

# TECNOLOGIA & CULTURA

Revista do Centro Federal de Educação Tecnológica Celso Suckow da Fonseca  
Cefet/RJ | Edição Especial | 2024

Trabalhos selecionados do  
“Encontro Regional de Matemática Aplicada  
e Computacional (ERMAC-2023) &  
Simpósio 1 Década do PPG-MCCT”

# TECNOLOGIA & CULTURA



CEFET/RJ - CENTRO FEDERAL DE EDUCAÇÃO  
TECNOLÓGICA CELSO SUCKOW DA FONSECA

Ministério da Educação (MEC)  
Secretaria de Educação Profissional e Tecnológica (Setec)

CEFET/RJ - CENTRO FEDERAL DE EDUCAÇÃO  
TECNOLÓGICA CELSO SUCKOW DA FONSECA

TECNOLOGIA & CULTURA - Revista do Cefet/RJ  
Edição Especial. 2024

Edição eletrônica disponível em: <https://www.cefet-rj.br/index.php/revista-tecnologia-cultura>

Av. Maracanã, 229 - Rio de Janeiro/RJ  
CEP 20.271-110

Telefone geral: (21) 2566-3022 r. 3160

Telefax: (21) 2284-6021

<http://www.cefet-rj.br>

E-mail: [revista@cefet-rj.br](mailto:revista@cefet-rj.br)

#### **Diretor-geral**

Maurício Saldanha Motta

#### **Vice-diretor**

Gisele Maria Ribeiro Vieira

#### **Diretora de Ensino**

Dayse Haime Pastore

#### **Diretor de Pesquisa e Pós-graduação**

Ronney Arismel Mancebo Boloy

#### **Diretora de Gestão Estratégica**

Célia Machado Guimarães e Souza

#### **Diretora de Extensão**

Renata da Silva Moura

#### **Diretor de Administração e Planejamento**

Bianca de França Tempone Felga de Moraes

#### **Editoria**

Taís Conceição dos Santos

#### **Biblioteca Central**

Mariana de Oliveira Caruso Carvalho

#### **Projeto Gráfico/Diagramação**

Divisão de Programação Visual (DPROV)

#### **Comité Científico, Editorial e Coordenadores de Áreas Temáticas do ERMAC**

Alexandre Santos Francisco (UFF - Universidade Federal Fluminense)

Amaury Alvarez Cruz (UFRJ - Universidade Federal de Rio de Janeiro)

André Ebling Brondani (UFF - Universidade Federal Fluminense)

Claudia Mazza Dias (UFRRJ - Universidade Federal Rural de Rio de Janeiro)

Dayse Haime Pastore (CEFET-RJ), Centro Federal de Educação Tecnológica

Celso Suckow da Fonseca)

Gustavo Benitez Alvarez (UFF - Universidade Federal Fluminense)

Isaac Pinheiro dos Santos (UFES - Universidade Federal de Espírito Santo)

Panters Rodríguez Bermudez (UFF - Universidade Federal Fluminense)

Reinaldo Rodriguez Ramos (Universidad de la Habana, Cuba e UFF - Universidade Federal Fluminense)

Sandra Malta (LNCC - Laboratório Nacional de Computação Científica)

Tiago Neves (UFF - Universidade Federal Fluminense)

Wagner Rambaldi Telles (UFF - Universidade Federal Fluminense)

Zochil Gonzalez Arenas (UERJ - Universidade do Estado do Rio de Janeiro)

#### **Revisores dos artigos selecionados do ERMAC para número especial da revista Tecnologia & Cultura**

Alex Tavares Silva (IFF-RJ, Instituto Federal Fluminense-RJ)

Allana Sthel Santos De Oliveira (UFF - Universidade Federal Fluminense)

André Ebling Brondani (UFF - Universidade Federal Fluminense)

André Renato Villela Da Silva (UFF - Universidade Federal Fluminense)

Cecília Toledo Hernández (UFF - Universidade Federal Fluminense)

Eliane Christo (UFF - Universidade Federal Fluminense)

Eliza Maria Ferreira (UFLA - Universidade Federal De Lavras / Universidad

Pedro de Valdivia - Chile)

Fábio Freitas Ferreira (UFF - Universidade Federal Fluminense)

Francisca Andrea Macedo França (UFF - Universidade Federal Fluminense)

Gustavo Semaan (UFF - Universidade Federal Fluminense)

Igor Campos De Almeida Lima (UERJ - Universidade do Estado do Rio de Janeiro)

Jorge Alberto Rodriguez Duran (UFF - Universidade Federal Fluminense)

Kelly Alonso Costa (UFF - Universidade Federal Fluminense)

Marcela Lima Santos Pereira (UFRRJ - Universidade Federal Rural de Rio de Janeiro)

Marcello Montillo Provenza (UERJ - Universidade do Estado do Rio de Janeiro)

Marcio Borges (LNCC - Laboratório Nacional de Computação Científica)

Mateus Pereira Martin (UFF - Universidade Federal Fluminense)

Mauricio Calvao (UFRJ - Universidade Federal de Rio de Janeiro)

Sérgio Souza Bento (UFES - Universidade Federal de Espírito Santo)

Thiago Jordem Pereira (UFF - Universidade Federal Fluminense)

Tecnologia & Cultura. \_ Edição Especial 2024) -  
Rio de Janeiro : Centro Federal de Educação  
Tecnológica Celso Suckow da Fonseca, 2019.  
v. : il.; 28 cms.

Semestral  
ISSN 1414-8498

I. Centro Federal de Educação Tecnológica Celso  
Suckow da Fonseca

#### **Observações**

Os conteúdos dos artigos publicados nesta revista são de inteira responsabilidade de seus autores. Proibida a reprodução total ou parcial desta obra sem autorização dos autores.

<b>PROMOVENDO REFLEXÕES SOBRE O USO DA MODELAGEM MATEMÁTICA ATRAVÉS DA LEI DE NEWCOMB-BENFORD .....</b>	<b>6</b>
Victor Augusto do Carmo Duarte Lavínia Beatriz Duarte da Silva Marcelo Ferreira Farias	
<b>PROPRIEDADES ESPECTRAIS DE GRAFOS BENZENO´IDES .....</b>	<b>14</b>
Allan José Amaral Ribeiro Francisca Andrea Macedo França André Ebling Brondani	
<b>MODELAGEM 3D DE CONCRETO SOBRE ESTRESSE TÉRMICO APLICANDO O MÉTODO DOS VOLUMES FINITOS .....</b>	<b>24</b>
João Victor de Mello Marinho Victor de Oliveira Carvalho Jorge Juárez Trujillo José Adilson de Castro	
<b>ROTEAMENTO DE MÉDICOS POR ESPECIALIDADE EM SITUAÇÃO DE PÓS - DESASTRE ATRAVÉS DE UM MODELO DE PROGRAMAÇÃO LINEAR INTEIRA MISTA .....</b>	<b>32</b>
Raila de Medeiros Melo Eliane da Silva Christo Tiago Araújo Neves	
<b>SEGMENTAÇÃO DE IMAGENS DE RESSONÂNCIA MAGNÉTICA DE TUMOR CEREBRAL UTILIZANDO MODELO 3D U-NET .....</b>	<b>40</b>
Erika Rocha de Araújo Vanessa da Silva Garcia Wesley Luiz da Silva Assis	
<b>O -ESPECTRO DO GRAFO BLOW-UP THE -SPECTRUM OF THE BLOW-UP GRAPH.....</b>	<b>50</b>
Uilton Cesar Peres Junior Carla Silva Oliveira André Ebling Brondani	
<b>ANÁLISE DO TRANSPORTE DE CONTAMINANTES EM UM RIO UTILIZANDO PROBLEMAS INVERSOS .....</b>	<b>59</b>
Anna Luisa Soares Castro Coimbra Fábio Freitas Ferreira Wagner Rambaldi Telles	
<b>ANÁLISE DO ROMPIMENTO DE BARRAGENS DA UHE BARRA DO BRAÚNA UTILIZANDO O SOFTWARE DE SIMULAÇÃO HIDRODINÂMICA IBER .....</b>	<b>67</b>
Sânzzia de Oliveira e Souza Figueira Thiago Jordem Pereira Ricardo Silveira Sousa Tibério Borges Vale Jader Lugon Junior Wagner Rambaldi Telles	
<b>ALGORITMOS DE AGRUPAMENTOS NÃO-SUPERVISIONADOS APLICADOS À TAXONOMIA E FILOGENIA DE LEISHMANIA spp. ....</b>	<b>78</b>
Antonio Revail Alves Pereira Marcello Goulart Teixeira Flávio Luis de Mello	
<b>PREVISÃO DE ENCHENTES NO RIO DOCE UTILIZANDO MODELO SARIMA .....</b>	<b>85</b>
Jean Sousa Mendes Eliane da Silva Christo Kelly Alonso Costa	
<b>APLICAÇÃO DE UMA FORMULAÇÃO MATEMÁTICA PARA MONTAGEM DE PLANOS DE ESTUDOS DOS ALUNOS DE UM CURSO DE ENGENHARIA .....</b>	<b>93</b>
Tiago Araújo Neves	
<b>METODO PRIMAL SIMPLEX ESPECIALIZADO EM PROBLEMA DA PRODUÇÃO DE RAÇÃO .....</b>	<b>103</b>
Marcio Lopes Gevegy Ricardo Silveira Sousa Wagner Rambaldi Telles	
<b>PREVISÃO DO CONSUMO DE ENERGIA ELÉTRICA ATRAVÉS DE MODELOS DE SÉRIES TEMPORAIS UNIVARIADOS E MULTIVARIADOS.....</b>	<b>113</b>
Rafael Santos de Oliveira Felipe Leite Coelho da Silva Josiane da Silva Cordeiro Coelho	

<b>UMA ANÁLISE ESTATÍSTICA DO LIMITE DE ESCOAMENTO DA LIGA INCONEL 690 TT A TEMPERATURA AMBIENTE .....</b>	<b>120</b>
Alexandre Santos Francisco Matheus Felipe Silva Oliveira	
<b>DESENVOLVIMENTO DE SOFTWARE PARA DIMENSIONAMENTO E SELEÇÃO DE ESTRUTURAS DE EDIFÍCIO EM CONCRETO ARMADO COM REDES NEURAIS ARTIFICIAIS OTIMIZADAS E AHP .....</b>	<b>128</b>
Adriano Ricardo Borges Costa Eliane da Silva Christo Tiago Araújo Neves	
<b>PROPOSTA METODOLÓGICA PARA A REVISÃO SISTEMÁTICA DA LITERATURA: ESTUDO DE CASO ASSOCIADO À TEORIA DAS FILAS E SIMULAÇÃO .....</b>	<b>139</b>
Kelly Cristine Franco de Azevedo Cecília Toledo Hernández Eliane da Silva Christo	
<b>PROBLEMAS DIFUSIVOS-CONVECTIVOS ABORDADOS PELO MÉTODO DAS DIFERENÇAS FINITAS CENTRADAS E UPWIND .....</b>	<b>148</b>
Victor de Souza Coelho Freitas Gustavo Benitez Alvarez Cleyton Senior Stampa Panthers Rodríguez Bermúdez	
<b>IMPLEMENTAÇÃO EM VOLUMES FINITOS DE UM MÉTODO TIPO GODUNOV PARA SISTEMAS 2 X 2 DE LEIS DE CONSERVAÇÃO .....</b>	<b>158</b>
Eron Oliveira do Nascimento Panthers Rodríguez Bermúdez Wagner Rambaldi Telles	
<b>COMPARAÇÃO ENTRE MODELOS PREDITIVOS EM SÉRIES TEMPORAIS HIERÁRQUICAS PARA O CONSUMO DE ENERGIA ELÉTRICA DO SETOR INDUSTRIAL BRASILEIRO .....</b>	<b>167</b>
Marlon Mesquita Lopes Cabreira Felipe Leite Coelho da Silva Josiane da Silva Cordeiro Coelho	
<b>CONSTRUÇÃO E CARACTERIZAÇÃO DE LOOPS DE CÓDIGO .....</b>	<b>175</b>
Rosemary Miguel Pires	

É com muita satisfação que apresentamos nesta edição especial da revista Tecnologia & Cultura alguns dos melhores trabalhos selecionados do “Encontro Regional de Matemática Aplicada e Computacional (ERMAC-2023) & Simpósio 1 Década do PPG-MCCT”, realizado entre os dias 30 de Outubro - 01 de Novembro de 2023 na cidade de Volta Redonda.

Todos os anos, a Sociedade Brasileira de Matemática Aplicada e Computacional (SBMAC) promove em suas Regionais os Encontros Regionais de Matemática Aplicada e Computacional, com a sigla ERMAC. Após várias edições em múltiplas cidades dos estados do Espírito Santo e Rio de Janeiro, A Regional 06 da SBMAC promoveu o evento de 2023 na cidade de Volta Redonda, no sul fluminense, onde foi organizado pelos professores ligados ao Programa de Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia (PPG-MCCT) e realizado nas instalações da Escola de Engenharia Industrial Metalúrgica de Volta Redonda (EIMVR), ligada à Universidade Federal Fluminense (UFF).

O objetivo deste evento é divulgar a produção, valorizar os esforços e fomentar iniciativas que envolvam a matemática aplicada e computacional particularmente junto às instituições de ensino e pesquisa dos estados do Rio de Janeiro e Espírito Santo, e destina-se aos alunos de graduação, pós-graduação, docentes e pesquisadores do tema e de suas áreas correlatas.

Junto ao ERMAC-2023, também foi realizado o “Simpósio 1 Década do PPG-MCCT”, evento comemorativo pelos primeiros 10 anos de funcionamento do programa. O PPG-MCCT é um programa interdisciplinar que visa trabalhar na interface entre as ciências tradicionais, buscando agregar áreas distintas do conhecimento. Ao longo dos 10 anos de funcionamento, o programa já formou mais de 90 mestres nas linhas de métodos computacionais, Otimização e Física Computacional, mostrando que o programa tem sido bem-sucedido em seus objetivos.

O evento recebeu 74 trabalhos submetidos, 50 na modalidade artigo e 24 na modalidade resumo estendido. Após a avaliação pelo comitê científico e a equipe de revisores, houve 68 trabalhos aprovados para apresentação no evento, 44 artigos e 24 resumos estendidos. Se inscreveram 156 participantes, sendo 71 estudantes de graduação, 38 estudantes de pós-graduação/professores da rede municipal ou estadual de ensino básico, 47 professores ou pesquisadores do ensino superior federal. Durante o evento houve 133 participantes credenciados.

Os artigos publicados na presente edição especial da revista Tecnologia & Cultura, mostram o caráter multidisciplinar e interdisciplinar do evento realizado, contendo trabalhos correspondentes a distintas áreas temáticas: Modelagem Matemática e Aplicações, Pesquisa Operacional, Otimização e Controle, Métodos Numéricos e Aplicações, Matemática Estocástica, Probabilidade e Estatística, Matemática Discreta, Matemática Aplicada à Engenharia, Biomatemática, Ensino de Matemática, entre outras. A seleção foi feita por indicação dos próprios revisores e pelo comitê científico do evento.

O evento mostrou o alcance e o potencial da área de matemática aplicada. Mais ainda, mostrou o quão dedicados são os pesquisadores e estudantes da área. Muitas das pesquisas apresentadas possuem impacto direto na sociedade ou nasceram de problemas enfrentados pela mesma. Isto mostra o quão importante é a motivação e o trabalho dos pesquisadores dos estados cobertos pelo evento, e de maneira mais ampla, dos profissionais da ciência. Esta paixão que move os pesquisadores e estudantes é contagiante e nos instiga a seguir na busca de soluções sustentáveis, eficientes e justas para os problemas do nosso tempo, sem perder o foco no futuro.

“Dedicado à memória do professor Diomar César Lobão”

Tiago Araújo Neves e Panters Rodríguez-Bermudez

## PROMOVENDO REFLEXÕES SOBRE O USO DA MODELAGEM MATEMÁTICA ATRAVÉS DA LEI DE NEWCOMB-BENFORD

### *PROMOTING REFLECTION ON THE USE OF MATHEMATICAL MODELING THROUGH THE NEWCOMB-BENFORD LAW*

Victor Augusto do Carmo Duarte

Lavínia Beatriz Duarte da Silva

Marcelo Ferreira Farias

**RESUMO:** A lei de Newcomb-Benford é um fenômeno estatístico que descreve a distribuição das frequências dos dígitos iniciais provenientes de uma ampla gama de conjuntos de dados. Essa lei tem sido aplicada com sucesso em auditorias financeiras, análise de dados e investigações forenses, mas certos abusos são cometidos, o que nos leva à discussão sobre o mau uso das ferramentas matemáticas na tomada de decisões em problemas reais. Neste trabalho, relatamos uma experiência na realização de uma atividade com uma turma de ingressantes do curso de Matemática, da UFRRJ, na qual apresentamos as frequências dos dígitos iniciais de dados numéricos de diferentes origens e sua conformidade à lei, discutimos os limites do uso de argumentos matemáticos na modelagem de problemas reais, em acordo com as preocupações da Matemática Crítica, e propomos um experimento feito pelos próprios alunos envolvendo a distribuição dos primeiros dígitos significativos no número de seguidores em redes sociais.

**Palavras-chave:** Ensino de Matemática. Educação Matemática Crítica. Modelagem Matemática. Probabilidade. Lei de Newcomb-Benford.

**ABSTRACT:** The Newcomb-Benford law is a statistical phenomenon that describes the distribution of initial digits in a wide range of datasets. While this law has been successfully applied in financial audits, data analysis, and forensic investigations, certain misuses have been observed, prompting a discussion on the inappropriate use of mathematical tools in decision-making for real-world problems. In this paper, we report an experience conducting an activity with a group of incoming Mathematics students at UFRRJ. We presented the frequencies of initial digits in numerical data from various sources and their compliance with the Newcomb-Benford law. We discussed the limitations of using mathematical arguments in modeling real-world problems, aligning with the concerns of Critical Mathematics. Additionally, we proposed an experiment for the students involving the distribution of the first significant digits in the number of followers on social media.

**Keywords:** Mathematics Education. Critical Mathematics Education. Mathematical Modeling. Probability. Newcomb-Benford Law.

## INTRODUÇÃO

A lei de Newcomb-Benford estabelece que, em muitos conjuntos de dados numéricos, as frequências relativas de ocorrência dos dígitos iniciais seguem um padrão específico, diferente de uma distribuição equiprovável, que seria o mais natural, *a priori*. De acordo com essa lei, o dígito “1” ocorre como o primeiro dígito significativo em cerca de 30,1% dos casos; o dígito “2” aparece menos, aproximadamente em 17,6% das ocorrências, e decresce sucessivamente até o dígito “9”, com cerca de 4,6% de frequência relativa. Assim, a probabilidade de ocorrência de um dígito inicial não é uniforme e segue uma distribuição logarítmica (Benford, 1938).

A observação dessa lei remonta ao trabalho de Simon Newcomb (1835-1909), um astrônomo e matemático canadense, que observou a tendência de certos dígitos ocorrerem com mais frequência em tabelas de logaritmos consultadas por diversos pesquisadores de diferentes áreas do conhecimento, pois as páginas que traziam números começados com o dígito “1” estavam mais desgastadas por serem as mais usadas, seguidas pelas páginas de números iniciados em “2”, e assim sucessivamente até o “9” (Newcomb, 1881). No entanto, foi o matemático Frank Benford (1883-1948) que, em 1938, popularizou essa observação empírica e formulou uma explicação teórica para o fenômeno. Desde então, a lei de Newcomb-Benford tem sido objeto de extenso estudo e aplicação em várias áreas do conhecimento.

Além da fundamentação teórica, tem sido confirmada empiricamente em uma ampla variedade de conjuntos de dados reais. Estudos têm mostrado que essa distribuição é observada em fenômenos tão diversos como populações de cidades, números contidos em jornais, valores de contas bancárias, comprimentos de rios, entre outros (Benford, 1938; Gonsalves, 2020).

Embora a lei de Newcomb-Benford tenha se mostrado eficaz em várias áreas, aplicações equivocadas e controversas têm conquistado grande repercussão nas mídias. Um caso emblemático refere-se a uma tentativa de imputar fraudes aos resultados das eleições presidenciais brasileiras de 2014, que foi utilizado para desacreditar o processo eleitoral de 2018 e também o de 2022. Um vídeo produzido em 2015, de poucos minutos, propagou-se nas redes sociais de apoiadores de um candidato marcado pelas constantes declarações contrárias às urnas eletrônicas, em épocas próximas aos pleitos de 2018 e 2022. A tática de desacreditar as elei-

ções tem sido usada recorrentemente por candidatos de extrema-direita em diversos países do mundo nos últimos anos, como Peru, Alemanha, Estados Unidos da América, Bolívia e Israel, que tiveram resultados de pleitos presidenciais questionados (Giovanaz, 2021).

A justificativa da probabilidade de fraude nas eleições de 2014 no Brasil, dada por um suposto especialista que posteriormente se mostrou associado ideologicamente ao espectro político do próprio candidato que alegava fraude no processo eleitoral, envolvia argumentos estatísticos e o uso da lei de Newcomb-Benford não explicitados. Diversas matérias veiculadas na imprensa brasileira apontaram a falsidade das alegações do vídeo (Mota, 2018; UOL, 2018). A ampla divulgação do vídeo de caráter científico fraudulento levou a maior instância da Justiça Eleitoral brasileira, responsável por organizar o sistema eleitoral do Brasil, a se manifestar oficialmente em seu sítio: “a simples análise por seção de votação ou por zonas eleitorais, que possuem poucas urnas, não é eficaz, pois o universo amostral é muito pequeno” (Tribunal Superior Eleitoral, 2022).

Sabemos como resultados matemáticos e estatísticos são parte da estrutura de argumentações de diversas naturezas. Ver a Matemática como um sistema infalível, livre de influências e interesses, onde se valoriza soluções matematizadas de modelos em detrimento a outras soluções obtidas por técnicas não matematizadas, limita a capacidade de interpretação da realidade em sua enorme complexidade. Bassanezi (2010) nos alerta para este fato ao considerar a modelagem como processo de criação de modelos, onde estão definidas as estratégias de ação do indivíduo sobre a realidade, carregada de interpretações e subjetividades do próprio modelador.

Diversos pensadores buscam refletir sobre a finalidade da educação matemática, promovendo posturas democráticas, debatendo questões sociais, econômicas, culturais, cidadania e estruturas de poder que, de certo modo, são perpetuadas com o uso da Matemática. Essas atitudes educacionais têm em Skovsmose (1996, 2000, 2001, 2007) um dos seus mais importantes proponentes. Para o autor, no trabalho pedagógico, professor e alunos devem assumir o papel de investigadores da sua realidade social, criando possibilidades múltiplas para a construção do conhecimento. Ele propõe uma exploração matemática na qual o objetivo principal é criar uma contraposição ao que denomina paradigma do exercício, que pressupõe uma organização da aula em que os exercícios matemá-



ticos apontam para uma resposta correta única e definitiva, fazendo com que a matemática não permita diálogo.

Há na obra de Skovsmose uma preocupação com o desenvolvimento da “materacia”, capacidade de interpretar e manejar sinais e códigos e de propor e utilizar modelos para encontrar soluções para os problemas enfrentados diariamente e, também, elaborar abstrações sobre as representações propostas para os sistemas retirados da realidade (D’Ambrosio, 2004). “Materacia não se refere apenas às habilidades matemáticas, mas também à competência de interpretar e agir numa situação social e política estruturada pela matemática.” (Skovsmose, 2000, p. 68).

Borba e Skovsmose (2001) chamam de “ideologia da certeza” a estrutura geral e fundamental de interpretação para questões que transformam a matemática em uma linguagem de poder. Para os autores, os alunos deveriam ser persuadidos contra ideias de que o argumento matemático é superior por sua própria natureza e convencidos de que a matemática é somente uma maneira possível de olhar o fenômeno.

Com essas mesmas preocupações em torno dos objetivos da educação matemática, desenvolvemos uma atividade com a turma de 25 ingressantes do segundo semestre letivo de 2023 do curso de Matemática, da Universidade Federal Rural do Rio de Janeiro, Campus Nova Iguaçu. Ela teve como objetivo propor uma reflexão sobre a aplicabilidade da matemática na tentativa de se estabelecer uma visão definitiva da realidade. Utilizamos a lei de Newcomb-Benford por enxergar um potencial para trazer tais reflexões a respeito dos limites do uso de argumentos matemáticos na tomada de decisões frente a um problema real.

Faremos o relato da nossa experiência, onde introduzimos o funcionamento dessa lei mediante o estudo de logaritmos, unindo conceitos matemáticos e tecnologia (programação na linguagem Python), explicitando situações nas quais a lei de Newcomb-Benford é aplicada com sucesso, mas também mostraremos os limites da aplicabilidade com uma compilação dos dados obtidos na atividade.

## A LEI DE NEWCOMB-BENFORD

A lei de Benford descreve uma distribuição de frequências dos dígitos iniciais diferente daquela esperada em uma distribuição uniforme

(com probabilidade  $\frac{11}{99}$  para cada dígito de 1 a 9).

Segundo a lei, dado um dígito significativo  $a$ , sua frequência de ocorrência se dá pela expressão:

$$P(a) = \log\left(\frac{a+1}{a}\right) = \log\left(1 + \frac{1}{a}\right),$$

para  $a \in \{1,2,3,4,5,6,7,8,9\}$  (Benford, 1938).

Khosravani e Rasinariu (2018) destacam que também existe uma expressão logarítmica para o segundo, terceiro e para o  $n$ -ésimo dígito. A regra geral usando logaritmos é basicamente a mesma utilizadas para o primeiro dígito. Tome  $a_1 a_2$  sendo o primeiro e  $a_2 a_2$  sendo o segundo dígito. Dessa forma, a regra geral logarítmica se dá por:

$$P(a_1 a_2) = \log\left(1 + \frac{1}{a_1 a_2}\right),$$

para  $a_1 \in \{1,2,3,4,5,6,7,8,9\}$  e

$a_2 \in \{0,1,2,3,4,5,6,7,8,9\}$ .

Ademais, é importante ressaltar que, diferentemente do caso específico para o primeiro dígito, quando aplicamos a lei de Benford para o segundo, terceiro, até o dígito na posição  $n$ , podemos utilizar o algarismo 0 como sendo uma das possibilidades para o dígito. Contudo, há de se ressaltar a observação de Newcomb (1881), segundo a qual “in the case of the third figure the probability will be nearly the same for each digit, and for the fourth and following ones the difference will be inappreciable”. Desse modo, embora a lei se mantenha válida, sua aplicação pode se tornar despropositada na análise de dígitos após a segunda posição.

A lei de Benford tem sido demonstrada e justificada através de diversas abordagens teóricas ao longo dos anos. Uma importante contribuição para a validação teórica dessa lei foi feita por Hill (1995), cuja prova parte do pressuposto de que uma distribuição de dígitos é invariante em relação à base se, e somente se, satisfaz à lei de Newcomb-Benford, resultado provado por Pinkham (1961). Desse modo, Hill fornece uma base matemática sólida para a lei de Benford, mostrando que a distribuição de dígitos iniciais é uma consequência natural da invariância de base. Santos *et al.* (2003), por sua vez, destrincha a prova de Hill.

Além da fundamentação teórica, a lei de Benford tem sido confirmada empiricamente em uma ampla variedade de conjuntos de dados reais. Estudos têm mostrado que essa distribuição é observada em fenômenos tão diversos como populações de cidades, números contidos em

jornais, endereços, valores de contas bancárias, comprimentos de rios, constantes matemáticas, entre outros (Benford, 1938; Gonsalves, 2020). O resultado de uma verificação empírica encontra-se na Tabela 1, que sumariza a frequência dos primeiros dígitos significativos dos 20.000 números iniciais da sequência de Fibonacci, bem como estes mesmos números multiplicados por algumas constantes (número áureo, número de Euler, Pi e aceleração da gravidade). Nota-se que, em cada coluna, a frequência observada muito se aproxima da previsão teórica, inclusive quando os números são multiplicados por constantes diferentes de 1 (uma propriedade da invariância de base da lei de Newcomb-Benford).

**Tabela 1.** Frequência teórica e observada de cada dígito nos números da sequência de Fibonacci.

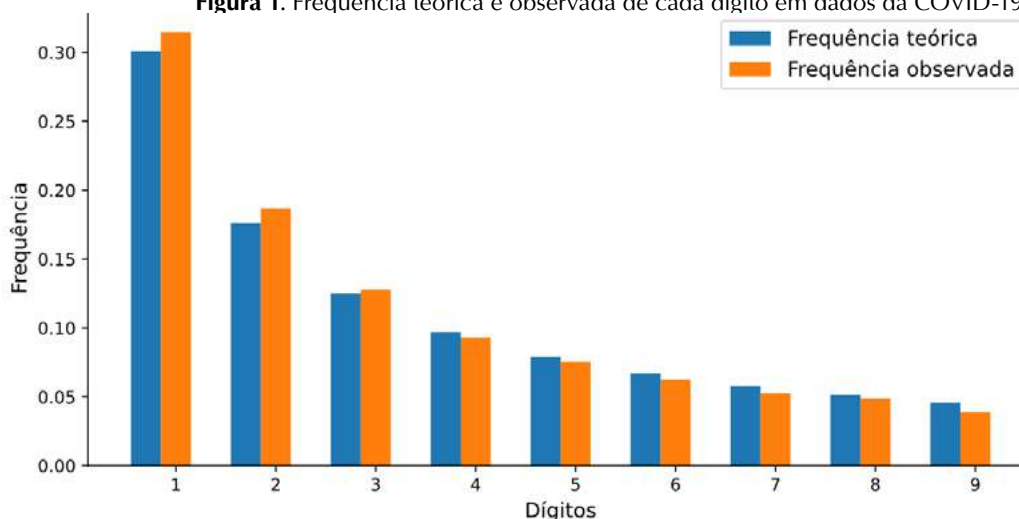
Dígitos	Frequência relativa para cada multiplicador					Previsão teórica
	1	$\varphi$	$e$	$\pi$	$G$	
1	30,10%	30,11%	30,10%	30,10%	30,10%	30,10%
2	17,61%	17,60%	17,62%	17,61%	17,61%	17,61%
3	12,50%	12,50%	12,49%	12,50%	12,50%	12,49%
4	9,69%	9,70%	9,69%	9,69%	9,68%	9,69%
5	7,92%	7,92%	7,93%	7,92%	7,92%	7,92%
6	6,69%	6,70%	6,68%	6,69%	6,70%	6,69%
7	5,80%	5,80%	5,81%	5,80%	5,80%	5,80%
8	5,13%	5,12%	5,12%	5,13%	5,11%	5,12%
9	4,58%	4,58%	4,59%	4,58%	4,59%	4,58%

## METODOLOGIA DA ATIVIDADE

A palestra “A Lei de Newcomb-Benford e Aplicações” foi planejada para ser realizada em um período de 50 minutos. Consistiu em uma apresentação de 30 minutos (com o auxílio de data show e slides) e uma atividade didática de 20 minutos. Na apresentação, abordamos resumidamente a história dos logaritmos; as propriedades operacionais dos logaritmos decimais (característica e mantissa); a disposição das mantissas dos logaritmos decimais em tabelas de logaritmos; as revoluções técnica e tecnológica propiciadas pela invenção; a observação das páginas mais usadas de uma tabela de logaritmos por Simon Newcomb e sua posterior conjectura; o enunciado da lei de Newcomb-Benford, e diversas aplicações em conjuntos de dados variados, bem como casos nos quais a lei não se aplica (por exemplo, na escolha do nono dígito que compõe o Cadastro da Pessoa Física-CPF) e casos nos quais a lei pode levar a conclusões falsas (o caso da suposta fraude nas urnas eletrônicas brasileiras).

Após a apresentação, abrimos um espaço para que os discentes mostrassem suas percepções e dúvidas. Demonstraram muita receptividade, interesse e espanto com as previsões de distribuições dos primeiros dígitos propostas pela lei de Newcomb-Benford em diversos conjuntos de dados (áreas de terrenos da União, populações das cidades brasileiras, números da sequência de Fibonacci, entre outros). Uma das figuras exibidas aos alunos durante a atividade foi a Figura 1, na qual a distribuição dos dígitos nos 294.563 números associados à ocupação de UTIs por pacientes acometidos pela COVID-19 em 2021 (Ministério da Saúde, 2022) é apresentada conforme encontrada empiricamente e prevista pela teoria. Como é possível observar, as distribuições são semelhantes. Porém, objetivando uma análise mais criteriosa que a mera impressão visual, foram calculados e apresentados aos discentes a norma euclidiana do vetor diferença entre os resultados encontrados e os esperados, e os dígitos com o maior e o menor erro relativo.

**Figura 1.** Frequência teórica e observada de cada dígito em dados da COVID-19.



Em seguida, passamos à explicação da atividade. A atividade didática teve como referência uma informação vista no documentário “A Era dos Dados” (2020), no episódio intitulado “Dígitos” (Temporada 1, Episódio 4), exibido através da plataforma de *streaming* Netflix. Uma das entrevistadas, a professora Jennifer Golbeck, da Faculdade de Estudos da Informação na Universidade de Maryland, EUA, afirmou que, utilizando como espaço amostral uma quantidade considerável de modelos na rede social Instagram, ou em qualquer outra, os números de seguidores de um conjunto aleatório de pessoas seguirão a lei de Newcomb-Benford.

Antes de começarmos a atividade, verificamos se a turma tinha acesso à internet, o que era essencial para a realização da atividade. Vimos que a grande maioria da turma possuía *smartphones* com acesso à internet. Aos discentes que não tinham acesso à internet, pedimos que se sentassem juntos aos colegas com acesso à rede. Distribuímos uma ficha contendo uma tabela na qual os estudantes marcariam nas linhas numeradas de 1 a 9, representando os primeiros dígitos significativos dos dados que seriam coletados.

Solicitamos aos alunos e às alunas que acessassem suas contas do Instagram (todos tinham acesso à rede social, por isso a escolha). Passamos às instruções da atividade: cada discente deveria escolher aleatoriamente o perfil de um seguidor seu e anotar o primeiro dígito significativo do número de seguidores do escolhido; em seguida, pedimos que escolhessem aleatoriamente um seguidor do seu seguidor e novamente anotassem o primeiro dígito do número de seguidores; procedendo assim, sucessivamente, até analisar 15 contas. Foram coletados ao todo 282 dados numéricos. Levamos pouco mais de 10 minutos para que todos os discentes anotassem os dados observados, mostrando como a atividade pode ser rapidamente realizada.

Em seguida, coletamos os dados registrados pela turma, obtivemos as frequências dos primeiros dígitos significativos e passamos à etapa final da atividade, que era a análise dos dados coletados. Realizamos uma implementação na linguagem de programação Python. O código, que analisamos abaixo, já estava pré-programado, faltando somente o preenchimento das frequências observadas de cada dígito significativo dentro do conjunto de dados. Antes do preenchimento, fizemos uma explicação de cada linha do código para a turma.

**Código 1.** Código em Python para tratamento de dados e comparação com a previsto pela lei de Benford.

```
1 import matplotlib.pyplot as plt #biblioteca para plotagem de gráficos
2 from math import * #biblioteca para operações matemáticas
3 import numpy as np #biblioteca para operações com vetores
4
5 b = [81, 44, 35, 28, 25, 25, 11, 20, 13] #quantidade de cada dígito
6
7 observed = np.array(b)/sum(b) #proporção observada em cada dígito
8 expected = np.array([log((i + 1)/i, 10) for i in range(1,10)]) #benford
9
10 digitos = ['1', '2', '3', '4', '5', '6', '7', '8', '9'] #lista de dígitos
11 largura = 0.3 #parâmetro para largura das barras no gráfico
12 posicoes = np.arange(9) #parâmetro para posições das barras no gráfico
13
14 plt.bar(posicoes, expected, largura, label='Frequência teórica')
15 plt.bar(posicoes+largura, observed, largura, label='Frequência observada')
16 plt.xticks(posicoes+largura, digitos)
17 plt.xlabel('Dígitos') #legenda do eixo x
18 plt.ylabel('Frequência') #legenda do eixo y
19 plt.legend()
20 plt.show()
```

Analisando o Código 1, temos a importação de bibliotecas pertinentes nas linhas 1-3. Elas são responsáveis, respectivamente, pela representação dos dados na forma de gráficos de barras, por operações matemáticas como o cálculo do logaritmo decimal de um número real e por operações envolvendo vetores. Na linha 5, é definido um vetor de 9 posições associadas aos 9 dígitos a serem analisados: cada posição receberá a quantidade de vezes que determinado dígito foi o dígito significativo no número de seguidores de uma conta no Instagram.

Na linha 7, o vetor com a frequência de cada dígito na primeira posição é efetivamente calculado com base nos resultados obtidos da amostra de dados analisada. Como indicado, a operação

é feita a partir da razão entre o número de vezes em que cada dígito ocupou a primeira posição e o número total de dígitos. Na sequência, a distribuição conforme a lei de Newcomb-Benford é calculada para cada dígito de 1 a 9 e é armazenada em outro vetor de 9 posições, que posteriormente será comparado ao primeiro.

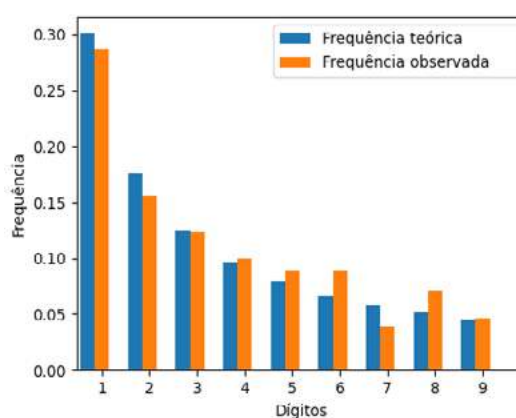
Das linhas 10 a 20, a representação gráfica dos dados encontrados é construída e exibida a partir de alguns parâmetros necessários para o uso da biblioteca importada na linha 1. Entre eles estão a largura das barras, as legendas associadas a cada uma, e outros. Essa sequência de comandos permite plotar o gráfico que associa a cada dígito duas barras: uma representando a frequência observada desse dígito na amostra de dados (na primeira posição) e outra representando a frequência teórica prevista pela lei de Newcomb-Benford.

Como pode ser percebido na Tabela 2 e na Figura 2, as frequências relativas dos primeiros dígitos foram próximas às previsões da lei de Newcomb-Benford, mas não tiveram o comportamento decrescente de frequência relativa previsto pelo modelo. Iniciamos uma discussão com os alunos sobre as incompatibilidades encontradas. O primeiro questionamento, feito pelos próprios alunos, foi sobre a obrigatoriedade do conjunto de todos os números de seguidores de uma rede social seguir a distribuição prevista para a lei, e, se segue, por qual motivo não encontramos a distribuição. A amostra de dados coletados poderia ser considerada pequena e, se continuássemos a coleta, a tendência era a convergência entre as previsões teóricas e as frequências relativas observadas, o que demonstra um raciocínio probabilístico baseado na chamada Lei Fraca dos Grandes Números (Ross, 2010). Segundo Kruger e Yavadalli (2017), não há consenso na literatura para que o tamanho de uma amostra seja adequado, mas que seja de pelo menos 50 a 100 para a lei ser observada (caso seja seguida), e que um tamanho de amostra de 500 ou mais pode ser preferido para uma análise adequada.

**Tabela 2.** Frequência relativa observada na atividade e a previsão teórica segundo a lei de Newcomb-Benford.

Dígito	Frequência relativa	Previsão teórica
1	28,72%	30,10%
2	15,60%	17,60%
3	12,41%	12,49%
4	9,92%	9,69%
5	8,86%	7,91%
6	8,86%	6,69%
7	3,9%	5,79%
8	7,09%	5,11%
9	4,60%	4,57%

**Figura 2.** Gráfico obtido a partir do Código 1, contendo a frequência relativa dos primeiros dígitos significativos obtidos na atividade.



Outros apontamentos feitos pela turma merecem menção: algumas contas de seguidores utilizadas pelos discentes eram privadas ou possuíam outra ferramenta que não permitia ver os seguidores, o que acarretaria uma ideia irreal do número de seguidores; outros levantaram a hipótese de contas terem um número falso de genuínos seguidores, tendo em vista que a rede social analisada permite a compra de pacotes de seguidores.

Apesar de a atividade ter sido aplicada a um grupo de alunos ingressantes em um curso de ensino superior, não existem limitações para a sua aplicação em uma turma de Ensino Médio, pelo contrário, é interessante a sua execução como uma atividade introdutória ao conteúdo de Logaritmos ou Funções Logarítmicas, uma vez que esta está relacionada às seguintes habilidades da Base Nacional Comum Curricular (BNCC): EM13MAT305, sobre “resolver e elaborar problemas com funções logarítmicas nos quais é necessário compreender e interpretar a variação das grandezas envolvidas” (Brasil, 2018, p. 528); e EM13MAT403, sobre “comparar e analisar as representações, em plano cartesiano, das funções exponencial e logarítmica para identificar as características fundamentais (domínio, imagem, crescimento) de cada uma, com ou sem apoio de tecnologias digitais, estabelecendo relações entre elas” (Brasil, 2018, p. 531).

Como os logaritmos são utilizados para simplificar expressões multiplicativas, transformando-as em expressões aditivas, a atividade facilitaria a compreensão dos estudantes na análise de grandes variações de dados. Ao analisar a distribuição dos primeiros (ou  $n$  primeiros) dígitos significativos, os alunos podem perceber a aplicação prática dos logaritmos na Lei, seja na compreensão de escalas ou em determinado uso arbitrário envolvendo variações exponenciais, como já exemplificadas. Uma vez que a Lei também é abrangida pelo conceito de probabilidade, sua análise estatística da distribuição dos dígitos em diferentes posições permite aos alunos a obtenção de uma perspectiva tangível sobre como os conteúdos de probabilidade e estatística podem ser aplicados além de problemas em urnas. Isto não apenas fortalece o interesse e a compreensão dos estudantes pelo assunto, mas também é um incentivo ao reconhecimento de padrões presentes em um conjunto de dados do mundo real, o que, de certa forma, vem a promover o estímulo do aluno à Lei ao ver que ela, de fato, lida com problemas concretos e recorrentes.

No fechamento da atividade, mostramos como o levantamento dessas questões torna evidente o fato de como uma tomada de deci-

são com o auxílio de ferramentas matemáticas poderosas tem suas limitações de uso, podendo nos levar a diversas conclusões. Finalizamos a atividade sem fazer apontamentos para as verdadeiras causas da (in)compatibilidade, como também defendem Alrø e Skovsmose (2002), ao realçar o elemento básico na realização de uma investigação: as atividades devem ser abertas, de tal forma que os caminhos a seguir e as possíveis conclusões a que se pode chegar não sejam conhecidos de antemão.

## CONCLUSÕES

Este trabalho busca combinar a análise empírica da lei de Newcomb-Benford em conjuntos de dados diversos com a proposta de uma atividade didática que promova a compreensão e o interesse dos alunos pelo tema, introduzindo o ensino de logaritmos ou probabilidade e fornecendo uma visão de sua aplicação tanto em um contexto de pesquisa quanto de educação matemática.

Borba e Skovsmose (2001) afirmam que “por meio de modelos matemáticos, também nos tornamos capazes de ‘projetar’ uma parte do que se torna realidade. Tomamos decisões baseados em modelos matemáticos e, dessa forma, a matemática molda a realidade”. Devemos desafiar a ideologia da certeza e esse poder formatador da matemática por meio de um currículo baseado na incerteza, com trabalhos em situações abertas nas quais surgem ambiguidades, evidenciando os limites da Matemática na solução de um problema real, questionando os interesses envolvidos na escolha de modelos matemáticos e não aceitando a neutralidade da matemática e suas soluções infalíveis. É assim que enxergamos a proposta de Borba e Skovsmose (2001) para propiciar uma formação crítica em Matemática. O fato de nosso experimento ter resultado em uma distribuição da frequência relativa distinta da prevista, podendo parecer inconsistente à primeira vista, realça exatamente o que queríamos discutir: os limites da aplicabilidade da Matemática.

Consideramos que a atividade foi atrativa, despertando interesses diversos na turma. Os alunos se interessaram pela coleta de dados das aplicações trabalhadas, pediram detalhes da programação do Código 1, solicitaram indicação do documentário, referências a livros, entre outros.

Ressaltamos que esta atividade pode ser desenvolvida e adaptada para turmas do Ensino Médio, sendo um tema motivador para se trabalhar História da Matemática, logaritmos e probabilidades.

## REFERÊNCIAS

- ALRØ, H.; SKOVSMOSE, O.** *Dialogue and learning in Mathematics Education: intention, reflection, critique*. Kluwer Academic Publishers, 2002<sup>o</sup> ed., 2006. ISBN: 978-1402009983
- BASSANEZI, R.** *Ensino-aprendizagem com modelagem matemática: uma nova estratégia*. Contexto, 4<sup>o</sup> ed., 2002. ISBN: 978-8572442077
- BENFORD, F.** The law of anomalous numbers. *Proceedings of the American Philosophical Society*, v.78, n.4, 551-572, 1938
- BORBA, M.; SKOVSMOSE, O.** A Ideologia da Certeza em Educação Matemática. In: SKOVSMOSE, O. *Educação Matemática Crítica – A Questão da Democracia*. Campinas: Papius, 2001.
- BRASIL.** *Ministério da Educação. Base Nacional Comum Curricular*. Brasília: MEC, 2018.
- D'AMBROSIO, U.** A relevância do projeto indicador nacional de alfabetismo funcional – INAF como critério de avaliação da qualidade do ensino de matemática. In: FONSECA, M. C. F. R. (Org.). *Letramento no Brasil: habilidades matemáticas*. São Paulo, Global, 2004.
- DÍGITOS** (temporada 1, ep. 4). *A Era dos Dados* [Série documental]. Direção: Nick Brigden. Produção: Christopher Collins, Lydia Tenaglia, Erik Osterholm, Latif Nasser, David Mettler. Estados Unidos: Zero Point Zero Production Inc., 2020. (46 min.), son., color.
- GIOVANAZ, D.** *Veja o que aconteceu em 7 países onde a direita alegou fraude eleitoral sem provas*. Brasil de Fato, 2021. Disponível em: <<https://www.brasildefato.com.br/2021/08/15/veja-o-que-aconteceu-em-7-paises-onde-a-direita-alegou-fraude-eleitoral-sem-provas>>. Acesso em: 29 set. 2023
- GONSALVES, R.** Benford's Law – a simple explanation. *Towards Data Science*, 2020. Disponível em: <<https://towardsdatascience.com/benford-s-law-a-simple-explanation-341e17abbe75>>. Acesso em: 13 jul. 2023
- KHOSRAVANI, A.; RASINARIU, C.** Emergence of Benford's Law in Music. *Journal of Mathematical Sciences: Advances and Applications*, v.54, 11-24, 2018. DOI: 10.18642/jmsaa\_7100122017
- HILL, T.** Base-invariance implies Benford's Law. *Proceedings of the American Mathematical Society*, v.123, n.3, 1995
- KRUGER, P.; YAVADALLI, V.** The power of one: Benford's law. *South African Journal of Industrial Engineering*, v.28(2), 1-13, 2017. DOI: 10.7166/28-2-1753
- MINISTÉRIO DA SAÚDE.** *Registro de Ocupação Hospitalar COVID-19*. Disponível em: <<https://opendatasus.saude.gov.br/dataset/registro-de-ocupacao-hospitalar-covid-19>>. Acesso em: 12 jul. 2023.
- MOTA, M. et al.** Mensagens com conteúdo #FAKE sobre fraude em urnas eletrônicas se espalham nas redes. *O Globo*, 2018. Disponível em: <<https://g1.globo.com/fato-ou-fake/noticia/2018/10/06/mensagens-com-conteudo-fake-sobre-fraude-em-urnas-eletronicas-se-espalham-nas-redes.ghtml>>. Acesso em: 08 set. 2023.
- NEWCOMB, S.** Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, v.4, n.1, 39-40, 1881
- PINKHAM, R.** On the distribution of first significant digits. *The Annals of Mathematical Statistics*, v.32, n.4, 1223-1230, 1961
- ROSS, S.** *Probabilidade : um curso moderno com aplicações*. Bookman, 8<sup>o</sup> ed., 2010. ISBN: 978-8577806218
- SANTOS, J.; TENÓRIO, J.; SILVA, L.** Uma aplicação da teoria das probabilidades na Contabilometria: a Lei de Newcomb-Benford como mediadora para análise de dados no campo da Auditoria Contábil. *Contabilidade, Gestão e Governança*, v.6, n.1, 35-54, 2003
- SKOVSMOSE, O.** Cenários para investigação. *BOLEMA – Boletim de Educação Matemática*, Rio Claro, n.14, 66-91, 2000.
- SKOVSMOSE, O.** Critical mathematics education: some philosophical remarks. In: *International Congress on Mathematics Education. Selected lectures*. Sevilha: S. A. E. M. 1996, 413-425.
- SKOVSMOSE, O.** *Educação Crítica: Incerteza, Matemática, Responsabilidade*. Tradução de Maria Aparecida Viggiani Bicudo. Cortez Editora, 1<sup>o</sup> ed., 2007. ISBN: 978-8524912948
- SKOVSMOSE, O.** *Educação Matemática Crítica – A Questão da Democracia*. Papius, 6<sup>o</sup> ed., 2001. ISBN: 978-8530806415
- TRIBUNAL SUPERIOR ELEITORAL.** *Fato ou boato: é mentira que Lei de Benford provou fraude nas eleições brasileiras*. Brasília, 2022. Disponível em: <<https://www.tse.jus.br/comunicacao/noticias/2022/Maio/fato-ou-boato-e-mentira-que-lei-de-benford-provou-fraude-nas-eleicoes-brasileiras>>. Acesso em: 27 ago. 2023
- UOL.** *Vídeo que põe suspeita sobre urnas cita lei matemática que não prova fraude*. São Paulo, 2018. Disponível em: <<https://noticias.uol.com.br/comprova/ultimas-noticias/2018/10/10/video-com-suspeitas-sobre-urnas-cita-lei-matematica-que-nao-prova-fraude.htm>>. Acesso em: 21 set. 2023

## DADOS DOS AUTORES

**Victor Augusto do Carmo Duarte.** Bacharelado em Matemática Aplicada e Computacional pela Universidade Federal Rural do Rio de Janeiro (UFRRJ), e-mail: victoracd2002@ufrj.br.

**Lavinia Beatriz Duarte da Silva.** Licencianda em Matemática pela Universidade Federal Rural do Rio de Janeiro (UFRRJ), e-mail: laviniabduarte@ufrj.br.

**Marcelo Ferreira Farias.** Doutor em Matemática pela Universidade Federal do Rio de Janeiro (UFRJ), docente da Universidade Federal Rural do Rio de Janeiro (UFRRJ), e-mail: mfarias@ufrj.br.

## PROPRIEDADES ESPECTRAIS DE GRAFOS BENZENOÍDES

### *SPECTRAL PROPERTIES OF BENZENOID GRAPHS*

Allan José Amaral Ribeiro

Francisca Andrea Macedo França

André Ebling Brondani

**RESUMO:** A teoria química dos grafos é uma área moderna e importante da matemática discreta e da química. Grafos benzenóides são redes obtidas pela disposição de hexágonos regulares congruentes no plano, de modo que dois hexágonos sejam disjuntos ou possuam uma aresta comum. Neste trabalho investigamos uma classe de grafos benzenóides, exibimos o seu espectro e expressamos a quantidade de seus autovalores principais em termos de sua estrutura.

**Palavras-chave:** Grafos benzenóides. Matriz de adjacência. Espectro. Espectro principal.

**ABSTRACT:** The Chemical Theory is a modern and important area of discrete mathematics and chemistry. Benzenoid graphs are the networks obtained by arranging congruent regular hexagons in the plane, so that two hexagons are either disjoint or possess a common edge. In this work we investigate a class of benzenoid graphs, display their spectrum and express the amount of their principal eigenvalues in terms of their structure.

**Keywords:** Benzenoid graphs. Adjacency matrix. Spectrum. Main spectrum..

## INTRODUÇÃO

A química matemática é um ramo da química teórica que discute a estrutura molecular por métodos matemáticos sem necessariamente se referir à mecânica quântica. A teoria química dos grafos é um ramo da química matemática que combina matemática, química e teoria dos grafos. Sendo que a teoria dos grafos é utilizada para modelar matematicamente moléculas, a fim de obter informações sobre as propriedades físicas desses compostos químicos. Para mais detalhes, sugerimos (DAS; GUTMAN, 2004) e (RANJINI; LOKESHA; CANGUL, 2011).

Um grafo molecular é um grafo em que os vértices correspondem aos átomos e as arestas às ligações químicas. Os índices topológicos e espectrais dos grafos são definidos e usados em muitas áreas para estudar propriedades de diferentes objetos, como átomos e moléculas. Tais índices são definidos como invariantes de grafos que correspondem e refletem diversas propriedades físicas, químicas, farmacológicas, farmacêuticas e biológicas das espécies químicas subjacentes, veja (GUTMAN, 2017).

A contagem de cadeias em grafos moleculares tem uma longa tradição na química teórica. O fato de que o número de cadeias fechadas é igual ao  $k$ -ésimo momento espectral do respectivo grafo molecular é importante na teoria orbital molecular de Hückel (GUTMAN; FURTULA, 2017). Contagens de cadeias fechadas iniciando em um determinado átomo são descritas em vários textos químicos; elas são idênticas às conectividades obtidas pelo Algoritmo de Morgan (MORGAN, 1965) e usadas para discriminar e enumerar os átomos em uma molécula orgânica. Além disso, têm um papel significativo em algoritmos para a percepção de simetrias topológicas (GUTMAN, 2017), que podem estar relacionadas a certos índices topológicos ou espectrais.

Neste trabalho, motivados pela aplicabilidade da contagem de cadeias em grafos moleculares, utilizamos técnicas espectrais para determinar o espectro e a quantidade de autovalores que são necessários para a contagem do número de cadeias em determinados hidrocarbonetos benzenóides. Tais autovalores foram denominados autovalores principais em (CVETKOVIC, 1978). Além disso, neste mesmo artigo o autor prova que um grafo possui exatamente 1 autovalor principal se e somente se o grafo é regular e propõe a investigação dos grafos com exatamente autovalores principais.

O trabalho está estruturado de modo que na próxima seção apresentamos algumas definições e resultados necessários para a compreensão do texto. Nas demais seções apresentamos as contribuições deste trabalho que envolvem a determinação do espectro de uma classe de grafos benzenóides e a solução do problema proposto por Cvetkovic (1978) para esta mesma classe de grafos. Por fim, apresentamos uma breve conclusão do trabalho.

## PRELIMINARES

Nesta seção, além de fixarmos algumas notações, são revistos alguns resultados e conceitos básicos sobre teoria de grafos e teoria de matrizes que são utilizados ao longo do texto. Para mais detalhes sugerimos (DIESTEL, 2000) e (HORN; JOHNSON, 2013).

O espaço vetorial formado pelas matrizes de ordem  $m \times n$  e com coeficientes reais é denotado por  $M_{m \times n}(\mathbb{R})$  ou, simplesmente,  $M_n(\mathbb{R})$  quando  $m = n$ . Dado uma matriz  $A$  denotamos a matriz transposta de  $A$  por  $A^T$ , sua matriz inversa, quando esta existir, por  $A^{-1}$ , e por  $A[[i], [j]]$  a submatriz obtida de  $A$  excluindo-se a linha  $i$  e a coluna  $j$ . Além disso, a matriz identidade e a matriz nula em  $M_n(\mathbb{R})$  são denotadas por  $I_n$  e por  $O_n$ , respectivamente.

Seja  $G = (V, E)$  um grafo simples com  $V = \{v_1, v_2, \dots, v_n\}$ . A cardinalidade de  $V$  é denominada *ordem* de  $G$ , a cardinalidade de  $E$  é o *tamanho* de  $G$  e dois vértices  $v_i$  e  $v_j$  em  $V$  são adjacentes quando  $\{v_i, v_j\} \in E$ . A *vizinhança* de um vértice  $v_i \in V$  é o conjunto  $N_G(v_i) = \{v_j \in V; \{v_i, v_j\} \in E\}$  e o *grau* do vértice  $v_i$ , denotado por  $d(v_i)$ , é a cardinalidade de  $N_G(v_i)$ . Denotamos os graus mínimo e máximo de  $G$  por  $\delta(G)$  e  $\Delta(G)$ , respectivamente. Uma cadeia de comprimento  $k$  em  $G$  é uma sequência de vértices, digamos  $(v_1, v_2, \dots, v_k, v_{k+1})$ , em que  $\{v_i, v_{i+1}\} \in E$ , onde  $0 \leq i \leq k$ . A *matriz de adjacência* de  $G$ , tem as entradas dadas por

$$a_{ij} = \begin{cases} 1, & \text{se } \{v_i, v_j\} \in E; \\ 0, & \text{caso contrário} \end{cases}$$

O *polinômio característico* de  $G$  é definido por  $\phi(G, x) = \det(xI - A)$  e suas raízes são denominadas *autovalores* de  $G$ . O multi-conjunto constituído pelos autovalores de  $G$  é denominado *espectro* de  $G$  e é denotado por  $\sigma(G)$ .



A seguir listamos alguns resultados importantes para as próximas seções que envolvem o cálculo de determinante para matrizes em blocos e a determinação de autovalores de certas matrizes tridiagonais. Finalizamos esta seção descrevendo a classe dos grafos benzenóides e definindo uma subclasse desta, que é o objeto de estudo desta pesquisa.

**Proposição 1.** (HORN; JOHNSON, 2013). Se  $A$  e  $D$  são matrizes quadradas, então

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{cases} \det(A) \det(D - CA^{-1}B), & \text{quando } A^{-1} \text{ existe;} \\ \det(D) \det(A - BD^{-1}C), & \text{quando } D^{-1} \text{ existe.} \end{cases}$$

**Proposição 2.** (NATH; PAUL, 2014). Se  $A$  e  $B$  são matrizes quadradas e  $P = \begin{bmatrix} A & B \\ B & A \end{bmatrix}$ , então  $\sigma(P) = \sigma(A + B) \cup \sigma(A - B)$ .

A Proposição 3, devido a Losonczi (1992), exhibe o espectro de um tipo especial de matriz tridiagonal.

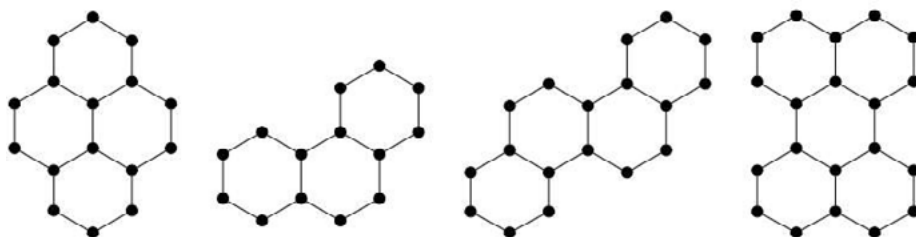
**Proposição 3.** Considere a matriz tridiagonal de ordem  $n$ ,

$$\text{Se } \alpha = \beta = \sqrt{ac} \neq 0, \text{ então } \sigma(T_n) = \left\{ b + 2\sqrt{ac} \cos\left(\frac{j\pi}{n}\right); 1 \leq j \leq n \right\}.$$

$$\text{Se } \alpha = \beta = 0, \text{ então } \sigma(T_n) = \left\{ b + 2c \sqrt{\frac{a}{c}} \cos\left(\frac{j\pi}{n+1}\right); 1 \leq j \leq n \right\}.$$

De acordo com Gutman (2017), grafos benzenóides são redes obtidas pelo arranjo de hexágonos regulares congruentes no plano de modo que dois hexágonos são disjuntos ou possuem uma aresta em comum. A Figura 1 ilustra alguns exemplos de grafos benzenóides.

Figura 1: Grafos benzenóides.



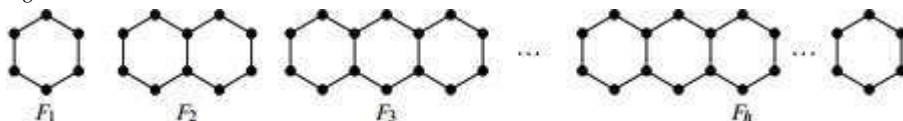
Fonte: Produção própria dos autores.

Considere a subclasse de grafos benzenóides

$$F = \{F_h \mid (h = \text{número de hexágonos}) \wedge (h \geq 1)\},$$

construída da seguinte forma:

Figura 2: Faixas benzenóides.



Fonte: Produção própria dos autores.

Os elementos de  $F$  são denominados *faixas benzenóides*. Além disso, dado  $h > 0$ , o grafo  $F_h \in F$  é bipartido, tem ordem  $n = 4h + 2$  e tamanho  $m = 5h + 1$ . Para  $h > 1$ ,  $F_h$  tem grau mínimo 2 e grau máximo 3.

### O espectro dos grafos da classe $F$

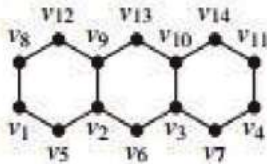
Nesta seção, apresentamos o espectro dos grafos benzenóides pertencentes a classe  $F$ . Dado  $h > 0$ , existe uma rotulação conveniente dos vértices de  $F_h \in F$  de modo que a sua matriz de adjacência seja dada por

$$A(F_h) = \begin{bmatrix} 0_{h+1} & C_{(h+1) \times h} & I_{h+1} & 0_{(h+1) \times h} \\ C_{h \times (h+1)}^T & 0_h & 0_{h \times (h+1)} & 0_h \\ I_{h+1} & 0_{(h+1) \times h} & 0_{h+1} & C_{(h+1) \times h} \\ 0_{h \times (h+1)} & 0_h & C_{h \times (h+1)}^T & 0_h \end{bmatrix}, \quad (1)$$

onde  $C = [c_{ij}] \in M_{(h+1) \times h}(\mathbb{R})$  é definida por  $c_{ij} = \begin{cases} 1, & \text{se } i \in \{j, j+1\}; \\ 0, & \text{caso contrário.} \end{cases}$

A Figura 3 exibe o grafo  $F_3$  rotulado de modo que a sua matriz de adjacência seja da forma dada em (1).

Figura 3: Grafo com a rotulação adotada.



Fonte: Produção própria dos autores.

**Lema 4.** Seja  $C$  a submatriz de  $A(F_h)$ , definida em (1). Se  $E_{h+1} = CC^T \in M_{h+1}(\mathbb{R})$  e  $Z_h = C^T C \in M_h(\mathbb{R})$ , então

$$\sigma(E_{h+1}) = \left\{ 2 + 2 \cos\left(\frac{j\pi}{h+1}\right); 1 \leq j \leq h+1 \right\} \text{ e } \sigma(Z_h) = \left\{ 2 + 2 \cos\left(\frac{j\pi}{h+1}\right); 1 \leq j \leq h \right\}$$

Demonstração. De fato, basta notar que as matrizes  $E_{h+1}$  e  $Z_h$  tem a forma da matriz dada na Proposição 3, com  $a = c = 1$  e  $b = 2$ . Para  $E_{h+1}$  tem-se  $\alpha = \beta = 1$  e para  $Z_h$ ,  $\alpha = \beta = 0$ .

**Teorema 5.** Para cada inteiro positivo  $h$ , o espectro do grafo  $F_h \in F$  é dado por

$$\sigma(F_h) = \left\{ -1, 1, \frac{1}{2} \left( 1 \pm \sqrt{9 + 8 \cos\left(\frac{j\pi}{h+1}\right)} \right), \frac{1}{2} \left( -1 \pm \sqrt{9 + 8 \cos\left(\frac{j\pi}{h+1}\right)} \right); 1 \leq j \leq h \right\}.$$

Demonstração. Inicialmente, note que a matriz  $A(F_h)$ , dada em (1), tem a forma

$$A(F_h) = \begin{bmatrix} M & N \\ N & M \end{bmatrix}, \text{ onde } M = \begin{bmatrix} 0_{h+1} & C_{(h+1) \times h} \\ C_{h \times (h+1)}^T & 0_h \end{bmatrix} \text{ e } N = \begin{bmatrix} I_{h+1} & 0_{(h+1) \times h} \\ 0_{h \times (h+1)} & 0_h \end{bmatrix}$$

Da Proposição 2, o espectro de  $A(F_h)$  e a união dos espectros das matrizes

$$R_1 = \begin{bmatrix} I_{h+1} & C_{(h+1) \times h} \\ C_{h \times (h+1)}^T & 0_h \end{bmatrix} \text{ e } R_2 = \begin{bmatrix} -I_{h+1} & C_{(h+1) \times h} \\ C_{h \times (h+1)}^T & 0_h \end{bmatrix}$$

Sejam  $E_{h+1} = CC^T$  e  $Z_h = C^T C$ . Do Lema 4,  $\det(Z_h) \neq 0$  e da Proposição 1,  $\det(R_1) = (-1) \det(Z_h)$  e  $\det(R_2) = -\det(Z_h)$ . Portanto,  $x = 0$  não é autovalor para  $R_1$  e nem para  $R_2$ . Aplicando a Proposição 1 e visto que o determinante de uma matriz simétrica corresponde ao produto de seus autovalores, para  $x \neq 0$ , obtemos:

$$\begin{aligned} \phi(R_1, x) &= \det \begin{bmatrix} (x-1)I_{h+1} & -C_{(h+1) \times h} \\ -C_{h \times (h+1)}^T & xI_h \end{bmatrix} = x^h \det \left[ (x-1)I_{h+1} - \frac{1}{x} E_{h+1} \right] \\ &= \frac{1}{x} \det \left[ x(x-1)I_{h+1} - E_{h+1} \right] \quad (2) \\ &= \frac{1}{x} \prod_{j=1}^{h+1} (x^2 - x - \mu_j), \end{aligned}$$

sendo que do Lema 4,  $\mu_j = 2 + 2 \cos\left(\frac{j\pi}{h+1}\right)$ ,  $1 \leq j \leq h+1$ , é um autovalor de  $E_{h+1}$ . Como  $x \notin \mathbb{Q}$ , da equação (2) obtemos

$$\sigma(R_1) = \left\{ 1, \frac{1}{2} \left( 1 \pm \sqrt{9 + 8 \cos\left(\frac{j\pi}{h+1}\right)} \right); 1 \leq j \leq h \right\}$$

Analogamente, mostra-se que

$$\sigma(R_2) = \left\{ -1, \frac{1}{2} \left( -1 \pm \sqrt{9 + 8 \cos\left(\frac{j\pi}{h+1}\right)} \right); 1 \leq j \leq h \right\}$$

e o resultado segue da Proposição 2.

### A cardinalidade do espectro principal dos grafos da classe F

Denotemos por  $\mathbf{1}_n$  o vetor-coluna cujas  $n$  coordenadas são iguais a 1 e seja  $\{x_1, x_2, \dots, x_n\}$  uma base ortonormal de autovetores da matriz de adjacência,  $A$  de um grafo  $G$ . O Teorema da Decomposição Espectral estabelece que  $A = \sum_{i=1}^n \lambda_i x_i x_i^T$ , onde  $\lambda_i$  é um autovalor de  $A$  correspondente ao autovetor  $x_i$ . Como a  $(i, j)$ -ésima entrada da matriz  $A^k$  corresponde ao número de cadeias de comprimento  $k$  que iniciam no vértice  $v_i$  e terminam no vértice  $v_j$ , concluímos que

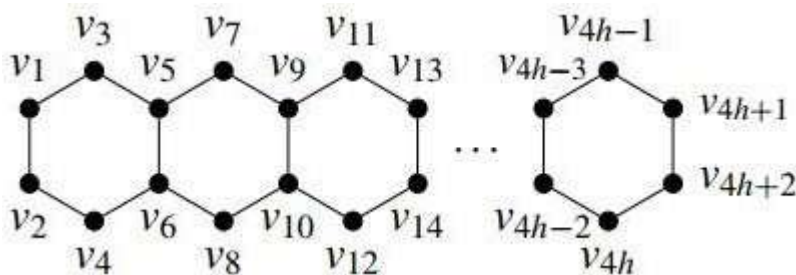
$$\mathbf{1}_n^T A^k \mathbf{1}_n = \sum_{i=1}^n \lambda_i^k (x_i^T \mathbf{1}_n)^T (x_i^T \mathbf{1}_n)$$

equivale ao número de cadeias de comprimento  $k$  em  $G$ . Os distintos autovalores que contribuem para a contagem do número de cadeias de comprimento  $k$  em um grafo são aqueles cujo autoespaço associado não é ortogonal ao vetor  $\mathbf{1}_n$ . Tais autovalores são denominados **autovalores principais** do grafo.

Seja  $G = (V, E)$  um grafo. Uma partição equilibrada  $\pi = \{V_1, V_2, \dots, V_k\}$  de  $V$  é uma partição de  $V$  com a seguinte propriedade: dadas duas células,  $V_i$  e  $V_j$  de  $\pi$ , existe uma constante  $q_{ij}$ , tal que cada vértice  $v \in V_i$  tem exatamente  $q_{ij}$  vizinhos em  $V_j$ . A matriz  $Q = [q_{ij}] \in M_k(\mathbb{R})$  é denominada matriz quociente de  $G$  com respeito a partição  $\pi$ .

Dado  $h \geq 1$ , considere a rotulação do grafo  $F_h$  dada na Figura 4.

Figura 4: Rotulação de  $F_h$



Fonte: Produção própria dos autores.

Sejam  $F_h = (V, E) \in \mathcal{F}$ ,  $V = \{v_1, v_2, \dots, v_{4h+2}\}$  e considere os subconjuntos de  $V$ , definidos por  $U_i = \{v_{2i-1}, v_{2i}\}$  e  $W_j = \{v_{4h-2j-1}, v_{4h-2j}\}$ , onde  $1 \leq j \leq h+1$  e  $0 \leq j \leq h-2$ . Definimos a partição  $\pi(F_h) = \{V_1, V_2, \dots, V_{h+1}\}$ , de acordo com a rotulação dada na Figura 4, em que

$$V_i = \begin{cases} U_1 \cup \{v_{4h+1}, v_{4h+2}\}, & \text{se } i = 1; \\ U_2 \cup W_0, & \text{se } i = 2; \\ U_i \cup W_{i-2}, & \text{se } 3 \leq i \leq h; \\ U_{h+1}, & \text{se } i = h+1. \end{cases}$$

Para posterior referência, enunciamos o lema a seguir que decorre diretamente da definição e cuja justificativa é imediata.

**Lema 6.** Para  $h \geq 2$ , a partição  $\pi(F_h)$  é uma partição equilibrada de  $F_h$ .

A matriz quociente de  $\pi(F_h)$ ,  $Q_h \in M_{h+1}(\mathbb{R})$ , é definida por  $Q_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$  e para  $h \geq 3$  é dada, recursivamente, por:

$$Q_3 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix} \text{ e } Q_h = \begin{bmatrix} Q_{h-1}[\{h\}, \{h\}] & P \\ P^T & S \end{bmatrix}$$

onde  $Q_{h-1}[\{h\}, \{h\}] \in M_{h-1}(\mathbb{R})$  é obtida de  $Q_{h-1}$  excluindo-se a linha  $h$  e a coluna  $h$ ;

$$S = \begin{bmatrix} a & 1 \\ 2 & 1-a \end{bmatrix} \quad \text{com} \quad a = \frac{1+(-1)^{h-1}}{2}; \quad P = [p_{ij}] \in M_{(h-1) \times 2}(\mathbb{R}) \text{ tal que}$$

$$p_{ij} = \begin{cases} 1, & \text{se } (i,j) = (h-1,1); \\ 0, & \text{caso contrário.} \end{cases}$$

Seja  $C$  uma matriz quadrada de ordem  $n$ . Denotamos por  $W(C) = [1 \ C \ C^2 \ \dots \ C^{n-1} \ 1]$  a matriz cadeia de  $C$ . Em particular, se  $C = A(G)$  é a matriz de adjacência de um grafo  $G$ , então  $W(G) = W(A(G)) = [w_{ij}]$  é denominada matriz cadeia de  $G$ , onde  $w_{ij}$  é o número de cadeias em  $G$  de comprimento  $j$  que começam no vértice  $v_i$ .

**Lema 7** (HAGOS, 2002). O posto da matriz cadeia de um grafo  $G$  é igual ao número dos seus autovalores principais.

**Lema 8** (HUANG *et al.*, 2015). Seja  $G$  um grafo. Se  $\pi(G)$  e  $Q$  denotam, respectivamente, uma partição equilibrada de  $G$  e a matriz quociente de  $G$  com respeito a  $\pi(G)$ , então o número de autovalores principais de  $G$  é igual ao posto de  $W(Q)$ .

Para  $h \geq 2$ , seja  $W_h = W(Q_h) = [1 \ Q_h \ 1 \ Q_h^2 \ 1 \ \dots \ Q_h^{h-1} \ 1] \in M_{(h+1)}(\mathbb{R})$  a matriz cadeia da matriz quociente  $Q_h$  associada a partição equilibrada  $\pi(F_h)$  do Lema 6. Para  $h=2$ , temos

$$W_2 = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 2 & 5 \\ 1 & 3 & 7 \end{bmatrix}. \text{ Para } h \geq 2 \text{ a matriz quociente de } W \text{ é dada de forma recursiva por}$$

$$W_3 = \begin{bmatrix} 1 & 2 & 4 & 9 \\ 1 & 2 & 5 & 11 \\ 1 & 3 & 7 & 18 \\ 1 & 2 & 6 & 14 \end{bmatrix}$$

e para  $h \geq 4$ , as entradas de  $W_h = [w_{ij}]$  são dadas por:

$$w_{ij} = \begin{cases} 1, & \text{se } j = 1; \\ w_{1(j-1)} + w_{2(j-1)}, & \text{se } (i = 1) \wedge (2 \leq j \leq h + 1); \\ w_{(i-1)(j-1)} + w_{i(j-1)} + w_{(i+1)(j-1)}, & \text{se } (2 \leq i \leq h) \wedge (i \text{ é ímpar}) \wedge (2 \leq j \leq h + 1); \\ w_{h(j-1)} + w_{(h-1)(j-1)} + w_{(h-2)(j-1)}, & \text{se } (i = h + 1) \wedge (i \text{ é ímpar}) \wedge (2 \leq j \leq h - 1); \\ 2w_{h(j-1)} + w_{(h+1)(j-1)}, & \text{se } (i = h + 1) \wedge (i \text{ é ímpar}) \wedge (h \leq j \leq h + 1); \\ w_{(i-1)(j-1)} + w_{(i+1)(j-1)}, & \text{se } (2 \leq i \leq h) \wedge (i \text{ é par}) \wedge (2 \leq j \leq h + 1); \\ w_{h(j-1)} + w_{(h-2)(j-1)}, & \text{se } (i = h + 1) \wedge (i \text{ é par}) \wedge (2 \leq j \leq h - 1); \\ 2w_{h(j-1)}, & \text{se } (i = h + 1) \wedge (i \text{ é par}) \wedge (h \leq j \leq h + 1). \end{cases}$$

Os dois lemas técnicos a seguir se referem a matriz cadeia  $W_h = [w_{ij}]$  cujas entradas estão definidas em (3), e são necessários para a demonstração do Teorema 11.

**Lema 9.** Para  $h \geq 3$ , considere a matriz cadeia  $W_h = [w_{ij}]$ . Dado o inteiro  $j$  tal que  $1 \leq j \leq h-1$ , temos

$$w_{ij} = w_{(i+2)j}, \text{ onde } j \leq i \leq h-1. \quad (4)$$

**Demonstração.** De fato, para  $h = 3$  segue que  $1 \leq i, j \leq 2$ . Por definição, temos:  $w_{11} = 1 = w_{31}$ ,  $w_{21} = 1 = w_{41}$  e  $w_{22} = 1 = w_{42}$ . Agora, suponhamos que a equação (4) é válida para  $h$  e verificaremos a permanência da veracidade de (4) para  $h+1$ .

Para  $1 \leq j \leq h$ , vamos verificar que  $w_{ij} = w_{(i+2)j}$ , onde  $j \leq i \leq h$ . Da hipótese de indução, se  $1 \leq j \leq h-1$ , temos  $w_{ij} = w_{(i+2)j}$  para todo  $i, j \leq i \leq h-1$ . Neste caso, falta verificar  $w_{ij} = w_{(i+2)j}$  para  $1 \leq j \leq h$  e  $i = h$ . Com efeito,  $w_{h1} = w_{(h+2)1}$ . Por conseguinte, se for par e  $2 \leq j \leq h-1$  então  $w_{(h+2)j} = w_{(h+1)(j-1)} + w_{h(j-1)}$ . Daí,

$w_{(h+2)j} - w_{hj} = (w_{(h+1)(j-1)} + w_{h(j-1)}) - (w_{(h-1)(j-1)} + w_{(h+1)(j-1)}) = 0$ . Por outro lado, temos por hipótese de indução que  $w_{(h+1)(j-1)} = w_{h(j-1)}$  o que implica em  $w_{(h+2)j} = w_{hj}$ , quando  $i = h$  for par e  $2 \leq j \leq h$ .

Analogamente, se verifica que  $w_{(h+2)j} = w_{hj}$ , quando  $i = h$  for ímpar e  $2 \leq j \leq h$ , o que conclui a demonstração.

**Lema 10.** Para  $h \geq 3$ , considere a matriz cadeia  $W_h = [w_{ij}]$ . Temos:

- (i)  $w_{(j+1)j} - w_{(j-1)j} = 1, 2 \leq j \leq h;$   
(ii)  $w_{(h+1)(h+1)} - w_{(h-1)(h+1)} = \begin{cases} \frac{h+4}{2}, & \text{se } h \text{ é par;} \\ \frac{h+3}{2}, & \text{se } h \text{ é ímpar.} \end{cases}$

**Demonstração.** Para o item (i), fazemos indução sobre  $h$ .

Para  $h = 3$ , temos  $w_{32} - w_{12} = 1$  e  $w_{43} - w_{23} = 1$ . Agora, suponhamos que a equação do item (i) é válida para  $h > 3$  e verificaremos que  $w_{(h+2)(h+1)} - w_{h(h+1)} = 1$ . Se  $h+2$  é ímpar então

$$w_{(h+2)(h+1)} = 2w_{(h+1)h} + w_{(h+2)h} \text{ e } w_{h(h+1)} = w_{(h-1)h} + w_{hh} + w_{(h+1)h}.$$

Assim,

$$w_{(h+2)(h+1)} - w_{h(h+1)} = [w_{(h+1)h} - w_{(h-1)h}] + [w_{(h+2)h} - w_{hh}].$$

Mas, da hipótese de indução, temos que  $w_{(h+1)h} - w_{(h-1)h} = 1$ . Logo, do Lema 9,  $w_{hh} = w_{(h+2)h}$ , pois,  $j \leq i \leq h$ . Consequentemente, para  $2 \leq j \leq h$ ,

$$w_{(h+2)(h+1)} - w_{h(h+1)} = 1.$$

Se é par, então

$$W_{(h+2)(h+1)} = 2W_{(h+1)h} \text{ e } W_{h(h+1)} = W_{(h-1)h} + W_{(h+1)h}$$

Assim,

$$W_{(h+2)(h+1)} - W_{h(h+1)} = W_{(h+1)h} - W_{(h-1)h}$$

Da hipótese de indução temos que  $W_{(h+1)h} - W_{(h-1)h} = 1$ , logo  $W_{(h+2)(h+1)} - W_{h(h+1)} = 1$  para  $2 \leq j \leq h$ . O que conclui a demonstração do item (i).

Para o item (ii), façamos indução sobre  $h$ .

Para  $h = 3$ , temos  $W_{44} - W_{24} = 14 - 11 = 3 = \frac{3+3}{2}$ . Agora, suponhamos que a equação do

item (ii) é válida para  $h > 3$  e verificaremos que  $W_{(h+2)(h+2)} - W_{h(h+2)} = \begin{cases} \frac{h+4}{2}, & \text{se } h \text{ é par;} \\ \frac{h+3}{2}, & \text{se } h \text{ é ímpar} \end{cases}$

Se  $h+2$  é par então

$$W_{(h+2)(h+2)} = 2W_{(h+1)(h+1)} \text{ e } W_{h(h+2)} = W_{(h-1)(h+1)} + W_{(h+1)(h+1)}$$

Assim,

$$W_{(h+2)(h+2)} - W_{h(h+2)} = W_{(h+1)(h+1)} - W_{(h-1)(h+1)}$$

Da hipótese de indução,  $W_{(h+1)(h+1)} - W_{(h-1)(h+1)} = \frac{h+4}{2}$ , pois  $h$  é par. Logo,

$$W_{(h+2)(h+2)} - W_{h(h+2)} = \frac{h+4}{2} = \frac{(h+1)+3}{2}$$

Se  $h+2$  é ímpar então

$$W_{(h+2)(h+2)} = 2W_{(h+1)(h+1)} + W_{(h+2)(h+1)} \text{ e } W_{h(h+2)} = W_{(h-1)(h+1)} + W_{h(h+1)} + W_{(h+1)(h+1)}$$

Assim,

$$W_{(h+2)(h+2)} - W_{h(h+2)} = W_{(h+1)(h+1)} + W_{(h+2)(h+1)} - W_{(h-1)(h+1)} - W_{h(h+1)}$$

Da hipótese de indução,  $W_{(h+1)(h+1)} - W_{(h-1)(h+1)} = \frac{h+3}{2}$ , onde  $h$  é ímpar. Logo,

$$W_{(h+2)(h+2)} - W_{h(h+2)} = \frac{h+4}{2} + [W_{(h+2)(h+1)} - W_{h(h+1)}]$$

Mas, pelo item (i), para  $j = h+1$  temos que  $W_{(h+2)(h+1)} - W_{h(h+1)} = 1$ . Assim,

$$W_{(h+2)(h+2)} - W_{h(h+2)} = \frac{h+3}{2} + 1 = \frac{(h+1)+4}{2}$$

Portanto,  $W_{((h+1)+1)((h+1)+1)} - W_{(h+1)((h+1)+1)}$  é par, o que conclui a demonstração.

Foi observado em testes computacionais, programado na linguagem python, o determinante e o posto da matriz cadeia  $W_h = [w_{ij}] \in M_{(h+1)}(\mathbb{R})$ , cujas entradas estão definidas em (3), a partir da quantidade de hexágonos nos grafos  $F_h$ , onde  $4 \leq h \leq 32$ . O grafo  $F_1$  possui exatamente um autovalor principal, pois é um grafo regular. Como  $\det(W_2) = -1$ , o grafo  $F_2$  possui três autovalores principais, que são os autovalores da matriz  $Q_2$ . Para  $h \geq 3$ , considere a matriz cadeia  $W_h = W(Q_h)$ . Substituindo a linha  $h+1$  de  $W_h$  pela subtração da mesma pela linha  $h-1$  e aplicando o Lema 9 e o Lema 10 obtemos

$$\det(W_h) = \begin{cases} -\det(W_{h-1}[\{h+1\},\{h\}]) + \frac{(h+4)}{2} \det(W_{h-1}), & h \text{ é par;} \\ -\det(W_{h-1}[\{h+1\},\{h\}]) + \frac{(h+3)}{2} \det(W_{h-1}), & h \text{ é ímpar.} \end{cases}$$

Aplicando indução sobre  $h \geq 3$ , se verifica por meio de operações elementares sobre as linhas da matriz cadeia  $W_h = [w_{ij}] \in M_{(h+1)}(\mathbb{R})$  que  $\det(W_h) = 0$  se e somente se  $h \equiv 1 \pmod{4}$ .

**Teorema 11.** O número de autovalores principais de  $F_h \in F$  é igual  $h$ , quando  $h \equiv 1 \pmod{4}$  e é igual a  $h+1$  nos demais casos.

Demonstração: Sejam  $F_h \in F$  e  $W_h = [w_{ij}]$  onde  $h \geq 3$ . Se  $h-1$  não é múltiplo de 4, então  $\det(W_h) \neq 0$ . Portanto, o posto de  $W_h$  é  $h+1$ . Do Lema 7, o número de autovalores principais de  $F_h$  é igual  $h+1$ .

Por outro lado, se  $h \equiv 1 \pmod{4}$  então  $\det(W_h) = 0$ . Porém, se  $h = 4q+2$  para algum inteiro  $q$ , temos:

$$\det(W_{4q+2}) = (2q+3) \det(W_{4q+1}) - \det(W_{4q+2}[\{4q+3\},\{4q+2\}]).$$

Como  $\det(W_{4q+2}) \neq 0$  e  $\det(W_{4q+1}) = 0$  obtemos que  $\det(W_{4q+2}[\{4q+3\},\{4q+2\}]) \neq 0$ . Consequentemente,  $W_h$  tem posto  $h$  e, do Lema 7, concluímos que  $F_h$  possui exatamente  $h$  autovalores principais.

## CONCLUSÕES

Neste trabalho estudamos propriedades espectrais de uma classe de grafos benzenóides. No Teorema 5, para cada  $h \geq 2$ , determinamos expressões explícitas para os autovalores de  $F_h$ . Por fim, no Teorema 11, mostramos que a quantidade de autovalores principais de tais grafos é igual a  $h$  quando  $h \equiv 1 \pmod{4}$  e é igual a  $h+1$  nos outros casos. Este resultado resolve a questão, proposta por Cvetković, sobre a classificação dos grafos com exatamente  $h$  autovalores principais para os grafos na classe  $F$ .

## AGRADECIMENTOS

Este trabalho é parcialmente financiado pelo MATH-AMSUD no âmbito do projeto GSA, equipe brasileira financiada pela CAPES no projeto 88881.694479/2022-01. O primeiro autor agradece o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- Brasil (CAPES). Os demais autores agradecem o apoio da FAPERJ, processo E-20/2022-284573.

## REFERÊNCIAS

- CVETKOVIC, D. M. The main part of the spectrum, divisors and switching of graphs. **Publications de l'Institut Mathématique (Beograd)**, v. 37, p. 31-38, 1978.
- DAS, K.C.; GUTMAN, I. Some properties of the second Zagreb index. **MATCH Communications in Mathematical and in Computer Chemistry**, v. 52, p. 103-112, 2004.
- DIESTEL, R. **Graph theory, in: Graduate Texts in Mathematics** 173. 2. ed. Springer-Verlag, Berlin Heidelberg, 2000.
- GUTMAN, I. **Selected Theorems in Chemical Graph Theory**. 1.ed. Univ. Kragujevac, Kragujevac, 2017.
- GUTMAN, I.; FURTULA, B. The Total  $\pi$ -Electron Energy Saga. **Croatica Chemica Acta**, v.90, n. 3, p. 359-368, 2017.
- HAGOS, E.M., Some results on graph spectra, **Linear Algebra and its Applications**, v. 356, p. 103–111, 2002.
- HORN, R. A.; JOHNSON, C. R. **Matrix Analysis**. 2. ed. Cambridge, 2013.
- HUANG, X.; HUANG, Q.; LU, L. Construction of graphs with exactly k main eigenvalues. **Linear Algebra and its Applications**, v. 486, p. 204–218, 2015.
- LOSONCZI, L. Eigenvalues and eigenvectors of some tridiagonal matrices. **Acta Mathematica Hungarica**, v. 60, p. 309-332, 1992.
- MORGAN, H. L. The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. **Journal of Chemical Documentation**, v. 5, n. 2, p. 107-113, 1965.
- NATH, M.; PAUL, S. On the distance Laplacian spectra of graphs. **Linear Algebra and its Applications**, v. 460, p. 97-110, 2014.
- RANJINI, P. S.; LOKESHA, V.; CANGUL, I. N., On the Zagreb indices of the line graphs of the subdivision graphs. **Applied Mathematics and Computation**, v. 218, p. 699-702, 2011.

## DADOS DOS AUTORES

**Allan José Amaral Ribeiro**. Universidade Federal Fluminense, allanjose@id.uff.br

**Francisca Andrea Macedo França**. Universidade Federal Fluminense, francisca\_franca@id.uff.br

**André Ebling Brondani**. Universidade Federal Fluminense, andrebrondani@id.uff.br



## MODELAGEM 3D DE CONCRETO SOBRE ESTRESSE TÉRMICO APLICANDO O MÉTODO DOS VOLUMES FINITOS

### 3D MODELING OF CONCRETE UNDER THERMAL STRESS APPLYING THE FINITE VOLUME METHOD

João Victor de Mello Marinho

Victor de Oliveira Carvalho

Jorge Juárez Trujillo

José Adilson de Castro

**Resumo:** O concreto civil de alta performance e os *castables* refratários apresentam o problema de lascado térmico, em vezes violentamente explosivo. Na construção civil, esse fenômeno pode se apresentar em casos de incêndio, aumentando os riscos para sua mitigação. Na indústria, os monólitos refratários apresentam esse problema durante sua instalação, precisando a definição de uma curva de aquecimento segura para eliminar a umidade presa nos poros o mais rápido possível. Atualmente, essas curvas são determinadas empiricamente. O presente estudo propõe a aplicação do método numérico dos volumes finitos para simular esse comportamento por computador, podendo ajustar os parâmetros envolvidos, seja para um concreto civil ou refratário. A solução é obtida através da linearização de um sistema de duas equações diferenciais parciais parabólicas. O modelo conseguiu descrever tridimensionalmente o comportamento transiente de um material cimentício aquecido até 450°C, resultando valores aceitáveis das temperaturas internas, pressões dos poros e eliminação da água livre.

**Palavras-chave:** Fenômenos de transporte. Método dos volumes finitos. Materiais cimentícios. Meio poroso.

**Abstract:** The high-performance civil concrete and refractory castables both exhibit the problem of thermal spalling, sometimes explosively. In civil construction, this phenomenon can occur in cases of fire, increasing risks for mitigation. On the other hand, industrial refractory monoliths present this risk during their installation and initial heating, where the definition of a carefully adjusted temperature curve is required to eliminate moisture trapped in the pores as quickly as possible. Currently, in the industry, these curves are mainly developed empirically. This study proposes the application of the numerical finite volume method to simulate this behavior on a computer, allowing the adjustment of involved parameters, whether for civil or refractory concrete. The solution is obtained through the linearization of a system of two parabolic partial differential equations. The model was able to three-dimensionally describe the transient behavior of a cementitious material subjected to thermal stress, resulting in acceptable values for internal temperatures, pore pressures, and elimination of free water up to a temperature of 450°C.

**Keywords:** Transport phenomena. Finite volume method. Cementitious materials. Porous media.

## INTRODUÇÃO

O concreto é uma mistura de cimento e agregados como pedra brita e areia sílica, amplamente utilizado em diversos setores como construção civil, túneis, ferrovias, usinas siderúrgicas e nucleares. Ao longo de recentes décadas, estudos na mistura de seus componentes desenvolveram técnicas como a otimização de tamanho de partícula, redução do conteúdo de água e o emprego de aditivos como o fumo de sílica que resultam em ganhos de propriedades como densidade, resistência mecânica e propriedades reológicas (MILLARD, 2019). Contudo, tais concretos, denominados concretos de alta performance, podem apresentar o fenômeno de lascado térmico ou explosivo, quando expostos ao estresse térmico.

Acredita-se que eventos catastróficos ocorrem principalmente devido às propriedades de permeabilidade e condutividade térmica oriundas das novas misturas. Nos anos de 1996 e 1998 ocorreu no incêndio do túnel da Mancha (ULM, 1999) que durou cerca de 10 horas, atingindo temperaturas de 700°C. O acidente envolveu a falha explosiva dos anéis de concreto que sustentam as estruturas dos túneis. Os danos são representados na figura 1 e foram estimados em 1.5 milhão de dólares por dia no período 6 meses e que o túnel permaneceu fechado para reparos.

**Figura 1:** Danos ao túnel da Mancha



Fonte: MANGABHAI et al., 2012.

Segundo Mindeguia (2010), as etapas de fragmentação do concreto podem ser definidas como: aquecimento rápido, que gera altas tensões unilaterais, eventualmente superando a resistência mecânica do material, mudanças rápidas de volume e estrutura dos agregados e pressão interna gerada pela liberação de vapores de água e gases do agregado e da massa de cimento que ficam presos.

No setor siderúrgico, tem-se desenvolvido misturas para concretos refratários, capazes de sustentar altas temperaturas de trabalho para serem utilizados em ambientes extremos, por exemplo, nos canais de corrida dos altos fornos para transportar a escória e o metal fundido até os carros torpedo. Para evitar o aparecimento dos fenômenos responsáveis pela fragmentação do concreto, são desenvolvidas curvas de aquecimento gradual (TRUJILLO, 2020). Palmer et al. (2014) ressaltam a importância econômica desses estudos, uma vez que as curvas de secagem industrial são conservadoras e definidas empiricamente, além disso a segurança é parte crucial nesse estudo, o fenômeno explosivo pode resultar na queda de estruturas e lançamento de detritos que se tornam projéteis perigosos. Além disso, Palmer et al. (2014) citam os diversos pesquisadores e seus esforços na compreensão dos diferentes fenômenos envolvidos na falha de lascado térmico dos concretos e nas particularidades envolvidas entre tipo de mistura, evitando falhas como a exemplificada na figura 2.

**Figura 2:** Concreto refratário com dano explosivo



Fonte: PALMER et al., 2014.

Atualmente, foram feitos grandes esforços para mensurar experimentalmente as propriedades envolvidas e a partir desses resultados propor relações constitutivas entre os parâmetros, com o intuito de prognosticar o comportamento em altas temperaturas do concreto, entendido como um meio poroso, mas, com baixa permeabilidade onde deve fluir vapor de água. Kalifa (2000) realizou medições experimentais de pressão, temperatura e perda de massa, aquecendo amostras de concretos de alta performance. Trujillo et al. (2023) desenvolveram uma solução computacional aplicada em concretos refratários, utilizando o método dos volumes finitos num domínio 3D (tridimensional) semelhante ao corpo de prova utilizado no experimento de Kalifa. Para adquirir maior conhecimento sobre o fenômeno, sem altos gastos, é possível desenvolver simulações numéricas,

testando diferentes curvas de aquecimento em concretos com diferentes propriedades intrínsecas, como permeabilidade e condutividade térmica. Assim, este artigo busca aplicar o modelo 3D resolvido com o método dos volumes finitos desenvolvido por Trujillo et al. (2023), com uma curva de aquecimento até 450°C para simular o comportamento dos campos de temperatura, pressão interna dos poros e variação de massa num bloco de concreto aquecido.

## MÉTODOS MATEMÁTICOS E NUMÉRICOS

### Modelo matemático

Este trabalho baseia-se no modelo matemático proposto por Bazant e Jirásek (2018), que consiste num sistema de duas equações diferenciais parciais parabólicas (EDP) que representam a conservação de massa (vapor de água) e a energia. Suas incógnitas são pressão  $P$  e temperatura  $T$ . O sistema representa o fluxo monofásico transiente de vapor de água num meio poroso que sofre alterações de permeabilidade em função do aumento de temperatura. A solução do sistema de equações é determinada através do método dos volumes finitos que converte as EDPs em equações lineares (CHAPRA, 2010) resolvidas por processos iterativos através de um código de programação em FORTRAN desenvolvido por Trujillo et al. (2023).

### Equações de conservação

A conservação da massa considera as variações da água livre e da água quimicamente ligada que são liberadas durante o aquecimento, acordo a equação (1). A de Darcy serve para descrever o transporte de umidade gerada pelo gradiente de pressão dentro dos poros de concreto, acordo a equação (2).

$$\frac{\partial W}{\partial t} = -\nabla \cdot J + \frac{\partial W_d}{\partial t} \quad (1)$$

$$J = -\frac{\alpha}{g} \nabla P \quad (2)$$

Sendo  $W_d$  a água liberada pela desidrataação,  $W$  é a quantidade de água livre dependente da umidade relativa dos poros  $h$ ,  $g$  é a gravidade,  $\alpha$  é a permeabilidade do concreto e  $J$  é o fluxo de vapor d'água proporcional ao gradiente de pressão  $\nabla P$ .

Em temperaturas abaixo de 100°C, o calor no concreto é conduzido de 30 a 100 vezes

mais rapidamente do que a umidade. Isso resulta em uma penetração de calor muito mais profunda. Essa diferença permite desacoplar os problemas, assim resolvendo primeiro o problema de transferência de calor, considerando trincas tão finas que sua contribuição para a permeabilidade efetiva seja insignificante. A transferência de umidade acelera após 100°C, especialmente entre 95-105°C, onde a permeabilidade ao vapor d'água aumenta consideravelmente. Além da condução, o calor pode ser transferido por convecção durante a difusão do vapor d'água, resultando no acoplamento bidirecional entre a transferência de calor e a umidade no concreto.

Já na equação (3), o balanço de energia considera as entalpias do sistema. A equação (4) é a lei de Fourier para condução de calor via gradiente de temperatura e convecção por massa de vapor. Como o aquecimento envolve transporte de massa e calor, a conservação de massa é complementada com a equação do balanço de energia.

Para fins práticos, é possível realizar ajustes e simplificações específicas, aonde a equação do balanço energético pode ser transformada na equação do calor. Nessa simplificação, a equação do calor aborda o concreto como um sistema multifásico, composto por uma estrutura sólida contendo poros ocupados por água e vapor.

$$\rho C \frac{\partial T}{\partial t} - C_a \frac{\partial W}{\partial t} = -\nabla \cdot q + C_w J \cdot \nabla T \quad (3)$$

$$q = -\lambda \nabla T + C_w J \cdot \nabla T \quad (4)$$

Sendo  $\rho$  a densidade da matriz de concreto;  $C$  é o calor específico;  $C_a$  é o calor latente de vaporização da água livre;  $C_w$  é o calor específico da água;  $q$  é o fluxo de calor, e por fim  $\lambda$  é a condutividade térmica do concreto.

Juntando as equações de balanço de massa e energia é criado um sistema de duas EDPs parabólicas com quatro variáveis de estado ( $W_d, W, T, P$ ). Sendo  $W_d$  calculada através da equação (5) que é uma função por partes obtida da análise termogravimétrica do concreto. Como  $W$  é uma função termodinâmica de  $P$  e  $T$ , sua variação no tempo pode ser substituída por sua variação sobre  $P$  e  $T$ . Com isso, duas funções de estado devem ser calculadas pelas EDPs, pressão e temperatura.

$$W_d(T) = \begin{cases} 0,0017 * T - 0,0171 & (0 < T < 200) \\ 0,0407 * T - 7,5143 & (200 \leq T < 400) \\ 0,0025 * T + 6,5253 & (400 \leq T \leq 800) \end{cases} \quad (5)$$

Este modelo matemático apresenta simplificações importantes. Por exemplo, o balanço de massa considera a umidade como um fluido que engloba o ar, a água líquida e o vapor. Embora cada uma dessas fases possa ter sua própria equação de balanço de massa, no caso do aquecimento inicial de concreto, após a cura, a presença de ar pode ser mínima. Somente em temperaturas acima de 250 ou 300°C, quando o vapor de água começa a sair dos poros, o ar entra em ação, substituindo o vapor de água na microporosidade.

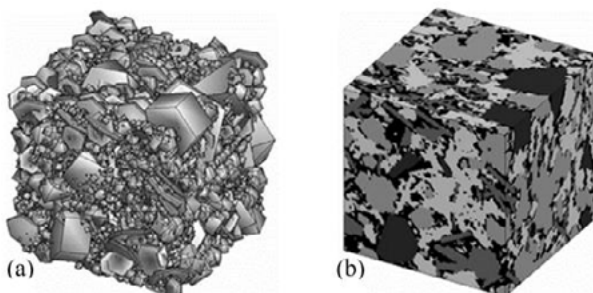
Na equação do balanço de energia, considerou-se o calor específico do concreto e da água líquida como constantes, embora suas contribuições possam variar com a temperatura e as possíveis contribuições de gases foram desconsideradas. Além disso, a equação de energia não contempla o calor fornecido por uma fonte de aquecimento por unidade de volume do concreto e tempo. Essa fonte é considerada apenas nas condições de contorno. A taxa de calor consumida pelo processo de desidratação também não está incluída neste modelo.

Outra simplificação envolve o primeiro aquecimento do concreto refratário, onde a estrutura porosa pode passar por transformações próprias, como solidificação (cristalização) de componentes específicos, permanecendo uma estrutura porosa ainda em estado transiente.

### Método de discretização

A discretização foi feita por um método totalmente implícito proposto por Patankar (1980). Neste método, as equações diferenciais foram manipuladas para se conformar a uma equação geral, que por sua vez foi discretizada em um domínio representativo do corpo de prova dos experimentos realizados. Esse corpo de prova foi dividido em uma série de volumes de controle não sobrepostos, como ilustrado na figura 3. Em seguida, cada nó da malha foi posicionado no centro de cada volume de controle.

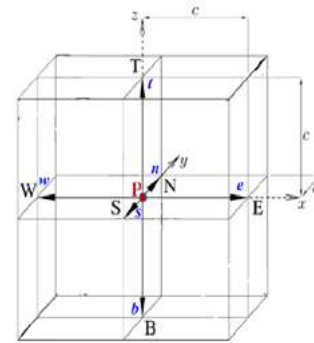
**Figura 3:** Volume de controle representativo do refratário (a) - representação próxima à realidade - e (b) - discretização para malha cartesiana.



Fonte: Adaptado de (ZEHMISCH et al., 2014).

O domínio foi um bloco de concreto de 300 x 300 x 120 mm dividido em volumes de controle regulares não sobrepostos. Cada volume forma um espaço tridimensional retangular, formando um estêncil de pontos na grade, representado pela Figura 4. Cada ponto da malha tem um nó, que pode representar, pressão, temperatura, ou conteúdo de água. Para se obter os valores dos nós, as equações (1) e (3) são integradas no volume de controle em intervalo de tempo definido.

**Figura 4:** Estêncil empregado para simulação 3-D.



Fonte original: (TRUJILLO, 2020).

Para a discretização espacial, as equações (6) e (7), fruto do desenvolvimento das equações de conservação de massa e energia, respectivamente, são integradas nos volumes de controle. Um perfil linear por partes é assumido com interpolação para P e T entre os nós, para calcular fluxos de massa. Para a discretização do tempo, foi usado um esquema implícito baseado em Patankar (1980). Com isso, foram desenvolvidos dois sistemas de equações lineares. Resolvendo primeiramente o campo de temperaturas, e com esses valores é calculado, em uma seguinte iteração, o campo de pressões.

$$\frac{\partial P}{\partial t} = \nabla \cdot \left( \frac{u}{A1} \nabla P \right) + \frac{1}{A1} \cdot \frac{\partial W_d}{\partial t} - \frac{A3}{A1} \cdot \frac{\partial T}{\partial t} \quad (6)$$

$$\frac{\partial T}{\partial t} = \nabla \cdot \left( \frac{\lambda}{A4} \nabla T \right) + \frac{C_w}{A4} \cdot \left( \frac{\alpha}{\rho} \cdot \nabla P \right) \cdot \nabla T - \frac{A2}{A4} \cdot \frac{\partial P}{\partial t} \quad (7)$$

### Condições de contorno

Para o fluxo térmico, somente a face superior foi aquecida por meio de radiação, seguindo uma curva de aquecimento próxima à utilizada por Kalifa (2000). A condição de radiação foi definida pela lei de Stefan-Boltzmann (BOLTZMANN, 1884), expressa na equação (8):

$$\vec{q} \cdot \vec{n} = \epsilon_c \cdot \sigma_{SB} (T^4 - T_{amb}^4) \quad (8)$$

Sendo  $n$  o vetor unitário normal ortogonal da superfície;  $\vec{q}$  é o fluxo de calor por radiação;  $T$  é a temperatura do corpo radiante;  $T_{amb}$  é a temperatura ambiente;  $\sigma_{SB} = 5,67 \times 10^{-8} \text{ W}/(\text{m}^2 \text{ K}^4)$  é a constante de Stefan-Boltzmann e  $\epsilon_c = 0,9$  é a emissividade de calor da superfície do concreto (JONES, 2000).

As faces laterais simulam estar em contato com blocos isolantes de baixa condutividade térmica, então, foram consideradas superfícies com uma condição de contorno do tipo Neumann, Equação (9):

$$\vec{q} \cdot \mathbf{n} = 0 \quad (9)$$

Na região inferior do bloco, para a condição de transferência de calor foi utilizada a lei de Newton de resfriamento, Equação (10):

$$\vec{q} \cdot \mathbf{n} = B_T(T - T_{amb}) \quad (10)$$

Sendo  $T$  a temperatura da superfície do domínio;  $B_T = 1,0 \text{ W}/(\text{m}^2 \text{ K})$  é o coeficiente de transferência de calor da superfície. Foi considerada convecção livre na superfície (CHAPMAN, 1987).

Para o fluxo de massa, o bloco de concreto foi definido com contornos permeáveis. Porém, a face inferior está em contato com uma chapa de aço, considerada impermeável, similar ao arranjo de Trujillo (2020), dessa forma, os fluxos normais a esse contorno desaparecem numa condição de contorno do tipo Neumann, acordo a equação (11).

$$\vec{j} \cdot \mathbf{n} = 0 \quad (11)$$

Sendo  $n$  o vetor unitário normal ortogonal da superfície de contorno;  $\vec{j}$  é o fluxo da massa de água. A face superior do corpo de prova está em contato com a atmosfera e as faces laterais estão em contato com tijolos refratários isolantes de alta porosidade, nesses contornos a pressão  $P$  foi definida como a pressão atmosférica  $P_{amb}$ . A troca de vapor entre a superfície e o meio ambiente não é instantânea, definindo que o fluxo normal é proporcional à diferença de pressões, conforme a eq. (12):

$$\vec{j} \cdot \mathbf{n} = B_w(P - P_{amb}) \quad (12)$$

Sendo  $B_w = 1 \times 10^{-6}$  o coeficiente de transferência de vapor da superfície dependente da movimentação do ar, a iluminação ou o isolamento parcial.

## Parâmetros iniciais

Em razão das propriedades não lineares, dependentes de temperatura, saturação, pressão e composição do concreto, os dados, relações e formulações das propriedades foram obtidas utilizando dados da literatura, como dispostos na Tabela 1:

**Tabela 1:** Principais parâmetros iniciais para a simulação

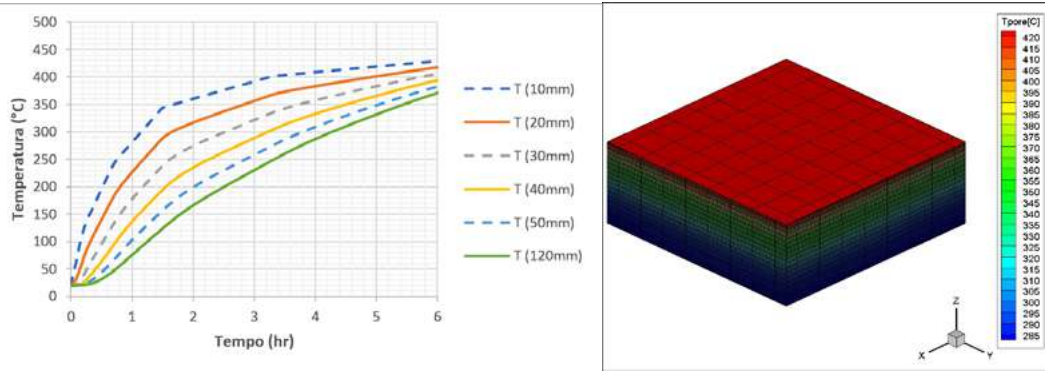
Propriedade	Nomenclatura	Origem	Valor	Unidade
Permeabilidade intrínseca	$\kappa$	(GUO, 2016)	$1,0 \cdot 10^{-17}$	$\text{m}^2$
Condutividade térmica	$\lambda$	(GUO, 2016)	1,64	$\text{W}/\text{m} \cdot \text{K}$
Duração da simulação	$t_r$	(Definição própria)	40	hr
Tamanho do passo espacial	$\Delta(x, y, z)$	(Definição própria)	0,25	mm
Tamanho do passo temporal	$\Delta t$	(Definição própria)	0,25	s
Temperatura inicial	$T_0$	(Definição própria)	20	$^{\circ}\text{C}$

## RESULTADOS E DISCUSSÃO

### Campo de Temperaturas

Na simulação foi usada uma curva de aquecimento desde  $20^{\circ}\text{C}$  até  $450^{\circ}\text{C}$  em um período de 6 horas, como mostra a Figura 5. Em todas as simulações foram considerados seis pontos dentro do material com profundidades de 10, 20, 30, 40, 50 e 120 mm. Se observa que quanto mais próximo da fonte quente, maior a temperatura do material, isso se dá pela condução de calor através do concreto. Vale ressaltar que, devido ao isolamento lateral, a simulação não considera a perda de calor do material para o ambiente. Porém, em uma situação real poderia haver uma saída mínima de calor para o ambiente, pelo que as curvas poderiam estar mais distantes entre si, o que mostra a necessidade dessa consideração.

**Figura 5:** Curvas de temperaturas em diferentes profundidades e imagem 3d na hora 4 de aquecimento.



Fonte: Gerado pelos autores

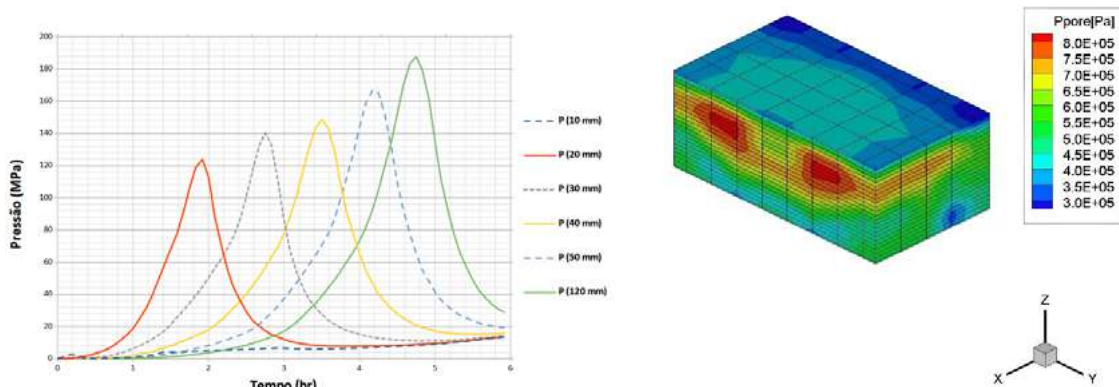
### Campo de Pressões internas

A figura 6 a seguir mostra os resultados das pressões internas. Nota-se que durante o aquecimento, ao longo do tempo há um aumento gradual com picos de pressão, com máximo de 188 MPa atingido na posição de 120 mm localizada na extremidade mais oposta ao aquecimento, as elevadas pressões atingidas estão relacionadas a baixa permeabilidade dos concretos de alta performance. Observa-se que a pressão em 10 mm atingiu valores de pressão por abaixo das demais, apresentando um comportamento linear ao longo da simulação, acredita-se que por estar mais próxima à face aquecida, não houve acúmulo de umidade suficiente para gerar pressões internas elevadas. É possível observar que

os picos de pressão, com exceção da posição de 10mm, são atingidos em intervalos de tempo que coincidem com níveis de temperatura em torno de 300 °C, e onde há aumento de conteúdo de água, este fenômeno ocorre em condições próximas à temperatura crítica da água.

Na figura 6 à direita, apresentam-se os resultados das pressões simuladas em 3 dimensões, no instante de tempo  $t = 0,5$  h em um corte de 50% do domínio. É possível observar no interior maiores níveis de pressão nas regiões centrais do domínio, a partir dos 20 mm de profundidade. Este fenômeno ocorreu ao início da simulação do aquecimento, quando esta região estava numa temperatura de 180 °C.

**Figura 6:** Curvas de pressões em diferentes profundidades e imagem 3d na hora 0,5 do aquecimento.

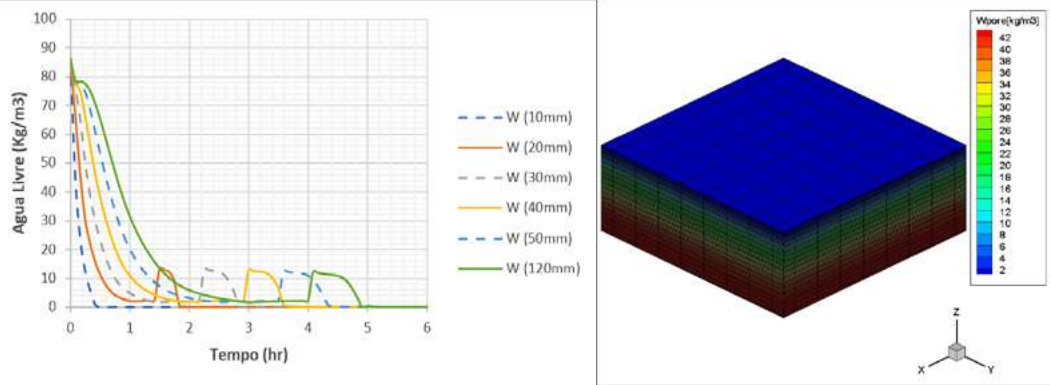


Fonte: Gerado pelos autores

### Conteúdo de água

A Figura 7 esquerda mostra o comportamento da saída de água livre; ao lado direito mostra o campo 3D da quantidade de água livre após 1h de aquecimento. Se percebe que quanto mais próximo da superfície quente, mais rápido a água livre é liberada, devido a esses pontos chegarem primeiro a temperaturas críticas, fazendo com que a água nessas regiões evapore primeiro. Em 1 hora a maior parte da umidade foi removida, em função do rápido aquecimento da fonte de calor. Em quase todas as profundidades simuladas ocorrem picos, acredita-se que podem ser da água quimicamente ligada que sai do concreto em maiores temperaturas. Às 5h se observa que toda água livre foi liberada do concreto.

Figura 7: Curvas do conteúdo de água em diferentes profundidades e imagem 3d na hora 1,0 de aquecimento.



Fonte: Gerado pelos autores

## CONCLUSÃO

Neste trabalho foi desenvolvida uma aplicação do método dos volumes finitos, caracterizado como conservativo por descrever os fenômenos de transporte levando-se em consideração a conservação de massa e energia. Foi possível simular o comportamento do vapor dentro de um concreto de alta performance submetido a estresse térmico. As curvas de temperatura apresentaram resultados satisfatórios, não entanto, é preciso um maior refinamento e ajustes nas condições de contorno, a fim de determinar de forma precisa condições com perdas de energia por saídas de calor que podem ocorrer devido a imperfeições do isolamento lateral nos cenários práticos. As curvas de pressão atingiram patamares consideravelmente elevados, isso é explicado pela baixa permeabilidade deste tipo de concreto, contudo, a sua validação somente seria possível com experimentos do caso específico, os quais são sofisticados de realizar. O comportamento do conteúdo de água nas diferentes profundidades mostrou uma queda abrupta ao início da simulação, porém, também mostrou a necessidade de maior refinamento, pois foi seguida de um aumento elevado em momentos posteriores.

## Referências

- BAZANT, Z. P.; JIRÁSEK, M. **Creep and hygrothermal effects in concrete structures**. Evanston: Springer. Solid Mechanics and its Applications, 2018. 960 p.
- BOLTZMANN, L. Derivation of Stefan's law: concerning the dependence of thermal radiation on the temperature of the electromagnetic light theory. **Annals of Physics and Chemistry**, v. 22, p. 291-294, 1884.
- CHAPMAN, A. J. **Fundamentals of heat transfer**. New York: Macmillan, 1987.
- CHAPRA, S.C.; Canale, R.P.; **Numerical Methods for Engineers**. 6. ed. New York: McGraw-Hill, 2010.
- GUO, F.; YUAN, Y.; MANG, H. A. Determination of the relative significance of material parameters for concrete exposed to fire. **International Journal of Heat and Mass Transfer**, v. 100, p. 191-198, 1 set. 2016.
- JONES, H. R. N. **Radiation heat transfer**. Oxford: Oxford University Press, 2000.
- KALIFA, P.; MENNETEAU, F. D.; QUENARD, D. Spalling and pore pressure in HPC at high temperatures. **Cement and Concrete Research**, v. 30, n. 12, p. 1915-1927, Agosto 2000.
- MANGABHAI, Raman & Hewlett, P & Taylor, G & Domone, P & Trout, Edwin & Berrie, I & Killoran, D. (2012). **ICT Yearbook 2012-13**. 17. ed. The concrete Society. 2012.
- MILLARD, A.; PIMIENTA, P. **Modelling of concrete behaviour at high temperature**. State-of-the-Art Report of the RILEM Technical Committee 227-HPB. Cham, Switzerland, p. 100. 2019. (30).
- MINDEGUIA, J. C.; PIMIENTA, P.; NOUMOWÉ, A.; KANEMA, M. Temperature, pore pressure and mass variation of concrete subjected to high temperature: Experimental and numerical discussion on spalling risk. **Cement and Concrete Research**, v. 40, n. 3, p. 477-487, 2010.
- PALMER, G.; COBOS, J.; MILLARD, J.; HOWES, T. The accelerated drying of refractory concrete: Part I: A review of current understanding. **Refractories WORLDFORUM**, v. 6, n. 2, p. 75 – 83, April 2014.
- PATANKAR, S. V. Numerical heat transfer and fluid flow. **Hemisphere publishing corporation**, 1980.
- TRUJILLO, J. J.; **Contribuições para a determinação de curvas de secagem de refratários monolíticos com auxílio experimental e de simulação numérica**. 2020, 377 f. Tese (Doutorado em Ciências) –Escola de Engenharia de Lorena Universidade de São Paulo, Lorena, 2020.
- TRUJILLO, J.J. CASTRO, A. INNOCENTINI, M. VERNILLI, F. Combined transient 3D simulation and experimental methods to assess a slow heat-up curve used to dry a refractory concrete. **International Journal of Thermal Sciences**, v. 185, 2023.
- ULM, F. J.; COUSSY, O.; BAZANT, Z. P. The “chunnel” fire. I: chemoplastic softening in rapidly heated concrete. **Journal of Engineering Mechanics**, v. 125, n. 3, mar. 1999

## DADOS DOS AUTORES

**João Victor de Mello Marinho**. Graduando em Engenharia Metalúrgica na Escola de Engenharia Industrial Metalúrgica da Universidade Federal Fluminense, UFF, joavmma@id.uff.br

**Victor de Oliveira Carvalho**. Graduando em Engenharia de Materiais na Escola de Engenharia Industrial Metalúrgica da Universidade Federal Fluminense, UFF, victoroliveiracarvalho@id.uff.br

**Jorge Juárez Trujillo**. Graduado em engenharia industrial pelo Instituto Politécnico Nacional do México, mestre em Engenharia Mecânica pela Universidade Federal Fluminense, doutor em Ciências pela Universidade de São Paulo no programa de Engenharia de Materiais. Docente no departamento de Engenharia Metalúrgica e de Materiais da Universidade Federal Fluminense, UFF, jorgejuareztrujillo@id.uff.br

**José Adilson de Castro**. Graduado em Engenharia Metalúrgica pela Universidade Federal de Ouro Preto, mestre em Engenharia Metalúrgica pela Universidade Federal Fluminense, Ph. D em Engenharia com ênfase em simulação de processos pela Tohoku University, no Japão. Professor do Departamento de Engenharia Mecânica da Universidade Federal Fluminense, UFF, joseadilsoncastro@id.uff.br



# ROTEAMENTO DE MÉDICOS POR ESPECIALIDADE EM SITUAÇÃO DE PÓS - DESASTRE ATRAVÉS DE UM MODELO DE PROGRAMAÇÃO LINEAR INTEIRA MISTA

Raila de Medeiros Melo

Eliane da Silva Christo

Tiago Araújo Neves

**Resumo:** O artigo apresenta um modelo matemático que pode auxiliar na tomada de decisão para garantir a minimização do atraso do tempo de atendimento das equipes médicas para distribuição de socorro na fase de resposta a um desastre, considerando as especialidades médicas da equipe e a demanda por estas especialidades. Para construção da abordagem exata é utilizado um modelo matemático similar ao usado para problemas de roteamento com janelas de tempo. Os resultados mostram que a abordagem exata entrega soluções viáveis aos decisores, porém, para grandes instâncias, é necessário um tempo computacional relativamente alto.

**Palavras-chave:** Logística Humanitária. Modelagem Matemática para Distribuição de Socorro Médico. Roteamento Médico. Modelagem Computacional. Programação Linear Inteira Mista.

**Abstract:** The article presents a mathematical model that support in decision-making to ensure the minimization of delays in medical response time for aid distribution in a disaster, considering the medical specialties and their demand. To Construct the exact approach, there is a mathematical model similar to used for routing problems with time windows. The results show that the exact approach delivers viable solutions to decision makers, however, for large instances, a relatively high computational time is required.

**Keywords:** Humanitarian Logistics. Mathematical Modeling for Medical Aid Distribution. Medical Routing. Computational modeling. Mixed Integer Linear Programming.

## INTRODUÇÃO

Furacões, enchentes, tempestades, rompimento de barragens entre muitos outros desastres afetam diversas pessoas anualmente e causam imensuráveis danos ao redor do mundo. De acordo com CRED (2021), desde 1990 houve 16.165 desastres naturais em todo o mundo, atingindo cerca de 8,3 bilhões de pessoas, gerando 32,4 milhões de mortes e um custo aproximado de 5,4 bilhões de dólares. O Brasil ocupa a 10ª posição no ranking de países afetados pelos desastres segundo o CRED (2021). Segundo a CNN Brasil (Janone, 2022), no início de 2022, aproximadamente 8 milhões de brasileiros foram afetados por desastres ambientais. Eventos recentes, como os rompimentos de barragens em Mariana (2015) e Brumadinho (2019) em Minas Gerais, deixaram um impacto duradouro na população e causaram a necessidade de auxílio humanitário. O acidente de Brumadinho, por exemplo, levou mais de dez mil pessoas a deixarem suas casas (Folha de S.Paulo, 2022). Diante desses desastres devastadores, a logística humanitária (LH) surgiu como uma forma de gerenciar crises e prestar ajuda em momentos críticos. A LH busca organizar, coordenar e efetuar assistência de maneira eficaz, abrangendo evacuações, distribuição de alimentos, água, suprimentos de emergência, medicamentos, voluntários, equipes médicas e a gestão de resíduos pós-desastre. Com base nesse cenário, este estudo apresenta um modelo matemático para reduzir atrasos nos atendimentos, considerando as especialidades médicas da equipe e as demandas específicas por essas especialidades.

## LOGÍSTICA HUMANITÁRIA

Logística Humanitária (LH) pode ser conceituada como processo de planejar, implementar e controlar de forma eficiente o fluxo e o armazenamento de bens, materiais e informações relacionadas do ponto de origem até o ponto de consumo, com o intuito de aliviar o sofrimento de pessoas em situações vulneráveis, de Oliveira (2019).

Thomas (2004) destaca também que a LH deve garantir o fluxo de suprimento e pessoas, assim como assegurar que esse fluxo ocorra de maneira eficiente e eficaz, ou seja, o auxílio deve chegar de maneira correta e no tempo oportuno.

Nogueira e Gonçalves (2009) representando a Federação Internacional da Cruz Vermelha, acrescenta ainda que a LH busca a pronta resposta, visando atender o maior número de

pessoas, evitar falta e desperdício, organizar as diversas doações que são recebidas nesse caso e, principalmente, atuar dentro de um orçamento limitado.

Diferentes estudos mostram que as principais tomadas de decisões em LH visando minimizar a situação de vulnerabilidade das pessoas afetadas tanto antes do desastre como após, se concentram em três principais áreas, sendo elas: localização de instalações, distribuição de socorro e evacuação em massa. O estudo concentra-se na fase de distribuição socorro, para melhor compreensão dela a seguir será apresentada com maiores detalhes a conceituação e principais estudos encontrados. É possível consultar os principais estudos relacionados a demais fases em Melo (2023). Os estudos, incluindo este serão de modelos determinísticos. Existem outras abordagens na literatura como modelos não determinísticos (modelagem estocástica e robusta), teoria dos conjuntos difusos, conjuntos neutrosóficos, entre outros. Contudo, o estudo e aplicação destas abordagens não faz parte o escopo deste trabalho e portanto, não serão detalhadas aqui. Os leitores interessados nestas abordagens podem encontrar mais informações em Leyva-Vázquez e Smarandache (2018) e Smarandache (2019).

A distribuição de socorro ou alívio compreende a alocação de itens físicos (pessoas, suprimentos, veículos, etc) ou não tais como informações. Pode ser classificado em três principais domínios: fluxo de rede, transporte e localização.

Geralmente as restrições associadas ao fluxo de rede consideram o limite de capacidade, as fontes de abastecimento sendo únicas ou múltiplas (*sourcing*), períodos (único ou multiperíodo), *commodity*, alocação de recursos entre outros.

Os problemas de transporte para distribuição de socorro adicionam mais restrições como: composição da frota, capacidade dos veículos, depósitos e modo de transporte.

Na sequência serão apresentados alguns estudos para modelos determinísticos ligados à distribuição de alívio/socorro. Leitores interessados em modelos não determinísticos podem encontrar mais informações em Mete e Zainsky (2010), Tofighi e Mansouri (2016), Ghasemi et al. (2020), Najafi et al. (2013) e Haghi et al. (2017).

Percebe-se que os problemas de distribuição de alívio algumas vezes se associam a problemas de roteamento e localização. Os autores Yi e Kumar (2007) aplicaram um modelo de fluxo de rede que minimiza o peso da soma

das demandas insatisfeitas, tanto de mercadorias como de pessoas feridas não atendidas. O modelo considerou um centro de informações que possibilita prever a demanda futura e assim realizar um balanço de fluxo de mercadorias e veículos nos nós de demanda.

Visando resolver o problema de distribuição de suprimentos para os necessitados, evacuar feridos para centros de triagem e transportar os trabalhadores humanitários para os locais necessários, os autores Al Theeb e Murray (2017) desenvolveram um modelo complexo em quatro níveis. Esse modelo utiliza um problema de roteamento de veículos com entregas divididas para agilizar a distribuição do socorro. Trata-se de um modelo multi-depósito, multi-período e multi-mercadoria para determinar a quantidade da mercadoria a recolher/ entregar em um nó, o número de trabalhadores a serem transferidos em um nó e o número de evacuados para pegar/ entregar em um nó.

Um Problema de Roteamento de Veículos (VRP) complexo aplicado à fase de resposta após um desastre natural foi desenvolvido por Penna e Santos (2018). Ele contempla a distribuição de última milha de bens de socorro após terremotos, é modelado como um VRP rico, envolvendo uma frota heterogênea de veículos, várias viagens, vários depósitos e dependências do veículo local. O método proposto é uma heurística híbrida genérica que utiliza uma formulação de particionamento para adicionar memória a uma estrutura de Busca Local Iterada com vários inícios. Para melhor atender aos requisitos de distribuição de última milha, o algoritmo foi avaliado em cenários reais do terremoto de Porto Príncipe. A heurística calcula rapidamente rotas eficientes ao determinar o número de veículos necessários e o subconjunto de depósitos a serem usados.

Recentemente Rabta et al. (2018) propuseram a utilização de drones (veículos aéreos não tripulados) para minimizar distância total (ou tempo/ custo). Os autores apontam em seu estudo que a utilização de tais meios de transporte vem recebendo maior atenção pelas organizações humanitária uma vez que podem superar problemas de inacessibilidade às regiões isoladas. O estudo propõe o transporte de itens de socorro leves, tais como vacinas, remédios, comprimidos de purificação de água, etc.

O estudo de Hezam et al. (2021) recentemente publicado e é utilizado como referência nesse trabalho, cita como oportunidade de pesquisa um foco maior nas fases de resposta e fases de preparação do desastre. O autor cita

também que raramente são vistos estudos com alocação multiperíodo. Devido à incerteza das demandas de socorro outro ponto apontado pelo autor é a transferência de itens de socorro ou estoque entre instalações para uma alocação mais eficiente desses recursos. Pensando em distribuição de alívio de itens de socorro os autores apontam que estão faltando estudos que contemplem a deterioração de itens de socorro ou itens de socorro de curta duração, reforçando a ideia de que o compartilhamento de itens entre centros de distribuição pode reduzir o potencial desperdício.

## CARACTERÍSTICAS E FORMULAÇÃO DO PROBLEMA

O problema tratara a distribuição de auxílio médico aos necessitados, buscando através da função objetivo minimizar o atraso global nos atendimentos de todos os pacientes. Conforme Hezam et al. (2021), esta área tem sido pouco considerada nos estudos científicos. Na revisão literária sobre LH desenvolvida em Hezam et al. (2021), apenas dois estudos são apontados para distribuição de socorro através de auxílio médico utilizando modelos determinísticos, que são Wang et al. (2018) e Lei et al. (2015). Duas particularidades de ambos os estudos chamam a atenção. Nenhum deles aborda as demandas por especialidades e conjunto médico disponível. Mediante ao que foi exposto, esse estudo busca contribuir com a literatura abordando estas duas lacunas. O presente trabalho considera que cada paciente tem demanda por uma única especialidade. Caso o paciente precise ser atendido por mais de uma especialidade ele deve ser inserido múltiplas vezes (uma para cada especialidade demandada). Todas as demandas devem ser atendidas. Considera-se que há uma central que possa concentrar as informações. Estas informações se referem à localização geográfica das demandas, suas necessidades médicas (em termos de especialidades e gravidade), a quantidade de médicos disponíveis e as especialidades (pediatria, ortopedia, etc) de cada médico (um médico pode possuir mais de uma especialidade), o tempo médio de atendimento por especialidade médica, o tempo limite para iniciar o atendimento de cada paciente (esse tempo indica a gravidade do paciente. Pacientes em situações críticas devem ter valores menores que pacientes em situações menos críticas), o tempo de deslocamento entre cada par de localidades, tempo de descanso médico entre rotas (este tempo pode variar em função de condições

individuais de cada médico como, por exemplo: idade, restrições físicas, cuidados especiais, etc) e tempo limite para a duração das rotas (estes valores serão usados para evitar a exaustão da equipe médica). Também se considera que existe um centro médico, onde os profissionais se preparam para atender as demandas (reúnem equipamento e suprimentos médicos) e que as rotas começam e terminam neste centro. De posse destas informações, a central determinará a rota a ser seguida por cada médico e repassar aos pontos de demanda os horários estimados para os atendimentos. A pesquisa construída pelos autores Kabcome e Mouktonglang (2015) foi utilizada como norteadora para a construção do modelo matemático proposto neste estudo. Abaixo as notações adotadas para a construção do modelo matemático:

**Conjuntos (índices):** Seja  $G(N,A)$  um grafo direcionado onde  $N$  é um conjunto de vértices e  $A$  é um conjunto de arcos  $(i, j)$  representando as conexões entre cada par de localidades (incluindo o centro médico), onde  $i \neq j$ . Seja:

$N = \{o, d, 1, 2, 3, \dots, n\}$  seja o conjunto de nós ( $o$  a origem e  $d$  o destino. No modelo ambos serão o mesmo local físico, que é o centro médico).

$N' = N \setminus \{o, d\}$ ,  $A = \{(i, j) : i, j \in N, i \neq j\}$ , para o conjunto de arcos.

$K = \{1, 2, 3, \dots, k\}$ , para todo conjunto de médicos.

$P = \{1, 2, 3, \dots, p\}$ , para todo conjunto de especialidade médica.

$R = \{1, 2, 3, \dots, r\}$ , para todo conjunto de rotas de cada médico.

#### Dados de entrada:

$B_i$  é o tempo limite para início do atendimento do paciente  $i$ ;

$T_{i,j}$  é o tempo de deslocamento entre as localidades onde estão os pacientes  $i$  e  $j$ ;

$S_i$  é o tempo de atendimento do paciente  $i$ ;

$C_k$  é o tempo necessário para descanso entre viagens do médico  $k$ ;

$Q_k$  é o limite de tempo para rotas do médico  $k$ ;

$$H_{kp} = \begin{cases} 1, & \text{se o médico } k \text{ possui habilitação na especialidade } p; \\ 0, & \text{caso contrário,} \end{cases}$$

$$D_{ip} = \begin{cases} 1, & \text{se o paciente } i \text{ demanda a especialidade } p; \\ 0, & \text{caso contrário,} \end{cases}$$

#### Variáveis de decisão:

$W_{k,r,i}$

$Y_{k,r,i}$  é o atraso no atendimento do paciente  $i$  na  $r$ -ésima viagem do médico  $k$ ;

$W_{k,r,i}$  é o tempo de início do atendimento do paciente  $i$  pelo médico  $k$  em sua  $r$ -ésima rota;

$X_{k,r,i,j} =$

$$\begin{cases} 1, & \text{se o paciente } j \text{ é atendido imediatamente após o paciente } i \text{ na } r\text{-ésima viagem do médico } k; \\ 0, & \text{caso contrário.} \end{cases}$$

**Modelo matemático:** O problema de distribuição de socorro médico por especialidade buscando atender diferentes localidades e suas demandas pode ser descrito matematicamente da seguinte forma:

$$\text{Min } Z = \sum_{i=1}^{N'} \sum_{k=1}^K \sum_{r=1}^R Y_{k,r,i} \quad (1)$$

s.a.

$$Y_{k,r,i} \geq W_{k,r,i} - B_i, \forall i \in N', \forall k \in K, \forall r \in R; \quad (2)$$

$$W_{k,r,i} + S_i + T_{i,j} - W_{k,r,j} \leq (1 - X_{k,r,i,j}) M, \forall k \in K, \forall r \in R, \forall i, j \in N, i \neq j; \quad (3)$$

$$W_{k(r+1),d} \geq W_{k,r,d} + C_k, \forall k \in K, \forall r \in \{1, \dots, (|R| - 1)\}; \quad (4)$$

$$W_{k,r,d} - W_{k,r,o} \leq Q_k, \forall k \in K, \forall r \in R; \quad (5)$$

$$W_{k,r,o} \leq W_{k,r,d}, \forall k \in K, \forall r \in R; \quad (6)$$

$$\sum_{j=1}^{N' \setminus \{d\}} X_{k,r,i,j} - \sum_{j=1}^{N' \setminus \{o\}} X_{k,r,i,j} = 0, \forall k \in K, \forall r \in R, \forall i \in N'; \quad (7)$$

$$\sum_{k=1}^K \sum_{r=1}^R \sum_{j=1}^{N' \setminus \{d\}} X_{k,r,i,j} = 1, \forall i \in N'; \quad (8)$$

$$\sum_{j=1}^{N' \setminus \{o\}} X_{k,r,o,j} \leq 1, \forall k \in K, \forall r \in R; \quad (9)$$

$$\sum_{j=1}^{N'} X_{k,r,o,j} = \sum_{j=1}^{N'} X_{k,r,j,d}, \forall k \in K, \forall r \in R; \quad (10)$$

$$X_{k,r,i,j} = 0, \forall i \in N', \forall p \in P, \forall k \in \{z : z \in K, D_{i,p} \neq 0\}, H_{k,p} = 0, \forall j \in N, \forall r \in R; \quad (11)$$

$$Y_{k,r,i}, W_{k,r,i}, X_{k,r,i,j} \geq 0 \text{ e } Y_{k,r,i}, W_{k,r,i}, X_{k,r,i,j} \in Z^*, \forall i, j \in N, \forall p \in P, \forall k \in K, \forall r \in R. \quad (12)$$

A Equação (1) representa a função objetivo do problema. Ela visa minimizar o somatório dos atrasos em todas as localidades, à medida que as restrições (2) permitirão que o problema calcule o atraso. No caso do valor de  $B$  ser maior que  $W$ , ou seja não há atraso, as restrições de não negatividade farão com que  $Y$  assumo zero. Por sua vez as restrições (3) certificam que, se o paciente  $j$  é atendido pelo médico  $k$  em sua  $r$ -ésima viagem imediatamente após o paciente  $i$ , então o tempo de início do atendimento do paciente  $j$  será igual ao tempo de início do atendimento do paciente  $i$  somado com o tempo de atendimento de  $i$  e com o tempo de deslocamento de  $i$  até  $j$  bem como as restrições (4) garantem que o médico  $k$  só poderá sair para uma nova viagem após o descanso obrigatório estipulado. As restrições (5) e (6) garantem que o tempo limite das rotas sejam respeitados o tempo de chegada ao destino seja maior (ou igual no caso de rota vazia) que o tempo de partida na origem. Em adição, as restrições (7) representam o equilíbrio de fluxo. Dizem que se um médico chegou a um paciente em uma viagem ele deve sair deste mesmo paciente na mesma viagem.

Por outro lado, as restrições (8) dizem que todos os pacientes devem ser visitados uma única vez, enquanto nas restrições (9) está limitando o número de partidas por rota a partir da origem. Já as restrições (10) certificam que se uma viagem foi iniciada, ela deve ser finalizada. Da mesma forma as restrições (11) garantirão que um paciente não seja atendido por um médico sem a especialidade requerida. Por fim, as restrições (12) são as restrições de não negatividade e integridade do problema.

Para testar os algoritmos construídos para o problema foram criadas instancias artificiais com dados de oferta e demanda. Contudo, para tornar os cenários mais realísticos, dados referentes às posições geográficas dos pacientes, centro médico, limites de tempo das rotas e o tempo de descanso entre rotas são reais (Melo, 2023). É válido ressaltar que as instâncias apresentadas no trabalho são maiores que encontradas na literatura. Há 4 possíveis tamanhos para o conjunto de pacientes: 5, 10, 15 e 20 pacientes. Cada paciente possui: 1) um tempo máximo para início de atendimento; 2) um tem-

po de atendimento, iguais aos usados por Lei et al. (2015); 3) uma posição geográfica (unidades básicas de saúde da cidade de Angra dos Reis, RJ-Brasil) e 4) uma demanda médica gerada de acordo com três distribuições estatísticas: Uniforme, Normal e Zipf. Os dados foram gerados usando Python (Numpy, 2022) e a ferramenta Google Colaboratory® (Borges, 2014). Também foram consideradas dois possíveis tamanhos para o conjunto de médicos: 4 e 6 médicos. Cada médico possui duas especialidades, escolhidas arbitrariamente para evitar a criação de instâncias inviáveis. Os tempos de descanso entre as rotas para cada médico, foram determinados com base em valores usados no dia a dia das equipes de resgate (entre 8 e 12 horas de descanso). O tempo máximo de cada rota também foi determinado com bases em dados reais (24 horas) e se estabeleceu que cada mé-

dico pode realizar, nas instâncias propostas, no máximo três rotas. Detalhes da criação das instâncias podem ser encontrados em Melo (2023). Forneceu-se ao software CPLEX® o tempo limite de execução de 10.800 segundos. O resolvidor de programação inteira foi executado em duas situações: gap (diferença relativa entre o melhor valor inteiro e o valor objetivo) igual a zero ou em casos em que não há memória suficiente no computador. Outra maneira e quando o usuário insere critérios de parada, tal como realizado neste estudo. A Tabela 1 apresenta uma síntese global de todos os resultados obtidos.

Os resultados apresentados na Tabela 1 demonstram que 54% das simulações alcançaram o gap igual a zero. Observa-se também que o resolvidor demandou um intervalo de tempo relativamente significativo para se adequar as condições propostas em grande parte das simulações.

Tabela 1: Síntese dos atrasos, gap do erro empírico e tempo de máquina para cada uma das distribuições de acordo com as instâncias.

4 médicos									
Localidades	Uniforme			Normal			Zipf		
	Atraso	Tempo máquina (segundos)	Gap do software	Atraso	Tempo máquina (segundos)	Gap do software	Atraso	Tempo máquina (segundos)	Gap do software
5	87	0,06	0,00%	110	0,05	0,00%	140	0,19	0,00%
10	288	3,25	0,00%	288	3,55	0,00%	382	26,08	0,00%
15	742	10807,19	71,73%	834	10257,56	78,11%	903	10803,48	84,34%
20	1678	10806,72	98,62%	1761	10804,19	99,80%	1795	10817,33	91,42%
6 médicos									
Localidades	Uniforme			Normal			Zipf		
	Atraso	Tempo máquina (segundos)	Gap do software	Atraso	Tempo máquina (segundos)	Gap do software	Atraso	Tempo máquina (segundos)	Gap do software
5	87	0,09	0,00%	87	0,13	0,00%	126	0,05	0,00%
10	193	2,48	0,00%	170	1,44	0,00%	242	5,45	0,00%
15	411	445,42	0,00%	675	10811,58	57,17%	1252	10811,91	93,46%
20	932	10807,90	74,69%	1636	10801,83	94,91%	2796	10812,44	99,11%

Fonte: Elaborado pelos autores.

## CONCLUSÕES

Um ambiente de desastre gera cenários de catástrofes, desordens e vulnerabilidade nas diversas esferas da sociedade ao longo dos séculos. Muitos são aqueles que nesse ambiente se sensibilizam e unem forças através da cooperação para minimizar os danos causados. A Logística humanitária de desastre é uma ferramenta eficaz nesse momento. Através dela a coordenação e gestão adequada podem ser implementadas, integrando organizações, ONG's, institui-

ções, voluntários e outros no principal objetivo a ser buscado, o auxílio adequado e eficaz aos necessitados e o restabelecimento da ordem.

Nos últimos dezessete anos houve um crescente interesse dos cientistas para o tema, todavia muitas lacunas ainda existem a serem estudadas. Apresentar essas lacunas com clareza é um dos objetivos desse trabalho. Nele são apresentados alguns dos principais estudos sobre o tema, jun-

tamente com as lacunas citadas por estes trabalhos. Algumas destas lacunas foram abordadas neste trabalho, mas ainda existem outras que devem ser analisadas em trabalhos futuros ou por outros pesquisadores. Este estudo também buscou apresentar o problema de distribuição de socorro médico e suas principais variáveis.

O modelo matemático proposto foi amplamente testado, abordando diferentes instâncias (demandas, disponibilidade de especialidades, médicos especialistas, etc). Os resultados apresentados na Tabela 1 indicam que quanto maior a quantidade de pacientes, maior a dificuldade do resolvidor para encontrar a solução ótima. Também é possível notar que o resolvidor apresenta muita dificuldade para solucionar as maiores instâncias visto que, mesmo após o limite de tempo, o melhor valor para o gap foi maior que 50%.

Com isto conclui-se que a abordagem alcançou soluções em um tempo computacional relativamente alto para as condições impostas para cenários de desastre. Com base nos resultados obtidos nesta pesquisa, é possível concluir também que é necessário considerar propostas que exijam menos tempo computacional. Todos os valores alcançados pelo modelo proposto estão disponíveis como pontos de referência para comparações.

O presente trabalho possui oportunidades de desenvolvimento para trabalhos futuros. Iniciando pela construção do código. A ele podem ser agregados itens fortemente abordados

pela literatura como por exemplo a distribuição de diferentes suprimentos como agulhas, luvas, medicamentos, alimentos, entre outros. Como eles são fabricados por diferentes fábricas e são entregues aos centros de distribuição em tempos distintos, podem ter suas entregas gerenciadas de maneira integrada com o suporte médico. Muitos estudos tratam esses temas como um único bloco, sem olhar para a heterogeneidade existente. A análise multiperíodo das entregas também se faz importante. A sugestão de integrar tais informações ao problema se dá pela oportunidade de gerar uma ferramenta robusta para auxiliar decisores em momentos de desastre. Uma oportunidade a ser abordada é a validação por dados realísticos de qual distribuição de probabilidade as demandas por especialidade se comportam. Neste trabalho foram apresentadas três distribuições e os cenários encontrados mediante a utilização de cada uma delas. Ao abordar o tema por essa ótica neste trabalho espera-se que pesquisadores busquem aprofundar-se ainda mais sobre o tema, buscando as distribuições, bem como os devidos parâmetros, que mais se assemelham aos casos reais.

Finalmente sugere-se que os próximos pesquisadores busquem contemplar nas propostas de solução maior abrangência das fases do desastre, quem sabe até a integração de todas as fases fornecendo assim uma ferramenta única e aproximando a proposta das realidades enfrentadas pelos gestores de desastre.

## REFERÊNCIAS

- Al Theeb, N., Murray, C. (2017). Vehicle routing and resource distribution in post disaster humanitarian relief operations. *International Transactions in Operational Research* 24, 6, 1253–1284.
- Borges, L. E. (2014). *Python para desenvolvedores: aborda Python 3.3*. Novatec Editora.
- CRED (2021). Database de desastres naturais. <https://www.www.emdat.be>. Acessado em 13 novembro de 2021.
- de Oliveira, E. F. e Soares, E. B. (2019). Logística humanitária: o desafio da gestão diante de desastres. *Brazilian Journal of Business*, 1(3):870–880.
- de S.Paulo, J. F. (2022). Cinco pessoas morrem em carro soterrado em Brumadinho (MG). <https://www1.folha.uol.com.br/cotidiano/2022/01/cincopeessoas-morrem-em-carro-soterrado-em-brumadinho-mg.shtml>. Acessado em 06 agosto de 2022.
- Ghasemi, P., Khalili-Damghani, K., Hafezalkotob, A., Raissi, S. (2020). Stochastic optimization model for distribution and evacuation planning (a case study of tehran earthquake). *Socio-Economic Planning Sciences*, 71, 100745.
- Haghi, M., Ghomi, S. M. T. F., Jolai, F. (2017). Developing a robust multi-objective model for pre/post disaster times under uncertainty in demand and resource. *Journal of Cleaner Production*, 154, 188–202.
- Hezam, I. M. et al. (2021). A systematic literature review on mathematical models of humanitarian logistics. *Symmetry*, 13(1):11.
- Janone, L. (2022). A cada desastre natural no brasil, em média, 3,4 mil pessoas são afetadas. CNN Brasil. Acessado em 06 agosto de 2022.
- Kabcome, P. e Mouktonglang, T. (2015). Vehicle routing problem for multiple product types, compartments, and trips with soft time windows. *International Journal of Mathematics and Mathematical Sciences*, 2015.
- Lei, L., Pinedo, M., Qi, L., Wang, S., e Yang, J. (2015). Personnel scheduling and supplies provisioning in emergency relief operations. *Annals of Operations Research*, 235(1):487–515.
- Leyva-Vázquez, M., Smarandache, F. (2018). Computación neutrosófica mediante Sympy. *Infinite Study*.
- Melo, R.M. (2023). Modelo Heurístico e Matemático para Roteamento de Médicos por Especialidade em Situação de Pós-Desastre. 2023. 95 f. Trabalho de Conclusão de Mestrado – Universidade Federal Fluminense, UFF, Rio de Janeiro.
- Mete, H. O., Zabinsky, Z. B. (2010). Stochastic optimization of medical supply location and distribution in disaster management. *International Journal of Production Economics* 126, 1, 76–84.
- Najafi, M., Eshghi, K., Dullaert, W. (2013). A multi-objective robust optimization model for logistics planning in the earthquake response phase. *Transportation Research Part E: Logistics and Transportation Review* 49, 1, 217–249.
- Nogueira, C., Gonçalves, M. (2009). A logística humanitária: apontamentos e a perspectiva da cadeia de assistência humanitária. Em *Anais do XXVIII Congresso de Pesquisa e Ensino em Transportes, ANPET, Vitória*.
- Numpy (2022). Random Generator Python. <https://numpy.org/doc/stable/reference/random/generator.html>. 2022.
- Penna, P. H. V., Santos, A. C., Prins, C. (2018). Vehicle routing problems for last mile distribution after major disaster. *Journal of the Operational Research Society* 69, 8, 1254–1268.
- Rabta, B., Wankmüller, C., Reiner, G. (2018). A drone fleet model for last-mile distribution in disaster relief operations. *International Journal of Disaster Risk Reduction* 28, 107–112.
- Smarandache, F. (2019). Grado de dependencia e independencia de los (sub) componentes de conjuntos borrosos y neutrosóficos. *Neutrosophic Computing and Machine Learning* 5, 6.
- Thomas, A. (2004). Elevating humanitarian logistics. *International Aid & Trade Review*, 102–106.
- Tofghi, S., Torabi, S. A., Mansouri, S. A. (2016). Humanitarian logistics network design under mixed uncertainty. *European Journal of Operational Research* 250, 1, 239–250.
- Wang, S., Liu, F., Lian, L., Hong, Y., e Chen, H. (2018). Integrated post disaster medical assistance team scheduling and relief supply distribution. *The International Journal of Logistics Management*.
- Yi, W., Kumar, A. (2007). Ant colony optimization for disaster relief operations. *Transportation Research Part E: Logistics and Transportation Review* 43, 6, 660–672

## DADOS DOS AUTORES

**Raila de Medeiros Melo**, mestre em Modelagem Computacional Ciência e Tecnologia (Escola de Engenharia Industrial Metalúrgica de Volta Redonda) da Universidade Federal Fluminense (UFF). [raila\\_medeiros@id.uff.br](mailto:raila_medeiros@id.uff.br).

**Eliane da Silva Christo**, doutora em Engenharia Elétrica, docente associada no Departamento de Engenharia de Produção (Escola de Engenharia de Petrópolis) da Universidade Federal Fluminense (UFF). [elianechristo@id.uff.br](mailto:elianechristo@id.uff.br).

**Tiago Araújo Neves**, doutor em Computação docente associado Departamento de Ciências Exatas (Escola de Engenharia Industrial Metalúrgica de Volta Redonda) da Universidade Federal Fluminense (UFF). [tneves@id.uff.br](mailto:tneves@id.uff.br).



# SEGMENTAÇÃO DE IMAGENS DE RESSONÂNCIA MAGNÉTICA DE TUMOR CEREBRAL UTILIZANDO MODELO 3D U-NET

## SEGMENTATION OF BRAIN TUMOR MAGNETIC RESONANCE IMAGES USING 3D U-NET MODEL

Erika Rocha de Araujo

Vanessa da Silva Garcia

Wesley Luiz da Silva Assis

**Resumo:** A evolução das análises por meio de técnicas de Imagens de Ressonância Magnética (IRM) desempenhou um papel crucial na medicina diagnóstica, promovendo avanços substanciais. Este estudo visa capacitar um modelo 3D-Unet com um conjunto de imagens de IRM, focalizando a detecção de tumores cerebrais, a geração de previsões e a avaliação de seus resultados. Para atingir esse propósito, são empregadas imagens de IRM no formato NIFTI, provenientes da competição BraTS2020, que compreende um conjunto de 369 grupos de imagens, cada grupo contendo 5 imagens. Essas imagens são fundamentais para o treinamento do modelo em consideração. As imagens resultantes oferecem uma representação aprimorada do tumor, destacando sua localização específica no cérebro. O modelo é iterativamente treinado, e a escolha da melhor iteração é fundamentada na avaliação de métricas de desempenho. Posteriormente, o modelo otimizado é empregado na geração de previsões, as quais demonstram uma correspondência notável com as máscaras reais dos tumores.

**Palavras-chave:** Segmentação de IRM. Modelo 3D U-NET . Imagens de Ressonância Magnética. Identificação de tumor cerebral

**Abstract:** The evolution of analyses through Magnetic Resonance Imaging (MRI) techniques has played a crucial role in diagnostic medicine, promoting substantial advancements. This study aims to train a 3D-Unet model with a set of MRI images, focusing on the detection of brain tumors, the generation of predictions, and the evaluation of their results. To achieve this purpose, MRI images in NIFTI format are used, originating from the BraTS2020 Competition, comprising a set of 369 image groups, each group containing 5 images. These images are essential for the training of the considered model. The resulting visualizations provide an enhanced representation of the tumor, highlighting its specific location in the brain. The model is iteratively trained, and the choice of the best iteration is based on the evaluation of performance metrics. Subsequently, the optimized model is employed in generating predictions, which demonstrate a notable correspondence with the actual tumor masks.

**Keywords:** MRI segmentation. 3D U-NET model. MRI. Brain tumor identification.

## INTRODUÇÃO

A Ressonância Magnética (RM), é uma técnica pioneira na medicina diagnóstica, que oferece imagens detalhadas dos tecidos moles do corpo humano. Na vanguarda dessa revolução, os avanços na visão computacional e na aprendizagem profunda de máquinas levaram à criação do modelo 3D U-Net (KRIZHEVSKY, A. et al. 2012), que integra a inteligência artificial a medicina diagnóstica. Por meio dessa tecnologia, é possível acelerar o processo de detecção de tumores e melhorar a precisão das análises, trazendo benefícios significativos aos pacientes.

A arquitetura do 3D U-Net consiste em uma rede neural profunda com uma estrutura em “U”. Ela possui uma seção de codificação (encoder) que captura características das imagens em várias escalas e uma seção de decodificação (decoder) que reconstrói a segmentação em alta resolução (ÇIÇEK, et al. 2016). O modelo utiliza convoluções tridimensionais e camadas de pooling tridimensionais para processar volumes de imagem completos.

Este artigo tem por objetivo realizar um estudo inicial para desenvolver uma aplicação de um conjunto de imagens de Ressonância Magnética (IRM) ao modelo 3D U-Net para detecção mais rápida e precisa de tumores cerebrais nas IRM e avaliar seus resultados. Para tanto, utiliza-se a biblioteca NiBabel para manipular os arquivos de imagens de IRM na linguagem python. Tais imagens estão no formato Neuroimaging Informatics Technology Initiative (NIFTI), que é utilizado para armazenar e compartilhar dados de imagens médicas no campo da neuroimagem. Estes dados são da Competição BraTS2020, um total de 369 grupos de imagens com 5 imagens do mesmo cérebro com imagens de T1, T1ce, T2, Flair e uma última sendo uma imagem do tumor segmentado para o treinamento do modelo. As imagens são públicas e utilizadas em campeonatos de pesquisas para as equipes participantes desenvolverem seus estudos contribuindo assim com a evolução do conhecimento científico.

## FUNDAMENTAÇÃO TEÓRICA:

### IRM (Imagem de Ressonância Magnética)

A Imagem de Ressonância Magnética é uma técnica de imagem médica não invasiva que utiliza campos magnéticos e ondas de rádio para gerar imagens detalhadas do corpo humano.

Segundo KNOLL, F. et al. (2019), durante a imagem, uma sequência de campos magnéticos variáveis, espacial e temporalmente, chamada “sequência de pulsos”, é aplicada pela máquina de RM. Isso induz o corpo a emitir campos de resposta eletromagnética ressonantes, que são medidos pela bobina receptora. As medições geralmente correspondem a pontos ao longo de um caminho prescrito através da representação de Fourier multidimensional do corpo discreto representado por imagens.

### Formato DICOM

O Formato DICOM (Digital Imaging and Communications in Medicine) é um padrão amplamente adotado para aquisição, armazenamento, troca e visualização de imagens médicas. Ele foi desenvolvido para garantir a interoperabilidade e a compatibilidade entre diferentes sistemas e dispositivos médicos, permitindo a comunicação eficiente e confiável de informações médicas (DICOM, 2023).

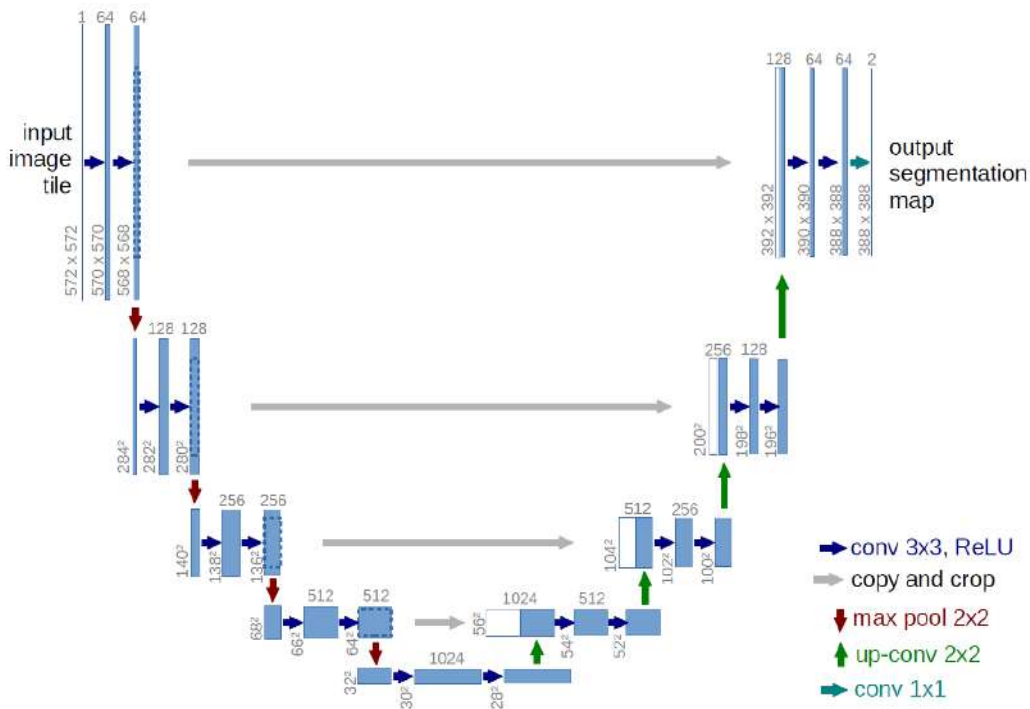
### NIFTI

O formato NIFTI consiste em um arquivo binário que contém informações sobre a aquisição e as características da imagem, juntamente com os próprios dados de imagem. Uma das principais vantagens do formato NIFTI é a capacidade de armazenar dados em três dimensões (3D) ou quatro dimensões (4D) para imagens dinâmicas, como sequências temporais (LARBINA, Michele; MURINO, Loredana, 2013). Ele permite que os pesquisadores acessem e analisem facilmente os dados de imagem usando várias ferramentas de processamento e análise.

A imagem é de padrão no formato DICOM e é pré-processada para o formato NIFTI para ser trabalhada. No caso deste trabalho, as imagens já estão no formato NIFTI.

### 3D UNET

Figura 1 - 3D U-NET



Fonte: (RONNEBERGER, et al. 2015).

Este modelo, inicialmente concebido para segmentação de imagens 2D, foi expandido para trabalhar em três dimensões, tornando-o ideal para processar volumes completos de imagem, como os obtidos em exames de IRM (RONNEBERGER, et al. 2015).

A U-Net 3D, arquitetura representada na Figura 1, mantém a estrutura codificador-decodificador da sua contraparte 2D, porém, opera em três dimensões. Isso significa que ela analisamos informações tridimensionais em vez de apenas fatias bidimensionais das imagens. Essa capacidade é fundamental para a detecção precisa de estruturas complexas em imagens volumétricas, como as obtidas em exames de IRM cerebral. Uma característica notável da U-Net é a utilização de conexões residuais, que permitem a preservação de detalhes de alta resolução durante o processo de decodificação (RONNEBERGER, et al. 2015).

Para compreender o funcionamento da U-Net 3D, é crucial entender o conceito de “encoder” e “decoder”. O “encoder” inicialmente reduz as dimensões da imagem e extrai características importantes, enquanto o “decoder” reverte esse processo, aumentando as dimensões e restaurando a resolução, permitindo a segmentação precisa. As conexões residuais mencionadas anteriormente permitem que informações detalhadas, que normalmente seriam

perdidas durante a redução de dimensões, sejam reintegradas no processo de decodificação (ÇIÇEK, et al. 2016).

A aplicação mais comum da U-Net 3D é a segmentação de estruturas médicas, como tumores cerebrais em imagens de IRM. Sua habilidade de aprender a forma tridimensional das estruturas torna-a inestimável para identificar regiões de interesse em imagens médicas complexas (ÇIÇEK, et al. 2016).

#### ETAPAS DO MÉTODO:

##### Pré-processamento dos dados

Para serem utilizadas no modelo 3D U-Net as imagens necessitam ainda de outros pré-processamentos.

- Gerar subvolumes: Nesta etapa é necessário gerar “patches” dos dados. Devido a necessidade de grande processamento, realizou-se a divisão em sub-volumes na região do tumor para ser possível extrair dados dessa região.
- Padronizar os dados: Para facilitar o aprendizado da rede neural artificial aplica-se a normalização para padronizar a imagem (média 0 e desvio padrão 1).

## Métricas

1) O Coeficiente Dice (Dice Similarity Coefficient - DSC)

É essencialmente uma medida de sobreposição entre duas amostras.

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (1)$$

A Equação (1) é uma primeira avaliação para calcular o Coeficiente Dice de uma única camada, porém como são 4 tipos imagens referentes ao mesmo cérebro: T1, T1ce, T2 e Flair, pode-se utilizar a média entre elas.

$$DSC(f, x, y) = \frac{1}{4} (DSC1(f, x, y) + DSC2(f, x, y) + DSC3(f, x, y) + DSC4(f, x, y)) \quad (2)$$

A Equação (2) realiza a média, onde o índice varia de 0 (incompatibilidade completa) para 1 (correspondência perfeita).

O Coeficiente Dice (DSC) é uma métrica comum usada na avaliação de segmentações médicas. Ela mede a sobreposição entre a máscara segmentada predita pela rede neural e a máscara de referência, segundo DICE, L. R. (1945). Quanto maior a sobreposição, melhor a segmentação. No entanto, o Coeficiente Dice (DSC) padrão é rígido e não leva em consideração a incerteza nas previsões. Isso pode ser problemático, pois as bordas das estruturas podem ser difíceis de definir nas imagens médicas.

## 2) Perda Entropia Cruzada Categórica

A função Perda Entropia Cruzada Categórica compara a distribuição de probabilidade prevista pelo modelo com a distribuição de probabilidade real (rótulo). Quanto mais próximas dessas distribuições, menor será a Perda (Chollet, 2020).

$$L(y, p) = \sum_{i=1}^C y_i \cdot \log(p_i) \quad (3)$$

Na Equação (3) temos:

- $L(y, p)$  é a entropia cruzada categórica.
- $y_i$   $y_i$  representa o valor real do rótulo da classe  $i$  (1 se a amostra pertence a essa classe, 0 caso contrário).
- $p_i$  é a probabilidade prevista pelo modelo para a classe.
- $C$  é o número total de classes.

## 3) Índice de Jaccard

O Índice de Jaccard, também conhecido como Coeficiente de Similaridade de Jaccard, é uma métrica fundamental para avaliar a sobreposição ou similaridade entre dois conjuntos (Jaccard, 1901).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (4)$$

A Equação (4)  $A$  e  $B$  são os conjuntos a serem comparados, e  $|A \cap B|$  e  $|A \cup B|$  representam, respectivamente, o tamanho da interseção e da união dos conjuntos.

#### 4) Acurácia

Esta métrica fornece uma medida direta da precisão geral do modelo em fazer previsões corretas.

$$\text{Acurácia} = \frac{\text{Número de previsões corretas}}{\text{Número total de amostras}} \quad (5)$$

Na Equação (5) a acurácia varia de 0 a 1, onde 1 indica uma precisão perfeita e 0 indica desempenho nulo.

### FLUXOGRAMA DAS ETAPAS DA PESQUISA :

Figura 2 – Fluxograma -



Fonte: Imagem gerada pelos autores

### CRIANDO E TREINANDO O MODELO 3D U-NET

O modelo 3D U-Net é treinado com pares de imagens e rótulos. Os rótulos são máscaras que indicam a localização dos tumores nas imagens. Durante o treinamento, o modelo aprende a mapear os patches de imagem para as máscaras de segmentação correspondentes.

Neste estudo é adotado a separação de três subconjuntos: de dados para treinamento, validação e teste deste modelo, separados 20% para validação, depois 15% para testes e o restante para treinamento.

### Avaliação do Treinamento

Para avaliação do desempenho dos resultados do treinamento é importante para escolha do melhor treinamento e será analisado:

- Função de Perda: avaliar como ele está se ajustando aos dados de treinamento
- Acurácia: verificar a proporção de previsões corretas em relação ao número total de previsões.
- Coeficiente Dice (DSC): avaliar a sobreposição entre a máscara prevista e a máscara de referência.
- Índice de Jaccard: verificar a sobreposição entre a máscara prevista e a máscara de referência para cada classe (ou região) de interesse.

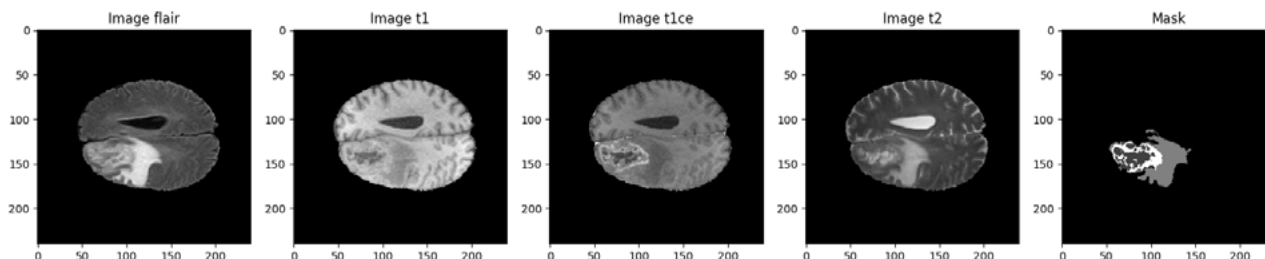
## RESULTADOS

Os resultados apresentados a seguir são referentes ao processamento de imagens de ressonância magnética de tumores cerebrais pré-processadas no formato NIFTI, disponíveis pela Competição BraTS2020, um total de 369 grupos de imagens com 5 imagens do mesmo cérebro com imagens de T1, T1c, T2, Flair e uma última sendo uma imagem do tumor segmentado para o treinamento do modelo.

As visualizações a seguir foram geradas a partir dos arquivos de treinamento do grupo de imagens BraTS20\_Training\_006.

Conhecendo os dados criando visualizações e examinando:

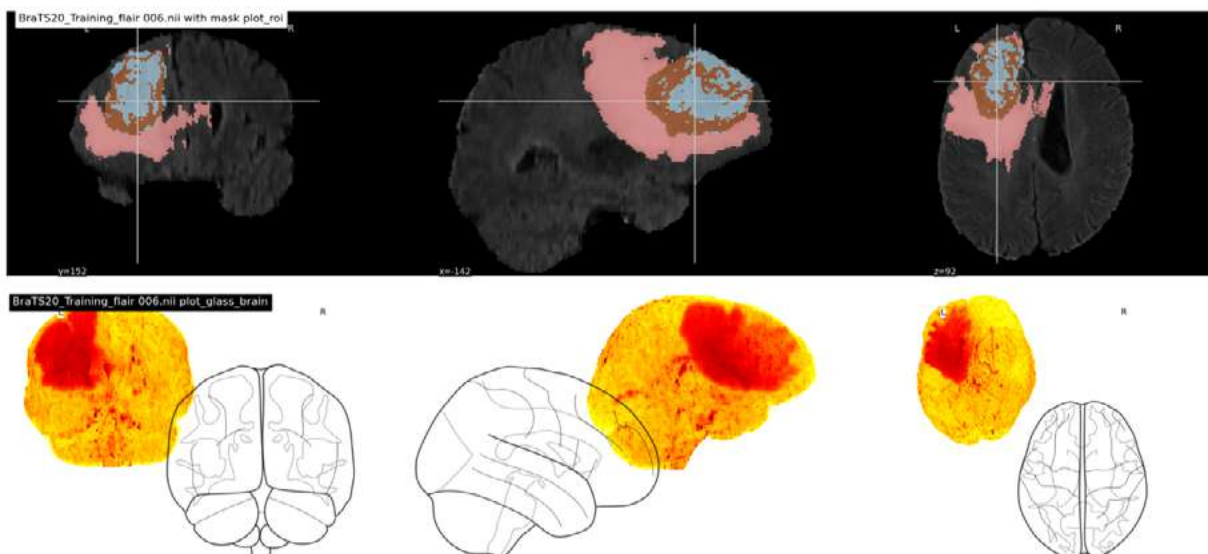
Figura 3 – 5 tipos de imagens do mesmo cérebro



Fonte: Imagem gerada pelos autores

A Figura 3 apresenta as imagens geradas do grupo de imagens BraTS20\_Training\_006 onde podemos observar 5 tipos de imagens referentes ao mesmo cérebro com a localização tumoral através da imagem chamada Mask. Foi utilizada a biblioteca NiBabel para realizar a interação com esse formato de dados NIFTI e a visualização das 5 tipos da esquerda para a direita: Flair, T1, T1ce, T2 e Mask.

Figura 4 - Imagens de ângulos diferentes



Fonte: Imagem gerada pelos autores

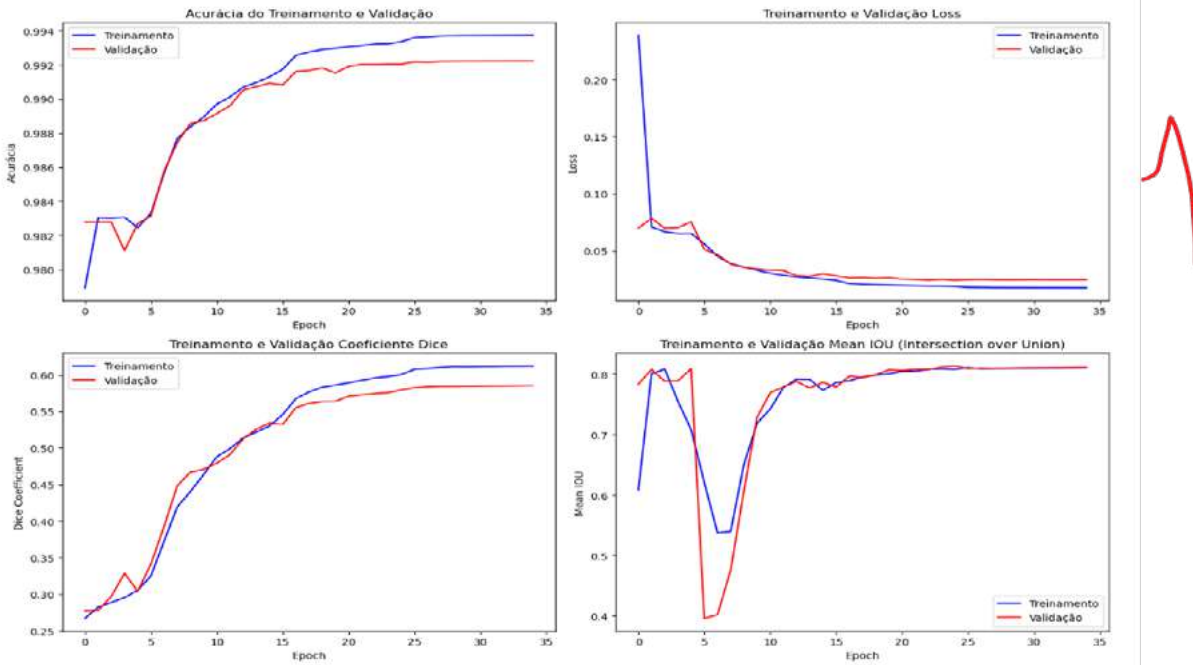
Na Figura 4, com o objetivo de explorar as imagens obtidas, utilizamos a biblioteca Nilearn que possui um conjunto de funções de plotagem para gerar visualizações de volumes cerebrais. No caso da Figura 4, aplicamos plot\_glass\_brain e plot\_roi, que gerou imagens destacando por meio de cores diferentes regiões de provável interesse já que as funções plot\_glass\_brain e plot\_roi não tem o objetivo de identificação tumoral. Pode-se observar nas imagens que é possível “girar” o ângulo para análise de outras perspectivas, aqui apresentadas, da esquerda para direita, como: Coronal, Sagital e Transversal.

## TREINANDO O MODELO 3D U-NET:

Foi realizado o treinamento do modelo consecutivas vezes com o objetivo de alcançar o modelo com melhor desempenho.

Foi observado que epoch=35 trazia resultados relevantes, então foi adotado como parâmetro nos demais treinamentos em busca de um bom desempenho.

Figura 5 – Métricas de desempenho do modelo



Fonte: Imagem gerada pelos autores

Na Figura 5 temos as métricas como o resultado para avaliação do desempenho deste modelo 3D U-Net treinado. Este resultado foi obtido de um modelo que apresentou melhor desempenho avaliando as métricas.

A avaliação do modelo revelou métricas de desempenho bastante relevantes. A acurácia, tanto no treinamento (0.9937) quanto na validação (0.9922), indica que o modelo está fazendo previsões precisas em ambos os conjuntos de dados. No entanto, é importante lembrar que uma alta acurácia não é garantia de um modelo excelente, e, portanto, a avaliação deve ser complementada com outras métricas.

O índice de Jaccard (IoU), que mede a sobreposição entre as máscaras previstas e as verdadeiras, também revelou resultados relevantes. Com um valor de 0.8101 de treinamento e 0.8107 na validação, o modelo demonstrou uma capacidade sólida de segmentação de objetos nas imagens.

Além disso, o coeficiente de Dice (Dice Coef), obteve no treinamento 0.6123. Embora ligeiramente mais baixo na validação (0.5856), ainda é considerado razoável. Isso sugere que o

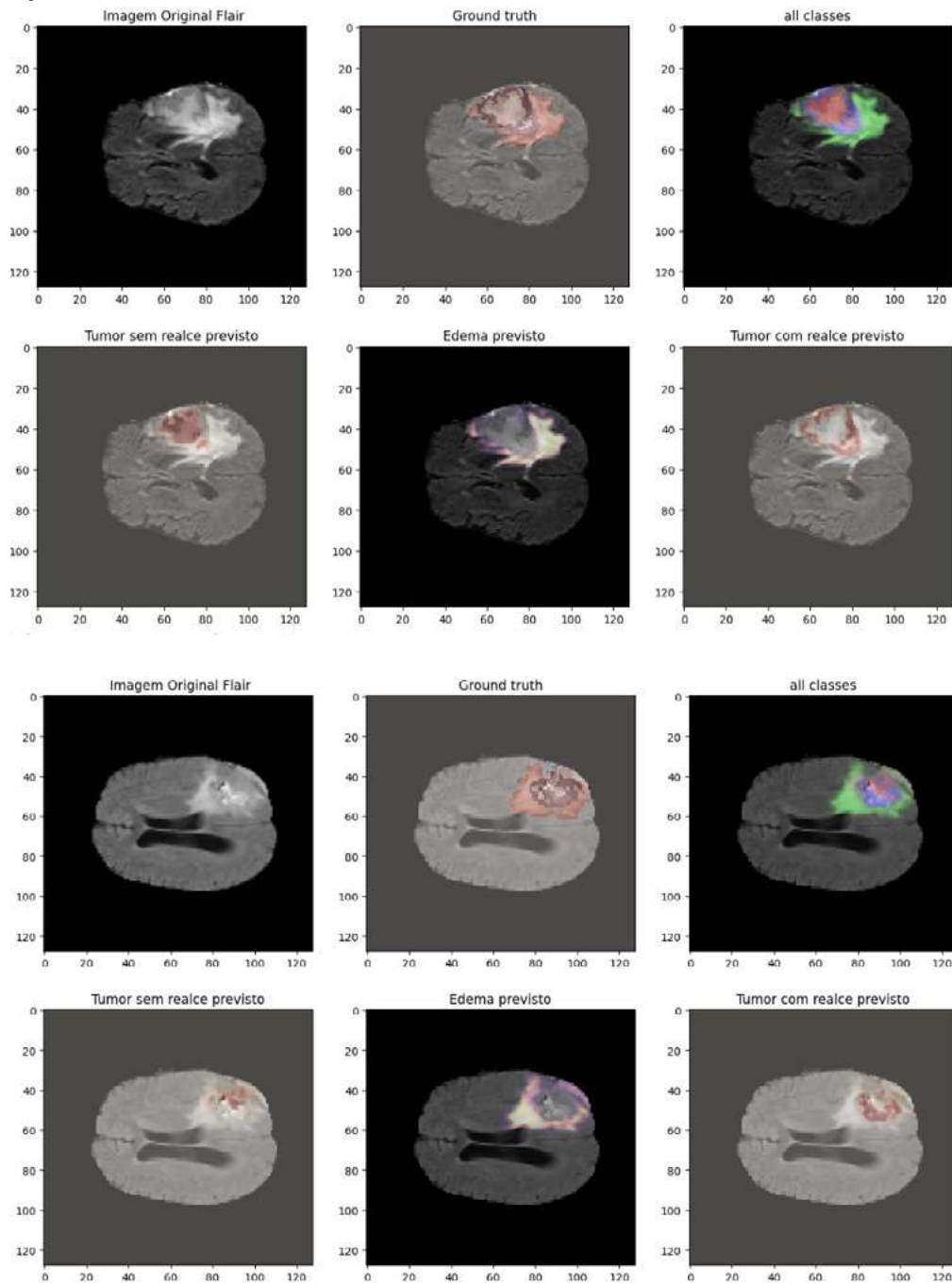
modelo está conseguindo uma boa concordância entre as máscaras segmentadas e as máscaras verdadeiras.

Em relação à perda (Loss), os resultados foram igualmente promissores. Com uma baixa perda de treinamento (0.0175), indica que o modelo está se ajustando bem aos dados. A ligeira diferença entre a perda de treinamento e a perda de validação (0.0245) não parece indicar super ajuste (overfitting).

### Realizando Previsões com o Modelo:

Foram utilizadas imagens de ressonância magnética Flair e T1ce, as quais foram redimensionadas e enviadas para o modelo previamente treinado para fazer previsões. O resultado é uma matriz de probabilidades para cada pixel na imagem. Em seguida são exibidas uma série de imagens em uma única figura, incluindo a imagem original, a máscara ground truth e as previsões para as diferentes classes. Isso permite visualizar como o modelo está se saindo na segmentação em comparação com a máscara ground truth.

Figura 6 – Previsões do modelo 3D U-Net -



Fonte: Imagem gerada pelos autores

Na figura 6 podemos observar as previsões do modelo 3D U-Net treinado para as imagens de 2 grupos, somente as imagens Flair e T1ce desses grupos de imagem, como saída temos a imagem original Flair, a máscara ground truth, uma contendo todas as classes, o tumor com realce (necrotic/core) previsto, o edema previsto e o tumor sem realce (enhancing) previsto.



## CONCLUSÕES

Este estudo buscou desenvolver uma aplicação inicial que emprega um conjunto de imagens de Ressonância Magnética no modelo 3D U-Net. O foco foi proporcionar uma detecção mais ágil e precisa de tumores cerebrais, com uma subsequente avaliação dos resultados obtidos.

As previsões geradas pelo modelo revelaram-se relevantes, conforme observado na comparação com as máscaras ground truth, evidenciada pelas métricas e imagens apresentadas na

Figura 6. Entretanto, é importante destacar que há margem para otimizações, tanto em termos de eficiência computacional quanto no que diz respeito ao coeficiente Dice, cujo aprimoramento pode ser alcançado por meio de ajustes no código.

Em resumo, os resultados obtidos ressaltam a eficácia do modelo, indicando sua capacidade notável na detecção de tumores cerebrais. No entanto, também apontam para oportunidades significativas de melhoria, destacando a importância de futuras otimizações visando aprimorar ainda mais o desempenho e a precisão da aplicação.

## Referências

- ABDULKADIR, A.; BROX, T.; LIENKAMP, S. S.; RONNEBERGER, O.; ÇIÇEK, Ö. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2016, pp. 424-432.
- CHOLLET, F. e. a. Keras Documentation. 2020. Disponível em: <https://keras.io/api/losses/probabilisticlosses/categoricalcrossentropy/>. Acesso em 28 de maio de 2023.
- DICOM Standard. Disponível em: <https://www.dicomstandard.org/>. Acesso em 28 de maio de 2023.
- DICE, L. R. Measures of the amount of ecologic association between species. In: Ecology, vol. 26, no. 3, 1945, pp. 297-302.
- JACCARD, P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Imprimerie Corbaz, 1901.
- KNOLL, F. et al. Fastmri: An open dataset and benchmarks for accelerated mri, 2019.
- KRIZHEVSKY, A. et al. ImageNet Classification with Deep Convolutional Neural Networks. In: Advances in Neural Information Processing Systems, 2012.
- LAROBINA, Michele; MURINO, Loredana. Medical Image File Formats. PMC – PubMed Central, 2013.
- RONNEBERGER, Olaf; BROX, Thomas; FISCHER, Philipp. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015.

## DADOS DOS AUTORES

**Erika Rocha de Araujo.** Bacharel em Ciência da Computação e Especialista em Big Data e Inteligência Competitiva. Mestranda em Modelagem Computacional Ciência e Tecnologia. Universidade Federal Fluminense ([erika.uerj@gmail.com](mailto:erika.uerj@gmail.com))

**Vanessa da Silva Garcia.** Doutora em Engenharia Nuclear. Docente do Departamento de Administração e Administração Pública. Universidade Federal Fluminense ([vanessagarcia@id.uff.br](mailto:vanessagarcia@id.uff.br))

## O -ESPECTRO DO GRAFO BLOW-UP

## THE -SPECTRUM OF THE BLOW-UP GRAPH

Uilton Cesar Peres Junior

Carla Silva Oliveira

André Ebling Brondani

**RESUMO:** Seja  $G$  um grafo conexo de ordem  $n$ ,  $A(G)$  sua matriz de adjacência e  $D(G)$  sua matriz diagonal dos graus. Em 2017, Nikiforov propôs uma combinação linear convexa que generaliza as matrizes  $A(G)$  e  $D(G)$  definida da seguinte maneira:

Neste artigo, exploramos o espectro da matriz  $A_\alpha$  em relação a operação *blow-up* de um grafo, onde exibimos seu polinômio característico, o seu espectro e tratamos do problema que refere-se à obtenção do menor valor de  $\alpha$  para o qual tal matriz é positiva semidefinida. **Palavras-chave:** Matriz  $A_\alpha$ .  $A_\alpha$ -espectro. Operação blow-up.

**Palavras-chave:** Matriz  $A_\alpha$ .  $A_\alpha$ -espectro. Operação blow-up.

**ABSTRACT:** Let  $G$  be a connected graph of order  $n$ ,  $A(G)$  its adjacency matrix and  $D(G)$  its diagonal matrix of degrees. In 2017, Nikiforov proposed a convex linear combination that generalizes the matrices  $A(G)$  and  $D(G)$  defined as follows:

In this article, we explore the spectrum of the matrix  $A_\alpha$  about the blow-up operation of a graph, where we display its characteristic polynomial and its spectrum, and deal with the problem that refers to obtaining the smallest value of  $\alpha$  for which such matrix is positive semidefinite.

**Keywords:**  $A_\alpha$  matrix.  $A_\alpha$ -spectrum. Blow-up operation.

## INTRODUÇÃO

Seja  $G = (V(G), E(G))$  um grafo simples de ordem  $n = |V(G)|$ . Quando  $\{v_i, v_j\} \in E$ ,  $v_i$  e  $v_j$  são ditos adjacentes. O grafo complementar de  $G$  é o grafo  $\bar{G} = (V, \bar{E})$  em que  $V = V(G)$  e se e somente se  $\{v_i, v_j\} \notin E$ . O conjunto de vizinhos de um vértice  $v$  é denotado por  $N_G(v)$ . O grau de um vértice  $v \in V$  é a sequência de graus de  $G$  é a sequência não crescente dos graus dos vértices de  $G$ , dada por  $d(G) = (d(v_1), d(v_2), \dots, d(v_n)) = (d_1, d_2, \dots, d_n)$ . Um subgrafo  $H$  de  $G$  é um grafo  $H = (W, F)$  em que  $W \subseteq V$  e  $F \subseteq E$ . Um grafo  $G$  é denominado  $r$ -regular se cada um de seus vértices tem grau  $r$ . O grafo  $kG$  é um grafo desconexo com  $k$  componentes isomorfas ao grafo  $G$ . Os grafos completo, caminho e ciclo de ordem  $n$  são denotados por  $K_n$ ,  $P_n$  e  $C_n$ , respectivamente. Um grafo  $G$  é dito bipartido quando  $V = X \cup Y$  tal que  $X \cap Y = \emptyset$  e dois vértices adjacentes não pertencem a mesma partição. Além disso, quando todos os vértices de  $X$  são adjacentes a todos os vértices de  $Y$ , tal grafo é dito bipartido completo e é denotado por  $K_{p,q}$ , onde  $|X| = p$  e  $|Y| = q$ .

Denotemos por  $M_n(\mathbb{R})$  o conjunto de todas as matrizes quadradas de ordem  $n$  com coeficientes em  $\mathbb{R}$ , por  $I_n$  a matriz identidade de ordem  $n$  e  $J_n$  a matriz de ordem  $n$  cujos elementos são todos iguais a 1. Dada uma matriz simétrica  $M(G)$  associada a um grafo  $G$  de ordem  $n$ , definimos o  $M$ -polinômio característico de  $G$  como  $p_M(G, x) = \det(xI_n - M(G))$ , e suas raízes são denominadas autovalores de  $M(G)$ . Como usual, indexamos os autovalores de  $M(G)$  de forma decrescente, e os denotamos por  $\lambda_1(M(G)) \geq \lambda_2(M(G)) \geq \dots \geq \lambda_n(M(G))$ . O  $M$ -espectro é o multiconjunto dos autovalores de  $M(G)$ , denotado por  $\text{Spec}(M(G)) = \{\lambda_1^{[s_1]}, \lambda_2^{[s_2]}, \dots, \lambda_r^{[s_r]}\}$ , onde  $s_i = m(\lambda_i)$  é a multiplicidade do autovalor  $\lambda_i(M(G))$ . Quando todos os autovalores de  $M$  são não negativos,  $M$  é dita positiva semidefinida.

A matriz de adjacência de  $G$  é definida por  $A(G) = [a_{ij}]$ , onde  $a_{ij} = 1$  se  $\{v_i, v_j\} \in E$  e  $a_{ij} = 0$  caso contrário. A matriz diagonal dos graus,  $D(G) = [d_{ij}]$ , é definida por  $d_{ii} = d(v_i)$ , e  $d_{ij} = 0$ ,  $\forall i \neq j$  e a matriz laplaciana sem sinal, denotada por  $Q(G)$ , é definida como  $Q(G) = D(G) + A(G)$  e o menor autovalor desta matriz é denotado por  $q_n(G)$ . Em 2017, Nikiforov (NIKIFOROV, 2017) definiu a matriz  $A_\alpha(G)$  como uma combinação linear convexa de  $A(G)$  e  $D(G)$  da seguinte forma  $A_\alpha(G) = \alpha D(G) + (1 - \alpha)A(G)$ , para  $\alpha \in [0, 1]$ .

É fácil ver que  $A_0(G) = A(G)$ ,  $A_1(G) = D(G)$  e  $2A_{1/2}(G) = Q(G)$ . Por conveniência, algumas vezes escrevemos  $A_\alpha(G) = A_\alpha$ . Os autovalores de  $A_\alpha(G)$  são denotados por  $\lambda_i(A_\alpha)$  onde  $1 \leq i \leq n$ , em ordem não crescente, ou seja,  $\lambda_1(A_\alpha) \geq \lambda_2(A_\alpha) \geq \dots \geq \lambda_n(A_\alpha)$ . Alguns artigos explicitam os autovalores da matriz  $A_\alpha$  de algumas classes de grafos, como podemos ver em (NIKIFOROV, 2017), (ROJO, 2017), (LIN; XUE; SHU, 2018), (BRONDANI; OLIVEIRA; FRANÇA, 2019), (DA SILVA JR; OLIVEIRA; DA COSTA; 2022), (BRONDANI; FRANÇA; OLIVEIRA, 2022a, 2022b), (BRONDANI; FRANÇA; LIMA, 2023). Nikiforov e Rojo (2017) definiram  $\alpha_0$  como o menor valor no intervalo  $[0, 1]$  tal que  $\lambda_n(A_{\alpha_0}(G)) \geq 0$ . Desta forma,  $A_\alpha(G)$  é positiva semidefinida se e somente se  $\alpha_0 \leq \alpha \leq 1$ . Neste mesmo artigo, foi levantado o seguinte problema "dado um grafo  $G$ , encontrar  $\alpha_0(G)$ ". Este problema foi tratado em artigos como (NIKIFOROV; ROJO, 2017) e (BRONDANI; FRANÇA; OLIVEIRA, 2022).

Os autovalores do grafo *blow-up*, definido na última seção deste artigo, e de seu grafo complementar para as matrizes de adjacência e laplaciana sem sinal já foram determinados na literatura como pode ser visto em (NIKIFOROV, 2006) e (LIMA; NIKIFOROV; OLIVEIRA, 2016). Os autores em (LIMA; NIKIFOROV; OLIVEIRA, 2016) usaram a operação *blow-up* para avaliar o comportamento assintótico do valor máximo do autovalor  $q_n(G)$  para grafos livres de subgrafos completos de ordem  $r + 1$ .

Neste trabalho, explicitamos os autovalores e o polinômio característico do grafo *blow-up* e de seu grafo complementar em termos dos autovalores da matriz  $A_\alpha$  do grafo.

O texto está organizado com os resultados preliminares na próxima seção e os resultados obtidos na seção seguinte.

## PRELIMINARES

Nesta seção, são apresentados os resultados preliminares necessários que são utilizados ao longo do texto, que podem ser vistos em (SILVESTER, 2000) e (NIKIFOROV, 2017). Conceitos gerais envolvendo as teorias de grafos e matrizes que são utilizados neste trabalho podem ser encontrados em (MERRIS, 2001), (BALAKRISHNAN; RANGANATHAN, 2012), (HARARY, 1969) e (HORN; JOHNSON, 2013).

A proposição a seguir relaciona o  $k$ -ésimo autovalor de  $A_\alpha$  com o grau do  $k$ -ésimo vértice da sequência de graus de  $G$ .

**Proposição 1** (NIKIFOROV, 2017). Se  $G$  é um grafo de ordem  $n$  com sequência de graus  $d(G) = (d_1, d_2, \dots, d_n)$ , então  $\lambda_k(A_\alpha) \leq d_k$ . Em particular,  $\lambda_1(A_\alpha) \leq \Delta(G)$  e  $\lambda_n(A_\alpha) < \delta\alpha$ .

A Proposição 2 explicita os autovalores da matriz  $A_\alpha$  de um grafo regular em função dos autovalores de  $A(G)$ .

**Proposição 2** (NIKIFOROV, 2017). Se  $G$  é um grafo  $r$ -regular de ordem  $n$  então  $\lambda_k(A_\alpha(G)) = \alpha r + (1 - \alpha)\lambda_k(A(G))$ ,  $1 \leq k \leq n$ .

As Proposições 3 e 4 explicitam completamente o  $A_\alpha$ -espectro do grafo completo e do grafo bipartido completo.

**Proposição 3** (NIKIFOROV, 2017). Os autovalores de  $A_\alpha(K_n)$  são  $\lambda_1(A_\alpha(K_n)) = n - 1$  e  $\lambda_k(A_\alpha(K_n)) = \alpha n - 1$ , onde  $2 \leq k \leq n$ .

**Proposição 4** (NIKIFOROV, 2017). Seja  $p \geq q \geq 1$ . Se  $\alpha \in [0, 1]$ , os autovalores de  $A_\alpha(K_{p,q})$  são

$$\begin{aligned} \lambda_1(A_\alpha(K_{p,q})) &= \frac{1}{2} \left[ \alpha(p+q) + \sqrt{\alpha^2(p+q)^2 + 4pq(1-2\alpha)} \right]; \\ \lambda_n(A_\alpha(K_{p,q})) &= \frac{1}{2} \left[ \alpha(p+q) - \sqrt{\alpha^2(p+q)^2 + 4pq(1-2\alpha)} \right]; \\ \lambda_k(A_\alpha(K_{p,q})) &= \alpha p, \text{ onde } 1 < k \leq q; \\ \lambda_k(A_\alpha(K_{p,q})) &= \alpha q, \text{ onde } q < k < p+q. \end{aligned}$$

A seguir, apresentamos o produto de Kronecker de matrizes e algumas de suas propriedades.

**Definição 1.** Sejam  $A = [a_{ij}] \in M_{m,n}(R)$  e  $B \in M_{p,q}(R)$ . O produto de Kronecker de  $A$  e  $B$ , denotado por  $A \otimes B$ , é uma matriz em  $M_{mp,nq}(R)$  definida por

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

Por definição, o produto de Kronecker é escrito como uma matriz em blocos, isto é,  $A \otimes B = [a_{ij}B]$ , onde  $a_{ij}B$  é o  $(i, j)$ -ésimo bloco de  $A \otimes B$ . Vale ressaltar que, como no produto usual de matrizes, o produto de Kronecker não é comutativo. A seguir citamos algumas propriedades que são válidas para o produto de Kronecker de matrizes.

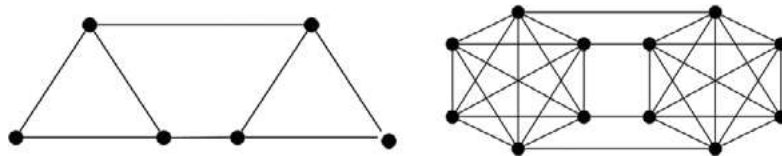
**Teorema 1** (BROXSON, 2006). Se  $A, B, C \in M_n(\mathbb{R})$ , então

- (i)  $(\lambda A) \otimes B = \lambda(A \otimes B) = A \otimes (\lambda B)$ ;
- (ii)  $(A \otimes B)^T = A^T \otimes B^T$ ;
- (iii)  $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ ;
- (iv)  $(A + B) \otimes C = A \otimes C + B \otimes C$ ;
- (v)  $A \otimes (B + C) = A \otimes B + A \otimes C$ ;
- (vi)  $A \otimes 0 = 0 \otimes A = 0$ ;
- (vii)  $J_n \otimes J_m = J_{nm}$ .

A seguir, apresentamos a definição do grafo  $\mathcal{H}_n^\ell$ .

**Definição 2.** Dados os inteiros  $n$  e  $\ell$ ,  $n \geq 3$  e  $1 \leq \ell \leq n$ , seja  $\mathcal{H}_n^\ell$  o grafo obtido de duas cópias do grafo completo  $K_n$  adicionando  $\ell$  arestas entre  $\ell$  vértices da cópia de  $K_n$  e os seus vértices correspondentes na outra cópia. Na Figura 1, apresentamos dois exemplos de grafos  $\mathcal{H}_n^\ell$ .

Figura 1: Grafos  $\mathcal{H}_3^2$  e  $\mathcal{H}_6^4$ , respectivamente.



Fonte: Produção própria dos autores.

Para concluir esta seção, exibimos a definição da operação *join* e de sua generalização, a operação *H-join*.

**Definição 3.** Sejam  $G$  e  $H$  grafos tais que  $V(G) \cap V(H) = \emptyset$ . O *join* de  $G$  e  $H$ , denotado por  $G \vee H$ , é o grafo com conjunto de vértices  $V(G \vee H) = V(G) \cup V(H)$ , e conjunto de arestas

**Definição 4.** Dado um grafo  $H$  com  $V(H) = \{1, 2, \dots, p\}$  e uma família de grafos  $\mathcal{F} = \{G_1, G_2, \dots, G_p\}$  tal que  $|V(G_j)| = n_j$ ,  $1 \leq j \leq p$ . O *H-join* de  $G_1, G_2, \dots, G_p$  é o grafo  $G = \vee_H \{G_j; 1 \leq j \leq p\}$  tal que

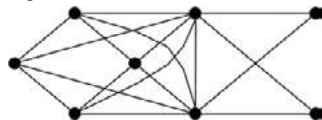
$$V(G) = \bigcup_{j=1}^p V(G_j) \text{ e } E(G) = \left( \bigcup_{j=1}^p E(G_j) \right) \cup \left( \bigcup_{\{r,s\} \in E(H)} E(G_r \vee G_s) \right)$$

**Observação 1.** Note que quando  $H = K_2$ , o *H-join* é a operação *join*, descrita na Definição 3.

A seguir, apresentamos um exemplo de *H-join*.

**Exemplo 1.** Na Figura 2, ilustramos o grafo *H-join* de  $\mathcal{F} = \{C_4, K_2, 2K_1\}$ , com  $H = P_3$ .

Figura 2: Grafo



Fonte: Produção própria dos autores.

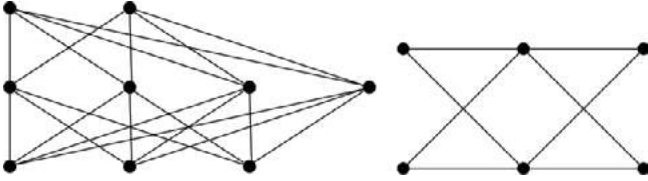
## RESULTADOS

Nesta seção, apresentamos o grafo *blow-up* de um grafo  $G$  e obtemos resultados acerca de seu  $A_\alpha$ -espectro.

**Definição 5.** Dado um grafo  $G$  com  $V(G) = \{v_1, \dots, v_n\}$  e um vetor  $m = [m_1, \dots, m_n]$  de inteiros positivos, denotamos por  $G \circ m$  o grafo obtido de  $G$  trocando cada vértice  $v_i \in V(G)$  por um conjunto independente de  $m_i$  vértices  $v_i^1, v_i^2, \dots, v_i^{m_i}$  e  $\{v_i^s, v_j^p\} \in E(G \circ m)$  se e somente se  $\{v_i, v_j\} \in E(G)$ . O grafo  $G \circ m$  é chamado de *blow-up* de  $G$ . Denotamos o grafo  $G \circ m$  por  $G^{(t)}$  quando  $m_i = m_j$ , onde  $1 \leq i, j \leq n$ .

A Figura 3 mostra os *blow-ups*  $P_3 \circ m$ , onde  $m = (3, 4, 2)$  e  $P_3^{(2)}$ , respectivamente.

Figura 3: Grafos blow-ups.



Fonte: Produção própria dos autores.

Podemos observar que o *blow-up* de um grafo  $G$  de ordem  $n$ ,  $G \circ m$ , onde  $m = [m_1, \dots, m_n]$ , é um caso particular do *H-join* quando tomamos  $H = G$  e  $\mathcal{F} = \{m_1 K_1, \dots, m_n K_1\}$ . Em (CARDOSO *et al.*, 2022), os autores expressaram os autovalores da operação *H-join* da matriz  $A_\alpha$  utilizando a matriz cadeia e autovalores principais do grafo. Neste trabalho, apresentamos o  $A_\alpha$ -espectro do *blow-up* do tipo  $G^{(t)}$  utilizando uma outra técnica de demonstração através do produto de Kronecker, além de explicitar o  $A_\alpha$ -espectro do grafo complementar da operação *blow-up*, o que ainda não havia sido feito.

Seja  $G$  um grafo de ordem  $n$  e tamanho  $m$ . O *blow-up* de  $G$ ,  $G^{(t)}$ , é de ordem  $tn$ , tamanho  $mt^2$  e sua matriz de adjacência de  $G^{(t)}$  é dada por

$$A(G^{(t)}) = \begin{bmatrix} A(G) & A(G) & \dots & A(G) \\ A(G) & A(G) & \dots & A(G) \\ \vdots & \vdots & \ddots & \vdots \\ A(G) & A(G) & \dots & A(G) \end{bmatrix} = J_t \otimes A(G)$$

Este comportamento é natural rotulando primeiramente os vértices  $v_i^1$ , com  $1 \leq i \leq n$  da primeira cópia de  $G$ , depois os vértices  $v_i^2$  da segunda cópia de  $G$ , e assim por diante. Desta forma, no bloco da linha  $i$ , coluna  $j$ , teremos um bloco que diz respeito a como a  $i$ -ésima cópia de  $G$  se comporta em relação à sua  $j$ -ésima cópia. A matriz dos graus de  $G^{(t)}$  é dada por

$$D(G^{(t)}) = \begin{bmatrix} tD(G) & 0 & \dots & 0 \\ 0 & tD(G) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & tD(G) \end{bmatrix} = tI_t \otimes D(G).$$

Desta forma, encontramos a matriz  $A_\alpha$  de  $G^{(t)}$ , dada por

$$A_\alpha(G^{(t)}) = \alpha tI_t \otimes D(G) + (1 - \alpha)J_t \otimes A(G).$$

Agora, para o grafo complementar de  $G^{(t)}$ , temos

$$A(\overline{G^{(t)}}) = \begin{bmatrix} J_n - I_n - A(G) & J_n - A(G) & \dots & J_n - A(G) \\ J_n - A(G) & J_n - I_n - A(G) & \dots & J_n - A(G) \\ \vdots & \vdots & \ddots & \vdots \\ J_n - A(G) & J_n - A(G) & \dots & J_n - I_n - A(G) \end{bmatrix} \\ = J_t \otimes J_n - I_t \otimes I_n - J_t \otimes A(G).$$

que podemos reescrever como  $A(\overline{G^{(t)}}) = J_t \otimes (J_n - A(G)) - I_t \otimes I_n$ . A matriz dos graus de  $\overline{G^{(t)}}$  é dada por

$$D(\overline{G^{(t)}}) = \begin{bmatrix} (tn-1)I_n - tD(G) & 0 & \dots & 0 \\ 0 & (tn-1)I_n - tD(G) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (tn-1)I_n - tD(G) \end{bmatrix} \\ = I_t \otimes ((tn-1)I_n - tD(G)).$$

Assim,

$$A_\alpha(\overline{G^{(t)}}) = \alpha D(\overline{G^{(t)}}) + (1-\alpha)A(\overline{G^{(t)}}) \\ = I_t \otimes [\alpha t(nI_n - D(G)) - I_n] + (1-\alpha)J_t \otimes (A(G))$$

O próximo teorema descreve o espectro das matrizes  $A_\alpha(\overline{G^{(t)}})$  e  $A_\alpha(G^{(t)})$ .

**Teorema 2.** Seja  $G$  um grafo de ordem  $n$ . Para cada  $i, 1 \leq i \leq n$ , as seguintes afirmações são válidas.

- Os autovalores de  $A_\alpha(G^{(t)})$  são  $t\lambda_i(A_\alpha(G))$  e  $\alpha t d(v_i)$ , onde  $\alpha t d(v_i)$  tem multiplicidade igual a pelo menos  $t-1$ .
- Os autovalores de  $A_\alpha(\overline{G^{(t)}})$  são  $t\lambda_i(A_\alpha(G)) + t-1$  e  $\alpha t(n-d(v_i)) - 1$ , onde  $\alpha t(n-d(v_i)) - 1$  tem multiplicidade igual a pelo menos  $t-1$ .

*Demonstração.* Seja  $G$  um grafo de ordem  $n$ . Dado um inteiro  $t \geq 2$ , temos

$$A_\alpha(\overline{G^{(t)}}) = \alpha t J_t \otimes D(G) + (1-\alpha)J_t \otimes A(G).$$

Como  $A_\alpha(G)$  é uma matriz simétrica, existe uma base ortonormal  $\{x_1, x_2, \dots, x_n\}$  formada por autovetores de  $A_\alpha(G)$ . Para cada  $i, 1 \leq i \leq n$ , consideremos o vetor  $j_t \otimes x_i$ , onde  $j_t$  é o vetor com  $t$  entradas iguais a 1. Desta forma,

$$A_\alpha(\overline{G^{(t)}})(j_t \otimes x_i) = \alpha t j_t \otimes D(G)x_i + (1-\alpha)[J_t j_t \otimes A(G)x_i] \\ = j_t \otimes \alpha D(G)x_i + t j_t \otimes (1-\alpha)A(G)x_i \\ = j_t \otimes A_\alpha(G)x_i = t\lambda_i(A_\alpha(G))(j_t \otimes x_i).$$

Assim, para cada  $i, t\lambda_i(A_\alpha(G))$  é um autovalor de  $A_\alpha(\overline{G^{(t)}})$ . Agora, seja  $p_t^{(i,j)}$  um vetor de ordem  $t$  onde a  $i$ -ésima entrada é igual a 1, a  $j$ -ésima entrada igual a  $-1$ , com  $i \neq j$  e todas as outras entradas iguais a 0, e, seja  $\{y_1, \dots, y_n\}$  uma base ortonormal de autovetores associados à  $D(G)$ . Desta forma,

$$A_\alpha(\overline{G^{(t)}})(p_t^{(i,j)} \otimes y_i) = \alpha t p_t^{(i,j)} \otimes D(G)y_i + (1-\alpha)[J_t p_t^{(i,j)} \otimes A(G)y_i] \\ = \alpha t p_t^{(i,j)} \otimes \lambda_i(D(G))y_i = \alpha t d(v_i)(p_t^{(i,j)} \otimes y_i).$$

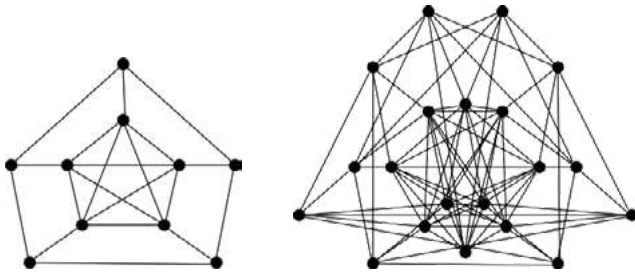
Portanto, para cada  $i, 1 \leq i \leq n, \alpha t d(v_i)$  é autovalor de  $A_\alpha(\overline{G^{(t)}})$  com multiplicidade ao menos  $t-1$ , pois podemos escolher  $p_t^{(i,j)}$  de  $t-1$  maneiras diferentes. Para obter o resultado para o item (ii), basta aplicar a mesma técnica de prova para  $\overline{G^{(t)}}$ .

Como aplicação do Teorema 2, apresentamos os Exemplos 2 e 3.

**Exemplo 2.** Na Figura 4 exibimos o grafo de Petersen  $G$  e seu *blow-up*  $\overline{G^{(2)}}$ .



Figura 4: Grafo de Petersen e seu *blow-up*, respectivamente.



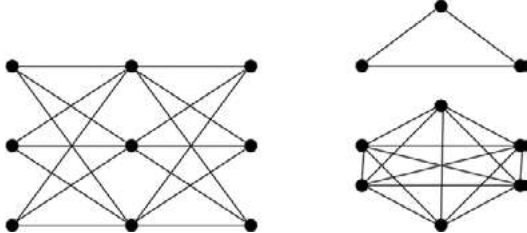
Fonte: Produção própria dos autores.

Como sabemos, o grafo de Petersen é um grafo 3-regular e o espectro da sua matriz de adjacência é dado por  $\text{Spec}(A(G)) = \{(-2)^{[4]}, 1^{[5]}, 3\}$ . Desta forma, pela Proposição 2, concluímos que seu  $A_\alpha$ -espectro é dado por  $\text{Spec}(A_\alpha(G)) = \{(5\alpha - 2)^{[4]}, (2\alpha + 1)^{[5]}, 3\}$  e portanto pelo Teorema 2 conclui-se que

$$\text{Spec}(A_\alpha(G^{(2)})) = \{(10\alpha - 4)^{[4]}, (4\alpha + 2)^{[5]}, 6, (6\alpha)^{[10]}\}.$$

**Exemplo 3.** Na Figura 5 ilustramos o grafo blow-up  $P_3^{(3)}$  e seu grafo complementar  $\overline{P_3^{(3)}}$ .

Figura 5: Grafos  $P_3^{(3)}$  e  $\overline{P_3^{(3)}}$ , respectivamente.



Fonte: Produção própria dos autores.

Como  $\text{Spec}(A_\alpha(P_3)) = \left\{ \alpha, \frac{1}{2}(3\alpha \pm \sqrt{9\alpha^2 - 16\alpha + 8}) \right\}$  e  $\text{Spec}(A_\alpha(\overline{P_3^{(3)}})) = \{0, 1, 2\alpha - 1\}$  do Teorema 2 obtemos

$$\text{Spec}(A_\alpha(P_3^{(3)})) = \left\{ (3\alpha)^{[5]}, (6\alpha)^{[2]}, \frac{3}{2}(3\alpha \pm \sqrt{9\alpha^2 - 16\alpha + 8}) \right\}.$$

$$\text{Spec}(A_\alpha(\overline{P_3^{(3)}})) = \{2, 5, (6\alpha - 1)^{[5]}, (3\alpha - 1)^{[2]}\}.$$

**Corolário 3.** Se  $G$  é um grafo de ordem  $n$ , então

$$p_{A_\alpha}(G^{(t)}, x) = p_{A_\alpha}(G, x/t) \prod_{i=1}^n (x - t\alpha d(v_i)).$$

$$\text{Em particular, } \lambda_n(A_\alpha(G^{(t)})) = t\lambda_n(A_\alpha(G)).$$

*Demonstração.* Do Teorema 2 temos o espectro de  $A_\alpha(G^{(t)})$  e, conseqüentemente, o seu polinômio característico. Da Proposição 1 obtemos  $\lambda_n(A_\alpha(G^{(t)})) < \alpha\delta(G^{(t)}) = \alpha t\delta(G)$ . Portanto,  $\lambda_n(A_\alpha(G^{(t)})) = t\lambda_n(A_\alpha(G))$ .

**Exemplo 4.** Como consequência do Corolário 3, podemos explicitar o polinômio característico de  $A_\alpha(K_n^{(t)})$  e  $A_\alpha(K_{p,q}^{(t)})$ . De acordo com as Proposições 3 e 4 temos,

$$p_{A_\alpha}(K_n^{(t)}, x) = (x - tn + t)(x - \alpha tn + t)^{n-1}(x - \alpha tn - \alpha t)^n$$

e

$$p_{A_\alpha}(K_{p,q}^{(t)}, x) = (x - \alpha tp)^{2q-1}(x - \alpha tq)^{2p-1} \left( x^2 - \alpha t(p+q)x - \frac{t^2}{4} \alpha^2(p+q)^2 + 4pq(1-2\alpha) \right)$$

Em particular,  $\lambda_n(A_\alpha(K_n^{(t)})) = t(\alpha n - 1)$  e

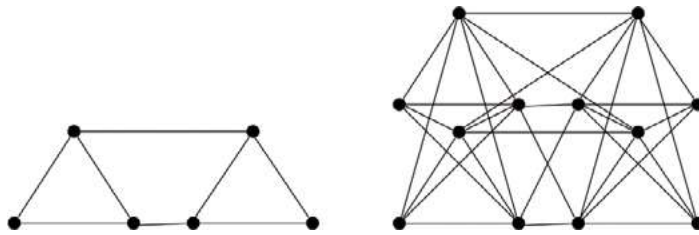
$$\lambda_n(A_\alpha(K_{p,q}^{(t)})) = \frac{t}{2} \left[ (\alpha(p+q) - \sqrt{\alpha^2(p+q)^2 + 4pq(1-2\alpha)}) \right]$$

Do Corolário 3 obtemos a observação a seguir.

**Observação 3.**  $A_\alpha(G^{(t)})$  é positiva semidefinida se e somente se  $A_\alpha(G)$  também for.

**Exemplo 5.** Na Figura 6 apresentamos o grafo  $\mathcal{H}_3^2$  e o seu *blow-up*  $\mathcal{H}_3^{2(2)}$ .

Figura 6: Grafo  $\mathcal{H}_3^2$  e o seu *blow-up*  $\mathcal{H}_3^{2(2)}$ , respectivamente.



Fonte: Produção própria dos autores.

De acordo com (BRONDANI; FRANÇA; OLIVEIRA, 2022),  $\alpha_0(\mathcal{H}_n^\ell) = \frac{2}{n+2}$ , para  $2 \leq \ell \leq n$ .

Logo  $\alpha_0(\mathcal{H}_3^2) = \frac{2}{5}$ . Da Observação 3, obtemos  $\alpha_0(\mathcal{H}_3^{2(2)}) = \frac{2}{5}$ . Para  $l \geq 2$ , obtemos  $G_{l+1}^{(t)}$  a partir de  $G_l^{(t)}$  aplicando a operação  $G^{(t)}$  ao grafo  $G_l^{(t)}$ , ou seja  $G_{l+1}^{(t)} = (G_l^{(t)})^{(t)}$  e consideramos  $G_1^{(t)} = G^{(t)}$ . Desta forma, a partir da Proposição 1 chegamos à Observação 4.

**Observação 4.** Se  $G$  é um grafo então  $\alpha_0(G) = \alpha_0(G^{(t)}) = \alpha_0(G_2^{(t)}) = \dots = \alpha_0(G_l^{(t)})$ , para  $l \geq 1$ .

## CONCLUSÕES

Neste trabalho, exibimos o  $A_\alpha$ -espectro do *blow-up* de um grafo  $G$  em termos do  $A_\alpha$ -espectro de  $G$  e concluímos que o menor valor de  $\alpha$  para o qual  $A_\alpha(G)$  é positiva semidefinida não se altera quando aplicamos a operação *blow-up*.

## AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001 e da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro - FAPERJ processo E-20/2022-284573. A pesquisadora Carla Oliveira é apoiada pelo CNPq Grant 304548/2020-0.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BALAKRISHNAN, R.; RANGANATHAN, K. **A Textbook of Graph Theory**. New York: Springer, 2012.
- BRONDANI, A. E.; OLIVEIRA, C. S.; FRANÇA, F. A. M.; de LIMA, L.  $\mathcal{A}_\alpha$ -spectrum of a firefly graph. **Electronic Notes in Theoretical Computer Science**, v. 346, p. 209–219, 2019. BRONDANI, A. E.; FRANÇA, F. A. M.; OLIVEIRA, C. S. Positive semidefiniteness of  $\mathcal{A}_\alpha(G)$  on some families of graphs. **Discrete Applied Mathematics**, v. 323, p. 113–123, 2022.
- BRONDANI, A. E.; FRANÇA, F. A. M.; OLIVEIRA, C. S.  $\mathcal{A}_\alpha$  and  $\mathcal{L}_\alpha$ -spectral properties of spider graphs. **Proyecciones Journal of Mathematics**, v. 41, p. 965–982, 2022.
- BRONDANI, A. E.; FRANÇA, F. A. M.; LIMA, R. G. O. J.  $\mathcal{A}_\alpha$ -spectrum of duplicate and corona operations in graphs. **Proyecciones Journal of Mathematics**, v. 42, p. 967–982, 2023.
- BROXSON, B. J. **The kronecker product**. Dissertação. – University of North Florida, Florida, 2006.
- CARDOSO, D. M.; GOMES, H.; PINHEIRO, S. J. The  $\mathcal{H}$ -join of arbitrary families of graphs. **Linear Algebra and its Applications**, v. 648, p. 160–180, 2022.
- DA SILVA JR., J. D. G.; OLIVEIRA, C. S.; DA COSTA, L. M. G. C. C. On the characteristic polynomial of the  $\mathcal{A}_\alpha$ -matrix for some operations of graphs. **arXiv**, 2022.
- HARARY, F. **Graph Theory**. Philippines: Addison-Wesley Plusbling Company, Inc., 1969. HORN, R. A.; JOHNSON, C. R. **Matrix Analysis**. 2. ed. Cambridge, 2013.
- LIN, H.; XUE, J.; e SHU, J. On the  $\mathcal{A}_\alpha$ -spectra of graphs. **Linear Algebra and its Applications**, v. 556, p. 210–219, 2018.
- LIMA, L.; NIKIFOROV, V.; OLIVEIRA, C. The clique number and the smallest  $\mathcal{Q}$ -eigenvalue of graphs. **Discrete Mathematics**, v. 339, p. 1744–1752, 2016. MERRIS, R. **Graph Theory**. Canada: JOHN WILEY SONS, INC., 2001.
- NIKIFOROV, V. Linear combinations of graph eigenvalues. **Electronic Journal of Linear Algebra**, v. 15, p. 329–336, 2006.
- NIKIFOROV, V. Merging the  $\mathcal{A}$ - and  $\mathcal{Q}$ -spectral theories. **Applicable Analysis and Discrete Mathematics**, v. 11, p. 81–107, 2017.
- NIKIFOROV, V.; ROJO, O. A note on the positive semidefiniteness of  $\mathcal{A}_\alpha(G)$ . **Linear Algebra and its Applications**, v. 519, p. 156–163, 2017.

## DADOS DOS AUTORES

**Uilton Cesar Peres Junior**. Mestrando em Engenharia de Produção e Sistemas. Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET-RJ). uilton.junior@aluno.cefet-rj.br.

**Carla Silva Oliveira**. Doutora em Engenharia de Produção. Docente e pesquisadora na Escola Nacional de Ciências Estatísticas (ENCE/IBGE). carla.oliveira@ibge.gov.br.

**André Ebling Brondani**. Doutor em Engenharia de Produção. Docente na Universidade Federal Fluminense (UFF). andrebrondani@id.uff.br.

# ANÁLISE DO TRANSPORTE DE CONTAMINANTES EM UM RIO UTILIZANDO PROBLEMAS INVERSOS

## ANALYSIS OF CONTAMINANT TRANSPORT IN A RIVER USING INVERSE PROBLEMS

Anna Luisa Soares Castro Coimbra

Fábio Freitas Ferreira

Wagner Rambaldi Telles

**Resumo:** É notável a existência de algum tipo de poluição na maioria dos rios no Brasil e, diante de situações como essa, o controle do comportamento de contaminantes nesses corpos hídricos é extremamente útil para processos de tomada de decisão, entre elas, determinar o melhor local para descarte de resíduos domésticos e industriais, evitando, assim, uma contaminação do meio ambiente e contribuindo para um melhor monitoramento da qualidade da água. Neste trabalho é analisado um modelo computacional para simular o transporte de contaminantes lançados em um trecho do rio São Pedro, localizado no município de Nova Friburgo, estado do Rio de Janeiro, visando estimar os valores ótimos para a velocidade média e o coeficiente de dispersão utilizando um método de otimização determinístico, mais especificamente, o método Levenberg-Marquardt.

**Palavras-chave:** Transporte de Contaminantes. Modelagem Computacional. Problemas Inversos. Método de Levenberg-Marquardt.

**Abstract:** It is notable that there is some type of pollution in most rivers in Brazil and, faced with situations like this, controlling the behavior of contaminants in these water bodies is extremely useful for decision-making processes, among them, determining the best place to dispose of domestic and industrial waste, thus avoiding contamination of the environment and contributing to a better monitoring of water quality. In this work, a computational model is analyzed to simulate the transport of contaminants released into a stretch of the São Pedro river, located in the municipality of Nova Friburgo, state of Rio de Janeiro, aiming to estimate the optimal values for the average speed and the dispersion coefficient using a deterministic optimization method, more specifically, the Levenberg-Marquardt method.

**Keywords:** Transport of Contaminants. Computational modeling. Inverse Problems. Levenberg-Marquardt method.

## INTRODUÇÃO

A água é um recurso natural essencial para a sobrevivência humana, desde criança aprendemos que nosso corpo é 70% água, e que precisamos nos manter hidratados e beber água. No entanto, por lidar diariamente com a água em abundância presente nos mares, rios, lagos e represas, a sociedade continua banalizando sua importância e usando-a como um bem descartável.

Entretanto, nem toda essa água presente no mundo é um recurso hídrico. Conforme a Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA, 2022), recurso hídrico é toda água proveniente da superfície ou subsuperfície da Terra, que pode ser empregada em um determinado uso ou atividade, podendo também passar a ser um bem econômico. Todo recurso hídrico é água, mas nem toda água é recurso hídrico. Logo, a água precisa possuir alguma finalidade de uso para a sociedade.

Assim, a poluição hídrica é causada por atividades humanas que geram algum impacto sobre esse recurso natural em qualidade e quantidade. A poluição se dá devido à alteração das características biológicas, físicas, químicas e sedimentar.

No Brasil, a Constituição da República Federativa do Brasil de 1988, documento oficial de maior importância nos países, prevê em seu art. nº 225 que todos têm direito ao meio ambiente ecologicamente equilibrado, bem de uso comum do povo e essencial à sadia qualidade de vida, impondo-se ao Poder Público e à coletividade o dever de defendê-lo e preservá-lo para as presentes e futuras gerações. Contudo, segundo o relatório da Fundação SOS Mata Atlântica divulgado em 2022, apenas 6,9% dos rios observados estão com qualidade da água em boa condição; 72,6% apresentaram qualidade da água regular; 17,8%, ruim; e 2,7%, péssima. O levantamento foi realizado em 146 pontos de coletas de 90 rios e corpos d'água de 65 municípios, em 16 estados do bioma Mata Atlântica, por 106 grupos de monitoramento da qualidade da água do Programa Observando os Rios (SOSMA, 2022).

Portanto, torna-se evidente a necessidade de ferramentas para auxiliar no monitoramento da qualidade das águas, uma vez que os órgãos gestores do meio ambiente esbarram em dificuldades logísticas, financeiras e operacionais: coletas e análises de água são caras, além de requererem material e pessoal especializado (Telles, 2009).

A relevância do presente trabalho se dá devido à grande mobilização mundial para um melhor controle da qualidade de água, em especial a água enquanto recurso hídrico presente nos rios, lagos, lençóis freáticos superficiais e atmosfera, a qual vem sendo atingida cada vez mais pelos crescentes índices de poluição. Diante deste cenário, é importante buscar subsídios para monitorar a qualidade da água de cursos d'água naturais.

Nesse sentido, a modelagem matemática aplicada ao problema de transporte de contaminantes em rios se mostra uma opção mais viável financeira e logisticamente ao desenvolver e aprimorar modelos computacionais que auxiliem no controle da poluição dos rios, pois ajudam a prever o comportamento e estimar o nível de danos que podem ser causados pela poluição despejada na água.

## ÁREA DE ESTUDO – RIO SÃO PEDRO E DESCRIÇÃO DO EXPERIMENTO

A área de estudo para a qual se aplicou a simulação analisada neste trabalho está situada na região serrana do Estado do Rio de Janeiro, dentro da bacia hidrográfica do rio Macaé, a qual possui uma área de drenagem de aproximadamente  $1.765 \text{ km}^2$  e engloba praticamente toda a área dos limites territoriais do município de Macaé, onde se localiza sua foz.

Em particular, o rio São Pedro, região escolhida para a realização do experimento, está situada em São Pedro da Serra, 7º Distrito de Nova Friburgo (RJ), localização (  $22^{\circ}19'17.7''\text{S } 42^{\circ}19'38.0''\text{W}$ ), local que foi afetado por um rápido processo de transformação espacial, social e econômica, com fortes impactos ambientais (LimaVerde, 2005).

A região tem como característica uma pequena profundidade e margens repletas de rochas conforme mostrado na Figura 1. Além de conter trechos sinuosos e grande biodiversidade, tem como principais atividades econômicas a agricultura familiar e o turismo, dado que a área engloba atrativos como as cachoeiras e a mata, possibilitando a prática de trilhas.

Figura 1: Vista parcial do trecho do rio São Pedro.



Fonte: Sousa (2009)

Os dados experimentais utilizados nesse trabalho, tem como base o experimento realizado e descrito por Sousa (2009), o qual ocorreu no dia 26 de janeiro de 2009, com início às 10:50:28 h e participação de 5 pessoas, onde buscou-se simular um cenário de poluição acidental no curso de água em questão. Para isso, foi utilizada uma solução de cloreto de sódio (NaCl) como traçador.

Para a preparação da solução salina, foram utilizados 2000 g de cloreto de sódio, diluídos em aproximadamente 15 l de água em um balde, os quais foram liberados instantaneamente em um ponto da seção de injeção, sobre a linha de corrente central do escoamento. A primeira seção foi escolhida a 50 m do ponto de injeção e, a segunda, a 100 m. Foram colhidas amostras de 200 ml da água a cada 15 s, totalizando 60 amostras, as quais, posteriormente, foram analisadas em laboratório. Na Tabela 1 são mostradas as configurações do rio.

Tabela 1: Características do rio no ponto de lançamento do traçador.

PARÂMETROS	VALORES
Vazão ( $Q = UAQ = UA$ )	0,43 m <sup>3</sup> /s
Área ( $AA$ )	1,05 m <sup>2</sup>
Largura ( $BB$ )	2,40 m
Profundidade ( $H = A/BH = A/B$ )	0,43 m
Velocidade ( $UU$ )	0,41 m/s
Declividade ( $SS$ )	0,03 m/m
Concentração inicial ( $C_0C_0$ )	15,50 mg/L

## MODELAGEM E SOLUÇÃO DO PROBLEMA PROPOSTO

O modelo matemático para o problema proposto neste trabalho leva em consideração o fato de o trecho do rio estudado ser suficientemente estreito e raso, fazendo com que as variações nas direções transversal e vertical sejam desprezíveis em relação às variações longitudinais. Logo, um modelo unidimensional que considera apenas as variações na referida direção, sendo expresso pela Equação de Advecção-Dispersão, Equação (1), que descreve o comportamento de um constituinte em relação ao seu deslocamento e sua dispersão ao longo do rio, com condições inicial e de contorno, descritas pelas Equações (2a) e (2b), respectivamente:

• Equação de Advecção-Dispersão:

$$\frac{\partial C}{\partial t} + U \frac{\partial C}{\partial x} = E_L \frac{\partial^2 C}{\partial x^2}, \quad -\infty < x < \infty, t > 0 \quad (1)$$

• Condições inicial e de contorno:

$$C(x, 0) = C_0 + \frac{M}{A} \delta(x), \quad -\infty < x < \infty \quad (2a)$$

$$C(\pm\infty, t) = C_0, \quad t > 0 \quad (2b)$$

em que  $C$  é a concentração ( $mg/l$ ),  $C_0$  é a concentração no domínio, exceto no ponto de lançamento do poluente ( $mg/l$ ),  $t$  é a variável temporal ( $s$ ),  $x$  é a variável espacial ( $m$ ),  $U$  é a velocidade no sentido longitudinal do escoamento ( $m/s$ ),  $E_L$  é o coeficiente de dispersão longitudinal ( $m^2/s$ ),  $M$  é a massa do constituinte ( $mg$ ),  $A$  é a área da seção transversal do rio ( $m^2$ ) e  $\delta(x)$  é a função *Delta de Dirac*.

A solução analítica para o modelo matemático descrito na Eq. (1), com condições de contorno e inicial, dadas pelas Eq. (2a) e (2b), é dada pela Eq. (3) (Sousa, 2009):

$$C(x, t) = C_0 + \frac{M}{A\sqrt{4\pi E_L t}} \exp\left(-\frac{(x - Ut)^2}{4E_L t}\right), \quad -\infty < x < \infty, t > 0 \quad (3)$$

A estimativa dos valores ótimos para a velocidade ( $U$ ) e o coeficiente de dispersão longitudinal ( $E_L$ ), de maneira que permita o melhor ajuste entre os dados experimentais e os obtidos analiticamente, Equação (3), é feita utilizando o método Levenberg-Marquardt, cujos resultados são apresentados na próxima seção.

## PROBLEMAS INVERSOS E O MÉTODO DE LEVENBERG-MARQUARDT

Para estimar a velocidade e o coeficiente de dispersão longitudinal da equação de advecção-dispersão, aplica-se as técnicas de problemas inversos para determinar os parâmetros de entrada. Ou seja, procura-se minimizar o funcional de resíduos quadrados,

$$S(\vec{Z}) = \frac{1}{2} \sum_{i=1}^M (C_{\text{calc}_i} - C_{\text{exp}_i})^2$$

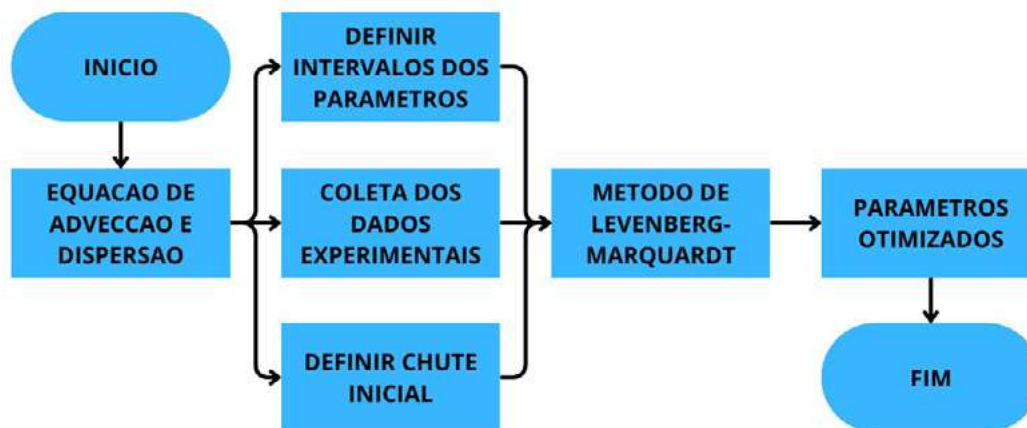
em que  $M$  é a quantidade de dados experimentais;  $\vec{Z} = (U, E_L)$  representa os parâmetros a serem estimados;  $C_{\text{calc}_i}$  é grandeza calculada e  $C_{\text{exp}_i}$  é a grandeza obtida experimentalmente.

Para estimar os parâmetros de entrada tem-se um problema de dimensão finita. Logo, utiliza-se métodos de estimação de parâmetros. Existem métodos estocásticos, determinísticos e híbridos. Para este trabalho escolheu-se trabalhar com o método de Levenberg-Marquardt. O método de Levenberg-Marquardt é um método determinístico que busca minimizar o funcional de resíduos quadrados, Equação (3), resolvendo a equação de gradiente igual a zero.

Para isso aplica-se o método de Newton não linear para criar um procedimento iterativo em busca da solução. Este procedimento gera uma matriz Hessiana mal condicionada. Para regularização do método, Marquardt propôs adicionar um valor na diagonal da matriz Hessiana tornando-a uma matriz com diagonal dominante e positiva definida. Espera-se que ao longo do procedimento iterativo essa contribuição tenda a zero.

Outro problema do método é que a sua solução depende intrinsecamente do chute inicial a candidato a solução, ou sejam, dependendo do chute inicial o método pode ir para mínimos locais

não desejados. Ferreira et. al. (2019) propôs um procedimento para criar e limitar o intervalo de busca do método de Levenberg-Marquardt, desenvolvendo um método híbrido, Levenberg-Marquardt com Restrições. Todas as vezes que o método de Levenberg-Marquardt gera um candidato a solução fora do intervalo de busca, o procedimento gera um novo chute inicial candidato a solução dentro do intervalo de busca, reiniciando o método. Assim, será iniciado um novo caminho em busca da solução. Neste trabalho foi utilizado apenas a limitação do intervalo de busca para definir novos chutes iniciais sem utilizar a estocasticidade que torna o método híbrido.



## RESULTADOS E DISCUSSOES

Os resultados apresentados foram obtidos usando o método Levenberg-Marquardt, em que o mesmo foi executado 100 vezes com o intuito de estimar e analisar a velocidade  $(U)(U)$  e o coeficiente de dispersão  $(E_L)(E_L)$  presentes no problema. Em cada uma dessas execuções, foi calculado o número de avaliações da função ob-

jetivo  $(S(\vec{Z}))(S(\vec{Z}))$ , imprescindível para que o algoritmo encontre os parâmetros de interesse, tomando como critério de parada o erro relativo entre duas iterações consecutivas expresso por  $|Z_i^{k+1} - Z_i^k| < 10^{-1}|Z_i^{k+1} - Z_i^k| < 10^{-1}$ . Para as estimativas iniciais  $\vec{Z}^0 \vec{Z}^0$  referentes a cada uma das 100 execuções, utilizou-se um valor tomado aleatoriamente nos intervalos  $0,1 \leq U \leq 1,00, 1 \leq U \leq 1,0$  e  $1,0 \leq E_L \leq 2,0, 1,0 \leq E_L \leq 2,0$ . Os resultados foram gerados utilizando o software scilab em um computador Intel Core i5 com 8 GB de memória RAM,

Ademais, para uma análise dos resultados obtidos pelo método Levenberg-Marquardt, foram calculados indicadores estatísticos, levando em conta o total de execuções, sendo esses:

melhor resultado (menor  $(S(\vec{Z}))(S(\vec{Z}))$ ), pior resultado (maior  $(S(\vec{Z}))(S(\vec{Z}))$ ), média aritmética simples e desvio padrão da velocidade  $(U)$

$(U)$  e do coeficiente de dispersão longitudinal  $(E_L)(E_L)$ , assim como para o valor e número de avaliações da função objetivo  $(S(\vec{Z})) (S(\vec{Z}))$ .

### Estimativa da Velocidade $(U)$

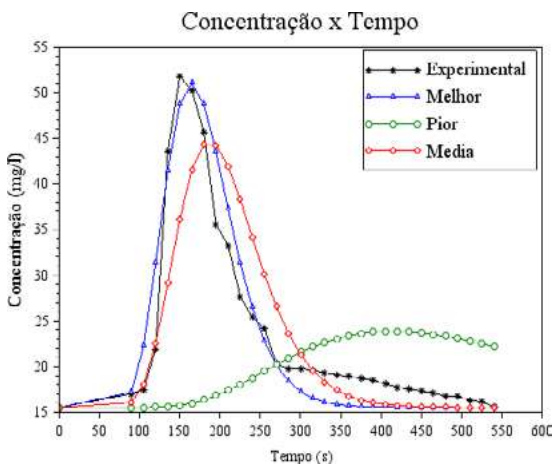
A seguir, são salientados os resultados para a estimativa da velocidade  $(U)(U)$  usando o método Levenberg-Marquardt. Na Tabela 2, é apresentado o resumo das informações estatísticas para o parâmetro de interesse e, na Figura 3, é mostrado o gráfico das curvas obtidas computacionalmente para o perfil de concentração ao longo do tempo, levando em consideração o melhor, pior e média dos valores obtidos para  $U$  após as 100 execuções do algoritmo. O coeficiente de dispersão longitudinal  $(E_L)(E_L)$  foi mantido fixo e igual a  $1,82 \text{ m}^2/\text{s}, 1,82 \text{ m}^2/\text{s}$ , da mesma maneira que os demais parâmetros citados nessa seção.



**Tabela 2:** Resultados obtidos para a velocidade ( $U$ ), ( $U$ ).

Parâmetros	$U(m/s)$	$S(\vec{Z})$
Melhor	0,590224	323,587388
Pior	0,223002	5425,668459
Médio	0,516569	1231,910495
Desvio Padrão	0,152267	1941,398293

**Figura 3:** Perfis das concentrações obtidas com a estimativa do parâmetro  $U$  pelo método Levenberg- Marquardt no melhor caso, pior caso, média e dados experimentais.



O baixo desvio padrão (0,152267) (0,152267) para o valor da velocidade ( $U$ ) ( $U$ ), encontrado ao longo das 100 execuções, evidencia a boa precisão e desempenho satisfatório do método de otimização Levenberg-Marquardt para esta série de problemas.

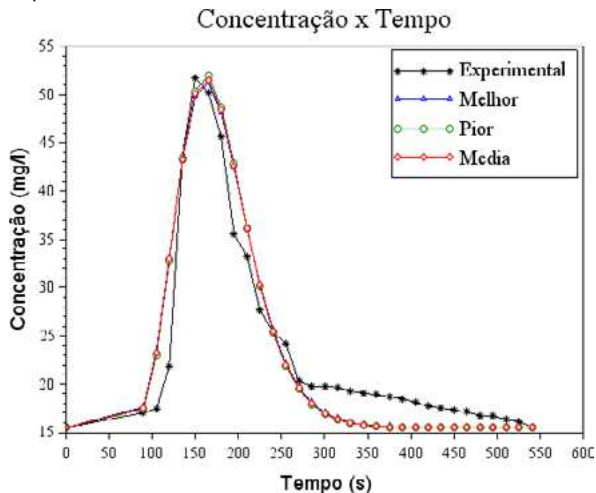
#### Estimativa do Coeficiente de Dispersão Longitudinal ( $E_L$ )

Na sequência, são evidenciados os resultados para a estimativa do coeficiente de dispersão longitudinal ( $E_L$ ), ( $E_L$ ). Assim, na Tabela 3 são mostrados os resultados obtidos pelo método Levenberg-Marquardt, tal como o gráfico, Figura 4, contendo as curvas para o perfil da concentração ao longo do tempo feitas com o melhor, pior e média dos valores obtidos após as 100 execuções do algoritmo. A velocidade ( $U$ ) ( $U$ ) permaneceu fixa e igual a  $0,60 m/s$ , do mesmo modo que os demais parâmetros abordados nesta seção.

**Tabela 3:** Resultados obtidos para coeficiente de dispersão longitudinal ( $E_L$ ), ( $E_L$ ).

Parâmetros	$E_L(m^2/s)$	$S(\vec{Z})$
Melhor	1,879151	336,417366
Pior	1,804638	337,893279
Médio	1,846805	336,839099
Desvio Padrão	0,023844	0,404773

**Figura 4:** Perfis das concentrações obtidas com a estimativa do parâmetro  $E_L$ , pelo método Levenberg-Marquardt no melhor caso, pior caso, média e dados experimentais.



Novamente, é observado um baixo desvio padrão (0,023844) (0,023844), achado no decorrer das 100 execuções, para o valor do coeficiente de dispersão Longitudinal ( $E_L$ ), ( $E_L$ ), o que evidencia o bom desempenho do método de otimização Levenberg-Marquardt para este tipo de problemática.

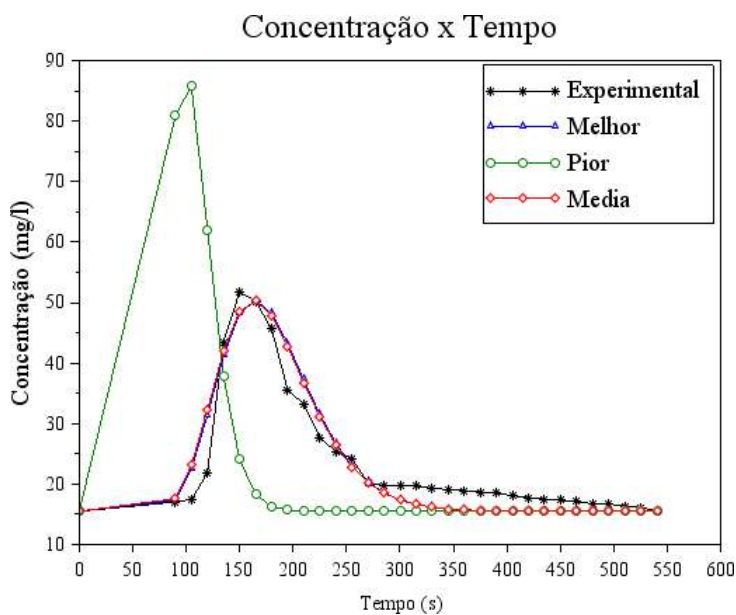
## Estimativa da Velocidade ( $U$ ) e do Coeficiente de Dispersão Longitudinal ( $E_L$ )

Nessa seção são mostrados os resultados obtidos pelo método Levenberg-Marquardt para a estimativa da velocidade ( $U$ ) e do coeficiente de dispersão longitudinal ( $E_L$ ), simultaneamente. As informações estatísticas envolvendo as estimativas dos parâmetros de interesse são descritas na Tabela 4, já o gráfico, Figura 5, exibe as curvas para o perfil da concentração ao longo do tempo feitas com o melhor, pior e média dos valores obtidos após as 100 execuções do algoritmo. Os outros parâmetros abordados no início desta seção foram mantidos constantes.

**Tabela 4:** Resultados obtidos para a estimativa da velocidade ( $U$ ) e do coeficiente de dispersão longitudinal ( $E_L$ ).

Parâmetros	$U(m/s)$	$E_L(m^2/s)$	$S(\vec{Z})$
Melhor	0,589341	1,876560	322,994827
Pior	0,988958	2,000000	14232,904809
Médio	0,593822	1,930021	1726,264413
Desvio Padrão	0,153576	0,071352	3497,073157

**Figura 5:** Perfis das concentrações obtidas com a estimativa dos parâmetros  $U$  e  $E_L$  pelo método Levenberg-Marquardt no melhor caso, pior caso, média dados experimentais.



Apesar de ser constatado um baixo desvio padrão (0,153576 e 0,071352) (0,153576 e 0,071352), ao longo das 100 execuções, para as estimativas da velocidade ( $U$ ) e do coeficiente de dispersão longitudinal ( $E_L$ ), o que enfatiza, outra vez, o bom desempenho do método de otimização Levenberg-Marquardt para este tipo de problemática.

## CONCLUSÕES

O presente trabalho buscou simular e investigar o transporte e comportamento de poluentes em corpos hídricos e sua consequente modelagem, utilizando conceitos matemáticos. Em seguida, para tal fim, buscou-se analisar a influência da variação do coeficiente de dispersão longitudinal e de velocidade no do rio São Pedro, afluente do rio Macaé, que está situado em São Pedro da Serra, 7º Distrito de Nova Friburgo, estado do Rio de Janeiro.

Por fim, foi realizada a estimativa dos parâmetros  $U$  e  $E_L$  usando o método de otimização irrestrita Levenberg-Marquardt. Onde, tanto de forma individual, quanto simultânea, pode-se concluir que, ao utilizar o método de otimização Levenberg-Marquardt com o objetivo de analisar o comportamento de uma pluma de contaminante em corpos hídricos ao longo do tempo, a partir de um ponto de interesse, foi possível representar o perfil numérico da concentração ao longo do tempo de maneira próxima aos dados experimentais, demonstrando um bom desempenho do referido método.

## REFERENCIAS

CAMPBELL, M. G. C. *Estimativa de parâmetros em corpos hídricos utilizando modelagem matemática e o método Levenberg-Marquardt*. 2019. Trabalho de Conclusão de Curso (Licenciatura em Matemática) – Instituto do Noroeste Fluminense de Educação Superior, Universidade Federal Fluminense, Santo Antônio de Pádua, RJ, 2019.

Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA). *Manejo de recursos hídricos*. Educação e ensino de ciências e matemática: Pesquisa, Aplicação e Novas Tendências, v. 2, n. 1, 2022. Disponível em: <<https://www.embrapa.br/tema-manejo-de-recursos-hidricos/perguntas-e-respostas#:~:text=Recurso%20h%C3%ADdrico%2C%20por%20sua%20vez,toda%20%C3%A1gua%20%C3%A9%20recurso%20h%C3%ADdrico.>>>. Acesso em: 17 de Fev. de 2023.

Ferreira, F. F., Neto, R. M. C & da Silva Neto, A. J. *Parâmetros da Equação de van Genuchten Obtidos pelo Método de Levenberg-Marquardt com Restrições*. 2019. XXII Encontro Nacional de Modelagem Computacional - XXII ENMCAt: Juiz de Fora, MG, 2019.

PLANALTO. *Constituição da República Federativa do Brasil de 1988*. 1988. Disponível em:

<[https://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm)>. Acesso em: 17 de Fev. de 2023.

SOS Mata Atlântica (SOSMA). *Observando os rios - o retrato da qualidade da água nas bacias hidrográficas da Mata Atlântica*. 2022. Disponível em: <[https://cms.sosma.org.br/wp-content](https://cms.sosma.org.br/wp-content/uploads/2022/03/SOSMA-Observando-os-Rios-2022.pdf)

[/uploads/2022/03/SOSMA-Observando-os-Rios-2022.pdf](https://cms.sosma.org.br/wp-content/uploads/2022/03/SOSMA-Observando-os-Rios-2022.pdf)>. Acesso em: 17 de Fev. de 2023.

SOUSA, E. P. *Avaliação de mecanismos dispersivos em rios através de problemas inversos*. 2009. Dissertação (Mestrado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, RJ, 2009.

TELLES, W. R. *Simulação do transporte horizontal bidimensional de substância conservativa*. 2009. Dissertação (Mestrado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, RJ, 2009.

## DADOS DOS AUTORES

**Anna Luisa Soares Castro Coimbra**. Mestre em Modelagem Computacional em Ciência e Tecnologia da Escola de Engenharia Industrial Metalúrgica de Volta Redonda - MCCT/EEIMVR/UFF. E-mail: [anna\\_luisa@id.uff.br](mailto:anna_luisa@id.uff.br)

**Fábio Freitas Ferreira**. Doutor em Modelagem Computacional pela Universidade do Estado do Rio de Janeiro (UERJ), professor associado no Departamento de Ciências da Natureza (RCN) da Universidade Federal Fluminense (UFF). E-mail: [fabiofreitasferreira@id.uff.br](mailto:fabiofreitasferreira@id.uff.br)

**Wagner Rambaldi Telles**. Doutor em Modelagem Computacional pela Universidade do Estado do Rio de Janeiro (UERJ), professor efetivo da Universidade Federal Fluminense, lotado no Departamento de Ciências Exatas, Biológicas e da Terra do Instituto do Noroeste Fluminense de Educação Superior - PEB/INFES/UFF. E-mail: [wtelles@id.uff.br](mailto:wtelles@id.uff.br)

# ANÁLISE DO ROMPIMENTO DE BARRAGENS DA UHE BARRA DO BRAÚNA UTILIZANDO O SOFTWARE DE SIMULAÇÃO HIDRODINÂMICA IBER

## ANALYSIS OF THE BARRA DO BRAUNA UHE DAM USING THE IBER HYDRODINAMIC SIMULATION SOFTWARE

Sânzzia de Oliveira e Souza Figueira

Thiago Jordem Pereira

Ricardo Silveira Sousa

Tibério Borges Vale

Jader Lugon Junior

Wagner Rambaldi Telles

**RESUMO:** O número de acidentes decorrentes do rompimento de barragens é cada vez maior, principalmente quando estas se localizam nas proximidades das regiões urbanizadas. Nos últimos anos, especialmente no Brasil, muitos deles geraram tanto impactos ambientais e socioeconômicos quanto de perdas humanas. Antes dessas construções, vários estudos são realizados como meio de minimizar ou mesmo não permitir que ocorram desastres futuros. Desse modo, em virtude de uma provável catástrofe envolvendo as barragens, foi realizada uma simulação da hipotética ruptura da barragem da Usina Hidrelétrica Barra do Braúna, situada no rio Pomba, tendo a cidade de Santo Antônio de Pádua-RJ como estudo de caso. Os resultados obtidos, utilizando o software de simulação hidrodinâmica IBER, mostram que tal evento, caso venha a ocorrer, representa significativo risco para a população e construções localizadas a jusante à barragem.

**Palavras-chave:** Rompimento de Barragem. Usina Hidrelétrica Barra do Braúna. Santo Antônio de Pádua. Software IBER.

**ABSTRACT:** The number of accidents resulting from dam failures is increasing, especially when they are located close to urbanized regions. In recent years, especially in Brazil, many of them have generated both environmental and socioeconomic impacts and human losses. Before these constructions, several studies are carried out as a means of minimizing or even preventing future disasters from occurring. Thus, due to a probable catastrophe involving the dams, a simulation was carried out of the hypothetical rupture of the dam at the Barra do Brauna Hydroelectric Plant, located on the Pomba River, using the city of Santo Antônio de Pádua, Rio de Janeiro state, as a case study. The results obtained, using the IBER hydrodynamic simulation software, show that such an event, if it were to occur, represents a significant risk for the population and buildings located downstream of the dam.

**Keywords:** Dam Rupture. Barra do Brauna Hydroelectric Plant. Santo Antonio de Padua. IBER Software.

## INTRODUÇÃO

A água é considerada um dos recursos naturais indispensáveis na manutenção da vida de todo o planeta, seja relacionada à vida humana ou ao meio ambiente. Quando vista em termos quantitativos, sua escassez prejudica no modo como é utilizada pelos seres vivos, bem como nos sistemas comportamentais e fisiológicos desenvolvidos a fim de se adequar às diversas condições ambientais trazidas pela gama da distribuição temporal e especial dos recursos hídricos (MENDONÇA, 2013).

Um dos maiores problemas enfrentados pelo mundo se baseia no planejamento e na execução desses recursos hídricos a fim de que a água não se torne escassa e chegue a toda população, sem exceção. Diante dessa situação, inúmeras barragens foram construídas, contendo diversas vantagens em seu uso, sendo a principal delas, o fornecimento de energia elétrica para toda a população.

Em contrapartida, devido à variedade de tamanhos, bem como seus usos, nem todas as barragens são mantidas em perfeitas condições, não atendendo a normas de segurança, o que pode resultar no rompimento da estrutura.

Diante disso, os estudos envolvendo a ruptura de barragens vêm aumentando, despertando o interesse de órgãos ambientais como forma de tornar eficaz as técnicas seguras de manutenção e de operação das mesmas (TSCHIEDEL, 2022).

Por meio dos avanços tecnológicos, Santos et al. (2003) estabelecem que os modelos matemáticos e físicos passaram a ser utilizados como modo de responder aos efeitos de ruptura da barragem. Assim, o uso de tais modelos simplifica no momento em que corre a simulação, o que, consequentemente, favorece no estudo das barragens.

De modo especial, a presente pesquisa tem como objeto de estudo a análise do rompimento hipotético da barragem da Usina Hidrelétrica (UHE) Barra do Braúna, localizada no estado de Minas Gerais, através do software IBER, tendo como estudo de caso a cidade de Santo Antônio de Pádua, situada no estado do Rio de Janeiro.

Para a estrutura do trabalho, são apresentados os resultados e discussões baseados na construção do modelo digital da região de interesse, da variação do coeficiente de rugosidade e da calibração e validação do modelo para o cenário de rompimento hipotético, assim como a análise dos resultados numéricos e suas respectivas conclusões.

## MATERIAIS E MÉTODOS

Nesta seção é caracterizada a UHE Barra do Braúna, bem como o processo de obtenção de dados hidrometeorológicos da região de interesse e o software de simulação hidrodinâmica IBER.

### UHE Barra do Braúna

A UHE Barra do Braúna situa-se no rio Pomba e o eixo de sua barragem localiza-se entre as cidades mineiras de Laranjal e Recreio, além de estar próxima de cidades fluminenses, como Santo Antônio de Pádua, Itaocara, Aperibé e Cambuci. As águas de seu reservatório também ocupam as cidades de Leopoldina e Cataguases, no estado de Minas Gerais. A máxima vazão de projeto dessa barragem é de 3.296 m<sup>3</sup>/s, o que significa que vazões excedentes representam o seu quesito de emergência.

A cidade escolhida como base para o estudo dos efeitos do rompimento hipotético da barragem da UHE Barra do Braúna, Santo Antônio de Pádua, deve-se ao fato da mesma ser a primeira cidade que se encontra a jusante da barragem e que possui o rio Pomba passando dentro de seus limites, dividindo-a ao meio. Um rompimento da barragem atingiria a área mais urbanizada da cidade, causando enormes danos a toda a população.

Santo Antônio de Pádua possui cinco pontes localizadas em sua área urbana, as quais ligam uma margem à outra do rio Pomba. Três delas são destinadas à passagem de veículos e duas pontes de pedestres. Além destas, há outras três outras pontes de interesse situadas no trecho do rio Pomba entre a represa e a referida cidade, mais especificamente, nos distritos de Paraoquena (ainda em Santo Antônio de Pádua), Cisneiros e Itapiruçu, os quais pertencem ao município de Palma, estado de Minas Gerais. A localização das referidas pontes é mostrada na Tabela 1. Por outro lado, na Figura 1 são representadas as pontes da zona urbana da cidade de Santo Antônio de Pádua.

Como forma de comparação dos resultados, foi tomado como base o nível d'água em cada uma das 6 pontes que há no trecho analisado, ou seja, entre a UHE Barra do Braúna e Santo Antônio de Pádua, as quais possibilitam o fluxo de veículos automotores. Nesse sentido, não foram consideradas as pontes Alberíades Gabry e Abel da Silva Malafaia, descrita na Tabela 1, uma vez que esta é destinada apenas a passagem de pedestres.

**Tabela 1:** Relação das pontes situadas entre a represa e a cidade de Santo Antônio de Pádua.

Descrição da Ponte	Cidade	Localização (UTM WGS84 23S)
Cisneiros	Palma-MG	773143,403; 7626548,468
Itapiruçu	Palma-MG	774530,326; 7624890,031
Paraoquena-Campelo	Santo Antônio de Pádua-RJ	784268,138; 7620655,88
Badih Chicralla	Santo Antônio de Pádua-RJ	789371,786; 7617911,681
Alberíades Gabry	Santo Antônio de Pádua-RJ	791782,059; 7615496,07
Abel da Silva Malafaia	Santo Antônio de Pádua-RJ	792016,947; 7615189,114
Raul de Moraes Veiga	Santo Antônio de Pádua-RJ	792395,971; 7614420,389
Paulino Padilha	Santo Antônio de Pádua-RJ	773143,403; 7626548,468

Fonte: A Autora (2023).

Figura 1 – Pontes localizadas na zona urbana de Santo Antônio de Pádua.



Fonte: Adaptado de ENGEORPS (2012).

## MODELAGEM MATEMÁTICA DO PROBLEMA PROPOSTO

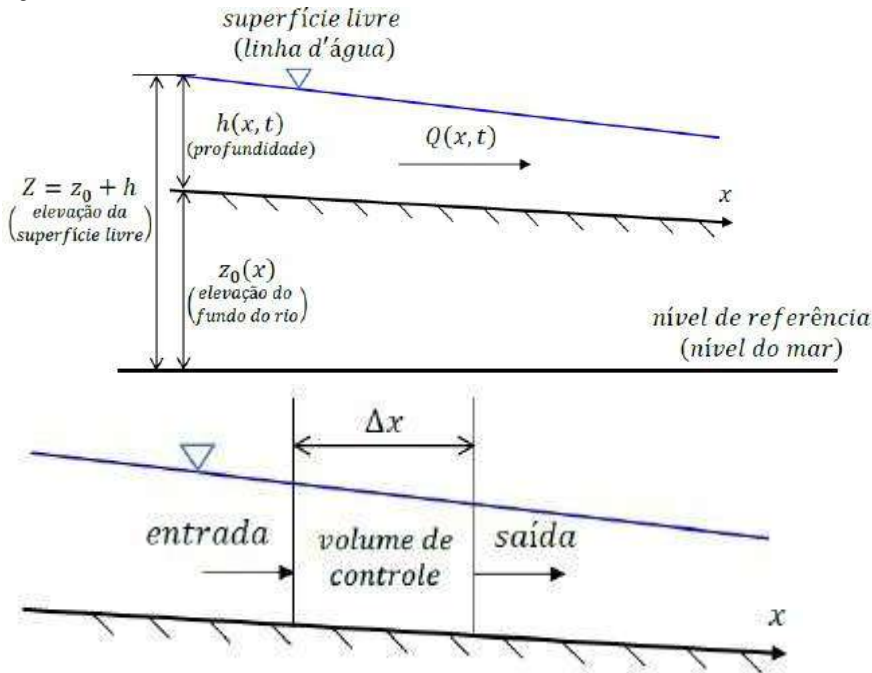
A melhor forma de compreender o processo de escoamento superficial em bacias e rios é analisar os cenários de intervenção, prever vazões, dentre outros fatores. Esses estudos são indispensáveis, por exemplo, na verificação das inundações e estiagens.

Nesse sentido, a propagação da vazão em cursos d'água é, de modo geral, caracterizada através de um grupo de duas equações, denominadas equações de Saint Venant, podendo também ser chamadas de equações de escoamento não permanente unidimensional (FAN; PONTES; PAIVA, 2014).

Além disso, Simões et al. (2011) enfatizam que as equações de Saint Venant são utilizadas em sua forma unidimensional a fim de prever enchentes e inundações, com destaque para o comportamento de ondas de cheias nos canais e rios e de correntes marítimas, formação de ressaltos hidráulicos, entre outros eventos.

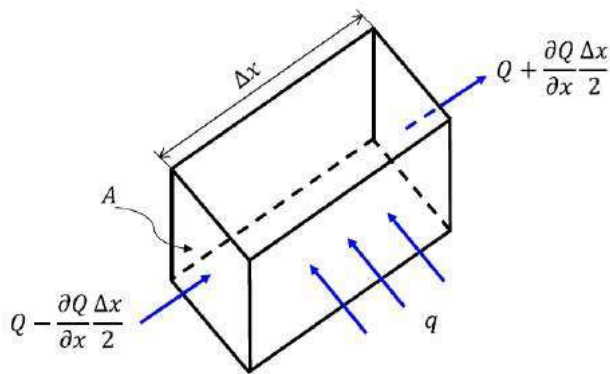
De acordo com Steinstrasser (2005), as equações de Saint Venant são deduzidas por meio das equações da Continuidade e da Quantidade de Movimento (Momentum), ambas aplicadas a um volume de controle, estabelecido dentro do fluido em movimento, mostrado na Figura 2. Em particular, a equação de Momentum é aquela que, no decorrer de todo o escoamento, representa a conservação da massa de líquido no conduto. Na Figura 3 é representado um volume de controle elementar, com ênfase ao balanço a partir do qual são derivadas as equações da Continuidade e Momentum, onde  $x$  refere-se à medida ao longo do canal.

Figura 2 – Volume de controle elementar dentro de um fluido em movimento.



Fonte: A Autora (2023).

Figura 3 – Volume de controle a partir do qual são derivadas as equações da Continuidade e Momentum.



Fonte: A Autora (2023).

Assim, as equações de Saint Venant são definidas, simultaneamente, a partir das equações da Continuidade (equação (1)) e da Quantidade de Movimento (equação (2)):

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0 \quad (1)$$

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left( \frac{Q^2}{A} \right) + gA \frac{\partial h}{\partial x} = gAS_0 - gAS_f \quad (2)$$

onde  $x$  e  $t$  são as variáveis do espaço e tempo, respectivamente;  $Q$  é a descarga do canal;  $A$  é a área da seção transversal;  $g$  é a aceleração da gravidade;  $h$  é a altura da lâmina d'água;  $S_0$  é a declividade do fundo do canal;  $S_f$  é a declividade da linha de energia, sendo expressas conforme as equações (3)-(4):

$$S_0 = -\frac{\partial z_0}{\partial x} \quad (3)$$

$$S_f = \frac{Q|Q|n^2}{A^2 R^{4/3}} \quad (4)$$

onde  $z_0$  é a elevação do fundo do canal;  $n$  é o coeficiente de rugosidade de Manning;  $R$  é o raio hidráulico.

Uma vez que  $A = Bh$  e  $B = uA$ , onde  $B$  é a largura da superfície livre e  $u$  é a velocidade média do escoamento, dividindo a equação (1) por  $B$ , bem como a equação (2) por  $A$ , as equações de Saint Venant podem ser representadas como nas equações (5)-(6):

$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} = 0 \quad (5)$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + g \frac{\partial h}{\partial x} = g(S_0 - S_f) \quad (6)$$

Para a análise do rompimento hipotético da barragem da UHE Barra do Braúna, neste trabalho, é utilizado o software IBER, o qual é um modelo que simula o fluxo turbulento, de superfície livre, em escoamento não permanente. Dentre alguns de seus campos de aplicação estão: ruptura de barragens; avaliação das zonas de inundação; cálculo hidráulico das redes de canais e de canalização; fluxo de marés em estuários; transporte de sedimentos; qualidade da água; e outros (SARTORI, 2018).

O IBER compõe-se em diferentes módulos ligados entre si, como por exemplo, o módulo hidrodinâmico, que determina características do escoamento, tais como a velocidade e a profundidade da água; módulo de turbulência; e módulo de transporte de sedimentos (CHARGEL, 2018).

Na solução numérica do modelo matemático presente no software IBER é usada a esquematização baseada no Método dos Volumes Finitos (MVF) no domínio físico, possibilitando a integração do modelo digital do terreno, bem como a construção da malha de cálculo não estruturada de volumes finitos.

## DADOS HIDROMETEOROLÓGICOS

O Brasil é um país que apresenta uma grande extensão territorial, todavia, os períodos de ocorrências das chuvas e estações de seca são bem definidos. Por isso, a enorme necessidade de se acompanhar todas as formas de chuva, realizando a gestão eficiente dos reservatórios, a fim de que os mesmos garantam que haja recursos hídricos, mesmo em épocas mais secas.

Nesse sentido, Sousa et al. (2012) definem o Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH) como a Política Na-

cional de Recursos Hídricos, inscrito na Lei nº 9433/1997, que coleta, trata, armazena e recupera todos os recursos hídricos.

No Brasil, existem dois bancos de dados ambientais de suma importância: Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP) e Portal Hidroweb, mantido pela Agência Nacional de Águas e Saneamento Básico (ANA). Ambos são de livre acesso e gratuito (ROSSETTI, 2008).

Por meio do aplicativo do Hidroweb, as estações hidrometeorológicas de interesse são selecionadas, sendo que cada uma delas apresenta um período diferente de funcionamento. É possível, por exemplo, fazer download dos dados das estações de interesse nos formatos do Access (.mdb), texto (.txt) e planilha eletrônica (.csv).

Ao selecionar uma estação de interesse é possível verificar o código, a bacia, a sub-bacia, o rio, o estado, o município, o responsável, a operadora e o tipo da estação. Quando há dados telemétricos disponíveis, é possível definir o período inicial e final da pesquisa, não excedendo 90 dias. Em particular, na região situada entre a barragem da UHE Barra do Braúna e a cidade de Santo Antônio de Pádua, há quinze estações, as quais dispõem de dados convencionais e/ou telemétricas.

No presente trabalho, como forma de calibrar e validar o modelo desenvolvido no software IBER, foram utilizados dados de duas estações de monitoramento fluviométrico localizadas a jusante da UHE Barra do Braúna, no período entre o dia 24 de janeiro de 2020 a 28 de janeiro de 2020:

- Estação 58788600: localizada próximo à barragem e utilizada como condição de contorno (dado de entrada) no modelo;
- Estação 58774000: localizada próximo à zona urbanizada de Santo Antônio de Pádua, de operacionalização do INEA, e utilizada para comparação com os resultados numéricos obtidos por meio das simulações no software IBER.

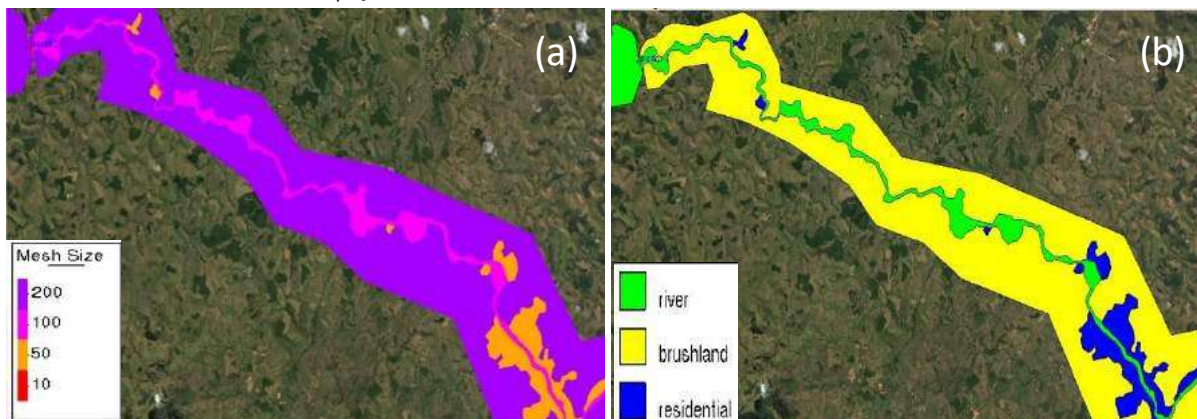
## RESULTADOS E DISCUSSÕES

Para as simulações realizadas nesse trabalho, foi definido um domínio de interesse entre a barragem e o município de Santo Antônio de Pádua, sendo adotada uma malha não estruturada com elementos triangulares de diferentes tamanhos de acordo com o uso e ocupação do solo. Na região que delimita o reservatório e planície,



foi utilizada uma malha não estruturada com elementos triangulares de tamanho 200 m enquanto na barragem, rio e área urbanizada, foram adotados, respectivamente, elementos de tamanhos iguais a 10 m, 100 m e 50 m. Na Figura 4 é mostrado domínio de interesse.

Figura 4 – (a) Representação das superfícies e seus respectivos elementos e (b) caracterização das superfícies da região de estudos de acordo com o uso e ocupação do solo.



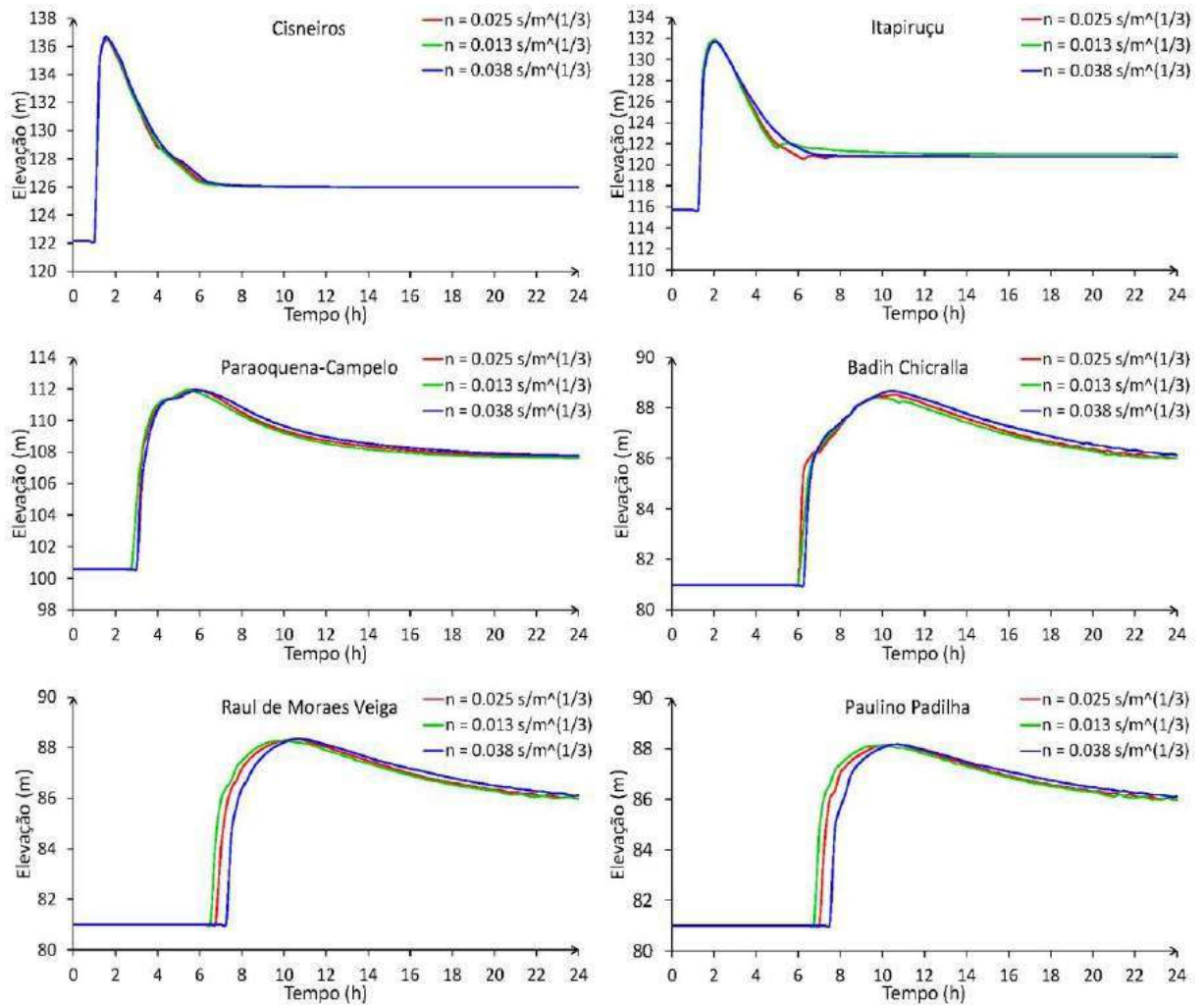
Fonte: A Autora (2023).

## VARIAÇÃO DO COEFICIENTE DE RUGOSIDADE

Foi realizada a análise do coeficiente de rugosidade ao longo do curso d'água do rio Pomba na região de estudos. Para isso, foi utilizada uma malha não estruturada com elementos triangulares de tamanho 200 m no reservatório e planície, enquanto na barragem, rio e área urbanizada, adotou-se, respectivamente, elementos de tamanhos iguais a 10 m, 100 m e 50 m. Aos coeficientes de rugosidade de Manning, foram atribuídos os valores de  $0,025 \text{ s/m}^{1/3}$  (reservatório, barragem e rio),  $0,050 \text{ s/m}^{1/3}$  (planície) e  $0,150 \text{ s/m}^{1/3}$  (área urbanizada).

Com base nisso, foi analisada a variação simultânea do coeficiente de rugosidade do reservatório, barragem e rio. Para isso, aplicou-se uma alteração de  $\pm 50\%$  do valor base ( $0,025 \text{ s/m}^{1/3}$ ), obtendo com isso, os valores de  $0,013 \text{ s/m}^{1/3}$  e  $0,038 \text{ s/m}^{1/3}$ , respectivamente. Por outro lado, os valores do coeficiente de rugosidade de Manning da planície ( $0,050 \text{ s/m}^{1/3}$ ) e área urbanizada ( $0,150 \text{ s/m}^{1/3}$ ), mantiveram-se inalterados. Na Figura 5, são mostrados os valores dos níveis d'água após a variação dos coeficientes.

Figura 5 – Perfis das elevações da lâmina d'água nas 6 pontes com base na variação do coeficiente de rugosidade do reservatório, barragem e rio.



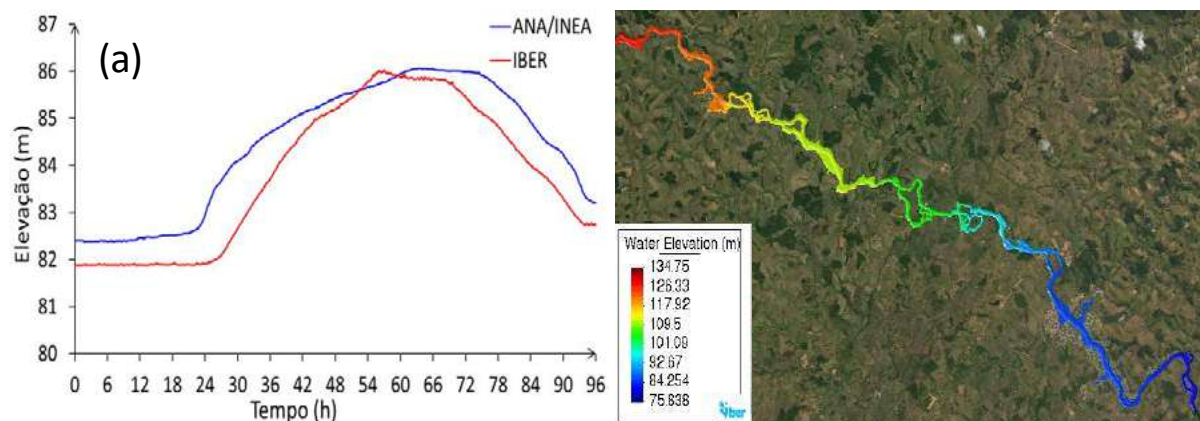
Fonte: A Autora (2023).

Verifica-se que a variação do coeficiente de rugosidade do reservatório, barragem e rio não ocasionou um impacto significativo no tempo de chegada da onda de cheia nos pontos onde se realizou a comparação das elevações da lâmina d'água. As maiores alterações ocorreram nas pontes localizadas mais à jusante da barragem.

## CALIBRAÇÃO E VALIDAÇÃO DO MODELO

Para a validação e calibração do software IBER para a região de interesse, foram utilizados os dados das estações de monitoramento da ANA, a fim de caracterizar o evento de cheia do rio Pomba no período compreendido entre 24 e 28 de janeiro de 2020. Na Figura 6, é apresentada a comparação entre os resultados numéricos obtidos no software IBER e os dados experimentais provenientes da estação de monitoramento da ANA/INEA.

Figura 6 – (a) Perfis da elevação da lâmina d'água na estação da ANA/INEA e (b) perfis da elevação máxima da lâmina d'água no rio Pomba na região analisada.



Fonte: A Autora (2023).

De acordo com a Figura 6, verifica-se que o modelo desenvolvido no software IBER simulou o pico da onda de cheia de forma satisfatória quando comparado com os dados experimentais obtidos no site da ANA/INEA, uma vez que representou o pior momento da onda de cheia na cidade de Santo Antônio de Pádua. Todavia, desconsiderou os níveis d'água antes e após o referido pico.

Afere-se, ainda, que a maior profundidade da lâmina d'água atingida tanto no software, quanto nos dados obtidos na estação da ANA/INEA, localizada próxima à zona urbanizada de Santo Antônio de Pádua, esteve próxima a 6 m, a qual ficou estabelecida entre um dado instante de tempo de 55 h a 65 h, aproximadamente, com elevação em torno de 86 m.

### SIMULAÇÃO DE CENÁRIOS HIPOTÉTICOS DE ROMPIMENTO DE BARRAGEM

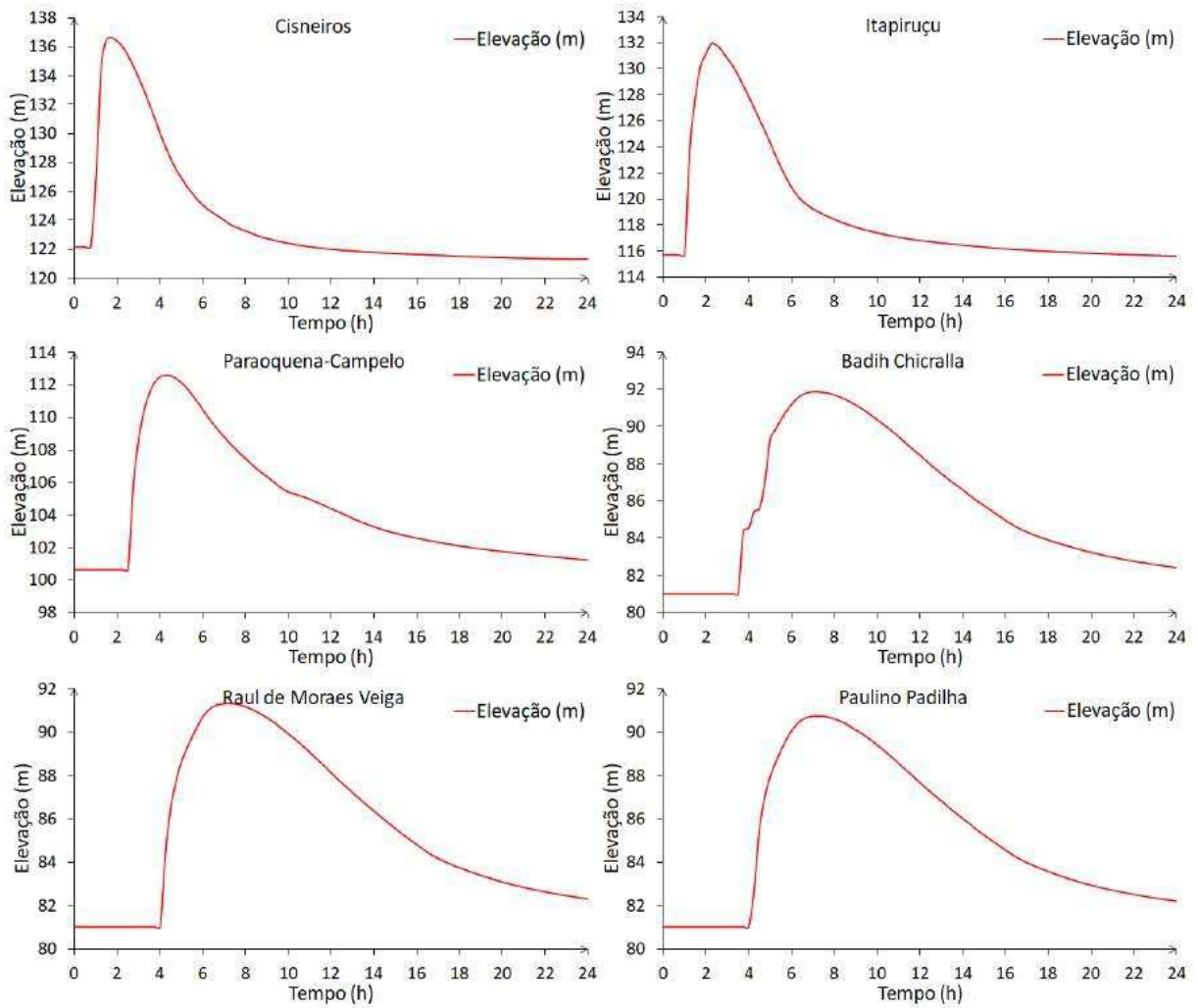
Baseado na configuração descrita anteriormente, foi simulado um cenário para o rompimento hipotético da barragem da UHE Barra do Braúna, para o qual foi tomado como referência os dados disponíveis no trabalho de Daru et al. (2013), o qual contém as seguintes informações: (i) largura média da brecha (m): 91,5; (ii) tempo de ruptura da barragem (h:min): 1:00; (iii) elevação da lâmina d'água no reservatório (m): 154,5.

Para a simulação desse cenário, foi considerado um tempo total de simulação igual a 24 horas. Já o rompimento total variou conforme o tempo de ruptura, contados a partir do instante inicial da simulação, o qual também foi considerado como o início da formação da brecha. Na Figura 7, são mostrados os perfis das elevações da lâmina d'água para o cenário simulado.

Além dos perfis das elevações da lâmina d'água nos locais onde as seis pontes de passagem de veículos ao longo do trecho analisado estão inseridas, levando em consideração o cenário simulado, também foi realizada uma análise dos mapas de máximos para elevações da lâmina d'água, representada na Figura 8.

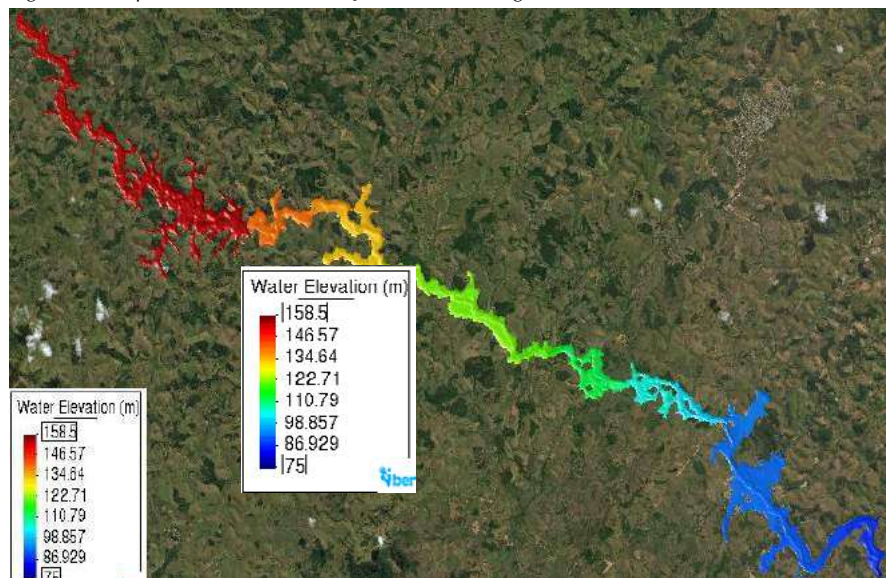
De acordo com as simulações referentes ao perfil da elevação da lâmina d'água nas pontes, foi possível constatar um aumento significativo nas elevações, independentemente da ponte investigada.

Figura 7 – Perfis das elevações da lâmina d'água no cenário simulado com base nas 6 pontes ao longo de 24 h.



Fonte: A Autora (2023).

Figura 8 – Mapa de máximo da elevação da lâmina d'água referente ao cenário simulado.



Fonte: A Autora (2023).

## CONCLUSÕES

A análise do rompimento hipotético da barragem UHE Barra do Braúna mostrou a necessidade indispensável da fiscalização dos responsáveis técnicos no decorrer e, posteriormente, à sua construção, bem como por órgãos competentes, visto que a sua ocorrência gera problemas irreversíveis, na maior parte dos casos, em especial, à zona urbana de Santo Antônio de Pádua.

Além disso, baseando-se no evento de inundação real simulado no software IBER (dados das estações da ANA/INEA), constata-se que o modelo o representou de modo satisfatório, o que indica uma boa robustez da metodologia aqui proposta para a região investigada

Perante as simulações realizadas no software IBER, é possível assegurar a viabilidade do modelo utilizado para o estudo, considerando-o de grande valia de cunho ambiental, em virtude de promover a simulação dos diferentes cenários e, com isso, fazer do seu uso uma ferramenta auxiliar no processo para a tomada de decisões, em se tratando de eventos de ordem catastrófica, como as ondas de cheias, ocasionadas pelo volume de água oriundo do rompimento de uma barragem.

Assim, nessa perspectiva, todo o trabalho aqui exposto é considerado uma ferramenta importante a fim de avaliar o impacto ambiental, social, econômico, dentre outros, decorrentes das inundações provenientes da ruptura da referida barragem.

## REFERÊNCIAS

- CHARGEL, L. T. **Avaliação do uso de modelos hidrodinâmicos, suas limitações, efeitos de escala e da representação topográfica**. 2018. 296 f. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil, 2018.
- DARU, R. L.; SILVEIRA, R. M.; HAINOSZ, F. S.; BURAS, M. Propagação de ondas de ruptura e plano de ações emergenciais - UHE Barra do Braúna. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, Bento Gonçalves (RS), 2013. Anais do **XX Simpósio Brasileiro de Recursos Hídricos**. 2013. p. 1-8.
- ENGEORPS. **Previsão de Eventos Críticos na Bacia do Rio Paraíba do Sul, RSC - C01 - Relatório dos serviços de campo de cartografia**. Brasília: ANA, SUM/SPR, 2012.
- FAN, F. M.; PONTES, P. R. M.; PAIVA, R. C. D. Avaliação de um método de propagação de cheias em rios com aproximação inercial das equações de Saint-Venant. **RBRH – Revista Brasileira de Recursos Hídricos**, Universidade Federal do Rio Grande do Sul, Rio Grande do Sul, v. 19, n. 04, p. 137–147, 2014.
- MENDONÇA, N. S. **Sustentabilidade e usos múltiplos da água da barragem Poção da Ribeira- SE**. 2013. 225 f. Dissertação (Mestrado em Desenvolvimento e Meio Ambiente) – Universidade Federal de Sergipe, Sergipe, 2013.
- ROSSETTI, E. K. **Gestão ambiental: ecoeficiência a caminho da sustentabilidade: o caso do Segmento Moveleiro da Serra Gaúcha**. 2008. 120 f. Dissertação (Mestrado em Administração) – Universidade de Caxias do Sul, Caxias do Sul, 2008.
- SANTOS, R. P.; FERREIRA, W. V. F.; PEREIRA, P. N.; MARTINS, J. R. S.; FADIGA JR, F. M.;
- SANTOS, R. C. P. Modelação da propagação de cheias ocasionadas por rompimento de barragens na cadeia de geração do rio Paranapanema. In: II CITENEL, Porto Alegre (RS), 2003. Anais do **II Congresso de Inovação Tecnológica em Energia Elétrica**. 2003. p. 1-4.
- SARTORI, R. Z. **Avaliação comparativa de modelos hidrodinâmicos para previsão de inundações: um estudo de caso do município de Getúlio Vargas-RS**. 2018. 148 f. Dissertação

## DADOS DOS AUTORES

**Sânzzia de Oliveira e Souza Figueira**. Doutoranda em Modelagem e Tecnologia para Meio Ambiente Aplicadas em Recursos Hídricos, Instituto Federal Fluminense, IFF. E-mail: sanzzi.f@gsuite.iff.edu.br.

**Thiago Jordem Pereira**. Doutor em Modelagem Computacional, Docente, Universidade Federal Fluminense, UFF/INFES. E-mail: tjordem@id.uff.br.

**Ricardo Silveira Sousa**. Doutor em Ciências da Computação e Matemática Computacional, Docente, Universidade Federal Fluminense, UFF/INFES. E-mail: rsousa@id.uff.br.

**Tibério Borges Vale**. Doutor em Ciências (com ênfase em Astrofísica), Docente, Universidade Federal Fluminense, UFF/INFES. E-mail: tiberio\_vale@id.uff.br.

**Jader Lugon Junior**. Doutor em Modelagem Computacional, Docente, Instituto Federal Fluminense, IFF. E-mail: jlugonjr@gmail.com.

**Wagner Rambaldi Telles**. Doutor em Modelagem Computacional, Docente, Universidade Federal Fluminense, UFF/INFES. E-mail: wtelles@id.uff.br.

## ALGORITMOS DE AGRUPAMENTOS NÃO-SUPERVISIONADOS APLICADOS À TAXONOMIA E FILOGENIA DE *LEISHMANIA* spp.

Antonio Revail Alves Pereira

Marcello Goulart Teixeira

Flávio Luis de Mello

**RESUMO.** O presente trabalho pretende apresentar resultados de agrupamentos de espécies de *Leishmania* ao nível de complexos. Além disso, serão discutidas as abordagens de vetorização utilizadas nos algoritmos de clusterização *k-means*, *OPTICS* e *Fuzzy c-means*.

**Palavras-chave.** Taxonomia, Filogenia, Aprendizado de Máquina, Clusterização, Lógica Fuzzy

**ABSTRACT.** The present work aims to present results of clustering species of *Leishmania* at the complex level. Additionally, the vectorization approaches used in the clustering algorithms *k-means*, *OPTICS*, and *Fuzzy c-means* will be discussed.

**Keywords.** Taxonomy, Phylogeny, Machine Learning, Clustering, Fuzzy Logic

## INTRODUÇÃO

A taxonomia biológica é parte da biologia sistemática, que se dedica a análise das relações de parentesco entre organismos. Uma vez que se decide a árvore filogenética dos organismos e se conhecem os ramos evolutivos, a taxonomia se encarrega de estudar as relações de parentesco (OLIVERA, 2011).

Aristóteles foi o primeiro pensador a classificar seres vivos, mas seu critério não é bem detalhado. Ele dividiu os animais em grandes grupos e, em certos casos, propôs algumas subdivisões. Um certo táxon era descrito inteiramente por seus aspectos físicos externos (RODRIGUES; HIDALGO, 2022). Apenas em 1735 o cientista sueco Carl N. Linnaeus (Lineu) propõe uma classificação mais refinada para seres vivos. Em seu trabalho intitulado *Systema Naturae* de 1758, ele estabelece como critério de classificação características estruturais e anatômicas. Lineu era criacionista e acreditava que o número de espécies era fixo e imutável assim como foram criados pela entidade divina (KLEPKA; CORAZZA, 2018).

O sistema criado por Lineu é rígido e não considera que os seres vivos estejam em constante evolução. A observação do mundo natural precisa admitir incerteza, instabilidade e mutabilidade porque os fenômenos naturais dificilmente apresentam contornos bem definidos, uma vez que se encontram em constantes mudanças.

Escolhemos, como nosso modelo inicial, um táxon parasito de grande importância médica, o gênero *Leishmania*, que ainda apresenta controvérsias em relação à posição taxonômica de algumas espécies.

Os morfotipos foram descritos em 1903 pelo médico britânico Ronald Ross, sendo o nome uma homenagem ao patologista escocês William Boog Leishman (ROSS, 1903) que, em maio do mesmo ano, descreveu pequenos corpos ovais encontrados no baço de um soldado morto após apresentar sintomas como febre alta, disenteria crônica e caquexia.

Várias tentativas de classificações foram propostas. Inicialmente tentou-se classificar as *leishmanias* usando as classificações lineanas monotéticas. Em 1908 Charles Nicolle separou a *L. infantum*, o agente etiológico da leishmaniose visceral do Mediterrâneo, da *L. donovani*, agente etiológico da kala-azar. Em 1982, (SAFJ'ANOVA, 1982) propõe dois subgêneros para a espécie baseando-se em características biológicas, chamando de *Leishmania* os parasitas de mamíferos e de *Sauroleishmania* os de répteis.

Uma proposta diferente foi enunciada em 1987 por Jafrey Shaw e Ralph Lainson que classificaram os subgêneros conforme o local de desenvolvimento dos parasitos no trato digestivo do flebotomíneo (inseto vetor)(GOSSAGE *et al.*, 2003).

A partir da década de 1980 a classificação fenética, que se baseia em múltiplas características de similaridades, aplicadas conjuntamente sem qualquer hipótese *a priori*, tem sido empregada na classificação de *Leishmania*. Em 2000, uma nova classificação foi proposta baseada em dados moleculares combinados, que dividem as espécies de *leishmanias* em duas principais linhagens filogenéticas chamadas de *Euleishmania* e *Paraleishmania* (CUPOLILLO *et al.*, 2000). Segundo (SCHÖNIAN *et al.*, 2018), dados moleculares baseados em sequências de diferentes alvos e em MLMT (Multilocus Microsatellite Typing) não suportam o conceito de complexos de espécies apresentadas por Laison e Shaw, e a classificação foi revisada, incluindo tanto a supressão de várias espécies como também a redução de algumas espécies ao nível de subespécies (AKHOUNDI *et al.*, 2016). As espécies analisadas neste trabalho foram as do gênero *Leishmania* subgêneros *Leishmania*, *Vianniae* *Mundia*.

Com o desenvolvimento das tecnologias de sequenciamento genético, aplicadas nos estudos em filogenia molecular que analisa diferenças moleculares, principalmente nas sequências de DNA, permite obter informações sobre as relações evolutivas entre organismos. Como consequência, revisões em taxonomias têm sido cada vez mais presentes na literatura. O que antes era feito observando aspectos morfológicos, atualmente consideram-se propriedades moleculares. Como os dados moleculares produzidos são muito volumosos o que dificulta seu processamento, algoritmos de aprendizado de máquina são cada vez mais utilizados em estudos moleculares. Não só pela velocidade que esses problemas são tratados como pela qualidade dos resultados obtidos.

Aprendizado de máquina (AM) é uma área de Inteligência Artificial que para (MONARD; BARANAUSKAS, 2003) tem o objetivo de desenvolver técnicas computacionais sobre aprendizado, bem como a construção de sistemas capazes de adquirir conhecimento de forma automática. Algoritmos de aprendizado não supervisionado identificam semelhanças nos dados automaticamente, sem a presença de um operador.

No contexto de AM a taxonomia é considerada um problema de clusterização. Clusterização consiste em, dado uma base de dados,



agrupar os objetos de modo que os objetos mais similares fiquem no mesmo grupo e objetos menos similares sejam alocados para grupos distintos (OCHI *et al.*, 2004).

Para este trabalho usamos os algoritmos *k-means*, *OPTICS* e *Fuzzy c-means*. Para todos esses foram apresentadas propostas de taxonomias. O *Fuzzy c-means* além de realizar agrupamentos, ele apresenta o grau de pertinência dos objetos em relação aos grupos, que incorporados à filogenia como uma maneira de representar a imprecisão presente no processo evolutivo, fazendo com que o conceito de espécies não tenham limites rígidos como é feito atualmente.

## PREPARAÇÃO DOS DADOS

O conjunto de dados foram coletados no site (NCBI, ) do Instituto Nacional de Saúde americano. Sequências de DNA dos cromossomos 10, 28 e 31 das *leishmanias* *L.aethiopica*, *L.arabica*, *L.braziliensis*, *L.donovani*, *L.enriettii*, *L.infantum*, *L.major*, *L.mexicana*, *L.panamensis* e *L.peruviana* foram adquiridos em formato *.fasta*. Os cromossomos 28 e 31 foram escolhidos por terem os genes *hsp70* e o *6pg*, característicos da espécie, e o cromossomo 10 foi escolhido aleatoriamente.

A fase seguinte consistiu em realizar o alinhamento das sequências, que significa deixar todas as sequências com a mesma quantidade de bases. Pelas características dos dados, precisávamos usar um programa que realizasse alinhamentos múltiplos de sequências longas. O Mauve foi o programa que atendeu a essas especificidades do conjunto de dados. Mauve é um sistema que realiza e integra a análise de eventos evolutivos em larga escala ao alinhamento tradicional de múltiplas sequências (DARLING *et al.*, 2004).

Por contarmos com apenas 10 sequências, uma para cada espécie, foram produzidas artificialmente uma quantidade maior de sequências. A produção dessas novas sequências implicava em termos que fazer mutações nas sequências-base. As mutações gênicas podem ser de três tipos, dentre elas a substituição, caracterizada pela troca de um ou mais pares de bases na sequência de nucleotídeos (PIERCE, 2016). O modelo de substituição Jukes-Cantor, escolhido para esse trabalho, considera que as frequências de equilíbrio de todas as bases seriam as mesmas e mudanças nucleotídicas ocorreriam na mesma taxa em qualquer sítio (SCHNEIDER, 2003).

Por conta dos alinhamentos, em alguns sítios apareciam gap representando um travessão

(-) ou a letra *N*. Como eles representam falta de informação, substituímos todos os gaps por *N*.

Na construção do algoritmo de substituição considerou-se a ocorrência de mutações, na natureza, em até 1% do genoma. Sorteou-se um valor de porcentagem entre 0, 1 e 1, que representa quantos sítios da sequência de entrada sofreriam substituição. Estabelecidos os sítios que sofreriam as mutações, o próximo passo foi, reconhecendo a base que está no sítio, substituir de maneira aleatória por outra base. Foram produzidas um total de quarenta sequências baseadas neste modelo.

## VETORIZAÇÃO

Com os cromossomos alinhados e tendo gerado uma maior quantidade de sequências, o próximo passo consistiu em transformar as bases *A*, *C*, *G*, *T* em números, procedimento chamado de vetorização. Chamamos de vetorização e não codificação devido o termo codificação (tradução) ser usado em genética.

Na vetorização em cruz, as bases e *N* foram associadas a coordenadas no plano cartesiano da seguinte forma:  $A = (0, 1)$ ;  $C = (1, 0)$ ;  $G = (1, 0)$ ;  $T = (0, 1)$ ;  $N = (0, 0)$ . Porém antes de cada coordenada acrescentamos uma enumeração em ordem crescente. Substituindo esta vetorização na sequência *TCGG*, por exemplo, teremos a vetor  $[1, (0, 1), 2, (1, 0), 3, (1, 0), 4, (1, 0), 5, (0, 0)]$ . Esta vetorização foi empregada inicialmente com um número reduzido de sequências com apenas um cromossomo. Não houve ganho significativo com essa vetorização visto que, a utilização dela no *k-means* retornou agrupamentos similares às taxonomias das *leishmanias* previstas na literatura.

A vetorização em graus é assim chamada porque associa cada base a um arco no círculo. Começando por  $1^\circ$ , a vetorização fica da seguinte forma:  $A = 73^\circ$ ,  $C = 145^\circ$ ,  $T = 217^\circ$ ,  $G = 289^\circ$ ,  $N = 1^\circ$ . De forma geral, essa é uma vetorização por números inteiros onde o arquivo de sequências de bases com inicialmente 1GB, após a vetorização passou a ocupar 2.7GB. Aplicando o *k-means* a esta base de dados, o resultado obtido apresentou dois *clusters* que não representam qualquer dos níveis da taxonomia convencional.

A vetorização *one-hot* consiste em transformar cada entrada em um grupo de bits, entre os quais as combinações legais de valores são apenas aquelas com um único bit alto (1) e todos

os outros baixos (0). As bases em ordem ficaram caracterizadas com seguintes valores:  $A = (1, 0, 0, 0, 0)$ ,  $T = (0, 1, 0, 0, 0)$ ,  $C = (0, 0, 1, 0, 0)$ ,  $G = (0, 0, 0, 1, 0)$ ,  $N = (0, 0, 0, 0, 1)$ . Esta vetorização foi abandonada, pois o arquivo com os vetores ficou com 10.2GB, que tornou sua utilidade em algoritmos de clusterização impraticável do ponto de vista de tempo computacional.

No método iCGR uma sequência de DNA é representada pela função iterada dos nucleotídeos e suas posições na sequência (YIN, 2019). Então, a sequência de DNA pode ser codificada de forma única e por três inteiros. Um inteiro é o comprimento da sequência e os outros dois inteiros representam as distribuições acumuladas de nucleotídeos na sequência. Como em nossas sequências tínhamos a letra  $N$  além das letras das bases, ao invés de termos um vetor de três posições passamos a ter um vetor de quatro posições. A implementação do iCGR encontrada nas referências trabalha apenas com sequências de até mil bases. Como nossas sequências contam com mais de nove milhões de bases, esta vetorização foi abandonada.

A vetorização pelo número de Gödel (NG) veio do fato dele ter usado sistema para codificar símbolos que representam fórmulas e para codificar sequências de fórmulas que representam as provas (GÖDEL, 1931). Adaptamos esse sistema para codificar as sequências de DNA que tínhamos. As sequências de bases de DNA foram distribuídas da seguinte forma:  $A = 1$ ,  $C = 2$ ,  $G = 3$ ,  $T = 4$ ,  $N = 5$ . Os primeiros cinco números primos 2,3,5,7 e 11 formam as bases das potências. Se quisermos determinar o NG para representar a sequência ATGCN, temos que calcular  $NG(ATGCN) = 2^1 \cdot 3^4 \cdot 5^3 \cdot 7^2 \cdot 11^5 = 159.802.854.750$ . A ideia de transformar uma sequência de DNA em um número inteiro é bem interessante, mas considerando nossas sequências de mais de nove milhões de bases, desistimos desta vetorização.

Com o avanço das técnicas de processamento de linguagem natural, a aplicação do *Word2vec* tem se mostrado robusta na descoberta de relações contextuais e semânticas. O *Dna2vec* utiliza a abordagem *Word2Vec* a toda sequência para que os *k-mers* do DNA sejam representados como vetores (YILMAZ, 2020). O problema com esse tipo de vetorização, pensando que cada sequência é considerada uma palavra, é que teríamos apenas cinquenta palavras para treinamento, o que é muito pouco para utilização do *Dna2vec*.

O EIIP (NAIR; SREENADHAN, 2006) representa a distribuição das energias dos elétrons livres ao longo de uma sequência de DNA.

Uma única sequência indicadora de EIIP é formada substituindo o EIIP dos nucleotídeos  $A = 0, 1260$ ,  $C = 0, 1340$ ,  $G = 0, 0806$ ,  $T = 0, 1335$  em uma sequência de DNA e, neste caso, consideramos  $N = 0$ . Diferentemente das anteriores, esta vetorização considera aspectos físicos de uma sequência de DNA. Esta vetorização gerou um arquivo de 6,7GB, que fez com que os algoritmos de clusterização usados além de demorarem bastante para apresentar um resultado, não separou bem os grupos taxonomicamente.

Na vetorização por número atômico as bases A,C,G,T e N foram convertidas em sequências numéricas dando a elas o número total de prótons em cada nucleotídeo (HOLDEN *et al.*, 2007):  $A = 70$ ,  $T = 66$ ,  $C = 58$ ,  $G = 78$ ,  $N = 0$ . Essa vetorização foi utilizada na obtenção dos resultados apresentados neste artigo. A vantagem desta vetorização em relação às outras, foi o tamanho do arquivo (1.1GB) contendo os vetores que representam as sequências após a vetorização.

## ALGORITMOS DE AGRUPAMENTO

Clusterização é um processo de agrupamento de dados não supervisionado em que a existência de padrões na base de dados é percebida pelo próprio algoritmo, que verifica a similaridade entre esses dados e os agrupam, para obter alta homogeneidade dentro dos grupos e alta heterogeneidade entre os grupos.

O *k-means* é um algoritmo simples e o mais comumente usado para clusterização de dados (JAIN *et al.*, 1999). Ele utiliza abordagem gananciosa para encontrar *clusters* que minimizem a soma dos quadrados das distâncias, convergindo para uma solução local (NUNES, 2016). A vantagem deste algoritmo é que ele é de fácil entendimento e implementação. Uma das suas desvantagens é a escolha do número de *clusters* (PHAM *et al.*, 2005). Para contornar esse problema rodamos o *k-means* com  $k$  variando entre 2 e 10, sendo validados pelo método do cotovelo (KASHANI; GRAETTINGER, 2015). Desistimos do algoritmo por dois motivos, considerando que uma vetorização nos retornou os mesmos agrupamentos previstos na literatura e com outra vetorização misturou espécies de subgêneros diferentes.

O *OPTICS* (Ordering Points to Identify the Clustering Structure) é um algoritmo baseado em densidade (ANKERST *et al.*, 1999). A grande vantagem é que não precisamos informar o número de *cluster*, isso é feito automaticamente. O tempo de processamento e o resultado do

algoritmo nos fizeram desistir do algoritmo. O melhor resultado é baseado no coeficiente de silhueta, que pode variar no intervalo [ 1, 1] sendo os valores mais próximos de 1. O melhor resultado separou os dados em quatro *clusters*. O problema com esses *clusters*, é que o algoritmo agrupou espécies de subgêneros diferentes em um mesmo *cluster* como exatamente feito pelo *k-means*. O algoritmo *Fuzzy c-means* (FCM) é um algoritmo baseado em lógica *fuzzy* que considera o grau de pertinência de cada ponto do conjunto de dados em relação aos diversos agrupamentos. O algoritmo FCM foi proposto primeiramente por (DUNN, 1973) e posteriormente generalizada por (BEZDEK *et al.*, 1984), que define este algoritmo como um programa que gera partições difusas e protótipos para qualquer conjunto de dados numéricos. De forma geral, FCM tenta dividir os dados em conjuntos minimizando a função objetivo abaixo.

$$J = \sum_{i=1}^n \sum_{j=1}^p \mu_{ij}^m d(x_i : c_j)^2$$

onde  $\mu_{ij}$  é o grau de pertinência da amostra  $x_i$  ao  $j$ -ésimo *cluster*,  $n$  é o número de instâncias,  $p$  é o número de *clusters* considerados no FCM (que deve ser decidido antes da execução),  $m > 1$  é o parâmetro Fuzzyficador (usualmente a escolha é por um número irracional no intervalo [1,25;2] (COX, 2005)),  $x_i$  um vetor de dados de treinamento onde  $i = 1, 2, \dots, n$  e representa um atributo do dado,  $c_j$ , onde  $j = 1, 2, \dots, p$ , é o centro de um agrupamento *fuzzy* e  $d(x_i : c_j)$  é a distância entre  $x_i$  e  $c_j$ .

## RESULTADOS E DISCUSSÕES

A taxonomia presente no trabalho de (AKHOUNDI *et al.*, 2017) apresenta a taxonomia atual dos subgêneros *Leishmania* e *Viannia* e *Mundinia*. Por considerarmos apenas as espécies que parasitam mamíferos e não aquelas que parasitam répteis, o subgênero *Sauroleishmania* ficou fora do escopo deste trabalho.

Os resultados obtidos através da aplicação dos algoritmos *k-means* e *OPTICS* não foram incluídos nesta seção devido à tendência de agrupar espécies pertencentes a subgêneros distintos em um único conjunto, ou por não apresentarem contribuições significativas em relação à taxonomia atualmente estabelecida.

O principal resultado deste trabalho foi produzido pelo algoritmo *Fuzzy c-means*. Fo-

ram utilizadas cinquenta sequências com a vetorização por número atômico. O FCM tem como base a associação de um objeto a todos os *clusters* usando a função de pertinência. A validação dos *clusters* é dado pelo coeficiente de partição *fuzzy*, assumindo valores no intervalo [0, 1], sendo 1 o melhor. O programa foi aplicado variando entre 2 e 10 *clusters*.

A Tabela 1 mostra que os valores de coeficiente de partição *fuzzy* variaram entre 0,40785 e 0,87481. Esses valores foram utilizados para avaliar a qualidade da clusterização pelo *Fuzzy c-means*. Assim, no agrupamento com 10 *clusters* foi observado o valor mais alto do coeficiente de partição *fuzzy*, atingindo a separação máxima. Porém, agrupamentos com dez *clusters* não são ideais porque significaria, por exemplo, colocarmos as *L. donovani* e *L. infantum* em *clusters* diferentes quando a literatura os colocam juntas.

Tabela 1: Tabela com valores dos coeficientes de partição *fuzzy*.

Quantidade de clusters	Coefficiente de Partição
2	0,52483
3	0,49749
4	0,43940
5	0,56039
6	0,61419
7	0,68551
8	0,44459
9	0,40785
10	0,87481

O agrupamento com sete *clusters*, além de apresentar o segundo maior valor de coeficiente de partição *fuzzy*, é a clusterização mais próxima da representação taxonômica clássica descrita para os complexos do gênero. O resultado com sete *clusters fuzzy* também manteve a divisão em sete complexos, porém com modificações no nível de espécies que formam os complexos.

Na primeira tabela da Tabela 2 temos as espécies distribuídas nos sete *clusters*. As espécies dos *clusters* 1, 2 e 4 aparecem agrupadas com outras espécies diferente do que ocorre na taxonomia convencional e mantendo a coerência em relação aos subgêneros.

A segunda tabela da Tabela 2 apresenta o *cluster* onde as espécies de *L.aethiopica* estão agrupadas, isto é verificado pelo maior valor das linhas. Os valores menores são os de pertinência da espécie em relação aos outros *clusters*.

Tabela 2: Tabelas com agrupamento e graus de pertinências.

Agrupamento com 7 clusters							
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
	<i>L.braziliensis</i>	<i>L.major</i>	<i>L.peruviana</i>	<i>L.donovani</i>	<i>L.enriettii</i>	<i>L.aethiopica</i>	<i>L.arabica</i>
	<i>L.panamensis</i>	<i>L.mexicana</i>		<i>L.infantum</i>			

Espécies	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
<i>L.aethiopica base</i>	0,002823538	0,001894672	0,002104146	0,003542764	0,003195455	0,982933219	0,003506206
<i>L.aethiopica 1</i>	0,015929514	0,010782328	0,011964421	0,019897718	0,017907856	0,903854268	0,019663894
<i>L.aethiopica 2</i>	0,003665983	0,002461349	0,002733374	0,00459825	0,004147011	0,977843394	0,00455064
<i>L.aethiopica 3</i>	0,008519494	0,005738579	0,006370907	0,010668743	0,009614474	0,948536541	0,010551261
<i>L.aethiopica 4</i>	0,010558704	0,007121699	0,007904724	0,01321117	0,011903164	0,936235029	0,013065509

A escolha pelo algoritmo FCM, que se baseia em conjuntos *fuzzy*, deve-se à vantagem de expressar o tipo de situação em que um objeto compartilha similaridade com vários grupos. Sendo assim, o FCM associa cada indivíduo parcialmente a todos os grupos.

Os graus de pertinências indicam que as espécies apresentam algum grau de pertencimento aos demais complexos. Se por um lado isso representa uma origem comum dentre todas as espécies, o que é esperado, por outro demonstra

que a definição de uma espécie não pode obedecer critérios tão rígidos, mas sim considerar graus de identidade entre diferentes espécies, uma representação mais fiel do contraste e ininterrupto fenômeno de evolução das espécies. Com esses valores podemos ainda quantificar a proximidade ou afastamento dessas espécies entre si e em relação a cada complexo.

Sendo assim, sugere-se o aprofundamento dos estudos de utilização do FCM para a classificação de espécies.

## REFERÊNCIAS

- AKHOUNDI, M.; DOWNING, T.; VOTYPKA, J.; KUHL, K.; LUKE, J.; CANNET, A.; RAVEL, C.; MARTY, P.; DELAUNAY, P.; KASBARI, M. *et al.* Leishmania infections: Molecular targets and diagnosis. **Molecular aspects of medicine**, Elsevier, v. 57, p. 1–29, 2017.
- AKHOUNDI, M.; KUHL, K.; CANNET, A.; VOTYPKA, J.; MARTY, P.; DELAUNAY, P.; SERENO, D. A historical overview of the classification, evolution, and dispersion of leishmania parasites and sandflies. **PLoS neglected tropical diseases**, Public Library of Science, v. 10, n. 3, p. e0004349, 2016.
- ANKERST, M.; BREUNIG, M. M.; KRIEGEL, H. P.; SANDER, J. Optics: Ordering points to identify the clustering structure. **ACM Sigmod record**, ACM New York, NY, USA, v. 28, n. 2, p. 49–60, 1999.
- BEZDEK, J. C.; EHRLICH, R.; FULL, W. Fcm: The fuzzy c-means clustering algorithm. **Computers & geosciences**, Elsevier, v. 10, n. 2-3, p. 191–203, 1984.
- COX, E. **Fuzzy modeling and genetic algorithms for data mining and exploration**. [S.l.]: Elsevier, 2005.
- CUPOLILLO, E.; MEDINA-ACOSTA, E.; NOYES, H.; MOMEN, H.; GRIMALDI, G. A revised classification for leishmania and endotrypanum. **Parasitology today**, Elsevier, v. 16, n. 4, p. 142–144, 2000.
- DARLING, A. C. E.; MAU, B.; BLATTNER, F. R.; PERNA, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. **Genome research**, Cold Spring Harbor Lab, v. 14, n. 7, p. 1394–1403, 2004.
- DUNN, J. C. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. Taylor & Francis, 1973.
- GÖDEL, K. Über formal unentscheidbare sätze der principia mathematica und verwandter systeme i. **Monatshefte für mathematik und physik**, Springer, v. 38, n. 1, p. 173–198, 1931.
- GOSSAGE, S. M.; ROGERS, M. E.; BATES, P. A. Two separate growth phases during the development of leishmania in sand flies: implications for understanding the life cycle. **International journal for parasitology**, Elsevier, v. 33, n. 10, p. 1027–1034, 2003.
- HOLDEN, T.; SUBRAMANIAM, R.; SULLIVAN, R.; CHEUNG, E.; SCHNEIDER, C.; JR, G. T.; FLAMHOLZ, A.; LIEBERMAN, D. H.; CHEUNG, T. Atcg nucleotide fluctuation of deinococcus radiodurans radiation genes. v. 6694, p. 669417, 2007.
- JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. **ACM computing surveys (CSUR)**, Acm New York, NY, USA, v. 31, n. 3, p. 264–323, 1999.
- KASHANI, A. G.; GRAETTINGER, A. J. Cluster-based roof covering damage detection in ground-based lidar data. **Automation in Construction**, Elsevier, v. 58, p. 19–27, 2015.
- KLEPKA, V.; CORAZZA, M. O essencialismo na classificação de lineu e a repercussão dessa controvérsia na biologia. **História da Ciência e Ensino: construindo interfaces**, v. 18, p. 73–110, 2018.
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes-Fundamentos e aplicações**, Manole Ltda, v. 1, n. 1, p. 32, 2003.
- NAIR, A. S.; SREENADHAN, S. P. A coding measure scheme employing electron-ion interaction pseudopotential (eiip). **Bioinformatics**, Bio-medical Informatics Publishing Group, v. 1, n. 6, p. 197, 2006.
- NCBI. **Site National Center for Biotechnology Information**. Online. <https://www.ncbi.nlm.nih.gov/>.
- NUNES, D. H. F. **Um breve estudo sobre o algoritmo K-means**. Dissertação (Mestrado), 2016.
- OCHI, L. S.; DIAS, C. R.; SOARES, S. S. F. Clusterização em mineração de dados. **Instituto de Computação-Universidade Federal Fluminense-Niterói**, v. 1, p. 46, 2004.
- OLIVERA, S. W. Taxonomia de bloom. **Universidad Cesar Vallejo**, v. 4, 2011.

## DADOS DOS AUTORES

**Antonio Revail Alves Pereira**. Universidade Federal do Rio de Janeiro . E-mail: antoniorevail@ppgi.ufrj.br

**Marcello Goulart Teixeira**. Universidade Federal do Rio de Janeiro . E-mail: marcellogt@dcc.ufrj.br

**Flávio Luis de Mello**. Universidade Federal do Rio de Janeiro . E-mail: fmello@poli.ufrj.br

## PREVISÃO DE ENCHENTES NO RIO DOCE UTILIZANDO MODELO SARIMA

Jean Sousa Mendes

Eliane da Silva Christo

Kelly Alonso Costa

**RESUMO:** As inundações fluviais representam uma ameaça grave em todo o mundo, com impactos financeiros e significativos para a sociedade. Este estudo concentrou-se na bacia do Rio Doce, especialmente na cidade de Governador Valadares, que enfrenta enchentes recorrentes durante o verão. O objetivo principal foi desenvolver um modelo estatístico de previsão de enchentes baseado no método SARIMA (*Seasonal Autoregressive Integrated Moving Average*). Uma característica fundamental deste estudo é a previsão com 6 horas de antecedência, para possibilitar uma resposta ágil às elevações do nível do rio e garantir tempo hábil para a implementação de medidas preventivas pelos órgãos competentes. O modelo obteve um erro médio absoluto percentual (MAPE) de 0,64% e um erro quadrático médio (RMSE) de 0,0224.

**Palavras-chave:** Previsão de Enchentes, SARIMA, Séries Temporais.

**ABSTRACT:** River floods pose a serious threat worldwide, with significant financial and societal impacts. This study focused on the Rio Doce basin, particularly in the city of Governador Valadares, which experiences recurrent floods during the summer. The main objective was to develop a statistical flood forecasting model based on the SARIMA (Seasonal Autoregressive Integrated Moving Average) method. A key feature of this study is the forecast with a 6-hour lead time, aiming to enable a prompt response to river level increases and ensuring sufficient time for the implementation of preventive measures by relevant authorities. The model achieved a Mean Absolute Percentage Error (MAPE) of 0.64% and a Root Mean Square Error (RMSE) of 0.0224.

**Keywords:** Flood Forecasting, SARIMA, Time Series.

## INTRODUÇÃO

As enchentes de rios são uma das catástrofes naturais mais custosas para a sociedade, tanto em termos financeiros quanto em perda de vidas (World Economic Forum, 2022). No século XX, as enchentes de rios causaram uma perda média global anual direta de US\$ 104 bilhões (UNDRR, 2015) e resultaram em cerca de sete milhões de vidas perdidas (Doocy et al., 2013). É compreensível, portanto, que as sociedades sempre tenham se esforçado para reduzir as catástrofes de inundações por meio de instrumentos de gerenciamento de enchentes.

O Rio Doce percorre extensas áreas dos estados de Minas Gerais e Espírito Santo, sendo uma importante bacia e fonte crucial de pesca e recursos para a cidade de Governador Valadares. No entanto, devido ao alto nível de assoreamento de suas margens, essa cidade enfrenta enchentes regulares com impacto significativo para os moradores das áreas na planície de inundação (Amorim et al., 2018).

Em 11 de janeiro de 2022, Governador Valadares registrou a terceira maior enchente histórica, onde o rio atingiu 4,35 metros. O maior nível do rio na cidade ocorreu em 1979, quando o rio Doce atingiu 5,18 metros, enquanto o segundo em 1997 com 4,77 metros. O alerta do risco de enchentes começa a partir de 3,2 metros, e a partir de 3,6 metros se caracteriza como inundação (SAAE, 2022).

Devido à gravidade dos danos causados pelas enchentes, a criação de ferramentas de previsão para o nível do rio tornou-se importante instrumento para mitigar impactos do fenômeno. Com base nessa necessidade, este artigo tem como objetivo desenvolver um modelo estatístico de previsão de enchentes do rio Doce na cidade de Governador Valadares, em Minas Gerais. Espera-se que essa abordagem proporcione informações mais precisas e oportunas para a gestão de situações críticas e contribua para a redução dos riscos associados às enchentes.

## REVISÃO BIBLIOGRÁFICA

Nesta seção, discutem-se algumas pesquisas relacionadas à aplicação de modelos para previsões do fluxo de rios e prevenções de enchentes.

De acordo com Sorooshian et al. (2008), um modelo é uma representação simplificada de um sistema do mundo real. O melhor mode-

lo é aquele que fornece resultados próximos à realidade com o uso do menor número de parâmetros e complexidade do modelo.

Na região Sul Fluminense, no interior do estado do Rio de Janeiro, Brasil, Meza (2021) estudou a previsão de inundações na cidade de Volta Redonda, que está localizada às margens do rio Paraíba do Sul e sofre com frequentes inundações. O ponto principal da pesquisa foi a utilização de um modelo de regressão linear generalizada, em inglês: “*generalized linear model*” (GLM), do tipo múltipla. As variáveis de entrada foram dados históricos de vazões de três pontos de medições diferentes do nível do rio. O autor também combinou a técnica de regressão múltipla com a otimização por enxame de partículas, “*particle swarm optimization*” (PSO), alcançando um menor erro estatístico de previsão. O erro percentual absoluto médio, “*mean absolute percentage error*” (MAPE), foi de 0,1758% para a combinação das técnicas. O MAPE é um tipo de cálculo de erro que expressa a porcentagem média dos desvios (em valor absoluto) entre os valores reais e as previsões. Além do disso, foi desenvolvido um aplicativo confiável para alertar sobre o risco de inundações do rio Paraíba do Sul em Volta Redonda.

Tadesse et al. (2017) estudaram o rio Waterval, na África do Sul, que é essencial para o planejamento e gerenciamento do sistema da operação do reservatório de Vaal Dam. Segundo o autor, não havia informações disponíveis na literatura sobre a previsão de fluxos de rios nessa região. Ele utilizou o modelo “*Seasonal Autoregressive Integrated Moving Average*” (SARIMA) de parâmetros  $(3,0,2) \times (3,1,3)[12]$  para previsões do fluxo médio mensal. O modelo obteve um erro estatístico “*root-mean-square error*” (RMSE) de 1,342. O RMSE é um tipo de erro que calcula a média das raízes quadráticas dos desvios entre os valores reais e as previsões. O modelo selecionado apresentou bom desempenho, com previsões próximas aos valores observados, sendo útil para o planejamento e gerenciamento do sistema do rio Waterval e operação do reservatório Vaal Dam.

## COLETA E TRATAMENTO DE DADOS

O estudo utilizou dados de medições de vazão horária do Rio Doce na estação de Governador Valadares (código 56850000, registrada como “UHE GOVERNADOR VALADARES”). Esses dados são gerenciados pela Agência Nacional das Águas (ANA) e estão acessíveis por

meio da plataforma “HidroWeb” (ANA, 2023). A análise abrangeu um período de quatro anos, de 2018 a 2022, excluindo os anos anteriores à 2017 devido à falta significativa de registros naquele ano. Os dados foram transformados em médias a cada seis horas para aprimorar a previsão de inundações, pensando em uma previsão com tempo hábil de resposta à mitigação de danos.

Uma análise exploratória envolveu a remoção dos horários sem valores (cerca de 1% dos registros), também se avalia a normalidade dos dados por meio de um histograma de frequência, além da aplicação teste de Shapiro-Wilk, o qual avalia a hipótese nula de que os dados seguem uma distribuição normal (Royston 1992). Por fim, verifica-se a estacionariedade da série por meio do teste de Dickey-Fuller (Dickey et al., 1979). Essas etapas têm o objetivo de garantir a qualidade dos dados para as análises subsequentes.

## SARIMA

O modelo estatístico “*seasonal autoregressive integrated moving average*” (SARIMA) foi escolhido para analisar os dados de fluxo de água do rio devido à sua capacidade de incorporar tendências e variações aleatórias nos dados, lidar com padrões sazonais ao longo do tempo e oferecer simplicidade em comparação com métodos mais complexos. Isso o torna uma escolha eficaz para a análise de séries temporais de fluxo de água de rios. O SARIMA é um modelo sazonal que incorpora tanto fatores não sazonais quanto sazonais em um modelo multiplicativo. A Equação (1), segundo Morettin (2006), apresenta a notação abreviada para um modelo SARIMA(p, d, q)(P, D, Q)s:

$$\Phi_P(B^s)\varphi(B)\nabla_B^D\nabla_s^d x_t = \Theta_Q(B^s)\theta(B)w_t \quad (1)$$

Onde  $\{w_t, w_t\}$  é a série temporal não estacionária,  $\{x_t, x_t\}$  é o processo de ruído branco gaussiano, “s” é o período da série temporal. Os componentes autoregressivos e de média móvel ordinários são representados por polinômios  $\phi(B)$  e  $\theta(B)$  de ordens “p” e “q”. Os componentes autoregressivos e de média móvel sazonais são representados por  $\Phi(B^s)$  e  $\Theta(B^s)$  de ordens P e Q. Os termos  $\nabla_d \nabla_d$  e  $\nabla_s^D \nabla_s^D$  são componentes de diferença ordinária e sazonal de ordens d e D, B é o operador de deslocamento ou atraso da observação no tempo.

Todo o desenvolvimento do modelo será realizado através do “software” R (R Core Team, 2021).

## TREINO E TESTE DO MODELO

Após a definição e tratamento da base de dados, é necessário dividi-la em duas partes: treinamento e validação. O período escolhido para a fase de treinamento compreende de 01/01/2018 até 08/01/2022. Essa escolha se deve ao fato de que, em 11/01/2022, a cidade de Governador Valadares enfrentou uma das maiores enchentes registradas no município (SAAE, 2022).

Portanto, é relevante testar o modelo para capturar os eventos ocorridos próximos dessa data. Para a etapa de teste, o modelo gerará previsões futuras que serão comparadas com os dados reais. Após cada previsão, o modelo será atualizado com o dado real correspondente e, em seguida, uma nova previsão será gerada e comparada com o próximo dado real. Esse processo será repetido durante o período de 10 dias na base de validação, com o objetivo de analisar o comportamento do modelo nos dias mais críticos da inundação ocorrida em 2022.

A Tabela 1 apresenta um resumo da estrutura da base de dados e suas divisões entre conjunto de treinamento e conjunto de teste.

Tabela 1: Resumo da estrutura da base de dados do Rio Doce na estação 56850000

Dados	Frequência	Período	Adaptação	Base de treino	Base de teste
Nível em cm	Horária	2018-2022	Médias de 6h	<< 07/01/22	>> 08/01/2022

## Definição de parâmetros

É crucial analisar a dependência temporal da série de dados usando a função de autocorrelação (FAC) e a função de autocorrelação parcial (FACP) para determinar os parâmetros do modelo SARIMA. A FAC ajuda a identificar a dependência temporal geral, enquanto a FACP isola a correlação direta entre valores presentes e passados, auxiliando na escolha do termo autorregressivo, “*autoregressive*” (AR) e do termo de média móvel, “*moving average*” (MA) do modelo (Morettin 2006).



Além da análise gráfica das funções de autocorrelação, também utiliza-se a função “auto.arima” do pacote “forecast” no R (Hyndman e Khandakar 2008). Essa função automatiza a seleção de parâmetros com base em critérios estatísticos como “Akaike information criterion” (AIC) e “Bayesian information criterion” (BIC).

O período sazonal escolhido para o SARIMA foi de 4 passos devido à natureza horária dos dados organizados em intervalos de 6 horas. Dessa forma, os padrões sazonais de curto prazo diários são capturados para o fluxo do rio.

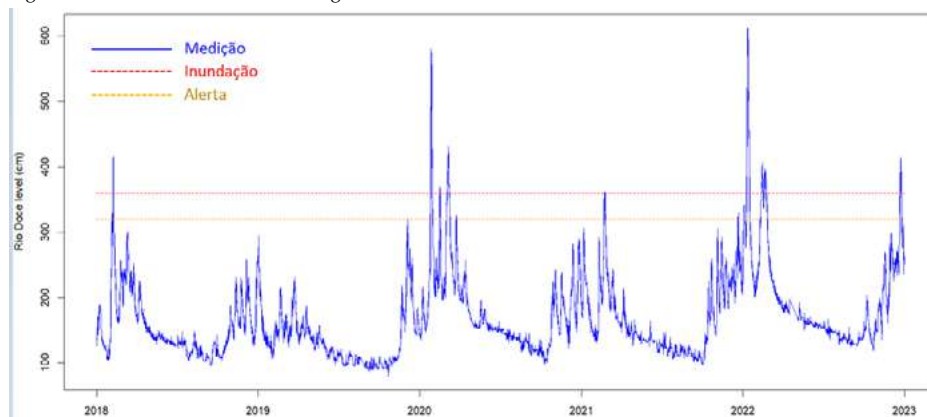
## VALIDAÇÃO DO MODELO

Para verificar a presença de autocorrelação serial nos resíduos para diagnosticar a adequação do modelo ajustado, aplica-se o teste Ljung-Box. Ele testa a hipótese nula de que não há autocorrelação nos resíduos até um determinado número de defasagens (Ljung et al. 1978).

## RESULTADOS

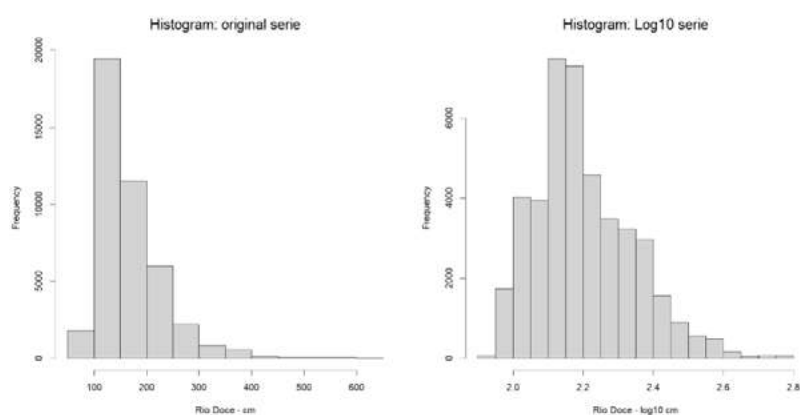
A Figura 1 traz a evolução do nível do rio Doce ao longo dos últimos anos, nota-se picos elevados durante os períodos de verão de cada ano, com destaque para o ocorrido no início de 2022.

Figura 1: Nível do Rio Doce ao longo dos últimos anos



A Figura 2 compara a série original com a série logaritimizada através da transformação logarítmica na base 10. Na primeira imagem é possível notar uma distribuição assimétrica e positiva (cauda direita) que foi reduzida após a transformação.

Figura 2: Comparação do nível do Rio Doce, dados reais versus simulação

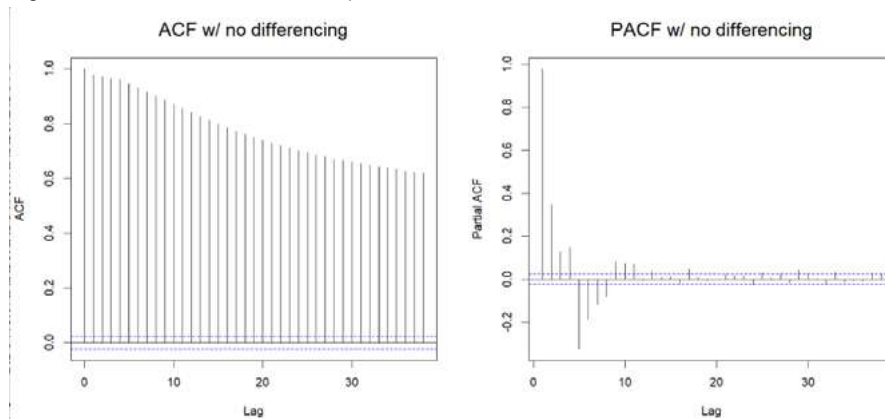


O Devido ao grande tamanho da amostra, o teste de Shapiro-Wilk na totalidade dos dados não é viável (o teste é limitado a 5000 observações). Contudo, o teorema “*central limit theorem*” (CLT) afirma que, em amostras grandes, a média se aproxima de uma distribuição normal, mesmo que as variáveis individuais não sejam normais. Isso significa que, em amostras suficientemente grandes, a soma ou proporção também se comporta de forma semelhante à distribuição normal (Anderson 2010).

Portanto, a confiança no CLT respalda a distribuição normal da série.

A Figura 3 traz as funções de autocorrelação (FAC) e autocorrelação parcial (FACP) geradas na análise.

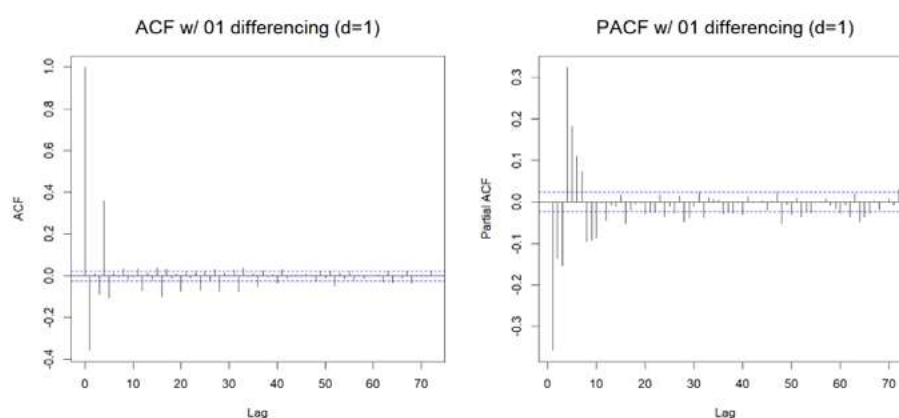
Figura 3: FAC e FACP da série temporal



O gráfico da FAC aponta um decaimento lento, o que indica que a série não é estacionária, propõe-se então uma diferenciação na própria série a fim de estabilizar a média.

A Figura 4 apresenta a FAC e FACP após a aplicação de uma defasagem na série.

Figura 4: FAC e FACP da série temporal após uma defasagem



Após a diferenciação, é possível observar que uma diferenciação de ordem  $d=1$  é suficiente. A estacionariedade é confirmada pelo teste de Dickey-Fuller, o qual resultou em um valor de  $p$  menor que 0,01, rejeitando a hipótese nula de não estacionariedade para os dados diferenciados.

A FAC também sugere uma ordem não sazonal  $q=2$  e uma ordem sazonal  $Q=1$ , isso por causa da presença de picos significativos apenas nas posições sazonais múltiplas de 4. Quanto à FACP, os resultados sugerem uma ordem não sazonal  $p=3$  e uma ordem sazonal  $P=2$ , considerando picos sazonais nas posições 4 e 8. A rápida taxa de decaimento nas funções FAC e FACP sugere um valor de  $D=0$ . Ao gerar o modelo SARIMA com esses parâmetros, obtém-se um erro MAPE de 0,6437%. No entanto, análises estatísticas mais precisas são necessárias para confirmar a escolha dos parâmetros em relação à análise gráfica.

Com base na função *auto.arima* do pacote “*forecast*” do R, o modelo SARIMA (2, 1, 3)(2, 0, 2)4 foi sugerido para a previsão dos fluxos do rio Doce. Este modelo apresentou um erro MAPE de

0,6428%, valor menor do que o modelo estimado pela análise gráfica, sendo, portanto, o critério para a escolha do modelo neste estudo.

A Tabela 2 apresenta os coeficientes do modelo SARIMA (2, 1, 3)(2, 0, 2)4.

Tabela 2: Modelo SARIMA gerado pelo software R

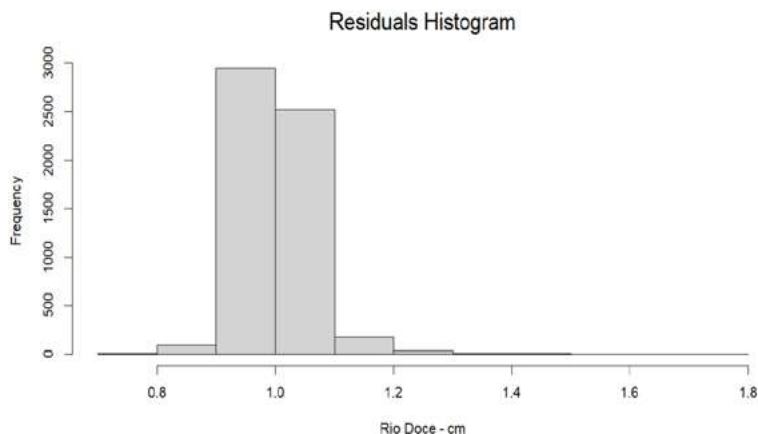
Parâmetros	$\phi_1$	$\phi_2$	$\theta_1$	$\theta_2$	$\theta_3$	$\Phi_1$	$\Phi_2$	$\Theta_1$	$\Theta_2$
Valor	0.7977	-0.3149	-1.2834	0.7000	-0.0570	1.3267	-0.4566	-1.1259	0.2388

Os termos  $\phi_1$  e  $\phi_2$  representam os coeficientes autorregressivos nas defasagens lag 1 e 2 dos componentes não sazonais, enquanto  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$  representam os de média móvel dos componentes não sazonais nas defasagens 1, 2 e 3 respectivamente. Já os termos  $\Phi_1$  e  $\Phi_2$  representam os coeficientes autorregressivos sazonais as defasagens 1 e 2 enquanto  $\Theta_1$  e  $\Theta_2$  são os de média móvel dos componentes sazonais nas defasagens 1 e 2.

O valor do BIC para o modelo foi de -27458,32, enquanto o AIC foi de -27524,96. Estes testes AIC são utilizados na avaliação de modelos O teste estatístico de Ljung-Box realizado no modelo resultou em um valor de p de 0,9913, indicando que não há autocorrelação significativa nos resíduos, pois, não é possível rejeitar a hipótese nula de independência.

A Figura 5 ilustra a distribuição dos resíduos deste modelo.

Figura 5: Distribuição dos resíduos do SARIMA (2, 1, 3)(2, 0, 2)4.



Na análise dos resíduos, notou-se concentração próxima ao nível do rio de 1 cm, indicando pequenas variações nas previsões em relação aos dados reais, em linha com o erro do modelo. A aplicação do teorema CLT também é válida neste caso, sugerindo uma distribuição normal dos resíduos, consideração respaldada pela amplitude da amostra, já que em conjuntos grandes, a distribuição amostral da média tende a se aproximar da normalidade.

A Tabela 3 permite a comparação do erro deste modelo em relação aos outros modelos discutidos na seção de revisão bibliográfica.

Tabela 3: Resumo dos resultados deste e de alguns outros trabalhos

Rio	Autor	Método	Previsão	MAPE	RMSE
Paraíba do Sul (BR)	Meza (2021)	GLM & PSO	Horária	0,18%	
Waterval (ZA)	Tadesse et al. (2017)	SARIMA	Mensal		1,3420
Rio Doce (BR)	Autores	SARIMA	6 horas	0,64%	0,0224

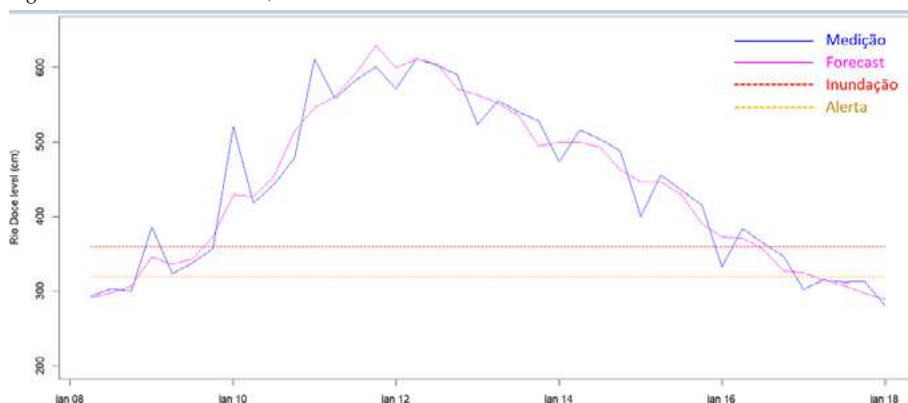
O estudo revelou um alto potencial de previsão ao adotar dados organizados em intervalos de 6 horas, uma escolha adequada para lidar com o tipo de elevação do nível do Rio Doce. Embora tenha apresentado um desempenho ligeiramente inferior em comparação com o modelo de previsão horária do Paraíba do Sul, é fundamental compreender que o comportamento do Rio Doce se difere desse rio, pois o aumento do volume se caracteriza por ser gradual e lento por causa do seu nível de assoreamento.

A escolha da organização dos dados (mensais, horários, etc.) depende dos objetivos da pesquisa. O estudo realizado na África do Sul optou por previsões mensais, uma vez que seu foco estava na avaliação do volume de água do reservatório local para geração de energia. Por outro lado, a previsão horária no Rio Paraíba do Sul se justifica devido à proximidade da cidade de Volta Redonda em relação ao leito do rio e sua rápida subida.

A previsão com 6 horas de antecedência possibilita uma resposta ágil à elevação do nível do rio, enfatizando a importância de selecionar modelos de previsão adequados para as particularidades de cada situação.

A Figura 6 ilustra o desempenho de validação do modelo de previsão horária durante os 10 dias próximos aos momentos mais críticos da inundação em 2022 na cidade de Governador Valadares.

Figura 6: Nível do Rio Doce, dados reais versus dados simulados



É evidente que o modelo demonstrou um ajuste notável em relação às medições reais, conseguindo prever com precisão as fases de ascensão, picos e declínio na curva.

nais, sendo adotada com base nas características específicas do Rio Doce e considerando o impacto das enchentes na cidade de Governador Valadares.

## CONCLUSÕES

Este estudo destaca a eficácia do modelo SARIMA na previsão hidrológica, apresentando baixo erro de modelagem (MAPE de 0,64% e RMSE de 0,0224) para o Rio Doce. Uma característica fundamental deste estudo é a previsão com 6 horas de antecedência. Essa abordagem difere dos modelos de previsão horária tradicio-

O SARIMA se revela uma ferramenta valiosa para prever enchentes de maneira eficaz. Recomenda-se expandir a pesquisa para incluir previsões horárias e mensais, permitindo comparações adicionais com estudos anteriores. Além disso, pretende-se explorar a implementação do método GLM com PSO nos dados do Rio Doce e comparar seus resultados com o desempenho do SARIMA, com o objetivo de aprimorar ainda mais as previsões em benefício da comunidade local.

## REFERÊNCIAS

- AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). **HIDROWEB**. 2023. Disponível em: < <https://www.ana.gov.br/> >. Acesso em: 7 set. 2023.
- ANDERSON, C. J.. **Central Limit Theorem**. The Corsini Encyclopedia of Psychology. 2010. DOI: <https://doi.org/10.1002/9780470479216.corpsy0160>.
- DICKEY, D. A., & FULLER, W. A.. **Distribution of the estimators for autoregressive time series with a unit root**. Journal of the American Statistical Association, v. 47, p. 427-431, 1979.
- DOOCY, S., DANIELS, A., MURRAY, S., and KIRSCH, T. D.. **The human impact of floods: a historical review of events 1980–2009 and systematic review**. PLoS Currents, v. 5, 2013.
- HYNDMAN RJ, KHANDAKAR Y.. **Automatic time series forecasting: the forecast package for R**. Journal of Statistical Software, v. 26, n. 3, p. 1-22, 2008.
- LJUNG, G. M., & BOX, G. E. P.. **On a measure of a lack of fit in time series models**. Biometrika, v. 65, n. 2, p. 297-303, 1978.
- MEZA, L. A. L.. **Modelo de Aprendizado de Máquina Combinado com Algoritmo de Exame de Partículas para a Emissão de Alerta de Inundações no Município de Volta Redonda**. 2021. 110 f. Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia–Uff, Volta Redonda, 2021. Dissertação de Mestrado.
- MORETTIN, P. A., & Tolo, C. M.. **Análise de Séries Temporais**. São Paulo: Blucher, 2006.
- R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Viena, 2021.
- SERVIÇO AUTÔNOMO DE ÁGUA E ESGOTO (SAAE). **Nível do Rio Doce**. 2022. Disponível em: <<https://www.saaegv.com.br>>. Acesso em: 12 dez. 2022.
- TADESSE, K. B., & DINKA, M. O.. **Application of SARIMA model to forecasting monthly flows in Waterval River, South Africa**. Journal of Water and Land Development, v. 35, p. 229-236, 2017.
- UNITED NATIONS OFFICE FOR DISASTER RISK (UNDRR). **Global Assessment Report on Disaster Risk Reduction 2015**. 2015. Disponível em: <<https://www.preventionweb.net>>. Acesso em: 8 set. 2023.
- WORLD ECONOMIC FORUM. **The Global Risks Report 2022**. 2022. Disponível em: <<https://www.weforum.org>>. Acesso em: 8 set. 2023.

## DADOS DOS AUTORES

**Jean Sousa Mendes**. Mestrando em Modelagem Computacional em Ciência e Tecnologia pela UFF, [jeansousa@id.uff.br](mailto:jeansousa@id.uff.br);

**Eliane da Silva Christo**. Doutora em Engenharia Elétrica pela PUC-RJ, é professora titular no departamento de Engenharia de Produção da UFF, [elianechristo@id.uff.br](mailto:elianechristo@id.uff.br);

**Kelly Alonso Costa**. Doutora em Engenharia Civil pela UFF, é professora titular no departamento de Engenharia de Produção da UFF, [kellyalonso@id.uff.br](mailto:kellyalonso@id.uff.br);

# APLICAÇÃO DE UMA FORMULAÇÃO MATEMÁTICA PARA MONTAGEM DE PLANOS DE ESTUDOS DOS ALUNOS DE UM CURSO DE ENGENHARIA

## APPLYING A MATHEMATICAL FORMULATION TO CREATE STUDY PLANS FOR STUDENTS OF AN ENGINEERING PROGRAM

Tiago Araújo Neves

Marcelo Lira Brandão

Kelly Alonso Costa

Thalya Salino Braga da Gama Furtado

Marco Antônio Moraes Pereira Júnior

**RESUMO:** A cada semestre, estudantes de graduação enfrentam a seleção de disciplinas, um processo que implica na análise de múltiplas turmas de disciplinas diversas, com vários conflitos de horário. Muitos alunos decidem com base em interesses pessoais e atividades não acadêmicas, aumentando assim, o risco de atrasar a formatura e de abandono devido à falta de um plano a longo prazo. Este trabalho apresenta os resultados da aplicação de um modelo matemático para auxiliar na criação de cronogramas de estudo, visando encurtar o tempo de graduação, usando dados reais. Os resultados indicam que o problema pode ser resolvido usando programação linear inteira, com tempos computacionais aceitáveis.

**Palavras-chave:** Otimização Combinatória. Plano de Estudos Personalizado. Formulação Matemática.

**ABSTRACT:** Each semester, undergraduate students face course selection, a process that involves analyzing multiple sections from different courses, with schedule conflicts. Many students make decisions based on personal interests and non-academic activities, thus increasing the risk of delaying graduation and dropping out due to the lack of a long-term plan. This work presents the results of applying a mathematical model to assist in creating study plan, aiming at reducing graduation time, using real data. The results indicate that the problem can be solved using integer linear programming, with acceptable computational times.

**Palavras-chave:** Combinatorial Optimization. Personalized Study Plan. Mathematical Formulation.

## 1. INTRODUÇÃO

A evasão escolar, caracterizada pela interrupção prematura dos estudos, é um desafio para nos sistemas educacionais de muitos países. Este processo resulta em indivíduos sem qualificação adequada, que tendem a enfrentar muitas dificuldades de entrada e permanência no mercado de trabalho, o que acaba se tornando um problema crítico para o próprio indivíduo e para a sociedade como um todo (OCDE, 2016).

Estudos indicam que o fenômeno da evasão possui diferentes causas (OCDE, 2016), e que o contexto socioeconômico tem forte influência na conclusão dos cursos, talvez até mais do que etnia e gênero. Estudantes vindos de contextos socioeconomicamente desfavorecidos possuem maior chance de abandonar os estudos devido às restrições financeiras ou problemas familiares.

Outra possível causa para esse fenômeno é o longo tempo necessário para concluir os programas de graduação, especialmente no ensino superior. Estudos mostram que menos de 40% dos alunos terminam o bacharelado dentro da duração teórica prevista (OCDE, 2019). O número se aproxima de 70% quando são considerados três anos adicionais, ou seja, um número considerável dos alunos leva cerca de 7 ou 8 anos para se formar. Outra preocupação são as taxas de conclusão, que estão diminuindo de acordo com os dados apresentados em 2016 e 2019 (OCDE, 2019), mesmo considerando o tempo adicional.

As causas para o longo tempo de conclusão podem ser várias, desde a formação inadequada no ensino médio até a necessidade de trabalhar e estudar simultaneamente. Contudo, uma das possíveis causas pode ser o mau planejamento da trajetória acadêmica. Algumas vezes, na tarefa de fazer seu plano de estudos, os alunos consideram apenas suas próprias preferências e atividades extraclasse, ignorando que algumas disciplinas são requisitos de muitas outras. Assim, quando tomam decisões sem

um plano de longo prazo, os alunos estão atrasando sua progressão no curso e possivelmente adiando sua formatura. Em casos críticos, os alunos (e às vezes também suas famílias) esgotam seus recursos financeiros e não conseguem continuar os estudos.

Considerando o cenário descrito, este trabalho apresenta um modelo matemático para resolver o Problema de Construção de Horário Personalizado (PCHP). O PCHP consiste em encontrar o melhor conjunto de turmas para um aluno, considerando as disciplinas que ele já concluiu e as disciplinas e turmas disponíveis no semestre atual, priorizando as disciplinas importantes para a sua progressão no programa.

Uma ferramenta computacional, que usa o modelo matemático, foi construída e testada em casos reais do problema usando dados de um curso de Engenharia da Universidade Federal Fluminense (UFF). O objetivo da ferramenta é ajudar os alunos a fazer um plano de estudos otimizado, que priorize disciplinas críticas nas cadeias do curso e com isso, tentar reduzir o tempo médio que os alunos levam para concluir seus estudos. Os resultados obtidos pela ferramenta mostram que é possível resolver o PCHP de maneira exata em um tempo computacional viável. Como os longos tempos necessários para concluir a graduação estão relacionados ao fenômeno da evasão, uma ferramenta que tente melhorar os índices de alunos que se formam dentro da duração teórica pode ajudar na redução das taxas de evasão. Assim, este trabalho aborda, indiretamente, uma das possíveis causas da evasão escolar, que tem um impacto tão negativo na sociedade em termos econômicos e sociais.

## 2. DESCRIÇÃO DO PROBLEMA E LITERATURA RELACIONADA

No Brasil, um estudante precisa concluir um conjunto de diferentes disciplinas para obter uma graduação e cada curso possui um conjunto diferente de disciplinas obrigatórias e optativas. Além disso, cada curso tem um conjunto de regras específicas (número

máximo de disciplinas por semestre, número máximo de semestres para concluir o curso, etc). Estas regras podem variar dependendo da universidade e também entre os cursos de uma mesma universidade. Contudo, muitos cursos apresentam grande semelhança em suas grades curriculares (conjunto de disciplinas obrigatórias e opcionais). Cursos com muitas afinidades, como por exemplo, engenharias e cursos na área de Tecnologia da Informação possuem muitas disciplinas em comum. Cada disciplina possui um conjunto de requisitos (outras disciplinas) que devem ser completadas antes que os estudantes possam cursá-la. Um requisito de uma disciplina pode ter um conjunto próprio requisitos criando assim, uma cadeia de disciplinas interligadas. Por exemplo, no curso analisado neste trabalho, a disciplina de Física 3 tem como requisito a disciplina de Física 2, que tem como requisito Física 1; A disciplina de Cálculo 1 é requisito da disciplina de Cálculo 2 e de Física 1. Observe que um curso pode ter muitas cadeias de disciplinas e que, em alguns casos, várias cadeias podem ter interseções disciplinas que estão presentes em mais de uma cadeia).

Cada disciplina pode ter múltiplas turmas. Cada turma possui um conjunto definido de professores e um conjunto de encontros semanais. Para otimizar os recursos, alunos de diferentes cursos podem compartilhar a mesma turma, e turmas diferentes de uma mesma disciplina podem apresentar diferenças na programação de seus encontros. Ressalta-se que, na universidade onde este estudo está sendo realizado (UFF), o aluno é responsável por fazer seu plano de estudos e, portanto, deve escolher em quais turmas das disciplinas elegíveis se inscrever. Uma disciplina é considerada elegível se ela está sendo ofertada no semestre em questão e se o aluno já tiver concluído seus requisitos. Nesse sentido, os alunos competem por aquelas turmas que melhor se adequam aos seus interesses.

Ao fazer o plano de estudos, os estudantes muitas vezes consideram apenas suas preferências e atividades não acadêmicas, ignorando que algumas disciplinas são requisitos de muitas outras. Assim, ao montar o plano de estudos sem uma visão de

longo prazo, os alunos possivelmente estão atrasando o percurso das cadeias e talvez adiando a conclusão do curso. Em casos críticos, os alunos (e suas famílias) esgotam seus recursos financeiros levando-os a abandonar os estudos. Além disso, quanto maior o atraso, maior a chance de atingir o limite de tempo dos cursos, que é 150% do tempo teórico na UFF. Assim, em um curso previsto para quatro anos, o estudante tem até seis anos para sua conclusão. Como o fenômeno da evasão também está relacionado aos longos tempos de graduação, melhorar as taxas de alunos que se formam dentro do prazo teórico pode acarretar uma redução nas taxas de evasão. Nesse sentido, auxiliar os alunos a construir um plano de estudos que considere as cadeias de disciplinas dos cursos pode reduzir ou evitar atrasos e, assim, aumentar as chances dos alunos se formarem dentro do prazo teórico.

Este trabalho aborda o Problema de Construção de Horário Personalizado (PCHP), que consiste em encontrar o melhor conjunto de turmas para um aluno cursar em um semestre, considerando as disciplinas já cursadas e as turmas disponíveis, priorizando disciplinas obrigatórias e críticas nas cadeias do curso, observando as restrições de carga horária máxima e questões de choque de horários.

Os problemas de alocação de horários são NP-Completo em quase todos os casos universitários (SCHAERF, 1999), o que significa que, mesmo com o auxílio de computadores, eles são difíceis de resolver. Existem três grandes problemas na literatura neste contexto (SCHAERF, 1999): O Problema de Alocação de Horários em Escolas (PAHE); Problema de Alocação de Cursos (PAC); e o Problema de Alocação de Exames (PAE). As abordagens mais comuns são meta-heurísticas ou baseadas em meta-heurísticas, que não podem garantir a obtenção de soluções ótimas. Abordagens exatas, como as formulações matemáticas, conseguem encontrar a solução ótima, mas, geralmente, demandam mais recursos computacionais, e são menos frequentes na literatura (TAN, 2021).

No PAHE deve-se determinar o horário semanal para todas as turmas de uma esco-



la, evitando que os professores estejam em duas turmas ao mesmo tempo, e vice-versa (SCHAERF, 1999); No PAC deve-se determinar a programação semanal de todas as aulas de um conjunto de cursos universitários, minimizando as sobreposições de aulas de cursos com alunos comuns; no PAE deve-se determinar o agendamento dos exames de um conjunto de cursos universitários, evitando sobreposição de exames de cursos com alunos comuns, e distribuindo ao máximo os exames para os alunos.

O trabalho de Chen *et al.* (2021) apresenta um levantamento sobre o PAC. Os autores comparam as restrições consideradas, os conjuntos de dados utilizados e as vantagens/desvantagens das abordagens propostas por vários artigos. Os autores afirmam que os problemas do mundo real variam em termos de requisitos devido a diferentes políticas, sistemas educativos e culturas. Assim, não existe uma solução geral (particularmente modelos matemáticos) que satisfaça todos os requisitos. Meta-heurísticas são as abordagens mais populares nos artigos pesquisados.

O trabalho de Tan (2021) concentram-se no levantamento de artigos para o PAHE. Os artigos pesquisados cobriram diversas abordagens de resolução, incluindo algoritmos matemáticos, meta-heurísticas, hiperheurísticas, matheurísticas, algoritmos de coloração de grafos e abordagens híbridas. Meta-heurísticas são a abordagem mais popular entre os artigos pesquisados. Segundo os autores, há uma mudança de popularidade das técnicas de busca local para métodos de otimização matemática (como Programação Inteira e Programação com Restrições) nos estudos mais recentes considerados. Considerando as abordagens meta-heurísticas, as mais utilizadas são aquelas que utilizam uma única solução (TALBI, 2009) devido ao longo tempo de execução necessário para encontrar boas soluções. Os métodos de base matemática são competitivos para instâncias pequenas e as meta-heurísticas são preferidas em instâncias de médio e grande porte.

Por sua vez, o trabalho de Kristiansen e Stidsen (2013) apresenta uma discussão

detalhada sobre o PAE. Os autores afirmam que a literatura sobre este problema baseada em métodos exatos (como a decomposição) é escassa. Os autores apresentam uma compilação de diversos artigos publicados ao longo dos anos e os organizam com base na principal técnica utilizada. De acordo com o seu relatório, a abordagem heurística é claramente dominante para o problema.

O problema da literatura mais próximo ao abordado neste artigo é o Problema de Particionamento de Estudantes (PPE) (SCHINDL, 2019). Neste problema o objetivo é alocar os alunos em turmas de forma que nenhum aluno seja alocado a turmas que se sobreponham no tempo (choque de horários). A preferência dos alunos por turmas e exigências de tempo livre também podem ser consideradas durante a otimização, bem como objetivos institucionais, como capacidades de turmas e distribuições de grupos de alunos. Embora o particionamento totalmente automatizado dos alunos obtenha bons resultados, essa abordagem tem sido criticada desde seus primeiros estudos (BUSAM, 1967). Entre as principais críticas estão o fato de os alunos perderem o poder de decisão, uma vez que eles são alocados pelo sistema, e o fato de que os objetivos institucionais são tratados com maior prioridade. É importante ressaltar que há poucos trabalhos que lidam com casos reais de aplicações do PPE na literatura (MÜLLER; MURRAY, 2010). Alguns destes trabalhos, como Müller e Murray (2010), Carter (2001) e Murray (2007) são estudos detalhados e merecem destaque.

Em Müller e Murray (2010), e Murray (2007) os autores descreveram um sistema usado para resolver a montagem do quadro de horários na Universidade de Purdue. Os autores afirmam que, em alguns casos, o quadro de um curso pode ser decomposto em uma série de problemas departamentais com pouca probabilidade de criar conflitos de alocação de horário de aula para os alunos. No entanto, o problema torna-se mais complexo quando os alunos frequentam cursos de várias unidades acadêmicas, dificultando a construção do do horário dos alunos sem considerar os dados de várias unidades. Eles afirmam que este é um problema impor-

tante para instituições que oferecem muitas disciplinas com múltiplas turmas, e que não é suficiente para considerar o particionamento de alunos apenas como um subproblema de do quadro de horários. Nestas referências, o particionamento é considerado durante a construção do quadro. No sistema proposto, os alunos solicitam tempo livre e inscrições em disciplinas de forma que cada solicitação tenha uma prioridade diferente. A cada solicitação, os alunos podem escolher disciplinas alternativas ou também entrar em lista de espera. Os autores modelaram o problema como um problema de satisfação de restrições e o objetivo é satisfazer o máximo de solicitações (de disciplinas e tempo livre) possível, maximizando a prioridade geral das atribuições. Os autores contam em detalhes como foram montados diversos mecanismos para tornar possível a resolução do problema e enfatizaram que desigualdades na qualidade da distribuição de horários e salas inviabilizou pelo menos uma tentativa de automatizar a construção do horário. Isso mostra a importância de equilibrar as demandas durante o processo.

Em Carter (2001) os autores descrevem o desenvolvimento de um sistema baseado em uma filosofia orientada por demanda para a Universidade de Waterloo. Segundo os autores, uma grande desvantagem do sistema é que ele foi projetado para resolver a instância específica do problema de Waterloo, portanto, não pode ser usado em outras instituições. No entanto, os autores afirmam que muitas das rotinas de otimização são projetadas para resolver problemas matemáticos mais gerais e podem ser usadas para resolver problemas parciais em outras instituições.

Em Waterloo, a atribuição do corpo docente às turmas das disciplinas normalmente é feita manualmente com antecedência. A administração central usa essas informações como entrada para um sistema de programação de horários orientado pela demanda: os alunos fazem uma pré-inscrição em disciplinas a partir de uma lista de ofertas publicada vários meses antes do início das disciplinas. A Universidade então constrói um quadro de horários que tenta minimizar potenciais conflitos para os estudantes. Para isso, os alunos com solicitações

de disciplinas semelhantes são divididos em os grupos que contêm aproximadamente o mesmo número de alunos. Os grupos são então atribuídos a turmas minimizando o número esperado de conflitos (com base nos horários preferenciais para cada turma). Na sequência, o problema é decomposto em problemas menores. A decomposição é necessária devido ao fato de um quadro de horários integrado para toda a Universidade estar sendo montado. O algoritmo de decomposição usa uma técnica baseada em grafos de conflito para criar subconjuntos de turmas de disciplinas devem ser analisadas em conjunto. Após a decomposição, um algoritmo guloso, que prioriza a preferência dos professores e os conflitos de horários dos alunos, é usado para alocar turmas a intervalos de tempo. Os resultados apresentados mostram que o sistema conseguiu alterar apenas cerca de 2% das atribuições do horário quando comparado com a solução construída manualmente, o que indica que o horário anterior era de fato bem construído. No entanto, essas poucas mudanças no horário foram capazes de reduzir em mais de 10% o número de conflitos nos horários dos alunos. O sistema também permite interação humana em tempo real. Segundo os autores, todos os departamentos têm responsáveis pelo quadro de horários que podem entrar em contato com seus pares de outros departamentos para discutir problemas mútuos, recomendar novos horários aos professores, etc. Os alunos também podem adicionar e descartar disciplinas ou fazer algumas alterações com a ajuda dos responsáveis departamentais. Embora a automação apresentada seja bastante sofisticada, o processo geral (inscrição antecipada, agendamento de horários, atribuição de sala, atribuição de alunos e adição/descarga) leva vários meses.

Como pode ser observado nos trabalhos de Müller e Murray (2010), Carter (2001) e Murray (2007) a resolução do PPE uma tarefa muito complexa e dar suporte aos responsáveis departamentais é importante, uma vez que eles têm um conhecimento muito maior dos detalhes envolvidos (necessidades das turmas, afinidade dos docentes por disciplinas específicas, etc.) (MURRAY, 2007). Além disso, manter o senso de pertencimento de cada departamento também é

um fator importante na aceitação das soluções produzidas pelo processo de agendamento (automatizado ou manual) (MURRAY, 2007).

A grande maioria dos trabalhos na literatura resolvem os problemas pela perspectiva institucional. Ou seja, problemas de horários e/ou particionamento são geralmente resolvidos em função das necessidades institucionais. Poucos trabalhos abordam os problemas na perspectiva dos alunos (BRANDÃO, NEVES e COSTA, 2022).

O PCHP, foco deste trabalho, é proposto por Oliveira Neto, Mariano e Farias (2017), onde autores apresentam um sistema que constrói planos de estudos personalizados visando reduzir os tempos de graduação. De acordo com o conhecimento dos autores, este é o primeiro artigo que aborda essa perspectiva. Os autores apresentam uma solução baseada em um Algoritmo Genético (TALBI, 2009) para sugerir as disciplinas/turmas que o(a) aluno(a) deve frequentar, com base nas disciplinas já concluídas pelo(a) estudante e no conjunto de disciplinas obrigatórias/opcionais do curso que estão disponíveis. Apesar de propor o problema de otimização juntamente com uma modelagem heurística para sua solução, os autores não propuseram uma formulação matemática para o problema.

No trabalho de Brandão, Neves e Costa (2022), foi proposto um modelo matemático para o PCHP. O modelo é baseado em conceitos de grafos e pode ser resolvida por Programação Linear Inteira. Também de acordo com o conhecimento dos autores, este é o primeiro artigo a resolver PCHP usando um método exato. Os autores aplicaram o modelo desenvolvido em dois casos específicos. Análises detalhadas para estes dois casos mostraram o grau de dificuldade do problema enfrentado pelos alunos e alunas a cada semestre. Os resultados apresentados, ainda que preliminares, indicavam que o PCHP poderia ser resolvido de maneira exata para casos reais. Este artigo é uma continuação do trabalho exposto em Brandão, Neves e Costa (2022) e apresenta os resultados da aplicação da modelagem matemática proposta para vários casos reais de alunos do curso de Engenharia de Produção da UFF. É importante en-

fatizar que o objetivo da ferramenta proposta é fazer sugestões de turmas, não atribuições. De acordo com a política da universidade, a decisão final é do(a) estudante.

### 3. MODELAGEM MATEMÁTICA

Na modelagem proposta a estrutura do curso foi modelada como um grafo  $G(V, E)$ , onde  $V$  é o conjunto de vértices que representam as disciplinas, e  $E$  é o conjunto de arcos que representam os requisitos. Assim, um arco partindo de  $a$  e chegando em  $b$  indica que  $a$  é requisito de  $b$ . Com isso, pode-se dizer que uma disciplina  $y$  é dependente de uma disciplina  $x$  se houver um caminho de  $x$  para  $y$  no grafo. Também pode-se definir a vizinhança de um vértice  $x$  como  $N(x) = \{y : y \in V, \exists(x, y) \in E\}$ . Cada vértice possui todas as informações sobre a disciplina que representa, como o número de turmas, seus horários, etc.

A Figura 1a mostra o algoritmo de cálculo da prioridade para cada vértice  $v$ .

que representa relevância desta disciplina no curso. Se a prioridade para  $v$  já estiver calculada (maior que zero), o algoritmo encerra. Caso contrário, a prioridade é calculado recursivamente com as prioridades de todos os vértices de  $v$  mais  $\alpha_v$  (valor arbitrário se a disciplina é obrigatória ou não). No trabalho,  $\alpha_v$  assume valores de 10 se a disciplina é obrigatória e 1 caso contrário. Os valores também podem ser usados para definir a prioridade de uma disciplina, da forma como está calculada, reflete se ela é obrigatório ou não, o número de outras disciplinas dependentes e quantas dessas disciplinas dependentes são obrigatórios. Quanto maior o número de disciplinas dependentes, maior o valor da prioridade. Além disso, quanto maior o número de disciplinas obrigatórias entre as disciplinas dependentes, maior o valor da prioridade. A Figura 1b apresenta o algoritmo para calcular a prioridade para todas as disciplinas de um curso. Primeiro, como fase de inicialização, a prioridade de todas as disciplinas é definida como zero. Posteriormente, a função de prioridade é chamada para cada disciplina. Como a

#### Algoritmo 1 prioridade (Vértice $v$ )

```

1: se ( $v.prioridade \leq 0$ ) então
2:    $aux = \alpha_v$ 
3:   para todo Vértice  $w \in N(v)$  faça
4:      $aux = aux + prioridade(w)$ 
5:   fim para
6:    $v.prioridade = aux$ 
7: fim se
8: retorne  $v.prioridade$ 

```

(a)

complexidade da função de prioridade é  $O(V^2)$ , a complexidade geral para calcular a prioridade para todos os vértices do grafo tem complexidade de pior caso  $O(V^3)$ , que é polinomial. Além disso, o algoritmo da Figura 1b não precisa ser usado para cada estudante, pois os valores de prioridade em um grafo que representa a estrutura de um

curso de graduação não variam de estudante para estudante e, portanto, podem ser calculados uma vez, armazenados e utilizados posteriormente. Observe que, em caso de eventos que provoquem mudanças no grafo (e conseqüentemente em seus valores), que são eventos de mudança na grade curricular e reestruturação de curso por exemplo, o algoritmo deve ser reutilizado para atualizar os valores de prioridade para cada disciplina. No entanto, eventos deste tipo não acontecem com frequência e, mesmo assim, o custo computacional para recalculas as prioridades é polinomial.

Figura 1: Algoritmos para o cálculo de prioridade utilizados

### Algoritmo 2 prioridadeGrafo (Grafo $G$ )

- 1: **para todo** Vértice  $v \in G.E$  **faça**
- 2:    $v.prioridade = 0$
- 3: **fim para**
- 4: **para todo** Vértice  $v \in G.E$  **faça**
- 5:   prioridade ( $v$ )
- 6: **fim para**

(b)

Fonte: Autores

A formulação matemática também utiliza as seguintes notações:

Seja  $S \subset V$  o conjunto de disciplinas já concluídos pelo(a) aluno(a). Seja  $R$  o conjunto de disciplinas elegíveis (disciplinas que possuem o conjunto de requisitos satisfeito) definido por  $R = \{i : i \in V - S, \exists (j, i) \in E, j \in V - S\}$ .

Seja também  $q$  a carga horária semanal máxima (em horas), determinada pelo curso, que um estudante pode cursar. Seja  $c_i$  a carga horária semanal da disciplina  $i$ . Seja  $L_i$  o conjunto de turmas da disciplina  $i$ . Seja  $T$  o horizonte de tempo. Seja  $h_{ikt}$  uma constante que assume 1 se a turma  $k$  da disciplina  $i$  tem um encontro

agendado para o intervalo  $t$ , e 0 caso contrário. Seja também  $p_i$  a prioridade da disciplina  $i$ , calculada usando os algoritmos da Figura 1. Por fim, seja  $x_{ik}$  a variável binária que assume 1 se a turma  $k$  da disciplina  $i$  é selecionada para compor a solução, e 0 caso contrário. Uma possível formulação matemática para o problema é descrita por (1)-(5).

$$\text{Max} \sum_{i \in R} \sum_{k \in L_i} p_i c_i x_{ik} \quad (1)$$

s.a.

$$\sum_{k \in L_i} x_{ik} \leq 1, \forall i \in R, \quad (2)$$

$$\sum_{i \in R} \sum_{k \in L_i} c_i x_{ik} \leq q, \quad (3)$$

$$\sum_{i \in R} \sum_{k \in L_i} h_{ikt} x_{ik} \leq 1, \forall t \in T, \quad (4)$$

$$x_{ik} \in \{0,1\}, \forall i \in R, \forall k \in L_i. \quad (5)$$

A função objetivo (1) maximiza a prioridade total e a carga horária das disciplinas sugeridas. As restrições (2) impõem que no máximo uma turma de cada disciplina pode ser utilizada. As restrições (3) asseguram que a carga horária semanal seja respeitada. As restrições (4) evitam choque de horários nas turmas selecionadas e (5) são as restrições de integralidade e não-nega-

tividade. A formulação prioriza disciplinas obrigatórias e disciplinas que são requisitos de muitas outras já que estes são fatores considerados no cálculo da prioridade de cada disciplina. A função objetivo também considera a carga horária das disciplinas pelo fato de que a carga horária cursada é um fator relevante para questões internas na universidade (como por exemplo, se dois alunos estão disputando uma vaga em uma turma, o estudante com mais horas cursadas tem preferência). Além disso, a formulação evita que mais de uma turmas de uma disciplina seja selecionada, trata os conflitos de tempo, e assegura a carga horária máxima.

Esta formulação foi usada em vários casos reais de estudantes de um curso da UFF e os resultados mostram que é possível resolver o PCHP usando a abordagem proposta. Estes resultados e uma breve discussão sobre eles estão expostos na próxima seção.

#### 4. RESULTADOS

Para os testes computacionais foram usados 16 casos reais de estudantes do curso presencial de Engenharia de Produção da Escola de Engenharia Industrial e Metalúrgica de Volta Redonda. Cada estudante corresponde a uma instância do problema e todos cederam seus dados voluntariamente. O curso conta com 54 disciplinas obrigatórias (cada uma com uma série de requisitos), divididas entre 10 períodos, além de disciplinas optativas, projetos finais e estágio. Ao todo, é preciso cumprir aproximadamente 3600 horas de atividades para se formar. Aqui cabe ressaltar que, apesar de usar um curso específico como estudo de caso, a formulação matemática criada é geral, e pode ser aplicada a qualquer curso que possua uma grade curricular estruturada. A escolha por este curso específico como estudo de caso se deu por conveniência, visto que os autores e voluntários estão ligados a este curso, o que facilita a obtenção de dados. Para a resolução das instâncias do problema, foi utilizado o software IBM Optimization Studio (IBM-OS) versão 12.10, que usa o CPLEX como motor de otimização. O computador utilizado tem processador i7 de terceira geração, 8 GB de memória principal e Sistema Operacional Linux.

Foram construídas duas versões da formulação proposta. Uma usando a interface do IBM-OS e outra usando diretamente as bibliotecas do CPLEX em Java, chamada de PHU neste trabalho. Um dos objetivos do estudo era verificar se existe algum *overhead* de tempo para o uso da interface do IBM-OS em relação ao

uso direto das bibliotecas em Java. É importante observar que medir os tempos de execução do IBM-OS com exatidão não é fácil, mesmo com recursos disponibilizados pela ferramenta. Para obter uma comparação justa, foi estabelecido que seria medido o tempo de relógio da execução de duas tarefas executadas sequencialmente: 1) montagem do modelo matemático; e 2) resolução até a obtenção da solução ótima. Para realizar a medição, foi feito um programa que faz uma chamada do executável `oplrunJava`, que é o mecanismo de uso do IBM-OS em interface de comando, passando como parâmetros os arquivos de configuração (somente opções padrão foram usadas). O `oplrunJava` foi executado cinco vezes para cada instância e as médias calculadas. O PHU realiza as mesmas tarefas listadas anteriormente e também foi executado 5 vezes. A medição de tempo, mostra que os dois programas, estão estatisticamente empatados, e as médias diferem em milésimos de segundos, sendo a média de todas as instâncias de 91 ms para o IBM-OS e 95 ms para o PHU. Como são medições de tempo de relógio, as medidas podem ter interferências do sistema operacional, mas o fato dos dois programas executarem as mesmas tarefas, e terem tempos próximos indicam que a interferência não foi significativa. Os detalhes podem ser vistos na Tabela 1.

A primeira coluna da Tabela 1 expõe as instâncias sendo tratadas. É importante ressaltar que cada instância equivale a um(a) aluno(a) real diferente. A segunda coluna expõe os tempos computacionais médios, em milissegundos, obtidos pelo IBM-OS. A terceira coluna mostra os tempos computacionais médios, também em milissegundos, obtidos pelo PHU. A quarta coluna apresenta o número de turmas elegíveis para a instância em questão. A última coluna expõe a quantidade de turmas sugeridas pela solução obtida pelas ferramentas propostas. Finalizando, a última linha apresenta as médias de cada coluna.

**Tabela 1:** Resultados computacionais para casos reais

Instância	Tempo IBM-OS (ms)	Tempo PHU (ms)	Total de Turmas	Turmas Sugeridas
1	92,2	95,2	67	8
2	92,4	94,6	64	9
3	94	95,2	67	8
4	95,4	95	46	8
5	91,8	94,4	39	8
6	88	90	18	7
7	96,2	93,4	46	9
8	93,8	94,8	42	8
9	87,4	98,2	26	8
10	89,8	96,6	44	8
11	90,6	94,8	42	9
12	89,6	94,6	42	9
13	92,8	100,6	43	8

14	89,6	99,4	41	será objeto de estudo futuro.
15	96,6	97,8	57	9
16	83	91	21	8
Médias	91,5	95,4	44,1	8,3

Fote: Autores

Observando a quarta coluna, é possível perceber que há alunos com muito mais turmas elegíveis do que outros. Isso ocorre devido ao fato de que alunos que estão mais adiantados no curso (caso das instâncias 6, 9 e 16) possuem menos disciplinas a serem cursadas e por isso o número de opções de turma a serem analisadas reduz consideravelmente. Essa diferença de número de possibilidades no conjunto de disciplinas entre as instâncias usadas se deve ao fato de que os dados utilizados, como já mencionado, vieram de alunos que se voluntariaram a ceder os seus dados. É importante ressaltar que foi possível encontrar as soluções ótimas para todas as instâncias e as soluções apresentadas pelas abordagens é sempre igual, visto que se tratam apenas de modos diferentes de construir uma mesma formulação matemática. Os resultados mostram que o uso do IBM-OS não tem *overhead* significativo quando comparado ao uso direto das bibliotecas em Java. Este resultado é importante devido ao fato de que construir um modelo de otimização usando o IBM-OS é muito mais simples do que usando as bibliotecas diretamente. Isto fica evidenciado pelo número de linhas necessárias para construir cada abordagem. A formulação construída utilizando o IBM-OS tem aproximadamente 40 linhas, enquanto o PHU possui aproximadamente 450 linhas. Isto mostra que, para cenários cuja fonte de dados sejam arquivos e que não tenham necessidade de interações sofisticadas com o resolvidor durante o processo de otimização, que é exatamente o caso deste projeto, o uso do IBM-OS é a melhor opção. Pelos resultados também percebe-se que a execução das ferramentas leva décimos de segundo, mostrando que a resolução de casos reais do PCHP via método exato é possível e viável. Neste ponto é importante ressaltar que foi possível provar a otimalidade das soluções para todas as instâncias e que o computador utilizado é uma máquina comum (apesar de ser um computador com um bom processador, a máquina usada tem mais de 10 anos), sem nenhum hardware especial. Também é importante mencionar que os horários sugeridos foram todos viáveis, ou seja, são horários em que não há choque de horários entre turmas, não há múltiplas turmas de uma mesma disciplina selecionadas e a carga horária máxima prevista pelo curso não é excedida. Contudo, não foi feita uma verificação da aceitação das soluções fornecidas pelos estudantes. Isso se deve ao fato de que o objetivo desta fase do projeto era verificar a viabilidade técnica da modelagem, que foi atingido com sucesso. Uma verificação da aceitação das soluções propostas pelos estudantes

também pode ser observado na Tabela 1 que a média de turmas disponíveis para escolha é de aproximadamente 44 turmas, enquanto a média de turmas sugeridas pela ferramenta é 8 turmas. Neste sentido, o número médio de possibilidades a serem analisadas por cada estudante pode ser descrito, de modo aproximado, por  $44! / (8! \times (44 - 8)!)$ , que totaliza mais de 5 mil possibilidades (considerando viáveis e inviáveis). Note que o período de inscrição em disciplinas na universidade é de apenas alguns dias, e que não existe, atualmente, uma ferramenta para auxiliar os alunos neste processo, o que torna esta uma tarefa árdua para ser executada. Mesmo considerando que não é possível se inscrever em múltiplas turmas de uma mesma disciplina, o número de possibilidades a serem analisadas ainda é muito grande, o que reforça a importância de uma ferramenta para auxiliar os estudantes neste processo.

## 5. CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho apresentou os resultados da aplicação de uma formulação matemática linear inteira, baseada em conceitos de grafos, para o PCHP em vários casos reais de estudantes de uma universidade brasileira. Em todos os casos foi possível encontrar a solução ótima para o problema em menos de um segundo, o que mostra que o uso da ferramenta proposta é viável. Os resultados também mostram que o número de possibilidades a ser analisada por um estudante a cada semestre é muito elevado e portanto, o uso de uma ferramenta para este fim pode ajudar na realização desta tarefa. A ferramenta proposta faz recomendações baseadas na estrutura curricular do curso, priorizando disciplinas obrigatórias, com mais carga horária e que são relevantes para o progresso do aluno nas cadeias de disciplinas do curso. Com isso, espera-se que, ao seguir as recomendações feitas pela ferramenta, os(as) alunos(as) possam concluir seus cursos mais rapidamente, reduzindo assim, os tempos médios para a conclusão de curso. Como o tempo para a conclusão é um dos fatores que está ligado à evasão, espera-se que ao reduzir os tempos médios, as taxas de evasão também diminuam.

Como trabalho futuro, pode-se propor a aplicação da mesma metodologia a outros cursos. Há interesse especial de aplicação da metodologia em cursos onde não há tantas relações de requisitos entre as disciplinas. Espera-se que para estes cursos, a ferramenta consiga encontrar as soluções ótimas, mas com um tempo computacional mais elevado,

visto que o número de possibilidades de turmas que podem se cursadas aumenta à medida que a quantidade de requisitos diminui. Também considera-se a inclusão de mais elementos na resolução do problema como por exemplo co-requisitos (que são disciplinas que precisam ser cursadas ao mesmo tempo), pedidos de tempo livre por parte dos estudantes (como é feito em algumas abordagens da literatura para o Problema de Particionamento de Estudantes), requisitos fracos (que são disciplinas sem a relação de requisito formal, mas que é desejável que tenham sido cursadas) e requisitos de outras naturezas (há exemplos de disciplinas que possuem como requisito uma carga horária mínima já cursada). Outro trabalho futuro é a análise da aceitação das soluções apresentadas pelo corpo discente.

## REFERÊNCIAS

- BRANDÃO, M. L.; NEVES, T. A.; COSTA, K. A. Modelagem matemática para auxiliar estudantes na montagem do horário visando reduzir o tempo de graduação. *In: LIV SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL (SBPO)*, 54, 2022. Anais... Rio de Janeiro, 2022.
- BUSAM, V. A. An algorithm for class scheduling with section preference. *Communications of the ACM*, v. 10, p. 567–569, 1967.
- CARTER, M. W. A Comprehensive Course Timetabling and Student Scheduling System at the University of Waterloo. *In: PRACTICE AND THEORY OF AUTOMATED TIMETABLING III*, 3, 2001. Anais... Berlin, 2001, p. 64-82.
- CHEN, M. C. *et al.* . A Survey of University Course Timetabling Problem: Perspectives, Trends and Opportunities. *IEEE Access*, v. 9, pp. 106515–106529, 2021.
- KRISTIANSEN, S.; STIDSEN, T. R. *A Comprehensive Study of Educational Timetabling - a Survey*. Kongens Lyngby: DTU Management Engineering, 2013. (DTU Management Engineering Report No. 8.2013).
- MÜLLER, T.; MURRAY, K. Comprehensive approach to student sectioning. *Annals of Operations Research*, v. 181, p. 249–269, 2010.
- MURRAY, K.; MÜLLER, T.; RUDOVÁ, H. Modeling and Solution of a Complex University Course Timetabling Problem. *In: PRACTICE AND THEORY OF AUTOMATED TIMETABLING VI – LNCS,6* , 2007. Anais... Berlin, 2007, p. 189-209.
- OCDE. *Education at a Glance 2016: OECD Indicators*. Paris: OECD Publishing, 2016.
- OCDE. *Education at a Glance 2019: OECD Indicators*. Paris: OECD Publishing, 2019.
- OLIVEIRA NETO, R. F.; MARIANO, C. C. L.; FARIAS, M. S. R. Horário Universitário Personalizado por Metaheurística: Um Estudo de Caso. *In: XLIX SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL (SBPO)*, 49, 2017. Anais... Rio de Janeiro, 2017.
- SCHAERF, A. A survey of automated timetabling. *Artificial Intelligence Review*, v. 13, p. 87–127, 1999.
- SCHINDL, D. Optimal student sectioning on mandatory courses with various sections numbers. *Annals of Operations Research*, v. 275, p. 209–221, 2019.
- TALBI, E.-G. *Metaheuristics: From Design to Implementation*. Hoboken: John Wiley & Sons, 2009.
- TAN, J. S. et al. A survey of the state-of-the-art of optimization methodologies in school timetabling problems. *Expert Systems With Applications*, v. 165, p. 1–13, 2021.

## DADOS DOS AUTORES

**Tiago Araújo Neves**. Graduação em Ciência da Computação pela Universidade Federal de Ouro Preto (2004), mestrado em Computação pela Universidade Federal Fluminense (2007) e doutorado em Computação pela Universidade Federal Fluminense (2011)

**Marcelo Lira Brandão** (marcelolira@id.uff.br). Universidade Federal Fluminense . Graduação em Engenharia de Produção pela UFF.

**Kelly Alonso Costa** (kellyalonso@id.uff.br). Universidade Federal Fluminense, Graduação, Mestrado e Doutorado em Engenharia Civil pela UFF.

**Thalya Salino Braga da Gama Furtado** (thalyafurtado@id.uff.br). Universidade Federal Fluminense, Graduação em Engenharia de Produção pela UniFOA e mestranda em Engenharia de Produção pela UFF.

**Marco Antônio Moraes Pereira Júnior** (marcomoraes@id.uff.br) Universidade Federal Fluminense, Graduação em Engenharia de Produção pela UFF e mestrando em Engenharia de Produção pela UFF.

# METODO PRIMAL SIMPLEX ESPECIALIZADO EM PROBLEMA DA PRODUÇÃO DE RAÇÃO

## PRIMAL SIMPLEX METHOD SPECIALIZED IN FEED PRODUCTION PROBLEM

Marcio Lopes Gevegy

Ricardo Silveira Sousa

Wagner Rambaldi Telles

**RESUMO:** O presente artigo apresenta o método Primal Simplex Canalizado (PSC) especializado na resolução de problemas na produção de rações. O trabalho tem como objetivo também utilizar a regra de Dantzig normalizada numa abordagem aprimorada para verificar seu desempenho frente à regra usual na escolha de variáveis. Para tal propósito, foram formulados 24 problemas (instâncias) que representam diferentes cenários da produção de ração, através de estudos preliminares e de uma visita técnica. Os problemas de produção de rações de aves e suínos são do Instituto Federal do Espírito Santo (IFES). No geral o PSC teve um desempenho bem satisfatório com as duas regras utilizadas. O método fornece soluções com uma ótima precisão numérica, executando os cálculos em um curto tempo computacional.

**Palavras-chave:** Programação Linear. Método Primal Simplex Canalizado. Problema da produção de Ração Aves e Suínos.

**ABSTRACT:** This article presents the Primal Simplex Channelized (PSC) method specialized in solving the feed problem. The aim of the work is to use the normalized Dantzig rule, an improved approach that aims to reduce the number of iterations necessary to obtain the optimal solution. For this purpose, 24 problems (instances) were formulated that represent different feed production scenarios, through preliminary studies and a technical visit. The feed production problems are the responsibility of the Federal Institute of Espírito Santo (IFES). However, for some instances the normalized Dantzig rule was superior to the usual rule, in general it had similar or inferior behavior when used to the Dantzig rule. Overall, the PSC performed very satisfactorily. The method provides solutions with excellent numerical precision, performing calculations in a short computational time.

**Keywords:** Linear Programming. Channeled Primal Simplex Method. Problem with the production of Poultry and Pig Feed.



## INTRODUÇÃO

Diversas áreas do conhecimento estudam formas mais eficazes de organizar os variados tipos recursos. Por exemplo, o tempo, a matéria prima, as máquinas, os operários, etc. (DOS SANTOS, 2020). No período da Segunda Guerra mundial, este tipo de pesquisa era empregado à tática para verificar qual a melhor estratégia a ser utilizada. Nesta época, iniciou-se a Pesquisa Operacional (PO) sendo desenvolvida principalmente pelas forças armadas dos Estados Unidos e Grã Bretanha (CHAVES, 2011).

As pesquisas eram aplicadas especificamente em operações militares, que buscavam a eficiência nos métodos utilizados, sendo necessário solucionar os problemas em um curto espaço de tempo. O Exército e a Marinha foram os pioneiros a usarem esta abordagem. Em 1941 Blackett foi considerado o pai da PO e atuou como conselheiro Científico e Diretor Naval (CHAVES, 2011).

Segundo Barbosa, os primeiros estudos envolvendo problemas de otimização com restrições ocorreu em 1826, nos estudos de Fourier relacionados a sistemas de inequações lineares (BARBOSA, 2014). Já em 1939, este tema ganhou popularidade quando iniciou o desenvolvimento dos primeiros Problemas de Programação Linear ao longo da Segunda Guerra Mundial.

Em 1947, George Dantzig formulou o Problema de Programação Linear (PPL) e desenvolveu o método Simplex. Este método é amplamente aplicado em problemas reais, inclusive na produção de ração animal, na literatura acadêmica é conhecido como Problema da Mistura.

Um dos problemas mais remotos de programação linear é o Problema da Mistura. Os ingredientes são compostos por custo e são utilizados em certa quantidade para compor o produto final. Este desafio consiste em encontrar a quantidade ideal a ser misturada de cada ingrediente de forma a obter a produção com o custo mínimo de acordo com restrições técnicas (ULBRICHT e RIBEIRO, 2012).

É fundamental formular corretamente uma ração balanceada em uma indústria alimentícia. Dar-se o nome de ração o alimento diário do animal, sendo que em sua formulação os ingredientes devem ser misturados e proporcionar a quantidade de nutrientes necessária para as diferentes fases da vida do animal. Para cada espécie há um quantitativo específico de nutrientes a serem usados. A ração deve ser aplicada suprindo as exigências nutricionais do animal (TONIN, 2019).

Tem-se que o Brasil ocupa a terceira posição entre os dez principais produtores de ração. O país é um dos maiores produtores de proteína animal do mundo. Em 2021 foram produzidas 80,094 milhões de toneladas métricas (FEED, 2022). Cabe destacar que a alimentação representa o maior percentual dos custos de produção animal com cerca de 70% dos custos, especialmente no segmento de aves e suínos. Neste cenário, é crucial utilizar ferramentas para auxiliar a tomada de decisão fornecendo informações para obter o custo mínimo de ração (NEME, 2000), como é o caso do Método Simplex especializado em problemas da produção de ração (GEVEGY, 2022).

Na indústria alimentícia animal, formular uma ração balanceada é fundamental para atender as exigências nutricionais dos animais (TONIN, 2019). No contexto da produção de rações para aves e suínos, é essencial combinar alimentos de forma adequada, levando em conta a disponibilidade de insumos, exigências nutricionais, custos e entre outros aspectos. Embora seja possível formular uma ração de qualidade por meio de métodos tradicionais, é necessário o uso de ferramentas de gestão mais refinadas, por exemplo, as técnicas de Programação Linear (CASTRO; BORGERT e DE SOUZA, 2015).

Um dos métodos mais empregados em Programação Linear é o método Primal Simplex Canalizado (PSC) que será brevemente apresentado na próxima seção. Para solucionar problemas reais é feito a formulação do modelo matemático com variáveis canalizadas. Vale ressaltar que atualmente é possível resolver problemas envolvendo milhares de restrições usando diversas variáveis aplicando o método Simplex computacionalmente (SALVADOR, 2016).

Neste sentido, para verificar o desempenho do método Simplex foi inserida a regra de Dantzig Normalizada. Conforme apresentado na literatura acadêmica, esta regra foi incorporada ao método buscando reduzir o número de iterações em alguns casos.

No presente trabalho, foram utilizados dados da fábrica de ração do Instituto Federal de Educação Ciência e Tecnologia do Espírito Santo, em Alegre - ES (IFES - Campus Alegre), para abordar o problema da produção de ração e estudado e implementado o Método Primal Simplex para problemas com restrições e variáveis canalizadas.

## MÉTODO PRIMAL SIMPLEX CANALIZADO

Cabe ressaltar que uma implementação bem estruturada do método Simplex consegue-se resolver problemas grandes e complexos computacionalmente de forma eficiente (MUNARI, 2009). Na literatura há diversas variações deste método, dentre elas destaca-se: encontrar melhores soluções em um curto tempo de execução, para fazer análise de sensibilidade ou para realizar pós-otimalidade no modelo matemático. É composto por procedimentos algébricos e por conceitos geométricos em sua estrutura (GEVEGY, 2022).

O Método Simplex foi um dos primeiros métodos utilizados na resolução de problemas de programação linear (DORNELLAS, 2014). É imprescindível buscar ferramentas na literatura para ter uma implementação do método eficiente e estável, principalmente envolvendo a estabilidade numérica. Neste caso devem ser utilizadas tolerâncias adequadas, representação eficiente de matriz básica e sua atualização que são componentes essenciais para uma boa implementação (MUNARI, 2009).

Um dos primeiros problemas usado por J. Laderman para validar o método Simplex foi o problema da dieta conforme apresentado por Namen e Bornstein (2004). O modelo era composto por 9 restrições por 27 variáveis de decisão usado no *National Bureau of Standards*. Para solucionar este problema foi utilizada calculadora de mesa, para resolução estimou-se o tempo de 12 homens/dia. Na época, os recursos computacionais estavam em um estágio inicial, já as implementações encontravam-se em uma fase bem básica.

O método Simplex era desenvolvido da mesma forma que era executado na teoria, isto é, o cálculo da inversa da base era realizado de forma explícita, as atualizações eram realizadas a cada iteração. Esta abordagem é muito custosa computacionalmente deixando o cálculo inviável para problemas de grande porte. Em 1954, ocorreu a primeira aparição significativa envolvendo a eficiência do método, que utilizou a denominada Forma Produto da Inversa (FPI). A partir de então, diversos trabalhos foram publicados utilizando este recurso e buscando formas mais eficientes para este algoritmo (MUNARI, 2009).

Cabe destacar que o método PSC se baseia no método Simplex tradicional que utiliza o formato padrão  $\text{Min } f(x)$  sujeito a  $Ax = b$  sendo  $x \geq 0$  conforme apresentado em Bazaraa (2010).

Neste caso, considere o modelo geral com restrições e variáveis canalizadas de otimização linear:

$$\begin{aligned} \text{Min } f(x) &= c^T x \\ \text{S. a. } d &\leq Ax \leq e \\ \alpha &\leq x \leq \beta \end{aligned} \quad (1)$$

Em que  $cc$  representa os custos do problema e o vetor  $xx$  as variáveis de decisão. Os vetores  $dd$  e  $ee$  representam os limitantes inferiores e superiores das restrições. Já  $\alpha\alpha$  e  $\beta\beta$  são os limitantes inferiores e superiores das variáveis utilizadas no problema. Para mais detalhes veja Gevegy (2022).

Sem perda de generalidade o método Primal Simplex Canalizado foi desenvolvido para o seguinte modelo:

$$\begin{aligned} \text{Min } f(x) &= c^T x \\ \text{S. a. } Ax &= b \\ L &\leq x \leq U \end{aligned} \quad (2)$$

Pressuponha que a matriz  $AA$  cujo posto é igual a  $mm$ , o vetor de custo  $cc$  e  $xx$  pertencentes ao conjunto  $\mathbb{R}^m \mathbb{R}^m$ , os vetores  $L, UL, U$  pertencentes ao conjunto  $\mathbb{R}^m \mathbb{R}^m$  e  $bb$  pertencente ao conjunto  $\mathbb{R}^m \mathbb{R}^m$ . Do modelo 1 para o 2 ocorre uma mudança de variável, transformando o problema em um equivalente. Os detalhes desta manipulação encontram-se em Gevegy (2022).

Dado estas considerações, é feito um particionamento nas colunas de  $AA$  da seguinte forma:

$$A = [B, N1, N2] \quad (3)$$

A matriz  $BB$  define a partição associada às variáveis básicas, as matrizes não básicas são representadas por  $N1N1$  e  $N2N2$ , no qual a primeira representa a partição associada às variáveis no limitante inferior ( $x_j = Lx_j = L$ ), e já a segunda o seu limitante superior ( $x_j = Ux_j = U$ ). De forma análoga, o vetor  $xx$  pode ser particionado em:

$$x = [x_B, x_{N1}, x_{N2}] \quad (4)$$

Na equação (4)  $x_{N1}x_{N1}$  representa as variáveis no limitante inferior, já  $x_{N2}x_{N2}$  representa as variáveis no limitante superior.

O mesmo ocorre com o vetor  $cc$  que é dado por:

$$c = [c_B, c_{N1}, c_{N2}] \quad (5)$$

Feito estas considerações, o modelo pode ser formulado da seguinte forma:

$$\begin{aligned} \text{Min } f(x) &= c_B x_B + c_{N1} x_{N1} + c_{N2} x_{N2} \\ \text{S. a. } Bx_B + N1x_{N1} + N2x_{N2} &= b \\ L_B &\leq x_B \leq U_B \\ x_{N1} &= L_{N1} \\ x_{N2} &= U_{N2} \end{aligned} \quad (6)$$

Para  $x_B, x_B$ , tem-se:

$$x_B = B^{-1}b - B^{-1}N1x_{N1} - B^{-1}N2x_{N2} \quad (7)$$

Após realizar estas atribuições, se  $L_B \leq x_B \leq U_B$   $L_B \leq x_B \leq U_B$ , tem-se uma solução básica factível, caso contrário, tem-se uma solução infactível. As variáveis básicas do problema são representadas pelos valores de  $x_B, x_B$ . Já que as variáveis não básicas  $x_{N1}, x_{N1}$  e  $x_{N2}, x_{N2}$  sendo a primeira em seu limitante inferior e a segunda no seu limitante superior respectivamente.

Ao substituir  $x_B, x_B$  no sistema, o modelo fica da seguinte forma:

$$\begin{aligned} \text{Min } f(x) &= c_B x_B + c_{N1} x_{N1} + c_{N2} x_{N2} \\ \text{S. a. } x_B &= B^{-1}b - B^{-1}N1x_{N1} - B^{-1}N2x_{N2} \\ L_B &\leq x_B \leq U_B \\ x_{N1} &= L_{N1} \\ x_{N2} &= U_{N2} \end{aligned} \quad (8)$$

Agora considere as seguintes definições conforme visto em Gevegy (2022):

- $I_B$  o conjunto de índices das variáveis básicas;
- $I_{N1}$  é o conjunto de índices das variáveis não básicas em seu limite inferior com  $n1$  elementos;
- $I_{N2}$  é o conjunto de índices das variáveis não básicas em seu limite superior com  $n2$  elementos;
- $N1$  é a partição (matriz) não básica associada às variáveis não básicas que estão em seu limite inferior, isto é,  $N1 = \{a_{.i} \mid i \in I_{N1}\}$ ;
- $N2$  é a partição (matriz) não básica associada às variáveis não básicas que estão em seu limite superior, isto é,  $N2 = \{a_{.i} \mid i \in I_{N2}\}$ .

Sendo assim, de acordo com o método Simplex é preciso realizar os cálculos envolvendo as matrizes e os vetores conforme descrito anteriormente da seguinte forma:

1. Cálculo da solução básica:  
 $x_B = B^{-1}b - B^{-1}N1x_{N1} - B^{-1}N2x_{N2}$ .
2. Cálculo do vetor multiplicador Simplex:  
 $\pi = c_B^T B^{-1}$ .
3. Cálculo dos custos reduzidos:  
 $\hat{c}_j = c_j - \pi^T a_{.j}, \forall j \in I_{N1} \cup I_{N2}$  (critério para escolha do índice que entra na base,  $q$ ).
4. Direção Simplex:  
 $d_B = B^{-1}a_{.q}, q \in I_{N1} \text{ ou } I_{N2}$  (índice não básico que entra na base).
5. Escolha do  $\delta$  (tamanho do passo) de tal forma que a nova solução seja factível, ou seja:  
 $L_{B_i} \leq x_{B_i} \pm \delta d_{B_i} \leq U_{B_i}, \quad i = 1, \dots, m$ .
6. Atualização da inversa da base:  
 $(\bar{B}^{-1}) = E B^{-1}$ , em que  $B^{-1}$  representa a inversa da base,  $(\bar{B}^{-1})$  é a nova inversa da matriz base e  $E$  a matriz de transformação linear, veja Bazaraa (2010).

Na inicialização do método Simplex as variáveis de decisão são iniciadas no seu limitante inferior ou superior, isto é,  $x_i = L_i, x_i = L_i$  sendo  $i$  variando de  $1, \dots, n1, 1, \dots, n1$  ou  $x_i = U_i, x_i = U_i$  sendo  $i$  variando de  $1, \dots, n2, 1, \dots, n2$ . Caso a solução básica inicial não seja factível, recorre-se à Fase I do método Simplex. Esta etapa objetiva buscar uma solução básica factível inicial. A Fase I consiste de mais uma implementação do próprio método. Logo, a Fase II inicia-se então com esta solução básica factível e determina (encontra) a solução ótima para o problema original.

Conforme apresentado por Sousa (2011) há diversas variações do método Simplex em Otimização Linear. Além do mais, existem diversas abordagens na seleção das variáveis que entram e saem da base.

No geral, métodos tipo Simplex, há flexibilidade na escolha da variável que entra na base. Por exemplo, no método Primal Simplex, qualquer variável do custo relativo negativo é candidata a ser escolhida para entrar na base. Existem diversas maneiras para escolher as variáveis que entram na base, a regra de Dantzig Normalizada é uma delas que será apresentada brevemente na próxima seção.

## REGRA DE DANTZIG NORMALIZADA

A regra de Dantzig usual (D) seleciona o índice da variável que possui o maior valor em módulo durante o cálculo dos custos, enquanto a regra de Bland faz o cálculo de todos os custos relativos das variáveis não básicas e seleciona o índice do valor mais negativo (MACULAN e FAMPA, 2004).

Os primeiros experimentos com a regra de Dantzig Normalizada (DN) foram realizados por Kuhn e Quandt (1963) e por Wolfe e Cuttler (1963). No início, esta regra foi rejeitada devido ao fato de sua implementação ser direta e a sua complexidade que a tornava impraticável. Entretanto, Goldfarb e Reid (1977), em seus experimentos com o método Simplex usando a regra de Dantzig Normalizada, realizaram a derivação de fórmulas de recorrência e realizaram testes computacionais que evidenciaram a superioridade desta regra em relação à regra usual de Dantzig para diversos problemas (SOUSA, 2011).

A regra usual seleciona a direção que possui o coeficiente mais negativo que reduz a função objetivo, porém, essa seleção não considera o comprimento das arestas que determina a escolha da direção, o que pode ser diferente. Por outro lado, a regra de Dantzig Normalizada toma o custo relativo normalizado mais negativo, levando em conta o quanto cada variável pode contribuir com a função objetivo em cada unidade desta direção. A equação a seguir apresenta este cálculo.

$$\text{Min} \left\{ \frac{\hat{c}_j}{d^j} \right\}, j \in J_N \quad (9)$$

O vetor  $\hat{c}_j$  representa o custo relativo com  $j$  pertencente ao conjunto não básico. Já o vetor  $d^j$  é a direção Simplex associada ao  $\hat{c}_j$ .

Na primeira iteração do método Simplex usando a regra de Dantzig Normalizada é feito o cálculo explícito de todas as direções, fato que consome tempo a mais somente na primeira iteração. Portanto, Forrest e Goldfarb (1992) utilizaram fórmulas recursivas para a atualização nas normas das direções a partir da 2ª iteração.

Na utilização da regra de Dantzig Normalizada, a escolha do índice que entra na base “ $qq$ ” é definida pela maior violação primal, isto para o limitante superior quanto o inferior (que são os custos relativos) dividido pela norma da direção Primal Simplex. Neste caso, escolhe-se o índice não básico que possui o maior valor em módulo

da razão. Este cálculo é executado conforme a seguinte fórmula:

$$\frac{\hat{c}_j}{\|d^j\|}, \quad j = 1, \dots, m \quad j \in J_{N1} \cup J_{N2} \quad (10)$$

Este cálculo é conhecido na literatura como cálculo das razões. Entretanto, são realizados outros cálculos extras, por exemplo, calcula-se o alfa ( $\alpha_j$ ), resolve-se o sistema dado por  $B^T v$  e em seguida é feito o produto interno dado por  $(\alpha_j v)$ . Na utilização da regra de Dantzig Normalizada é necessária a execução dos seguintes passos:

Cálculo da 1ª iteração:

- Direção Simplex – Calcula-se para todas as direções  $d^j = B^{-1}a_j$ .
- Cálculo das Normas dado por  $\|d_B^j\|, j \in N1 \cup N2$ .
- Cálculo das razões  $\frac{\hat{c}_j}{\|d^j\|}$ , toma a maior razão em módulo ( $q$ ).
- Calcula  $v = (B^{-1})^T d_B$  (é um dos cálculos extras).

A partir da 2ª iteração usam-se as fórmulas de recorrência conforme visto em Forrest e Goldfarb (1992).

Neste sentido, para calcular recursivamente (em função da anterior) as normas das direções, na segunda iteração em diante usa-se o quadrado da norma de forma a facilitar os cálculos algébricos, isto é, considere apenas das componentes básicas da direção Simplex  $d_B^j$ .

Assim, considere  $\gamma^j = \|d_B^j\|^2$ , para calcular o quadrado das normas das novas direções Primal Simplex é feito da seguinte forma:

$$\bar{\gamma}^j = \gamma_j - 2\bar{\alpha}_j v + \bar{\alpha}_j^2 \gamma_q, \quad j \in J_{N1} \cup J_{N2}, \quad j \neq q \quad (11)$$

Caso  $j = q$ , tem-se:

$$\bar{\gamma}^q = \frac{\gamma_q}{\alpha_j^2}, \quad j = q \quad (12)$$

Para mais detalhes consulte Sousa (2011) e Forrest e Goldfarb (1992).

## RESULTADOS E DISCUSSÕES

Para desenvolver o método e executar os testes foi utilizado um notebook com 8 GB de memória RAM, processador Intel Celeron CPU 1000M @1.8 Ghz. O sistema operacional utilizado foi o Windows 8.1 Single Language de 64 bits. Como linguagem de programação usou-se o C++, além disso, para programar usou-se o IDE Visual Studio Community com o compilador GCC (MingGW). Os problemas solucionados foram classificados como sendo de pequeno porte segundo Gevegy (2022), baseados no fato das dimensões do problema serem inferiores a 100.

A execução do programa conta com um menu onde o usuário seleciona o tipo de ração e digita o nome do arquivo (no formato *txt*) onde esta formulado o modelo do problema a ser carregado. Feito isso, o método é executado e os resultados podem ser visualizados através da Figura 1 ao lado.

Figura 1 – Tela de solução da *instAves6*.

conta com a solução ótima, os custos relativos, o tempo computacional e o valor ótimo da instância *instAves6*. As variáveis  $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9$  representam os ingredientes presentes nesta ração. Além disso, é apresentada a quantidade de ração a ser produzida, no caso, 100 kg. Tanto na fase I quanto na fase II são apresentados o total de iterações em cada uma destas fases. Por último, o sistema informa o custo em reais para produzir 100 kg desta ração, isto é, R\$ 226,9209 reais.

Cabe destacar que durante a execução das instâncias, o tempo registrado foi de 0,000000000001 segundos. Sendo assim, o tempo computacional é praticamente nulo.

A ração é feita realizando a mistura destes ingredientes levando em consideração as especificações diárias de quantidade conforme exigências nutricionais dos diversos tipos de animais (aves e suínos, por exemplo) (ROSTAGNO et. al, 2011). O processamento destes ingredientes é feito de forma a satisfazer os critérios de nutrientes empregados na mistura, isto é, atendendo a quantidade exigida de cálcio, proteínas, carboidratos, vitaminas e etc.

Neste trabalho, verifica-se o desempenho do PSC na otimização dos problemas de produção de ração do Instituto Federal do Espírito Santo (IFES). Para tal propósito, foram formulados 24 problemas (instâncias) que representam diferentes cenários da produção de ração. Por meio de estudos preliminares e de uma visita técnica, foi feito um estudo de caso no IFES em Alegre, ES. A formulação utilizada nas instâncias

```
SOLUCAO OTIMA:
-----
Variaveis      valores      custo relativo
x[1] =         61.66886640      0.0000000000
x[2] =         28.59825547      0.0000000000
x[3] =          8.4000000000     -0.3561791967
x[4] =          1.032879206     0.0000000000
x[5] =          0.0000000000     26.54382080
x[6] =          0.3000000000     -0.2061791967
x[7] =          0.0000000000     17.53382080
x[8] =          0.0000000000     35.34382080
x[9] =          0.0000000000    -102.3680639

Total          100.000 Kg

-----
VALOR OTIMO:
Funcao Objetivo Iteracoes  Tempo (seg)
0.0000000000000 46         0.000000000000 (FASE I)
226.9209234847   5         0.000000000000 (FASE II)
Total           51         0.000000000000 (TOTAL)

FUNCAO OBJETIVO R$: 226.9209234846942
```

Fonte: (Gevegy, 2022).

O sistema apresentado na Figura 1

a seguir é uma representação dos problemas de produção de rações expressos em kg, mas a formulação mais adequada é tratada em termos de % conforme Gevegy (2022). A solução encontrada pelo método PSC foi comparada com a do software CPLEX. As instâncias de 1 até 15 são equivalentes às rações de aves. Já as instâncias de 16 a 24 equivalem às rações para suínos.

A esparsidade é uma característica de um Problema de Programação Linear que pode ser definida em função do número de elementos não-nulos da matriz de coeficientes do problema. Assim, a esparsidade pode ser calculada tomando o total de elementos não-nulos pelo total de elementos da matriz (SANTOS, 2020). A esparsidade representa a baixa densidade de elementos não-nulos presentes na matriz do problema.

No geral, os problemas formulados têm como características uma média de 21,88 linhas por 10,75 colunas. Além disso, apresenta um percentual de esparsidade de 70,56%. As instâncias relacionadas às aves possuem em média 22,67 linhas por 9,47 colunas e uma esparsidade de 69,36%. Já as instâncias relacionadas a suínos possuem uma média de 20,67 linhas por 12,89 colunas apresentando um percentual de esparsidade de 72,57%. Percebe-se que os problemas apresentam um percentual de 65% a 75% dos componentes com valores nulos. A instância *instAves15* possui a maior esparsidade com 83,16%.

A Tabela 1 apresenta o número de iterações do método PSC usando a regra de Dantzig usual (D) e Dantzig Normalizada (DN) e compara com o CPLEX para os problemas que envolvem aves.

**Tabela 1:** Número de iterações - PSC e CPLEX - Aves.

Instância	PSC D		PSC DN		CPLEX
	Fase I	Fase II	Fase I	Fase II	
<i>instAves1</i>	43	4	45	3	5
<i>instAves2</i>	40	4	40	5	4
<i>instAves3</i>	40	4	41	3	5
<i>instAves4</i>	42	4	28	7	6
<i>instAves5</i>	44	5	41	5	4
<i>instAves6</i>	46	5	41	3	8
<i>instAves7</i>	34	4	35	7	1
<i>instAves8</i>	44	4	48	2	7
<i>instAves9</i>	44	3	43	2	6
<i>instAves10</i>	43	5	46	3	7
<i>instAves11</i>	18	4	19	7	9
<i>instAves12</i>	38	4	40	13	1
<i>instAves13</i>	46	5	47	7	4
<i>instAves14</i>	46	4	46	4	3
<i>instAves15</i>	65	5	66	3	6
MÉDIA	42,20	4,27	41,73	4,93	5,07

Fonte: Adaptado Gevegý (2022).

**Tabela 2:** Número de iterações - PSC e CPLEX - Suínos.

Instância	PSC D		PSC DN		CPLEX
	Fase I	Fase II	Fase I	Fase II	
<i>inst1Suínos</i>	34	4	34	3	3
<i>inst2Suínos</i>	37	4	38	5	3
<i>inst3Suínos</i>	37	4	36	5	4
<i>inst4Suínos</i>	38	8	41	6	5
<i>inst5Suínos</i>	24	5	23	4	3
<i>inst6Suínos</i>	36	5	36	5	3
<i>inst7Suínos</i>	40	4	40	5	6
<i>inst8Suínos</i>	37	5	36	6	4
<i>inst9Suínos</i>	37	7	37	8	4
MÉDIA	35,56	5,11	35,66	5,22	3,89

Fonte: Gevegý (2022).

Observa-se que a regra de Dantzig usual (D) teve uma média de 42,20 iterações na fase I e 4,27 na fase II. Já usando a regra de Dantzig Normalizada (DN) teve uma média de 41,73 iterações na fase I e 4,93 na fase II. No geral, as duas regras tiveram desempenho semelhante.

Nota-se que para algumas instâncias de aves, como *instAves15*, exigem um número maior de iteração como visto na fase I do método PSC. Esta instância aparenta ser mais complexa, o que pode necessitar de estratégias adicionais com o intuito de fazer o método convergir mais rápido.

A regra de DN teve um desempenho melhor que a regra D na instância *instAves4* na fase I. No entanto, em muitos casos, a regra DN não reduziu o número de iterações para os problemas propostos.

Embora o número de iterações na fase I seja muito maior que na fase II, mas na média foi inferior a  $2m$ , em que  $m$  é o número de restrições (linhas) técnicas do problema. Este desempenho está de acordo com muitos trabalhos na literatura.

Na fase II de ambas as regras teve um número de iterações na casa das unidades, enquanto que na fase I na casa das dezenas.

Nota-se que somando o esforço requerido pelo PSC nas fases I e II é muito maior que o CPLEX, mas considerando-se apenas a fase II o desempenho é similar. Vale ressaltar que o software foi desenvolvido há muito tempo e sofre diversas atualizações para melhorar a sua performance, em especial neste caso para obter uma solução básica inicial factível.

Agora, a Tabela 2 apresenta o desempenho do PSC frente ao software CPLEX para problemas relacionados às rações suínas.

Os resultados mostram que a fase I usando a regra usual apresentou uma média de 35,56 iterações e na fase II 5,11, entretanto usando a regra de Dantzig Normalizada obteve-se uma média de 35,66 na fase I e 5,22 na fase II. No geral, pode-se observar que tanto a fase I quanto a fase II (PSC D) as regras tiveram desempenho similar. Em três instâncias da fase I a regra de Dantzig Normalizada conseguiu fazer menos iterações que a regra usual (*inst3Suínos*, *inst5Suínos* e *inst8Suínos*).

No geral, nota-se um desempenho muito parecido para PSC fase II e CPLEX com relação aos problemas das aves (Tabela 1) e dos Suínos (Tabela 2) e uma pequena redução na fase I para o PSC.

Durante a execução do método PSC, observou-se que nos casos onde há a fixação de alguns alimentos, reduzindo o espaço de busca, o PSC passa a ter um desempenho melhor. Note que o desempenho do CPLEX e PSC foram muito próximos aos resultados obtidos com os problemas da ração para aves e para suínos.

Portanto, novamente comparando a fase II do PSC com o CPLEX observam-se valores iguais ou superiores ao software comercial. Percebe-se que a iteração na fase II é equivalente à encontrada no CPLEX.

Cabe destacar que o PSC é uma opção ao CPLEX para solucionar problemas de produção de ração. Já em casos que demandem uma convergência mais rápida, o CPLEX pode ser recomendado. É necessário explorar estratégias de inicialização e refinamento para otimizar ainda mais o desempenho do PSC.

A Tabela 3 apresenta os valores ótimos encontrados para todos os problemas com relação às instâncias de aves para verificar a precisão numérica do PSC com o software CPLEX,

**Tabela 3:** Qualidade da solução PSC e CPLEX – Aves/IFES.

Instância	PSC D FO	PSC DN FO	CPLEX FO	Diferença PSC - CPLEX
instAves1	317,938263624112	317,938263626184	317,938263610	0,00000001411
instAves2	312,737313637549	312,737313576780	312,737313510	0,00000012755
instAves3	238,200271280072	238,200271280072	238,200271370	-0,00000008993
instAves4	254,587177621759	254,587177602686	254,587177600	0,00000002176
instAves5	300,102859384186	300,102859235882	300,102859460	-0,00000007581
instAves6	226,920918645802	226,920916594807	226,920916580	0,00000206580
instAves7	255,155999999823	255,156000000001	255,156000000	-0,00000000018
instAves8	254,613130338601	254,613130280827	254,613130270	0,00000006860
instAves9	250,549454619190	250,549454639079	250,549454600	0,00000001919
instAves10	268,578678347934	268,578678328143	268,578678360	-0,00000001207
instAves11	252,897099737040	252,897100778399	252,897100990	-0,00000125296
instAves12	216,202000003176	216,202000001759	216,202000000	0,00000000318
instAves13	217,630286507885	217,630287263048	217,630286490	0,00000001788
instAves14	259,373527128053	259,373527014805	259,373527030	0,00000009805
instAves15	266,925810178974	266,925810290845	266,925810130	0,00000004897

Fonte: Geveg (2022).

Destaca-se que nas instâncias *instAves3*, *instAves5*, *instAves7*, *instAves10*, *instAves11* o método PSC encontrou um valor ótimo com uma precisão melhor que a do CPLEX. Isto pode ter ocorrido devido ao valor de tolerância usado na implementação do PSC, sendo que, para estes problemas foram obtidos valores com uma ótima precisão numérica.

No geral, nota-se que os valores ótimos encontrados pelo PSC estão muito próximos dos resultados do software CPLEX, iguais até a nona casa decimal. Assim, tem-se uma implementação bem estável do PSC.

Para verificar que a precisão numérica do PSC com relação às instâncias de suínos, apresenta-se a solução ótima encontrada para todos os problemas na Tabela 4 a seguir, para PSC e CPLEX.

**Tabela 4:** Qualidade da solução PSC e CPLEX – Suínos/IFES.

Instância	PSC D FO	PSC DN FO	CPLEX FO	Diferença PSC - CPLEX
instSuínos1	254,604283816199	254,604283816199	254,60428378	0,000000036199
instSuínos2	244,478869885988	244,478869885988	244,47886981	0,000000075988
instSuínos3	222,937456386946	222,937456382652	222,93745644	-0,000000053054
instSuínos4	277,972291084490	277,972291027866	277,97229105	0,000000034490
instSuínos5	282,104672523379	282,104672515090	282,10467251	0,000000013379
instSuínos6	295,548511621892	295,548511569412	295,54851155	0,000000071892
instSuínos7	244,931766947135	244,931766909286	244,93176693	0,000000017135
instSuínos8	236,974388647334	236,974388647334	236,97438853	0,000000117334
instSuínos9	223,612240931667	224,131361436225	223,61224102	-0,000000088333

Fonte: Geveg (2022).

Neste sentido, o método obteve o valor ótimo com uma precisão similar ao encontrado nas instâncias de aves. Cabe destacar que nas instâncias *instSuínos3* e *instSuínos9* o método PSC encontrou valor ótimo com uma precisão melhor do que foi apresentado no software CPLEX. Nota-se que a qualidade da solução ótima do PSC se iguala aos resultados do software CPLEX até a sétima casa decimal.

Assim, verifica-se novamente que a implementação do PSC tem uma ótima estabilidade numérica para os problemas tratados.

## CONCLUSÃO

O PSC foi estudado e implementado para solucionar problemas que retratem a produção de ração de aves e suínos da fábrica de ração do IFES/Alegre. O método foi usado em 24 problemas, sendo 15 deles relacionados à ração de aves e os outros 9 para suínos. A solução foi comparada com o software CPLEX e demonstrou uma ótima precisão da solução. O método PSC apresentou um baixo número de iterações (fase I e Fase II). Frente ao CPLEX foi inferior, mas o software é uma implementação comercial desenvolvida e aprimorada há muito tempo.

Durante a pesquisa foram realizados diversos testes exaustivos de forma a aprimorar a implementação. Para os testes executados encontraram-se soluções com uma ótima precisão numérica.

Para validar o método foi realizado comparações com relação à solução do software

CPLEX. Na maioria dos casos o PSC encontrou solução que se iguala até a nona casa decimal quando comparado com o software comercial.

O método PSC auxilia na tomada de decisão na produção da fábrica. Além disso, informa o quanto cada ingrediente contribui tanto na quantidade em quilo de ração quanto no custo.

Assim, o método se mostrou eficaz na solução dos problemas reais da produção de ração de forma precisa e rápida. O presente trabalho evidencia que o PSC é uma ferramenta que pode contribuir muito no setor de produção de ração.

Como trabalhos futuros pretende-se adicionar estratégias para reduzir a quantidade de iteração na fase I. Implementar uma interface gráfica para facilitar a gestão dos ingredientes e utilizar o método para fazer a formulação da ração de aves e suínos considerando a formulação em termos de porcentagem como tratada em diversos trabalhos na literatura.

Além do mais, pretende-se trabalhar com outros tipos de ração utilizando o PSC. Por exemplo, o método pode ser empregado na fabricação de ração para cães e gatos, bovinos, caprinos e peixes.

Pretende-se aprimorar uma interface que está sendo desenvolvida de forma a facilitar a formulação e o gerenciamento da produção de ração. Para isso, é necessário integrar o PSC com a interface de forma a facilitar a formulação e auxiliar na tomada de decisão conforme disponibilidade de insumos (ingredientes).



## REFERÊNCIAS

- BARBOSA, G. M. Utilização da programação linear na otimização de resultados de produção na empresa. **Revista Integração - Universidade São Judas Tadeu**, São Paulo, 66 (2014), 49-58.
- BAZARAA, M. S.; JARVIS, J.; SHERALI H. D. **Linear programming and network flows**. New Jersey: John Wiley & Sons, 2010.
- CASTRO, L. Y.; BORGERT, A; DE SOUZA, F, R. Definição do mix de produção em uma indústria de lácteos com uso da programação linear: um estudo de caso. **Anais do Congresso Brasileiro de Custos-ABC**. Foz do Iguaçu, PR, 2015.
- CHAVES, V. H. C. **Perspectivas históricas da pesquisa operacional**. Dissertação de Mestrado, Universidade Estadual Paulista (UNESP), Rio Claro (SP), 2011.
- DORNELLAS, C. A. **Análise de tópicos relevantes em programação linear e aplicações no ensino de engenharia**. Dissertação de Mestrado, Universidade Estadual Paulista (UNESP), Ilha Solteira, 2014.
- DOS SANTOS, L. F. **Método dual simplex para problemas canalizados com estruturas de dados eficientes**. Dissertação de Mestrado, Universidade Federal Fluminense (UFF), Volta Redonda, 2020.
- FEED, ALL ABOUT. **World feed production increases to over 1.2 billion tons**. 2022. Disponível em: <https://www.allaboutfeed.net/market/market-trends/world-feed-production-increases-to-over-1-2-billion-tons/>. Acesso em 27/03/2022.
- FORREST, J. J., GOLDFARB, D. Steepest-edge simplex algorithms for linear programming. **Mathematical programming** 57, 1 (1992), 341-374.
- GEVEGY, M. L. **Método Simplex Canalizado para Produção de Ração: Um Estudo de Caso**. Dissertação de Mestrado, Universidade Federal Fluminense (UFF), Volta Redonda, 2022.
- MACULAN, N., FAMPA, M. H. **Otimização linear**. Rio de Janeiro: Editora Universidade da UFRJ, 2004.
- MUNARI JUNIOR, P. A. **Técnicas computacionais para a implementação eficiente e estável de métodos tipo simplex**. Dissertação de Mestrado, Universidade São Paulo (USP), São Carlos, São Paulo, 2009.
- NAMEN, A. A., BORNSTEIN, C. T. Uma ferramenta para avaliação de resultados de diversos modelos de otimização de dieta. **Pesquisa Operacional** 24, 3 (2004), 445-465.
- NEME, R. **Digestibilidade verdadeira e biodisponibilidade da lisina sulfato e da lisina hcl determinadas em aves**. Universidade Federal de Viçosa (2000).
- ROSTAGNO, H. S., ALBINO, L. F. T., DONZELE, J., GOMES, P., OLIVEIRA, R. d., LOPES, D., FERREIRA, A., BARRETO, S. d. T., EUCLIDES, R. **Tabelas brasileiras para aves e suínos. Composição de alimentos e exigências nutricionais** 2 (2011), 186.
- SALVADOR, T. R. **Um estudo do método simplex e sua aplicação na radioterapia**. Dissertação de Mestrado, Universidade Federal de São Carlos, Sorocaba, 2016.
- SOUSA, R. S. Método dual simplex com a regra de Dantzig normalizada. **XLIII Simpósio Brasileiro de Pesquisa Operacional (2011)**, 2447-2457.
- TONIN, H. N. **Modelo de programação linear aplicado a redução de custo na formulação de ração para nutrição de bovinos**. 2019. 49 f. Dissertação de Mestrado, Universidade de Uberaba (Uniube), Uberaba, Minas Gerais, 2019.
- ULBRICHT, G., RIBEIRO, R. Otimização em formulação de rações utilizando técnicas de programação linear: Um estudo de caso em processos industriais. Em **II Congresso Brasileiro de Engenharia de Produção – CONBREPRO**. Ponta Grossa, PR, Brasil (2012), vol. 28.

## DADOS DOS AUTORES

**Marcio Lopes Gevegy**. Mestre em Modelagem Computacional, Tutor, Universidade Federal Fluminense, UFF/CEDERJ. E-mail: marciogevegy@id.uff.br.

**Ricardo Silveira Sousa**. Doutor em Ciência da Computação e Matemática Computacional, Docente, Universidade Federal Fluminense, UFF/INFES. E-mail: rsousa@id.uff.br.

**Wagner Rambaldi Telles**. Doutor em Modelagem Computacional, Docente, Universidade Federal Fluminense, UFF/INFES. E-mail: wtelles@id.uff.br.

# PREVISÃO DO CONSUMO DE ENERGIA ELÉTRICA ATRAVÉS DE MODELOS DE SÉRIES TEMPORAIS UNIVARIADOS E MULTIVARIADOS

Rafael Santos de Oliveira

Felipe Leite Coelho da Silva

Josiane da Silva Cordeiro Coelho

**RESUMO:** O setor industrial é o maior consumidor de energia elétrica no Brasil e notavelmente é um dos mais importantes pelo seu grande impacto econômico e social. O objetivo deste trabalho é comparar os modelos de previsão autoregressivos vetoriais e o regressão dinâmica aplicados aos dados de energia elétrica do setor industrial. Para o modelo de regressão dinâmica será utilizado como variável exógena o PIB industrial. A capacidade preditiva dos modelos será avaliada pelas métricas MAPE e RMSE. Os resultados indicam que o modelo de regressão dinâmica apresentou a melhor capacidade preditiva.

**Palavras-chave:** Séries Temporais. Energia Elétrica. Modelos de Previsão. Produto Interno Bruto.

**ABSTRACT:** The industrial sector is the largest consumer of electrical energy in Brazil and is notably one of the most important due to its great economic and social impact. The objective of this work is to compare vector autoregressive forecasting models and dynamic regression applied to electrical energy data from the industrial sector. For the dynamic regression model, industrial GDP will be used as an exogenous variable. The predictive capacity of the models will be evaluated using the MAPE and RMSE metrics. The results indicate that the dynamic regression model presented the best predictive capacity.

**Keywords:** Time Series. Electricity. Forecast Models. Gross Domestic Product.

## INTRODUÇÃO

Nos últimos anos, o consumo de energia elétrica e o produto interno bruto (PIB) do setor industrial brasileiro vem sendo investigados devido a sua importância no desenvolvimento nacional em termos sociais e econômicos. Dessa forma, a análise e previsão desses conjuntos de dados são importantes para o planejamento do sistema elétrico e, também, para o desenvolvimento de setores da indústria.

Neste contexto, a análise de séries temporais é importante para investigar e prever o consumo de energia elétrica (Sadownik; Barbosa, 1999). Esta análise permite identificar padrões e tendências nos dados históricos, o que é fundamental para a previsão de eventos futuros. Também é possível identificar valores atípicos e anomalias nos dados, que podem indicar problemas ou eventos excepcionais que requerem atenção especial. Neste trabalho, os conjuntos de dados tratados consideram períodos trimestrais para poder adequar-se a alguns pressupostos dos modelos de séries temporais. Primeiramente, será realizada uma análise exploratória dos dados referentes a cada um dos seus conjuntos, depois a verificação dos pressupostos dos modelos e por fim, a predição dos mesmos.

A previsão do consumo de energia elétrica no setor industrial brasileiro tem sido realizada por diferentes modelos de séries temporais, tais como os modelos de redes neurais artificiais (Silva et al., 2022), modelos de regressão dinâmica integrados a uma abordagem bottom-up (Silva, et al. 2018; Silva et al., 2019.) e os modelos de amortecimento exponencial e de Box-Jenkins (Mendes et al., 2023).

De acordo com Amaro et al. (2017), a previsão da demanda de energia de curto prazo é difícil de ser realizada porque envolve uma grande influência de vários fatores sociais e econômicos. Mas, sugerem que os modelos univariados podem ser utilizados para prever a demanda de energia elétrica de forma eficiente. Por outro lado, Castro et al. (2016) mostraram que a inclusão de variáveis que se inter-relacionam com a variável de interesse, aumenta a capacidade de previsão, sugerindo a utilização de modelos multivariados.

O objetivo deste trabalho é comparar os modelos autorregressivos vetoriais (VAR) e o de regressão dinâmica para a previsão do consumo de energia elétrica do setor industrial. No modelo de regressão dinâmica, será utilizado o PIB como variável regressora. A capacidade predi-

tiva dos modelos foi avaliada pelas métricas do erro médio absoluto percentual (MAPE) e a raiz do erro quadrático médio (RMSE).

## METODOLOGIA

Neste trabalho, foram utilizados os dados do PIB (em reais) e o consumo de energia elétrica (em MWh) do setor industrial brasileiro, considerando o período de janeiro de 2002 até dezembro de 2021. Esses dados estão disponíveis no sítio eletrônico do Instituto de Pesquisa Econômica e Aplicada (IPEA, 2022). Para testar o poder de ajuste e as generalizações das previsões, os dados foram divididos em dois conjuntos: o conjunto de treino, com intuito de exploração dos dados, ajustes e validação dos modelos e o conjunto de teste para avaliar a capacidade preditiva.

Na primeira parte do estudo foi apresentada uma análise exploratória das séries temporais para compreender seu comportamento e suas características de nível, tendência, ciclo, sazonalidade e ruído. Estas informações são importantes para identificar o tipo de modelo a ser ajustado.

O conjunto com as séries temporais do consumo de energia elétrica industrial e PIB foi dividido em duas partes: as primeiras 68, 72 e 76 observações trimestrais de cada série foram utilizadas como conjunto de treinamento, correspondendo aos anos de 2002 até 2018, 2002 até 2019 e 2002 até 2020, respectivamente, para estimar os parâmetros e as últimas 4 observações como o conjunto de teste, ou seja, para obter como previsão 2019, 2020 e 2021.

Dessa forma, foram ajustados os modelos VAR e o de regressão dinâmica. Em seguida, foram obtidas as previsões para cada uma das séries temporais considerando 4 passos à frente. Para selecionar o modelo mais adequado foram utilizadas as métricas de acuracidade do erro médio absoluto percentual (MAPE) e a da raiz do erro quadrático médio (RMSE).

## MODELO AUTORREGRESSIVOS VETORIAIS - VAR

Os modelos autorregressivos vetoriais (VAR) são amplamente utilizados para análise de séries temporais multivariadas e têm sido aplicados para avaliar modelos macroeconômicos, investigar os efeitos de política monetária, e ob-

ter as projeções de variáveis de interesse. Esses modelos foram apresentados por Sims (1980). Entre os modelos multivariados de séries temporais, os modelos VAR são os mais utilizados na prática, devido a sua semelhança com os modelos de regressão e a relativa facilidade de ajustar esses modelos às séries temporais reais (Box *et al.*, 2015). Por exemplo, os parâmetros podem ser estimados usando métodos de mínimos quadrados que geram expressões de forma fechada para as estimativas. Esses modelos são utilizados em econometria para descrever o comportamento dinâmico de séries temporais econômicas (ou financeiras) e para produzir previsões.

O modelo vetorial autorregressivo de ordem  $p$ , VAR( $p$ ), para  $n$  variáveis é dado por

$$X_t = \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + a_t, \quad (1)$$

em que  $X_t$  é um vetor de ordem  $(n \times 1)$ , composto pelo  $t$ -ésimo valor das variáveis analisadas,  $\phi_0$  é um vetor de ordem  $(n \times 1)$ , composto por constantes,  $\phi_k$  é uma matriz de ordem  $(n \times n)$  composta por elementos  $\psi_{ijk}$ ,  $i = 1, 2, \dots, n, j = 1, 2, \dots, n, k = 1, 2, \dots, p$  e  $a_t$  é um vetor de ordem  $(n \times 1)$ , composto pelos  $t$ -ésimos erros aleatórios.

## MODELOS DE REGRESSÃO DINÂMICA

Os modelos de regressão dinâmica combinam as características de variáveis explicativas com a dinâmica de séries temporais. Deve-se atentar que vocábulo “regressão dinâmica” não estabelece que os parâmetros do modelo evoluem com o tempo. Todavia, a palavra “dinâmica” caracteriza um modelo de regressão agregado pela estrutura de dependência de uma série temporal (Cochrane, 1949).

A aplicação dos modelos de regressão dinâmica fica condicionada a existência da estrutura de dependência entre a variável de interesse e variáveis causais e, concomitantemente a estrutura de correlação da série dependente (série a ser explicada) sugerir que não se pode supor a independência dos erros.

Os modelos de regressão dinâmica podem ser descritos por

$$\phi(B)Y_t = \beta x_t + \epsilon_t \quad (2)$$

onde  $Y_t$  é a variável dependente (endógena) no instante  $t$ ;  $\beta$  é o vetor de coeficientes das variáveis causais;  $x_t$  é o vetor de variáveis causais

(exógenas) no instante  $t$ ;  $\epsilon_t$  é o ruído aleatório associado ao modelo, onde supõem-se que os  $\epsilon_t$  são independentes e identicamente distribuídos com densidade  $N(0, \sigma^2)$ ;  $\phi(B)$  é o polinômio autorregressivo de ordem  $p$ , dado por

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (3)$$

sendo  $B$  o operador de atraso.

Assim, a estrutura do modelo de regressão dinâmica possibilita assumir como elementos  $x_t$  variáveis causais e juntamente suas defasagens.

O polinômio  $\phi(B)$  presente no modelo proporciona uma grande flexibilidade desta classe de modelos, porém, dificulta a procura por um modelo adequado.

O modelo de regressão dinâmica (1), pode ser considerado como um caso particular do que é conhecido na literatura como modelo de regressão de Cochrane e Orcutt generalizado que é dado por

$$\phi(B)y_t = \beta x_t + w_t \quad (4)$$

$$R(B)w_t = \epsilon_t \quad (5)$$

onde  $R(B)$  é um polinômio auto-regressivo.

A equação (4) pode ser reescrita isolando-se o erro  $w_t$ , da seguinte forma

$$w_t = \phi(B)y_t - \beta x_t \quad (6)$$

Substituindo a expressão (6) na equação (5), tem-se que

$$R(B) \cdot [\phi(B)y_t - \beta x_t] = \epsilon_t \quad (7)$$

Pode-se notar pela equação (7) que o modelo de regressão generalizado de Cochrane & Orcutt introduz defasagens tanto na variável dependente ( $Y_t$ ) quanto nas causais. Também é estabelecido por esta equação que a introdução do polinômio autorregressivo  $R(B)$  não afeta a relação de causalidade entre  $y_t$  e  $x_t$ . A equação (7) pode ser escrita em termos de novas variáveis  $Y_t^*$  e  $x_t^*$  dadas por

$$Y_t^* = R(B)y_t \quad (10)$$

$$x_t^* = R(B)x_t \quad (11)$$

onde  $R(B)$  é denominado de “fator comum” e representa a estrutura de correlação presente no erro  $w_t$ .

Com as novas variáveis obtêm-se a seguinte equação

$$\phi(B)Y_t^* = \beta x_t^* + \epsilon_t \quad (12)$$

Assim, o modelo de regressão generalizado de Cochrane & Orcutt se reduz ao modelo de regressão dinâmica usual ao se considerar as novas variáveis  $Y_t^*$  e  $X_t^*$ .

## MÉTRICAS PARA OS MODELOS DE PREVISÃO

Para selecionar o modelo de previsão mais adequado foram utilizadas as métricas de raiz do erro quadrático médio (RMSE) e o erro médio absoluto percentual (MAPE). Seguem as definições abaixo, considerando  $N$  a quantidade de valores previstos,  $\hat{y}_t$  é o valor previsto e  $y_t$  é o valor observado ambos para o instante  $t$ :

A raiz do erro quadrático médio é definido por:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

O erro médio absoluto percentual (MAPE) é definido por:

$$MAPE = 100 \times \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

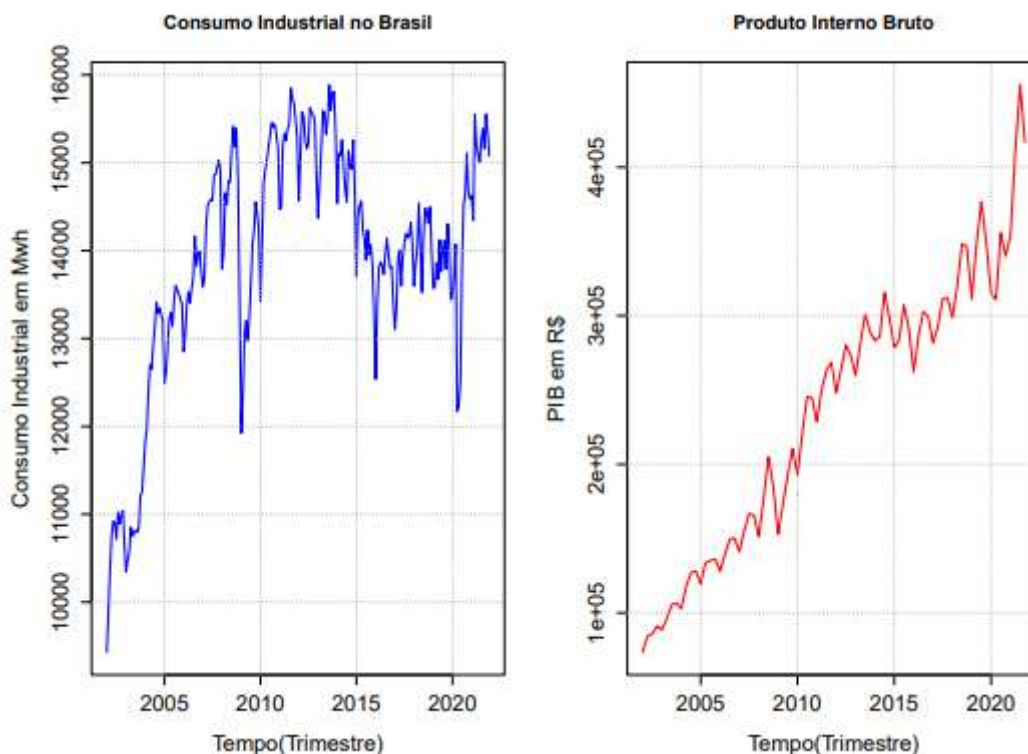
## RESULTADOS E DISCUSSÕES

Nesta seção, primeiramente será apresentada uma análise exploratória para as séries temporais do consumo de energia elétrica e PIB do setor industrial brasileiro. Em seguida, serão apresentados os resultados dos modelos VAR e o de regressão dinâmica.

### Análise Exploratória

A Figura 1 mostra as séries temporais do consumo de energia elétrica e o PIB do setor industrial brasileiro. Pode-se observar nesta figura que ambas as séries temporais possuem tendência e sazonalidade.

Figura 1: Séries Temporais do Consumo de energia elétrica industrial do Brasil e o PIB industrial brasileiro.



A série temporal do consumo de energia elétrica industrial possui bastante mudança de nível, com vales mais significativos em 2008, 2016 e 2020. Em 2008 teve a crise econômica mundial e em 2020 e 2021 a pandemia da COVID-19 e por estes motivos refletiram diretamente no comportamento da série temporal. No PIB também houve uma queda, porém menos acentuada. Nota-se também após a pandemia um crescimento significativo em ambas as séries temporais.

Uma maneira de confirmar a intuição anterior é utilizando a decomposição STL (*Seasonal and Trend Decomposition using Loess*). Essa decomposição divide os dados da série temporal nas componentes de tendência, sazonalidade e resíduo. Através disso, confirma-se a existência de tendência e a sazonalidade para o período trimestral.

Na Tabela 1, pode-se observar as medidas estatísticas para os dados de consumo de energia elétrica da indústria brasileira e para o PIB industrial brasileiro para o período de 2011 a 2021.

**Tabela 1:** Tabela de medidas estatísticas anuais do gasto energético da indústria e do PIB dos últimos dez anos.

	Média	Variância	Desvio padrão	Mínimo	Máximo	Co. Variação
Consumo	43494,45	5303194,35	2302,87	36943	47020	5,29
PIB	308146,31	2131700042,47	46170,34	228636,05	455358,93	14,98

A correlação entre os dados de consumo de energia elétrica da indústria do Brasil e do PIB industrial é 0,63. Com isto, nos dá evidências de que uma série temporal possa interferir e melhorar o modelo no comportamento da outra e podemos aplicar vários modelos multivariados para tentar melhorar a previsão da série temporal endógena.

#### Modelo VAR

Antes de ajustarmos o modelo VAR(4), foi verificado pelos critérios de parcimônia AICc, BIC e AIC qual era o mais baixo. Para a adequabilidade do modelo proposto foi aplicado o teste

de Portmanteau para a correlação serial. Aplicando este teste de correlação serial, apresentou valor-p aproximadamente de 0,05 indicando a ausência de correlação serial a 95% de confiança. Pelo teste da raiz unitária, calculando os autovalores associados a matriz de defasagens, todos os valores deram menores que 1. Logo, o modelo VAR(4) estimado é estável e portanto o modelo é adequado.

A Tabela 2, apresenta os valores das métricas para cada período dos conjuntos de teste. O modelo VAR(4) apresentou uma capacidade preditiva tendo um MAPE e RMSE médio de 4,37 e 2177,31.

**Tabela 2:** Tabela com as métricas MAPE e RMSE para o conjunto de teste do consumo de energia elétrica industrial para o modelo VAR(4).

Teste	2019		2020		2021		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
VAR(4)	1,33	690,60	5,79	3076,67	5,99	2764,68	4,37	2177,31

#### Modelo de Regressão Dinâmica

Na tabela 3 abaixo apresenta as métricas obtidas para o modelo de regressão dinâmica, considerando os diferentes conjuntos de teste. O modelo de regressão utilizou o PIB como variável regressora. Esse modelo apresentou um MAPE e RMSE médio de 4 e 2015,51, respectivamente.

**Tabela 3:** Tabela com as métricas MAPE e RMSE para o conjunto de teste do consumo de energia elétrica industrial considerando o modelo de Regressão Dinâmica.

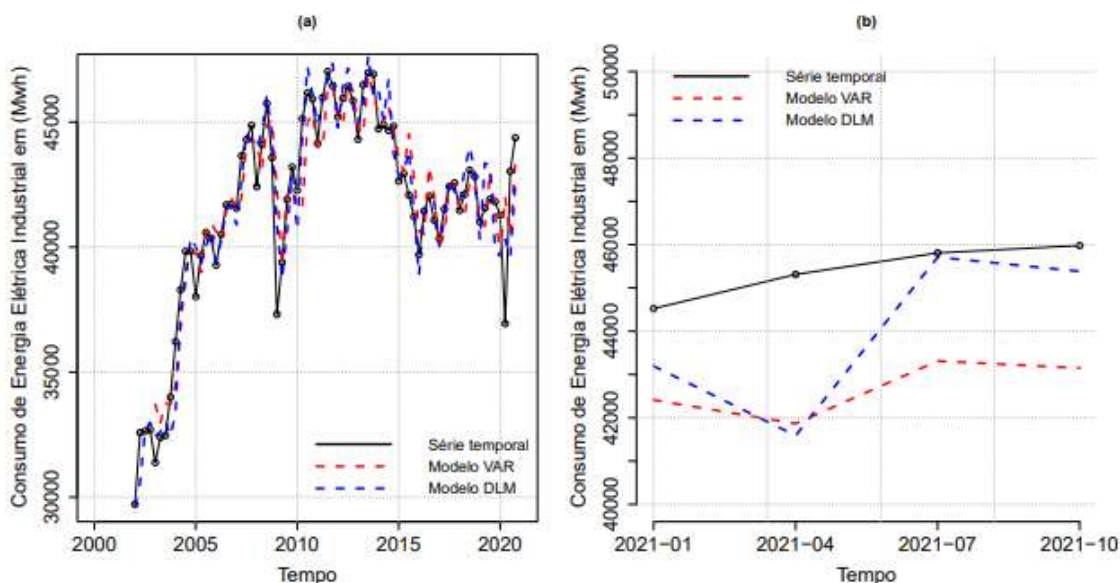
Teste	2019		2020		2021		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
DLM	2,68	1166,31	6,14	2881,89	3,18	1998,36	4	2015,52

## Comparação entre os modelos VAR e de Regressão Dinâmica

A Tabela 4 apresenta as métricas de comparação para os conjuntos de treinamento e de teste para os modelos considerados neste estudo. O modelo de regressão dinâmica apresentou menores valores de MAPE e RMSE tanto para o conjunto de teste e treinamento (Tabela 4). Portanto, o modelo de regressão dinâmica apresentou a melhor capacidade preditiva.

Observa-se na Figura 2(a) que ambos os modelos capturaram bem as variações ao longo da série temporal nos anos de 2000 a 2021, porém, nota-se que no ano de 2020 os modelos não se ajustaram bem no vale abrupto existente. Na Figura 2(b) observa-se a série predita pelo modelo de regressão dinâmica está mais próxima aos dados observados para a previsão trimestral de 2021, corroborando com as métricas apresentadas na Tabela 4.

**Figura 2:** Ajuste (para o conjunto de treinamento) (a) e previsão (para o conjunto de teste) (b) do consumo de energia elétrica do Brasil utilizando os modelos de regressão dinâmica e VAR(4).



## CONCLUSÕES

Neste trabalho, aplicaram-se os modelos de regressão dinâmica e o VAR, a fim de prever o consumo de energia elétrica no Brasil. Os resultados mostraram que estes modelos podem contribuir para obtenção de projeções do consumo de eletricidade no horizonte de curto prazo. Os dados trimestrais analisados foram de janeiro de 2002 até dezembro de 2021, obtidos no sítio eletrônico do IPEA. As métricas de pre-

**Tabela 4:** Tabela com as métricas MAPE e RMSE para os modelos VAR e de regressão dinâmica.

Conjuntos	Treinamento		Teste	
	MAPE	RMSE	MAPE	RMSE
DLM	2,12	1173,57	3,18	1998,36
VAR(4)	2,3	1368,8	5,99	2764,68

visão MAPE e RMSE apontaram que o modelo de regressão dinâmica foi aquele com a melhor capacidade de predição e de ajuste dentre os modelos analisados. Como uma extensão deste trabalho, pretende-se para pesquisas futuras, investigar e comparar os resultados deste trabalho usando os modelos VAR-VEC e de redes neurais para realizar as previsões das séries temporais, investigar comportamentos das séries e verificar a relação contemporânea entre elas.

## REFERÊNCIAS

- AMARO, R. S., CERETTA, P. S., CORONEL, D. A., BENDER, R., Filho, OLIVEIRA, M. G. (2017). Comparação da capacidade preditiva de modelos ARIMA para o consumo de energia elétrica no Brasil, Revista Unemat de Contabilidade, 6(11), 173-192. DOI: <https://doi.org/10.30681/ruc.v6i11.1848>
- BOX, G. E. P., JENKINS, G. M., e REINSEL, G. C. (2015). Time Series Analysis: Forecasting and Control. Prentice-Hall.
- BOX, G. E. P., JENKINS, G. M., e REINSEL, G. C. (1994). Time series analysis: Forecasting and control. 3a ed. New Jersey: Printice Hall.
- BOX, G. E. P., JENKINS, G. M. (1970). Time series analysis: forecasting and control.
- CASTRO N. R., GILIO I., SILVA, A. F., OZAKI, V. G. (2016). Modelos univariados e multivariados aplicados à previsão de valores de exportação: uma análise comparativa para o complexo soja, Revista Espacios. 37(3). Recuperado em 10 abril, 2020, de: <https://www.revistaespacios.com/a16v37n03/16370315.html>
- COCHRANE, D.; ORCUTT, G. H. (1949). "Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms". Journal of the American Statistical Association. 44 (245): 32–61. doi:10.1080/01621459.1949.10483290.
- GUJARATI, D. N., PORTER, D. C. (2011). Econometria Básica. 5a ed. São Paulo, SP: AMGH Editora.
- HAMILTON, J. D. (1994). Time Series Analysis. Princeton University Press.
- IPEA (2022). Ipea data. URL <http://www.ipeadata.gov.br/Default.aspx>.
- LEITE, S. Q. (2006). Projeções para a demanda por energia elétrica no Brasil, 2006-2015. Dissertação de Mestrado Profissionalizante, Faculdades Ibemec, Rio de Janeiro, RJ, Brasil. Recuperado em junho, 2018, de: <http://s3.amazonaws.com/public-cdn.ibmec.br/portalibmeccontent/public/arquivos/di/dis\2006\12\sidimar\quezada\leite.pdf>
- MENDES, R. F. da S., da COSTA, K., da SILVA, F. L. C., COELHO, J. da S. C., VERA-TUDELA, C. A. R., & PINTO, R. V. (2023). Forecasting models for the electricity consumption of the cement industry in Brazil. OBSERVATÓRIO DE LA ECONOMÍA LATINOAMERICANA, 21(7), 6016–6031. <https://doi.org/10.55905/oelv21n7-011>
- MORETTIN, P., A., TOLOI, C. M. C. (2020). Análise de Séries Temporais: Modelos multivariados e não lineares. Editora: Edgard Blucher Ltda, Volume 2.
- MORETTIN, P., A., TOLOI, C. M. C. (2018). Análise de Séries Temporais: Modelos lineares univariados 3a edição. Editora: Edgard Blucher Ltda, Volume 1.
- SILVA, F. C. *et all* (2022). Statistical and Artificial Neural Networks Models for Electricity Consumption Forecasting in the Brazilian Industrial Sector. Energies. 15. 588. 10.3390/en15020588.
- SILVA, F. C. *et all* (2018). A bottom-up methodology for long term electricity consumption forecasting of an industrial sector - Application to pulp and paper sector in Brazil. Energy. 144. 10.1016/j.energy.2017.12.078.
- SILVA, F. C. *et all* (2018). A bottom-up bayesian extension for long term electricity consumption forecasting. Energy. 167. 10.1016/j.energy.2018.10.201.
- HYNDMAN, R.J., & ATHANASOPOULOS, G. (2021) Forecasting: principles and practice, 3rd edição, OTexts: Melbourne, Australia. OTexts.com/fpp3. Acessado em 10/08.

## DADOS DOS AUTORES

**Rafael Santos de Oliveira** ([fael.rso1@gmail.com](mailto:fael.rso1@gmail.com)). Universidade Federal Rural do Rio de Janeiro (UFRRJ)

**Felipe Leite Coelho da Silva** ([felipeleite@gmail.com](mailto:felipeleite@gmail.com)). Universidade Federal Rural do Rio de Janeiro (UFRRJ)

**Josiane da Silva Cordeiro Coelho** ([josicordeiro@gmail.com](mailto:josicordeiro@gmail.com)). Universidade Federal Rural do Rio de Janeiro (UFRRJ)



# UMA ANÁLISE ESTATÍSTICA DO LIMITE DE ESCOAMENTO DA LIGA INCONEL 690 TT A TEMPERATURA AMBIENTE

## A STATISTICAL ANALYSIS ON THE YIELD STRENGTH OF THE INCONEL 690 TT ALLOY AT ROOM TEMPERATURE

Alexandre Santos Francisco

Matheus Felipe Silva Oliveira

**Resumo.** Este artigo apresenta uma análise estatística do limite de escoamento da liga Inconel 690 TT, utilizada em tubos de gerador de vapor. O objetivo é determinar a distribuição de probabilidade que melhor representa os dados de limite de escoamento. Esta análise inclui medidas estatísticas, bem como testes gráficos e de qualidade de ajuste. Os resultados mostram que os dados apresentam uma distribuição simétrica, semelhante a distribuição Normal. O intervalo de confiança para a média amostral do limite de escoamento é calculado com um nível de confiança de 95%. Os resultados demonstram que a média amostral está dentro dos limites especificados por norma, em temperatura ambiente e elevada. Um teste de hipótese confirma que a média amostral pode ser considerada a média populacional. Esta análise contribui para a compreensão da propriedade mecânica da liga sob diversas condições de temperatura e garante o controle de qualidade dos tubos de gerador de vapor.

**Palavras-chave:** Liga Inconel 690 TT. Limite de escoamento. Análise estatística. Tubos de gerador de vapor. Usina nuclear.

**Abstract.** This article presents a statistical analysis on the yield strength of the Inconel 690 TT alloy, used in steam generator tubes. The goal is to determine the probability distribution that best represents the yield strength data. This analysis includes measures of statistical, as well as graphical and goodness of fit tests. The results show that the data exhibits a symmetric distribution, resembling a Normal distribution. The confidence interval for yield strength sample mean is calculated with a confidence level of 95%. The results demonstrate that the sample mean is within the limits specified by norm, at room and elevated temperatures. A hypothesis test confirms that the sample mean can be considered the population mean. This analysis contributes to understanding the mechanical property of the alloy under varying temperature conditions and ensures the quality control of the steam generator tubes.

**Keywords:** Inconel 690 TT alloy. Yield strength. Statistical analysis. Steam generator tubes. Nuclear power plants.

## I - INTRODUCTION

Steam generators are heat exchangers that convert the feedwater into steam using the heat produced by the nuclear reactor. They are used in nuclear power plants with pressurized water reactor (PWR), the most common type of nuclear power plant in the world. In these plants, there are two different water circuits, the primary and the secondary. The primary circuit contains the pressurized water that circulates through the reactor core, where it is heated due to the nuclear fission of uranium atoms. The heated radioactive water is taken to the steam generator, where it transfers its heat to the secondary circuit, without having direct contact with it. The secondary circuit contains pure water that is vaporized due to the heat received from the primary circuit. The steam is directed to the main turbine, where it makes the electric generator work, producing electricity. The extra steam is condensed in the condenser, where it loses its heat to the refrigerant water. The condensed water is pumped back to the steam generator, to restart the cycle (NRC, 2018).

The steam generator tubes are components that allow the heat exchange between the two water circuits, without mixing them. They are made from special metal alloys that should resist high temperatures, high pressures and corrosion involved with the process. The steam generator tubes are essential for safety and efficiency of PWR nuclear power plants. They prevent the radioactive contamination of the secondary circuit and the external environment, besides ensuring the quality of the produced steam. The steam generator tubes should be inspected and monitored periodically, to detect possible failures or faults that may interfere with their performance (NRC, 2018).

According to ASME SB-163 (ASME, 2022), ASME III NB-2000 (ASME, 2000) and ASME Code Case N-20-3 (ASME, 1998), the Inconel 690 TT should be used in the steam generator tubes. The Inconel 690 TT is an iron-nickel-chromium alloy with high resistance to corrosion and creep, which is used in the Angra 1 nuclear power plant steam generator tubes. The yield strength of this alloy is a very important mechanical property to guarantee the performance under operation conditions. In this article, a statistical analysis of the mechanical data on yield strength of the Inconel 690 TT alloy was conducted, obtained from tensile tests. Additionally, a statistical analysis of the chemical composition of the Inconel 690 TT alloy was also conducted to better understand its properties and behavior.

To do this analysis, the measurements of statistical location and dispersion of the mecha-

nical data were calculated, such as the mean, median, standard deviation, variance, mode and the coefficient of variation (LEWIS, 1996). After that, the graphical method was used to compare the probability density histograms of the mechanical data with the curves of the Normal, Lognormal, Weibull and Exponential probability functions. This method allows a visual analysis of the probability function fit to the mechanical data but doesn't offer any quantitative criteria to choose the best distribution (LEWIS, 1996). Due to that, the Chi-squared, Anderson-Darling and Kolmogorov-Smirnov goodness of fit tests were applied, that evaluate the null hypothesis that the mechanical data follows a certain distribution based on the measures of the distance between the empirical and theoretical cumulative distribution functions (MONTGOMERY, 2002).

After the definition of the probability distribution that better represents the mechanical data, the confidence interval for the yield strength mean was calculated, with confidence level of 95%. The confidence interval represents the interval in which the population mean is expected to be in, based on the sample mean and on the variability of the mechanical data. Besides that, a box plot (graphic that shows the median, quartiles and the extreme values of the data) was made to identify possible outliers, values that deviate from the central tendency of the mechanical data (MONTGOMERY, 2002).

The statistical analysis realized in this article contributes to the knowledge of mechanical properties of the Inconel 690 TT alloy and for the quality control of the steam generator tubes of the Angra 1 nuclear power plant.

## II - METHODOLOGY

The goal of this analysis was to identify the probability distribution that better represents the mechanical data, as well as to calculate the confidence interval for the mean and to detect possible outliers. Once the best probability function is found, it will be used for the reliability analysis of the steam generator tubes.

This paper contains statistical data from the steam generator tubes yield strength. The statistical information includes the mean ( $\mu$ ), the standard deviation ( $\sigma$ ), the coefficient of variation ( $COV$ ) and the probability distribution that better represents the mechanical properties of the Inconel 690 TT alloy. Also, other important statistical information will be provided, such as the median, the bias and others.

According to ASME SB-163, the Inconel 690 TT alloy should have a minimum value of yield strength of 276 MPa, at room temperature (24 °C), and a minimum value of yield strength of 215 MPa at elevated temperature (350 °C). Additionally, a statistical analysis of the chemical composition of the Inconel 690 TT alloy was also conducted to better understand its properties and behavior. The mechanical data provided contain 102 measurements of yield strength at room and elevated temperatures obtained from a tensile test (ASME, 2000).

The first step to start the statistical analysis of these mechanical data is to build a frequency histogram that enables analyzing the distribution of the yield strength values obtained from the tensile test. To build the frequency histogram, a frequency table must be built, and for that, the class interval, or bins, may be defined by (LEWIS, 1996)

$$\Delta S = r[1 + 3,3\log_{10}(N)^{-1}] \quad (1)$$

where  $rr$  is the range (difference between the maximum and minimum values of the yield strength data at room temperature, and  $NN$  is the Number of the yield strength data points.

The next step after building the frequency histogram is to use the graphical method to define the probability distribution that better fits the mechanical data. This method consists of comparing the histogram with the curves of the theoretical probability density functions of each distribution. In this article, the Normal, Lognormal, Weibull and Exponential probability distributions will be used. The distribution that presents the best visual fit with the frequency histogram can be chosen to represent the mechanical data (MONTGOMERY, 2002).

After performing the graphical analysis, goodness of fit tests was also performed to confirm the graphical method results and to obtain more accurate results. The goodness of fit test consists of comparing the mechanical data obtained from the tensile test with the expected data according to the chosen probability function. As previously said, the Normal, Lognormal, Weibull and Exponential probability distributions will be used in this test. The goodness of fit test provides a numerical value of how well the data fits the probability distribution, as well as the confidence level that indicates the probability of rejecting the null hypothesis that the mechanical data fits a certain distribution. In this article, the Chi-squared, the Anderson-Darling and the Kolmogorov-Smirnov goodness of fit tests will be used to verify which distribution better fits the mechanical data (D'AGOSTINO; STEPHENS, 1986).

Once the graphical method and the goodness of fit tests are done, and the probability distribution that better fits the mechanical data is chosen, a confidence interval test for the mean is needed to ensure that all the calculations that were made previously are valid, and that all calculations that will be made in the future, involving the reliability analysis, will be using the right values of the mechanical properties of the Inconel 690 TT alloy. A level of confidence of 95% will be used for the confidence interval test for the mean (MONTGOMERY, 2002).

### III – STATISTICAL ANALYSIS OF THE MECHANICAL DATA

In this section, the results of the statistical characterization of the yield strength of the Inconel 690 TT alloy will be discussed. To characterize this data distribution, measurements of statistical location and dispersion must be calculated based on the data obtained from the tensile test. A box plot will also be plotted to show the maximum and minimum values of the mechanical data, as well as possible outliers. The measurements of statistical location calculated include the sample size ( $N$ ), the sample mean ( $MM$ ), the median ( $MdMd$ ) and the mode ( $MoMo$ ). The measurements of statistical dispersion calculated include the sample standard deviation ( $SS$ ), the sample variance ( $S^2S^2$ ) and the coefficient of variation ( $COV$ ) (SEKULSKI, 2019). Uncertainty measurements as the difference bias and ratio bias were calculated as well. The results of the calculations are shown in the Table 1.

Table 1. Statistical data of the yield strength at room and elevated temperatures.

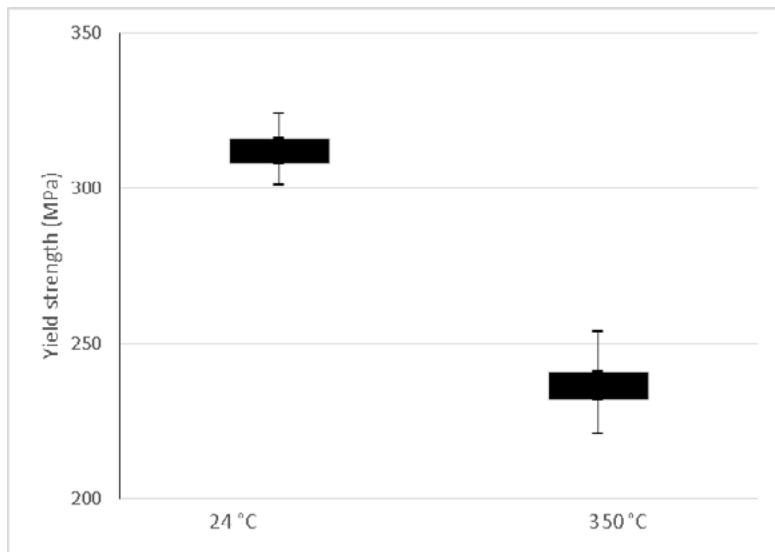
Temperature:	24 °C	350 °C
Sample size:	102	102
Sample mean:	312.58 MPa	236.56 MPa
Median:	313.00 MPa	236.00 MPa
Mode:	310.00 MPa	236.00 MPa
Standard deviation:	5.52 MPa	6.63 MPa
Coefficient of variation:	1.77 %	2.80 %
Variance:	30.44 MPa <sup>2</sup>	43.99 MPa <sup>2</sup>
(average) difference bias:	11.58 MPa	15.56 MPa
(average) ratio bias:	1.04	1.07

Source: Authors (2023)

The results infer that the mechanical data has a symmetric distribution, with sample mean, median and mode with close values, which is characteristic of a Normal distribution (LEWIS, 1996). The value of the sample mean at both room and elevated temperature are also above the minimum

specified by ASME-SB-163. Both standard deviation and coefficient of variation got low values, which indicate the low variability of the mechanical data. The box plot in the Figure 1 doesn't show any outlier, that means that there are no values that differ a lot from the sample mean.

Figure 1. Yield strength box plot at room and elevated temperatures.



Source: Authors (2023)

These results suggest that all the mechanical data obtained from the tensile test of the Inconel 690 TT alloy were consistent and representative. They can be used for further analysis.

Once the data characterization of the Inconel 690 TT mechanical data was done, a frequency table was built and presented as the Table 2. The frequency table must cover all the data range, from the minimum to the maximum value.

Table 2. Frequency table for yield strength at room and elevated temperatures.

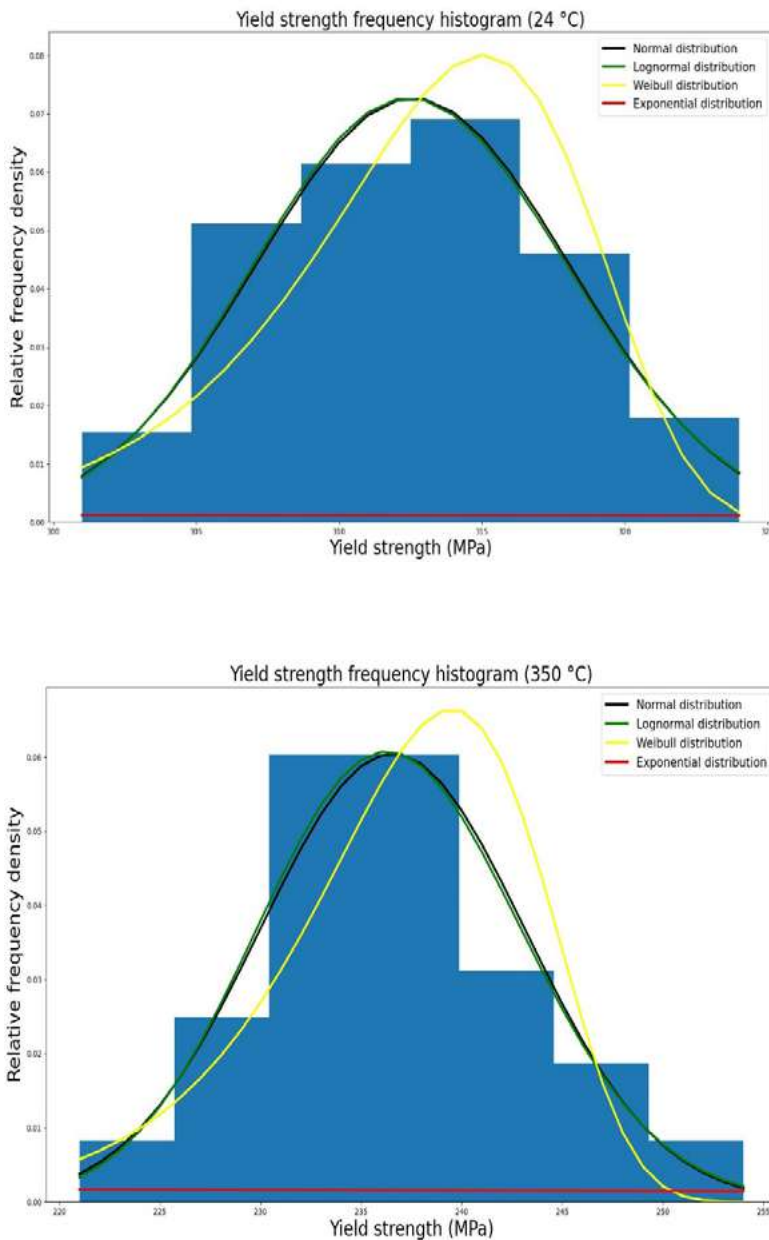
24 °C			
Bins	Frequency	Relative freq.	Relative freq. dens.
[301, 304[	6	0.0588	0.0147
[305, 308[	20	0.1961	0.0490
[309, 312[	24	0.2353	0.0588
[313, 316[	27	0.2647	0.0662
[317, 320[	18	0.1765	0.0441
[321, 324[	7	0.0686	0.0172
350 °C			
Bins	Frequency	Relative freq.	Relative freq. dens.
[221, 225[	4	0.0392	0.0098
[226, 230[	12	0.1176	0.0294
[231, 235[	29	0.2843	0.0711
[236, 240[	30	0.2941	0.0735
[241, 250[	15	0.1471	0.0368
[246, 250[	8	0.0784	0.0196
[251, 255[	4	0.0392	0.0098

Source: Authors (2023)

The frequency table was built to assist in the creation of a frequency histogram in the Figure 2 to compare the mechanical data with the Normal, Lognormal, Weibull and Exponential distributions. The graphical method plots the theoretical distribution curves over the frequency histogram of the actual data and analyzes which theoretical distribution has the best fit to the shape of the mechanical data frequency histogram. The results presented in Figure 2 show that the Normal and Lognormal distributions have a good fit to the sample data, followed by the Weibull distribution. The Weibull distribution has many characteristics in the reliability design (SUNGHO *et al.*, 2010). The exponential distribution has not fit for the data.

Due to the proximity of Normal and Lognormal distribution curves using the graphical method, it wasn't possible to determine which of these distributions have a better fit to the mechanical data of the Inconel 690 TT alloy. The results were inconclusive because the graphical method is based in a visual and subjective analysis of the fitting. A goodness of fit test was necessary to get more precise and conclusive results. The goodness of fit tests are statistical methods that allow to compare the probability distributions and choose the one that better fit the data according to the null hypothesis (MONTGOMERY, 2002).

Figure 2. Frequency histogram at room and elevated temperatures.



Source: Authors (2023)

The Chi-squared, Kolmogorov-Smirnov and Anderson-Darling goodness of fit tests were performed to test which of the distributions has a better fit to the sample data of the mechanical properties of the Inconel 690 TT alloy. These tests are based on the difference between the empirical distribution of the sample data (test statistic) and the theoretical distribution function tested (critical value). The tests return a p-value, that indicates the probability of the data to fit in a certain distribution assumed by the null hypothesis. The higher the value of p-value, the higher is the evidence favoring the null hypothesis. The results of the tests show that the Normal distribution presents higher values in the Chi-squared and Anderson-Darling tests. The difference between the results of the Normal and Lognormal distribution in the Kolmogorov-Smirnov tests is negligible as shown in Table 3 (D'AGOSTINO; STEPHENS, 1986).

Table 3. Goodness of fit test results for yield strength at room and elevated temperatures.

24 °C				350 °C		
Chi-squared						
	$\chi_0^2$	$\chi_{crit}^2$	p-value (%)	$\chi_0^2$	$\chi_{crit}^2$	p-value (%)
Exponential	55.56	9.49	0.00	32.09	11.07	0.00%
Normal	0.02	7.81	98.40	0.12	9.49	90.44%
Lognormal	0.16	7.81	87.28	0.11	9.49	91.24%
Weibull	0.184	7.81	85.40	0.113	9.49	91.00%
Kolmogorov-Smirnov						
	$D_n$	$D_{crit}$	p-value (%)	$D_n$	$D_{crit}$	p-value (%)
Exponential	0.608	0.135	54.30	0.599	0.135	54.92%
Normal	0.072	0.135	94.26	0.121	0.135	90.37%
Lognormal	0.061	0.135	95.14	0.115	0.135	90.84%
Weibull	0.129	0.135	89.74	0.185	0.135	0.00%
Anderson-Darling						
	$A^*$	$A_{crit}$	p-value (%)	$A^*$	$A_{crit}$	p-value (%)
Exponential	45.517	0.752	0.00	44.654	0.752	0.00%
Normal	0.415	0.752	67.81	0.903	0.752	36.65%
Lognormal	1.85	0.752	6.43	0.795	0.752	42.66%
Weibull	23.06	0.752	0.00	14.06	0.752	0.00%

Source: Authors (2023)

Once the distribution that better fits the data is chosen, a confidence interval test for the mean was performed. This test was done to determine the true value of the population mean with a certain confidence level, based on the sample mean and standard deviation. The confidence interval provides a range of values for the yield strength, that contains the true value of the mean. The test is very useful to assess how precise the goodness of fit tests was and to analyze the variability of the data, to see how far the mean from the sample mean be.

The minimum yield strength values were already established by ASME-SB-163 in the results of the tensile test. For these values, some statistics were so low, to the point that they couldn't even be found in a standard Normal distribution table. Three confidence intervals for the mean were performed: 90%, 95% and 99% confidence levels. They are shown in Table 4.

Table 4. Confidence interval for mean at room and elevated temperatures.

Confidence level	Confidence interval (MPa)	
	24 °C	350 °C
90%	311.68 < $\mu$ < 313.48	235.47 < $\mu$ < 237.64
95%	311.51 < $\mu$ < 313.65	235.27 < $\mu$ < 237.85
99%	311.17 < $\mu$ < 313.99	234.86 < $\mu$ < 238.26

Source: Authors (2023)

The results of the confidence interval test for the mean got a very low range in every confidence level tested. This is evidence that the yield strength mean is according to the limits established by ASME-SB-163, which got higher ranges for the yield strength to be worked.

A lot of problems involving engineering may require accepting or rejecting a hypothesis about some statistical parameter. In this article, a hypothesis test for the mean was performed. Through this test, it is possible to check how strong is the evidence that the yield strength mean is according to the limits established by ASME-SB-163.

It was stated that the yield strength data follows a Normal distribution, with a test statistic  $Z_0$ . The resulting probability is compared to a pre-defined significance level (usually  $\alpha$  is equal to 0.05) to determine whether the null hypothesis should be rejected or not. The observed value of  $Z_0$  is compared to tabulated values  $Z_{crit}$  that consider the significance level. This allows engineers to make informed decisions about whether the material meets the requirements set by standards such as ASME-SB-163.

Two hypotheses, a null ( $H_0$ ) and an alternative ( $H_1$ ), are established regarding the mean of the yield strength at room temperature. The goal is to determine if these values are above a minimum value set by technical standards. The hypotheses are formulated by

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases} \quad (2)$$

where  $\mu_0$  is the supposed true value of the yield strength mean.

The results shown in Table 5 show that  $Z_0$  values are greater than  $Z_{crit}$  is equal to 1.645, so the null hypothesis must be rejected. Therefore, there is strong statistical evidence to assert that the mean of the yield strength of the Inconel 690 TT alloy at room and elevated temperatures will be greater than the minimum values established by ASME-SB-163.

Table 5. Result of the hypothesis test for the mean at room and elevated temperatures.

Temperature	24 °C		350 °C	
Parameter	$Z_0$	$Z_{crit}$	$Z_0$	$Z_{crit}$
Yield strength mean	66.92	1.645	32.84	1.645

Source: Authors (2023)

Once the characterization and statistical analysis of the mechanical property data of the Inconel 690 TT alloy at room and elevated temperatures were completed, a statistical analysis of the chemical composition of the Inconel 690 TT alloy was also conducted to ensure that its composition is in accordance with the standards.

The outcome of the chemical composition analysis, conducted using the spectrometry technique on the Inconel 690 TT alloy, is presented in Table 6. It can be observed that the average compositions of the constituent chemical elements of the alloy are satisfactory, in accordance with the limits set by the International Atomic Energy Agency (IAEA) (IAEA, 2011).

The comprehensive analysis of the chemical composition of the Inconel 690 TT alloy shows that it is evident that the alloy's constituents align with the standards set by the IAEA. The spectrometry technique has proven effective in accurately determining the composition of the alloy.

Table 6. Inconel 690 TT alloy chemical composition.

Element	Composition (%)	
	Mean	Standard deviation
Niquel	59.2815	0.2525
Chromium	29.3377	0.1291
Iron	10.33	1.02743
Carbon	0.021	0.00117
Silicon	0.1982	0.01926
Manganese	0.2786	0.01592
Cobalt	0.0069	0.00071
Copper	0.01	6.9E-18
Sulfur	0.0017	0.00051
Phosphor	0.0068	0.00069
Nitrogen	0.0289	0.00271
Aluminum	0.0152	0.00268
Boro	0.0004	2.7E-19
Titanium	0.2871	0.02109
Molybdenum	0.0101	0.00099
Niobium+ Tantalum	0.0004	2.7E-19

Source: Authors (2023)

#### IV - CONCLUSIONS

The statistical analysis conducted in this study provides valuable insights into the mechanical properties of the Inconel 690 TT alloy used in the steam generator tubes of the Angra 1 nuclear power plant. The analysis determined that the mechanical data exhibits a symmetric distribution, with the sample mean, median, and mode close in value, suggesting a Normal distribution. The standard deviation and coefficient of variation indicated low variability in the mechanical data.

Through graphical and goodness of fit tests, it was determined that the Normal and Lognormal distributions best fit the mechanical data. However, the confidence interval test for the mean revealed that the sample mean falls within the limits specified by ASME-SB-163, supporting the use of the Normal distribution for further analysis.

Furthermore, the hypothesis test for the mean confirmed that the sample mean can be considered as the population mean, validating all calculations and tests performed in this study. This statistical analysis contributes to the understanding of the Inconel 690 TT alloy's mechanical properties and quality control of the steam generator tubes in the nuclear power plant.

In addition, the chemical composition of the Inconel 690 TT alloy adheres to the standards set by the IAEA. This ensures the reliability of the alloy's composition and its compliance with the required norms.

By ensuring the reliability and safety of the steam generator tubes, the findings of this analysis can aid in maintaining the efficiency and effectiveness of PWR nuclear power plants, promoting the continued use of nuclear energy as a clean and sustainable energy source. The adoption of Inconel 690 TT alloy in steam generator tubes demonstrates its suitability for with-

tanding the harsh conditions of the nuclear power plant environment, ensuring the prevention of radioactive contamination, and enhancing the safety of the plant and the surrounding environment.

## ACKNOWLEDGEMENTS

This study was financed in part by Eletrobrás Termonuclear S/A through the scholarship for Matheus Felipe Silva Oliveira.

## REFERENCES

- ASME - American Society of Mechanical Engineers. **ASME Code Case N-20-3**: Alternative rules for qualification of welders and welding procedures for welding nickel alloy stainless steels. New York: ASME, 1998.
- ASME - American Society of Mechanical Engineers. **ASME III NB-2000**: Rules for construction of nuclear facility components. New York: ASME, 2000.
- ASME - American Society of Mechanical Engineers. **ASME SB-163**: Specification for seamless nickel and nickel alloy condenser and heat-exchanger tubes. New York: ASME, 2022.
- D'AGOSTINO, R. B.; STEPHENS, M. A. **Goodness-of-fit techniques**. New York: Marcel Dekker Inc, 1986.
- IAEA - International Atomic Energy Agency. **IAEA-TECDOC-1668**: Report on the Assessment and management of ageing of major nuclear power plant components important to safety steam generators. Vienna: IAEA, 2011.
- LEWIS, E. E. **Introduction to Reliability Engineering**. 2ª edição, New York: John Wiley Sons, Inc, 1996.
- MONTGOMERY, R. **Applied Statistics and Probability for Engineers**. New York: John Wiley Sons, Inc., 2002.
- NCR - U.S. Nuclear Regulatory Commission. **Pressurized Water Reactor (PWR) Systems**. Washington: NCR, 2018. Available on: <https://www.nrc.gov/reading-rm/basic-ref/students/for-educators/04.pdf>. Accessed at: 05 jan. 2024.
- SEKULSKI, Z. Statistical Properties of the Yield Strength of Normal Strength Hull Structural Steel Plates. **Annals of Dunarea de Jos**. Shipbuilding: University of Galati, 2019. Fascicle XI.
- SUNGHO, P. *et al.* A statistical study on tensile characteristics of stainless steel at elevated temperatures. **Journal of Physics.: Conference Series**, Bristol, v. 240, Article ID 012083, 2010. DOI 10.1088/1742-6596/240/1/012083. Available on: <https://iopscience.iop.org/reading-rm/article/10.1088/1742-6596/240/1/012083>. Accessed at: 05 jan. 2024.

## AUTORES

**Alexandre Santos Francisco** (afrancisco@id.uff.br). Doutor em Engenharia Nuclear; Docente, Departamento de Engenharia Mecânica de Volta Redonda, UFF.

**Matheus Felipe Silva Oliveira** (matheusfsoliveira@outlook.com). Mestrando em Modelagem Computacional, Pesquisador, Programa de Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia, UFF.



## DESENVOLVIMENTO DE SOFTWARE PARA DIMENSIONAMENTO E SELEÇÃO DE ESTRUTURAS DE EDIFÍCIO EM CONCRETO ARMADO COM REDES NEURAIIS ARTIFICIAIS OTIMIZADAS E AHP

## SOFTWARE DEVELOPMENT FOR DIMENSIONING AND SELECTION OF REINFORCED CONCRETE BUILDING STRUCTURES WITH OPTIMIZED ARTIFICIAL NEURAL NETWORKS AND AHP

Adriano Ricardo Borges Costa

Eliane da Silva Christo

Tiago Araújo Neves

**Resumo:** A Engenharia evoluiu com a Computação, e algoritmos de aprendizado de máquina como as Redes Neurais Artificiais (RNAs) vêm sendo aplicadas. As RNAs ajudam na resolução de problemas complexos. Os projetos de edifícios estão ficando mais complexos e dimensionar e orçar opções de projetos, mesmo com computação convencional, é tarefa complexa, logo, a Engenharia é ideal para aplicações de RNAs. Geradas as opções focando “resistência x suficiência”, selecionou-se os melhores projetos “qualidade x menor custo” para o investidor. Surgiu então este trabalho, gerar um protótipo para dimensionamento de lajes via RNA e classificação por projeto de melhor ganho no orçamento via AHP. Os resultados foram promissores atendendo aos objetivos e essa combinação pode ser uma opção na engenharia de edifícios.

**Palavras-chave:** RNA. AHP. Lajes. Concreto Armado.

**Abstract:** Engineering evaluated together Computing, and machine learning algorithms as Artificial Neural Nets (ANN) have been used. ANN helps solving complex problems. Building projects are becoming more complex (design and rules). Dimensioning and costing projects options, since by conventional computing, are complex tasks, so, Engineering is an ideal area to apply ANNs. Generated alternatives pointing to “strength x sufficiency”, select best projects on “quality x low costs” to investor. Many revisions tighten engineers’ prizes until contracting and building. Then, raised this work idea, generating a prototype to concrete slabs dimensioning by ANN and classify the best gain budget project slabs by AHP. Results has potential and compliance objectives, so, this combination can be another good option to civil engineering.

**Palavras-chave:** RNA. AHP. Slabs. Reinforced Concrete.

## INTRODUÇÃO

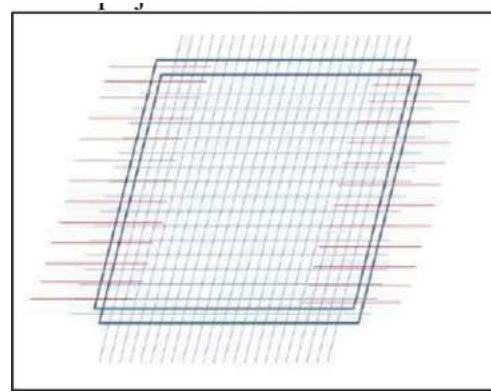
Na Engenharia de Edifícios é necessário criar vários projetos, e dimensioná-los demanda muitos cálculos e análises, uma tarefa complexa e repetitiva (BOTELHO, 2013). Estes têm muitos elementos a dimensionar e orçar, como as lajes, com armaduras principais (+) e (-) e suas secundárias, por pavimentos, e sem falar nas vigas e pilares (UBERABA, 2009). As redes neurais já vêm sendo usada nas engenharias, e surgiu a ideia de aplicar RNAs tipo perceptron multicamada para dimensionar lajes de edifícios combinado com a técnica AHP para classificar os orçamentos destes projetos. Com a topologia, quantidade certa de neurônios, parametrização, funções apropriadas e de aceleradores de aprendizado corretos, a RNA gera uma função-solução em sua estrutura (KOVÁCS, 1996). Vários pré-treinamentos são feitos para se chegar à melhor topologia e parâmetros, (SILVA, 2010) e este trabalho, além de produzir um protótipo, também testou na prática o potencial da combinação RNA+AHP, tendo gerado resultados promissores.

## MATERIAIS E METODOLOGIA

### Dimensionamento e Orçamento de Estrutura de Edifícios

Os engenheiros projetam edifícios para investidores, baseados no dimensionamento e orçamento da estrutura, acabamento e lajes (Fig.1). Os requisitos são: Investimento; Topografia; Sondagem do Terreno; Cartas ambientais. Logo após, são geradas as opções de plantas baixas de pré-projeto (estrutura e design) e os orçamentos (Fig.2) para aprova do cliente. A complexidade de nº de revisões dependerá dos projetos.

Fig.1: Ex. (simplif.) Arms. Laje



Fonte: Autor

Fig.2: Ex. Planta Baixa (simplif.) Projeto

Planta Baixa do Projeto 6		4,0	10,0	13,0	18,0	21,0	27,0	
4,0	6,0x4,0 Quarto	Copa	3,0x4,0 Sala	5,0x4,5 A.Serv	Swite	3,0x4,0 Sala	6,0x4,0 Copa	Quarto
	Coz.			Coz.	WC		Coz.	
8,0	6,0x4,5 Swite	Coz	3,0x4,5 Escada / Hall	5,0x4,5 Sala	Copa	3,0x4,5 Escada / Hall	6,0x4,5 Coz.	Swite
	Wc	A.Serv	Vent			Vent	A.Serv	Wc
12,5	6,0x4,5 Copa	Sala	3,0x4,5 Hall /Elevador	5,0x4,5 Copa	Sala	3,0x4,5 Hall /Elevador	6,0x4,5 Sala	Copa
	Wc			WC			Wc	
17,0	6,0x4,0 Quarto	Coz.	3,0x4,0 A.Serv	5,0x4,0 Swite	Coz.	3,0x4,0 A.Serv	6,0x4,0 Coz.	Quarto
				WC	A.Serv.			
21,0								

Fonte: Autor

### Redes Neurais Artificiais - Estrutura e Funcionamento

A RNA PCS (1 neurônio) é capaz de separar 2 classes contínuas, mas não resolve problemas mais complexos, (LUDWIG, 2007). Já a PMC tem uma arquitetura complexa de neurônios, com topologia escalável conforme o problema, (BRAGA, 2000), como sempre 3+ camadas (1 Entr., Intermediária(s), 1 Saída), (AZEVEDO, 2000). Sua construção e processamento são os seguintes:

#### Etapa 0 do Treinamento - Carga e Parametrização

1. Faz-se uma análise prévia dos atributos com alta correlação com a sistemática do problema, e RNA terá melhores chances de gerar a função-objetivo. Os grupos de atributos formam amostra, que terão entradas  $\mathbf{x}_{\alpha, t}^{(k)}$ , passam por pré-processamento ajustando-as ao domínio da função de ativação, (Silva, 2010), ex.:  $\mathbf{D}_R = (5, \dots, 55, \dots, 105) \rightarrow \mathbf{D}_{\text{Tanh}} = (-1, \dots, 0, \dots, 1)$ .

2. **Configuração da Execução e Pré-Processamento:** A RNA tem vários parâmetros de processamento (Nº iterações, Topologia, Aceleradores  $\eta$  e  $\mu$ , EQM Admissível, Amostras) que são configurados. Cada amostra  $\alpha$  de entradas  $\mathbf{x}_{\alpha, t}^{(k)}$ , onde  $\{i = (1, n) \mid n \in R\}$ , são injetadas na RNA pela camada de entrada  $\mathbf{k} = \mathbf{0}$ , a cada iteração  $t$  até a última amostra  $p$ , via um dos métodos de injeção. A experiência mostra que a injeção sequencial de amostras  $\alpha$  faz a RNA ser tendenciosa e divergir da função-objetivo. *Os métodos:* (aleatório, particionado,

k-means), sendo aleatório o menos tendencioso, (MANZAN, 2017), (SILVA, 2010).

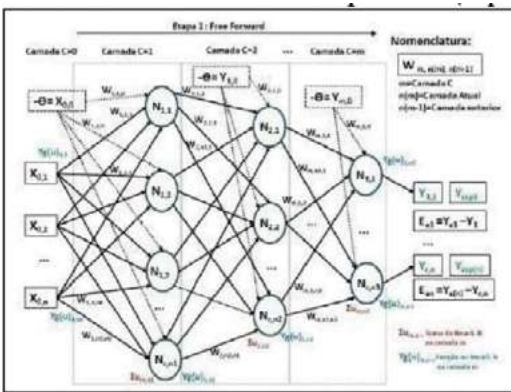
Além das entradas  $x_{\alpha,i}^{(k)}$ , também são injetadas sua(s) saída(s) esperada(s)  $y_{\alpha,j}^{(k=m)}$ . Estas foram geradas convencionalmente previamente com as entradas, (MANZAN, 2017). As saídas  $y_{\alpha,j}^{(k=m)}$  para a camada  $m$  também passaram por *pré-processamento*.

**3. Vetor de Pesos  $w$ :** O Algoritmo também gera uma matriz de pesos  $w_{i,c}^{(k)}$  onde  $i = (1, j)$ , com  $j$  neurônios e  $c = (1, i)$  com  $(i = \epsilon)$  entradas. Os vetores  $w_{i,c}^{(k)}$  além de se conectarem com todas as  $x_{\alpha,i}^{(k-1)}$  da 1ª camada, também se conectam a todos os  $j$  neurônios na camada  $k$ . Seus valores iniciais são aleatórios, antes e só na iteração  $(t = 0)$ , dentro da faixa de domínio da função de ativação usada na(s) camada(s) intermediária(s), conforme Silva (2010) e Manzan (2017)

**Etapa 1 - Treinamento - Freeforward (Fig.3):**

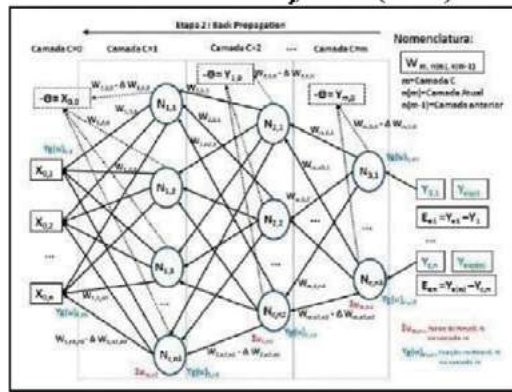
**4.** Cada neurônio possui 4 processadores. O 1º é o *Combinador* (Fig.3), que calcula o somatório dos produtos das entradas  $x$  com os pesos  $w$  (Eq.1), entrantes em cada neurônio  $j$ , Silva (2010).

Fig.3: Etapa 1 - Freeforward (Arranque a Frente)



Design: Autor - Fonte: Silva

Fig.4: Etapa 2 - Retro-Propagação (Reverso).



Design: Autor - Fonte: Silva

$$u_{\alpha,j}^{(k)} = \sum_{i=1}^n (w_{ji}^{(k)} \cdot x_{\alpha,i}^{(k-1)} - \theta) \quad ; \quad y_{\alpha,j}^{(k)} = g(u_{\alpha,j}^{(k)}) \quad \text{com } u \text{ e } y \in \mathbb{R} \quad (1) (2)$$

Sendo:

$w_{j,i}^{(k)}$  = peso neurônio  $j$  da camada  $k$  conectado na entrada  $i$  da camada  $k - 1$  ;

$x_{\alpha,i}^{(k-1)}$  = entrada  $i$  na camada  $k - 1$  da amostra  $\alpha$  ;  $u_{\alpha,i}^{(k)}$  = Função somador dos  $(w \cdot x)$  ;

$y_{\alpha,j}^{(k)}$  = saída calculada do neurônio  $j$  camada  $m$  ;  $g(u_{\alpha,j}^{(k)})$  = função de ativação ;

$\theta$  = Bias ;

Obs.: O termo  $\theta$  (Bias), termo de translação da função objetivo, e pode se incluso como  $x_{\alpha,0}^{(k-1)}$

**5.** Logo em seguida, o 2º Processador (Eq.2) (Fig.3) *Aplicador da Função de Ativação* o sobre o produto  $u_{\alpha,j}^{(k)}$  gera as saídas  $y_{\alpha,j}^{(k)}$  em cada neurônio da camada  $k$ . Estas se tornam as entradas  $(x_{\alpha,i}^{(k)})$  para a próxima camada  $(k + 1$  de neurônios  $j$ , até a última,  $n$  as saídas calculadas  $y_{\alpha,j}^{(k=m)}$ , (SILVA, 2010) e (LUDWIG, 2007).

**Etapa 2 - Treinamento - Back Propagation (Fig.4):**

6. Após o Freeforward, é lógico que as saídas  $y_{\alpha,j}^{(k=n)}$ , da camada de saída  $k = m$  não coincidam com as esperadas  $y_{\alpha,j}^{(k=n)}$ , requerendo reajuste de todas  $y_{\alpha,j}^{(k)}$ . Passa-se para uma 2ª parte, a Back Propagation (Fig.4), reverso do Freeforward. O 3º Processador, (Avaliador) de erro do neurônio  $j$  camada ( $k = n$ ) calcula o erro  $\delta_j^{(k=n)}$  (Eq.3) entre as saídas  $y_{\alpha,j}^{(m)}$  e  $y_{\alpha,j}^{(m)}$ , (LUDWIG, 2007).

$$\delta_{\alpha,j}^{(k)} = (y_{\alpha,j}^{(m)} - y_{\alpha,j}^{(k)}) \cdot g'(I_{\alpha,j}^{(k)}) \quad ; \quad w_{j,i}^{(k)}(t+1) = w_{j,i}^{(k)}(t) + \mu \cdot \Delta w + \eta \cdot \delta_{\alpha,j}^{(k)} \cdot y_{\alpha,i}^{(k-1)}$$

**Sendo:**  $w_{j,i}^{(k)}(t+1) = peso$ ;  $\delta_j^{(k)}$  = ErroDelta camada  $k$ ;  $g'(I_{\alpha,j}^{(k)})$  = derivada de  $y_{\alpha,j}^{(k)}$  relação a  $u_{\alpha,j}^{(k)}$ ;  $\mu$  = (propulsor de avanço) e  $\eta$  = fator de aprendizagem, baseado na teoria do da menor distancia ao mínimo absoluto onde a convergência ( $y_e - y_c$ ) de cada amostra é um mínimo e absoluto.

7. O 4º Processador (**Difusor de Erros**) (Fig.4) reajusta os pesos  $w_{j,i}^{(k)}$  (Eq.4) neurônio / neurônio  $j$ , camada  $k$  em camada ( $k - 1$ ) (Eqs.3 e 4), ajustando proporcionalmente de ( $\Delta w$ ) os  $w_{j,i}^{(k)}$  as parcelas  $(\mu \cdot \Delta w)$  e  $\eta$  do **gradiente** ( $\delta_{\alpha,j}^{(k)} \cdot y_{\alpha,i}^{(k-1)}$ ), entre iterações  $t$  (Eq.4), (SILVA, 2010).

8. As etapas Freeforward e Back Propagation são encapsuladas num loop de repetição, (processamento em 2 etapas). Este se repetirá até que um método de parada ocorra.

**Calcula-se o Erro Médio EQM:** A cada iteração  $t$ , o erro quadrático médio amostral  $E_{\alpha}(t)$  é reduzido (Eq.5) desde as saídas  $y_{\alpha,j}^{(k=m)}$  dos neurônios  $j$  da última camada  $m$ , para cada conjunto de amostras até a última, (MORAES, 2010), sendo acumulado no erro do conjunto de amostras  $E_{\alpha}(t)$  (Eq.5). Após todas processadas é gerado o erro quadrático geral  $E_M(t)$  (Eq.6) e o  $\Delta E_M$  (Eq.7) entre 2 iterações, e se  $\Delta E_M \leq \Delta E(adm)$ , pára-se o algoritmo com as convergências na faixa % requerida e a RNA treinada.

$$E_{\alpha}(t) = \frac{1}{2} \sum_{\alpha,1}^{\alpha,n} [y_{\alpha,j}^{(k)} - y_{\alpha,j}^{(k)}]^2 \quad ; \quad E_M(t) = \frac{1}{p} \sum_1^p [E_{\alpha}(t)] \quad ; \quad \Delta E_M = E_M(t) - E_M(t-1) \quad (5) (6) (7)$$

**Etapa 3 Testes Operacionais da RNA (Fig.3):**

9. Após a fase de treinamento da RNA segue a fase dos testes preditivos da RNA, os valores dos melhores parâmetros e pesos  $w$  de melhor convergência são carregados e, para cada amostra  $x_{\alpha,i}^{(k)}$  das amostras para teste, é executado o algoritmo Freeforward em apenas 1 iteração. Os resultados com convergências altas comprovarão que se a RNA foi bem treinada.

**Prospecções e Parametrizações para o Treinamento da RNA:**

Segundo Silva (2010) e Azevedo (2000), e comprovado na prática, os parâmetros da RNA devem ser pré-ajustados em treinos prospectivos, pois atuam no % e velocidade da convergência. Destacam-se 5 parâmetros atuantes na convergência: 1º) **Nº Atributos (significativos) Entrada:** RNAs exigem uma quantidade de atributos no domínio do problema, Manzan (2016). 2º) **Função de Ativação:** Funções adequadas geram melhor convergência (Manzan, 2016), o programa *Orange*

*Canvas™* T comprova que Tg.Hip. [ $g(u) = \frac{1-e^{-\beta \cdot u}}{1+e^{-\beta \cdot u}}$ ] melhor que a Logística. [ $g(u) = \frac{1}{1+e^{-\beta \cdot u}}$ ]. 3º) **Nº Neurônios nas Intermediária(s):** O nº Neurônios/camada é empírico e ajustes nos parâmetros de execução vão da experiência do pesquisador, (MANZAN, 2016) e (LUDWIG, 2007), (SILVA, 2010) há 3 técnicas para o cálculo: Witten-Flank (Eq.7), Fletcher-Gloss (Eq.8) e Kolmogorov (Eq.9).

$$n_1 = \frac{n + n_c}{2} \quad ; \quad 2\sqrt{n} + n_2 \leq n_1 \leq 2n + 1 \quad ; \quad 2n + 1 \quad (7) (8) (9)$$

Sendo:

$n = n^2$  entradas,  $n_1 = qt. Nros$  na cam. interm.,  $n_2 = qt. Nros$  na cam. saída,  $n_c = n^2$  classes

**4º) Faixa dos Aceleradores:** Os aceleradores de aprendizado  $\eta$  e de avanço  $\mu$  (momentum) são importantíssimos, (MANZAN, 2016). O  $\eta$  atua no gradiente  $\eta \cdot \delta_{af}^{(k)} \cdot y_{af}^{(k-1)}$  como uma bússola direcionando a convergência no caminho oposto ao sinal do gradiente (SILVA, 2010). A 2ª parcela momentum ( $\mu \cdot \Delta w$ ) atua na velocidade da convergência até o erro mínimo, (SILVA, 2016). O erro  $\Delta w$  é repassado aos pesos  $w_{ji}^{(k)}$  das camadas a cada *Back-Propagation*, (LUDWIG, 2007). Essas parcelas levam o EQM ( $E_M$ ) entre saídas (Eq. 5) para o mínimo de sua variação  $\Delta E_M$  (Eq.7) entre iterações, (SILVA, 2010).

**5º) Convergência Mínima x Nº de Iterações:** Via de regra, (SILVA,2010), há um mínimo de iterações para altas convergências, evitando o *underfitting* (baixa convergência por parada antecipada), ou o *overfitting* (convergência degradada por excesso de iterações). Normalmente, altas convergências são mais demoradas e baixas mais rápidas.

**6º) EQM:** São necessários muitos treinos, pois vários são descartados por desviarem da solução, e segundo (Gomes 2015), algoritmos meta-heurísticos não garantem a obtenção da solução.

### Apoio Multicritério com AHP (Analytic Hierarchy Process)

Orçar propostas de projetos é complexo e demorado devido às muitas variações em “resistência x menor custo”. Há várias técnicas de apoio a gestão e uma delas é a “apoio à decisão por Multicritério” (ADMC-AHP) criada nos 90’s por Saaty. O objetivo, melhor projeto, possui alternativas e critérios (Fig.5), (GOMES, 2015), contendo ou não riscos. Cria-se então, uma tabela hierárquica contendo: objetivo (1º nível), as alternativas (2º nível) e os critérios (3º). Os critérios são as características comuns às alternativas, definidos por experts, (Fernandez, 2020). Estes são normalizados no mesmo domínio (0 a 1) tornando-se utilidades. Recebem um peso (relevância %) definido por especialistas e calculado na tabela cruzada 1 (Eqs. 10 e 11), são recalculados por fator arbitrário ou qualitativo (AHP) agora por alternativa, na tabela 2, recarregando-os na tabela 1. O objetivo (Gan-Proj) é calculado (Eq.12) para cada alternativa. A coluna Gan-Proj (Tab.1) indicará a alternativa de maior ganho agregado. Esse tipo de cálculo é chamado ADCM-AHP, (SANTOS, 2013).

Tabela 1: Ex. Matriz de Decisão Multicritério

Proj.	CustConstru	UCstConst	PCstConst	NvAccto	UNvAccto	PNvAccto	MgLucr	UMgLucr	PMgLucr	ROI	UROI	PROI	Gan - Proj
8	1.957.311,70	0,48	0,23	1,05	0,33	0,40	0,72	0,32	0,20	1.577.593,23	0,50	0,18	0,39
9	990.230,46	0,74	0,23	1,08	0,34	0,40	0,78	0,34	0,20	913.388,58	0,29	0,18	0,42
10	836.619,46	0,78	0,23	1,03	0,33	0,40	0,77	0,34	0,20	688.621,47	0,22	0,18	0,41
	3.784.161,61	Peso->	0,7	3,16	Peso->	0,4	2,27	Peso->	0,8	3.179.603,28	Peso->	0,9	

Design: Autor - Fone: (SANTOS, 2013)

$$P_{c_{ij}} = P_{c_{ij}} : \sum_{i=1}^n \left[ \sum_{j=1}^n (P_{c_{ji}}) \right] ; P_{Nc_i} = MédiaGeom(P_{c_{ji}}) ; G_{aP_{f_i}} = \sum_{i=1}^n (P_{c_i} \cdot U_{c_i}) \quad (10) (11) (12)$$

Tabela 2: Matriz Cruzada de Pesos/Alternativa

	PCstConst	PNvAccto	PMgLucr	PROI	Crit	PCstConst	PNvAccto	PMgLucr	PROI		PesosNrm
PCstConst	1	0,5714286	1,14286	1,2857143	4	0,23	0,23	0,23	0,23	PCstConst	0,23
PNvAccto	1,75	1	2	2,25	7	0,40	0,40	0,40	0,40	PNvAccto	0,40
PMgLucr	0,875	0,5	1	1,125	3,5	0,20	0,20	0,20	0,20	PMgLucr	0,20
PROI	0,7777778	0,4444444	0,88889	1	3,1111111	0,18	0,18	0,18	0,18	PROI	0,18
	4,4027778	2,515873	5,03175	5,6607143	17,611111						

Design: Autor - Fone: (SANTOS, 2013)

Na aferição das alternativas (Ex.Tabela 1), se o peso do critério ROI for mais relevante ao investidor, o Proj.8 seria a escolha, mas como foi o acabamento, o Proj.9 foi a escolha com maior ganho.

### 3. MÉTODOS

- Desenvolvidos os módulos do Protótipo: 2.1-Cadastro de Projetos; 2.2-Mapeamento de Plantas baixas; 3.1-Dimensionamento Convencional; 3.2-Dimension. (RNA) com base pseudocódigos de Silva (2010) e Ludwig (2007); 3.3-Orçamentos e Classificação (AHP).
- Gerados 6 conjuntos de projetos e mapeadas as lajes, injetando-as no sistema (Tabela 3). Prospec-tadas entradas (9 a 15 atributos significativos). Dos conjuntos 3 a 6, separou-se 7 projetos p/ treino e 3 p/ teste), à exceção dos iniciais 1 e 2 (4 p/ treino, 2 p/ teste).
- Dimensionadas as lajes (módulo 3.1) injetaram-se no 3.2, os datasets de entrada das lajes dos projetos 1 a 7) de apenas 3 pavimentos (Térreo, Garagem, Pavto Tipo).

Tabela 3: Mapeamento das Plantas Baixas das Lajes dos Conjuntos de Edifícios

Conj.	Treino	Treino – (Nº Lajes / Projeto) x 3 (Térreo, Garagem, Tipo)	Teste - Proj. (Nº Lajes x 3)
1	1	$(6x3 + 12x3 + 9x3 + 25x3) = 156$	$5(6x3)$
2	2	$(6x3 + 12x3 + 9x3 + 25x3) = 156$	$5(6x3) + 6(20x3)$
3	3	$1(12x3)+2(12x3)+3(25x3)+4(25x3)+5(20x3)+6(20x3)+7(20x3)= 402$	$8(20x3) + 9(25x3) + 10(25x3)$
4a	4	$1(20x3)+2(20x3)+3(25x3)+4(25x3)+5(20x3)+6(20x3)+7(25x3) = 465$	$8(25x3) + 9(20x3) + 10(20x3)$
4b	5	$1(20x3)+2(20x3)+3(25x3)+4(25x3)+5(20x3)+6(20x3)+7(25x3) = 465$	$8(25x3) + 9(20x3) + 10(20x3)$
4b	6	$1(20x3)+2(20x3)+3(25x3)+4(25x3)+5(20x3)+6(20x3)+7(25x3) = 465$	$8(25x3) + 9(20x3) + 10(20x3)$

Fonte: Autor

- Na topologia, a saída (última camada) composta de 5  $y=5$ bits em 5 neurônios, gerando 6 diâmetros de aço (5,6,8,10,12,16 mm). Desbinarizando este 1 Byte na camada final, cada *Byte=1 diâmetro* de uma das armaduras da laje dos pavimentos. A complexidade variou conforme: posição, dimensão, quantidade, carga, área de aço e espaçamento.
- Foram executadas ~460 prospecções de treino/teste, conforme Silva (2010), para encontrar as topologias e a parametrizações de convergências ideais (Tabela 4).

5.1. **Topológica (Faixa) melhor Convergência:** (8, 10 e 12) e a adotada foi **14-8-5** de maior %, conforme (Eq.7) a técnica de Witten-Flank mencionada em Silva (2010);

Tabela 4: Conjuntos de Projetos vs Topologias Prospec-tadas vs Topologias Escolhidas

Conjunto	Treino / Teste	Projs .Treino / Teste	Nº ents. / Nº nros / Nº saidas	Top. Melhor Converg.
1	1	Tr(1 a 4) / Ts(5)	9 / 8 a 12 / 5	9-12-5
2	2	Tr(1 a 4) / Ts(5 e 6)	9 a 13 / 8 a 12 / 5	13-12-5
3	3	Tr(1 a 7) / Ts(8 a 10)	13 a 14 / 10 a 12 / 5	14-10-5
4a	4	Tr(1 a 7) / Ts(8 a 10)	14 / 10 a 12 / 5	14-10-5
4b	5	Tr(1 a 7) / Ts(8 a 10)	15 / 10 a 12 / 5	15-10-5
4b	6	Tr(1 a 7) / Ts(8 a 10)	15 / 10 a 14 / 5	15-12-5

Fonte: Autor

5.2. **Atributos de Entrada:** Posições 4 lajes vizinhas, Tp. cômodo, Lado x, Lado y, Cargatotal/m<sup>2</sup>, Altura; Direções das barras, Fck, Fyd, Área e Espaçamento das barras;

5.3. **Função de Ativação:** Tangente Hiperbólica na camada intermediária (Gráf.2);

5.4. **Aceleradores (Faixa):** Das prospecções resultaram  $\eta \leq 0,1$  e  $\mu \leq 0,05$ , sugeridas por Silva

(2010) e utilizadas por Abdalla (2013);  $\eta$  e  $\mu$  maiores geram erro  $\left(\frac{k}{\mu}\right)$ .

5.5. **Converg. (%) Mín. e Máx.Iterações:** 80% de converg. em até 100 iterações.

5.6. **EQM:**  $\leq 10^{-2}$ . Processamento em 4 etapas (pré-processamento; Freeforward; BackPropaga-tion; pós-processamento), (LUDWIG, 2010).

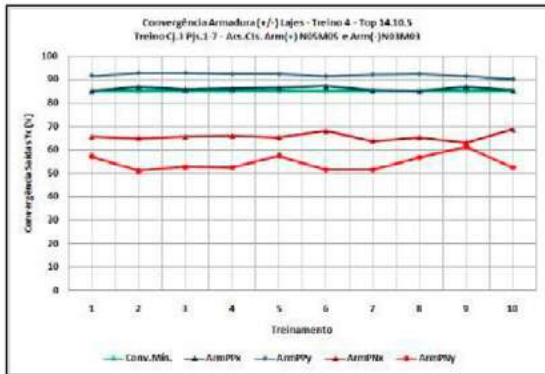
- Dimensionadas as lajes, os resultados passaram ao módulo 3.3 gerando o orçamento (projs. 8 a 10)obtendo o mais vantajoso (Critérios: Custo, Acabamento, Margem de Lucro, ROI) via AHP.

## RESULTADOS E DISCUSSÕES

### Treino 2 Conj.2 (Pjs.1-5) Teste(Pjs.6-7) 13 Atributos Acs.: $\eta 0.1\mu 0.05 (+/-)$ ver.335:

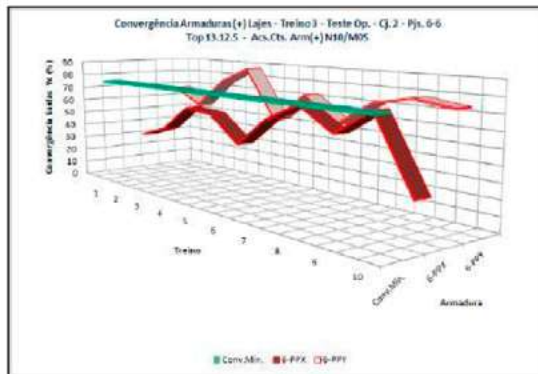
- Iniciando com projetos de 1 a 4 (treinamento) e 6 (teste), até que foram interessantes os resultados das armaduras (+), com convergências médias  $> 75\%$  (PPX 76,2% e PPY 78,2%) no treino (Gráf.1), com PPY 76,2% mais rápida que PPX, e iterações entre 1 e 2. Convergências no Teste Pj.6 (PPX 49,7% e PPY 68,4%) (Gráf.2).
- As convergência das negativas (PNX 41,8% e PNY 4%) ficaram  $<< 75\%$ .

Gráf.1: Cj2 Treino 2 Conv. Arms.(+) Pjs.1 a 4



Fonte: Autor

Gráf.2: Cj2 Testes 2 Conv. Arms.(+) Pj.6



Fonte: Autor

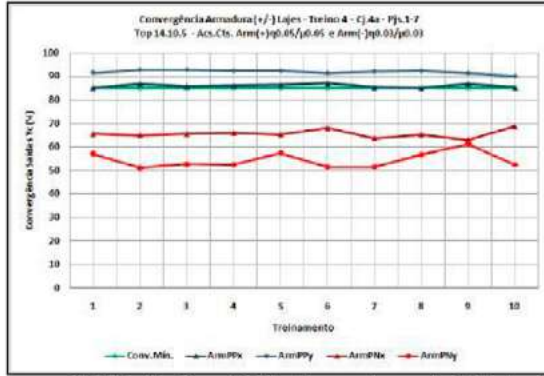
### Treino 3 Conj.3 (Pjs.1-7) Teste(Pjs.8-10) 13 Atribs. Acs: $\eta 0.1\mu 0.05 (+/-)$ ver.336:

- Passaram-se as amostras (de 195 p/ 402), tentando diminuir a heterogeneidade nas armaduras (-), tanto no treino como nos testes. As convergências médias de treino (Projs.8 a 10) arms. PPX (+)  $> 80\%$ , considerado bom, com potencial de maiores.
- As PPY convergiram mais rápido que as PPX, 70% dos treinos entre 1 e 2 iterações.
- As (PNX 86,0% e PNY 86,0%) na faixa dos 60%, ainda baixas ( $<$  mínima 75%), mas maior que o conjunto 2 composta por 13 atributos e 5 projetos.
- Testes das convergências da PPX e PPY dos projetos 8 a 10 ainda ficaram  $< 75\%$ .

### Treino 4 Conj.4(Pjs.1-7) Teste(Pjs.8-10) 14 Atribs. Acs.:(+) $\eta 05\mu 05/(-)\eta 03\mu 03$ ver.339:

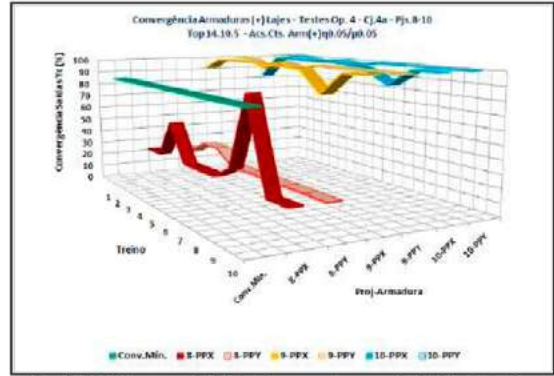
- As convergências médias dos Treinos com (Projs. 1 a 7 - Gráf.5) (465 lajes), foram (PPX 86,0% e PPY 91,9%)  $> 85\%$ , mostrando aumento em relação ao conjunto 3. Já as negativas (PNX 65,5% e PNY 54,5%) (Gráf.5) ainda abaixo da mínima 75%.
- As convergências médias de Teste (Fig.6), Pj8 (PPX 65,5%, PPY 54,5%), Pj9 (PPX 89%, PPY 99%) e Pj10 (PPX 83%, PPY 99%) (Gráf.6) (+), considerado rápido.
- Só no projeto 8 (Gráf.6) a convergência foi  $< 75\%$  (mínima), levando a mais estudos e concluindo que, o maior n° de classes das negativas teve alta influência na dificuldade das (-) convergirem (comprovado no software *Orange Canvas*<sup>TM</sup>).
- Incluindo mais 1 atributo significativo (13  $\rightarrow$  14), ocorreu ganho de convergência nas armaduras (+) e (-). Convergências Arms.(+)  $> 93\%$  foram consideradas ótimas.
- As convergências das (-) deste conjunto ainda foram baixas, mas reduzindo-se os aceleradores,  $\eta = 0,05$  e  $\mu = 0,05$  para  $\eta = 0,03$  e  $\mu = 0,03$ , essas se elevaram 100% comparado ao conj.3, tal é a importância dos aceleradores na convergência.
- O índice de Losses armaduras (+) foi 0%, e negativas (-) foi de 30%.
- Velocidades de convergência médias nos Treinos: Arms. positivas (PPX 2 e PPY 3 em 90% das iterações) e negativas (PNX 11 e PNY 11 em 100% das iterações).

Gráf.5: Treino 4 Cj4a Conv. Arms.(+) Pjs.1 a 7



Fonte:Autor

Gráf.6: Testes 4 Cj4a Conv. Arms.(+) Pjs.8a10

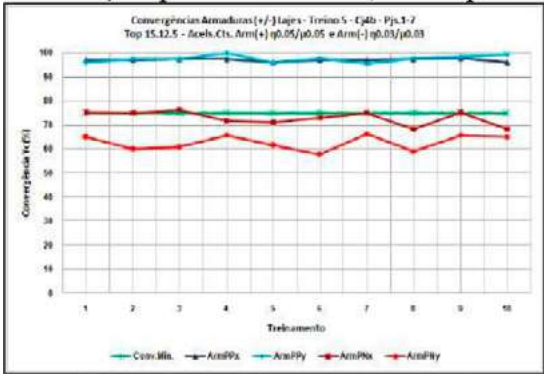


Fonte:Autor

**Treino 5 Conj.4b (Pjs.1-7) Teste(Pjs.8-10) 15 Atribs. Acs.:  $\eta 0.03 \mu 0.03$  (+/-) ver.339:**

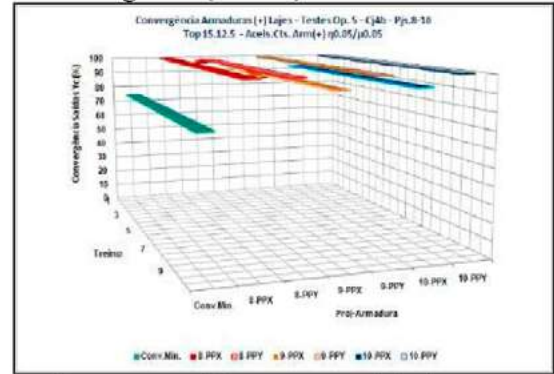
- Encontrando e incluindo mais 1 atributo relevante, (14 → 15), surtiu efeito, elevando as convergências médias das amaduras positivas e negativas (Gráf.7).
- As convergências médias nos treinos dos Projetos 1 a 7 subiram as positivas (PPX 86,0% e PPY 91,9%) e as negativas (PNX 86,0% e PNY 91,9%) (Gráf.7).
- As convergências médias nos Testes das (+) subiram Pj8 (PPX 100%, PPY 100%), Pj9 (PPX 92,7%, PPY 100%) e Pj10 (PPX 93,0%, PPY 100%) (Gráf.8).
- As convergências médias dos Testes das (-) foram Pj9 (PNX 81,4%, PNY 80%) e Pj10 (PNX 68%, PNY 79,5%) ficando menor que a mínima apenas no Pj8 (PNX 65,1%, PNY 57,9%), (Gráf.10) subiram 40% nos projetos 9 e 10 comparados ao Cj.3,mas ao custo do nº de LOSTs das negativas ter aumentou ~30% acima das (+).
- As velocidades médias de convergência dos Treinos: Armaduras (-) aumentaram para 22 iterações parecendo ser esta, a faixa p/ terem alta convergência (Gráf. 9).

Gráf.7: Treino 5 Cj4b Conv. Arms.(+) Pjs.1 a 7



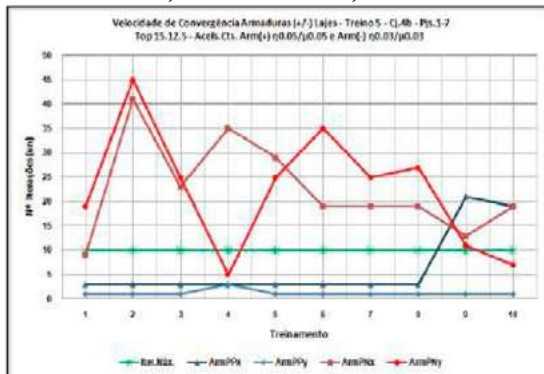
Fonte:Autor

Gráf.8: Teste 5 Cj4b Conv. Arms.(+) Pjs.8 a 10



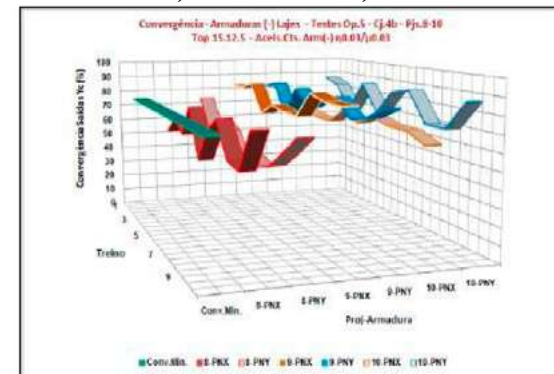
Fonte:Autor

Gráf.9: Treino 5 Cj4b Vel. Conv.(+/-) Pjs.1 a 7



Fonte:Autor

Gráf.10: Teste 5 Cj4b Conv. Arms.(-) Pjs.8 a 10



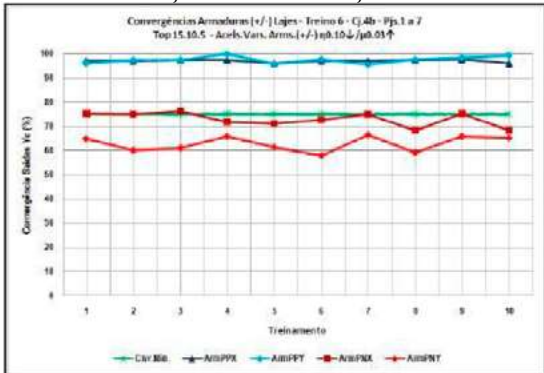
Fonte:Autor



**Treino 6 Conj.4b (Pjs.1-7) Teste(Pjs.8-10) 15 Atribs. Acs.: $\eta$ 0.1 $\downarrow$  $\mu$ 0.03 $\uparrow$ (+/-) ver.339:**

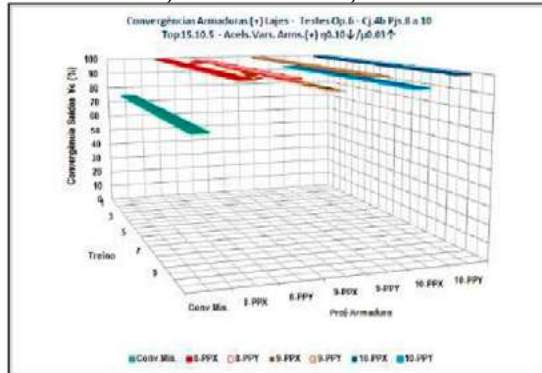
- Os Treinos as armaduras (PPX 97,1% e PPY 97,6%) quase cravaram 100% de convergência média (Gráf.11) mostrando que os aceleradores variáveis aumentaram a convergência. As negativas treinaram até (PNX 73,0% e PNY 63,0%) (Gráf.11).
- A velocidade média das (+) em 3 iterações, mas as (-) subiram para 22, (Gráf.13).
- Observou-se que, as convergências médias das armaduras (+/-) do Conj.6 tiveram convergências altas com aceleradores variáveis, confirmada em Braga (2009).
- As convergências das (-) mais lentas requerem mais pesquisas com aceleradores variáveis,mas no compito geral, estes têm ótimo efeito elevador da convergência.

Gráf.11: Treino 6 Cj4b Conv.Arms.(+) Pjs.1 a 7



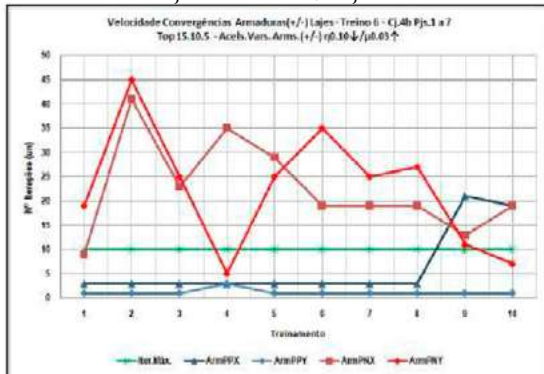
Fonte:Autor

Gráf.12: Teste 6 Cj4b Conv.Arms.(+) Pjs.8 a 10



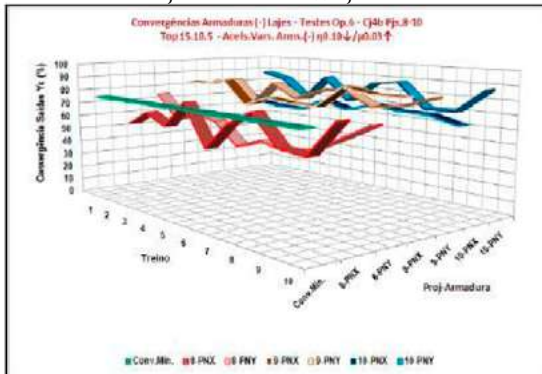
Fonte:Autor

Gráf.13: Treino 6 Cj4b Vel.Conv.(+/-) Pjs.1 a 7



Fonte:Autor

Gráf.14: Teste 6 Cj4b Conv.Arms.(-) Pjs.8 a 10

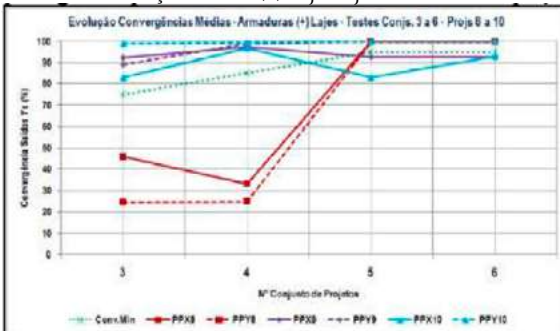


Fonte:Autor

**Evolução das Convergências das Topologias Treinadas (ver. 336 a 339):**

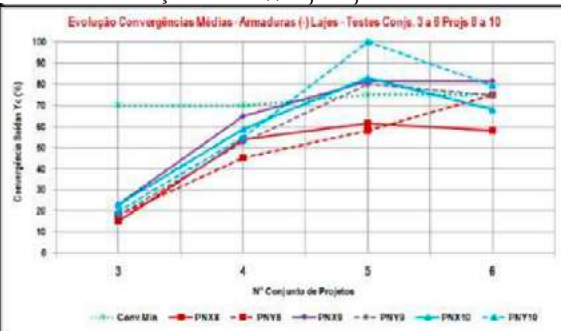
Nos Gráfs.15 e 16, é notável a evolução dos % das convergências das (+/-) dos projs. 8 a 10, conformes as topologias e parâmetros, e ainda indicando espaço para melhorias futuras.

Gráf.15: Evol. Orçam. Arms.(+) Lajes Cj 3 a 6



Fonte: Autor

Gráf.16: Evol. Orçam. Arms.(-) Lajes Cj 3 a 6



Fonte: Autor

## Evolução do Orçamento - Conjs. de Projetos 4 a 6 Projs. 8 a 10 testados via RNA:

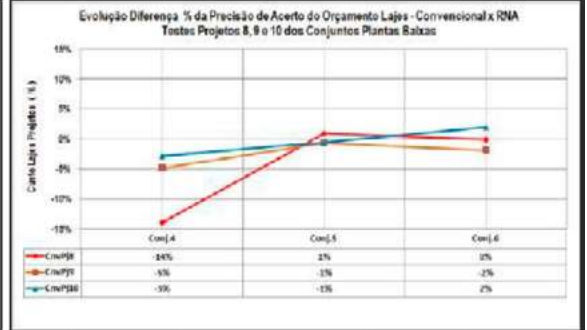
- As diferenças dos custos entre orçamento das lajes dimensionadas, por método Convencional e por RNA, reduziram-se conforme esperado, do conjunto 4 {8(14%) 9(5%) 10(3%)} para o 6 {8(0%) 9(2%) 10(2%)}, Gráfs. 17 e18.
- Mesmo com a precisão atual, o protótipo ainda tem espaço para aperfeiçoar e reduzir o erro de orçamento abaixo de 1%, mas a evolução conseguida é notável.

Gráf.17: Evol. Orçam. Arms.(+) Lajes Cj 3 a 6



Fonte: Autor

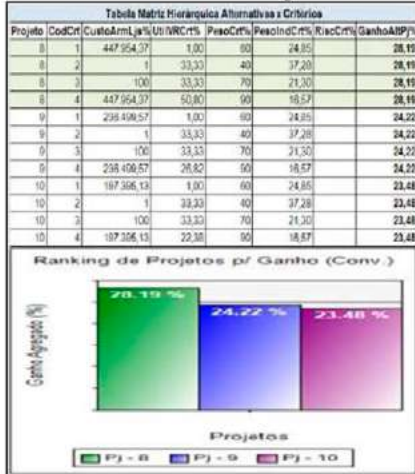
Gráf.18: Evol. Orçam. Arms.(-) Lajes Cj 3 a 6



Fonte: Autor

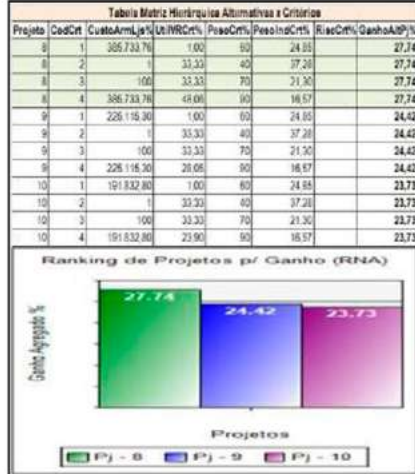
- Na classificação do melhor orçamento das lajes dos Projs. 8 a 10 (um exemplar nas Figs.5 e6), a variação dos % dos critérios revezou os projetos conforme ajustes do operador, como esperado. O módulo 3.3 foi aderente à literatura e pesquisas, e gerou orçamentos das lajes com rapidez e flexibilidade no ajuste dos critérios.

Fig.5: Tab. Hieraq. e Gráf. Class. Pjs.8a10 Cj.4



Fonte: Autor

Fig.6: Tab. Hieraq. e Gráf. Class. Pjs.8a10 Cj.4



Fonte: Autor

## CONCLUSÕES E TRABALHOS FUTUROS

- Normalmente, algoritmos de RNAs exigem muito processamento, (DINIZ, 2017), (FÉLIX,2015), (MANZAN, 2016), mas o este protótipo teve ótimo desempenho.
- A menor/menor dificuldade em convergir é diretamente proporcional ao menor/menor nº de classes a classificar, inclusive corroborado p/ software *OrangeCanvas*<sup>TM</sup>.
- Verificou-se nos Cjs.2 e 3 as amostras formam matrizes com maior heterogeneidade(numérica e posicional) (SPERANDIO, 2003) o que reduz o aprendizado.
- Conclui-se que, uma RNA pode aprender rápido a sistemática da função objetivo, se:
  - Adotando-se atributos ( $x$  e  $y_c$ ) com alta correlação com o sistema equacional.
  - Efetuar treinos com expressivo nº de amostras (Ex.: > 400 Lajes).
  - Adotando-se **aceleradores menores**. Para aceleradores *fixos*, recomenda-se a faixa: de **0.05** <

$\eta(\text{aprendizado}) < 0.05$  (arms.+) e,  $0.03 < \mu$  (propulsor)  $< 0.10$  arms.(.). Para os variáveis, sugere-se:  $0.10 < \eta < 0.05$  e  $0.03 < \mu < 0.10$ .

- A técnica AHP funcionou como esperado no orçamento e classificação do melhor projeto, confirmando que o dimensionamento por RNA está no rumo certo.
- Ficou demonstrado (Gráfs.17 e 18) que a precisão do orçamento no Conj.5 foi a melhor, e a técnica AHP confirmou a precisão, consequência da evolução da RNA.
- Confirmou-se (Gráfs.15 e 16) que o algoritmo de RNA PMC construído funcionou corretamente e pôde gerar altas convergências. Comparando com outros trabalhos, o protótipo atendeu aos objetivos e tem potencial na engenharia de edifícios.

## REFERÊNCIAS

- ABDALLA, L. S.; VOLOTÃO C. S. Estudo da configuração de diferentes arquiteturas de redes neurais artificiais MLP para classificação de imagens ópticas. In: Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR, Foz do Iguaçu, 13 a 18 de abril de 2013, INPE. 2013. p. 8200-8207.
- AZEVEDO, F. M.; BRASIL, L. M.; OLIVEIRA R. C. L. **Redes Neurais com Aplicações em Controle e em Sistemas Especialistas**. 1ª. ed. Florianópolis: Visual Books, 2000.
- BOTELHO, M. H. C.; MARCHETTI O. **Concreto Armado em Ti Amo - Volume I**. 7ª. ed. São Paulo: Editora Blucher, 2013.
- FERNANDEZ, J. C. **Uso de Redes Neurais Artificiais e AHP na Redução da Subjetividade na Seleção de Áreas com Potencial para Coleta de Água da Chuva**. 2020. 68 f. Dissertação (Mestrado em Engenharia Civil)–UFU, Uberlândia - MG, 2020.
- GOMES, L. A. M.; GOMES C.F.S.; ALMEIDA A.T. . **Tomada de Decisão Gerencial - Enfoque Multicritério** - 3ª. ed. São Paulo: Editora Atlas, 2009.
- GOMES, F. M; PEREIRA F. M.; SILVA M. B.; MARINS F. A. S.. Aplicação da Metaheurística Algoritmo Genético na Otimização de Problemas com Múltiplas Respostas. In: XXXV ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO Perspectivas Globais para a Engenharia de Produção, Fortaleza, CE, Brasil, 13 a 16 de outubro de 2015.
- KOVÁCS, Z. L. **Redes Neurais Artificiais - Fundamentos e Aplicações**. 2. ed. São Paulo: Editora Collegium Cognitione e Edição Acadêmica, 1996.
- LUDWIG JR, O. L.; MONTGOMERY. E. **Redes Neurais Artificiais para engenharia e ciências aplicadas**. 1. ed. São Paulo: Editora Ciência Moderna, 2007.
- MANZAN, J. R. G. **Análise de Desempenho de Redes Neurais Artificiais do Tipo Multilayer Perceptron por meio do Distanciamento dos Pontos do Espaço de Saída**. 2016. 129 f. Tese(Doutorado em Engenharia Elétrica)–UFU, Uberlândia - MG, 2016.
- MORAES, E. C. **Reconhecimento de Padrões e Redes Neurais Artificiais em Predição de Estruturas Secundárias de Proteínas**. 2016. 135 f. Tese (Doutorado em Ciências e Engenharia de Sistemas de Computação)–UFRJ, Rio de Janeiro - RJ, 2010.
- SANTOS, L. F; CRUS R. B. C.. O Uso do Método AHP na Tomada de Decisão para Seleção de Sistemas de Lajes de Edifícios Comerciais. In: Engenharia Estudo e Pesquisa. ABPE, v. 13 - n. 1 - p. 39-52 - jan./jun. 2013.
- SILVA, I. N.; SPATTI D, H.; FLAUZINO, R. A. **Redes Neurais Artificiais para engenharia e ciências aplicadas**. 1ª. ed. São Paulo: Artliber Editora, 2010.
- SPERANDIO, D.; MENDES. J. T.; SILVA L. H. M. **Cálculo Numérico**. 1ª.ed. São Paulo: Pearson Prentice Hall, 2003.
- UBERABA, L. R. H. **Projeto de Estruturas de Concreto Armado - Projetos de Estruturas de Edifícios à Luz da Informática**. 1ª. ed. São Paulo: Editora e Gráfica Publi, 2009.

## DADOS DOS AUTORES

**Adriano Ricardo Borges Costa**. Graduação em Engenharia Civil - Centro Universitário de Volta Redonda – RJ, Mestrado Profissional em Materiais (Stricto Sensu) - Centro Universitário de Volta Redonda – RJ.

**Eliane da Silva Christo**. Graduada em Engenharia Elétrica pela Universidade Federal de Juiz de Fora, mestrado em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro, doutorado em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro.

**Tiago Araújo Neves**. Graduação em Ciência da Computação pela Universidade Federal de Ouro Preto (2004), mestrado em Computação pela Universidade Federal Fluminense (2007) e doutorado em Computação pela Universidade Federal Fluminense (2011)

# PROPOSTA METODOLÓGICA PARA A REVISÃO SISTEMÁTICA DA LITERATURA: ESTUDO DE CASO ASSOCIADO À TEORIA DAS FILAS E SIMULAÇÃO

## METHODOLOGICAL PROPOSAL FOR A SYSTEMATIC LITERATURE REVIEW: CASE STUDY ASSOCIATED WITH QUEUING THEORY AND SIMULATION

Kelly Cristine Franco de Azevedo

Cecília Toledo Hernández

Eliane da Silva Christo

**RESUMO:** O artigo oferece proposta de revisão sistemática abordando teoria das filas e simulação, utilizando as bases de dados *Scopus* e *Web of Science* da CAPES. Estas áreas são examinadas para entender a evolução e tendências das publicações, avaliando a eficiência dos processos logísticos. Com a análise metodológica, a pesquisa explora a quantidade anual de artigos, destacando extremos de atividade. A fusão dos dados proporciona insights sobre os padrões de pesquisa, evidenciando anos de alta produção acadêmica e flutuações naturais. Os resultados sublinham anos com maior publicação, revelando dinâmicas na pesquisa de teoria das filas e simulação e oferecendo visão das tendências. Ao restringir a busca à logística, embora reduza a quantidade de artigos, destaca o uso dessas ferramentas nessa área. Isso demonstra a relevância do tópico, sugerindo oportunidades para futuras pesquisas e projetos. A pesquisa oferece um panorama sobre a dinâmica de publicações, promovendo a continuidade da investigação nestas áreas.

Palavras-chave: Teoria das Filas. Simulação. Logística. Revisão sistemática.

**ABSTRACT:** The present article provides a systematic review addressing queueing theory and simulation, utilizing the Scopus and Web of Science databases from CAPES. These areas are examined to comprehend the evolution and trends of publications, assessing the efficiency of logistical processes. Through methodological analysis, the research explores the annual quantity of articles, highlighting peaks and valleys of activity. The merging of data provides insights into research patterns, emphasizing years of high academic production and natural fluctuations. The results underscore years with higher publication rates, revealing dynamics in queueing theory and simulation research and offering a view of trends. By narrowing the search to logistics, although it reduces the quantity of articles, it still highlights the use of these tools in this field. This demonstrates the ongoing relevance of the topic, suggesting opportunities for future research and projects. The study offers an overview of publication dynamics, promoting continued investigation in these areas.

Keywords: Queueing Theory. Simulation. Logistics. Systematic Review.

## INTRODUÇÃO

Num ambiente globalizado a logística surge como uma vantagem competitiva utilizada como estratégia de criação de valor. Nesse sentido a avaliação de desempenho tem o intuito de promover informações sobre a performance das atividades logísticas (BALLOU, 2012). Quando se fala de desempenho logístico, os indicadores abrangem todas as atividades logísticas de transporte, aquisição de materiais e peças, gestão de estoque, gestão de armazenagem, gestão da manutenção, engenharia de infraestrutura, e distribuição física, sendo o tempo e a qualidade aspectos desafiadores da produção nos mercados transnacionais, por este motivo o indicador “Qualidade e competência logística” está entre os definidos em 2018 pelo Banco Mundial (MACHADO; DOS SANTOS, 2021).

Para Tavares (2018) na avaliação dos cenários sobre desempenho logístico é necessário a adoção de ferramentas de apoio à decisão que contribuam para a redução de custos e ganhos operacionais que por sua vez aumentam a capacidade de competição das empresas. Entre estas ferramentas se destacam a simulação computacional e a teoria das filas (DA SILVA et al., 2021).

A teoria das filas é um ramo da Pesquisa Operacional que tem como foco principal utilizar conceitos básicos de processos estocásticos e de matemática aplicada com o intuito de estudar o processo de formação de filas e quais são as características da mesma permitindo o dimensionamento adequado de instalações, equipamentos e infraestrutura adequada para a quantidade de demanda (DE BRUNS; SONCIM; DE SINAY, 2001).

Outra ferramenta utilizada para estudar e avaliar os processos de fila em sistemas de produção e serviços é a simulação computacional. Da Silva et al. (2021) realizaram um estudo sobre a otimização dos processos logísticos do porto de Itajaí utilizando simulação e indicadores da teoria das filas, mostrando as particularidades de uso de cada uma das ferramentas. Neste contexto aparecem diferentes questões de pesquisa a serem respondidas: Qual dessas duas ferramentas, teoria de filas ou simulação, dá uma melhor resposta na otimização de processos logísticos relacionados com filas de espera em processos produtivos ou de serviços? Quais vantagens ou desvantagens estão associadas a cada ferramenta?

Neste sentido este artigo apresenta uma proposta metodológica para conduzir a revisão

bibliográfica tendo como temática a utilização da teoria das filas e da simulação computacional de eventos discretos para avaliar indicadores de eficiência logística. Como resultado espera-se que a proposta metodológica auxilie no levantamento do estado da arte e no entendimento dos *gaps* ou lacunas que permitam desenvolver uma pesquisa aprofundada na temática com aplicação prática em estudo de caso nas operações portuárias.

## REFERENCIAL TEÓRICO

### Revisão Sistemática

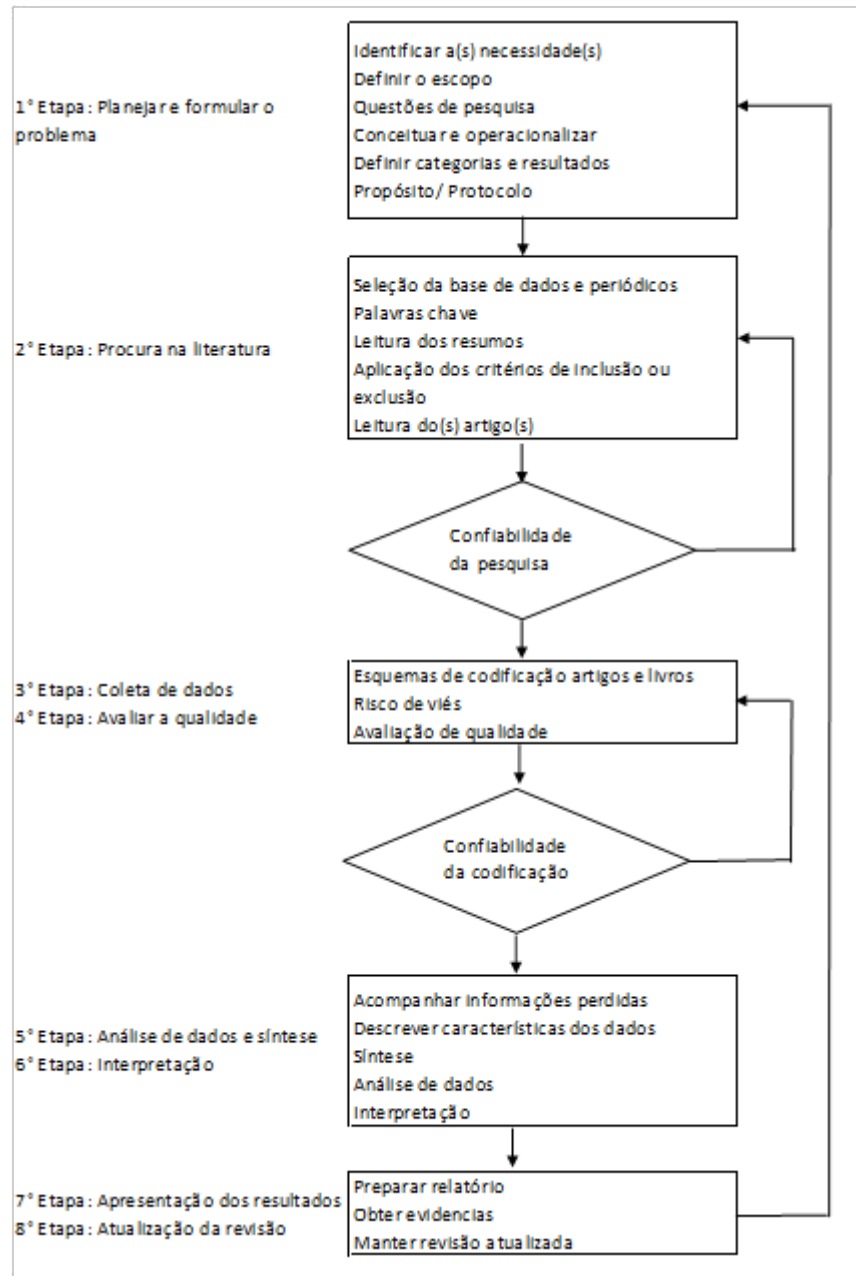
Revisão sistemática é o nome dado a uma abordagem de pesquisa que tem como objetivo coletar, avaliar e sintetizar as evidências relevantes sobre um determinado assunto. É uma ferramenta essencial para consolidar o conhecimento existente, oferecer insights confiáveis e embasar decisões para realizar uma pesquisa. Conforme destacado por Biolchini (2005) as principais etapas que compõem o desenvolvimento de uma revisão sistemática são: planejamento, revisão, execução, revisão e análise dos resultados.

O planejamento é o ponto de partida, onde é definida a pergunta de pesquisa e os critérios de inclusão e exclusão dos estudos. Isso estabelece os parâmetros para a busca e seleção dos artigos relevantes. A etapa de revisão envolve a busca em bases de dados acadêmicas e outras fontes relevantes para identificar todos os estudos pertinentes ao tema.

Na fase de execução, os estudos são avaliados, seguindo critérios pré-definidos. A análise dos resultados implica na síntese das informações extraídas dos estudos selecionados, identificando padrões, tendências e relações entre os resultados. A revisão adicional pode ocorrer após a análise, permitindo uma validação ou ajuste das etapas anteriores.

Para Rothstein e Hopewell (2009) a revisão sistemática da literatura utiliza-se de critérios muito bem definidos e de modo rigoroso para identificar, avaliar e sintetizar a literatura. Estes próprios autores propõem que para desenvolver e garantir a confiabilidade, devem ser seguidas oito etapas. Thomê et al. (2016) elaboram o fluxograma da Figura 1 explicando o que deve ser incluído em cada uma delas.

Figura 1: Etapas para desenvolver a revisão sistemática da literatura.



Fonte: Thomê et al. (2016).

Independentemente da quantidade de etapas propostas por cada autor, na revisão sistemática o pesquisador pode se fundamentar nas contribuições dos autores de análises aprofundadas apresentadas nas obras, nutrindo sua própria exploração por meio da compreensão e ressonância com o pensamento preexistente (SEVERINO, 2014). Muitas vezes propostas metodológicas mais simples que contenham todas as etapas podem ser de interesse para o pesquisador.

### Teoria das Filas

A teoria das filas busca estudar a avaliação da espera em filas, empregando medidas representativas de desempenho do sistema (TAHA,

2008). As filas se formam quando a procura por um serviço é superior a capacidade de atendimento desse sistema (CAMELO et al., 2010). Portanto o foco dessa teoria é analisar quanto tempo um usuário espera em fila até o atendimento usando como instrumento modelos diversos de filas (HILIER LIEBERMAN, 2013).

O sistema de filas pode ser descrito através de um usuário chegando e esperando ser atendido, passando pelo atendimento e logo após ser atendido saindo do sistema (BRONSON, 1985). O principal objetivo de um sistema de filas é relacionar as variáveis de chegada, atendimento e capacidade de serviço, mostrando como deve ser o funcionamento desse sistema

(HILLIER; LIEBERMAN, 2013). As análises sobre os sistemas de filas buscam então descobrir um ponto de equilíbrio entre a capacidade de serviço disponível aos usuários e os custos operacionais que são gerados até o atendimento (ALVES; MENEZES, 2010).

Segundo Fraga (2012), a notação de Kendall, é uma maneira de descrever os modelos para simplificação de sua análise, utilizando cinco características: (a/b/c):(d/e/f) onde: “a” significa o processo de chegada, onde essa chegada é representada pelo intervalo entre eles; “b” o processo de atendimento que é representado pelo tempo em serviço; “c” o número de servidores, ou seja, a quantidade da fonte de serviço; “d” é a regra da fila representando como é o processo de chegada e atendimento do usuário no sistema; “e” é o número máximo de clientes no sistema; “f” o tamanho máximo da população que pode ser atendida.

### Simulação

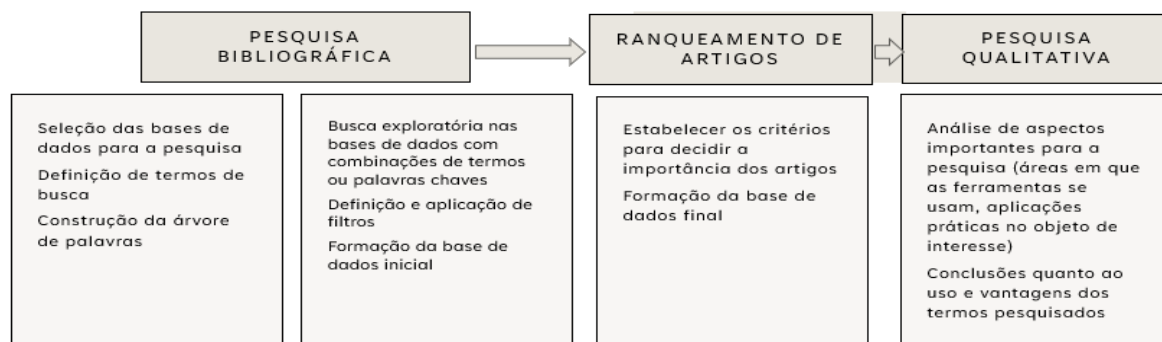
A simulação pode ser definida como sendo uma técnica utilizada em diversas áreas com o objetivo de modelar e estudar sistemas não triviais, isso porque esta tem a habilidade de fazer a iteração entre problemas do meio corporativo e a academia (FELIX; PEIXOTO; EDWIGES, 2023), permitindo que se estude o comportamento, desempenho e interações em um ambiente ideal e virtual através da criação de

modelos computacionais que buscam a semelhança do que acontece na vida real, proporcionando então detalhes de como esses sistemas podem evoluir ao longo do tempo. A simulação é cabível a tudo aquilo que é real, ou seja, significa reproduzir a performance de um sistema (CAMELO et al., 2010). É um método que emprega um modelo matemático para viabilizar a exploração e avaliação do comportamento e desempenho de um sistema, verificando a antecipação de comportamentos futuros envolvendo a criação de um ambiente virtual que espelha o sistema em análise. Isso proporciona a capacidade de explorar o comportamento e o desempenho do sistema sob diversas condições, sem a necessidade de realizar modificações no sistema real. Nesse contexto, a solução de integração se manifesta como um sistema complexo, cujo modelo é caracterizado por elementos estocásticos, dinâmicos e discretos (WIESNER, 2016).

### PROPOSTA METODOLÓGICA

A proposta metodológica para a revisão sistemática da literatura sobre os temas de interesse pode ser apreciada na Figura 2. Esta é considerada uma proposta simplificada que contempla três etapas, mas inclui todos os aspectos que outros pesquisadores como Biolchini (2005), Rothstein e Hopewell (2009) e Thomê et al. (2016) propuseram em seus estudos.

Figura 2: Etapas para a revisão sistemática



Fonte: Elaborado pelos autores

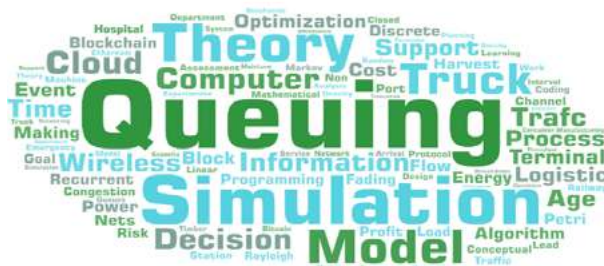
Para realizar uma revisão sistemática se parte de uma pergunta de pesquisa e de critérios de inclusão e exclusão de documentos e bibliografias para poder realizar o estudo. Nesta pesquisa pretende-se estudar aplicações de teoria de filas e simulação, para estabelecer qual das duas ferramentas dá uma melhor resposta na otimização de processos logísticos relacionados com filas de espera em processos produtivos ou de serviços.

Assim, os passos realizados na etapa de pesquisa bibliográfica foram:

- **Seleção das bases de dados para a pesquisa:** Isto compreende a escolha da plataforma de busca. No contexto deste estudo, a *Web of Science* e a *Scopus* foram escolhidas para serem as duas plataformas de informações para a identificação e consolidação de estudos relacionados ao tema estipulado. Essas bases de dados merecem destaque devido à sua amplitude temática, abarcando variados campos do saber. Além disso, as plataformas em questão apresentam ferramentas de pesquisa avançadas, que concedem a capacidade de refinar os resultados de acordo com critérios específicos o que simplifica a seleção de estudos pertinentes e assegura um maior grau de validade e solidez à revisão bibliográfica.

Para iniciar a pesquisa foram formulados critérios de busca e empregou-se a árvore de palavras (Figura 3) como primeiro recurso para o desenvolvimento da pesquisa.

Figura 3: Árvore de palavras



Fonte: Elaborado pelos autores

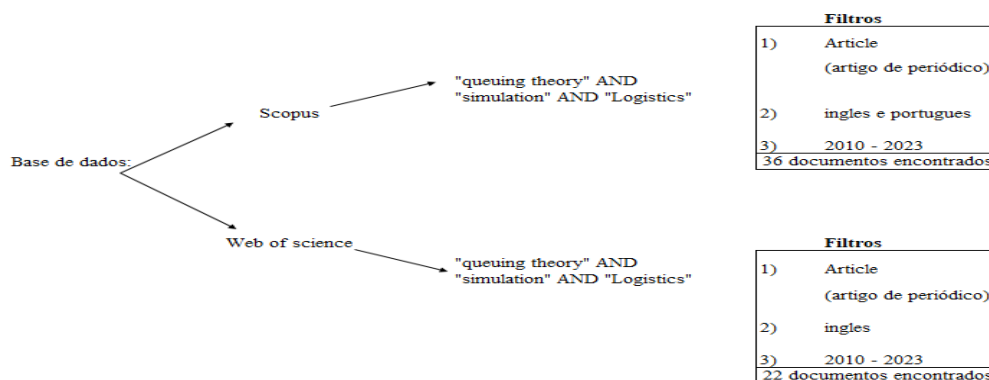
Essa abordagem permitiu uma representação das palavras mais frequentes relacionada a temática do trabalho, auxiliando na identificação de tendências e conteúdos predominantes.

- **Definição dos termos de busca:** Com a árvore de palavras construídas é possível a formação das palavras-chave. Neste estudo as palavras chaves usadas foram: *"queuing theory"*; *queuing theory* AND *"simulation"* e *"queuing theory"* AND *"simulation"* AND *"Logistics"*.

- **Busca exploratória de documentos nas bases de dados:** Esta busca exploratória foi realizada a partir da estruturação dos termos utilizando lógica booleana, uma abordagem que possibilita a organização precisa das palavras-chave com base nas particularidades das bases de dados. Esse método permitiu criar consultas de pesquisa refinadas, combinando os termos de maneira estratégica para obter resultados mais pertinentes e alinhados com os objetivos da investigação. A aplicação da lógica booleana proporciona uma abordagem sistemática para moldar a busca, maximizando a relevância das informações recuperadas e otimizando o processo de descoberta acadêmica.

- **Definição e aplicação de filtros:** Após a seleção das bases de dados e a definição dos termos de pesquisa, foram implementados filtros com o intuito de assegurar a qualidade dos dados adquiridos. Essa abordagem permitiu discernir e excluir informações que pudessem ser desatualizadas, imprecisas, antiquadas ou desprovidas de relevância para a investigação em questão. Um exemplo dos filtros utilizados é representado na Figura 4 para um conjunto de palavras-chaves.

Figura 4: Filtros base de dados



Fonte: Elaborado pelos autores

Na Tabela 1, é possível observar os resultados das pesquisas booleanas em cada uma das bases de dados.



Tabela 1: Triagem dos artigos bases “Web of Science” e “Scopus”

Triagem dos artigos	Scopus	Web of science
“queuing theory”	4222	5252
“queuing theory” AND “simulation”	1537	1686
“queuing theory” AND “logistics”	126	112
“queuing theory” AND “simulation” AND “Logistics”	52	46
“queuing theory” AND “simulation” AND “Logistics” (filtros: 2010-2023; inglês e português; artigo de periódico)	13	28

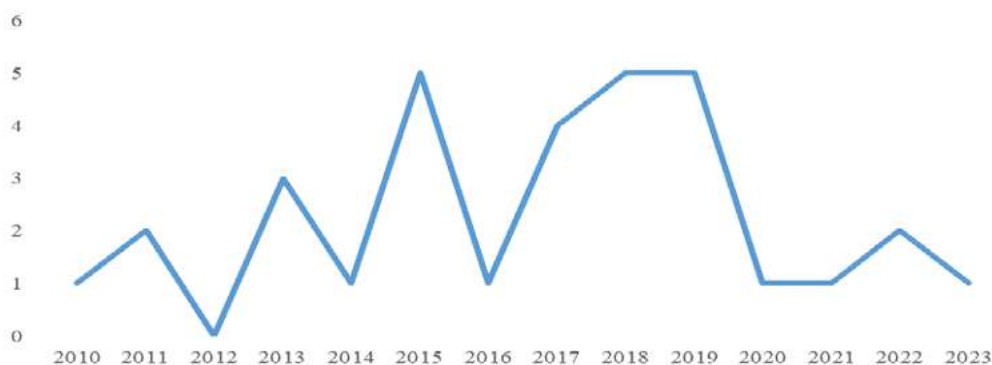
Fonte: Elaborado pelos autores

A predominância de artigos é notável quando se empregam os termos “queuing theory” e “queuing theory” AND “simulation”. Isso reflete a abrangência e relevância dessas ferramentas em uma ampla gama de campos científicos. A riqueza de pesquisa nessas áreas sugere sua importância contínua no estudo de sistemas de espera e análise de processos.

No entanto, é interessante observar que ao incluir o termo “logistics” juntamente com os outros dois, a quantidade de artigos se reduz. Esse fenômeno sugere que, embora a incorporação da logística possa trazer foco e aplicabilidade específicos, ela também pode limitar a diversidade de tópicos abordados. Isso ilustra como a interseção entre esses termos influencia a amplitude das publicações e aponta para a influência direta que a logística exerce na contextualização e escopo da pesquisa.

Na Figura 5, é visível a representação gráfica da quantidade de artigos abordando o tema estabelecido, publicados no intervalo de tempo entre 2010 e 2023.

Figura 5: Quantidade de artigos por ano



Fonte: Elaborado pelos autores

Pode-se apreciar que os anos com a maior quantidade de publicações são 2015, 2018 e 2019, cada um com um total de 5 artigos publicados. De maneira geral houve uma constância nos estudos sobre teoria das filas, simulação e logística ao longo dos anos. O decréscimo nos últimos anos, ainda pode ser um reflexo da pandemia. Isso mostra que o tema é atual com aplicações em diversos campos.

**- Formação da base de dados inicial:** Os conjuntos de dados provenientes do *Scopus* e da *Web of Science* foram combinados, e os artigos duplicados foram eliminados, o que resultou em um novo arquivo no formato de planilha. A base de dados da *Web of Science* apresentou 28 artigos, enquanto a base da *Scopus* continha 13 artigos. Após a remoção das duplicações, o total ficou em 32 artigos únicos. Após a leitura dos resumos decidiu-se utilizar todos os artigos previamente selecionados e os 32 artigos vão compor a base de dados inicial.

A segunda etapa proposta é o ranqueamento dos artigos segundo seu grau de importância para finalmente obter a base de dados final. A seguir será explicitado os aspectos destacados desta etapa:

**- Definição de critérios de ranqueamento:** Para o ranqueamento dos artigos é necessário definir critérios e pesos. Existem diferentes métricas para avaliar o desempenho das pesquisas e, portanto, a qualidade das publicações. A Tabela 2 mostra diferentes critérios associados a tópicos de destaque para o estudo em questão.

Tabela 2: Tópicos e critérios para ranqueamento

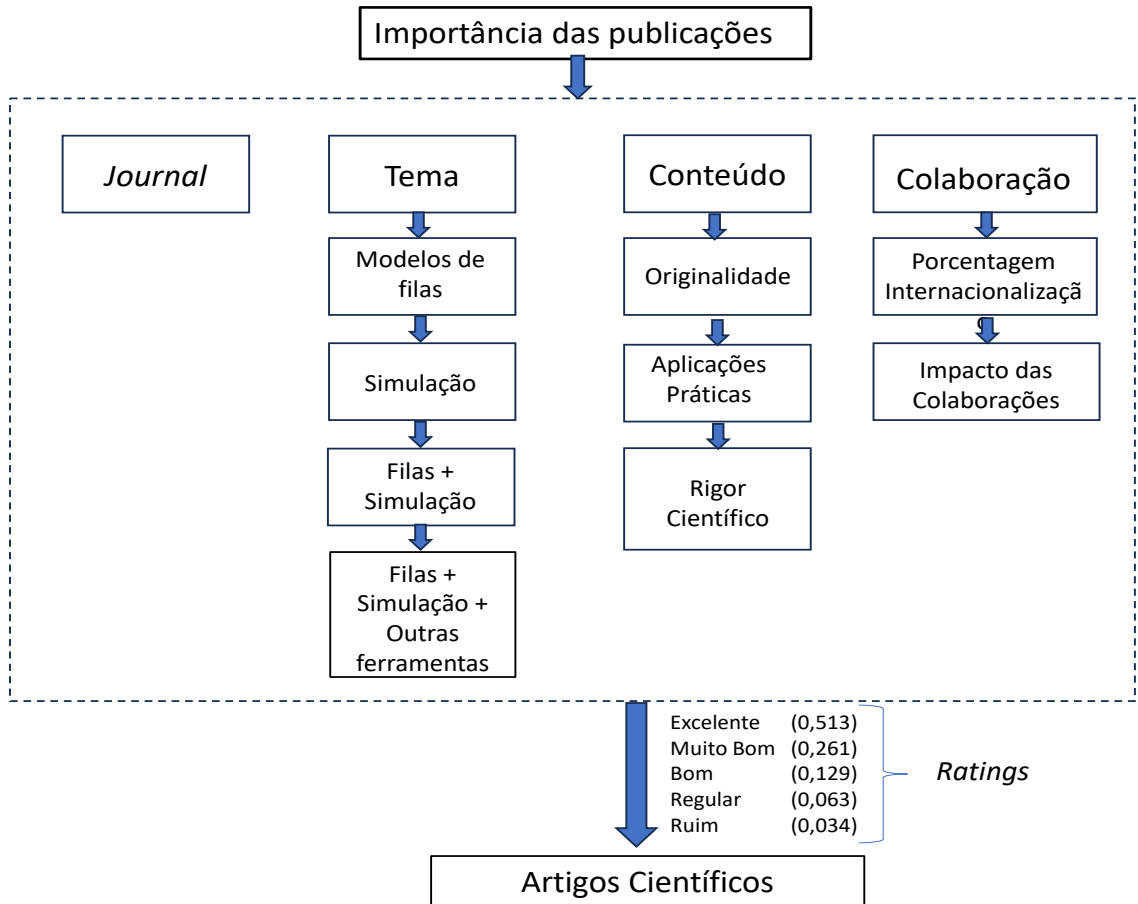
Tópicos	Crítérios
<b>Journal</b> (Tem como objetivo identificar a relevância e excelência do jornal em que o artigo foi publicado)	Valor dos índices JCR/SJR/SNIP do jornal em que o artigo foi publicado.
<b>Tema</b> (Tem como objetivo identificar a relevância do artigo para a pesquisa. Para isso, avalia-se através dos resumos e palavras-chave dos artigos, o quanto que estão alinhados aos objetivos desta pesquisa acadêmica em questão)	Nota dada ao artigo em relação ao seu alinhamento à pesquisa em função dos seguintes temas: - Modelos de filas aplicados às operações logísticas - Simulação de eventos discretos aplicado às operações logísticas - Simulação de eventos discretos e modelos de filas aplicados de forma conjunta em operações logísticas - Modelos de filas e simulação associados a outras ferramentas de tomada de decisão
<b>Conteúdo do artigo</b>	- Originalidade - Aplicações práticas - Rigor científico
<b>Colaboração</b> (Indica o número de publicações de uma instituição, grupo ou pesquisador produzidas em coautoria internacional, nacional ou institucional, e autoria única)	- Porcentagem de internacionalização - Impacto das colaborações

Fonte: Elaborado pelos autores

Outros critérios ainda podem ser definidos no transcurso da pesquisa. Para estabelecer as prioridades podem ser usados diversos métodos como matrizes de prioridades ou métodos de tomada de decisão com múltiplos critérios.

É uma proposta deste estudo o uso do *Analytic Hierarchy Process* (AHP), desenvolvido por Saaty (1980). Neste caso será AHP com *ratings* dada a quantidade de artigos a comparar ser superior a 9 elementos que é uma limitante estabelecida para esta ferramenta. A Figura 6 mostra a estrutura hierárquica que será usada para aplicar o AHP.

Figura 6: Estrutura Hierárquica para ranqueamento dos artigos



Fonte: Elaborado pelos autores

A hierarquia se inicia pelo objetivo global “Importância das publicações”. No segundo nível, os critérios (identificados como “tópicos” na Tabela 2), no terceiro nível encontram-se os subcritérios (variáveis que definem cada tópico da Tabela 2) e por último as alternativas (artigos científicos que serão avaliados, os 32 que fazem parte da base de dados inicial para esta aplicação). Para avaliar critérios e subcritérios será realizada uma comparação pareada usando a Escala Fundamental proposta por Saaty (1980). Para comparar as alternativas (artigos científicos) perante os subcritérios será realizada uma avaliação com a escala absoluta definida na Figura 6 (excelente, muito bom, bom, regular e ruim) esses níveis de intensidade definem as categorias dos *ratings*. As prioridades destas cinco categorias ou intensidades foram determinadas usando o processo de comparações par a par do método AHP (BASTOS, 2015).

A nota dada ao artigo em relação ao alinhamento do tema e o conteúdo será estabelecida pelos autores do trabalho, todos os outros indicadores têm como fonte o motor de busca acadêmica das bases de dados usadas.

Para aplicar o AHP com *ratings*, os próprios autores e pesquisadores atuarão como especialistas para estabelecer as comparações pareadas usando tanto a escala relativa como a absoluta, necessárias quando utilizado o AHP com *ratings*. Como resultado será obtido um peso ou importância para cada artigo.

- **Formação da base de dados final:** Com esse ranqueamento finalmente poderá se decidir quais artigos vão compor a base de dados final que se usará como referência para a pesquisa em questão.

A última etapa proposta no procedimento sistemático de revisão da literatura é a avaliação qualitativa dos artigos contidos na base de dados final. A mesma consiste em:

- **Avaliação qualitativa:** Os artigos contidos na base de dados final serão usados para construir a teoria. Será realizado uma leitura dos documentos e identificação de pontos comuns entre eles, a partir do tema considerado.

Até o momento, segundo o estado da pesquisa, os achados nos artigos pesquisados mostram que ambas as ferramentas, modelos de filas e simulação computacional, se mostram pertinentes para resolver problemas logísticos e avaliar cenários para melhoria de processos com gargalos (FELIX; PEIXOTO; EDWIGES, 2023; DA SILVA et al. 2021). Vantagens e desvantagens po-

dem ser analisadas para casos específicos e estabelecer modelos que permitam aproveitar as vantagens de cada ferramenta por separado ou implementar o uso de forma conjunta em projetos de pesquisa, mostra-se um assunto pertinente para um estudo mais detalhado.

## CONCLUSÕES

Diferentes propostas de revisão sistemática foram analisadas, algumas com menos etapas do que outras, mas contendo os mesmos elementos. Com base na proposta metodológica adotada neste estudo, foi realizada uma revisão dos temas de teoria das filas e simulação com aplicações na logística. A proposta do estudo estabelece a realização de três etapas, das quais somente foi concluída totalmente a primeira delas, a revisão bibliográfica. As etapas de ranqueamento e avaliação qualitativa estão em fase de execução.

O mapeamento inicial permitiu a construção da árvore de palavras-chave destacando as mais frequentes relacionadas ao tema. A estruturação dos termos com lógica booleana possibilitou a formulação de consultas alinhadas com os critérios de pesquisa das bases de dados. A aplicação de filtros foi importante para garantir a qualidade dos dados, eliminando informações desatualizadas e irrelevantes.

Como resultado inicial obteve-se uma base formada por 32 artigos que serão estudados e organizados segundo as etapas da proposta metodológica. Com o aprofundamento no estudo serão identificadas lacunas do conhecimento e a pesquisa será dirigida a um setor importante da economia, o portuário. Isto permitirá dirigir o estudo para uma proposta sólida de tema de pesquisa que será usada como tema de dissertação.

## REFERÊNCIAS

- ALVES, A.C.B.; MENEZES, M.A.F. Introdução à Pesquisa Operacional. 1 ed. – Goiânia: PUC Goiás, 2010.
- BALLOU, R. Logística empresarial. Transportes. Administração de materiais. Distribuição física. São Paulo: Atlas, 2012.
- BASTOS M. H. R. Um estudo sobre o desenvolvimento do perfil empreendedor em estudantes de Instituições de Ensino Superior da cidade de Volta Redonda. Dissertação de mestrado – UFF, 2015.
- BIOLCHINI, J. et al. Systematic review in software engineering. System engineering and computer science department COPPE/UFRJ, Technical Report ES, v. 679, n. 05, p. 45, 2005.
- BRONSON, R. Pesquisa operacional. São Paulo: McGraw-Hill do Brasil, 1985.
- CAMELO, G. R.; COELHO, A. S.; BORGES, R. M.; DE SOUZA, R. M. Teoria das filas e da simulação aplicada ao embarque de minério de ferro e manganês no terminal marítimo de ponta da madeira. Cadernos do IME-Série Estatística, v. 29, n. 2, p. 1, 2010.
- CASTRO NETO, L. R. Modelagem e simulação da cadeia produtiva do minério de ferro. 2006. Tese de Doutorado. Universidade de São Paulo.
- DA SILVA, M. R. F. et al. Proposta de otimização dos processos logísticos do porto de Itajaí utilizando simulação e indicadores da teoria das filas. Revista Produção Online, v. 21, n. 1, p. 231-258, 2021.
- DE BRUNS, R.; SONCIM, S. P.; DE SINAY-PHD, M. C. F. Pesquisa operacional: uma aplicação da teoria das filas a um sistema de atendimento. 2001.
- DE OLIVEIRA, G. T. A relação dos indicadores de desempenho da logística portuária com os indicadores de desempenho da logística internacional. Revista Eletrônica de Estratégia & Negócios, v. 11, 2018.
- FELIX, J. C.; PEIXOTO, C. S.; EDWIGES, C. A. O planejamento de suprimentos utilizando-se a teoria de filas e simulação computacional em uma oficina de recuperação de componentes de vagões. Exacta, v. 21, n. 2, 2023.
- HILLIER, F. S.; LIEBERMAN, G. J. Introdução à pesquisa operacional. McGraw Hill Brasil, 2013.
- MACHADO, L. K. C.; DOS SANTOS, A. C. Índice de Desempenho Logístico (LPI): Uma análise da eficiência logística e da importância relativa dos seus indicadores. Ciências da Administração, v. 23, n. 60, p. 53-72, 2021.
- ROTHSTEIN, H.R.; HOPEWELL, S. Grey Literature. In The Handbook of Research Synthesis and Meta-Analysis. 2. ed. edited by Cooper H.; Hedges, L.V.; Valentine, J.C., p. 103–125. New York: Russell Sage Foundation, 2009.
- SAATY, T. L The Analytic Hierarchy Process. New York: McGraw Hill, 1980.
- TAHA, Hamdy A. Pesquisa operacional. Pearson Educación, 2008.
- THOMÉ, A.M.T; SCAVARDA, L.F.; SCAVARDA, A.J. Conducting systematic literature review in operations management, Production Planning & Control, Londres, v. 27, n. 5, p. 408-420, jan. 2016.
- WIESNER, A. K. Modelagem e simulação de uma solução de integração para identificação de gargalos de desempenho baseadas em formalismo matemático: uma abordagem orientada à Teoria das Filas. 2016.

## DADOS DOS AUTORES

**Kelly Cristine Franco de Azevedo** (kellyfranco@id.uff.br). Engenheira de Produção (Universidade Federal Fluminense), mestranda em Modelagem Computacional (Universidade Federal Fluminense)

**Cecília Toledo Hernández** (ctoledo@id.uff.br). Doutora em Engenharia Mecânica (Universidade Estadual de São Paulo). Professora Associada do Departamento de Engenharia de Produção da Universidade Federal Fluminense.

**Eliane da Silva Christo** (elianechristo@id.uff.br). Doutora em Engenharia Elétrica (Pontifícia Universidade Católica do Rio de Janeiro). Professora Associada do Departamento de Engenharia de Produção da Universidade Federal Fluminense.

## PROBLEMAS DIFUSIVOS-CONVECTIVOS ABORDADOS PELO MÉTODO DAS DIFERENÇAS FINITAS CENTRADAS E *UPWIND*

### DIFFUSIVE-CONVECTIVE PROBLEMS ADDRESSED BY THE CENTERED-DIFFERENCE AND UPWIND METHOD

Victor de Souza Coelho Freitas

Gustavo Benitez Alvarez

Cleyton Senior Stampa

Panters Rodríguez Bermúdez

**Resumo:** O presente trabalho tem como objetivo abordar numericamente problemas governados pela equação de transporte difusiva-conectiva em regime permanente. Mais especificamente, foram estudados cenários bidimensionais nos quais há o predomínio do fenômeno da convecção, uma vez que desta forma tem-se um problema de perturbação singular. Para a abordagem numérica foi escolhido o método de diferenças finitas centradas e o método de diferenças finitas *upwind*. Foram implementados problemas contendo ou não camadas-limite em suas soluções. Ambos os métodos foram avaliados, com seus resultados comparados para determinar suas vantagens e desvantagens no contexto deste estudo.

**Palavras-chave:** Diferenças finitas centradas. Diferenças finitas *upwind*. Problema difusivo-conectivo. Oscilações espúrias.

**Abstract:** The present work aims to numerically address problems governed by the discovery of steady-state diffusive-convective transport. More specifically, two-dimensional scenarios were trained in which the specific convection predominates, since this creates a unique disturbance problem. For the numerical approach, the centered finite difference method and the upwind finite difference method were chosen. Problems containing or not limited layers were posed in their solutions. Both methods were evaluated, with their results compared to determine their advantages and advantages in the context of this study.

**Keywords:** Centered finite differences. Upwind finite differences. Diffusive-convective problem. Spurious oscillations.

## INTRODUÇÃO

Problemas difusivos-convectivos dizem respeito à cenários nos quais tem-se um fenômeno físico no qual partículas, energia, ou outra quantidade física é transportada dentro de um domínio por meio de dois processos acoplados: difusão e convecção. A difusão é o mecanismo de transporte no qual a quantidade do elemento físico de interesse se desloca de um meio mais concentrado para o meio menos concentrado. Já a convecção se dá pela translação desse elemento através do domínio no qual ele está presente.

Para casos nos quais tem-se dominância do fenômeno convectivo, é observado o surgimento de oscilações espúrias na solução numérica, não sendo representativas do fenômeno físico estudado. Sendo este um fenômeno físico governado por uma Equação Diferencial Parcial (EDP), não temos para todos os casos uma solução analítica explícita. Assim sendo, é preciso abordar o problema numericamente, de forma que se possa aproximar a solução dentro de um domínio discreto, tanto para o caso unidimensional, quanto para o caso bidimensional.

O interesse generalizado em entender e obter soluções para esta classe de problemas está, além do seu significado físico, na constatação que esta equação modelo relativamente simples pode estar presente em outros problemas mais complexos, envolvendo, por exemplo, aplicações na mecânica dos fluidos. Desde meados dos anos 1980, diversos métodos foram propostos com o intuito de eliminar, ou ao menos reduzir tais oscilações sem demaziado *smearing*, ou suavização, das camadas presentes. Neste sentido, o método de diferenças finitas *upwind* se apresenta como uma alternativa, uma vez que devido sua formulação com derivadas avante ou atrasadas, promove maior estabilização em regiões de alto gradiente.

Por outro lado, a metodologia das diferenças finitas é um dos mais bem estabelecidos métodos de solução numérica para equações diferenciais, no qual as derivadas em uma dada equação são substituídas por diferenças aproximadas, resultando assim em um sistema algébrico, que pode ser resolvido para encontrar a solução numérica.

Iniciando com a Equação 1 abaixo, que representa a forma geral da equação difusiva-convectiva, na qual  $U$  é a quantidade física de interesse. Sendo  $\vec{w}$  um campo vetorial de velocidade e  $f$  representa uma fonte ou sumidouro dentro do domínio estabelecido. Para esta equação, temos que a

solução é uma função  $U(x, y, t)$  que varia de acordo com o tempo, dentro de um domínio  $\Omega \subset \mathbb{R}^2$ . Com  $\Omega$  sendo aberto e delimitado, cujo limite  $\Gamma$  é suave e definido.

$$\frac{\partial U}{\partial t} + \vec{w} \cdot \nabla U = \nabla \cdot (\nu \nabla) U + f \#1$$

Considerando que  $U$  não varia com o tempo, ou seja, se encontra em regime permanente e depende apenas de suas variáveis espaciais, tem-se que  $\frac{\partial U}{\partial t} = 0$ . Assim, pode-se rescrever a Equação 1 conforme demonstrado pela Equação 2 abaixo, na qual  $k$  denota o coeficiente difusivo. Se caracterizando como uma equação diferencial parcial (EDP) elíptica.

$$-k \Delta U + \vec{w} \cdot \nabla U = f \#2$$

Sabendo que o número de Péclet ( $P_e$ ) é dado como a razão entre o coeficiente convectivo e o coeficiente difusivo, pode-se definir  $\epsilon_n$  como o inverso de  $P_e$ , com  $\vec{w} = (a_1, a_2)$ , tal que  $\epsilon_n = \frac{k}{a_n}$ , para  $n = 1, 2$ . É esperado que, para casos de convecção dominante, ou seja,  $\epsilon_n \rightarrow 0$ , a Equação 2 tende a se reduzir a uma equação diferencial parcial de primeira ordem, que só permite uma condição de contorno.

Entretanto, como  $\epsilon_n > 0$ , ainda que muito pequeno, tem-se uma equação de segunda ordem que necessita de duas condições de contorno. Logo, é esperado um panorama de solução anormal próximo às condições de contorno, já que para  $\epsilon_n \rightarrow 0$  o problema está super-especificado.

Conforme discutido por Carmo e Benitez (2003), assim como Volker e Knobloch (2007), este panorama anormal é o que se chamam as oscilações espúrias, uma vez que a solução da equação difusiva-convectiva possui camadas-limite, que são pequenas regiões do domínio que possuem alto gradiente. Haja vista que o tamanho das malhas normalmente utilizadas em soluções numéricas é significativamente maior que a espessura da camada-limite, essa região não é resolvida de forma adequada. Dando assim, origem à tais oscilações que não são físicas, ou seja, não representam a solução do problema posto, mas sim instabilidades no método numérico.

Como discutido por Zhao (2018), problemas convectivos-difusivos, nos quais existe dominância da convecção, apresentam para os métodos clássicos de resolução, como o método das diferenças finitas (MDF) e elementos finitos, instabilidades em seu panorama de solução. Como mostrado por Frerichs e John (2021), pode-se contornar esse problema com artifícios para melhorar a estabilidade, como o método *upwind* para a metodologia de diferenças finitas. Neste caso, é adicionado uma difusividade numérica na solução, proveniente da formulação matemática do método. Entretanto, esse procedimento tende a prejudicar a precisão dos resultados obtidos, com uma difusão excessiva em alguns casos. Atualmente, não se tem uma formulação que seja ao mesmo tempo estável e precisa, para soluções com a presença de camadas-limite.

### Método das diferenças finitas centradas e *upwind*

De acordo com Tanehill, Anderson e Pletcher (2012), a abordagem do método das diferenças finitas se caracteriza como a discretização de um problema inicialmente contínuo, de forma que as variáveis de interesse existam unicamente em pontos discretos. Por essa metodologia, as derivadas presentes são substituídas por diferenças entre tais pontos, logo tem-se uma representação algébrica de uma EDP, na qual tem-se ao final do processo um sistema linear que pode ser facilmente resolvido.

Partindo da Equação 3 abaixo, como mostrado por Maliska (2017) e Pantakar (2018), pode-se substituir as derivadas de segunda ordem (termo difusivo) pela formulação centrada de 2<sup>a</sup> ordem. Já as derivadas de primeira ordem (termo convectivo) são substituídas pela formulação centradas de 2<sup>a</sup> ordem, obtendo-se assim a discretização pelo método das diferenças centradas, considerando um campo de velocidades  $\vec{w} = (a_1, a_2)$ , tem-se, por fim, a Equação 4.

$$\begin{aligned}
 & -k \left( \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \right) + \vec{w} \cdot \left( \frac{\partial U}{\partial x} \hat{i} + \frac{\partial U}{\partial y} \hat{j} \right) = f \quad 3 \\
 & -k \left[ \left( \frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{\Delta x^2} \right) + \left( \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{\Delta y^2} \right) \right] + \dots \#4 \\
 & \dots + \left[ \left( a_1 \frac{U_{i+1,j} - U_{i-1,j}}{2\Delta x} \right) + \left( a_2 \frac{U_{i,j+1} - U_{i,j-1}}{2\Delta y} \right) \right] = f
 \end{aligned}$$

Organizando a Equação 4 acima, concatenando os termos de mesmos índices:

$$\begin{aligned}
 & \left( \frac{-k}{\Delta y^2} - \frac{a_2}{2\Delta y} \right) U_{i,j-1} + \left( \frac{-k}{\Delta x^2} - \frac{a_1}{2\Delta x} \right) U_{i-1,j} + \dots \\
 & \dots + \left( \frac{2k}{\Delta x^2} + \frac{2k}{\Delta y^2} \right) U_{i,j} + \dots \#5 \\
 & \dots + \left( \frac{-k}{\Delta x^2} + \frac{a_1}{2\Delta x} \right) U_{i+1,j} + \left( \frac{-k}{\Delta y^2} + \frac{a_2}{2\Delta y} \right) U_{i,j+1} = f
 \end{aligned}$$

Já para o MDF *upwind*, como mostrado por Thomas (2013), a discretização do termo convectivo irá depender da direção do campo de velocidade. Para um valor de velocidade positivo, utiliza-se a formulação atrasada de 1<sup>a</sup> ordem; já para um valor de velocidade negativo, utiliza-se a formulação avançada de 1<sup>a</sup> ordem. De maneira que a formulação *upwind* para velocidades positivas e negativas estão representadas, respectivamente, pelas Equações 6 e 7 abaixo.

$$-k \left[ \left( \frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{\Delta x^2} \right) + \left( \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{\Delta y^2} \right) \right] + \dots \#6$$

$$\dots + \left[ \left( a_1 \frac{U_{i,j} - U_{i-1,j}}{\Delta x} \right) + \left( a_2 \frac{U_{i,j} - U_{i,j-1}}{\Delta y} \right) \right] = f$$

$$-k \left[ \left( \frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{\Delta x^2} \right) + \left( \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{\Delta y^2} \right) \right] + \dots \#7$$

$$\dots + \left[ \left( a_1 \frac{U_{i+1,j} - U_{i,j}}{\Delta x} \right) + \left( a_2 \frac{U_{i,j+1} - U_{i,j}}{\Delta y} \right) \right] = f$$

Para ambos métodos, após realizada a discretização apropriada, concatena-se os termos de mesmo índice de maneira que se tenha os coeficientes  $A, B, C, D, E, A, B, C, D, E$  na equação matricial  $MU = f$   $MU = f$ , conforme mostrada pela Equação 8 abaixo, sendo esta a equação a partir da qual foram construídos os algoritmos computacionais.

$$AU_{i,j-1} + BU_{i-1,j} + CU_{i,j} + DU_{i+1,j} + EU_{i,j+1} = f \#8$$

Na qual  $M$  é a matriz dos coeficientes resultantes da formulação de diferenças finitas utilizada, sendo uma matriz pentadiagonal esparsa de tamanho  $n^2 \times n^2$ , correspondendo à uma malha com  $n$  nós em cada eixo. Especificamente para a metodologia das diferenças centradas, temos os coeficientes mostrados abaixo pelas Equações 9 a 13.

$$A = \left( \frac{-k}{\Delta y^2} - \frac{a_2}{2\Delta y} \right) \#9$$

$$B = \left( \frac{-k}{\Delta x^2} - \frac{a_1}{2\Delta x} \right) \#10$$

$$C = \left( \frac{2k}{\Delta x^2} + \frac{2k}{\Delta y^2} \right) \#11$$

$$D = \left( \frac{-k}{\Delta x^2} + \frac{a_1}{2\Delta x} \right) \#12$$

$$E = \left( \frac{-k}{\Delta y^2} + \frac{a_2}{2\Delta y} \right) \#13$$

## METODOLOGIA

Para o contexto da equação difusiva-convectiva aplicada a um domínio  $\Omega \subset \mathbb{R}^2$ , não se obteve, para este trabalho, soluções exatas para que se possa validar os códigos criados. Nesse caso optou-se por utilizar casos clássicos da literatura. Assim, os panoramas das soluções são conhecidos de forma que é possível avaliar os resultados obtidos por meio dos códigos criados, para o método das diferenças finitas centradas e *upwind*.

Os casos de estudo implementados têm a finalidade de promover situações de convecção dominante, de forma que se possa constatar a presença, ou não, de oscilações espúrias (instabilidades). Os problemas escolhidos se dividem em dois grupos: os que possuem camada-limite em sua solução e os que possuem solução suave. Com os resultados obtidos, será possível constatar qual método forneceu os melhores resultados para cada um dos cenários propostos.



## Resultados numéricos

Para o primeiro problema estudado foi escolhido um cenário no qual não há camadas-limite na solução, de forma que para este foi definido um domínio  $(-0.5 < x < 0.5, -0.5 < y < 0.5)$ , o campo de velocidades é definido como  $\vec{w} = (-y, x)$ , a condição de contorno é zero ao redor de todo o domínio, já o coeficiente difusivo é dado por  $k = 10^{-10}$ . Para este estudo foi também imposta a condição interna dada pela função seno  $f(0, y) = \sin(2\pi y) \forall y \in [-0.5, 0]$ , como mostrado na Figura 1 abaixo. Por fim, foram utilizadas as malhas quadradas de tamanhos  $n_x = n_y = 20$  e  $n_x = n_y = 60$ .

Figura 1: Domínio do primeiro problema. Fonte: Adaptado de Carmo e Benitez (2003)

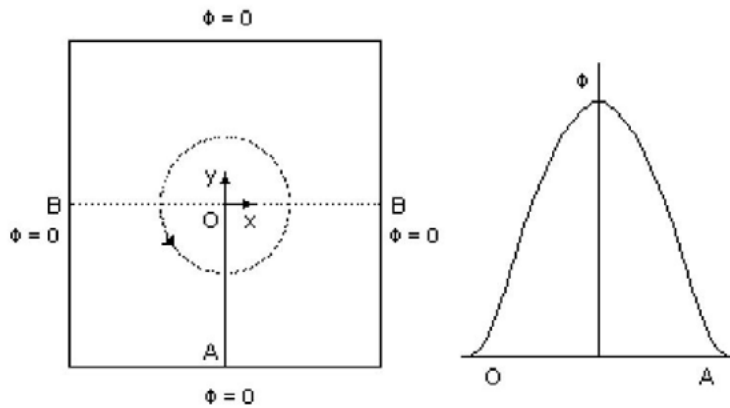


Figura 2: Solução do primeiro problema para  $n_x = n_y = 20$

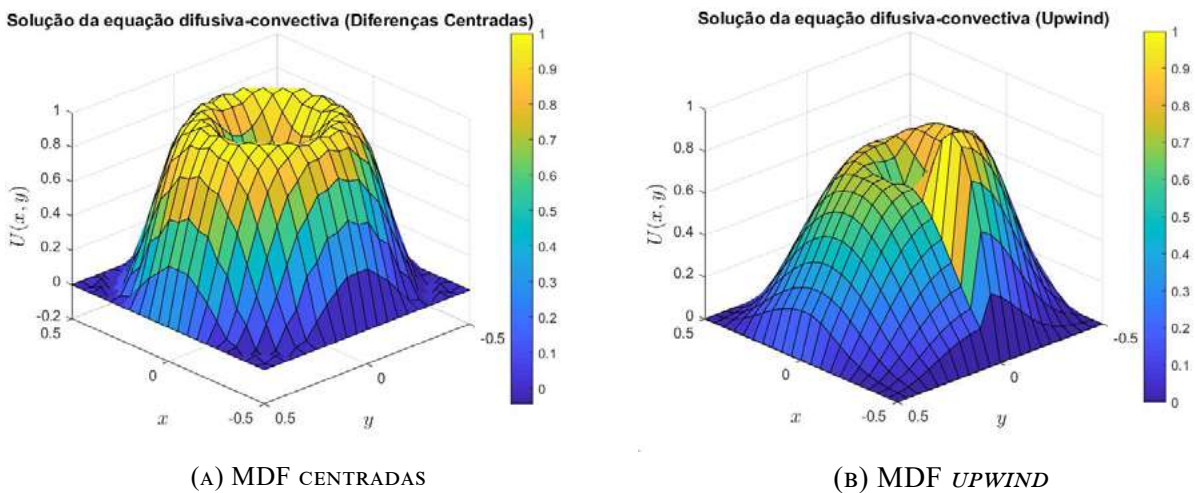
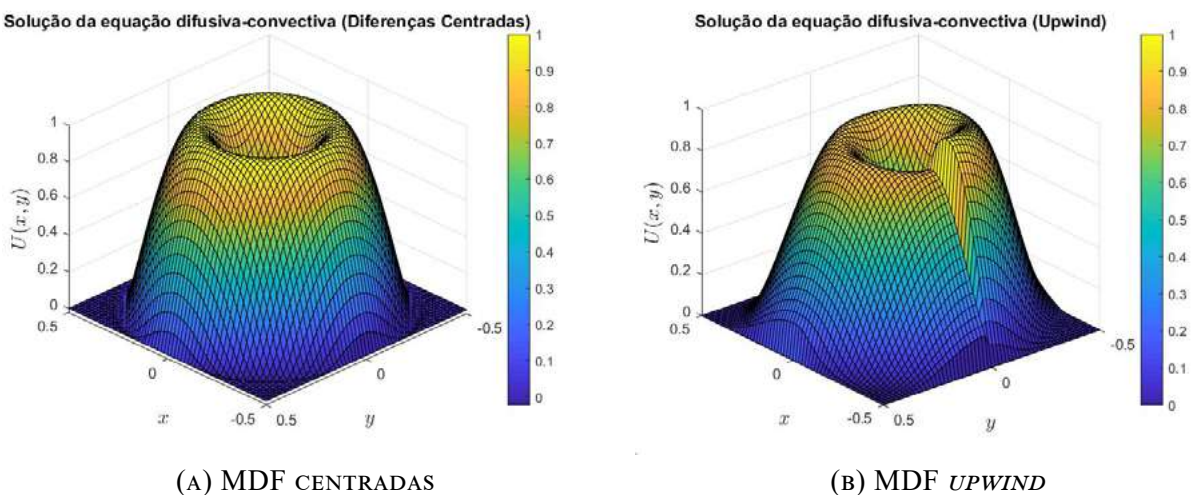


Figura 3: Solução do primeiro problema para  $n_x = n_y = 60$



Para este problema pode-se perceber que as soluções apresentadas pelo MDF *upwind* apresentam grande difusão numérica introduzida pelo método, constatada pelo degrau próximo à condição interna imposta. Analisando as Figuras 2 e 3 é visto que conforme a malha é refinada, a diferença entre os valores da solução próximo aos pontos com  $y \in [-0.5, 0]$  é reduzida, mas não é eliminada.

Em contraste, as soluções obtidas com o MDF centradas apresentaram um panorama coerente com o esperado, sem a presença de tal degrau. Isso se deve à maior precisão da formulação centrada, e à tendência do MDF *upwind* de suavizar as soluções, que neste caso resultou em uma diferença significativa em relação ao que era esperado.

Para o segundo problema, as condições de contorno foram consideradas zero, já o termo fonte foi definido como a função Gaussiana representada pela Equação 14 abaixo. O campo de velocidades considerado será variável dentro do domínio, de maneira que  $\vec{w} = (y, -x)$ . Já o coeficiente difusivo será  $k = 10^{-1}$ ,  $k = 10^{-3}$  e  $k = 10^{-6}$ , com uma malha  $n_x = n_y = 60$ .

$$f(x, y) = \frac{1}{0.02\pi} e^{-\frac{1}{0.02}[(x-0.5)^2 + (y-0.5)^2]} \quad \#14$$

Figura 4: Solução do segundo problema para  $k = 10^{-1}$ .

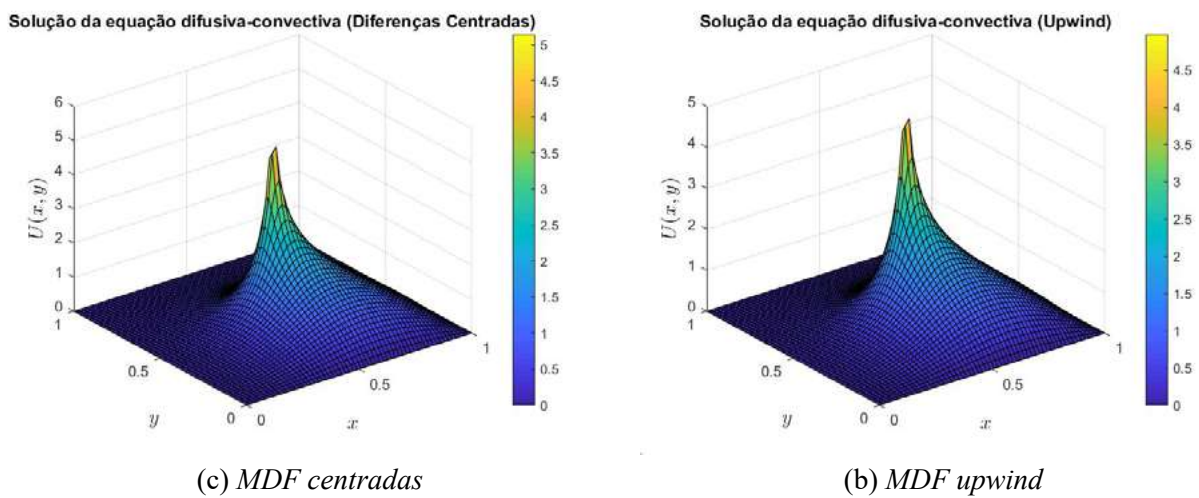


Figura 5: Solução do segundo problema para  $k = 10^{-3}$ .

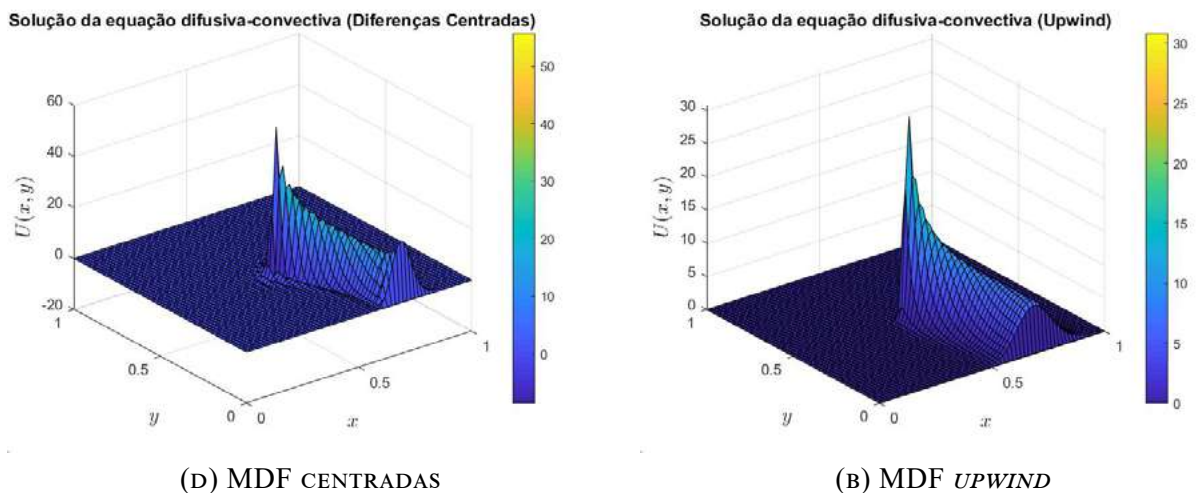
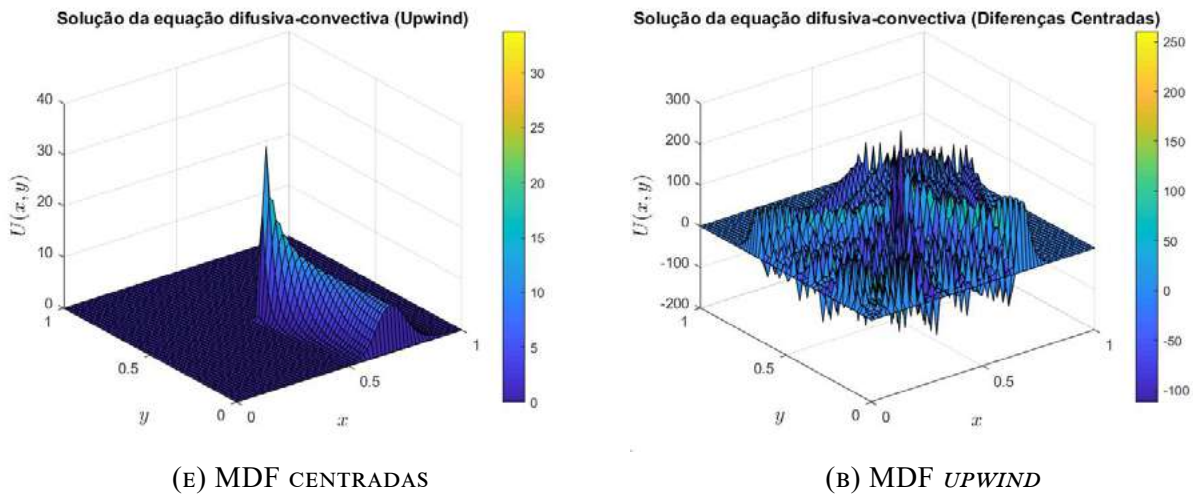


Figura 6: Solução do segundo problema para  $k = 10^{-6}$ .



Observa-se pelas Figuras 4 a 6 acima, que conforme o coeficiente difusivo é reduzido, o panorama da solução se torna dominado pelo fenômeno convectivo e desta forma tem-se uma concentração maior de  $U$  na direção do fluxo de velocidade circular imposto. Nota-se também que para  $k = 10^{-3}$  o MDF centradas apresenta instabilidade na forma de oscilações espúrias na base da função Gaussiana e próximo à região de condição de contorno, já para  $k = 10^{-6}$  o panorama está povoado de oscilações, sendo coerente com esperado de acordo com a literatura, uma vez que  $P_g \gg 1$ .

O terceiro e último problema analisado contém as condições de contorno dada pelas Equações 15 a 22. O coeficiente difusivo considerado é  $k = 10^{-2}$ , já o campo de velocidade adotado é  $\vec{w} = (y, -x)$ . Por fim, as malhas consideradas são  $n_x = n_y = 20$  e  $n_x = n_y = 60$ , para que se possa observar como a solução se altera com o maior número de pontos discretos. Abaixo podem ser vistos as soluções numéricas obtidas, representas pelas Figuras 7, 8 e 9.

$$\begin{aligned}
 U(x, 0) &= 0 \quad \forall x \in [0,1] \#15 \\
 U(x, 1) &= 1 \quad \forall x \in [0,1] \#16 \\
 U(1, y) &= 0 \quad \forall x \in [0,1] \#17 \\
 U(0, y) &= 0 \quad \forall x \in [0,0.6] \#18 \\
 U(0, y) &= y - 0.6 \quad \forall x \in [0.6,0.65] \#19 \\
 U(0, y) &= 18(y - 0.65) + 0.05 \quad \forall x \in [0.65,0.70] \#20 \\
 U(0, y) &= (y - 0.70) + 0.95 \#21 \\
 U(0, y) &= 1 \quad \forall x \in [0.75,1] \#22
 \end{aligned}$$

Figura 7: Solução do terceiro problema para  $n_x = n_y = 20$

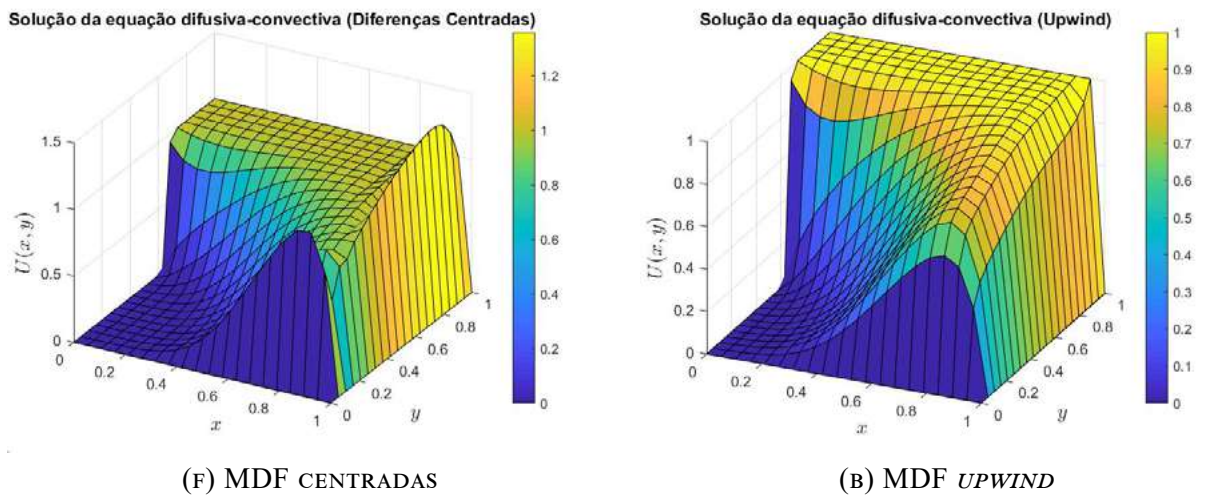


Figura 8: Solução do terceiro problema para  $n_x = n_y = 40$

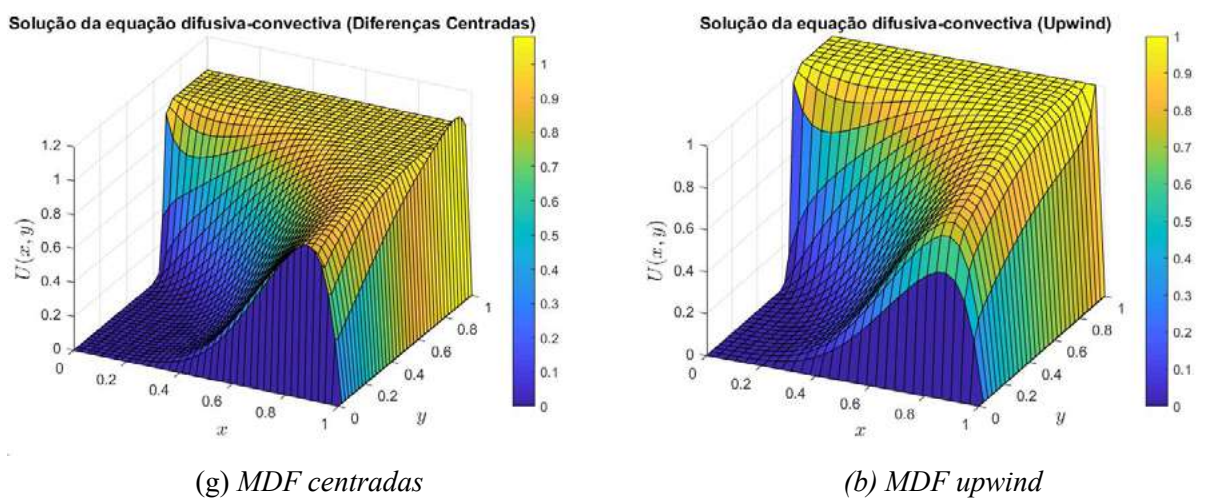
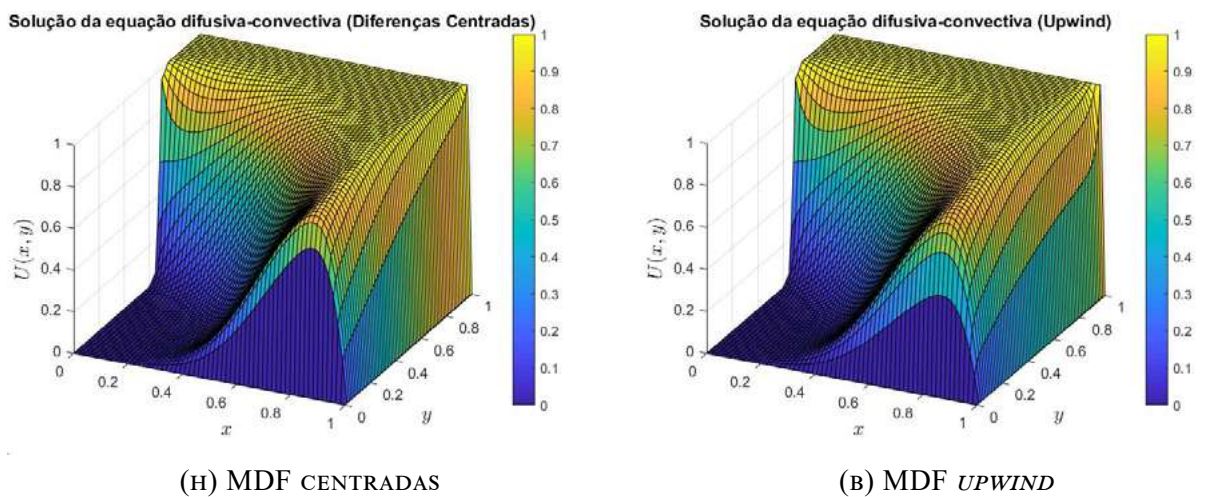


Figura 9: Solução do terceiro problema para  $n_x = n_y = 60$



Considerando os panoramas mostrados, percebe-se que conforme a malha se torna mais densa em relação ao número de nós, a solução apresentada pelo MDF centradas se torna mais suave, com a solução apresentada por  $n_x = n_y = 60$  se mostrando livre de instabilidades, com um panorama similar ao apresentado pelo MDF *upwind*. Ainda, pode se notar que a solução apresentada nesse caso pelo MDF *upwind* indica maior difusão numérica, evidenciada pela suavização da região adjacente aos pontos próximos à condição de contorno  $x \in [0.6, 1]$ , em comparação ao observado com a formulação centrada.

## CONCLUSÕES

Para o primeiro problema, com solução sem camadas-limite, a formulação centrada é mais precisa que o *upwind*, se apresentando livre de instabilidades, sendo esta uma constatação coerente com o esperado uma vez que o MDF centradas é uma metodologia de segunda ordem de precisão, e o *upwind* é de primeira ordem de precisão, tendo esta última apresentado excessiva suavização da solução.

Para o segundo e terceiro problema, que possuem camadas-limite em suas soluções, a formulação centrada apresenta oscilações espúrias, resultantes da instabilidade que surge quando o método é submetido à um cenário de convecção dominante. Neste contexto, percebe-se que o refino da malha tende à reduzir o número de oscilações percebidas. Entretanto, o aumento do número de nós na malha, acarreta o aumento do uso de memória e tempo de processamento, recursos estes escassos.

Ainda, para o segundo problema, mesmo nas malhas mais refinadas produzidas neste estudo foram observadas oscilações espúrias quando o coeficiente difusivo utilizado é inferior à  $10^{-3}$ , para o MDF centradas. Neste contexto de convecção dominante, a formulação *upwind* se mostrou estável, uma vez que não apresentou oscilações espúrias para os problemas contendo camada-limite interna ou externa. Por outro lado, por ser um método estabilizado, o MDF *upwind* induziu excessiva suavização para os problemas aqui abordados. Prejudicando assim, a precisão da solução obtida, para soluções com ou sem camada-limite.

Por fim, constata-se que para a obtenção de soluções estáveis, livres de oscilações, para problemas difusivos-convectivos nos quais existe dominância da convecção, o MDF *upwind* apresenta melhor desempenho quando comparado à formulação centrada. Entretanto, deve-se levar em consideração a difusão numérica introduzida na solução.

## REFERÊNCIAS BIBLIOGRÁFICAS

- DO CARMO, Eduardo Gomes Dutra; ALVAREZ, Gustavo Benitez. A new stabilized finite element formulation for scalar convection–diffusion problems: the streamline and approximate upwind/Petrov–Galerkin method. **Computer methods in applied mechanics and engineering**, v. 192, n. 31-32, p. 3379-3396, 2003.
- FRERICHS, Derk; JOHN, Volker. On reducing spurious oscillations in discontinuous Galerkin (DG) methods for steady-state convection–diffusion equations. *Journal of Computational and Applied Mathematics*, v. 393, p. 113487, 2021.
- LEVEQUE, Randall J. **Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems**. Society for Industrial and Applied Mathematics, 2007.
- MALISKA, Clovis Raimundo. **Transferência de calor e mecânica dos fluidos computacional**. Grupo Gen-LTC, 2017.
- PATANKAR, Suhas. **Numerical heat transfer and fluid flow**. Taylor & Francis, 2018.
- PLETCHER, Richard H.; TANNEHILL, John C.; ANDERSON, Dale. **Computational fluid mechanics and heat transfer**. CRC press, 2012.
- THOMAS, James William. **Numerical partial differential equations: finite difference methods**. Springer Science & Business Media, 2013.
- JOHN, Volker; KNOBLOCH, Petr. On spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part I–A review. **Computer methods in applied mechanics and engineering**, v. 196, n. 17-20, p. 2197-2215, 2007.
- ZHAO, Shubo et al. Two types of spurious oscillations at layers diminishing methods for convection–diffusion–reaction equations on surface. **Numerical Heat Transfer, Part A: Applications**, v. 74, n. 7, p. 1387-1404, 2018.

## DADOS DOS AUTORES

**Victor de Souza Coelho Freitas**. Mestre em Modelagem Computacional, Aluno do Programa de Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia, Escola de Engenharia Industrial Metalúrgica de Volta Redonda, UFF, victorcoelho@id.uff.br

**Gustavo Benitez Alvarez**. Doutor em Engenharia Nuclear, Docente do Programa de Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia, Escola de Engenharia Industrial Metalúrgica de Volta Redonda, UFF, gustavobenitez@id.uf.br

**Cleyton Senior Stampa**. Doutor em Engenharia Mecânica, Docente do Programa de Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia, Escola de Engenharia Industrial Metalúrgica de Volta Redonda, UFF, cleytons@id.uf.br

**Panters Rodríguez Bermúdez**. Doutor em Matemática, Docente do Programa de Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia, Escola de Engenharia Industrial Metalúrgica de Volta Redonda, UFF, pantersrb@id.uf.br

# IMPLEMENTAÇÃO EM VOLUMES FINITOS DE UM MÉTODO TIPO GODUNOV PARA SISTEMAS 2 X 2 DE LEIS DE CONSERVAÇÃO

## *FINITE VOLUME IMPLEMENTATION OF A GODUNOV'S TYPE METHOD FOR 2 X 2 SYSTEMS OF CONSERVATION LAWS*

Eron Oliveira do Nascimento

Panters Rodríguez Bermúdez

Wagner Rambaldi Telles

**RESUMO:** No presente trabalho apresenta-se uma implementação computacional do método de Godunov, o qual utiliza uma discretização em volumes finitos em 1D, e é direcionado para a resolução de sistemas 2 X 2 de equações diferenciais parciais hiperbólicas denominadas leis de conservação. O desenvolvimento é feito a partir de um processo de linearização para obter-se o valor do fluxo para os volumes, em conjunto, de um *Riemann solver* aproximado e condições de correção de entropia. A validação do código foi realizada tanto para as equações de águas rasas como para um modelo de escoamento horizontal e vertical de fluidos imiscíveis em meios porosos, para isto foram utilizados resultados analíticos e numéricos encontrados em distintos artigos científicos de impacto internacional. Neste artigo, mostramos casos de escoamento trifásico horizontal (sem efeitos gravitacionais) em meios porosos utilizando exemplos tomados de Azevedo et al. (2010); além de casos de escoamento trifásico vertical (considerando efeitos gravitacionais), cujas soluções numéricas obtidas podem ser consideradas as primeiras aproximações numéricas das soluções exatas semi-analíticas, encontradas e descritas qualitativamente em Rodríguez-Bermúdez (2010) e Rodríguez-Bermúdez e Marchesin (2013).

**Palavras-chave:** Volumes finitos. Método de Godunov. Leis de conservação. Implementação computacional.

**ABSTRACT:** In this work, we present a computational implementation of Godunov's method, which utilizes a finite volume discretization in 1D, aiming to solve 2 X 2 systems of hyperbolic partial differential equations called conservation laws. The development started with a linearization process to obtain the flux's value to which volume, combining an approximate Riemann solver and entropy correction conditions. The code validation was performed for shallow water equations and a model for immiscible fluid flows in horizontal/vertical porous media, utilizing analytical and numerical results from papers with international impact. In this work, we only show cases for horizontal tri-phase flow (without gravitational effects) with results taken from Azevedo et al. (2010) and cases for vertical tri-phase flow (with gravitational effects), which numerical solutions obtained considered the first approximations of semi-analytical results found and described qualitatively in Rodríguez-Bermúdez (2010) and Rodríguez-Bermúdez e Marchesin (2013).

**Keywords:** Finite volume. Godunov's method. Conservation laws. Computational implementation.

## INTRODUÇÃO

Sistemas de leis de conservação estão presentes em diversos modelos de problemas da engenharia e fenômenos naturais. Estas são um conjunto de equações diferenciais parciais (comumente não lineares) de primeira ordem de caráter hiperbólico e representam fenômenos de transporte de distintos tipos. Devido à não linearidade destas equações diferenciais, a resolução das mesmas costuma apresentar dificuldades matemáticas e computacionais que, em geral, limitam a compreensão e a obtenção das soluções destes problemas da engenharia. Uma destas dificuldades matemáticas presente, é o conhecido fenômeno de formação de ondas de choque (frentes descontínuas que se transportam a velocidade constante), ondas de rarefação, etc., que se produzem como consequência do cruzamento entre as curvas características do sistema hiperbólico. Para poder estudar tais fenômenos, o conceito de solução de EDP deve ser estendido (soluções generalizadas ou fracas). Porém, soluções fracas costumam não ser únicas, e critérios de entropia adicionais para selecionar as soluções fisicamente corretas precisam ser considerados.

Por outro lado, para sistemas de duas ou mais leis de conservação, a teoria existente de Lax (1973) e Liu (1976) sobre problemas de Riemann é muito mais complexa que para modelos escalares, o que dificulta enormemente a obtenção de soluções analíticas ou exatas para esses tipos de problemas. Complicações que surgem nos problemas físicos como heterogeneidades no meio, problemas multidimensionais no espaço (2D ou 3D), variações de condições como temperatura e pressão, efeitos de capilaridade ou interação entre partículas, entre outras, acrescentam as dificuldades enfrentadas. Deste modo, a implementação de métodos numéricos robustos e precisos, se faz necessário para obter as soluções numéricas de problemas de Riemann para leis de conservação.

Durante anos a comunidade científica tem dedicado esforços ao estudo, criação, aprimoramento de esquemas numéricos acurados e à implementação de códigos computacionais (simuladores) para resolver estes problemas. Segundo LeVeque (2004), existem basicamente três tipos de formulações para implementar esses métodos: diferenças finitas, volumes finitos, elementos finitos. Neste artigo, é apresentada a implementação de um método tipo Godunov, baseado em volumes finitos, para a solução de problemas unidimensionais de sistemas  $2 \times 2$  de leis de conservação. Utiliza-se a construção de matrizes lineares nas interfaces dos volumes, em substituição da matriz jacobiana das funções de fluxo. Como resultado, o problema é transformado em diversos sistemas locais, resolvidos pelo *Riemann solver* de Roe (1981). A condição de entropia para correção das soluções é definida pelo critério de Harten e Hyman (1983).

O código construído foi utilizado para resolver alguns problemas de escoamento trifásico em meios porosos, já previamente publicados na literatura, dadas as condições iniciais de Riemann. Os casos sem influência da gravidade foram baseados no artigo de Azevedo et al. (2010) e os casos com gravidade, baseados na tese de Rodríguez-Bermúdez (2010) e no artigo de Rodríguez-Bermúdez e Marchesin (2013). Os resultados consistem em perfis numéricos para as saturações das fases cuja descrição qualitativa é validada.

A perspectiva de construção do simulador visou a possibilidade de expandir o código para outros tipos de problemas e números de equações envolvidas, desde que modeladas por sistemas hiperbólicos.

## VOLUMES FINITOS

A equação diferencial da lei de conservação pode ser formulada através da discretização do domínio em volumes finitos e usando o princípio da conservação de massa. Segundo LeVeque (2004), para o caso unidimensional:

$$\frac{d}{dt} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} Q(x, t) dx = f \left( Q \left( x_{i-\frac{1}{2}}, t \right) \right) - f \left( Q \left( x_{i+\frac{1}{2}}, t \right) \right) \quad (1)$$

Onde  $Q$  é o vetor que agrupa as  $m$  variáveis conservadas e  $f(Q)$  o vetor das  $m$  funções de fluxo do modelo. Ou seja, a variação no tempo da quantidade  $Q$  no volume  $\Delta x$  é igual à diferença entre



os valores do fluxo nas interfaces de saída  $x_{i+\frac{1}{2}}^e$  e entrada  $x_{i-\frac{1}{2}}^e$ .

Reduzindo o intervalo  $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$  em uma escala infinitesimal e assumindo certas condições de diferenciabilidade, a relação (1) é simplificada para a forma diferencial, sendo:

$$Q_t + f(Q)_x = 0 \quad \#(2)$$

Em que os subíndices  $x, t$  indicam as derivadas parciais em relação ao espaço e ao tempo, respectivamente. A expressão (2) descreve um sistema de EDP hiperbólicas na sua forma conservativa.

Retornando à forma integral (1), em intervalos infinitesimais, a função  $Q(x, t)$  pode ser aproximada por um valor constante em cada volume  $\Delta x$ , tornando as soluções das integrais apenas o produto do valor de  $Q(x, t)$  pela variação temporal ( $\Delta t$ ) ou espacial ( $\Delta x$ ). Obtém-se assim uma expressão discretizada em que o cálculo do valor de cada  $Q_i$  no volume, na iteração de tempo seguinte  $n + 1$ , é dado apenas por termos da iteração atual  $n$ :

$$Q_i^{n+1} = Q_i^n - \frac{\Delta t}{\Delta x} \left( F_{i+\frac{1}{2}}^n - F_{i-\frac{1}{2}}^n \right) \quad \#(3)$$

Onde  $F_{i-\frac{1}{2}}^n$  e  $F_{i+\frac{1}{2}}^n$  são os valores que a função de fluxo assume nas interfaces. A expressão (3) é a base usada no código implementado.

## MÉTODO DE GODUNOV

Segundo apresenta Godunov e Bohachevsky (1959), o sistema (2) pode ser resolvido através de um algoritmo no qual para cada volume  $\Delta x$ : aproxima-se  $Q(x, t)$  por uma função polinomial, resolve o sistema para as condições delimitadas nas interfaces e calcula a média do valor obtido ao longo do próprio volume, processo esse chamado *REA*: *reconstruct-evolve-average*. Já LeVeque (2004) propõe uma adaptação deste conceito para calcular o valor dos fluxos  $F_{i-\frac{1}{2}}^n$  e  $F_{i+\frac{1}{2}}^n$  na expressão (3).

Fazendo a aproximação da função  $Q(x, t)$  por um valor constante, os diversos sistemas locais em cada volume agora têm como condição

inicial específica um problema de Riemann. Sendo a descontinuidade (interface) considerada a origem,  $x_{i+\frac{1}{2}} = 0$  para esse problema local, os valores constantes  $Q_E$  e  $Q_D$  dos volumes adjacentes correspondem à definição desta condição. Os valores  $F_{i-\frac{1}{2}}^n$  e  $F_{i+\frac{1}{2}}^n$  obtêm-se a partir da solução do problema de Riemann para as respectivas interfaces, e depois avaliando na função de fluxo; denomina-se fluxo de Godunov (numérico).

Para a implementação desenvolvida neste trabalho é usada a forma não conservativa (4), requerido pelo *Riemann solver* aplicado, em que reescreve-se (2) sendo  $A(Q)$  a matriz jacobiana do sistema hiperbólico.

$$Q_t + A(Q)Q_x = 0 \quad \#(4)$$

Dependendo do modelo estudado, a matriz  $A(Q)$  pode apresentar complexidades, dificultando a resolução direta do problema de Riemann. Para a implementação, entretanto, foi usado um solucionador numérico aproximado para o problema de Riemann (*Riemann solver*). De acordo com Roe (1981), a matriz  $AA$  pode ser reescrita com suas componentes sendo valores constantes, apenas necessário atender às condições: (i) preservar o caráter hiperbólico do sistema, (ii) ser consistente à matriz original e (iii) satisfazer a condição de choque de Rankine-Hugoniot. Isto transforma o sistema em um caso linear, com maior facilidade de solução.

Uma nova matriz  $\tilde{A}$  é definida em cada interface, em substituição de  $A(Q)$ . A definição de  $\tilde{A}$  é abrangente e apenas se faz necessário que atenda aos critérios acima mencionados. Como sugerido por Roe (1981) e Toro (2009), foi usado apenas a média simples entre os valores que a matriz assume para  $Q$  nas adjacências, facilitando a implementação e reduzindo o trabalho computacional:  $\tilde{A} = \frac{1}{2} [A(Q_E) + A(Q_D)]$ .

Assim, os sistemas resolvidos nas interfaces serão:

$$\begin{cases} Q_t + \tilde{A}Q_x = 0 \\ Q(x, 0) = \begin{cases} Q_E, & x < 0 \\ Q_D, & x > 0 \end{cases} \end{cases} \quad \#(5)$$

A solução dos sistemas (5), conforme desenvolve Toro (2009), é obtida pela expansão dos estados  $Q$  em função dos autovalores  $\lambda_j$  e autovetores  $\kappa^{(j)}$  de  $\hat{A}$  e avaliada para  $x = 0$  (interface). Aplicando à função de fluxo, a forma geral do fluxo de Godunov para um sistema de  $m$  equações é:

$$F_{i+\frac{1}{2}} = \frac{1}{2} \left[ f(Q_i) + f(Q_{i+1}) - \sum_{j=1}^m (\beta_j - \alpha_j) h(\lambda_j) \kappa^{(j)} \right] \#(6)$$

$$F_{i-\frac{1}{2}} = \frac{1}{2} \left[ f(Q_{i-1}) + f(Q_i) - \sum_{j=1}^m (\beta_j - \alpha_j) h(\lambda_j) \kappa^{(j)} \right] \#(7)$$

Onde  $\alpha_j$  e  $\beta_j$  são os coeficientes da combinação linear de  $Q_i$  e  $Q_{i+1}$ , respectivamente.

A utilização do *Riemann solver* de Roe (1981), introduz erros decorrentes da aproximação da matriz  $\hat{A}$  em linear e gera soluções apenas na forma de choques, devido à condição de Rankine-Hugoniot. O critério de correção de entropia de Harten e Hyman (1983) foi escolhido e aplicado segundo a definição da função (8):

$$h(\lambda_j) = \begin{cases} \delta_j, & \text{se } |\lambda_j| < \delta_j \\ |\lambda_j|, & \text{se } |\lambda_j| \geq \delta_j \end{cases} \#(8)$$

$$\delta_j = \max\{0, \lambda_j - \lambda_j(Q_E), \lambda_j(Q_D) - \lambda_j\} \#(9)$$

Sendo o fator  $\delta_j$  um parâmetro para mensurar o quanto a solução aproximada desvia da solução correta, adicionando um valor intermediário que suaviza os choques para formar rarefações corretas.

Tratando-se de um método explícito, está sujeito a um critério de estabilidade para a convergência da solução no código. Courant et al. (1967) define que para valores entre  $0 < v < 1$  o método é convergente e estável e este princípio foi usado no código para calcular o intervalo de tempo  $\Delta t$  de cada iteração, conforme (10).

$$\Delta t = \frac{v \Delta x}{\lambda_{max}^n} \#(10)$$

Nos sistemas em estudo, para cada matriz  $\hat{A}$  na interface, existem dois autovalores tais que  $\lambda^+ > \lambda^-$  indicando os grupos de onda *fast-family* e *slow-family*, respectivamente. A determinação do maior valor para cada iteração pode ser feita, segundo LeVeque (2004), pela escolha (11):

$$\lambda_{max}^n = \max_i \left( \left| \lambda_{i-\frac{1}{2}} \right|, \left| \lambda_{i+\frac{1}{2}} \right| \right) \#(11)$$

## SIMULAÇÕES NUMÉRICAS

Para demonstrar a acurácia e robustez do método de Godunov implementado, foram selecionados alguns exemplos de escoamento trifásico em meios porosos, assim como exemplos do modelo de águas rasas (*shallow water*). Neste artigo, apenas apresentamos alguns exemplos de escoamento em meios porosos.

As variáveis conservadas de interesse neste problema são as saturações de cada fase:  $Q = (q_1, q_2, q_3)$ . Consideramos um escoamento com algumas simplificações, tais como: permeabilidade constante do meio (rocha homogênea), as densidades e viscosidades das fases são constantes, temperatura constante, entre outras, veja Rodríguez-Bermúdez (2010) e Rodríguez-Bermúdez e Marchesin (2013).

Pelo fato de considerarmos o escoamento imiscível e o meio poroso totalmente saturado pelas fases, teremos que o problema pode ser descrito por duas variáveis e suas funções de fluxo correspondentes,

tornando-se um sistema 2 X 2. Neste estudo foram escolhidos:  $q_1$  e  $q_3$  como variáveis conservadas e  $q_2$  calculada pela relação  $q_2 = 1 - q_1 - q_3$ .

O modelo proposto tem funções de fluxo obtidas a partir da lei de Darcy para infiltração e relaciona as permeabilidades relativas das fases com as suas respectivas saturações, dadas segundo uma expressão quadrática pelo modelo de Corey et al. (1956). As funções em (12) e (13) são descritas por Rodríguez-Bermúdez (2010):

$$f_1(q_1, q_3) = \frac{q_1^2 \left( \alpha + \frac{q_3^2}{\mu_3} \rho_{13} + \frac{q_2^2}{\mu_2} \rho_{12} \right)}{\frac{q_1^2}{\mu_1} + \frac{q_2^2}{\mu_2} + \frac{q_3^2}{\mu_3}} \quad \#(12)$$

$$f_3(q_1, q_3) = \frac{q_3^2 \left( \alpha + \frac{q_1^2}{\mu_1} \rho_{31} + \frac{q_2^2}{\mu_2} \rho_{32} \right)}{\frac{q_1^2}{\mu_1} + \frac{q_2^2}{\mu_2} + \frac{q_3^2}{\mu_3}} \quad \#(13)$$

Onde  $\alpha$  é o fator adimensional que indica a razão entre efeitos de convecção por gradientes de pressão e convecção por gravidade. Quando esse valor é pequeno, os efeitos gravitacionais são predominantes, enquanto um valor maior, a gravidade é desprezível.

A diferença entre as densidades das fases  $i$  e  $j$  (nesta ordem) é dado por  $\rho_{ij}$  e  $\mu_i$  representa a viscosidade dinâmica. O desenvolvimento das funções, das variáveis utilizadas e do tempo da solução estão em unidades adimensionalizadas.

A implementação tem como estrutura um código principal onde se calcula de forma iterativa a variável  $Q_i$  usando a expressão (3) junto do fornecimento, pelo usuário, das propriedades das fases e dos parâmetros numéricos (valor  $\mathcal{V}$ , tempo de simulação, condições iniciais). As funções de fluxo (12) e (13), a formação da matriz linearizada  $\hat{A}$  e os cálculos necessários para o valor do fluxo de Godunov (6) e (7) são escritos em formas de funções independentes, em um arquivo de código separado, permitindo a fácil alteração (adaptação) para outros tipos de problemas.

A malha utilizada apresenta 1000 pontos correspondentes aos volumes em domínio computacional unitário, o número de estabilidade CFL utilizado é de  $\nu = 0.1$ , parâmetros estes determinados por tentativa. As condições iniciais são de tipo de Riemann (constantes por partes), fornecendo os valores das saturações das fases presentes na

descontinuidade inicial (poço de injeção). Como condição de contorno, embora não necessária do modelo físico real, é usada a repetição dos valores calculados nas adjacências do contorno computacional. O código foi construído e simulado utilizando a linguagem (*software*) MATLAB.

## RESULTADOS E DISCUSSÕES

Consideramos dois casos de escoamento trifásico 1D: (i) escoamento horizontal e (ii) escoamento vertical puramente gravitacional. Foram realizadas simulações com base em resultados previamente publicados (apresentados em conjunto da simulação correspondente) à fim de validação qualitativa da implementação.

### (i) Escoamento trifásico sem gravidade:

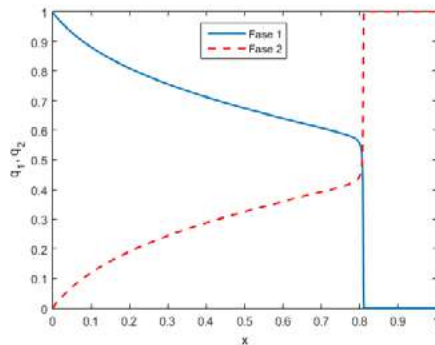
Escoamentos que ocorrem em uma direção horizontal, em que os efeitos gravitacionais são desconsiderados. Este tipo de escoamento ocorre devido às diferenças entre os gradientes de pressão entre as fases. Nestes casos, o estado inicial à esquerda  $Q_E$  corresponde a uma proporção de mistura das fases que é injetada, em um meio já preenchido com uma outra proporção de mistura, correspondente ao estado à direita  $Q_D$ . Para obter as soluções de Riemann da figura 1 são usados os seguintes parâmetros:  $\alpha = 1$ ;  $\rho_1 = \rho_2 = \rho_3 = 1$ ;  $\mu_1 = 1$ ,  $\mu_2 = 2$ ,  $\mu_3 = 0.5$  e tempo da solução de  $t = 0.59$ . Exemplos tomados do artigo de Azevedo et al. (2010) cujos resultados obtidos estão de acordo com o esperado, embora tenham-se utilizado outros tipos de esquemas numéricos

A figura 1(a) mostra a solução de Riemann quando apenas a fase 1 é injetada, num reservatório horizontal preenchido apenas pela fase 2 (escoamento bifásico), a solução de Riemann consiste numa sequência de uma onda de rarefação seguida por uma onda de choque adjacente, avançando para direita.

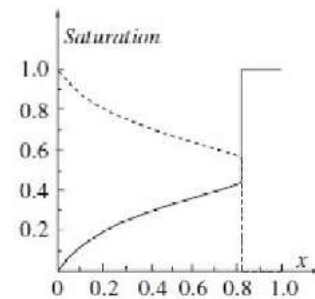
Ao inserir a fase 3 no estado de injeção, na figura 1(c), o grupo de ondas terá uma maior vantagem em avançar, porém, a menor saturação relativa à fase 1 impede que a fase 3 prossiga além de  $x = 0.7$  e forma a primeira interface de choque com a fase 2; agora existe uma sequência de ondas de rarefação seguida de choque para os grupos de onda *slow-family*. A fase 1 continua a avançar resultando em uma região de saturação constantes, seguida da interface com a fase 2, e consequentemente a formação de uma nova onda de choque para o grupo de onda *fast-family*.

Por fim, igualando as saturações na injeção, como na figura 1(e), os efeitos descritos anteriormente para os grupos de onda se repetem, mas agora a fase 1 é que tem o avanço interrompido na formação da onda de choque, enquanto a fase 3 prossegue, com efeito de formar uma região constante mais extensa.

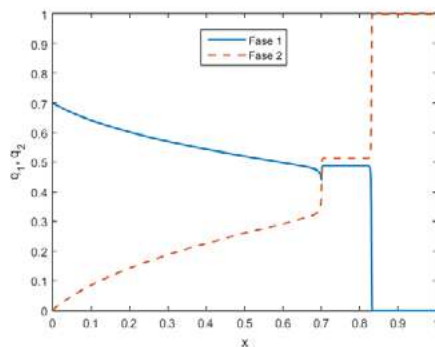
Figura 1: Perfis de saturação para as fases 1 e 2 obtidas para problemas de Riemann aplicados a escoamentos horizontais desconsiderando a gravidade (à esquerda), com a base de exemplo correspondente de Azevedo et al. (2010) (à direita).



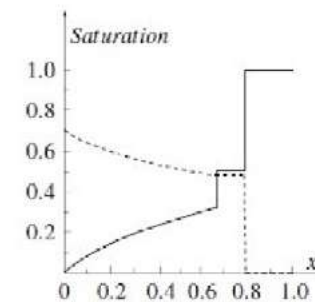
(a)  $Q_E = (1,0,0)$   $Q_F = (1,0,0)$  e  $Q_D = (0,1,0)$   $Q_D = (0,1,0)$



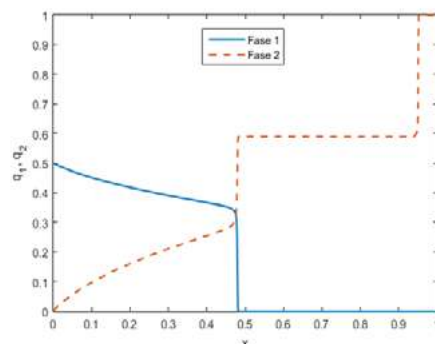
(b) Resultado esperado



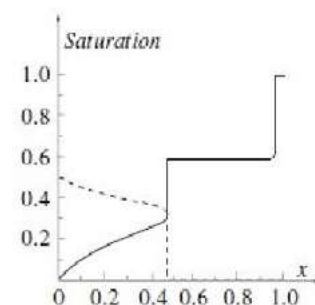
(c)  $Q_E = (0,7,0,0,3)$   $Q_F = (0,7,0,0,3)$  e  $Q_D = (0,1,0)$   
 $Q_D = (0,1,0)$



(d) Resultado esperado



(e)  $Q_E = (0,5,0,0,5)$   $Q_F = (0,5,0,0,5)$  e  $Q_D = (0,1,0)$   
 $Q_D = (0,1,0)$



(f) Resultado esperado

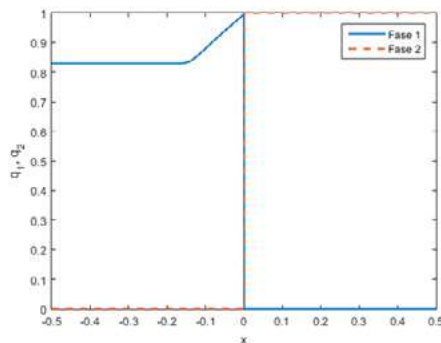
**(ii) Escoamento trifásico com gravidade:** Escoamentos que ocorrem em uma direção vertical (ou inclinada), os efeitos da gravidade são considerados a partir dos distintos valores de densidade entre as fases. Considera-se um problema onde inicialmente três fases diferentes, presentes no reservatório, estão separadas por uma membrana fixada na origem do meio em estudo. Os estados iniciais serão as saturações destas fases que estão à esquerda (acima) e à direita (abaixo) da membrana,  $Q_E$  e  $Q_D$ , respectivamente. Ao iniciar a simulação a membrana é retirada e a acomodação das fases se inicia; o efei-

to gravitacional pode ser exemplificado como se acontecesse na vertical. Para obter as soluções de Riemann da figura 2 são usados os seguintes parâmetros:  $\alpha = 0$ ;  $\mu_1 = \mu_2 = \mu_3 = 1$ , a densidade varia caso a caso, a aceleração da gravidade atua no sentido positivo do eixo  $x$  e tempo da solução foi de  $t = 2.15$ . São simulados alguns dos exemplos presentes em Rodríguez-Bermúdez (2010) e Rodríguez-Bermúdez e Marchesin (2013), onde os autores obtiveram (descreveram) qualitativamente as soluções de Riemann utilizando métodos semi-analíticos e a teoria de leis de conservação para sistemas. As simulações numéricas obtidas no presente artigo estão de acordo com as soluções descritas nas referências citadas e podem ser consideradas como as primeiras aproximações numéricas das soluções.

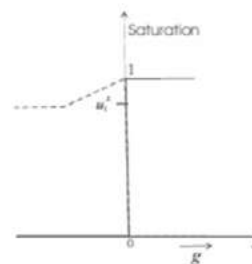
Nos casos das figuras 2(a) e 2(c), as densidades foram de  $\rho_1 = \rho_2 = 1$ ,  $\rho_3 = 0.8$ . As fases 1 e 3 estão na região “superior” do meio, a tendência será da fase 3 (menos densa) ir ficando acima da fase 1, este fenômeno é representado por uma onda de rarefação no sentido de  $x \leq 0$  no qual a saturação da fase 1 aumenta (da esquerda para direita). A formação da onda choque que precede a onda de rarefação na solução da figura 2(c) vem da maior saturação inicial da fase 3, o que resulta no surgimento dessa descontinuidade característica. Como a fase 2 só está presente na região “inferior” do meio, sua saturação não se altera já que não se desloca.

Nos casos das figuras 2(e) e 2(g), as densidades foram de  $\rho_1 = 1$ ,  $\rho_2 = \rho_3 = 0.8$ . São utilizadas condições iniciais semelhantes às dos exemplos anteriores. Agora a fase 2 tem menor densidade e também se deslocará para as regiões “superiores”. A estrutura do grupo de ondas *slow-family* que viaja para acima é chamada de *double-characteristic-shock*, isto é, um choque com rarefações adjacentes a ambos lados, e é uma estrutura atípica nas soluções de Riemann para os modelos estudados na literatura. O grupo de ondas *fast-family* que viaja pra abaixo é composta por uma onda de choque no qual essencialmente a fase 1 desloca-se para o fundo do reservatório.

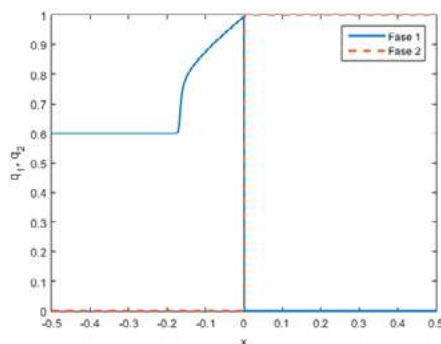
Figura 2: Perfis de saturação para as fases 1 e 2 obtidas para problemas de Riemann em escoamentos verticais produzidos pela ação da gravidade (à esquerda), com a resposta semi-analítica correspondente de Rodríguez-Bermúdez (2010) (à direita).



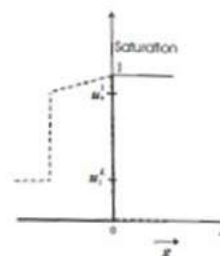
(a)  $Q_E = (0.83, 0, 0.17)$   $Q_E = (0.83, 0, 0.17)$  e  $Q_D = (0, 1, 0)$   
 $Q_D = (0, 1, 0)$



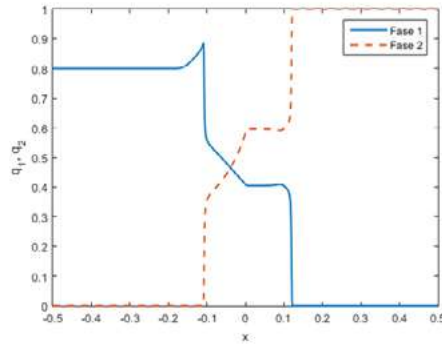
(b) Resultado semi-analítico esperado



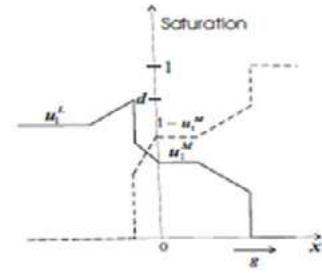
(c)  $Q_E = (0.6, 0, 0.4)$   $Q_E = (0.6, 0, 0.4)$  e  $Q_D = (0, 1, 0)$   
 $Q_D = (0, 1, 0)$



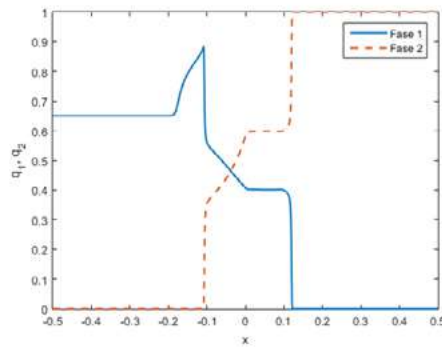
(d) Resultado semi-analítico esperado



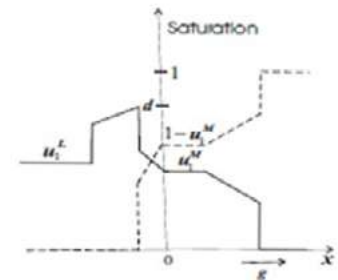
(e)  $Q_E = (0.8, 0, 0.2)$  e  $Q_D = (0.1, 0)$



(f) Resultado semi-analítico esperado



(g)  $Q_E = (0.65, 0, 0.35)$  e  $Q_D = (0.1, 0)$



(h) Resultado semi-analítico esperado

## CONCLUSÕES

Foi implementado um método tipo Godunov em volumes finitos para sistemas 2 X 2 de leis de conservação e que pode ser estendido para o caso de um maior número de equações. Foi validado o código para distintos modelos descrevendo fenômenos de transporte, e logo, utilizado para obter soluções numéricas de um problema ainda em estudo como o caso de escoamento trifásico com gravidade em meios porosos.

Na implementação utilizou-se a construção de matrizes lineares nas interfaces dos volumes, em substituição da matriz jacobiana das funções de fluxo. Como resultado, o problema é transformado em diversos sistemas locais, resolvidos pelo *Riemann solver* de Roe (1981) e a condição de entropia para correção das soluções é definida pelo critério de Harten e Hyman (1983). Do ponto de vista computacional, esta técnica reduziu o custo ao obter as soluções.

O tempo computacional de execução do código está relacionado ao número de iterações, à complexidade das funções de fluxo e ao valor  $\mathcal{V}$  usado.

Nestes exemplos, os resultados para escoamentos horizontais têm maior tempo de simulação, comparado aos verticais, devido à característica das soluções percorrerem uma maior distância horizontalmente, requerendo maior número de iterações no código; mesmo que possuam a mesma função de fluxo e valor  $\mathcal{V}$  usado nas simulações verticais.

Os resultados fornecidos pelo código e descritos na literatura se correlacionaram adequadamente tanto para as simulações em escoamentos horizontais bem como nos escoamentos verticais com gravidade, como demonstrado nas figuras 1 e 2. Os perfis das soluções de Riemann, apresentam uma formação bem definida dos grupos de onda (estruturas de choques e rarefações) e ausência de oscilações espúrias, mostrando a correta escolha dos parâmetros numéricos.

Além disso, a utilização do *Riemann solver* utilizando-se da configuração da matriz  $\hat{A}$  não influenciou negativamente na obtenção da solução e a técnica de correção de entropia se mostrou necessária e importante para o ajuste no fluxo calculado permitindo que o método convergisse para os resultados esperados.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AZEVEDO, A. V. et al. The solution by the wave curve method of three-phase flow in virgin reservoirs. **Transport in Porous Media**, v. 83, n. 01, p. 99–125, 2010. Disponível em: < <https://link.springer.com/article/10.1007/s11242-009-9508-9> >. Acesso em: 2 fev. 2023.
- COREY, A. et al. Three-Phase Relative Permeability. **Journal of Petroleum Technology**, v. 8, n. 11, p. 63–65, 1956. Disponível em: < <https://onepetro.org/JPT/article/8/11/63/161021/Three-Phase-Relative-Permeability> >. Acesso em: 27 dez. 2022.
- COURANT, R.; FRIEDRICHS, K.; LEWY, H. On the partial difference equations of mathematical physics. **IBM journal of Research and Development**, v. 11, n. 02, p. 215–234, 1967. Disponível em: < <https://ieeexplore.ieee.org/abstract/document/5391985> >. Acesso em: 22 nov. 2022.
- GODUNOV, S. K.; BOHACHEVSKY, I. Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics. **Matematičeskij sbornik**, v. 47, n. 03, p. 271–306, 1959. Disponível em: < <https://hal.science/hal-01620642/> >. Acesso em: 14 nov. 2022.
- HARTEN, A; HYMAN, J. M. Self adjusting grid methods for one-dimensional hyperbolic conservation laws. **Journal of computational Physics**, v. 50, n. 02, p. 235–269, 1983. Disponível em: < <https://www.sciencedirect.com/science/article/abs/pii/0021999183900669> >. Acesso em: 12 fev. 2023.
- LAX, P. D. **Hyperbolic systems of conservation laws and the mathematical theory of shock waves**. 1. ed. SIAM, 1973.
- LEVEQUE, R. J. **Finite-Volume Methods for Hyperbolic Problems**. 1. ed. New York: Cambridge University Press, 2004.
- LIU, T. -P. The entropy condition and the admissibility of shocks. **Journal of Mathematical Analysis and Applications**, v. 53, n. 01, p. 78–88, 1976. Disponível em: < <https://www.sciencedirect.com/science/article/pii/0022247X76901463> >. Acesso em: 6 out. 2023.
- RODRÍGUEZ-BERMÚDEZ, P. **Buoyancy driven three-phase flow in porous media**. 2010. 112 f. Tese (Doutorado em Matemática) - IMPA, Rio de Janeiro, 2010.
- RODRÍGUEZ-BERMÚDEZ, P.; MARCHESIN, D. Riemann Solutions for vertical flow of three phases in porous media: simple cases. **Journal of Hyperbolic Differential Equations**, v. 10, n. 02, p. 335-370, 2013. Disponível em: < <https://www.worldscientific.com/doi/abs/10.1142/S0219891613500124> > Acesso em: 19 jun. 2023.
- ROE, P. L. Approximate riemann solvers, parameter vector, and difference schemes. **Journal of Computational Physics**, v. 43, n. 02, p. 357–372, 1981. Disponível em: < <https://www.sciencedirect.com/science/article/abs/pii/0021999181901285> >. Acesso em: 11 out. 2022.
- TORO, E. F. **Riemann solvers and numerical methods for fluid dynamics: a practical introduction**. 3. ed. Berlin: Springer Science and Business Media, 2009.

## DADOS DOS AUTORES

**Eron Oliveira do Nascimento** (eronoliveira@id.uff.br). Mestre em Modelagem Computacional, Aluno do Programa de Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia, Universidade Federal Fluminense, Rio de Janeiro, Brasil.

**Panters Rodríguez Bermúdez** (pantersrb@id.uff.br). Doutor em Matemática, Docente do Programa de Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia, Universidade Federal Fluminense, Rio de Janeiro, Brasil.

**Wagner Rambaldi Telles** (wtelles@id.uff.br). Doutor em Modelagem Computacional, Docente do Programa de Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia, Universidade Federal Fluminense, Rio de Janeiro, Brasil.

# COMPARAÇÃO ENTRE MODELOS PREDITIVOS EM SÉRIES TEMPORAIS HIERÁRQUICAS PARA O CONSUMO DE ENERGIA ELÉTRICA DO SETOR INDUSTRIAL BRASILEIRO

## COMPARISON BETWEEN FORECAST MODELS IN HIERARCHICAL TIME SERIES FOR ELECTRICITY CONSUMPTION IN THE BRAZILIAN INDUSTRIAL SECTOR

Marlon Mesquita Lopes Cabreira

Felipe Leite Coelho da Silva

Josiane da Silva Cordeiro Coelho

**RESUMO:** Análise e previsão do consumo de energia elétrica do setor industrial brasileiro vêm sendo realizada ao longo dos últimos anos devido a sua importância para economia brasileira e ao meio ambiente. Esse setor é o maior consumidor de energia elétrica no Brasil. Os dados de consumo de energia elétrica do setor industrial podem ser organizados em uma estrutura hierárquica composta pelas regiões (Norte, Nordeste, Centro-Oeste, Sudeste e Sul), no nível inferior, e o Brasil no nível superior. Este trabalho tem o intuito de avaliar a capacidade preditiva dos modelos de suavização exponencial, Box-Jenkins, redes neurais autorregressivas (RNA) e de *perceptron* multicamadas (MLP). Estes modelos foram integrados as abordagens *bottom-up* e *top-down*. Os resultados mostram que os modelos de suavização exponencial forneceram as melhores previsões em comparação com os outros métodos analisados.

**Palavras-Chave:** Séries temporais hierárquicas. Setor industrial brasileiro. Energia elétrica. Suavização exponencial. Redes neurais.

**ABSTRACT:** Analysis and forecasting of electrical energy consumption in the Brazilian industrial sector have been carried out over the last few years due to its importance for the Brazilian economy and the environment. This sector is the largest consumer of electricity in Brazil. Electric energy consumption data from the industrial sector can be organized in a hierarchical structure composed of regions (North, Northeast, Central-West, Southeast and South), at the lower level, and Brazil at the upper level. This work aims to evaluate the predictive capacity of exponential smoothing, Box-Jenkins, autoregressive neural networks (ANN) and multilayer perceptron (MLP) models. These models were integrated into bottom-up and top-down approaches. The results show that the exponential smoothing models provided the best predictions compared to the other analyzed methods.

**Keywords:** Hierarchical time series. Brazilian industrial sector. Electricity. Exponential smoothing. Neural networks.



## INTRODUÇÃO

O setor industrial é o maior consumidor de energia elétrica entre os setores que compõem a matriz energética brasileira. Esse setor tem importância para os indicadores do produto interno bruto (PIB) e para sociedade brasileira. Neste contexto, o planejamento energético deste setor é importante para o desenvolvimento econômico, ao meio ambiente e para eficiência energética. Dessa forma, análise e previsão de séries temporais hierárquicas do consumo de energia elétrica podem contribuir para o planejamento e desenvolvimento do setor industrial brasileiro.

As séries históricas do consumo de energia elétrica do setor industrial podem ser organizados em uma estrutura de hierárquica. Em que no topo (nível 0), se encontra os dados do Brasil, e na base (nível 1), os dados de cada uma das cinco regiões geográficas (Norte, Nordeste, Centro-Oeste, Sudeste e Sul).

Os modelos preditivos utilizados neste trabalho foram o de suavização exponencial, Box-Jenkins, redes neurais autorregressivas e de *perceptron* multicamadas (MLP - *Multilayer Perceptrons*). Cada uma desses modelos foram acoplados a duas abordagens em séries temporais hierárquicas, conhecidas como *bottom-up* e *top-down*. Na abordagem *bottom-up*, são geradas as previsões independentes para cada série no nível mais baixo da hierarquia e, em seguida, elas são agregadas para produzir a previsão dos níveis superiores da hierarquia. Por sua vez, a abordagem *top-down* gera as previsões da série temporal totalmente agregada e, em algumas situações, desagrega tais previsões para os níveis exigidos em proporções históricas.

O objetivo deste trabalho é comparar os modelos preditivos de suavização exponencial, Box e Jenkins, redes neurais autorregressivas e de redes neurais *perceptron* multicamadas para obtenção das projeções mensais do consumo de energia elétrica do setor industrial brasileiro acoplados as abordagens *bottom-up* e *top-down*. A capacidade preditiva das abordagens será avaliada pelas métricas do erro médio absoluto percentual (MAPE) e do erro quadrático médio (RMSE).

No setor industrial, Silva et al. (2018) introduziram uma abordagem *bottom-up* para obter as projeções de longo prazo do consumo de eletricidade para a indústria de papel e celulose no Brasil, considerando as medidas de eficiência energética. Silva et al. (2019) propõem um modelo que combina os modelos lineares hierárquicos com abordagem *bottom-up* para o consumo de energia elétrica do setor industrial no Brasil.

Com a evolução recente dos algoritmos de aprendizado de máquina, eles têm sido utilizados para previsão hierárquica de eletricidade. Um estudo comparativo usando os modelos de redes neurais artificiais para previsão do consumo de eletricidade do setor industrial brasileiro é apresentado em Silva et al. (2022). Maiçara et al. (2015) realizaram a projeção do consumo de energia elétrica para o setor residencial no Brasil. Spiliotis et al. (2021) propõem uma abordagem de previsão hierárquica baseada em aprendizado de máquina. Cartinteiro et al. (2007) apresentou um modelo neural híbrido hierárquico para previsão de carga de longo prazo. Flidner e Lawrence (1995) desenvolveram um estudo empírico com uma análise comparativa sobre o desempenho de previsão de séries temporais hierárquicas.

## METODOLOGIA

Neste trabalho, foram utilizados os dados do consumo mensal de energia elétrica do setor industrial do Brasil e por região geográfica para o período de janeiro de 1979 até dezembro de 2022. Os dados estão disponíveis no sítio eletrônico da Empresa de Pesquisa Energética (EPE, 2022). Os dados não possuem variância estacionária e muitas mudanças de nível, então utilizou uma transformação Box-Cox. Para avaliar a capacidade preditiva dos modelos de suavização exponencial, Box e Jenkins, Redes Neurais Autorregressivas e de *perceptron* multicamadas acopladas as abordagens *bottom-up* e *top-down*, os dados foram divididos em dois grupos, o conjunto de treinamento e de teste. O primeiro conjunto de treinamento foi do período de janeiro de 1979 a dezembro de 2018, o segundo de 1979 até 2019 e assim sucessivamente, até considerar o último conjunto que finalizava em dezembro de 2022. Como observado na tabela 1:

**Tabela 1:** Os cinco conjuntos de treinamento e teste.

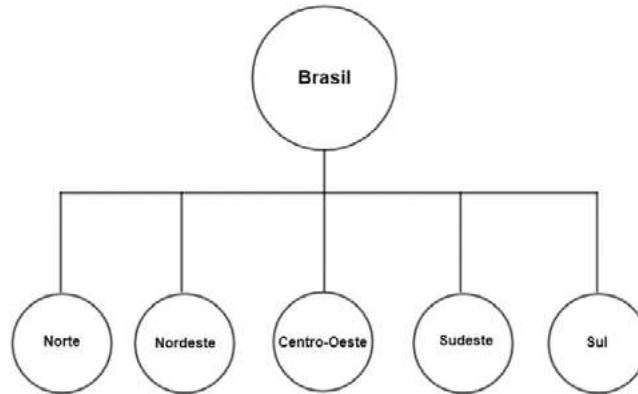
Cenário	Treinamento	Teste
1°	1979-2017	2018
2°	1979-2018	2019
3°	1979-2019	2020
4°	1979-2020	2021
5°	1979-2021	2022

Fonte: Desenvolvido pelos autores.

## NOTAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

Os dados de consumo de energia elétrica da indústria brasileira podem ser desagregados em cinco grupos (regiões do Brasil), apresentando uma estrutura hierárquica, onde a base da estrutura é formada por cinco séries temporais (Norte, Nordeste, Centro-Oeste, Sudeste e Sul) e no topo da hierarquia se encontra a série totalmente agregada que representa os dados do Brasil. Como apresentado na figura 1:

**Figura 1:** Esquema da estrutura hierárquica de um nível para o consumo de energia elétrica do setor industrial brasileiro.



Fonte: Desenvolvidos pelos autores.

Seja  $Y_t$  o consumo de energia elétrica observado no mês  $t$ , para  $t = 1, \dots, T$ , com  $T$  representando a quantidade de observações na série temporal. De acordo com Hyndman et al. (2018), uma série temporal hierárquica pode ser representada através de uma equação matricial representada na equação (1):

$$Y_t = S b_t \#(1)$$

em que  $S$  é uma matriz soma de ordem  $n \times m$ , em que  $n$  é a quantidade de séries em toda hierarquia e  $m$  é a quantidade de séries na base da hierarquia (quantidade de regiões, em particular),  $b_t$  é o vetor de observações na base da hierarquia de dimensão  $m$  no tempo  $t$ . Em particular, a equação matricial utilizada neste trabalho está apresentada na equação (2):

$$\begin{pmatrix} Y_{B,t} \\ Y_{BN,t} \\ Y_{BR,t} \\ Y_{BC,t} \\ Y_{BD,t} \\ Y_{BS,t} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} Y_{BN,t} \\ Y_{BR,t} \\ Y_{BC,t} \\ Y_{BD,t} \\ Y_{BS,t} \end{pmatrix} \#(2)$$

em que  $n = 6$  é o total de séries temporais na hierarquia,  $m = 5$  é o total de séries temporais no nível inferior da hierarquia, e  $Y_{B,t}$ ,  $Y_{BN,t}$ ,  $Y_{BR,t}$ ,  $Y_{BC,t}$ ,  $Y_{BD,t}$  e  $Y_{BS,t}$  representam as séries temporais do consumo de energia elétrica para o Brasil e para as regiões Norte, Nordeste, Centro-Oeste, Sudeste e Sul, respectivamente. Previsões utilizando uma estrutura de série temporal hierárquica podem ser calculadas através da equação (3):

$$\tilde{Y}_h = S G \tilde{Y}_h \#(3)$$

em que  $\tilde{Y}_h$  representa um conjunto de previsões combinadas  $h$  passos à frente,  $\tilde{Y}_h$  são as previsões de base (ou seja, é o nível inferior da estrutura hierárquica), que são previsões independentes de todas as séries na hierarquia e  $G$  é uma matriz referente ao tipo de abordagem de previsão hierárquica que será utilizada.

## A ABORDAGEM *BOTTOM-UP*

Esta abordagem consiste em utilizar modelos preditivos em cada série no nível inferior e da hierarquia, em seguida, agregá-las para produzir previsões para toda estrutura hierárquica. Uma vantagem dessa abordagem é que prevendo nos níveis inferiores, nenhuma informação perdida devido a agregação, ou seja, a abordagem é capaz de captar características particulares das séries inferiores, assim, podendo contribuir para a capacidade preditiva. Pode-se representar a abordagem *bottom-up* utilizando a equação (4):

$$\begin{pmatrix} \hat{Y}_{B,T,h} \\ \hat{Y}_{BN,h} \\ \hat{Y}_{BR,h} \\ \hat{Y}_{BC,h} \\ \hat{Y}_{BD,h} \\ \hat{Y}_{BS,h} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_S \cdot \underbrace{\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_G \cdot \begin{pmatrix} \hat{Y}_{B,T,h} \\ \hat{Y}_{BN,h} \\ \hat{Y}_{BR,h} \\ \hat{Y}_{BC,h} \\ \hat{Y}_{BD,h} \\ \hat{Y}_{BS,h} \end{pmatrix} \quad \#(4)$$

Onde  $G = [0_{m \times (n-m)} | I_m]$  é uma matriz de zeros de ordem  $m \times (n - m)$  acoplada a uma matriz identidade de ordem  $m$ .

## A ABORDAGEM *TOP-DOWN*

A abordagem *top-down* envolve, primeiramente, previsões para a série no topo da hierarquia  $Y_{B,t}$  em seguida, desagregando-as para níveis inferiores utilizando um conjunto de proporções  $p_1, p_2, \dots, p_m$ , onde  $m$  é a quantidade de séries na base da hierarquia. Estas proporções ditam como a previsão deve ser distribuída nas séries inferiores. Uma vantagem desta abordagem é a simplicidade, já que prevendo somente com uma série podemos ter previsões para toda a hierarquia. Pode-se representar a abordagem *top-down* utilizando a equação (5):

$$\begin{pmatrix} \hat{Y}_{B,T,h} \\ \hat{Y}_{BN,h} \\ \hat{Y}_{BR,h} \\ \hat{Y}_{BC,h} \\ \hat{Y}_{BD,h} \\ \hat{Y}_{BS,h} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} p_1 & 0 & 0 & 0 & 0 & 0 \\ p_2 & 0 & 0 & 0 & 0 & 0 \\ p_3 & 0 & 0 & 0 & 0 & 0 \\ p_4 & 0 & 0 & 0 & 0 & 0 \\ p_5 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_G \cdot \begin{pmatrix} \hat{Y}_{B,T,h} \\ \hat{Y}_{BN,h} \\ \hat{Y}_{BR,h} \\ \hat{Y}_{BC,h} \\ \hat{Y}_{BD,h} \\ \hat{Y}_{BS,h} \end{pmatrix} \quad \#(5)$$

Onde  $G = [p | 0_{m \times (n-1)}]$  é uma matriz com um vetor coluna de proporções  $p$  acoplado a uma matriz de zeros de ordem  $m \times (n - 1)$ . Como este trabalho tem o interesse em comparar previsões para a série de topo, não será desagregada as previsões para níveis inferiores, assim não sendo necessário o cálculo de proporções.

## MODELO DE SUAUIZACÃO EXPONENCIAL

O modelo de suavização exponencial foi introduzido por Brown (1959), Holt (1957) e Winters (1960). Neste modelo as previsões são geradas a partir de uma média ponderada das observações passadas, sendo que o peso das observações decresce exponencialmente à medida que elas envelhecem. Existem vários modelos de suavização exponencial, eles variam pelas características da série temporal.

Hyndman et al. (2018) propuseram uma classificação que depende de três elementos: erro, tendência e sazonalidade da série temporal. O erro pode ser aditivo ( $A$ ) ou multiplicativo ( $M$ ), a tendência pode ser aditiva ( $A$ ), aditiva com amortecimento ( $A_d$ ) ou no caso de inexistência, nenhuma ( $N$ ), por sua vez, a sazonalidade, pode ser aditiva ( $A$ ), multiplicativa ( $M$ ) ou nenhuma ( $N$ ). Dessa forma, cada modelo de

suavização exponencial utilizada neste trabalho pode ser classificada como ETS (Erro, Tendência, Sazonalidade). Neste trabalho foi utilizado o algoritmo *ETS* proposto por Hyndman et al. (2008) para ajustar os modelos de suavização exponencial.

## MODELO DE BOX E JENKINS

O modelo de Box e Jenkins utiliza as funções de autocorrelação da série temporal para obter os parâmetros do modelo. Estes modelos foram estudados por Box e Jenkins (1976) e busca integrar modelos autorregressivos (AR) e de médias móveis (MA) a um conjunto de observações. O modelo utilizado foi o ARIMA sazonal (SARIMA), para incorporar a componente de sazonalidade que está presente nos dados. A estrutura do modelo SARIMA de ordem  $(p, d, q) \times (P, D, Q)_s$  é dado pela equação (6):

$$\phi(B)\Phi(B^s)\Delta^d\Delta_s^D Z_t = \theta(B)\Theta(B^s)a_t \quad (6)$$

em que  $a_t$  é o ruído branco;  $\phi(B)$  é o operador autorregressivo de ordem  $p$ ;  $\theta(B)$  é o operador de médias móveis de ordem  $q$ ;  $\Phi(B^s)$  é o operador autorregressivo sazonal de ordem  $P$ ;  $\Theta(B^s)$  é o operador de médias móveis sazonal de ordem  $Q$ ;  $\Delta^d$  é o operador da diferença simples e;  $\Delta_s^D$  é o operador da diferença sazonal;  $s$  é o número de observações por ano (período). O modelo de Box e Jenkins foi obtido através do algoritmo proposto por Hyndman e Khandakar (2018).

## REDES NEURAIS

As redes neurais artificiais foram apresentadas pela primeira vez por McCulloch e Pitts (1943). Desde então, observa-se aplicações de redes neurais nas diversas áreas da ciência inclusive em séries temporais. Toda a rede neural artificial tem uma camada de entrada e outra de saída. Também pode haver camadas intermediárias contendo neurônios ocultos, a qual chama-se de camada oculta. Ambas as camadas contêm uma quantidade de neurônios artificiais, em que são conectados entre si formando uma rede.

A camada de entrada é responsável por receber informações oriunda de um meio externo, no caso de uma série temporal, em geral, são observações passadas, já a camada oculta, tem a tarefa de aplicar pesos as observações passadas e utilizam funções, que são chamadas

de funções de ativação, para introduzir a não linearidade nas saídas das camadas ocultas. A camada de saída, em geral, entrega o valor da previsão no tempo futuro subsequente Hyndman et al. (2018).

Neste trabalho foram utilizados dois tipos de redes neurais, a autorregressiva e a MLP ambas são do tipo *feed-forward*, ou seja, a informação que entra na rede flui através da camada de entrada até a camada de saída sem ciclos ou retroalimentação, além disso, foi implementado uma camada oculta.

## CAPACIDADE PREDITIVA

Em cada conjunto teste, foi avaliado a capacidade preditiva do método, para tal, é necessário a escolha de uma métrica. Este trabalho utilizou as métricas apresentadas nas equações (6) e (7):

Erro médio absoluto percentual (MAPE).

$$MAPE = 100 \cdot \frac{1}{N} \cdot \sum_{t=1}^N \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (6)$$

Raiz do erro quadrático médio (RMSE).

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{t=1}^N (y_t - \hat{y}_t)^2} \quad (7)$$

considerando-se que  $N$  é quantidade de valores previstos,  $\hat{y}_t$  é o valor previsto e  $y_t$  é o valor observado ambos para o instante  $t$ . Com os valores de capacidade preditiva em cada conjunto teste, foi selecionado o modelo com menor média entre as métricas.

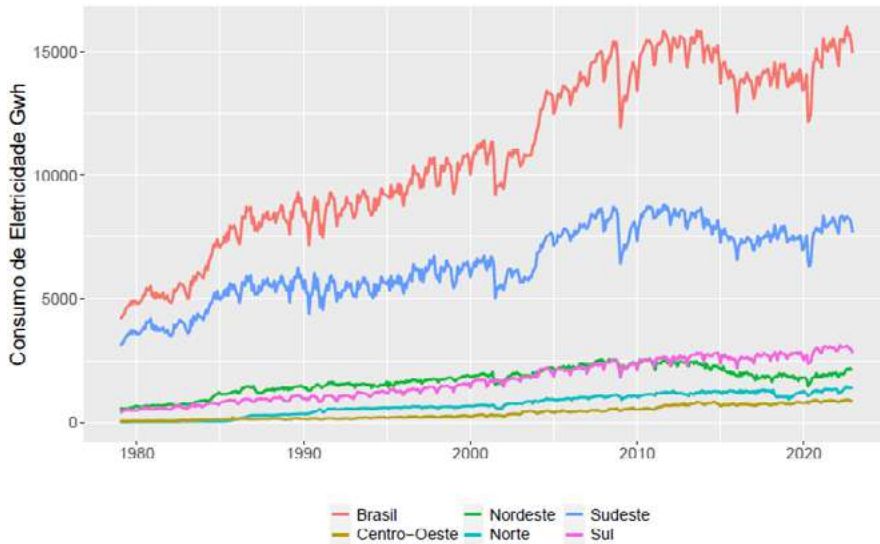
## RESULTADOS

Nesta seção será apresentado as séries temporais do consumo de energia elétrica do setor industrial brasileiro e de cada região. Além disso, será exibido os resultados de capacidade preditiva.

### Descrição das Séries Temporais

Na figura 2 é apresentado o consumo de energia elétrica do setor industrial brasileiro comparando as cinco regiões geográficas.

**Figura 2:** Consumo de energia elétrica no setor industrial do Brasil e região.



Fonte: Desenvolvidos pelos autores.

O comportamento da série da região Sudeste é semelhante a série do Brasil além disso, é a região que mais consome energia elétrica, com contrapartida, a região Centro-Oeste é a que menos consome. As regiões Sul e Norte possuem clara tendência de crescimento, aquela, ocupando a segunda região que mais consome energia elétrica por parte de suas industriais, esta, a penúltima. A região Nordeste teve uma tendência de queda depois de 2010. Em todas as séries é visível mudanças de nível e alguma sazonalidade.

### Comparação entre Modelos Preditivos

Nesta seção serão apresentados os resultados das métricas MAPE e RMSE das previsões dos modelos de suavização exponencial, Box-Jenkins, Redes Neurais Autorregressivas e Redes Neurais MLP acoplados as abordagens *bottom-up* e *top-down*. Em todas as séries, foi aplicado uma transformação Box-Cox para controlar a variância dos dados, já que as séries de consumo de energia elétrica do setor industrial possuem sazonalidade multiplicativa e muitas mudanças de nível.

Na tabela 2, pode-se verificar que os valores do MAPE médio são menores que 3,7% entre os modelos preditivos acoplados a abordagem *top-down* adotados neste trabalho. O modelo de suavização exponencial via abordagem *top-down* exibiu o menor valor para ambos os modelos médias, logo, obteve a melhor capacidade preditiva em comparação com as outras abordagens.

**Tabela 2:** Desempenho dos métodos preditivos nos dados de consumo de eletricidade do setor industrial brasileiro usando a abordagem *top-down*.

Teste	2018		2019		2020	
Suav. Exponencial	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	2,20%	361,59	3,54%	547,72	6,10%	913,20
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,80%	356,99	1,53%	264,25	<b>3,04%</b>	<b>488,75</b>
Teste	2018		2019		2020	
Box e Jenkins	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,47%	271,66	2,46%	399,18	6,38%	945,95
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	3,31%	625,30	2,19%	375,53	<b>3,16%</b>	<b>523,52</b>
Teste	2018		2019		2020	
RNA	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,47%	266,39	1,57%	266,87	5,89%	910,29
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	2,67%	445,81	4,62%	755,97	<b>3,24%</b>	<b>529,06</b>
Teste	2018		2019		2020	
MLP	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,32%	273,89	3,01%	499,27	6,76%	991,07
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	4,03%	751,82	2,98%	539,15	<b>3,62%</b>	<b>611,04</b>

Fonte: Desenvolvida pelos autores.

Na tabela 3, os valores de MAPE médios são menores que 4,6\% entre os modelos preditivos acoplados a abordagem *bottom-up*. O modelo de suavização exponencial via abordagem *bottom-up* exibiu o menor valor médio em ambas as métricas, logo, obteve a melhor capacidade preditiva.

**Tabela 3:** Desempenho dos métodos preditivos nos dados de consumo de eletricidade do setor industrial brasileiro usando a abordagem *bottom-up*.

Teste	2018		2019		2020	
Suav. Exponencial	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,44%	268,17	3,01%	408,05	5,65%	892,20
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	2,45%	449,84	1,57%	262,89	<b>2,83%</b>	<b>470,63</b>
Teste	2018		2019		2020	
Box e Jenkins	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,30%	249,01	2,58%	424,99	6,50%	977,58
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	3,84%	677,18	2,10%	362,86	<b>3,27%</b>	<b>538,33</b>
Teste	2018		2019		2020	
RNA	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,78%	308,97	1,91%	360,19	5,89%	1019,85
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	8,64%	1398,48	3,78%	704,15	<b>4,57%</b>	<b>758,33</b>
Teste	2018		2019		2020	
MLP	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	2,39%	410,96	3,56%	584,06	7,10%	1068,06
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	4,01%	841,99	3,51%	630,57	<b>4,11%</b>	<b>707,13</b>

Fonte: Desenvolvida pelos autores.

## CONCLUSÕES

Neste trabalho, considerou-se que o consumo de energia elétrica do setor industrial brasileiro pode ser organizado em uma estrutura hierárquica, onde a série histórica do Brasil se encontra no topo, e na base as séries históricas das cinco regiões geográficas. Neste contexto, foram utilizados os modelos preditivos de suavização exponencial, de Box-Jenkins, de redes neurais autorregressivas e de redes neurais MLP integrados as abordagens *bottom-up* e *top-down*, com o intuito de avaliar as previsões do consumo de energia elétrica do setor industrial no Brasil em cinco cenários de conjunto teste diferentes. Antes de realizar as previsões foi utilizado uma transformação Box-Cox nos dados afim de controlar a variância das séries temporais, para avaliar a capacidade preditiva entre as abordagens foram utilizadas as médias das métricas MAPE e RMSE.

As abordagens consideradas neste trabalho apresentaram um MAPE médio inferior a 4,6%. O modelo de suavização exponencial acoplado a abordagem de *bottom-up* apresentou melhor capacidade entre todos os modelos analisados. Portanto, os resultados mostraram que o modelo de suavização exponencial pode contribuir para obtenção de projeções do consumo de energia elétrica no horizonte de curto prazo.

Para pesquisas futuras é interessante comparar os resultados deste trabalho usando modelos de redes neurais de memória longa.

## REFERÊNCIAS

- Athanasopoulos, G., Ahmed, R. A., e Hyndman, R. J. Hierarchical forecast for Australian domestic tourism. **International Journal of Forecasting**, 25(1):146-166. ISSN 0169-2070. 2009. URL: <https://www.sciencedirect.com/science/article/pii/S0169207008000691>. Acesso em: 10 de out. de 2023.
- Box, G. E. P., Jenkins, G. M., e Reinsel, G. C. **Time Series Analysis: Forecasting and Control**. Prentice-Hall 2015.
- Carpinteiro, O. A., Leme, R. C., de Souza, A. C. Z., Pinheiro, C. A., e Moreira, E. M. Long-term load forecasting via a hierarchical neural model with time integrators. **Electric Power Systems Research**, 77(3):371-378. ISSN 0378-7796. 2017. Disponível em: <https://doi.org/10.1016/j.epr.2006.03.014>. Acesso em: 9 de out. de 2023.
- Empresa de Pesquisa Energética. Séries estudos de energia, anuário de energia elétrica. 2023 URL: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/anuario-estatistico-de-energia-eletrica>. Acesso em: 20 Jul. 2023.
- Fliedner, E. B. e Lawrence, B. Forecasting system parent group formation: An empirical application of cluster analysis. **Journal of Operations Management**, 12(2):119-130. ISSN 0272-6963. 1995. Disponível em: [https://doi.org/10.1016/0272-6963\(94\)00009-4](https://doi.org/10.1016/0272-6963(94)00009-4). Acesso em: 9 de out. de 2023.
- Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., Shang, H. L. Optimal Combination forecast for hierarchical time series. **Computational Statistics Data Analysis**, 55(9):2579-2589. ISSN 0167-9473. 2011. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0167947311000971>. Acesso em: 11 de out. de 2023.
- Hyndman, R.J., e Athanasopoulos, G. **Forecasting: principles and practice**, 2nd edition, OTexts: Melbourne, Australia. 2018. URL: <https://otexts.com/fpp2/index.html>. Acesso em: 20 Jul. 2023
- Maçaira, P. M., Souza, R. C., e Oliveira, F. L. C. Modelling and forecasting the residential electricity consumption in Brazil with Pegels exponential smoothing techniques. **Procedia Computer Science**, 55:328-335. 2015. Disponível em: <https://doi.org/10.1016/j.procs.2015.07.057>. Acesso em: 11 de out. de 2023.
- Mcculloch, W. e Pitts, W. **Bulletin of Mathematical Biophysics**, p. 127-147. 1943.
- Sadownik, R., Barbosa, E. P. Short-term forecasting of industrial electricity consumption in Brazil. **Journal of Forecasting**, 18(3), p.215-224. 1999. Disponível em: [https://doi.org/10.1002/\(SICI\)1099-131X\(199905\)18:3<215::AID-FOR719>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1099-131X(199905)18:3<215::AID-FOR719>3.0.CO;2-B). Acesso em: 11 de out. de 2023.
- Silva, F. L., Cyrino Oliveira, F. L., e Souza, R. C. A bottom-up bayesian extension for long term electricity consumption forecasting. **Energy**, 167:198-210. ISSN 0360-5442. 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0360544218321984>. Acesso em: 11 de out. de 2023.
- Silva, F., da Costa, K., Rodrigues, P., Salas, R., e Linkolk, J. Statistical and artificial neural networks models for electricity consumption forecasting in the Brazilian industrial sector. **Energies**, 15:588. 2022. Disponível em: <https://doi.org/10.3390/en15020588>. Acesso em: 11 de out. de 2023.
- Silva, F. L., Souza, R. C., Cyrino Oliveira, F. L., Lourenço, P. M., e Calili, R. F. A bottom-up methodology for long term electricity consumption forecast of an industrial sector - application to pulp and paper sector in Brazil. **Energy**, 144:1107-1118. ISSN 0360-5442. 2018. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0360544217321217>. Acesso em: 10 de out. de 2023.
- Spiliotis, E., Abolghasemi, M., Hyndman, R. J., Petropoulos, F., e Assimakopoulos, V. (Hierarchical forecast reconciliation with machine learning. **Applied Soft Computing**, 112:107756. ISSN 1568-4946. 2021. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1568494621006773>. Acesso em: 11 de out. de 2023.

## DADOS DOS AUTORES

**Marlon Mesquita Lopes Cabreira** ([marlonmlc23@gmail.com](mailto:marlonmlc23@gmail.com)). Mestre em Modelagem Matemática pela UFRRJ. Docente do Departamento de Matemática na Universidade Federal Rural do Rio de Janeiro, Seropédica, UFRRJ.

**Felipe Leite Coelho da Silva** ([felipeleite@ufrj.br](mailto:felipeleite@ufrj.br)). Doutor em Engenharia Elétrica, na área de Métodos de Apoio à Decisão, pela PUC-Rio. Docente do Departamento de Matemática na Universidade Federal Rural do Rio de Janeiro, Seropédica, UFRRJ.

**Josiane da Silva Cordeiro Coelho** ([josicordeiro@gmail.com](mailto:josicordeiro@gmail.com)). Doutora em Estatística pela Universidade Federal do Rio de Janeiro. Docente do Departamento de Matemática na Universidade Federal Rural do Rio de Janeiro, Seropédica, UFRRJ.

# CONSTRUÇÃO E CARACTERIZAÇÃO DE LOOPS DE CÓDIGO

## *CONSTRUCTION AND CHARACTERIZATION OF CODE LOOPS*

Rosemary Miguel Pires

**RESUMO:** A construção de loops de código, como apresentamos neste trabalho, foi introduzida por Griess (1986). O desenvolvimento da teoria de loops de código tem bastante aplicabilidade no estudo das álgebras não associativas. Griess (1986) demonstrou que um loop de código é um loop de Moufang. Em (CHEIN; GOODAIRE, 1990) foi demonstrado que loops de código tem um único comutador não trivial, um único associador não trivial e um único quadrado não trivial. Neste trabalho, apresentamos como os loops de código são construídos. Para isto, desenvolvemos uma introdução inicial sobre códigos (lineares) pares, loops e loops de Moufang. Por fim, mostramos como determinar quando um certo loop finito pode ser caracterizado como loop de código.

**Palavras-chave:** Loops de Moufang. Códigos lineares pares. Loops de Código.

**ABSTRACT:** The construction of code loops, as presented in this work, was introduced by Griess (1986). The development of code loop theory has considerable applicability in the study of non-associative algebras. Griess (1986) demonstrated that a code loop is a Moufang loop. In (CHEIN; GOODAIRE, 1990) it was demonstrated that code loops have a unique nonidentity commutator, a unique nonidentity associator and a unique nonidentity square. In this work, we present how code loops are constructed. For this, we developed an initial introduction to doubly even (linear) codes, loops and Moufang loops. Finally, we show how to determine when a certain finite loop can be characterized as a code loop.

**Keywords:** Moufang Loops. Even linear codes. Code Loops.



## INTRODUÇÃO

Duas das quatro operações matemáticas básicas com números naturais não são associativas, o que por si só já justifica o estudo de estruturas não associativas. O surgimento de uma álgebra não associativa ocorreu no fim do século XIX, entretanto somente a partir da década de 30 do século XX, o estudo de álgebras não associativas, tais como as álgebras de Lie, as álgebras de Jordan e as álgebras alternativas, ganhou força e avançou em diversas direções.

A teoria de quasigrupos e loops também é jovem, tendo o seu início formal nos trabalhos pioneiros da matemática alemã Ruth Moufang, nos anos 30. Uma maneira intuitiva de “explicar” o que é um loop é dizer que ele é um grupo não associativo. Vale ressaltar que alguns resultados de investigações realizadas nas últimas décadas têm revelado que quasigrupos e loops possuem um importante papel em diversas especialidades da matemática, a exemplo de análise combinatória, teoria de códigos, geometria (projetiva, diferencial, hiperbólica, esférica, web), teoria de grupos (finita e de Lie) e teoria de nós. Esta teoria se enquadra no campo de estudo das Álgebras não Associativas, dentro da área de Álgebra.

A teoria de loops, em especial de loops de Moufang, é muito ampla e tem diversas aplicações. Dentre elas, uma das aplicações mais recentes é na Teoria de Loops de Código.

Griess (1986) deu início ao estudo dos Loops de Código para a construção de um grupo conhecido na literatura como “monstro”. Ele também demonstrou que um loop de código é um loop de Moufang. Os matemáticos Chein e Goodaire (1990) demonstraram que os loops de código tem uma propriedade especial: possuem um único comutador não trivial, um único associador não trivial e um único quadrado não trivial. Loops de Moufang, com uma ou mais destas propriedades de “unicidade”, têm um papel importante no estudo de loops que tem anéis de loop alternativos. Por exemplo, Chein e Goodaire provaram que um loop de Moufang, não-associativo com um único quadrado não trivial, deve ter um anel de loop alternativo sobre qualquer anel de característica dois.

Neste trabalho, apresentamos, primeiramente, alguns conceitos e resultados básicos da teoria de loops de Moufang. Também mostramos como construir loops de código e, provamos com detalhes que estes são loops de Moufang que satisfazem as propriedades de “unicidade” mencionadas acima. A seguir, mos-

tramos que loops de código possuem um único elemento central não trivial, como consequência de uma caracterização dada em (CHEIN; GOODAIRE, 1990). Por fim, determinamos uma outra caracterização relacionada com E-loops. De modo geral, mostramos como determinar quando um certo loop finito pode ser caracterizado como loop de código.

## LOOPS DE MOUFANG

Nesta seção, apresentamos algumas definições e resultados preliminares sobre loops. Para mais detalhes e demonstrações veja (GOODAIRE; JASPERS; POLCINO, 1996) e (PFLUGFELDER, 1990).

Um **loop** é um conjunto  $L$ , munido com uma operação binária (denotada por justaposição ou por  $\cdot$ ) tal que, dados quaisquer elementos  $x, y, z$ , a equação  $x \cdot y = z$  determina o terceiro elemento de modo único e tal que existe um elemento identidade bilateral (denotado por 1).

Um loop  $L$  é um loop com a propriedade inversa (denotado por loop P.I.) se cada  $x \in L$  tem um único inverso bilateral, que denotamos por  $x^{-1}$ , e se, para todo  $x, y \in L$ , o loop satisfaz  $x^{-1}(xy) = y$  e  $(yx)x^{-1} = y$ , ou seja, as propriedades de inverso à esquerda e inverso à direita, respectivamente.

**Teorema 1.** Em qualquer loop  $L$ , para quaisquer elementos  $x, y, z$  de  $L$ , as seguintes identidades (chamadas identidades de Moufang) são equivalentes:

$$((xy)x)z = x(y(xz)), \quad (1)$$

$$((xy)z)y = x(y(zy)), \quad (2)$$

$$(xy)(zx) = (x(yz))x. \quad (3)$$

Se  $L$  é um loop satisfazendo qualquer uma dessas identidades, então  $L$  é um loop P.I. que também satisfaz

$$(yx)x = yx^2, \quad x(xy) = x^2y, \quad (xy)x = x(yx).$$

Um loop  $L$  é chamado um *loop de Moufang* se ele satisfaz qualquer uma das três identidades de Moufang.

Veremos a seguir um modo equivalente de determinar se um conjunto não vazio finito munido de uma operação binária é um loop. Para isto, precisamos definir a noção de quasigrupo. Seja  $G$  um conjunto não-vazio munido de uma operação binária “ $\cdot$ ” (denotado por  $(G, \cdot)$ ), ou sim-

plesmente, por  $G$ , e chamado de grupóide) e seja a qualquer elemento fixado em  $G$ . As aplicações de translação  $L_a : G \rightarrow G$  e  $R_a : G \rightarrow G$ , são definidas respectivamente por  $L_a(x) = a.x$  e  $R_a(x) = x.a$ , para todo  $x \in G$ .

Um grupóide  $(G, \cdot)$  é chamado um quasigrupo se as aplicações  $L_a : G \rightarrow G$  e  $R_a : G \rightarrow G$  são bijeções para todo  $a \in G$ . A definição de quasigrupo é equivalente a dizermos que dados quaisquer dois elementos de  $x, y, z$  em  $G$  o terceiro pode ser unicamente selecionado em  $G$  de modo que  $x.y = z$ . Logo, temos que um loop  $(L, \cdot)$  é um quasigrupo com elemento identidade (bilateral) 1.

**Teorema 2.** Seja  $(G, \cdot)$  um grupóide finito. São equivalentes:

- (i)  $(G, \cdot)$  é um quasigrupo.
- (ii)  $L_a : G \rightarrow G$  e  $R_a : G \rightarrow G$  são injetoras para todo  $a \in G$ .
- (iii)  $L_a : G \rightarrow G$  e  $R_a : G \rightarrow G$  são sobrejetoras para todo  $a \in G$ .
- (iv) As leis de cancelamento à esquerda e à direita valem para  $(G, \cdot)$ .
- (v) Cada elemento em  $G$  aparece somente uma vez em cada linha e em cada coluna da tabela de multiplicação para  $(G, \cdot)$ .

**Definição 1.** Sejam  $x, y$  e  $z$  três elementos de um loop  $L$ . O *comutador* de  $x$  e  $y$  é o único elemento  $[x, y]$  de  $L$  que satisfaz  $xy = (yx)[x, y]$  e o *associador* de  $x, y$  e  $z$  é o único elemento  $(x, y, z)$  de  $L$  que satisfaz  $(xy)z = (x(yz))(x, y, z)$ .

Seja  $L$  um loop. O *núcleo à esquerda*, o *núcleo à direita* e o *núcleo do meio* de  $L$  são definidos respectivamente por:

$$N_l = \{a \in L \mid (a, x, y) = 1, \forall x, y \in L\}, \quad N_r = \{a \in L \mid (x, y, a) = 1, \forall x, y \in L\}$$

$$\text{e } N_m = \{a \in L \mid (x, a, y) = 1, \forall x, y \in L\}.$$

O núcleo de  $L$  é  $N(L) = N_l \cap N_r \cap N_m$  e o centro de  $L$  é  $Z(L) = \{x \in N(L) \mid [a, x] = 1, \forall a \in L\}$ .

**Proposição 1.** Todo núcleo de um loop é um subloop associativo e portanto um grupo. O centro de um loop é um grupo abeliano.

As demonstrações dos próximos resultados podem ser encontradas em (CHEIN; GOODAIRE, 1990).

**Lema 1.** Se  $L$  é um loop de Moufang,  $z \in L$  é um elemento que comuta com todo elemento de  $L$  e  $z^2 \in N(L)$ , então  $z \in Z(L)$ .

O próximo corolário será aplicado para dar uma caracterização dos loops de código (Ver Teorema 7).

**Corolário 1.** Se  $L$  é um loop de Moufang,  $z \in L$  é um elemento que comuta com todo elemento de  $L$  e  $z^2 = 1$ , então  $z \in Z(L)$ .

**Lema 2.** Se  $L$  é um loop de Moufang com um único comutador não trivial,  $\varepsilon$ , então  $\varepsilon^2 = 1$ ,  $\varepsilon \in Z(L)$ , e, para quaisquer  $x, y, z \in L$ ,  $x^2$  comuta com todo elemento de  $L$ , e  $(x, y, z)^3 = 1$  ou  $\varepsilon\varepsilon$ .

Para o nosso estudo, é muito importante saber quando o comutador e o associador (não triviais) de um loop são iguais e qual classe de loops satisfaz esta propriedade. Uma classe destes loops será apresentada no Teorema 6.

**Proposição 2.** Se  $L$  é um loop de Moufang loop com um único associador não trivial,  $f$ , e um único comutador não trivial,  $\varepsilon$ , então  $f = \varepsilon$ .

**Teorema 3.** Se  $L$  é um loop de Moufang com um único quadrado não trivial,  $e$ , então  $e^2 = 1$  e  $L$  é um grupo abeliano ou  $[L, L] = (L, L, L) = L^2 = \{1, e\} \subseteq Z(L)$ .

**Teorema 4.** Se  $L$  é um loop de Moufang com  $L^2 = \{1, \varepsilon\}$ , então, para todo  $w, x, y, z \in L$ ,

$$[xy, z] = [x, z][y, z](x, y, z)$$

$$(wx, y, z) = (w, y, z)(x, y, z).$$

**Teorema 5.** (Moufang) Se  $a, b, c$  são elementos de um loop de Moufang e  $a.b.c = a.b.c$ , o subloop gerado por  $\{a, b, c\}$  é um grupo.

Do Teorema 5, se  $(x, y, z) \neq 1$ , então  $(x', y', z') \neq 1$  para qualquer permutação  $x', y', z'$  de  $x, y, z$ . Portanto se  $L$  tem um único quadrado não trivial e desta forma, um único associador não trivial, então  $(x', y', z') = (x, y, z)$ . Em outras palavras, o associador de três elementos é independente da ordem de seus elementos. Assim, a segunda propriedade do Teorema 4 também é válida para  $(x, wy, z)$  e  $(x, y, wz)$ .

**Proposição 3.** Seja  $F$  um loop de Moufang.

1. Se  $(x, y, z)^2 = 1$  e todos os comutadores e associadores de  $F$  são centrais, então

$$[xy, z] = [x, z][y, z](x, y, z).$$

2. Se os comutadores e associadores de  $F$  são centrais, então

$$(wx, y, z) = (w, y, z)(x, y, z).$$

Usando a proposição anterior se prova que  $(x, wy, z) = (x, w, z)(x, y, z)$  e  $(x, y, wz) = (x, y, w)(x, y, z)$ . De fato, observe que  $[xy, z] = [z, xy]$ . Daí,  $(x, y, z) = (z, x, y)$ . Logo,

$$(x, y, wz) = (wz, x, y) = (w, x, y)(z, x, y) = (y, w, x)(y, z, x) = (x, y, w)(x, y, z).$$

Analogamente temos o outro caso.

Como consequência dos resultados acima, provamos o próximo resultado.

**Proposição 4.** Se os quadrados e comutadores de um loop de Moufang  $F$  são centrais, então

$$(xy)^2 = x^2y^2[x, y].$$

**Demonstração.** Temos que  $(xy)^2 = (xy)((xy)[x, y])$ . Como  $[x, y] \in Z(F)$ , obtemos  $(xy)^2 = (xy)(yx)[x, y]$ . Usando a identidade de Moufang  $xy.zx = (x.yz)x$  ficamos com  $(xy)^2 = (xy^2)x[x, y]$ , mas  $y^2 \in Z(F)$ , logo obtemos a igualdade desejada.

## LOOPS DE CÓDIGO: CONSTRUÇÃO E ALGUMAS PROPRIEDADES

Nesta seção, apresentamos uma introdução à Teoria de Loops de Código. Para isto, seja

$\mathbb{F}_2^n$  o espaço vetorial  $n$ -dimensional sobre o corpo de 2 elementos  $\mathbb{F}_2 = \{0, 1\}$ . Para vetores  $u$  e  $v$  em  $\mathbb{F}_2^n$ ,  $|v|$  (o peso de  $v$ ) denota o número de coordenadas não-nulas de  $v$  e  $|u \cap v|$  (peso de  $u$  intersecção  $v$ ) denota o número de posições nas quais as coordenadas de  $u$  e  $v$  são ambas não-nulas, ou seja, se  $u = (u_1, \dots, u_n)$  e  $v = (v_1, \dots, v_n)$ , com  $u_i, v_i \in \mathbb{F}_2$  para todo  $i = 1, \dots, n$ , então

$$|v| = |\{i \mid v_i = 1\}| \text{ e } |u \cap v| = |\{i \mid u_i = v_i = 1\}|.$$

**Definição 2.** Um código par é um subespaço  $V \subseteq \mathbb{F}_2^n$  tal que  $|v| \equiv 0 \pmod{4}$  e  $|u \cap v| \equiv 0 \pmod{2}$  para quaisquer  $u, v \in V$ .

Como exemplo de código par, consideremos o subespaço  $V \subseteq \mathbb{F}_2^7$  com base  $B = \{v_1, v_2, v_3\}$  onde

$$v_1 = (1, 1, 1, 1, 0, 0, 0), \quad v_2 = (1, 1, 0, 0, 1, 1, 0), \quad v_3 = (1, 0, 1, 0, 1, 0, 1).$$

**Definição 3.** Seja  $V$  um código par. Chamamos de "factor set" a função  $\phi: V \times V \rightarrow \{1, -1\}$  definida para todo  $u, v, w \in V$ , da forma:

$$\phi(v, v) = (-1)^{\frac{|v|}{4}}; \tag{4}$$

$$\phi(v, w) = (-1)^{\frac{|v \cap w|}{2}} \phi(w, v); \tag{5}$$

$$\phi(0, v) = \phi(v, 0) = 1; \tag{6}$$

$$\phi(v + w, u) = \phi(v, w + u) \phi(v, w) \phi(w, u) (-1)^{|v \cap w \cap u|}. \tag{7}$$

Agora denotamos por  $L(V)$  o conjunto  $V \cup (-V)$ . Vamos definir uma operação binária  $\cdot$  sobre  $L(V)$ , chamada de produto e denotada, quando conveniente, por justaposição. Sejam  $v, w \in V$ , definimos

$$v.w = \phi(v, w)(v + w), \text{ onde } \phi(v, w) \in \{1, -1\}, \\ v.(-w) = (-v).w = -(v.w), \\ (-v).(-w) = v.w. \tag{8}$$

Com o produto definido acima, se prova que  $(L(V), \cdot)$ , ou simplesmente  $L(V)$ , tem uma estrutura de loop. Loops construídos desta forma, são chamados de *loops de código*. Como a dimensão do código par  $V$  é finita, o loop de código também é finito. Logo, definimos o posto de  $L(V)$  como sendo a dimensão de  $V$ .

Como exemplo, consideremos  $V$  o código par do exemplo anterior e  $L(V) = V \cup (-V)$ , com a operação de multiplicação definida como acima. Claramente temos que  $L(V)$  é um loop de código de posto 3 com 16 elementos. Existem, a menos de isomorfismo, apenas 5 loops de código de posto 3 (PIRES, 2011).

Pretendemos provar agora que  $L(V)$  é de fato um loop. Definimos que dois elementos  $x, y$  de  $L(V)$  da forma  $x = au$  e  $y = bv$ , com  $a, b \in \{1, -1\}$ , são iguais se  $a = b$  e  $u = v$ . Sempre que denotarmos um elemento de  $L(V)$  da forma  $x = av$  significa que  $a \in \{1, -1\}$  e  $v \in V$ . Escreva o produto de dois elementos quaisquer  $x = au$  e  $y = bv$  de  $L(V)$  da forma  $xy = (ab\phi(u, v))(u + v)$ .

Seja  $L = L(V)$ . Como  $L$  é um conjunto finito, podemos aplicar o Teorema 2 para provar que  $L$  é loop. Para isto, basta provarmos que para todo  $a \in L$  que as aplicações de translação  $L_a: L \rightarrow L$  e  $R_a: L \rightarrow L$ , definidas respectivamente por  $L_a(x) = a.x$  e  $R_a(x) = x.a$ , para todo  $x \in L$  são injetoras e que  $L$  tem elemento identidade.

Com efeito, sejam  $x = au \in L$  e as aplicações de translação  $L_x: L \rightarrow L$  e  $R_x: L \rightarrow L$  definidas como antes. Sejam  $y = bv$  e  $z = cw$  elementos de  $L$  tais que  $L_x(y) = L_x(z)$ . Pela definição de  $L_x$  obtemos  $xy = xz$ , ou seja,  $(au)(bv) = (au)(cw)$ . Pela

definição do produto em  $L$  temos que  $(ab\varphi(u, v))(u + v) = (ac\varphi(u, w))(u + w)$ . Desta igualdade, vemos que  $ab\varphi(u, v) = ac\varphi(u, w)$  e que  $u + v = u + w$ , o que nos dá  $v = w$  e  $\varphi(u, v) = \varphi(u, w)$ , e, finalmente  $b = c$ . Logo,  $y = bv = cw = z$ , como queríamos. Portanto,  $L_x$  é injetora, para todo  $x \in L$ . Analogamente, prova-se que  $R_x$  é injetora, para todo  $x \in L$ . Agora,  $0$  é o elemento identidade de  $L$ , pois  $0.v = \varphi(0, v)(0 + v) = 1v = v$  e  $0.(-v) = -0.v = -1v = -v$ .

Consideramos através do lema abaixo algumas propriedades relativas ao código par  $V$ . Estas propriedades são obtidas com base num resultado mais geral apresentado no artigo (GRIESS, 1986, Lema 5, p. 225).

**Lema 3.** Sejam  $x, y, z$  quaisquer elementos de um código par  $V$ . São válidas as seguintes propriedades:

$$|x \cap y \cap (z + x)| \equiv |x \cap y \cap z| \pmod{2}; \quad (9)$$

$$\frac{1}{2}|y \cap (x + z)| \equiv \frac{1}{2}|y \cap x| + \frac{1}{2}|y \cap z| + |y \cap x \cap z| \pmod{2}; \quad (10)$$

$$\frac{1}{2}|x \cap (x + y + z)| \equiv \frac{1}{2}|x \cap y| + \frac{1}{2}|x \cap z| + |x \cap y \cap z| \pmod{2}. \quad (11)$$

Além de mostrar a existência de loops de código em seu trabalho, Griess (1986) também demonstrou que um loop de código é um loop de Moufang. Além disso, Chein e Goodaire (1990) demonstraram que um loop de código satisfaz uma propriedade bem especial relacionada aos quadrados, comutadores e associadores de seus elementos. Em outras palavras, em um loop de código existe um único quadrado não trivial, um único comutador não trivial e um único associador não trivial. Estes resultados são apresentados no teorema seguinte. No trabalho destes autores foi introduzida apenas uma ideia da demonstração. Aqui, apresentamos uma demonstração incluindo todos os detalhes necessários para o entendimento.

**Teorema 6.** O loop  $L(V)$  é loop de Moufang e valem as seguintes propriedades, para quaisquer  $u, v, w \in V$ :

$$v^2 = (-1)^{\frac{|v|}{4}} 0, \quad (12)$$

$$[u, v] = u^{-1}v^{-1}uv = (-1)^{\frac{|u \cap v|}{2}} 0, \quad (13)$$

$$(u, v, w) = ((uv)w)((u(vw))^{-1}) = (-1)^{|u \cap v \cap w|} 0. \quad (14)$$

Demonstração. Seja  $L(V) = \{1, -1\} \times V$ , onde  $V$  é um código par. Temos que provar que  $L(V)$  satisfaz qualquer uma das identidades de Moufang. Vamos provar que vale a Identidade 3 do Teorema 1:

$$(ax)(by).(cz)(ax) = [(ax).(by)(cz)](ax), \quad \forall ax, by, cz \in L(V) \quad (15)$$

Pela definição do produto de  $L(V)$ , para verificar esta identidade de Moufang basta provarmos que

$$\varphi(x, y)\varphi(z, x)\varphi(x + y, z + x) = \varphi(y, z)\varphi(x, y + z)\varphi(x + y + z, x), \quad (16)$$

onde  $\varphi: V \times V \rightarrow \{1, -1\}$  é a função *factor set* (conforme Definição 3).

De fato, por um lado

$$(ax)(by).(cz)(ax) = (abac\varphi(x, y)\varphi(z, x)\varphi(x + y, z + x))(y + z).$$

Por outro lado,

$$[(ax).(by)(cz)](ax) = (abac\varphi(y, z)\varphi(x, y + z)\varphi(x + y + z, x))(y + z).$$

Agora, para demonstrar a Equação (16), primeiramente, observamos que pelas Equações (7) e (9),

$$\varphi(x, y)\varphi(x + y, z + x) = \varphi(y, z + x)\varphi(x, x + y + z)(-1)^{|x \cap y \cap z|}. \quad (17)$$

A partir da Equação (5), obtemos:  $\varphi(x, x + y + z) = \varphi(x + y + z, x)(-1)^{|x \cap (x + y + z)|}$ . Usando este resultado juntamente com a Equação (11), podemos reescrever a Equação (17) da forma:

$$\phi(x, y)\phi(x+y, z+x) = \phi(y, z+x)\phi(x+y+z, x)(-1)^{\frac{|x\cap y|}{2} + \frac{|x\cap z|}{2}} \quad (18)$$

Multiplicando cada termo da última equação por  $\phi(z, x)$ , obtemos o seguinte resultado:

$$\phi(x, y)\phi(z, x)\phi(x+y, z+x) = \phi(y, z+x)\phi(z, x)\phi(x+y+z, x)\phi(z, x)(-1)^{\frac{|x\cap y|}{2} + \frac{|x\cap z|}{2}} \quad (19)$$

Agora, observe que a partir das Equações (5) e (7) obtemos:

$$\phi(x, z)\phi(x, z+y) = \phi(x+z, y)\phi(z, y)(-1)^{|x\cap y\cap z|};$$

$$\phi(y, x+z) = \phi(x+z, y)(-1)^{\frac{x\cap y}{2} + \frac{y\cap z}{2} + |x\cap y\cap z|}.$$

Assim encontramos a seguinte equação:

$$\phi(z, x)\phi(y, z)\phi(x, y+z)\phi(y, z+x)(-1)^{\frac{|x\cap y|}{2} + \frac{|x\cap z|}{2}} = 1. \quad (20)$$

Para concluirmos esta primeira parte da demonstração do teorema, basta ver que a Equação (20) é equivalente a equação seguinte:

$$\phi(y, z+x)\phi(z, x)\phi(x+y+z, x)\phi(z, x)(-1)^{\frac{|x\cap y|}{2} + \frac{|x\cap z|}{2}} = \phi(y, z)\phi(x, y+z)\phi(x+y+z, x). \quad (21)$$

Logo, a Equação (16) segue diretamente das Equações (19) e (21) e, portanto,  $L(V)$  é um loop de Moufang. Agora iremos provar as propriedades de "unicidade" relacionadas ao quadrado (comutador, associador) enunciadas no Teorema.

**Propriedade do quadrado:** Como  $v^2 = \phi(v, v)(v+v)$ , e por definição  $\phi(v, v) = (-1)^{\frac{|v|}{4}}$ , então

$$v^2 = (-1)^{\frac{|v|}{4}} 0.$$

**Propriedade do comutador:** Por definição,

$$\begin{aligned} u.v &= \phi(u, v)(u+v) = (-1)^{\frac{|u\cap v|}{2}} \phi(v, u)(v+u) = \\ &= (-1)^{\frac{|u\cap v|}{2}} 0.\phi(v, u)(v+u) = (-1)^{\frac{|u\cap v|}{2}} 0.(v.u). \end{aligned}$$

Desta forma, pela unicidade do comutador, obtemos que

$$[u, v] = (-1)^{\frac{|u\cap v|}{4}} 0.$$

**Propriedade do associador:** Por definição, por um lado

$$(uv)w = \phi(u, v)\phi(u+v, w)(u+v+w) = (-1)^{|u\cap v\cap w|} \phi(v, w)\phi(u, v+w)(u+v+w).$$

Por outro lado,

$$(-1)^{|u\cap v\cap w|} 0.(uvw) = (-1)^{|u\cap v\cap w|} 0.(\phi(v, w)\phi(u, v+w)(u+v+w)),$$

ou seja,

$$(-1)^{|u\cap v\cap w|} 0.(uvw) = (-1)^{|u\cap v\cap w|} \phi(v, w)\phi(u, v+w)\phi(0, u+v+w)(u+v+w).$$

Logo,  $(-1)^{|u\cap v\cap w|} 0.(uvw) = (uv)w$ . Desta forma, pela unicidade do associador, concluímos que

$$(u, v, w) = (-1)^{|u\cap v\cap w|} 0.$$

## CARACTERIZAÇÕES DE LOOPS DE CÓDIGO

**Definição 4.** Um loop de Moufang  $L$  é chamado *E-loop* se existe um subloop central  $Z$  de 2 elementos tal que  $L/Z \in A$ , onde  $A$  é a variedade de grupos com identidade  $x^2 = 1$ .

**Teorema 7.** Seja  $L$  um loop de Moufang. O loop  $L$  é um E-loop se, e somente se,  $|L^2| \leq 2$ .

Demonstração. Seja  $L$  um E-loop. Então existe um subloop central  $Z = \{1, x\}$  tal que  $\bar{y}^2 = \bar{1}$ , para todo  $\bar{y} \in L/Z$ . Vamos mostrar que  $L^2 \subseteq Z$ . Se supormos que  $e \in L^2$  seja tal que  $e \neq 1$  e  $e \neq x$ , então existirá  $u \in L$  tal que  $u^2 = e$ . Daí teremos  $\bar{u}^2 \neq \bar{1}$  contradizendo a definição de  $L$ . Logo  $u^2 = 1$  ou  $u^2 = x$ , ou seja  $L^2 \subseteq Z$ . Portanto,  $|L^2| \leq 2$ . Reciprocamente, suponhamos que  $|L^2| \leq 2$ . Caso  $|L^2| = 1$ , então  $L$  é um loop de Moufang comutativo. De fato, para quaisquer  $x, y \in L$ ,  $[x, y] = x^{-2}(xy^{-1})^2y^2 = 1$ . Logo, pelo Corolário 1, qualquer elemento de  $L$  está no centro de  $L$ . Considere então um subloop central qualquer de dois elementos, digamos  $Z = \{1, x\}$ , onde  $x \in L$  é qualquer. Assim,  $L/Z \in A$ , onde  $A$  é a variedade de grupos tais que  $y^2 = 1$ , para todo  $y$ . Agora suponhamos que  $L^2 = \{1, e\}$  e que  $L^2 = Z$ . Pelo Teorema 3,  $L^2$  é um subloop central e assim, podemos ver com simples cálculos que  $\bar{y}^2 = \bar{1}$ , para todo  $\bar{y} \in L/Z$ .

**Teorema 8.** (CHEIN; GOODAIRE, 1990) Um loop finito  $L$  é isomorfo a um loop de código se e somente se  $L$  é um loop de Moufang com  $|L^2| \leq 2$ .

Como consequência desta caracterização e do Teorema 3, observamos que se  $L$  é um loop de código não associativo, então  $|Z(L)| = 2$ .

Finalizamos este trabalho dando uma outra caracterização dos loops de código. A demonstração segue diretamente do Teorema 7 e do Teorema 8.

**Teorema 9.** Um loop de Moufang  $L$  é um loop de código se, e somente se,  $L$  é um E-loop finito.

Vários dos resultados deste trabalho são aplicados no estudo dos problemas de classificação e representação de loops de código. Para mais detalhes, veja os trabalhos de (PIRES, 2011) e (GRISHKOV; PIRES, 2018).

## CONCLUSOES

Neste trabalho, apresentamos uma introdução e construção de loops de código. Provamos com detalhes que loops de código são loops de Moufang com um único comutador não trivial, um único associador não trivial e um único quadrado não trivial. Por fim, exibimos uma caracterização destes loops dada em (CHEIN; GOODAIRE, 1990) e, provamos uma outra caracterização relacionada com os E-loops.

## REFERÊNCIAS

- CHEIN, O.; GOODAIRE, E. G. Moufang Loops with a Unique Nonidentity Commutator (Associator, Square). **J. Algebra** 130, 369-384, 1990.
- GOODAIRE, E. G.; JASPERS, E.; POLCINO, C. **Alternative Loop Rings**. North Holland Math, Studies N.184, Elsevier, Amsterdam, 1996.
- GRIESS, R. L. Code loops. **J. Algebra** 100, 224-234, 1986.
- GRISHKOV A.; PIRES, R. M. Variety of loops generated by code loops. **International Journal of Algebra and Computation**, Volume 28, No. 1, 163-177, 2018.
- PFLUGFELDER, H. O. **Quasigroups and Loops: An Introduction**. Berlin, 1990.
- PIRES, R. M. **Loops de código: automorfismos e representações**. 2011. Tese (Doutorado em Matemática) – IME/USP, São Paulo, 2011.

## DADOS DA AUTORA

**Profa. Dra. Rosemary Miguel Pires** (rosemarypires@id.uff.br). Universidade Federal Fluminense

