esforço cognitivo requerido na fase de modelagem.

5. RELAÇÃO DE SUBORDINAÇÃO NO ELECTRE TRI

A relação de subordinação é construída para tornar possível a comparação de uma alternativa a com um limite padrão b_h . A afirmação de que aSb_h , significa que “ a não tem um desempenho pior do que o limite b_h ”.

Na validação da afirmação aSb_h devem-se verificar duas condições:

- Concordância: para que aSb_h (ou b_hSa) seja aceita, uma maioria suficiente de critérios deve ser a favor desta afirmação.
- Não-discordância: quando na condição de concordância esperada, nenhum dos critérios na minoria deve se opor à afirmação aSb_h (ou b_hSa).
- Os seguintes passos são seguidos na obtenção da relação de subordinação:
- Computar o índice de concordância parcial $c_j(a, b_h)$ e $c_j(b_h, a)$,
- Computar o índice de concordância global $c(a, b_h)$,
- Computar o índice de discordância parcial $d_j(a, b_h)$ e $d_j(b_h, a)$,
- Computar a relação de subordinação fuzzy conforme o índice de credibilidade $\sigma(a, b_h)$,
- Determinar um plano de corte λ , referente a relação fuzzy para obter uma relação de subordinação. Isto é: se $\sigma(a, b_h) \geq \lambda \Rightarrow aSb_h$.

O grau de credibilidade da relação de subordinação $\sigma(a, b_h)$ expressa com que intensidade se pode “acreditar” que “ a subordina b_h ” de acordo com o índice de concordância global $c_j(a, b_h)$ e com o índice de discordância $d_j(a, b_h)$.

A relação de subordinação fuzzy S é obtida com base em um nível de corte λ . Esse nível é considerado como o menor valor do índice de credibilidade compatível com a afirmação de que “ a subordina b_h ”. Isto é: aSb_h se e somente se $\sigma(a, b_h) \geq \lambda$.

6. PROCEDIMENTOS DE CLASSIFICAÇÃO

O ELECTRE TRI classifica as alternativas seguindo dois passos consecutivos:

construção de uma relação de subordinação S , que caracteriza como as alternativas são comparadas aos limites das classes; e, exploração (através de procedimentos de classificação) da relação S .

A regra do procedimento de exploração é realizada para analisar o modo em que uma alternativa a é comparada com os limites padrão determinados para a classe na qual a deve ser enquadrada. Dois procedimentos de classificação são avaliados.

O procedimento de classificação descendente é descrito a seguir:

- Compare a sucessivamente com b_i , para $i = p, p - 1, \dots, 1$.
- Encontre um b_h que seja o primeiro limite tal que aSb_h .
- Classifique a na classe C_{h+1} . Ou seja: na Classe limitada inferiormente pelo limite b_h .
- O procedimento de classificação ascendente é descrito a seguir:
- Compare a sucessivamente com b_i , $i = 1, 2, \dots, p$.
- Encontre o primeiro b_h para o qual $b_h > a$.
- Classifique a na classe limitada superiormente por este limite. Ou seja: classifique a na classe C_h .

Sendo estes dois procedimentos diferentes, consequentemente, pode ocorrer a classificação de algumas alternativas em diferentes classes, pois: o procedimento otimista tende a classificar as alternativas em classes mais altas; o procedimento pessimista tende a classificar as alternativas nas categorias mais baixas.

No caso do ELECTRE TRI, uma divergência entre estas classificações indica uma incapacidade do sistema em comparar a alternativa a algum dos perfis das classes de equivalência utilizadas.

Tradicionalmente, considera-se que essa incapacidade pode ser causada tanto por incoerência do avaliador, quanto do modelo de classificação (incluindo o conjunto de critérios) ou pelo sistema de coleta de dados. A divergência das classificações é comum em situações em que haja critérios conflitantes (por exemplo: qualidade e tempo), sendo neste caso inerente ao problema e não devendo ser considerada como uma falha na modelagem.

Assim, quando ocorre divergência entre as classificações pessimista e otimista, o classificador deve adotar uma das duas classificações de acordo com o seu perfil: (mais exigente ou menos exigente). Para futuros detalhes do conceito e do algoritmo deste método, veja [Yu 1992], [Bouyssou and Roy 1993], [Bouyssou and Roy 1993], [Mousseau and Dias 2004] e [Figueira et al. 2005, 2010].

7. CASO DE ESTUDO

A região metropolitana de Fortaleza é composta por 15 municípios (ver mapa 1), inserida nesta região metropolitana está a cidade de Fortaleza, caso de estudo. O

prefeito da cidade tem realizado um trabalho de levantamento e cadastro dos dados pertinentes as ZEIS, através do PLHISFOR – Plano Local de Habitação de Interesse Social de Fortaleza e a base de dados foi disponibilizada pelo IPLANFOR – Instituto de Planejamento de Fortaleza, possuindo 127 registros sobre as ZEIS, estes registros estão no formato espacial e representado através de polígonos (ver mapa 2). Este mapa, apresenta as principais vias em conjunto das ZEIS, delimitada pela área da cidade de Fortaleza e apresenta uma divisão estrutural de seis secretarias regionais e mais uma centro.

Região Metropolitana de Fortaleza - Mapa 1

Cidade de Fortaleza- Mapa 2

A localização geográfica das ZEIS (ver mapa 3) e a lista de atributos disponibilizados no banco de dados do IPLANFOR (ver tabela 2). Diante dos atributos disponibilizados, foram selecionados 7 (sete) que serviram como critérios do caso de estudo. Os atributos selecionados são: (área_m2, densidade, num_imov, renda_med, dom_munic, dom_partic, a_vazios), considerados como critérios: (Área, Densidade, Número de

Imóveis, Renda, Domínio Municipal, Domínio Particular, Área Vazias). Estes critérios serão abordados neste estudo.

As ZEIS são classificadas em 3 (três) tipologias instituídas no Plano Diretor, ou seja, as 45 (quarenta e cinco) ZEIS 1 (ZEIS de Ocupação), as 56 (cinquenta e seis) ZEIS 2 (ZEIS de Conjuntos, Mutirões e Loteamentos irregulares) e as 34 (trinta e quatro) ZEIS 3 (ZEIS de Vazio).

ZEIS da cidade de Fortaleza - Mapa 3

A equipe de Tecnologia de Informação e Comunicação – TIC do IPLANFOR, realizou um trabalho conjunto com os órgãos internos e externos da prefeitura municipal de

Fortaleza e disponibilizou os dados para o caso de estudo. A tabela 2 demonstra todos os atributos da base de dados.

Critérios - Tabela 2

Nome	Órgão de Origem	Usado Por	Formato Original
zeis	PLHISFOR	vw_zeis_fm	PostgreSQL
Estrutura			
Nome do Campo	Formato dos Dados	Informação	
gid	integer	Identificador	
name	character varying(254)	Nome da zeis	
zeis	character varying(1)	Tipo de zeis (só número)	
tipo	character varying(254)	Tipo de zeis	
cod	character varying(254)	Código PLHIS	
regional	character varying(254)	Regional	
area_m2	character varying(254)	Área	
popul_ass	character varying(254)	População do assentamento	
qtd_ass	character varying(254)	Quantidade de assentamentos	
popul_zeis	character varying(254)	População da zeis	
num_imov	character varying(254)	Número de imóveis	
densidade	character varying(254)	Densidade populacional na zeis	
renda_med	character varying(254)	Renda média dos moradores	
dom_munic	character varying(254)	Área de domínio municipal	
dom_estado	character varying(254)	Área de domínio estadual	
dom_uniao	character varying(254)	Área de domínio federal	
dom_partic	character varying(254)	Área de domínio particular	
dom_religi	character varying(254)	Área de domínio de organizações religiosas	
dom_seminf	character varying(254)	Área sem informação de domínio	
a_vazios	character varying(254)	Área Vazios	
uh_horizon	character varying(254)	Unidades Habitacionais Horizontais	
uh_vertica	character varying(254)	Unidades Habitacionais Verticais	
obs	character varying(254)	Observações	
geom	geometry(MultiPolygon,31984)	Geometria	

O governo municipal demonstrou interesse em realizar o caso de estudo para classificar / ordenar as ZEIS com o intuito de conceber os parâmetros de urbanização e realizar a tomada de decisão quanto à ZEIS que serão estudadas conforme apresentação da classificação nas categorias.

8. ELICITAÇÃO DOS PARÂMETROS

As preferências do tomador de decisão foram obtidas de forma verbal [Dias et al. 2002], para a utilização do método ELECTRE TRI neste trabalho de classificar as alternativas conforme as categorias predefinidas, são categorizadas da seguinte maneira: Categoria 1 – Baixa adequação, Categoria 2 – Média adequação, 3 – Moderada adequação e Categoria 4 – Alta adequação.

O método avaliará a adequação de acordo com os múltiplos critérios considerados pela comparação de cada perfil das categorias, a fim de classificar as ZEIS mais adequadas para os estudos dos parâmetros urbanísticos, com os valores dos parâmetros subjetivos

sendo especificado por especialistas, legislação e outros estudos.

Os parâmetros subjetivos mais importantes são compostos por três perfis de referência (B_1 , B_2 e B_3). O parâmetro B_3 de referência define o conjunto de valores dos critérios de desempenho que separam a adequação moderada da alta, o B_2 de referência define o conjunto de valores dos critérios de desempenho que separam a adequação média da moderada, enquanto o B_1 define os limites de referência que separam a adequação Baixa da Média.

Outros parâmetros subjetivos são: os três limiares indiferença, preferência e veto, que podem ser associados a cada critério e estes parâmetros foram todos considerados iguais pelo tomador de decisão, entretanto, o parâmetro pesos foi preferido de acordo com a tabela 3.

Os critérios considerados estão presentes na tabela 3.

Tabela 3

Critérios		Área	Densidade	Domínio Particular	Domínio Municipal	Número de Imóveis	Renda	Área Vazias
Custo		Min	Max	Min	Max	Min	Max	Max
Unidade		M²	Hab/M²	M²	M²	Unitário	Salário	M²
Pesos		5	20	5	35	15	10	10
Classes	b ₁	100000	0,02	50000	600	1200	2,5	1000
	b ₂	60000	0,03	30000	5300	750	3,0	8000
	b ₃	20000	0,04	10000	10001	300	3,5	15000
Nível de corte		0,55						
Indiferença		5	5	5	5	5	5	5
Preferencia		15	15	15	15	15	15	15
Veto		20	20	20	20	20	20	20

Foi utilizado o servidor de algoritmos [Alçada-Almeida 2013] para obtenção dos resultados das categorias e o aplicativo QGIS [QGIS Development Team 2013], para exibir os mapas.

9. ANÁLISES DE SENSIBILIDADE

Começamos a utilização do método por aplicar a classificação dos polígonos para os

dois cenários Otimista e Pessimista. Se os resultados obtidos na classificação são os mesmos em ambos os cenários, isto significa que podemos utilizar estes locais na categoria particular, e assim a sua classificação é obtida independentemente da localização da ZEIS dentro deste local adequado. No final, as alternativas mais adequadas são identificadas nos mapas (4 e 5).

Pessimista - Mapa 4

Otimista - Mapa 5

Foram detectadas 4 ZEIS mudando de categoria do cenário pessimista para otimista.

10.CONCLUSÃO

Este trabalho apresentou uma abordagem multicritério para a classificação das ZEIS. Os resultados obtidos demonstram que esta é uma alternativa que pode ser considerada no tratamento deste tipo de problema. Este trabalho torna-se inédito no Brasil e com impacto social elevado para as comunidades mais carentes do nosso país. Esta é a contribuição central do presente trabalho.

Os dados utilizados na construção da aplicação aqui reportada foram obtidos de forma direta junto ao especialista no problema abordado, após a extensa leitura do valioso

trabalho realizado pelo órgão competente. No entanto, mesmo com estas características, trata-se de uma modelagem particular e, sendo assim, ressalta-se, que os resultados são particulares e não devem ser extrapolados para situações distintas da investigada. O ponto a ser destacado é que a opção pelo ELECTRE TRI, em detrimento do emprego de outros métodos multicritérios deve-se ao fato de que este é um método nativo para o tratamento de classificação ordenada.

No futuro esta pesquisa poderá expandir para aplicação de outros problemas reais, em outros contextos. Esta metodologia pode ser expandida para o domínio de tomada de decisão em grupos por considerar a opinião de grupos mais diversificada. Como apresentado em [Silva et al. 2014] a limitação

da linguagem no ambiente comercial, podemos construir ou reprogramar em outras linguagens de desenvolvimento de aplicações em ambiente com informações espaciais.

BIBLIOGRAFIA

- [1] Alçada-Almeida, L. (2013). webapp Matrix. <http://mad.dec.uc.pt/accounts/register/>, [accessed on Apr 19].
- [2] Almeida, A. T. De (2013). Processo de Decisão nas Organizações: Construindo modelos de decisão multicritério.
- [3] Bana, E., Costa, C. A. and Vansnick, J. C. (1995). Uma nova abordagem ao problema de construção de uma função de valor cardinal: MACBETH. *Investigação operacional*, v. 15, p. 15–35.
- [4] Bezerra, R. C. R., Lucena, G., Santana, E. W. De, Fracalossi, M. and Perdigão, L. de M. M. e S. (2015). Relatório das zeis.
- [5] Bouyssou, D. and Roy, B. (1993). Aide multicritere a la decision: Methodes et cas,..., "Economica", Paris,
- [6] Brans, J.-P. and Vincke, P. (1985). Note—A Preference Ranking Organisation Method: (The PROMETHEE Method for Multiple Criteria Decision-Making). *Management science*, v. 31, n. 6, p. 647–656.
- [7] Devaud, J. M., Groussaud, G. and Jacquet-Lagrange, E. (1980). UTADIS: Une méthode de construction de fonctions d'utilité additives rendant compte de jugements globaux. *European Working Group on Multicriteria Decision Aid*, Bochum,
- [8] Dias, L., Mousseau, V., Figueira, J. and Clímaco, J. (2002). An aggregation/disaggregation approach to obtain robust conclusions with ELECTRE TRI. *European Journal of Operational Research*, v. 138, n. 2, p. 332–348.
- [9] Edwards, W. and Barron, F. H. (1994). SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement. *Organizational behavior and human decision processes*, v. 60, n. 3, p. 306–325.
- [10] Figueira, J., Greco, S. and Ehrgott, M. (2005). Multiple criteria decision analysis: state of the art surveys. *Springer Science & Business Media*. v. 78
- [11] Figueira, J. R., Greco, S., Roy, B. and Slowinski, R. (2010). ELECTRE methods: main features and recent developments. *Handbook of multicriteria analysis*. Springer. p. 51–89.
- [12] Gomes, L. F. A. M., Gomes, C. F. S. and De Almeida, A. T. (2012). Tomada de decisão gerencial: enfoque multicritério. 4. ed. - S ed. Atlas.
- [13] Gomes, L. F. A. M. and Others (2009). An application of the TODIM method to the multicriteria rental evaluation of residential properties. *European Journal of Operational Research*, v. 193, n. 1, p. 204–211.
- [14] Keeney, R. L. and Raiffa, H. (1993). Decisions with multiple objectives: preferences and value trade-offs. *Cambridge university press*.
- [15] Mousseau, V. and Dias, L. (2004). Valued outranking relations in ELECTRE providing manageable disaggregation procedures. *European Journal of Operational Research*, v. 156, n. 2, p. 467–482.
- [16] Mousseau, V., Slowinski, R. and Zielniewicz, P. (2000). A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. *Computers & operations research*, v. 27, n. 7, p. 757–777.
- [17] Pardalos, P. M., Siskos, Y. and Zopounidis, C. (1995). *Advances in Multicriteria Analysis*. Springer US.
- [18] QGIS Development Team (2013). QGIS User Guide Release 1.8.0. QGIS Project, p. 267.
- [19] Silva, S., Alçada-Almeida, L. and Dias, L. C. (oct 2014). Development of a Web-based Multicriteria Spatial Decision Support System for the assessment of environmental sustainability of dairy farms. *Computers and Electronics in Agriculture*, v. 108, p. 46–57.
- [20] Vincke, P. (1992). *Multicriteria decision-aid*. John Wiley & Sons.
- [21] Yu, W. (1992). ELECTRE TRI(aspects méthodologiques et manuel d'utilisation). Document- Université de Paris-Dauphine, LAMSADE,

Capítulo 9

JOGOS COOPERATIVOS NA ALOCAÇÃO DE CUSTOS DE ESTOQUES DE PEÇAS COMPARTILHADOS

*Natália Nogueira Ferreira Souza,
Bernardo Santos Aflalo,
Takashi Yoneyama*

Resumo: Este artigo tem o objetivo de estudar diferentes métodos de alocação de custos em um compartilhamento de estoques de componentes entre operadores aeronáuticos. Três modelos de alocação de custos de estoques foram testados: (1) proporcional ao potencial de consumo, (2) proporcional à quantidade de aeronaves e (3) utilizando o valor de *Shapley*, derivado da teoria dos jogos cooperativos. Com o objetivo de mostrar os ganhos relativos ao compartilhamento de estoques, a situação sem compartilhamento, na qual cada operador dimensiona separadamente seu próprio estoque, também foi simulada. Os resultados apontam que a alocação dos custos do compartilhamento de peças pelo método de *Shapley Value* apresenta vantagens frente aos outros métodos, permitindo ganhos individuais proporcionais à contribuição de cada operador, o que aumenta a atratividade do compartilhamento para grandes operadores.

Palavras Chave: Controle de Estoque; Teoria dos Jogos Cooperativos; Alocação de Custos;

1. INTRODUÇÃO

Atualmente, as empresas aéreas chegam a terceirizar até 65 % de todo o serviço referente à manutenção e reparo [Berger, 2014]. Este serviço tem um escopo de cobertura do avião cada vez mais completo, englobando desde itens de linha até itens estruturais.

O dimensionamento dos estoques de peças reparáveis é um ponto que merece muita atenção de seus gestores, pois um estoque maior que o necessário para suportar a frota operacional pode representar um desperdício de recursos, enquanto um estoque subdimensionado pode provocar atrasos e paradas não programadas das aeronaves. Todos os custos envolvidos no processo são de interesse direto das companhias aéreas e fabricantes de aeronaves, e uma alocação adequada destes é ponto importante para a dinâmica do negócio.

Uma alternativa econômica para gestão de estoque é o conceito de estoque *pool* de peças, ou seja, um conjunto de componentes que é compartilhado entre os operadores que operam aeronaves semelhantes, de modo a reduzir a quantidade de peças média para suportar cada aeronave [Karsten et al. 2010]. Dessa maneira, os operadores não precisam investir diretamente nessas peças e gastos diretos como armazenamento e seguro; e indiretos como depreciação e custo de oportunidade são, então, compartilhados entre os participantes deste *pool* de peças.

Nesse contexto, o objetivo deste trabalho é discutir diferentes maneiras de compartilhar os custos envolvidos em estoques compartilhados entre operadores. Três modelos, além da situação sem compartilhamento, são testados: (1) proporcional ao potencial de consumo, (2) proporcional à quantidade de aeronaves e (3) utilizando o valor de *Shapley*, derivado da teoria dos jogos cooperativos [Shoham e Leyton-Brown, 2009].

2. MÉTODOS

Para testar os modelos de compartilhamento, foram simulados 4 operadores com diferentes quantidades de aeronaves e/ou número de horas voadas por ano (HV/Ano). Os estoques de um único componente foram considerados

nas análises, conforme descrito nas subseções a seguir.

2.1 SIMULAÇÃO DAS REMOÇÕES

Para estimativa do MTBUR, considera-se um componente hipotético com as seguintes características:

- Quantidade de componentes por aeronave: 2;
- Densidade de probabilidade de falha de acordo com distribuição *Weibull*, representativa de componentes reais.

Uma simulação baseada em eventos foi realizada utilizando o R, da qual as remoções foram geradas aleatoriamente. Ao final, o tempo médio entre remoções não programadas, mais conhecido como MTBUR (do inglês *Mean Time Between Unscheduled Removal*) foi de 3652. Esse número, apesar de ser bem menor que a média dos componentes no mundo aeronáutico, foi utilizado pois resulta em um estoque relativamente grande de peças, o que permite exercitar a alocação de custos de uma maneira mais efetiva.

2.2 CÁLCULO DO CUSTO TOTAL DE ESTOQUES

Com as informações de MTBUR do componente e o consumo total de cada operador (representado, neste trabalho, pela letra grega λ), é possível iniciar as simulações de dimensionamento de estoque.

2.3 DISTRIBUIÇÃO DE POISSON

A distribuição de Poisson [Miguel 2005] é utilizada para calcular o número de peças necessárias para atingir o nível de pelo menos 95% de disponibilidade de estoque para cada coalizão possível entre os operadores. Como essa distribuição admite apenas valores discretos como entrada (devido ao termo fatorial no denominador), a disponibilidade exata de 95% seria algo fortuito. Na prática, os valores de disponibilidade são, no mínimo, 95%, mas podem atingir valores consideravelmente mais altos. Apesar de isso representar maior disponibilidade de peças para o operador, também implica em gastos indesejados bem mais altos. A

Figura ilustra a alteração do estoque médio por aeronave com o aumento de λ de uma

coalizão hipotética:

Figura 1: Estoque médio por aeronave utilizando a função de Poisson.

Note que, em pontos específicos, tem-se um aumento do estoque médio por aeronave ao aumentar o consumo da mesma. Essa é uma característica gerada pela flutuação de disponibilidade que, por sua vez, é causada pela parte discreta da distribuição de Poisson. Esse tipo de jogo é chamado de não-convexo [Shoham e Leyton-Brown, 2009]: nestas situações, a formação de grandes coalizões são prejudicadas. Isso pode distorcer o resultado da alocação de recursos, anulando alguns efeitos benéficos do uso do valor de *Shapley*.

3. DISTRIBUIÇÃO DE POISSON MODIFICADA

Com a intenção de atingir o valor exato de 95% de confiabilidade e evitar os efeitos demonstrados na Figura 3, há a possibilidade de dimensionar o estoque necessário de cada coalizão modificando a restrição de número inteiro de peças. Isso não significa que, na prática, o estoque deveria ser composto de um número não inteiro de peças, mas apenas que, dado o estoque real, a alocação de custos será realizada considerando um número não inteiro de componentes.

Isso pode ser feito substituindo a função fatorial na distribuição de Poisson pela função

Gama, pois ela representa uma extensão dessa função para o conjunto de números reais [Weisstein et al 2015]. A função Gama, representada por Γ , é definida como [Weisstein 2015]:

$$\Gamma(x) = \int_0^{\infty} \omega^{x-1} e^{-\omega} d\omega \quad (1)$$

A relação com a função fatorial pode ser representada da seguinte maneira:

$$\Gamma(x) = (x-1)! \quad (2)$$

Em que:

- x é um número real ou complexo
- $(x-1)!$ é a fatorial de $x-1$

Essa função gera uma probabilidade contínua, permitindo ao estudo inferir um valor mais preciso do número de peças que atende a disponibilidade exata de 95%, como citado anteriormente. A Figura 4 demonstra o comportamento da função de Poisson com a modificação de Gama, onde é possível perceber a monotonicidade da função, conforme desejado.

Figura 2: Estoque médio por aeronave utilizando a função de Poisson modificada

3.1 RECOMENDAÇÃO DE PEÇAS

O compartilhamento dos custos utilizadas no modelo do valor de *Shapley* será, portanto, uma modificação da equação de Poisson, na

$$f(k; \lambda) = \frac{\exp^{-\lambda} \lambda^k}{\Gamma(k+1)} \quad (3)$$

Com essa adaptação alteramos algumas propriedades da função original: apesar da distribuição de probabilidade ser a mesma nos 'k's inteiros, a distribuição de probabilidade cumulativa, por exemplo, é levemente alterada quando comparada à distribuição de Poisson tradicional. Por esse motivo, essa distribuição será chamada de "Poisson modificada" ao longo do trabalho.

3.2 COMENTÁRIOS ADICIONAIS SOBRE OS CUSTOS DE ESTOQUE

O objetivo do trabalho é verificar qual a melhor maneira de compartilhar os custos do estoque: tanto a aquisição quanto a manutenção do mesmo. Tipicamente, a manutenção envolve seguro, depreciação/obsolescência, aluguel de armazém, homem-hora, custo de oportunidade e outros. Esses custos geralmente são especificados como uma fração do custo da peça armazenada (por exemplo, 10% do valor da peça nova, por ano).

Ao longo deste trabalho, a quantidade de peças será utilizada como uma maneira de representar a alocação de custos. Essa abordagem visa a facilitar as comparações e o entendimento das metodologias de

qual o termo fatorial é substituído pelo equivalente da função Gama:

alocação de custos. O estudo não tem como objetivo fazer o dimensionamento do estoque: para isso, existem metodologias consagradas (dentre as quais, a distribuição de Poisson é talvez uma das mais relevantes).

3.3 MÉTODOS DE ALOCAÇÃO DE CUSTO

Nas subseções a seguir, as diferentes técnicas de alocação de custos de estoques compartilhados são listadas.

3.4 CÁLCULO DO VALOR DE SHAPLEY

A alocação de custos utilizando o valor de *Shapley* procura atribuir, a cada participante, benefícios proporcionais à contribuição individual à coalizão total. Considerando N o número de jogadores e $c(S)$, $\forall S \subseteq N$ cada combinação entre jogadores (Coalizão), o valor alocado para cada indivíduo x_i , $i \in N$ é calculado como:

$$x_i = \frac{1}{N!} \sum_{R} [v(P_i^R \cup \{i\}) - v(P_i^R)] \quad (4)$$

Sendo que a soma inclui todos os termos de $N!$ na ordem R e P_i^R é o set de jogadores que precedem i em R . O número de termos da equação cresce exponencialmente com o

número de jogadores [Moreira; Luna e Guedes 2002], podendo se tornar um cálculo extremamente caro, dependendo da quantidade de participantes (por ex., o número de termos deste cálculo para uma coalizão de 25 indivíduos é da ordem de $1E+25$).

O valor de *Shapley* é uma distribuição “justa” de *payoffs* entre os participantes da coalizão. De maneira geral, jogadores que contribuem mais para os ganhos da coalizão, recebem mais vantagens por isso. Nesse sentido, grandes operadores tendem a ser mais beneficiados pela adoção de uma alocação de custos dessa natureza. Considerando que $v(A) = \sum_{i \in A} x_i$, essa distribuição “justa” é garantida pelas seguintes propriedades [Shoham e Leyton-Brown, 2009]:

- - Simetria – Se dois jogadores i e j são equivalentes de tal modo que $v(S \cup \{i\}) = v(S \cup \{j\})$ para qualquer coalizão S que não envolva i ou j , o *payoff* recebido por i e j são iguais;
- - Jogador *Dummy* – um jogador é considerado *dummy* se o *payoff* que ele contribui na Coalizão é exatamente igual ao *payoff* que ele conseguiria sozinho, ou seja, $v(S \cup \{i\}) - v(S) = v(\{i\})$. Nesse caso, como o jogador não acrescenta para a coalizão, ele recebe exatamente o que receberia se estivesse sozinho;
- - Aditividade – considere dois jogos cooperativos com distribuição de *payoffs* diferentes v_1 e v_2 . A propriedade da aditividade diz que, se combinarmos esses dois jogos em um só, o *payoff* desse novo jogo deve ser igual à soma dos *payoffs* dos jogos separados $\phi_i(v_1 + v_2) = \phi_i(v_1) + \phi_i(v_2)$.

Uma propriedade adicional, garantida pelas três mencionadas acima, é a eficiência, ou seja, o custo total é dividido entre os participantes $\sum_{i \in N} \phi_i(v) = v(N)$ [Shoham e Leyton-Brown, 2009]. Essa propriedade é importante para o problema proposto, já que garante que os custos totais de estoque serão cobertos pelos operadores.

3.5 ALOCAÇÃO DE CUSTOS PROPORCIONAL AO POTENCIAL DE CONSUMO

Nesta técnica, o custo atribuído a cada operador é calculado como sendo diretamente proporcional ao próprio consumo

na coalizão total. Isto está ilustrado na equação 5:

$$PC_i = \frac{GC \cdot \lambda_i}{\sum_i \lambda_i} \quad (5)$$

Onde:

- - PC_i é o valor da recomendação Proporcional ao potencial de consumo;
- - GC é a quantidade recomendada para a combinação de todos os operadores, (1+2+3+4, também chamada de grande coalizão);
- - λ_i é o consumo das aeronaves do operador i , ou seja, $\frac{n_{i\text{comp}} \cdot \text{HV/ano}}{\text{MTBUR}}$, em que $n_{i\text{comp}}$ é o número de componentes instalados na aeronave, HV/ano é o número de horas voadas em um ano.

3.6 PROPORCIONAL AO NÚMERO DE AERONAVES

Ao calcular o custo relativo de cada operador pode-se assumir que todos os operadores que compartilham a coalizão tenham o mesmo número de horas voadas por ano. Dessa forma, a alocação de custo seria proporcional ao número de aeronaves de cada operador. Segue abaixo:

$$PAC_i = \frac{nAC_i}{\sum_{j=1} nAC_j} \text{Poisson}(\sum_j \lambda_i \cdot nAC_j) \quad (6)$$

Em que:

- PAC_i é o número de peças com as quais o operador i deve arcar;
- nAC_i é o número de aeronaves do operador j .

Note que, nesse método, o operador i arca com um número de peças proporcional à sua frota, mas considerando que a demanda total pode ser calculada considerando seu próprio perfil operacional. Devido a essa simplificação, uma desvantagem desse

método é que a soma dos custos de cada operador não necessariamente é igual aos custos real do estoque. A vantagem desse método é a facilidade de aplicá-lo em campo, já que, para calcular os custos do operador i , apenas o perfil de uso das aeronaves do operador i é necessário;

3.7 OPERADORES INDEPENDENTES

A fim de permitir uma melhor comparação entre os diferentes modelos, será considerada também a situação sem compartilhamento de estoques, no qual cada operador é responsável pelos custos completos de seu próprio estoque. Com o objetivo de manter coerência com os outros métodos utilizados, o dimensionamento também utilizou a equação de Poisson modificada o que permite, portanto, uma quantidade não inteira de itens no estoque. Com isso, não é objetivo sugerir que seja possível tal dimensionamento de peças, ao contrário, esse valor será apenas uma referência útil para avaliarmos vantagens e desvantagens dos métodos de alocação de custos ora testados.

4 RESULTADOS E DISCUSSÕES

A aplicação dos diferentes métodos de alocação de custos, bem como os comentários sobre vantagens e desvantagens de cada modelo, está descrita nas subseções a seguir. Para os testes, foram consideradas 4 situações:

- A situação Base, em que todos os operadores possuem mesmo perfil de horas voadas;
- A situação 2, com operadores com perfis de horas voadas diferentes;
- A situação 3, com operadores com perfis radicalmente diferentes, situação na qual o uso do valor de *Shapley* mostra uma visível vantagem em relação aos outros métodos;
- A situação 4, que apresenta apenas um operador com perfil operacional muito diferente do restante.

4.1 COMPARAÇÃO ENTRE MODELOS – SITUAÇÃO BASE

A situação base considera que todas as aeronaves têm a mesma quantidade de horas voadas por ano. A diferença está na quantidade de aeronaves, conforme descrito na Tabela 1:

Tabela 1: Frota dos operadores

Operador	Aeronaves	HV/Ano
1	1	3000
2	5	3000
3	10	3000
4	30	3000

Todos os resultados obtidos das simulações iniciais com o método de “*Shapley*”, “Operadores independentes”, “Proporcional

ao potencial de consumo” e “Proporcional ao número de aeronaves” são apresentados na Tabela 2.

Tabela 2: Alocação de custos para Situação Base

Operadores	1	2	3	4	Soma
Sem compartilhamento	4	13	23	61	101
Compartilhado (<i>Shapley</i>)	2,42	10,36	19,94	57,28	90
Compartilhado (Proporcional ao consumo)	1,96	9,78	19,57	58,7	90
Compartilhado (Número de aeronaves)	1,96	9,78	19,57	58,7	90

O conceito de alocação de custos baseados em jogos cooperativos pode parecer abstrato até o ponto em que existem exemplos que direcionam o raciocínio para o principal objetivo da cooperação, ou seja, chegar aos valores que sejam satisfatórios a todos os participantes [Moreira et al 2002]. Note que o compartilhamento de estoques (linhas 2, 3 e 4 da Tabela 2) sempre reduz os custos individuais e totais, quando comparados à situação sem compartilhamento. Com o aumento do número de operadores e suas respectivas frotas, aumenta a possibilidade de mais peças

circulando nesse estoque para o uso coletivo, incrementando a probabilidade de todas as aeronaves tenham a confiabilidade de 95 % de disponibilidade atendida.

A diferença é que o uso do valor de *Shapley* fornece economias maiores ao quarto operador, o maior desta coalizão. Isso acontece porque este operador de fato contribui mais para os ganhos conjuntos da coalizão e, portanto, ele recebe mais benefícios que os modelos proporcionais. Tanto o método de compartilhamento intitulado como “Proporcional ao potencial de consumo” e “Proporcional ao número de aeronaves” apresentam nesse caso as mesmas recomendações, já que o número de horas voadas é o mesmo.

4.2 COMPARAÇÃO ENTRE MODELOS – SITUAÇÃO 2

Os perfis adotados para as simulações da situação 2 constam na Tabela 3.
Tabela 3

Tabela 3 : Simulação de perfis para a situação 2

Operador	Aeronaves	HV/Ano
1	3	300
2	10	1000
3	20	1500
4	50	3000

Nesse caso é possível observar que o operador 4 tem uma influência alta nos jogos cooperativos para alocação de custos, pois tanto as suas horas voadas como sua frota

são maiores comparadas aos operadores 1, 2 e 3. A Tabela 4, com os valores dos custos alocados, mostra os resultados da situação 2:

Tabela 4: Alocação de custos para Situação 2

Operadores	1	2	3	4	Soma
Sem compartilhamento	2	10	23	97	132
Compartilhado (<i>Shapley</i>)	0,95	7,15	20,01	93,89	122
Compartilhado (Proporcional ao consumo)	0,58	6,39	19,17	95,86	122
Compartilhado (Número de aeronaves)	0,72	6,87	19,76	93,98	121,33

Nenhum dos operadores nesse conjunto possui o mesmo número de horas voadas ou frota, o que permite perceber a diferença de resultados entre o compartilhamento proporcional ao consumo e o proporcional ao número de aeronaves.

Neste exemplo, o operador 1 possui um impacto baixo na coalizão total e, por isso, também recebe benefícios bastante reduzidos utilizando o valor de *Shapley*, quando comparado aos outros modelos. Tabela 4 é a soma dos custos para o modelo proporcional ao número de aeronaves: o valor não soma 122, que é a quantidade total de peças no estoque compartilhado. Ao assumir que todas as aeronaves da coalizão têm o mesmo perfil operacional, é introduzido um erro que resulta em um valor total diferente do real.

proporcionais de alocação de custo. O proporcional ao consumo é ainda mais baixo que o proporcional à quantidade de aeronave, pois as horas voadas por ano por esse operador são bem menores que as das outras aeronaves na coalizão.

Novamente, o valor de *Shapley* reduz o valor apenas para o grande operador da coalizão, ao custo da elevação dos custos para todos os outros operadores. Outro fato interessante explicitado na

4.3 COMPARAÇÃO ENTRE MODELOS – SITUAÇÃO 3

Os perfis adotados para essa simulação estão na Tabela 5:

Tabela 5: Simulação de perfis para a Situação 3

Operador	Aeronaves	HV/ Ano
1	1	300
2	1	300
3	2	200
4	50	3000

Os operadores 1, 2 e 3 têm poucas diferenças na operação, assim como no λ resultante. Porém o operador 4 tem um aumento muito significativo dessa variável, já que possui a maior frota e maior número de horas voadas ao ano. O uso de perfis tão extremos, nos quais a diferença de horas

voadas e frotas são tão grandes, é simplesmente para ilustrar um fenômeno que pode acontecer em configurações que envolvam grandes/ pequenos operadores em uma coalizão. Na Tabela 6 seguem os resultados da alocação de custos para a Situação 3.

Tabela 6: Alocação de custos – Situação 3

Operadores	1	2	3	4	Soma
Sem compartilhamento	1	1	1	97	100
Compartilhado (<i>Shapley</i>)	0,52	0,52	0,61	96,35	98
Compartilhado (Proporcional ao consumo)	0,19	0,19	0,26	97,35	98
Compartilhado (Número de aeronaves)	0,26	0,26	0,37	96,3	97,19

Interessante notar que o modelo de compartilhamento proporcional ao consumo gerou um custo para o operador 4 ainda maior que o caso sem compartilhamento, o que eliminaria o incentivo de fazer parte desta coalizão. Isto acontece possivelmente por um efeito parecido ao demonstrado na Figura 3, pelo fato da técnica “proporcional ao consumo” utilizar o Poisson tradicional (não modificado) para cálculo dos estoques. Imagine a seguinte situação: uma coalizão A, formada por 2 aeronaves com mesmo perfil operacional, utiliza uma única peça para garantir 95,1% de disponibilidade da mesma. Nesta configuração, cada aeronave é responsável por 0,5 peça. Se a esta coalizão se juntar outra aeronave, pode haver a

necessidade de comprar mais uma peça para garantir uma disponibilidade maior que 95%. Nesse caso, a nova coalizão, formada por 3 aeronaves, precisa de 2 peças (ou seja, cada aeronave paga, em média, 0,67 peça). Ou seja, a adição de aeronaves parece ter feito a coalizão menos atrativa, aumentando os custos para cada operador.

4.4 COMPARAÇÃO ENTRE MODELOS – SITUAÇÃO 4

No caso ilustrado na Tabela , existe um operador que possui a maior frota do conjunto, porém possui um perfil de horas voadas bem inferior.

Tabela 7: Simulação de perfis para a Situação 4

Operador	Aeronaves	HV/ Ano
1	1	3000
2	2	3000
3	2	2000
4	50	450

Mesmo que o operador 4 seja responsável pela maior parte dos custos, é possível perceber uma maior participação de todos os

outros operadores comparando com a Situação 3, como será visto na Tabela 8:

Tabela 8: Alocação de custos para Situação 4

Operadores	1	2	3	4	Soma
Sem compartilhamento	4	7	7	18	36
Compartilhado (<i>Shapley</i>)	2,54	4,7	4,7	16,06	28
Compartilhado (Proporcional ao consumo)	2,24	4,48	4,48	16,8	28
Compartilhado (Número de aeronaves)	1,93	3,85	3,85	18,18	27,82

Note que, nesta situação, o modelo proporcional ao número de aeronaves resultou em um custo maior que a situação sem compartilhamento para o operador 4 (similar ao que aconteceu para o modelo proporcional ao consumo na situação 3). Esse fenômeno aconteceu pelos mesmos motivos descritos na seção anteriormente: a restrição de números inteiros da equação de Poisson, ilustrado na Figura 3. De maneira geral, é interessante notar os ganhos marginais de operadores maiores em coalizões de compartilhamento de estoques, principalmente ao se utilizar os modelos proporcionais (ao consumo e ao número de aeronaves).

4.5 COMENTÁRIOS GERAIS

Todos os cálculos e simulações realizadas assumem a distribuição direta dos custos do estoque entre os operadores, sem nenhuma margem aplicada sobre os valores. Nessas condições, pôde-se propor a distribuição de custos de modo a garantir o interesse em todos os jogadores na grande coalizão.

No caso de o compartilhamento ser oferecido por um agente externo no formato de um serviço, as margens de lucro aplicadas podem alterar os incentivos dos operadores em participar da grande coalizão: casos como o da simulação 3 possivelmente mostrariam que o operador com a maior

quantidade de aeronaves não teria incentivo em participar da mesma. Note que, na simulação apresentada, o incentivo de participação já é marginal. Se fossem adicionadas taxas e/ou margens de lucro, possivelmente o operador não teria mais incentivos em participar da grande coalizão.

Ainda sim, o uso do valor de *Shapley* para alocação de custos tende a aumentar o interesse dos grandes operadores nesse tipo de serviço, quando comparado às outras formas descritas neste trabalho. Para ilustrar os incentivos que cada operador possui para entrar na coalizão, diversas simulações foram realizadas, variando o perfil operacional do operador 4 da seguinte forma: partindo de perfis operacionais iguais (1 aeronave, voando 1000 horas por ano), cada simulação acrescentou uma aeronave a mais para o operador 4, de modo que este passe a ter maior porcentagem do consumo total de peças da coalizão. Com isso, o objetivo é avaliar os ganhos de operadores pequenos e grandes ao entrar na coalizão. A Figura 5 a seguir mostra como o ganho do operador 4 ao se juntar à coalizão evolui, à medida em que se aumenta sua frota de aeronaves (os pontos "Modelos Proporcionais" representam os modelos "Proporcional ao potencial de consumo" e "Proporcional à Quantidade de Aeronaves", já que as horas voadas por ano são iguais para todos os operadores).

Figura 1: Ganho do operador 4 com o aumento da porcentagem de consumo da coalizão

No gráfico, o eixo y é a redução de custos que o operador 4 percebe, em relação à situação sem compartilhamento. O eixo x representa qual porcentagem do consumo total da coalizão é requerido pelo operador 4.

Com o aumento do número de aeronaves do operador 4, o mesmo se torna cada vez mais responsável pela maior parte dos custos envolvidos. No limite, se o consumo de um único operador fosse aumentado até que o consumo de todos os outros possa ser considerado desprezível, esse operador atua como se estivesse sozinho. Por isso, à medida que se aumenta a quantidade de

aeronaves, os ganhos de se participar da coalizão diminuem. Note que o uso do valor de *Shapley* para o compartilhamento dos custos faz com que o operador maior tenha uma redução menor de ganhos (ou seja, maior incentivo de participação na coalizão) à medida que a porcentagem de consumo total da coalizão aumenta.

O gráfico análogo para o operador 2 da simulação, que mantém o tamanho da frota constante enquanto o operador 4 aumenta o número de aeronaves, pode ser visualizado na Figura 4:

Figura 4: Ganho do operador 2 com o aumento da porcentagem de consumo da coalizão.

Figura , o eixo x também representa a porcentagem de consumo da coalizão que o operador 4 representa. Note que, no caso do operador 2, representando um operador pequeno em uma grande coalizão, o ganho sempre aumenta com o aumento do consumo

Figura) é repassado para todos os operadores, independentemente de quem contribuiu mais para esses ganhos globais. Ainda sim, utilizando o valor de *Shapley*, os pequenos operadores têm grande benefício ao se juntarem à coalizão (por volta de 40%).

Pode-se concluir que pequenos operadores sempre terão grandes incentivos para participar de compartilhamentos de peças, enquanto grandes operadores tendem ter

de outros operadores. Nos modelos proporcionais, esse aumento é muito maior que no modelo com o valor de *Shapley*, uma vez que todo benefício de redução de estoques por aeronaves com o aumento do consumo da coalizão (ilustrado na

cada vez menos incentivo, à medida que se aumenta a participação do mesmo na coalizão. Uma vantagem do uso do valor de *Shapley* é distribuir os custos de forma a garantir que os grandes operadores mantenham um maior interesse no compartilhamento de estoques, através da redução das suas taxas pagas no *pool*. Ainda sim, o incentivo de participação desses pequenos operadores é bem maior que o dos grandes operadores, o que possivelmente indica que o uso de *Shapley* ao invés de

métodos de compartilhamento de custos proporcionais não tiraria o interesse destes operadores de participar do *pool* de peças.

5. CONCLUSÕES

Diferentes técnicas de alocação de custos para compartilhamento de estoques foram testadas, com foco especial ao método de *Shapley*, baseado na teoria dos jogos cooperativos. Esta metodologia se mostrou importante para atrair grandes operadores aos programas de compartilhamento de estoques, uma vez que a alocação de custos é realizada de uma maneira mais justa e proporcional à contribuição de cada operador à coalizão. Os custos dos operadores menores aumentam um pouco em relação aos modelos proporcionais, mas ainda sim o incentivo destes em participarem de tais

coalizões é muito grande (com redução de custos da ordem de 40%). De fato, a prática mostra que grandes operadores têm muito menos incentivos a participarem de programas de compartilhamentos de estoque. O uso de teoria de jogos cooperativos permite harmonizar os incentivos a essa participação.

O trabalho focou em mostrar uma maneira justa de compartilhar os estoques, dada a coalizão. Como trabalhos futuros, sugere-se estudar sob quais condições a grande coalizão (ou seja, com todos os operadores) é formada, o que acontece quando a alocação de custos está no *core* do jogo cooperativo [Karsten 2010]. Para aplicações práticas do estudo ora desenvolvido, faz-se necessário também verificar a influência da adição de custos adicionais e margens de lucro, o que aumentaria o custo total da coalizão.

REFERÊNCIAS

- [1] Berger, J. M. (2014). MRO Market Overview & Trends. Annual Transportation Review Board Meeting, Washington DC.
- [2] Karsten, F.J.P., Slikker M., van Houtum, G.J. (2010). Spare parts inventory pooling games. International Workshop on Supply Chain Models for Shared Resource Management, Brussels.
- [3] Miguel, M. I. R. (2005). Ensino e aprendizagem do modelo de Poisson: Uma experiência com modelagem. Doutorado em Educação matemática apresentada a Pontifícia Universidade Católica de São Paulo, São Paulo.
- [4] Moreira, R. C.. (2002). Teoria dos Jogos Cooperativos: Conceitos Fundamentais. In: Anais do XXXIV SBPO, Rio de Janeiro, SOBRAPO.
- [5] Moreira, R. C., Luna, H. P. L., Guedes, P. G. S. (2002). Um estudo comparativo entre a teoria dos jogos cooperativos e uma heurística aplicada a um problema real de alocação de custos. *Pesquisa operacional*, vol. 22, nº1.
- [6] Shoham, Y., Leyton-Brown, K. (2009). Multiagent Systems: Algorithmic, Game-Theoretic and Logical Foundations. Cambridge University Press. ISBN-13: 9780521899437
- [7] Weisstein, E. W. (2015). Gama Function, MathWorld--A Wolfram Web Resource. <http://mathworld.wolfram.com/GamaFunction.html>. Acessado: 20-de dezembro de 2015.

Capítulo 10

MATHEURÍSTICA BASEADA EM PSO PARA O PROBLEMA DE CUSTO DE DISPONIBILIDADE DE RECURSOS COM MÚLTIPLOS MODOS

Lettiery D'Lamare Portela Procópio

Lucídio dos Anjos Formiga Cabral

Marco Cesar Goldbarg

Resumo: Este artigo descreve a construção de uma Matheurística baseada na Otimização por Enxame de Partículas (PSO) afim de solucionar o Problema de Custo de Disponibilidade de Recursos com Múltiplos Modos. Inspirado na necessidade de balancear a utilização de recursos renováveis com o tempo total (makespan), através do escalonamento das atividades com seus diversos modos de execução presente no projeto. Testes realizados em instâncias da literatura comprovam a eficácia da utilização da programação matemática adaptada com a Metaheurística PSO.

Palavras Chave. Disponibilidade de Recursos, PSO, Matheurística.

1. INTRODUÇÃO

O processo de tomada de decisão é uma etapa crítica para diversos tipos de projetos reais e já é pressuposta por alguns autores como uma ciência própria, a Ciência da Gestão (*Management Science*) Taylor [2004]; Calazans [2012]. Certamente, o ato de refletir e planejar os passos de um determinado projeto impacta fortemente na sua execução, destinando-o para o sucesso ou fracasso em casos de omissão desta etapa. Métodos científicos apresentam-se como ferramentas para avaliar as possíveis alternativas.

Por isso, o planejamento de projetos necessita da coordenação de atividades relacionadas para atingir objetivos pré-definidos. Entre os objetivos mais comuns estão o tempo de conclusão e o custo total do projeto. Um recente exemplo de aplicação pode ser encontrado no trabalho de [Yamashita e Morabito, 2012], o qual fala sobre o Problema de Custos de Disponibilidade de Recursos com Múltiplos Modos (PCDRMM), que tem o objetivo de otimizar o custo da alocação de recursos nas atividades considerando o tempo final do projeto.

Problemas como o PCDRMM permitem ganhos concretos nos negócios empresariais através da otimização de processos. Suas soluções são medidas por duas grandezas distintas: a quantidade de recursos que serão utilizados, e o tempo final do projeto. Métodos exatos se apresentam como uma abordagem naturalmente desejada, onde soluções ótimas forneceriam cenários atrativos para a indústria, mas a necessidade de minimização do tempo total do projeto, e a alocação das atividades com diversos modos distintos, caracterizam um problema de otimização combinatória pertencente a classe *NP-Hard* (como desmostrado por Möhring [1984]; Blazewicz et al. [1983]). Essas propriedades nos remetem a escolha de métodos heurísticos ou heurísticos com hibridização exata, para permitir menor tempo de processamento computacional e resultados de boa qualidade.

A aplicação desse problema pode ser encontrada em vários cenários reais como na construção civil onde máquinas, matéria prima e equipes de trabalhadores são recursos escassos e o tempo final do projeto necessita ser otimizado. Assim como o compartilhamento de recursos na computação em nuvem, a produção industrial

e até projetos de desenvolvimento de softwares podem ser solucionados utilizando o PCDRMM.

Assim, este trabalho é motivado pela busca de soluções de boa qualidade para o Problema de Custo de Disponibilidade de Recursos com Múltiplos Modos através de exemplos da literatura que obtiveram sucesso com heurísticas para resolver problemas onde a utilização de recursos é cara para o projeto como um todo ou que ainda existe uma escassez dos recursos utilizados.

2. TRABALHOS RELACIONADOS

Os trabalhos da literatura que abordam o PCDRMM apresentam vários métodos de solução para o problema. É atribuída a um decisor (gestor do projeto) a responsabilidade da escolha que respeite a data final de entrega do projeto. Uma avaliação entre o custo e o tempo é feita para escolher entre perder em tempo e ganhar nos custos do projeto, ou adiantar o prazo e aumentar custo.

O primeiro trabalho abordando o tema foi apresentado por Hsu e Kim [2005] e trata de uma adaptação do PCDR acrescentando a característica dos múltiplos modos às atividades. O modelo foi denominado *Multi-Mode Resource Investment Problem* (MMRIP). Foi proposta uma regra de decisão que considera simultaneamente restrições de data de finalização do projeto (*deadline*) e a utilização de recursos para eleger em qual modo atividade será executada. Os modos das atividades foram selecionados de forma a possibilitar a redução do tempo final e o custo total.

Segundo Chen et al. [2012], o trabalho de Hsu e Kim [2005] foi o primeiro a idealizar o PCDRMM. Um método heurístico de enumeração baseada em árvore de precedência que seleciona a ordem em que as atividades são executadas foi proposto por Chen et al. [2012]. A heurística realiza três passos para alcançar os objetivos: seleção das atividades; atribuição dos modos; e decisão dos tempos de início. A aplicação é o aluguel de recursos computacionais na nuvem para processamento, uma vez que essa alocação é paga apenas uma vez, caracterizando os recursos renováveis.

O PCDRMM foi assim chamado pela primeira vez por Yamashita e Morabito [2012]. Neste trabalho foi proposto um algoritmo exato que utiliza *Branch-and-Bound*. Seus resultados

foram comparados à solução da formulação matemática por um *solver* comercial. A comparação mostrou que o algoritmo proposto por Yamashita e Morabito [2012] exibiu desempenho melhor que o *solver* comercial, produzindo o resultado em tempo computacional inferior. Os testes foram realizados em um conjunto de 12 instâncias geradas pelo software Progen. Até o momento, este foi o único trabalho da literatura que propõe um método exato para solucionar o PCDRMM.

Um esquema de geração de agendamento de atividades denominado FAL (*forward activity list*) foi apresentado por Qi et al. [2015]. O esquema foi usado com heurísticas para reparar soluções inviáveis. Foi proposta uma heurística de otimização por enxame de partículas combinada com *scatter search* a qual foi comparada com um algoritmo genético implementado pelos autores em instâncias com até 30 atividades.

Por fim, uma adaptação do PCDRMM é encontrada em Nadjafi [2014] com o nome de *Multi-mode Resource Availability Cost Problem with Recruitment and Release Dates for Resources* (MRACP-RR), onde o custo de disponibilização dos recursos é contabilizado apenas no período entre a primeira e a última utilização do recurso, análogo ao custo por hora de utilização. Foi proposta uma heurística *Simulated Annealing* para a construção das soluções tendo o tempo de processamento como critério de parada. Como não existia registro anterior do problema adaptado, foram realizados testes computacionais e comparados com o *solver* Lingo 11. A heurística se mostrou eficaz para instâncias com até 90 atividades.

3. DESCRIÇÃO DO PROBLEMA

O PCDRMM pode ser definido como um problema que balanceia a minimização da utilização dos recursos renováveis e o tempo de duração do projeto, contudo, suas atividades podem assumir diversos modos de execução que variam o tempo e a quantidade de recursos utilizados. Elas ainda possuem uma relação de precedência temporal que determina o tempo mínimo de início.

O custo de utilização dos recursos é contabilizado por todo o horizonte do projeto, e os recursos podem ser reutilizados por outras atividades sem adicionar quaisquer gastos à função objetivo. Este problema equilibra a utilização dos recursos de forma a não penalizar o tempo final do projeto, por isso sua aplicabilidade a projetos reais se apresenta de forma imediata ou de fácil adaptação.

Ressalte-se que a principal tarefa desse problema é organizar as atividades a fim de minimizar o tempo final do projeto conhecido como *makespan*, e o custo atribuído à disponibilização dos recursos para as atividades. Uma das primeiras características que torna o problema mais aplicável a projetos reais é a possibilidade de uma atividade ser realizada de várias maneiras distintas.

Um exemplo de dados de entrada para o problema pode ser observado na Figura 1 e na Tabela 1. Para mapear a precedência das atividades descritas na Figura 1, temos um conjunto $H = \{(s, 1), (s, 2), (s, 4), (1, 3), (2, 3), (3, t), (4, t)\}$, no qual a existência do par ordenado $(1, 3)$ implica que a atividade 1 precede a atividade 3. s e t são atividades virtuais inseridas no problema para demarcar o início e fim respectivamente do projeto.

Figura 1 Exemplo de Precedências.

São demonstradas na Tabela 1 as variáveis d_{jj} demarcam a duração da atividade j no

modo i , e r_{ijk} quantifica a necessidade da atividade j no modo i pelo recurso do tipo k .

Tabela 1 Exemplo de dados de Entrada

Atividade j	Modo i	Duração d_{ji}	Recurso 1 r_{ji1}	Recurso 2 r_{ji2}	Recurso 3 r_{ji3}
s	$M_s=1$	1	0	0	0
1	$M_1=1$	1	3	2	1
2	$M_2=1$	1	4	1	2
3	$M_3=2$	1	2	2	2
		2	5	1	0
4	$M_4=1$	1	2	0	2
t	$M_t=1$	1	0	0	0

A Tabela 1 exemplifica os valores referentes às atividades $j=\{s, 1, \dots, t\}$, que executam nos possíveis modos i , e duram d_{ji} unidades de tempo. Também requerem $r_{ji1}, \dots, 3$ quantidades de recursos.

A propriedade de *múltiplos modos* torna-se visível na atividade 3 ($j = 3$), a qual pode ser executada em dois modos diferentes. O primeiro modo ($i = 1$) gastará 2 unidades de tempo, 2 recursos do tipo 1, 2 recursos do tipo 2, e 3 recursos do tipo 3. No seu segundo modo ($i = 2$) necessitará de 5 unidades de tempo, 1 recurso do tipo 1, e 1 recurso do tipo 3.

A Tabela 1 também demonstra a não influência das atividades virtuais s e t para o projeto, pois não necessitam de unidade de tempo nem quantidade de recurso.

Para obter uma solução viável para esse problema é necessário definir o tempo de início de cada atividade. Um exemplo de solução para o exemplo é ilustrado na Figura 2. Nesta solução, as atividades 1 e 4 iniciam no tempo 0, a atividade 2 no tempo 3 e a atividade 3 no tempo 7. As quatro atividades são representadas por retângulos numerados. As disponibilidades dos recursos tipo 1, 2 e 3 são ilustradas pelas barras verdes, azuis e vermelhas, respectivamente.

Figura 2 Exemplo de solução.

O eixo horizontal do gráfico mostra o tempo e o eixo vertical a quantidade do recurso. A linha pontilhada vermelha mostra o tempo limite para a entrega do projeto, $D = 10$, e a linha pontilhada azul mostra o *makespan* da solução, neste caso igual a 9.

O custo final de utilização dos recursos é dado pela soma das maiores quantidades necessárias de cada recurso vezes o custo unitário c_k . Os recursos tipo 1 e 2 (verde e azul) assumem os maiores valores no primeiro período de utilização com a solicitação das atividades 1 e 4. O recurso 3 assumiu seu

maior valor para atender a atividade 3 executada no modo 1 ($i=1$). O custo total do projeto é calculado como: $2 \cdot a_1 + 4 \cdot a_2 + 3 \cdot a_3$, onde os valores que multiplicam a_k são os custos dos respectivos tipos de recursos, $2 \cdot 2 + 4 \cdot 4 + 3 \cdot 3 = 29$.

Embora o tempo de término do projeto seja considerado como uma restrição, em projetos reais o tempo de finalização é um fator importante na escolha do processo. Assim, alterações que reduzam o tempo, sem elevar o custo, são bem-vindas. Neste caso, a solução de mesmo custo com um prazo

menor para a entrega do projeto é melhor que a anterior, e pode-se dizer que é dominante. Um exemplo é mostrado na Figura 3. Nela, as atividades 1 e 2 começam no tempo 0 e as

atividades 3 e 4 no tempo 4. O *makespan* é 6 e o custo permanece o mesmo da solução mostrada na Figura 2.

Figura 3 Exemplo de solução com redução de Makespan.

3.1 FORMULAÇÃO MATEMÁTICA

Uma formulação matemática para o PCDRMM foi apresentada por Yamashida e Murabito [2012]. O problema é considerado como um projeto com n atividades, das quais a primeira e a última atividade (1 e n) são equivalentes às atividades virtuais s e t . O tempo do projeto

é discretizado em instantes $t = 0, 1, \dots, D$, onde D é a data limite para finalizar o projeto. O modelo de Yamashida e Murabito [2012], baseado em Talbot [1982], possui os parâmetros e variáveis listados a seguir.

$$-x_{jit} = \begin{cases} 1, & \text{se a atividade } j \text{ for executada no modo } i \text{ e termina no tempo } t. \\ 0, & \text{caso contrário.} \end{cases}$$

$$a_k = \{\text{quantidade utilizada do recurso } k.$$

Os dados de entrada do problema são:

- - c_k = custo atribuído a disponibilização de um recurso do tipo k .
- - D = data de entrega do projeto.
- - H = conjunto de relações de precedência.
- - d_{ji} = duração da atividade j quando executada no modo i .
- - r_{jik} = quantidade de recursos do tipo k que a atividade j utiliza ao ser executada no modo i .
- Objetivo:

- - LF_j = instante de término mais tarde que a atividade j pode ser finalizada, considerando as relações de precedência entre as atividades. Esse valor é obtido pela alocação regressiva das atividades, iniciando da atividade t alocada no instante D .
- - EF_j = instante de término mais cedo que a atividade j pode ser completada, também considera as relações de precedência entre as atividades. O EF_j é obtido pela alocação progressiva das atividades, tendo início na atividades.

$$\text{Minimizar } \sum_{k=1}^m c_k a_k \quad (1)$$

Sujeito a:

$$\sum_{i=1}^{M_j} \sum_{t=EF_j}^{LF_j} x_{jit} = 1 \quad j = 1, \dots, n \quad (2)$$

$$\sum_{i=1}^{M_h} \sum_{t=EF_h}^{LF_h} t \cdot x_{hit} \leq \sum_{i=1}^{M_j} \sum_{t=EF_j}^{LF_j} (t - d_{ji}) \cdot x_{jit} \quad j = 1, \dots, n, \forall h | (h, j) \in H \quad (3)$$

$$\sum_{j=1}^n \sum_{i=1}^{M_j} \sum_{b=1}^{t+d_{ij}-1} r_{jik} \cdot x_{jib} \leq a_k \quad k = 1, \dots, m, t = 0, \dots, D \quad (4)$$

$$x_{jit} \in \{0, 1\} \quad j = 1, \dots, n, t = 0, \dots, D, i = 1, \dots, M_j \quad (5)$$

$$a_k \geq 0, a_k \in \mathbb{Z} \quad k = 1, \dots, m \quad (6)$$

A função objetivo (1) minimiza o custo de disponibilização dos recursos. A restrição (2) garante que toda atividade só será executada exclusivamente uma vez em um único modo dentro da sua janela de tempo (LF_j - EF_j). A precedência das atividades é respeitada pela restrição (3), onde se existe o par (h, j) no conjunto de precedência H, então o tempo de término da atividade predecessora h, ($t \cdot r_{jik}$) não pode ser maior do que o tempo de início de sua atividade sucessora j dado por $(t - d_{ji}) \cdot x_{jit}$. A restrição (4) garante que durante o tempo de execução das atividades ($\sum_{b=1}^{t+d_{ij}-1}$), a_k assume o maior valor de utilização do recurso do tipo k. Esta restrição garante a propriedade de recursos renováveis. As restrições (5) e (6) definem a variável de decisão binária x_{jit} e variável a_k inteira e positiva.

3.2 MATHEURÍSTICA

As matheurísticas são heurísticas híbridas que mesclam meta-heurísticas e técnicas de

programação matemática. Esta hibridização pode ser realizada segundo diversas abordagens. Uma revisão de tais abordagens foi apresentada por Ball[2011].

Motivado pelos promissores resultados em [Qi et al., 2015], a Meta-heurística *Particle Swarm Optimization* (PSO) apresentada por Kennedy e Eberhart [1995] com base no comportamento de pássaros e outros animais que andam em grupo, foi possível mapear um padrão para a movimentação desses indivíduos em busca do objetivo do bando.

No PSO cada indivíduo (representado por uma solução) de uma população (enxame), quantifica sua experiência local (vizinhos locais), e ideais do enxame (melhores globais). Assim, dar-se o aprendizado individual de cada partícula de um enxame. O fluxograma da Figura 3 resume a adaptação dos passos para o processo evolucionário do algoritmo proposto.

Figura 4 Fluxograma do PSO Proposto.

A criação do enxame inicial influencia diretamente no número de interações

necessárias para alcançar o objetivo do algoritmo e na diversificação das regiões do

espaço de busca. Obter enxames iniciais que permitam uma boa eficiência na busca e implicitamente representem regiões de busca diversificadas são características almejadas por três algoritmos de criação de soluções iniciais:

- **Melhor makespan:** são selecionados os modos das atividades com a menor duração do tempo de execução, desconsiderando a utilização dos recursos. A escolha do tempo de início das atividades é exatamente o instante de término da sua última predecessora $EF_j - \min(d_{ji})$.
- **Modos aleatório:** onde os modos das atividades são selecionados de maneira aleatória e são alocadas de acordo com sua ordem de prioridade (precedência).
- **Melhor modo por custo:** que tem como a escolha do modo m_{ij} da atividade j é realizada verificando o custo de utilização dos recursos do modo i , isso possibilita uma tentativa de redução do custo do projeto.

Uma partícula do PSO é um escalonamento que respeita todas as restrições da formulação matemática apresentada na seção 3.1, o qual pode ser avaliado tanto pelo seu tempo *makespan*, quanto pelo seu custo de utilização dos recursos $c_k a_k$. A minimização desses valores são as metas a serem alcançadas, porém, a função objetivo da modelagem matemática do problema, tenta minimizar apenas o custo, deixando o tempo como uma restrição complementar do problema.

Para melhor quantificar a localização das partículas no algoritmo, foi inserido o valor do *makespan* na função de avaliação da solução $f(s) = p_t * t + p_c * c$, onde p_t é a proporção do tempo final da solução t , e p_c é a proporção do custo por disponibilidade c . A utilização de $p_t = p_c$ balanceia a necessidade do custo e o tempo da solução, essas

constantes servem também de parâmetros de ajuste do algoritmo de acordo com a necessidade do decisor. Para melhorar a avaliação dos indivíduos os valores de t e c são normalizados, uma vez que são grandezas distintas. A função de avaliação ainda atribui uma penalidade à solução que não cumprir o requisito da data final do projeto D .

A melhor experiência conhecida de uma partícula é seu *pbest*, e o melhor escalonamento conhecido de todo o enxame (melhor global) é valor de referencia do *gbest*. Assim a fase de avaliação do *gbest* e *pbest*, quantifica os escalonamentos para verificar se foi gerado um novo *gbest* ou *pbest* na iteração atual.

A representação da solução do PCDRMM no PSO é baseada em [Qi et al., 2015], que mapeia uma partícula em dois vetores: $t_j = (t_1, t_2, \dots, t_n)$, demarca o tempo de início t_j de cada atividade j , e $m_j = (m_1, m_2, \dots, m_n)$, enumera o modo m_j de todas as n atividades. Essas duas informações necessárias para representar um escalonamento são concatenados em apenas um vetor $x = t_j \cup m_j$, com $x_i = (q_1, q_2, \dots, q_n, q_{n+1}, q_{n+2}, \dots, q_{2n})$. Os primeiros n números $(q_1, q_2, \dots, q_n)$ representam os tempos de início das atividades e, logo em seguida $(q_{n+1}, q_{n+2}, \dots, q_{2n})$ os modos são apresentados. Como uma solução do PCDRMM não pode ser representada no domínio contínuo e o PSO gera soluções com essas características, adotamos o maior número inteiro $\lfloor x \rfloor$ para as variáveis contínuas.

As posições das partículas e velocidades calculadas com base na própria experiência e pela observação da sua vizinhança. Elas se movem no espaço de busca de acordo com as equações a seguir:

$$v_{i(e+1)} = w \cdot v_{ie} + c_1 \cdot r_1 (pbest_{ie} - x_{ie}) + c_2 \cdot r_2 (gbest_e - x_{ie}) \quad (7)$$

$$x_{i(e+1)} = x_{ie} + v_{i(e+1)} \quad (8)$$

$$e = 1, \dots, E \quad (9)$$

$$i = 1, \dots, m \quad (10)$$

A equação (7) representa a atualização da velocidade para a nova época $(e + 1)$, onde

temos v_{ie} como a velocidade da partícula i na época e , w representa a inércia da direção

que estava sendo seguida, r_1 e $r_2 \in [0, 1]$ são constantes aleatórias, c_1 e c_2 são os coeficientes que representam a aceleração cognitiva que levam os indivíduos em direção ao $pbest$ e ao $gbest$ respectivamente. A equação (8) atualizar a posição do indivíduo i

na época x_{ie} , essas operações são realizadas por todas as m partículas do enxame A . O pseudo-código descrito a seguir esclarece os passos do Algoritmo de Enxame de Partículas proposto para a solução do problema:

Algoritmo 1 Pseudo-código do PSO

```

Entrada: A – conjunto de partículas
        pbest – melhores locais
        gbest – melhor global
        E – limite de épocas
1:  A ← iniciarEnxameDeParticulas()
2:  para todo e = 1 até E faça
3:      atualizar(pbest, gbest, A)
4:      para todo a ∈ A faça
5:          atualizarVelocidades(a)
6:          a ← MoverParticula(a, e)
7:      fim
8:      B ← Seleção(A)
9:      BuscaLocalExata(B)
10:     atualizar(pbest, gbest, B)
11: fim
12: CriaçãoTradeOff(pbest)

```

No algoritmo existe uma etapa de busca local exata no processo, onde são selecionadas 10% das soluções da população para serem alteradas, sendo que 5% são constituídos pelos melhores indivíduos conhecidos, e os outros 5% são eleitos de forma aleatória, esse processo apenas modifica a solução se for possível a minimização nos valores de tempo ou custo.

Essa busca local é feita utilizando o modelo de programação matemática de Yamashita e Morabito [2012] adaptado como descrito nas equações (11), (12) e (13) e implementado pelo *framework* CPLEX da IBM. São aplicadas algumas alterações no modelo original para possibilitar a otimização de apenas parte da solução em um tempo computacional aceitável.

A adaptação feita nessa etapa adiciona ao modelo original do PCDRMM o sub-conjunto de atividades fixas F que não serão alteradas pelo modelo, o qual é formado pelas atividades de um escalonamento viável $F = \{ \forall j \in s' \mid t'_j - d'_{ji} \in a_{kt} \}$, onde t'_j é o tempo em que a atividade j foi escalonada, d'_{ji} é a duração da atividade j no modo i , e a_{kt} é o

conjunto de intervalos de tempo em que a demanda a_k do recurso do tipo k foi alcançada.

O conjunto F é formado pelas atividades j que não alocaram novos recursos para sua conclusão, assim as atividades que serão alteradas pelo novo modelo, são consideradas críticas para a solução em relação ao custo. O conjunto F deve conter mais de 60% no número total de atividade, caso isso não ocorra, novas atividades são escolhidas de forma aleatória para que a solução criada s'' (solução gerada após a matheurística) não seja descaracterizada da solução original s' . Todas as outras variáveis descritas na formulação matemática, também são utilizadas nessa adaptação e os novos conjuntos criados são:

- $s' =$ escalonamento viável com todas as atividades j alocadas no tempo t no modo i .
- $a_{kt} =$ conjunto de intervalos de tempo t que foram utilizados toda a demanda do recurso do tipo k .
- $F =$ conjunto de atividades de uma solução s' que satisfazem a condição

$$t'_j - t'_{ji} \notin a_{kt}.$$

Objetivo:

$$\text{Minimizar } \sum_{k=1}^m c_k a_k \quad (11)$$

Sujeito a:

$$\sum_{i=1}^{M_j} \sum_{t \in EF_j}^{LF_j} x_{jit} = 1 \quad j = 1, \dots, n, \forall j \notin F \quad (12)$$

$$x_{jit} = 1 \quad \forall j \in F, \forall j, i, t \in s' \quad (13)$$

$$\text{Restrições (3), (4), (5) e (6)} \quad (14)$$

A descrição acima, modifica a restrição (2) do modelo original removendo as atividades pertencentes ao conjunto F , e adiciona a restrição (13) que fixa o tempo t o modo i da atividade $j \in F$ na solução s' . Todas as outras restrições (14) permaneceram idênticas ao modelo original.

Ao final da busca local exata, as soluções escolhidas para serem modificadas pelo novo modelo, são reavaliadas em busca de novos valores para seus $pbest$. Após essa etapa uma verificação de quantas épocas já ocorreram é efetivada, e caso o número de gerações não tenha alcançado 50, o algoritmo retorna ao processo de busca do novos $pbest$ e $gbest$, caso contrário o algoritmo entra na fase final para a criação da curva de *trade-off*, baseada na heurística da fronteira de Pareto, onde uma solução s^1 é dita dominante da solução s^2 , se ambas as condições a seguir forem satisfeita:

- A solução s^1 não é pior do que a solução s^2 em nenhuma das grandezas (custo e tempo), ou seja, $custo_{s^1} \leq custo_{s^2}$ e $tempo_{s^1} \leq tempo_{s^2}$.

- A solução s^1 é melhor do que a solução s^2 em no mínimo uma grandeza, ou seja, $custo_{s^1} < custo_{s^2}$ ou $tempo_{s^1} < tempo_{s^2}$.

O conjunto final de soluções serão formados por indivíduos as quais suas dominantes não foram conhecidas pelo algoritmo. Esse processo entrega o conjunto de soluções que formarão o *trade-off* para o decisor, que deve eleger entre uma solução balanceada que atenda suas expectativas de tempo e custo do projeto.

4. RESULTADOS

O modelo e sua adaptação, foram implementados no software *ILOG CPLEX*. Todo código deste trabalho foi desenvolvido na linguagem de programação C++, e os testes executados em um computador com processador Intel Core i5 2,3 GHz, com 4 GB de memória Ram. Os parâmetros de configurações do PSO com a Matheurística (PSOM) apresentados na Tabela 2, foram obtidos com 50 iterações em instâncias excluídas dos resultados finais através da ferramenta IRACE Ibáñez et al. [2011].

Tabela 2: Parâmetros da Matheurística

Coeficientes	Valores
E (Quantidade de épocas)	60
A (quantidade de partículas)	50
w (inércia da velocidade)	0,35
c_1 (coeficiente da experiência local)	0,30
c_2 (coeficiente da experiência global)	0,45

Nas tabelas abaixo os custos das soluções são apresentados pelas colunas com o "(C)", e o tempo de execução que é dado em segundos demarcadas com "(T)". Os índices das instâncias estão presentes na coluna de "*Inst*" cada tabela. Os valores das soluções da Matheurística são identificados pelas colunas "*PSOM*", e comparados com os resultados do "*solver*". Para obter os valores das colunas do *PSOM* foram necessárias 50 execuções da Matheurística.

Um outro conjunto de 12 instâncias criadas no software Progen para o PCDRMM foi disponibilizado por Yamashita e Morabito [2012]. Nelas, todas as 12 atividades possuem: 3 modos de execução distintos, 4 tipos de recursos disponíveis. Porém o tempo e a quantidade de recursos da primeira e da última atividade são iguais a zero, o que caracteriza as atividades virtuais do problema.

Tabela 3: Resultados do solver e PSOM com instâncias do PCDRMM

Inst	solver (C)	PSOM (C)	solver (T)	PSOM (T)
1	38,76	44,78	1,03	2,46
2	117,10	117,10	2,86	2,01
3	277,29	292,56	39,44	2,95
4	376,10	378,31	22,28	5,83
5	93,49	101,75	4,49	3,87
6	121,52	132,36	1,57	2,74
7	105,52	111,36	7,08	2,38
8	148,75	156,67	79,42	4,71
9	21,31	21,31	1,73	2,94
10	172,07	194,34	4,64	3,82
11	104,31	104,31	90,19	3,93
12	159,84	166,22	102,67	5,54

Uma diferença de 4,90% separa os resultados obtidos do *PSOM* das soluções ótimas. Contudo, o *PSOM* consegue alcançar em 3 das 12 instâncias o valor obtido pelo *solver*. Esse cenário se inverte quando quando observado o tempo de processamento computacional do *solver* retardando em média de 73,05% dos valores obtidos pelo *PSOM*.

Para os resultados da tabela abaixo, foi necessária uma adaptação para utilizar os dados de entrada disponível na biblioteca PSPLIB. Nestas instâncias existe a inclusão dos recursos não renováveis, porém para no PCDRMM é considerar todos os recursos como renováveis. Esses dados, são diferenciados pela quantidade de atividades com 10, 20 e 30 identificadas como (J10),

(J20) e (J30) respectivamente. Cada conjunto de instância (J10, J20 e J30) tem 51 arquivos com dados diferentes, porém com a mesma complexidade.

Com o objetivo de analisar a melhora da etapa da Matheurística aplicada na solução proposta, foi inserida na Tabela 4 a coluna *PSO* que ignora a etapa da Matheurística do algoritmo descrito na seção 4.

Na tabela abaixo os resultados apresentados nas colunas *solver*, *PSO* e *PSOM* são dados pela média total do conjunto de instâncias. Porém, o *solver* foi capaz de executar apenas 2 das 51 instâncias do conjunto J20 com um tempo limite de 5 horas de processamento. Esse cenário é agravado no conjunto J30 onde o *solver* não é capaz de encontrar nenhuma solução ótima para as instâncias.

Tabela 4: Resultados do solver e PSOM com instâncias do PCDRMM

Inst	solver (C)	PSOM (C)	PSO (C)	solver (T)	PSOM (T)	PSO (T)
J10	63,67	72,43	90,55	113,34	6,78	2,55
J20	54,27	56,76	76,46	16751,20	4,15	2,36
J30	*	116,54	128,31	*	11,97	6,46

Os resultados exibidos na Tabela 4 apresentam uma aproximação de 4,14% do *PSOM* para as soluções exatas (*solver*) no

conjunto J10, em um tempo de processamento computacional 94,01% menor. Já na média duas soluções calculadas

pelo *solver* no conjunto J20, observamos um afastamento de 4,55% em relação ao PSOM. Contudo, a Matheurística finaliza sua execução 99,98% mais rápido do que o *solver*.

Já ao confrontar os resultados do *PSO* com sua versão sem a etapa da Matheurística (*PSOM*), temos uma melhora de 25,02% em relação a média dos custos das soluções do conjunto J10. Isso acontece em um tempo de processamento 62,39% maior do que no *PSO*. Essa melhora no custo das soluções devido a inserção da etapa da Matheurística, proporciona uma melhor de 34,71% em relação ao *PSO* no conjunto J20 em um tempo 75,85% mais lento. E no conjunto J30, *PSOM* apresenta soluções 10,1% melhores do que o *PSO*, em um tempo de processamento 85,29% mais lento do que o *PSO*. Esses valores expõem uma melhora no total de 16,03% com a inserção da etapa da Matheurística no algoritmo.

É notável com o crescimento do número de atividades, que o tempo de processamento para a criação de uma solução do *solver* torna-se exponencial e inviabiliza o uso em aplicações reais. Já o *PSOM* se mantém com a mesma proporção de tempo de processamento por quantidade de atividades nas instâncias. Esses testes apontam a qualidade da Matheurística como melhor solução para o problema, uma vez que conserva o tempo de processamento

computacional reduzido, e otimiza o custo da solução final.

5.CONCLUSÕES

A Matheurística *PSOM*, constitui-se como uma boa alternativa para solucionar o Problema de Custo de Disponibilidade de Recursos com Múltiplos Modos, uma vez que cria soluções com diferença de apenas 4% dos valores ótimos em um tempo 94% vezes menor do que os obtidos pela implementação matemática do problema. O *PSOM* obteve um desempenho conclusivamente superior na solução das instâncias examinadas. Em decorrência, sugere-se um potencial promissor para a solução do PCDRMM conforme presentemente adaptado ou mesmo sua aplicação em problemas com estrutura matemática correlata.

Pela experiência adquirida no teste computacional, sugere-se que a modelagem dos movimentos dos recursos entre as atividades, bem como o uso de aperfeiçoamento heurístico baseada em grafo de dependência descrito em Costa et al. [2013]. São alternativas promissoras de melhoria de performance do algoritmo proposto. Adicionalmente aconselha-se o desenvolvimento de heurísticas construtivas como o GRASP ou o uso dos controles de memória típicos da Busca Tabu para hibridizar as etapas de busca local.

REFERÊNCIAS

- [1] Ball, Michael O. "Heuristics based on mathematical programming." *Surveys in Operations Research and Management Science* 16.1 (2011): 21-38.
- [2] Blazewicz, J., Lenstra, J., e Rinnooy Kan, A. (1983). Scheduling subject to resource constraints: classification and complexity. *Discrete Applied Mathematics*, 5:11-24.
- [3] Calazans, A. T. S. (2012). Qualidade da informação: conceitos e aplicações. *TransInformação*, 20 (1).
- [4] Chen, L., Li, X., e Cai, Z. (2012). Heuristic methods for minimizing resource availability costs in multi-mode project scheduling. *Systems, Man, and Cybernetics (SMC)*, 2012 IEEE International Conference on.
- [5] Hsu, C. C. e Kim, D. S. (2005). A new heuristic for the multi-mode resource investment problem. *Journal of the Operational Research Society*.
- [6] Ibáñez, M. L., Dubois-Lacoste, J., Stutzle, T., e Birattari, M. (2011). The irace package, iterated race for automatic algorithm configuration. (TR/IRIDIA/2011-004). URL <http://iridia.ulb.ac.be/IridiaTrSeries/IridiaTr2011-004.pdf>.
- [7] Kennedy, J. e Eberhart, R. (1995). Particle swarm optimization. *International Conference on Neural Networks*.
- [8] Möhring, R. (1984). Minimizing costs of resource requirements, project networks subject to a fixed completion time. *Operations Research*, 32:89-120.
- [9] Nadjafi, B. A. (2014). Multi-mode resource availability cost problem with recruitment and release dates for resources. *Applied Mathematical Modelling*, 38.
- [10] Costa, W., Procópio, L. D. P., Filho, G. F. S., Cabral, L. F. A., e Silva, G. C. (2013). Uma abordagem híbrida aplicada ao problema da alocação dinâmica de espaços. 11th Brazilian Congress on Computational Intelligence (BRICSCCI and CBIC).

[11] Qi, J.-J., Liu, Y.-J., Jiang, P., e Gou, B. (2015). Schedule generation scheme for solving multi- mode resource availability cost problem by modified particle swarm optimization. *Journal of Scheduling*, 18:285–298.

[12] Talbot, F. B. (1982). Resource-constrained project scheduling problem with time-resource trade-offs: The nonpreemptive case. *Management Science*, 28(10):1197 – 1210.

[13] Taylor, F. W. (2004). *Scientific management*. Routledge.

[14] Yamashita, D. e Morabito, R. (2012). Um algoritmo branch-and-bound para o problema de programação de projetos com custo de disponibilidade de recursos e múltiplos modos. *Gestão e Produção*.

Capítulo 11

RECOMENDAÇÃO DE GRUPOS DE FACEBOOK A PARTIR DE TEXTOS E IMAGENS DE USUÁRIOS

Gustavo Ramos Lima

Patrick Marques Ciarelli

Resumo: Este trabalho tem como foco a recomendação de grupos do Facebook a usuários, baseada no perfil de postagem de cada usuário do Facebook. Para realizar esta tarefa, os textos e as imagens divulgadas pelos usuários são usados como fonte de informação, e os experimentos indicaram que o uso conjunto dos dois tipos de informações fornece resultados melhores ou iguais ao uso das mesmas informações de forma separada. Este trabalho foi resultado de uma iniciação científica, na qual o aluno realizou as tarefas de revisão de literatura, pré-processamento das bases de dados, experimentos de classificação, implementação de parte do código e uso de códigos de terceiros, conforme indicado no trabalho. O aluno também contribuiu parcialmente na escrita deste artigo.

Palavras chave. Recuperação da Informação, Redes Sociais, Classificação.

1 INTRODUÇÃO

O crescimento das redes sociais online, como Facebook, WhatsApp, Twitter e Instagram, tem sido imenso na última década, graças ao acesso cada vez mais fácil à internet e ao barateamento de computadores e dispositivos móveis que permitem conexão a essas redes. Atualmente, existem mais de 2 bilhões de usuários no mundo conectados em ao menos uma rede social e a perspectiva é que este número aumente a cada ano [Statista, 2016].

Várias redes sociais e muitos usuários deixam as mensagens, imagens e vídeos para acesso público, podendo qualquer pessoa ler o conteúdo e, inclusive, interagir com o usuário que a divulgou. O grande número de usuários somada a grande quantidade de mensagens divulgadas por minuto nas redes sociais (só no Facebook é estimado cerca de 500 mil comentários postados por minuto [Zephoria, 2018]), indicam que a quantidade de dados disponíveis não só é gigantesca mais também abrange grande parte da população mundial.

Com esses dados é possível identificar hábitos, interesses, padrões de comportamento e outras informações sobre um usuário que podem ser utilizados para modelar o seu perfil. Com perfis de usuários é possível oferecer uma variedade de aplicações, tais como mecanismos personalizados de busca, recomendação de amigos e conteúdo, identificação de usuários com comportamento malicioso e indicação de produtos e serviços direcionados aos interesses e necessidades dos usuários [Kosinski et al., 2013; Jin et al., 2013; Mislove et al., 2010]. Na área de marketing, é cada vez mais comum as empresas buscarem saber mais sobre os seus consumidores de forma a oferecer produtos de interesse e elevar seus lucros. Abordagens desta natureza já são usadas por grandes empresas, como a Amazon e o Submarino. Para além dos usos comerciais, pesquisas científico-acadêmicas contribuem igualmente através da exploração de múltiplos processos, colocando em questão métodos e propondo análises diversificadas.

Uma abordagem comum para identificar informações relevantes dos usuários é através dos textos que são publicados nas redes sociais, pois são muito informativos e é o meio primário de comunicação na maioria das redes. No entanto, existem alguns fatos que dificultam a extração de informação dos textos, alguns deles são os erros de

ortografia, palavras com vários significados semânticos, sinônimos, os diversos idiomas no mundo e o imenso dicionário de palavras que cada idioma possui. Uma opção para obter mais informações dos usuários, e tornar os perfis mais precisos, é utilizar dados provenientes de imagens e fotografias. Essa opção é viável, pois é muito comum usuários não só divulgarem mensagens como também imagens nas redes sociais. Em [Alqhtani et al., 2015] é utilizada uma abordagem que combina as informações de texto e imagem, e foi obtida uma ligeira melhora na qualidade dos resultados quando combinada esses dois tipos de informações em comparação a usá-las separadamente. Esta observação motiva uma melhor análise da combinação desses dados.

Neste contexto, este trabalho analisa o uso de textos e imagens na tarefa de recomendar grupos de Facebook para usuários. Para realizar esta tarefa, foi utilizada uma base de dados capturada de grupos de Facebook, do qual foram selecionados alguns dos usuários com maior frequência de postagem nestes grupos. Baseado no histórico de postagens realizadas pelos usuários e numa modelagem de cada grupo, são recomendados os grupos os quais as pessoas têm maior afinidade. As informações foram extraídas tanto dos textos quanto das imagens, sendo testados três tipos de abordagens: uma usando somente informação textual, outra usando informação das imagens e uma última combinando ambas as informações. Os resultados obtidos indicaram que o uso dos dois tipos de dados combinados pode fornecer resultados melhores do que usando eles de forma separada.

Este trabalho possui a seguinte estrutura. Na Seção 2 são apresentadas as técnicas utilizadas para extração de informação em textos e imagens. Na Seção 3 é apresentada a base de dados. O método para a recomendação de grupos e os algoritmos de aprendizado de máquina usados nos experimentos são descritos na Seção 4. Os resultados experimentais e a sua análise são realizados na Seção 5. Finalmente, a Seção 6 termina o trabalho com as conclusões e trabalhos futuros.

2 TÉCNICAS DE EXTRAÇÃO DE INFORMAÇÃO

2.1 REPRESENTAÇÃO DE TEXTOS

Quando se trabalha com textos é necessário realizar uma série de procedimentos para facilitar a recuperação de informação. A seguir são descritos os passos seguidos neste trabalho.

O primeiro passo a ser realizado é o pré-processamento do texto. Nesta etapa são removidos pontuação, caracteres não alfanuméricos e números. A seguir, as letras são passadas para caixa baixa e termos com baixo valor semântico, conhecidos como *stopwords* [Baeza-Yates e Ribeiro-Neto, 1998], são removidos, pois além de não contribuírem com informações relevantes, eles podem prejudicar a tarefa de recuperação de informação. Alguns exemplos de *stopwords* são artigos, preposições e interjeições.

Após este procedimento, é utilizada uma etapa para obter a forma canônica das

palavras, isto é, sem as flexões verbais, diminutivos, aumentativos e plural. Esta etapa é importante, pois o computador não é capaz de entender, sem este processamento, que palavras como, professor e professores estão se referindo a um mesmo tipo de profissão. Neste trabalho foi utilizado o software Cogroo (cogroo.sourceforge.net) que permite, entre outras coisas, obter a forma canônica de palavras da língua portuguesa.

Um próximo passo é adotar uma forma eficiente para representar um texto tal que possa ser processado por um computador de forma adequada. Uma das formas mais populares de representação de textos é através do modelo de espaço vetorial apresentado por Salton e Yang [1975]. Neste modelo, cada documento (ou texto) d_j é representado por um vetor de n dimensões, sendo que cada posição t_i deste vetor representa um termo que está associado a um peso $w_{i,j}$. Neste trabalho, $w_{i,j}$ é o número de vezes que o termo t_i ocorreu dentro do documento d_j . Desta forma, o documento d_j é representado conforme a Equação 1:

$$d_j = \{w_{1,j}, w_{2,j}, \dots, w_{n,j}\} \quad (1)$$

Para um conjunto de D documentos, é obtido um conjunto de termos distintos que formam o dicionário de termos de D . Quando um termo não ocorre dentro de um documento é associado um peso zero a ele. Com a representação vetorial é possível utilizar métricas como distância euclidiana e similaridade de cosseno para medir o grau de semelhança entre diferentes documentos. Na tarefa realizada neste trabalho, cada documento está associado a uma categoria presente no conjunto C , que são os grupos do Facebook.

Dois passos adicionais são realizados sobre o conjunto de documentos D . O primeiro é a remoção de termos pouco discriminativos, que neste trabalho são considerados os termos que apareceram em apenas um documento de D ou que apareceram em todos os documentos de D . No primeiro caso, a remoção é motivada, pois pode ser um erro

de ortografia ou palavra muito rara que não influencia nos resultados, mas que aumenta o número de termos presentes nos documentos, aumentando o custo computacional na etapa de classificação. No segundo caso, a presença de termos em todos os documentos pode indicar ser uma *stopword* que não foi removida dos textos.

O segundo passo adicional corresponde ao uso da técnica de ponderação dos pesos dos termos, conhecida como *Inverse Document Frequency* (idf), que é utilizada para aumentar a importância dos termos que aparecem em poucos documentos e diminuir daqueles termos que aparecem em muitos documentos [Salton e Yang, 1975]. Essa ponderação é combinada com a frequência de ocorrência de cada termo em cada documento, sendo obtido o novo peso $\widehat{w}_{i,j}$ do termo t_i no documento d_j . Este novo peso é calculado pela Equação 2:

$$\widehat{w}_{i,j} = w_{i,j} \times idf_i, \quad idf_i = \log \left(\frac{|D|}{nd_i} \right) \quad (2)$$

onde $w_{i,j}$ é a frequência do termo t_i no documento d_j , idf_i é o valor do idf para o termo

t_i , sendo $|D|$ o número de documentos do conjunto D e nd_i é o número de documentos

que o termo t_i apareceu, ou seja, o número de documentos em que $w_{ij} > 0$.

Assim, ao final de todos os procedimentos, é obtida uma matriz numérica M com dimensões $|D| \times n$, representando os documentos, e um vetor N com dimensões $|D| \times 1$, representando o vetor de categorias associados aos documentos, sendo n o número de termos distintos no conjunto D .

2.2 REPRESENTAÇÃO DE IMAGENS

Assim como os textos, as imagens também precisam passar por uma etapa de pré-processamento para extrair informações relevantes e reduzir a dimensionalidade dos dados.

Existem várias formas propostas para representar imagens em tarefas de categorização, porém muitas delas são lentas, complexas, necessitam de calibrar muitos parâmetros ou precisam de um conjunto de imagens com vários elementos rotulados [de Souza Gazolli, 2014]. Neste trabalho foi adotado a abordagem proposta por [de Souza Gazolli e Salles, 2012], por ser rápida, simples de implementar, sem parâmetros para calibração e apresentar

resultados semelhantes aos de outras abordagens.

A abordagem usada neste trabalho, chamada de Transformada Census Modificada Contextual (CMCT), representa cada imagem na forma de um vetor. Nesta abordagem, inicialmente é obtida a imagem em escala de cinza. Sobre cada pixel x da imagem é centralizada uma pequena máscara 3×3 que realiza uma operação sobre os pixels vizinhos ao pixel x . Primeiro é calculada a intensidade média $\bar{I}_x$ dos pixels embaixo da máscara. A seguir, as intensidades dos pixels vizinhos ao pixel x (I_y) são comparadas a $\bar{I}_x$. Se o valor da intensidade de um pixel y (I_y) for maior ou igual a $\bar{I}_x$, um bit 1 é gerado no pixel da máscara 3×3 que está sobre o pixel y . Caso contrário, é gerado um valor 0 no mesmo local da máscara. A Equação 3 apresenta este cálculo. Após passar por todos os pixels da vizinhança de x , a máscara 3×3 estará com valores binários em todas as posições, exceto no pixel central. Esses valores binários formam uma palavra de 8 bits que é convertida em um número inteiro entre 0 e 255. Esta operação inicial é chamada de Transformada Census Modificada (MCT), ou MCT8 por formar palavras de 8 bits.

$$T_x = \bigotimes_{y \in N_x} \zeta(I_y, \bar{I}_x), \quad \zeta(m, n) = \begin{cases} 1, & m \geq n \\ 0, & \text{caso contrário.} \end{cases} \quad (3)$$

O número obtido na Equação 3 é armazenado numa posição relacionada ao pixel x para gerar uma nova imagem. Ou seja, após passar a máscara em todos os pixels da imagem original, uma nova imagem é obtida. Dessa nova imagem é calculado o seu histograma, ou seja, o número de vezes que cada valor ocorreu na nova imagem. Este histograma possui $2^8 = 256$ elementos.

Após este passo, o MCT é passado sobre essa nova imagem e obtido o histograma da

imagem resultante, novamente com 256 elementos. Os histogramas da primeira e da segunda imagem gerada são concatenados formando um vetor h de 512 elementos. Para evitar uma diferença muito grande entre os valores dos elementos, e obter um melhor desempenho de classificação, é usada uma operação de logaritmo sobre os valores não nulos do vetor h , conforme indicada na Equação 4. Após este passo, o novo vetor $\tilde{h}$ é normalizado pela Equação 5.

$$\tilde{h}_i = 1 + \log(h_i), \quad \forall i \in \{1, \dots, 512\} | h_i > 0. \quad (4)$$

$$\hat{h}_i = \frac{\tilde{h}_i}{\sum_{i=1}^{512} \tilde{h}_i}, \quad \forall i \in \{1, \dots, 512\}. \quad (5)$$

A Figura 1 ilustra os passos realizados para representar a imagem em um vetor de características pelo método CMCT [de Souza Gazolli, 2014]. O processo começa pela

imagem e termina com o CMCT normalizado, sendo que o CMCT possui 512 elementos, numerados de 0 a 511.

Figura 1: Processo de extração do CMCT. Versão modificada da figura em [de Souza Gazolli, 2014].

O procedimento anterior é aplicado para cada imagem da base de dados. Ao final, é obtida uma matriz numérica F com dimensões $W \times 512$ representando as imagens e um vetor N com dimensões $W \times 1$, representando o vetor de categorias associadas às imagens, sendo W o número de imagens na base de dados.

um campo, seja de mensagem ou de link de imagem, estivesse vazio. Também é considerado um link de imagem vazio se houver um link, mas não foi coletada a imagem. Finalmente, após estas etapas, obteve-se a base de dados usada neste trabalho.

3. BASE DE DADOS

3.1 COLETA DOS DADOS

Para a realização dos experimentos foram coletados posts de seis grupos da rede social Facebook, postados no período entre 01/04/2011 a 15/07/2015, usando a API em java Facebook4j (facebook4j.org). Os grupos são sobre diferentes assuntos, como natureza, religião, política, etc. De cada post foi coletado um conjunto de campos, como mensagem, data de postagem, identificação do post, nome do grupo, link da imagem postada no Facebook, entre outros.

Uma vez coletados os posts, o segundo passo foi a coleta das imagens. Nesta etapa, somente as imagens presentes no Facebook foram coletadas, visto que links que apontavam para páginas externas poderiam estar quebrados ou terem sido atualizados, tal que poderiam mostrar imagens diferentes das originais postadas. Além disso, só foram coletadas imagens que tivessem pelo menos uma dimensão mínima de 15×15 pixels e tamanho mínimo de 15 kB.

Após a coleta das imagens, um procedimento de filtragem sobre os dados coletados foi aplicado. Este procedimento teve como função remover todos os posts cujo ao menos

3.2. PREPARAÇÃO DOS DADOS

Para realizar a tarefa de recomendar grupos aos usuários de acordo com os seus perfis de postagens, os seguintes procedimentos foram realizados. Inicialmente, foram identificados os 20 usuários com o maior número de postagens na base de dados. Todos os posts desses 20 usuários foram separados da base de dados, e com eles serão obtidos informações dos perfis de postagens desses usuários para poder recomendar grupos a eles. Estes dados serão chamados de base de usuários.

Removidos todos os posts desses 20 usuários da base de dados, o passo seguinte foi selecionar aleatoriamente, de cada grupo, 1000 posts. Esse número foi selecionado para que o número de posts por grupo fosse igual e evitar que um grupo tenha mais dados que os outros, já que a quantidade de posts por grupo varia muito. Esses posts serão usados para obter informações de cada grupo. Esses dados serão chamados de base de grupos.

Para cada post da base de usuários e de grupos foi obtida a representação em forma de vetor da mensagem, conforme o procedimento descrito na Seção 2.1. Assim, foi obtida uma matriz de mensagens da base

de usuários de 4830×6183 , ou seja, 4830 posts e 6183 termos diferentes, enquanto que as mensagens para a base de grupos formaram uma matriz de 6000×6183 . Esse procedimento foi seguido tomando o cuidado de garantir que os índices de palavras em ambas as matrizes de dados fosse o mesmo. Por exemplo, se na primeira matriz a palavra “descansar” estivesse na posição 45, na segunda matriz também seria 45.

Também foi obtida para cada base de dados a representação vetorial de cada imagem incluída nos posts. Para isso, o procedimento apresentado na Seção 2.2 foi utilizado. Para a base de dados de usuários, foi obtida uma matriz de imagens de 4830×512 e, para a base de grupos, uma matriz de 6000×512 .

Assim, cada post de cada base de dados é representado por dois vetores: um vetor representando a mensagem e outro vetor representando a imagem. Com essas matrizes serão realizados os experimentos, conforme a descrição na Seção 4.

4. MÉTODO DE RECOMENDAÇÃO DE GRUPOS

A abordagem proposta para recomendar os grupos aos usuários se baseia no perfil dos posts de cada usuário. Um algoritmo de classificação será inicialmente treinado com uma das matrizes obtidas da base de grupos, que pode ser a de mensagens ou a de imagens. Após o treinamento, será apresentada ao classificador a matriz equivalente da base de usuários. Ou seja, se o classificador foi treinado com uma matriz de mensagens, a matriz apresentada ao classificador será também uma matriz de mensagens.

Sendo assim, o classificador retornará o grupo (classe) mais provável a ser associado ao vetor avaliado. Ao final do processo, todos os vetores da matriz de usuários estarão associados a um grupo pelo classificador. Após esta etapa, é identificado para cada usuário o número de posts associado a cada um dos grupos, de tal forma que será obtido um vetor com g dimensões, sendo g o número de grupos, que neste trabalho é igual a 6. A partir desse vetor são recomendados os grupos para o usuário. A Tabela 1 mostra um exemplo fictício desse resultado para dois usuários e três grupos.

Tabela 2: Exemplo da categorização de posts de dois usuários.

Usuários	G1	G2	G3
usuário_1	23	6	10
usuário_2	18	5	33

Uma abordagem simples para a recomendação seria recomendar os grupos cujo usuário tenha ao menos um post rotulado neste grupo. Porém, como pode ser visto na Tabela 1, é comum o usuário ter posts rotulados em mais de um grupo (ou em todos, no caso do exemplo), e essa abordagem iria falhar ao recomendar todos os grupos. Uma alternativa proposta é aplicar um limiar τ sobre os números de posts. Todos os grupos que possuem um número de posts maior ou igual a $\tau \times \max_{post}$ é recomendado ao usuário, onde $\max_{post}$ é o número de posts do grupo com o maior número de posts. Para o caso da Tabela 1, se $\tau = 0,5$, o usuário_1 é associado ao grupo G1 e o usuário_2 aos grupos G1 e G3.

Existem vários classificadores que poderiam ser usados neste trabalho, sendo neste trabalho usado dois: o kNN (k Vizinhos mais

Próximos) e SVM (*Support Vector Machine*) (em ambos os casos foram usadas implementações de terceiros). O kNN foi usado por questão de simplicidade de implementação e entendimento. O kNN é um classificador que calcula a distância entre cada vetor de treinamento e o vetor que será classificado. A partir deste resultado, são selecionados os k vizinhos mais próximos do vetor de teste, que são os k vetores de treinamento mais próximos do vetor de teste. A classe (grupo) com mais vetores entre os k vizinhos mais próximos é associada ao vetor de teste [Duda et al., 2001]. A medida mais comum a ser aplicada no kNN para textos é a similaridade cosseno [Baeza-Yates e Ribeiro-Neto, 1998], sendo portanto a utilizada neste trabalho. A Equação 6 mostra esse cálculo, onde x e y são dois vetores e d é o número de elementos nos vetores.

$$\cos(x, y) = \frac{\sum_{i=1}^d x_i y_i}{\sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2}} \quad (6)$$

O SVM é mais complexo, mas foi usado porque obteve melhores resultados que o kNN para classificação de imagens, segundo o observado em [de Souza Gazolli, 2014]. O SVM encontra hiperplanos que conseguem separar, com uma margem de separação maximizada, exemplos (vetores) de diferentes classes (grupos). Caso os exemplos não sejam linearmente separáveis, é necessário usar um fator de custo c para definir o efeito da separação errada dos exemplos. Além disso, pode-se mapear os exemplos em um novo espaço para facilitar a separação, geralmente um espaço de maior dimensão. Para isso, um método chamado de kernel é usado [Bennett e Campbell, 2000]. Neste trabalho foi utilizado o kernel de interseção de histograma [Barla et al., 2003], o mesmo utilizado por [de Souza Gazolli, 2014]. Para o SVM utilizado é necessário definir o valor de c .

Para avaliar a qualidade dos resultados do método proposto foi utilizada a métrica acurácia, que consiste no número de usuários cujos grupos foram recomendados corretamente dividido pelo número de usuários, multiplicado por 100%. Esse valor varia entre 0 a 100%, e quanto maior for o valor, melhor.

5. EXPERIMENTOS E RESULTADOS

Antes de usar a abordagem proposta, é necessário definir o limiar τ , e os valores de k

e c do kNN e SVM, respectivamente. Foram testados um conjunto de valores em uma etapa de validação de forma a maximizar a acurácia do método. Para isso, foram testados valores de k na faixa entre 1,2,...,15 e limiares de τ entre 0,4 e 1,0, aumentando em passos de 0,1. Os diferentes valores de k e τ foram avaliados obtendo a acurácia sobre 5 dos 20 usuários e usando os 6000 posts dos grupos para treinar o kNN. Os valores de k e τ que retornaram a maior acurácia foram usados para recomendar os grupos para os demais 15 usuários. Foram encontrados valores de k e τ para os dados de mensagens e de imagens. Procedimento semelhante foi usado para o SVM: os mesmos 5 usuários foram usados para obter os melhores valores de c e τ , tanto para os dados de mensagens como de imagens, sendo o SVM treinado com os dados dos grupos. Os melhores valores foram usados para recomendar os grupos para os demais 15 usuários. Os valores de c avaliados foram entre 1 e 15, aumentando em passos de 2.

A segunda linha da Tabela 2 mostra a acurácia obtida sobre os 15 usuários quando usando somente a informação das mensagens, enquanto que a terceira linha mostra os resultados quando usando somente as imagens como informação. A segunda coluna mostra os resultados obtidos pelo kNN e a terceira mostra os resultados do SVM.

Tabela 3: Resultados dos experimentos com os dados separados.

Tipo de Dados	kNN	SVM
Mensagens	73,33%	66,67%
Imagens	53,33%	66,67%

Como pode ser observado, o kNN obteve um resultado melhor do que o SVM para recomendar grupos a partir das mensagens, enquanto o SVM foi melhor que o kNN quando é usada a informação proveniente das imagens. Também é observado que as mensagens foram mais informativas que as imagens, pois as acurácias obtidas com as mensagens foram maiores ou iguais às obtidas pelas imagens. No entanto, as imagens mostraram ser uma fonte de informação útil, pois só com esse tipo de informação foi possível recomendar corretamente os grupos, no pior dos casos,

para 53,33% dos usuários, ou o equivalente a 8 usuários.

Um novo conjunto de experimentos foram realizados combinando as informações de texto e imagens, sendo essa combinação realizada de duas formas. Na primeira são treinados dois kNN, um com os dados de mensagens e outro com os dados de imagens, com os valores de k encontrados na etapa anterior. Assim, cada post é classificado duas vezes: o texto e a imagem, e é somado o número de posts associados a cada grupo. Novamente é encontrado um valor de τ sobre os 5 usuários para ser

aplicado sobre os outros 15 usuários. O mesmo procedimento é usado para o SVM.

Na outra forma, os vetores de mensagens e imagens de cada post são concatenados, formando um único vetor por post. Os parâmetros de kNN e SVM são selecionados conforme explicado anteriormente, e são obtidos os resultados sobre os 15 usuários.

A Tabela 3 indica os resultados obtidos, sendo a segunda linha o resultado quando se usa dois classificadores e a terceira linha quando os vetores são concatenados. Os resultados do kNN e do SVM são mostrados na segunda e terceira coluna, respectivamente.

Tabela 3: Resultados dos experimentos com os dados juntos.

Tipo de Dados	kNN	SVM
Dois Classificadores	73,33%	66,67%
Mensagens + Imagens	73,33%	80,00%

Como visto na tabela, os resultados de ambos os métodos foram iguais ou superiores ao uso de somente um tipo de informação, e para o caso da concatenação dos vetores, o SVM obteve uma melhora em relação ao uso dos dados de forma separada. Este resultado realça a importância de usar ambos os tipos de informação em tarefas de recomendação e classificação.

6. CONCLUSÃO

Neste trabalho foi analisado o uso de informações extraídas de textos e imagens divulgados pelos usuários para recomendar grupos do Facebook. Os resultados experimentais obtidos indicam que tanto textos como imagens carregam informações relevantes para realizar esta tarefa, e que a combinação desses dois tipos de dados pode fornecer melhores resultados.

Trabalhos futuros incluem o estudo de métodos para não indicar nenhum grupo, caso o usuário de fato não pertença e não tenha afinidade a nenhum dos grupos analisados. Também serão estudadas formas de usar outros campos coletados na base de dados original, como o campo tempo, com intuito de aprimorar os resultados. Por fim, o aumento do número de usuários avaliados permitirá uma análise mais precisa dos resultados.

7. AGRADECIMENTOS

Os autores gostariam de agradecer à Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES) pelo apoio financeiro, com o projeto nº 526/2015.

Patrick Marques Ciarelli agradece ao apoio financeiro fornecido pelo CNPq (bolsa 312032/2015-3).

REFERÊNCIAS

- [1] Alqhtani, S. M., Luo, S., e Regan, B. (2015). Fusing text and image for event detection in twitter. *The International Journal of Multimedia & Its Applications (IJMA)*, 7(1):27 – 35.
- [2] Baeza-Yates, R. e Ribeiro-Neto, B. (1998). *Modern Information Retrieval*. Addison-Wesley, New York, first edition edition.
- [3] Barla, A., Odone, F., e Verri, A. (2003). Histogram intersection kernel for image classification. *Proceedings of 2003 International Conference on Image Processing*, 3:513–516.
- [4] Bennett, K. P. e Campbell, C. (2000). Support vector machines: Hype or hallelujah? *SIGKDD Explorations Newsl.*, 2(2):1–13.
- [5] de Souza Gazolli, K. A. (2014). Utilizando Contexto na Representação de Imagens para a

Classificação de Cenas. PhD thesis, Universidade Federal do Espírito Santo.

- [6] de Souza Gazolli, K. A. e Salles, E. O. T. (2012). A contextual image descriptor for scene classification. *Online Proceedings on Trends in Innovative Computing*, p. 66–71.
- [7] Duda, R. O., Hart, P. E., e Stork, D. G. (2001). *Pattern Classification*. Wiley-Interscience, New York, 2 edition.
- [8] Jin, L., Chen, Y., Wang, T., Hui, P., e Vasilakos, A. V. (2013). Understanding user behavior in online social networks: a survey. *IEEE Communications Magazine*, p. 144–150.
- [9] Kosinski, M., Stillwell, D., e Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences of the United States of America*, p. 1–4.

[10] Mislove, A., Viswanath, B., Gummadi, K. P., e Druschel, P. (2010). You are who you know: inferring user profiles in online social networks. Proceedings of the third ACM international conference on Web search and data mining, p. 251–260.

[11] Salton, G. e Yang, C. S. (1975). A vector space model for automatic indexing. Communications of the ACM, p. 613–620.

[12] Statista (2016). Number of social network users worldwide from 2010 to 2019 (in billions). Disponível em: <http://www.statista.com/statistics/278414/number-of-worldwide-social-networkusers/>.

[13] Zephoria (2018). The Top 20 Valuable Facebook Statistics - Updated June 2018. Disponível em: <https://zephoria.com/top-15-valuable-facebook-statistic>

Capítulo 12

MODELO MULTIOBJETIVO DE ANÁLISE ENVOLTÓRIA DE DADOS COMBINADO COM O DESENVOLVIMENTO DE FUNÇÕES EMPÍRICAS E OTIMIZAÇÃO VIA SIMULAÇÃO MONTE CARLO

Aneirson Francisco da Silva

Fernando Augusto Silva Marins

Erica Ximenes Dias

Marcelo Vilela Figueiredo

Rafael de Carvalho Miranda

Resumo: O controle de qualidade é um dos principais pilares para um bom rendimento de uma linha de produção, visando garantir maior eficiência, eficácia e redução dos custos de produção. A identificação das causas dos defeitos e seu controle são atividades relativamente complexas devido às muitas variáveis presentes em certos processos produtivos. Este trabalho foi desenvolvido em uma grande indústria siderúrgica do Brasil, que atua na produção de componentes ferroviários e industriais, com o objetivo de reduzir defeitos de fundição de peças ferroviárias. A partir da base de dados, disponível na empresa estudada, foi avaliada a eficiência do processo de produção envolvendo sete produtos e trinta e oito variáveis de processo desses produtos. Nesta análise de eficiência, foi adotado o modelo Bi-objetivo de Análise de Envoltória de Dados (BiO-MCDEA) e foram evidenciadas as variáveis de processo de entrada (*input*) e saída (*output*) que são importantes para a melhoria da eficiência do processo produtivo. Com base neste conjunto de variáveis identificadas pelo modelo BiO-MCDEA, foram desenvolvidas funções empíricas por meio da análise de regressão não linear múltipla para representar o processo produtivo de fundições industriais e ferroviárias. Finalmente, com a transformação de cada coeficiente das funções objetivo em variáveis aleatórias, foi realizada a otimização via simulação Monte Carlo para determinar o melhor ajuste nas variáveis selecionadas como sendo relevantes pelo modelo BiO-MCDEA. Os resultados obtidos para o processo de produção foram interessantes e foram validados estatisticamente.

Palavras-chave: Análise Envoltória de Dados Multicritério, Simulação Monte Carlo, Regressão Não Linear, Eficiência do processo produtivo, Problemas Industriais.

1 INTRODUÇÃO

Desperdício é definido por Ghinato (2002) como atividade completamente desnecessária, que gera custo e não agrega valor, devendo ser totalmente eliminada. Dentre os desperdícios encontrados em processos produtivos, destacam-se os defeitos, os quais são definidos por Menegon et al. (2003) como produtos com alguma de suas características fora da especificação e que, por esta razão, não satisfazem aos requisitos de aplicação.

Há diversas aplicações de ferramentas estatísticas e de simulação na determinação de eficiência de processos produtivos, que são utilizadas para avaliar o nível de qualidade de determinado produto ou processo, e, neste trabalho, foram exploradas:

- MCDEA - *Multi Criteria Data Envelopment Analysis Models* (LI; REEVES, 1999, GHASEMI; IGNATIUS; EMROUZNEJAD, 2014).
- - Funções empíricas (WÖLK et al., 2002).
- OvMCS - *Optimization via Monte Carlo Simulation* (KROESE et al., 2011, KAZANTSEV et al., 2015 and KORYTKOWSKI; KARKOSZKA, 2016).

Ghasemi, Ignatius & Emrouznejad (2014) afirmam que a MCDEA tem ampla aplicação prática na determinação de eficiência relativa de DMUs - *Decision Making Units*, e tem sido uma das áreas da Pesquisa Operacional (PO) com maior ritmo de crescimento, tornando-se uma importante ferramenta para análise de eficiência em empresas públicas e privadas. Longo (2006) define OvSMC como uma técnica para resolver problemas matemáticos estocásticos. Em Kandel, Laerke & Elsgaard (2016), Tan, Awaiye & Trabelsi (2014) e Behara et al. (2011) pode-se encontrar a descrição e análise de trabalhos que

utilizaram a técnica estatística de análise de regressão para obter funções empíricas associadas à diversos problemas industriais.

As Figuras 1 e 2 mostram os resultados relativos às publicações e citações, obtidos na *Web of Knowledge*, utilizando, respectivamente, as Palavras-chave "*Multiple Criteria Data Envelopment Analysis*" e "*Optimization via Monte Carlo Simulation*". Estes resultados atestam que tais ferramentas têm sido bastante adotadas nas soluções de diversos problemas em variados contextos.

Foi realizada uma busca nas bases de dados *Scopus*, *Web of Knowledge* e *Emerald*, combinando nesta busca as palavras chaves: "MCDEA - *Multiple Criteria Data Envelopment Analysis*", "OvMCS - *Optimization via Monte Carlo Simulation*", "Regression", "Freight Wagon Parts" e "Foundry". Percebeu-se que não houve nenhuma ocorrência de artigos com essa combinação de palavras-chave, e, portanto, identificou-se lacunas nesta área de pesquisa relativa a avaliação da eficiência de processos produtivos na área siderúrgica. Isto gerou algumas questões de pesquisa que este trabalho buscou responder:

- - Como combinar a MCDEA com a OvSMC para avaliar eficiência de processos produtivos na área siderúrgica?
- - Há vantagens de se utilizar o modelo MCDEA na seleção de variáveis de entrada (*input*) e de saída (*output*) na análise de eficiência de processos produtivos na área siderúrgica?
- - Como validar estatisticamente os resultados da OvSMC quando aplicada na análise de eficiência de processos produtivos na área siderúrgica?

Figura 1. Publicações e citações utilizando a palavra-chave “Multiple Criteria Data Envelopment Analysis”

Source: Web of Knowledge

Figura 2. Publicações e citações utilizando a palavra-chave “Optimization via Monte Carlo Simulation”

Source: Web of Knowledge

Assim, o objetivo geral deste trabalho foi propor uma abordagem inovadora para avaliar eficiência de processos produtivos na área siderúrgica, combinando um modelo Multiobjetivo de Análise Envoltória de Dados, o desenvolvimento de funções empíricas e a Otimização via Simulação Monte Carlo. Os objetivos específicos, que dizem respeito a aplicação da nova abordagem proposta numa indústria siderúrgica brasileira que produz componentes ferroviários e industriais, incluíram:

- Avaliar a eficiência da produção de diferentes peças ferroviárias por meio de um modelo MCDEA.
- Determinar as variáveis (de *input* e de *output*) relevantes na produção de peças ferroviárias por meio de um modelo MCDEA.
- Desenvolver funções empíricas que representem adequadamente o processo produtivo de fundidos ferroviários e industriais.
- Validar estatisticamente os resultados gerados pela OvMCS no ajuste proposto às variáveis identificadas

como relevantes por um modelo MCDEA na produção de fundidos ferroviários e industriais.

Como delimitação desta pesquisa, para a obtenção das funções empíricas foi utilizado o algoritmo OLS - *Ordinary Least Squares* (LIU et al., 2016), na OvMCS foi utilizado o otimizador *Optquest* (YANG, 2010) e o modelo MCDEA adotado foi o BiO-MCDEA (GHASEMI; IGNATIUS; EMROUZNEJAD, 2014), pois, dentre os modelos MCDEA, ele tem gerado bons resultados com relação ao aumento da discriminação entre as DMUs, quando não é satisfeita a Regra de Ouro da DEA (BANKER et al., 1989).

Considerando estas características e objetivos do trabalho realizado, com base em Bertrand & Fransoo (2002), ele pode ser classificado como sendo uma pesquisa aplicada, com objetivos empíricos descritivos, pois o modelo desenvolvido descreve de forma adequada as relações causais que podem existir na realidade, favorecendo a compreensão de processos reais. A forma de abordar o problema foi a quantitativa, sendo o método de pesquisa a Modelagem e Simulação.

O artigo está estruturado em mais três seções. A Seção 2 apresenta o referencial teórico, com informações sobre MCDEA, OvSMC, e desenvolvimento de funções empíricas. Já na Seção 3 descreve-se o problema, o método e materiais utilizados, a modelagem realizada, os resultados obtidos e suas análises. Por fim, na Seção 4 são apresentadas as considerações finais, seguidas das referências bibliográficas.

2. MCDEA, OVSMC E DESENVOLVIMENTO DE FUNÇÕES EMPÍRICAS

Sueyoshi & Goto (2012) afirmam que a *Data Envelopment Analysis* (DEA) tem sido um método quantitativo amplamente utilizado na avaliação de eficiência de empresas nos setores públicos e privados. Neste contexto, Friesner, Mittelhammer and Rosenman (2013) afirmam que a DEA é uma das mais populares ferramentas para se medir eficiência em relação a custos e produtividade em indústrias. Os mesmos autores citam algumas vantagens em se utilizar DEA na determinação de eficiência, tais como:

- - Não incorpora regras rígidas, permitindo aos pesquisadores coletarem apenas dados de entrada (*inputs*) e saída

(*outputs*) que forem relevantes para as DMUs.

- - É técnica não paramétrica, correlacionando *inputs* e *outputs* mesmo sem o atendimento da hipótese de ocorrerem distribuições de probabilidade normal.
- Outras características importantes da modelagem DEA são citadas no trabalho de Azadeh, Sheikhalishahi & Koushan (2013), tais como:
 - - Identifica as DMUs que são *benchmarks* para as demais DMUs analisadas.
 - - Os dados de entrada e saída são representados na forma numérica.
 - - Associa o valor 1 (ou 100%) para as DMUs eficientes e valores menores que 1 para as DMUs ineficientes.
 - - Cria um *ranking* para DMUs ineficientes, porém o mesmo não é feito para as DMUs eficientes.

Dentre os vários modelos da DEA, há o modelo DEA CCR proposto por Charnes, Cooper & Rhodes (1978), com *constant returns to scale* (CRS), podendo ser formulado por:

$$\max w_0 = \sum_{r=1}^s u_r y_{r0} \quad (1)$$

Sujeito a:

$$\sum_{i=1}^m v_i x_{i0} = 1 \quad (2)$$

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0, \quad j = 1, 2, \dots, n \quad (3)$$

$$u_r \geq 0, r = 1, 2, \dots, s \quad (4)$$

$$v_i \geq 0, i = 1, 2, \dots, m \quad (5)$$

sendo w_0 o valor da eficiência da DMU que está sob análise (DMU_0), y_{r0} e x_{i0} , respectivamente, o valor da r -ésima saída e i -ésima entrada para a DMU_0 , y_{rj} o valor da r -ésima saída para a j -ésima DMU, x_{ij} o valor da i -ésima entrada para a j -ésima DMU, u_r a variável de decisão (peso) associada à r -ésima saída, v_i a variável de decisão (peso) associada à i -ésima entrada.

Os modelos DEA podem ser orientados aos *inputs* – buscando reduzir os valores dos

recursos utilizados, mantendo os resultados obtidos constantes, ou orientados aos *outputs* – buscando aumentar os valores dos resultados obtidos, mantendo constantes os valores dos recursos utilizados. A partir do modelo DEA CCR, Banker, Charnes & Cooper (1984) propuseram o modelo DEA BCC, com retornos variável de escala, com o acréscimo da variável c_0 , que é livre ou irrestrita, na função objetivo (1) e na restrição (3).

Uma importante questão discutida por Banker *et al* (1989) é que, a fim de evitar problemas com poder de discriminação dos modelos DEA, o número de DMU's deverá ser maior ou igual a três vezes a soma no número de variáveis (*inputs* e *outputs*), ou maior ou igual ao produto do número de variáveis de *input* pelo número de variáveis de *output*, adotando o critério que fornecer a maior quantidade de DMU's. Isto ficou conhecido como a Regra de Ouro da DEA.

De fato, existem dois problemas que podem ocorrer com o uso da DEA, segundo Li e Reeves (1999): poder discriminante fraco -

ocorre quando se identificam muitas DMUs como eficientes e a distribuição de pesos irrealista - refere-se à situação em que DMUs podem ser eficientes, mas apresentam valores altos de pesos para uma única saída, ou valores baixos de pesos para uma única entrada.

Para tentar evitar estes problemas citados, Li & Reeves (1999) propuseram o modelo de DEA com Múltiplos Critérios - *Multicriteria Data Envelopment Analysis* (MCDEA), em que cada critério de eficiência é uma função objetivo independente, dado por (6) - (14):

$$\min d_0 \text{ or } \max \sum_{r=1}^s u_r y_{r0} \quad (6)$$

$$\min M \quad (7)$$

$$\min \sum_{j=1}^n d_j \quad (8)$$

Sujeito a:

$$\sum_{i=1}^m v_i x_{i0} = 1 \quad (9)$$

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} + d_j \leq 0, j = 1, 2, \dots, n \quad (10)$$

$$M - d_j \geq 0, j = 1, 2, \dots, n \quad (11)$$

$$u_r \geq 0, r = 1, 2, \dots, s \quad (12)$$

$$v_i \geq 0, i = 1, 2, \dots, m \quad (13)$$

$$d_j \geq 0, j = 1, 2, \dots, n \quad (14)$$

Neste modelo MCDEA há três funções objetivo, sendo que primeira função objetivo d_0 está associada à maximização da eficiência para a DMU analisada. Sendo d_j a função objetivo que visa minimizar a ineficiência para a DMU analisada, e M é a função objetivo que visa minimizar o máximo valor de ineficiência. Os demais parâmetros e variáveis são análogos aos modelos anteriores.

No modelo MCDEA de (6) - (14), os três objetivos são analisados separadamente, sem ordem de preferência definida. Os ganhos de

eficiência definidos no âmbito dos critérios de mínimo e máximo são mais restritivos do que os definidos no DEA clássico, e estes dois critérios geralmente identificam menos DMU eficientes. Assim, o objetivo do modelo MCDEA não é extrair uma solução ótima, mas encontrar uma solução Pareto-ótima, deixando ao analista a decisão final. Já Ghasemi; Ignatius & Emrouznejad (2014) propuseram o modelo Bi-Objective Weighted DEA model (BiO-MCDEA) que busca tratar dois objetivos simultaneamente, considerando uma dada ponderação destes, segundo a formulação dada por (15) - (21):

$$\text{Min } h = \left(w_2 M + w_3 \sum_{j=1}^n d_j \right) \quad (15)$$

Sujeito a:

$$\sum_{i=1}^m v_i x_{i0} = 1 \quad (16)$$

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} + d_j = 0, \quad j = 1, 2, \dots, n \quad (17)$$

$$M - d_j \geq 0, \quad j = 1, 2, \dots, n \quad (18)$$

$$u_r \geq \varepsilon, \quad r = 1, 2, \dots, s \quad (19)$$

$$v_i \geq \varepsilon, \quad i = 1, 2, \dots, m \quad (20)$$

$$d_j \geq 0, \quad j = 1, 2, \dots, n \quad (21)$$

Neste artigo objetivou-se avaliar a eficiência da produção de diferentes peças ferroviárias de uma indústria siderúrgica no Brasil, bem como identificar as entradas e saídas variáveis que são importantes neste processo produtivo, para tanto adotou-se um modelo BiO-MCDEA, com $\varepsilon = 0,00001$ (GHASEMI; IGNATIUS; EMROUZNEJAD, 2014), para esta finalidade.

Outra técnica de Pesquisa Operacional que, de acordo com Yang (2010), tem sido aplicada em todas as áreas do conhecimento

é Simulação Monte Carlo (SMC). Neste trabalho a SMC foi utilizada em conjunto com técnicas de otimização, num procedimento conhecido por Otimização via Simulação Monte Carlo (OvSMC), que são indicados para solucionar problemas complexos que apresentem vários ótimos locais, envolvendo um *mix* de variáveis contínuas e discretas (KROESE; TAIMRE; BOTEV, 2011). Neste contexto, destaque-se que vários problemas estocásticos, que são de interesse neste trabalho, como o problema (22) proposto por Shapiro (2001).

$$\min_{x \in X} \{ \mathcal{G}(x) := E_A[\mathcal{G}(x, \omega)] \}, \quad (22)$$

sendo $G : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$, a expectativa em relação à medida de probabilidade P definida em um espaço amostral $(\Omega, \mathcal{F})$, e $X \subset \mathbb{R}^n$. Assume-se para cada $x \in X$ a expectativa de $\mathcal{G}(x)$ e bem definida. A função $G(x, \omega)$ em si pode ser definida por um problema de otimização. Uma etapa importante do trabalho aqui relatado envolveu a obtenção de funções empíricas que representassem adequadamente o processo produtivo de fundidos ferroviários e industriais.

Sabe-se que uma das maneiras de estimar uma função empírica é por meio da análise de regressão, que, segundo Helene (2013), é amplamente utilizada com esta finalidade em diversas áreas. Regressões múltiplas são utilizadas para estudar a relação entre variáveis dependentes e variáveis independentes (ou explicativas), por meio da obtenção de uma equação linear que represente as relações entre as mesmas, tal como expresso em (23):

$$y_l = \beta_0 + \beta_1 x_{1l} + \beta_2 x_{2l} + \dots + \beta_k x_{kl} + \tau_l, \quad l = 1, 2, \dots, p \quad (23)$$

sendo y_l uma variável dependente, $x_1, x_2, \dots, x_k$ as variáveis independentes, τ_l o fator de erro, e p o número de amostras observadas. Dentre as diversas técnicas de análise regressão, linear ou não linear, optou-se, aqui,

pelo algoritmo OLS - *Ordinary Least Squares* (HELENE, 2013, FERRARO; GIORDANI, 2012), que consiste em se desenvolver um modelo por meio de regressão linear ou não

linear e refinar os parâmetros com base em sucessivas iterações (LIU *et al.*, 2016).

Na próxima seção será descrito o problema tratado, as etapas do trabalho, incluindo a modelagem MCDEA, a obtenção das funções empíricas para os processos produtivos e a aplicação da OvSMC para a simulação de cenários de interesse da empresa em estudo.

3 MODELAGEM DO PROBLEMA E ANÁLISES DOS RESULTADOS

No desenvolvimento do trabalho foram seguidas as seguintes etapas:

- - Identificação do problema: Após entrevistas com os gestores de uma empresa brasileira de grande porte, que produz componentes ferroviários e industriais feitos à base de aço, foi estabelecida como questão principal a ser enfocada: Como melhorar a eficiência de processos produtivos industriais?

Num primeiro momento, para ser testada a nova metodologia proposta neste trabalho, foram escolhidos, pelos gestores da empresa estudada, os processos produtivos de sete peças (que foram as DMUs analisadas no modelo BiO-MCDEA) que compõem sistemas de acoplamento entre vagões e sustentação dos mesmos, tais peças pesam entre 300 kg a 1.000 kg, passando por diversos processos produtivos (ver Figura 3).

Após o acompanhamento do processo produtivo destas peças, que se dá conforme o fluxograma da Figura 3, os engenheiros da empresa estudada definiram entradas e saídas que seriam objeto de análise pelo modelo BiO-MCDEA, visando a identificação de melhorias neste processo.

Seguem descrições técnicas das variáveis de entrada:

- x_1 Tempo de Vazamento: Tempo decorrido entre a abertura e o fechamento da válvula de vazamento para a transferência do aço líquido da panela de vazamento para o molde, onde o aço irá solidificar no formato da peça;
- x_2 -Tempo de Argônio: Tempo de injeção de Argônio, durante a preparação do

aço líquido no forno. Argônio é um gás inerte e sua injeção é necessária para misturar o aço líquido com os demais elementos de liga adicionados, de forma a tornar a solução homogênea;

- x_3 Quantidade de Escória: Escória é a camada de materiais que ficam na parte de cima da panela de vazamento e tem a função de isolar termicamente o aço líquido, de forma a evitar troca de calor com o ambiente;

- x_4 Tempo entre Moldagem e Vazamento: Tempo decorrido entre a confecção do molde e o vazamento de aço no molde para a confecção da peça. Tempos muito longos podem ocasionar defeitos nas peças;

- x_5 Tempo entre Vazamento e Desmoldagem: Tempo decorrido entre o vazamento de aço líquido no molde e a desmoldagem. Tempos muito curtos podem ocasionar defeitos de conformação nas peças, tais como defeitos dimensionais, trincas, rechupes, dentre outros;

- x_6 Tempo com Forno Ligado: Tempo em que o Forno a Arco Elétrico fica ligado para fundir a sucata de aço;

- x_7 Tempo na Aciaria: Tempo de todo o processo de confecção de aço líquido, desde o carregamento de sucata no forno, adição de liga, fusão do aço e basculamento (transferência do aço líquido do forno para a panela de vazamento); - x_8 Tempo para Oxidação: Tempo de injeção de oxigênio para a oxidação de elementos de liga presentes no aço líquido;

- x_9 Última temperatura no Forno: Última temperatura do aço líquido registrada no forno, antes do basculamento;

- x_{10} Valor do Rendimento Energético: Calculado a partir da divisão do KWh gasto para fundir o aço pela tonelada de aço fundida;

- x_{11} Quantidade de Sucatas: Total de sucata fundida no forno;

- x_{12} Carbono Equivalente (C_{eq}): Calculado por meio de (24)

$$Ceq = C + \frac{Mn}{6} + \frac{(Cr+Mo+V)}{5} + \frac{(Ni+Cu)}{15} \quad (24)$$

sendo C a quantidade de Carbono, Mn a quantidade de Manganês, Cr a quantidade de Cromo, Mo a quantidade de Molibdênio, V a quantidade de Vanádio, Ni a quantidade de Níquel e Cu a quantidade de Cobre

- x_{13} Quantidade de Alumínio;
- x_{14} Quantidade de Carbono: Tem influência nas propriedades de Resistência Mecânica, Dureza, Modelo de Elasticidade, Soldabilidade e Fragilidade das peças produzidas; - x_{15} Quantidade de Cobre;
- x_{16} Quantidade de Cromo;
- x_{17} Quantidade de Enxofre: Influencia negativamente no surgimento de trincas (defeito de fundição);
- x_{18} Quantidade de Fosforo: Influencia negativamente no surgimento de trincas (defeito de fundição);
- x_{19} Quantidade de Manganês;
- x_{20} Quantidade de Molibdênio; - x_{21} Quantidade de Nióbio;
- x_{22} Quantidade de Níquel;
- x_{23} Quantidade de Silício;
- x_{24} Quantidade de Titânio;
- x_{25} Quantidade de Vanádio;
- x_{26} Limite de Resistência: Valor da tensão necessária para a ruptura do corpo de prova; x_{27} Limite de Escoamento: Valor da tensão necessária para iniciar o processo de deformação plástica do corpo de prova;
- x_{28} Alongamento: Utilizada para a medição da elasticidade do aço por meio da comparação do comprimento do corpo de prova antes e depois da ruptura ocasionada no ensaio de tração;

- x_{29} Redução de Área: Utilizada para a medição da ductilidade do aço, porém realizada por meio da comparação entre a área da seção transversal da ruptura com a seção transversal inicial do corpo de prova;

- x_{30} Medida de Dureza (HB): Utilizada para medir a dureza do aço resultante do processo.

As variáveis de saída x_{31} a x_{37} são consideradas como outputs indesejáveis e, portanto, no modelo BiO-MCDEA foram consideradas como entradas. Seguem as descrições técnicas das variáveis de saídas analisadas: - x_{31} Número de Trincas Externas: Quantidade de trincas na superfície externa da peça; - x_{32} Número de Trincas Internas: Quantidade de trincas na cavidade interna da peça; - x_{33} número de Entradas: Entradas de maçarico. Tal processo é realizado para se retirar as partes indesejadas da peça. Por falha humana ou inclusão de areia do molde na peça pode ocorrer o defeito “entrada de maçarico”, onde ocorre a retirada de aço além do necessário. Tal fato pode ocasionar o retrabalho da peça, preenchendo os vazios com solda ou então o refugo da peça, quando se trata de zona crítica sem a permissão de recuperação; - x_4 Número de Inclusões Externas: inclusão de materiais indesejáveis dentro do aço líquido que, ao se solidificar, permanecem na superfície externa da peça; - x_{35} Número de Inclusões Internas: Inclusões presentes nas cavidades internas das peças; - x_{36} Número de Defeitos Identificados pelo Ultrassom: Defeitos de inclusão ou trinca presente no interior da peça (fora da superfície); - x_7 Número de Rejeições pelo Calibre: Rejeições de calibres utilizados para avaliar o dimensional das peças; - y_1 Preço de Venda [\$].

Destaque-se que, segundo informações dos engenheiros da empresa estudada, as etapas produtivas que mais contribuem para a incidência de defeitos são: Confeção de Moldes, Preparação e Fusão do Aço, Enchimento dos Moldes, Desmoldagem e Tratamento Térmico.

Figura 3. Fluxograma de Produção..

Fonte: Empresa estudada

É interessante observar que, como o total de variáveis de entrada e variáveis de saídas (TN/OV) foi trinta e oito, pela DEA Golden Rule (BANKER et al., 1989), seriam necessárias 114 ($= 3 \cdot TN/OV$) DMUs, para se utilizar os modelos clássicos da DEA e manter o poder de discriminação da técnica. Este fato também justifica o uso do modelo BiO-MCDEA neste caso, pois há apenas sete DMUs para serem analisadas.

- Coleta de dados: As informações sobre os inputs e outputs para cada uma das sete

peças (DMUs) estudadas foram obtidas na base de dados da própria empresa. Como o volume de dados é muito grande, por uma questão de conveniência, eles não foram incluídos neste texto, mas encontram-se disponíveis no Apêndice A.

- Modelagem e solução do problema: O procedimento proposto neste trabalho para propor melhorias no processo produtivo em questão seguiu o fluxograma da Figura 4.

Fig. 4. Fluxograma de execução do projeto

Para formular e otimizar o modelo BiO-MCDEA visando a identificação dentre as sete DMUs (peças) aquelas com processos produtivos mais eficientes, bem como as variáveis de input e output mais importantes em cada processo produtivo, utilizou-se a programação VBA em Excel® e o algoritmo Simplex da Programação Linear (HILLIER; LIEBERMAN, 2001), disponível no Solver do Excel®.

Com as variáveis de entrada e saída importantes para cada DMU, foi aplicado o algoritmo OLS, disponível nas funções avançadas do Excel®, para o desenvolvimento das funções empíricas associadas aos respectivos processos produtivos. Para o ajuste das funções empíricas foi executada a OvSMC por intermédio do software Crystal Ball® em conjunto com seu otimizador Optquest (LI;

WANG, 2012). Nestas aplicações de métodos quantitativos foi utilizado uma máquina com Intel core i7, processador 2.8 de GHZ com 16 GB de memória RAM e sistema operacional IOS. O tempo computacional na OvSMC foi em média de 5 minutos para cada instância testada.

Passa-se a mostrar os resultados de cada etapa do procedimento e a comentá-los. O ranqueamento dos processos produtivos pelo modelo BiO-MCDEA está na Tabela 1, e percebe-se, que apenas as DMUs 5 e 7 (em negrito) foram eficientes, sendo que, a DMU 6 teve o pior desempenho em termos de eficiência. Atendeu-se, assim, o primeiro objetivo específico “Avaliar a eficiência da produção de diferentes peças ferroviárias por meio de modelo MCDEA”.

Tabela 1. Eficiência das DMUs

DMU	Eficiência
1	0,6719
2	0,6616
3	0,7916
4	0,5187
5	1,0000
6	0,3386
7	1,0000

O próximo passo consistiu em identificar, pelo modelo BiO-MCDEA, quais variáveis de entrada e de saída foram significativas para seus respectivos processos produtivos. Para tanto foram obtidos pelo algoritmo simplex (Solver do Excel®), aplicado ao modelo BiO-MCDEA, os valores dos pesos das variáveis de entrada (v_i) e de saída (u_j) para cada DMU e observou-se quais foram diferentes de zero.

Desta maneira foram identificadas como significativas as variáveis de entrada: Tempo de vazamento (v_1), Quantidade de Molibdênio (v_{20}), Quantidade de Nióbio (v_{21}), Quantidade de Vanádio (v_{25}), Número de Trincas Externas (v_{31}), Número de Trincas Internas (v_{32}), Número de Inclusões Internas (v_{35}), Número de Defeitos Identificados pelo Ultrassom (v_{36}) e Número de Rejeições pelo Calibre (v_{37}). Com isto foi atendido o segundo objetivo específico do artigo “Determinar as variáveis (de *input* e de *output*) relevantes na produção de peças ferroviárias por meio de modelo MCDEA”. Na sequência do trabalho, para as variáveis significativas de cada processo produtivo (DMU) já identificadas, foram construídas as respectivas funções empíricas por meio do algoritmo OLS. Cabe destacar que as saídas indesejáveis na aplicação da BiO-MCDEA passaram a ser as variáveis

dependentes (respostas) na etapa de desenvolvimento das funções empíricas.

Não foram estimadas as funções empíricas para as variáveis respostas “Número de Inclusões Internas (v_{35})” e “Número de Defeitos Identificados pelo Ultrassom (v_{36})”, pois estes defeitos são causados por variáveis de processo que não são bem controladas na empresa objeto de estudo, e não foram consideradas na análise realizada. Por razões similares, também não foi estimada a função empírica para a variável dependente (resposta) “Número de Rejeições pelo Calibre (v_{36})”, pois defeitos dimensionais ocorrem devido ao excesso ou pela falta de material onde se passa o calibre para a devida verificação dimensional.

Portanto foram consideradas, nesta etapa de obtenção de funções empíricas para processos produtivos em função das variáveis significativas, apenas as variáveis respostas “Número de Trincas Externas” e “Número de Trincas Internas”; após várias tentativas, chegou-se às formulações genéricas que apresentaram valores de R^2 -ajustado (HELENE, 2013) aceitáveis, ou seja, maiores que 70%, tendo sido utilizada uma amostra com 250 elementos para cada uma das variáveis analisadas:

$$\hat{Y}_1 = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + \tilde{\beta}_2 x_{20} + \tilde{\beta}_3 x_{21} + \tilde{\beta}_4 x_{25} + \tilde{\beta}_5 x_1^2 + \tilde{\beta}_6 x_{20}^2 + \tilde{\beta}_7 x_{21} x_{25} \quad (25)$$

$$\hat{Y}_2 = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 x_{20} + \tilde{\beta}_2 \frac{x_{21} x_{20}}{x_1^2 x_{25}} + \tilde{\beta}_3 \frac{\cos(x_{20})}{x_{21}} + \tilde{\beta}_4 \frac{\sin(x_{20})}{x_{25} x_{21} x_1} - \tilde{\beta}_5 \frac{x_1}{x_{21}} - \tilde{\beta}_6 \sqrt{x_1} \frac{x_{25}^2}{x_{20}^2} \quad (26)$$

sendo, $\hat{Y}_1$ e $\hat{Y}_2$, respectivamente, os valores estimados para o “Número de Trincas Externas” e “Número de Trincas Internas”.

As Tabelas 2 e 3, respectivamente, exibem os coeficientes (estatísticas) dos termos matemáticos componentes das funções empíricas de “Número de Trincas Externas” e de “Número de Trincas Internas”, obtidos pela aplicação do algoritmo OLS.

Como os valores de R^2 -ajustados para $\hat{Y}_1$ e $\hat{Y}_2$ foram, respectivamente, 73% e 77%, as funções empíricas foram consideradas adequadas por apresentarem R^2 -ajustado

acima de 70%, p -value menor que 5% e por apresentarem resíduos normais, mostrando assim, que não há tendenciosidade nestas funções empíricas adotadas (HELENE, 2013). Destaque-se que, portanto, ficou satisfeito o terceiro objetivo específico deste trabalho “Desenvolver funções empíricas que representem adequadamente o processo produtivo de fundidos ferroviários e industriais”.

Tabela 2. Estatísticas da Função Empírica obtida para “Número de Trincas Externas”

Termos Matemáticos na Expressão (23)	Valores dos Coeficientes (CV)	p-value [%]	Limitantes Inferiores (95%) (LB)	Limitantes Superiores (95%) (UB)
Intercepto ($\tilde{\beta}_0$)	20,9	0,0000	14,6	27,2
x_1	-0,30093	0,3596	-0,5025	-0,0993
x_{20}	-59,3	0,3475	-98,9	-19,7
x_{21}	-674,7	0,0005	-958,2	-391,3
x_{25}	-2,119,2	1,4215	-3,8095	-429
x_1^2	0,00165	4,0324	0,0001	0,0032
x_{20}^2	622,8	0,0000	465,6	779,9
$x_1 \cdot x_{25}$	39,1	4,2177	1,4	76,8

Source: Excel®

Tabela 3. Estatísticas da Função Empírica obtida para “Número de Trincas Internas”.

Termos Matemáticos na Expressão (24)	Valores dos Coeficientes (CV)	p-value [%]	Limitantes Inferiores (95%) (LB)	Limitantes Superiores (95%) (UB)
Intercepto ($\tilde{\beta}_0$)	1,98042	0,0000	1,3109	2,6500
$x_1 x_{20}$	0,19418	0,0000	0,1486	0,2398
$x_1 x_{20} / (x_{21}^2 x_{25})$	0,0000001	4,2526	0,0000005	0,000002
$\frac{\cos(x_{20})}{x_{21}}$	0,00084	1,0038	0,0002	0,0015
$\frac{\sen(x_{20})}{x_{25} x_{21} x_1}$	0,06978	0,8677	0,0181	0,1215
$\frac{x_1}{x_{21}}$	-0,14871	4,4500	-0,3065	0,0091
$\sqrt{x_1} \frac{x_{25}^2}{x_{20}^2}$	-26,38232	2,0000	-46,6642	-6,1000

Source: Excel®

Para realizar a OvSMC das funções empíricas (25) e (26), considerou-se cada coeficiente ($\tilde{\beta}$) nestas expressões como sendo uma variável aleatória $x \sim T [LB, CV, UB]$, com distribuição de probabilidade triangular

(KOKONENDJI; KIESSÉ; ZOCCHI, 2007), atendendo (27) e ilustrada na Figura 5; observe-se que os valores dos parâmetros estão nas Tabelas 3 e 4:

$$f(x) = \begin{cases} \frac{2(x-LB)}{(UB-LB)(CV-LB)} & LB \leq x < CV \\ \frac{2}{(UB-LB)} & x = CV \\ \frac{2(UB-x)}{(UB-LB)(UB-CV)} & CV < x \leq UB \\ 0 & \text{para qualquer outro valor} \end{cases} \quad (27)$$

sendo $f(x)$ a função densidade de probabilidade triangular para a variável

aleatória x , LB o limitante inferior, CV o valor mais provável e UB o limitante superior.

Figura 5. Distribuição de probabilidades triangular para a variável aleatória $x \sim T[LB, CV, UB]$

O problema para o qual foi aplicada a OvSMC para encontrar os valores ajustados das variáveis de entrada importantes já

identificadas, e que podem ser adotados pelos gestores nos processos produtivos na prática, está formulado por (28) – (29):

$$\min (\hat{Y}_1 + \hat{Y}_2) \quad (28)$$

s.t:

$$x_i \in X \text{ (Conjunto Discreto de Soluções viáveis)} \quad (29)$$

sendo V o conjunto de combinações de valores das variáveis de entrada utilizados pela Empresa objeto de estudo.

Destaque-se que, antes de aplicar a OvSMC ao problema (28) – (29), foi testada a existência de correlações entre as variáveis

de entrada, os resultados do Teste de Correlação de Pearson (MONTGOMERY, 2005) estão na Tabela 4, e constata-se que há correlação alta e positiva (em negrito) entre os pares de variáveis x_{21} - x_{20} , x_{25} - x_{20} e x_{25} - x_{21} , sendo as demais correlações não significativas estatisticamente.

Tabela 4. Pearson Correlation Values (p-values) para variáveis de entrada.

Variables	x_1	x_{20}	x_{21}
x_{20}	-0,028 (<i>p-value</i> 0,661)		
x_{21}	0,008 (<i>p-value</i> 0,932)	0,737 (<i>p-value</i> 0,000)	
x_{25}	-0,052 (<i>p-value</i> 0,414)	0,818 (<i>p-value</i> 0,000)	0,626 (<i>p-value</i> 0,000)

Source: Minitab 17®

Após a sequência de etapas seguidas na execução da OvSMC (tendo sido utilizado o Crystal Ball e o módulo Optquest), considerando estas correlações significativas

entre as variáveis de entrada; obteve-se os resultados que estão nas Tabelas 5 – 6 e ilustrados nas Figuras 6 – 7.

Tabela 5. Intervalos $[LB, UB]$ obtidos na OvSMC para variáveis de resposta “Número de Trincas Externas” e de “Número de Trincas Internas”, considerando existência das correlações significativas entre as variáveis de entrada.

Valor final das funções objetivo	LB	UB
$\hat{Y}_1$	3.3	16.9
$\hat{Y}_2$	2.5	8.7

Figura 6. Resultado para a variável resposta “Número de Trincas Externas” com a OvSMC considerando existência de correlação

Fonte: Crystal Ball®

Figura 7. Resultado para a variável resposta “Número de Trincas Internas” com a OvSMC considerando existência de correlação

Fonte: Crystal Ball®

Tabela 6. Valores ajustados obtidos na OvSMC para variáveis de entrada, considerando existência das correlações significativas entre elas.

Variável de Entrada	Valor
x_1	26
x_{20}	0.01
x_{21}	0.003
x_{25}	0.001

O quarto objetivo específico era “Validar estatisticamente os resultados gerados pela OvMCS no ajuste proposto às variáveis identificadas como relevantes por modelo MCDEA na produção de fundidos ferroviários

e industriais”. Para a consecução desta meta, comparou-se os resultados da OvSMC (Tabelas 5 e 6) com os valores reais ocorridos para as variáveis envolvidas disponibilizados no banco de dados da Empresa objeto de

estudo. Para fins de ilustração, na Tabela 7 estão alguns valores extraídos deste banco de dados que se situam dentro do intervalo

de confiança (95%) gerado pela OvSMC (Tabelas 5 e 6), ficando, portanto, validado o procedimento aqui proposto.

Tabela 7. Valores extraídos da base de dados da empresa estudada para validação dos resultados da OvSMC

Tipo	Variável	Valor
Valores finais das funções objetivo	$\hat{Y}_1$	6
	$\hat{Y}_2$	5
Valores de ajuste (recomendados) para as variáveis de entrada significativas	X_1	26
	X_{20}	0.01
	X_{21}	0.003
	X_{25}	0.001

4 CONCLUSÕES E RECOMENDAÇÕES PARA FUTURAS PESQUISAS

Passa-se a responder às questões de pesquisa propostas na Seção 1:

- Como combinar a MCDEA com a OvSMC para avaliar eficiência de processos produtivos na área siderúrgica? Foi descrito neste trabalho um procedimento para avaliar eficiência de processos produtivos na área siderúrgica que propiciou resultados interessantes e úteis para a Empresa estudada. Ele combina um Modelo BiO-MCDEA, com a geração de funções empíricas e a aplicação da OvSMC.
- Há vantagens de se utilizar o modelo MCDEA na seleção de variáveis de entrada e de saída na análise de eficiência de processos produtivos na área siderúrgica? O modelo BiO-MCDEA adotado permitiu a identificação das quatro variáveis relevantes (significativas) de entrada e saída para cada processo produtivo, o que facilita a tarefa dos gestores em atuar na melhoria do desempenho destes processos, ajustando os valores destas quatro variáveis conforme Tabela 6.
- Como validar estatisticamente os resultados da OvSMC quando aplicada na análise de eficiência de processos produtivos na área siderúrgica? Os resultados obtidos pela OvSMC, para as duas funções empíricas investigadas, foram validados estatisticamente, utilizando-se uma comparação dos dados reais (banco de dados da Empresa) com os intervalos de

confiança (95%) gerados pela OvSMC para as variáveis respostas “Número de Trincas Externas” e Número de Trincas Internas”.

Outro comentário importante se refere a constatação que realizar a OvSMC utilizando toda a informação gerada pela análise de regressão (coeficientes originais e os intervalos de confiança), bem como a inserção da correlação entre as variáveis de entrada, mostrou-se interessante pois possibilitou melhorar a qualidade dos resultados da otimização, e também permitiu a sua validação estatística.

Optou-se pela estratégia de considerar os coeficientes das funções empíricas como sendo variáveis aleatórias (com função de distribuição de probabilidade triangular), o que se mostrou interessante para resolver problemas complexos e reais. Destaque-se, também, que o algoritmo OLS gerou funções empíricas com boa aderência à realidade dos processos produtivos analisados.

Finalmente, pode-se concluir que todos os objetivos propostos neste trabalho foram plenamente atendidos, e que o procedimento proposto é exequível, inovador e pode propiciar informações de muita valia para os gestores de processos produtivos.

Para futuros trabalhos pretende-se:

- Aplicar a abordagem propostas em outros tipos de indústrias.
- Aplicar a abordagem em outros tipos de problemas, como, por exemplo, econômicos e financeiros.

AGRADECIMENTOS

FAPESP (Processos FAPESP - 2015/12711-4; FAPESP- 2015/24560-0), CNPq (Processos

CNPq – 306214/2015-6, CNPq - 431758/2016-6), e FAPEMIG (Processo APQ-01188-16).

REFERENCIAS

- [1]. Azadeh, A.; Sheikhalishahi, M.; Koushan, M., 2013. An integrated fuzzy MCDEA-Fuzzy simulation approach for optimization of operator allocation with learning effects in multi products CMS. *Applied Mathematical Modeling*, 37, 9922-9933.
- [2]. Charnes A.; Cooper, W.W; Rhodes, E. 1978. Measuring the efficiency of decision-making units, *Eur. J. Operational Research*, 2, 429-444.
- [3]. Banker, R. D.; Charnes, A.; Cooper, W. W.; Swarts, J.; Thomas, D. A., 1989. An Introduction to Data Envelopment Analysis with Some of its Models and Their Uses. *Research in Governmental and Non-Profit Accounting*, 5, 125-163.
- [4]. Behara, S. R. B.; Kippax, P.; Larson, L.; Morton, D. A. V.; Stewart, P., 2011. Kinetics of emitted mass – A study with three dry powder inhaler devices. *Chemical Engineering Science*, 66, 5284-5292.
- [5]. Bertrand, J. W. M., Fransoo, J. C., 2002. Operations management research methodologies using quantitative modeling. *International Journal of Operations and Production Management*, 22, p. 241-264.
- [6]. Ferraro, M. B.; Giordani, P. A., 2012. Multiple linear regression model for imprecise information. *Metrika*, 75, 1049-1068.
- [7]. Friesner, D.; Mittelhammer, R.; Rosenman, R., 2013. Inferring the incidence of industry inefficiency from MCDEA estimates. *European Journal of Operational Research*, 224, 414-424.
- [8]. Ghasemi, M.; Ignatius, J.; Emrouznejad, A., 2014. A bi-objective weighted model for improving the discrimination power in MCDEA. *European Journal of Operational Research*, 233, 640-650.
- [9]. Ghinato, P. Lições Práticas para a Implementação da Produção Enxuta. EDUCS - Editora da Universidade de Caxias do Sul, 2002.
- [10]. Helene, W., 2013. Rethinking the linear regression model for spatial ecological data. *Ecology*, 94, 2381-2391.
- [11]. Hillier, F. S.; Lieberman, G. J.: Introduction to Operations Research, 8th Edition. McGraw-Hill, 2001.
- [12]. Kandel, T. P.; Laerke, P. E.; Elsgaard, L., 2016. Effect of chamber enclosure time on soil respiration flux: comparison of linear and non-linear flux calculation methods. *Atmospheric Environment*, 141, 245-254.
- [13]. Kazantsev, D. M.; Akhundov, I. O.; Karpov, A. N.; Shwartz, N. L.; Alperovich, V. L.; Terekhov, A. S.; Latyshev, A. V. 2015. Monte Carlo simulation of GaAs (0 0 1) surface smoothing in equilibrium conditions. *Applied Surface Science*, 333, 141-146.
- [14]. Kokonendji, C. C., Kiessé, T. S.; Zocchi, S. S., 2007. Discrete triangular distributions and non-parametric estimation for probability mass function. *Journal of Nonparametric Statistics*, 19, 241-257.
- [15]. Korytkowski, P.; Karkoszka, R., 2016. Simulation-based efficiency analysis of an in-plant milk-run operator under disturbances. *The International Journal of Advanced Manufacturing Technology*, 82, 827-837.
- [16]. Kroese, D. P.; Taimre, T.; Botev, Z. I. *Handbook of Monte Carlo Methods*. John Wiley & Sons, New York, 2011.
- [17]. Li, X.; Reeves, G. R., 1999. A multiple criteria approach to data envelopment analysis. *European Journal of Operational Research*, 115, 507-517.
- [18]. Li, S.; Wang, C., 2012. Analysis for Quick Response Strategy using OptQuest Simulation. *Applied Mechanics and Materials*, 220-223, 263-267.
- [19]. Liu, H.; Wei, G.; Xu, Z.; Liu, P.; Li, Y., 2016. Quantitative analysis of Fe and Co in Co-substituted magnetite using XPS: The application of non-linear least squares fitting (NLLSF). *Applied Surface Science*, 389, 438-446.
- [20]. Longo, S., 2006. Monte Carlo simulation of charged species kinetics in weakly ionized gases. *Plasma Sources Science and Technology*, 15, 5181-5188.
- [21]. Menegon, D.; Nazareno, R. R.; Rentes, A. F. Relacionamento entre desperdícios e técnicas a serem adotadas em um sistema de produção enxuta. XXIII Encontro Nacional de Engenharia de Produção, 2003.
- [22]. Montgomery, D. C. *Design and Analysis of Experiments*. 6th. ed., New York: Wiley, 2005.
- [23]. Mosegaard, K.; Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, 100(B7), 12,431-447.
- [24]. Ramón, J.; Cristóbal, S. A., 2011. Multi criteria data envelopment analysis model to evaluate the efficiency of the Renewable Energy technologies. *Renewable Energy*, 36, 2742-2746.
- [25]. Shapiro, A., 2001. Monte Carlo Simulation Approach to Stochastic Programming. *Proceedings of the 2001 Winter Simulation Conference* B. A.

Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, eds.

[26]. Sueyoshi, T.; Goto, M., 2012. Efficiency-based rank assessment for electric power industry: A combined use of Data Envelopment Analysis (DEA) and DEA-Discriminant Analysis (DA). *Energy Economics*, 34, 634 - 644.

[27]. Wölk, J.; Strey, R.; Heath, C. H.; Wyslouzil, B. E., 2002. Empirical function for homogeneous water nucleation rates. *The Journal of Chemical Physics*, 117, 4954-4960.

[28]. Yang, X. S., 2010. *Engineering optimization. an introduction with metaheuristic applications*. Wiley.

Capítulo 13

HORÁRIO UNIVERSITÁRIO PERSONALIZADO POR META-HEURÍSTICA: UM ESTUDO DE CASO

Rosalvo Ferreira de Oliveira Neto

Caique César Leite Mariano

Max Santana Relemberg Farias

Resumo: Este estudo busca auxiliar estudantes que não estão regulares em seus cursos a construírem seu horário levando em consideração suas restrições no curso. A proposta tem como objetivo maximizar o horário do aluno levando em consideração o histórico de disciplinas cursadas e a grade de ofertas disponível. O estudo utilizou como meta-heurística o Algoritmo Genético e Colônia de Formigas para verificar a eficácia da proposta. A validação do estudo foi realizada com uma base de dados disponibilizada pela Universidade Federal do Vale do São Francisco (UNIVASF) para os alunos do curso de Engenharia da Computação. O teste t-Student emparelhado unicaudal mostrou que a proposição de um horário personalizado a partir de uma meta-heurística é estatisticamente superior a um horário montado pelo aluno de forma manual com o nível de confiança de 95%.

Palavras Chave. Problema de Programação de Horários, Meta-heurística, Otimização. Tópicos: PO na Educação; Meta-heurísticas; Otimização Combinatória.

1. INTRODUÇÃO

Ao longo de décadas a comunidade científica busca encontrar uma solução computacional para o problema de agendamento de horário acadêmico, que é a tarefa de criar horários exequíveis e válidos com base em algumas restrições. Acredita-se que este problema é NP-Completo, ou seja, que não é possível encontrar em tempo computacional razoável ou viável uma solução ideal [Cooper e Kingston, 1996]. É por isso que muito dos esforços apresentados na literatura têm sido feito utilizando meta-heurística, que é o nome dado a uma classe de algoritmos aproximativo com algum mecanismo que lhe fornece a capacidade de escapar de ótimos locais. Estes mecanismos de escape em sua grande maioria utilizam algum grau de aleatoriedade para encontrar soluções satisfatórias para um problema, no entanto, sem garantia de otimalidade [Glover e Kochenberger, 2003].

O horário acadêmico de um universitário é elaborado durante a matrícula. No entanto, a maioria das universidades brasileiras buscam oferecer nas matrículas disciplinas que satisfazem os alunos regulares. Ou seja, alunos que nunca perderam disciplinas e que poderão cursar todas as disciplinas ofertadas para o seu semestre sem choque de horário. Esse processo de matrícula funciona bem para alunos regulares, mas na prática, a maior parcela dos discentes dos cursos universitários não é considerada regular.

Durante o processo de matrícula, que ocorre a cada semestre, os discentes devem efetuar matrícula nas disciplinas que pretendem cursar respeitando pré-requisitos e choque de horários, no entanto, existe a possibilidade dos alunos construírem suas grades de horários sem respeitar necessariamente a grade curricular do curso, seja por motivos pessoais ou devido às reprovações. Embora esse processo de matrícula seja um procedimento fácil e bem compreendido pela maioria dos discentes, ele passa a ser muito difícil para os alunos que não acompanham as disciplinas de um bloco semestral do curso, pois, eles não conseguem encaixar os blocos de disciplinas, ofertadas na matrícula, de forma que permita avançar no curso. É por conta dessa dificuldade, que alguns alunos acabam construindo seus horários sem algumas disciplinas e prolongando o seu tempo de permanência na universidade. Uma

alternativa seria a abordagem utilitarista no momento da construção da oferta de disciplinas, ou seja, construir a grade de horários de forma a beneficiar a maior quantidade de alunos possíveis, no entanto, esta abordagem seria inviável por parte das universidades, pois as instituições de ensino superior devem ofertar as disciplinas para que o curso seja concluído dentro do prazo que foi estabelecido pelo ministério da educação, não importando se apenas um aluno é regular.

Nesse contexto, esta pesquisa identificou uma oportunidade para minimizar este problema, que é a sugestão de grade de horário personalizada para o aluno, a partir da grade de disciplinas ofertada pela universidade e de seu histórico escolar utilizando meta-heurística. Dessa forma, a abordagem proposta pode favorecer o aluno fazendo com que ele aproveite melhor seus horários e permita que este possa concluir o curso de forma a amortizar os atrasos causados por eventuais problemas.

Para demonstrar o potencial da solução proposta, construção de horários universitários personalizados, foi utilizado como estudo de caso os horários dos alunos matriculados no curso de Engenharia de Computação, da Universidade Federal do Vale do São Francisco (Univasf), onde cerca de 90% dos discentes ativos não acompanham as disciplinas de um bloco semestral, oferecido pela coordenação do curso. A solução proposta fornece uma grade de horário mais eficiente para o aluno, como será discutido no fim deste artigo. O restante do artigo está dividido como segue. A seção 2 apresenta a definição do problema. A seção 3 apresenta os trabalhos relacionados. A seção 4 detalha a modelagem do problema. A seção 5 exibe a metodologia experimental adotada no estudo. A seção 6 apresenta os resultados obtidos para validação da solução proposta. Por fim, a seção 7 conclui o trabalho e propõe trabalhos futuros.

2. DEFINIÇÃO DO PROBLEMA

O problema de montagem de grade de horário personalizada por aluno, assim como outros problemas de arranjo e combinação, pode ser considerado um problema de otimização do tipo timetabling. Esse tipo de

problema, em geral, não se conhece uma solução algorítmica eficiente que possa resolvê-lo em tempo computacional viável devido a sua complexidade. Para o problema de escolha de grade de horários, no contexto do timetabling, podemos dizer que cada aluno representa um agente que deve organizar suas disciplinas (eventos) em t horários válidos.

Em nosso estudo de caso, os alunos precisam se matricular em disciplinas obrigatórias, optativas e/ou eletivas de sua preferência, sendo restringidos apenas pelos pré-requisitos, choques de horários e o limite de 495h de aula (40h de aula por semana). O curso de Engenharia da Computação da Univasf possui 3.930h obrigatórias, sendo que, 1.350h (cerca de 34 disciplinas) podem ser ofertadas por outros cursos de engenharia da própria instituição (ciclo básico das engenharias).

A partir destas informações, é possível expressar o problema de elaboração de horário personalizado para o curso de Engenharia da Computação da Univasf como sendo um problema de otimização, onde o objetivo é maximizar uma função f que depende da quantidade de horas da grade de horário, da quantidade de disciplinas selecionadas e a quantidade de pré-requisitos. Essa função está sujeita às seguintes restrições:

- A carga horária total não deve exceder 495 horas;
- Não pode existir conflito entre disciplinas (choque de horários);
- Não é possível cursar uma disciplina sem antes ter cursado todos os seus pré-requisitos.

3. TRABALHOS RELACIONADOS

Diversas abordagens de meta-heurística têm sido aplicada com sucesso para problemas de agendamento educacional. Dentre os algoritmos aplicados a esse problema podemos citar: Colônia de Abelhas com Hill Climber [Bolaji et al., 2014], Tabu Search [Di Gaspero e Schaerf, 2001], Simulated Annealing [Basir et al., 2013], Colônia de Formigas (CF) [Socha et al., 2003] e Algoritmo Genético (AG) [Pillay e Banzhaf, 2010]. As abordagens propostas encontradas na literatura visam solucionar dois problemas:

agenda de horários de provas e agenda de horários de curso.

O problema de horários de provas pode ser definido como a atribuição de um número de provas a um dado número de períodos de tempo sujeito às restrições, como por exemplo, exames de alunos comuns não podem ser agendados ao mesmo tempo. Uma solução factível é uma agenda em que todos os exames sejam atribuídos a intervalos de tempo viáveis sem violar as restrições.

O problema do horário de curso universitário pode ser definido como atribuir um determinado número de cursos a um determinado número de intervalos de tempo e salas sujeitos às restrições: 1) nenhum professor pode ser designado para mais de um curso ao mesmo tempo, 2) o número de alunos que frequentam o curso deve ser menor ou igual à capacidade da sala e 3) não mais do que um curso é permitido em um mesmo intervalo de tempo em cada sala.

A seguir serão descritos seis trabalhos encontrados na literatura que utilizam meta-heurística para solucionar o problema de agendamento educacional.

Ciscon et al. [2008] apresentam em seu trabalho uma proposta de solução para o problema de geração de grade de horários. Além dos objetivos e restrições usuais para esse tipo de problema, é levado em consideração a eliminação de janelas (horários vagos em um determinado dia) e aulas isoladas, que segundo os autores é um dos principais motivos da falta de aceitação por parte dos professores para as grades de horários. A técnica usada para resolver esse problema foi AG e para representar o problema foi usada uma estrutura de dados tridimensional (número de turmas X número de dias X número de horários). De acordo com os autores, os resultados obtidos foram satisfatórios e foi possível atestar a capacidade do algoritmo de eliminar janelas e aulas isoladas, além disso, os horários gerados satisfaziam a diretoria, os professores e reduzia o tempo gasto para gerar o mesmo de forma manual.

Hamawaki et al. [2005] aplicaram AG para resolver o problema da grade de horários. O objetivo foi a geração de soluções visando obter grades de horários com múltiplos cursos para uma instituição de ensino. De acordo com os autores, é possível obter uma

melhor qualidade das soluções com uma quantidade grande de indivíduos por geração e utilizando 500 gerações, porém o custo computacional é elevado, desta forma uma análise de custo benefício seria necessária.

Silva et al. [2014] apresentam em seu artigo uma solução para grades de horários com múltiplos cursos onde a obtenção da solução é feita através do uso de um AG. Segundo os autores, horários válidos são obtidos. No entanto, quando populações pequenas são utilizadas, o algoritmo apresenta soluções inválidas, o que pode ser esperado uma vez que o espaço de busca é menor para populações pequenas. Simulações com 300 indivíduos e 300 gerações apresentaram soluções onde não existiam colisões de docentes/salas.

Cassemiro et al. [2014] apresentam um modelo híbrido para solucionar o problema de construção de grade de horários. Esse modelo híbrido propõe a utilização da Busca Tabu (BT) para encontrar um indivíduo bom para ser colocado na primeira geração de um AG. Duas escolas públicas foram utilizadas no experimento, que foi composto por 3 (três) métodos diferentes para configurar a população inicial do AG. A primeira iniciava a população aleatoriamente, a segunda iniciava com um indivíduo obtido através de uma BT e a última usava a grade de horário da escola que é considerado uma solução boa como um indivíduo a ser inserido na solução inicial. Os autores concluíram que nos casos em que era inserido na população inicial um bom indivíduo, as soluções encontradas pelo AG foram consideravelmente melhores. Outra solução híbrida foi proposta por Sá et al. [2015]. Em sua proposta, a meta-heurística BT foi utilizada no ciclo evolutivo do AG para otimizar o melhor indivíduo de cada geração.

Cirino et al. [2015] apresentam dois métodos para alocação de disciplinas em salas. O primeiro é um modelo de programação inteira e o segundo utilizando a meta-heurística AG. O estudo de caso foi realizado com uma base de dados fornecida pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (ICMC - USP). Os resultados obtidos

mostraram que as duas abordagens proporcionam soluções promissoras. No entanto, os autores concluíram que a meta-heurística funciona melhor para instâncias maiores do problema e que o modelo de programação inteira obtém melhores resultados para instâncias menores do problema.

De acordo com o levantamento bibliográfico realizado nesta pesquisa, não foram encontradas propostas de trabalho semelhantes ao nosso artigo, grades de horários personalizadas por alunos. No entanto, não é possível afirmar que não existem propostas, pois a quantidade de trabalhos acadêmicos nesta área é extenso. Uma afirmação plausível seria de que é difícil encontrá-los, como foi o caso desta pesquisa que não encontrou artigos com esse propósito. Outro dado que o levantamento bibliográfico evidencia é que o algoritmo genético tem sido uma das principais heurísticas aplicada nesta área.

4. MODELAGEM DO PROBLEMA

Para a proposta de oferta de grade de horários personalizada este trabalho selecionou duas meta-heurísticas: AG e CF. A escolha do AG foi pelo fato de ser amplamente utilizado com sucesso para problemas de agendamento educacional. O CF foi escolhido por identificarmos que o mesmo apresenta uma estrutura de dados baseada em grafos que melhor representa o espaço de busca para o problema de oferta de grade de horário personalizada. Esta seção apresenta os detalhes da modelagem realizada neste estudo. Inicialmente é apresentada a função objetivo que foi definida para avaliar a grade de horário de cada aluno, em seguida é apresentado como o problema foi modelado nas duas meta-heurísticas selecionadas.

4.1. FUNÇÃO OBJETIVO

A qualidade da grade de horário para cada aluno é definida pela maximização da função objetivo descrita na equação (1).

$$\alpha \sum_{i=1}^N CH_i * X_i + \beta \sum_{i=1}^N CP_i * X_i - \gamma \sum_{i=1}^N OP_i * X_i + \delta \left(\sum_{i=1}^N X_i \right) - MP + NCO * CO \quad (1)$$

Onde:

- N é o número de disciplinas que foram ofertadas e o discente pode cursar;
- CHi é a carga horária da disciplina i;
- CPi é a carga horária que a disciplina i pode desbloquear nos semestres seguintes;
- Xi é 1 se a disciplina está na grade de horários e 0 se não está;
- OPi é 1 se a disciplina é optativa e 0 se não for;
- MP é a quantidade máxima de disciplinas do mesmo período;
- NCO é 1 se existir alguma disciplina matriculada sem seu co-requisito, e 0 caso contrário;
- CO é uma constante punitiva em caso de existir alguma disciplina matriculada sem seu co-requisito tornando a solução infactível;
- α , β , γ e δ são variáveis de ponderamento de seus respectivos termos.

Esta função leva em consideração a carga horária de cada grade proposta e também as restrições de disciplinas. A função está sujeita às seguintes restrições: 1) A carga horária total não deve exceder 495 horas, 2) Não pode existir conflito entre disciplinas (choque de horários) e 3) Não é possível

cursar uma disciplina sem antes ter cursado todos os seus pré-requisitos. Logo, disciplinas que são pré-requisito devem ter um peso maior, pois não incluí-las na grade de horários pode bloquear o curso para o aluno no futuro, conforme análise apresentada na seção de definição do problema.

4.2. MODELAGEM DO ALGORITMO GENÉTICO

O Algoritmo Genético (AG) é um algoritmo de busca inspirado na teoria da evolução das espécies [Holland, 1992]. A Figura 1 ilustra o fluxograma de execução do AG. Inicialmente, é gerada uma população inicial, onde cada indivíduo é representado por um cromossomo, e este representa uma possível solução para o problema [Davis e Mitchell, 1991]. Em seguida, cada cromossomo é avaliado por uma função de aptidão e são selecionados os indivíduos mais aptos ao problema. Por fim, os indivíduos selecionados são combinados uns com os outros para gerar filhos que evoluem durante sucessivas gerações até atingir um critério de parada. Cada cromossomo descendente pode ou não sofrer mutação.

Figura 1: Fluxograma de Execução do AG

Para o problema definido nesta pesquisa, cada cromossomo foi representado em uma lista. Cada item da lista representa uma

disciplina que pode ser cursada pelo aluno. A Figura 2 ilustra como o problema foi modelado para o AG.

Figura 2: Modelagem do Problema no AG

As disciplinas ofertadas pelo curso e as disciplinas cursadas pelo aluno são utilizadas para gerar as Disciplinas Admissíveis, que são as disciplinas que o aluno pode cursar. A partir das Disciplinas Admissíveis é gerado a população inicial limitando cada cromossomo a ter uma carga horária máxima de 495 horas (limite estabelecido pelo curso).

A aptidão de cada cromossomo é igual ao valor da função objetivo para o problema modelado. O método de seleção adotado foi o torneio, onde são selecionados de forma aleatória dois indivíduos e o que possui maior valor de aptidão é selecionado para cruzamento. O cruzamento é aplicado com uma dada probabilidade denominada taxa de cruzamento. O método escolhido para o cruzamento dos indivíduos foi o cruzamento de ponto único. Se o cruzamento é aplicado,

os pais trocam a parte final de seus cromossomos gerando dois filhos, caso contrário os dois filhos serão cópias exatas dos pais.

4.3. MODELAGEM DO ALGORITMO COLÔNIA DE FORMIGAS

O Algoritmo Colônia de Formigas (CF) é um algoritmo de busca baseado na maneira como as formigas se comunicam indiretamente durante a procura por alimento [Dorigo e Di Caro, 1999]. Essa comunicação indireta é realizada através das trilhas de feromônio que são formadas enquanto a formiga procura por alimento. A Figura 3 ilustra o fluxograma de execução do CF. O CF utiliza grafos para representar o problema e os caminhos nos grafos representam possíveis soluções ao problema.

Figura 3: Fluxograma de Execução do CF

Inicialmente, é definido o tamanho da população de formigas para o algoritmo. Cada formiga é aleatoriamente colocada em um nó do grafo. A formiga vai escolher o caminho no grafo de acordo a probabilidade de cada vértice. As probabilidades associada a cada vértice são atualizadas pela taxa de feromônio, que avalia de acordo com a função de aptidão o quão boa é aquela solução. Desta forma, novas formigas tendem a ter mais chances de escolher o melhor caminho encontrado (solução incumbente). No entanto, o algoritmo também prevê uma taxa de evaporação do feromônio, que

diminui a probabilidade de escolha do caminho no decorrer das iterações.

Para o problema definido nesta pesquisa, cada nó do grafo representa uma disciplina ofertada, na qual contém os seu pré-requisitos, sua carga horária, seu horário e também a informação de quais disciplinas ela é pré-requisito. O grafo para cada aluno é formado apenas pelas Disciplinas Admissíveis. A Figura 4 ilustra como o problema foi modelado para o CF. Cada caminho do grafo representa uma solução possível ao problema, ou seja, uma agenda de horário para o aluno.

Figura 4: Modelagem do Problema no CF

5. METODOLOGIA EXPERIMENTAL

Um estudo experimental com alunos do curso de Engenharia da Computação da Univasf foi realizado para validar a eficácia da geração da grade de horários personalizada. Foram selecionados 97 alunos do curso que realizaram matrícula de forma manual pelo sistema da universidade (SIG@) no primeiro semestre de 2016. O critério de seleção foi o aluno está há mais de dois semestres matriculado no curso, faltar mais de dois semestres para se formar e não ter sido aprovado em todas as disciplinas até o momento. São necessários 10 semestres para concluir o curso. O objetivo deste critério de

seleção foi excluir amostras em que o aluno não teria dificuldade em escolher as disciplinas. O estudo compara a carga horária das matrículas realizadas pelos alunos contra a carga horária proposta pela meta-heurística.

As meta-heurísticas AG e CF foram selecionadas para o estudo. Um estudo experimental foi realizado para identificar quais os melhores parâmetros para cada meta-heurística e foram identificados os parâmetros que estão descritos na Tabela 1. Os melhores valores encontrados para os parâmetros da função objetivo foram: $\alpha = 2$, $\beta = 1$, γ e $\delta = 50$.

Tabela 1: Parâmetros Usados no AG e CF

Técnica	Parâmetros
Algoritmo Genético	População: 50 indivíduos Número máximo de gerações: 1500 Taxa de crossover: 90% Taxa de mutação: 2%
Colônia de Formigas	Feromônio inicial: 10000 Quantidade máxima de iterações: 200 Quantidade de formigas: 10 Taxa de evaporação do feromônio: 1%

A comparação foi realizada através da diferença entre as duas cargas horárias. O teste t-Student emparelhado unicaudal foi aplicado para averiguar se existe diferença estatisticamente significativa na média da diferença. A configuração do teste utilizada neste estudo é descrita a seguir: Hipótese nula: $\mu_1 - \mu_2 = 0$ e Hipótese alternativa: $\mu_1 > \mu_2$, onde: μ_1 representa a média da carga horária utilizando uma meta-heurística e μ_2 representa a média da carga horária utilizando a grade de horários realizada manualmente pelo aluno.

6. RESULTADOS E DISCUSSÕES

As simulações foram realizadas de acordo com a configuração experimental descrita anteriormente para cada uma das meta-heurísticas selecionada. A Figura 5 ilustra o gráfico com a diferença em horas entre a grade de horário sugerida pelo AG e a carga de horário realizada na matrícula pelo aluno. Como pode ser observado, dos 97 alunos selecionados para o estudo apenas 4 realizaram uma matrícula com carga horária maior do que a sugerida pela meta-heurística e em 5 casos foram iguais as cargas horárias. No entanto, em 88 casos a carga horária sugerida pela abordagem proposta foi superior.

Figura 5: Comparativo AG vs Aluno

A Tabela 2 exibe o resumo dos resultados obtidos no teste t-Student emparelhado. A diferença média entre a carga horária sugerida pelo AG foi de 144 horas. Uma vez que o p-value é menor que 0.05, concluímos que a carga horária sugerida pelo AG fornece resultados diferentes da carga horária de uma

matrícula manual. Especificamente, os dados indicam que a carga horária proposta pelo AG é, em média, maior do que a carga horária de uma matrícula realizada de forma manual pelo aluno com um nível de confiança de 95%.

Tabela 2: Resumo dos Resultados do Teste de Hipótese

μ_1	μ_2	Limite Inferior	$\mu_1 - \mu_2$	Limite superior	p-value
AG	Aluno	125.555	144	∞	2.2e-16
CF	Aluno	140.4195	159	∞	2.2e-16
CF	AG	9.12762	15	∞	3.263e-05

A Figura 6 ilustra o gráfico com a diferença em horas entre a grade de horário sugerida pelo CF e a carga de horário realizada na matrícula pelo aluno. Como pode ser observado, dos 97 alunos selecionados para o estudo apenas 1 realizou matrícula com

carga horária maior do que a sugerida pela meta-heurística e em 5 casos foram iguais as cargas horárias. No entanto, em 91 casos a carga horária sugerida pela abordagem proposta foi superior.

Figura 6: Comparativo CF vs Aluno

A diferença média entre a carga horária sugerida pelo CF foi de 159 horas. Uma vez que o p-value é menor que 0.05, concluímos que a carga horária sugerida pelo CF fornece resultados diferentes da carga horária de uma matrícula manual. Especificamente, os dados indicam que a carga horária proposta pelo CF é, em média, maior do que a carga horária de uma matrícula realizada de forma manual pelo aluno com um nível de confiança de 95%.

A Figura 7 ilustra o gráfico com a diferença em horas entre a grade de horário sugerida pelo CF e o AG. Como pode ser observado, em 6 casos o AG propôs uma carga horária superior ao CF. Não houve diferença em 54 casos, e em 37 casos a carga horária

proposta pelo CF foi superior ao AG. Uma vez que o p-value é menor que 0.05, concluímos que a carga horária sugerida pelo CF fornece resultados diferentes da carga horária sugerida pelo AG.

Os dados apresentados indicam que a carga horária proposta pelo CF é, em média, maior do que a carga horária sugerida pelo AG com um nível de confiança de 95%. Outro fator importante observado nos experimentos foi a diferença do custo computacional entre o AG e CF. O tempo de execução dos experimentos com todos os 97 alunos selecionados foi de 5 horas e 25 minutos para o AG e 3 minutos para o CF.

Figura 7: Comparativo CF vs AG

7. CONCLUSÃO

Este artigo apresentou uma abordagem de proposta de grades de horários personalizada para cada aluno a partir dos horários disponibilizados pelos colegiados acadêmicos. A abordagem leva em consideração o histórico escolar do aluno e também a importância de cada disciplina para a conclusão do curso dentro de um prazo menor. A proposta foi testada utilizando as meta-heurísticas AG e CF. O estudo experimental mostrou que a abordagem proposta produz uma grade de horário melhor, em média, do que a grade construída pelo aluno de forma manual.

A principal contribuição deste estudo é abordar o problema de construção de horários utilizando a perspectiva do aluno, uma vez que a perspectiva da gestão da instituição de ensino superior tem sido amplamente estudada (recursos = salas disponíveis x horários dos professores x disciplinas). Nesta linha de raciocínio, este trabalho inclui a outra parte interessada, o aluno, na pergunta de pesquisa "Como montar uma grade de horário eficiente?".

A partir da abordagem proposta neste estudo, o aluno pode reduzir o tempo de permanência

no curso ao priorizar a realização de matrículas em disciplinas que são pré-requisitos e sem deixar de maximizar a sua carga horária no semestre. Outra contribuição deste estudo é que ele indica que a utilização do CF pode ser mais eficiente em termos de custo computacional do que a do AG, obtendo resultados melhores como foi observado no experimento. Uma explicação plausível para essa maior eficiência computacional do CF em relação ao AG é a geração de grade de horário não válida quando são aplicados os operadores de cruzamento e mutação do AG, pois só é possível verificar se o horário é viável após a aplicação destes operadores, enquanto no CF, que utiliza um grafo completamente conectado, a medida que a formiga percorre o grafo são removidos os nós que podem gerar choque de horários com as disciplinas que foram percorridas, desta forma, o CF não gera soluções infactíveis.

Como trabalho futuro, pretendemos disponibilizar a abordagem proposta como uma ferramenta WEB permitindo assim que um número maior de alunos possam utilizar essa proposta.

REFERÊNCIAS

- [1] Basir, N., Ismail, W., e Norwawi, N. M. (2013). A simulated annealing for tahmidi course timetabling. *Procedia Technology*, 11:437 – 445.
- [2] Bolaji, A. L., Khader, A. T., Al-Betar, M. A., e Awadallah, M. A. (2014). University course timetabling using hybridized artificial bee colony with hill climbing optimizer. *Journal of Computational Science*, 5(5):809 – 818.
- [3] Cassemiro, M. V., Miranda, D. S., e Wanner, E. F. (2014). Desenvolvimento de um modelo híbrido baseado em algoritmo genético e busca tabu para resolução do problema de quadro de horários escolar. In *Anais do XLVI SBPO*, p. 1870–1878.
- [4] Cirino, R. B. Z., Santos, M. O., e Delbem, A. C. B. (2015). Aplicação da metaheurística agc para o problema de alocação de aulas à salas. In *Anais do XLVII SBPO*, p. 1745–1755.
- [5] Ciskon, L. A., Oliveira, A. C., Hipólito, T. R., Alvarenga, G. B., e Roullier, A. C. (2008). O problema de geração de horários: Um foco na eliminação de janelas e aulas isoladas. In *Anais do XXVII SBPO*, p. 1724–1732. SOBRAPO.
- [6] Cooper, T. B. e Kingston, J. H. (1996). The complexity of timetable construction problems. In *Selected Papers from the First International Conference on Practice and Theory of Automated Timetabling*, p. 283–295. Springer-Verlag.
- [7] Davis, L. D. e Mitchell, M. (1991). *Handbook of Genetic Algorithms*. Van Nostrand Reinhold.
- [8] Di Gaspero, L. e Schaerf, A. (2001). Tabu search techniques for examination timetabling. In Burke, E. e Erben, W., editors, *Practice and Theory of Automated Timetabling III*, volume 2079, p. 104–117. Springer Verlag, Berlin-Heidelberg, Germany.
- [9] Dorigo, M. e Di Caro, G. (1999). New ideas in optimization. chapter *The Ant Colony Optimization Meta-heuristic*, p. 11–32. McGraw-Hill Ltd., UK.
- [10] Glover, F. e Kochenberger, G. A. (2003). *Handbook of metaheuristics*. Kluwer Academic Publishers.
- [11] Hamawaki, C. D. L., Yamanaka, K., Oliveira, R. C., e Hamawaki, O. t. (2005). Alocação de grade horária em instituições de ensino superior utilizando algoritmos genéticos. In *Anais do IV SIRC/RS*.
- [12] Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. MIT Press, Cambridge, MA, USA.
- [13] Pillay, N. e Banzhaf, W. (2010). An informed genetic algorithm for the examination timetabling problem. *Applied Soft Computing*, 10(2):457 – 467.
- [14] Silva, A. B., Betemps, C. M., e Heinen, M. R. (2014). Algoritmos genéticos na obtenção de uma grade de horários com múltiplos cursos para uma instituição de ensino. In *Anais do V Encontro Anual de Tecnologia da Informação (5o EATI)*, p. 239–246.
- [15] Socha, K., Sampels, M., e Manfrin, M. (2003). *Ant Algorithms for the University Course Timetabling Problem with Regard to the State-of-the-Art*. Springer Berlin Heidelberg.
- [16] Sá, M. A. C., Wanner, E. F., e Sarubbi, J. F. M. (2015). Quadro de horários escola: Uso de algoritmo memético para acelerar a convergência do método evolutivo. In *Anais do XLVII SBPO*, p. 1958– 1969.

Capítulo 14

DETECÇÃO DE MINAS TERRESTRES VIA ANÁLISE MULTIVARIADA DE IMAGEM

Leandro Vitral Andraos

Paulo Henrique Coelho Maranhão

Resumo: Apesar dos avanços tecnológicos, as ações de minas terrestres ainda afetam a vida de muitas pessoas e a busca de métodos eficientes para a detecção dessas minas torna-se um tema relevante. Dessa forma, este trabalho propõe a técnica de Análise Multivariada de Imagens para a detecção de minas terrestres usando imagens infravermelha (IR). Adicionalmente, é proposto o monitoramento dessas minas por meio do gráfico de controle T^2 de Hotelling. Resultados experimentais mostram que os métodos propostos foram eficientes tanto para a detecção, quanto para o monitoramento de minas terrestres.

Palavras Chave. Mina terrestre, detecção, análise multivariada de imagens.

1. INTRODUÇÃO

A utilização de imagens térmicas infravermelhas (IR) para a detecção de minas terrestres tornou-se um tema de grande interesse nos últimos tempos. De acordo com os dados do Serviço de Ação de Minas (Nações Unidas, 2014), dez pessoas são mortas ou mutiladas por uma mina terrestre a cada dia e existe, ainda hoje, uma mina para cada 17 crianças no mundo, apresentando um sério risco para a população. Segundo Hong et al. (2001), estima-se que ainda existam pelo menos 100 milhões de minas terrestres espalhadas por mais de 60 países, o que torna a busca de métodos mais eficientes para detecção de minas um tema ainda atual e importante. Além disso, de acordo com o monitoramento global de munição e minas terrestres (Landmine Monitor, 2015), 61 países foram classificados como afetados pelas minas no ano de 2015.

O princípio da termografia deixa claro que a assinatura do solo é alterada pela presença de objetos superficialmente enterrados. Lopez et al. (2003) relaciona a detecção de minas terrestres com suas propriedades térmicas, já que isso resulta em perturbações do padrão térmico esperado, podendo então ser medidas por sensores IR. Tal detecção vale para cada tipo de mina e mesmo se houver pouca quantidade de metal, sua identificação pode ser realizada.

A utilização da termografia IR para a detecção de minas terrestres enterradas é uma aplicação recente, embora a teoria básica tenha sido desenvolvida por Watson (1975). Além disso, a utilização de imagens infravermelhas térmicas para detectar minas terrestres baseia-se no pressuposto de que as minas têm propriedades térmicas diferentes do solo onde está inserida, (Janssen et al., 1996). Muito embora, sua aplicação esteja restrita no período do por do sol até o nascer do sol, pois nesse momento, solo e mina estarão em equilíbrio térmico, não criando assim um gradiente de temperatura na superfície da terra.

DePersia et al. (1995) foram percussores em trabalhos que relataram assinaturas térmicas em áreas com minas terrestres. As diferenças de temperatura entre a superfície do solo e a mina são atribuídas a diferenças em suas capacidades de calor, bem como a perturbação da superfície durante sua instalação. Estudos realizados por Simunek et

al. (2001) avaliaram como a textura do solo e a água afetam as assinaturas térmicas de minas terrestres. Já Messelink et al. (2002) utilizou classificadores baseados em distância de Mahalanobis e Fisher para encontrar objetos de forma circular baseadas em imagens infravermelhas.

Com o desenvolvimento da teoria de problemas inversos, López et al. (2003) realizaram uma análise não destrutiva interna do solo, a fim de detectar a presença de minas enterradas. Já Sendur e Baertlein (2009) descreveram técnicas baseadas em sensores e em processamento de sinais para melhorar a detecção de minas com o uso de imagens de infravermelhos térmicos. Em Macdonald e Lockwood (2003) é discutida a questão do detector de metais ainda ser o dispositivo popular mais utilizado na detecção de minas terrestres metálicas, no entanto é pouco útil no caso mais comum, mina não metálica (PVC e polietileno). Estudos realizados por Mende et al. (2004) investigaram como a absorção de temperatura cria um gradiente de concentração térmico na superfície do solo.

Atualmente novas investigações em detecção de minas estão sendo avaliadas. Estudos baseados em radares holográficos (métodos que discriminam alvos de objetos enterrados) são conduzidas por Zea e Cadernas (2013); Szymanik (2014) introduziu um método novo envolvendo termografia infravermelha ativa com microondas; Song et al (2016) intercalou radares holográficos com simulação numérica e Hameed (2016) desenvolveu algoritmos capazes de escanear sistematicamente um campo de minas através de robôs. Esse algoritmo permite um planejamento sistemático dos movimentos dos robôs para detectar as minas terrestres. A trajetória permite que o robô escaneie a área minada no menor tempo possível, evitando qualquer mina terrestre ao longo do processo. Permite também que o robô funcione de forma totalmente autônoma com mínima ou nenhuma intervenção humana, reduzindo drasticamente os riscos de ferimentos no local de trabalho e maximizando a eficiência da operação.

A desminagem de terrenos é, ainda, uma tarefa árdua e perigosa com um custo de remoção excessivamente alto e de acordo com Smith (2014) os métodos tradicionais de detecção de minas são pouco eficazes e

perigosos, assim estudos vêm sendo realizados para aperfeiçoar as técnicas existentes ou implementar novas metodologias para melhorar sua qualidade em termos de sua identificação.

Dessa forma, o objetivo deste artigo é avaliar a eficácia de uma técnica de análise estatística multivariada para detecção de minas terrestres. A partir da obtenção de uma série de imagens e da aplicação dessa técnica espera-se classificar e detectar os objetos na imagem. Em seguida, serão desenvolvidos gráficos de controle baseados na estatística T^2 de Hotelling para identificar eventuais divergências nas imagens baseado em um grupo de imagens controle.

O restante do capítulo inclui mais cinco Seções. A segunda apresenta os métodos abordados no trabalho, com algumas definições de processamento e análise multivariada de imagens incluindo as principais características da técnica de componentes principais utilizada. A terceira Seção explica a metodologia abordada e os resultados e discussões levantados no trabalho aparecem na Seção quatro. Por fim, as conclusões são apresentadas na Seção cinco.

2. MÉTODOS

2.1. ANÁLISE MULTIVARIADA DE IMAGEM

Ao longo dos anos o uso da inspeção visual tem sido comum em processos de controle e monitoramento em diversas áreas, e apesar de serem realizadas por pessoas qualificadas para a função, tornou-se um sistema limitado e sujeito a falhas humanas. De acordo com Crósta (1993) o objetivo principal do processamento de imagens é o de remover essas barreiras, inerentes ao sistema visual humano, facilitando a extração de informações a partir de imagens. Em Gonzalez (2000) a definição de imagem é feita como uma função bidimensional $f(x,y)$ em que x e y são coordenadas espaciais e a amplitude f em qualquer par de coordenadas (x,y) é denominado de intensidade da imagem nesse ponto. Dessa forma, uma imagem é composta por um número finito de elementos, com localização e valores específicos, sendo chamado de pixel cada representação do elemento da imagem. Cada pixel poderá conter mais de uma medida, ou seja, diferentes canais de cor representam

distintas medições para um determinado pixel.

Uma técnica muito difundida de análise de imagens em monitoramento de processos é a Análise Multivariada de Imagens (MIA) (Geladi e Grahn, 1996), (Prats-Montalbán et al. 2011). A MIA envolve conjuntamente ferramentas de análise multivariada de dados com processamento de imagens para manipular simultaneamente uma grande quantidade de dados. Assim, a metodologia é baseada em estruturar as imagens a partir de um processamento e então aplicar alguma técnica estatística específica para análise, sendo muito utilizada a análise de componentes principais.

De acordo com Geladi e Grahn (1996) a MIA pode ser aplicada em diversas áreas e em inúmeras imagens produzidas por processos complexos de monitoramento. Além disso, diversos métodos de processamento de imagens foram desenvolvidos e utilizados para extrair características ou informações das imagens e o uso da MIA permite a monitoração e detecção de falhas do processo de forma mais ágil (Bharati e Macgregor, 1998).

Bharati et al. (2004) utilizaram a MIA para uma análise de textura de imagens para avaliação da qualidade de um conjunto de chapas de aço. Moeller et al. (2006) utilizaram imagens satélites da cidade de Phoenix, e a partir da análise multivariada de imagens conseguiram identificar mudanças urbanas, tais como edificações ou alterações na superfície, ao longo do tempo. Gonzalez e Sanchez (2006) estudaram a indústria de embalagens aplicadas a alimentos e medicamentos e utilizando a MIA concluíram que a contaminação em embalagens pode passar despercebida em determinados lotes. Ghasemzadeh et al (2014) utilizaram a análise multivariada de imagens para avaliar as diferentes camadas de películas na composição das embalagens de plástico para evitar ao máximo riscos de contaminações no produto.

Além disso, Kruse et al. (2014) compararam métodos de classificação dos pixels das imagens de folhas expostas à poluição e o uso da MIA resultou na detecção precoce destes sintomas, no monitoramento ambiental e na qualidade dos vegetais. Com isso, Everard et al. (2014) avaliaram imagens de folhagens de espinafre com contaminações

fecais e também propôs o uso da MIA para a detecção precoce destas contaminações. Mais recente, Damasceno et al. (2015) utilizaram a MIA para construir curvas de calibração e assim determinar o pH de águas potáveis.

Outro fato importante da Análise Multivariada de Imagens é sua possibilidade de monitoramento em tempo real. Neste caso, consegue-se identificar irregularidades ou quaisquer outros detalhes antes de se concluir o processo final de produção. Dessa forma, a MIA tem uma grande vantagem em termos de sua aplicabilidade, podendo ser utilizada em diferentes áreas e processos de monitoramento. Uma revisão sobre a aplicação do MIA pode ser encontrada em Duchesne e Macgregor (2012).

Utilizando a MIA, Wise e Geladi (2000) discutem a reorganização de uma matriz original de três ou mais dimensões e através de um processo de "desdobramento" cada pixel se transforma numa linha da matriz. Em outras palavras, uma imagem originalmente formada com $I \times J$ por J pixels com K bandas espectrais é convertida para duas dimensões $I \times J \times K$ por $K \times K$. A partir desta nova matriz utiliza-se a técnica de Análise de Componentes Principais (PCA) para reduzir a dimensão da imagem através das componentes e em seguida gráficos de controle com a estatística T^2 para o

monitoramento e detecção de possíveis distúrbios nas imagens podem ser construídos. O processo será detalhado no item a seguir.

2.2. PROCESSAMENTO DIGITAL DE IMAGEM

As imagens utilizadas no artigo estão no modelo RGB (*Red, Green, Blue*), onde cada cor aparece em seus componentes espectrais primários em tons de vermelho, verde e azul, sendo baseado no sistema de coordenadas cartesianas. Assim, cada imagem apresenta três bandas, uma para cada cor primária e a combinação dessas cores primárias formam as imagens em cores compostas.

Como descrito anteriormente, um passo importante para a construção do modelo PCA é a transformação dos dados originais da imagem numa matriz bidimensional. O desenvolvimento do processo pode ser visto em Prats-Montalbán et al. (2011), Prats-Montalbán e Ferrer (2014). A estrutura dos dados originais das imagens utilizadas são dispostas em tensores tridimensionais (generalização do conceito de matriz), pois apresentam uma camada para cada uma das três bandas de cor. Após a transformação passa-se a trabalhar com uma matriz bidimensional, condição necessária para aplicação das componentes principais, ou seja, os pixels da imagem ficam dispostos em um vetor de apenas uma banda.

Figura 1: Imagem RGB (3 bandas de cor) com a matriz transformada, n pixels e 3 colunas.

O processo consiste, então, em transformar o cubo numa matriz de duas dimensões, ou seja, os eixos relativos à localização são

agregados formando um vetor para cada pixel nas linhas e este é cruzado com o número de bandas nas colunas. Assim, a

reorganização da imagem leva a formação de um vetor de três elementos para cada conjunto de pixels correspondentes nas imagens, como pode ser visto na Figura 1 desenvolvida em Gonzalez e Woods (1992).

2.3. ANÁLISE DE COMPONENTES PRINCIPAIS

Devido à grande combinação de pixels numa imagem, independentemente das escalas de cores, uma técnica muito utilizada na MIA é a Análise de Componentes Principais (PCA). O objetivo deste método é comprimir dados altamente correlacionados e projetá-los num subespaço de dimensões reduzidas através de combinações lineares dos dados originais Bharati e MacGregor (1998). Dessa forma, a estrutura de variância e covariância de um vetor aleatório, composto de p -variáveis aleatórias, serão explicadas a partir das combinações lineares das variáveis originais. Tais combinações são chamadas de componentes principais, também chamadas de variáveis latentes, visto que não podem ser medidas diretamente. Além disso, essas componentes são não correlacionadas entre si.

Ao gerar o modelo PCA, o número de componentes principais será igual ao número de variáveis originais, e tais componentes serão arranjadas por ordem decrescente do seu poder de explicação em relação à variância ou covariância total. A qualidade da aproximação pode ser medida através da avaliação da proporção da variância total explicada por essas combinações. Além disso, a variabilidade do vetor aleatório com p variáveis originais é aproximado com a variabilidade do vetor aleatório que contém as k componentes principais. Maiores

informações e variedades de exemplos podem ser encontrados em Vines (2000) e Jackson (2003).

A vantagem deste método é selecionar as primeiras componentes e dado seu alto grau de explicação, reduzir a dimensão da matriz dos dados fazendo com que poucas componentes expliquem grande parte da variabilidade dos dados originais (Mingoti, 2005). Assim, pode-se utilizar, por exemplo, somente 2 ou 3 componentes, ao invés das 25 variáveis originais obtidas de uma imagem. A técnica PCA será aplicada a todas as imagens envolvidas no estudo.

O modelo PCA é descrito por Mingoti (2005) como:

$$Y_j = a'_j X \quad Y_j = a'_j X \quad (1)$$

Onde X é a matriz de dados originais, j é o número de componentes principais do modelo PCA, a'_j a' é o vetor corresponde aos coeficientes das combinações lineares e Y o vetor correspondente às combinações lineares das variáveis aleatórias do vetor X .

Cada autovalor λ_j (cada componente tem um valor diferente) obtido a partir da matriz de covariância representa a variância de uma componente principal de Y_j . Y_j . Como há uma ordenação dos autovalores em ordem decrescente, a primeira componente é responsável pela maior variabilidade e a j -ésima a de menor. Assim, a proporção de variância total que é explicada pela j -ésima componente pode ser dada por

$$\frac{Var(Y_j)}{Variancia\ total\ de\ X} = \frac{\lambda_j}{\sum_{i=1}^j \lambda_i} \quad (2)$$

Com isso, espera-se que a quantidade de informação contida nas j componentes seja tão alta quanto às contidas nas p variáveis originais, reduzindo assim a dimensão das imagens analisadas.

2.4. CONTROLE ESTATÍSTICO DE PROCESSO

Para estabelecer o modelo de controle estatístico de qualidade, primeiramente é necessário identificar as imagens consideradas ideais, sem imperfeições ou anomalias. A partir dessas imagens consideradas perfeitas são determinados os limites de controle, que serão adicionadas aos gráficos de controle das imagens com algum tipo de perturbação (mina e/ou outros objetos).

Em relação às estatísticas de controle, as aplicações mais comumente utilizadas convergem ao uso da estatística de T^2 de Hotelling (conhecida como estatística de Hotelling). Ferramenta muito utilizada junto do

$$T^2 = (X_i - \bar{X})S^{-1}(X_i - \bar{X}) \quad (3)$$

onde os pixels de cada banda são representados por X_i , $\bar{X}$ é a média dos pixels de cada banda e S^{-1} é a matriz de covariância amostral. Os limites superiores de controle são dados por:

$$LSC = \frac{(m-1)^2}{m} B_{\left(1-\frac{\lambda}{2}, \frac{p(m-p-1)}{2}\right)} \quad (4)$$

em que p representa o número de bandas, m o número total de pixels, α o nível de significância e B indicam os quantis correspondentes à distribuição Beta. O nível de confiança adotado foi de 95%. A partir desta estatística gráficos de controles serão construídos com seus respectivos limites para o caso das imagens em controle. Em seguida, tal limite será adicionado às imagens com a inserção de alguns objetos e espera-se que haja uma quantidade maior de pixel acima do LSC, o que significa dizer que as componentes estão captando bem esses objetos. Já o limite inferior de controle (LIC) é definido como zero. Para Mason et. al (1992) isto ocorre porque qualquer mudança na média leva a um aumento na estatística T^2 e assim o LIC pode ser ignorado.

Alguns exemplos da aplicação destes gráficos podem ser vistos em Prats-Montalbán et al. (2011), Kruse et al. (2014) e Prats-Montalbán e Ferrer (2014).

modelo PCA, pois consegue medir o quão distante as observações estão em relação ao centro do espaço de variáveis latentes. A T^2 é obtida da seguinte forma (Montgomery 2004):

3. METODOLOGIA

O experimento consistiu em captar imagens IR (infravermelho) de uma caixa preenchida com areia para diferentes temperaturas. Para medição do contraste térmico no experimento foi utilizada uma câmera de imagem térmica da série T da FLIR. A experimentação foi toda realizada em laboratório com o objetivo de haver um melhor controle dos parâmetros e ter agilidade nos resultados.

As temperaturas foram controladas através de aquecimento por lâmpada ultravioleta. A mesma foi posicionada a 30 cm da caixa na vertical. A lâmpada aquecia a areia até que alcançasse temperatura de 40°C e depois era desligada (Figura 2), o experimento terminava assim que a temperatura decaísse para 28°C. Para acompanhar a temperatura foram usados dois medidores. O primeiro foi um termômetro simples de mercúrio colocado dentro da caixa e o segundo através da medida do sensor térmico da câmera. Ambas as medições de temperatura foram muito semelhantes. Conforme ocorria o esfriamento, as imagens eram captadas. As imagens foram captadas conforme a temperatura decrescia 4°C.

A mina utilizada é o modelo M 409 - *Training & Practice* (material plástico) usada pelas forças armadas brasileiras com finalidade antipessoal, conforme pode ser vista na Figura 3. É uma das minas terrestres mais produzidas e deve ser enterrada a no máximo três centímetros do solo para que possa ser ativada. Sua explosão inicia-se quando alguém pisa no prato de pressão e aciona involuntariamente o detonador. No experimento, a mina foi enterrada na caixa a 2 cm de profundidade.

Figura 2: Foto real da mina utilizada.

Figura 3: Realização do experimento.

Alguns objetos, além da mina, foram incorporados ao experimento para simular as assinaturas de objetos termais semelhantes.

Os objetos foram: Pedra, clipe e moeda (ver Figura 5).

Figura 4: Apenas mina na caixa.

Figura 5: Objetos inseridos junto da mina

As etapas do experimento foram divididas em três partes, todas com o objetivo de simular a ausência ou presença de objetos enterrados. As etapas são descritas da seguinte forma:

- **Etapa 1:** Ausência de objetos e somente areia (imagem perfeita, sem perturbação, usada como controle);
- **Etapa 2:** Presença da mina somente na caixa (Figura 4);
- **Etapa 3:** Inserção de objetos (pedra, clipe e moeda) junto à mina na caixa (Figura 5).

Após a obtenção de todas essas imagens foram utilizados conjuntamente os *softwares* Matlab e R para a construção das componentes principais e o cálculo da estatística T^2 de Hotelling. Então, primeiramente o modelo de componentes principais foi elaborado para a Etapa 1 a fim

de verificar se o processo estava realmente em controle. Dado que o processo estava em controle, construiu-se os modelos para a Etapa 2 e Etapa 3 e em seguida, com as componentes foi possível analisar a assinatura térmica da mina individualmente e posteriormente em conjunto com outros objetos.

4. RESULTADOS

Após a realização de todas as etapas do experimento as imagens nas faixas de temperatura estudadas (28°C, 32°C, 36°C, 40°C) puderam ser obtidas, como pode ser visto na Figura 6. As diferenças de coloração são devido às diferenças de temperatura entre os pixels e a régua de temperatura (vide Figura 7) fornece as faixas alcançadas, com azul para valores menores e tons claros para temperaturas mais altas.

Figura 6: Imagens originais obtidas com a câmera infravermelho.

A Etapa 1 é composta apenas pela areia e apresenta imagens como uma grande massa uniforme, variando pouco conforme o aumento de temperatura. Na temperatura mais baixa o tom avermelhado cobre quase toda a imagem, e conforme o aumento de temperatura começam a surgir alguns pontos com tons mais claros, como pode ser visto nas imagens da primeira linha da Figura 6. Ao adicionar a mina (Etapa 2 - Figura 6) é possível perceber algumas diferenças nas imagens em relação à Etapa 1, mesmo nas temperaturas mais altas. Mas com o decréscimo de temperatura a mina torna-se cada vez mais branca. A 40°C tanto a mina quanto a areia estão quentes logo há menos diferenças visuais em termos de comparação.

Ou seja, dada a diferença de condutividade térmica entre a areia seca e a mina (composição de PVC), esta última absorve melhor a temperatura e é capturada na imagem conforme a temperatura diminui.

Na Etapa 3 tem-se outros objetos adicionados a imagem, porém nem todos são captados pela imagem infravermelha. Na verdade, somente a pedra e a mina podem ser identificadas visualmente, já que a moeda e o clipe de metal não foram captados. A mina continua bem distinguível (mancha mais branca na parte superior esquerda) com uma temperatura mais alta e a pedra com uma temperatura mais baixa passa também a ser identificada (parte inferior esquerda).

Figura 7: Régua de Temperatura.

A partir das imagens da Figura 6 o modelo de componentes principais e os limites superiores de controle (nível $\alpha = 5\%$) puderam ser construídos. Assim, para cada temperatura foi desenvolvido um modelo

diferente. Apesar de todas as imagens terem sido analisadas apenas a de temperatura de 32°C foi adicionada ao artigo como exemplo. A Figura 8 mostra as imagens das três componentes principais obtidas.

Figura 8: Modelo de Componentes Principais para todos os objetos ($T = 32^{\circ}\text{C}$).

A primeira componente principal (CP1) é a de maior explicação e a partir dela parece ser fácil a identificação de cada etapa. Na areia a CP1 explica grande parte da variabilidade, seguida por valores baixos de CP2 e CP3. Com a inclusão da mina a CP1 aumenta ainda mais seus valores e apresenta baixos valores de CP2 e CP3, mostrando que a mina está sendo explicada grande parte pela primeira componente. Já a Etapa 3 (mina e objetos) tem valores menores para CP1, um valor

maior de CP2 e pequeno valor de CP3. Nesse caso a variabilidade é melhor explicada considerando as duas primeiras componentes. Isso ocorreu para todas as opções de temperatura estudadas. O Quadro 1 mostra tais resultados, como efeito de exemplificação, para o caso da temperatura $T = 32^{\circ}\text{C}$. A Figura 8 corrobora os valores das CP's, pois visualmente é perceptível a massa uniforme na areia, a mina na primeira e a pedra na segunda componente.

Quadro 1: Explicação das Componentes (%) para cada etapa do experimento em $T = 32^{\circ}\text{C}$.

Itens\ CP's (%)	CP1	CP2	CP3
Areia	83.80	14.50	1.70
Mina	93.00	6.70	0.30
Mina e objetos	84.20	14.80	1.00

Uma das características mais importantes das componentes principais na análise multivariada de imagem é a redução do armazenamento da imagem, já que uma parcela significativa da variabilidade está contida na CP1 e diminui drasticamente a partir das demais. O motivo é devido aos autovalores, já que indicam as variabilidades em termos de contraste de cada uma das bandas da imagem. Assim, imagens formadas a partir das componentes do vetor correspondente aos maiores autovalores apresentem o maior contraste. Logo, devido a

grande explicação da CP1 pode-se manter somente a principal componente da imagem o que reduz significativamente a espaço de armazenamento e/ou transmissão.

Após a construção e análise das componentes o passo seguinte é analisar o T^2 Hotelling para os pixels e tentar identificar diferenças entre as imagens. O Quadro 2 apresenta o resultado da porcentagem dos valores de T^2 acima do limite superior de controle (LSC) para cada temperatura e etapa de coleta das imagens.

Quadro 2: T^2 Hotelling (%) acima do LSC para os elementos da imagem.

Itens\T (°C)	28	32	36	40
Areia	0.048	0.067	0.058	0.031
Mina	6.094	3.855	6.490	3.675
Mina e objetos	9.3369	10.987	9.563	10.590

Sabe-se que pela teoria dos gráficos de controle T^2 , espera-se que ocorra em torno de 1 descontro para cada 90 pontos analisados, ou seja, 1% de pontos fora de controle. Sendo assim, percebe-se pelo Quadro 2, que o valor baixo do T^2 para areia para cada temperatura se encontra realmente em controle. Ao adicionar a mina o valor da estatística se eleva, já que há um número

maior de pixels acima do LSC, o mesmo ocorrendo quando se coloca outros objetos. Para se observar isto, basta verificar o valor do T^2 no Quadro 2. Gráficos de T^2 que retratam os valores obtidos no Quadro 2 foram construídos para cada temperatura, como exemplificação apenas $T=32^\circ\text{C}$ será mostrado.

Figura 9: T^2 Hotelling – Etapa 1 ($T=32^\circ\text{C}$)

Figura 10: T^2 Hotelling – Etapa 2 ($T=32^\circ\text{C}$).

Figura 11: T^2 Hotelling – Etapa 3 ($T=32^\circ\text{C}$).

Pelo resultado das Figuras 9-11 percebe-se a diferença no valor do T^2 encontrado. Na Etapa

1 (caso controle) os pixels apresentam quase a totalidade dentro do limite de controle. Já ao

adicionar a mina (Etapa 2), a estatística de Hotelling consegue captar sua inserção, pois muitos pontos ultrapassam o LSC. Na Etapa 3, além da mina, o T^2 capta também a inserção da pedra, conforme pode-se verificar na Figura 11. Além disso, outro fato interessante é que o T^2 reflete bem a posição dos elementos dentro da caixa, com a mina do lado esquerdo e a pedra do lado direito.

5. CONCLUSÃO

A existência de minas terrestres é uma preocupação na questão da segurança, tanto para veículos quanto para os indivíduos. Ter uma leitura a priori do terreno permitiria reduzir os riscos de acidentes decorrentes de explosões. Assim, esse trabalho propôs a utilização da Análise Multivariada de Imagens para detecção de minas terrestres com base na captação de imagens com infravermelho térmico. Além disso, para auxiliar na identificação dessas minas foi proposto um esquema de controle baseado no gráfico T^2 de Hotelling.

Analisando os resultados foi verificado a eficácia do método proposto para todas as temperaturas medidas, embora só tenha sido mostrado o resultado para a temperatura de 32°, implicando que a MIA pode ser uma ferramenta importante para a detecção de minas terrestres. Os resultados apontaram ainda para uma grande explicação da primeira componente principal, a qual foi responsável por praticamente toda a variabilidade das imagens, indicando que as

imagens podem ser reduzidas para uma nova dimensão, minimizando tempo e custo de processamento. Verificou-se também que houve uma melhor identificação da mina e dos objetos à medida que a temperatura diminui, uma consequência direta da diferença de condutividade térmica entre os elementos da imagem.

O esquema de controle proposto também foi bem-sucedido, sendo observado que a quantidade de pixels acima do LSC aumenta à medida que novos objetos são inseridos a imagem. Assim, para o caso do experimento considerando somente a areia, o valor do número de pontos fora de controle do gráfico T^2 foi baixo, visto que nesse caso não há perturbação no sistema, e, portanto, o processo é considerado sob controle. No entanto, ao adicionar a mina e outros objetos o número de descontroles aumenta significativamente, indicando a presença dos mesmos.

Por fim, para estudos futuros, uma possibilidade seria criar um sistema de monitoramento online de imagens, pois com a redução dimensional e otimização no processamento dos dados, consegue-se obter processos automatizados que ampliam e agilizam o monitoramento, podendo gerar estatísticas instantaneamente. Outra possibilidade é a de construir outras estatísticas para identificação dos objetos na imagem considerando a análise da forma dos objetos, determinando uma assinatura para os diversos tipos de objetos na imagem.

REFERÊNCIAS

- [1] Bharati, M. and MacGregor, J. (1998). Multivariate image analysis for real-time process monitoring and control. *Industrial & Engineering Chemistry Research*, 37(12):4715-4724.
- [2] Bharati, H, Jay L., e Macgregor, J. (2004) Image texture analysis: methods and comparisons. *Chemometrics and intelligent laboratory systems* (Elsevier) 72, n. 1 (2004): 57-71.
- [3] Crósta, A. P. (1993). *Processamento Digital de Imagens de Sensoriamento Remoto*. Ed. Rev. - Campinas, SP: IG/UNICAMP.
- [4] DePersia, A. T, A, Bowman, P, Lucey, and E, M. Winter(1995). Phenomenology considerations for hyperspectral mine detection, *Proceedings of SPIE*, 2496:159-167.
- [5] Duchesne C, Liu J. J, MacGregor J. F. (2012). Multivariate image analysis in the process industries: a review. *Chemom Intell Lab Syst*;117(1):116-28.
- [6] Everard, C. D, KIM, M, S. e LEE, H. (2014) "A comparison of hyperspectral reflectance and fluorescence imaging techniques for detection of contaminants on spinach leaves." *Journal of Food Engineering* (Elsevier) 143 (2014): 139-145.
- [7] Geladi P, Grahn H. (1996). *Multivariate image analysis*. Chichester, England: Wiley.
- [8] Ghasemzadeh-Barvarz, M., Rodrigues, D., e Duchesne, C., 2014. Multivariate image analysis for inspection of multilayer. *Polymer Testing*, 40:196-206.
- [9] Gonzalez, R. C.; Woods, R. E. (1992). *Digital image processing*. Reading, MA: Addison-Wesley.

- [10] Hameed, I. A. (2016). Motion planning for autonomous landmine detection and clearance robots. *Engineering and Natural Sciences*, Norwegian University of Science and Technology NTNU in Alesund, Postboks 1517, NO-6025 Ålesund, Norway.
- [11] Hong, S., Miller, T., Tobin, H., Borchers, B., Hendrickx, J. M., Lensen, H.A., Schwering, P. B., Baertlein, B. A. (2001). Impact of soil water content on landmine detection using radar and thermal infrared sensors. *Proc. SPIE*, 4394:409 – 416.
- [12] Jackson JE. (2003). *A user's guide to principal components*. New York: Wiley.
- [13] Janssen, Y. H. L, Jong, A.N, Winkel, H. and Putten, F.J.M.V (1996). Detection of surface laid and buried mines with IR and CCD cameras, an evaluation based on measurements. In A. C. Dubey, R. L. Barnard, and C. J. Lowe, editors, *Proc. SPIE Vol. 2765*. Orlando (FL), USA.
- [14] Kruse, O. M. O., Prats-Montalbán, J. M., Indahl, U. G., Kvaal, K., Ferrer, A., and Futsaether, C. M. (2014). Pixel classification methods for identifying and quantifying leaf surface injury from digital images. *Computers and Electronics in Agriculture*, 108:155{165}.
- [15] Landmine Monitor (2015). International Campaign to Ban Landmines-Cluster Munition Coalition(ICBL-CMC). ISBN:978-2-8399-1707-0. http://www.themonitor.org/media/2152583/Landmine-Monitor-2015_finalpdf.pdf. Acessado em 09/06/2017.
- [16] Lopez P., H. Sahli, D. L. Vilarino, and D. Cabello (2003). Detection of Perturbations in Thermal IR Signatures: An Inverse Problem for Buried Landmine Detection. Brussels, Belgium. Available at http://www.clearfast.vub.ac.be/publications_files/Lopez_spie_2003.pdf
- [17] Macdonald, J., Lockwood, J. R. (2003). *Alternatives for Landmine Detection*. RAND.
- [18] Mason, R. L.; Tracy, N. D.; Young, J. C. (1992). Multivariate control charts for individual observations. *Journal of Quality Technology*, v. 24, n.2, p. 88-95.
- [19] Mende, H., Dej, B., Khanna, S., Apps, R., Boyle, M., Addison, F. (2004). Microwave enhanced IR detection of landmines using 915 MHz and 2450 MHz. *Defence Research Reports*, no.drdc-ottawa-tm-2004-266. Ottawa (Canada): Defence R&D.
- [20] Messelink, W. A. C. M., Schutte, K., Vossepoel M. A, Cremer F., Schavemaker J. G. M (2002). Feature-based detection of land mines in infrared images. *Proc. SPIE 4742*, Detection and Remediation Technologies for Mines and Minelike Targets VII, 108.
- [21] Mingoti, S. A. (2005). *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Editora UFMG.
- [22] Moeller, M. S, e Blaschke, T. "Urban change extraction from high resolution satellite image." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vienna, Austria (Citeseer) 36, n. 2 (2006): 151-156
- [23] Montgomery, D. C. (2004). *Introdução ao Controle Estatístico da Qualidade*. 4 ed. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A.
- [24] Nações Unidas (2014). Web page: <<https://nacoesunidas.org/onu-10-pessoas-morrem-ou- ficam-mutiladas-por-dia-no-mundo-devido-as-minas-terrestres/>>. Acesso: 21 de março de 2017.
- [25] Prats-Montalbán JM, Ferrer A, de Juan A. (2011). Multivariate Image Analysis: a review with applications. *Chemom Intell Lab Syst*;107:1–23.
- [26] Prats-Montalbán, J. M. and Ferrer, A. (2014). Statistical process control based on multivariate image analysis: A new proposal for monitoring and defect detection. *Computers & Chemical Engineering*, 71:501{511}.
- [27] Sendur I, K, and Baertlein A, B. (1999). Techniques for improving buried mine detection in thermal IR imagery, *Proc. SPIE 3710*, Detection and Remediation Technologies for Mines and Minelike Targets IV, 1272; doi:10.1117/12.357009.
- [28] Song, X., Su, J., Huang, M., Lu, S. and Zhu, P. (2016). Landmine detection with holographic radar, *Ground Penetrating Radar (GPR) 16th International Conference on*, pp. 1-4, 2016.
- [29] Simunek, J., Hendrickx, J. M. H, and Borchere, B.(2001). Modeling transient temperature distributions around landmines in homogeneous bare soils. *Proc. of SPIE*, This Volume, Detection and Remediation Technologies for Mines and Minelike Targets IV.
- [30] Szymanik, B. (2014). Inverse Problem Solution in Landmines Detection Based on Active Thermography. *Radioengineering*.
- [31] Vines, S. K. (2000). Simple principal components. *Appl. Stat.* 49, 441 – 451.
- [32] Watson, K. (1975). Geological application of thermal infrared images. *Proc. IEEE*, vol. 63, pp. 128–137.
- [33] Wise, B. M. and Geladi, P. (2000). A brief introduction to multivariate image analysis (MIA). *Eigenvector Research*.
- [34] Zea, Q., and Cardenas M. (2013). RASCAN Type Radar Image Resolution Enhancement for Non-metallic Landmine Detection, *XVIII Symposium of Image Signal Processing and Artificial Vision*, pp. 1-4.

Capítulo 15

ESTUDO DA PROPAGAÇÃO DOS ERROS EM PROCESSOS PRODUTIVOS MULTIVARIADOS EM ESTÁGIOS

Josiane da Silva Jesus

Paulo Henrique Coelho Maranhão

Resumo: Atualmente, uma parte considerável dos processos produtivos contém dois ou mais estágios. Assim a qualidade final de um produto está relacionada de como possíveis erros se propagam nos diversos estágios. Além disso, não é raro que as variáveis de interesse desses processos sejam correlacionadas, de forma que um erro em uma das variáveis afeta todas as outras e em consequência, todo o processo é afetado. Este trabalho propõe um modelo multivariado para estudar a propagação dos erros entre estágios. O estudo foi realizado por meio de simulações, que mostrou resultados bastante promissores.

Palavras chave. Processo produtivo, propagação dos erros, multivariado.

1. INTRODUÇÃO

O mundo atual, competitivo e globalizado, provoca uma busca constante por qualidade, eficiência, diminuição de custos e maior retorno financeiro em toda e qualquer organização. Esse conjunto de objetivos se transformou em necessidade básica e crucial para sobrevivência de empresas e indústrias.

Um produto ou serviço pode ser afetado negativamente pela variabilidade durante o seu processo de fabricação ou execução. Essa variabilidade são desvios que representam o afastamento entre o resultado obtido e o resultado desejado em um processo qualquer. Sabe-se ainda, que grande parte dos processos produtivos possuem dois ou mais estágios, e muitas vezes esses estágios são correlacionados. Dessa forma, estudar a propagação dos erros entre estágios correlacionados pode ser uma contribuição importante para se verificar qual a etapa do processo é responsável pela maior transmissão da variância, de forma que se possa agir preventivamente nessa etapa.

Geralmente, a análise de um processo leva em consideração mais de uma variável de interesse. Porém, não é raro ocorrer casos em que as variáveis de interesse mensuradas em cada estágio também tenham um grau elevado de correlação e técnicas multivariadas são necessárias para o desenvolvimento de modelos que consigam monitorar a propagação do erro para esses casos. O principal objetivo deste trabalho é desenvolver um estudo sobre a propagação de erros em processos mecânicos em estágios, cujas características de qualidade de interesse estejam correlacionadas.

Os primeiros estudos sobre a transmissão da variância foram realizados em meados dos anos noventa. Dessa forma, em 1998 foi realizado um estudo da variabilidade em um sistema flexível de montagem de portas de veículos com auxílio de layout fornecido pelo programa CAD e uma técnica estatística multivariada, denominada de Análise de Componentes Principais, a qual descreve a variabilidade do processo através de uma transformação linear das variáveis observadas originalmente [Ceglarek 1998]. Em 1999, foi proposto um modelo autorregressivo de ordem 1 (AR1) para estimar a propagação da variação em estágios de um processo produtivo em relação a uma única variável de interesse,

identificando qual etapa produtiva contribui significativamente para a variabilidade do processo [Lawless et al. 1999]. Posteriormente, foi desenvolvido um método para avaliar a propagação da variação em um processo de usinagem por meio de um modelo de estado-espço, baseado nas informações do projeto do produto e processo, também foi utilizado o conceito virtual de operação para isolar as falhas e determinar as causas das mesmas [Huang et al. 2002].

Já em 2004, avaliou-se um processo de usinagem através de uma matriz homogênea de transformação, cujos elementos representam a diferença espacial entre dois sistemas de coordenadas e foi empregado o método de Análise de Componentes Designados (DCA). O método DCA é similar ao método das Componentes Principais, a diferença encontra-se no fato de que o DCA é capaz de identificar falhas de fixação simultâneas em chapas metálicas e não considera qualquer tipo de interação entre elas e o processo de fabricação [Camelio et al. 2004]. No mesmo ano, implementou-se a metodologia Seis Sigma em um processo produtivo de automóveis, com o objetivo de reduzir os desvios gerados. Para isso, verificou-se a folga presente nas portas traseiras de 17 veículos em dois estágios de fabricação, Chaparia e Montagem [Gaio e Sá 2004].

No ano seguinte, [Liao e Wang 2005] desenvolveram um método novo a partir da geometria fractal, a qual caracteriza fenômenos espaciais ou temporais contínuos de corpos não rígidos, e do método de elementos finitos, com o objetivo de analisar a variação na superfície micro geométrica dos componentes desses corpos. A geometria fractal foi expressa por uma função chamada de Weierstrass-Mandelbrot, capaz de representar as características da superfície micro geométrica das peças utilizadas no processo de montagem. O método de elementos finitos foi usado para analisar a deformação dos componentes do sistema produtivo. Essa variação merece atenção em processos de montagem de alta precisão. Subsequentemente, [Zhang et al. 2007] apresentaram a metodologia do fluxo de variação baseada na descrição do projeto (CAD) e do processo. Essa metodologia utiliza a modelagem de espaço-estado (linear e que estabelece relação entre os erros e

suas causas), análise de sensibilidade para verificar o grau de impacto dos desvios na qualidade de cilindros automotivos e estabelecer a otimização do processo a fim de minimizá-los. O estudo da decomposição do erro auxilia engenheiros a identificar qual estágio contribui mais para a variabilidade do processo.

Em 2012, foi realizado um estudo objetivando a minimização da propagação de desvios. A ideia era avaliar estágio por estágio do processo minimizando o somatório ao quadrado das variações em todas as fases [Yang et al. 2012]. No ano seguinte foi proposto um modelo de previsão de qualidade (CQPM), capaz de lidar com variáveis complexas presentes em sistema multi-estágio de fabricação, através de técnicas de dados como Componentes Principais (ACP), que extrai informações de diversas variáveis inter-relacionadas e as expressa em um novo conjunto de variáveis ortogonais [Arif et al. 2013].

O artigo proposto encontra-se organizado em 5 seções. Na seção 2, são tratados os principais conceitos e definições de ferramentas da Estatística Multivariada e o Modelo utilizado. A seção 3 descreve a metodologia aplicada nas simulações. A quarta seção apresenta a análise de resultados. Por fim, a quinta seção contém a conclusão do trabalho.

2. MÉTODOS

2.1 MODELO AUTORREGRESSIVO DE PROPAGAÇÃO

Em 1999, [Lawless et. al, 1999], propuseram um modelo autorregressivo (AR(1)) para estudar a transmissão da variação em processos de múltiplos estágios. A escolha deste modelo é justificada visto que se desejava estudar a influência da variância entre etapas subsequentes. Dessa forma, a equação a seguir ilustra o modelo AR(1):

$$Y_i = \alpha_i + \beta_i Y_{i-1} + \varepsilon_i \quad i = 2, \dots, k \quad (1)$$

em que Y_i é a variável que representa a medida observada na etapa i e Y_{i-1} é a medida observada

na etapa imediatamente anterior, onde Y_i segue uma distribuição Normal com média μ_i e variância σ_i^2 . Os termos α_i e β_i representam respectivamente, os coeficientes linear e angular do modelo. Os resíduos ε_i 's representam os erros aleatórios do modelo e também seguem uma distribuição Normal com média 0 e variância constante σ_{ε}^2 . Além disso, supõe-se que os erros e Y_i apresentam

covariância igual à zero, ou seja, são independentes.

O principal atributo dos modelos AR (1) é que o valor atual está fortemente relacionado com o imediatamente anterior, ou seja, há uma correlação na primeira defasagem, entre Y_i e $Y_{(i-1)}$. Isso contribui para que os erros possam ser obtidos em cada defasagem considerada.

Dessa maneira, a variabilidade transmitida entre dois estágios é estabelecida pela equação a seguir:

$$\sigma_i^2 = \beta_i^2 \sigma_{i-1}^2 + \sigma_{\varepsilon,i}^2 \quad (2)$$

onde σ_i^2 é a variância total da etapa i ; $\sigma_{\varepsilon,i}^2$ é a variância adicionada na etapa atual; σ_{i-1}^2 é a variância na etapa anterior; β_i^2 é o coeficiente

responsável pela transmissão da variabilidade entre os dois estágios.

Caso tenhamos k etapas, a equação (2) pode ainda ser generalizada pela seguinte expressão:

$$\sigma_k^2 = \beta_k^2 \sigma_{k-1,A}^2 + \dots + \beta_k^2 \beta_{k-1}^2 \dots \beta_2^2 \sigma_1^2 + \sigma_{k,A}^2 \quad (3)$$

2.2 ANÁLISES DE COMPONENTES PRINCIPAIS

Análise de Componentes Principais (ACP) é uma técnica multivariada, em que um número de variáveis relacionadas é transformado em um conjunto menor de variáveis não correlacionadas, chamadas de Componentes Principais (CP). Nesse método são geradas tantas componentes quanto variáveis, mas a grande vantagem da técnica é que, em geral, poucas componentes explicam a maior parte da variabilidade do conjunto de variáveis originais. Assim, se p componentes representam a variabilidade total de um conjunto de p variáveis originais, a maior

parte desta variabilidade pode ser explicada por k componentes ($k < p$). Logo, k componentes substituem as p variáveis iniciais, reduzindo a dimensão do conjunto de dados. Além disso, a primeira CP é uma combinação linear com variância máxima.

Assim seja X um vetor aleatório com vetor de médias μ e matriz de covariâncias ou de correlação, $\Sigma_{p \times p}$, com λ_i autovalores e e_i autovetores (normalizados) que obedecem às condições: $e_i' e_j = 0$ ($i \neq j$); $e_i' e_i = 1$ (para todo i) e $\Sigma_{p \times p} e_i = \lambda_i e_i$ (para todo i). Logo, a j -ésima componente principal é dada pela expressão a seguir:

$$Y_j = e_j' X = e_{j1} X_1 + e_{j2} X_2 + \dots + e_{jp} X_p \quad (4)$$

A Figura 1 abaixo retrata um esquema de como funciona o método.

Figura 1: Esquema de funcionamento do método ACP.

Em geral, observa-se que grande parte da variabilidade das variáveis originais são explicadas por poucas componentes principais. Sendo assim, pode-se dizer que se mais de 80% da variância total pode ser explicada por duas ou três componentes, então estas componentes podem substituir as variáveis originais sem perda de informação [Johnson e Wichern 2007].

2.3 MODELO AUTORREGRESSIVO NAS COMPONENTES PRINCIPAIS

O modelo autorregressivo de ordem um, descrito na seção anterior, servirá de base para determinação do modelo autorregressivo nas componentes principais. Dessa forma, a componente principal do estágio i pode ser modelada em função da componente principal da etapa anterior ($i-1$), conforme pode ser visto na equação a seguir:

$$CP_i = \alpha_i + \beta_i CP_{i-1} + \varepsilon_i \quad (5)$$

Aplicando a variância na i -ésima componente principal, tem-se:

$$Var(CP_i) = \sigma_i^2 = \beta_i^2 \sigma_{i-1}^2 + \sigma_{\varepsilon_i}^2 \quad (6)$$

Por analogia, todas as suposições e definições aplicadas ao modelo autorregressivo dado pela equação (1) continuam válidas para a equação (5). Assim, os termos α_i e β_i representam também os coeficientes linear e angular do modelo e os termos ϵ_i 's são os resíduos. Considera-se também que os resíduos e as CP's são independentes.

Na prática, as equações (5) e (6) podem ser estimadas por meio das equações de (7) a (11), que representam respectivamente os estimadores da covariância das CP's entre os estágios, da variância das CP's, dos coeficientes angular e lineares e da variância adicionada no próprio estágio.

$$S_{i-1,i} = \frac{\sum_{j=1}^n (CP_{j,i-1} - \bar{CP}_{i-1})(CP_{j,i} - \bar{CP}_i)}{n} \quad (7)$$

E a variância é dada pela equação (8):

$$S_{ii} = \frac{\sum_{j=1}^n (CP_{j,i} - \bar{CP}_i)^2}{n} \quad (8)$$

O parâmetro de propagação é calculado através da equação (9).

$$\beta_i = \frac{S_{i-1,i}}{S_{i-1,i-1}} \quad (9)$$

O coeficiente linear e a variância adicionada entre etapas são obtidos pelas equações (10) e (11), respectivamente.

$$\alpha_i = \bar{CP}_i - \beta_i \bar{CP}_{i-1} \quad i=2, \dots, k \quad (10)$$

$$\sigma_{i,A}^2 = S_{ii} - \beta_i (S_{i-1,i}) \quad i=2, \dots, k \quad (11)$$

3. METODOLOGIA

Neste tópico, apresenta-se a metodologia utilizada em um estudo de simulação. Dessa forma, a ideia é simular um processo produtivo formado por dois estágios, em que cada estágio é composto por quatro variáveis conforme a Figura 2, em que as mesmas são correlacionadas em diferentes níveis. Assim, para cada estágio foram geradas 100 amostras dessas quatro variáveis, todas seguindo uma distribuição normal com média zero e variância um.

Considerou-se ainda, três tipos de correlação entre as variáveis, a saber: forte, moderada e fraca, e mais três tipos de relacionamento

entre os estágios, conforme pode ser visto na Figura 3. Assim, para a primeira simulação, figura 3, assume-se que as variáveis em cada estágio estão fortemente correlacionadas entre si, considerando também três tipos de correlação entre os estágios. Para a segunda e terceira simulação, a mesma rotina é seguida. Posteriormente, 100 números aleatórios são gerados, todos seguindo distribuição normal com média zero e variância 3, apenas na primeira variável do estágio 1, para avaliar como os erros dessa variável afetam as demais variáveis. Portanto, as três simulações foram divididas com a variância 1 e com a variância 3.

Figura 2: representação de variáveis nos estágios.

Figura 3: Esquema dos casos simulados considerando variáveis em cada estágio fortemente correlacionadas.

A partir de cada grupo de variáveis em cada estágio foram geradas componentes principais, selecionando a componente que mais explica a variabilidade das variáveis originais. Em seguida, é aplicado o modelo

proposto entre as componentes de cada estágio, onde a componente do segundo estágio é regredida em função da componente do primeiro estágio, conforme Figura 4.

Figura 4: representação de componentes principais nos estágios.

4. RESULTADOS

Nesta seção, são apresentados os resultados referentes ao estudo de simulação descrito na seção 3.

4.1 RESULTADOS SIMULAÇÃO 1

Depois de obter as variáveis fortemente correlacionadas em cada estágio, de acordo com a metodologia, obtivemos as componentes principais para primeira e segunda etapas, considerando as três relações entre as etapas. As Tabelas 1 e 2 mostram as variância explicadas pelas componentes principais no estágio 1.

Tabela 1: Percentual da variância explicada pelas CP's simulação 1 Estágio 1– $N(0,1)$.

% CP1	% CP2	% CP3	%CP4
95.8%	4.20%	0.00%	0.00%

Tabela 2: Percentual da variância explicada pelas CP's simulação 1 Estágio 1– $N(0,3)$.

% CP1	% CP2	% CP3	%CP4
99.6%	0.40%	0.00%	0.00%

Os valores relativos às proporções das componentes principais em cada caso no estágio 2 são mostrados nas Tabelas 3 e 4.

Tabela 3: Percentual da variância explicada pelas CP's simulação 1 Estágio 2 – N(0,1).

Caso	% CP1	% CP2	% CP3	%CP4
1 (90%)	83.10%	8.70%	5.00%	3.20%
2 (50%)	46.40%	24.20%	17.90%	11.50%
3 (30%)	39.10%	26.50%	21.00%	13.40%

Tabela 4: Percentual da variância explicada pelas CP's simulação 1 Estágio 2 – N(0,3).

Caso	% CP1	% CP2	% CP3	%CP4
1 (90%)	98.00%	1.00%	0.60%	0.40%
2 (50%)	82.90%	7.90%	5.90%	3.30%
3 (30%)	63.50%	16.70%	12.50%	7.30%

Analisando as Tabelas 1, 2, 3 e 4, observa-se que quando as variáveis originalmente geradas são fortemente correlacionadas, as primeiras componentes de cada estágio são responsáveis por quase toda a variabilidade das variáveis originais, como nos estágios 1 com as variâncias iguais a 1 e 3; no primeiro

caso do estágio 2 com variância 1 e em todo estágio 2 com variância 3. Desta forma, aplicamos o modelo proposto considerando as primeiras componentes em cada estágio. As Figuras 5, 6 e 7 mostram como a variância se propaga entre os estágios.

Figura 5: Variabilidade entre estágios Simulação 1 caso 1.

Figura 6: Variabilidade entre estágios Simulação 1 caso 2.

Figura 7: Variabilidade entre estágios Simulação 1 caso 3.

A Figura 5 revela que 91,99% e 99,24% da variabilidade, presentes nos estágios 2 com as variâncias 1 e 3, respectivamente, foram originados no estágio anterior. Isso significa que um erro no estágio anterior é quase totalmente propagado para a próxima etapa. Ainda assim, é possível ver por meio das Figuras 6 e 7 que à medida que diminui a relação entre os estágios, diminui a transmissão do erro. No entanto, para a variância igual a 3, a propagação dos desvios

foi bastante alta em relação à variância igual a 1.

4.2 RESULTADOS SIMULAÇÃO 2

As variâncias explicadas pelas componentes principais no estágio 1 para a simulação 2 são apresentadas nas Tabelas 5 e 6. Já, as Tabelas 7 e 8 apresentam os valores das variâncias explicadas no estágio 2 para cada nível de correlação entre os estágios.

Tabela 5: Percentual da Variância explicada pelas CP's simulação 2 Estágio 1– N(0,1).

% CP1	% CP2	% CP3	%CP4
78.60%	21.40%	0.00%	0.00%

Tabela 6: Percentual da Variância explicada pelas CP's simulação 2 Estágio 1 - N(0,3).

% CP1	% CP2	% CP3	%CP4
92.00%	8.00%	0.00%	0.00%

Tabela 7: Percentual da Variância explicada pelas CP's simulação 2 Estágio 2 – N(0,1).

Caso	% CP1	% CP2	% CP3	%CP4
1 (90%)	68.10%	22.30%	5.40%	4.20%
2 (50%)	34.60%	28.00%	21.60%	15.80%
3 (30%)	31.00%	26.00%	24.60%	18.40%

Tabela 8: Percentual da Variância explicada pelas CP's simulação 2 Estágio 2 – N(0,3).

Caso	% CP1	% CP2	% CP3	%CP4
1 (90%)	87.50%	8.30%	2.30%	1.90%
2 (50%)	59.40%	18.40%	13.80%	8.40%
3 (30%)	43.00%	25.10%	19.30%	12.60%

Observe que, embora as variáveis nos estágios sejam moderadamente correlacionadas, as primeiras componentes principais do estágio 1 com as variâncias 1 e 3 ainda são capazes de um alto grau de explicação. Para os estágios 2 com as variâncias 1 e 3, a variância explicada pelas

primeiras componentes tem menores valores. As Figuras 8, 9 e 10 mostram a propagação de erros entre os dois estágios em termos da primeira componente principal, uma vez que a mesma representa a maior parte da variabilidade das variáveis no estágio 1.

Figura 8: Variabilidade entre estágios Simulação 2 caso 1.

Pela Figura 8, observa-se que a propagação da variabilidade entre as etapas foi de 92,10% para o caso 1 com variância 1 e de 98,28% para o caso 1 com variância 3 da simulação 2 e a Figura 9 mostra que no segundo estágio, 34,80% da variância foi derivada do estágio anterior e 65,20% dos erros foram gerados no próprio estágio 2 com

variância 1. Entretanto, para a Figura 9 com variância 3, a transmissão dos desvios entre as etapas foi de 85,87%. Para a correlação entre os estágios de 30%, tem-se que 98,04% e 27,21% dos erros presentes no segundo estágio com as variâncias 1 e 3, respectivamente, foram gerados pelo próprio estágio, conforme mostra a Figura 10.

Figura 9: Variabilidade entre estágios Simulação 2 caso 2.

Figura 10: Variabilidade entre estágios Simulação 2 caso 3.

4.3 RESULTADOS SIMULAÇÃO 3

Com a análise de componentes principais, obtiveram-se os valores da variabilidade máxima e mínima fornecidas pelas

componentes principais no estágio 1, conforme as Tabelas 9 e 10.

Tabela 9: Percentual da Variância explicada pelas CP's simulação 3 Estágio 1 – N(0,1).

% CP1	% CP2	% CP3	%CP4
75.10%	24.90%	0.00%	0.00%

Tabela 10: Percentual da Variância explicada pelas CP's simulação 3 Estágio 1 – N(0,3).

% CP1	% CP2	% CP3	%CP4
77.30%	22.70%	0.00%	0.00%

As Tabelas 11 e 12 fornecem os valores da variabilidade explicada por cada componente principal no estágio 2 em cada situação.

Tabela 11: Percentual da Variância explicada pelas CP's simulação 3 Estágio 2- N(0,1).

Caso	% CP1	% CP2	% CP3	%CP4
1 (90%)	62.50%	25.60%	6.70%	5.20%
2 (50%)	31.90%	28.00%	22.40%	17.70%
3 (30%)	29.10%	26.10%	24.60%	20.20%

Tabela 12: Percentual da Variância explicada pelas CP's simulação 3 Estágio 2- N(0,3).

Caso	% CP1	% CP2	% CP3	%CP4
1 (90%)	66.90%	22.40%	6.10%	4.60%
2 (50%)	37.80%	24.40%	19.70%	18.10%
3 (30%)	31.10%	26.50%	21.80%	20.50%

Analisando a Figura 11 observa-se que a propagação de erros entre os estágios foi de 91,28% e 93,91% para o caso 1 com variância 1 e caso 1 com variância 3. Além disso, a Figura 12 mostra que a transmissão das variâncias entre os estágios foi de

38,95% e 53,12% da primeira componente nos casos 2 com as variâncias 1 e 3 e no caso 2 com variância 1 (Figura 13) a propagação das variâncias foi ainda menor, cerca de 0,5 %. Para o caso 3 com variância

3 da simulação 3, os desvios transmitidos

foram de 21,70%.

Figura 11: Variabilidade entre estágios CP 1 Simulação 3 caso 1.

Figura 12: Variabilidade entre estágios CP1 Simulação 3 caso 2.

Figura 13: Variabilidade entre estágios CP1 Simulação 3 caso 3.

5. CONCLUSÃO

Após a análise dos casos simulados, conclui-se que o modelo autorregressivo nas Componentes Principais foi capaz de capturar a variação oriunda da etapa anterior, mostrando-se promissor para aplicação em

casos reais. Além disso, observou-se que quanto maior a correlação entre as variáveis de interesse, maior a capacidade do modelo proposto em reduzir a dimensão do modelo, devido ao alto percentual de variância explicado pelas componentes principais, em

particular a primeira componente. Em consequência, maior será a capacidade do modelo de prever a transmissão da variabilidade entre os estágios.

Em resumo, pode-se inferir que os modelos autoregressivos nas Componentes Principais

são válidos para avaliar o comportamento da propagação de erros em processos com dois ou mais estágios. Através deles, é possível identificar qual estágio é responsável pela maior variabilidade do processo.

REFERÊNCIAS

- [1] Ceglarek, D.J. Multivariate analysis and evaluation of adaptive sheet metal assembly systems. The university of michigan. 1998.
- [2] Lawless, J.F., Mackay, R.J., Robinson, J.A. Analysis of variation transmission in manufacturing processes – part 1. Journal of quality technology, v.31, n.2, 1999.
- [3] Huang, Q., Zhou, S., Shi, J. Diagnosis of multi-operational machining processes through variation propagation analysis. Robotics and computer integrated manufacturing 18, p.233-239, 2002.
- [4] Camelio, J., Hu, S. J., Zhong, W. Diagnosis of multiple fixture faults in machining processes using designated component analysis. Journal of manufacturing systems, v.23, n.4, 2004.
- [5] Gaio, A. P.C., Sá, F.C. O uso da metodologia seis sigma em um processo da indústria automotiva por meio de um estudo de caso. Trabalho graduação em engenharia de produção. UERJ. Resende. 2004.
- [6] Liao, X.G., Wang, G. Employing fractals and FEM for detailed variation analysis of non-rigid assemblies. International Journal of Machine Tools & Manufacture 45, p.445-454, 2005.
- [7] Zhang, M., Djurdjanovic, D., Ni, J. Diagnosibility and sensitivity analysis for multi-station machining processes. International Journal of Machine Tools & Manufacture 47, p.646-657, 2007.
- [8] Yang, Z., Popov, A.A., McWilliam, S. Variation propagation control in mechanical assembly of cylindrical components. Journal of manufacturing systems 31, p 162-176, 2012.
- [9] Arif, F., Suryana, N., Hussin, B. Cascade quality prediction method using multiple pca+id3 for multi-stage manufacturing system. Internacional conference on electronic engineering and computer science. Ieri procedia 4, p.201-207, 2013.
- [10] Johnson, R.A., Wichern, D.W. Applied multivariate statistical analysis. 6ª ed. Pearson prentice hall. 2007.
- [11] Montgomery, D.C. Introdução ao controle estatístico de qualidade. 4ªed. Gen/ltc. Rio de janeiro. 2013.
- [12] Mingoti, S.A. Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada. 2ª reimpressão. Editora UFMG. Belo horizonte. 2013.

Capítulo 16

ALGORITMO DE BUSCA LOCAL BASEADO EM INDICADOR APLICADO AO PROBLEMA DE ROTEAMENTO DE VEÍCULOS MULTIOBJETIVO COM COLETA OPCIONAL

Ramon Rocha Leite

Luciana Pereira de Assis

Alessandro Vivas Andrade

Marcelo Ferreira Rego

Cristiano Grijó Pitangui

Resumo: Uma considerável parte do custo operacional das empresas está relacionado aos custos logísticos para a distribuição de bens, principalmente no transporte de mercadorias. Assim, esse trabalho aborda o Problema de Roteamento de Veículos Multiobjetivo onde o atendimento da demanda de coleta é opcional. São considerados os objetivos de minimizar a distância total percorrida e a quantidade de itens não coletados. Devido à natureza combinatória desse problema, métodos heurísticos são utilizados para gerar um conjunto de soluções em um tempo computacional aceitável. O algoritmo utilizado neste trabalho é a busca local multiobjetivo que utiliza um indicador binário o qual possibilita ao tomador de decisão apontar a suas preferências de otimização. Foi comparado o desempenho do algoritmo utilizando três indicadores diferentes (Epsilon, Hipervolume e Fonseca) em 12 instâncias de problemas. Os resultados mostram que não há um indicador que seja melhor em todas as métricas avaliadas.

1. INTRODUÇÃO

Sob qualquer ponto de vista: econômico, político e militar - [o transporte] é, inquestionavelmente, a indústria mais importante do mundo" (BALLOU, 1993). As empresas e indústrias estão cada vez mais preocupadas em propiciar, com custo mínimo, melhor serviço e escoamento das suas mercadorias. O custo de escoamento tem influência direta no custo final de qualquer tipo de produto e pode ser a chave do sucesso ou fracasso de uma empresa. Em 1927, Henry Ford já afirmava que "o limite real de uma empresa é o transporte" (FORD, 1927).

Alguns problemas de roteamento encontrados na literatura envolvem tanto serviço de entrega quanto coleta de mercadorias. Estudos envolvendo problemas de coleta e entrega têm crescido nos últimos anos por se tratar de um problema básico da logística reversa. Em geral, este problema consiste em definir rotas de menor custo que satisfaçam a todas as demandas de coleta e entrega dos consumidores, respeitando a capacidade limitada dos veículos. Entretanto, em diversas situações reais, esses problemas não se limitam a apenas um único objetivo, havendo outros critérios de desempenho que precisam ser avaliados, tais como satisfação dos clientes, satisfação dos funcionários ou atendimento da legislação vigente.

Em outras situações reais, a realização da coleta pode não ser obrigatória, permitindo a escolha de quais coletas realizar. Este problema é conhecido como problema de roteamento de veículos com coletas opcionais. Um exemplo desta situação pode ser observado nas indústrias de agrotóxicos. No Brasil, as empresas produtoras e comercializadoras de agrotóxicos têm um prazo de um ano para coleta e destinação das embalagens (BRASIL, 2002). Com isso, não é necessário atender a todas demandas de coletas, já que tem-se o prazo de um ano para efetuá-las. Apesar da realização das demandas de coleta não serem obrigatórias, a sua realização é desejada.

Em diversos trabalhos da literatura (BRUCK; DOS SANTOS; ARROYO, 2012; GRIBKOVSKAIA et al., 2007; GRIBKOVSKAIA; LAPORTE; SHYSHOU, 2008; GUTIÉRREZ-JARPA et al., 2010) coleta não é obrigatória, o problema é tratado como mono-objetivo. Na maioria destes trabalhos a função objetivo é

dada pelo custo das rotas menos o lucro gerado pelas coletas realizadas. Entretanto, as características do problema possibilitam a elaboração de uma modelagem multiobjetivo onde, além da minimização do custo da rota, deve-se maximizar o número de coletas realizadas. Apesar de muitos problemas reais desta natureza possuírem características multiobjetivo, esta abordagem ainda é pouco estudada dentre os problemas de roteamento de veículos (BOFFEY et al., 1995; JOZEFOWIEZ; SEMET; TALBI, 2008).

Assim sendo, este trabalho apresenta um algoritmo de busca local multiobjetivo para solucionar o problema de roteamento de veículos no qual os objetivos

são minimizar o custo de transporte, bem como maximizar o número de coletas realizadas. Essa variante é chamada de Problema de Roteamento de Veículos Multiobjetivo com Coleta Opcional, proposto por (ASSIS, 2013).

Em problemas de otimização multiobjetivo, geralmente, tem-se objetivos que são conflitantes entre si, e por isso, não existe uma única solução para o problema, mas sim um conjunto de soluções que é chamada fronteira Pareto-ótima. Em algumas situações, definir o que é uma boa aproximação da fronteira não é uma tarefa muito simples e irá depender do tomador de decisão e do cenário a ser otimizado (BASSEUR et al., [s.d.]).

Em geral, os algoritmos populacionais são mais facilmente adaptáveis para resolver problemas multiobjetivos, pois, o conceito de população utilizado nestes algoritmos, se assemelham ao conceito de conjunto de soluções presentes nos problemas multiobjetivos. Por isso, este trabalho explora as principais funcionalidades do algoritmo de busca local multiobjetivo baseado em indicador (IBMOLS, *Indicator-Based Multiobjective Local Search*), proposto por (BASSEUR et al., [s.d.]). Este método é baseado no IBEA (ZITZLER; KUNZLI, 2004) que trabalha com uma população de indivíduos que representam soluções que irão compor a fronteira de Pareto. Além disso, o IBMOLS apresenta uma fase de busca local, para explorar o espaço de soluções do problema MOVRPOC, abordado neste trabalho.

O algoritmo implementa um indicador binário que avalia a qualidade das soluções encontradas, e possibilita ao tomador de decisões apontar suas preferências de otimização. Neste trabalho foram analisados três indicadores diferentes (Épsilon, Hipervolume, Fonseca) (BASSEUR et al., [s.d.]) para se averiguar qual deles melhor se adequa ao problema tratado.

O presente artigo está organizado como segue: na seção 2 são apresentadas a definição do problema e sua modelagem matemática multiobjetivo. Em seguida (seção 3), é proposto o algoritmo baseado no IBMOLS para solucionar o problema em questão. Na seção 4 são apresentados os resultados computacionais obtidos pelo algoritmo, e por fim, na seção 5 são feitas as conclusões deste trabalho.

2. DEFINIÇÃO DO TRABALHO

O problema abordado neste trabalho é denominado: Problema de Roteamento de Veículos Multiobjetivo com Coleta Opcional (MOVRPOC, *Multiobjective Vehicle Routing Problem with Optional Collections*). Seguindo a classificação proposta por (JOZEFOWIEZ; SEMET; TALBI, 2008) para problemas de roteamento multiobjetivo, este problema é uma generalização do problema de roteamento de veículos com coleta e entrega simultânea (VRPSPD). A generalização se dá pela transformação de uma restrição em uma função objetivo, assim sendo, neste problema a restrição de atendimento de todas as demandas de coletas é transformada em uma função objetivo. Para facilitar a resolução do problema, este objetivo é transformado em uma função de minimização no qual deseja-se minimizar o número total de itens não coletados. Assim, o problema de roteamento de veículos multiobjetivo com coleta opcional apresenta duas funções objetivo: (i) minimização do custo total das rotas ou distância total percorrida e (ii) minimização do número total de itens não coletados.

No MOVRPOC todas as demandas de entrega devem ser satisfeitas, porém, o atendimento das demandas de coleta é opcional, mas desejável. Então, o objetivo é construir rotas que minimizem o custo de se realizar o trajeto, atendendo a todas as demandas de entrega, e ao mesmo tempo,

minimizar o número de demandas de coleta não realizadas.

Para atendimento das demandas, existe uma frota de veículos homogêneos e de capacidade limitada, no qual esta capacidade não deve ser extrapolada em nenhum momento do trajeto. Outra restrição é que um consumidor deve ser visitado uma única vez e por um único veículo, ou seja, se este consumidor tiver a demanda de coleta atendida, esta deverá ser realizada simultaneamente com a demanda de entrega.

Os dois objetivos a serem otimizados são conflitantes, ou seja, a melhora em um provoca a piora do outro. Portanto, não existe uma solução que seja ótima para ambos os objetivos simultaneamente. Isso leva ao conceito de dominância de Pareto, gerando um conjunto de soluções não dominadas (ANGEL; BAMPIS; GOURVÈS, 2007).

O MOVRPOC pode ser reduzido em tempo polinomial ao VRPSPD. Utilizando o método tradicional de resolução de problemas multiobjetivo, por exemplo o método ε -Restrito (TAKAHASHI, 2007), cada problema modelado seguindo o método citado recai ao VRPSPD, considerando que a função objetivo referente ao atendimento das demandas de coleta seja transformada em restrição. Como o MOVRPOC pode ser reduzido ao VRPSPD e sendo o VRPSPD um problema NP-Difícil (DELL'AMICO; RIGHINI; SALANI, 2006), logo o MOVRPOC também pertence a classe NP-Difícil.

2.1 MODELAGEM MATEMÁTICA

A modelagem matemática para o problema proposto é apresentada em (ASSIS et al., 2012), a qual se trata de uma adaptação da modelagem proposta por (MONTANÉ; GALVÃO, 2006) para o problema de roteamento de veículos com coleta e entrega simultânea. Esta adaptação se diferencia da modelagem original dado que: (i) uma segunda função objetivo, com a finalidade minimizar o número de coletas não realizadas, é adicionada e; (ii) é removida a restrição que todas as demandas de coleta devem ser realizadas.

O problema de roteamento de veículos multiobjetivo com coleta opcional pode ser definido como: dado um grafo completo $G = (V, A)$, onde $v = \{v_0, v_1, v_2, \dots, v_n\}$ é o

conjunto de vértices e $A = \{(v_i, v_j): i, j \in V \text{ e } i \neq j\}$ o conjunto de arestas. O vértice v_0 representa o depósito e os demais vértices representam os consumidores. Cada aresta (v_i, v_j) tem um valor $c_{ij} \geq 0$ associado que representa o custo de se alcançar o vértice v_j a partir do vértice v_i . Cada consumidor v_i tem uma demanda d_i de entrega e uma demanda p_i de coleta. Tem-se disponível uma frota de k_{max} veículos homogêneos de capacidade Q

para atendimento das demandas. O parâmetro y_{ij} é o somatório das cargas coletadas entre o depósito e o nó v_i (inclusive) dirigida ao nó v_j enquanto z_{ij} é o somatório das cargas entregues entre o nó v_i (exclusive) e o depósito dirigida ao nó v_j .

As variáveis de decisão do problema consistem em:

$$x_{ij}^k = \begin{cases} 1, & \text{se o arco } (v_i, v_j) \text{ faz parte da rota trafegada pelo veículo } k, \\ 0, & \text{caso contrário} \end{cases}$$

$$l_j = \begin{cases} 1, & \text{se a demanda de coleta do consumidor } v_j \text{ é satisfeita,} \\ 0, & \text{caso contrário} \end{cases}$$

Dada a definição formal e as variáveis de decisão do problema, a modelagem

matemática para o MOVRPOC é definida a seguir.

$$\text{Min } f_1(x) = \sum_{k=1}^{k_{max}} \sum_{i=0}^n \sum_{j=0}^n c_{ij} x_{ij}^k \quad (1)$$

$$\text{Min } f_2(l) = \sum_{j=1}^n p_j (1 - l_j) \quad (2)$$

Sujeito a:

$$\sum_{i=0}^n \sum_{k=1}^{k_{max}} x_{ij} = 1, \quad j = 1, \dots, n \quad (3)$$

$$\sum_{i=0}^n x_{ij}^k - \sum_{i=0}^n x_{ji}^k = 0, \quad j = 0, \dots, n \text{ e } k = 0, \dots, k_{max} \quad (4)$$

$$\sum_{j=1}^n x_{0j}^k \leq 1, \quad k = 1, \dots, k_{max} \quad (5)$$

$$\sum_{i=0}^n z_{ij} - \sum_{i=0}^n z_{ji} = d_j \quad \forall j \neq 0 \quad (6)$$

$$\sum_{i=0}^n y_{ji} - \sum_{i=0}^n y_{ij} = p_j l_j, \quad \forall j \neq 0 \quad (7)$$

$$y_{ij} + z_{ij} \leq Q \sum_{k=1}^{k_{max}} x_{ij}^k \quad i, j = 0, \dots, n \quad (8)$$

$$x_{ij}, l_j \in \{0, 1\}, \quad i, j = 0, \dots, n \quad (9)$$

$$y_{ij} \geq 0, \quad i, j = 0, \dots, n \quad (10)$$

$$z_{ij} \geq 0, \quad i, j = 0, \dots, n \quad (11)$$

A função objetivo, dada na equação 1, visa a minimização do custo (distância percorrida), já a equação 2, apresenta a segunda função

objetivo, que visa a minimização do somatório das coletas não realizadas.

Estas funções objetivo devem satisfazer a um conjunto de restrições. A primeira restrição

(equação 3) indica que cada ponto de demanda deve ser visitado por um único veículo. A equação 4 representa a restrição de conservação do fluxo. A equação 5 indica que k_{max} veículos, no máximo, podem ser utilizados. A equação 6 obriga a satisfação de todas as demandas de entrega. A equação 7 garante o atendimento da demanda de coleta do consumidor j quando a variável de decisão l_j assumir valor unitário. A equação 8 define que as demandas devem ser transportadas nos arcos incluídos na solução e ainda impõem um limite para a carga total transportada pelo veículo. A equação 9 representa a restrição de integralidade e as equações 10 e 11 representam as restrições de não-negatividade para demandas de coleta e entrega, respectivamente.

2.2 ALGORITMO DE BUSCA LOCAL MULTI OBJETIVO BASEADO EM INDICADOR (IBMOLS)

Os algoritmos para resolução de problemas multiobjetivo fazem uso de mecanismos para minimizar a distância da fronteira de Pareto ótima e maximizar a diversidade das soluções contidas na fronteira. (ZITZLER; KUNZLI, 2004) propuseram o *Indicator-Based Evolutionary Algorithm* (IBEA) a fim de formalizar preferências em termos da relação de dominância, ou seja, os autores assumiram que a preferência do tomador de decisão é dada em termos de um indicador binário I de qualidade que irá associar um número real para cada solução do conjunto de soluções não dominadas. Este indicador mede a qualidade de cada indivíduo da fronteira. Muitos indicadores têm sido propostos com a intenção de capturar diferentes preferências de otimização. Dessa forma o objetivo do problema é encontrar um conjunto de soluções não dominadas que minimize o valor de I .

(ASSIS, 2013) mostraram que o MOILS (*Multiobjective Iterated Local Search*)

apresentou resultados superiores ao IBMOLS e, portanto, propõe algumas modificações no algoritmo original de (BASSEUR et al., [s.d.]), inserindo algumas das principais características do ILS (LOURENÇO; MARTIN; STÜTZLE, 2003) dentre elas a fase de Perturbação e a fase de Busca Local.

Os passos básicos do algoritmo baseado no IBMOLS, utilizado neste trabalho, são mostrados no algoritmo 1. Inicialmente o algoritmo gera uma população inicial P . Dada uma população de tamanho N , parte dos indivíduos são obtidos pela aplicação da heurística $P\epsilon$ (ASSIS et al., 2012). Os indivíduos restantes da população são gerados aleatoriamente (linha 3). Dada a população P , o valor da aptidão de cada indivíduo é obtido aplicando um indicador binário I (linha 4). O cálculo do valor da aptidão é apresentado na subseção 2.3. Assim, para cada solução s pertencente a P , aplica-se uma perturbação e um mecanismo de busca local (linhas 9 -- 10). Os mecanismos de perturbação e busca local aplicados são similares aos métodos utilizados no algoritmo MOGVNS, apresentados em (ASSIS, 2013). Em seguida, se a solução encontrada s'' for igual a solução s , considera-se, então, apenas a perturbação feita no indivíduo s (linhas 11 -- 13). Neste caso, o indivíduo que passou pela perturbação é considerado com o intuito de preservar a diversidade da população. Assim ela poderá ser novamente selecionada e novamente perturbada podendo gerar outros pontos da fronteira. O algoritmo verifica, então, se o indivíduo gerado é melhor que o pior indivíduo, então s é inserido e o pior é removido da população. Dessa forma, a aptidão de todos os indivíduos é atualizada (linhas 14 -- 17). A variável "*continue*" garante que o processo de modificação de um indivíduo s permanece enquanto a busca local estiver gerando novos indivíduos. Caso contrário, passa-se a explorar outro indivíduo.

Algoritmo 1 Algoritmo IBMOLS modificado**Entrada:** N (Tamanho da População), I (Indicador Binário)

```

1:  $PO \leftarrow \emptyset$ ;
2: repetir
3:    $P \leftarrow \text{População Inicial}(N)$ ;
4:    $Aptidao[] \leftarrow \text{Calcular Aptidão}(P, I)$ 
5:    $s_w \leftarrow \text{Menor}(Aptidao)$ ;
6:   para todos solução  $s \in P$  fazer
7:      $continue \leftarrow \text{TRUE}$ ;
8:     enquanto  $continue$  fazer
9:        $s' \leftarrow \text{Perturbação}(s)$ ;
10:       $s'' \leftarrow \text{Busca Local}(s')$ ;
11:      se  $s = s''$  então
12:         $s'' \leftarrow s'$ ;  $continue \leftarrow \text{FALSE}$ ;
13:      fim se
14:      se  $Aptidao(s'') > Aptidao(s_w)$  então
15:         $P \leftarrow P \cup \{s''\} - \{s_w\}$ ;
16:         $\text{Calcular Aptidão}(s'', s_w, P, I)$ ;
17:         $s_w \leftarrow \text{Menor}(Aptidao)$ ;
18:      se não
19:         $continue \leftarrow \text{FALSE}$ ;
20:      fim se
21:    fim enquanto
22:  fim para
23:   $PO \leftarrow \text{Soluções Não Dominadas}(PO \cup P)$ ;
24:  até Critério de Parada
25: retornar  $PO$ ;

```

Nos trabalhos apresentados por (ASSIS et al., 2012; ASSIS, 2013), apenas um indicador binário I_ε é implementado. Este trabalho implementa e avalia outros dois indicadores: Hipervolume, Fonseca e Flemining para solucionar o MOVRPOC. Detalhes dos indicadores estão na subseção a seguir.

2.3 INDICADORES BINÁRIOS (I)

Seja X o espaço de busca de um determinado problema de otimização e Z o espaço de objetivos correspondentes ao mesmo. Nesse caso, pode-se assumir que $Z = R^n$, onde n é o número de objetivos a serem minimizados. A aptidão de cada solução pode ser definida a partir do indicador binário, de modo que dados dois vetores de solução normalizados A e B a qualidade das soluções de uma população P pode ser avaliada por um indicador binário de diversas maneiras. Uma delas consiste em somar os valores de indicador de cada indivíduo x em relação aos demais membros

pertencentes a Z , $I(A, B)$ é um valor real que representa a diferença de qualidade entre esses dois conjuntos de vetores de objetivos (BASSEUR et al., [s.d.])

Para se calcular essa diferença, todos os vetores objetivos são normalizados, a fim de evitar valores negativos (BASSEUR et al., [s.d.]). Nessa equação, m_i e M_i são, respectivamente, o menor e maior valor da função objetivo i dos indivíduos contidos na população P . Cada função objetivo i de cada

Indivíduo x de P é normalizado conforme a equação 12:

$$F_i(x) = \frac{f_i(x) - m_i}{M_i - m_i} \quad (12)$$

z da população, conforme equação 13. Nessa abordagem, todos os indivíduos são considerados para se mensurar a qualidade de cada solução.

$$I(P \setminus \{x\}, x) = \sum_{z \in P \setminus \{x\}} I(z, x) \quad (13)$$

Outra maneira consiste em, ao invés de se obter o somatório, considerar apenas o melhor valor de indicador obtido de uma determinada solução x em relação a qualquer outra solução z pertencente à população de soluções. Essa maneira define que a qualidade de uma determinada solução se dê a partir da existência de soluções similares ou melhores na população (BASSEUR et al., 2011). Tal estratégia está definida na equação 14.

$$I(P \setminus \{x\}, x) = \min_{z \in P \setminus \{x\}} I(z, x) \quad (14)$$

Uma outra alternativa, proposta por (ZITZLER; KUNZLI, 2004), consiste em fazer uma

combinação entre essas duas abordagens, além de ampliar a influência das soluções dominantes sobre aquelas que são dominadas, através de um fator de escala $k > 0$, conforme equação 15:

$$I(P \setminus \{x\}, x) = \sum_{z \in P \setminus \{x\}} -e^{-I(z, x)/k} \quad (15)$$

Desse modo, o cálculo de aptidão do IBMOLS, é obtido por: $I(P \setminus \{x\}, x)$, para todo $x \in P$ (linha 4 algoritmo 1, sendo que a estratégia de utilização do indicador pode se dar de acordo com qualquer uma das equações 13, 14 e 15 supramencionadas. Por sua vez, a atualização dos valores de aptidão após inserção de uma nova solução e remoção pior solução (linha 16 do mesmo algoritmo) se dá de acordo com o apresentado no algoritmo 2 (ASSIS, 2013):

A seguir, são apresentados os três indicadores binários analisados nes trabalho: I_ϵ , I_{HD} , e I_{FD} .

Algoritmo 2 Calcula Aptidão - IBMOLS

Entrada: s'' (solução inserida), s_w (solução removida), P (população) and I (indicador binário)

- 1: $Aptidao(s'') \leftarrow I(P \setminus \{s''\}, s'')$;
- 2: $Aptidao(z) \leftarrow Aptidao(z) + I(s'', z), \forall z \in P$;
- 3: $Aptidao(z) \leftarrow Aptidao(z) - I(s_w, z), \forall z \in P$;

2.3.1 INDICADOR BINÁRIO EPSILON I_ϵ

Um dos indicadores utilizado para calcular a aptidão dos indivíduos da população é o Indicador Epsilon I_ϵ , o qual apresenta bons resultados para diferentes problemas de otimização combinatória (BASSEUR et al., [s.d.]). Dados dois vetores de soluções x_1 e x_2 pertencentes à fronteira de Pareto, o indicador $I_\epsilon(x_1, x_2)$ mensura a distância entre os dois pontos no espaço dos objetivos $Z = R^n$. Esta distância é calculada a partir da diferença de valor das soluções em cada objetivo do problema, de acordo com a equação 16, apresentada abaixo:

$$I_\epsilon = \max_{i \in \{1, \dots, n\}} (f_i(x_1) - f_i(x_2)) \quad (16)$$

O cálculo de aptidão do indivíduo é realizado de acordo com a equação 15, sendo portanto um indicador que busca potencializar a influência das soluções dominantes sobre as dominadas.

2.3.2 INDICADOR BINÁRIO HIPERVOLUME I_{HD}

O hipervolume de uma solução x_1 , denotado por $H(x_1)$, representa o espaço de objetivos que x_1 domina. Dadas duas soluções x_1 , e x_2 , o indicador binário de hipervolume, denotado por $I_{HD}(x_1, x_2)$, representa o volume do espaço de objetivos que é dominado por x_1 , mas não é dominado por x_2 , (BASSEUR et al., [s.d.]). Seu valor é definido de acordo com a equação 17:

$$I_{HD} = \begin{cases} H(x_1) - H(x_2) & \text{se } x_1 < x_2 \\ H(x_1 + x_2) - H(x_2) & \text{caso contrário.} \end{cases} \quad (17)$$

Assim, como o Indicador Épsilon, é utilizada a equação 15 para o cálculo de aptidão do indivíduo no Indicador Hipervolume. Desse modo, é um indicador que também preserva a influência das soluções dominantes.

2.3.3 INDICADOR BINÁRIO FONSECA E FLEMING (I_{FON})

Um outro indicador, adaptado por (BASSEUR et al., 2011) de um trabalho de (FONSECA; FLEMING; OTHERS, 1993), utiliza um método de ranqueamento originalmente aplicado com um algoritmo genético multiobjetivo. O valor de indicador é aplicado a partir da verificação da relação de dominância entre duas soluções.

Todavia, diferentemente dos indicadores I_ϵ e I_{HD} , o cálculo de aptidão dos indivíduos da população é realizado através do somatório dos valores de fitness de cada indivíduo com relação aos demais membros da população, conforme a equação 13. Portanto, trata-se de um indicador que leva em conta todos os membros da população para se determinar a qualidade de cada solução, porém não preserva a influência das soluções dominantes em mesmo grau que os indicadores anteriores. Seu valor é definido de acordo com a equação 18:

$$I_{Fon}(x_1, x_2) = \begin{cases} -1 & \text{se } x_1 < x_2, \\ 0 & \text{caso contrário} \end{cases} \quad (18)$$

A lógica desse indicador se baseia em um ranqueamento do indivíduo considerando o número de outras soluções dominadas por ele. Todos iniciam com o valor de fitness igual a zero. Ao ser comparado aos demais membros, para cada outra solução que um indivíduo domina é acrescido -1 ao seu valor de fitness. Caso contrário, esse valor não sofre alteração. As soluções que possuírem o menor valor são as melhores, pois são aquelas que dominam uma quantidade maior de outras soluções.

3. RESULTADOS COMPUTACIONAIS

Os testes computacionais deste trabalho utilizaram 12 instâncias de 50 a 150 consumidores mais o depósito, propostas por (SALHI; NAGY, 1999). São instâncias definidas para o problema de roteamento de veículos com coleta e entrega simultâneas (VRPSPD), no qual há demanda de entrega e de coleta por parte dos consumidores, sendo portanto possível de ser utilizada para o MOVPOP, problema ora abordado. Utilizando-se métodos exatos, apenas as instâncias com 50 consumidores são resolvidas em tempo computacional aceitável (ASSIS, 2013).

O custo de transporte é dado pela distância euclidiana entre os pontos (consumidores ou depósito). Nas instâncias da classe CMT1 a CMT4, os consumidores estão distribuídos de forma aleatória no espaço do problema, em torno do depósito. Já nas instâncias CMT11 e CMT12 eles aparecem agrupados, de modo que o depósito não está centralizado entre os consumidores, mas sim entre os grupos (ASSIS et al., 2012). As instâncias armazenam as seguintes informações: capacidade do veículo, número de consumidores, bem como suas respectivas demandas de coleta e entrega. Ademais, a localização de cada nó é dado pelas coordenadas (x, y), a partir das quais se calcula a distância entre os mesmos.

No intuito de verificar a qualidade das soluções e avaliar o desempenho dos indicadores utilizados no IBMOLS, foram utilizadas três métricas, as quais são descritas a seguir:

- Cardinalidade: número de soluções presentes na melhor fronteira de Pareto encontrada.
- Hipervolume: consiste em avaliar a área total de um hipercubo de dimensão igual ao número de objetivos do problema.
- Cobertura de dois conjuntos: essa métrica é utilizada para comparar dois conjuntos de soluções, auferindo em que medida uma fronteira de Pareto domina outra, o que não pode ser realizado através apenas do hipervolume.

Tendo em vista que foram testados três indicadores, para facilitar a comparação de desempenho entre os mesmos, ao invés de se determinar a cobertura dos indicadores aos pares, foi utilizada uma forma de cálculo proposta por (BATISTA et al., 2011) que compara cada indicador em relação à união

dos conjuntos de soluções não dominadas finais dos outros indicadores, generalizando a cobertura que passa a ser definida de um para muitos.

$$C(A, B) = \frac{|\{x_1 \in B: \exists x_2 \in A, x_2 \preceq x_1\}|}{|B|} \quad (19)$$

A Cardinalidade e o Hipervolume é mensurada em números absolutos, ao passo que a Cobertura é fornecida em números relativos.

3.1 ANÁLISE DOS RESULTADOS

Os parâmetros do algoritmo foram definidos a partir dos experimentos realizados em (ASSIS et al., 2012). Na primeira geração do algoritmo, a população inicial é gerada pelo algoritmo P_ε . O tamanho da população inicial foi definido como 10 e a probabilidade de realizar o mecanismo de perturbação igual a

30%. O fator de escala k para os indicadores foi determinado como 0.1 e o número de gerações 50.

Foram obtidas 180 amostras para análise, pois foram testados 3 indicadores em 12 diferentes instâncias do problema, executados 5 vezes cada ($3 \times 12 \times 5$). A média e o desvio padrão dos resultados avaliados pelas métricas são apresentadas nas tabelas 1, 2, 3 e 4.

O indicador I_ε obteve os menores tempos de execução em 10 das 12 das instâncias testadas, 83% do total. Os indicadores I_{FON} e I_{HD} obtiveram um tempo de execução menor em uma instância cada. No entanto, considerando o melhor e o pior tempo de execução, em 75% das instâncias o intervalo formado pelo desvio-padrão do indicador com pior tempo contém o valor correspondente ao tempo do melhor indicador, ou vice-versa. Isso não ocorre apenas nas instâncias 12X, 12Y e 2Y.

TABELA 1: Média e Desvio-Padrão dos Resultados de tempo (em segundos) obtidos

Instâncias	I_ε		I_{FON}		I_{HD}	
	Média	DP	Média	DP	Média	DP
CMT1X	4783,25	241,31	5221,07	448,8	5140,48	202,24
CMT1Y	7700,22	234,36	7246,09	964,65	7366,17	477,8
CMT2X	11696,6	1009,23	11868,65	1620,46	11772,4	649,95
CMT2Y	27499,58	2073,65	32398,84	1607,76	32438,68	4541,82
CMT3X	152092,6	6979,93	162727,6	11602,19	160291,6	7381,23
CMT3Y	203925,6	7675,61	207824,2	10057,66	207873,4	4173,17
CMT4X	574257,6	18919,79	588513,2	16569,53	574362,2	37086,5
CMT4Y	784814,4	45346,66	801915,6	16155,43	790977	57675,27
CMT11X	497025,4	38917,17	506154,8	44709,15	507026	52041,12
CMT11Y	327682,2	34860,79	345287,6	19075,34	342428,4	18764,02
CMT12Y	129340,2	6356,13	138568,2	5581,87	128980,2	5395,27
Média	235070,58	13927,38	242803,70	11161,39	239548,74	16097,82

Com relação à cardinalidade das soluções, houve maior equilíbrio. O indicador I_{FON} obteve os maiores valores em 5 instâncias, seguido pelos indicadores I_ε e I_{HD} , com 4 e 3 melhores resultados, respectivamente. Ressalta-se que os melhores valores de cada indicador estão distribuídos entre instâncias de diferentes números de consumidores, de tal modo que não é possível concluir que o uso de um indicador obteve melhor resultado em instâncias com um certo número de consumidores.

Tratando-se da métrica de hipervolume, os indicadores I_{FON} e I_ε tiveram desempenho

semelhante. O indicador I_{FON} obteve melhor desempenho em 6 instâncias, contra 5 do I_ε , embora este último tenha obtido um valor melhor na média geral do hipervolume de todas as instâncias juntas. Porém, assim como no tempo de execução, considerando os indicadores com o menor e maior valor de hipervolume, em 75% das instâncias o intervalo que se forma com o desvio-padrão a partir da média do menor resultado contém o valor do hipervolume do melhor indicador, ou vice-versa. Somente nas instâncias 1X, 1Y e 3Y isso não ocorre.

Tabela 2: Média e Desvio-Padrão dos Valores de Cardinalidade das Soluções

Instâncias	I_E		I_{FON}		I_{HD}	
	Média	DP	Média	DP	Média	DP
CMT1X	2,6	0,49	2,8	0,4	3,6	0,8
CMT1Y	7,4	1,02	11,4	1,02	11,2	1,17
CMT2X	5	1,41	4,4	1,62	3,2	0,75
CMT2Y	22,4	3,01	29,8	3,06	27,8	5,08
CMT3X	5,2	1,47	6,2	2,14	7	1,1
CMT3Y	17,2	1,6	19,8	3,12	18	3,03
CMT4X	7	1,26	7,2	1,94	6,4	1,5
CMT4Y	30,8	3,06	29	3,1	30,2	2,93
CMT11X	37,4	7,66	35,6	6,89	37	8,37
CMT11Y	20,8	1,6	24,6	4,08	24,8	2,93
CMT12X	14,6	2,73	16	3,58	14,8	2,99
CMT12Y	27,2	4,45	24,8	4,02	24,2	3,87
Média	16,47	2,48	17,63	2,91	17,35	2,88

Tabela 3: Média e Desvio-Padrão do valores de Hipervolume das soluções

Instâncias	I_E		I_{FON}		I_{HD}	
	Média	DP	Média	DP	Média	DP
CMT1X	391,61	16,62	405,05	17,78	413,48	3,61
CMT1Y	2833,78	244	3119,52	80,61	3016,16	107,51
CMT2X	2166,34	247,25	1886,7	351,62	1956,23	266,36
CMT2Y	54779,43	745,53	54355,08	837,48	54242,71	1371,66
CMT3X	970,92	155,08	1123,73	26,12	1116,95	5,63
CMT3Y	9247,71	93,23	9348,28	40,43	9106,29	219,12
CMT4X	1890,53	285,22	2296,68	634,4	1820,75	348,09
CMT4Y	38618,37	866,12	38369,73	1103,39	38393,53	1582,99
CMT11X	36891,88	1131,33	36843,05	614,28	36393,13	912,33
CMT11Y	19246,31	1584,18	18610,46	1585,78	18651,36	1355,23
CMT12X	10735	70,94	10736,53	98,45	10616,73	222,36
CMT12Y	20577,81	1193,77	20770,49	535,56	20574,07	871,64
Média	16529,14	552,77	16488,78	493,83	16358,45	605,54

Já na cobertura, o I_{HD} , que obteve o pior desempenho em todas as avaliações anteriores, se destaca sendo o melhor em 8 instâncias (em uma delas com o mesmo valor do I_E), 67% do total, seguido pelo I_{FON} com 3 melhores resultados e depois o I_E com apenas

2. Nesse conjunto de resultados, o intervalo que se forma com o desvio-padrão a partir da média do menor resultado contém o valor de cobertura do melhor indicador ou vice-versa, em 50% das instâncias.

Tabela 4: Média e Desvio-Padrão do valores de cobertura das soluções

Instâncias	I_E		I_{FON}		I_{HD}	
	Média	DP	Média	DP	Média	DP
CMT1X	0,68	0,19	0,74	0,25	0,73	0,22
CMT1Y	0,32	0,09	0,50	0,08	0,51	0,18
CMT2X	0,46	0,45	0,36	0,34	0,46	0,20
CMT2Y	0,50	0,16	0,34	0,15	0,50	0,09
CMT3Y	0,36	0,08	0,52	0,17	0,37	0,13
CMT4X	0,32	0,14	0,64	0,26	0,37	0,23
CMT4Y	0,42	0,16	0,36	0,10	0,48	0,11
CMT11X	0,50	0,17	0,45	0,12	0,44	0,27
CMT11Y	0,40	0,21	0,37	0,12	0,40	0,13
CMT12X	0,45	0,15	0,40	0,03	0,51	0,17
CMT12Y	0,37	0,11	0,31	0,08	0,48	0,10
Média	0,42	0,18	0,46	0,15	0,49	0,16

O I_E apresentou a maior quantidade dos menores tempos de execução, além de um desempenho próximo ao I_{FON} na cardinalidade e hipervolume. Contudo, obteve o pior desempenho na cobertura. Por sua vez, o I_{FON} apresentou o maior número de melhores resultados nas métricas de cardinalidade e hipervolume, entretanto trata-se de uma diferença muito pequena sobre o indicador I_E . Já com relação aos valores de cobertura de cada indicador com relação aos demais, o I_{HD} se sobressaiu. Como as diferenças entre as médias dos indicadores estão dentro do intervalo do desvio-padrão na maioria dos casos, nota-se que é necessário uma quantidade maior de testes para, a partir de uma análise estatística, auferir se de fato há diferença de desempenho estatisticamente significativa entre os indicadores, em todas as métricas utilizadas.

Além da análise dos resultados em cada instância separadamente, foram comparados os resultados por grupos de instâncias. Um grupo foi composto pelas instâncias menores (CMT1, CMT2 e CMT3), outro grupo composto pelas instâncias maiores (CMT4) e por fim um terceiro grupo que abrange as instâncias que possuem uma configuração diferente na disposição dos consumidores, se comparadas com os outros grupos (CMT11 e CMT12). Mesmo nessa análise agrupada, não houve um indicador que se destacasse e apresentasse resultado significativamente distinto em relação ao obtido separadamente em cada instância.

REFERÊNCIAS

- [1] Angel, E.; Bampis, E.; Gournès, L. Approximation in Multiobjective Problems. In: Gonzalez, T. F. (Ed.). [s.l.] Chapman & Hall/CRC computer & information science, 2007. p. 28.1--28.15.
- [2] Assis, L. P. et al. Multiobjective vehicle routing problem with fixed delivery and optional collections. *Optimization Letters*, v. 7, n. 7, p. 1419–1431, 2012.
- [3] Assis, L. P. D. E. Investigação de metaheurísticas aplicadas ao problema de roteamento de veículos multiobjetivo com coleta opcional. [s.l.: s.n.].
- [4] Ballou, R. Logística Empresarial: Transportes, Administração de Materiais, Distribuição Física. 1. ed. São Paulo: Atlas, 1993.
- [5] Basseur, M. et al. The efficiency of

4. CONCLUSÕES

O presente trabalho apresentou uma análise de diferentes indicadores aplicados ao IBMOLS para solucionar o MOVRPOC. Assim, o algoritmo foi testado com três diferentes indicadores encontrados na literatura: o indicador Epsilon (I_E), indicador de Hipervolume (I_{HD}) e o indicador baseado no método de ranqueamento de (I_{FON}) (FONSECA; FLEMING; OTHERS, 1993).

O algoritmo foi aplicado às instâncias de teste com 50 a 150 consumidores. Para avaliação de cada indicador, as métricas utilizadas foram cardinalidade, hipervolume e cobertura.

Os resultados obtidos de acordo com as métricas avaliadas demonstram um desempenho muito próximo entre os indicadores. Além da análise das instâncias separadamente, os resultados foram analisados por grupos de instâncias: até 100 consumidores (CMT1, CMT2 e CMT3), 150 consumidores (CMT4) e instâncias que apresenta características diferentes na disposição dos consumidores (CMT11 e CMT12). Mesmo nessa análise de grupos, pode-se observar que não existe diferença significativa nos resultados obtidos pelos três indicadores.

Assim sendo, trabalhos futuros poderão fazer uma investigação mais profunda destes indicadores e estudo de novos indicadores para o problema abordado.

indicator-based local search for multi-objective combinatorial optimisation problems. *J Heuristics*, [s.d.].

- [6] Basseur, M. et al. The efficiency of indicator-based local search for multi-objective combinatorial optimization problems. *Journal of Heuristics*, v. 18, n. 2, p. 263–296, 2011.
- [7] Batista, L. S. et al. Pareto Cone e-Dominance: Improving Convergence and Diversity in Multiobjective Evolutionary Algorithms. In: [s.l.] Springer Berlin Heidelberg, 2011. p. 76–90.
- [8] Boffey, B. et al. Multiobjective routing problems. *Top*, v. 3, n. 2, p. 167–220, 1995.
- [9] Brasil. Decreto no 4074, de 4 de janeiro de 2002, jan. 2002.
- [10] Bruck, B. P.; Dos Santos, A. G.; Arroyo, J. E. C. Hybrid metaheuristic for the single vehicle routing problem with deliveries and selective pickups2012 IEEE Congress on Evolutionary

Computation. Anais...IEEE, jun. 2012. Disponível em:
<<http://ieeexplore.ieee.org/document/6256456/>>

[11] Dell'amico, M.; Righini, G.; salani, M. A Branch-and-Price Approach to the Vehicle Routing Problem with Simultaneous Distribution and Collection. *Transportation Science*, 2006.

[12] Fonseca, C. M.; Fleming, P. J.; Others. Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization. ICGA. Anais...1993

[13] Ford, H. Hoje e Amanhã. [s.l.] São Paulo, 1927.

[14] Gribkovskaia, I. et al. General solutions to the single vehicle routing problem with pickups and deliveries. *European Journal of Operational Research*, v. 180, n. 2, p. 568–584, 2007.

[15] Gribkovskaia, I.; Laporte, G.; Shyshou, A. The single vehicle routing problem with deliveries and selective pickups. *Comput. Oper. Res.*, v. 35, n. 9, p. 2908–2924, 2008.

[16] Gutiérrez-Jarpa, G. et al. A branch-and-price algorithm for the Vehicle Routing Problem with Deliveries, Selective Pickups and Time Windows.

European Journal of Operational Research, v. 206, n. 2, p. 341–349, 2010.

[17] Jozefowiez, N.; Semet, F.; Talbi, E.-G. Multi-objective vehicle routing problems. *European Journal of Operational Research*, v. 189, n. 2, p. 293–309, 2008.

[18] Lourenço, H. R.; Martin, O. C.; Stützle, T. Iterated Local Search. In: [s.l.] Kluwer Academic Publishers, 2003. p. 321–353.

[19] Montané, F. A. T.; Galvão, R. D. A tabu search algorithm for the vehicle routing problem with simultaneous pick-up and delivery service. *Computers & Operations Research*, v. 33, p. 595–619, 2006.

[20] Salhi, S.; Nagy, G. A cluster insertion heuristic for single and multiple depot vehicle routing problems with backhauling. *Journal of the Operational Research Society*, v. 50, p. 1034–42., 1999.

[21] Takahashi, R. H. C. Notas de Aula: Otimização Escalar e Vetorial, 2007.

[22] Zitzler, E.; Kunzli, S. Indicator-Based Selection in Multiobjective Search. In: [s.l.: s.n.]. p. 832–842.

Capítulo 17

MENSURAÇÃO DA EFICIÊNCIA EM SERVIÇOS ESPECIALIZADOS TERCEIRIZADOS A PARTIR DAS DIMENSÕES DE QUALIDADE VIA ANÁLISE ENVOLTÓRIA DE DADOS: UMA APLICAÇÃO NUMA ORGANIZAÇÃO DO SETOR AUTOMOTIVO DA REGIÃO SUL-FLUMINENSE

Daniel Maia de Souza

Pauli Adriano de Almada Garcia

Pítias Teodoro

Resumo: Ao longo dos últimos anos, as organizações vêm ampliando a utilização da terceirização de serviços para além de processos operacionais rotineiros, sendo cada vez mais utilizado para serviços especializados, que requerem dos prestadores de serviços habilidades altamente qualificadas, pensamento analítico e capacidade de julgamento. Entretanto, a avaliação de desempenho para os serviços especializados é mais restrita na literatura, pois apresenta maiores desafios devido à maior complexidade para identificação de medidas para sua avaliação. O presente trabalho apresenta a proposição de um procedimento para avaliação de serviços especializados terceirizados, por meio da aplicação de um modelo de análise envoltória de dados (DEA) adaptado para o tratamento de dados imprecisos, a fim de considerar as dimensões de qualidade de serviços da escala SERVPERF. A aplicação do procedimento foi realizada em uma organização do setor automotivo da região Sul-fluminense, que utiliza a terceirização para serviços especializados em diversas de suas áreas administrativas.

Palavras chave. Terceirização, SERVPERF, DEA.

1. INTRODUÇÃO

Mesmo quando uma operação produtiva é projetada, e suas atividades se encontram planejadas e controladas, a tarefa do gestor não está encerrada, pois todas as operações podem ter seu desempenho melhorado, uma vez que este pode ser identificado por meio de sua eficiência, que mede a utilização adequada dos recursos necessários para uma tarefa [Slack et al. 1999].

Para a melhor eficiência de um sistema econômico, as organizações podem organizar a produção de diferentes modos, de forma que parte da produção será realizada internamente, por meio da criação de estruturas hierárquicas, e outra parte será obtida por meio de transações junto ao mercado, onde a terceirização aparece como alternativa [Barney e Hesterly 1999]. A terceirização é uma forma de gestão do processo produtivo da empresa, por meio de um tipo de aliança em que uma empresa contratante transfere a realização de serviços específicos, executada pela empresa contratada com organização própria e autonomia técnica e jurídica [CNI 2014].

Para Mclvor [2013], o processo de terceirização pode proporcionar três benefícios para o negócio: 1) redução de custos; 2) especialização e 3) qualidade dos serviços. Para Gunasekaran et al. [2014], a terceirização reduz os custos de ativos, os custos de produção, e permite à empresa contratante o foco em suas competências centrais, possibilitando uma flexibilidade estratégica e redução de custos administrativos e de horas extras.

Hatonen e Eriksson [2009] informam que a adoção da estratégia de terceirização pelas organizações, a partir da década de 1980, possibilitou a evolução das práticas adotadas até o final da década de 2000. Este fenômeno pode ser dividido em duas fases principais: a primeira envolvia processos de manufatura com mão-de-obra intensiva, a segunda, e atual, expandiu de forma a incluir serviços especializados, que necessitam de profissionais intensivos em conhecimento, para a execução de atividades como pesquisa e desenvolvimento, por exemplo [Mclvor 2016], [Srivastava et al. 2012].

Dessa forma, a terceirização moveu-se para além de motivos de redução de custos, considerando seu uso como meio para

acesso a uma reserva mão-de-obra qualificada [Mclvor, 2016], [Hatonen e Eriksson, 2009]. Para Edvardsson e Durst [2014], a terceirização de serviços especializados – encontradas em processos de pesquisa e desenvolvimento, engenharia, gerenciamento de risco, planejamento financeiro, entre outros – requer dos prestadores de serviços habilidades altamente qualificadas, pensamento analítico e capacidade de julgamento.

Contudo, de acordo com um levantamento realizado junto a diversas empresas que utilizaram estratégias de terceirização, 44% dos respondentes não obtiveram redução de custos e apenas 34% estavam satisfeitos com a inovação das empresas terceirizadas, dentre outros resultados não desejáveis [Handley e Benton Jr. 2009]. Para Gunasekaran et al. [2014], muitas organizações experimentam dificuldades com a disponibilidade de produtos e serviços de alta qualidade a custos mínimos por não possuírem medidas apropriadas de desempenho.

Segundo Mclvor [2013], a gestão da terceirização de serviços especializados é mais complexa e desafiadora em comparação à gestão da terceirização de processos de manufatura, pois os serviços apresentam características únicas, como o envolvimento do cliente na entrega do serviço, exigindo a personalização e o contato físico, enquanto os processos de manufatura são mais suscetíveis à padronização e modularização. Para Srivastava et al. [2012], os dados operacionais gerados durante a terceirização de serviços complexos são relativamente desestruturados e não numéricos por natureza, o que requer um alto nível de envolvimento e iniciativas operacionais. Dessa forma, limitada atenção tem sido dada em como as organizações medem a qualidade do serviço em contratos de terceirização de serviços especializados [Mclvor 2013].

Dentro desse contexto, a presente pesquisa tem como intuito responder a seguinte pergunta problema: Como sistematizar a análise dos desempenhos de serviços especializados terceirizados?

Dessa forma, este artigo propõe um procedimento estruturado para avaliação do desempenho de serviços terceirizados especializados, capaz de identificar as empresas terceirizadas referências e os

pontos de melhoria para as empresas terceirizadas menos eficientes, visando permitir um ponto de partida para um processo de melhoria contínua das prestações de serviços especializados.

A seguir, o trabalho está organizado nas seguintes seções: a seção 2 consiste em uma análise das medidas disponíveis na literatura para avaliação de desempenho para serviços terceirizados especializados. A seção 3 consiste em uma análise dos modelos utilizados na literatura para avaliação de desempenho que podem ser aplicados aos serviços terceirizados especializados. Na seção 4 são apresentados os procedimentos metodológicos para a realização desta pesquisa. Na seção 5 são apresentados os resultados alcançados. Por fim, a seção 6 são realizadas as conclusões da pesquisa.

2. MEDIDAS PARA AVALIAÇÃO DE DESEMPENHO PARA SERVIÇOS TERCEIRIZADOS ESPECIALIZADOS

Para a análise de desempenho de processos de produção, as medidas de eficiência operacionais são aplicadas pois possuem maior tangibilidade, facilitando a definição de medidas quantitativas [Aguzzoul et al. 2014], [McIvor 2013], [Ho et al. 2010]. Na terceirização de tarefas rotineiras, as entregas podem ser descritas em termos de transações realizadas, enquanto que, em tarefas mais complexas, as entregas necessitam de maior profundidade e amplitude [Srivastava et al. 2012]. A identificação da qualidade de serviço é mais devido às quatro características distintivas dos serviços: intangibilidade, heterogeneidade, perecibilidade e inseparabilidade [Parasuraman et al. 1988]. Dessa forma, as medidas de qualidade de serviços são aplicadas para análise de desempenho de serviços que exigem maior contato entre fornecedor e cliente, o que confere um caráter intangível ao processo [Gunasekaran et al. 2014].

A mais conhecida e utilizada medida de qualidade de serviços é a escala SERVQUAL, desenvolvida em 1985 por Parasuraman, Zeithaml e Berry, seguida da escala SERVPERF, desenvolvida em 1992 por Cronin e Taylor [Faria et al. 2015], [Wang et al. 2015], [Ladhari 2009], sendo que, segundo McIvor [2013], alguns autores alegam que estas

medidas têm potencial para aplicação na avaliação em serviços especializados terceirizados.

A escala SERVQUAL [Parasuraman et al. 1988] mede a qualidade do serviço por meio da diferença entre as expectativas dos consumidores relacionados ao desempenho esperado do serviço e sua real percepção do serviço prestado. Para isso, são utilizados dois questionários, um para expectativa e um para desempenho, contendo vinte e duas questões distribuídas nas cinco dimensões da qualidade descritas a seguir, onde os respondentes indicam suas opiniões por meio de uma escala *Likert*, que varia de (1) “Discordo totalmente” a (7) “Concordo totalmente”:

- Tangíveis: aparência física da infraestrutura, equipamentos e pessoal;
- Confiabilidade: habilidade em executar o serviço prometido com acurácia e no tempo;
- Presteza: atitude em ajudar os clientes e prover rápido serviço;
- Garantia: conhecimento e cortesia dos funcionários, e sua habilidade em inspirar credibilidade;
- Empatia: nível de cuidado e atenção individualizada da empresa para os clientes.

Por outro lado, a escala SERVPERF [Cronin e Taylor, 1992] visa mensurar a qualidade dos serviços sem considerar a mensuração das expectativas, utilizando apenas a percepção do consumidor acerca do desempenho do serviço fornecido, que justificam ser o fator de maior influência nas intenções de compra, tornando os processos de avaliação, coleta e tabulação de dados com o emprego do SERVPERF mais rápidos em relação ao uso do SERVQUAL, visto que, embora considere as mesmas dimensões e questões da escala SERVQUAL, utiliza-se um único questionário.

Recentes pesquisas demonstram, por meio do crescente quantitativo de artigos publicados mundialmente entre o final da década de 1980 e o início da década de 2010, que o tema qualidade em serviços, especialmente com o emprego das escalas SERVQUAL ou SERVPERF, ainda é objeto de interesse de pesquisadores de relevância internacional [Faria et al. 2015], [Wang et al.

2015]. Segundo Ladhari [2009], entretanto, não há consenso na literatura a respeito da generalização das dimensões adotadas pelas escalas SERVQUAL e SERVPERF para a avaliação da qualidade de serviços, sendo uma das principais críticas relacionada à universalidade das cinco dimensões, para as quais seria mais adequado adaptá-las conforme as exigências da aplicação. Dentre as diferentes abordagens para as dimensões apresentadas, as que mais se diferenciam das dimensões apresentadas nas escalas SERVQUAL / SERVPERF são as que consideram os custos da prestação de serviços [Kang et al. 2012], [Handley e Benton Jr. 2009].

3. MODELO PARA AVALIAÇÃO DE DESEMPENHO DE SERVIÇOS TERCEIRIZADOS ESPECIALIZADOS

A fim de medir a qualidade de serviços, diversos modelos multicritérios de apoio à decisão (MCDM – Multi-criteria decision making) têm sido desenvolvidos, pois buscam agregar diferentes medidas em uma medida única capaz de auxiliar a tomada de decisão [Aguzzoul et al. 2014], [Gunasekaran et al. 2014], [Ho et al. 2010]. Para Lee e Kim [2014], dentre esses diferentes modelos, a análise envoltória de dados (DEA – Data Envelopment Analysis) não somente é capaz de agregar diferentes dimensões em uma única medida, mas possui como diferencial a possibilidade de prover direcionais de benchmarking para cada DMU ineficiente.

A Análise Envoltória de Dados tem origem no conceito de eficiência de Farrell [1957], onde um único insumo (*input*) gera um único produto (*output*) [Garcia 2013]. Com base neste conceito, Farrell [1957] identificou um entrave ao aumento da produtividade das organizações devido à inexistência de uma medida satisfatória de eficiência que pudesse combinar múltiplas entradas (*inputs*) e saídas (*outputs*) de diferentes métricas. Anos depois, Charnes et al. [1978] apresentaram uma técnica que atendesse esta necessidade, que futuramente seria intitulada como análise envoltória de dados (DEA – *Data Envelopment Analysis*). Fundamentado na programação linear, DEA é uma técnica não paramétrica para calcular a eficiência relativa de um conjunto comparável de unidades tomadoras de decisão (DMUs – *Decision Making Units*), os quais se distinguem no arranjo de entradas

e saídas de seus processos, sendo possível identificar aquelas com melhores produtividades, que formarão a fronteira eficiente [Cook e Seiford 2009], [Soares de Mello et al. 2005]. Além disso, segundo Soares de Mello et al. [2005], a técnica permite a medição do nível de eficiência das unidades que não estão na fronteira de eficiência, e a identificação das unidades referência (*benchmarks*) com as quais deverão ser comparadas para aumentarem seu desempenho.

Devido às oportunidades identificadas para práticas do DEA, diferentes modelos foram desenvolvidos na literatura ao longo das últimas décadas para se adaptarem a diversas aplicações [Cook e Seiford 2009].

Em geral, os modelos dispõem de duas orientações radiais possíveis: a *inputs*, que busca minimizar os recursos utilizados mantidos inalterados os níveis de produção; e a *outputs*, que implica no aumento da produção sem alterar as quantidades de insumos utilizadas.

Além disso, conforme característica dos modelos de programação linear, os modelos DEA dispõem de formulações primais e duais. O modelo primal, também conhecido como modelo de multiplicadores, caracteriza-se por obter, como resultado, um conjunto de pesos ótimos de cada *input* e *output*. Já na formulação dual, também conhecida como modelo de envelope, é possível identificar quais são as DMUs *benchmarks* para a DMU sob análise.

O modelo original é conhecido como CCR [Charnes et al. 1978] ou CRS (Constant *Return to Scale*), que denota proporcionalidade entre os valores de *inputs* e *outputs* [Soares de Mello et al. 2005]. Esse modelo está representado por (1).

$$\begin{aligned}
 & \text{Max } h_0 \\
 & \text{sujeito a:} \\
 & x_{i0} - \sum_{k=1}^n x_{ik} \lambda_k \geq 0, \forall i \\
 & -h_0 y_{j0} + \sum_{k=1}^n y_{jk} \lambda_k \geq 0, \forall j \\
 & \lambda_k \geq 0, \forall
 \end{aligned} \tag{1}$$

As variáveis de decisões, no modelo de envelope são h_0 e λ_k . h_0 é a eficiência ($h_0 = 1/\text{Eff}_0$) e λ_k é a contribuição da DMU_k na formação do alvo da DMU₀. h_0 representa por quanto todos os produtos devem ser multiplicados, mantendo-se constantes os recursos, para a DMU atingir a fronteira eficiente.

$$\begin{aligned} & \text{Max } h_0 \\ \text{sujeito a:} \\ & -h_0 y_{j0} + \sum_{k=1}^n y_{jk} \lambda_k \geq 0, \forall j \\ & \lambda_k \geq 0, \forall k \end{aligned} \quad (2)$$

Onde h_0 é o inverso da eficiência da DMU em análise (DMU₀); y_j é o j -ésimo *output* ($j=1, \dots, s$) da DMU_k ($k=1, \dots, n$); e λ_k é a contribuição de cada DMU na formação do alvo da DMU₀.

Entretanto, cabe ressaltar que a escala *Likert* utilizada no modelo SERVPERF trata-se de dados imprecisos. Dessa forma, é necessário a adaptação do modelo DEA para utilização desses dados.

Para a avaliação de dados imprecisos, Chen et al. [2015] desenvolveram o modelo DEA para aplicação com escala *Likert*, representado por (3):

$$\begin{aligned} & \text{Max } \frac{1}{s} \sum_{j=1}^s h_j \\ \text{sujeito a:} \end{aligned}$$

Para a análise da eficiência de unidades prestadoras de serviços, Lee e Kim [2014] aplicaram um modelo DEA de *input* unitário orientado a *outputs*, que considera as dimensões de qualidade de serviços. Este modelo DEA foi desenvolvido por Lovell e Pastor [1999], que, na formulação do envelope, é representado por (2):

$$\begin{aligned} & -\tilde{y}_{j0} + \sum_{k=1}^n y_{jk} \lambda_k \geq 0, \forall j \\ & x_{i0} - \sum_{k=1}^n x_{ik} \lambda_k \geq 0, \forall i \\ & \tilde{y}_{j0} \geq h_j y_{j0} \quad 1 \leq \tilde{y}_{j0} \leq L \\ & \tilde{y}_{j0} \text{ inteiro}, \quad j_{Lik} \in O_{Lik} \\ & \lambda_k \geq 0, \forall k \\ & h_j \geq 1, \forall j \end{aligned} \quad (3)$$

Onde $\tilde{y}_{j0}$ é variável inteira, O_{Lik} é a escala *Likert*, e L representa o número de níveis da escala *Likert* ($L = 7$ para a escala de sete pontos)

4. PROCEDIMENTOS METODOLÓGICOS

O procedimento proposto para análise de desempenho de serviços terceirizados especializados foi desenvolvido em cinco etapas, conforme Figura 1. Este procedimento foi aplicado em uma organização do setor automotivo da região Sul-fluminense, que utiliza a terceirização para serviços especializados em diversas de suas áreas administrativas.

Figura 1 – Procedimento para análise de desempenho de serviços especializados

Fonte: Elaborado pelo autor (2017)

Para a Etapa 1, foram selecionados os serviços especializados que participarão da análise. Para isso, optou-se pela seleção

intencional dos participantes. Segundo Creswell [2010], uma ideia subjacente à pesquisa qualitativa é a seleção intencional

de participantes que melhor ajudarão o pesquisador a entender problema e a questão de pesquisa. Dessa forma, os serviços serão selecionados por conveniência, de acordo com a facilidade de acesso aos gestores de contrato e às equipes de usuários.

Para a Etapa 2, a medida de desempenho de serviços especializados a ser utilizada na pesquisa será uma combinação entre as dimensões do modelo SERVPERF e os custos associados à prestação de serviço. O uso das dimensões de qualidade de serviços dos modelos SERVPERF se justifica por serem escalas mais utilizadas para avaliação de serviços [Ladhari 2009], podendo ser aplicado para serviços terceirizados [McIvor 2013]. A adição de critérios que consideram os custos da prestação de serviços se justificam por ser utilizado com frequência na literatura sobre serviços terceirizados [Kang et al. 2012], [Handley e Benton Jr. 2009].

Para a Etapa 3, foi disponibilizado aos gestores dos serviços terceirizados selecionados um questionário contendo as 22 questões do questionário da escala SERVPERF, além da questão adicional sobre custos, apresentada por Handley e Benton Jr. [2009]: “Atingiu ou excedeu as expectativas para os custos totais anuais”.

Para a Etapa 4, optou-se pelo desenvolvimento de um modelo DEA que considera o modelo de *input* unitário orientado a *outputs*, desenvolvido por Lovell e Pastor [1999], e utilizado por Lee e Kim [2014] para medidas de qualidade de serviços por meio do modelo SERVPERF. De forma a adaptar o modelo DEA de *input* unitário às condições de incertezas associadas à utilização de escala *Likert*, será proposta uma adaptação ao modelo (3) desenvolvido por Chen et al. [2015]. Para isso, foi eliminada a restrição referente ao somatório dos *inputs*.

$$\begin{aligned} & \text{Max } \frac{1}{s} \sum_{j=1}^s h_j \\ & \text{sujeito a:} \\ & -\tilde{y}_{j0} + \sum_{k=1}^n y_{jk} \lambda_k \geq 0, \forall j \\ & \tilde{y}_{j0} \geq h_j y_{j0} \\ & 1 \leq \tilde{y}_{j0} \leq L \\ & \tilde{y}_{j0} \text{ inteiro}, \quad j_{Lik} \in O_{Lik} \\ & \lambda_k \geq 0, \forall k \\ & h_j \geq 1, \forall j \end{aligned} \tag{4}$$

As DMUs a serem analisadas pelo modelo DEA proposto serão os contratos dos prestadores de serviços terceirizados obtido na Etapa 1.

O modelo DEA proposto considera como *outputs* as medidas de desempenho dos prestadores de serviços terceirizados, obtidos na Etapa 2 e seus respectivos valores obtidos na Etapa 3. São eles:

- Y1 Tangíveis
- Y2 Confiabilidade
- Y3 Presteza
- Y4 Segurança
- Y5 Empatia
- Y6 Custos

A Etapa 5 consiste na análise dos resultados descritos a seguir, obtidos após a aplicação do modelo DEA proposto na Etapa 4:

- Valores de eficiência para cada uma das DMUs em análise;
- Identificação das metas de melhorias para cada critério de cada DMU;
- Identificação das DMUs referências para cada DMU considerada menos eficiente

5. RESULTADOS OBTIDOS

Após aplicação das Etapas 1, 2 e 3, os valores do desempenho percebidos dos prestadores de serviços terceirizados, conforme avaliação de seus respectivos clientes estão descritos na Tabela 1.

Tabela 1 – Desempenho dos prestadores de serviços terceirizados segundo avaliação dos clientes.

DMU	Contrato	Y1	Y2	Y3	Y4	Y5	Y6
		Tangíveis	Confiabilidade	Presteza	Segurança	Empatia	Custos
1	Planejamento Logístico	6	4	1	2	3	3
2	Disposição de Materiais	2	5	5	6	6	2
3	Análise de Inventário	7	3	3	7	5	5
4	Custos Logísticos	5	6	4	5	2	6
5	Engenharia de Processos	1	5	6	5	6	3
6	Instalações Industriais	2	6	4	3	2	5
7	Custos de Manufatura	6	1	7	6	7	3
8	Sistemas Eletroeletrônicos	2	5	6	4	5	6
9	Desenho e Simulação	6	1	3	3	1	6
10	Homologação	6	2	5	4	4	1
11	Desenvolvimento de Fornecedores	5	3	1	7	3	5
12	Auditoria de Processos	2	1	2	6	1	5
13	Governança	5	4	2	3	1	7
14	Gestão de Projetos	4	3	4	7	5	5
15	Custeio Técnico	2	1	1	5	3	2

A partir da aplicação do modelo DEA proposto na Etapa 4, foi possível a obtenção dos seguintes resultados:

1 - Eficiência dos prestadores de serviço terceirizados

Os resultados encontrados para cada DMU estão descritos na Figura 2. Das 15 DMUs analisadas, 9 (60%) foram consideradas

eficientes, enquanto que 6 (40%) foram consideradas não eficientes. Nota-se que as DMUs 2, 3, 4, 5, 7, 8, 9, 13 e 14 foram consideradas eficientes, visto que possuem eficiência de 100%. Já as DMUs 1, 6, 10, 11, 12 e 15 são consideradas relativamente menos eficientes, visto que possuem eficiência menor que 100%.

Figura 2 – Eficiência dos Prestadores de Serviço Terceirizados

Fonte: Elaborado pelo autor (2017)

2 - Identificação de metas de melhoria

Outro resultado extraído por meio do modelo matemático é a identificação de metas de melhorias, que são os valores de alvo para as DMUs consideradas menos eficientes. Estes resultados são apresentados na Tabela 2.

Para as DMUs consideradas eficientes, uma vez que os valores de alvos são idênticos aos seus respectivos valores originais, seus dados foram excluídos da Tabela 2.

Para as DMUs consideradas menos eficientes, é necessário que os valores de

alvo destacados na Tabela 2 sejam atingidos para que possam ser consideradas eficientes. Para o estudo de caso, os seguintes resultados foram obtidos:

Para a DMU 1, deve-se melhorar os *outputs* Y3, Y4, Y5 e Y6 em 300%, 200%, 33% e 67%, respectivamente.

- Para a DMU 6, deve-se melhorar os *outputs* Y1, Y4, Y6 em 150%, 67% e 20%, respectivamente.

- Para a DMU 1, deve-se melhorar os *outputs* Y3, Y4, Y5 e Y6 em 300%, 200%, 33% e 67%, respectivamente.

3 - Identificação das DMUS referências

A partir dos valores encontrados para λ_k , é possível identificar a contribuição de cada DMU_k na formação do alvo da DMU_0 . Dessa forma, a partir dos valores destacados na Tabela 3, são identificadas as DMUs de referência para cada DMU considerada menos eficiente. Nota-se que as DMUs

consideradas eficientes possuem elas mesmas como referência.

- As DMU 1 e 10 possuem como referência as DMUs 3, 4 e 7.
- A DMU 6 possui como referência a DMU 4.
- As DMUs 11 e 12 possuem como referência a DMU 3.
- A DMU 15 possui como referência as DMUs 3, 4, 6, 7 e 8.

Tabela 2 – Identificação de alvos para as DMUs

DMU	Y1 Tangíveis	Y2 Confiab.	Y3 Presteza	Y4 Segurança	Y5 Empatia	Y6 Custos
DMU 1 ORIGINAL	6	4	1	2	3	3
DMU 1 ALVO	6	4	4	6	4	5
DMU 1 %	0,0%	0,0%	300,0%	200,0%	33,3%	66,7%
DMU 6 ORIGINAL	2	6	4	3	2	5
DMU 6 ALVO	5	6	4	5	2	6
DMU 6 %	150,0%	0,0%	0,0%	66,7%	0,0%	20,0%
DMU 10 ORIGINAL	6	2	5	4	4	1
DMU 10 ALVO	6	3	5	6	5	4
DMU 10 %	0,0%	50,0%	0,0%	50,0%	25,0%	300,0%
DMU 11 ORIGINAL	5	3	1	7	3	5
DMU 11 ALVO	7	3	3	7	5	5
DMU 11 %	40,0%	0,0%	200,0%	0,0%	66,7%	0,0%
DMU 12 ORIGINAL	2	1	2	6	1	5
DMU 12 ALVO	7	3	3	7	5	5
DMU 12 %	250,0%	200,0%	50,0%	16,7%	400,0%	0,0%
DMU 15 ORIGINAL	2	1	1	5	3	2
DMU 15 ALVO	5	4	5	5	4	5
DMU 15 %	150,0%	300,0%	400,0%	0,0%	33,3%	150,0%

Fonte: Elaborado pelo autor (2017)

Tabela 3 – Identificação de DMUs referências

DMU	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}
1	0,00	0,00	0,43	0,43	0,00	0,00	0,14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
6	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
7	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
8	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
9	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00
10	0,00	0,00	0,29	0,29	0,00	0,00	0,43	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
11	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
12	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
13	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00
14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
15	0,00	0,00	0,16	0,38	0,00	0,03	0,27	0,17	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Fonte: Elaborado pelo autor (2017)

6. CONCLUSÕES

Com base nos resultados obtidos nesta pesquisa, foi possível atingir o propósito de desenvolvimento de um procedimento para análise de serviços terceirizados especializados. Para isso, um modelo de análise envoltória de dados (DEA) foi adaptado, de forma a considerar a imprecisão de dados qualitativos existentes nas dimensões consideradas nas dimensões utilizadas de qualidade de serviços e custos.

Dentre os principais benefícios com a abordagem proposta, espera-se poder contribuir para o processo de melhoria contínua da prestação de serviços especializados junto às organizações, em específico às associadas ao setor automotivo brasileiro. Além disso, foi possível contribuir para a ainda restrita literatura sobre o tema de terceirização de serviços especializados, especialmente com referência a procedimentos de avaliação de desempenho.

De forma a superar as limitações da pesquisa, sugere-se, como recomendações futuras, o desenvolvimento de modelo DEA para melhor discriminação das eficiências, visto que grande parte das DMUs analisadas foram consideradas eficientes. Além disso, recomenda-se a aplicação em outros casos práticos que possam aprimorar o desenvolvimento do procedimento proposto. De forma complementar, sugere-se avaliar, junto aos gestores dos contratos e usuários finais, suas práticas de gestão sobre as empresas prestadoras de serviços, que podem ser analisados em conjunto com o levantamento da qualidade de serviços, de forma a identificar os fatores direcionadores do desempenho das empresas terceirizadas. Por fim, comparando estes resultados com os pontos de melhoria identificados para as prestadoras de serviço menos eficientes, poderá ser proposta a adequação das práticas de gestão por parte dos clientes.

REFERÊNCIAS

- [1] Aguezoul, A. Third-party logistics selection problem: A literature review on criteria and methods. *Omega*, v. 49, p. 69–78, 2014.
- [2] Barney, J. B.; Hesterly, W. Economia das organizações: entendendo a relação entre as organizações e a análise econômica. In: CLEGG, S. R. et al. *Handbook de Estudos Organizacionais*. São Paulo: Atlas, 1999.
- [3] Charnes, A.; Cooper, W. W.; Rhodes, E. Measuring efficiency of decision making units. *European Journal of Operational Research*, v. 2, n. 6, p. 429–444, 1978.
- [4] Chen, Y.; Cook, W.; Du, J.; Hu, H.; Zhu, J. Bounded and discrete data and Likert scales in data envelopment analysis: Application to regional energy efficiency in China. *Annals of Operations R*
- [5] Confederação Nacional da Indústria (CNI). Terceirização: o imperativo das mudanças. In: *Propostas da indústria eleições 2014*, Brasília. Vol. 8, 67 p.
- [6] Cook, W. D; Seiford, L. M. Data envelopment analysis (DEA) – Thirty years on. *European Journal of Operational Research*, v. 192, p. 1–17, 2009.
- [7] Creswell, J. W. *Projeto de pesquisa: métodos qualitativo, quantitativo e misto*. Porto Alegre: Artmed, 3.ed., 2010.
- [8] Cronin, J. J.; Taylor, A. S. Measuring service quality: A reexamination and an extension. *Journal of Marketing*, v. 56, n. 3, p. 55–67, 1992.
- [9] Edvardsson, I. R.; Durst, S. Outsourcing of knowledge processes: a literature review. *Journal of Knowledge Management*, v. 18, n. 4, p. 795–811, 8 jul. 2014.
- [10] Faria, L. F. V.; Freitas, A. L. P.; Molina-Palma, M. A. Qualidade em serviços sob a perspectiva dos métodos SERVQUAL e SERVPERF: um estudo bibliométrico. *GEPROS. Gestão da Produção, Operações e Sistemas*, Bauru, Ano 10, nº 3, jul-set/2015, p. 53-67
- [11] Farrell, M. J. The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, v. 120, n. 3, p. 253-290, 1957.
- [12] Garcia, P. A. A. Uma abordagem via análise envoltória de dados para o estabelecimento de melhorias em segurança baseadas na FMEA. *Gestão & Produção*, v. 20, n. 1, p. 87-97, 2013.
- [13] Gunasekaran, A. et al. Performance measures and metrics in outsourcing decisions: A review for research and applications. *International Journal of Production Economics*, v. 161, p. 153–166, 2015.
- [14] Handley, S. M.; Benton, W. C. Unlocking the business outsourcing process model. *Journal of Operations Management*, v. 27, n. 5, p. 344–361, 2009.
- [15] Hatonen, J.; Eriksson, T. 30+ years of research and practice of outsourcing - Exploring the past and anticipating the future. *Journal of International Management*, v. 15, n. 2, p. 142–155, 2009.

- [16] HO, W.; XU, X.; DEY, P. K. Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of Operational Research*, v. 202, n. 1, p. 16–24, 2010.
- [17] Kang, M. et al. Aligning organizational control practices with competitive outsourcing performance. *Journal of Business Research*, v. 65, n. 8, p. 1195–1201, 2012.
- [18] LADHARI, R. A review of twenty years of SERVQUAL research. *International Journal of Quality and Service Sciences*, v. 1, n. 2, p. 172–198, 2009.
- [19] Lee, H; Kim, C. Benchmarking of service quality with data employment analysis. *Expert Systems with Applications*, v. 41, p. 3761-3768, 2014.
- [20] Lovell, C. A. K.; Pastor, J. T. Radial DEA models without inputs or without outputs. *European Journal of Operational Research*, v. 118, n. 1, p. 46–51, 1999.
- [21] Mcivor, R. What do we know about services outsourcing? Edinburgh: ICAS Insight, 2013.
- [22] Mcivor, R. An Analysis of the Application of Process Improvement Techniques in Business Process Outsourcing. *International Journal of Quality & Reliability Management*, v. 33, n. 3, 2016.
- [23] Parasuraman, A.; Zeithaml, V. A.; Berry, L. L. Servqual: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, v. 64, n. 1, p. 12-40, 1988.
- [24] Slack, N., Chambers, S.; Johnston, R., *Administração da Produção*. 1º Ed. São Paulo: Atlas S.A, 1999.
- [25] Soares de Mello, J.C.C.B. et al. Curso de Análise Envoltória de Dados. In: XXXVII SBPO – Simpósio Brasileiro de Pesquisa Operacional, 2005, Gramado. Anais, p.2520-2547.
- [26] Srivastava, V.; Sharfuddin, A.; Datta, S. Managing quality in outsourcing of high-end services: a conceptual model. *Total Quality Management & Business Excellence*, v. 23, n. 11-12, p. 1315–1327, 2012.
- [27] Wang, L.Y; Luor, T.; Luarn, P.; LU, H. P. Contribution and Trend to Quality Research - a literature review of Servqual model from 1998 to 2013. *Informatica Economică*, v. 19, n. 1, p. 34–46, 2015.

Capítulo 18

ANÁLISE NUMÉRICA EM UM PROBLEMA DE PROGRAMAÇÃO LINEAR COM A UTILIZAÇÃO DO MÉTODO DOS MULTIPLICADORES DE LAGRANGE

Rafaela Boeira Cechin

Leandro Luís Corso

Resumo: Os Multiplicadores de Lagrange (ML), dentro do processo de otimização, apresentam a possibilidade de realizar análises de sensibilidade pós-otimização, presente no conceito de Programação Linear. O objetivo foi analisar os possíveis ganhos dentro de otimização com a utilização das informações obtidas com os ML. Para isso, foi estudado um artigo do ramo alimentício, que visa utilizar a Programação Linear a fim de reduzir o custo total do misturado milheto-soja, e que foi publicado em um periódico científico. Então foi sugerido alterações no modelo matemático, utilizando o método dos ML, que desta forma, foi possível reduzir mais de 16% do custo total do produto estudado no artigo.

Palavras chave: Pesquisa Operacional, Programação Linear, Multiplicadores de Lagrange, Alimentação.

1. INTRODUÇÃO

Os estudos da Pesquisa Operacional (PO) iniciaram-se na década de 40 e a sua primeira contribuição ocorreu durante a Segunda Guerra Mundial (CHAVES, 2011). Após este momento, a utilização da PO foi tão satisfatória que decidiram dar continuidade nos estudos e expandiram para as mais diversas áreas. Kersting, Mladenov e Tokmakov (2017) comentam que há dois critérios que possibilitaram no avanço da PO: o método simplex para resolução de problemas na Programação Linear, e o surgimento dos computadores, que tornou possível a resolução de cálculos complexos com um grande volume de dados.

Dentro da Pesquisa Operacional, há os estudos de Programação Linear (PL), que, conforme Ignácio (2009) e Kersting, Mladenov e Tokmakov (2017) comentam, são problemas que possuem constantes não negativas e somente funções matemáticas lineares, além de serem utilizados para maximizar os lucros ou minimizar os custos, por exemplo. Para estes problemas de PL, o método simplex é um procedimento eficiente e altamente utilizado para a resolução, pois ele permite a utilização de milhares de variáveis.

Após a obtenção dos resultados com a aplicação do método simplex, é possível fazer a análise de sensibilidade, em que Bre et al. (2016) define como sendo uma alteração nos parâmetros do problema, dentro de certo limite, sem alterar a solução encontrada. Para esta análise, pode ser utilizado os Multiplicadores de Lagrange (ML), método utilizado neste artigo para a análise numérica.

Como citado, o primeiro passo da análise numérica deste artigo é a escolha de uma PL apresentada na literatura científica. Em seguida, foi gerada a programação no software MatLab, a fim de replicar e analisar os resultados apresentados em um modelo matemático apresentado por Balasubramanian et al. (2012). A partir do modelo de PL, foram calculados os valores dos ML, assim como os limites de variação das constantes do lado direito das restrições, além de interpretar o significado destes valores no problema. Para finalizar, foram sugeridas modificações no modelo matemático, visando um melhoramento do artigo escolhido na literatura.

2. MULTIPLICADORES DE LAGRANGE

De acordo com a sugestão de Hillier e Lieberman (2005), o método dos Multiplicadores de Lagrange pode ser utilizado quando há a necessidade de realizar uma otimização condicionada com diversas restrições de igualdade. Para determinar a função lagrangeana, antes é necessário definir a função objetivo e as restrições, respectivamente representadas nas equações 1 e 2.

$$Z = c_1x_1 + c_2x_2 + \dots + c_nx_n \quad (1)$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m \\ x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 \end{cases} \quad (2)$$

Onde, Hillier e Lieberman (2005) consideram que:

- a) $Z \rightarrow$ valor da medida do desempenho total;
- b) $x_n \rightarrow$ variáveis de decisão;
- c) $c_n, b_m, a_{mn} \rightarrow$ constantes de entrada ou ainda, parâmetro do modelo;
- d) restrição $m \rightarrow$ restrição funcional; e
- e) $x_n \geq 0 \rightarrow$ restrição não negativa.

Para os autores, as restrições também podem ser representadas de forma padronizada, como mostra a equação 3.

$$\begin{cases} g_1(x) = b_1 \\ g_2(x) = b_2 \\ \vdots \\ g_m(x) = b_m \\ \text{para } m < n \end{cases} \quad (3)$$

Deste modo, a função lagrangeana é apresentada na equação 4.

$$L(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i * [g_i(x) - b_i] \quad (4)$$

Para esta equação, Hillier e Lieberman (2005) *conceituam que* $L(x, \lambda)$ representa a Função de Lagrange e λ , o Multiplicador de Lagrange. Sendo $g_i(x) - b_i = 0$ para todos valores de i . Para descobrir o valor de cada λ_i , necessita-se derivar a equação $L(x, \lambda)$.

Para exemplificar, consideram-se somente duas variáveis (x_1 e x_2) para a função objetivo. Ou seja, $f(x_1, x_2)$ será a função objetivo a ser maximizada, por exemplo. Ainda, será arbitrada somente uma restrição: $g(x_1, x_2) = 0$. Utilizando-se desta definição, a função lagrangeana ficará de acordo com a equação 5.

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda * g(x_1, x_2) \quad (5)$$

De forma que, ao derivar parcialmente a função lagrangeana, para encontrar os valores de x_1, x_2, λ , obtêm-se os resultados apresentados na equação 6.

$$\begin{cases} \frac{\delta L}{\delta x_1} = \frac{\delta [f(x_1, x_2)]}{\delta x_1} - \frac{\delta [\lambda * g(x_1, x_2)]}{\delta x_1} = 0 & (6) \\ \frac{\delta L}{\delta x_2} = \frac{\delta [f(x_1, x_2)]}{\delta x_2} - \frac{\delta [\lambda * g(x_1, x_2)]}{\delta x_2} = 0 & (7) \\ \frac{\delta L}{\delta \lambda} = \frac{\delta [f(x_1, x_2)]}{\delta \lambda} - \frac{\delta [\lambda * g(x_1, x_2)]}{\delta \lambda} = 0 & (8) \end{cases}$$

Onde $\delta L[x_1, x_2, \lambda] = 0$ e os resultados de x_1 e x_2 representam graficamente o ponto ótimo da função objetivo.

Ainda, os ML possuem uma interpretação econômica, porém com o mesmo significado matemático, chamado de Preço Sombra. Boussemart, Leleu e Shen (2017) afirmam que este conceito é usado em análise econômica de custo/benefício. Estes autores também comentam sobre Preço Sombra ao citar que, após as alterações nos valores das restrições, a solução otimizada originalmente passará a ser super otimizada.

3. CÁLCULOS E ANÁLISE NUMÉRICA

A fim de realizar a análise numérica, foi escolhido um artigo publicado no endereço

eletrônico do PubMed, que faz parte da Biblioteca Nacional de Medicina Norte Americana e há milhares de artigos disponíveis para pesquisa e referentes a diversos assuntos, por exemplo biomedicina, saúde, ciências químicas, agricultura e alimentação, bioengenharia, entre outras áreas. A escolha foi realizada com base na Programação Linear, as variáveis, os parâmetros do problema e os resultados apresentados, de forma que fosse possível a reprodução dos resultados.

Assim, foi selecionado o artigo dos autores Balasubramanian, Singh, Patil e Onkar, e seu título é Avaliação da qualidade da extrusão do misturado de milho-soja, formulado por meio da Programação Linear. Este artigo foi publicado em 2012 na revista Journal of Food Science and Technology, e refere-se ao estudo feito de uma Programação Linear, a fim de reduzir os custos do produto final, porém mantendo diversos fatores de aceitação dos clientes, como por exemplo, densidade, variação da cor, dureza, textura, sabor, entre outros (BALASUBRAMANIAN et al., 2012). A função objetivo do problema descrito pelos autores está representada na equação 9, e representa a minimização do custo da mistura. As variáveis do modelo são a quantidade do misturado milho-soja, que consiste na combinação de três diferentes produtos: mexoeira, grão de soja decortado e capim pé-de-galinha.

$$\text{Minimizar } Z = C_i X_i \quad (9)$$

Os autores fazem as seguintes correlações:

Z é a função matemática que representa o custo total da mistura;

C_i é o custo por quilograma de cada ingrediente; e

X_i é o vetor que representa a proporção de cada ingrediente utilizado na mistura

Os autores basearam-se nas informações da Tabela 1 para formular sua Programação Linear. Está apresentado o custo unitário, quantidade de proteínas, energia e cálcio para cada uma das variáveis.

Tabela 1 – Quantidades dos ingredientes para elaboração da PL

Variável	Ingrediente	Custo ¹ (Rs/kg)	Proteína (%)	Energia (kcal)	Cálcio (mg)
x_1	Mexoeira	7	11,6	361	42
x_2	Grão de soja decorticado	20	43,2	432	240
x_3	Capim pé-de-galinha	8	7,3	328	344

¹ Preços de Dezembro de 2009

Fonte: Adaptado de Balasubramanian et al. (2012)

É importante ressaltar que o custo é medido em rupias por quilograma (Rs/kg), que é a moeda oficial da Índia, país onde o artigo foi publicado. Sendo assim, o modelo matemático da PL é apresentado na Tabela 5,

sendo que os coeficientes da função objetivo foram utilizados conforme sugestão dos autores.

Tabela 2 – Programação Linear apresentado por Balasubramanian et al. (2012)

Função objetivo: $\text{Min } Z = 0,7x_1 + 2x_2 + 0,8x_3$ (10)			
Restrições	Equações		Resultados testes
Proteínas (g/d)	$11,6x_1 + 43,2x_2 + 7,3x_3 = 20$	(11)	Verificado ($20 = 20$)
Valor calorífico (kcal/d)	$361x_1 + 432x_2 + 328x_3 \geq 800$	(12)	Verificado ($843 \geq 800$)
Cálcio (mg/d)	$42x_1 + 240x_2 + 344x_3 = 135$	(13)	Verificado ($135 = 135$)
Tamanho do produto	$1x_1 + 1x_2 + 1x_3 = 1,5$	(14)	Verificado ($1,5 = 1,5$)
Limitações	$0x_1 + 1x_2 + 0x_3 = 0,15$	(15)	Verificado ($0,1053 \geq 0,15$)
			$Z = 1,2039$
			$x_1 = 1,2253$
			$x_2 = 0,1053$
			$x_3 = 0,1694$

Fonte: Adaptado de Balasubramanian et al. (2012)

Ao observar a Tabela 2, é possível perceber que todas as equações e inequações são lineares, atendendo à primeira característica imposta pelos modelos de otimização linear. A segunda propriedade também é atendida, já que nenhuma restrição é não negativa. No MatLab, a função linprog permite cálculos de inequações, porém com sinal de menor ou igual. Assim, a equação 12 é modificada para a equação 16, apresentada a seguir.

$$-361x_1 - 432x_2 - 328x_3 \leq -800 \quad (16)$$

A segunda etapa deste artigo consiste na reprodução destes valores no software MatLab. Porém, algumas alterações nas restrições são necessárias para possibilitar os resultados apresentados por Balasubramanian et al. (2012). Isso se deve pelo motivo de as equações apresentadas pelos autores violarem as restrições que os mesmos estabeleceram no artigo, ou seja, as respostas não estão de acordo com o modelo matemático apresentado. Por esse motivo, as equações 16 e 15 devem ser modificadas para, respectivamente, as equações 17 e 18, para estarem de acordo.

$$-361x_1 - 432x_2 - 328x_3 \leq -500 \quad (17)$$

$$0x_1 + 1x_2 + 0x_3 \leq 0,15 \quad (18)$$

Desta forma, o modelo matemático se torna possível de solução para os parâmetros apresentados. O modelo de Programação Linear reescrito está representado na equação 19.

$$\text{Minimizar } Z = 0,7x_1 + 2x_2 + 0,8x_3$$

Sujeito às restrições:

$$\begin{cases} 11,6x_1 + 43,2x_2 + 7,3x_3 = 20 \\ -361x_1 - 432x_2 - 328x_3 \leq -500 \\ 42x_1 + 240x_2 + 344x_3 = 135 \\ 1x_1 + 1x_2 + 1x_3 = 1,5 \\ 0x_1 + 1x_2 + 0x_3 \leq 0,15 \end{cases} \quad (19)$$

Como terceira etapa, após as devidas alterações, foi possível reproduzir os valores de Z , x_1 , x_2 , e x_3 encontrados no artigo dos autores. A Tabela 3 apresenta os valores calculados com a otimização.

Tabela 3 – Valores encontrados com a otimização

X		FVAL		Multiplicadores de Lagrange				
1,2253	0,1053	0,1694	1,2039	Inequações		Equações		
				2,03E-15	6,27E-15	-0,0359	-8,42E-04	-0,2486

Fonte: Adaptado do MatLab (2015)

Os valores das variáveis (x) que a Tabela 6 apresenta significam a quantidade necessária, em kg, de cada ingrediente, sendo que x_1 corresponde para a quantidade de mexoeira, x_2 para o grão de soja decortado e x_3 para o capim pé-de-galinha; e o valor da função objetivo ($fval$) representa o custo, em rupias, para a produção do produto final. Estes resultados podem ser interpretados como a solução otimizada para o problema, levando em consideração as restrições impostas.

Ao resolver o modelo matemático linear, é possível também se obter os valores para os Multiplicadores de Lagrange. Desta forma, obtendo-se cinco valores, pois há um valor dos ML para cada restrição da programação,

e neste artigo, havia duas inequações e três equações. Para as restrições das inequações, o valor dos ML obtido foi abaixo da tolerância de $e-8$ utilizada por linprog, os quais se assumiram como zero para este trabalho. Por isso, pode-se considerar que estas restrições são inativas. Para as restrições das equações, como os ML são diferentes de zero, estas são consideradas ativas.

Para a aplicação da teoria do Preço Sombra, é necessário conhecer os limites inferiores e superiores das constantes do lado direito das restrições, apresentados na Tabela 4. Estes valores foram obtidos com o software SciPy e são os limites inferiores e superiores do lado direito das restrições (também chamados de valor de segundo membro).

Tabela 4 – Variação das constantes do lado direito das restrições

Restrição	Limite inferior	Valor atual	Limite superior
1	16,7374	20	21,53715
2	-543,3	-500	$+\infty$
3	79,292	135	242,997
4	1,37394	1,5	1,79719
5	0,10533	0,15	$+\infty$

Fonte: Adaptado de SciPy (2015)

Os valores encontrados na Tabela 6 para os ML significam o quanto impactaria o lucro máximo se fosse adicionado uma unidade de cada recurso, ou seja, exemplificando com a primeira restrição, para cada grama a menos de proteína disponível como recurso, o custo da função objetivo diminui em 0,0359 rupias. Porém, este conceito só se aplica se a variação dos valores das constantes do lado direito das restrições, variar dentro dos limites inferiores e superiores, que foram apresentados na Tabela 3. Da mesma forma que, no exemplo anterior, a grama reduzida de proteína só reduzirá o custo em 0,0359 rupias se a quantidade de proteína não diminuir além de 16,7374g.

Como explicado anteriormente, as inequações possuem restrições inativas. Por este motivo, os limites superiores da segunda

e da quinta restrição são $+\infty$. Quanto aos coeficientes da função objetivo, eles podem ser alterados para $-\infty$ e $+\infty$, sem alterar o valor ótimo das variáveis. Isso ocorre devido às igualdades nas restrições, as quais limitam o espaço de soluções.

O valor da função objetivo da Tabela 6 representa o valor otimizado do problema. Para finalizar a análise numérica deste artigo, é sugerido algumas alterações, para tornar o resultado super otimizado, que o seu impacto no custo está apresentado na Tabela 5. Como o objetivo do trabalho de Balasubramanian et al. (2012) é a minimização do custo do produto, então foi sugerido que alterassem as constantes do lado direito das restrições ativas para os seus respectivos limites inferiores, conforme também mostrado na tabela a seguir.

Tabela 5 – Modificações sugeridas

Restrição	Valor dado no problema	Valor sugerido	ML	Impacto no custo
1	20	16,7374	0,0359	0,11712734
2	-500	Restrição inativa	0	0
3	135	79,292	0,00084178	0,04689388
4	1,5	1,37394	0,2486	0,031338516
5	0,15	Restrição inativa	0	0
			Total:	0,195359736

Fonte: Elaborado pelos autores (2015)

Assim, a Programação Linear ficou de acordo com a equação 20.

$$\begin{aligned}
 &\text{Minimizar } Z = 0,7x_1 + 2x_2 + 0,8x_3 \\
 &\text{Sujeito às restrições:} \\
 &\begin{cases} 11,6x_1 + 43,2x_2 + 7,3x_3 = 16,7374 \\ -361x_1 - 432x_2 - 328x_3 \leq -500 \\ 42x_1 + 240x_2 + 344x_3 = 79,292 \\ 1x_1 + 1x_2 + 1x_3 = 1,37394 \\ 0x_1 + 1x_2 + 0x_3 \leq 0,15 \end{cases} \quad (20)
 \end{aligned}$$

Então, foi calculado o quanto impactaria no custo, ou seja, foi calculada a diferença de valores entre o dado no problema e o sugerido, e multiplicado pelos seus respectivos ML. Para as inequações, não foram sugeridas mudanças, pois como seus

ML são iguais à zero, estas restrições não impactam no custo final do produto.

O total deste impacto calculado foi de aproximadamente 0,1954 rupias, ou seja, o custo do produto foi reduzido para 1,0085 rupias. Concluindo, o custo da solução otimizada ficou de 1,2039 rupias; e custo da solução super otimizada é de 1,0085 rupias; significando uma redução de 16,23% no custo.

4. CONCLUSÃO

É possível utilizar a PO em complexas aplicações e, uma das vantagens é a possibilidade de testar teoricamente diversos cenários, por simulações computacionais. Assim, evitam-se gastos desnecessários para uma empresa, já que uma análise

computacional não possui riscos econômicos, podendo ser viável analisar a melhor alternativa para determinado problema da organização e só então, aplicá-lo. A PO fornece o melhor cenário possível para um problema, e a exploração dos ML podem ainda melhorar este cenário. Em contrapartida, este método ainda não é muito utilizado nas empresas, infelizmente.

Além, foi evidenciada a assistência que os *softwares* matemáticos proporcionaram para este tipo de modelagem matemática, uma vez que validaram os resultados do artigo e ainda forneceram informações adicionais. Pelas alterações realizadas na PL, foi possível perceber a importância de se ter uma

programação bem formulada, pois somente assim pode-se confiar no resultado obtido.

Utilizando-se do exemplo estudado neste artigo, pode-se observar uma redução no custo do produto de mais de 16%, alterando somente os parâmetros da Programação Linear. Considerando a aplicação deste método em uma empresa real, ainda há diversos dados que poderiam ser alterados, como por exemplo, custo da matéria-prima, layout do setor de fabricação, estoque, entre outros; parâmetros em que os ML também podem ser aplicados e otimizados.

REFERÊNCIAS

- [1] BALASUBRAMANIAN, S. et al. Quality evaluation of millet-soy blended extrudates formulated through linear programming. *Journal of food science and technology*, v. 49, n. 4, p. 450-458, 2012.
- [2] BOUSSEMARY, Jean-Philippe; LELEU, Hervé; SHEN, Zhiyang. Worldwide carbon shadow prices during 1990–2011. *Energy Policy*, v. 109, p. 288-296, 2017.
- [3] BRE, Facundo; SILVA, Arthur Santos; GHISI, Enedir; FACHINOTTI, Victor. Residential building design optimization using sensitivity analysis and genetic algorithm. *Energy and Buildings*, v. 133, p. 853-866, 2016.
- [4] CHAVES, Viviane Hengler Corrêa. Perspectivas históricas da Pesquisa Operacional. 2011, 118 f. Dissertação (Mestrado em Educação

Matemática) – Universidade Estadual Paulista, Programa de Pós-Graduação em Educação Matemática, Rio Claro, 2011.

- [5] HILLIER, Frederick; LIEBERMAN, Gerald. *Introduction to Operations Research*. 8ª ed. Nova Iorque: McGraw-Hill.

[6] IGNÁCIO, Bruno Alvarez Ferreira. Desenvolvimento de um modelo de Programação Linear para apoio à tomada de decisão em uma cadeia de suprimentos. 2009, 195 f. Dissertação (Mestrado Acadêmico em Engenharia Mecânica) – Centro Universitário da FEI, São Bernardo do Campo, 2009.

- [7] KERSTING, Kristian; MLADENOV, Martin; TOKMAKOV, Pavel. Relational linear programming. *Artificial Intelligence*, v. 244, p. 188-216, 2017.

Capítulo 19

APLICAÇÃO DO CONTROLE ESTATÍSTICO NOS PROCESSOS DECOLAGENS DE VOOS REGULARES DOS AEROPORTOS BRASILEIROS.

Luiz Antonio Figueiredo

Jose Elmo Menezes

Ricardo Luiz Machado

Resumo: A utilização do controle estatístico de processo – CEP - possibilita monitorar o desempenho das operações aeroportuárias. Melhores processos com menos variabilidade propiciam níveis melhores de qualidade nos resultados das operações aeroportuárias. No ambiente aeroportuário, podem ocorrer atrasos, antecipações ou cancelamentos enquanto voos estão no ar ou em solo, propagando-se ao longo do sistema devido às interdependências entre diferentes operações regulares. Este estudo teve como objetivo aplicar técnicas de controle estatístico de processo na análise de dados referentes a operações de decolagens no período de 2013 a 2015 em aeroportos brasileiros. Os dados analisados foram obtidos no banco de dados da Agência Nacional de Aviação Civil – ANAC, através de análises em relatórios regulares ativos de voos – VRA e horários de transportes de voos – HOTRAN. Para alcançar o objetivo este trabalho foi dividido em quatro seções. Os resultados obtidos nos Gráficos de Controle referente aos anos de 2013 a 2015, a variável do mês de maio se localizam fora dos limites de controle três-sigma superior em 2015.

*Artigo originalmente publicado nos anais do Congresso Nacional de Excelência em Gestão 2016

1 INTRODUÇÃO

1.1 O SISTEMA AEROPORTUÁRIO

De acordo com Horonjeff & Mckelvey (2010, p. 01), desde o início do século XX, a aviação civil tornou-se um dos setores mais fascinantes, importantes e complexos do mundo. No Brasil, segundo o Relatório da Agência Nacional de Aviação Civil (ANAC 2015, p. 4), foram transportados cerca de 85 milhões de passageiros no mercado doméstico em 2015. No mercado internacional foram transportados aproximadamente 7 milhões de passageiros por empresas brasileiras em 2015.

O estudo do sistema aeroportuário é geralmente dividido em duas categorias: aéreo e terrestre. A categoria do espaço aéreo envolve o planejamento e a administração da acomodação do movimento de aeronaves no aeroporto, tratando das chegadas e partidas de voos. A categoria do espaço terrestre do aeroporto lida com o planejamento e a administração da acomodação do movimento de veículos, passageiros e cargas em solo, abrangendo o terminal aeroportuário seus componentes de acesso. A figura 1 apresenta as duas categorias (YOUNG, 2014, p. 239).

Figura 01: Os componentes de um Aeroporto

Fonte: Young (2014, p. 100)

Na Figura 01 observa-se, no conjunto de operações com passageiros, que o fluxo logístico inicia com o acesso ao aeroporto, passando por seu estacionamento e prosseguindo até as instalações do terminal aeroportuário. A partir deste ponto inicia-se o conjunto de operações com as aeronaves, concluindo o processo com a decolagem. O fluxo reverso caracteriza a chegada de passageiros ao aeroporto pela rota aérea.

Segundo Young (2014, p. 410), a movimentação eficiente de aeronaves e passageiros nos aeroportos é de extrema importância e depende de duas características-chave das operações: a demanda por serviço por parte dos passageiros e de operadores de aeronaves e a capacidade do aeroporto, tanto em espaço aéreo quanto em ambiente local. Para mensurar os níveis de atividades, de serviços e de investimentos nos aeroportos, comumente são utilizados indicadores

relativos ao número de passageiros atendidos, quantidade de carga transportada e o número de operações aéreas horárias praticadas no aeroporto.

1.2 CAPACIDADE AEROPORTUÁRIA

Segundo Ashford (2011, 234), a capacidade do aeroporto é um elemento essencial da análise de plano diretor de um aeroporto, que é um componente importante do processo e uma parte crítica do planejamento aeroportuário e avaliação do desempenho das operações do terminal como um todo. Young (2014, p. 412) define a capacidade aeroportuária como uma quantidade prática máxima de operações que o sistema é capaz de atender dentro de determinado período de tempo.

O termo capacidade, em seu sentido mais geral, refere-se à medida quantitativa para prestação de serviço de uma instalação de processamento para acomodar a demanda sustentada, ou seja, o máximo rendimento ou prestação de serviço durante um período especificado de tempo sob condições de serviço (ASHFORD, 2011, 235).

De acordo com Stevenson (2011, p. 157) por uma série de razões as decisões sobre capacidade estão entre as mais fundamentais que gerentes têm que tomar como as decisões que envolvem a capacidade e o impacto sobre como a organização terá que atender a demanda futura para produtos e serviços, decisões sobre a capacidade que afeta os custos operacionais, decisões se a capacidade é um determinante fundamental do custo inicial, as decisões sobre a capacidade que envolve frequentemente o comprometimento de recursos a longo prazo, as decisões sobre se capacidade pode influenciar na competitividade da empresa.

Segundo Young (2014, p. 413), embora a capacidade aeroportuária normalmente se relacione à capacidade de lidar com operações de aeronaves, há outras áreas de operação em um aeroporto em que parâmetros alternativos de capacidade são igualmente importantes. Por exemplo, a movimentação eficiente de passageiros por pontos dentro do terminal aeroportuário é determinada, em parte, pela capacidade de processamento de passageiros dos locais dentro dele e pelo número de automóveis capazes de desembarcar passageiros nessa

entrada, mensurado pela capacidade veicular.

Para Ashford (2011, p. 234,), as análises de capacidade do aeroporto são realizadas para atender a três propósitos: o comportamento da demanda, a capacidade operacional dos componentes do sistema aeroportuário para lidar com fluxos de passageiros e aeronaves projetados e a estimativa dos atrasos e outras manifestações de capacidade-constrangimento experimentadas no sistema em diferentes níveis de demanda.

Para Young (2014, P. 413), existem na verdade, duas definições comumente usadas para descrever a capacidade aeroportuária: a capacidade máxima e a capacidade teórica. A capacidade máxima é a definida como a taxa máxima em que as operações com aeronaves podem ser praticadas sem levar em conta pequenos atrasos que possam ocorrer como resultado de imperfeições operacionais ou pequenos eventos aleatórios. A capacidade prática é entendida como o número de operações que podem ser acomodadas ao longo do tempo sem que ultrapasse uma quantidade finita de atrasos, expressas geralmente em termos de média máxima de atrasos aceitáveis.

A FAA – Federal Aviation Administration define dois parâmetros de capacidade prática para avaliar a eficiência das operações aeroportuárias: capacidade prática horária (PHOCAP – *practical hourly capacity*) e a capacidade prática anual (PANCAP – *practical anual capacity*). Estes parâmetros são definidos pela FAA, como as quantidades máximas de operações que podem ser realizadas em um aeroporto, em termos horários e anuais, possuindo como restrição para aceitação no máximo quatro minutos de atraso, em média, durante duas horas de maior congestionamento, chamadas de picos, por hora e por ano, respectivamente (YOUNG, 2014, p. 413).

O efeito direto do desequilíbrio entre a capacidade e a demanda é o congestionamento do aeroporto e o atraso de voos. Segundo Aisling & Kenneth (1999, p. 01), muitos aeroportos possuem problemas significativos de atraso, como consequência da falta de equilíbrio entre capacidade e demanda. Ashford (2011) discute o tema, sugerindo que os atrasos de voos são obviamente frustrantes para os viajantes do transporte aéreo e caros para as companhias,

que representam os clientes mais importantes para os terminais aeroportuários.

Desta forma, este estudo tem por objetivo aplicar a ferramenta de Controle Estatístico de Processo (CEP) nos dados de decolagens dos anos de 2013 a 2015 dos Aeroportos brasileiros através das informações contidas de voo realizados, cancelamentos, atrasos e antecipações nos horários de decolagens apresentados com justificativa no VRA- Voo Regular Ativo dos anos de 2013, 2014 e 2015 e comparar sua evolução quanto a variabilidade do processo através das cartas de controle, que são as ferramentas principais utilizadas no controle estatístico de processo e têm como objetivo detectar desvios de parâmetros representativos do processo.

Com a finalidade de alcançar o objetivo este trabalho foi dividido em quatro seções. A primeira seção com o levantamento da literatura a respeito das operações aeroportuárias, capacidade, Fatores que impactam a capacidade e nível de serviço aeroportuário, a segunda seção com o levantamento literário a respeito de controle estatístico de processo e cartas de controle e estabilidade do processo. A terceira seção com a Metodologia, aplicação das cartas de controle nos dados de decolagens dos voos regulares dos aeroportos brasileiros e a quarta seção com a análise dos dados e a conclusão.

1.3 FATORES IMPACTANTES NA CAPACIDADE AEROPORTUÁRIA

A capacidade de um aeródromo não é constante. Ela varia consideravelmente com base em determinados fatores, incluindo a utilização das pistas de pouso, o tipo de aeronave operada, o percentual de operações de decolagem e pouso realizadas, as condições climáticas e as regulamentações que prescrevem o uso de pistas de pouso (YOUNG, 2014, p. 414).

Segundo Ashford (2011, 238), fatores que influenciam a capacidade de um sistema de pista são numerosos e estão relacionados comumente com o controle de tráfego aéreo, as características da demanda, as condições ambientais nas proximidades do aeroporto e o layout e design do sistema de pista.

1.4 NÍVEL DE SERVIÇO

Segundo Ballou (1993, p.73), o nível de serviço logístico é a qualidade com que o fluxo de bens e serviços é gerenciado. É o resultado líquido de todos os esforços logísticos da firma no desempenho oferecido pelos fornecedores aos seus clientes no atendimento dos pedidos. O nível de serviço prestado ao cliente é um fator chave do conjunto de valores logísticos que empresas oferecem aos seus clientes para assegurar sua fidelidade. Fleury (2000, p.56) sustenta que o produto do serviço logístico é o atendimento ao cliente, ou seja, o serviço prestado. As organizações investem em pessoal e equipamentos, selecionam fornecedores e usam tecnologia de informação, com o objetivo de utilizar em prática um processo logístico capaz de diferenciá-las criando valor para os clientes por meio de um serviço superior. Além disso, existe atualmente maior ênfase no serviço devido ao maior interesse pelo consumidor. Muitas firmas agora apoiam o conceito de marketing que coloca o foco principal nas necessidades e desejos do cliente, sendo o nível de serviço um dos mais importantes elementos da satisfação do consumidor. Dessa forma, segundo Ballou (2004, p.79), o serviço oferecido representa grande número de fatores individuais, muitos dos quais estão sob controle da função logística. Estes fatores foram classificados de acordo com sua relação e a transação do produto, ou seja, elementos de pré-transação, transação e pós-transação conforme apresentado na Tabela 01:

Tabela 01: Elementos do nível de serviço

NÍVEL DE SERVIÇO PRESTADO AO CLIENTE		
Elementos de Pré-transação	Elementos de Transação	Elementos de pós-transação
Política posta por escrito Política nas mãos dos cliente Estrutura organizacional	Nível de estoque Habilidade no trato de atrasos Elementos do ciclo do pedido Transbordo Precisão Conveniência do pedido Substituibilidade do produto	Instalação, garantias, reparos, peças de reposição, Rastreamento do produto Queixas e reclamações

Fonte: Ballou (2004, p. 79)

No contexto específico da avaliação do nível de serviço em uma operação aeroportuária, o Conselho Internacional de Aeroportos (2000) e Seneviratne e Martel (1991, p. 03) sugeriram a utilização das seguintes variáveis: tempo de espera, tempo de processamento, tempo de caminhada, distância de caminhada, mudança de nível, nível de orientação/informação, disponibilidade de espaço, disponibilidade de espaço e números de lugares para carros.

Segundo Kazda (2007, p. 251), a FAA – *Federal Aviation Administration* (1988) e a IATA – *International Air Transport Association* (2004) desenvolveram procedimentos e modelos para o cálculo necessário de áreas para cada atividade dentro do terminal. A necessidade refere-se à quantidade de espaço disponível por passageiro na hora

ocupada do espaço aeroportuário, que resulta em um determinado nível de serviço (LOS - *Level of Service*) com uma classificação de nível de conforto variando de “A” a “F”, com base no fluxo de hora ocupado e padrões aceitáveis para o tempo de enfileiramento e aglomeração, medidos por meio do número total de passageiros tratados em um determinado momento, que chegam, partem ou são transferidos. No cálculo é imprescindível ser específico em relação à taxa de fluxo de passageiros em qualquer atividade ou instalação. As diretrizes: para a quantidade de espaço por passageiro na hora ocupada e o resultado em termos do nível de serviço (LOS- *Level of Service*) são apresentados na Tabela 2

Tabela 2: Nível de Serviço – Espaço a ser fornecido para passageiros em diferentes funções

Atividade	Simulação	Nível de serviço padrão (LOS) [M²/Passageiros]					
		A	B	C	D	E	F
Esperando e circulando	Movimentação livre	2.7	2.3	1.9	1.5	1.0	Menos
Aérea de reivindicação de bagagem	Movimentação com bagagens	2.0	1.8	1.6	1.4	1.2	Menos
Filas no Check-in	Filas, com bagagens	1.8	1.6	1.4	1.2	1.0	Menos
Portão da sala: área de inspeção do governo	Filas, sem bagagens	1.4	1.2	1.0	0.8	0.6	Menos

Fonte: KASDA (2007, p. 242)

Para a construção do terminal como um todo, a FAA– *Federal Aviation Administration* recomenda 14 metros quadrados por passageiro no período de pico, considerando a área bruta de construção de um terminal doméstico. A IATA *International Air Transport Association* recomenda uma área de 25 a 30 metros quadrados, considerando terminais domésticos e fretados (KAZDA, 2007, p. 252).

2 CONTROLE ESTATÍSTICO DE PROCESSO

O controle estatístico de processo – CEP, ligado ao controle de qualidade moderno, surgiu nos Estados Unidos nos anos 20 na Companhia Telefônica Bell. Em 1924, Walter Shewhart elaborou a primeira carta de controle e um uso racional para o processo de monitoramento e controle (MACGREGOR, 1995, p. 403).

A ideia do CEP é que melhores processos de produção com menos variabilidade propiciam níveis melhores de qualidade nos resultados da produção. Nesse sentido, o controle de um processo é totalmente diferente de sua inspeção, que tem como objetivo a eliminação de peças de baixa qualidade que não alcançam as expectativas dos consumidores. Para se controlar e estabilizar os processos da empresa utilizam-se ferramentas de CEP, que necessitam apenas de pequenas amostras. Assim, as investigações estarão orientadas para as grandes irregularidades da linha de produção (CARVALHO, 2005, p. 261).

Segundo Filho apud Arient et al. (2005, p. 1976), a verificação de que o processo está ou não sob controle é feita pelo exame de amostras extraídas periodicamente. No processo sob controle, a característica de qualidade do conjunto dos itens produzidos possui distribuição normal. Quando a variabilidade se torna anormal, as amostras indicarão que o processo de fabricação se modificou e ficou fora de controle.

Torne-se necessário, portanto, sustentar este esforço com técnicas que possam facilitar a análise e o processo de tomada de decisão. Neste ambiente se enquadram as ferramentas para aprimoramento da qualidade – FAQ, como meio de facilitar o trabalho daqueles que são responsáveis pela condução de um

processo de planejamento ou análise e solução de problemas (OLIVEIRA, 1995, p. 07). Segundo Montgomery (2004, p. 96), o CEP é um grupo de ferramentas de resolução de problemas de grande utilidade na obtenção da estabilidade do processo e na melhoria da capacidade.

2.1 CARTAS DE CONTROLE

Vieira et al. (2002, p. 01) afirmam que dois produtos nunca serão idênticos em função da variabilidade de produção e que esta característica leva à definição de qualidade. Dessa forma, a qualidade pode ser atingida com a redução da variabilidade e pode ser analisada por meio do controle estatístico do processo.

O gráfico de controle (GC), também conhecido como gráfico de controle de Shewhart, é uma técnica estatística usada para medição e análise do comportamento de um processo (MONTGOMERY, 2004, p. 403). Este controle visa garantir a qualidade do produto através de um conjunto de atividades executadas sobre o processo (JURAN, 1993, p. 285). O principal objetivo do gráfico de controle, segundo Gaither (2002, p.14) é apontar quando os processos de produção podem ser alterados o suficiente para afetar a qualidade do produto.

Nas palavras de Devore (2015, p 601), os gráficos de controle são empregados extensivamente como um conceito técnico de diagnóstico para o monitoramento dos processos de produção e de serviços a fim de identificar instabilidade e circunstâncias não usuais.

As cartas ou gráficos de controle consistem em uma linha central (LC), um par de limites de controle, um dos quais se localiza abaixo (LIC) e outro acima da linha central (LSC). (KUME, 1993; VIEIRA, 1999). Os valores característicos marcados no gráfico representam amostras coletadas em um processo, conforme apresentado na figura 02. Os elementos das cartas de controle são determinados pelas equações 1,2 e 3, sendo μ a média do processo e σ o desvio padrão.

Figura 2. Representação gráfica de processos sob controle e fora de controle

Fonte: Carneiro Neto (2003)

Na análise dos parâmetros das cartas de controle, ao se considerar a média amostral para uma amostra aleatória de tamanho n , selecionada em um instante de tempo particular, é altamente provável que, em um processo sob controle, a média amostral, apresente valores dentro do intervalo de ± 3 desvios padrão (σ) da média do processo μ (DEVORE, 2015, p. 603).

O gráfico $(\bar{X})$ baseado em valores de parâmetros conhecidos tendo a incerteza em relação ao valor aleatório (v.a.). Para um processo sob controle, X tem uma distribuição normal com valor médio μ e desvio padrão σ sendo que $(\bar{X})$ indica a média amostral para uma amostra aleatória de tamanho n , sabendo que:

$$E(\bar{X}) = \mu$$

$$\mu = \sigma/\sqrt{n}$$

$(\bar{X})$ Tem distribuição normal

$$\bar{P}(\mu - 3\sigma \leq \bar{X} \leq \mu + 3\sigma) = P(-3,00 \leq Z \leq 3,00) = 0$$

Se todos esses valores lançados na carta de controle estiverem dentro dos limites de controle sem apresentarem qualquer tendência particular e a disposição dos pontos dentro dos limites for aleatória, o processo é considerado sob controle. Entretanto, se os pontos incidirem fora dos limites de controle ou apresentarem uma disposição atípica, o processo é julgado fora de controle. Segundo as palavras de Ramos (2000, p. 130), em um processo

estatisticamente estável os pontos nos gráficos de controle devem distribuir-se aleatoriamente em torno da linha média sem que haja padrões estranhos como tendências crescentes ou decrescentes, ciclos, estratificações ou misturas, além de pontos fora dos limites de controle. Segundo Devore (2015, p. 602), é importante compreender que processo sob controle não é sinônimo de atingir as especificações de projeto ou tolerância. Um processo sob controle é

simplesmente aquele cujo comportamento em relação às variações é estável durante o tempo.

É importante destacar que um gráfico de controle não permite a identificação de quais são as causas especiais de variação que estão atuando em um processo fora de controle estatístico, mas ele processa e dispõe informações que podem ser utilizadas na identificação destas causas (WERKEMA, 1995, p. 197). No processo de construção desses limites são usados comumente três desvios padrões, o que pode ser justificado pelos bons resultados obtidos na prática (KONRATH, 2002, p. 23).

Os gráficos de controle podem ser classificados em dois tipos gerais. Se a característica da qualidade pode ser expressa como um número em alguma escala contínua de medida, ela é usualmente, chamada de uma variável e seus gráficos são chamados de gráficos de controle para

variáveis. Apresentados como controle da média do processo, usualmente construídos através das médias, o gráfico X ou a variabilidade do processo pode ser monitorado tanto através do gráfico de controle para o desvio padrão, chamado gráfico S como para o gráfico de amplitude, chamado gráfico R. (LACHMAN et al., 2001, p. 1379). Mas quando as características das variáveis não são medidas em escala contínua ou em uma escala quantitativa, nestes casos, pode-se julgar cada unidade de produto como conforme ou não conforme. Os gráficos com tais características são chamados gráficos de controle para atributos (MONTGOMERY, 2004, p. 177). Nos gráficos de controle X-barra e S-barra não se aproximam do correto se a distribuição das variáveis não for normal. Na prática em geral os valores de μ e σ não são conhecidos. Sendo m o número de amostras então esses devem ser estimados conforme a equação 5:

$$\begin{aligned} x - barra(i) &= \frac{\pm x_1 + x_2 + x_3 \dots + x_m}{m} \\ s - barra(i) &= \frac{\pm x_1 + x_2 + x_3 \dots + x_m}{m} \end{aligned} \quad (5)$$

Neste caso para a construção de gráficos em amostras para testar a hipótese de controle de processo segundo Montgomery (2004,

p.131), são utilizadas as fórmulas com base em dados passados, sem nenhum valor de referência conforme a tabela 03.

Tabela 03: formulas de gráfico de controle baseado em dados passados sem valor de referência

Gráfico X-Barra	Gráfico S-Barra
Limites	Limites
Limite Superior: LSC = X-barra + A ₂ * R-barra	Limite Superior: LSC = B ₄ * S-barra
Linha de centro: LC = R-barra	Linha de centro: LC = S-barra
Limite Inferior: LSC = X-barra – A ₂ * R-barra	Limite Inferior: LSC = B ₃ * S-barra

Fonte: Montgomery (2004, p. 131)

Muitas características da qualidade não podem ser representadas numericamente, em tais casos são classificados como itens conformes ou não conformes, ou seja, defeituosos ou não-defeituosos sendo tais características chamadas de atributos. Segundo Devore (2014, p 616) a expressão

atributos é utilizada na literatura para descrever duas situações: uma que cada item produzido é defeituoso ou não-defeituoso; e a segunda situação que um único item pode ter um ou mais defeitos e o número de defeitos é determinado.

Desta forma, Montgomery (2004, p. 177) apresenta alguns gráficos de controle de atributos, como o gráfico chamado p , ou Gráfico de Controle, para a fração não-conforme, que trabalha com a proporção de itens não conformes e o gráfico chamado

gráfico C , de controle para não conformidades por unidade, quando se é conveniente trabalhar com o número de defeitos ou não conformidades, apresentados na tabela 04, com as formulas para cálculos de seus respectivos limites.

Tabela 04: Formulas para cálculo dos LSC, LIC e LC dos gráficos ' p ' e ' C '

Gráfico C de Controle	Gráfico P de Controle
Limites	Limites
Limite Superior: $LSC = \bar{c} + 3 * \sqrt{\bar{c}}$	Limite Superior: $LSC = \bar{p} + 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$
Linha de centro: $LC = \bar{c}$	Linha de centro: $LC = \bar{p} = \frac{\sum_{i=1}^m D_i}{\bar{n}}$, onde D_i é número de itens não conforme na i -ésima amostra e $\bar{n}$ é o número médio de amostra.
Limite Inferior: $LIC = \bar{c} - 3 * \sqrt{\bar{c}}$	Limite Inferior: $LIC = \bar{p} - 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$

Fonte: Montgomery (2004, p. 158)

2.2 ESTABILIDADE DO PROCESSO

Sendo sigma a letra de origem Grega que representa o desvio padrão conforme Downing (2006).

Para Montgomery (2004), uma amostra é considerada fora de controle se:

- Um dos pontos se localiza fora dos limites de controle três-sigma;
- Dois, em três pontos consecutivos, se localizam além dos limites de alerta de dois sigmas;
- Quatro, em cinco pontos consecutivos, se localizam a uma distância de um sigma ou mais em relação à linha central;
- Oito pontos consecutivos se localizam de um mesmo lado da linha central.

Após verificar a estabilidade do processo, é necessário empregar os índices de capacidade, que, segundo Montgomery (2004), as principais utilizações da análise de dados de uma análise de capacidade são: definir até que ponto o processo manterá as tolerâncias; auxiliar a elaboração do produto ou modificações; especificar exigências de desempenho; planejar a sequência de produção e reduzir variabilidade. A verificação da capacidade do processo é verificada pelo cálculo do parâmetro C_p , onde $\hat{\sigma}$ é o desvio padrão da distribuição, $\bar{R}$ a amplitude média e d_2 a média da distribuição relativa, definido pelas equações 6 e 7:

$$C_p = \frac{LSC - LIC}{6\hat{\sigma}} \quad (5)$$

$$\hat{\sigma} = \frac{\bar{R}}{d_2} \quad (6)$$

$$\hat{\sigma} = \frac{\bar{R}}{d_2} \quad (7)$$

O índice C_p avalia a distância da média do processo aos limites da especificação, tomando aquela que for menor, e, portanto, mais crítica em termos de chances de serem produzidos itens fora de especificação. Se $C_p > 1$ temos um processo capaz.

3. METODOLOGIA

Considerando que este trabalho se refere ao Controle Estatístico de Processo procurou-se empreender uma pesquisa que abrangesse os campos teórico e prático da estatística, com enfoque na aplicação da análise nos atrasos de voos regulares no Brasil.

Como sendo método nas palavras de Lakatos (2010, p. 06) o grupo das atividades sistemáticas e coerentes que, com maior segurança e eficiência, permite alcançar o objetivo, em concordância com Filho (2012, p. 106) que define como sendo um curso ou ordem que se submete qualquer espécie de atividade, com a finalidade de alcançar um fim determinado.

Foi utilizado, assim, o método quantitativo quando foram analisados elementos mensuráveis, garantindo a precisão dos resultados e o método qualitativo devido à capacidade deste método de avaliar a complexidade de um problema. Segundo Regino (2003, p. 72), “o método qualitativo é uma forma de entender a natureza de um fenômeno social. Fenômenos que não podem ser investigados por meio da metodologia quantitativa”.

Conforme Regino (2003, p. 71), deve-se utilizar dados quantitativos quando a pesquisa lidar com fatores mensuráveis passíveis de tratamento numérico, sendo o método quantitativo melhor indicação, pois se caracteriza pelo emprego da quantificação, tanto na coleta quanto no tratamento de informações.

Para Creswell (2009, p. 18), as vantagens de combinar as abordagens quantitativa e qualitativa são: Proporcionar vantagens que compensam os pontos fracos de ambas as abordagens; Prover evidências mais abrangentes para o estudo de um problema de pesquisa do que cada abordagem isoladamente; Ajudar a responder a questões que não podem ser respondidas por abordagem separadamente; Encorajar os pesquisadores a colaborarem, superando

relacionamentos contraditórios quantitativos e qualitativos; Encorajar o uso de ponto de vista múltiplos ou concepções metodológicas mais do que uma simples associação típica das concepções dos pesquisadores quantitativos e qualitativos; Ser prática no sentido de que o pesquisador é livre para usar todos os métodos possíveis para solucionar o problema de pesquisa.

Para elaboração deste trabalho foi utilizado dados disponibilizados pela ANAC nos relatórios regulares ativos de voos – VRA e horários de transportes de voos – HOTRAN dos anos de 2013, 2014 e 2015. Dados disponíveis com a movimentação diária dos voos regulares agrupados mensalmente, com data, hora revista de decolagem, hora da decolagem, companhia aérea e justificativas de atrasos, cancelamentos e antecipações de voos.

As seguintes etapas foram realizadas para a conclusão deste trabalho, iniciando-se pelo levantamento dos dados de 2013, 2014 e 2015 disponibilizados pela ANAC e após com o levantamento literário a respeito das operações aeroportuária, controle estatístico de processos e legislações a respeito de atrasos, na sequência com a análise e a classificação dos dados, separação dos dados mês a mês dos voos com atrasos e identificação dos atrasos superiores a 4 horas, aplicação do controle estatístico de processo através das cartas de controle, análise das cartas de controle e Conclusão.

3.1 APLICAÇÃO DO CEP NA ANÁLISE DE ATRASOS DE DECOLAGENS EM AEROPORTOS BRASILEIROS

As aplicações não industriais não diferem substancialmente das aplicações industriais do CEP. Como exemplo, os gráficos de não conformidade podem ser aplicados para reduzir os erros de faturamento nas operações de cartão de crédito como também os gráficos de controle de variáveis podem ser utilizados para monitorar e controlar o tempo de circulação em contas em operações financeiras (MONTGOMERY, 2004, 190).

Foram utilizadas neste trabalho dados fornecidos pela ANAC através de relatórios de voo dos anos de 2015, 2014 e 2013 VRA- Voo Regular Ativo com informações contidas provenientes do cruzamento dos dados

constantes na base de dados de Horário de Transporte – HOTRAN, que contém a relação diária dos voos previstos para serem realizados, com informações de atrasos. Para construção das cartas de controle de Atributos ‘P’, foram utilizados dados retrospectivos dos anos de 2015, 2014 e 2013 apresentados na tabela 05, 06 e 07, nos meses de junho e julho do ano de 2014 utilizadas as médias dos respectivos meses dos anos de 2015 e 2013 devido à realização de auditoria nos dados pela ANAC. E através

do software Excel e o Suplemento *Action* nos referidos dados apresentado com as formulas da tabela 04 para cartas de controle “P”, como também a formula (6) para determinação da capacidade, resultando nas informações da tabela 08 e 09 e Gráficos de Controle “P” 01, 02 e 03, visando assim avaliar o estado de controle do processo e obtenção dos limites de controle destes para utilização dos mesmos em comparações futuras.

Tabela 05: Descrição dos voos referente ao ano de 2015

Voos do ano 2015	Mês Jan.	Mês Fev.	Mês Mar.	Mês Abr.	Mês Mai	Mês Jun.	Mês Jul.	Mês Ago.	Mês Set.	Mês Out.	Mês Nov.	Mês Dez.
Voos Realizados no Horário Previsto	78916	68769	76904	73330	74724	63488	66054	65480	61262	62761	58902	64982
Voos Cancelados	11345	10993	14373	14316	14430	11438	8917	9643	9334	8938	8135	7492
Voos de manutenção, retorno etc	195	182	204	180	209	168	203	183	225	234	222	229
Voos Antecipados	6131	6922	7345	6896	7482	15390	16360	18771	17467	18338	16595	15259
Voos com Atraso Superior a 4 horas	10	4	9	12	22	17	20	15	8	15	7	27
Voos com Atraso Inferior a 4 horas	18039	15123	12172	12186	11487	15197	20600	13952	19122	18860	20522	23827
Total de voos	114636	101993	111007	106920	108354	105698	112154	108044	107418	109146	104383	111816

Fonte: VRA – ANAC

Tabela 06: Descrição dos voos referente ao ano de 2014

Voos do ano 2014	Mês Jan.	Mês Fev.	Mês Mar.	Mês Abr.	Mês Mai	Mês Jun.	Mês Jul.	Mês Ago.	Mês Set.	Mês Out.	Mês Nov.	Mês Dez.
Voos Realizados no Horário Previsto	72825	66950	70647	67316	68577	65183	69852	76042	74598	77948	71100	73178
Voos Cancelados	8615	9505	11852	11572	12996	11102	8930	10388	10208	10814	11203	11332
Voos de manutenção, retorno etc	192	160	198	163	222	175	180	210	162	245	226	213
Voos Antecipados	6574	4779	6550	5960	6186	10079	11038	7226	7299	7156	6735	5865
Voos com Atraso Superior a 4 horas	10	7	3	7	5	13	15	14	11	4	12	15
Voos com Atraso Inferior a 4 horas	22801	15677	17694	19449	19207	17606	20642	13453	12443	14640	17366	23161
Total de voos	111017	97078	106944	104467	107193	104158	110657	107333	104721	110807	106642	113764

Fonte: VRA – ANAC

Tabela 07: Descrição dos voos referente ao ano de 2013

Voos do ano 2013	Mês Jan.	Mês Fev.	Mês Mar.	Mês Abr.	Mês Mai	Mês Jun.	Mês Jul.	Mês Ago.	Mês Set.	Mês Out.	Mês Nov.	Mês Dez.
Voos Realizados no Horário Previsto	75150	68636	72128	72014	72080	66879	73651	75284	71640	80655	66539	68011
Voos Cancelados	6954	7056	8932	8098	9578	10766	8943	9134	8292	8401	7969	9377
Voos de manutenção, retorno etc	101	72	131	99	119	182	158	149	119	136	118	171
Voos Antecipados	5681	6176	8068	6729	5600	4769	5717	7476	7162	9156	6619	5550
Voos com Atraso Superior a 4 horas	9	7	8	2	3	9	10	8	12	1	2	10
Voos com Atraso Inferior a 4 horas	21739	17161	16206	14686	18685	20016	20684	13311	13057	16680	17208	25365
Total de voos	109634	99108	105473	101628	106065	102621	109163	105362	100282	115029	98455	108484

Fonte: VRA – ANAC

Gráfico 1: Gráfico 'P' de Controle para o Proporção de voos com atrasos inferiores a 4 horas em 2013

Gráfico 2: Gráfico 'P' de Controle para o Proporção de voos com atrasos inferiores a 4 horas em 2014

Gráfico 3: Gráfico 'P' de Controle para o Proporção de voos com atrasos inferiores a 4 horas em

Gráfico de proporções para voos com atrasos inferiores a 4 horas

2015

Fonte: Autor

Tabela 8: LSC, LIC, LC apresentados nos gráficos de controle para proporção de voos com atrasos inferiores a 4 horas, anos 2013, 2014 e 2015.

Resultados			
Gráfico "P" de Controle voos com atrasos inferiores a 4 horas			
Limites	Ano 2013	Ano 2014	Ano 2015
Limite Superior: $LSC = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$	0,173776	0,170091	0,157788
Linha de centro: $LC = \bar{p} = \frac{\sum_{i=1}^m D_i}{\bar{n}}$	0,170298	0,166673672	0,154496
Limite Inferior: $LIC = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$	0,16682	0,163256722	0,151204

Fonte: Autor

Tabela 9: Itens Defeituosos

Proporção de atrasos			
Meses	2013	2014	2015
Janeiro	0,206823	0,212964	0,166313
Fevereiro	0,163269	0,146425	0,139429
Março	0,154183	0,165264	0,112222
Abril	0,139722	0,181656	0,112351
Maio	0,177768	0,179396	0,105906
Junho	0,190431	0,164442	0,140111
Julho	0,196786	0,192799	0,189925
Agosto	0,12664	0,125653	0,128633
Setembro	0,124223	0,116219	0,176298
Outubro	0,158692	0,136739	0,173883
Novembro	0,163716	0,162201	0,189206
Dezembro	0,241321	0,216327	0,219677

Fonte: Autor

Pelo fato da Resolução nº 141, de 9 de março de 2010 da ANAC, que Dispõe sobre as Condições Gerais de Transporte aplicáveis aos atrasos e cancelamentos de voos e às hipóteses de preterição de passageiros que passageiros e dá outras providências, cita que, o transportador, ao constatar que o voo irá atrasar em relação ao horário originalmente programado, sendo que em

caso de atraso no aeroporto de partida ou de escala por mais de 4 (quatro) horas, o transportador deverá oferecer acomodação ou reembolso do valor integral pago pelo bilhete de passagem não utilizado, incluídas as tarifas aos passageiros. Faz-se necessário construir as cartas de controles P de atributos nos dados relativos aos atrasos superiores a 4 horas.

Gráfico 4: Gráfico de Atributos “p” dos atrasos com mais de 04 horas do ano de 2015

Gráfico 5: Gráfico de Atributos “P” dos atrasos com mais de 04 horas do ano de 2014

Gráfico 6: Gráfico de Atributos “p” dos atrasos com mais de 04 horas do ano de 2013

Fonte: Autor

Tabela 10: LSC, LIC, LC Gráfico P de Controle atrasos com mais de 4 horas apresentados nos gráficos de controle para não-conformidades dos anos 2013, 2014 e 2015

Resultados			
Gráfico P de Controle atrasos com mais de 4 horas			
Limites	Ano 2013	Ano 2014	Ano 2015
Limite Superior: $LSC = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$	0,000138	0,000177	0,00023
Linha de centro: $LC = \bar{p} = \frac{\sum_{i=1}^m D_i}{\bar{n}}$	6,42E-05	9,03E-05	0,000128
Limite Inferior LIC $= \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$	0	3,17E-06	2,47E-05

Tabela 11: Resultado da Capacidade de Processo dos anos 2013, 2014 e 2015

Resultados			
Capacidade do Processo atrasos com mais de 4 horas			
Capacidade	Ano 2013	Ano 2014	Ano 2015
$CP = LSC - LIC / 6\sigma$	0,004342	0,004937	0,00747

Fonte: Autor

4. ANALISE DOS RESULTADOS

Analisando os resultados apresentados nos Gráficos de Controle 01, 02 e 03 referente aos atrasos inferiores a 4 horas dos anos de 2013 a 2015, como também a tabela 08 e 9 de resultados, as variáveis se localizam fora dos limites de controle três-sigma, sendo que praticamente todas as variáveis ultrapassam os LSC e LIC com a exceção do mês de março, junho e novembro de 2014. Desta forma segundo Montgomery (2004), no *Western Electric Handbook* (1956), que apresenta um conjunto de regras de decisão para detecção de padrões não aleatórios em gráficos de controle citados na seção 2.2 deste trabalho, o processo está fora de controle. Analisando também os gráficos 04, 05 e 06 atributos referente a atrasos acima de 4 horas dos anos de 2015 o limite superior de controle (LSC) foi ultrapassado no mês de dezembro o que indica que o processo está fora de controle neste período, mas sob controle nos demais períodos. Na tabela 10 de cálculo da Capacidade do Processo os resultados são inferiores a 1 o que indica que o processo usa mais de 100% da faixa de tolerância, sendo um processo muito sensível

sujeito a um grande número de não conformidades.

5. CONCLUSÃO

Os resultados apresentados demonstraram a importância da utilização do CEP nos processos, a utilização desta ferramenta trará como consequência principal o aumento da qualidade das operações. O controle estatístico de processo (CEP), embora pouco utilizado nas operações aeroportuárias, é uma ferramenta de grande utilidade, além de abastecer com informações imprescindíveis para a validação de processos, permite também uma investigação detalhada de todos os pontos críticos de controle, diagnosticando as possíveis não conformidades em todas as etapas do processo.

É essencial enfatizar que um gráfico de controle de processos não permite a identificação de quais são as causas especiais de variação que estão agindo em um processo fora de controle estatístico, mas ele processa e dispõe informações que podem ser utilizadas na identificação destas causas. Sendo que neste trabalho foi possível verificar através de um conjunto de regras de

decisão para detecção de padrões não aleatórios em gráficos de controle, sugerindo especificamente que o processo está fora de controle quando analisamos as variáveis referente aos atrasos, indicando também que o mês de dezembro é um mês crítico. Como a capacidade de um aeródromo não é constante. Ela varia consideravelmente com base em várias considerações, incluindo a utilização das pistas de pouso, o tipo de aeronave sendo operadas, o percentual de operações de decolagem e pouso sendo realizadas, as condições climáticas e as regulamentações que prescrevem o uso de pistas de pouso com base nessas considerações sugiro que as unidades aeroportuárias planejem e controle de melhor maneira suas operações com a finalidade de

melhorar o serviço prestado aos clientes e atraso operacionais.

Outro fator a ser considerado é que se deve tomar cuidado na interpretação de pontos que se localizam abaixo do limite inferior, estes pontos podem não representar melhoria real de qualidade no processos, frequentemente são causados por erros de processo de inspeção que no caso erros de informação, outro fator também segundo Montgomery (2004), nota-se também que as não-conformidade na sua maior parte são atribuídas a poucos tipos de defeitos, uma abordagem útil para uma análise adicional de não-conformidades seria o diagrama de causa e efeito que é usado para indicar várias fontes de não-conformidades.

REFERÊNCIAS

- [1] Agência Nacional de Aviação Civil, Anuário do Transporte Aéreo 2014, volume único, 1ª edição 2015.
- [2] Agência Nacional de Aviação Civil, Resolução Nº 141, dispõe sobre as Condições Gerais de Transporte aplicáveis aos atrasos e cancelamentos de voos e às hipóteses de preterição de passageiros e dá outras providências. 2010 p. 1-4.
- [3] AISLING, R.; KENNETH J. B. An assessment of the capacity and congestion levels at European airports, European Regional Science Association 1999.
- [4] ASHFORD, N. J. et al. Airport engineering: planning, design, and development of 21st century airports /.-4th ed. 2011 p. 234-241.
- [5] BALLOU, Ronald H. Gerenciamento da Cadeia de Suprimentos. São Paulo: Bookman, 4 ed. São Paulo. 2004.
- [6] BALLOU, Ronald H. Logística Empresarial, Administração de Materiais e Distribuição Física. São Paulo: Atlas, 1993.
- [7] CARNEIRO Neto W. Controle estatístico de processo CEP [CDROM]. Recife: UPE-POLI; 2003.
- [8] CARVALHO, M. M. et al., Gestão da Qualidade: Teoria e Casos, Rio de Janeiro; Elsevier 2005.
- [9] CRESWELL, John W Research design: Qualitative, quantitative, and mixed methods approaches/John W Creswell. -3rd ed. 2009.
- [10] DEMOVILE Jay L., Probabilidade e estatística para engenharia e Ciências, 8ª ed. Cengage Learning, São Paulo. 2015, (4).
- [11] FILHO P. DOMINGOS; SANTOS J. A. Metodologia Científica –São Paulo: Futura, 1998.
- [12] FLEURY, Paulo Fernando et al. Logística Empresarial. São Paulo, 1ª ed. Atlas. 2000.
- [13] GAITHER Norman, FRAZIER Greg, Administração da produção e operações, tradução José Carlos Barbosa dos Santos; Revisão Petrônio Garcia Martins, São Paulo: Pioneira Thomson Learner. 2002.
- [14] HORONJEFF, R.; MCKELVEY, F. Planning and Design of Airports. 4. Ed., McGraw-Hill, New York, 2010.
- [15] JURAN JM. Na Liderança pela Qualidade. 2a ed. São Paulo. Livraria editora Pioneira, 1993.
- [16] KAZDA A.; CAVES R. E. Airport Design and Operações 2ª edição Elsever. 2007.
- [17] KONRATH AC. Decomposição da estatística do gráfico de controle multivariado T2 de Hotelling por meio de um algoritmo computacional. [Dissertação] Florianópolis: Universidade Federal de Santa Catarina. 2000.
- [18] KUME H. Métodos estatísticos para melhoria da qualidade. São Paulo: Ed. Gente; 1993.
- [19] LACHMAN L. et al. Teoria e prática na indústria farmacêutica. Lisboa: Fundação Coloutre Gulbekian, 2001.
- [20] LAKATOS, Eva Maria; MARCONI, Marina de Andrade. Fundamentos de metodologia científica. 7. ed. São Paulo: Atlas, 2010. 297 p.
- [21] MACGREGOR, J. F.; KOURTI, T. *Statistical Process Control of Multivariate Process, Control Eng. Practice*, v. 3, n. 3, 1995.

- [22] MONTGOMERY Douglas C., Introdução ao Controle Estatístico da Qualidade, 4^o ed. Rio de Janeiro, LTC.2004, p. (1), (2), (3), (5), (6), (7) 95-209 .
- [23] Oliveira S. T., Ferramentas Para o Aprimoramento da Qualidade, 2^o ed. São Paulo. Pioneira. 1996.
- [24] RAMOS AW. CEP para processos contínuos e em bateladas. São Paulo: Ed Edgard Blucher; 2000.
- [25] REGINO, Sueli Maria et al. Metodologia Científica: guia para elaboração e apresentação de trabalhos acadêmicos. Goiânia: Faculdades Alves Faria, 2003.
- [26] SENEVIRATNE, P.N.; MARTEL, N. VARIABLES Influencing Performance of Air Terminal Buildings. *Transportation Planning and Technology*. 1991.
- [27] STEVENSON, Willian J. Administração das Operações de Produção, 6^a ed. – Rio de Janeiro, LTC 2001.
- [28] VIEIRA S. Estatística para a qualidade: como avaliar com precisão a qualidade em produtos e serviços. Rio de Janeiro: Elsevier; 1999, p. 49-126.
- [29] VIEIRA, S.R.; et al. Handbook for geoestatal analysis of variability in soil and climate data. In: ALVAREZ V.V.H.; SCHAEFER, C.E.G.R.; BARROS, N.F.; MELLO, J.W.V.; COSTA, L.M. (Ed.). Tópicos em ciência do solo. Viçosa: Sociedade Brasileira de Ciência do Solo, 2002, v.2, p.1-45.
- [30] WERKEMA MCC. Ferramentas estatísticas básicas para o gerenciamento de processos. Belo Horizonte: Fundação Christiano Ottoni, Escola de Engenharia da UFMG; 1995, v.2, p.197 -284.
- [31] YOUNG, S., WELLS A. Aeroportos e Gestão; tradução: Ronaldo Saraiva de Menezes; Kétnes Ermelinda de Guimarães Lopez. – 6^a ed.. – Porto Alegre: Bookman. 2014.

Capítulo 20

MODELAGEM MATEMÁTICA NO SEQUENCIAMENTO DE PROJETOS COM RESTRIÇÃO DE RECURSOS: UM ESTUDO DE CASO

Clarisse da Silva Vieira

Natanael Salgado Coutinho

Gustavo José de Aguiar Gomes Costa

Vinícius de Souza Bandeira

Vinícius Rossi Oliveira

Resumo: Diversos projetos, sejam eles industriais ou meramente de um processo manual comum, dependem de atividades que podem ocorrer de forma sequencial ou simultânea. No cenário de competitividade atual, necessitamos de trabalhar da forma mais eficiente possível e muitas vezes é necessário reduzir o tempo desses projetos através do sequenciamento otimizado dessas atividades. Este artigo tem como objetivo estudar o modelo do Problema de Sequenciamento em Projetos com Restrição de Recursos (PSPRR) que trata do estudo do menor tempo de execução das atividades de acordo com suas precedências e a limitação de recursos. Para isso, foi realizada uma revisão em trabalhos anteriores para mostrar alguns tipos de relações de precedências entre atividades e tipos de recursos existentes em um projeto. Aplicamos então o modelo estudado em um estudo de caso proposto pelos autores desse mesmo artigo. Os resultados encontrados mostram a melhora esperada, razão da importância do tema abordado.

Palavras Chave. Sequenciamento, Projeto, Recurso.

1. INTRODUÇÃO

Diante da intensa competitividade das empresas e a busca incessante por melhores resultados, as organizações investem e estudam cada vez mais para reduzir o tempo de sua produção. Com isso, se uma empresa deseja estar, de algum modo, entre as melhores do mercado, ela deve ter um planejamento ótimo. O “Planejamento e Controle da Produção” e a “Pesquisa Operacional” são áreas responsáveis por coordenar e aplicar de forma eficiente os recursos produtivos, fazendo com que as decisões deixem de ser tomadas de forma intuitiva e passem a seguir critérios de melhoria racionais e computacionais. A complexidade dos problemas reais e empresariais, em contrapartida a possibilidade de redução do tempo na realização dos projetos através da aplicação de métodos eficientes de solução, é o motivo do presente estudo.

Já existem uma grande variedade de modelos matemáticos que são implementados em programas computacionais capazes de ler as informações colocadas por um programador. Portanto, cabe a esse programador entender o problema real da empresa para transformá-lo na linguagem digital do programa. Diversos problemas, dentro do gerenciamento de projetos e processos, devido a sua complexidade e dificuldade de obtenção da solução, devem ser solucionados por esses meios. Dentre estes problemas, está o Problema de Sequenciamento em Projeto com Restrição de Recurso (PSPRR)

O Problema de Sequenciamento em Projeto com Restrição de Recurso (PSPRR) é um problema clássico da Literatura, composto por um conjunto de atividades que devem seguir uma precedência exigida no processo real para serem realizadas. Porém, devemos analisar que, para viabilizar essa execução, as quantidades disponíveis de recurso no projeto devem ser respeitadas, visando diminuir o tempo total de duração do projeto.

Com o intuito de explorar esse tema e aplicá-lo em um Estudo de Caso, este artigo apresenta, na seção 2, uma revisão bibliográfica sobre o Problema de Sequenciamento em Projetos com Restrição de Recursos. A seção 3 explicita o Estudo de Caso em uma etapa produtiva de uma fábrica de impressoras. A seção 4 mostra o modelo de Programação Linear utilizado. Os resultados e as discussões estão na seção 5.

E por fim a seção 6 dispõe as conclusões obtidas seguidas das referências.

2. O PROBLEMA DE SEQUENCIAMENTO EM PROJETO COM RESTRIÇÃO DE RECURSO

Segundo [Brucker *et al.* 1998], o Problema de Sequenciamento em Projetos com Restrição de Recursos (PSPRR) é definido, genericamente, por um conjunto V de atividades, ($i = 1, 2, \dots, n$), e um conjunto R de recursos renováveis, ($r = 1, \dots, m$). Cada tipo de recurso está disponível em uma quantidade constante de unidades (R_k). Cada atividade i do projeto necessita de uma quantidade constante do recurso (rik) para ser executada. A atividade i é executada sem ser interrompida em pi unidades de tempo. Os valores de R_k , rik e pi são valores positivos ou nulos. Algumas relações de precedência entre as atividades são definidas a priori. O objetivo é minimizar a duração total do projeto (*makespan*), determinando a data de início de cada atividade do projeto, respeitando as limitações de recursos e atendendo todas as relações de precedência definidas a priori.

O PSPRR é relevante tanto no meio acadêmico quanto no meio industrial. Devido à complexidade do problema e dificuldade na obtenção da solução ótima, muitas variações da formulação original podem ser observadas. Diversos autores, ao longo de vários anos, propõem modelos de Programação Matemática para o Problema, como por exemplo, em [Balas 1967], [Christofides *et al.* 1987], [Mingozzi *et al.* 1998], [Klein e Scholl 1999], [Brucker e Knust 2000], [Carlier e Néron 2003], [Artigues *et al.* 2003], [Deblaere *et al.* 2007], [Damay *et al.* 2007], [Bianco e Caramia 2013] e [Artigues *et al.* 2013].

[Balas 1967] introduz uma abordagem inovadora, na época, para o problema, considerando que as atividades que inicialmente poderiam ser executadas simultaneamente, não podem ser realizadas plenamente em paralelo, mesmo que não existam relações de precedência entre elas. Isso porque tais atividades compartilham o mesmo recurso. Essa abordagem foi considerada em muitas outras formulações propostas posteriormente, como por exemplo, a apresentada por [Artigues *et al.* 2003].

A formulação proposta por [Artigues *et al.* 2003] considera ainda o PSPRR como um problema de fluxo e como uma extensão do

modelo matemático clássico do problema de sequenciamento do *Job Shop*. O que também pode ser constatado em [Deblaere *et al.* 2007].

[Christofides *et al.* 1987], [Mingozzi *et al.* 1998], [Klein e Scholl 1999], [Brucker e Knust 2000], assim como [Carlier e Néron 2003], modificam a formulação do problema original relaxando algumas restrições. [Klein e Scholl 1999], [Brucker e Knust 2000] e [Carlier e Néron 2003] também propõem o cálculo de limites inferiores para o PSPRR.

[Brucker e Knust 2000] considera que é possível dividir o horizonte de planejamento em sucessivos intervalos de tempos e que as atividades do projeto podem ser particionadas. Nele, as partes de atividades diferentes podem ser executadas simultaneamente. No entanto, as partes de uma mesma atividade devem ser executadas de uma maneira sequencial. Sendo assim, o problema consiste em determinar quais as partes das atividades que poderão ser executadas simultaneamente dentro de um mesmo intervalo, obedecendo às restrições de recursos e precedência e com o intuito de minimizar o tempo total de execução do projeto.

O modelo de Programação Matemática proposto por [Carlier e Néron 2003] também considera que o horizonte de planejamento pode ser dividido em sucessivos intervalos de tempo e que as atividades de um projeto podem ser particionadas. No entanto, as partes das atividades são executadas em intervalos consecutivos de tempo, não podendo mais ser executadas simultaneamente, sendo essa uma das principais diferenças em relação ao modelo de [Brucker e Knust 2000]. A quantidade total de recurso disponível, bem como os tempos de processamento e as quantidades de recurso para cada atividade, são valores definidos no início do projeto.

Outros autores, como [Damary *et al.* 2007], [Bianco e Caramia 2013] e [Artigues *et al.* 2013], propõem modelos matemáticos de Programação Linear Inteira Mista para o Problema, alterando algumas considerações e relaxando algumas restrições da definição original, como por exemplo, a duração das atividades, quantidade de recurso, tipo de recurso utilizado e relações de precedências.

Devido a grande proposição de formulações para os problemas classificados como PSPRR, em contrapartida, um número muito

grande de métodos de resolução também são apresentados.

Segundo [Morillo *et al.* 2015], definido o PSPRR, a solução pode ser encontrada com a utilização da otimização clássica ou métodos heurísticos.

Os métodos exatos são agrupados em algoritmos que usam técnicas analíticas ou matemáticas, direcionando a resolução para uma solução ótima, caso ela exista. Estes métodos se baseiam em características específicas, como por exemplo, continuidade, diferenciabilidade, espaço de busca, entre outras. Os métodos exatos podem ser aplicados a uma grande variedade de problemas. No entanto, segundo [Fahmy *et al.* 2014], o tempo de análise se torna impraticável como o aumento do tamanho do problema.

Alguns desses métodos mais representativos são: Busca exaustiva (*Exhaustive Search*), Programação Linear Inteira Mista (*Mixed Integer Programming*), Dividir e Conquistar (*Divide and Conquer*), Programação Dinâmica (*Dynamic Programming*) e, Ramificação e Dimensionamento (*Branch and Bound*). [Morillo *et al.* 2015], ressaltam que o melhor dos métodos exatos é o *Branch and Bound*, que usa soluções parciais, construídas heurísticamente, e que resolve o problema de forma a garantir a otimização. Porém, essa metodologia não pode ser aplicável a problemas muito complexos.

Ainda conforme [Morillo *et al.* 2015], os métodos heurísticos, se baseiam em um processo de obtenção de uma boa solução para o problema. A maior parte das abordagens das heurísticas consiste em dois elementos: uma regra de prioridade (*Priority Rule*) para determinar a prioridade de cada atividade baseada em critérios pré-definidos e um esquema de geração de cronograma (*Schedule Generation Scheme*) para criar um cronograma usando a lista de atividades priorizadas. Alguns dos métodos heurísticos mais utilizados são: Busca Local (*Local Search*), Algoritmos Gulosos (*Greedy Algorithms*), Métodos Truncados de Derivação e Limitação, métodos baseados em arcos disjuntivos e algoritmos baseados em sequenciamento por blocos.

Em [Herroelen e Leus 2004], um sistema matemático de programação é desenvolvido visando à diminuição das discrepâncias entre os modelos e a realidade. [Valls *et al.* 2005] adicionaram outro elemento, a técnica de

Justification para melhorar a qualidade dos cronogramas gerados. [Fahmy *et al.* 2014] também apresentam uma nova técnica, a *Stacking Justification* (empilhamento), e sua implantação no proposto PSO (*Particle Swarm Optimization*) de Múltiplo *Justification* (ou MJPSO) com o objetivo de melhorar a qualidade das soluções dos PSPRR. Os resultados experimentais da pesquisa dos autores ilustraram que a combinação de *Stacking Justification* com o original Double *Justification* alcançou melhora considerável na qualidade das soluções. Além disso, mostram que a combinação das duas técnicas foi testada usando *Particle Swarm Optimization* e o algoritmo desenvolvido, MJPSO, que superou muitos dos algoritmos de melhores desempenhos na literatura, e alcançou os melhores resultados em respeito às aplicações da técnica de PSO para os PSPRR.

Para solucionar o impacto da complexidade do problema nas soluções dos algoritmos, outras metodologias foram desenvolvidas inspiradas na natureza e em unidades evolutivas como o Algoritmo Imune (AI). Os fundamentos do AI são baseados na resposta do sistema imunológico humano quando um antígeno desconhecido é reconhecido por anticorpos [Castro e Timmis 2002]. Um exemplo desse tipo de algoritmo é o CBIA (*Chaos-based improved immune algorithm*) que, segundo [Wu *et al.* 2011], tem como objetivo reduzir custo computacional e melhorar a eficiência da busca pela solução sem ser restringido pela dificuldade do problema.

Outras proposições de metodologia podem ser encontradas na literatura. Como por exemplo, [Bhaskar *et al.* 2011] que aconselham um método *fuzzy* com tempos aproximados das atividades para resolver o problema de sequenciamento e [Damay *et al.* 2007] que resolvem o problema de PSPRR usando um modelo de programação linear, tomando como base as atividades serem ou não preemptivas.

3. ESTUDO DE CASO

[Almeida 2009] elaborou uma problemática envolvendo o sequenciamento da produção em uma fábrica de impressoras. O autor buscou analisar o processo de montagem final de impressoras, o qual representa uma importante etapa da produção, pois apresentam recursos limitados, o que diferencia das demais etapas do processo produtivo. O autor ainda afirma que a etapa de montagem é o gargalo e que esta diretamente ligada à demanda final. Daí torna-se relevante uma ferramenta que auxilie as tomadas de decisões as quais permitem elaborar um planejamento novo com novas estratégias para a adequação das exigências do cenário externo.

Para um maior entendimento do problema proposto por [Almeida 2009], a Figura 1 destaca as etapas as quais foram modeladas e adaptadas ao modelo proposto por este estudo.

Figura 1 - Etapas do processo de montagem das impressoras

Fonte: [Almeida 2009]

De acordo com a figura é possível observar que existem diferentes caminhos a serem seguidos para realizar o processo de montagem da impressora. Como no trabalho do mesmo autor é ressaltada a existência de quatro tipos de impressoras que devem fluir dentro do processo, o *leadtime* pode ser

prejudicado, uma vez que cada impressora dependera dos recursos das mesmas atividades. Ao final do estudo, [Almeida 2009] apresenta o tempo total de produção unitário das impressoras que podem passar por duas linhas de montagem. A Tabela 1, demonstra os tempos obtidos.

Tabela 1- Tempo unitário de produção da impressora x linha.

Tempo unit. (horas)	Linha 1	Linha 2
Impressora 1	55	49
Impressora 2	53	46
Impressora 3	35	39
Impressora 4	45	47

Fonte: [Almeida 2009]

Para desenvolver o estudo, consideramos o processo de fabricação da impressora 1, fluindo nas duas linhas propostas por [Almeida 2009], já que ela apresenta o maior tempo produtivo em ambas as linhas. Logo, os dados de relevância são:

a) para impressora 1 na linha 1: 55 horas;

b) para impressora 1 na linha 2: 49 horas.

A Tabela 2 apresenta as relações de precedência e as durações de cada atividade do processo produtivo.

Tabela 2 - Dados do problema

Atividade	Predecessores	Duração (horas)
1	-	1
2	1	3
3	2	4
4	3	1,5
5	4	0,5
6	5, 3	2
7	6	3
8	7	5
9	8	4
10	9	1
11	10	2
12	11	5
13	12	3
14	13	5
15	14	6
16	15	0,5
17	16	2,5
18	17	1
19	18, 15	5

Portanto, conhecendo os dados que serão manipulados, o problema pode ser ilustrado através de um Diagrama. Cada atividade foi

enumerada de um a dezenove. A Figura 2 representa o Diagrama mencionado.

Figura 2 - Diagrama de Atividades

De acordo com a Figura 2, é possível identificar as linhas de produção da impressora. A linha 1 apresenta o maior número de atividades possível. Já a linha 2 é

caracterizada pelo menor número de atividades. A Tabela 3 identifica as atividades descritas no Diagrama de Atividades.

Tabela 3 - Nomenclatura das atividades

Atividade	Nomenclatura
1	Chegada de matéria-prima
2	Recebimento
3	Almoxarifado de componentes
4	Processo SMT
5	Reparação das placas defeituosas
6	Processo PTH
7	Soldagem dos componentes PHT
8	Reparação das placas defeituosas
9	Acabamento
10	Teste ICT
11	Reparação das placas defeituosas
12	Teste funcional
13	Reparação das placas defeituosas
14	Embalagem
15	Almoxarifado de placas de circuito impresso
16	Processo de montagem do produto final no BOX BUILD
17	Reparação dos produtos defeituosos
18	Almoxarifado de produtos do BOX BUILD
19	Expedição

Fonte: [Almeida 2009]

A Tabela 4 destaca as atividades pertencentes a cada linha de produção.

Tabela 4 - Definição das linhas de produção

Linha 1	Linha 2
Atividade 1	Atividade 1
Atividade 2	Atividade 2
Atividade 3	Atividade 3
Atividade 4	-
Atividade 5	-
Atividade 6	Atividade 6
Atividade 7	Atividade 7
Atividade 8	Atividade 8
Atividade 9	Atividade 9
Atividade 10	Atividade 10
Atividade 11	Atividade 11
Atividade 12	Atividade 12
Atividade 13	Atividade 13
Atividade 14	Atividade 14
Atividade 15	Atividade 15
Atividade 16	-
Atividade 17	-
Atividade 18	-
Atividade 19	Atividade 19

4. PROPOSIÇÃO DO PROBLEMA E METODOLOGIA

O problema consiste em definir a duração total do projeto, de tal forma que as atividades das duas linhas de produção sejam executadas ao mesmo tempo compartilhando os recursos disponíveis. Dessa forma, as atividades para a produção da impressora 1 devem ser realizadas tanto na linha 1 e 2 para que o projeto seja totalmente concluído. Para isso, é preciso definir as datas de início para cada uma das atividades considerando a execução das duas linhas de produção e as

limitações referentes à disponibilidade dos recursos. Considera-se também que cada uma das 19 atividades necessitam de um tipo diferente de recurso disponível exatamente na quantidade necessária para a sua execução e que há compartilhamento de recurso entre as mesmas atividades de linhas diferentes, ou seja, todas as atividades compartilham recursos, exceto as atividades 4, 5, 16, 17 e 18 que pertencem somente a linha de produção 1.

O modelo matemático implementado foi apresentado inicialmente em [Artigues *et al.* 2003] e está reescrito a seguir:

$\min t_{n+1}$		(1)
s.a.		
$y_{ij} = 1$	$\forall (i,j) \in E$	(2)
$y_{ji} = 0$	$\forall (i,j) \in E$	(3)
$t_j - t_i - y_{ij} \sum_{k \in V} p_i \geq p_j - \sum_{k \in V} p_i$	$\forall i \in V \cup \{s\}, \forall j \in V \cup \{t\}, i \neq j$	(4)
$f_{ijk} - y_{ij} \max(r_{ik}) \leq 0$	$\forall i \in V \cup \{s\}, \forall j \in V \cup \{t\},$ $\forall k \in R, i \neq j$	(5)
$\sum_{j \in V \cup \{t\}} f_{ijk} = r_{ik}$	$\forall i \in V \cup \{s\}, \forall k \in R, i \neq j$	(6)
$\sum_{i \in V \cup \{s\}} f_{ijk} = r_{jk}$	$\forall j \in V \cup \{t\}, \forall k \in R, i \neq j$	(7)
$t_i \geq 0$	$\forall i \in V \cup \{s\} \cup \{t\}$	(8)
$f_{ijk} \geq 0$	$\forall i \in V \cup \{s\}, \forall j \in V \cup \{t\}, \forall k \in R$	(9)
$y_{ij} \in \{0, 1\}$	$\forall i \in V \cup \{s\}, \forall j \in V \cup \{t\}$	(10)

Onde:

- $V = \{1, 2, \dots, n\}$: conjunto das n atividades do projeto;
- $R = \{1, 2, \dots, m\}$: conjunto dos m recursos renováveis;
- E : conjunto das atividades cuja relação de precedência é conhecida *a priori*, ou seja, a atividade j só será iniciada após a conclusão da atividade i ;
- s e $n+1$: atividades fictícias que representam, respectivamente, o início e o fim do projeto;
- r_{ik} : quantidade do recurso k que a atividade i necessita para ser executada;
- p_i : tempo de processamento da atividade i ;
- f_{ijk} : fluxo do recurso k da atividade i para a atividade j ;
- y_{ij} : variável binária que estabelece se a atividade i precede j ;
- t_i : data de início da atividade i ;
- t_{n+1} : data de início da atividade fictícia $n+1$.

Neste modelo, a função objetivo (1) é minimizar t_{n+1} , pois é a data de início da última atividade (fictícia) e, minimizando essa data, consequentemente a duração do projeto é minimizada. As restrições (2) e (3) estão relacionadas às relações de precedência. Para as atividades cuja relação de precedência é conhecida *a priori*, o valor da

variável $y_{ij} = 1$. Isso implica que j só será iniciada após a conclusão da atividade i , logo $y_{ji} = 0$. A restrição (4) define as relações de sequenciamento entre as atividades que não possuem a relação de precedência definida *a priori*. A restrição (5) estabelece a existência de fluxo entre as atividades através das variáveis f_{ijk} e x_{ij} . As restrições (6) e (7) mostram que o fluxo do recurso k , em uma atividade i , deve ser igual à quantidade requerida deste recurso pela atividade ser executada, isto é, deve chegar e sair da atividade a mesma quantidade de recurso que esta atividade necessita (lei de conservação do fluxo). As restrições (8), (9) e (10) definem o domínio das variáveis t_i , f_{ijk} e y_{ij} .

O problema apresentado foi modelado através da metodologia acima e para a implementação computacional do problema, utilizou-se a linguagem MPL (*Mathematical Programming Language*) e o otimizador GLPK 4.4 (*GNU Language Program Kit*) que é um pacote livre de otimização para problemas de Programação Linear. O computador utilizado na implementação e execução do modelo possui a seguinte configuração: sistema operacional *Windows 7 Home Premium*, processador *Intel Core i5*, 2.8 GHz e memória RAM instalada de 4 Gb. Para a solução deste problema, relativamente simples, foram gastos 67,2 segundos para se encontrar a solução, utilizando 32,3 MB de memória.

Por meio desta ação, pudemos verificar que o tempo total gasto nos procedimentos é, de 59 horas, conforme mostra a Tabela 5:

Tabela 5 - Datas de início das atividades

Atividade Linha 1	Data de Início	Atividade Linha 2	Data de Início
1	1	1	0
2	4	2	1
3	8	3	4
4	12	-	-
5	13,5	-	-
6	14	6	8
7	16	7	10
8	19	8	13
9	24	9	18
10	28	10	22
11	29	11	23
12	31	12	25
13	36	13	30
14	39	14	33
15	44	15	38
16	50	-	-
17	50,5	-	-
18	53	-	-
19	54	19	45

5. RESULTADOS E CONCLUSÕES

Observa-se que, no citado modelo implementado para o problema apresentado, existem dezenove tipos de recursos disponíveis em uma unidade. Cada tipo de recurso está alocado para cada tipo de atividade do projeto. Dessa forma, as atividades 1, 2, 3, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 e 19 são comuns às duas Linhas (essas atividades ocorrem tanto na Linha 1 e na Linha 2) e compartilham os mesmos recursos (recursos tipo 1, 2, 3, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 e 19). Apenas as atividades 4, 5, 16, 17 e 18 não compartilham recursos (recursos tipo 4, 5, 16, 17 e 18). Pois, essas atividades fazem parte apenas da Linha 1.

A proposta de melhoria seria a inclusão de recursos para a execução das atividades que pertencem as duas Linhas. Assim, as atividades 1, 2, 3, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 e 19 que inicialmente seriam executadas sequencialmente, porque utilizam do mesmo tipo de recurso, agora poderiam ser realizadas simultaneamente, em paralelo. Após a inclusão de quantidades adicionais de recursos, o tempo total de conclusão do projeto seria agora definido pelo maior tempo da linha de produção, ou seja, seria igual à 55, tempo da Linha 1.

Conclui-se, portanto, que com a inclusão dos recursos tipo 4, 5, 16, 17 e 18, o tempo estimado pode ser melhorado. A redução no tempo gasto do início ao final do processo foi de 6,78%, o que representa uma redução significativa quando somamos a produção feita em um ano.

Entretanto, deve-se observar que a inclusão de novas quantidades de recursos traz gastos extras à produção. Assim sendo, atenta-se para a necessidade de que um cálculo dos ganhos com a redução do tempo e dos gastos com recursos a mais seja feito, de forma que se verifique se vale a pena a aquisição dos recursos, do ponto de vista econômico.

Atente-se para o fato de que outros testes foram feitos com o modelo melhorado, consistindo na inclusão de mais recursos. Porém, notou-se que nenhuma melhoria no tempo foi alcançada com a inclusão de tais recursos, já que o menor tempo total de produção é definido pelo maior tempo entre as linhas de montagem já otimizado. Dessa forma, a melhoria pôde ser verificada apenas com a inclusão de uma unidade aos recursos comuns entre as atividades que os compartilhavam. Trata-se, portanto, do melhor rearranjo possível segundo o modelo proposto.

REFERÊNCIAS

- [1] Almeida, J.F.F. (2009). Planejamento da produção de impressoras por dimensionamento de lotes e sequenciamento: um estudo de caso. Dissertação de Mestrado em Engenharia de Produção - Escola de Engenharia, Universidade Federal de Minas Gerais.
- [2] Artigues, C., Michelon, P. e Reusser, S. (2003). Insertion techniques for static and dynamic resource constrained project scheduling. *European Journal of Operational Research*, 149: 249-267.
- [3] Artigues, C. Leus, R. e Nobibon, F. (2013). Robust optimization for resource-constrained project scheduling with uncertain activity durations. *Flexible Services and Manufacturing Journal*, 25: 175-205.
- [4] Balas, E. (1967). Project scheduling with resource constraints. *Operational Research*, 15: 915-957.
- [5] BHASKAR, T., PAL, M. E PAL, A. (2011). A HEURISTIC METHOD FOR RCPSP WITH FUZZY ACTIVITY TIMES. *EUROPEAN JOURNAL OF OPERATIONAL RESEARCH*, 208: 57-66.
- [6] Bianco, L. e Caramia, M. (2013). A new formulation for the project scheduling problem under limited resources. *Flexible Services and Manufacturing Journal*, 25: 6-24.
- [7] Brucker, P., Knust, S., Schoo, A. e Thiele, O. (1998). A branch and bound algorithm for the resource-constrained project scheduling problem. *European Journal of Operational Research*, 107: 272-288.
- [8] Brucker, P. e Knust, S. (2000). A linear programming and constraint propagation-based lower bound for the RCPSP. *European Journal of Operational Research*, 127: 355-362.
- [9] Carlier, J. e Néron, E. (2003). On linear lower bounds for the resource constrained project scheduling problem. *European Journal of Operational Research*, 149: 314-324.
- [10] Castro, L. and Timmis, J. (2002). An artificial immune network for multimodal function optimization. *Proceedings of the IEEE Congress on Evolutionary Computation*, 1: 699-674.
- [11] Christofides, N., Alvarez-Valdes, R. e Tamarit, J. M. (1987). Project scheduling with resource constraints: a branch and bound approach. *European Journal of Operational Research*, 29: 262-273.
- [12] DAMAY, J., QUILLIOT, A. E SANLAVILLE, E. (2007). LINEAR PROGRAMMING BASED ALGORITHMS FOR PREEMPTIVE AND NON-PREEMPTIVE RCPSP. *EUROPEAN JOURNAL OF OPERATIONAL RESEARCH*, 182: 1012-1022.
- [13] Deblaere, F., Demeulemeester, E. L., Herroelen, W. S. e Vonder, S. V. (2007). Robust resource allocation decisions in resource constrained projects. *Decision Science*, 38: 5-37.
- [14] Fahmy, A., Hassan, T. e Bassion, H. (2014). Improving RCPSP solutions quality with stacking justification – application with particle swarm optimization. *Expert Systems with Applications*, 41: 5870-5881.
- [15] Herroelen, W. e Leus, R. (2004). Robust and reactive project scheduling: a review and classification of procedures. *International Journal of Production Research*, 42: 1599-1620.
- [16] Klein, R. & Scholl, A. (1999). Computing lower bounds by destructive improvement: an application to resource-constrained project scheduling. *European Journal of Operational Research*, 112: 322-346.
- [17] Mingozzi, A.; Maniezzo, V.; Ricciardelli, S. & Bianco, L. (1998). An exact algorithm for the resource-constrained project scheduling problem based on a new mathematical formulation. *Management Science*, 44: 714-729.
- [18] Morillo, D., Moreno, L. e Serna, F. (2015). A branch and bound hybrid algorithm with four deterministic heuristics for the resource constrained project scheduling problem (RCPSP). *DYNA*, 82.
- [19] Valls, V., Ballestin, F. e Quintanilla, S. (2005). Justification and RCPSP: a technique that pays. *European Journal of Operational Research*, 165: 375-386.
- [20] WU, S., WAN, H. SHUKLA, S. E LI, B. (2011). CHAOS-BASED IMPROVED IMMUNE ALGORITHM (CBIIA) FOR RESOURCE-CONSTRAINED PROJECT SCHEDULING PROBLEMS. *EXPERT SYSTEMS WITH APPLICATIONS*, 38: 3387-3395.

Autares

ALESSANDRO VIVAS ANDRADE

Possui graduação em Engenharia Elétrica pela Universidade Federal de Minas Gerais (UFMG). Possui Mestrado e Doutorado em Engenharia Elétrica pela Universidade Federal de Minas Gerais na área de Redes de Computadores. É professor Adjunto do Departamento de Computação da Faculdade de Ciências Exatas da Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM). Possui experiência em Redes de Computadores, Otimização, Aprendizado de Máquina e Inteligência Artificial.

ANEIRSON FRANCISCO DA SILVA

Pós – doutorando na (Universidade de São Paulo- USP, 2017-2018), Doutor em Engenharia de Produção (Universidade Estadual Paulista-Unesp, 2013), Mestre em Engenharia de Produção (Universidade Federal de Itajubá- UNIFEI, 2009). Atualmente professor na Faculdade de Engenharia de Guaratinguetá da Universidade Estadual Paulista “Júlio de Mesquita Filho”. Atua na Pós-graduação Lato e Stricto Sensu nas linhas Pesquisa Operacional, Gestão da Produção e Operações, Métodos de Simulação e Lógica Fuzzy.

ANTONIO AUGUSTO GONÇALVES

É Doutor em engenharia de produção pela Universidade Federal do Rio de Janeiro. Possui mestrado em engenharia de produção pela UFRJ. Atualmente é professor do programa de Mestrado em Administração e chefe da área de Tecnologia de Informação do Instituto Nacional do Câncer. É membro ativo do European Working Group on Operational Research Applied Health Services (ORAHs).

ANTONIO HENRIQUEZ DE ARAUJO JR.

Finalizou, em Setembro de 2015, na Universidade do Minho, Portugal, pesquisa de pós-doutoramento sobre eficiência relativa entre portos brasileiros e europeus com suporte CAPES Grant BEX-1445/14-6. Graduado em Engenharia de Transportes (Ingenieurbaugesellschaft) pela Bergische Universität Wuppertal (Alemanha, 1976), mestre em Controladoria e Contabilidade pela Universidade de São Paulo (1990) e doutor em Engenharia de Produção pela Universidade de São Paulo (2004). Realizou pesquisa de pós-doutorado no Instituto Tecnológico de Aeronáutica (divisão de Mecânica Aeronáutica - Departamento de Produção, 2006-2007). Atuou na Gestão Econômica e Estratégica da indústria automobilística (Daimler-Benz e Mercedes-Benz do Brasil) e aeronáutica (Empresa Brasileira de Aeronáutica - EMBRAER). É, atualmente, Professor Adjunto IV, da Universidade do Estado do Rio de Janeiro, no Departamento de Engenharia de Produção (Faculdade de Tecnologia), onde é responsável pelas disciplinas Engenharia Econômica I e II, Custos Industriais e Metodologia da Pesquisa. Compõe o corpo editorial das revistas: Journal of Aerospace Technology and Management, editoria de Sistemas de Gestão (ISSN: 2175-9146), Science Journal of Business Management (ISSN: 2276-6316), Rio's International Journal on Sciences of Industrial and Systems Engineering and Management (ISSN: 1982-6443), Revista Carioca de Produção (ISSN: 1984-4743) e Sistemas & Gestão (ISSN: 1980-5160). Revisor dos periódicos: Journal of Air Transport Management (ISBN: 0969-6997), Journal of Business Management (ISSN: 2276-6316), Tecnologia em Metalurgia, Materiais e Mineração (ISSN: 1807-300X), Revista de Contabilidade e Controladoria, UFPR (ISSN: 1984-6266), e Sistemas & Gestão (ISSN: 1980-5160). É, ainda, avaliador de artigos dos seguintes simpósios: Encontro Nacional dos Cursos de Graduação em Administração (ENANGRAD), Encontro Nacional de Engenharia de Produção (ENEGEP), Congresso Brasileiro de Custos (CBC) e do Congresso da Sociedade Brasileira de Pesquisa Operacional (SBPO). Atua nas áreas de pesquisa: Produtividade Industrial, Eficiência e Eficácia do Transporte Aéreo, Produtividade em Transportes, Economia do Transporte Aéreo, Gestão da Qualidade e da Produtividade e Economia do Meio Ambiente. É avaliador do Sistema de Acreditação Regional de Cursos de Graduação do MERCOSUL, avaliador nacional de cursos e Institucional e de Cursos do Ministério de Educação e Cultura (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira -

INEP) além de compor o Banco de Revisores do Banco Nacional de Itens do ENADE. É Sócio fundador da Red Ibero Americana de Investigación en Transporte Aereo.

ARTUR DA SILVA SIQUEIRA DE NOVAIS

Engenheiro de Produção formado pelo Centro Universitário de Volta Redonda - UniFOA. Possui também dois cursos técnicos: eletrônica e meio ambiente. Ao longo da vida acadêmica realizou muitos projetos de iniciação científica em diversas áreas como por exemplo qualidade, metalúrgica, ciências dos materiais, biomedicina, entre outros. Atualmente dedica-se na pesquisa de áreas correlatas a qualidade e biotecnologia. Atua no Centro Universitário de Volta Redonda - UniFOA.

BERNARDO SANTOS AFLALO

Cientista de dados com mais de 10 anos de experiência, possui graduação em Eng. Mecânica pela UFMG, mestrado e atualmente estudante de doutorado de Computação do ITA. Trabalhou em empresas como Embraer e TOTVS e atualmente é o líder de ciência de dados na XP Investimentos.

BYANCA PORTO DE LIMA

Possui graduação em engenharia de produção pela Universidade Federal Fluminense (2007). É Mestre em Engenharia de Produção com Ênfase em Qualidade e Produtividade pela Universidade de Taubaté (2010). Experiência em Projetos na área automobilística (PSA Peugeot Citroen - 2006 a 2013). Atua como professora do Curso de Engenharia de Produção no Centro Universitário de Volta Redonda (UNIFOA) desde 2010, sendo Coordenadora do Curso de Engenharia de Produção no UNIFOA de 2015 a 2017. Atualmente é aluna de Doutorado da UNESP. Tem experiência na área de Engenharia de Produção, com ênfase em Gerência do Projeto e do Produto, atuando principalmente nos seguintes temas: montadora de automóveis, engenharia simultânea., desdobramento da função qualidade., desenvolvimento de embalagens, desenvolvimento de fornecedores e pesquisa operacional. Atua no Centro Universitário de Volta Redonda - UniFOA.

CAIQUE CÉZAR LEITE MARIANO

Possui curso técnico profissionalizante em Informática, incompleto, pelo Instituto Federal do Sertão Pernambucano. Atualmente é Bacharel em Engenharia da Computação pela Universidade Federal do Vale do São Francisco.

CARLA REGINA DE MORAES

Tecnóloga em secretariado executivo pela Fatec.

CAROLINA RIUL DE FREITAS

Tecnóloga em secretariado executivo pela Fatec.

CAROLINE CRISTINA DE CARVALHO

Graduada em Engenharia de Produção pelo Centro Universitário de Volta Redonda - UniFOA. Durante a formação acadêmica realizou alguns Projetos de Iniciação Científica, dentre os quais o artigo cujo tema é "A teoria de filas como ferramenta de apoio na análise de um serviço de atendimento", o qual está relacionado aos conceitos de Pesquisa Operacional e Métodos Quantitativos. Pós-Graduada em Gestão da Qualidade e Produtividade Industrial, interessada e dedicada nas áreas correlatas a qualidade. Atua no Centro Universitário de Volta Redonda - UniFOA.

CLARISSE DA SILVA VIEIRA

Possui graduação em Engenharia Civil pela Universidade Federal de Minas Gerais (2001), mestrado em Engenharia de Produção pela Universidade Federal de Minas Gerais (2004) e doutorado em Engenharia Mecânica (2010). Tem experiência na área de Engenharia de Produção, com ênfase em Programação Linear, Não-Linear, Mista e Dinâmica, atuando principalmente nos seguintes temas: Programação Linear, Restrição de Recurso, Sequenciamento.

CRISTIANO GRIJÓ PITANGUI

Possui graduação em Ciência da Computação pela Universidade Federal de Juiz de Fora (2004). É mestre (2007) e doutor (2013) pela COPPE/UFRJ em Engenharia de Sistemas e Computação na linha de Inteligência Artificial. É professor Adjunto do Departamento de Tecnologia e Eng. Civil, Computação e Humanidades (DTECH) da Universidade Federal de São João Del Rei (UFSJ) como também da Pós-Graduação em Gestão de Instituições Educacionais (UFVJM). Possui experiência em Aprendizado de Máquina com ênfase em Computação Bioinspirada.

DANIEL MAIA DE SOUZA

Mestre em Administração com ênfase em Estratégia e Operações pela Universidade Federal Fluminense (UFF). Pós-graduado em Gestão de Negócios pela Fundação Getúlio Vargas (FGV). Engenheiro de Produção com ênfase em Mecânica pela Universidade do Estado do Rio de Janeiro (UERJ). Analista de Planejamento Logístico na MAN Latin America. Atuação acadêmica e profissional em temas relacionados a planejamento e operação logística, gestão de terceirização, processos lean e Indústria 4.0.

ERICA XIMENES DIAS

Graduada em Engenharia Mecânica (2009), Mestre em Engenharia Mecânica (Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP, 2014), Doutora em Engenharia Mecânica (Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP, 2018).

FERNANDO AUGUSTO SILVA MARINS

Graduado em Engenharia Mecânica (1976), Mestre em Pesquisa Operacional (Instituto Tecnológico de Aeronáutica - ITA, 1981), Doutor em Engenharia Elétrica (Universidade Estadual de Campinas - UNICAMP, 1987). Fez pós-doutorado na Brunel University - Reino Unido e atualmente é Professor Titular na Faculdade de Engenharia do Campus de Guaratinguetá da Universidade Estadual Paulista "Julio de Mesquita Filho". Atua na Pós-graduação Lato e Stricto Sensu nas linhas de pesquisa envolvendo Pesquisa Operacional, Logística e Gerenciamento da Cadeia de Suprimentos.

FERNANDO LUIZ GOLDMAN

D. Sc. em Políticas Públicas, Estratégias e Desenvolvimento pelo Instituto de Economia da UFRJ, é engenheiro eletricista pela UFRJ, com ênfase em Sistemas Elétricos de Potência, tendo atuado por mais de 35 anos em grandes empreendimentos do Setor Elétrico Brasileiro. Tem especialização em Gestão Empresarial pela FGV e Mestrado Acadêmico em Engenharia de Produção pela UFF, com linha de pesquisa em Conhecimento e Inovação Tecnológica. É professor das disciplinas de Sistemas de Apoio à Decisão e Gestão da Inovação Tecnológica no Curso de Engenharia de Produção do Unifeso. Foi Diretor Regional do Pólo-RJ da Sociedade Brasileira de Gestão do Conhecimento – SBGC (2007-2011).

FRANCISCO SANTOS SABBADINI

É Doutor em engenharia de produção pela Universidade Federal do Rio de Janeiro (COPPE), professor universitário em cursos de graduação e pós graduação. Possui mestrado em administração e desenvolvimento empresarial pela UNESA. Coordena o Laboratório de Inovação em Produtos, Processos e Serviços na Faculdade de Tecnologia na Universidade do Estado do Rio de Janeiro. É membro ativo do European Working Group on Operational Research Applied Health Services (ORAHS).

FRANCISCO SEVERO DE LIMA JUNIOR

Possui graduação em Análise e Desenvolvimento de Sistemas - FATEC CARAPICUÍBA (SP) (2015). Monitor da FATEC na matéria de Estrutura de Dados. Estudante/pesquisador, participando da elaboração e desenvolvimento do Núcleo de Iniciação Científica da FATEC Carapicuíba. Participante do III Fórum Mundial de Educação Tecnológica realizado em março de 2015 no Estado de Pernambuco com a exposição do trabalho "A Escala da Ideologia da Racionalidade Tecnológica e o Tecnólogo: A Educação Tecnológica à luz da Teoria Crítica.

GUSTAVO AUGUSTO LIMA DE CAMPOS

Is an associate professor at UECE, with an expertise in the field of Artificial Intelligence, mainly on Intelligent Agents to solve decision problems in complex environment, based on systematic and local search methods, neural networks, logic and fuzzy systems. He obtained a Doctor degree on Electrical Engineering at the Systems Engineering Department, at the Electrical and Computing Engineering Faculty of the State University of Campinas, São Paulo, in 2003.

GUSTAVO JOSÉ DE AGUIAR GOMES COSTA

Possui graduação em Engenharia de Produção pela Universidade Federal de Ouro Preto (2018). Tem experiência na área de Engenharia de Produção como coordenador e supervisor de Produção e nas áreas de Contabilidade e Finanças. Atuou como diretor no Centro Acadêmico de Engenharia de Produção da Universidade Federal de Ouro Preto (2013).

GUSTAVO RAMOS LIMA

Aluno de graduação do curso de Engenharia Elétrica da Universidade Federal do Espírito Santo (UFES). Foi aluno de Iniciação Científica entre 2016 e 2017.

IGOR MAKIYAMA DE SOUZA BARBOSA

Graduando em Engenharia de Produção Mecânica (2014 - 2018) na Faculdade de Engenharia do Campus de Guaratinguetá da Universidade Estadual Paulista "Julio de Mesquita Filho", bolsista PIBIC na área de Pesquisa Operacional (2015-2017). Atualmente estagiário na empresa IOCHPE Maxion - unidade Cruzeiro-SP, no setor de PCP e em projetos de redução de custos no transporte de carga.

JOSÉ ELMO DE MENEZES

Possui graduação em Licenciatura em Matemática pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP-1993), mestrado em Matemática pela Universidade de Brasília (UNB-1998) e doutorado em Estatística pela Universidade de São Paulo (USP-2005). Atualmente é professor adjunto da Universidade Católica de Goiás(UCG) e professor E1 do IF-GO campus Inhumas. Tem experiência na área de Matemática Aplicada e Probabilidade Aplicada, com ênfase em Teoria da Confiabilidade e Análise de Sobrevivência, atuando principalmente nos seguintes temas: processo estocásticos, compensator process, martingales, reliability importance measure e algoritmo.

JOSIANE DA SILVA JESUS

Possui graduação em Engenharia de Produção pela Universidade Federal Fluminense de Volta Redonda (2014) e mestrado em Engenharia Mecânica pelo Instituto Militar de Engenharia -IME (2017) na área de concentração de Mecânica aplicada. Green Belt em Lean Seis Sigma pela Voitto Treinamento e Desenvolvimento. Atuou como estagiária na Eletrobrás Termonuclear S.A, participando no desenvolvimento de indicadores de desempenho do setor de Manutenção de Canteiros e Vilas e no projeto de melhoria contínua na logística de resíduos sólidos.

LARISSA MELO NEGRÃO

Graduada em Engenharia de Produção pelo Centro Universitário de Volta Redonda - UniFOA. Durante a formação acadêmica realizou alguns Projetos de Iniciação Científica, relacionados aos conceitos de Pesquisa Operacional. Coursou técnico profissionalizante em Administração de Empresas pelo Instituto de Cultura Técnica - ICT e é pós-graduanda em Gestão da Qualidade e Produtividade Industrial. Possui experiência como estagiário na área de Engenharia de Manutenção, nas empresas Votorantim Siderurgia e Companhia Siderúrgica Nacional - CSN. Atualmente atua como Técnico de Abastecimento na área de Planejamento e Controle da Manutenção, na Siderúrgica ArcelorMittal. Atua no Centro Universitário de Volta Redonda - UniFOA.

LEANDRO LUÍS CORSO

Graduação em Engenharia Mecânica pela Universidade de Caxias do Sul, mestrado (2006) e doutorado (2010). Pós-doutorado na Monash University (2012-2013) na área de Pesquisa Operacional. Pós-doutorado no setor militar - Naval Postgraduate School Califórnia EUA (2013). Pesquisador/professor de áreas de Otimização de Processos do Mestrado Profissional em Engenharia Mecânica da Universidade de Caxias do Sul.

LEANDRO VITRAL ANDRAOS

Possui Graduação Interdisciplinar no Bacharelado de Ciências Exatas pela Universidade Federal de Juiz de Fora, Graduação em Estatística pela Universidade Federal de Juiz de Fora e Graduação em Engenharia de Produção pela Universidade Salgado de Oliveira. Possui MBA em Administração e Finanças pela UNINTER e é Mestre em Engenharia no Instituto Militar de Engenharia (IME). Foi Professor Assistente no Curso de curta duração de Análise de dados Longitudinais ministrado no Instituto Brasileiro de Geografia e Estatística (IBGE). Atualmente trabalha como Estatístico no IBGE (aprovado em Concurso Público) lotado na Gerência de Métodos Estatísticos e é Professor/Tutor da Fundação Getúlio Vargas (FGV) e no curso ADONAI - Preparatório da Marinha.

LETTIERY D'LAMARE PORTELA PROCÓPIO

Possui graduação em Licenciado em Ciências da Computação pela Universidade Federal da Paraíba (2014). Mestrado em Computação Distribuída na área de Pesquisa Operacional pela Universidade Federal da Paraíba (2016). Atualmente é Doutorando em Ciências da Computação - Algoritmos Experimentais pela Universidade Federal do Rio Grande do Norte.

LIDIO MAURO LIMA DE CAMPOS

E Professor Adjunto da Universidade Federal do Pará - UFPA na Faculdade de Computação (Cursos de Sistemas de Informação e Ciência da Computação). Doutor em Engenharia Elétrica com ênfase em Computação Aplicada pelo Programa de Pós-Graduação em Engenharia Elétrica - PPGEE na mesma instituição (2016), Mestre em Ciência da Computação, Área de Concentração: Sistemas do Conhecimento, pela Universidade Federal de Santa Catarina - UFSC (2001), Graduado em Engenharia Elétrica pela Universidade Federal do Pará - UFPA (1998) e em Tecnólogo em Processamento de Dados pela

Universidade da Amazônia - UNAMA. É Certificado PMP-Project Management Professional pelo Project Management Institute (PMI-Pennsylvania-EUA). Na área de pesquisa e desenvolvimento atua em Computação Natural, Aprendizagem de Máquina, Inteligência Computacional, Mineração de Dados e Gerência de Projetos.

LUCIANA PEREIRA DE ASSIS

É Professora do Curso de Sistemas de Informação da Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM). Também atua como Professora do Programa de Pós-Graduação em Educação (PPGED/UFVJM). Graduiu-se em Ciência da Computação (UNIBH), cursou Mestrado em Ciência da Computação (UFMG) e Doutorado em Engenharia Elétrica (UFMG). Tem interesses nas áreas de Otimização, Inteligência Artificial e Análise de Redes Sociais.

LUCÍDIO DOS ANJOS FORMIGA CABRAL

Possui graduação em Ciência da Computação pela Universidade Federal do Ceará (1990), mestrado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (1993) e doutorado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (2001). Atualmente é professor Titular do Departamento de Computação Científica do Centro de Informática da Universidade Federal da Paraíba.

LUÍS MIGUEL ALÇADA TOMÁS ALMEIDA

A 5-year first-degree in Informatics Engineering (1989) and a Ph.D. in Management Sciences (Information Management) at the University of Coimbra (2007). He is an Assistant Professor at the Faculty of Economics - University of Coimbra. In addition, he has been a researcher at the Institute for Systems Engineering and Computers (INESC) since 1990. His lecturing and research interests include Information Systems Analysis and Algorithm Development, Multimedia Systems, Geographical Information Systems, Databases and Algorithm Servers. He has published in scientific journals such as Decision Support Systems, Geographical Analysis and Socio-Economic Planning Sciences.

LUIZ ANTONIO FIGUEIREDO

Mestre em Engenharia De Produção graduado pela Pontifícia Universidade Católica PUC, Graduado em Gestão da Produção e Pós graduado em Gestão da Qualidade e Produtividade pela Faculdade Oswaldo Cruz e pela Faculdade de Tecnologia Gerencial do SENAI em Logística Empresarial, Formado também pela Escola de Especialistas de Aeronáutica como Técnico em Manutenção de Aeronaves nas habilitações Aviônicos, Motores e Célula, Piloto de Avião e Helicóptero, Instrutor de voo pelo Instituto de Aviação Civil. Atualmente Coordenador de curso da Faculdade de Tecnologia SENAI Ítalo Bologna.

LUIZ OTAVIO RIBEIRO AFONSO FERREIRA

Doutorando em Ciências Aplicada a Decisão na Universidade de Coimbra, possui mestrado em Informática Aplicada, MBA em Gerenciamento de Projetos e graduação em Ciências Econômicas pela Universidade de Fortaleza (2003, 2010 e 1995) e especialização em informática pela Universidade Federal do Ceará (1997). Competências em Sistema de Suporte à Decisão, Sistema de Informação Geográfica e Ciência da Computação, com ênfase na Matemática da Computação, em Análise Estatística e Gerência de Projetos. Área de pesquisa: Pesquisa Operacional, Apoio à Decisão Multicritério, Ciências de Dados e Inteligência Empresarial. Atuação profissional como: Docente Universitário, Cientista de Dados, Gerente de Projetos e Analista de Sistemas com vasta experiência em gerenciamento de projetos (CPM, PERT, EVA, métricas) em TIC, banco de dados (oracle, postgresql, mysql, mariadb, mongodb, nosql), integração de dados com ETL, modelagem de sistemas e análise

estatísticas (R, SSPS, Stata e Excel avançado). Conhecimento em linguagens como: python, php, delphi, java e javascript.

MARCELO FERREIRA REGO

Graduado em Sistemas de Informação pela Universidade Estadual de Montes Claros (2006), Especialista em Administração de Sistemas de Informação pela Universidade Federal de Lavras (2008), Mestre em Ciência da Computação pela Universidade Federal de Ouro Preto (2013). Foi Analista de Tecnologia da Informação da Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM) por 8 anos, atuando na área de Desenvolvimento de Sistemas. Desde de 2015 é professor assistente do curso de Sistemas de Informação da UFVJM. Interesse em pesquisas na área de metaheurísticas, otimização multiobjetivo e pesquisa operacional. Experiência profissional em programação WEB, Framework MVC, Banco de Dados, Engenharia de Software, Gerência de Projetos.

MARCELO VILELA FIGUEIREDO

Marcelo Vilela Figueiredo, formado em Engenharia de Produção pela UFOP (2010), Especialização em Gestão da Produção pela FEG-UNESP (2012), Mestrado em Engenharia de Produção pela FEG-UNESP (2017) e doutorando em Engenharia Mecânica pela FEG-UNESP. Atuação profissional como Engenheiro de Processos Pleno no setor de Melhoria Contínua (Seis Sigma e Lean Manufacturing) na AmstedMaxion Fundição e Equipamentos Ferroviários SA de 01/2011 a 12/2017 e como Gerente de Performance Improvement na Ernst & Young de 12/2017 até o momento. Experiência profissional com foco em Controle Estatístico de Processos e Pesquisa Operacional.

MARCO CESAR GOLDBARG

Possui graduação na Academia Militar de Agulhas Negras (1975), graduação em Engenharia de Fortificação e Construção pelo Instituto Militar de Engenharia (1982), mestrado em Sistemas e Computação pelo Instituto Militar de Engenharia (1987), doutorado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (1990), pós-doutorado pela Universidade Federal de Minas Gerais (1999). Atualmente é Professor Titular da Universidade Federal do Rio Grande do Norte.

MARIA CRISTINA VENDRAMETO

Possui graduação em História pela Universidade de São Paulo (2001) e mestrado em Educação pela Universidade de São Paulo (2005), doutorado em andamento na Pontifical Universidade Católica de São Paulo - PUC/SP. Foi coordenador pedagógico curso Logística e Transporte até 2012 e professor titular na FATEC CARAPICUÍBA Centro Estadual de Educação Tecnológica Paula Souza. Tem experiência na área de Educação, com ênfase em História e Historiografia da Educação, atuando principalmente nos seguintes temas: fontes documentais, arquivologia, história da educação, educação profissional e memória. Pesquisa atualmente áreas de Políticas Públicas, Sociologia do Trabalho, Economia, Formação Tecnológica.

MÁRIO JORGE FERREIRA DE OLIVEIRA

Professor Titular do Instituto de Matemática da UFRJ. Possui doutorado em Pesquisa Operacional - University of Strathclyde e pós doutorado em Sistemas de Informação- London School of Economics and Political Sciences. Trabalha concomitantemente no departamento de matemática aplicada do Instituto de Matemática da UFRJ. É membro ativo do European Working Group on Operational Research Applied Health Services (ORAHS).

MAX SANTANA RELEMBERG FARIAS

Bacharel em Ciência da Computação pela Universidade Tiradentes (2001), mestre em Modelagem Computacional de Conhecimento pela Universidade Federal de Alagoas (2008) e doutor em Ciência da Computação pela Universidade Federal de Pernambuco (2014). Em 2003 ingressou na carreira acadêmica, como professor substituto, no Centro Federal de Educação Tecnológica do Rio Grande do Norte (CEFET-RN) e em 2004 foi aprovado em concurso público para professor efetivo da Universidade Federal do Vale do São Francisco (UNIVASF). Tem experiência na área de Ciência da Computação, com ênfase em Linguagens de Programação e Arquitetura de Sistemas de Computação, atuando principalmente nos seguintes temas: algoritmos e programação, algoritmos evolucionários, otimização multiobjetivo, arquitetura e organização de computadores e MPSOCs.

MURILO VIEIRA STRONGOLI

graduando em Engenharia de Produção Mecânica (2014 - 2018) na Faculdade de Engenharia do Campus de Guaratinguetá da Universidade Estadual Paulista "Julio de Mesquita Filho", bolsista PIBIC na área de Pesquisa Operacional (2015-2017). Atualmente estagiário na empresa IOCHPE Maxion - unidade Cruzeiro-SP, no setor de PCP e em processos de manufatura de chassis.

MYLENA CRISTINA REZENDE PACHECO

Mylena Cristina Rezende Pacheco é graduada em Engenharia de Produção pelo Centro Universitário Serra dos Órgãos - UNIFESO (2015). Na área da Pesquisa Operacional publicou os artigos: AHP como um modelo matemático: uma análise de sensibilidade simples - XII Simpósio de Excelência e Gestão em Tecnologia (SEGeT 2016); Modelos Multicriteriais de Apoio à Decisão: O Método AHP como auxílio à seleção de fornecedores em uma confecção - XXXVI Encontro Nacional de Engenharia de Produção (ENEGEP 2016), cujo tema também obteve menção honrosa por sua apresentação e discussão no I Congresso Acadêmico Científico do UNIFESO (CONFESO - 2016). Na área da Gestão do Produto publicou: A utilização de ferramentas de planejamento de produto como fator determinante para mudança de concepção do seu desenvolvimento: Projeto carrega chuva - XXI Simpósio de Engenharia de Produção (SIMPEP 2014).

NATÁLIA NOGUEIRA FERREIRA SOUZA

Mestre em Engenharia pelo ITA e formada em Engenharia Química pela Escola de Engenharia de Lorena - USP, cursou o PEE, Programa de Especialização de Engenharia e hoje trabalha como engenheira de vendas da área Comercial na indústria aeronáutica brasileira, atuando na parte de suporte de vendas de materiais.

NATANAEL SALGADO COUTINHO

Possui graduação em Engenharia de Produção pela Universidade Federal de Ouro Preto (2017). Tem experiência na área de Engenharia de Produção, atuando principalmente nos seguintes temas: recurso, resource, sequenciamento, scheduling e project.

NILSON BRANDALISE

Possui Doutorado em Engenharia Civil pela Universidade Federal Fluminense (2017), Mestrado em Administração pela Universidade Federal do Paraná (2001), Especialização em Comércio Exterior (1996), MBA Executivo (1998), Graduação em Administração pela Faculdade Católica de Administração e Economia (1981). Atualmente é professor adjunto da Universidade Federal Fluminense. Tem experiência na área de Administração, com ênfase em Administração Financeira, atuando principalmente nos seguintes temas: administração,

administração financeira, engenharia econômica, matemática financeira, comportamento organizacional, orçamento e qualidade.

PÂMELA BRAGA BERTAHONE

Possui graduação em Engenharia de Produção pela Universidade Federal Fluminense (2017). Atualmente cursa Mestrado Profissional em Engenharia de Produção pela mesma instituição.

PATRICK MARQUES CIARELLI

Possui graduação em Engenharia Elétrica pela Universidade Federal do Espírito Santo (2006), mestrado e doutorado em Engenharia Elétrica pela Universidade Federal do Espírito Santo em 2008 e 2012, respectivamente. Atualmente é professor da Universidade Federal do Espírito Santo. Tem experiência na área de Engenharia Elétrica, com ênfase em Inteligência Artificial, Recuperação de Informação, Reconhecimento de Padrões e Processamento de Imagens.

PAULI ADRIANO DE ALMADA GARCIA

Graduado em Matemática pela UFF (1999), Mestre em Sistemas e Computação pelo IME (2001) e Doutor em Engenharia Nuclear pela UFRJ - COPPE (2006). Membro da SOBRAPO e da ABRISCO. Atualmente é Professor Associado do Departamento de Administração e Administração Pública da UFF em Volta Redonda, onde coordena o curso de graduação em Administração e o grupo de pesquisa ARCADE. Atua como colaborador/pesquisador do Grupo LOGIS junto ao Programa de Engenharia de Produção da UFF/Niterói, onde também atua em orientação de mestrado e doutorado. Tem experiência em análise de confiabilidade de sistemas, análise probabilística de riscos em Engenharia de Poço e Nuclear.

PAULO HENRIQUE COELHO MARANHÃO

Possui graduação em Bacharelado em Estatística pela Universidade Federal do Ceará (1994), mestrado em Estatística pela Universidade Federal de Pernambuco (1998) e Doutorado em Engenharia de Produção pela PUC-RJ (2012). Atualmente é professor do Instituto Militar de Engenharia (IME). Tem experiência na área de Estatística e Engenharia de Produção, em especial nas áreas de Controle Estatístico de Qualidade, Estatística Aplicada e Estatística Multivariada.

PÍTIAS TEODORO

Doutorado em Engenharia de Transportes pela Universidade Federal do Rio de Janeiro - PET/COPPE/UFRJ (2015); Mestrado em Administração pela Universidade Federal de Lavras - PRPG/DAE/UFLA (2005); Especialização em Gestão Financeira pela Universidade Federal de Juiz de Fora FEA/UFJF (2001) e Graduação em Administração pela Universidade Federal de Juiz de Fora - FEA/UFJF (2000). Professor Adjunto da Universidade Federal Fluminense - Instituto de Ciências Humanas e Sociais de Volta Redonda. Tem experiência na área de Administração em Finanças, com ênfase na elaboração e análise de projetos para investimentos públicos (viabilidade socioeconômica) e privados (viabilidade técnico-econômico-financeira).

RAFAEL DE CARVALHO MIRANDA

Rafael de Carvalho Miranda, Professor Adjunto da Universidade Federal de Itajubá (UNIFEI) no Instituto de Engenharia de Produção e Gestão (IEPG). Possui graduação em Engenharia de Produção-Mecânica pela UNIFEI (2008) e mestrado (2012) e doutorado (2015) em Engenharia de Produção pela mesma instituição. É pesquisador do Núcleo de Estudos Avançados para Auxílio a Decisão (NEAAD), tendo experiência profissional e acadêmica nas áreas de Engenharia de Produção, Pesquisa Operacional, Simulação e Eventos Discretos,

Métodos Quantitativos para Apoio a Tomada de Decisão, Otimização e Engenharia Econômica.

RAFAEL LUSTOSA DA COSTA

Graduado em Logística

RAFAELA BOEIRA CECCHIN

Possui graduação em Engenharia de Produção e mestrado em Engenharia Mecânica pela Universidade de Caxias do Sul (UCS), com foco em Pesquisa Operacional e Otimização.

RAMON ROCHA LEITE

Bacharel em Sistemas de Informação pela Universidade Federal dos Vales do Jequitinhonha e Mucuri - UFVJM. Especialista em Tecnologias Digitais e Inovação na Educação pela Universidade Cândido Mendes - UCAM. Atualmente é Analista de Tecnologia da Informação na UFVJM.

RICARDO LUIZ MACHADO

Doutor em Engenharia de Produção pela Universidade Federal de Santa Catarina, na área de Planejamento de Operações (2003) e Mestre em Engenharia de Produção pela mesma instituição, na área de Gestão da Qualidade e Produtividade (1997). É Engenheiro Civil pela Universidade Federal de Goiás (1992). É coordenador do Programa de Pós-Graduação em Engenharia de Produção e Sistemas da PUC Goiás, ministrando disciplinas e orientando dissertações na linha de pesquisa de Gestão de Operações e Logística. É docente em cursos de pós-graduação lato sensu, ministrando disciplinas de Gestão de Operações Industriais e de Serviços, Logística Empresarial e Planejamento e Controle da Produção.

ROSALVO FERREIRA DE OLIVEIRA NETO

Graduado em Sistemas de Informação pela Faculdade Integrada do Recife (2004), com Mestrado em Ciências da Computação pela Universidade Federal de Pernambuco (2008) e Doutorado em Ciências da Computação pela Universidade Federal de Pernambuco (2015). Atualmente é Professor e Pesquisador na Universidade Federal do Vale do São Francisco. Tem experiência na área de Ciência da Computação, com ênfase em Mineração de Dados, Inteligência Artificial, Banco de Dados e Engenharia de Software, atuando principalmente nos seguintes temas: processo de descoberta de conhecimento em bases de dados, redes neurais, algoritmos de buscas, data warehouse e desenvolvimento dirigido por modelos.

ROSINEI BATISTA

Pós Doutorando no Instituto Tecnológico de Aeronáutica (ITA) na área de corrosão em aços aeronáuticos no Instituto de Estudos Avançados - IEAv, Departamento de Ciência e Tecnologia Aeroespacial - DCTA - 2018-2020. Pós-Doutorado em Engenharia Mecânica pela Universidade Estadual Paulista Júlio de Mesquita Filho, Campus de Guaratinguetá - FEG-UNESP (2011-2014). Doutorado em Engenharia Mecânica pela UNESP- FEG (2004). Pró-Reitor de Pesquisa, Pós-Graduação e Extensão do Centro Universitário Teresa D'Ávila - UNIFATEA. Docente Permanente no Programa de Pós-Graduação Mestrado Profissional em Design, Tecnologia e Inovação - PPG-DTI. Docente do Curso Superior de Tecnologia em Gestão da Produção Industrial - CSTGPI na Faculdade de Tecnologia do Estado de São Paulo - FATEC - Unidade de Cruzeiro. Docente Permanente no Programa de Pós-graduação - Mestrado Profissional em Engenharia de Materiais e no Programa de Pós-Graduação - Mestrado Acadêmico em Desenvolvimento, Tecnologias, Sociedade - UNIFEI. Possui experiência nas atividades ensino, pesquisa e extensão nas áreas de Desenho Industrial, Engenharia de Materiais, Mecânica, Qualidade e Produção, com ênfase em Novos Materiais,

Tecnologias Sociais, Planejamento de Experimentos, Processos de Fabricação, Corrosão, Caracterização Microestrutural e Gerenciamento da Cadeia de Suprimentos.

TAKASHI YONEYAMA

Concluiu Engenharia Eletrônica pelo Instituto Tecnológico de Aeronáutica - ITA (1975), Medicina pela Universidade de Taubaté (1993)., Mestrado em Engenharia Eletrônica pelo ITA (1979) e Doutorado em Engenharia Elétrica pela Universidade de Londres (1983). Desde 1990 é Professor Titular do ITA. Foi Presidente (2004-2006) da Sociedade Brasileira de Automática. É co-autor de 4 livros. Possui mais de 99 artigos em periódicos e 350 em eventos científicos. Orientou, 31 doutores, 84 mestres e 67 bolsistas de iniciação científica, bem como 159 trabalhos de conclusão de curso.

VINÍCIUS BANDEIRA

Possui graduação em Engenharia de Produção pela Universidade Federal de Ouro Preto (2016). Atualmente é Gestor Administrativo Jr. - Epar Business Expert. Tem experiência na área de Engenharia de Produção, com ênfase em Planejamento, Projeto e Controle de Sistemas de Produção.

VINÍCIUS ROSSI OLIVEIRA

Possui graduação em Engenharia de Produção pela Universidade Federal de Ouro Preto (2016). Atualmente é Engenheiro de Produção da Universidade Federal do Espírito Santo. Tem experiência na área de Engenharia de Produção.

Agência Brasileira do ISBN

ISBN 978-85-7042-007-7

9 788570 420077