



# EAMC2020

XIII Encontro Acadêmico de Modelagem Computacional

Laboratório Nacional de Computação Científica - LNCC/MCTIC

17 a 21 de Fevereiro de 2020.

## Livro de resumos



### **XIII Encontro Acadêmico de Modelagem Computacional**

Laboratório Nacional de Computação Científica - LNCC/MCTIC

17 a 21 de Fevereiro de 2020. Petrópolis - RJ.

#### **Comitê Organizador**

Ana Luiza Martins Karl - LNCC

Bárbara Raquel Mendoça Rezende - UFF

Guilherme Guilhermino Neto - UFJF

João Vitor de Oliveira - LNCC

Juan Diego dos Santos Heringer - IPRJ/UERJ

Luis Alonso Mansilla Alvarez - LNCC

Matheus Muller Pereira da Silva - LNCC

#### **Comitê Científico**

Caio César Graciani Rodrigues - LNCC

Grazione de Souza Boy - IPRJ/UERJ

Gregório Kappaun Rocha - IFF

João Antônio Recio da Paixão - UFRJ

Juliana Vianna Valério - UFRJ

Lucas dos Santos Fernandez - LNCC

Luziane Ferreira de Mendonça - UFRJ

Rafael Alves Bonfim de Queiroz - UFJF

Yoisell Rodríguez Núñez - UFF

### XIII Encontro Acadêmico de Modelagem Computacional

Laboratório Nacional de Computação Científica - LNCC/MCTIC

17 a 21 de Fevereiro de 2020. Petrópolis - RJ.

#### Apresentação

O Encontro Acadêmico de Modelagem Computacional do LNCC (EAMC/LNCC) é um evento dedicado à Modelagem Computacional, que busca promover maior integração entre os alunos, docentes, pesquisadores e demais profissionais da área de Ciência e Tecnologia. O EAMC é inteiramente organizado por discentes do LNCC e de outras instituições colaboradoras, sendo composto por minicursos, mesas redondas, palestras, além de apresentações de trabalhos em formato pôster e apresentações orais.

Criado em 2007, o EAMC tem como principal objetivo proporcionar um ambiente científico que promova a divulgação científica, estimule a cooperação entre os profissionais da área, além de fomentar a multidisciplinaridade na formação de novos profissionais. O EAMC abrange diferentes áreas do conhecimento, tais como Computação Científica, Controle e Filtragem de Sistemas Dinâmicos, Modelagem de Biossistemas e Bioinformática, Modelagem de Circulação e Transporte, Modelagem de Equilíbrio e Otimização e, mais recentemente, Ciência de Dados e áreas correlatas, como Inteligência Artificial.

Em 2020, o XIII Encontro Acadêmico de Modelagem Computacional (EAMC) foi realizado no Laboratório Nacional de Computação Científica, em Petrópolis/RJ, nos dias 17 a 21 de fevereiro. Em sua 13a edição, o EAMC comemorou os 40 anos de funcionamento do Laboratório Nacional de Computação Científica (LNCC/MCTIC) e os 20 anos da Pós-Graduação em Modelagem Computacional do LNCC.

### **XIII Encontro Acadêmico de Modelagem Computacional**

Laboratório Nacional de Computação Científica - LNCC/MCTIC

17 a 21 de Fevereiro de 2020. Petrópolis - RJ.

#### **Agradecimentos**

O Comitê Organizador do XIII EAMC gostaria de agradecer à comunidade do LNCC, entre pesquisadores, professores, corpo discente e funcionários pela participação e dedicação ao evento.

Em particular, gostaríamos de expressar profunda gratidão à professora Sandra Malta, maior incentivadora deste evento, pela ajuda durante a organização, bem como à secretaria do Programa de Pós Graduação do LNCC, pelo apoio concedido – em particular à secretária Roberta Machado pela ajuda ativa e constante no decorrer do encontro.

Agradecemos de forma especial a disponibilidade e contribuição dos membros do Comitê Científico, dos palestrantes convidados e professores dos minicursos. Agradecemos também também aos pós-doutorandos do LNCC: Caio Graciani, Karina Baptista dos Santos, Lucas dos Anjos e Lucas Fernandez pela valiosa ajuda na avaliação dos trabalhos apresentados durante o evento.

Finalizamos os agradecimentos esperando que o XIII EAMC tenha sido de grande valia à todos os participantes, sem os quais o sucesso dessa edição não seria possível.

Comitê Organizador

XIII Encontro Acadêmico de Modelagem Computacional - XIII EAMC  
Laboratório Nacional de Computação Científica - LNCC

## Trabalhos completos

	Page
<i>Modelagem epidemiológica da dispersão da tuberculose pulmonar</i> Amanda Araújo Serrão de Andrade, Gabrieli Dutra Silva, Vinícius Prata Klôh, José Karam Filho	7
<i>Método primal-dual na otimização agrícola</i> Angel Ramon Sanchez Delgado, Sergio Drummond Ventura, Pedro Marcio Ferreira	18
<i>Predição para recomendação de filmes com base no agrupamento pelo conteúdo do script</i> Henrique Matheus Ferrira da Silva, Rafael Silva Pereira	30
<i>Proposta de uma heurística construtiva para o projeto de circuitos combinacionais</i> Hugo Assis de Oliveira, Francisco Augusto Lima Manfrini	39
<i>Segmentação de imagens digitais via modelos variacionais com restrições reladas</i> Italo Messias Felix Santos, Gilson Antonio Giraldi, Abimael Dourado Loula *Avaliado como melhor trabalho na seção oral	47
<i>Otimização biobjetivo de um modelo para controle de qualidade de processos dinâmicos com tamanho de amostra variável</i> João Gabriel Malaguti, João Vitor de Lima, Rafael de Souza Terra	57
<i>Análise de sensibilidade topológica aplicada ao processo de fraturamento hidráulico</i> Jorge Morvan Marotte Luz Filho, Antonio André Novotny	65
<i>Simplifications of Navier-Stokes in journal bearing simulation</i> Júlia de Araújo Mota, Juliana Vianna Valério, Flavio Medeiros Rangel	76
<i>Deep neural network for single event effects classification</i> Juliano Alves de Oliveira, Marco Antonio Assis de Mello, Marcilei Aparecida Guazzelli da Silveira, Rudolf Theoderich Buhler, Vitor Angelo Paulino de Aguiar, Nilberto Heder Medina, Renato Camargo Giacomini	86
<i>Análise de características na tarefa de segmentação interativa de imagens digitais</i> Thiago Amorim de Faccio, Marilaine Colnago, Wallace Casaca	103
<i>Filtros ópticos otimizados por algoritmos genéticos associados à lógica nebulosa</i> Wilton Moreira Ferraz Junior, Carlos Henrique da Silva Santos, Marcos Sergio Gonçalves	113

---

Resumos	Page
<i>Análise do perfil de mutações de resistência em paciente HIV positivo multirresistente aos antirretrovirais</i> Ana Luiza Martins Karl, Adriana de Oliveira Afonso	123
<i>Estudo comparativo de métodos de limitação e fluxo para leis de conservação</i> Ana Caroline Francisca da Silva, Daniel Gregorio Alfaro Vigo	124
<i>Redes neurais artificiais: modelagem computacional da previsão de volume de uma usina hidrelétrica</i> Bárbara Raquel Mendonça Rezende, Eliane da Silva Christo, Fernando Tadeu Pereira de Medeiros	125
<i>Boa colocação e estabilidade via semigrupos lineares para o problema termoelástico</i> Bruna Thais Silva Sozzo, Jaime Edilberto Muñoz Rivera	127
<i>Simulação da produção de óleo por poços horizontais hidraulicamente fraturados em reservatórios de baixa permeabilidade</i> Camila Lima Chung, Ralph Alves Bini da Silva Almeida, Leonardo Figueira Werneck, Grazione de Souza, Helio Pedro Amaral Souto	128
<i>Aplicação de redes complexas para estudo da correlação geográfica de doenças</i> Caroline de Oliveira Costa Souza Rosa, Artur Ziviani	129
<i>Emprego do OpenFlow em SmartGrid: Uma abordagem utilizando o ambiente GENI</i> Cássia da Silva Pereira, Raul Vitor Arantes Monteiro, Ailton Akira Shinoda	130
<i>Problem formalization of embedding methods for dynamic graphs</i> Claudio Daniel Tenório de Barros, Artur Ziviani	131
<i>Programação genética para problemas de classificação</i> Gianni Ribeiro Silvério da Conceição, Camila Silva de Magalhães	132
<i>Modelagem computacional da transmissão da dengue através do modelo epidêmico compartmental SIRS</i> Gicelle Martins da Cruz, João David Mangabeira Barbosa, Magna Pereira Amorim, Yoisell Rodríguez Núñez, Vanessa da Silva Garcia	133
<i>Does the Ataxia group have genetic similarities?</i> Gustavo Simões Carnival, Sérgio Vale Aguiar Campos	135
<i>Análise mutacional via implementação do algoritmo soma com transporte</i> José dos Santos Fernandes, Mariane Moreira de Souza, Anderson José de Oliveira	136
<i>Modelagem computacional do câncer de mama: Estudo comparativo via fatores de tratamento</i> Josué Cleber de Azevedo Custódio, Felipe Bezerra de Azevedo, Yoisell Rodríguez Núñez, Vanessa da Silva Garcia	137

---

<i>Numerical simulation of non-isothermal water-oil flow in petroleum reservoirs</i>	
Juan Diego dos Santos Heringer, Grazione de Souza, Helio P. Amaral Souto	139
<i>Neuroevolução aplicada ao trading de criptomoedas</i>	
Lincon Vidal	
*Avaliado como melhor trabalho na seção de pôster	140
<i>Heat transfer solution for circular channels with axial heat diffusion using integral transforms</i>	
Lívia Mantuano Corrêa, Daniel José Nahid Mansur Chalhub	142
<i>Identificação de bolsões específicos como potenciais inibidores da proteína NS2B-NS3 dos principais flavivírus de importância médica</i>	
Luan Antônio de Moraes, Gabriela de Lima Menezes, Roosevelt Alves da Silva	143
<i>Modelagem da vascularização arterial do rim humano</i>	
Luís Fernando Mendes Cury, Caio Sergio Bianchi Reis Guimarães, Mauricio Younes-Ibrahim, Pablo Javier Blanco	144
<i>Controle ótimo aplicado a um modelo epidemiológico compartmental de transmissão da dengue</i>	
Magna Pereira Amorim, Yoisell Rodríguez Núñez, Panteros Rodríguez Bermúdez, Vanessa da Silva Garcia	145
<i>Modelagem computacional de processos de contaminação de aquíferos considerando biodegradação e adsorção</i>	
Marcelo Lemos da Silva, João Gabriel de Souza Debossam, Grazione de Souza	147
<i>Study of Ideonella sakaiensis PETase enzyme via computational molecular biophysics</i>	
Marcio Fernandes da Silva, Stéphano Andrade de Vasconcelos, Marcos Serrou do Amaral	148
<i>Modelagem computacional do crescimento tumoral mamário com quimioterapia</i>	
Matheus Fonseca da Silva, Paulo Raul Cardoso Faria, Yoisell Rodríguez Núñez, Vanessa da Silva Garcia	150
<i>Engenharia virtual da PETase termoestável da Ideonella sakaiensis para degradação altamente eficiente do PET</i>	
Stéphano Andrade de Vansoncelos, Sérgio Leandro Espindola Preza, Marcio Fernandes da Silva, Marcos Serrou do Amaral	152
<i>Evolução diferencial para classificação de dados</i>	
Taciano dos Passos Ferreira Pinheiro, Camila Silva de Magalhães	154
<i>Atenuação da dimensionalidade de dados para detecção de anomalias através da utilização de mapas auto organizáveis</i>	
Thiago Moeda Sant'Anna, Fabio Porto	155



# Modelagem Epidemiológica da Dispersão da Tuberculose Pulmonar

Amanda Araújo Serrão de Andrade<sup>1</sup>, Gabrieli Dutra Silva<sup>1</sup>, Vinícius Prata Klôh<sup>1</sup>  
e José Karam Filho<sup>1</sup>

<sup>1</sup> Laboratório Nacional de Computação Científica, Petrópolis/RJ, Brazil

---

## Resumo

A Tuberculose Pulmonar (TB) é uma doença infecto-contagiosa grave que afeta os pulmões. Apesar de ser uma doença com um potencial alto de prevenção e cura, ainda é um grande problema de saúde pública no Brasil. Desta forma, o objetivo deste trabalho é identificar o padrão de dispersão da TB e o possível impacto da administração de quimioprofilaxia\*secundária. Para tanto, foram formulados dois modelos matemático-epidemiológicos SEIR (Suscetíveis-Expostos-Infectados-Removidos) que tiveram como base para simulações os dados oficiais da propagação da doença na cidade do Rio de Janeiro durante o ano de 2018. Os modelos foram simulados em três cenários diferentes onde não existe o uso da quimioprofilaxia, a quimioprofilaxia é administrada após o estabelecimento da doença e onde a quimioprofilaxia é administrada nos primeiros casos. Concluímos que os modelos apresentados auxiliaram na identificação do impacto positivo de medidas de quimioprofilaxia secundária na interrupção dos padrões de dispersão da doença na população simulada, estando em concordância com estudos clínicos e as recomendações de muitos países. Além disso, os modelos criados, apesar de simples, têm um alto valor preditivo e a enorme capacidade de serem adaptados e aplicados a outros possíveis cenários, que englobem diferentes aspectos da dispersão da TB.

**Palavras-chave:** Modelagem epidemiológica, tuberculose pulmonar, modelo SEIR

---

## 1 INTRODUÇÃO

A tuberculose (TB) é uma doença infecto-contagiosa grave causada pela bactéria *Mycobacterium tuberculosis*, que afeta principalmente os pulmões, mas também pode ocorrer em outros órgãos do corpo, como ossos, rins e meninges (Souza et al., 2019) a. Em 2018, no *Global Tuberculosis Report* (WHO, 2016), estimou-se que 10 milhões de pessoas adoeceram por TB no mundo e que a doença tenha causado 1,3 milhão de

\*Quimioprofilaxia: Medida terapêutica para a prevenção da infecção pelo e para evitar o desenvolvimento da doença nos indivíduos infectados.

mortes, o que mantém a TB entre as 10 principais causas de morte no planeta. No Brasil, no mesmo período, ocorreram 72.788 casos novos de TB, sendo registrados 4.534 óbitos (da Saúde, 2019).

A transmissão da TB acontece de forma direta, de pessoa a pessoa, principalmente através do ar sendo a principal fonte de infecção o indivíduo com a forma pulmonar da doença, que elimina bactérias para o exterior (Nogueira et al., 2012). Com isso, um dos fatores mais importantes para a transmissão do *Mycobacterium tuberculosis* é a aglomeração de pessoas em espaços mal ventilados, estando intimamente ligada às condições de vida da população (WHO, 2016).

Além do tratamento medicamentoso que acontece após o início dos sintomas, indivíduos expostos à bactéria ou em risco de exposição podem receber a quimioprofilaxia, que consiste na administração de isoniazida para prevenir a infecção (quimioprofilaxia primária) ou para evitar o desenvolvimento da doença nos indivíduos expostos (quimioprofilaxia secundária), desta forma, interrompendo a cadeia de transmissão (Petersen et al., 2019).

Apesar de ser uma doença com um potencial alto para sua prevenção e cura, a TB é ainda hoje um grande problema de saúde pública (WHO, 2016). O Brasil segue a proposta da Organização Mundial da Saúde em relação às prioridades referentes à detecção precoce de casos de TB, ao tratamento do paciente e à sua conclusão com fechamento propício: a cura e o fim do ciclo de transmissão (MS, 2016). Desse modo, faz-se necessário ações de vigilância epidemiológica e de entendimento da formação de redes de contágio e disseminação da doença em uma população suscetível, além de estudar de modo dinâmico a transmissão dessa patologia do ponto de vista quantitativo, com o intuito de decidir quais as medidas imprescindíveis para a sua contenção (Brauer, 2017).

Os modelos matemáticos e as simulações computacionais desses modelos são ferramentas úteis para construir e testar teorias e conjecturas de avaliação quantitativa (Brauer, 2017). No presente trabalho, temos como principal hipótese que a aplicação do modelo matemático epidemiológico SEIR (Oliveira et al., 2008), no estudo da dinâmica da dispersão de Tuberculose Pulmonar, nos permita compreender as características da dispersão dessa doença infecciosa na cidade do Rio de Janeiro, Estado do Rio de Janeiro - Brasil.

Portanto, o objetivo deste trabalho é identificar o padrão de dispersão de Tuberculose Pulmonar. São propostos dois modelos conceituais descritos em linguagem matemática e computacional que utilizam dados epidemiológicos reais e atuais, os quais refletem a realidade da propagação da doença na cidade do Rio de Janeiro (capital do Estado do Rio de Janeiro - Brasil), no ano de 2018, para gerar simulações computacionais acerca da epidemiologia da Tuberculose Pulmonar. Além disso, este trabalho visa analisar o impacto do uso da quimioprofilaxia secundária na dispersão da Tuberculose Pulmonar.

## 2 TRABALHOS RELACIONADOS

Foi realizado um estudo na literatura científica vigente de publicações comprendendo o período de 2008 à 2019, onde (Souza et al., 2019), (WHO, 2016),(da Saúde, 2019), (Nogueira et al., 2012), (Petersen et al., 2019), (MS, 2018), (Soares et al.,



2017) e (MS, 2016) analisaram diferentes aspectos da TB, incluindo: patogênese, transmissão, resposta imune, e principalmente, epidemiologia, incluindo dados estatísticos atuais a respeito desse tema. A pesquisa na literatura auxiliou a identificação do problema e motivou a aplicação da modelagem computacional. Assim, foram estudados trabalhos a respeito da Modelagem Matemático-Epidemiológica e sua aplicação às doenças infecciosas ((Brauer, 2017), (Huppert and Katriel, 2013), (Kretzschmar, 2016), (Oliveira et al., 2008), (Aparicio and Castillo-Chavez, 2009)). Portanto, com esses estudos foi possível selecionar o modelo que melhor se enquadra na dinâmica do problema.

### 3 METODOLOGIA

Nesta seção são apresentados os modelos conceituais e matemáticos, as considerações e hipóteses adotadas e a implementação e simulação computacionais.

#### 3.1 Modelos

Os modelos epidemiológicos clássicos usam abordagens compartimentais para descrever a dispersão de doenças. Nesse tipo de modelo, a população é dividida em diferentes compartimentos, que refletem o estado em que os indivíduos se encontram, e são feitas suposições sobre a natureza e a taxa de tempo de transferência de um compartimento para outro, obedecendo o princípio de conservação de massa/equilíbrio, cada indivíduo pode estar apenas em um compartimento por unidade de tempo (Huppert and Katriel, 2013).

##### 3.1.1 Modelo Conceitual

A partir das características epidemiológicas da Tuberculose Pulmonar, o modelo a ser utilizado no presente trabalho é o modelo SEIR adaptado, constituído por um sistema de equações diferenciais ordinárias, em que a variável independente é o tempo (Oliveira et al., 2008). Neste modelo, a espécie humana é dividida em quatro compartimentos:

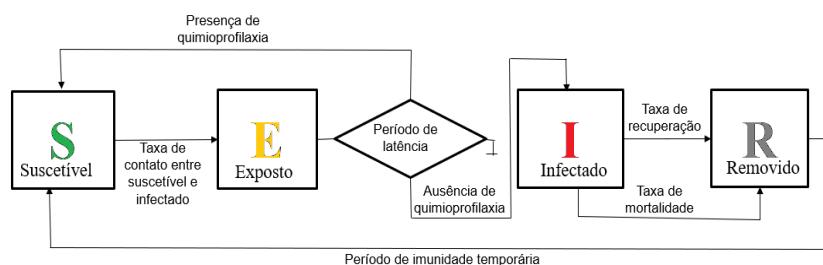
- **Susceptíveis (S):** inclui todos os indivíduos que podem contrair a infecção, caso sejam expostos a ela;
- **Expostos (E):** inclui os indivíduos que estão no período de latência, ou seja, aqueles que já foram infectados mas que ainda não são capazes de transmitir a doença; na nossa adaptação do modelo, foram incluídas taxas de ausência ou presença de tratamento profilático. Se a pessoa recebeu o tratamento profilático, ela volta ao grupo de suscetíveis, pois o tratamento impede o desenvolvimento da doença. O oposto também é verdadeiro, se não recebeu a profilaxia, é transferido para o compartimento dos Infectados.
- **Infectados (I):** indivíduos capazes de transmitir a doença através de contato com um indivíduo suscetível;
- **Removidos (R):** classe de indivíduos que se recuperam ou morrem. Os recuperados, após o fim do período de imunidade temporária, voltam para a classe dos suscetíveis, podendo ter novamente a doença.

No Modelo 1, as interações entre os compartimentos acima descritos foram modeladas assumindo as seguintes hipóteses:

- i)* A população está isolada, isto é, não possui contato com membros de outras populações e nem sofre a influencia de migrações;
- ii)* Os indivíduos infectados recebem o tratamento imediato;
- iii)* Os indivíduos recuperados ficam imunes por um período de tempo, tornando-se suscetíveis novamente;
- iv)* Não se considera a resistência do patógeno ao tratamento;
- v)* O tamanho da população é constante. Isto significa que a taxa de natalidade e a taxa de mortalidade por causas naturais são iguais e não variam com o tempo;
- vi)* A população é homogênea. Não se leva em consideração idade, gênero, estrutura social ou diferentes níveis de suscetibilidade imunológica;

A Figura 1 apresenta os fatores que motivam a troca de estados dos indivíduos. Inicialmente, as pessoas que pertencem ao compartimento dos suscetíveis, com o passar do tempo e de acordo com a taxa de contato com infectados, passam por um período latente (compartimento dos expostos) em que: caso tenham recebido a quimioprofilaxia, não desenvolvem a doença e voltam a ser suscetíveis a ela; caso não tenham recebido a quimioprofilaxia, desenvolvem a doença e são transferidos para o compartimento dos infectados, tornando-se capazes de transmitir a doença. A mudança de indivíduos infectados para o estado dos removidos se dá devido a cura pelo tratamento (que segue a taxa de recuperação) e a mortalidade devido à doença. Após o período de imunidade temporária, um indivíduo recuperado volta a ser suscetível à doença.

**Fig. 1:** Modelo 1 - Fatores de troca de estado. Fonte própria.



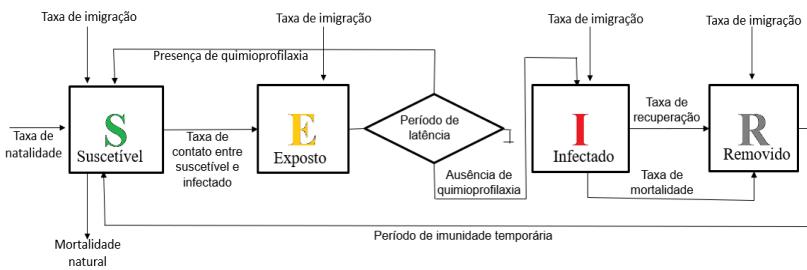
O Modelo 1 apresentado, ainda que simples, atende as características fundamentais da epidemiologia da doença, como será mostrado na Seção 4. A simplicidade do modelo permite o relaxamento das suposições e modificações dos parâmetros a fim de adaptar diferentes novas hipóteses. Para demonstrar a possibilidade de adaptação do Modelo 1, também apresentamos um segundo modelo, no qual consideraremos as interações entre os diferentes compartimentos assumindo, além das hipóteses *i* a *iv* do Modelo 1, as duas novas hipóteses a seguir:



- i) A população sofre influência da migração. O indivíduo migra para um dos quatro compartimentos do modelo;
- ii) Só há morte natural de indivíduos suscetíveis.

Para o Modelo 2, a dinâmica de indivíduos (Figura 2) entre os diferentes compartimentos ocorre de forma similar ao Modelo 1 inicialmente apresentado; entretanto, os indivíduos imigram em diferentes estágios, podendo compor o grupo dos suscetíveis, expostos, infectados ou removidos.

**Fig. 2:** Modelo 2 - Fatores de troca de estados, natalidade, mortalidade natural e imigração. Fonte própria.



### 3.1.2 Modelos Matemáticos

O modelo SEIR, aplicado ao problema proposto, pode ser representado através de um Sistema de Equações Diferenciais, em que a variável independente é o tempo ( $t$ ). Esse sistema de equações visa traduzir para a linguagem matemática as suposições epidemiológicas apresentadas no desenho do modelo conceitual. Para o Modelo 1, a população  $N$  é dada pela Equação 1 e o modelo SEIR é dado por um sistema de equações diferenciais (Sistema 3). Vale ressaltar que nos compartimentos há variação temporal, enquanto que a população total não varia.

Para o Modelo 2, levando em consideração as hipóteses apresentadas anteriormente, a população  $N$  varia no tempo e é dada pela Equação 2. Assim, o modelo SEIR para a dispersão da TB também é dado por um sistema de equações diferenciais (Sistema 4). Porém, vale ressaltar que devido a taxa de imigração de cada compartimento e as taxas de natalidade e mortalidade dos suscetíveis, tanto a população quanto os compartimentos variam no tempo.

$$N = S(t) + E(t) + I(t) + R(t) \quad (1)$$

$$\begin{cases} \frac{\partial S}{\partial t} = -\beta SI + \epsilon\alpha_a E + \gamma(\delta I) \\ \frac{\partial E}{\partial t} = \beta SI - (\epsilon\alpha_a + \epsilon\alpha_b)E \\ \frac{\partial I}{\partial t} = -\delta I - \tau I + \epsilon\alpha_b E \\ \frac{\partial R}{\partial t} = \delta I + \tau I - \gamma(\delta I) \end{cases} \quad (3)$$

$$N(t) = S(t) + E(t) + I(t) + R(t) \quad (2)$$

$$\begin{cases} \frac{\partial S}{\partial t} = -\beta SI + \epsilon\alpha_a E + \gamma(\delta I) + \varphi_s S - \lambda S + v S \\ \frac{\partial E}{\partial t} = \beta SI - (\epsilon\alpha_a + \epsilon\alpha_b)E + \varphi_e E \\ \frac{\partial I}{\partial t} = -\delta I - \tau I + \epsilon\alpha_b E + \varphi_i I \\ \frac{\partial R}{\partial t} = \delta I + \tau I - \gamma(\delta I) + \varphi_r R \end{cases} \quad (4)$$

onde:

$\beta$  = Taxa de contato prolongado por unidade de tempo (quando existe o encontro entre um indivíduo suscetível e um infectado);

- $\epsilon$  = período de latência (Taxa, por unidade de tempo em que os indivíduos expostos se tornam infecciosos);
- $\alpha_a$  = presença de quimioprofilaxia;
- $\alpha_b$  = ausência de quimioprofilaxia;
- $\gamma$  = período de imunidade temporária;
- $\delta$  = taxa de recuperação por unidade de tempo;
- $\tau$  = taxa de mortalidade por unidade de tempo, devido à doença;
- $\lambda$  = taxa de mortalidade natural;
- $v$  = taxa de natalidade;
- $\varphi_s$  = taxa de imigração de indivíduos na condição de suscetível;
- $\varphi_e$  = taxa de imigração de indivíduos na condição de exposto;
- $\varphi_i$  = taxa de imigração de indivíduos na condição de infectado;
- $\varphi_r$  = taxa de imigração de indivíduos na condição de removido.

Para o estudo de caso, consideramos os dados estatísticos relatados no Boletim Epidemiológico “Brasil Livre da Tuberculose: evolução dos cenários epidemiológicos e operacionais da doença” (da Saúde, 2019), publicado pelo Ministério da Saúde, a partir dos dados do Sistema de Informação de Agravos de Notificação (SINAN), do Sistema de Informação de Tratamentos Especiais de Tuberculose (SITE-TB) e do Sistema de Informações sobre Mortalidade (SIM). Os indicadores operacionais utilizados no presente trabalho corresponderam a dados obtidos através de vigilância epidemiológica da primeira à última semana de 2018 na cidade Rio de Janeiro, capital do estado do Rio de Janeiro (RJ), onde foram registrados 6.014 casos de TB, o coeficiente de incidência foi de 89,9 casos/100 mil habitantes, coeficiente de mortalidade foi de 4,2 casos/100 mil habitantes e a proporção de cura com confirmação laboratorial foi de 59,7 por cento (da Saúde, 2019). Os dados obtidos no levantamento epidemiológico e na revisão bibliográfica serão incluídos no modelo matemático, da seguinte forma:

$\beta$  = Taxa de contato. O valor associado a este parâmetro será o coeficiente de incidência, dado por (da Saúde, 2019).

$\epsilon$  = Período de latência. O valor utilizado será gerado através de uma triangulação dos períodos de latência mais encontrados na literatura: de 4 a 10 semanas.

$\alpha_a$  e  $\alpha_b$  = Presença e ausência, respectivamente, de quimioprofilaxia. Valores binários. Quando um dos parâmetros for 1, o outro deve ser zero. Não existe nenhuma situação possível em que os dois parâmetros tenham o mesmo valor.

$\gamma$  = Período de imunidade temporária. Dado, em meses, da triangulação dos períodos mais encontrados na literatura: 1 a 6 meses.

$\delta$  = Taxa de recuperação. O valor associado à proporção de cura por individuo infectado, dado por (da Saúde, 2019).

$\tau$  = Taxa de mortalidade. Valor associado ao coeficiente de mortalidade pela doença, dado por (da Saúde, 2019).

Para os valores acima, serão feitas as devidas conversões e análises proporcionais considerando o número amostral de 500 indivíduos no tempo de 365 dias.



### 3.2 Implementação e simulações computacionais

A simulação computacional foi implementada utilizando o software *AnyLogic* v.8.4.0<sup>1</sup>. O *Anylogic* é um software de ferramentas de simulação dinâmica Multi-Métodos que integra simulação por Sistemas Dinâmicos, Eventos Discretos e Baseada em Agentes em linguagem de modelagem, incluindo baixos níveis de construções de modelagem (variáveis, equações, parâmetros e eventos), formas de apresentação (formas geométricas), ferramentas de análise (conjuntos de dados, histogramas e gráficos), ferramentas de conectividade, imagens e tela de experimentos. O software é focado em simulações de diferentes áreas: mercado, saúde, manufatura, logística, defesa, gestão de processos, TI, dinâmica de populações, entre outras. Quando o Anylogic inicia a simulação, são construídas as Equações 1 e 2 que, durante a simulação, são resolvidas pelo método numérico de Runge-Kutta 4<sup>a</sup> ordem (RK4)<sup>2</sup>.

## 4 RESULTADOS E DISCUSSÕES

Para todos os cenários simulados foi utilizada uma amostra de 500 indivíduos da população, sendo 3 inicialmente infectados, durante o período de 365 dias ( $T_0 = 0$  e  $T_f = 365$ ). Os parâmetros utilizados nas simulações computacionais foram:  $N_0 = 500$  indivíduos,  $T_0 = 0$ ,  $T_f = 365$  dias,  $\Delta t = 1$  dia, taxa de incidência = 10 contatos por dia, incubação  $\epsilon = 28$  a 70 dias, imunidade temporária  $\gamma = 30$  a 180 dias, taxa de recuperação  $\delta = 95\%$  dos indivíduos, taxa de mortalidade  $\tau = 5\%$  dos indivíduos infectados, taxa de natalidade  $v = 0,3$  indivíduos suscetíveis ao ano, taxa de mortalidade natural  $\lambda = 0,05$  indivíduos suscetíveis ao ano, taxa de imigração dos susceptíveis  $\varphi_s = 150$  por ano, taxa de imigração dos expostos  $\varphi_e = 30$  por ano, taxa de imigração dos infectados  $\varphi_i = 5$  por ano e taxa de imigração dos removidos  $\varphi_r = 20$  por ano. Para discretizações no domínio do tempo, com valores de  $\Delta t < 1$  dia, não foram obtidas alterações significativas nos resultados.

### 4.1 Cenário simulado 1

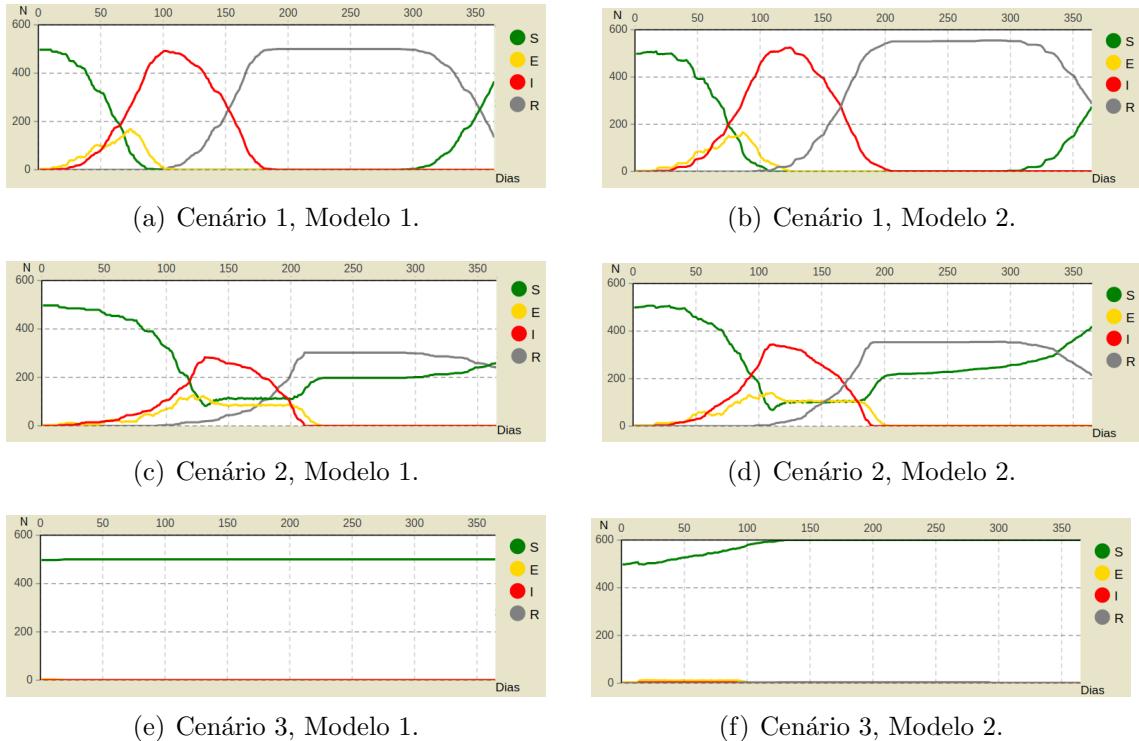
O primeiro cenário simulado analisou a dispersão da TB em uma população suscetível sem o uso da quimioprofilaxia, representando um surto epidemiológico da doença. As Figuras 3(a) e 3(b) apresentam os resultados das simulações computacionais para o cenário 1, utilizando os modelos 1 e 2 respectivamente. Podemos perceber que a TB apresentou um quadro de elevado surto em que os indivíduos suscetíveis rapidamente se tornaram infectados, devido a não utilização da quimioprofilaxia. Para ambos os modelos utilizados os comportamentos mantiveram-se semelhantes. Sabemos que um surto surge somente se o número de indivíduos infectados aumenta e que segundo o modelo matemático SEIR significa  $\frac{\partial I}{\partial t} > 0$ .

### 4.2 Cenário simulado 2

O segundo cenário analisou o impacto que o uso da quimioprofilaxia secundária tem na dispersão da doença. Neste cenário, a quimioprofilaxia secundária foi inserida

<sup>1</sup><https://www.anylogic.com/>

<sup>2</sup><https://help.anylogic.com/index.jsp?topic=%2Fcom.anylogic.help%2Fhtml%2Fexperiments%2FNumerical+Methods.html>



**Fig. 3:** Evolução da população de células.

durante um possível surto da doença. Este cenário é apresentado pela Figura 3(c) para o Modelo 1, onde podemos perceber que a inserção da prevenção com quimioprofilaxia, em dois instantes diferentes, conferiu imunidade aos indivíduos suscetíveis evitando o aumento de casos da TB e consequentemente um aumento de surto da doença. O comportamento se repetiu para o Modelo 2, como podemos observar na Figura 3(d).

#### 4.3 Cenário simulado 3

Assim como o segundo cenário, o terceiro cenário analisou o impacto do uso da quimioprofilaxia secundária, logo após os primeiros casos de infecção, com sua inserção antes de ocorrer um surto da doença. As Figuras 3(e) e 3(f) mostram o cenário ideal no combate a TB, no qual a doença é erradicada logo no início do seu aparecimento. Para o Modelo 1, a população permaneceu constante e para o Modelo 2, é possível perceber o aumento da população por conta dos fatores de nascimento e de imigração. Neste caso, as condições iniciais são  $S \cong N$  e  $I \cong 0$  e utilização da quimioprofilaxia. Este cenário é semelhante aos cenário em que não ocorre o aparecimento da TB, ou seja, com a condição inicial  $S = N$ ,  $E = 0$  e  $I = 0$ . Após o período de análise, identifica-se o fim da dispersão da doença, pois não existem mais sujeitos infectados, apenas suscetíveis.

O Modelo 1 aqui apresentado, mesmo que simples, atende as características fundamentais da epidemiologia da doença. Tal modelo, provê predições válidas acerca da dispersão da TB em uma população homogênea, isolada e igualmente suscetível. Quando aplicado a simulações computacionais de três possíveis cenários de dispersão



da doença, o modelo gera resultados condizentes com a realidade e permite a análise da interferência de um fator que limita a dispersão: a quimioprofilaxia secundária. No primeiro cenário simulado (que está mais próximo à dispersão atual da doença na cidade estudada), identifica-se um período de surto e alta exposição de indivíduos suscetíveis, que ou se tornarão infectados e, após o período de tratamento podem vir a óbito ou se tornar recuperados e aptos a se tornarem suscetíveis novamente; O segundo cenário simulado para o Modelo 1 também identificou um período de surto, entretanto, após a inserção da quimioprofilaxia secundária como um mecanismo de evitar a infecção, ficou clara a interrupção da cadeia de dispersão da doença, pois após a sua inserção, diminuiu o número de sujeitos infectados, assim, impedindo a formação de novos períodos de surtos; No terceiro, e último, cenário simulado, encontra-se o melhor cenário possível que, eventualmente, acaba com a dispersão da doença na população. Nesse cenário, a quimioprofilaxia secundária é utilizada em todos os indivíduos expostos antes de qualquer período de surto, e consegue rapidamente diminuir o número de indivíduos infectados presentes na população, ao impedir que novos sujeitos se infectem. É importante mencionar que, em todos os cenários e modelos, as flutuações observadas nos gráficos deram-se ao fato de o tempo de transição entre os estados dos modelos serem diferentes (como o tempo de incubação diferente dos tempos de recuperação e imunidade temporária) e não à necessidade de um aumento da discretização no domínio temporal ( $\Delta t < 1$  dia).

Considerando as novas hipóteses adotadas para o Modelo 2, podemos observar que os resultados foram semelhantes ao Modelo 1 para os três cenários apresentados. A simplicidade do modelo inicial (Modelo 1) permitiu o relaxamento das suposições e modificações de parâmetros com o objetivo de testar diferentes novas hipóteses, permitindo-nos partir de premissas mais simples para modelarmos o fenômeno epidemiológico estudado.

O presente estudo demonstrou, em uma população simulada, que a intervenção através do uso da quimioprofilaxia secundária apresenta um impacto significativo na transmissão da TB, especialmente quando administrada a casos pontuais antes de um período de surto, concordando com estudos recentes sobre o assunto e com as recomendações de muitos países. Entretanto, sabe-se que embora as normas internacionais oficiais (WHO, 2016) recomendem que todos os contatos de indivíduos infectados com TB sejam examinados, a escassez de recursos destinados à saúde, própria dos países em desenvolvimento, onde é elevada a incidência de TB, faz com que se dê maior atenção ao tratamento dos casos mais severos da doença ou após um período de grave surto.

## 5 CONCLUSÕES

Uma importante vantagem no uso de modelagem computacional na epidemiologia de doenças infecciosas é que a representação matemática de um processo biológico permite transparência e acurácia nas suposições epidemiológicas, dentro das limitações do modelo, nos permitindo adquirir conhecimento acerca da epidemiologia e comparar os diferentes resultados modelados e os padrões observados.

No presente trabalho, foram formulados dois modelos matemático-epidemiológicos SEIR capazes de identificar o padrão de dispersão de Tuberculose Pulmonar em uma

população homogênea. Os modelos puderam auxiliar na identificação do impacto positivo de medidas de quimioprofilaxia secundária na interrupção dos padrões de dispersão da doença na população simulada, estando em concordância com estudos clínicos e as recomendações de muitos países. Além disso, os modelos criados, apesar de simples, têm um alto valor preditivo e a enorme capacidade de serem adaptados e aplicados a outros possíveis cenários, que englobem diferentes aspectos da dispersão da doença.

Os modelos tiveram como base dados epidemiológicos da doença na cidade do Rio de Janeiro, RJ - Brasil, no ano de 2018, isto torna-os relevantes e acurados para a população brasileira nesse mesmo período, não os impedindo de serem adaptados e aplicados a populações de outras regiões ou países.

## 6 Agradecimentos

Os autores agradecem à Pós graduação em Modelagem Computacional do Laboratório Nacional de Computação Científica, ao orientador do trabalho Prof. Dr. José Karam Filho, aos Profs. Dra. Mariza Ferro, Dr. Bruno Schulze e Dra. Ana Teixeira Ribeiro de Vasconcelos, e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico.

## REFERÊNCIAS

- [1] Aparicio, J. P. and Castillo-Chavez, C. (2009). "Mathematical modelling of tuberculosis epidemics". *Math Biosci Eng*, 6(2):209–37.
- [2] Brauer, F. (2017). "Mathematical epidemiology: Past, present, and future". *Infectious Disease Modelling*, 2(2):113–127.
- [3] da Saúde, M. (2019). "Brasil livre da tuberculose: evolução dos cenários epidemiológicos e operacionais da doença". <http://portalarquivos2.saude.gov.br/images/pdf/2019/marco/22/2019-009.pdf>.
- [4] Huppert, A. and Katriel, G. (2013). "Mathematical modelling and prediction in infectious disease epidemiology". *Clinical Microbiology and Infection*, 19(11):999–1005.
- [5] Kretzschmar, M. (2016). "Measurement and modeling: Infectious disease modeling". *Reference Module in Biomedical Sciences*.
- [6] MS (2016). "Perspectivas brasileiras para o fim da tuberculose como problema de saúde pública". 47. <http://portalarquivos2.saude.gov.br/images/pdf/2016/marco/24/2016-009-Tuberculose-001.pdf>.
- [7] MS (2018). "Protocolo de vigilância da infecção latente pelo mycobacterium tuberculosis no brasil". [http://bvsms.saude.gov.br/bvs/publicacoes/protocolo\\_vigilancia\\_infeccao\\_latente\\_mycobacterium\\_tuberculosis\\_brasil.pdf](http://bvsms.saude.gov.br/bvs/publicacoes/protocolo_vigilancia_infeccao_latente_mycobacterium_tuberculosis_brasil.pdf).
- [8] Nogueira, A. F., Facchinetto, V., de Souza, M. V. N., and Alves, T. R. (2012). "Tuberculose: uma abordagem geral dos principais aspectos". *Rev. Bras. Farm*, 93(1):3–9.
- [9] Oliveira, I. M. d. et al. (2008). "Modelos epidemiológicos seir".
- [10] Petersen, E., Chakaya, J., Jawad, F. M., Ippolito, G., and Zumla, A. (2019). "Latent tuberculosis infection: diagnostic tests and when to treat". *The Lancet Infectious Diseases*, 19(3):231–233.
- [11] Soares, M. L. M., Amaral, N. A. C. d., Zacarias, A. C. P., and Ribeiro, L. K. d. N. P. (2017). "Sociodemographic, clinical and epidemiological aspects of tuberculosis treatment abandonment in pernambuco, brazil, 2001-2014". *Epidemiologia e Serviços de Saúde*, 26(2):369–378.
- [12] Souza, C. D. F. d., Matos, T. S., Santos, V. S., and Santos, F. G. B. (2019). "Tuberculosis surveillance in an endemic area of northeastern brazil. what do the epidemiological indicators reveal?" *Jornal Brasileiro de Pneumologia*, 45(2).
- [13] WHO, W. H. O. (2016). "Global tuberculosis report 2016". <https://apps.who.int/iris/bitstream/handle/10665/250441/9789241565394-eng.pdf;jsessionid=9EE8574795A4A41D171A36A6100BC415?sequence=1>.



# Método primal-dual na otimização agrícola

Angel Ramon Sanchez Delgado<sup>1</sup>, Sergio Drumond Ventura<sup>1</sup> e Pedro Marcio Ferreira<sup>1</sup>

<sup>1</sup> Universidade Federal Rural do Rio de Janeiro/ICE/DEMAT/PPGCTIA/PPGMMC

---

## Abstract

Considerando a importância que tem a geração de esquemas ótimos de produção e comercialização de insumos agrícolas no agronegócio, se faz preciso o desenvolvimento de procedimentos computacionais robustos e simples, que permitam determinar a otimalidade de problemas agrícolas modeláveis matematicamente. Nesse trabalho, apresentamos um procedimento computacional de pontos interiores primal-dual, que permite selecionar as culturas e os meses de plantio, em um perímetro irrigado, proporcionando a maximização da receita líquida com a melhor utilização dos recursos disponíveis e a comercialização ótima dos insumos, água e terra, assim como uma experiência numérica que tem como cenário agrícola o perímetro irrigado do Gorutuba, localizado no município de Nova Porteirinha, Norte do estado de Minas Gerais.

**Keywords:** Agronegócio, Produção, Receita Líquida, Pontos Interiores, Programação Matemática.

---

## 1 INTRODUÇÃO

Na otimização das decisões agrícolas temos que escolher entre as diferentes políticas de produção a mais eficiente, em relação às metas e condições de viabilidade. As decisões baseadas no julgamento e na intuição podem ser satisfatórias quando o número de fatores do problema é limitado e suas relações são claras. Mas em situações onde esse número cresce é necessária a utilização de modelos matemáticos que representem ou simulem condições reais.

A otimização agrícola (OA) representa a classe de problemas ligados a situações agrícolas, tais como: maximização da produção e receita líquida de culturas agrícolas (Frizzone et al. (2005), Carvalho et al. (2000), Carvalho et al. (2009), Delgado and Ventura (2014)) rotação de culturas (Santos et al. (2007), Galdino et al. (2019), Alfandaria et al. (2015)), manejo da movimentação animal em um perímetro agropecuário (Ventura et al. (2019)), alocação de recursos agrícolas (Silva et al. (1998)), controle ótimo de irrigação (Delgado et al. 2010), entre outros; e que podem, em geral, ser modelados como um problema de programa matemática (PPM).

Na otimização agrícola contemporânea o problema desafiador é como irrigar para maximizar o lucro. Em geral, ao produtor agrícola lhe interessa a determinação de um padrão de cultivo ótimo das culturas disponíveis, de tal maneira que a receita líquida seja máxima quando é feita uma racionalização da água e da terra (problema primal). Além disso, lhe interessa saber qual é a taxa de mudança na receita líquida, quando é feita uma variação nos volumes de água disponíveis e nas áreas irrigáveis por mês, ou seja, qual é o esquema ótimo da comercialização dos insumos (problema dual).

Uma classe de métodos de otimização muito explorada nos dias de hoje é denominada de pontos interiores. Os métodos de pontos interiores (MPI) têm sido amplamente investigados e utilizados principalmente na resolução de problemas de programação linear (PPL) e, mais recentemente, em problemas de programação quadrática e não-linear, com bom desempenho em problemas de grande porte. O objetivo central do presente trabalho, é apresentar um procedimento conceitual e computacional, tipo pontos interiores primal-dual, que permita selecionar as culturas e meses de plantio, que proporcionem a maximização da receita líquida do agricultor, assim como a determinação dos preços dos insumos (água e terra irrigável), de modo de fornecer a oferta mínima para que se aceite um acréscimo deles.

## 2 METODOLOGIA

No planejamento agrícola com limitações de água e terra, procura-se selecionar as culturas e os meses de plantio em um perímetro irrigado que proporcionem a maximização da receita líquida e a melhor utilização dos recursos disponíveis. O problema é determinar um padrão de cultivo ótimo das culturas de tal maneira que a receita líquida seja máxima quando é feita uma racionalização dos recursos. Esquematicamente, o modelo de programação linear para representar o problema anterior, pode ser equacionado como segue:

$$\text{maximizar } RL(x) = \sum_{i=1}^m \sum_{j=1}^{12} (p_{ij}\tau_i - c_i)x_{ij} = \sum_{i=1}^m \sum_{j=1}^{12} r_{ij}x_{ij}, \quad (1a)$$

$$\text{sujeito a } \sum_{i=1}^m v_{ij}x_{ij} \leq v_j, \quad j = 1, \dots, 12, \quad (1b)$$

$$\sum_{i=1}^m x_{ij} \leq a_j, \quad j = 1, \dots, 12, \quad (1c)$$

$$x_{ij} \geq 0, \quad i = 1, \dots, m \text{ e } j = 1, \dots, 12, \quad (1d)$$

onde

$i$  – índice referente às culturas agrícolas,  $i = 1, \dots, m$ ,

$j$  – índice referente ao mês do ano,  $j = 1, \dots, 12$ ,

$p_{ij}$  – preço da cultura  $i$  no mês  $j$  ( $\text{R\$} \cdot \text{kg}^{-1} \cdot \text{ha}^{-1}$ ),

$\tau_j$  – produtividade média da cultura  $i$  (kg),



$c_i$  – média dos custos médios da cultura  $i$  incluindo a tarifa de água para irrigação ( $\text{R\$} \cdot \text{ha}^{-1}$ ),

$r_{ij} = p_{ij}\tau_i - c_i$  – receita líquida marginal ( $\text{R\$} \cdot \text{ha}^{-1}$ ),

$v_{ij}$  – volume de água utilizado com a cultura  $i$  durante o mês  $j$  ( $\text{m}^3 \cdot \text{ha}^{-1}$ ),

$x_{ij}$  – área cultivada com a cultura  $i$  no mês  $j$  (ha),

$v_j$  – volume de água disponível no mês  $j$  ( $\text{m}^3$ ),

$a_j$  – área total disponível no mês  $j$  (ha).

A equação (1a) representa a função objetivo a ser maximizada e as equações (1b)-(1c)-(1d) correspondem às restrições de volume de água e área, respectivamente. Adiante, esse PPL será chamado de problema primal (PPL-primal). Vale destacar que para o produtor agrícola, além da solução ótima do problema primal é importante saber qual é a taxa de mudança na receita líquida quando é feita uma variação nos volumes de água disponíveis e nas áreas irrigáveis por mês, ou seja, qual é o esquema ótimo da comercialização dos insumos (água e terra). Esta informação interessa também ao modelo PPL que determina os preços destes insumos, de modo a fornecer a oferta mínima para que se aceite um acréscimo de água e/ou de terra. Neste caso, deseja-se o menor valor total de produção se é aumentado em uma unidade o volume total de água disponível e em um hectare a área irrigada. Matematicamente, procura-se uma solução ótima do PPL conhecido como problema dual (PPL-dual) associado ao problema primal:

$$\text{minimizar } f(u, \zeta) = \sum_{j=1}^{12} v_j u_j + \sum_{j=1}^{12} a_j \zeta_j, \quad (2a)$$

$$\text{sujeito a } v_{ij} u_j + \zeta_j \geq r_{ij}, \quad i = 1, \dots, m \text{ e } j = 1, \dots, 12, \quad (2b)$$

$$u_j, \zeta_j \geq 0, \quad j = 1, \dots, 12, \quad (2c)$$

onde

$u_j$  – preço de uma unidade do volume total de água disponível no mês  $j$  ( $\text{R\$} \cdot \text{m}^{-3}$ ),

$\zeta_j$  – preço de um hectare de terra do perímetro irrigado (ou de arrendamento de um hectare de terra) ( $\text{R\$} \cdot \text{ha}^{-1}$ ).

É possível observar que as restrições (2b), indicam que, para cada cultura e mês do ano, o volume de água consumido pela cultura  $i$  no mês  $j$  multiplicado pelo o preço de uma unidade do volume total de água, mais o preço do arrendamento de um hectare de terra irrigada, deve ser maior ou igual à receita líquida marginal  $r_{ij} = p_{ij}\tau_i - c_i$ . Senão, não se estaria fazendo o melhor uso possível dos recursos disponíveis. Entre as vantagens das soluções duais destacam-se as possibilidades de fornecer informações econômicas sobre a utilização de recursos, como por exemplo, auxiliar na tomada de decisão para a aquisição de recursos adicionais.

A estratégia que apresentamos a continuação para resolver os problemas primal e dual, segue o conhecido método função barreira (Fiacco and Cormick (1968), Bazaar

et al. (1979)). Resumidamente, associa-se ao problema restrito dual (2a)-(2b)-(2c) uma sequência de problemas irrestritos (simples de resolver) cujas soluções ótimas convergem a uma solução ótima do problema dual. Essa associação é feita com a integração das restrições do problema dual na função objetivo utilizando um parâmetro de penalidade e uma barreira logarítmica que penaliza o não cumprimento das restrições duais. A abordagem metodológica escolhida baseou-se na procura de um procedimento computacional que resolve os problemas primal e dual utilizando uma técnica diferente ao SIMPLEX e de fácil implementação computacional.

É conhecido que o problema barreira logarítmica associada ao problema (2a)-(2b)-(2c), para cada  $\mu > 0$ , está dado por:

$$\text{minimizar } \phi_\mu(u, \zeta, z) = f(u, \zeta) - \mu B(u, \zeta, z), \quad (3a)$$

$$\text{sujeito a } v_{ij}u_j + \zeta_j - z_{ij} = r_{ij}, \quad i = 1, \dots, m \text{ e } j = 1, \dots, 12, \quad (3b)$$

onde  $z$  representa as variáveis de folga associadas a cada restrição do problema dual e

$$B(u, \zeta, z) = \sum_{i=1}^m \sum_{j=1}^{12} \ln z_{ij} + \sum_{j=1}^{12} \ln u_j + \sum_{j=1}^{12} \ln \zeta_j.$$

Note que na realidade, (3a)-(3b) representa uma família de problemas irrestritos não-lineares, para cada  $\mu > 0$ ; mas ainda, o mínimo de cada (3a)-(3b) é alcançado em um ponto interior da região de viabilidade do dual e quando  $\mu$  tende a zero, o ponto interior se movimenta a um ponto próximo da solução ótima do problema dual. O conjunto de soluções ótimas dos problemas barreira logarítmica em função de  $\mu$ , resulta em uma curva (ou trajetória) contida no interior do conjunto de viabilidade de (2a)-(2c) e conhecida como trajetória central. Utilizando o operador Lagrangeano (Bazaraa et al. (1979)) definido por:

$$L_\mu(u, \zeta, x, z) = \phi_\mu(u, \zeta, z) - \sum_{i=1}^m \sum_{j=1}^{12} x_{ij}g_{ij}(u, \zeta, z),$$

onde  $x_{ij}$  está representando os multiplicadores de Lagrange só para manter a relação notacional entre os problemas primal e dual em relação ao sistema de otimalidade e  $g_{ij}(u, \zeta, z) = v_{ij}u_j + \zeta_j - z_{ij} - r_{ij}$ . Resolve-se (3a)-(3b) para cada  $\mu > 0$ , resolvendo o programa não-linear irrestrito:

$$\text{minimizar } L_\mu(u, \zeta, x, z). \quad (4)$$

Agora, as condições de otimalidade de primeira ordem de (4) ficam definidas pelo sistema de equações não-lineares:

$$\frac{\partial L_\mu}{\partial u_j} = v_j - \frac{\mu}{u_j} - \sum_{i=1}^m v_{ij}x_{ij} = 0, \quad j = 1, \dots, 12,$$

$$\frac{\partial L_\mu}{\partial \zeta_j} = a_j - \frac{\mu}{\zeta_j} - \sum_{i=1}^m x_{ij} = 0, \quad j = 1, \dots, 12,$$

$$\frac{\partial L_\mu}{\partial x_{ij}} = v_{ij}u_j + \zeta_j - z_{ij} - r_{ij} = 0, \quad i = 1, \dots, m \text{ e } j = 1, \dots, 12,$$

$$\frac{\partial L_\mu}{\partial z_{ij}} = -\frac{\mu}{z_{ij}} + x_{ij} = 0, \quad i = 1, \dots, m \text{ e } j = 1, \dots, 12.$$



Definindo  $\eta_j = \frac{\mu}{u_j} > 0$  e  $\gamma_j = \frac{\mu}{\zeta_j}$ , temos que para todo  $i = 1, \dots, m$  e  $j = 1, \dots, 12$ :

$$\sum_{i=1}^m v_{ij}x_{ij} + \eta_j - v_j = 0, \quad (5a)$$

$$\sum_{i=1}^m x_{ij} + \gamma_j - a_j = 0, \quad (5b)$$

$$v_{ij}u_j + \zeta_j - z_{ij} - r_{ij} = 0, \quad (5c)$$

$$x_{ij}z_{ij} - \mu = 0 \quad (5d)$$

$$x_{ij}, z_{ij}, u_j, \zeta_j > 0. \quad (5e)$$

O sistema (5a)-(5d) está composto por  $24(m+1)$  equações lineares e não-lineares e  $24(m+1)$  variáveis. De forma matricial, o sistema pode ser escrito como:

$$\begin{aligned} Ax + w - b &= 0, \\ A^T y - z - r &= 0, \\ XZ e_{12m} - \mu e_{12m} &= 0, \end{aligned}$$

onde  $A = \begin{pmatrix} V_1 & \cdots & V_m \\ I_{12} & \cdots & I_{12} \end{pmatrix}$  é uma matriz  $24 \times 12m$ , onde  $V_j = \text{diag}(v_{i1}, \dots, v_{i12})$ , para  $i = 1, \dots, m$  e  $I_{12}$  é a matriz identidade  $12 \times 12$ . Suponha que existem pelo menos duas colunas  $i$  e  $k$  de  $A$  tal que  $v_{ij} \neq v_{kj}$ , para todo  $j = 1, \dots, 12$ , isto é, o volume de água necessário para o desenvolvimento da cultura  $i$  é diferente do volume necessário para o desenvolvimento da cultura  $k$  durante qualquer mês do ano. Com esta suposição  $A$  tem posto completo e definimos:

$$\begin{aligned} x &= (x_{11}, \dots, x_{112}, x_{21}, \dots, x_{212}, \dots, x_{m1}, \dots, x_{m12})^T \in \mathbb{R}^{12m}, \\ w &= (\eta_1, \dots, \eta_{12}, \gamma_1, \dots, \gamma_{12})^T \in \mathbb{R}^{24}, \\ b &= (v_1, \dots, v_{12}, a_1, \dots, a_{12})^T \in \mathbb{R}^{24}, \\ y &= (u_1, \dots, u_{12}, \zeta_1, \dots, \zeta_{12})^T \in \mathbb{R}^{24}, \\ z &= (z_{11}, \dots, z_{112}, z_{21}, \dots, z_{212}, \dots, z_{m1}, \dots, z_{m12})^T \in \mathbb{R}^{12m}, \\ r &= (r_{11}, \dots, r_{112}, r_{21}, \dots, r_{212}, \dots, r_{m1}, \dots, r_{m12})^T \in \mathbb{R}^{12m}, \\ X &= \text{diag}(x_{11}, \dots, x_{112}, x_{21}, \dots, x_{212}, \dots, x_{m1}, \dots, x_{m12})^T \in \mathbb{R}^{12m \times 12m}, \\ Z &= \text{diag}(z_{11}, \dots, z_{112}, z_{21}, \dots, z_{212}, \dots, z_{m1}, \dots, z_{m12})^T \in \mathbb{R}^{12m \times 12m}, \\ e_{24} &= (1, \dots, 1) \in \mathbb{R}^{24}, e_{12m} = (1, \dots, 1) \in \mathbb{R}^{12m}. \end{aligned}$$

O sistema anterior pode ser resolvido utilizando o método de Newton-Raphson. Para facilitar o cálculo da direção dentro do método, acrescentamos a restrição trivialmente cumprida:  $YWe_{24} - \mu e_{24} = 0$ , onde:

$$\begin{aligned} Y &= \text{diag}(u_1, \dots, u_{12}, \zeta_1, \dots, \zeta_{12})^T \in \mathbb{R}^{24 \times 24}, \\ W &= \text{diag}(\eta_1, \dots, \eta_{12}, \gamma_1, \dots, \gamma_{12})^T \in \mathbb{R}^{24 \times 24}. \end{aligned}$$

Consideremos então o sistema:

$$Ax + w - b = 0, \quad (6a)$$

$$A^T y - z - r = 0, \quad (6b)$$

$$XZ e_{12m} - \mu e_{12m} = 0, \quad (6c)$$

$$YWe_{24} - \mu e_{24} = 0. \quad (6d)$$

Conceitualmente o procedimento aqui proposto funciona da seguinte maneira: suponha que  $(x, y, w, z)$  é um ponto inviável primal-dual, estritamente positivo e  $\mu > 0$ . O método gera iterativamente a chamada direção de Newton  $(\Delta x, \Delta y, \Delta w, \Delta z)$  tal que o ponto  $(x + \Delta x, y + \Delta y, w + \Delta w, z + \Delta z)$  satisfaz o sistema (6a)-(6d), isto é, procura-se um vetor tal que:

$$A\Delta x + \Delta w = \rho, \quad (7a)$$

$$A^T\Delta y - \Delta z = \sigma, \quad (7b)$$

$$Y\Delta w + W\Delta y = f_{yw}, \quad (7c)$$

$$X\Delta z + Z\Delta x = f_{xz}, \quad (7d)$$

onde  $\rho = b - Ax - w$ ,  $\sigma = r - A^T y + z$ ,  $f_{yw} = \mu e_{24} - YWe_{24}$  e  $f_{xz} = \mu e_{12m} - XZe_{12m}$  são os chamados: resíduo primal, resíduo dual e folgas complementárias respectivamente. A solução única do sistema anterior se segue do fato de que  $A$  tem posto completo. Manipulando a primeira e a segunda equação do sistema, obtemos que:

$$\Delta w = \rho - A\Delta x \quad \text{e} \quad \Delta z = A^T\Delta y - \sigma.$$

e, substituindo (7c) e (7d), temos

$$Y(\rho - A\Delta x) + W\Delta y = f_{yw} \quad \text{e} \quad X(A^T\Delta y - \sigma) + Z\Delta x = f_{xz},$$

de modo que

$$-YA\Delta x + W\Delta y = f_{yw} - Y\rho \quad (8a)$$

$$Z\Delta x + XA^T\Delta y = f_{xz} - X\sigma. \quad (8b)$$

Logo,  $\Delta x + Z^{-1}XA^T\Delta y = Z^{-1}(f_{xz} - X\sigma)$  e  $\Delta x = Z^{-1}Z^{-1}(f_{xz} - X\sigma - XA^T\Delta y)$ . Assim,

$$-YAZ^{-1}(f_{xz} - X\sigma - XA^T\Delta y) + W\Delta y = f_{yw} - Y\rho$$

e, portanto,

$$(AZ^{-1}XA^T + Y^{-1}W)\Delta y = Y^{-1}(f_{yw} - Y\rho) + AZ^{-1}(f_{xz} + X\sigma). \quad (9)$$

Resolvendo a equação acima, finalmente obtemos

$$\begin{aligned} \Delta x &= Z^{-1}(f_{xz} + X\sigma - XA^T\Delta y) \\ \Delta z &= A^T\Delta y - \sigma, \\ \Delta w &= \rho - A\Delta x. \end{aligned}$$

O procedimento conceitual para resolver simultaneamente os problemas primal e dual, pode ser resumido da seguinte maneira: Dado um ponto inviável primal-dual, mas estritamente positivo  $(x, y, w, z)$  e um parâmetro  $\mu > 0$ , em cada iteração calculamos a direção Newton  $(\Delta x, \Delta y, \Delta w, \Delta z)$  que satisfaz (7a)-(7d) e um tamanho de passo apropriado  $\theta$  para manter as variáveis estritamente positivas. Seguidamente determinamos o novo iterado  $x = x + \theta\Delta x$ ,  $y = y + \theta\Delta y$ ,  $w = w + \theta\Delta w$ ,  $z = z + \theta\Delta z$ . O procedimento finaliza quando a folga de dualidade e a inviabilidade primal e dual sejam menores que uma tolerância prefixada.



### 3 PROCEDIMENTO COMPUTACIONAL PRIMAL-DUAL

A continuação o procedimento primal-dual em pseudocódigo, pronto para ser implementado em uma plataforma computacional tipo Matlab, com os respectivos dados escolhidos pelo usuário. Inicialmente devem ser conhecidos: os preços de cada cultura em cada mês do ano ( $\text{R\$} \cdot \text{kg}^{-1}$ ), a produtividade média de cada cultura ( $\text{kg}$ ), o custo médio de cada cultura, incluindo a tarifa de água para irrigação ( $\text{R\$} \cdot \text{ha}^{-1}$ ), o volume de água utilizado por cada cultura, durante cada mês ( $\text{m}^3 \cdot \text{ha}^{-1}$ ), o volume de água disponível em cada mês ( $\text{m}^3$ ), a área total disponível a cada mês ( $\text{ha}$ ) e uma solução inicial estritamente positiva, não necessariamente viável,  $(x_0, y_0, w_0, z_0) \in \mathbb{R}^{24(m+2)}$ .

---

**Algoritmo 1:** Primal-dual

---

**Entrada:**  $p_{ij}, \tau_j, c_i, v_{ij}, v_j, a_j \geq 0$ ,  $\epsilon, \xi \in (0, 1)$ ,  $\mu_0 > 0$ ,  $x_0 \in \mathbb{R}_{++}^{12m}$ ,  $y_0 \in \mathbb{R}_{++}^{24}$ ,  
 $w_0 \in \mathbb{R}_{++}^{24}$ ,  $z_0 \in \mathbb{R}_{++}^{24}$ .

$A = \begin{pmatrix} V_1 & \cdots & V_m \\ I_{12} & \cdots & I_{12} \end{pmatrix}$  onde  $V_j = \text{diag}(v_{i1}, \dots, v_{i12})$ , para  $i = 1, \dots, m$  e  $I_{12}$  é a matriz identidade  $12 \times 12$ ;

$$r_{ij} = p_{ij}\tau_j - c_i, \quad i = 1, \dots, m, \quad j = 1, \dots, 12;$$

$$r = (r_{11}, \dots, r_{12}, r_{21}, \dots, r_{212}, \dots, r_{m1}, \dots, r_{m12})^T \in \mathbb{R}^{12m};$$

$$b = (v_1, \dots, v_{12}, a_1, \dots, a_{12})^T \in \mathbb{R}^{24};$$

$$e_{24} = (1, \dots, 1) \in \mathbb{R}^{24}, \quad e_{12m} = (1, \dots, 1) \in \mathbb{R}^{12m};$$

$$\mu = \mu_0, \quad x = x_0, \quad y = y_0, \quad w = w_0, \quad z = z_0$$

$$X = \text{diag}(x), \quad Z = \text{diag}(z), \quad Y = \text{diag}(y), \quad W = \text{diag}(w);$$

$$\rho = b - Ax - w \in \mathbb{R}^{24}, \quad \sigma = r - A^T y + z \in \mathbb{R}^{12m};$$

$$f_{xz} = \mu e_{12m} - XZe_{12m} \in \mathbb{R}^{12m}, \quad f_{yw} = \mu e_{24} - YWe_{24} \in \mathbb{R}^{24};$$

**enquanto**  $\max\{\|\rho\|, \|\sigma\|, \|f_{xz}\|, \|f_{yw}\|\} > \epsilon$  **faça**

$$\left. \begin{aligned} &\text{Resolver } (AZ^{-1}XA^T + Y^{-1}W)\Delta y = Y^{-1}(f_{yw} - Y\rho) + AZ^{-1}(f_{xz} + X\sigma); \\ &\Delta x = Z^{-1}(f_{xz} + X\sigma - XA^T\Delta y); \\ &\Delta z = A^T\Delta y - \sigma; \\ &\Delta w = \rho - A\Delta x; \\ &\theta = \min \left\{ \frac{\|x\|}{\|\Delta x\|}, \frac{\|w\|}{\|\Delta w\|}, \frac{\|y\|}{\|\Delta y\|}, \frac{\|z\|}{\|\Delta z\|} \right\}; \\ &x = x + 0,999\theta\Delta x; \quad y = y + 0,999\theta\Delta y; \\ &w = w + 0,999\theta\Delta w; \quad z = z + 0,999\theta\Delta z; \\ &X = \text{diag}(x), \quad Z = \text{diag}(z), \quad Y = \text{diag}(y), \quad W = \text{diag}(w); \\ &\mu = \xi\mu \\ &\text{Recalcular } \rho, \sigma, f_{xz}, f_{yw}; \end{aligned} \right.$$

**fim**

---

Seguidamente devem ser montadas a matriz  $A$  e a matrizes diagonais  $V_i$ . Fazendo agora:  $\mu = \mu_0$ ,  $x = x_0$ ,  $y = y_0$ ,  $w = w_0$ ,  $z = z_0$  e calculando as matrizes diagonais:  $X = \text{diag}(x)$ ,  $Y = \text{diag}(y)$ ,  $W = \text{diag}(w)$   $Z = \text{diag}(z)$ , e é possível conhecer o “gap” ou brecha de viabilidade primal da solução inicial ( $\rho = b - Ax - w$ ), a brecha de viabilidade dual ( $\sigma = r - A^T y + z$ ), o gap primal das folgas complementares ( $f_{xz} = \mu e_{12m} - XZe_{12m}$ ) e o gap dual das folgas complementares ( $f_{yw} = \mu e_{24} - YWe_{24}$ ). Finalmente, enquanto

$\max\{\|\rho\|, \|\sigma\|, \|f_{xz}\|, \|f_{yw}\|\} > \epsilon$ , para  $\epsilon \in (0, 1)$ , isto é, o iterado atual ainda não satisfaz proximamente o sistema dado por (6a)-(6d), calcula-se uma direção Newton-Raphson, um tamanho de passo ( $\theta$ ), e um novo ponto  $(x, y, w, z)$  é gerado iterativamente. A cada iteração, os parâmetros  $\mu, \rho, \sigma, f_{xz}, f_{yw}$  são atualizados.

## 4 EXPERIÊNCIA NUMÉRICA

O ensaio numérico realizado para determinar o quanto bem formulado está o problema considerado, assim como quanto eficiente é o algoritmo proposto para resolver tal problema, utilizou dados do perímetro irrigado do Gorutuba, localizado no município de Nova Porteirinha, Norte do estado de Minas Gerais, descritos em Carvalho et al. (2000). O perímetro está formado por nove lotes e foram considerados os plantios de quatro culturas anuais: abóbora, feijão, milho, quiabo; e quatro culturas perenes: banana prata, banana nanica, limão e manga. A Tabela 1 representa os preços, produtividade média e custos das culturas anuais e perenes.

**Tabela 1:** PREÇO DOS PRODUTOS AGRÍCOLAS ( $R\$ \cdot KG^{-1}$ ), PRODUTIVIDADE MÉDIA ( $R\$ \cdot KG^{-1}$ ) E CUSTOS MÉDIOS DE PRODUÇÃO ( $R\$ \cdot HA^{-1}$ ) PARA AS CULTURAS CONSIDERADAS.

Mês	Abóbora	Feijão	Milho	Quiabo	Banana Prata	Banana Nanica	Limão	Manga
Preço dos Produtos								
Jan		0,662			0,401	0,084	0,307	
Fev		0,500	0,180	0,400	0,484	0,085	0,188	
Mar			0,154		0,375	0,158	0,180	
Abr		0,750	0,148		0,404	0,126	0,180	
Mai		0,831	0,156		0,431	0,091	0,364	
Jun		0,830			0,601	0,090		
Jul		0,829	0,161		0,592	0,092		
Ago		0,756	0,164	0,455	0,413	0,219	0,402	1,00
Set					0,315	0,182	0,751	
Out		0,818			0,214	0,269	0,751	
Nov	0,200	0,752	0,156		0,215	0,158	0,384	0,294
Dez					0,198	0,092	0,429	0,219
Produtividade								
	18,000	1,800	5,500	15,000	26,000	44,000	22,000	18,000
Custo de Produção								
	2674,20	965,00	651,50	3611,50	1951,5	2676,5	996,40	1615,90

O volume de água necessário para irrigação dessas culturas foi estimado considerando-se o fornecimento rotativo de água, em função da demanda evapotranspirométrica local, do tipo de solo e época de plantio para cada cultura (ver Carvalho et al. (2000)). O volume de água disponível para cada mês e a área irrigável de cada lote do perímetro é mostrado na Tabela 2.



**Tabela 2:** VOLUMES MENSAIS DE ÁGUA  $v_j$  EM 1000M<sup>3</sup> E ÁREA IRRIGÁVEL  $a_j$  EM HA PARA OS LOTES CONSIDERADOS.

Lotes	A	B	C	D	E	F	G	H	I
Mês	Volumes Mensais								
Jan	53,10	11,09	12,35	5,15	5,29	6,35	13,36	6,55	12,35
Fev	85,82	17,42	8,86	17,93	8,91	9,34	17,82	7,88	17,82
Mar	59,80	11,09	5,74	11,38	05,81	5,74	12,17	3,92	11,48
Abr	60,26	14,26	7,06	7,16	7,06	7,06	14,11	6,88	14,44
Mai	88,02	17,42	8,78	8,78	17,68	8,78	17,93	8,68	15,92
Jun	75,29	15,84	15,84	7,92	8,10	7,92	9,70	8,10	13,08
Jul	71,14	15,84	15,84	7,92	8,35	8,35	15,60	7,85	14,19
Ago	81,00	16,63	7,92	7,87	8,30	8,30	14,61	9,11	14,95
Set	76,32	16,42	7,35	7,49	7,49	7,49	14,98	8,93	13,33
Out	52,81	12,10	10,62	5,18	2,66	5,31	10,26	6,55	8,32
Nov	19,08	4,32	3,96	1,98	1,80	1,99	3,96	2,59	2,31
Dez	18,00	6,05	4,54	2,27	1,95	2,95	5,29	2,23	2,89
	Área								
	38,00	11,00	8,46	3,51	5,94	2,95	8,16	5,90	7,97

## 5 RESULTADOS E DISCUSSÃO

Para evitar a ocupação excessiva de área de determinadas culturas, restringimos a área das culturas anuais em: 1,65; 1,82; 2,83 e 0,20 na área cultivada imposta às culturas abóbora, feijão, milho e quiabo, respectivamente a fim de refletir as condições reais de plantio do perímetro. Na Tabela 3 são apresentadas as áreas a serem cultivadas para cada cultura, entre as várias épocas de plantio, para todos os lotes considerados.

**Tabela 3:** ÁREAS POSSÍVEIS DE SEREM IRRIGADAS  $x_i$  HA, RESULTANTES DO PROBLEMA PRIMAL, PARA: TODOS OS LOTES ESTUDADOS, COM RESTRIÇÕES DE ÁREA CULTIVADA NAS CULTURAS ANUAIS. OS NÚMEROS ENTRE PARÊNTESIS APÓS OS NOMES DAS CULTURAS REPRESENTAM OS MESES DE PLANTIO DAS CULTURAS ANUAIS.

Culturas	A	B	C	D	E	F	G	H	I
Abóbora (6)	0,627	0,1815	0,139	0	0,026	0,048	0	0,09	0,131
Feijão (1)	0	0	0	0	0	0	0	0	0
Feijão (2)	0	0	0	0	0	0	0	0	0
Feijão (3)	0	0	0	0	0	0	0	0	0
Feijão (4)	0	0,200	0,153	0	0,108	0,053	0	0,107	0,145
Feijão (5)	0	0	0	0	0	0	0	0	0
Feijão (7)	0,6916	0	0	0	0	0	0	0	0
Feijão (8)	0	0	0	0	0	0	0	0	0
Feijão (10)	0	0	0	0	0	0	0	0	0
Feijão (11)	0	0	0	0	0	0	0	0	0
Milho (1)	0	0	0	0	0	0	0	0	0
Milho (3)	0	0	0	0	0	0	0	0	0
Milho (4)	0	0,311	0,239	0	0,168	0,083	0	0,166	0,225
Milho (7)	1,075	0	0	0	0	0	0	0	0
Milho (10)	0	0	0	0	0	0	0	0	0
Milho (11)	0	0	0	0	0	0	0	0	0
Milho (12)	0	0	0	0	0	0	0	0	0
Quiabo (4)	0	0,022	0,016	0	0,011	0,0059	0	0,011	0,0015
Quiabo (1)	0,076	0	0	0	0	0	0	0	0
Banana P.	22,138	6,198	1,165	3,51	4,06	1,99	8,16	0,593	6,47
Banana N.	0	0	0	0	0	0	0	0	0
Manga	0	0	0	0	0	0	0	0	0
Limão	4,537	3,062	4,144	0	0,747	0,585	0	3,70	0
Área Irrigável	38,00	11,00	8,46	3,51	5,94	2,95	8,16	5,90	7,97

Com relação às culturas anuais, o plantio do feijão e do milho pode ser recomendado nos meses de abril e julho e, para o quiabo e a abóbora, áreas de plantios praticamente desconsideravam (salvo o lote A) foram obtidas em razão da pequena percentagem de ocupação, citada anteriormente. A cultura de abóbora é recomendável em todos os lotes durante o mês de junho, salvo nos lotes D e G. Na Tabela 4 são apresentados os resultados duais correspondentes à variação da taxa da receita líquida quando se incrementa em uma unidade o volume de água disponível e em um hectare a área irrigada no respectivo lote.

**Tabela 4:** VARIAÇÃO DE PREÇOS DOS INSUMOS ÁGUA  $u_j$  (R\$·m<sup>3</sup>) E TERRA  $\zeta_j$  (R\$·ha<sup>-1</sup>).

Preços	A	B	C	D	E	F	G	H	I
$u_1$	0	0	0	0	0	0	0	0	0
$u_2$	0	0	0	0	0	0	0	0	0
$u_3$	0	4,50	4,61	0	4,87	2,62	0	4,93	0
$u_4$	2,614	0	0	0	0	0	0	0	0
$u_5$	0	0	0	0	0	0	0	0	0
$u_6$	0	0	0	0	0	0	0	0	0
$u_7$	0	0	0	0	0	0	0	0	0
$u_8$	0	0	0	0	0	0	0	0	0
$u_9$	0	0	0	0	0	0	0	0	0
$u_{10}$	0	0	0	0	0,835	0	0	0	0
$u_{11}$	0,597	1,76	1,03	0	0,147	0,57	0	1,96	1,83
$u_{12}$	0	0	0	0	0	0	0	0	0
$\zeta_1$	0	0	0	0	0	0	0	0	0
$\zeta_2$	0	0	0	0	0	0	0	0	0
$\zeta_3$	0	0	0	314,4	0	0	312,30	0	0
$\zeta_4$	0	0	0	147,33	0	0	147,61	0	0
$\zeta_5$	0	0	0	1,73	0	0	1,26	0	0
$\zeta_6$	0	0	0	312,98	0	0	311,73	0	0
$\zeta_7$	0	0	0	2621,84	0	0	2620,31	0	0
$\zeta_8$	0	0	0	0	0	0	0	0	0
$\zeta_9$	0	0	0	0	0	0	0	0	0
$\zeta_{10}$	0	0	0	314,8	0	0	312,94	0	0
$\zeta_{11}$	0	0	0	0	0	0	0	0	0
$\zeta_{12}$	0	0	0	2852,65	0	0	2850,54	0	0

Da Tabela 4, pode-se observar que não existe variação de renda ao incrementar uma unidade de insumo nos lotes A, B, C, E, H e I. O maior custo de produção pelo aumento de uma unidade do insumo terra é alcançado no lote D no mês de dezembro (R\$ 2.852,65). Também não foi observada a variação de receita líquida no mês de janeiro, nos lotes D e G. Os meses de março e novembro continuam sendo os meses, para todos os lotes (salvo o lote G), em que se tem custos ligeiramente aumentados pelo incremento de unidade de volume de água. Finalmente, na Tabela 5, são apresentadas as rendas líquidas obtidas para cada lote.

**Tabela 5:** RENDAS LÍQUIDAS OBTIDAS PARA CADA LOTE DO PERÍMETRO IRRIGADO DO GORUTUBA.

LOTES	Receitas Líquidas (R\$ 1.000,00)
A	16.9936,79
B	57.963,69
C	30.883,26
D	2.3045,71
E	3.029,93
F	16.281,42
G	53.502,59
H	24.614,83
I	42.648,61

## 6 CONCLUSÃO

Apresentamos um procedimento computacional de pontos interiores para resolver os problemas primal-dual associados a maximização da receita líquida e à comercialização ótima dos insumos, água e terra. A geração de esquemas ótimos primal–dual para produção e comercialização da renda agrícola com limitações de água e terra representa uma visão ampla no planejamento agrícola com limitações de insumos, em geral. Finalmente, os resultados numéricos obtidos com a aplicação do procedimento aqui apresentado para o cenário agrário considerado, resultam compatíveis com os conhecidos na literatura, o que indica sua validade.

## REFERENCES

- [1] Alfandaria, L., Plateau, A., and Schepeler, X. (2015). “A branch-and-price-and-cut approach for sustainable crop rotation planning”. *European Journal of Operational Research*, 241(3):872–879.
- [2] Bazaarra, M. C., Sherali, H. D., and Shetty, H. D. (1979). *Nonlinear Programming: Theory and Algorithms*. John Wiley.
- [3] Carvalho, D. F., Delgado, A. R. S., de Oliveira, R. F., da Silva, W., and do Forte, V. L. (2009). “Maximização da produção e da receita agrícola com limitações de água e nitrogênio utilizando método de pontos interiores”. *Revista Engenharia Agrícola*, 29(2):321–327.
- [4] Carvalho, D. F., Soares, A., Ribeiro, C. A., Sediama, C. G., and Pruski, F. (2000). “Otimização do uso da água no perímetro irrigado do Gorutuba, utilizando-se a técnica da programação linear”. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 4(2):203–209.
- [5] Delgado, A. R. S. and Ventura, S. D. (2014). “Experiência numérica na maximização da produção e receita agrícola do meloeiro e da alfaca americana”. *Revista Pesquisa Operacional para o Desenvolvimento*, 6(3):332–342.
- [6] Fiacco, A. V. and Cormick, M. C. (1968). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Wiley.
- [7] Frizzone, J. A., Andrade, A. S., de Souza, J. L. M., and Zocoler, J. L. (2005). *lanejamento de irrigação. Análise de decisão de investimentos*. EMBRAPA.
- [8] Galdino, D., Delgado, A. R. S., Ventura, D. S., Dib, C. M., and Rodrigues, P. (2019). “Binary programming for the simulation of crop rotation and animal transit in an integrated crop-livestock system”. *Revista Ciéncia Agronómica*, 50(1):123–130.
- [9] Santos, L. M. R., Santos, R. H., Arenales, M. N., and Raggi, L. A. (2007). “Um modelo para a programação de rotação de culturas”. *Pesq. Operacional*, 27(3):535–547.
- [10] Silva, E. M. et al. (1998). *Pesquisa Operacional: Programação Linear*. 3a ed.
- [11] Ventura, D. S., Delgado, A. R. S., Teixeira, V., et al. (2019). “Optimum Control of Grazing and Cattle Movement in the Silvopastoral Establishment “La Aguada”(Córdoba-Argentina)”. *Journal of Agricultural Studies*, 7(3):9–28.



# Predição para recomendação de filmes com base no agrupamento pelo conteúdo do script

Henrique Matheus Ferreira da Silva<sup>1</sup> and Rafael Silva Pereira<sup>1</sup>

<sup>1</sup> Laboratório Nacional de Computação Científica, Petrópolis/RJ, Brasil

---

## Abstract

sistemas de recomendação se tornaram extremamente populares hoje em dia, utilizados nos mais diferentes setores e são estudados tanto na academia quanto na indústria. Neste trabalho discutimos um processo de recomendação utilizando análise de sentimento sobre scripts de filmes e agrupando filmes de sentimentos similares. Foi utilizado o algoritmo de VAT e estatística de Hopkins para validar a clusterização além do método da silhueta para identificar melhor número de grupos.

**Keywords:** Clusterização, Filmes, análise de Sentimento

---

## 1 Introdução

Neste trabalho será apresentado um processo de clusterização de filmes para identificar que filmes são próximos entre si. Para comparar diferentes filmes, devemos extrair features associadas a este. Neste trabalho as features selecionadas serão extraídas através de um algoritmo de análise de sentimento implementado em (Pereira and Porto, 2019), obtendo estas features para cada *script* extraído, estas são normalizadas e utilizaremos uma implementação do algoritmo k-means para clusterizarmos os filmes pelas features extraídas, de forma a identificar filmes similares pelo mesmo *cluster*, a até mesmo identificação de gêneros em que um filme está localizado.

A área de sistemas de recomendações sobre arquivos textuais é explorada em várias formas. Mais especificamente são encontradas análises de livros utilizando empréstimos de bibliotecas Yang et al. (2009); Tsuji et al. (2014); Maneewongvatana (2010); Pera and Ng (2014), Estilometria Stamatatos (2009); Vaz et al. (2013, 2012); Manning et al. (2008), Comparações mult-camadas Zhang and Chow (2015); Rahman et al. (2007), Opiniões

de usuários J. Mooney and Roy (1999); Garrido et al. (2014, 2013) neste trabalho estamos seguindo a linha de análise de sentimento como Priyanka et al. (2015); Bhure (2015); Aggarwal (2018)

## 2 Metodologia

Neste trabalho temos como objetivo utilizar um conjunto de dados textuais de forma a comparar seus conteúdos e poder retornar ao usuário, dado que este gostou de filme a, deve gostar de filmes b,c,d , utilizando como referência os *scripts* destes. Este tipo de sistema de recomendação pode ser utilizado em serviços de *streaming*, livrarias, entre outros.

Para este fim utilizaremos o algoritmo de análise de sentimento, que recebe como entrada uma *string* e retorna 10 séries, uma para cada sentimento medido. Dadas as séries, podemos compará-las, de forma que as mais próximas entre si representam textos com conteúdo mais próximo entre si.

### 2.1 Material e métodos

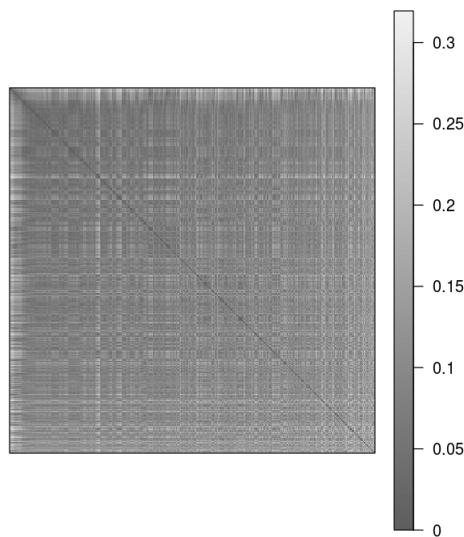
Para a análise, foi construído um *script* para mineração dos dados contidos em [https://www.springfieldspringfield.co.uk/movie\\_scripts.php](https://www.springfieldspringfield.co.uk/movie_scripts.php), o *script* em questão recebe um *link* de entrada, encontra todos os *sublinks* que contém *scripts* armazenados, e isola a parte da página que contém o texto, anexando este a uma lista com índice nomeado pelo nome do filme referenciado.

Tendo então o conjunto de scripts associados em uma lista nomeada, salvamos este no arquivo em anexo *Textos.Rdata*, a seguir para cada um destes, rodamos o algoritmo de análise de sentimento de forma a construir o conjunto de dados que analisaremos, o algoritmo calcula os sentimentos de trechos equi-espacados do texto, faz o somatório, e normaliza para criar um vetor de taxas associado a cada filme. assim é construído um dataset de 15895 linhas e 11 colunas, a coluna 1 determina qual filme a linha está associada.

Tendo este dataset então verificamos se este é clusterizável, utilizando o algoritmo de VAT Bezdek and Hathaway (2002), além da estatística de hopkins. Após validarmos que é clusterizável, selecionamos o melhor número de clusters pelo método de silhueta Peter (1987), para o algoritmo de *kmeans*, e assim geramos os grupos, assim filmes de mesmo grupo pertencem a gêneros similares, e quanto mais próximos entre si, maior a chance que o fâ de um se interesse pelo outro.

## 3 Resultados

Uma descrição completa do processo está descrita no notebook em anexo 'Movie Clustering Large Dataset', caso o usuário tenha interesse em inclusive ter sugestões de filmes pode expandi-o, modificando a seção de *queries* Primeiramente buscamos identificar se existe *clusters* neste dataset, para isto rodamos o algoritmo de vat com o pacote Hahsler et al. (2008), com resultado demonstrado em 1, após isto calculamos a estatística de hopkins utilizando o pacote *clustertend* YiLan and RuTong (2015) obtermos um resultado de 0.22, como referência a estatística de hopkins sobre o dataset iris é 0.17, um dataset comum para exemplos de clusterização.

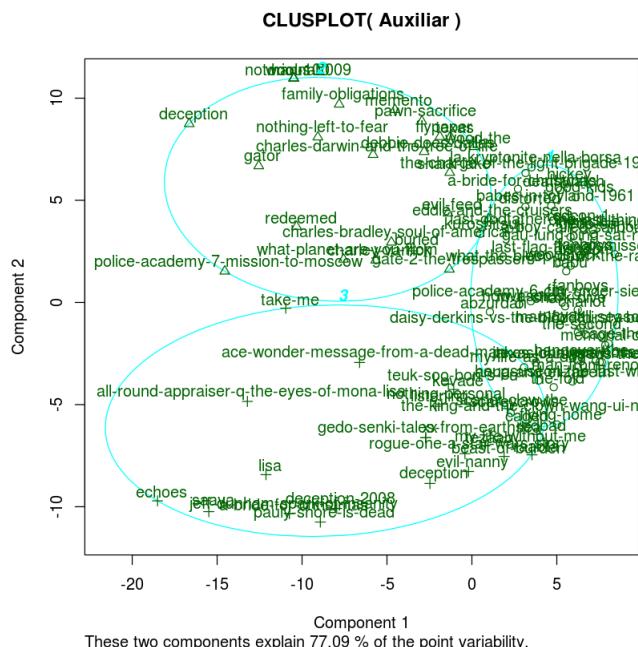


**Fig. 1:** Matriz de dissimilaridade reordenada

Assim podemos visualizar um subset destes clusters e procurar informações que elementos de mesmo cluster tem, a visualização do cluster é apresentada em 2, enquanto informações estatísticas sobre diferentes grupos são descritas nas tabelas 4,5,6.

Para verificarmos então se estes clusters realmente representam grupos diferentes para cada um dos clusters rodamos um teste de *Kolmogorov-Smirnov* sobre cada par de features entre diferentes clusters, temos que se o valor p deste for menor que um threshold (aqui usaremos 10%), rejeitamos a hipótese nula dizendo que os atributos de diferentes clusters não compartilham distribuição.

Os resultados foram:

**Fig. 2:** Imagem de amostragem do cluster**Tabela 1:** P VALOR PARA O TESTE DE *Kolmogorov* PARA PAR CLUSTERS 1,3

Raiva	0.08682838
Antecipação	0.02461349
Nojo	0.66195836
Medo	0.47032043
Felicidade	0.99140958
Negatividade	0.62381944
Positividade	0.73369080
Tristeza	0.86910391
Surpresa	0.73600792
Confiança	0.22847012



**Tabela 2:** P VALOR PARA O TESTE DE *Kolmogorov* PARA PAR CLUSTERS 2,3

Raiva	0.15055066
Antecipação	0.08220645
Nojo	0.13464480
Medo	0.99137804
Felicidade	0.73640654
Negatividade	0.71656641
Positividade	0.28842648
Tristeza	0.59943281
Surpresa	0.55340315
Confiança	0.21247235

**Tabela 3:** P VALOR PARA O TESTE DE *Kolmogorov* PARA PAR CLUSTERS 1,2

Raiva	0.7347548
Antecipação	0.7452620
Nojo	0.4082325
Medo	0.7317382
Felicidade	0.2848410
Negatividade	0.9459201
Positividade	0.1310267
Tristeza	0.3692732
Surpresa	0.9544680
Confiança	0.7787255

Isto implica que Raiva e antecipação vêm de diferentes distribuições no cluster 1 e 3 enquanto o 2 e 3 possuem antecipações diferentes, mostrando que estes 2 sentimentos são os maiores influenciadores para a separação dos gêneros.

#### 4 Conclusões e Trabalhos futuros

Neste trabalho foi possível observar que é possível clusterizar textos em função de suas proporções de sentimentos, com este método é possível gerar sistemas de recomendação onde indicamos que o sentimento possuído pelo indivíduo num determinado filme pode ser equivalente a outros, ou seja, ter uma alta probabilidade de "gostar" do outro determinado filme.

Trabalhos futuros poderiam explorar outras formas de comparar estes textos utilizando scores de sentimentos, como por exemplo, comparação de séries geradas pelo processo de análise de sentimento de forma que séries similares representam textos similares.

#### REFERENCES

- [1] Aggarwal, C. (2018). pages 413–434.
- [2] Bezdek, J. and Hathaway, R. (2002). “Vat: A tool for visual assessment of (cluster) tendency”. pages 2225 – 2230.
- [3] Bhure, P. (2015). “Book recommendation system using opinion mining technique”. *International Journal of Research in Engineering and Technology*, 04:332–334.

- [4] Garrido, A., Buey, M., Escudero, S., Ilarri, S., Mena, E., and B. Silveira, S. (2013). “Tm-gen: A topic map generator from text documents”. pages ??
- [5] Garrido, A., Pera, M., and Ilarri, S. (2014). “Sole-r: A semantic and linguistic approach for book recommendations”.
- [6] Hahsler, M., Hornik, K., and Buchta, C. (2008). “Getting things in order: An introduction to the r package seriation”. *Journal of Statistical Software*, 25(3):1–34.
- [7] J. Mooney, R. and Roy, L. (1999). “Content-based book recommending using learning for text categorization”. *Proceedings of the Fifth ACM Conference on Digital Libraries*.
- [8] Maneewongvatana, S. (2010). “A recommendation model for personalized book lists”. pages 389 – 394.
- [9] Manning, C., Raghavan, P., and Schütze, H. (2008). “Introduction to information retrieval, chapt”. *Evaluation in Information Retrieval*, pages 151–175.
- [10] Pera, M. and Ng, Y.-K. (2014). “Automating readers’ advisory to make book recommendations for k-12 readers”. *RecSys 2014 - Proceedings of the 8th ACM Conference on Recommender Systems*, pages 9–16.
- [11] Pereira, R. S. and Porto, F. A. (2019). “Dss: A data science suite”. *SBBR 2019 Proceedings*.
- [12] Peter, R. (1987). “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis”. *J. Comput. Appl. Math.*, 20.
- [13] Priyanka, K., Shanker Tewari, A., and Barman, A. (2015). “Personalised book recommendation system based on opinion mining technique”. pages 285–289.
- [14] Rahman, M., Pi Yang, W., Chow, T., and Wu, S. (2007). “A flexible multi-layer self-organizing map for generic processing of tree-structured data”. *Pattern Recognition*, 40:1406–1424.
- [15] Stamatatos, E. (2009). “A survey of modern authorship attribution methods”. *JASIST*, 60:538–556.
- [16] Tsuji, K., Takizawa, N., Sato, S., Ikeuchi, U., Ikeuchi, A., Yoshikane, F., and Itsumura, H. (2014). “Book recommendation based on library loan records and bibliographic information”. *Procedia - Social and Behavioral Sciences*, 147.
- [17] Vaz, P. C., Martins de Matos, D., and Martins, B. (2012). “Stylometric relevance-feedback towards a hybrid book recommendation algorithm”. pages 13–16.
- [18] Vaz, P. C., Ribeiro, R., and Martins de Matos, D. (2013). “Book recommender prototype based on author’s writing style”. pages 227–228.
- [19] Yang, X., Zeng, H., and Huang, Y. (2009). “Artmap-based data mining approach and its application to library book recommendation”. pages 26 – 29.
- [20] YiLan, L. and RuTong, Z. (2015). “clustertend: Check the clustering tendency”. R package version 1.4.
- [21] Zhang, H. and Chow, T. (2015). “Organizing books and authors by multilayer som”. *IEEE Transactions on Neural Networks and Learning Systems*, 27:1–14.



## 5 Apendice

**Tabela 4:** DISTRIBUIÇÃO DE DADOS DO *cluster1*

	anger	anticipation	disgust	fear
Min.	0.03204	0.04826	0.01681	0.04669
1st Qu.	0.06107	0.09724	0.04714	0.07663
Median	0.07092	0.10906	0.05480	0.08778
Mean	0.07180	0.10913	0.05674	0.09000
3rd Qu.	0.08068	0.12036	0.06448	0.10164
Max.	0.12915	0.17506	0.11757	0.16267
	joy	negative	positive	sadness
Min.	0.03070	0.09635	0.1047	0.03851
1st Qu.	0.07770	0.13333	0.1656	0.07095
Median	0.09150	0.14616	0.1814	0.07941
Mean	0.09195	0.14743	0.1814	0.08077
3rd Qu.	0.10553	0.16086	0.1972	0.08931
Max.	0.15492	0.22851	0.2634	0.14079
	surprise	trust		
Min.	0.02391	0.06507		
1st Qu.	0.04604	0.10486		
Median	0.05249	0.11708		
Mean	0.05292	0.11784		
3rd Qu.	0.05929	0.12899		
Max.	0.09418	0.19756		

**Tabela 5:** DISTRIBUIÇÃO DE DADOS DO *cluster* 2

	anger	anticipation	disgust	fear
Min.	0.007752	0.04286	0.007752	0.007752
1st Qu.	0.038515	0.10487	0.030896	0.054031
Median	0.046925	0.12443	0.038206	0.066943
Mean	0.052355	0.12070	0.042550	0.074981
3rd Qu.	0.058501	0.13953	0.050633	0.087279
Max.	0.144099	0.21978	0.118705	0.182401
	joy	negative	positive	sadness
Min.	0.01983	0.05912	0.06667	0.01149
1st Qu.	0.08772	0.09976	0.20348	0.05573
Median	0.11877	0.11111	0.22152	0.06084
Mean	0.11318	0.12250	0.21237	0.07167
3rd Qu.	0.13924	0.13223	0.24184	0.08219
Max.	0.26437	0.26629	0.32471	0.18130
	surprise	trust		
Min.	0.00000	0.0000		
1st Qu.	0.04545	0.1113		
Median	0.05655	0.1379		
Mean	0.05616	0.1335		
3rd Qu.	0.06651	0.1609		
Max.	0.09766	0.2045		

**Tabela 6:** DISTRIBUIÇÃO DE DADOS DO *cluster* 3

	anger	anticipation	disgust	fear
Min.	1	0	0	0
1st Qu.	1	0	0	0
Median	1	0	0	0
Mean	1	0	0	0
3rd Qu.	1	0	0	0
Max.	1	0	0	0
	joy	negative	positive	sadness
Min.	0	0	0	0
1st Qu.	0	0	0	0
Median	0	0	0	0
Mean	0	0	0	0
3rd Qu.	0	0	0	0
Max.	0	0	0	0
	surprise	trust		
Min.	0	0		
1st Qu.	0	0		
Median	0	0		
Mean	0	0		
3rd Qu.	0	0		
Max.	0	0		



# Proposta de uma heurística construtiva para o projeto de circuitos combinacionais

Hugo Assis de Oliveira<sup>1</sup> and Francisco Augusto Lima Manfrini<sup>2</sup>

<sup>1</sup> IFSEMG - Campus Juiz de Fora, Juiz de Fora/MG, Brasil

<sup>2</sup> IFSEMG - Campus Juiz de Fora, Juiz de Fora/MG, Brasil

---

## Abstract

Evolutionary algorithms are a set of strategies to find optimal solutions in many different problems. When applied in the development of digital circuits, these techniques are often referred to as evolvable hardware. This paper presents the proposal of a new technique based on cartesian genetic programming, one form of genetic algorithm. The technique consists in a constructive heuristic where circuits are built up by grouping smaller circuits. By creating a network of multiplexers, it seeks to reduce the problem's complexity with the selective addressing of the circuit's logic gates and inputs. A group of functions from some common circuits were tested with this algorithm. As a result, the technique was able to optimize the number of logic gates in some cases. Thus, the algorithm's capacity on finding solutions was verified, besides the advantage of using a constructive heuristic in circuit design.

**Keywords:** Evolutionary Algorithm, Cartesian Genetic Programming, Multiplexer, Circuit Design

---

## 1 INTRODUÇÃO

Com o aumento da complexidade dos dispositivos digitais, o projeto de circuitos lógicos se torna uma tarefa cada vez mais difícil. Utilizando técnicas convencionais de simplificação por álgebra booleana ou mapas de Karnaugh, especialistas são capazes de desenvolver circuitos funcionais. Porém, em muitos casos surge a necessidade de se implementar circuitos ainda menores e mais eficientes. Com isso, diferentes algoritmos foram desenvolvidos para a busca de soluções otimizadas (1).

Algoritmos evolutivos consistem em uma das estratégias de busca em problemas de otimização. Nessa técnica, a busca é feita tendo como base a teoria da evolução das espécies proposta por Darwin. Essa abordagem possibilita o descobrimento de configurações inovadoras tendo em vista a capacidade de processamento dos computadores e o número de possibilidades que podem ser testadas num curto período de tempo (1).

Contato: Hugo Assis, hugo.4600@gmail.com

Entre estes algoritmos estão as técnicas de programação genética cartesiana (CGP) (2). Esta técnica surgiu a partir de métodos de evolução de circuitos digitais. Em CGP os indivíduos são representados por um conjunto de funções dispostas numa matriz de nós. As conexões entre os nós dependem de sua posição na matriz enquanto as funções dos nós seguem uma lista definida pelo usuário, podendo representar diferentes portas lógicas no caso de circuitos digitais (2).

Neste trabalho, é apresentada a proposta de uma nova heurística baseada em CGP para o projeto de circuitos lógicos combinacionais (CLC's). Pela evolução sequencial de circuitos menores tendo como base multiplexadores, foi possível desenvolver uma heurística construtiva onde o foco está em encontrar as menores configurações de circuito que satisfaçam a saída desejada. Assim, o tamanho do circuito aumenta somente enquanto a solução não é encontrada ou até atingir um limite superior definido.

Uma técnica semelhante é mostrada em (5) e (1), onde os autores se utilizam de multiplexadores como base para a evolução dos circuitos. A premissa vêm de que qualquer expressão booleana pode ser solucionada utilizando-se apenas multiplexadores de duas entradas. Algumas das funções testadas aqui são as mesmas do trabalho citado e foram escolhidas para critério de comparação entre as técnicas.

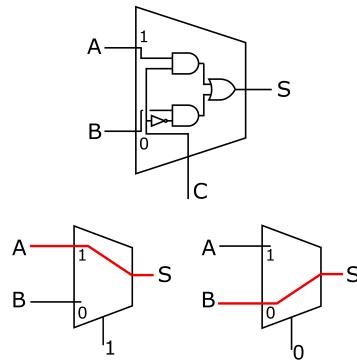
## 2 HEURÍSTICA CONSTRUTIVA

Os multiplexadores são dispositivos capazes de direcionar apenas um dos dois sinais de entrada à saída dependendo do sinal recebido no seu terminal de controle. Como representado na Fig. 1, o sinal da saída é igual ao sinal de entrada de uma das entradas do multiplexador cuja escolha é feita pelo circuito de controle. O objetivo dos multiplexadores neste algoritmo é dividir a tabela verdade até que apareçam na entrada conjuntos de linhas contendo apenas uns ou zeros. Dessa forma, os sinais de entrada são "filtrados" onde são indesejados e são conectados diretamente à saída quando representam o sinal desejado.

Para que haja a construção dos circuitos durante o processo evolutivo, um circuito contendo um multiplexador e um conjunto de portas lógicas em seu controle são gerados aleatoriamente e evoluídos de forma a dividir ao máximo a saída lógica entre "uns" e "zeros". Ao final desse processo um teste é realizado entre a saída desejada e a obtida com o circuito gerado e, caso ainda haja erros, novos circuitos são conectados à entrada deste multiplexador e são evoluídos de forma a dividir ainda mais a tabela verdade. Este novo conjunto de multiplexadores adicionado corresponde a um novo nível. Novos níveis são inseridos até que a solução seja encontrada ou o número máximo de níveis seja atingido.

### 2.1 Codificação

A representação dos genótipos é feita numa lista unidimensional contendo todas as características daquele indivíduo. Como mostrado na Fig. 2 cada indivíduo possui um conjunto de variáveis de entrada, um conjunto de multiplexadores contendo um circuito de controle próprio e uma única saída. Também mostra-se a tabela de codificação utilizada para as funções lógicas utilizadas. Ela indica quais são as possíveis portas lógicas que podem ser inseridas no circuito durante a evolução. No algoritmo, cada multiplexador não é declarado explicitamente, sendo caracterizado no vetor pelo conjunto de portas lógicas que



**Fig. 1:** Funcionamento dos multiplexadores. O circuito superior mostra o circuito interno do multiplexador enquanto as demais representam a conexão entre a saída com as entradas A e B dependendo do valor lógico presente na entrada de controle C.

podem ou não estar conectadas a seu controle, pelo índice da porta que é efetivamente conectada à entrada do multiplexador e pela sua pontuação. As portas lógicas dos circuitos de controle são representadas por um conjunto de nós unidimensional com base na programação genética cartesiana. A cada multiplexador adicionado calcula-se sua pontuação com relação às linhas que aquele multiplexador tem acesso. Ela indica o quanto este circuito está acertando e é utilizada para análise do desempenho daquele conjunto de portas lógicas. Na imagem, a pontuação de cada multiplexador é indicada por P1, P2 e P3.

## 2.2 Inicialização

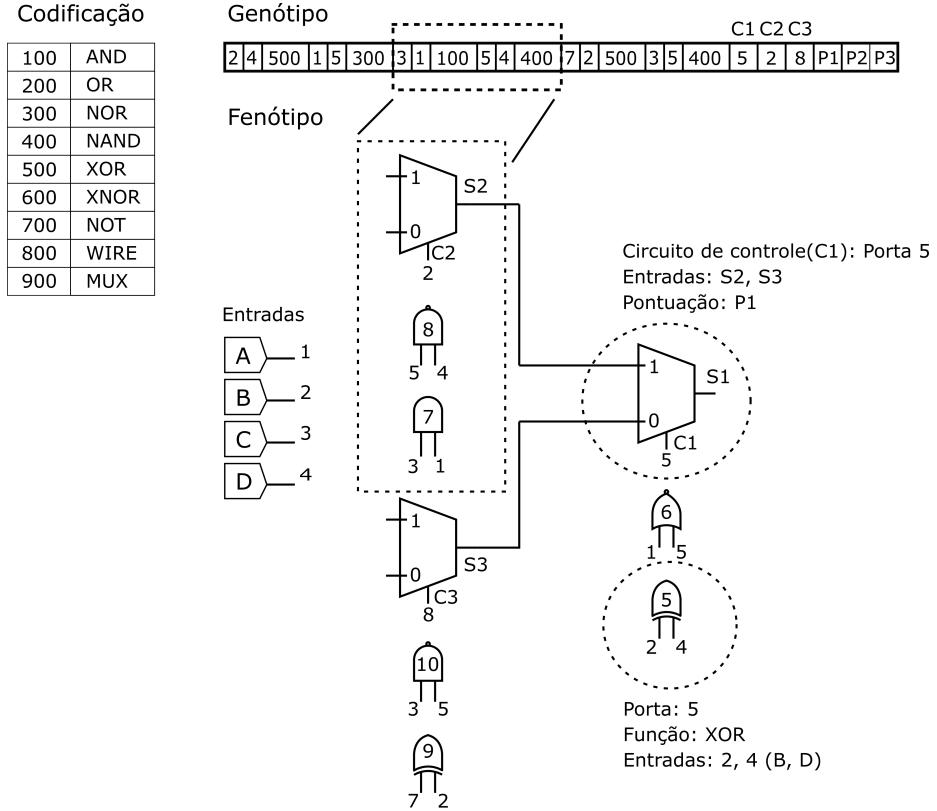
No genótipo, o algoritmo começa pela evolução do controle do primeiro multiplexador. Ao encontrar um circuito capaz de dividir a tabela em dois conjuntos compostos majoritariamente por "uns" ou "zeros" de acordo com a saída da tabela verdade, um teste é feito para avaliar se a solução pode ser obtida. Caso ainda haja erro, são adicionados outros circuitos nas entradas que ainda precisam de divisão.

O número de circuitos adicionados depende do número máximo de níveis determinado pelo usuário. O primeiro nível consiste no primeiro multiplexador com seu circuito de controle, o segundo nível são os dois multiplexadores que podem ser ligados às entradas do primeiro e os próximos níveis são aqueles conjuntos de circuitos adicionados às entradas de multiplexadores de um nível anterior. O número total de multiplexadores possíveis dado o número máximo de níveis é calculado pela Eq. (1).

$$N_{mux} = 2^{N_{niveis}} - 1 \quad (1)$$

Onde  $N_{mux}$  representa o número de multiplexadores de base e  $N_{niveis}$  é o número máximo de níveis definido. Já o número máximo de portas lógicas para um genótipo também depende do número de portas no controle de cada multiplexador e pode ser calculado pela Eq. (2).

$$N_{portas} = N_{mux} + N_{mux}N_{tamControle} \quad (2)$$



**Fig. 2:** Codificação proposta do genótipo para um circuito com dois níveis e quatro entradas A, B, C e D. À direita nota-se a função e as entradas de uma das portas do primeiro multiplexador, enquanto este possui uma porta ligada à entrada de controle C1 com pontuação P1 no genótipo. A tabela à esquerda indica as possíveis portas lógicas e sua codificação no genótipo sendo WIRE uma conexão direta sem portas e MUX um multiplexador. A imagem também ressalta a parte do circuito que representa um pedaço do genótipo.

Onde  $N_{tamControle}$  é o número máximo de portas para cada circuito de controle ligado a cada multiplexador.

### 2.3 Mutação

Nos circuitos de controle, os genes que representam as portas são declarados como uma linha de nós onde um só é capaz de se conectar às entradas e aos nós anteriores. Durante a evolução, as mutações são feitas aleatoriamente e pontualmente nas conexões dos nós ou em suas funções. Em seguida, avalia-se o quanto bem aquele multiplexador está dividindo a tabela verdade sendo que o melhor da população é utilizado para a próxima geração, comportamento conhecido como elitismo. Da Eq. (3), o tamanho da população consiste na soma do número de pais e de filhos gerados durante as mutações.

$$(1 + \Lambda) \quad (3)$$



Em que 1 representa o número de pais e  $\lambda$  o número de filhos. Em CGP, o número de filhos melhora a eficiência da evolução quando fixado em quatro ao passo que o número de pais se mantém em um.

Ao final da última geração, se o circuito encontrado não conseguir dividir todas as linhas então outros circuitos são adicionados às suas entradas e evoluídos até a obtenção de uma solução aceitável. A exceção está nos casos em que se atinge o último nível e a solução não é obtida.

## 2.4 CGP

Nos casos onde o circuito atinge o último nível e não encontra solução ou quando um circuito adicionado não é capaz de dividir ainda mais a tabela verdade, utilizou-se a programação genética cartesiana como uma alternativa para alguns ramos do circuito. Nestas situações a solução encontrada foi inserir um conjunto de portas lógicas utilizando CGP na posição destes multiplexadores.

No algoritmo, essa matriz é chamada por uma função ao final do processo evolutivo do circuito principal. Para cada chamada da função, um circuito é evoluído seguindo os padrões da programação genética cartesiana e tendo em vista a solução das linhas que não foram divididas pelo multiplexador daquela posição. Na cartesiana, o indivíduo consiste num conjunto de nós dispostos numa matriz cujas colunas limitam a conexão entre os mesmos. As entradas de um nó só podem estar conectadas às variáveis de entrada e à saída de nós de colunas anteriores de acordo com a chamada taxa de retorno (1).

### 2.4.1 Mutação em CGP

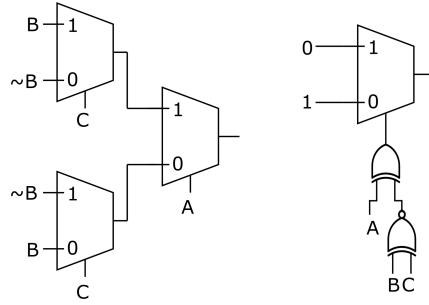
Na literatura, diferentes estudos mostram que a evolução em CGP é favorecida pela deriva gênica (3). A presença de nós inativos, ou seja, nós que não influenciam nas saídas, permite que os indivíduos escapem de ótimos locais durante a evolução. Como apresentado em (3), as mutações ocorrem ao longo de todo o genótipo em alelos de entrada e de funções pontualmente a cada geração até que um nó ativo seja mutado, o que garante que haverá alterações no fenótipo a cada indivíduo gerado.

## 3 EXPERIMENTOS COMPUTACIONAIS

A análise do desempenho da heurística proposta foi feita comparando com três exemplos *benchmarks* estudados por vijayakumari (5) que serão apresentados a seguir. Entre eles estão os circuitos geradores e verificadores de paridade, utilizados na identificação de erro durante a transmissão de informações entre dispositivos. Nesses sistemas, o número de 1's presente num conjunto de bits é fixado como sendo par ou ímpar. Caso haja alterações, o que é identificado pelos circuitos verificadores, é possível saber que houve falha na comunicação.

### 3.1 Função 1: $F = \sum(1, 2, 4, 7)$

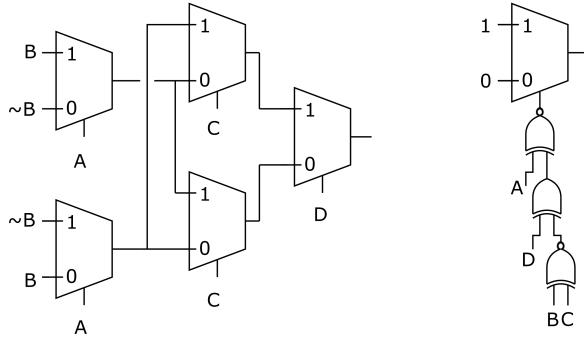
A primeira função corresponde ao verificador de paridade ímpar de três bits. A Fig. 3 mostra a solução encontrada pelos autores em comparação com esta heurística. Como é mostrado, o número de portas lógicas mínimo encontrado pelos dois algoritmos é equivalente e possui 4 portas a menos do que é comumente utilizado (5).



**Fig. 3:** Soluções encontradas para a função 1. À esquerda é mostrado o circuito encontrado pelo método de Vijayakumari enquanto o circuito à direita representa uma das melhores soluções encontradas por esta técnica.

### 3.2 Função 2: $F = \sum(1, 2, 4, 7, 8, 11, 13, 14)$

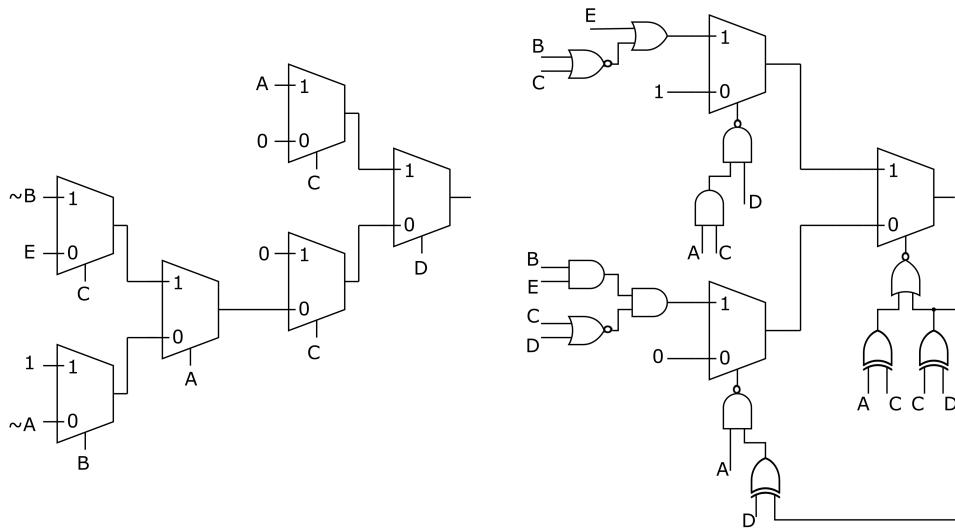
A segunda função é referente ao verificador de paridade ímpar de quatro bits. Como representado na Fig. 4 o circuito apresenta apenas quatro portas lógicas em comparação com as 15 utilizadas em média. A solução encontrada apresentou uma porta lógica a menos do que a encontrada no método I em (5) e uma mais do que foi encontrado pelo método II.



**Fig. 4:** Circuitos obtidos para a função 2, sendo o primeiro o encontrado pelos outros autores e o segundo o que foi encontrado neste.

### 3.3 Função 3: $F = \sum(0, 1, 9, 22, 23, 25, 30, 31)$

Por fim, neste último exemplo a solução encontrada apresentou 15 portas lógicas. Da Fig. 5, o circuito apresenta 16 portas a menos do que é comumente utilizado. Também possui 10 portas a mais do que o encontrado pelo método I de (5) e 11 a mais do que o método II dos mesmos autores.



**Fig. 5:** Circuitos obtidos para a função 3. À esquerda é mostrada uma solução encontrada em (5) e à direita uma encontrada por esta técnica.

## 4 RESULTADOS E DISCUSSÕES

Após várias simulações, observou-se que o número mínimo de níveis e portas lógicas em cada controle depende da complexidade da saída desejada. Verificou-se que em alguns casos uma solução é obtida com poucas gerações e com poucas portas para que uma solução funcional possa ser encontrada. Já para outras funções, a inicialização deve ser feita com um número alto de níveis e com muitas portas nos circuitos de controle.

As simulações foram realizadas com objetivo de encontrar soluções para algumas das funções em dispositivos digitais. Aqui três exemplos foram testados mas o algoritmo pode de ser facilmente aplicado a outros problemas.

Como mostrado, os circuitos de controle podem possuir um número diferente de portas limitados apenas pelo número máximo inserido na inicialização. Outra característica importante e que pode ser observada nos circuitos obtidos é de que o controle de multiplexadores de níveis posteriores acabam utilizando portas de multiplexadores de níveis anteriores em alguns casos. Esse comportamento acaba permitindo o desenvolvimento de circuitos ainda menores, característica fundamental no projeto de CLC's.

## 5 CONCLUSÃO

A partir dos resultados obtidos verifica-se a eficiência do algoritmo na solução de algumas funções e com um conjunto bem reduzido de portas lógicas. Porém, para funções mais complexas, mudanças no número de portas lógicas e na estratégia de evolução poderiam levar a uma otimização na busca pelo algoritmo.

Com este trabalho também foi possível avaliar a eficiência da programação genética cartesiana no projeto de circuitos lógicos reduzidos para uma variedade de funções. Pela implementação da heurística, observou-se uma redução no tamanho do circuitos assim

como no tempo de projeto se comparado a técnicas convencionais. Novas estratégias de evolução e otimizações nas configurações dos circuitos poderiam levar a resultados melhores e em menor tempo.

## REFERÊNCIAS

- [1] MANFRINI, Francisco; BARBOSA, Helio JC; BERNARDINO, Heder S. *Optimization of combinational logic circuits through decomposition of truth table and evolution of sub-circuits*. In: 2014 IEEE Congress on Evolutionary Computation (CEC). IEEE, 2014. p. 945-950.
- [2] MILLER, Julian F. *Cartesian genetic programming* Springer, Berlin, Heidelberg, p. 17-34, 2011.
- [3] TURNER, Andrew James; MILLER, Julian Francis. *Neutral genetic drift: an investigation using Cartesian Genetic Programming*. Genetic Programming and Evolvable Machines, v. 16, n. 4, p. 531-558, 2015.
- [4] MILLER, Julian Francis. *The alchemy of computation: designing with the unknown*. Natural Computing, p. 1-12, 2019.
- [5] VIJAYAKUMARI, C. K. et al. *Genetic algorithm based design of combinational logic circuits using universal logic modules*. Procedia Computer Science, v. 46, p. 1246-1253, 2015.



# Segmentação de imagens digitais via modelos variacionais com restrições relaxadas

Italo Messias Felix Santos<sup>1</sup>, Gilson Antonio Giraldi<sup>1</sup> e Abimael Dourado Loula<sup>1</sup>

<sup>1</sup> LNCC - Laboratório Nacional de Computação Científica, Petrópolis/RJ, Brasil

---

## Resumo

Este trabalho tem como objetivo a segmentação de imagens digitais a partir do modelo de Ambrosio e Tortorelli que se trata de uma aproximação contínua do modelo de Mumford e Shah utilizando artifícios de  $\Gamma$ -convergência. Essa simplificação pode exigir o uso de funções descontínuas como a função Heaviside e/ou distribuições como o delta de Dirac para possibilitar computação de propriedades em regiões específicas da imagem. Essa forma de tratar o problema é utilizada em métodos consagrados como o Modelo de Chan e Vese. Nossa proposta é evitar os problemas decorrentes do uso das funções acima no cálculo das integrais associadas ao modelo. Com este propósito, adicionamos termos quâdricos acoplados aos principais termos do modelo variacional de Ambrosio e Tortorelli de maneira relaxada. Assim, embora o novo termo não garanta as propriedades desejadas para o mínimo do funcional, esta abordagem evita os problemas mencionados. A metodologia produziu resultados semelhantes à formulação proposta por Chan e Vese com uma expressão simples que possibilita o uso de métodos numéricos mais eficientes.

**Palavras-chave:** Métodos variacionais, Processamento de imagens, Método de contornos ativos, Segmentação de imagens.

---

## 1 INTRODUÇÃO

Segmentação é uma área de intensa pesquisa em processamento de imagens uma vez que é uma etapa fundamental para a análise de imagens. Segmentar uma imagem significa dividi-la em regiões disjuntas e homogêneas em relação a um determinado critério [4]. Nesta área, algoritmos baseados em morfologia matemática, contornos ativos, técnicas baseadas em projeções, técnicas de limiarização e modelos variacionais vêm sendo usados para extração de objetos de interesse.

Contato: Italo, italof@lncc.br

No contexto de segmentação de imagens desejamos separar uma imagem em duas regiões distintas. Para isso, a partir da formulação de Ambrosio e Tortorelli [2] (uma aproximação para o modelo de Mumford e Shah [6] que visa aproximar uma dada função por outra contínua por partes) exploramos o caso em que a resposta do método é uma função constante por partes, mais precisamente queremos que o resultado possua somente dois valores distintos em duas regiões do domínio da imagem.

O método de Chan e Vese [3] explora esta abordagem no modelo de Mumford e Shah através de contornos ativos utilizando funções descontínuas como a Heaviside e/ou distribuição delta de Dirac, garantindo propriedades como comprimento de arco, área dentre outras. Nossa proposta é obter resultados semelhantes a partir de relaxações, evitando o uso de funções que impossibilitam a aplicação de métodos numéricos mais eficientes para aproximar a solução do problema.

O texto está organizado como segue. Na seção 2, apresentamos os fundamentos da técnica em foco. Em seguida, na seção 3 desenvolvemos o método proposto, condições de mínimo, relaxações e método de diferenças finitas utilizado. Os experimentos computacionais são discutidos na seção 4. As conclusões são apresentadas na seção 5.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste trabalho, uma imagem digital com domínio  $\Omega \subset \mathbb{R}^2$  é representada por uma função  $g : \Omega \rightarrow [0, 1] \subset \mathbb{R}$ , sendo  $g$  uma função quadrado integrável ( $g \in L^2(\Omega)$ ). A premissa do modelo de Ambrosio e Tortorelli é determinar  $u_{min} \in H^1(\Omega) = \left\{ f \in L^2(\Omega); \frac{\partial f}{\partial x} \in L^2(\Omega) \right.$   
 $\left. \text{e } \frac{\partial f}{\partial y} \in L^2(\Omega) \right\}$  (que representa uma versão contínua por partes de  $g$ ), que minimiza o funcional  $E(u)$  proposto por Mumford e Shah [6]. Para isto, em [2] utilizam-se recursos de  $\Gamma$ -convergência para definir um novo funcional  $E_c : H^1(\Omega) \times H^1(\Omega) \rightarrow \mathbb{R}$ , dado por

$$E_c(u, v) = \frac{1}{2} \int_{\Omega} \{ \beta(u - g)^2 + v^2 \|\nabla u\|^2 + \alpha M_c(v) \} d\Omega, \quad (1)$$

onde

$$M_c(v) = c \|\nabla v\|^2 + \frac{(1-v)^2}{4c},$$

$c$  é um número pequeno,  $\alpha > 0$  e  $\beta > 0$  são parâmetros e  $g \in L^2(\Omega)$  é a imagem original. Além disso,  $E_c \rightarrow E$  quando  $c \rightarrow 0$ , de tal forma que as sequências minimizantes  $u_c, v_c$  de  $E_c$  convergem para  $(u_{min}, v_{min}) \in H^1(\Omega) \times H^1(\Omega)$ .

No funcional (1) os pontos de borda de uma imagem  $u$  são representados pelo conjunto  $C = \{\mathbf{x} \in \Omega; v(\mathbf{x}) = 0\}$ . No caso particular em que

$$u(\mathbf{x}) = \begin{cases} u_1, & \text{se } \mathbf{x} \in \Omega_1 \\ u_2, & \text{se } \mathbf{x} \in \Omega_2 \end{cases}, \quad (2)$$

onde  $u_1$  e  $u_2$  são constantes,  $\Omega = \Omega_1 \cup \Omega_2 \cup C$  com  $\Omega_1 \cap \Omega_2 \cap C = \emptyset$  uma boa forma de representar os subconjuntos de  $\Omega$  é utilizar o métodos dos conjuntos de níveis e tratar  $C$  como a curva de nível zero de  $v$ , em outras palavras, desejamos que  $\Omega_1 = \{\mathbf{x} \in \Omega; v(\mathbf{x}) < 0\}$ ,  $\Omega_2 = \{\mathbf{x} \in \Omega; v(\mathbf{x}) > 0\}$  e  $C = \{\mathbf{x} \in \Omega; v(\mathbf{x}) = 0\}$ . Apesar de simples a formulação



original Ambrosio e Tortorelli não permite esta escolha uma vez que  $v(\mathbf{x}) \in [0, 1]$ . De acordo com [1] a alteração no funcional (1) dada por

$$E_c(u, v) = \frac{1}{2} \int_{\Omega} \left\{ \beta(u - g)^2 + \left( \frac{1}{2} - v \right)^2 \|\nabla u\|^2 + \alpha M_c(v) \right\} d\Omega, \quad (3)$$

onde

$$M_c(v) = c \|\nabla v\|^2 + \frac{(3v(1-v))^2}{c},$$

permite que  $C = \{\mathbf{x} \in \Omega; v(\mathbf{x}) = \frac{1}{2}\}$ , desta forma, podemos considerar  $\Omega_1 = \{\mathbf{x} \in \Omega; v(\mathbf{x}) < \frac{1}{2}\}$ ,  $\Omega_2 = \{\mathbf{x} \in \Omega; v(\mathbf{x}) > \frac{1}{2}\}$ . Além disso, o segundo termo é nulo uma vez que  $u$  é constante por partes,  $(\frac{1}{2} - v)^2$  é nulo em  $C$ . Este fato, juntamente com a expressão (2) permite simplificar o funcional (3) como

$$E_c(u_1, u_2, v) = \frac{\beta}{2} \left( \int_{\Omega_1} (u_1 - g)^2 dx + \int_{\Omega_2} (u_2 - g)^2 dx \right) + \frac{\alpha}{2} \int_{\Omega} \left( c \|\nabla v\|^2 + \frac{(3v(1-v))^2}{c} \right) dx. \quad (4)$$

Para calcular as integrais em  $\Omega_1$  e  $\Omega_2$  é natural utilizar as aproximações contínuas

$$H_{\epsilon}(t) = \frac{1}{2} \left( 1 + \frac{2}{\pi} \arctan \left( \frac{t}{\epsilon} \right) \right), \quad (5)$$

$$\delta_{\epsilon}(t) = \frac{\epsilon}{\pi(\epsilon^2 + t^2)}, \quad (6)$$

para as funções

$$H(t) = \begin{cases} 1, & \text{se } t \geq 0 \\ 0, & \text{se } t < 0 \end{cases}, \quad \delta(t) = \begin{cases} \infty, & \text{se } t = 0 \\ 0, & \text{caso contrário} \end{cases}. \quad (7)$$

conhecidas como Heaviside e distribuição delta de Dirac, respectivamente.

Chan e Vese [3] utilizam  $H$  para efetuar o cálculo nos domínios específicos e  $\delta$  para computar o comprimento da curva  $C$  considerando todo o domínio  $\Omega$ . Na formulação de Ambrosio e Tortorelli [2] o comprimento de arco é aproximado com auxílio do terceiro termo do funcional (1) para  $c$  pequeno. Utilizando as funções acima o funcional (4) pode ser reescrito como

$$\begin{aligned} E_c(u_1, u_2, v) = & \frac{\beta}{2} \left( \int_{\Omega} (u_1 - g)^2 H_{\epsilon}(0.5 - v) dx + \int_{\Omega} (u_2 - g)^2 (1 - H_{\epsilon}(0.5 - v)) dx \right) \\ & + \frac{\alpha}{2} \int_{\Omega} \left( c \|\nabla v\|^2 + \frac{(3v(1-v))^2}{c} \right) dx. \end{aligned} \quad (8)$$

Para determinarmos as condições de mínimo precisamos calcular a primeira variação de Gateux para o funcional (8), o que recai em termos envolvendo a equação (9).

$$\delta_{\epsilon}(t) = \frac{\partial}{\partial t} H_{\epsilon}(t) = H'_{\epsilon}(t). \quad (9)$$

### 3 MODELO PROPOSTO

Nesta seção, iremos propor uma abordagem simples para evitar o uso das funções  $H_\epsilon$  e  $\delta_\epsilon$  permitindo boas aproximações através do método das direções alternadas (ADI - *Alternating Direction Implicit*) [5]. Para isso, supomos que  $v$  permite representar as regiões  $\Omega_1$  e  $\Omega_2$  na forma  $\Omega_1 = \{\mathbf{x} \in \Omega; v(\mathbf{x}) \approx 1\}$  e  $\Omega_2 = \{\mathbf{x} \in \Omega; v(\mathbf{x}) \approx 0\}$  onde  $v(\mathbf{x}) = 0.5$  é o limiar que separa ambos os conjuntos  $\Omega_1$  e  $\Omega_2$  pelas propriedades da função  $v$ . Efetuadas as considerações acima, apresentamos o modelo proposto dado por

$$\begin{aligned} E_c(u_1, u_2, v) &= \frac{\beta}{2} \left( \int_{\Omega} (u_1 - g)^2 v^2 dx + \int_{\Omega} (u_2 - g)^2 (1 - v)^2 dx \right) \\ &\quad + \frac{\alpha}{2} \int_{\Omega} \left( c \|\nabla v\|^2 + \frac{(3v(1-v))^2}{c} \right) dx. \end{aligned} \quad (10)$$

Em relação ao funcional (10) temos as seguintes observações:

- O acoplamento de  $v^2$  no primeiro termo torna o cálculo de  $\int_{\Omega} (u_1 - g)^2$  desprezível para  $v$  próximo de 0.
- O acoplamento de  $(1 - v)^2$  no segundo torna o cálculo de  $\int_{\Omega} (u_2 - g)^2$  desprezível para  $v$  próximo de 1.

A relaxação do primeiro e do segundo termo não garantem que

$$\int_{\Omega_1} (u_1 - g)^2 dx = \int_{\Omega} (u_1 - g)^2 v^2 dx \quad \text{e} \quad \int_{\Omega_2} (u_2 - g)^2 dx = \int_{\Omega} (u_2 - g)^2 (1 - v)^2 dx,$$

mas permite obter uma aproximação para o mínimo de (3) através da expressão (10). Nas próximas seções iremos utilizar somente o funcional  $E_c$  presente na expressão (10).

#### 3.1 Primeira variação do funcional relaxado

Primeiramente, note que para  $v \in H^1(\Omega)$  fixado,  $E_c$  é uma função em  $\mathbb{R}^2$  logo a condição necessária de mínimo ocorre quando o gradiente de  $E_c$  é nulo com respeito a  $u_1$  e  $u_2$ , daí

$$\begin{aligned} \nabla(E_c(u_1, u_2)) &= \beta \left( \int_{\Omega} (u_1 - g)v^2 dx, \int_{\Omega} (u_2 - g)(1 - v)^2 dx \right) = (0, 0), \\ \Rightarrow & \begin{cases} u_1 = \frac{\int_{\Omega} gv^2 dx}{\int_{\Omega} v^2 dx} \\ u_2 = \frac{\int_{\Omega} g(1 - v)^2 dx}{\int_{\Omega} (1 - v)^2 dx} \end{cases}. \end{aligned} \quad (11)$$

Conhecendo os valores de  $u_1$  e  $u_2$  para cada  $v$  podemos determinar as condições necessárias, em relação a  $v$ , para um mínimo de (10). Utilizamos para isso a diferencial de



Gateaux e verificamos que  $v \in H^1(\Omega)$  deve satisfazer

$$\begin{aligned} & \int_{\Omega} \left( \beta(u_1 - g)^2 v \hat{v} + \beta(u_2 - g)^2 v \hat{v} - \beta(u_2 - g) \hat{v} \right. \\ & \quad \left. + c\alpha \nabla v \cdot \nabla \hat{v} + \frac{9\alpha}{c} v(1-v)(1-2v) \hat{v} \right) dx = 0, \quad \forall \hat{v} \in H^1(\Omega) \\ \Rightarrow & \quad \int_{\Omega} \beta \left( ((u_1 - g)^2 + (u_2 - g)^2) v + \frac{9\alpha}{c} v(1-v)(1-2v) \right) \hat{v} dx \\ & \quad + \int_{\Omega} c\alpha \nabla v \cdot \nabla \hat{v} dx = \int_{\Omega} \beta(u_2 - g)^2 \hat{v} dx, \quad \forall \hat{v} \in H^1(\Omega). \end{aligned} \quad (12)$$

Para efetuar a discretização baseada em um esquema de diferenças finitas precisamos extrair a forma forte da equação (12) (aplicando o Teorema do Divergente no segundo termo da equação e utilizando o fato da equação ser válida para qualquer  $\hat{v} \in H^1(\Omega)$ ), dada por

$$\begin{cases} \beta((u_1 - g)^2 + (u_2 - g)^2) v + \frac{9\alpha}{c} v(1-v)(1-2v) - c\alpha \Delta v = \beta(u_2 - g)^2 & \text{em } \Omega \\ \frac{\partial v}{\partial \mathbf{n}} = 0 & \text{em } \partial\Omega \end{cases}, \quad (13)$$

onde  $\mathbf{n}$  é o vetor unitário ortogonal da fronteira  $\partial\Omega$  de  $\Omega$ .

### 3.2 Solução numérica

Para facilitar o cálculo pelo método das diferenças finitas iremos adicionar a dependência artificial de  $v$  ao tempo. Desta forma, a aproximação para o mínimo de (10) será dada em um ponto estacionário de  $v$ , isto é, quando  $\frac{\partial v}{\partial t} = 0$ . Desta forma, temos a nossa equação final dada por

$$\begin{cases} \frac{\partial v}{\partial t} + \beta \gamma_1(u_1, u_2)v + \frac{\alpha}{c} \eta(v)(1-2v) - c\alpha \Delta v = \beta \gamma_2(u_2) & \text{em } \Omega \\ \frac{\partial v}{\partial \mathbf{n}} = 0 & \text{em } \partial\Omega \end{cases}, \quad (14)$$

onde  $\gamma_1(u_1, u_2) = (u_1 - g)^2 + (u_2 - g)^2$ ,  $\gamma_2(u_2) = (u_2 - g)^2$  e  $\eta(v) = 9v(1-v)$ .

Se tratando de um problema de evolução escolhemos um esquema de diferenças finitas baseado no método das direções alternadas [5], que além de ser uma discretização de ordem quadrática no tempo e no espaço, possui ordem de complexidade linear. Os termos explícitos são dados pelas funções  $\gamma_1$ ,  $\gamma_2$  e  $\eta$ .

$$\begin{cases} \frac{v_{i,j}^{n+\frac{1}{2}} - v_{i,j}^n}{\Delta t/2} - 2c\alpha \partial_x^- \partial_x^+ v_{i,j}^{n+\frac{1}{2}} + \left( -2\eta_{i,j}^n \frac{\alpha}{c} + \beta \gamma_1^n \right) v_{i,j}^{n+\frac{1}{2}} &= 2c\alpha \partial_y^- \partial_y^+ v_{i,j}^n \\ &+ \beta \gamma_2^n - \frac{\alpha}{c} \eta_{i,j}^n \\ \frac{v_{i,j}^{n+1} - v_{i,j}^{n+\frac{1}{2}}}{\Delta t/2} - 2c\alpha \partial_y^- \partial_y^+ v_{i,j}^{n+1} + \left( -2\eta_{i,j}^{n+\frac{1}{2}} \frac{\alpha}{c} + \beta \gamma_1^{n+\frac{1}{2}} \right) v_{i,j}^{n+1} &= 2c\alpha \partial_x^- \partial_x^+ v_{i,j}^{n+\frac{1}{2}} \\ &+ \beta \gamma_2^{n+\frac{1}{2}} - \frac{\alpha}{c} \eta_{i,j}^{n+\frac{1}{2}} \end{cases}, \quad (15)$$

onde  $\partial_x^-$  e  $\partial_x^+$  denotam operadores usuais de diferenças finitas atrasada e adiantada, respectivamente (idem para  $\partial_y^-$  e  $\partial_y^+$ ), sendo  $v_{i,j}^n \equiv v^n(i, j)$ . Para forçar a condição de contorno para  $v_n$  em  $\partial\Omega$  fazemos:  $v_{-1,j}^n = v_{0,j}^n$ ;  $v_{N+1,j}^n = v_{N,j}^n$ ;  $v_{i,-1}^n = v_{i,0}^n$ ;  $v_{i,M+1}^n = v_{i,M}^n$ , onde  $N$  e  $M$  são a quantidade de pixeis da imagem de entrada  $g$  nas direções  $x$  e  $y$  respectivamente, além disso, note que  $v^0$  deve ser fornecido para a inicialização do algoritmo. O critério de parada é computado pela norma do máximo via expressão:

$$\max_{0 \leq i \leq N; 0 \leq j \leq M} |v_{i,j}^{n+1} - v_{i,j}^n| < \varepsilon, \quad (16)$$

onde  $\varepsilon$  é uma tolerância previamente definida.

Observe que precisamos resolver  $M$  sistemas tridiagonais com  $N$  variáveis para a primeira expressão e  $N$  sistemas tridiagonais com  $M$  variáveis para a segunda, como cada sistema tridiagonal pode ser resolvido com ordem de complexidade linear através do *Algoritmo de Thomas* [7] a complexidade do método é de ordem  $O(N \cdot M)$  (linear) por iteração.

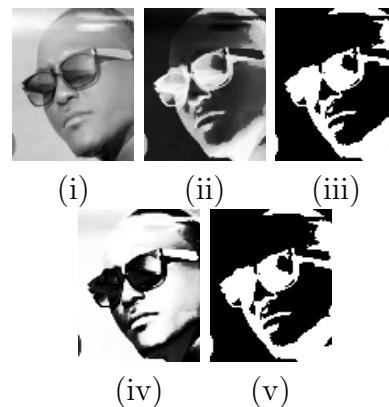
## 4 Experimentos Computacionais

Nesta sessão iremos efetuar experimentos e comparações com o método de Chan e Vese, mais precisamente, iremos considerar a segmentação via método de Chan e Vese como o resultado desejado. O motivo da escolha se deve ao fato de ser um método conhecido e popular na literatura que busca fazer a mesma simplificação deste trabalho para o modelo de Mumford e Shah com um termo adicional que pondera a área formada pela curva fechada  $C$ . Uma vez que o modelo de Ambrosio e Tortorelli é uma aproximação para o modelo de Mumford e Shah esperamos obter resultados semelhantes, porém, com uma expressão que requer menos custo computacional e permite o uso de outros métodos para aproximar a solução.

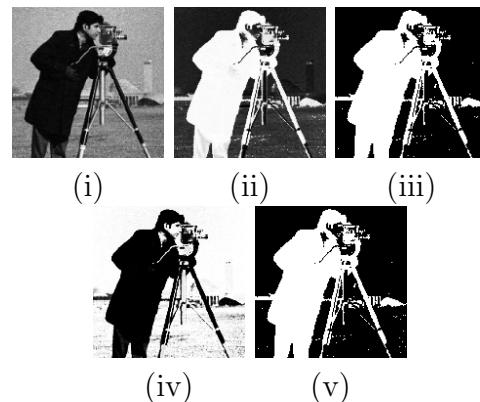
A seguir, nas figuras 1, 2, 3, 4, 5 e 6, ilustramos alguns experimentos efetuados com o método proposto utilizando os parâmetros  $\Delta t = 0.1$ ,  $c = 0.1$ ,  $\alpha = 0.5$ ,  $\beta = 500.0$  e tolerância  $\varepsilon = 10^{-1}$  na equação (15) em comparação aos resultados provindos do método de Chan e Vese. Como condição inicial, para cada imagem de entrada  $g$ , utilizamos a função

$$I_g(\mathbf{x}) = \begin{cases} 1, & \text{se } \mathbf{x} \in \Omega_g - \{R_g\} \\ 0, & \text{se } \mathbf{x} \in R_g \end{cases}, \quad (17)$$

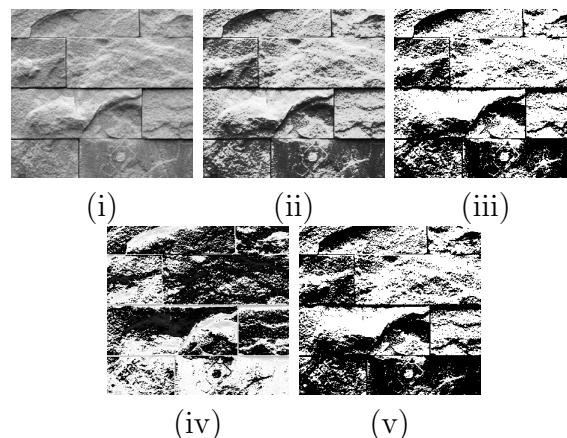
onde  $R_g \subset \Omega_g$  e por simplicidade escolhemos  $R_g$  como um retângulo centrado em  $\Omega_g$ .



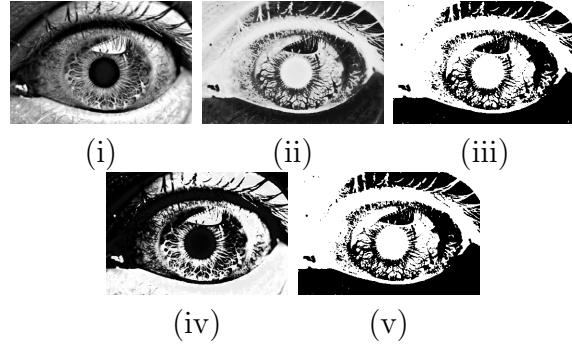
**Fig. 1:** (i) Imagem original. (ii) Segmentação via método de Chan e Vese. (iii) Resultado da binarização da imagem anterior com limiar  $v = 0$ . (iv) Segmentação via método proposto. (v) Resultado da binarização da imagem anterior com limiar  $v = 0.5$ .



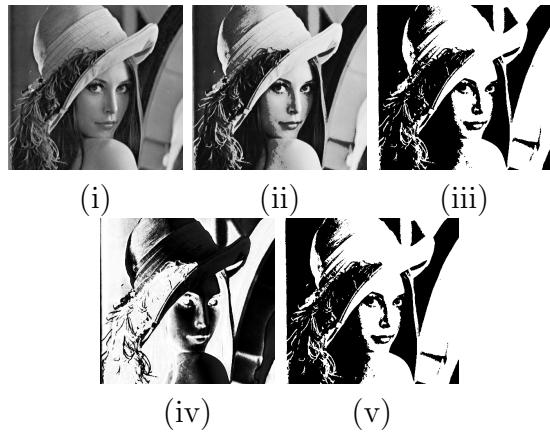
**Fig. 2:** (i) Imagem original. (ii) Segmentação via método de Chan e Vese. (iii) Resultado da binarização da imagem anterior com limiar  $v = 0$ . (iv) Segmentação via método proposto. (v) Resultado da binarização da imagem anterior com limiar  $v = 0.5$ .



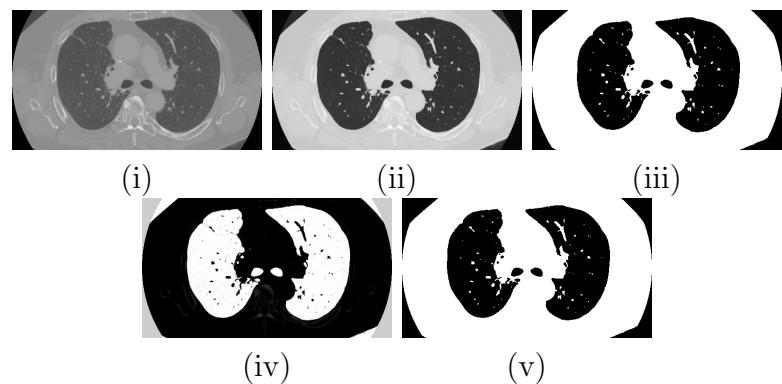
**Fig. 3:** (i) Imagem original. (ii) Segmentação via método de Chan e Vese. (iii) Resultado da binarização da imagem anterior com limiar  $v = 0$ . (iv) Segmentação via método proposto. (v) Resultado da binarização da imagem anterior com limiar  $v = 0.5$ .



**Fig. 4:** (i) Imagem original. (ii) Segmentação via método de Chan e Vese. (iii) Resultado da binarização da imagem anterior com limiar  $v = 0$ . (iv) Segmentação via método proposto. (v) Resultado da binarização da imagem anterior com limiar  $v = 0.5$ .



**Fig. 5:** (i) Imagem original. (ii) Segmentação via método de Chan e Vese. (iii) Resultado da binarização da imagem anterior com limiar  $v = 0$ . (iv) Segmentação via método proposto. (v) Resultado da binarização da imagem anterior com limiar  $v = 0.5$ .



**Fig. 6:** (i) Imagem original. (ii) Segmentação via método de Chan e Vese. (iii) Resultado da binarização da imagem anterior com limiar  $v = 0$ . (iv) Segmentação via método proposto. (v) Resultado da binarização da imagem anterior com limiar  $v = 0.5$ .

Note que apesar dos resultados distintos as imagens binárias referentes aos métodos estão muito próximas mostrando o potencial do método proposto. A seguir na tabela 1



apresentamos a similaridade entre as respectivas imagens binárias via métrica SSIM [8] com truncamento a partir de 4 dígitos decimais.

**Tabela 1:** TABELA DE VALORES PARA A MÉTRICA SSIM COMPUTADA PARA AS IMAGENS (III) E (V) DAS FIGURAS 1, 2, 3, 4, 5 E 6.

Figura	Métrica SSIM
Fig. 1.(iii)-(v)	0.9974
Fig. 2.(iii)-(v)	0.9929
Fig. 3.(iii)-(v)	0.7378
Fig. 4.(iii)-(v)	0.9575
Fig. 5.(iii)-(v)	0.9421
Fig. 6.(iii)-(v)	0.9895

Considerando que quanto mais próximo de 1.0 melhor o resultado, conseguimos alcançar precisões acima de 94% na maior parte dos casos mesmo com uma tolerância alta de 0.1 no critério de parada dado pela equação (16). Alcançamos 99% de precisão no melhor e 77% no pior caso. Em relação ao pior caso, a imagem possui textura e mesmo com a baixa precisão, em termos visuais, as diferenças não são tão perceptíveis como podemos notar na Figura 3.

Apesar da eficiência do método necessitamos de valores altos para o parâmetro  $\beta$  devido ao relaxamento. Nos experimentos utilizamos  $\beta = 500.0$  em todos os casos, pois, verificamos através de testes que valores altos para  $\beta$  tornam a convergência mais rápida. Porém é possível obter resultados semelhantes com  $\beta > 50$ , embora neste caso o método necessite de mais iterações para atingir o critério de parada.

## 5 CONCLUSÃO

Neste trabalho propomos uma simplificação para o modelo de Ambrosio e Tortorelli com o objetivo de separar uma dada imagem em duas regiões distintas. Desta maneira, supomos que a função  $u$  da forma variacional (1) era constante por partes e utilizamos termos de relaxação para evitar o uso das funções  $H_\epsilon$  e  $\delta_\epsilon$ , que além de caras computacionalmente impossibilitam métodos numéricos eficientes na aproximação do problema.

Mesmo utilizando relaxações que não garantem algumas propriedades desejadas, o método produziu resultados com precisão de até 99% comparados com a segmentação via método de Chan e Vese devido às características da função  $v$  presentes no modelo de Ambrosio e Tortorelli.

## 6 Agradecimentos

Agradecemos a CAPES pelo apoio financeiro que permitiu a realização deste trabalho.

## REFERÊNCIAS

- [1] R. Alicandro, A. Braides, and J. Shah. Free-discontinuity problems via functionals involving  $l^1$ -norm of the gradient and their approximations. *Interfaces and Free Boundaries - INTERFACE FREE BOUND*, 1:17–37, 1999.

- [2] L. Ambrosio and V. Tortorelli. Approximation of functionals Depending on Jumps by Elliptic Functionals via  $\Gamma$ -Convergence. *Communications on Pure and Applied Mathematics*, 43:999–1036, 1990.
- [3] T. Chan and L. Vese. An Active Contour Model without Edges. *Scale-Space Theories in Computer Vision*, 1682:141–151, 1999.
- [4] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 1992.
- [5] J. Douglas Jr. and H. H. Rachford Jr. On the Numerical Solution of Heat Conduction Problems in Two and Three Space Variables. *Transactions of the American Mathematical Society*, 82(2):421–43, 1956.
- [6] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problem. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.
- [7] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, New York, NY, USA, 3 edition, 2007.
- [8] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. on image processing*, 13(4):600–612, 2004.



# Otimização Biobjetivo de um Modelo para Controle de Qualidade de Processos Dinâmicos com Tamanho de Amostra Variável

João Gabriel Malaguti<sup>1</sup>, João Vitor de Lima<sup>1</sup> e Rafael de Souza Terra<sup>2</sup>

<sup>1</sup> Universidade Federal de Juiz de Fora/Departamento de Estatística, Juiz de Fora/MG, Brasil

<sup>2</sup> Universidade Federal de Juiz de Fora/Departamento de Ciência da Computação, Juiz de Fora/MG, Brasil

---

## Resumo

Na literatura de controle de qualidade, geralmente se estuda o caso em que o sistema controlado tem valor fixo para probabilidade de choque aleatório. No entanto, quando se leva em conta as alterações que um sistema sofre ao longo do tempo (por exemplo reposição de peças do maquinário) este deveria ser considerado dinâmico com a probabilidade de choque decrescente. Decorre então que o tamanho de amostra também pode ser variável. Criamos para o tamanho da amostra um modelo com base em inspeção de qualidade e equações diferenciais a ser otimizado minimizando dois objetivos: tamanho médio de amostra e tempo esperado até o sinal. Utilizando busca em grade e simulação de Monte Carlo chegamos a um conjunto de possíveis equações ótimas melhores que o caso com amostras fixas.

**Palavras-chave:** Controle de qualidade, Amostragem, Otimização

---

## 1 INTRODUÇÃO

Controle de qualidade de um processo industrial é definido por uma variável a qual se deseja manter entre determinados limites chamados limites de controle (LIC e LSC, sendo respectivamente o limite inferior de controle e o limite superior de controle) ao longo do tempo (Costa et al., 2005; Cano et al., 2015). Esta variável pode ser uma medida de centralidade, dispersão ou até número de não-conformidades. Utilizaremos como exemplo um processo interessado em controlar a média. O processo pode se encontrar controlado, quando a média real do processo é a média desejada ou fora de controle quando a média real é diferente da desejada. A saída de controle pode acontecer por um choque aleatório,

com probabilidade  $\lambda$ . O número de amostras necessárias até que a média estimada esteja fora dos intervalos de controle dado que o processo esteja descontrolado é denominado tempo esperado até o sinal (Costa et al., 2005). Quando a estimativa se encontra fora do intervalo, o processo industrial é de alguma forma regenerado, ou seja, retorna à média desejada por alguma intervenção (por exemplo, uma limpeza no equipamento, um retreino do manejo ou ainda a troca de peças defeituosas do maquinário).

Consideramos que os processos industriais são dinâmicos, isto é, sofrem variações em seu estado ao decorrer do tempo (Katok and Hasselblatt, 1995). Esses processos são geralmente modelados por meio de equações diferenciais, podendo ser ordinárias ou parciais. Neste estudo, a variável que sofre alteração direta do tempo é a probabilidade de choque aleatório, isto é,  $\lambda$  é uma função de  $t$ , sendo resultado da equação diferencial ordinária vista em Eq. (1)

$$\begin{aligned}\lambda' &= -\alpha * \lambda + \gamma \\ \lambda &= \frac{\gamma}{\alpha} + C_1 * e^{-\alpha*t}\end{aligned}\tag{1}$$

onde  $\alpha$  é o coeficiente de proporcionalidade,  $\gamma$  é um fator somado tal que quando  $t$  tende a infinito  $\lambda$  tende a um valor diferente de zero e  $C_1$  a constante resultante da solução da equação diferencial ordinária.

Baseando-se nos tipos de amostragem presentes na teoria de inspeção de qualidade (Costa et al., 2005), criamos um modelo para o tamanho de amostra de modo que existam dois tipos de amostragem para o processo dinâmico: amostragem normal e amostragem atenuada (com tamanho de amostra menor). Quando o número de amostras consecutivas dentro do intervalo ( $N$ ) é maior do que dez o sistema passa para a amostragem atenuada, com o seu tamanho de amostra ( $n$ ) sendo função de  $N$ , e consequentemente função de  $\lambda$  (Eq. (2))

$$n = \begin{cases} n_0, & \text{para amostragem normal} \\ \frac{S}{K} + C_2 * e^{-K*(N-11)}, & \text{(arredondado) para amostragem atenuada} \end{cases}\tag{2}$$

na qual  $K$  é o coeficiente de proporcionalidade,  $S$  é um fator somado tal que quando  $N$  tende a infinito  $n$  tende a um valor diferente de zero e  $C_2$  a constante resultante da solução da equação diferencial ordinária.

Os valores para  $C_2$  e  $S/K$  são definidos a partir dos valores escolhidos para  $n$  quando  $N$  é igual a onze (primeiro valor para amostragem atenuada) e quando  $N$  tende a infinito, seguindo a Eq. (3).

$$\begin{aligned}n_{11} &= \frac{S}{K} + C_2 \\ n_\infty &= \frac{S}{K}\end{aligned}\tag{3}$$



Temos como objetivo otimizar a Eq. (2) em relação ao coeficiente  $K$  de modo que o tamanho médio das amostras ao longo do processo e o tempo esperado até o sinal (TES) sejam mínimos.

## 2 METODOLOGIA

### 2.1 Busca em grade

A busca em grade ou varredura é um processo de busca exaustiva por um subconjunto previamente especificado, sendo bastante utilizada para o *tuning* de hiper-parâmetros em aprendizado de máquinas (Huang et al., 2007; Bergstra and Bengio, 2012; Larochelle et al., 2007). Os valores definidos pela busca em grade são avaliados na função a qual se quer otimizar. O método permite uma paralelização simples apesar de ser ingênuo, dependendo dos valores determinados pelos pesquisadores.

Mas Hsu et al. (2003) apontam que uma busca em grade com poucos valores pode ser seguida com uma busca mais densa, buscando melhorar a estimativa otimizada. Dependendo o número de valores esse método também é computacionalmente mais leve. Especificamente para a linguagem de programação utilizada, a linguagem R (R Core Team, 2019), temos que com uma grade muito densa, o número de avaliações se torna impraticável (Cortez, 2014).

### 2.2 Otimização biobjetivo

Um problema em que a função nos retorna dois resultados é definido como um problema biobjetivo. Geralmente esses objetivos são conflitantes e não temos um resultado ótimo global, ou seja, não existe um único valor que otimize essa função e sim, um conjunto de soluções ótimas (denominado curva de Pareto).

Para encontrar o conjunto de valores que otimizam a função, que nesse estudo são os valores que a minimizam, utilizamos de um método denominado Eficiência de Pareto (Miettinen, 2012). Para que um valor pertença à fronteira de Pareto (conjunto ótimo) ele deverá ser não-dominado por nenhum outro valor, onde a dominância é definida pela Eq. (4).

$$\begin{cases} f_i(x) \leq f_i(y) | i = 1, 2, \text{ temos que } x \text{ domina } y \\ f_i(x) \geq f_i(y) | i = 1, 2, \text{ temos que } x \text{ é dominado por } y \\ f_i(x) \not\leq f_i(y) \text{ e } f_j(x) \leq f_j(y) | i \neq j, \text{ temos que } x \text{ é indiferente a } y \end{cases} \quad (4)$$

Neste estudo, as funções  $f_1$  e  $f_2$  são as funções responsáveis pelo cálculo do tamanho médio de amostra e tempo esperado até o sinal, que não possuem equações formais e serão consideradas como caixas-pretas, sendo seus resultados obtidos por métodos computacionais.

### 2.3 Estudo de simulação

Estudos de simulação Monte Carlo são métodos estatísticos computacionais baseados na geração de diferentes amostras independentes para se obter respostas aproximadas para problemas estocásticos, sendo utilizado desde o século XX (Lewis and Orav, 1989). A precisão deste método está relacionada ao número de amostras geradas (iterações ou

rodadas do estudo), seguindo a lei dos grandes números que descreve que, com um grande número de ensaios, o valor esperado dos resultados observados converge para a média verdadeira deles.

Para que possamos realizar simulações Monte Carlo, primeiro precisamos fazer certas definições, como os valores para  $\alpha$ ,  $C_1$ ,  $\gamma$  e qual será a variável controlada pelo processo industrial. Para o processo foram escolhidos os valores vistos na Eq. (5) de modo que  $\lambda$  decaia suavemente entre 0, 1 (para  $t$  igual a zero) e 0,0001 (para  $t$  tendendo a infinito).

$$\begin{cases} C_1 = 0,09999 \\ \gamma = 5 * 10^{-8} \\ \alpha = 0,00005 \end{cases} \quad (5)$$

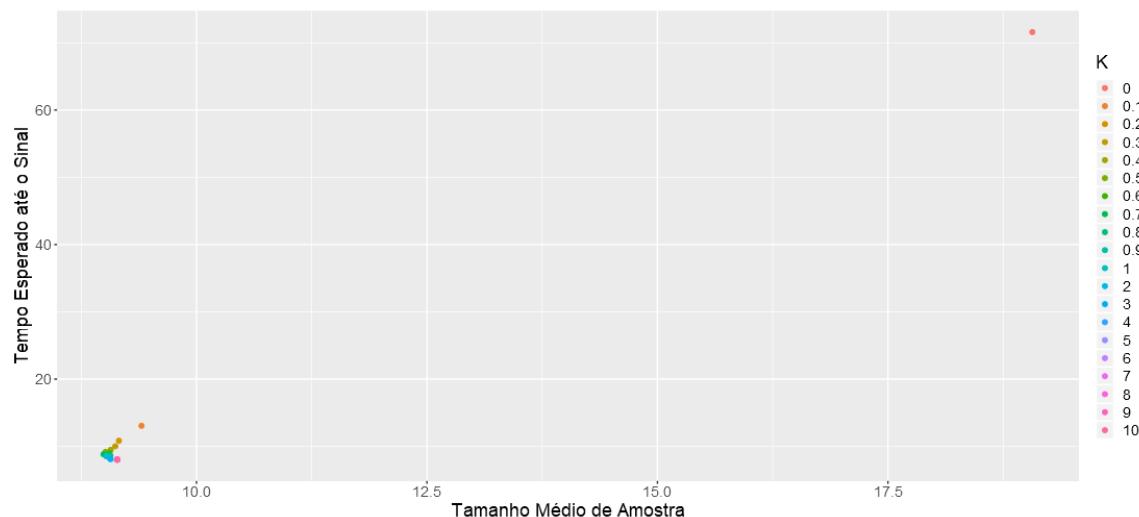
Definimos como variável de interesse a média e seguimos o exemplo apresentado por Costa et al. (2005) tendo então média original igual a mil, e intervalo de controle [997, 1003]. Analisamos três diferentes cenários, sendo cada um caracterizado pelo valor do deslocamento da média pertencentes ao conjunto {1002, 1003, 1004}. Esse deslocamento é causado pelo choque aleatório a qual o processo pode sofrer. Foram escolhidos também os valores para  $n_0$ ,  $n_{11}$  e  $n_\infty$ , respectivamente iguais a vinte, dezenove e dois. Feitas as devidas substituições, a Eq. (2) se encontra atualizada na Eq. (6).

$$n = \begin{cases} 20, \text{ para amostragem normal} \\ 2 + 17 * e^{-K*(N-11)}, \text{ (arredondado) para amostragem atenuada} \end{cases} \quad (6)$$

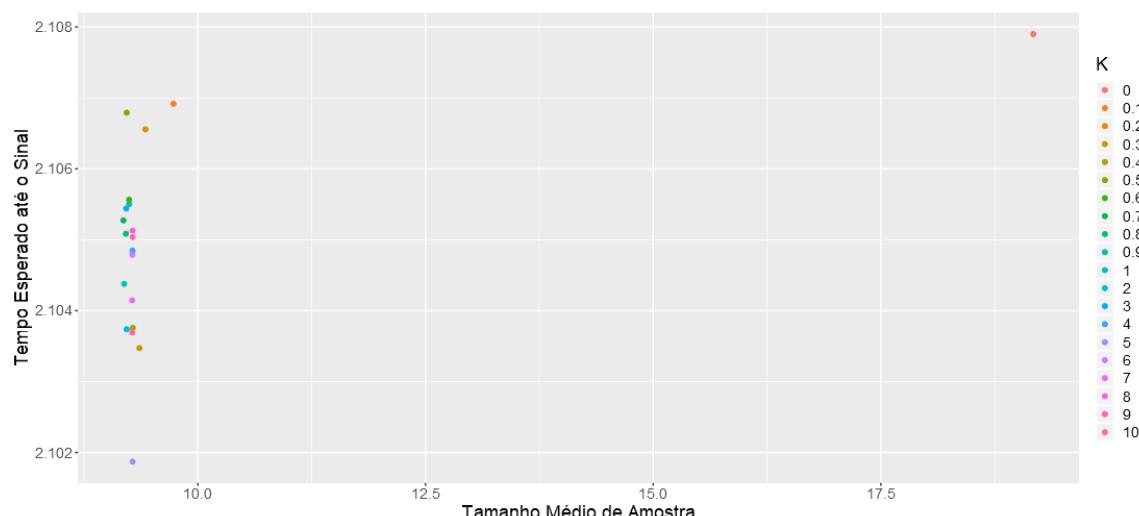
Ainda definimos também o número de iterações Monte Carlo ( $A = 10.000$ ) e os valores de  $K$  a serem testados, sendo os valores pertencentes a união dos conjuntos  $\{0; 0,1; \dots; 0,9; 1\}$  e  $\{2; 3; \dots; 9; 10\}$  visto que para valores muito maiores,  $n$  converge rapidamente para 2.

### 3 RESULTADOS

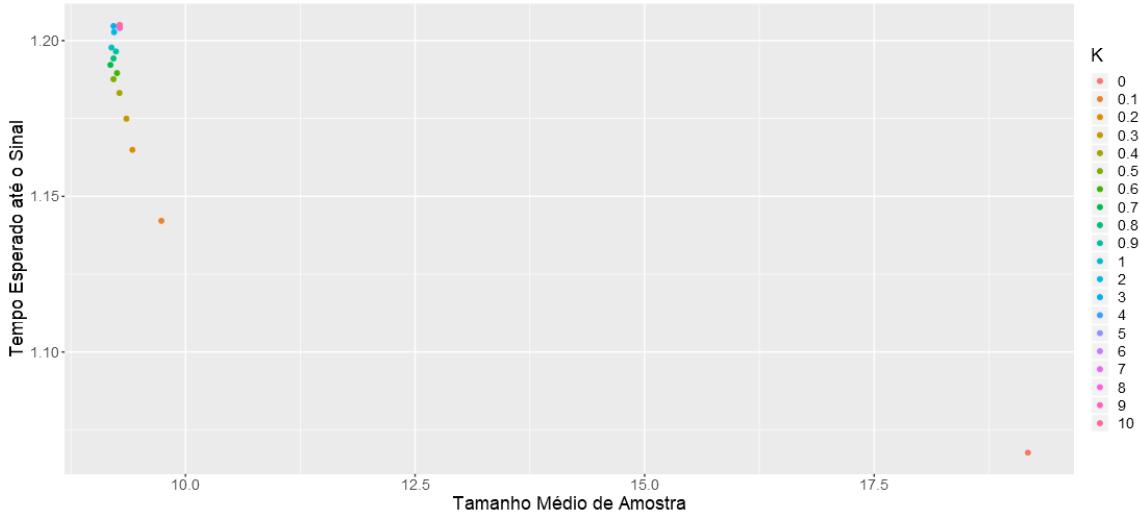
Podemos representar os resultados de maneira visual a partir de um diagrama de dispersão de tamanho médio de amostra (TMA) por tempo esperado até o sinal onde cada ponto representa um valor de K para um determinado deslocamento, como nas imagens Fig. 1, Fig. 2 e Fig. 3.



**Fig. 1:** TMA por TES para Deslocamento 1002



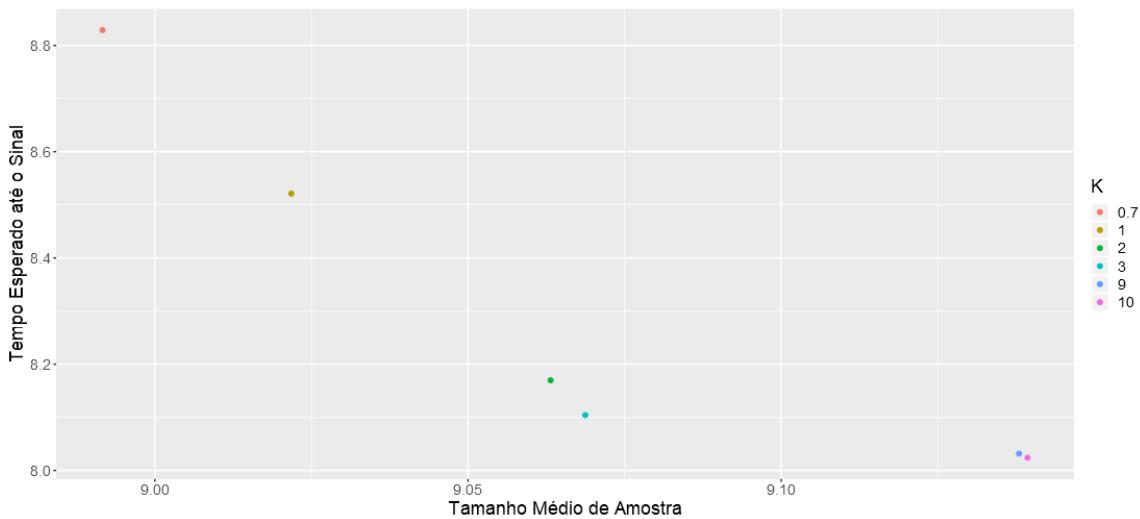
**Fig. 2:** TMA por TES para Deslocamento 1003



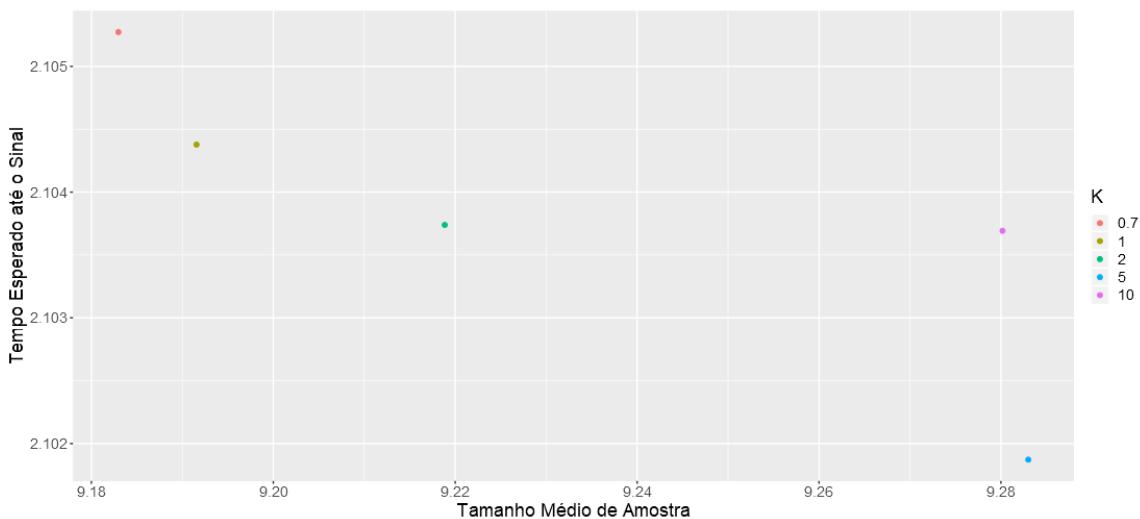
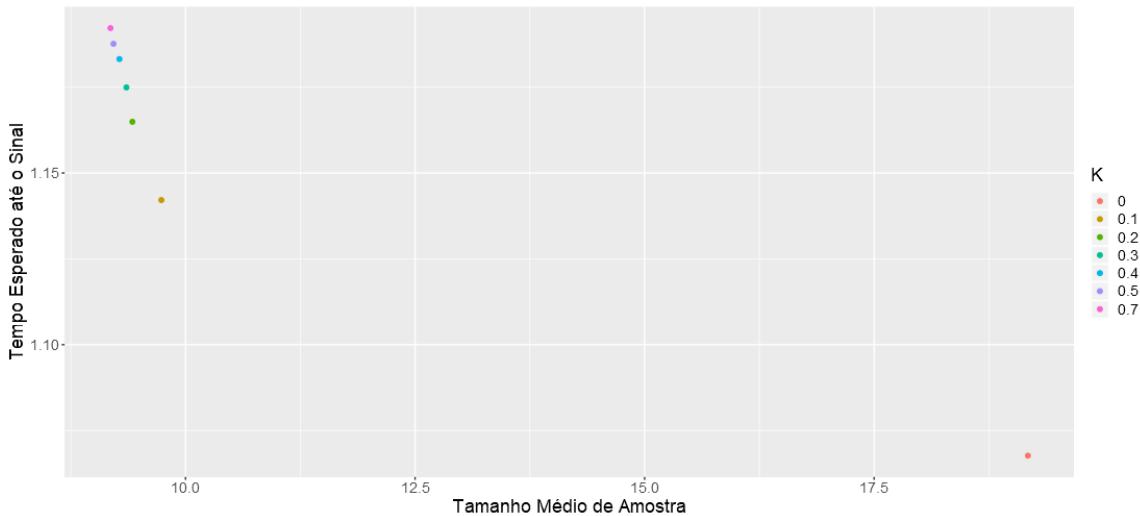
**Fig. 3:** TMA por TES para Deslocamento 1004

Notamos, em um primeiro momento, que em todos os diagramas existe um ponto muito distante dos outros, sendo este  $K=0$ , que implica numa amostra atenuada com valor fixo em 19 ao invés de decrescente conforme o número de amostras consecutivas dentro do intervalo ( $N$ ). Podemos perceber também que cada deslocamento causa um padrão diferente para os pontos: na Fig. 1 eles se encontram agrupados e nas outras eles estão mais dispersos, ambos no canto esquerdo.

Computando as curvas de Pareto para cada gráfico, visíveis nas Fig. 4, Fig. 5 e Fig. 6, temos então nossos conjuntos de valores ótimos:  $\{0,7; 1; 2; 3; 9; 10\}$  para o deslocamento 1002,  $\{0,7; 1; 2; 5; 10\}$  para o deslocamento 1003 e  $\{0; 0,1; 0,2; 0,3; 0,4; 0,5; 0,7\}$  para o deslocamento 1004. Para o caso com amostras fixas, temos que nenhum destes valores pertenceria à curva de Pareto, sendo dominados por um valor (para o deslocamento igual a 1004) e por todos os valores (para os deslocamentos iguais a 1002 e 1003).



**Fig. 4:** Curva de Pareto para Deslocamento 1002

**Fig. 5:** Curva de Pareto para Deslocamento 1003**Fig. 6:** Curva de Pareto para Deslocamento 1004

## 4 CONSIDERAÇÕES FINAIS

Apesar de não se poder determinar um valor ótimo para otimização biobjetivo, seria negligente não mencionar que apenas um valor dentre o subconjunto pesquisado pertence a fronteira de Pareto para os três cenários analisados ( $K = 0, 7$ ) e existem outros valores que pertencem à duas das três fronteiras encontradas: {1, 2, 10}.

O processo analisado pode ser ainda mais complexo, permitindo a alteração não só da média, mas também da variância, além de mudanças nos valores utilizados no modelo original para  $\lambda$  estudando outros formatos de decaimento.

Futuramente, se pretende otimizar a função utilizando algoritmo genético, assim como alguns bioinspirados, como *Social Spider Algorithm* ou ainda *Bat Algorithm*. Recomendamos então, caso se queira a adoção do modelo para tamanho de amostra antes destes

trabalhos futuros, a utilização do valor 0,7 para  $K$ .

## REFERÊNCIAS

- [1] Bergstra, J. and Bengio, Y. (2012). “Random search for hyper-parameter optimization”. *Journal of Machine Learning Research*, 13(Feb):281–305.
- [2] Cano, E. L., Moguerza, J. M., and Corcoba, M. P. (2015). “Quality control with r”. *An ISO*.
- [3] Cortez, P. (2014). *Modern optimization with R*. Springer.
- [4] Costa, A. F. B., Epprecht, E. K., and Carpinetti, L. C. R. (2005). *Controle estatístico de qualidade*. Atlas São Paulo.
- [5] Hsu, C.-W., Chang, C.-C., Lin, C.-J., et al. (2003). “A practical guide to support vector classification”.
- [6] Huang, C.-M., Lee, Y.-J., Lin, D. K., and Huang, S.-Y. (2007). “Model selection for support vector machines via uniform design”. *Computational Statistics & Data Analysis*, 52(1):335–346.
- [7] Katok, A. and Hasselblatt, B. (1995). *Introduction to the modern theory of dynamical systems*, volume 54. Cambridge university press.
- [8] Larochelle, H., Erhan, D., Courville, A., Bergstra, J., and Bengio, Y. (2007). “An empirical evaluation of deep architectures on problems with many factors of variation”. ACM, pages 473–480.
- [9] Lewis, P. and Orav, E. (1989). “Simulation methodology for statisticians, operations analysts, and engineers: vol. 1, wadsworth publ”. Co., Belmont, CA.
- [10] Miettinen, K. (2012). *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media.
- [11] R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.



# Análise de Sensibilidade Topológica Aplicada ao Processo de Fraturamento Hidráulico

Jorge Morvan Marotte Luz Filho<sup>1</sup> and Antonio André Novotny<sup>1</sup>

<sup>1</sup> Laboratório Nacional de Computação Científica - LNCC, Petrópolis/RJ, Brazil

---

## Resumo

A análise de sensibilidade topológica fornece uma função escalar, denominada derivada topológica, que mede a sensibilidade de um dado funcional de forma em relação a uma perturbação singular infinitesimal no domínio. Tal conceito vem sendo amplamente utilizado nas mais diversas aplicações e tem se demonstrado de grande relevância em problemas de engenharia. Neste trabalho é apresentada a derivada topológica no contexto do fraturamento hidráulico. Para tanto, primeiramente é introduzido o modelo de fraturamento aqui considerado, bem como a derivada topológica a este associada. Por fim, algumas simulações numéricas ilustram o efeito da presença da tensão *in situ* durante a propagação da falha e a aplicabilidade da metodologia aqui proposta.

**Palavras-chave:** Fraturamento hidráulico, Análise de sensibilidade topológica, Derivada topológica

---

## 1 INTRODUÇÃO

O fraturamento hidráulico, também conhecido como *fracking*, é um processo industrial que visa aumentar de forma significativa a superfície de produção de um meio poroso onde gás pode ser coletado através de uma fratura pré-existente. A fratura é então propagada através do bombeamento em alta pressão de um fluido de fraturamento, permitindo assim que o gás aprisionado seja coletado na superfície [1]. Por questões econômicas e pelo seu impacto ambiental, o fraturamento hidráulico tem sido ultimamente objeto de estudo de diversos trabalhos.

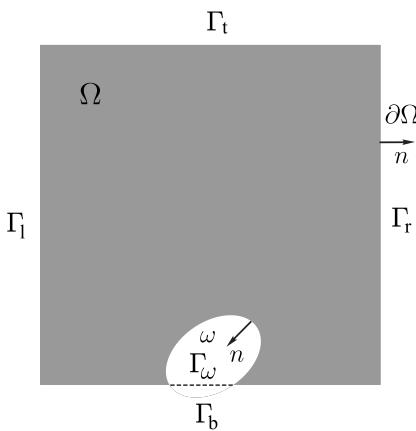
Nesse contexto apresenta-se uma extensão do modelo baseado no conceito da derivada topológica proposto por Xavier [2] ao incorporar o estado de tensões iniciais e o efeito da inércia do fluido durante o fenômeno de propagação da falha. O modelo aqui

apresentado resulta da aplicação dos modelos de Francfort-Marigo [3] e de Biot [4]. Ademais, considera-se uma idealização bidimensional em que o processo de fraturamento é ativado através de um campo de pressão não-constante distribuído por todo o domínio. A análise de sensibilidade é então realizada levando-se em consideração a incorporação das hipóteses acima. De posse da derivada topológica associada ao modelo aqui proposto, um algoritmo de otimização topológica é desenvolvido para simular os processos de nucleação e propagação.

O trabalho se organiza da seguinte forma: O modelo de fraturamento hidromecânico proposto é apresentado na Seção 2. A fórmula da derivada topológica associada, bem como as hipóteses utilizadas na análise de sensibilidade topológica são apresentadas na Seção 3. Na Seção 4 o algoritmo de otimização topológica desenvolvido é ilustrado na forma de pseudocódigo e comentado em detalhes, ao passo que um conjunto de experimentos numéricos são conduzidos na Seção 5. Finalmente, algumas considerações finais são apresentadas na Seção 6.

## 2 MODELO DE FRATURAMENTO HIDROMECÂNICO

Neste trabalho propõe-se uma extensão do modelo apresentado por Xavier [2] que é resultado da adaptação do modelo de Francfort-Marigo [3], bem como da aplicação da teoria de Biot [4]. Toma-se o reservatório como sendo composto somente pelo poço de pressurização e pela região a ser fraturada. Ainda, um único bloco do reservatório é assumido como domínio de referência para representar a rede de fraturas. Uma de suas limitações é não considerar o estado de tensões iniciais (*in situ*) existente devido à presença da camada rochosa acima do bloco de referência. Visando eliminar esta limitação, o modelo aqui proposto é definido no seguinte cenário: considera-se uma matriz porosa saturada (linearmente elástica e isotrópica) submetida a um fluxo de propante em regime quase estático. Portanto, o problema não é dependente do tempo, sendo assim introduzido um pseudotempo  $t_i$  referente aos incrementos realizados na pressão. A matriz é representada por um domínio geométrico  $\Omega \subset \mathbb{R}^2$ , contendo o subdomínio  $\omega \subset \Omega$  que representa a falha geológica.



**Fig. 1:** Bloco saturado contendo falha geológica.

A caracterização da região ocupada pelo dano é realizada através da introdução do



parâmetro  $\rho$ , dado por

$$\rho = \rho(x) := \begin{cases} 1, & \text{se } x \in \Omega \setminus \bar{\omega}, \\ \rho_0, & \text{se } x \in \omega, \end{cases} \quad (1)$$

onde  $0 < \rho_0 \ll 1$ . A região  $\Omega \setminus \bar{\omega}$  representa a área saudável da matriz porosa, enquanto que  $\omega$  representa a falha geológica.

De posse dessa configuração, tem-se o funcional de forma  $\mathcal{F}_\omega(u)$  a ser minimizado com relação a  $\omega$  para cada passo de tempo  $t_i$ ,

$$\mathcal{F}_\omega(u) = \mathcal{J}(u) + \kappa|\omega|, \quad (2)$$

onde  $\kappa$  é um parâmetro de liberação de energia e  $\mathcal{J}(u)$  é a energia potencial total, dada por

$$\mathcal{J}(u) = \frac{1}{2} \int_{\Omega} \sigma(u) \cdot \nabla u^s - \int_{\Omega} \alpha p \operatorname{div}(u) + \int_{\Omega} \sigma_0 \cdot \nabla u^s, \quad (3)$$

onde  $u$  representa o campo de deslocamentos e  $\sigma_0$  o estado de tensões iniciais. Além disso,  $u$  é solução do seguinte problema variacional: Encontrar  $u \in \mathcal{U}$ , tal que

$$\int_{\Omega} \sigma(u) \cdot \nabla \eta^s = \int_{\Omega} \alpha p \operatorname{div}(\eta) - \int_{\Omega} \sigma_0 \cdot \nabla \eta^s, \quad \forall \eta \in \mathcal{V}, \quad (4)$$

onde  $\alpha$  é conhecido como o coeficiente de Biot e  $p = p(x)$  é a pressão atuante na matriz porosa, solução do seguinte problema variacional: Encontrar  $p \in \mathcal{P}$  tal que

$$\int_{\Omega} k \nabla p \cdot \nabla \varphi = 0, \quad \forall \varphi \in \mathcal{Q}, \quad (5)$$

onde  $k$  é o coeficiente de permeabilidade dado pelo seguinte

$$k = k(x) := \begin{cases} k_r, & \text{se } x \in \Omega \setminus \bar{\omega}, \\ k_f, & \text{se } x \in \omega, \end{cases} \quad (6)$$

com  $k_r \ll k_f$ . O tensor tensão,  $\sigma(\varphi)$  é definido da seguinte maneira

$$\sigma(\varphi) = \rho \mathbb{C} \nabla \varphi^s. \quad (7)$$

Uma vez que a matriz porosa é considerada isotrópica, então o tensor de elasticidade  $\mathbb{C}$  pode ser escrito em termos dos coeficientes de Lamé,  $\mu$  e  $\lambda$ , da seguinte forma

$$\mathbb{C} = 2\mu \mathbb{I} + \lambda(\mathbb{I} \otimes \mathbb{I}), \quad (8)$$

onde  $\mathbb{I}$  e  $\mathbb{I}$  são os tensores de quarta e segunda ordem, respectivamente. Ademais, para o estado plano de deformações os coeficientes de Lamé são dados por

$$\mu = \frac{E}{2(1+\nu)}, \quad \lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}, \quad (9)$$

onde  $E$  e  $\nu$  representam o módulo de elasticidade e o coeficiente de Poisson, respectivamente. O conjunto  $\mathcal{U}$  e o espaço  $\mathcal{V}$  são definidos da seguinte forma

$$\mathcal{V} = \mathcal{U} := \left\{ \varphi \in H^1(\Omega; \mathbb{R}^2) : \varphi|_{\Gamma_b} = 0, \varphi \cdot n|_{\Gamma_r \cup \Gamma_l} = 0 \right\}, \quad (10)$$

e o conjunto  $\mathcal{P}$  e o espaço  $\mathcal{Q}$  são dados por

$$\mathcal{P} := \left\{ \varphi \in H^1(\Omega) : \varphi|_{\Gamma_t} = 0, \varphi|_{\Gamma_w} = p_i \right\}, \quad \mathcal{Q} := \left\{ \varphi \in H^1(\Omega) : \varphi|_{\Gamma_t} = 0, \varphi|_{\Gamma_w} = 0 \right\}, \quad (11)$$

em que  $p_i$  é a pressão prescrita sobre  $\Gamma_w$  no instante  $t_i$ , definida por

$$p_i = p_{i-1} + \Delta p_i, \quad (12)$$

e  $\Delta p_i$  é o incremento da pressão. Assim sendo, o problema de minimização fica definido da seguinte forma: para cada incremento de tempo  $t_i$

$$\underset{\omega \subset \Omega}{\text{Minimizar}} \mathcal{F}_\omega(u), \text{ sujeito à (4).} \quad (13)$$

Por fim, o parâmetro de liberação de energia  $\kappa$ , é substituído pelo parâmetro modificado  $\kappa_\delta$ , que é definido através de uma nova propriedade do material  $\kappa_s$  e um fator de escala  $\delta$  dado pela largura da falha inicial. Portanto, tem-se

$$\kappa = \kappa_\delta := \frac{\kappa_s}{\delta}. \quad (14)$$

### 3 ANÁLISE DE SENSIBILIDADE TOPOLOGICA

Esta seção tem por objetivo introduzir o conceito da derivada topológica. Para tanto, considera-se um domínio aberto e limitado  $\Omega \subset \mathbb{R}^d$ , com  $d \geq 2$  sujeito a uma perturbação confinada numa pequena região  $B_\varepsilon(\hat{x})$  de tamanho  $\varepsilon$  centrada em um ponto arbitrário  $\hat{x}$ . Introduz-se então a função característica  $x \mapsto \chi(x)$ ,  $x \in \mathbb{R}^d$ , associada ao domínio original, denotada por  $\chi := \mathbf{1}_\omega$ , tal que

$$|\Omega| = \int_{\mathbb{R}^d} \chi, \quad (15)$$

onde  $|\Omega|$  é a medida de Lebesgue de  $\Omega$ . Agora, é definida a função característica associada ao problema topologicamente perturbado da forma  $x \mapsto \chi_\varepsilon(\hat{x}; x)$ ,  $x \in \mathbb{R}^d$ . Para o caso de uma perfuração, por exemplo,  $\chi_\varepsilon(\hat{x}) := \mathbf{1}_\Omega - \mathbf{1}_{B_\varepsilon(\hat{x})}$  e o domínio perturbado é obtido como  $\Omega_\varepsilon = \Omega \setminus \overline{B_\varepsilon}$ . Finalmente, assume-se que um dado funcional de forma  $\psi(\chi_\varepsilon(\hat{x}))$ , associado ao domínio topologicamente perturbado, admite uma expansão da forma

$$\psi(\chi_\varepsilon(\hat{x})) = \psi(\chi) + f(\varepsilon)D_T\psi(\hat{x}) + \mathcal{R}(\varepsilon), \quad (16)$$

onde  $\psi(\chi)$  é o funcional de forma associado ao domínio original (não perturbado),  $f(\varepsilon)$  é uma função de correção de primeira ordem positiva de  $\psi(\chi)$ , tal que  $f(\varepsilon) \rightarrow 0$  quando  $\varepsilon \rightarrow 0$ , e  $\mathcal{R}(\varepsilon)$  é o termo residual, ou seja,  $\mathcal{R}(\varepsilon)/f(\varepsilon) \rightarrow 0$  quando  $\varepsilon \rightarrow 0$ . A função



$\hat{x} \mapsto D_T\psi(\hat{x})$  é chamada *derivada topológica* de  $\psi$  em  $\hat{x}$ . Portanto, o produto  $f(\varepsilon)D_T\psi(\hat{x})$  representa uma correção de primeira ordem sobre  $\psi(\chi)$  para aproximar  $\psi(\chi_\varepsilon(\hat{x}))$ .

Neste trabalho, considera-se que somente as propriedades mecânicas são sensíveis à nucleação da perturbação topológica  $B_\varepsilon(\hat{x})$ , diferentemente do que é realizado por [2], onde a permeabilidade também é sensível à perturbação topológica. A escolha de não perturbar a permeabilidade traz consigo a introdução da hipótese de que a propagação da fratura ocorre de forma instantânea, não sendo o fluido sensível a este fenômeno. Além disso, o estado de tensões iniciais  $\sigma_0$  também não é considerado sensível à perturbação topológica.

De posse do exposto acima, a derivada topológica do funcional de forma dado por (2), com respeito a nucleação de uma pequena inclusão  $B_\varepsilon(\hat{x})$  com propriedades mecânicas distintas da matriz, é dada por

$$D_T\mathcal{F}_\omega(x) = D_T\mathcal{J}(x) + \kappa_\delta D_T|\omega| \quad \forall x \in \Omega, \quad (17)$$

onde os termos  $\kappa_\delta D_T|\omega|$  e  $D_T\mathcal{J}(x)$  são, respectivamente, dados por

$$\kappa_\delta D_T|\omega|(x) = \begin{cases} +\kappa_\delta, & \text{se } x \in \Omega \setminus \overline{\omega}, \\ -\kappa_\delta, & \text{se } x \in \omega, \end{cases} \quad (18)$$

$$\begin{aligned} D_T\mathcal{J}(\hat{x}) &= \mathbb{P}_\gamma \sigma(u)(\hat{x}) \cdot \nabla u^s(\hat{x}) \\ &\quad + (1 - \gamma^\alpha) \frac{1+a}{1+a\gamma} \alpha p(\hat{x}) \operatorname{div}(u)(\hat{x}) - \frac{(1 - \gamma^\alpha)^2}{2\rho\mu(1+a\gamma)} \alpha^2 p(\hat{x})^2, \end{aligned} \quad (19)$$

na qual  $\mathbb{P}_\gamma$ , denominado tensor de polarização, é definido como

$$\mathbb{P}_\gamma = -\frac{1}{2} \frac{1-\gamma}{1+b\gamma} \left( (1+b)\mathbb{I} + \frac{1}{2}(a-b) \frac{1-\gamma}{1+a\gamma} \mathbf{I} \otimes \mathbf{I} \right), \quad (20)$$

e os coeficientes  $a$  e  $b$  são dados por

$$a = \frac{\lambda + \mu}{\mu} \quad \text{e} \quad b = \frac{\lambda + 3\mu}{\lambda + \mu}. \quad (21)$$

Maiores detalhes acerca da dedução da fórmula da derivada topológica aqui apresentada podem ser encontrados na seção 4.2.1 de [2], onde é realizada a análise assintótica considerando contrastes aplicados somente às propriedades mecânicas. Apesar de não considerar o estado de tensões iniciais em sua análise, a fórmula da derivada topológica obtida por [2] coincide com a fórmula aqui apresentada uma vez que este não é sensível à perturbação topológica.

## 4 ALGORITMO DE OTIMIZAÇÃO

O algoritmo de otimização aqui apresentado baseia-se no fato de que é possível reduzir o valor do funcional associado através da introdução de uma inclusão no interior da região  $\omega^*$ , onde a derivada topológica é negativa. Ou seja,

$$\omega^* := \{x \in \Omega : D_T\mathcal{F}_\omega(x) < 0\}. \quad (22)$$

O algoritmo é então projetado a fim de nuclear inclusões com tamanho compatível da região previamente danificada através da introdução de um parâmetro  $\beta \in (0, 1)$  de modo que  $\beta = 0$  representa a nucleação de uma inclusão somente no ponto onde a derivada topológica é mais negativa e  $\beta = 1$  em toda a região onde esta é negativa. Portanto, é definida a seguinte quantidade:

$$D_T\mathcal{F}_\omega^* := \min_{x \in \omega} D_T\mathcal{F}_\omega(x), \quad (23)$$

que permite estabelecer a inclusão a ser nucleada  $\omega^\beta \subset \omega^*$  como

$$\omega^\beta := \{x \in \omega^* : D_T\mathcal{F}_\omega(x) \leq (1 - \beta)D_T\mathcal{F}_\omega^*\}, \quad (24)$$

tal que  $\beta \in (0, 1)$  é escolhido de modo que  $|\omega^\beta| \approx \pi\delta^2/4$  (e  $|\omega^\beta| \leq \pi\delta^2/4$ ). Dessa forma o tamanho da inclusão fica então relacionado à largura  $\delta$  do dano inicial.

**Algoritmo 1:** Algoritmo de evolução do dano.

---

<b>Entrada:</b> $\Omega, \omega, \delta, N, \bar{p}_0, \Delta\bar{p}_i$	
<b>Saída :</b> Topologia ótima $\omega^*$	

```

1 para  $i = 1 : N$  faça
2   resolver o problema de pressão (5);
3   resolver o problema de elasticidade (4);
4   avaliar da derivada topológica  $D_T\mathcal{F}_\omega$  de acordo com (17);
5   computar a região  $\omega^*$  de (22);
6   enquanto  $|\omega^*| \geq \pi\delta^2/4$  faça
7     intensificar a malha na ponta da trinca;
8     resolver o problema de elasticidade;
9     avaliar a derivada topológica  $D_T\mathcal{F}_\omega$ ;
10    computar a região  $\omega^*$  de (22);
11    computar a região  $\omega^\beta$  de (24);
12    nuclear a nova inclusão  $\omega^\beta$  dentro de  $\omega^*$ ;
13    atualizar a região danificada:  $\omega \leftarrow \omega \cup \omega^\beta$ ;
14    resolver o problema de elasticidade e avaliar  $D_T\mathcal{F}_\omega$ ;
15    computar a região  $\omega^*$ ;
16  fim enqto
17 fim para

```

---

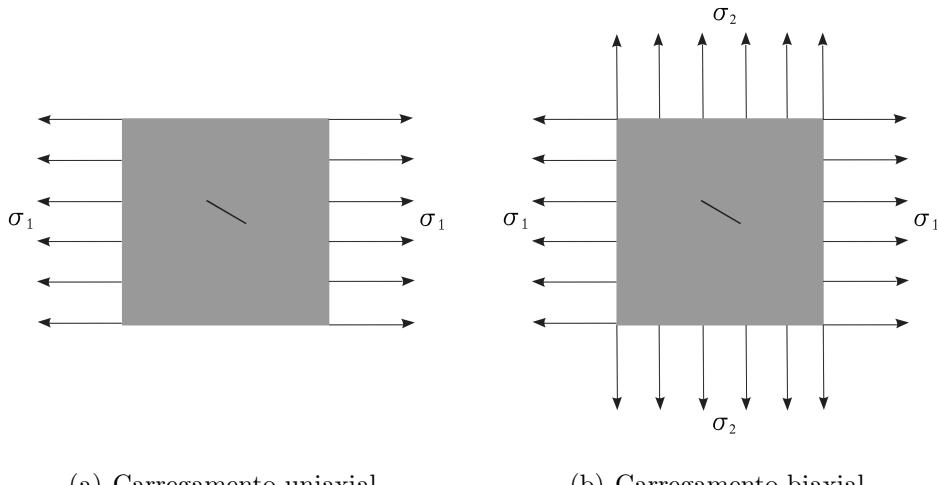
Em linhas gerais, o funcionamento do algoritmo é da seguinte forma: Dada a solução do sistema escalar (5), resolve-se o sistema (4) e o campo da derivada topológica (17) é avaliado. Caso a derivada topológica seja positiva em todo o domínio ou se  $|\omega^*| < \pi\delta^2/4$  a inclusão de uma perturbação de tamanho  $\pi\delta^2/4$  em qualquer ponto do domínio é susceptível ao aumento do valor do funcional associado. Portanto, nesse caso o algoritmo não irá propagar a falha, mas sim incrementar a pressão  $p_i$  sobre  $\Gamma_w$  e repetir este procedimento. No entanto, caso o campo da derivada topológica seja negativo numa região não danificada e  $|\omega^*| \geq \pi\delta^2/4$ , então uma inclusão  $\omega^\beta$  é nucleada dentro da região  $\omega^*$ . A nucleação de  $\omega^\beta$  modifica a topologia, de modo que deve-se computar novamente a solução do sistema

(4). Então, avalia-se o novo campo derivada topológica e o processo deve ser repetido até que a condição  $|\omega^*| \geq \pi\delta^2/4$  não mais seja satisfeita. Então, a pressão  $p_i$  é incrementada e o procedimento é repetido. As etapas aqui descritas são sintetizadas na forma do pseudocódigo apresentado no Algoritmo 1.

## 5 RESULTADOS E DISCUSSÕES

### 5.1 O efeito da tensão *in situ*

Conforme mencionado anteriormente uma das limitações do modelo proposto por Xavier [2] é a ausência do estado de tensões iniciais. Os exemplos aqui propostos ilustram como a presença da tensão *in situ* pode alterar a direção de propagação de uma falha. Sabe-se que durante o fraturamento hidráulico uma falha isolada irá se reorientar até se tornar ortogonal à direção de menor tensão principal [5]. Serão apresentados aqui dois experi-



**Fig. 2:** Tipos de carregamento.

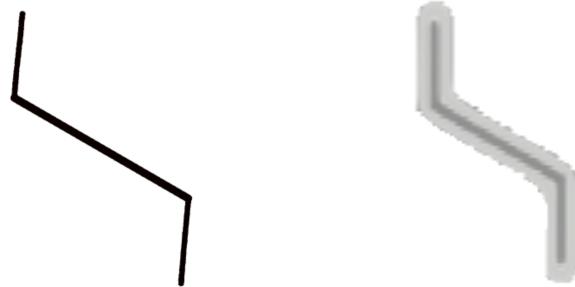
mentos, propostos por [5], nos quais considera-se um bloco de  $8m \times 8m$  na presença de uma falha de largura  $\delta = 0,025$  m e comprimento  $h = 0.8$  m em seu centro com uma inclinação de  $30^\circ$  em relação ao eixo horizontal. No primeiro, o bloco está sujeito a um carregamento uniaxial,  $\sigma_1 = 50$  MPa, ao passo que no segundo este encontra-se sob carregamento biaxial,  $\sigma_1 = 50$  MPa e  $\sigma_2 = 25$  MPa, conforme ilustra a Fig. 2. Os parâmetros utilizados durante as simulações estão summarizados a seguir.

**Tabela 1:** EFEITO DA TENSÃO *in situ*: PARÂMETROS.

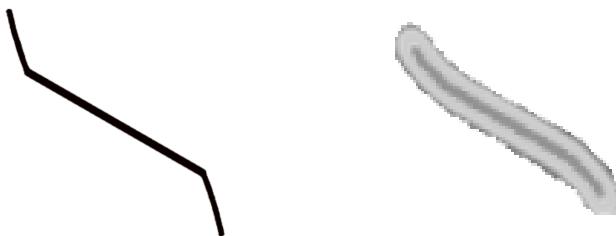
Parâmetro	Valor	Parâmetro	Valor
$h$	0.8 m	$E$	10 GPa
$\delta$	0.3125 m	$\rho_0$	$10^{-6}$
$l$	$(2/3)\delta$	$\nu$	$1/3$

Para o caso de carregamento uniaxial, a propagação da falha se dá na direção vertical. Já sob carregamento biaxial observa-se uma maior tendência para a direção vertical,

porém menos acentuada do que a observada no caso uniaxial. As Fig. 3 e 4 ilustram os resultados obtidos para os carregamentos uniaxial e biaxial (esquerda) comparados com aqueles obtidos por [5] (direita). Os parâmetros utilizados durante as simulações estão sumarizados na Tabela 1.



**Fig. 3:** Ampliação do resultado do primeiro exemplo.



**Fig. 4:** Ampliação do resultado do segundo exemplo.

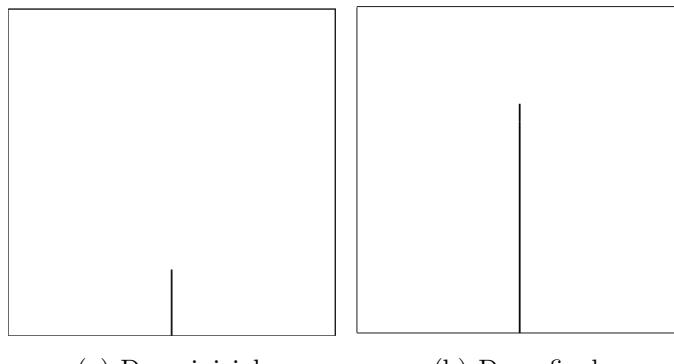
### 5.2 Exemplo Base

Depois de apresentado como a presença do estado inicial de tensões pode afetar a direção de propagação da falha, é proposto aqui um exemplo semelhante ao realizado por [2], porém com a presença de um carregamento de 3,5 MPa distribuído na face superior de um bloco de  $5m \times 5m$ . Ainda, o bloco se encontra sob o estado plano de deformações e apresenta uma falha geológica de largura  $\delta$  e comprimento  $h$  orientada na vertical e localizada no centro de sua base. Os valores atribuídos às propriedades materiais, apresentadas na Tabela 2, são utilizados por [6] e correspondem às propriedades de um campo real presente na Bacia de Campos. Os parâmetros  $\alpha^m$  e  $\alpha^f$  referem-se ao coeficiente de Biot na região saudável e na falha, respectivamente, e  $\gamma^f$  é o contraste aplicado na permeabilidade.

O modelo de dano de Francfort-Marigo não é adequado para a determinação de carga crítica em problemas com singularidades, como por exemplo, o caso limite em que  $\delta \rightarrow$

**Tabela 2:** EXEMPLO BASE: PARÂMETROS.

Parâmetro	Valor	Parâmetro	Valor
$h$	1.0 m	$E$	30 GPa
$\delta$	0.025 m	$\rho_0$	$10^{-6}$
$l$	$(2/3)\delta$	$\nu$	0,3
$p$	8 MPa	$\kappa_s$	20,0 J/m
$k$	1 mD	$\alpha^m$	0,75
$\gamma^f$	$10^6$	$\alpha^f$	1,0



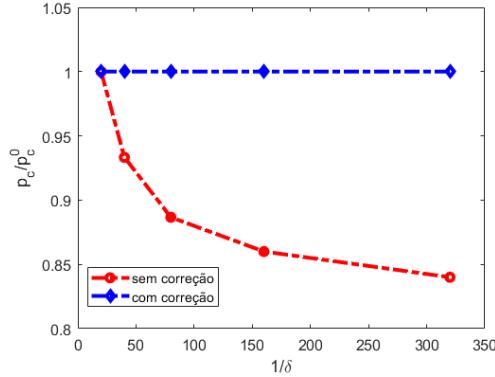
(a) Dano inicial. (b) Dano final.

**Fig. 5:** Resultados do exemplo base.

0. Neste caso, qualquer carga trativa não nula é suficiente para elevar infinitamente a densidade de energia na ponta da trinca. Nesse sentido, para caracterizar a carga crítica é introduzido o parâmetro  $\kappa_\delta$  que aumenta a medida que a largura  $\delta$  é reduzida, competindo assim contra o aumento da densidade de energia elástica na ponta do dano. Para verificar tal afirmação, são realizados cinco testes com diferentes valores para a largura  $\delta$  do dano inicial, a saber,  $\delta \in \{\frac{1}{20}, \frac{1}{40}, \frac{1}{80}, \frac{1}{160}, \frac{1}{320}\} [m]$ . A pressão crítica  $p_c$  é o valor corrente da pressão  $p_i$  prescrito sobre  $\Gamma_\omega$  capaz de realizar a nucleação da primeira inclusão. A Fig. 2(a) ilustra o comportamento da pressão crítica com e sem a introdução do fator de correção, onde os valores estão normalizados de acordo com a primeira estimativa para a pressão crítica  $p_c^0$ . Já a Fig. 2(b) representa o dano ao final de 225 iterações do processo de otimização. A pressão crítica de ativação observada foi de 3.68 MPa.

### 5.3 Meio heterogêneo

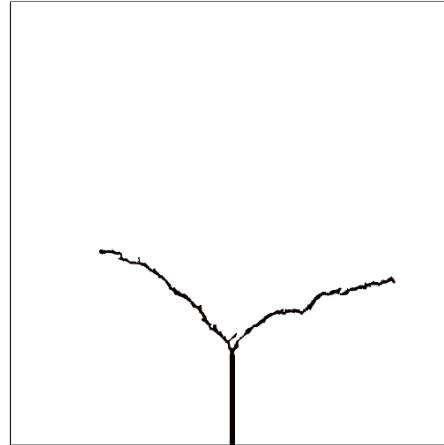
Aqui propõe-se um exemplo similar ao exemplo base e sujeito ao mesmo estado de tensões iniciais, porém considera-se que o meio não mais é homogêneo, ou seja, suas propriedades podem variar espacialmente dentro de determinada faixa. Para simular tal condição, o módulo de Young é corrompido com ruído gaussiano branco (WGN) de média nula e desvio padrão  $\tau$ . Dessa forma,  $E$  é substituído por  $E_\tau = E(1 + s\tau)$ , onde  $s : \omega \rightarrow \mathbb{R}$  é uma função que assume valores randômicos no intervalo  $(0,1)$  e  $\tau = 2$  corresponde ao nível de ruído. Os parâmetros utilizados nessa simulação são apresentados na Tabela 3 e a pressão crítica de ativação observada foi  $p_c = 6.64$  MPa. O processo foi interrompido após 140 iterações do processo de otimização e o dano final obtido é apresentado na Fig.



**Fig. 6:** Teste da pressão crítica de ativação.  
**Tabela 3:** MEIO HETEROGRÉNEO: PARÂMETROS.

Parâmetro	Valor	Parâmetro	Valor
$h$	1.0 m	$E$	17 GPa
$\delta$	0.025 m	$\rho_0$	$10^{-6}$
$l$	$(2/3)\delta$	$\nu$	0,2
$p$	8 MPa	$\kappa_s$	150,0 J/m
$k$	1 mD	$\alpha^m$	0,75
$\gamma^f$	$10^3$	$\alpha^f$	1,0

7.



**Fig. 7:** Meio heterogêneo: Dano final.

## 6 CONCLUSÕES

O trabalho aqui apresentado visa estender o modelo proposto por Xavier [2] para cenários mais próximos da realidade. Nesse sentido, o estado de tensões iniciais foi levado em consideração e sugeriu-se uma análise de sensibilidade mais simples visando incorporar o efeito da inércia do fluido durante a propagação da falha. Uma das principais vantagens do algoritmo aqui utilizado é a capacidade de capturar bifurcações, o que pode ser observado



no exemplo em que o meio é heterogêneo. A análise de sensibilidade aqui desenvolvida foi realizada de forma que possa ser utilizada para o futuro estudo e implementação do caso transiente.

## 7 Agradecimentos

Os autores agradecem o apoio das seguintes agências de fomento: CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e FAPERJ (Fundação Carlos Chagas de Amparo à Pesquisa do Estado do Rio de Janeiro) .

## REFERÊNCIAS

- [1] A. Cheng. *Poroelasticity*. Springer, 2016.
- [2] M.D.S. Xavier. *Análise de sensibilidade topológica na mecânica do dano e da fratura*. Ph.D. Thesis, Laboratório Nacional de Computação Científica, Petrópolis, Brazil, 2018.
- [3] G. A. Francfort J. J. Marigo. Stable damage evolution in a brittle continuous medium. *European Journal of Mechanics, A/Solids*, 12(2):149–189, 1993.
- [4] M. A. Biot. General theory of three-dimensional consolidation. *Journal of Applied Physics*, 12(2):155–164, 1941.
- [5] C. Chukwudzie, B. Bourdin, and K. Yoshioka. A variational approach to the modeling and numerical simulation of hydraulic fracturing under in-situ stresses. *in: Proceedings of the 38th Workshop on Geothermal Reservoir Engineering*, 2013.
- [6] L. C. Pereira, L. J. N. Guimarães, B. Horowitz, and M. Sánchez. Coupled hydro-mechanical fault reactivation analysis incorporating evidence theory for uncertainty quantification. *Computers and Geotechnics*, 56:202–215, 2014.



# Simplifications of Navier-Stokes in Journal Bearing Simulation

Júlia de Araújo Mota<sup>1</sup>, Juliana Vianna Valério<sup>2</sup> e Flavio Medeiros Rangel<sup>2</sup>

<sup>1</sup> Programa de Pós-Graduação em Informática, UFRJ, RJ, Brazil

<sup>2</sup> Departamento de Ciência da Computação, UFRJ, RJ, Brazil

---

## Abstract

Reynolds equations are widely used to understand pressure generation in lubricant theory, but there are differences in operating parameters as well as in the material used that justify the need for a review of the basic fluid film lubrication equation. For this reason, the present work aims, through a dimensionless analysis, to map the plots of the Navier-Stokes equation in this first study in journal bearings, ordering them by importance and, at the limit of simplifications, to become the Poisson equation.

**Keywords:** Navier-Stokes, Reynolds Equation, Fluid Flow, Journal Bearing, Numerical Solution

---

## 1 INTRODUCTION

The Reynolds equations is widely used to understand the generation of pressure in lubricant theory, but different situations resulted in different type of Reynolds equations. This multiplicity can mask some important effect, according to Dowson (1962). Already at this time, Dowson pointed out differences in operations parameters as well as in the material used which justified re-examination of the basic equation of fluid-film lubrication. Due to this analysis, Dowson presented a generalized form of Reynolds equation, that permits the variation of relevant quantities across, as well as along, the lubricant film.

Nowadays, the operating conditions are even more discrepant suggesting a close investigation again in the models used to describe the velocity and pressure into rotor components. Dowson (1962) generalized Reynolds equation for fluid-film lubrication. In this paper, Navies-Stokes will be revisited also evaluating the effects of inertia. A dimensionless analysis will be performed for the flow between two cylinders with a small thickness of lubricating oil film. Due to the forces and rotation, the cylinders become eccentric. Taking advantage of the geometric characteristics, especially for very small

clearance, the terms of the equations describing the flow of the fluid film can be ordered by their importance and, at the limit of simplifications, become the Poisson equation. The objective turns to be find numerically the pressure field.

In order to illustrate the success of this approach, two known situations of a journal bearing are presented: the short and the infinitely long journal bearing approximations. The pressure field is calculated in both cases and compared to the literature. It is interesting to note that, in the literature, this approximations are made neglecting terms of the simplified model which allowed analytic calculations of pressure field. In this work, these components are not neglected and the pressure field is calculated numerically from the same model even for the short and infinitely long journal bearing.

## 2 FROM NAVIER-STOKES TO POISSON

Dowson (1962) alerted that Reynolds equation, derived from lubrication theory, was no longer enough to understand important effects in bearing performances, so he presented a generalization of it.

Nowadays, even with the models improvements, these conditions are very discrepant suggesting a close investigation again in the models used to describe the velocity and pressure into rotor components. In this work, Navies-Stokes will be revisited and with a carefully dimensionless analysis of each component of the equation. This study will allow all components, and specially the neglected ones, to be ordered in importance. From this, previously neglected components can be reincorporated into the equation so that the model became able to describe some specific physical effect.

Typically, in the literature, the simplifications into the equations are made as reasonable assumptions, like in Dowson (1962), who assume that inertia and body-force terms in the equations of motion are small compared with the viscous and pressure terms. In this work this simplifications will be made by a dimensionless analysis and the magnitude of all terms will be compared between them. So, to neglect a term it has to be quantitatively smaller than the others. Assuming a Newtonian, incompressible fluid in a steady state laminar flow, using dimensionless analysis and taking advantage of the geometry since it occurs in an annular space with a small gap, the complete Navies Stokes equation is simplified allowing that velocity can be derived analytically in function of pressure gradient. So, these simplifications result in a Poisson equation that governs the pressure distribution in a film-flow annular space, as in bearings and seals. Note that the Poisson equation can be numerically solved with a low computational effort while solving the system of original equations requires a very high computational cost.

This article presents a complete deduction of simplification of Navies-Stokes sorting the equations terms in type of phenomenon and magnitude to be neglected. So, depending on the operations parameters, different terms can be considered into the model equation.

The momentum and continuity equations which govern the fluid flow are:

$$\begin{aligned} \rho \left( \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) &= \nabla \cdot \boldsymbol{\sigma}, \\ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) &= 0, \end{aligned} \tag{1}$$

where  $\rho$  is the density of the fluid,  $\mathbf{v}$  the velocity field and  $\boldsymbol{\sigma}$  is Cauchy's tensor is  $\boldsymbol{\sigma} =$



$-p\mathbf{I} + \tau$ , where  $p$  is the pressure field,  $\tau$  the stress tensor, and  $\mathbf{I}$  identity tensor.

Assuming a Newtonian fluid, the proportionality constant between deformation rates and the shear stress is the viscosity  $\mu$ :

$$\tau = \mu(\nabla \mathbf{v} + (\nabla \mathbf{v})^T),$$

since it is an incompressible flow, its density  $\rho$  is constant. And also considering a steady-state, the equation (1) became equation (2):

$$\begin{aligned} \rho(\mathbf{v} \cdot \nabla \mathbf{v}) &= -\nabla p + \mu \nabla^2 \mathbf{v}, \\ \nabla \cdot \mathbf{v} &= 0. \end{aligned} \quad (2)$$

Even with these hypotheses, we also have velocity and pressure as unknowns of a non-linear partial differential equation, that with the appropriate boundary conditions can be numerically solved, but with a high computational effort. The main goal lies on taking advantage of the geometric characteristics of these rotors components as bearings and seals. It occurs in an annular space with a *very small* gap between the inner and outer cylinders.

The smaller this gap, compared to the radius of each cylinder and component length, the more terms are neglected and the model becomes simpler. These geometric characteristics are responsible for huge simplifications in the governing equations that will be detailed presented.

Following the formulation presented by Pina and Carvalho (2006) and Andrade (2008) the cylindrical coordinates will be used and also considered the curvature effects. A common assumption in the literature is that the radius of curvature of the solids bounding the oil film is large compared with the thickness of the lubricant film neglecting the curvature effects.

Direction  $z$ :

$$\rho \left( u \frac{\partial u}{\partial z} + v \frac{\partial u}{\partial r} + \frac{w}{r} \frac{\partial u}{\partial \theta} \right) = -\frac{\partial p}{\partial z} + \mu \left( \frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial u}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2} \right). \quad (3)$$

Direction  $r$ :

$$\begin{aligned} \rho \left( v \frac{\partial v}{\partial r} + u \frac{\partial v}{\partial z} + \frac{w}{r} \frac{\partial v}{\partial \theta} - \frac{w^2}{r} \right) &= \\ -\frac{\partial p}{\partial r} + \mu \left( \frac{\partial}{\partial r} \left[ \frac{1}{r} \frac{\partial(rv)}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} - \frac{2}{r^2} \frac{\partial w}{\partial \theta} + \frac{\partial^2 v}{\partial z^2} \right). \end{aligned} \quad (4)$$

Direction  $\theta$ :

$$\begin{aligned} \rho \left( v \frac{\partial w}{\partial r} + u \frac{\partial w}{\partial z} + \frac{w}{r} \frac{\partial w}{\partial \theta} + \frac{vw}{r} \right) &= \\ -\frac{1}{r} \frac{\partial p}{\partial \theta} + \mu \left( \frac{1}{r} \frac{\partial}{\partial r} \left[ \frac{1}{r} \frac{\partial(rw)}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 w}{\partial \theta^2} + \frac{2}{r^2} \frac{\partial v}{\partial \theta} + \frac{\partial^2 w}{\partial z^2} \right). \end{aligned} \quad (5)$$

Continuity:

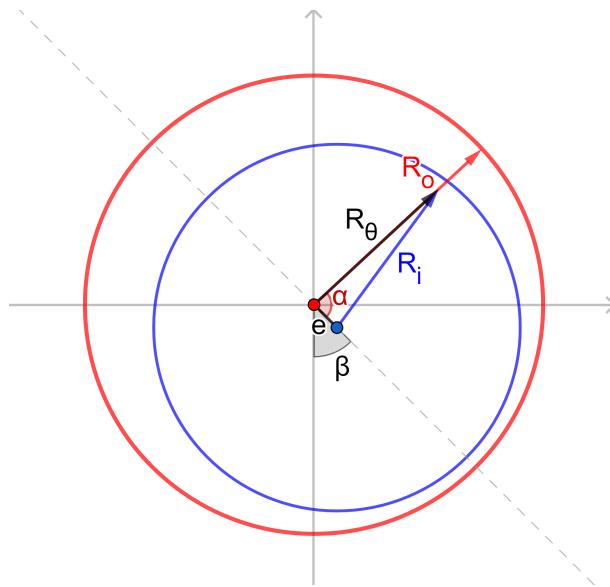
$$\frac{1}{r} \frac{\partial(rv)}{\partial r} + \frac{1}{r} \frac{\partial w}{\partial \theta} + \frac{\partial u}{\partial z} = 0 \quad (6)$$

## 2.1 Geometry

The flow occurs in an annular space with between the inner and outer cylinders. In the journal bearing situation, they called rotor and stator as shown in the Fig.(1). Due to eccentricity, the rotor description from the stator center depends on  $\theta$  and is given by the equation (7).

$$R_\theta = e \cos \alpha + \sqrt{R_i^2 - e^2 \sin^2 \alpha}, \quad (7)$$

$$\text{where } \alpha = \begin{cases} \left| \theta + \left( \frac{\pi}{2} - \beta \right) \right| & \text{if } 0 < |\theta + (\pi - \beta)| < \pi \\ \left| 2\pi - \theta - \left( \frac{\pi}{2} - \beta \right) \right| & \text{if } \pi < \left| \theta + \left( \frac{\pi}{2} - \beta \right) \right| < 2\pi \end{cases}.$$



**Fig. 1:** Schematic drawing of an eccentricity journal bearing

The gap between stator and rotor radius ( $F = R_o - R_\theta$ ) is small compared with the radius of each cylinder ( $R_o, R_i$ ) and length of the component ( $L$ ).

## 2.2 Dimensionless analysis

As commented before, taking advantage of the geometric characteristics of these rotors components, which occur in an annular space with a *very small* gap between the inner and outer cylinders, significant simplifications are achieved.

Comparing the magnitude of the terms it is possible to order the equations components in importance and decide which effects could be neglected. These geometric characteristics are responsible for huge simplifications in the governing equations that will be detailed presented.

Considering  $U$  typical velocity,  $L$  characteristic length with the same order of magnitude as the outer and inner radius. Although the gap  $(R_o - R_\theta) = F$  is much smaller than characteristic quantity,  $(R_o - R_\theta) = F \ll L$ .



Realizing that in the gap occurs a gradient of pressure, so in the magnitude of the pressure the  $F$  is considered instead of  $L$ :  $P = \mu UL/F^2$ .

The hat means dimensionless variable:  $u = U\hat{u}$ ,  $v = U\hat{v}$ ,  $w = U\hat{w}$ ,  $p = P\hat{p}$ ,  $z = L\hat{z}$ ,  $r = [(R_o - R_\theta)]\hat{r} = F\Delta\hat{r}$ ,  $(r\theta) = L(\hat{r}\theta)$ .

The equation (3) becomes:

$$\begin{aligned} \rho \left( U\hat{u} \frac{\partial U\hat{u}}{\partial L\hat{z}} + U\hat{v} \frac{\partial U\hat{u}}{\partial F\hat{r}} + \frac{U\hat{w}}{L\hat{r}} \frac{\partial U\hat{u}}{\partial \theta} \right) = \\ - \frac{\partial P\hat{p}}{\partial L\hat{z}} + \mu \left( \frac{1}{L\hat{r}} \frac{\partial}{\partial F\hat{r}} \left[ L\hat{r} \frac{\partial U\hat{u}}{\partial F\hat{r}} \right] + \frac{1}{L^2\hat{r}^2} \frac{\partial^2 U\hat{u}}{\partial \theta^2} + \frac{\partial^2 U\hat{u}}{\partial L^2\hat{z}^2} \right) \end{aligned} \quad (8)$$

Rearranging equation (8) representing the  $z$ -direction motion and highlighting Reynolds number  $\left( Re = \frac{\rho UL}{\mu} \right)$

$$\begin{aligned} Re \left( \left( \frac{F^2}{L^2} \right) \hat{u} \frac{\partial \hat{u}}{\partial \hat{z}} + \left( \frac{F}{L} \right) \hat{v} \frac{\partial \hat{u}}{\partial \hat{r}} + \left( \frac{F^2}{L^2} \right) \frac{\hat{w}}{\hat{r}} \frac{\partial \hat{u}}{\partial \theta} \right) = \\ - \frac{\partial \hat{p}}{\partial \hat{z}} + \left( \frac{1}{\hat{r}} \frac{\partial}{\partial \hat{r}} \left[ \hat{r} \frac{\partial \hat{u}}{\partial \hat{r}} \right] + \left( \frac{F^2}{L^2} \right) \frac{1}{\hat{r}^2} \frac{\partial^2 \hat{u}}{\partial \theta^2} + \left( \frac{F^2}{L^2} \right) \frac{\partial^2 \hat{u}}{\partial \hat{z}^2} \right) \end{aligned} \quad (9)$$

Note that in the inertial part of the equation there are terms smaller than others, since  $(F^2/L^2) < (F/L)$ . So, this term  $\hat{v} \frac{\partial \hat{u}}{\partial \hat{r}}$  is a candidate to return to model in order to consider some inertial effects.

In a traditional lubrication theory all  $(F^2/L^2)$  and  $(F/L)$  terms are neglected, so the equation becomes:

$$- \frac{\partial p}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) = 0 \quad (10)$$

Doing the same in other directions, neglecting all  $(F^2/L^2)$  and  $(F/L)$  terms:

$$- \frac{\partial p}{\partial r} = 0, \quad (11)$$

the pressure does not vary in the radial direction.

And finally, the equation in the tangential direction  $\theta$ , has to:

$$\frac{\partial}{\partial r} \left( \frac{1}{r} \frac{\partial(rw)}{\partial r} \right) - \frac{1}{r} \frac{\partial p}{\partial \theta} = 0. \quad (12)$$

After the simplifications, it is clear that the pressure in the radial direction is constant since its gradient is zero and the radial velocity has disappeared. For this reason it can be considered negligible compared to other velocities. Notice that to include the turbulence effect, a candidate term to return to equation, among other terms, is the radial velocity. This fact makes the radial velocity susceptible to investigations in disturbance cases.

### 2.3 Velocities and Poisson equation

With the typical simplifications it is possible integrate the equations to find the velocities. It is going to change if the ( $F/L$ ) terms were not neglected. Integrating (10) e (12) to find the velocities:

$$u = \left( \frac{\partial p}{\partial z} \right) \frac{1}{4} r^2 + c_1 \ln r + c_2 \quad (13)$$

$$w = \frac{1}{2} \frac{\partial p}{\partial \theta} r \left( \ln r - \frac{1}{2} \right) + c_3 r + \frac{c_4}{r} \quad (14)$$

The boundary conditions, remembering that the radius  $R_o$  of the stator and  $R_\theta$  the description of the rotor from the stator center:

- $u(R_o) = 0$  e  $u(R_\theta) = 0$
- $w(R_o) = 0$  e  $w(R_\theta) = W$

$$u = \frac{\partial p}{\partial z} \frac{R_\theta^2}{4} \left[ \left( \frac{r}{R_\theta} \right)^2 - \frac{(R_o^2 - R_\theta^2)}{R_i^2 \ln(R_o/R_\theta)} \ln \left( \frac{r}{R_\theta} \right) - 1 \right] \quad (15)$$

$$w = \frac{1}{2} \frac{\partial p}{\partial \theta} \left[ r \left( \ln r - \frac{1}{2} \right) + Kr - \frac{R_o^2}{r} \left( \ln R_o + K - \frac{1}{2} \right) \right] + \frac{WR_\theta}{(R_\theta^2 - R_o^2)} \left( r - \frac{R_o^2}{r} \right). \quad (16)$$

where

$$K = \frac{1}{R_o^2 - R_\theta^2} \left[ R_\theta^2 \left( \ln R_\theta - \frac{1}{2} \right) - R_o^2 \left( \ln R_o - \frac{1}{2} \right) \right] \quad (17)$$

With velocities the continuity equation is used and integrated in the annular region of interest.

$$\int_{R_\theta}^{R_o} \left( \frac{\partial rv}{\partial r} + \frac{\partial w}{\partial \theta} + \frac{\partial ru}{\partial z} \right) dr = 0 \quad (18)$$

Using the expressions (15) and (16) in (18):

$$\frac{\partial}{\partial \theta} \left( C_1 \frac{\partial p}{\partial \theta} \right) + \frac{\partial}{\partial z} \left( C_2 \frac{\partial p}{\partial z} \right) = \frac{\partial}{\partial \theta} C_{0W} \quad (19)$$

where:

$$C_{0W} = -WR_\theta \left[ \ln \left( \frac{R_o}{R_\theta} \right) \left( 1 + \frac{R_\theta^2}{(R_o^2 - R_\theta^2)} \right) - \frac{1}{2} \right] \quad (20)$$

$$C_1 = \frac{R_\theta}{2} \left[ \frac{1}{2R_\theta} [R_o^2 \ln R_o - R_\theta^2 \ln R_\theta - (R_o^2 - R_\theta^2)(1 + K)] \right] - \frac{R_\theta^2}{2\mu} \left[ \left( \ln R_\theta - \frac{1}{2} + K \right) \ln \left( \frac{R_o}{R_\theta} \right) \right] \quad (21)$$



$$C_2 = \frac{-R_\theta^2}{8} \left[ R_o^2 - R_\theta^2 - \frac{(R_o^4 - R_\theta^4)}{2R_\theta^2} \right] + \left( \frac{R_o^2 - R_\theta^2}{R_r^2 \ln(R_o/R_\theta)} \right) \left[ R_o^2 \ln\left(\frac{R_o}{R_\theta}\right) - \frac{(R_o^2 - R_\theta^2)}{2} \right] \quad (22)$$

This is a Poisson equation and a numerical method is necessary to solve it. With the pressure field approximation, the velocities can be calculated.

A inner and out pressure is imposed.

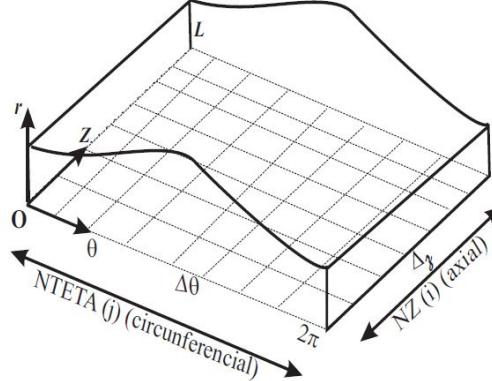
$$p(z = 0) = P_{in} \quad p(z = L_b) = P_{out}.$$

### 3 NUMERICAL SOLUTION

The transformation of the three-dimensional Navier-Stokes equations into a Poisson equation of two dimensions allows a considerable computational gain. Even with the gain a numerical scheme is necessary to calculate an approximation of the pressure field in the domain of interest. A centered-difference formula (see Burden and Faires (2010)) was chosen in order to describe locally the gradient of pressure.

#### 3.1 Discretization

A rectangular grid with equally spaced  $NZ$  nodes along the axial direction and  $N\theta$  nodes along the azimuthal direction was used to discretize the domain, as shown in Fig.(2).



**Fig. 2:** Figure from Andrade (2008)

The functions  $F_1$  and  $F_2$  were evaluated at each node position in order to calculate the appropriate derivatives.

So  $F_1 = C_1(z, \theta) \frac{\partial p}{\partial \theta}$  and  $F_2 = C_2(z, \theta) \frac{\partial p}{\partial z}$  became:

$$\frac{\partial F_1}{\partial \theta} = \frac{F_1^E - F_1^W}{\Delta \theta} \quad \frac{\partial F_2}{\partial z} = \frac{F_2^N - F_2^S}{\Delta z}$$

The term independent of pressure is also discretized:

$$\frac{\partial C_{0W}}{\partial \theta} = \frac{C_{0W}^E - C_{0W}^W}{\Delta \theta} \quad \frac{\partial C_{0U}}{\partial z} = \frac{C_{0U}^N - C_{0U}^S}{\Delta z}$$

Using the discretization scheme the results is an algebraic equation for each node  $i, j$  relates the pressure there  $P_{i,j}$  with its four neighbors, equation (23):

$$\begin{aligned} P_{i-1,j} \frac{(C_{2(i-1,j)})}{\Delta z^2} + P_{i,j-1} \frac{(C_{1(i,j-1)})}{\Delta \theta^2} - P_{i,j} \left( \frac{(C_{1(i,j)} + C_{1(i,j-1)})}{\Delta \theta^2} + \frac{(C_{1(i,j)} + C_{1(i,j-1)})}{\Delta z^2} \right) + \\ P_{i,j+1} \frac{(C_{1(i,j)})}{\Delta \theta^2} + P_{i+1,j} \frac{(C_{2(i,j)})}{\Delta z^2} = \frac{1}{\Delta \theta} [C_{0W(i,j)} - C_{0W(i,j-1)}] \end{aligned} \quad (23)$$

The boundary conditions are:

$$\begin{aligned} p(z=0) &= p(1,j) = P_{in} \\ p(z=L_b) &= p(NZ,j) = P_{out} \\ p(\theta=0) &= p(\theta=2\pi) = p(i,1) = p(i,N\theta) \end{aligned}$$

The system is then solved using Numpy's Linear Algebra Package<sup>1</sup>.

## 4 RESULTS AND DISCUSSION

In the literature, several works use the Reynolds equation after a series of simplifications to find the pressure field in journal bearings. However, since the equation found can only be solved numerically, they use the strategy of approximating the equation for the cases of short and infinitely long journal bearings, finding reduced models with analytical resolution.

In order to verify the simplifications that have been made in this work, we are going to compare the results obtained with the current approximations and the model presented here. The relationship used to determine the classification of a journal bearing by its length was presented in Frêne *et al.* (1990). According to the author, a journal bearing may be considered short if the ratio  $L/(2R_o)$  is less than or equal to 1/8 and long if it is greater than 4.

### 4.1 Short journal bearing

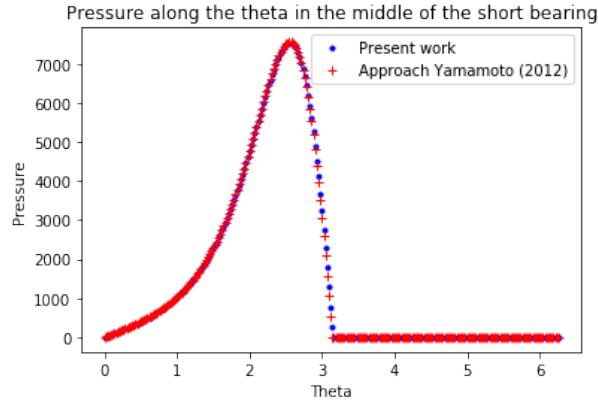
According to Ishida and Yamamoto (2012), to analyze the short journal bearing, the pressure variation in the  $z$  direction is considered much larger than in the  $\theta$  direction, i.e.  $\partial p / \partial \theta \ll \partial p / \partial z$ . Thus, the first term in the equation (15) is neglected and the approach formula for short journal bearing is obtained as the equation (24).

$$p = \frac{-3\mu\epsilon\omega \sin \theta}{(R_\theta - R_i)^2 (1 + \epsilon \cos \theta)^3} \left[ \left( z - \frac{L}{2} \right)^2 - \frac{L^2}{4} \right] \quad (24)$$

where  $\epsilon = \frac{e}{R_o - R_i}$  is the eccentricity ratio and  $\omega$  the rotor speed.

A numerical simulation was made for a rotation of  $\omega = 10,472 [\text{rad/s}]$  on a radius rotor journal bearing  $R_i = 0.2 [\text{m}]$ , with a radial clearance  $F = 0.0001 [\text{m}]$ , length  $L = 0.4 [\text{m}]$ , operating with a viscosity oil  $\mu = 0.015 [\text{Pa.s}]$  and density  $\rho = 860 [\text{Kg/m}^3]$ , needed a  $NZ = 16$  and  $N\theta = 264$  to obtain a relative error  $E \approx 10^{-3}$ . Pressure behavior is reported with the literature result in Fig.(3).

<sup>1</sup><https://docs.scipy.org/doc/numpy/reference/routines.linalg.html>



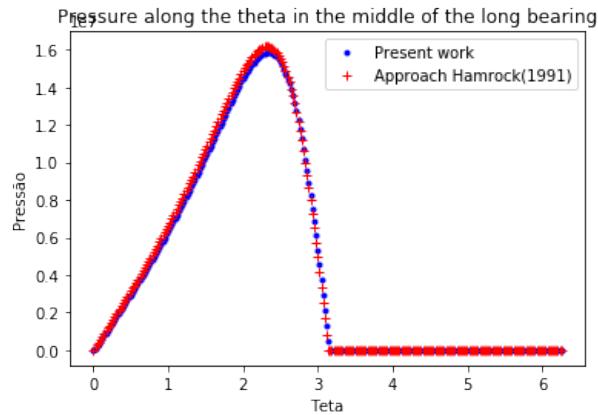
**Fig. 3:** Pressure calculated on a short journal bearing

#### 4.2 Infinitely long journal bearing

Another verification will be performed by comparing the results obtained with those of the long journal bearing approach developed by Hamrock (1991). In this context, the pressure variation in the  $\theta$  direction is considered much larger than in the  $z$  direction, a term of the equation is neglected, and the long journal bearing approximation formula presented by the author is given in the equation 25. Compared to the solution of the present work it has a relative error  $E \approx 10^{-2}$ .

$$p = \frac{6\mu\omega \left(\frac{R_i}{F}\right)^2 \epsilon \sin \theta (2 + \epsilon \cos \theta)}{(2 + \epsilon^2)(1 + \epsilon \cos \theta)^2} + P_{in} \quad (25)$$

The numerical simulation for the infinitely long journal bearing was made with the same parameters as above except for the length  $L = 1.6 [m]$ . The pressure behavior for the infinitely long journal bearing can be seen in Fig.(4).



**Fig. 4:** Pressure calculated on a long journal bearing

## 5 CONCLUSIONS

The preliminary results are quite promising. Once compared the model according to the literature, the following investigation will include new operating parameters. Taking advantage of the mapping done through dimensionless analysis, the impacts of the incorporation of terms will be understood within the traditional model, aiming to contemplate existing physical effects that are not currently captured.

Another possibility that the present work provides is the advance in the studies on rotordynamic coefficients in journal bearings, besides the possibility of obtaining results for a different geometry, which allows to adapt the calculations for other rotary machine components such as seals.

## 6 Acknowledgements

The code used for the simulations presented in this paper is part of ROSS<sup>2</sup>, an open source library written in Python for rotordynamic analysis, by Timbó et al. (2019).

## REFERENCES

- [1] Andrade, S. (2008). *Asymptotic Model for Monophasic Flow Through Progressive Cavities Pump*. Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.
- [2] Burden, R. L. and Faires, J. D. (2010). *Numerical Analysis*. Cengage Learning, 9 ed.
- [3] Dowson, D. (1962). “A generalized reynolds equation for fluid-film lubrication”. *International Journal of Mechanical Sciences*, 4(2):159–170.
- [4] Frêne, J., Nicolas, D., Degueurce, B., Berthe, D., and Godet, M. (1990). “Static characteristics of journal bearings”. In: *Hydrodynamic Lubrication - Bearings and Thrust Bearings*, volume 33 of *Tribology Series*, Elsevier, Amsterdam.
- [5] Hamrock, B. J. (1991). “Hydrodynamic journal bearings - analytical solutions”. In: *Fundamentals of fluid lubrication*, NASA, Washington.
- [6] Ishida, Y. and Yamamoto, T. (2012). “Flow-induced vibrations”. In: *Linear and nonlinear rotordynamics*, Wiley Online Library, Weinheim.
- [7] Pina, E. P. and Carvalho, M. S. (2006). “Three-dimensional flow of a newtonian liquid through an annular space with axially varying eccentricity”. *Journal of Fluids Engineering*, 128(2):223–231.
- [8] Timbó, R., Martins, R., Bachmann, G., Rangel, F., Mota, J., Valério, J., and Ritto, T. (2019). “Ross - rotordynamic open source software”. *Journal of Open Source Software*. Submitted.

<sup>2</sup><https://github.com/ross-rotordynamics/ross>



# Deep Neural Network for Single Event Effects Classification.

Juliano Alves de Oliveira<sup>1</sup>, Marco Antônio Assis de Mello<sup>1</sup>, Marcilei Aparecida Guazzelli da Silveira<sup>1</sup>, Rudolf Theoderich Buhler<sup>1</sup>, Vitor Ângelo Paulino de Aguiar<sup>2</sup>, Nilberto Heder Medina<sup>2</sup> e Renato Camargo Giacomini<sup>1</sup>

<sup>1</sup> Centro Universitário da FEI, São Bernardo do Campo/SP, Brazil

<sup>2</sup> Instituto de Física da USP, São Paulo/SP, Brazil

---

## Abstract

This work proposes the use of deep neural networks, specifically a DeepConvLSTM network, for the classification of single effect events (SEE) that occur in electronic devices. Also proposed are three architectural variations derived from the original proposal, DeepConvGRU, DeepConvBiLSTM and DeepConvBiGRU, with the purpose of improving the system classification capacity. Finally, a simple neural network interpretability algorithm is proposed to observe the importance of each signal feature to better understand the phenomenon and possible generations of insights into the electrical characteristics of the phenomenon.

**Keywords:** Deep Neural Networks, Single Event Effects, LSTM, GRU, classification

---

## 1 INTRODUCTION

The study of ionizing particles is relevant in several areas of knowledge such as aerospace, military, medicine, radio-protection, among others, because these applications can suffer from the effects of ionizing radiation on their electronic circuits (20; 14; 22). The detection and classification of particles have also a notable application in high energy physics research, which presents in its experiments a large data volume. Such data must be processed in real time, in order to define whether it should be stored or not, for further analysis. The smart detectors use classification algorithms, which identify known particle patterns and discards them, thus avoiding unnecessary storage of data (8). There are several approaches and circuits used in particle detection, some are based on CMOS pixel sensors (CPS) that rely on circuits and sensor elements on the same substrate (33). Other applications involve Monolithic Active Pixel Sensors (MAPS) which are a set of sensors

forming a binary activation grid, in this way, for each collision that reaches a certain area, a set of pixels are disturbed causing its activation, and signaling that there was activity in that region (29). These examples of detectors are based on a set of sensors that allows great event coverage and have good spatial discrimination of the particle. In this work, we use a single sensor element, a commercial off-the-shelf transistor (COTS) PMOS 3N163. Although a set of sensor elements has a broader response to the event, the convenience of having a single sensor element with a single signal, in response to the event, was chosen because it allows further signal processing without the need to synchronize the multiple answers that would be obtained in other configurations.

Single Event Effect (SEE) is a short-lived phenomena that occur when an ionizing particle collides with a semiconductor device, creating in its path a cloud of electron-hole pairs, which are assimilated by the electric current flowing in the target device. It raises an anomaly in its operation. Within the various sub-classifications of the SEE, the Single Event Transient (SET) is characterized as the one that generates a momentary current peak in the device (15; 34). This type of event will be the object of study of this work. As its name indicates, the SET has a transient nature. In this way, this phenomenon presents temporal dependence, unlike a Single Event Upset (SEU), for example, that occurs in digital devices and can be interpreted as a logical inversion in the device (memory or register). Because of this transitory nature of SET (see fig. 10), a classification system that as input such data must be able to deal with this temporal variable. Traditional statistical methods present difficulties in the classification of problems with temporal dependence (30), but there are several advances in recent years in machine learning area, leveraged mainly by an increase in computational performance and access to large volumes of data. The strategy of bringing all these advances for the field of particle classification and detection was adopted in this work. One of the techniques that has been showing great gain in recent research on several non trivial problems such as image processing is the use of deep neural networks (DNN). It allows the discovery and extraction of characteristics of the input signals, and with each layer added, these characteristics are stacked and combined, creating new characteristics, increasing the information diversity and bringing new ways for better classifiers. It also presents the advantage of allowing the use of raw data with the need of little pre-processing. The different configurations of the DNN allow expressive gains depending on the addressed problem. For this particular case two DNN architectures based on (35) working together were adopted. The first one is known as Convolutional Neural Network (CNN), that allows the extraction of characteristics in a spatial way through filters that are obtained during the learning process. Such filters run through the entire input vector searching for the pattern to which each filter was designed to detect. The second configuration that works in conjunction with CNN is known as Long Short Term Memory (LSTM) which are cells able to learn characteristics and relate them over time, and unlike Recurrent Neural Networks (RNN) that present problems with information that presents a long dependence in time, the LSTM's are able to deal with this problem. They are frequently used in the processing of texts and sound, which are the main cases that present temporal relation (37). The adoption of machine learning techniques is already widely used in the field of high particle physics (HEP) in several fields. For example, the use of Generative Adversarial Networks (GANs) in the simulation of high energy phenomena. With a computational cost lower than traditional simulation

methods it furnishes closer results (19). Other successful applications include, but are not limited to: hit reconstruction (1), particle tracking finding in individual detector system (13), multiple particles detection using information provided by multiple systems (4; 26; 3), and also using machine learning on initial trigger system (27) and early Higgs boson searches and discovery (26) e (2).

This paper explores the application of machine learning techniques for the classification of electrical signals generated from the SEE, as well as the use of the algorithm to analyze second order effects, present in the measured signals.

Section 2 better describes the DNN used architecture, pre-processing of data, the experimental setup used to obtain the data, as well as its operability. Section 3 will discuss the obtained results. Lastly, in section 4, conclusions and future work will be exposed.

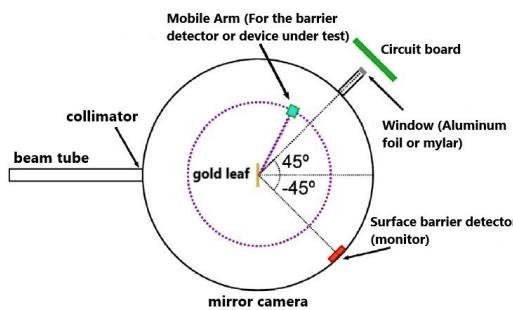
## 2 Methods and techniques

In this section, the materials and methods used during the experiments, as well as the methods adopted in the event acquisition process, will be described.

### 2.1 Experimental setup

The Pelletron 8UD accelerator is an electrostatic accelerator of the Tandem category, which can achieve energies above 100 MeV for experiments that require the use of heavier ions. This category of accelerator uses a mechanism where an ion with negative charge is accelerated by a positive electric field until a region where the charge is inverted, being accelerated again after this inversion.

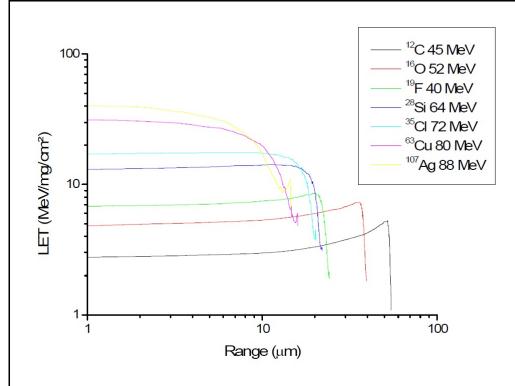
The possible configurations allow, in addition as well as to experiments inside the vacuum spreading chamber, irradiation in equipment outside this environment, since some targets are difficult to be used in vacuum. Using a window with a thin sheet of aluminum or mylar, it is possible to route the ions from within the chamber to the device in question. In this configuration, the window can be positioned at 15 or 45 degree scattering angles (Figure 1), but limited to external proton beams up to  $^{48}Ti$ , due to high energy loss in the air and in the windows about to leave.



**Fig. 1:** Schematic of experimental setup used in radiations.

At the date of the experiment, the available ions in the *Laboratório Aberto de Física Nuclear*(LAFN - Open Laboratory of Nuclear Physics) are:  $^1H$ ,  $^{6,7}Li$ ,  $^{10,11}B$ ,  $^{12,13}C$ ,

$^{16,17,18}O$ ,  $^{19}F$ ,  $^{28,29,30}Si$ ,  $^{35,37}Cl$ ,  $^{48}Ti$ ,  $^{63,65}Cu$  and  $^{107,109}Ag$ . In figure 2 it is possible to consult the LET values used in the function experiments of the silicon penetration depth.



**Fig. 2:** Heavy ions used in experiments. LET as function of the depth in the silicon.

Table 1 discriminates the ions used in the study in question, which include  $^{12}C$ ,  $^{16}O$ ,  $^{19}F$ ,  $^{28}Si$ ,  $^{35}Cl$ ,  $^{63}Cu$  and  $^{107}Ag$ . These ions were scattered at 15 degrees through a gold foil of  $275 \mu\text{g.cm}^{-2}$ , which provides LETs of 2 to 45 MeV/mg/cm<sup>2</sup>.

**Table 1:** BEAM ENERGY, PENETRATION DEPTH AND LET IN SILICON.

Ion	Energy (MeV)	Depth in Si(μm)	LET in Si(MeV/mg/cm <sup>2</sup> )
$^{12}C$	45.0	55.4	2.6
$^{16}O$	52.5	38.9	4.6
$^{19}F$	42.0	23.8	6.8
$^{28}Si$	66.0	21.6	12.8
$^{35}Cl$	75.0	19.5	17.1
$^{63}Cu$	82.0	14.1	30.7
$^{107}Ag$	97.5	13.1	39.8

A second experiment was carried out to obtain new data to evaluate predicted effects in simulation. In this second round the newest line present in the LAFN named SAFIIRA (See figure 3) was used.

This new line was built to meet the international standards that govern ionizing radiation studies. Some of these requirements, such as low ion flow, high coverage area, high uniformity, sample handling and various combinations of ions and energies (5). In this second round of the experiment a single ion beam of O was used using energy of 63 MeV. Table 2 contains information about the beam used in the second round.

**Table 2:** BEAM ENERGY, PENETRATION DEPTH AND LET IN SILICON.

Ion	Energy (MeV)	Depth in Si(μm)	LET in Si(MeV/mg/cm <sup>2</sup> )
$^{16}O$	63	50.4	4.6

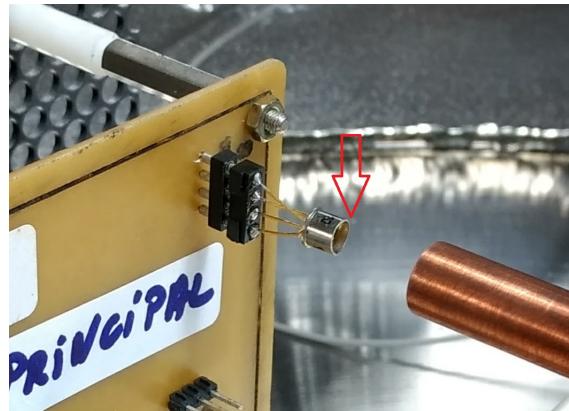


**Fig. 3:** Overview of the new line SAFIIRA.

## 2.2 Device under test (DUT)

The PMOS 3N163 transistor was used. The use of this type of device is due to several benefits shown in the study of ionizing radiations, the ease of acquiring the device, biasing and manipulation of the component. Its metal encapsulation is easy to remove thus facilitating the exposure of the silicon to the particles. Previous studies with such component available, showing the responsiveness of this component to events involving radiation, as a good dosimeter (12) and also as SEE detector (32).

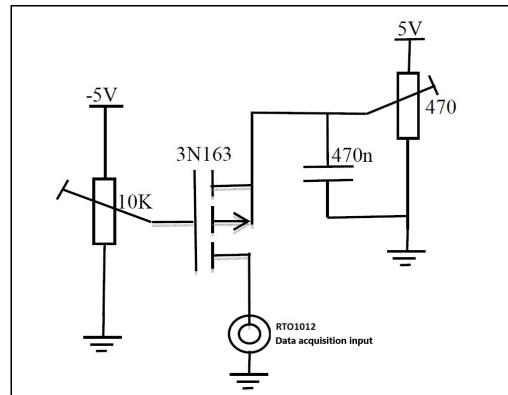
In all setups the device package is removed, in order to expose the sensitive regions of the device directly to the irradiated beams. In figure 4 it is possible to observe the decapsulated component.



**Fig. 4:** Detail of the decapsulated device.

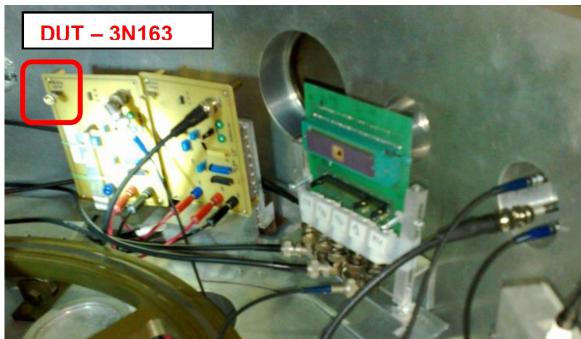
The first configuration used in the tests adopted bias of a polarization 0.13V between drain and source (VDS) and  $-4.5V$  between gate and source (VGS). The device operates in its linear region, making the response to the nearest event without distortion by the amplifying properties that could be imparted to the device due to its polarization. The drain current was amplified, and the current measured through a successive approximation 12-bit A/D converter preceded by a Sample-and-Hold circuit. Thus any change in current flowing through the device is recorded and associated with an effect caused by the emission of radiation in the device. Due to the low sampling rate in relation to the duration of a

SEE event, only 0.02% of the possible events were recorded (31), resulting in the need to create a new system to perform data collection. The second configuration implemented a series of improvements in the experiment. The first one was the polarization circuit (Figure 5), which gave the experiment the ability to polarize the device under test in several operating regions. The sampling rate has increased significantly with the use of a Rohde & Schwarz model 1-GHz RTO1012 oscilloscope, which provides the sampling capability of 10G samples/s, which is equipped with a real-time digital trigger feature. This monitoring system proved to be of great value because of the memory economy it made possible, since it only recorded the events in which the voltage values exceeded a threshold, rejecting possible spurious signs.

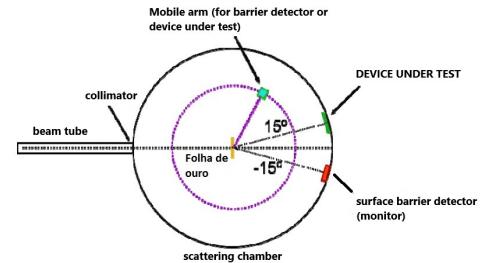


**Fig. 5:** Schematic of the 3N163 transistor bias system.

The device under test was positioned inside the vacuum chamber (Figure 6), using a plate where the polarization circuit, connectors and supports for the transistor were located. The device was irradiated at 15 degrees, as shown in figure 7 (38).



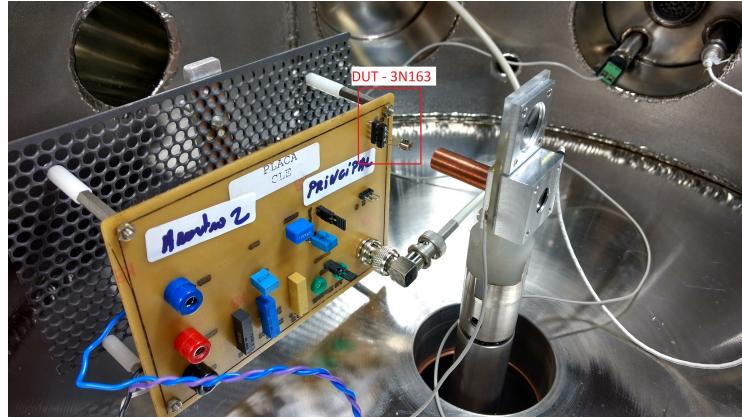
**Fig. 6:** Device under test in the vacuum chamber.



**Fig. 7:** Schematic view of the arrangement used to perform the SEE tests.

During the radiation simulations, other waveforms were detected resulting from the action of the ionizing particles in the port region of the simulated 3N163 device, in this way a new round of experiments was necessary to obtain more data to confront the simulated results. For this new experiment, the same configuration of polarizing circuit described previously was used, as it proved to be a good apparatus to obtain the pulses from the radiation, but the arrangement of the plate in the camera occurred differently from the

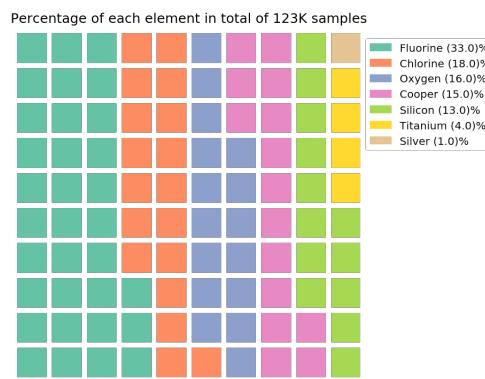
previous one, in this new line the device was directly positioned in the direction of the beam, there was no scattering of the beam made by a gold foil, in this way it was possible to record a larger number of events during the experiment. The figure 8 contemplates the positioning of the device to be tested and all the polarization apparatus necessary for its operation.



**Fig. 8:** Disposition of the DUT in the camera used in the second experiment.

### 2.3 Data analysis

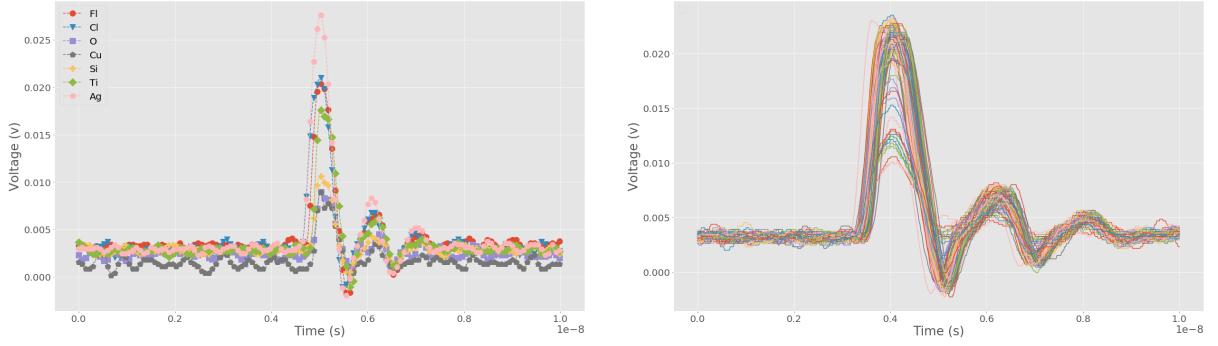
The data obtained during experiments in the Pelletron are of a transient nature, have a peak which records the moment when the electron-hole pairs generated during the traversal of the ionizing particle by the device are assimilated to the main current of the device. Each recorded pulse has 2000 points, spaced in  $10\mu s$ , which results in a  $20\mu s$  event capture window. Approximately 123K samples were collected, being 7 different elements, as can be seen in figure 9, there are a number of different samples for each ion, this is due to several factors such as the fluency of the particle during the experiment.



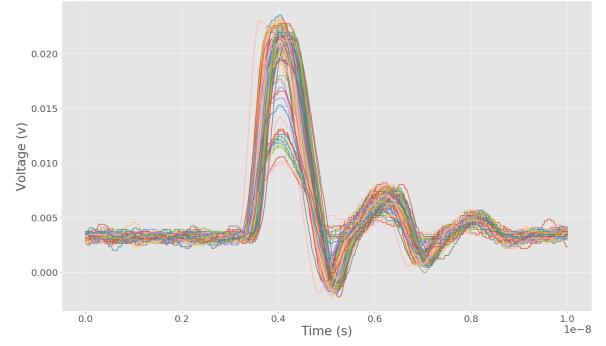
**Fig. 9:** Percentage of each element in total of 123k samples. Each square represents approximately 1.5K samples.

The challenge of classifying these different ions is that, given a certain event for an element, different responses are obtained because the characteristics of the signal depend,

besides the type of ion and its energy, also depend on the angle of arrival and depth of penetration of the particle, which are parameters difficult to control and which make the classification task more complex. In figure 10 a sample of each ion is observed and in figure 11 the distribution of 100 ions of the same type.

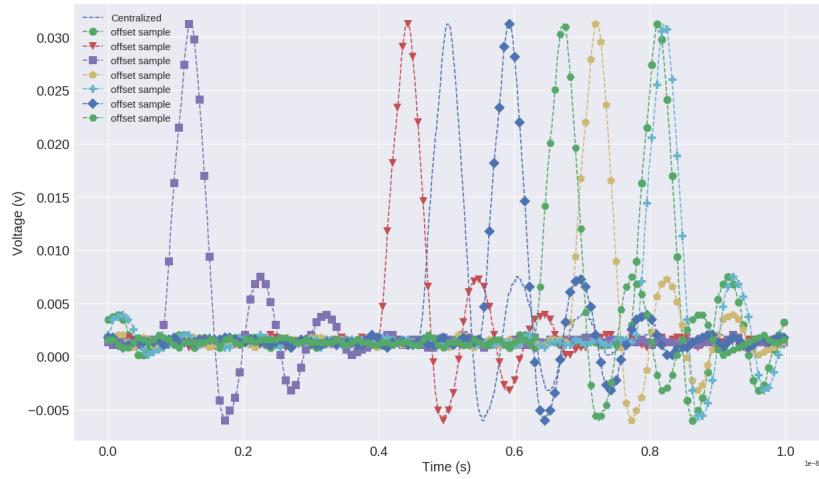


**Fig. 10:** One sample for each ion.



**Fig. 11:** Multiple samples for the same ion.

A first problem that must be faced is the unbalance of classes, which ends up hampering the obtaining of a classification model (24), to deal with this problem was created a method of data augmentation, which consists of generating synthetic data based on the data real. The process randomly selects a sample and promotes a horizontal displacement, based on the peak value, in 10% to 90% of the total signal window. This technique allows the generation of more data for the classes with a reduced number of samples besides promoting greater diversity in the signal, since the acquisition method centralizes the peak at the moment of triggering, which eliminates the bias of always having the maximum centralized value. Figure 12 shows the generation of some samples by the date augmentation method. As generated 41k samples (including the originals) for training purpose.



**Fig. 12:** Data augmentation for same ion sample.



## 2.4 Neural Network Architecture

There are several jobs that use DNNs for time series classification. The first approach taken was to implement the proposed neural network in (25), in this work if the performance of the fully convolutional network (FCN)(41) architecture is extended by extending its architecture with the use of Long Short Term Recurrent Neural Networks (LSTM RNN) modules named new LSTM-FCN module, or the use of LSTM RNN with attention, called ALSTM-FCN, the use of these two new approaches (LSTM-FCN and ALSTM-FCN) improved, according to the work, the classification performance in the datasets used to benchmark this type of task, which are provided by the University of California Riverside (UCR)(9). The use of this architecture was not efficient in this work due to difficulties in the training phase, since the partial results obtained with many training epochs did not show great advances compared to the results obtained by the baseline model. The second approach was inspired by the architectures used by the Mcfly(40) framework of the Netherlands eScience center (NLeSC) research group, that systematically generates neural networks based on two base architectures, CNN and DeepConvLSTM. From these two base architectures, the framework performs a search for the best hyperparameters that optimize the ranking result. In a preliminary test, with a portion of the data, the DeepConvLSTM architecture was selected because it presented better performance in a representative sample of the data, thus only the hyperparameters for this architecture will be detailed. The following parameters are optimized by the framework: the number of Conv layers ( $N$ ), number of LSTM Layers ( $M$ ), number of filters for each Conv Layer ( $\#ConvFilters$ ) and the size of each hidden layer of each LSTM layer ( $\#LSTMCells$ ).

The first layer of the proposed model is batch normalization, the technique that is applied in this step has the purpose of normalizing the distribution of the data that is received, allowing a speed gain in neural network training because it avoids the internal covariate shift that occurs, mainly in the inner layers of the neural network, and allows to increase the learning rate used by the network (23). The batch normalization method occurs in three steps: The first one calculates the average of each data batch, then the batch variance is calculated and the data distribution is scaled and shifted. The algorithm 1 describes the actions performed:

---

**Algorithm 1** Batch Normalizing Transform, applied to activation  $x$  over a mini-batch.  
Adapted from (23).

---

**Require:** Values of  $x$  over a mini-batch:  $B = \{x_{1\dots m}\}$ ;

**Ensure:**  $y_i = BN_{\gamma, \beta}(x_i)$

$$\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad \{\text{mini-batch mean}\}$$

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad \{\text{mini-batch variance}\}$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad \{\text{normalize}\}$$

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma, \beta}(x_i) \quad \{\text{scale and shift}\}$$


---

In the algorithm 1  $\gamma$  and  $\beta$  are parameters learned during the neural network training step.

Moving on to the next layers of our neural network we come to the first major convolutional block of our architecture, which consists of a convolutional layer, a batch normalization layer and finally a ReLU (Retified Linear Unit) activation layer. Convolutional networks are typically applied to image problems (2d) because they can learn patterns and look for them in various regions of an image, thus surpassing a fully connected conventional neural network (6). The power of convolutional neural networks lies in the sharing of trained weights, which decreases the number of weights to be trained and consequently makes training faster, its time sub-sampling and local receptive fields (which can be translated as the size of the convolution filters). The idea behind of a convolutional network is simple, the goal is to acquire a  $n$  number of filters, each filter being responsible for obtaining a characteristic of the input data, these characteristics, which are simple a priori, are combined in deeper layers into the network thus generating new more complex features and so on. The convolutional term comes from the fact that each filter traverses all the input data in search of this strained feature as a result of a map from which data pattern manifests itself in the data under analysis.

Another element that is part of the convolutional block is the ReLU which is responsible for transforming your entry or set or set of entries to a specific domain, which in this case is to zero entries that have negative values and replicate, in its output, positive values. This activation layer is defined in eq. 1, and is widely used because it reduces gradient fade problems, which occurs with sigmoid (16) activations, and has a reduced computational cost because it is just comparisons, summation, and multiplications.

$$f(x) = \max(0, x) \quad (1)$$

After the  $N$  convolutional blocks are composed and the blocks containing the LSTM cells are introduced into the structure. This type of architecture has applications in a number of fields of study because of its data sequence processing capabilities (18; 36; 28).

The basic concepts behind of LSTM is memory cells and port units, that were originally introduced in (21). The heart of the LSTM operation is based on the state of the cell, which is changed as the data continues through the gates, each gate being responsive by performing an operation on the main data stream of the cell. The first gate is responsible for the acceptance or not of the data that were proposed by the network exit. This is known as the “forget gate layer” and has as a decision function a sigmoid( $\sigma$ ), which has output values between 0 and 1, being the value 1 indicative of keeping the data forward and 0 representing the total removal of the data. This gate is governed by Eq.2:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2)$$

Where:

- $W_f$  = Weight matrix
- $h_{t-1}$  = Previous output bias vector
- $x_t$  = Input



- $b_f$  = Bias vector

Continuing in the state of the cell, the data now passes through the gate known as the “input gate layer” which is responsible for determining what information is to be stored in the state of the cell. First, the sigmoid layer determines which values will be updated, and then a tanh layer generates a vector of possible new candidate values,  $\tilde{C}_t$ , that will be added to the state. There are two equations in this part of the cell Eq.3 and Eq.4 that work together to derive the result from this section:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

$$\tilde{C} = \tanh(W_C[h_{t-1}, x_t] + b_C) \quad (4)$$

At this point, the old cell state  $C_{t-1}$  is updated to the current state  $C_t$ . The update process consists in multiplying the old state by  $f_t$ , such that the cell “forgets” the values previously stored from the decision, and finally add  $i_t * \tilde{C}_t$ , the result of these two operations presents the new candidates scaled by how much we decide to update each state value. In this way we have:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

Finally, the output value will be selected based on the state of the cell, which should still pass through a filtering process of the values. The first step of this filtering is to run a sigmoid function that delimits which parts of the cell state go to the output. This pre-selected value passes through a tanh function that will be responsible for mapping the cell state to a range between -1 and 1, such that only the selected values will be part of the final response. This layer is known as “selection gate layer”. Eq.6 and Eq.7 refer to the process of filtering and selecting output values.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t * \tanh(C_t) \quad (7)$$

The next element of the structure is known as Dropout. This cell is responsible for randomly eliminating a percentage of the elements contained in a given layer during the training phase. This type of procedure is used to avoid overfitting deep nets by generating variations of the same net, which during the validation stage benefits from this variation and obtaining a better average result (39).

The time distributed layer is used to assign each slice of the time series a layer, which in this case is a fully connected network.

The softmax layer applies Eq.8 to a vector of K real numbers, and normalizes them to a distribution probability that is proportional to the exponential value of the input numbers.

$$\sigma(Z)_i = \frac{e^{Z_i}}{\sum_{j=1}^K e^{Z_j}} \quad (8)$$

Finally, the Take Last layer is used to capture the last sequence response returned by the softmax layer as the final response.

Derived from this initial architecture, in this work, we propose some modifications in order to improve the model accuracy as well as the performance, so some modifications will be proposed. The first of these is the addition of Gated Recurrent Unit (GRU) cells, which have been proposed in (10) instead of LSTM cells, as they proved to be as efficient as LSTM in similar scenarios, but with better performance because they perform fewer mathematics operations within its (11). The equations governing the cell GRU are:

$$z_t = \sigma_g(W_z x_t + U_z h_{t-1} + b_z) \quad (9)$$

$$r_t = \sigma_g(W_r x_t + U_r h_{t-1} + b_r) \quad (10)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \sigma_h(W_h x_t + U_h(r_t \cdot h_{t-1}) + b_h) \quad (11)$$

Where:

- $x_t$  = input vector
- $h_t$  = output vector
- $z_t$  = update gate vector
- $r_t$  = reset gate vector
- $W$ ,  $U$  and  $b$  = parameter matrices and vector
- $\sigma_g$  = Activation function sigmoid
- $\sigma_h$  = Activation function hyperbolic tangent

The second proposal is to use bidirectional layers with both LSTM and GRU cells, this type of approach creates the same sequence of cells in parallel with the original but in the opposite direction of data flow, with the purpose of using future data in predicting the next values. This type of architecture has obtained satisfactory results in Natural Language processing applications (17).

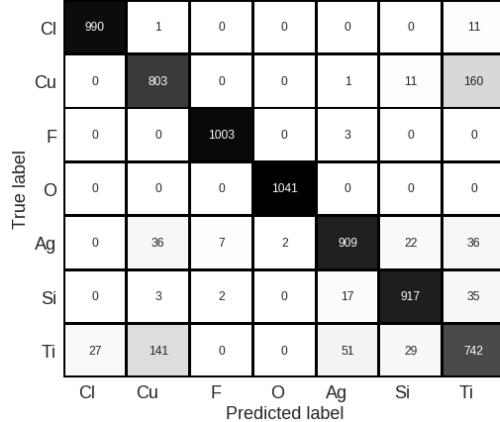
## 2.5 Training model

The training was performed with 80% of the total available data, and 10% of this portion was used to validate and monitor the evolution of the neural network. The remaining 20% was reserved for the model testing stage. An SVM classifier was obtained with the same data for the purpose of comparing the results obtained by the trained model. In total 100 epochs were performed in the training phase. To avoid stagnation of neural network training on a plateau, a loss monitoring function has been added to each epoch. This function evaluates whether there has been improvement in function loss by monitoring whether the loss of the neural network decreases, if there are no improvements the function reduces the learning rate of the neural network to the next epoch. This method allows the optimization function to be able to minimize the function to find the global minimum (7).

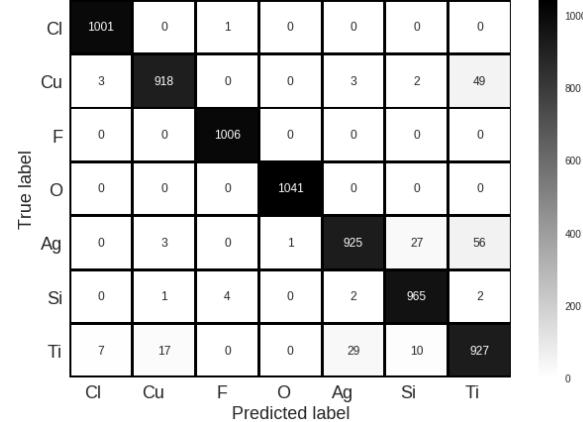


### 3 RESULTS AND DISCUSSION

In Fig. 13 the confusion matrix obtained by the SVM model in the test data is shown, showing the classification for each class.



**Fig. 13:** Confusion matrix for SVM baseline model on test data.



**Fig. 14:** Confusion matrix for neural network on test data.

The result obtained in Fig. 13 shows the classification for the test base. Although it presents satisfactory results for some classes of ions, like Oxygen and Carbon, it presents many errors of classification between the Copper (radii 135pm) and Titanium (radii 140pm), both metals of transition, that by presenting atomic rays similar effects, making classification difficult in these cases.

In Fig. 14 the resulting confusion matrix obtained for the DeepConvLSTM network after the training process.

The results obtained by the LSTM network for the data set, indicated in the Fig. 14, present better results when compared with the base model (SVM). It is possible to notice that the disambiguation between the Titanium and Copper classes, and also among other classes, showing better performance in relation to the base model.

**Table 3:** SUMMARY OF RESULTS FOR EACH MODEL IN DIFFERENT METRICS

Classes	SVM			DeepConvLSTM-NN			Samples
	Precision	Recall	F1-Score	Precision	Recall	F1-Score	
Cl	0.97	0.99	0.98	0.99	1.00	1.00	1002
Cu	0.82	0.82	0.82	0.98	0.95	0.96	975
F	0.99	1.00	0.99	1.00	1.00	1.00	1006
O	1.00	1.00	1.00	1.00	1.00	1.00	1041
Ag	0.93	0.90	0.91	0.96	0.92	0.94	1012
Si	0.94	0.94	0.94	0.96	0.99	0.97	974
Ti	0.75	0.75	0.75	0.91	0.93	0.92	990
Summary	0.91	0.91	0.91	0.97	0.97	0.97	7000

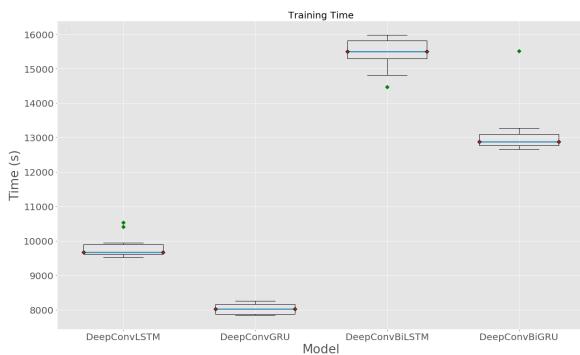
Other metrics (Precision, Recall and F1-Score) were used to evaluate, on a case-by-case basis, the performance between the two proposed models and are demonstrated in the Table 3. For the majority metrics used, the model obtained through the use of LSTM cells presented equal or higher results, bringing the importance of the temporal relation

of the input data, because there were gains of prediction performance in the model that uses this relation.

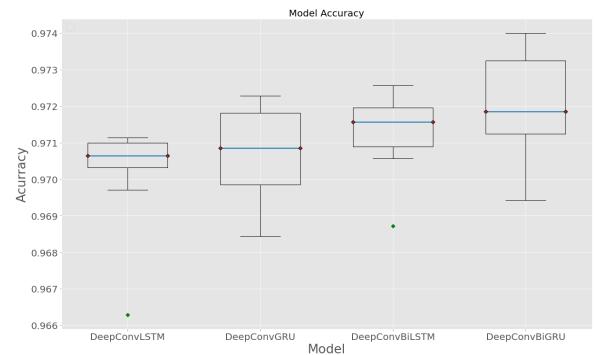
### 3.1 Proposed models

The variations proposed in this work that can be named DeepConvGru, DeepConvBiLSTM and DeepConvBiGRU were trained in 10 rounds with random weight initialization to avoid weight initialization bias, and each round featured 300 epochs. In fig. 15 shows the average training time of each model.

Looking at fig.15 it is possible to notice that the training times for architectures with GRU cells, both architectures with bidirectional structures present longer training time. Accuracies for each model were performed on the same test dataset and can be seen in fig.16.



**Fig. 15:** Training time for each model.



**Fig. 16:** Training time for each model.

Looking at the results obtained in fig.16, both GRU cell models are significantly better than their counterpart with LSTM cells, but these average differences are in the order of  $10^{-2}$ . The table 4 shows the ANOVA test results comparing the results obtained by each model in relation to DeepConvLSTM.

**Table 4:** ANOVA TEST RELATED TO DEEPCONVLSTM ACCURACY

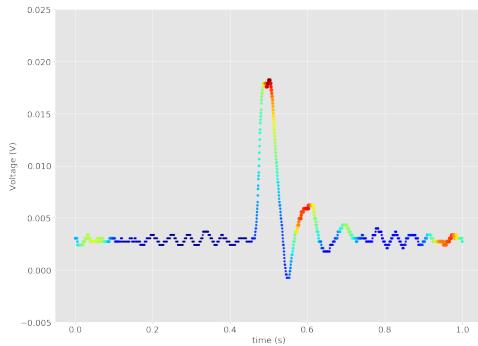
Model	F-statistic	p-value
DeepConvGRU	0.4600	0.5062
DeepConvBiLSTM	3.5233	0.0768
DeepConvBiGRU	8.6162	0.0088

Looking at the table 4 you can see that only the DeepConvBiGru model has a p-value less than 0.05, showing a significant improvement over the original model.

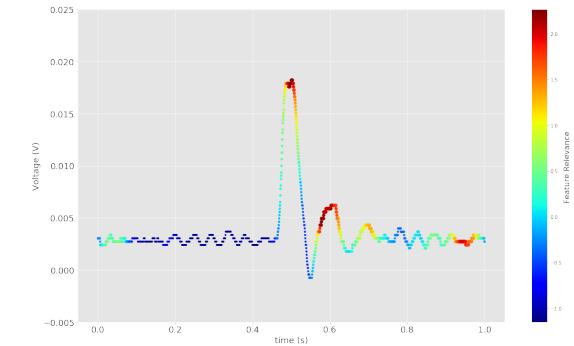
### 3.2 Model Interpretability

Model interpretability is important for understanding how the model makes decisions, and it is possible to gain insight into the problem being addressed. In Fig. 17 and Fig. 18 show how each model addresses a single signal.

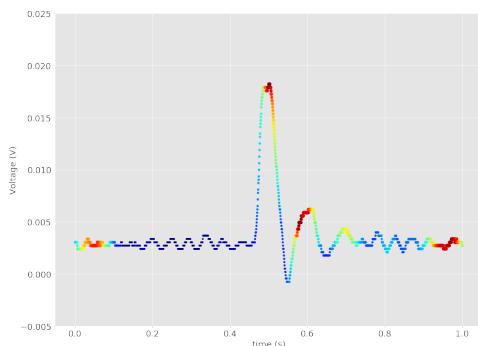
In Fig. 17 and Fig. 18 show the most important regions for correct signal classification. There are no major differences between the regions of attention of the model, but the



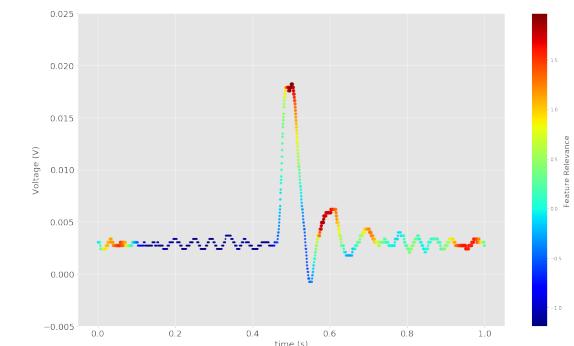
(a) DeepConvLSTM feature importance.



(b) DeepConvGRU feature importance.

**Fig. 17:** Feature importance without bidirectional.

(a) DeepConvBiLSTM feature importance.



(b) DeepConvBiGRU feature importance.

**Fig. 18:** Feature importance with bidirectional.

model with GRU cells consider the region after the signal peak as important, which can be explained by having less information retention mechanisms, but most of the noise of the signal. Signal is considered as regions not important for classification.

## 4 CONCLUSIONS

The proposed goal was successfully achieved, the neural network developed using LSTM cells provided better classification and disambiguation results between close classes compared to the previous model (SVM) used as baseline. The DeepConvGRU, DeepConvBiLSTM and DeepConvBiGRU networks showed approximate results from the original model, but the DeepConvBiGRU model showed significant improvements in the statistical test with p-value less than 0.05, and despite having a longer training time, this does not affect its application final use. With the models created it was possible to define levels of relevance for each region of the curve, which allows to have new ideas about the analysis of the phenomenon and its mechanisms. The next steps involve migration of this model to embedded systems, allowing the creation of low cost particle detectors or systems capable of dealing with the anomalies imposed by the event.

## ACKNOWLEDGEMENT

We are grateful to the CITAR Project, Centro Universitário da FEI and University of São Paulo for providing the facilities for the conduction of the experiments and data analysis. This publication is financially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## REFERENCES

- [1] Aad, G., Abbott, B., Abdallah, J., Khalek, S. A., Abdinov, O., Aben, R., Abi, B., Abolins, M., and AbouZeid, e. a. (2014). “A neural network clustering algorithm for the ATLAS silicon pixel detector”. *Journal of Instrumentation*, 9(9).
- [2] Aaltonen, T., Adelman, J., Álvarez González, B., Amerio, S., Amidei, D., Anastassov, A., Annovi, A., Antos, J., and Apollinari, e. a. (2010). “Search for the higgs boson using neural networks in events with missing energy and  $b$ -quark jets in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV”. *Phys. Rev. Lett.*, 104:141801.
- [3] Abazov, V. M., Abbott, B., Abolins, M., Acharya, B. S., Adams, M., Adams, T., Agelou, M., and Agram, e. a. (2005). “Measurement of  $\sigma(p\bar{p} \rightarrow z) \cdot \text{Br}(z \rightarrow \tau\tau)$  at  $\sqrt{s} = 1.96$  TeV”. *Phys. Rev. D*, 71:072004.
- [4] Abramowicz, H., Caldwell, A., and Sinkus, R. (1995). “Neural network based electron identification in the zeus calorimeter”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 365(2-3):508–517.
- [5] Aguiar, V., Medina, N., Added, N., Macchione, E., Aguirre, F., Ribas, R., Nascimento, S., Escudeiro, R., Allegro, P., Perego, C., Fagundes, L., Duarte, J. G., Scarduelli, V., Morais, O. B., Almeida, E. A., Joaquim, P. M., Souza, M., Cecotte, A., Martins, R., Brage, J., Leistenschneider, E., Oliveira, R., Assis, R., Leite, A., Terassi, J., Abreu, J., Simoes, R., Joaquim, A., Servelo, W. A., Silva, S. C., Minas, J., Silva, M., Silva, V., Kshinskiy, D., and Silveira, M. (2017). “New Setup for SEE Measurements in South America”. *RADECS 2017*.
- [6] Bengio, Y. (1997). “Convolutional Networks for Images, Speech, and Time-Series Parsing View project MoDeep View project”. Technical report n°.
- [7] Bengio, Y. *Practical recommendations for gradient-based training of deep architectures*. title.
- [8] Casadei, D., Aracena, I., Banerjee, S., Beauchemin, P.-H., Calvet, S., Cranmer, K., Damazio, D., Djilkibaev, R., Eifert, T., Hillier, S., Idarraga, J., Johns, K., Kaushik, V., Konoplich, R., Kowalewski, R. V., Lei, X., Mann, A., Mermod, P., Mincer, A., Morel, J., Nemethy, P., Pinder, A., Taylor, R. P., Watson, A., and Zhao, L. (2011). “ATLAS NOTE The implementation of the ATLAS missing E T triggers for the initial LHC operation”. Technical report n°.
- [9] Chen, Y., Keogh, E., Hu, B., Begum, N., Bagnall, A., Mueen, A., and Batista, G. *The UCR Time Series Classification Archive*. title. [www.cs.ucr.edu/~eamonn/time\\_series\\_data/](http://www.cs.ucr.edu/~eamonn/time_series_data/).
- [10] Cho, K., Van Merriënboer, B., Bahdanau, D., and Bengio, Y. (2014). “On the properties of neural machine translation: Encoder-decoder approaches”. *arXiv preprint arXiv:1409.1259*.
- [11] Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). “Empirical evaluation of gated recurrent neural networks on sequence modeling”. *arXiv preprint arXiv:1412.3555*.
- [12] de Oliveira, J. A., de Melo, M. A. A., da Silveira, M. A. G., and Medina, N. H. (2014). “Electronic system for data acquisition to study radiation effects on operating MOSFET transistors”. *AIP Conference Proceedings - RTFNB 2013*, 1265:130–134.
- [13] Denby, B. H. (1988). “Neural Networks and Cellular Automata in Experimental High-energy Physics”. *Comput. Phys. Commun.*, 49:429–448.
- [14] Ding, L., Gerardin, S., Bagatin, M., Bisello, D., Mattiazzo, S., and Paccagnella, A. (2016). “Radiation tolerance study of a commercial 65 nm CMOS technology for high energy physics applications”. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 831:265–268.
- [15] Edwards, R., Dyer, C., and Normand, E. (2004). “Technical standard for atmospheric radiation single event effects,(see) on avionics electronics”. IEEE, pages 1–5.
- [16] Glorot, X., Bordes, A., and Bengio, Y. (2011). “Deep Sparse Rectifier Neural Networks”. Technical report n°.
- [17] Graves, A., Jaitly, N., and Mohamed, A.-r. (2013). “Hybrid speech recognition with deep bidirectional lstm”. IEEE, pages 273–278.
- [18] Graves, A., Liwicki, M., Fernández, S., Bertolami, R., Bunke, H., and Schmidhuber, J. “A Novel Connectionist System for Unconstrained Handwriting Recognition”.
- [19] Guest, D., Cranmer, K., and Whiteson, D. (2018). “Deep Learning and its Application to LHC Physics”.
- [20] Gutiérrez, O., Prieto, M., Sánchez-Reyes, A., and Gómez, A. (2019). “TID characterization of COTS parts using radiotherapy linear accelerators”. *IEICE Electronics Express*, 16(7):20190077–20190077.
- [21] Hochreiter, S. and Schmidhuber, J. (1997). “Lstm can solve hard long time lag problems”. pages 473–479.
- [22] Hodson, R. F., Morgan, D., Ladbury, R. L., Chen, Y., Bay, M., and Zinchuk, J. (2019). “Radiation single event effects (see) impact on complex avionics architecture reliability”.
- [23] Ioffe, S. and Szegedy, C. (2015). “Batch normalization: Accelerating deep network training by reducing internal covariate shift”. In: *32nd International Conference on Machine Learning, ICML 2015*, pages 448–456.
- [24] Iosifidis, V. and Ntoutsi, E. (2018). “Dealing with bias via data augmentation in supervised learning scenarios”.
- [25] Karim, F., Majumdar, S., Darabi, H., and Chen, S. (2018). “Lstm fully convolutional networks for time series classification”. *IEEE Access*, 6:1662–1669.
- [26] Khachatryan, V., Sirunyan, A. M., Tumasyan, A., Adam, W., Bergauer, T., Dragicevic, M., Erö, J., Fabjan, C., and Friedl, e. a. (2014). “Observation of the diphoton decay of the Higgs boson and measurement of its properties”. *European Physical Journal C*, 74(10):1–49.
- [27] Kohne, J. K. et al. (1997). “Realization of a second level neural network trigger for the H1 experiment at HERA”. *Nucl. Instrum. Meth.*, A389:128–133.
- [28] Li, X. and Wu, X. (2015). “CONSTRUCTING LONG SHORT-TERM MEMORY BASED DEEP RECURRENT NEURAL NETWORKS FOR LARGE VOCABULARY SPEECH RECOGNITION”. Technical report n°.



- [29] Mager, M. (2016). "ALPIDE, the Monolithic Active Pixel Sensor for the ALICE ITS upgrade".
- [30] Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2018). "Statistical and Machine Learning forecasting methods: Concerns and ways forward". *PLoS ONE*, 13(3):1–26.
- [31] Medina, N. H., Silveira, M. A. G., Added, N., Aguiar, V. A. P., Aguirre, F., Giacomini, R., Macchione, E. L. A., de Melo, M. A. A., Oliveira, J. A., Santos, R. B. B., Seixas, L. E., and Tabacniks, M. H. (2013). "Brazilian facilities to study radiation effects in electronic devices". *2013 14th European Conference on Radiation and Its Effects on Components and Systems (RADECS)*, pages 1–7.
- [32] Medina, N. H., Silveira, M. A. G., Added, N., Aguiar, V. A. P., Giacomini, R., Macchione, E. L. A., De Melo, M. A. A., Santos, R. B. B., and Seixas, L. E. (2015). "First successful SEE measurements with heavy ions in Brazil". *IEEE Radiation Effects Data Workshop*, 2015-Janua(January):1–3.
- [33] Morel, F., Hu-Guo, C., Bertolone, G., Claus, G., Colledani, C., Dorokhov, A., Dozì, G., Dulinski, W., Fang, X., Goffe, M., Himmi, A., Jaaskelainen, K., Senyukov, S., Specht, M., Szelezniak, M., Pham, H., Valin, I., Wang, T., and Winter, M. (2013). "Recent citations MISTRAL & ASTRAL: two CMOS Pixel Sensor architectures suited to the Inner Tracking System of the ALICE experiment". pages 23–27.
- [34] Normand, E. (1996). "Single-event effects in avionics". *IEEE Transactions on nuclear science*, 43(2):461–474.
- [35] Ordóñez, F. J. and Roggen, D. (2016). "Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition". *Sensors*, 16(1).
- [36] Sak, H. H., Senior, A., and Google, B. "Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling". Technical report n°.
- [37] Sepp, H. and J?urgen, S. (1997). "LONG SHORT-TERM MEMORY". *Neural Computation*, 9(8):1–32.
- [38] Silveira, M. A. G., Melo, M. A. A., Aguiar, V. A. P., Rallo, A., Santos, R. B. B., Medina, N. H., Added, N., Seixas, L. E., Leite, F. G., Cunha, F. G., Cirne, K. H., Giacomini, R., and de Oliveira, J. A. (2015). "A Commercial off-the-shelf pMOS Transistor as X-ray and Heavy Ion Detector". *Journal of Physics: Conference Series*, 630:012012.
- [39] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). "Dropout: a simple way to prevent neural networks from overfitting". *The journal of machine learning research*, 15(1):1929–1958.
- [40] van Kuppevelt, D., Meijer, C., van Hees, V., Bos, P., Spaaks, J., Kuzak, M., Huber, F., Hidding, J., and van der Ploeg, A. *mcfly: deep learning for time series*. title.
- [41] Wang, Z., Yan, W., and Oates, T. (2017). "Time series classification from scratch with deep neural networks: A strong baseline". pages 1578–1585.



# Análise de Características na Tarefa de Segmentação Interativa de Imagens Digitais

Thiago Amorim de Faccio<sup>1</sup>, Marilaine Colnago<sup>2</sup> e Wallace Casaca<sup>2</sup>

<sup>1</sup> UNESP - Faculdade de Ciências e Tecnologia, Presidente Prudente-SP, Brasil

<sup>2</sup> UNESP - Câmpus de Rosana, Rosana-SP, Brasil

---

## Abstract

Este trabalho tem como principal finalidade otimizar a tarefa de segmentação de imagens através da análise de *features*, que são fornecidas como dados de entrada no cálculo da matriz de pesos associada à segmentação via Coordenadas de Laplace. A ideia é analisar diferentes combinações de *features* (bandas) que permitam realçar os objetos e as bordas de uma imagem, sendo que essas bandas são resultados da aplicação de técnicas como o modelo de Mistura Gaussiana, o algoritmo *Expectation-Maximization*, o detector de bordas *Canny*, as funções de distância local e, finalmente, a conversão de modelos de cores. Através de métricas do estado-da-arte em segmentação e da análise qualitativa dos resultados obtidos, foi possível constatar que o estudo apresentou um ganho considerável com relação às *features* clássicas tradicionalmente utilizadas no algoritmo de segmentação das Coordenadas de Laplace.

**Keywords:** Segmentação de Imagens, Análise de *Features*, Processamento Digital de Imagens

---

## 1 INTRODUÇÃO

A data de seus primeiros usos no início do Século XX, técnicas de Processamento Digital de Imagens (PDI) evoluíram passando pelo surgimento dos primeiros grandes computadores até a era digital, e se tornando cada vez mais versáteis e comuns no dia-a-dia de pessoas em todo o mundo.

Hoje, programas com grande capacidade de tratamento de imagens estão cada vez mais próximos dos usuários, em computadores pessoais e celulares, tendo ainda amplo protagonismo nas redes sociais por meio da utilização de filtros e *apps* de edição de fotografias. Entretanto, não é só para uso em meios de disseminação da informação que se sustentam as pesquisas baseadas em PDI. Tal campo de estudo pode oferecer soluções para diversas outras tarefas, em áreas como, por exemplo, da medicina, auxiliando no

Contato: Faccio, thiagodefaccio@gmail.com

estudo detalhado de raios-X, ou ainda, na contagem de glóbulos brancos em amostras de sangue. Também, em áreas como da segurança e monitoramento, há inúmeras aplicações como na análise de imagens de satélites e até mesmo em frentes como geologia, arqueologia, entre outras mais.

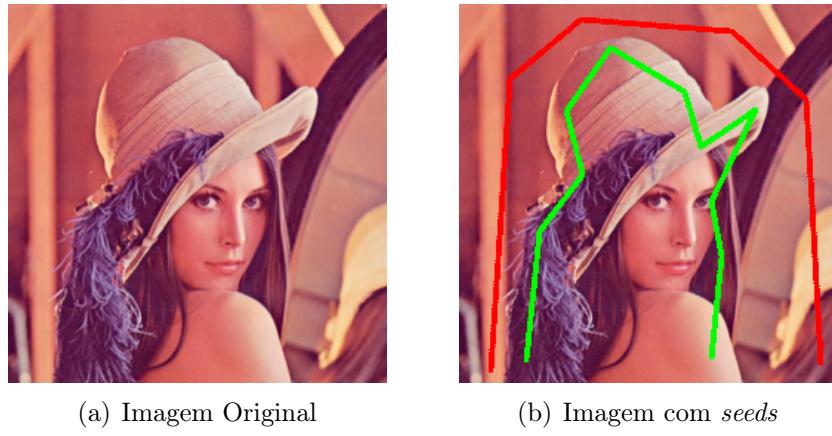
No caso da aplicação de segmentação de imagens, esta busca promover a subdivisão da imagem em agrupamentos (*clusters*) de pixels, que se relacionam através de diferentes características, especialmente a intensidade e o gradiente, com a intenção de aperfeiçoar o reconhecimento da informação visual na imagem para a interpretação tanto por um agente humano quanto para fins de aprendizado de máquina.

Desta forma, visando subsidiar o tratamento de imagens relacionadas a programas e aplicativos de edição de fotografias, este trabalho tem como propósito melhorar a precisão na tarefa de segmentação interativa, isto é, na modalidade supervisionada de segmentação em que *seeds* (*brushes/pinceladas*) devem ser fornecidas pelo usuário a fim de orientar o segmentador de imagens. Assim, neste trabalho, foram utilizados diferentes filtros de PDI, em que cada filtro produz uma *feature*, que é fornecida como um parâmetro para o cálculo de uma matriz de pesos, utilizada no processo de segmentação da imagem.

## 2 METODOLOGIA

### 2.1 Segmentação via Coordenadas de Laplace

Utilizando uma técnica chamada de Coordenadas de Laplace, Casaca et al. (2014) desenvolveram um algoritmo de segmentação que utiliza o conceito de *seeds/marcas*, isto é, informações de entrada dadas pelo usuário de maneira a definir qual objeto deverá ser segmentado na imagem (vide a Figura 1 para uma ilustração). Tal abordagem atua na imagem através da execução dos seguintes passos: (i) definição das *seeds* por parte do usuário, (ii) construção de um grafo de afinidade, (iii) formulação e minimização do funcional de energia das Coordenadas de Laplace, e (iv) obtenção da segmentação (caracterizada pelos pixels do objeto e do fundo da imagem).



**Fig. 1:** Etapa do provimento das *seeds/marcas* fornecidas pelo usuário.

Conforme mencionado acima, a atribuição das *seeds* à imagem é uma tarefa que deve ser realizada pelo usuário, já que a resolução do funcional de energia na etapa (iii) do método depende desse tipo de informação. No caso da construção do grafo de afinidade



$G = G(V, A, W)$ , este é obtido a partir de um *stencil* (de conectividade 8), em que cada nó de  $G$  é representado por um pixel  $p_i$  da imagem (aqui representada por  $I$ ). Já as *features* são dadas como entrada para a matriz de pesos  $W$ , que armazena os pesos das arestas do referido grafo. Os elementos de  $W$  são definidos da seguinte forma:

$$w_{ij}^k = \exp\left(-\frac{\beta \|I_i^k - I_j^k\|_\infty^2}{\sigma}\right), \quad \sigma = \max_{(i,j) \in A} \|I_i^k - I_j^k\|_\infty, \quad (1)$$

em que  $k$  representa a  $k$ -ésima *feature* computada,  $A$  é o conjunto de arestas do grafo base, e  $\beta$  é um escalar que representa um parâmetro de ajuste.

Sejam as informações fornecidas pelo usuário (vide Figura 1(b) para um *insight* visual), isto é, o conjunto de índices  $O = \{i : p_i \in I$  rotulado como objeto} e  $F = \{i : p_i \in I$  rotulado como fundo}. Assim, a partir dos valores da matriz de pesos  $W$  (Eq. (1)), podemos definir o funcional de energia das Coordenadas de Laplace (Casaca et al., 2014):

$$E(x) = c_1 \sum_{i \in F} \|x_i - x_F\|_2^2 + c_2 \sum_{i \in O} \|x_i - x_O\|_2^2 + c_3 \sum_{i \in V} \left\| d_i x_i - \sum_{j \in N(i)} w_{ij} x_j \right\|_2^2 \quad (2)$$

em que  $V$  representa o conjunto de pixels da imagem  $I$ ,  $d_i$  é a valência associada ao pixel  $p_i \in I$ ,  $N(i)$  corresponde à vizinhança unitária de  $p_i$ , e  $c_1$ ,  $c_2$  e  $c_3$  são constantes positivas.

A solução de  $E(x)$  retorna o vetor real  $x = (x_i)$  em que cada  $x_i$  está associado a um pixel  $p_i$  da imagem, gerando assim um mapa escalar para a imagem. Finalmente,  $x_O$  e  $x_F$  são valores reais já pré-definidos no funcional de energia, onde deve-se impor  $x_O \neq x_F$ . Tais valores correspondem aos pixels que foram rotulados através dos conjuntos  $O$  e  $F$ . Por exemplo, se assumirmos  $x_O \geq x_F$ , podemos então atribuir os rótulos  $y_i \in \{x_O, x_F\}$  a cada pixel  $p_i$  da imagem tal que:

$$y_i = \begin{cases} x_O, & \text{se } x_i \geq \frac{x_O + x_F}{2} \\ x_F, & \text{caso contrário} \end{cases} \quad (3)$$

obtendo, assim, a segmentação definitiva da imagem a partir do mapa escalar  $x = (x_i)$ , e dos valores de referência  $x_O$  e  $x_F$ .

Em nosso trabalho, a ideia é utilizar a técnica de segmentação das Coordenadas de Laplace como protótipo de teste, de forma a possibilitar melhorias na segmentação através da adição de novas *features* na matriz de pesos  $W$ . Esse tipo de ajuste reflete diretamente na solução do funcional  $E(x)$ , visto que o mesmo emprega os pesos  $w_{ij}$  em sua composição. Assim, estamos interessados em modificar a Eq. (1) a partir do uso de novas bandas  $I^k$ .

Na sequência, são descritas as novas bandas (*features*) que foram analisadas e incorporadas na composição dos pesos  $w_{ij}$ .

## 2.2 Modelo de Mistura Gaussiana

Modelos de misturas são formados pelas componentes de mistura e seus respectivos pesos. Essas componentes são dadas pelas distribuições de probabilidade, denotadas aqui por  $D_i$ ,  $i = 1, 2, \dots$ , e os pesos são as probabilidades  $\pi_i$ , ligadas a essas distribuições. A densidade resultante do modelo é dada pela combinação ponderada das densidades de cada

uma das componentes, tendo como fator de ponderação da combinação a probabilidade  $\pi_i$  relacionada a  $D_i$ , de forma que:

$$\sum_{i=1}^K \pi_i = 1, \quad 0 \leq \pi_i \leq 1, \quad (4)$$

em que  $K$  indica o número de componentes do modelo, e o conjunto  $\{\pi_i, i = 1, 2, \dots, K\}$  pode ser também interpretado como um conjunto de probabilidades, *a priori*. O agrupamento recorrente dos modelos de misturas pode ser categorizado como um algoritmo de otimização iterativo, no qual o objetivo principal é a maximização do critério de agrupamento visando otimizar a função de verossimilhança (Bishop, 2006). Para o caso de modelos de misturas baseados na combinação linear de gaussianas, podemos descrever os grupos pela média  $\mu_i$  e pela matriz de covariância  $C_i$ , e assim escrever a função de densidade de probabilidade da seguinte forma:

$$p(x) = \sum_{i=1}^K \pi_i N_i(x|\mu_i, C_i), \quad (5)$$

em que  $N_i(x|\mu_i, C_i)$  denota a distribuição normal relacionada ao vetor  $\mu_i$  e a matriz de covariância  $C_i$ , como visto em (Fu and Wang, 2012). Neste caso, o problema é que os valores dos parâmetros  $\pi_i$ ,  $\mu_i$  e  $C_i$  não são conhecidos de tal forma que, para problemas de agrupamento (*clustering*), estes valores precisam ser definidos, no entanto, existem alguns métodos para encontrá-los. Como o objetivo é a maximização da função de verossimilhança, neste trabalho, empregamos o algoritmo EM para determinar os valores desconhecidos.

### 2.3 Algoritmo EM

O Algoritmo *Expectation Maximization* (EM), introduzido por (Dempster et al., 1977) vem sendo estudado e aprimorado desde então, e tem como principal função o cálculo iterativo de estimativas de máxima verossimilhança, mais especificamente: a *Maximum Likelihood Estimate* (MLE) (McLachlan and Krishnan, 2008).

O algoritmo EM funciona em dois passos (McLachlan and Peel, 2000). A primeira etapa pode ser chamada de passo E, *expectation*, em que nessa etapa é definido um conjunto de dados completo, o qual é baseado nas informações observadas para calcular o valor esperado da função de máxima verossimilhança. Em seguida, tem-se a etapa M, *Maximization*, que usa os dados obtidos da função de verossimilhança para definir os novos valores para os parâmetros desconhecidos. As etapas E e M se repetem até que se atinja a convergência do método iterativo.

Caso soubéssemos todos os dados  $x_k$  de um Modelo de Mistura Gaussiana com  $M$  componentes Gaussianas, então seria mais simples calcular os parâmetros desse modelo, representado por  $(\pi_1, \pi_2, \dots, \pi_M, \theta_1, \theta_2, \dots, \theta_M)$ , onde  $\theta_i = (\mu_i, C_i)$  são as médias e as matrizes de covariância. Ou ainda, se os parâmetros  $(\pi_1, \pi_2, \dots, \pi_M, \theta_1, \theta_2, \dots, \theta_M)$  fossem conhecidos, então também seria possível gerar os valores dos dados  $x_k$ . Como nenhum desses valores são conhecidos em situações práticas, tem-se então como saída estimar tais parâmetros utilizando o Algoritmo EM, que como mencionado anteriormente, estima a máxima verossimilhança mesmo com dados incompletos.



Sendo assim, a etapa E de um Modelo de Mistura Gaussiana com  $M$  componentes Gaussianas e com  $N$  dados  $x_k$  pode ser calculada da seguinte forma:

$$Q(\theta|\theta^{(t)}) = \sum_{l=1}^M \sum_{i=1}^N \gamma(Z_{il}) \log(\pi_l) + \sum_{l=1}^M \sum_{i=1}^N \gamma(Z_{il}) \log(N(x_i|\theta_l)), \quad (6)$$

em que  $N(x_i|\theta_l)$  representa a probabilidade da distribuição para a  $l$ -ésima componente Gaussiana de  $x_i$ , e  $\gamma(Z_{il})$  é a probabilidade a posteriori da  $l$ -ésima componente Gaussiana, que segundo (Bishop, 2006), é representada por:

$$\gamma(Z_{il}) = \frac{\pi_l N(x_i|\theta_l)}{\sum_{j=1}^M \pi_j N(x_i|\theta_j)}. \quad (7)$$

Após a realização dos passos da etapa E, os dados obtidos são então usados na reestimativa dos parâmetros da próxima iteração pela etapa M, que pode ser resumidamente descrita pelos seguintes passos:

$$\pi_l^{(t+1)} = \frac{1}{N} \sum_{i=1}^N \gamma(Z_{il}) \quad (8)$$

$$\mu_l^{(t+1)} = \frac{1}{N \pi_l^{(t+1)}} \sum_{i=1}^N x_i \gamma(Z_{il}) \quad (9)$$

$$C_l^{(t+1)} = \frac{1}{N \pi_l^{(t+1)}} \sum_{i=1}^N (x_i - \mu_l^{(t+1)})^2 \gamma(Z_{il}). \quad (10)$$

Finalmente, quando a etapa M é concluída, os parâmetros reestimados são usados na próxima iteração da etapa E, e assim as etapas *Expectation* e *Maximization* se alternam até a convergência.

#### 2.4 Conversão do Sistema de Cores RGB para CIEL\*a\*b\*

A conversão entre estes dois modelos de cores é intermediado pelo modelo XYZ, e é dividida em duas etapas. Na primeira etapa, deve ser realizada a conversão do RGB para o XYZ da seguinte maneira:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (11)$$

No segundo passo, deve ser feita a conversão do modelo XYZ para o modelo L\*a\*b\*, e neste caso, os valores de L\*, a\* e b\* podem ser definidos por funções, através das seguintes equações:

$$L^* = 116 \cdot f\left(\frac{Y}{Y_n}\right) - 16 \quad (12)$$

$$a^* = 500 \cdot \left( f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right) \quad (13)$$

$$b^* = 200 \cdot \left( f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right) \quad (14)$$

onde a função  $f$  é definida como segue:

$$f(t) = \begin{cases} \sqrt[3]{t}, & \text{se } t \geq \delta^3 \\ \frac{t}{3\delta^2} + \frac{4}{29}, & \text{caso contrário} \end{cases} \quad (15)$$

em que os valores de  $X_n$ ,  $Y_n$  e  $Z_n$  são referentes ao ponto de cor branca do modelo, que muda de acordo com variações do modelo XYZ (Gonzalez and Woods, 2017).

## 2.5 Distância Local de uma Imagem

Seja  $I$  uma imagem e  $p_i$  um pixel de  $I$ . Considere também  $N_r(p_i)$  a vizinhança  $r \times r$  em torno do pixel  $p_i$ , sendo  $r$  um número ímpar, e seja  $x_i$ , com  $i = 1, \dots, n$  ( $n \leq r^2$ ), os elementos dessa vizinhança. Então, a função de distância local da imagem  $I$ , também conhecida por *rangefilt* na *plataforma MATLAB*, atribui ao pixel da matriz de saída da função, o valor resultante da seguinte relação:

$$p_i = \max_{1 \leq i \leq n}(x_i) - \min_{1 \leq i \leq n}(x_i) \quad (16)$$

Ou seja, para cada elemento da matriz  $I$ , é avaliada a maior distância entre dois pixels de sua vizinhança, e este valor é então atribuído ao elemento que está sendo analisado.

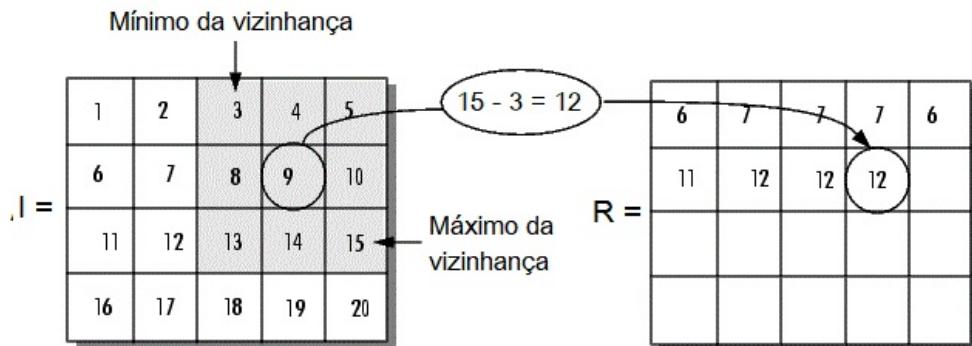


Fig. 2: Representação do processo da função *rangefilt* descrito pela Eq. (16).

## 2.6 Detector de Bordas Canny

Esta técnica de detecção de bordas, desenvolvida por (Canny, 1986), foi concebida com a intenção de melhorar algumas falhas em técnicas anteriores como, por exemplo, diminuir a taxa de erro na detecção de bordas, fazer com que os pontos considerados como borda estivessem bem localizados, reduzindo assim ruídos causados quando mais de um ponto em uma região era considerado uma borda.

Encontrar uma solução que satisfaça todas as necessidades supracitadas não é uma tarefa trivial, mas Canny concluiu que uma boa aproximação poderia ser calculada pela derivada de uma Gaussiana (Gonzalez and Woods, 2017) da seguinte forma:

$$\frac{d}{dx} \left( e^{-\frac{x^2}{2\sigma^2}} \right) = -\frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \quad (17)$$



Se considerarmos a imagem a ter as bordas extraídas por este algoritmo como  $f(x, y)$ , e a função Gaussiana como  $G(x, y)$ , teríamos então a função a ser derivada do lado esquerdo da Eq. (17) reescrita da seguinte forma:

$$G(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (18)$$

Nesse sentido, foi desenvolvido um algoritmo *multi-stage*, que tem seus passos expressos de formas variadas por diferentes autores da literatura. No entanto, uma forma usualmente adotada na literatura para descrever os passos é a seguinte:

- **Primeiro Passo - Suavização:** como forma de diminuir os ruídos da imagem, é aplicada uma suavização por convolução, ou ainda, o *Gaussian Filter* (Gonzalez and Woods, 2017), ou *Gaussian Blur* (Gedraite and Hadad, 2011).
- **Segundo Passo - Gradiente:** este é calculado através de um filtro o qual retorna a primeira derivada de  $G(x, y)$  quando aplicado na horizontal ( $G_x$ ) e na vertical ( $G_y$ ). Cada um desses filtros gera uma imagem, e o gradiente  $G$ , sempre perpendicular à borda, pode ser expresso da seguinte forma:

$$G = \sqrt{G_x^2 + G_y^2} \quad (19)$$

- **Terceiro Passo - Supressão de não-Máximos:** Após o passo anterior, a imagem resultante terá bordas grossas, entretanto, o objetivo é que essas bordas sejam finas e bem definidas. Portanto, deve-se realizar uma supressão por máximo local, onde cada ponto de borda será avaliado com outros dois do gradiente e apenas os pontos avaliados que representarem o máximo dos três serão mantidos.
- **Quarto Passo - Threshold:** Aqui é preciso determinar pelo menos dois valores de *Threshold*, um valor mínimo  $a$ , e um valor máximo  $b$ , com  $a, b \in (0, 1)$ . Uma escolha adequada para os valores de  $a$  e  $b$  é importante, e esta pode ser computada pelo clássico método de Otsu (1979).
- **Quinto Passo - Hysteresis:** Neste passo, todos os pontos remanescentes dos testes anteriores são definidos como elementos pertencentes a borda, ou não. Os pixels com valores maiores do que  $b$  são considerados elementos da borda; contrariamente aqueles que possuem valor menor do que  $a$  são descartados, e os que estão entre  $a$  e  $b$  devem ser avaliados com relações a suas conexões.

### 3 RESULTADOS E DISCUSSÃO

#### 3.1 Resultados Quantitativos

Os resultados dos testes foram calculados com base em 5 métricas de qualidade: *Rand Index* (RI) Rand (1971), *Global Consistency Error* (GCE) Iancu et al. (2011), *Variation of Information* (VI) Benes and Zitova (2014), *Boundary Displacement Error* (BDE) Unnikrishnan et al. (2005), e *Dice Coefficient* (DICE) Benes and Zitova (2014). Além disso, os tempos de processamento também foram relacionados. A fim de transmitir uma melhor compreensão dos valores das métricas, aquelas que apresentam resultados mais acurados

quando são ascendentes foram representadas pelo símbolo (+) e, por outro lado, aquelas que apresentam melhores resultados quando forem decrescentes estão indicadas com (-). Todos os cálculos das métricas foram realizados em uma base de 50 imagens coloridas, que foram avaliadas de acordo com a análise das diferentes combinações de *features* utilizadas para determinar os pesos  $w_{ij}$  do grafo base  $G$ , referente à cada imagem dessa base.

A base de dados empregada em nosso estudo foi extraída de (Rother et al., 2004), sendo que ela apresenta imagens obtidas em diferentes contextos como fotos de pessoas, animais, paisagens, objetos, entre outros elementos. O objetivo é então aplicar o método de segmentação das Coordenadas de Laplace para diferentes configurações de *features*, e testar seus resultados para cada um desses cenários.

**Teste 1:** O primeiro teste baseou-se nos resultados inicialmente obtidos no trabalho de (Casaca, 2015), sendo esses os valores de referência para realizar as demais comparações.

**TABELA 1:** RESULTADOS INICIAIS, EM QUE  $r$  É O PARÂMETRO DO *rangefilt*.

Métricas		$r = 3$	$r = 5$
RI	(+)	0.9048	0.9257
GCE	(-)	0.0765	0.0636
VI	(-)	0.4488	0.3801
BDE	(-)	8.5205	6.6576
DICE	(+)	0.9222	0.9394
Tempo (s)		166	170

Os próximos 4 testes foram realizados com diferentes composições de *features* na construção da matriz de pesos  $W$ , em que foram utilizadas todas as técnicas extratoras de bandas mencionadas na seção anterior. No caso, estas foram representadas por Comp (a), Comp (b), Comp (c) e Comp (d), em que:

- **Comp (a):** combinação das bandas da imagem em formato  $L^*a^*b^*$  com uma média das bandas de uma imagem resultante da segmentação pelo Algoritmo EM.
- **Comp (b):** combinação a qual conta com duas novas *features* com relação a Comp (a), que são resultantes do tratamento de uma imagem através do detector de bordas *Canny*. Foram extraídas as três bordas da imagem no formato *RGB*, e também do formato  $L^*a^*b^*$ . Na sequência, foi realizada as médias das bandas de cada formato.
- **Comp (c):** além da utilização de todas as *features* da Comp (b), embasado na ideia de criar novas amostras de entradas a partir da média entre outras, foi então criada uma nova *feature*, resultante da média das duas que foram utilizadas na composição anterior.
- **Comp (d):** por fim, como a função *rangefilt* tem se mostrado uma boa aliada na melhoria dos resultados dos testes realizados neste trabalho, o teste da Comp (c) foi novamente realizado, no entanto, a partir da aplicação da função *rangefilt* com parâmetro  $r = 3$  na *feature*, que só foi introduzida na Comp (c).

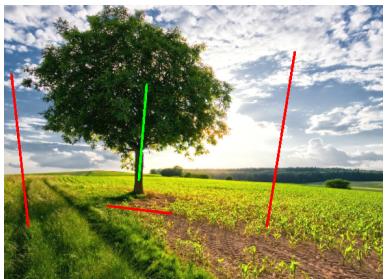
**TABELA 2:** RESULTADOS INTRODUZINDO NOVAS COMPOSIÇÕES DE BANDAS.

Métricas		Inicial	(a)	(b)	(c)	(d)
RI	(+)	0.9257	0.9216	0.9352	0.9357	0.9386
GCE	(-)	0.0636	0.0663	0.0575	0.0570	0.0550
VI	(-)	0.3801	0.3979	0.3492	0.3479	0.3356
BDE	(-)	6.6576	6.5259	5.6398	5.6731	5.2602
DICE	(+)	0.9394	0.9364	0.9479	0.9480	0.9508
Tempo (s)		170	190	285	319	327

Ainda que na Coluna (a) da Tabela 2 os *scores* não foram melhores que os resultados iniciais obtidos por (Casaca, 2015), nos demais testes, reportados a partir da Coluna (b), é possível observar que houve uma melhora significativa nos *scores* das métricas utilizadas apesar do aumento no tempo de processamento.

### 3.2 Resultados Qualitativos

Após o estudo e implementação das novas *features* ao código original, a segmentação apresentou resultados mais satisfatórios do que anteriormente, tanto para imagens com poucas *seeds* de entrada, quanto para imagens com um maior número de *seeds*. Para ilustrar esse ponto, as Figuras 3 e 4 apresentam os resultados obtidos pela técnica original, e com as novas *features* adicionadas ao modelo.



(a) Seeds



(b) Resultados Originais



(c) Novos resultados

**Fig. 3:** Resultados ilustrativos de segmentação.

## 4 CONCLUSÃO

O estudo realizado com foco direcionado às novas configurações de entrada na matriz de pesos no modelo de segmentação testado trouxe resultados bastante satisfatórios. De fato, a ideia de introduzir novas *features*, as quais visam destacar os limites dos objetos em uma imagem, aliada também com a variação de parâmetros, pode resultar em melhorias ainda mais significativas na tarefa de segmentação.

O propósito é dar prosseguimento no referido estudo, buscando novas combinações de *features* que não só agreguem nos resultados, mas que também apresente um tempo computacional aceitável.



**Fig. 4:** Resultados ilustrativos de segmentação.

## REFERÊNCIAS

- [1] Benes, M. and Zitova, B. (2014). “Performance evaluation of image segmentation algorithms on microscopic image data”. *Journal of microscopy*, 257.
- [2] Bishop, C. (2006). *Pattern Recognition and Machine Learning*, volume 16. Springer.
- [3] Canny, J. (1986). “A computational approach to edge detection”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698.
- [4] Casaca, W. (2015). *Graph Laplacian for spectral clustering and seeded image segmentation*. Instituto de Ciências Matemáticas e de Computação - ICMC-USP, São Carlos, São Paulo.
- [5] Casaca, W., Nonato, L. G., and Taubin, G. (2014). “Laplacian coordinates for seeded image segmentation”. IEEE Computer Society, pages 384–391.
- [6] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). “Maximum likelihood from incomplete data via the em algorithm”. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38.
- [7] Fu, Z. and Wang, L. (2012). pages 61–66. In: *Color Image Segmentation Using Gaussian Mixture Model and EM Algorithm*, Springer.
- [8] Gedraite, E. S. and Hadad, M. (2011). “Investigation on the effect of a gaussian blur in image filtering and segmentation”. In: *Proceedings ELMAR-2011*, IEEE.
- [9] Gonzalez, R. C. and Woods, R. E. (2017). *Digital Image Processing (4th Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- [10] Iancu, A., Popescu, B., Brezovan, M., and Ganea, E. (2011). “Quantitative evaluation of color image segmentation algorithms”. *IJCSA*, 8:36–53.
- [11] McLachlan, G. J. and Krishnan, T. (2008). *The EM algorithm and extensions*. 2 ed.
- [12] McLachlan, G. J. and Peel, D. (2000). *Finite mixture models*. Wiley Series in Probability and Statistics.
- [13] Otsu, N. (1979). “A threshold selection method from gray-level histograms”. *Systems, Man and Cybernetics, IEEE Transactions on*, 9:62–66.
- [14] Rand, W. M. (1971). “Objective criteria for the evaluation of clustering methods”. *Journal of the American Statistical Association*, 66(336):846–850.
- [15] Rother, C., Kolmogorov, V., and Blake, A. (2004). “Grabcut: interactive foreground extraction using iterated graph cuts”. *ACM TRANS. GRAPH*, pages 309–314.
- [16] Unnikrishnan, R., Pantofaru, C., and Hebert, M. (2005). “A measure for objective evaluation of image segmentation algorithms”. pages 34–34.



# Filtros Ópticos Otimizados por Algoritmos Genéticos Associados à Lógica Nebulosa

Wilton Moreira Ferraz Junior<sup>1,2</sup>, Carlos Henrique da Silva Santos<sup>2</sup> e  
Marcos Sergio Gonçalves<sup>1</sup>

<sup>1</sup> Faculdade de Tecnologia da Universidade Estadual de Campinas - FT/Unicamp/SP, Brazil

<sup>2</sup> Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – IFSP, Itapetininga/SP, Brasil

---

## Abstract

A combinação de simulações eletromagnéticas com técnicas de otimização heurísticas possibilita o projeto de estruturas complexas, como as requeridas por dispositivos ópticos. Assim, este trabalho apresenta a integração do método dos elementos finitos (*Finite Element Method*) - FEM associado a um Algoritmo Genético (GA) combinado com Lógica Nebulosa (*Fuzzy*) para o desenvolvimento de um filtro óptico destinado a sistemas DWDM (*Dense Wavelength Division Multiplexing*). Os resultados numéricos demonstram a eficiência na otimização, primeiro quanto uma melhor convergência de 86% do GA com *Fuzzy* contra 76% com apenas GA. Como segunda contribuição, também observa-se a redução na quantidade de iterações necessárias nas otimizações que, por consequência, reduz o tempo de processamento.

**Keywords:** DWDM, Algoritmo Genético, Lógica Nebulosa, Filtro Óptico.

---

## 1 INTRODUÇÃO

A complexidade para a fabricação e desenvolvimento de novas tecnologias ópticas associada à redução de seus custos e dimensões dos dispositivos têm demandado recursos computacionais cada vez mais sofisticados, desde as simulações até a infraestrutura experimental de validação. Para exemplificar, ao se considerar o cenário disruptivo de Internet das Coisas e as iniciativas de sua democratização por meio de tecnologias de baixo custo, são necessários meios de comunicação mais eficientes [1]. Neste cenário, destacam-se as fibras ópticas, por apresentarem baixo custo de fabricação e extensa largura de banda, sendo os sistemas de Multiplexação Densa por Divisão de Comprimento de Onda (*Dense Wavelength Division Multiplex* - DWDM) a técnica mais empregada para otimizar a capacidade de transmissão dessas fibras [2].

Os sistemas DWDM são tecnologias que permitem a transmissão de vários canais ópticos por meio de uma única fibra óptica. Ademais, com o intuito de manter a padronização dos canais ópticos, os sistemas DWDM devem atender às normas da *International Telecommunication Union* (ITU), distinguindo-os em canais de 12,5 GHz, 25 GHz, 50 GHz e 100 GHz [3][4]. Em um sistema DWDM, os filtros ópticos são dispositivos essenciais, sendo responsáveis pela

seleção de um determinado canal óptico e também na atenuação de ruídos[2]. Ainda são necessários considerar outros requisitos como, por exemplo, uma largura de banda compatível com os espaçamentos entre os canais e sua geometria (tamanho, forma e materiais) que possibilitem a fabricação de vários filtros em um único circuito óptico integrado.

Todavia, constata-se que essa é uma tarefa complexa e que, usualmente, não há soluções analíticas para a abstração teórica de sua estrutura, requerendo métodos numéricos que possibilitem realizar análises envolvendo os campos eletromagnéticos. Neste contexto, observa-se que o desempenho computacional e técnicas de otimização são importantes requisitos a serem considerados. Diante desta crescente demanda, este trabalho apresenta a integração da Lógica Nebulosa (*Fuzzy*) associada a um Algoritmo Genético (*Genetic Algorithm - GA*) para a otimização desses dispositivos modelados, utilizando método dos elementos finitos. O GA é uma heurística inspirada na natureza e que é bastante utilizada em projetos de dispositivos fotônicos, mimetizando as relações evolutivas baseadas na Teoria da Seleção de Natural de Charles Darwin e associando a conceitos genéticos, sendo usualmente chamado de Neo-Darwinianos [5].

Para isso, uma estrutura de dados vetorial representa um cromossomo e cada elemento do vetor associa-se a um gene. Portanto, um conjunto de cromossomos (soluções candidatas) compõe um conjunto de soluções candidatas representadas por uma matriz. Essa matriz é modificada ao longo das iterações (gerações) por meio de operadores genéticos chamados de recombinação e mutação. A cada iteração os cromossomos são avaliados por uma função objetivo e os melhores são selecionados para compor a população da próxima geração. Neste trabalho, a integração da Lógica *Fuzzy* tem por objetivo auxiliar na pontuação da função objetiva que modela os filtros, visando potencializar resultados aceitáveis. Essa associação visa reduzir a quantidade de iterações necessárias à otimização e melhorar as soluções retornadas para o filtro já descrito em [4].

Este artigo apresenta na segunda seção, um breve resumo sobre FEM e, na terceira seção, alguns dos fundamentos GA e *Fuzzy* além da forma como foram integrados neste trabalho. A seção 4 apresenta a arquitetura do filtro óptico otimizado, sendo sucedido pela seção 5 que por sua vez, expõe a metodologia adotada nesta pesquisa e os resultados obtidos nas otimizações realizadas. Por fim, a sexta seção apresenta as conclusões e trabalhos em andamento.

## 2 MÉTODO DOS ELEMENTOS FINITOS - FEM

O método dos elementos finitos (*Finite Element Method - FEM*) é uma técnica numérica para obter soluções de equações diferenciais parciais em domínios finitos. Ao longo do tempo, ele foi aplicado extensivamente em problemas de elasticidade e análise estrutural e, cada vez mais, a problemas de outros ramos, como dinâmica de fluidos e eletromagnetismo [6].

A equação de onda escalar em duas dimensões (2D) para campos eletromagnéticos no domínio da frequência, considerando que não há variação dos campos na direção  $x$ , é dada por (1).

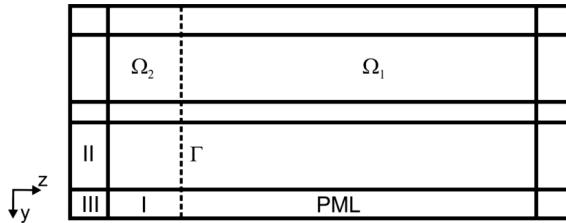
$$S_y \frac{\partial}{\partial y} \left( p \frac{s_y}{s} \frac{\partial \Phi}{\partial y} \right) + S_z \frac{\partial}{\partial z} \left( p \frac{s_z}{s} \frac{\partial \Phi}{\partial z} \right) + k_0^2 s^2 q \Phi = 0 \quad (1)$$

Onde  $\Phi = H_x$ ,  $p = 1/n^2$  e  $q=1$  para o campo magnético ou  $\Phi = E_x$ ,  $p = 1$  e  $q = n^2$  para o campo elétrico, sendo  $n$  o índice de refração e  $k_0$  o número de onda. Com a finalidade de truncar o domínio computacional, foram utilizadas camadas de PML - *Perfectly Matched Layer* [7], sendo necessário utilizar operadores diferenciais modificados e o parâmetro  $s$ , definido em [6, 7] dado por (2).

$$s = \begin{cases} 1 - j \frac{3c^2}{2\omega_0 n d} (\frac{\rho}{d})^2 \ln(\frac{1}{R}) & \text{para regiões dentro da PML} \\ 1 & \text{para regiões fora da PML} \end{cases} \quad (2)$$



Onde  $\omega_0$  é a frequência angular,  $\rho$  é a distância do início da PML,  $d$  é a espessura da PML e  $R$  é o coeficiente teórico de reflexão [6, 7]. A Fig.1 ilustra um exemplo do domínio computacional utilizado nas análises. Nas PML para as regiões I, perpendiculares ao eixo  $y$ , adotou-se que  $s_y = 1$  e  $s_z = s$ , para as regiões II perpendiculares ao eixo  $y$ ,  $s_y = s$  e  $s_z = 1$  e para os cantos III,  $s_y = s_z = 1$ . [6, 7].



**Fig. 1:** Exemplo de domínio computacional

Neste trabalho, o domínio computacional  $\Omega$  é encerrado por PEC ou PMC. Este domínio é dividido em dois subdomínios  $\Omega_1$  e  $\Omega_2$ , com fronteira em  $\Gamma$  (plano de incidência), conforme mostrado na Fig.1 Considerando o método de Galerkin [6] para a discretização de (1), tem-se para o subdomínio  $\Omega_1$  em (3),

$$\iint \left\{ p \frac{s_y^2}{s} \frac{\partial \{N\}}{\partial y} \frac{\partial \Phi}{\partial y} + p \frac{s_z^2}{s} \frac{\partial \{N\}}{\partial z} \frac{\partial \Phi}{\partial z} - k_0^2 s^2 q \Phi N \right\} dy dz = \int_{\Gamma} p \{N\} \frac{\partial \Phi_{in,1}}{\partial z} dy, \quad (3)$$

onde  $N$  são as funções de peso para um elemento triangular quadrático [6],  $\Gamma$  é a superfície do plano de incidência. Para as regiões fora da PML,  $\Phi_{in,1}$  é dado por (4). Para as regiões dentro da PML,  $\Phi_{in,1} = 0$ , Em (4)  $\Phi_1$  é a amplitude do campo incidente e corresponde ao modo fundamental de propagação. Aplicando (4) em (3), obtém-se (5).

$$\Phi_{in,1} = \Phi_1(y) e^{-j\beta z} \quad (4)$$

$$\iint \left\{ p \frac{s_y^2}{s} \frac{\partial \{N\}}{\partial y} \frac{\partial \Phi}{\partial y} + p \frac{s_z^2}{s} \frac{\partial \{N\}}{\partial z} \frac{\partial \Phi}{\partial z} - k_0^2 s^2 q \Phi N \right\} dy dz = - \int_{\Gamma} j\beta p \{N\} \Phi_1 dy. \quad (5)$$

Para o subdomínio  $\Omega_2$ , adotando-se o campo incidente como  $\Phi_{in,2} = \Phi_2(y) e^{j\beta z}$ , realizando-se o mesmo procedimento feito para  $\Omega_1$ , levando-se em consideração a continuidade dos campos na fronteira  $\Gamma$  [8] e expandido o campo espalhado como  $\Phi = \{N\}^T \{\phi_e\}$  obtém-se o sistema matricial de equações em (6).

$$[K]\phi = [M]\{\Phi_{in}\}. \quad (6)$$

Neste caso, as matrizes elementares são descritas em (7), (8) e (9)

$$[K_e] = \iint_e \left\{ p \frac{s_y^2}{s} \frac{\partial \{N\}}{\partial y} \frac{\partial \{N\}^T}{\partial y} + p \frac{s_z^2}{s} \frac{\partial \{N\}}{\partial z} \frac{\partial \{N\}^T}{\partial z} - k_0^2 s^2 q N \{N\}^T \right\} dy dz; . \quad (7)$$

$$[M_e] = \int_{\Gamma} p \{N\}_{\Gamma} \{N\}_{\Gamma}^T dy; \quad (8)$$

$$\{\Phi_{in,e}\} = -j2\beta\Phi_{\Gamma}(y). \quad (9)$$

### 3 ALGORITMOS GENÉTICOS E FUZZY

Os Algoritmos Genéticos (GA) são abstrações de conceitos associados à Teoria da Seleção Natural de Darwin e à genética Mendeliana [9]. O GA possui uma natureza estocástica, envolvendo um conjunto de soluções candidatas (população) para satisfazerem uma ou mais funções objetivas (*fitness*) [10], em que há um conjunto solução (população) que é iterada (gerações) até que uma condição de parada seja satisfeita.

Ao longo desse processo, alguns indivíduos (genes) se relacionam com outros (recombinação), formando novas soluções candidatas. Além disso, há operações de alterações aleatórias (mutações) nos atributos (alelos) de algumas dessas soluções candidatas, que associadas a recombinação, diversifica as soluções candidatas para facilitar a busca por resultados esperados. Esse processo iterativo com as soluções candidatas possui um mecanismo de armazenamento dos melhores resultados, que aqui é chamado de elitismo por associar ao contexto dos melhores indivíduos adaptados ao ambiente (espaço de busca), sendo medidos por uma função objetivo chamada de *fitness* [10, 11].

Com isso, pode-se representar o GA por meio de uma matriz (10), em que  $n$  é a quantidade de indivíduos na população e  $L$ , a quantidade de atributos necessários para representar cada indivíduo. Nesta aplicação, cada atributo é do tipo real de ponto flutuante duplo e seus respectivos valores representam características fenotípicas de cada um dos indivíduos que são medidas pela função objetivo (*fitness*), para se avaliar a adaptação do indivíduo ao meio [11].

$$P_{ga} = \begin{bmatrix} P_{1,1} & P_{1,2} & \dots & P_{1,L} \\ P_{2,1} & P_{2,2} & \dots & P_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n,1} & P_{n,2} & \dots & P_{n,L} \end{bmatrix} \quad (10)$$

A manutenção e controle da diversidade dos atributos da população são tarefas complexas que requerem conhecer os operadores genéticos (mutação e recombinação), os critérios de seleção e a definição da função objetivo (*fitness*) do problema [12, 13]. Essa afirmação vem associada à demanda pela redução na quantidade das iterações nas otimizações em que há elevado consumo de memória e processamento, por conseguinte, tempo de execução computacional, para que seja calculado o *fitness* de cada indivíduo. Isso é comum em problemas de engenharia que utilizam métodos numéricos para a modelagem de novos dispositivos [14].

Por essas razões, diferentes técnicas computacionais vêm sendo adaptadas e integradas ao AG para que seu desempenho seja melhorado como, por exemplo, a associação de Redes Neurais Artificiais para variar a forma de busca [14], controlar a geração da população inicial com outros algoritmos como o de Abelhas [15] ou o controle nebuloso por lógica nebulosa por conjuntos Fuzzy [16]. Este último é bastante interessante e objeto de estudo neste trabalho por não requerer sofisticado poder computacional para a classificação e por proporcionar flexibilidade descritiva para pontuar regiões de interesse nas otimizações.

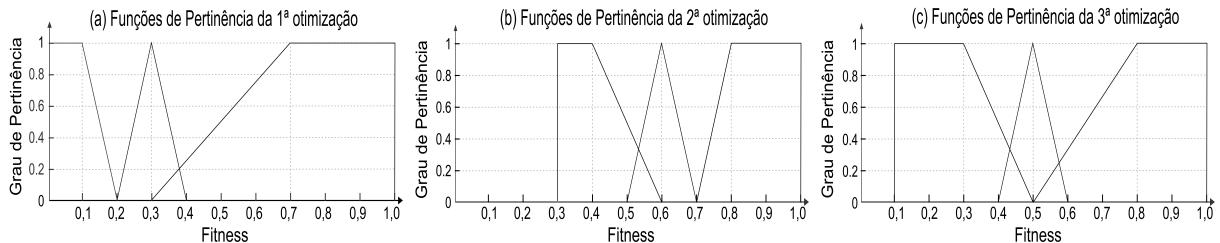
Os fundamentos da lógica nebulosa, conhecidos como conjuntos nebulosos (*Fuzzy Sets*), foram apresentados à comunidade científica por Lotfi Zadeh em 1965 [17] e permitiu reproduzir variáveis linguísticas humanas de maneira mensurável pela matemática e a execução de operações entre conjuntos [18].

Um conjunto nebuloso é usualmente representado por:  $\langle U, M, \mu_{\tilde{A}} \rangle$ ,  $x \in X$ , onde  $U$  é o universo de discurso,  $\mu_{\tilde{A}}(x) : X \rightarrow M$  e  $\mu_{\tilde{A}}(x)$  é o grau de pertinência  $X$  de  $x$  em  $A$  -  $M$  número no intervalo  $[0, 1]$  [19].

Para isso, inicialmente é preciso representar os valores de entrada utilizando conjuntos nebulosos. Cada conjunto é figurado por uma função de pertinência caracterizada por valores modais e eles representam a pertinência de um valor de entrada no intervalo em estudo. As funções de

pertinência possuem vários formatos, sendo mais comuns o Triangular, Trapezoidal, Gaussiano e função de Bell. [20]

Neste trabalho, foram utilizadas as funções Triangular e Trapezoidal, por serem aproximações válidas para muitas funções de maior complexidade. Além disso, estes tipos de funções são relativamente simples de serem implementadas e para se ajustar seus parâmetros. Para determinar os melhores conjuntos de valores para as funções, foram executados 3 ciclos de otimizações com diferentes parâmetros para cada função, conforme apresentado na Fig. 2.

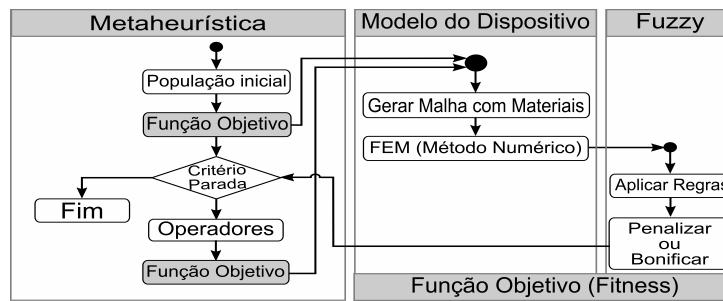


**Fig. 2:** Funções de pertinência: (a) funções da primeira otimização (b) funções da segunda otimização (c) funções da terceira otimização.

Assim, com fluxo geral, o GA gera uma população inicial (conjunto de soluções candidatas) aleatoriamente e, no passo seguinte os, indivíduos desta população são analisados pela função objetivo que inicialmente gera a malha com os materiais, para logo após por meio do FEM determinar o (*fitness*) de cada indivíduo, retornando o valor da função objetivo por regras matemáticas abordadas na seção 2.

Com o resultado escalar obtido pela regras matemáticas da função objetivo, são aplicadas as regras *Fuzzy* para bonificar os indivíduos com grau de pertinência superiores a 0,6 (seria o equivalente a 60% de acoplamento) no intervalo de (*fitness*) de 0,7 a 1,0, conforme apresentado na Fig. 3. Esse processo iterativo de análise da função objetivo integrada ao *Fuzzy* é realizado em cada uma das iterações por todos os indivíduos da população.

Os conjuntos de valores da Fig. 2 (c) utilizados na terceira otimização foram os que apresentaram melhores resultados com perda de inserção de 1.3 dB. Já os conjuntos de valores utilizados na Fig. 2 (a) apresentaram perda de inserção de 2.1 dB e, por fim, os da Fig. 2 (b) apresentaram perda de inserção de 2,7 dB. Observou-se também, que, em todos os casos que a velocidade de convergência das otimizações que utilizaram o GA+*Fuzzy*, foram superiores as das que utilizaram somente o GA. Demais resultados obtidos do filtro são mostrados no seção 5.



**Fig. 3:** Fluxograma da integração entre Método dos Elementos Finitos e GA+*Fuzzy*.

#### 4 ARQUITETURA DO FILTRO ÓPTICO

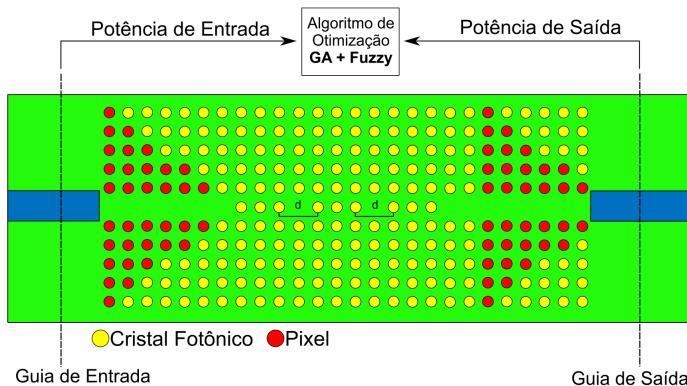
Sistemas de Multiplexação Densa por Divisão de Comprimento de Onda - DWDM (*Dense Wavelength Division Multiplex*) são tecnologias que permitem a transmissão de múltiplos canais

ópticos em uma fibra óptica. Estes sistemas podem valer-se de diferentes tecnologias, como filtros de interferência de filmes, fibras ópticas com grade de Bragg, guias de onda planares do tipo AWG (*Arrayed Waveguide Gratings*) e ressonadores em anéis [4].

Dentre estas tecnologias, os filtros ópticos se configuram como uma opção atrativa, pois podem ser fabricados com um baixo custo, apresentam perda baixa de inserção entre 1 dB e 2 dB e boa resposta espectral para separação convencional entre canais ópticos de 100 GHz [4]. Contudo, podem conter alguns aspectos indesejáveis como dimensões na ordem de centenas de micrômetros ou dificuldades no ajuste da frequência de ressonância, como ocorre nos filtros de filme fino, [21] dificultando a alta integração de vários filtros em um único circuito óptico [4].

Uma opção para o desenvolvimento de filtros ópticos em pequenas dimensões é a utilização de cristais fotônicos, que são estruturas periódicas que obedecem a uma determinada simetria e de *pixels* que são colunas dielétricas com raio igual  $0,1044 \mu\text{m}$  e separação de  $0,58 \mu\text{m}$ . Estas estruturas se destacam pela presença de regiões de bandas proibidas em determinadas faixas de frequências, nas quais a onda eletromagnética não consegue se propagar. A introdução de um defeito na periodicidade é uma técnica de construção nesta estrutura, que cria modos de propagação com frequências de operação dentro da banda proibida [4, 22].

A modelagem do filtro levou em consideração a facilidade na fabricação e baseou-se nas colunas utilizadas em estruturas de cristais fotônicos, conforme apresentado na Fig.4. Para tanto, o projeto deste filtro considerou duas microcavidades inseridas em um guia de onda contínuo. Por consequência, o dispositivo é formado por dois acopladores ópticos e por microcavidades em série. O dispositivo é formado por cristais fotônicos estruturados por colunas dielétricas com índice de refração de 3,4 e raio de  $0,1044 \mu\text{m}$ . A simetria dos cristais é retangular com periodicidade de  $0,58 \mu\text{m}$ . Nestas condições, há a formação de uma região de banda proibida para os modos TE entre os comprimentos de onda de  $1,3 \mu\text{m}$  até  $1,93 \mu\text{m}$  [23].



**Fig. 4:** Arquitetura utilizada para o projeto do filtro óptico.

Com a finalidade de se obter um guia de ondas, foi retirada uma linha de colunas dielétricas do centro da estrutura periódica.

Foram incluídas no interior deste guia de onda, conjuntos com três colunas que constituem as microcavidades ressonantes, sendo que a distância  $d$  entre os conjuntos de colunas é o que define o comprimento de onda de ressonância.

Para os acopladores com base em [4] buscou-se por soluções que maximizassem o acoplamento dos feixes ópticos, considerando-se que as variações aleatórias na periodicidade das colunas dielétricas. Para isso, para cada uma das posições das colunas dielétricas foi estabelecido um conceito de pixel de projeto, com as mesmas dimensões dessas, mas considerando que seus valores de índice refração podem ser 1,0 ou 3,4, diferente de [4] e que cada pixel poderia assumir três valores de índices de refração. Consequentemente, há uma simplificação em um eventual

processo de fabricação. O valor de cada pixel será determinado no processo de otimização que levará em consideração o melhor acoplamento obtido e irá sobrepor a coluna dielétrica com o valor estabelecido aleatoriamente pela meta-heurística, que neste caso considera-se o GA e também o GA+*Fuzzy*.

## 5 METODOLOGIA DE DESENVOLVIMENTO E RESULTADOS

Com o propósito de encontrar a melhor solução para o problema apresentado, utilizou-se inicialmente uma integração do FEM + GA. O FEM permite verificar a propagação da onda eletromagnética no dispositivo.

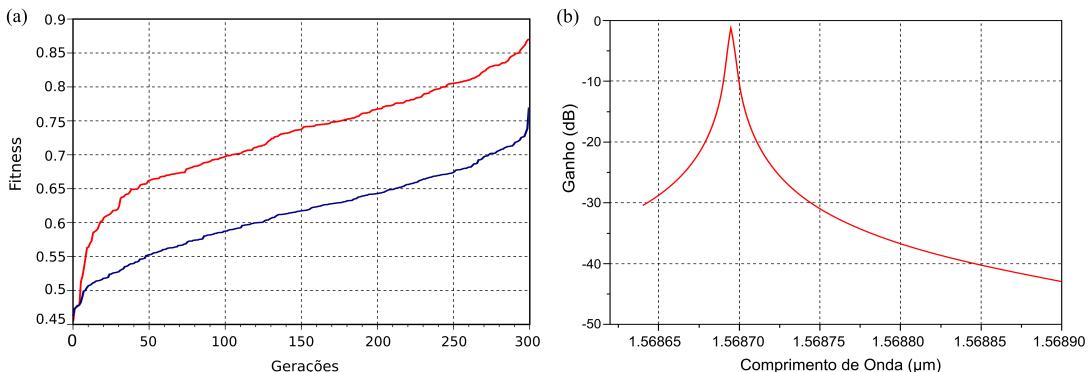
No guia de onda de entrada, aplica-se o modo fundamental TE<sub>10</sub>, determinando a potência óptica de entrada. Depois de propagar-se pelo filtro óptico, a onda eletromagnética é acoplada no guia de onda de saída e novamente é efetuada a média da potência óptica. Ambas são enviadas ao GA+*Fuzzy* que efetuará a otimização dos *pixels*, com objetivo de maximizar o acoplamento óptico entre os guias de onda no comprimento de onda de ressonância.

Para este propósito, utilizou-se uma probabilidade de recombinação de 60% e de mutação de 20%, uma população inicial de 100 indivíduos e 300 gerações. A probabilidade que cada indivíduo possui para ser a solução é dada por (11),

$$P_i = \frac{P_i}{\sum_{k=1}^N P_k} \quad (11)$$

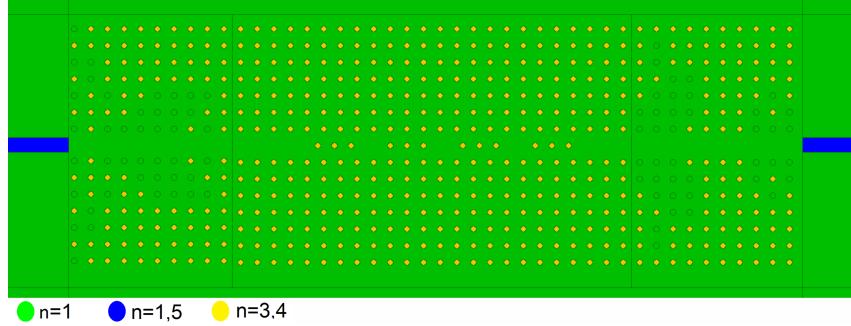
na qual  $N$  é o número de indivíduos da população e  $p$  é a relação entre a potência óptica de saída pela potência óptica de entrada. Os dois indivíduos que apresentarem a maior probabilidade serão escolhidos para compor a próxima população. Como apresentado na Fig.5 (a) é possível notar que a eficiência de convergência do FEM e GA foi de 0,76 em uma escala de 0 a 1 enquanto a otimização FEM e GA+*Fuzzy* utilizando os mesmos parâmetros descritos anteriormente alcançou o resultado de 0,86. Outro fator importante a se observar é a quantidade de iterações na otimização: FEM e GA precisou de 265 para alcançar seu melhor resultado, já o FEM e GA+*Fuzzy* alcançou seu melhor resultado com 58 iterações.

Já a Fig.5 (b) demonstra a curva do ganho do filtro óptico em função do comprimento de onda. Conforme observado, a ressonância do filtro ocorre no comprimento de onda de 1,5687 μm, valor este compatível com a grade ITU. Neste comprimento de onda, o filtro óptico apresenta uma perda de inserção, considerando-se as perdas dos dois acopladores e das microcavidades, de 1.3 dB e uma largura de banda de 366 MHz.



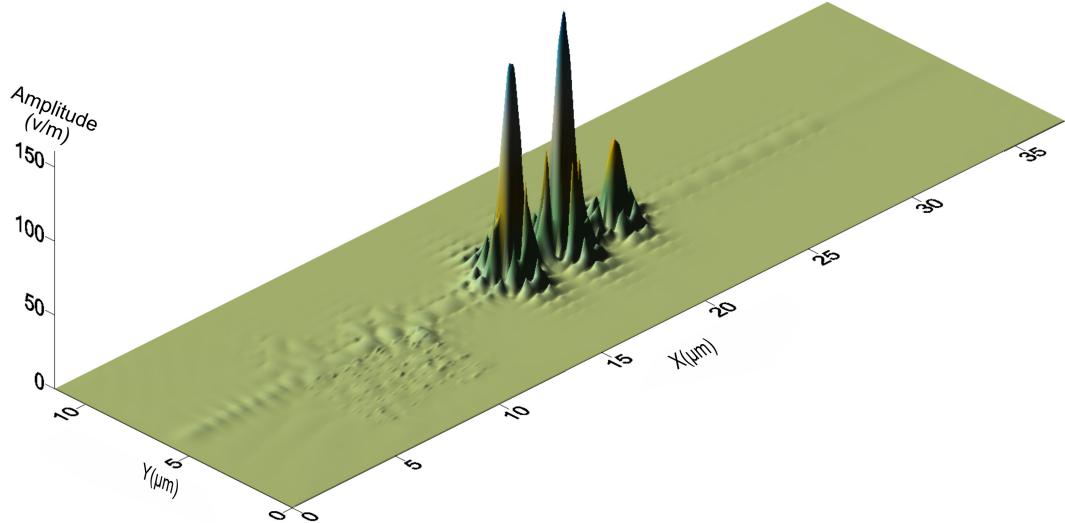
**Fig. 5:** (a) Comparação dos Fitness obtidos nas otimizações utilizando MEF e GA e MEF e GA+*Fuzzy*. (b) Ganho do filtro óptico em função do comprimento de onda para a operação no comprimento de onda de 1,5687 μm.

A Fig.6 apresenta a distribuição resultante dos *pixels* ao final da realização das etapas de otimizações dos acopladores de entrada e saída. A geometria do filtro óptico configura-se com três microcavidades obtida após as otimizações para  $\lambda = 1,5687 \mu\text{m}$ . Nessa figura, é possível verificar a distribuição dos *pixels* dos acopladores de entrada e saída. A distância  $d$  entre as colunas foi de  $1,364 \mu\text{m}$ .



**Fig. 6:** Projeto do filtro óptico para o comprimento de onda de  $1,5687 \mu\text{m}$ .

A Fig.7 apresenta a componente  $x$  do campo elétrico no interior do filtro óptico no comprimento de onda de operação de  $1,5687 \mu\text{m}$ . Nesse caso, é possível verificar a ação dos acopladores nos acoplamentos entre os guias de onda contínuos e de cristais fotônicos. Observa-se, também, os elevados valores alcançados pela amplitude do campo elétrico no comprimento de onda de ressonância.



**Fig. 7:** Distribuição da componente  $x$  do campo elétrico no interior do filtro óptico.

## 6 CONCLUSÃO

A combinação do FEM com o GA+*Fuzzy* demonstrou ser eficiente na otimização, apresentando uma de convergência 86% contra 76% do FEM e GA, além de requerer menor quantidade de iterações na otimização, conforme já indicado na Fig.5 (a). Esses resultados foram obtidos em três rodadas de otimizações com diferentes conjuntos para as funções de pertinência, visando-se estabelecer os melhores parâmetros para as funções.

Com isso, observou-se que todas otimizações que utilizaram o GA+*Fuzzy* são computacionalmente mais eficientes, tanto no quesito de maior agilidade na convergência da solução e,



consequente, redução nas iterações de otimização, o que demanda menor tempo de processamento. Da perspectiva física da aplicação, as otimizações utilizando GA+*Fuzzy* possibilitaram obter um filtro óptico que demonstra uma perda de potência de sinal resultante da inserção do dispositivo em uma linha de transmissão, considerando-se as perdas dos dois acopladores e da microcavidade, de 1,3 dB e uma largura de banda de 366 MHz no comprimento de onda de 1,5687  $\mu\text{m}$ . Em suma, os resultados numéricos utilizando a otimização com pontuação *Fuzzy* apresentaram uma melhoria de cerca de 10% na eficiência de acoplamento do filtro; um menor tempo de processamento devido a menor quantidade de iterações do GA.

Essa equipe de pesquisa tem atuado na implementação e testes de outras heurísticas inspiradas na natureza, combinadas com *Fuzzy* para se verificar seu desempenho em diferentes otimizações de dispositivos fotônicos, visando distribuir futuramente uma ferramenta computacional gratuita à comunidade.

## 7 AGRADECIMENTOS

Os autores agradecem ao IFSP, campus Itapetininga, (sob o edital PRP-226/2016) e à Faculdade de Tecnologia da Unicamp pelo suporte técnico.

## Referências

- [1] Flavio Mania, Carlos Henrique da Silva Santos, and Alexandre Alvaro. Outlining low costs and open embedded systems for rfid in internet of things applications. In *2014 IEEE Brasil RFID*, pages 16–18. IEEE, 2014.
- [2] Yin-Hai Li, Wen-Tan Fang, Zhi-Yuan Zhou, Shi-Long Liu, Shi-Kai Liu, Zhao-Huai Xu, Chen Yang, Yan Li, Li-Xin Xu, Guang-Can Guo, et al. All-optical quantum signal demultiplexer. *arXiv preprint arXiv:1803.00206*, 2018.
- [3] RG ITU-T. Series g: Transmission systems and media digital systems and networks. *Ginebra, Suíça: ITU*, 1997.
- [4] Joao Victor Marangoni and Marcos Sergio Goncalves. Numerical analysis of a novel optical filter for dwdm systems. *IEEE Latin America Transactions*, 14(4):1943–1947, 2016.
- [5] Melanie Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.
- [6] Jian-Ming Jin. *The finite element method in electromagnetics*. John Wiley & Sons, 2015.
- [7] Masanori Koshiba, Yasuhide Tsuji, and Masafumi Hikari. Finite element beam propagation method with perfectly matched layer boundary conditions. *IEEE transactions on magnetics*, 35(3):1482–1485, 1999.
- [8] Yasuhide Tsuji and Masanori Koshiba. Finite element method using port truncation by perfectly matched layer boundary conditions for optical waveguide discontinuity problems. *Journal of lightwave technology*, 20(3):463–468, 2002.
- [9] Thomas Bäck, David B Fogel, and Zbigniew Michalewicz. *Evolutionary computation 1: Basic algorithms and operators*. CRC press, 2018.
- [10] Carlos H Silva-Santosa, Kleucio Claudiob, Hugo E Hernández-Figueroac, and Marcos Sérgio Gonçalves. Design of photonic devices using bio-inspired algorithms. In *2009 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*, pages 122–126. IEEE, 2009.

- [11] Wei Yu, Baizhan Li, Hongyuan Jia, Ming Zhang, and Di Wang. Application of multi-objective genetic algorithm to optimize energy efficiency and thermal comfort in building design. *Energy and Buildings*, 88:135–143, 2015.
- [12] Francisco Herrera and Manuel Lozano. Adaptation of genetic algorithm parameters based on fuzzy logic controllers. *Genetic Algorithms and Soft Computing*, 8:95–125, 1996.
- [13] Nicol N Schraudolph and Richard K Belew. Dynamic parameter encoding for genetic algorithms. *Machine learning*, 9(1):9–21, 1992.
- [14] Adriano da Silva Ferreira, Carlos Henrique da Silva Santos, Marcos Sergio Gonçalves, and Hugo Enrique Hernández Figueroa. Towards an integrated evolutionary strategy and artificial neural network computational tool for designing photonic coupler devices. *Applied Soft Computing*, 65:1–11, 2018.
- [15] Denis Tavares da Silva and Carlos Henrique da Silva Santos. Controlando população inicial de algoritmos genéticos para a otimização de funções. *Revista Brasileira de Iniciação Científica*, 6(2):68–91, 2019.
- [16] Tamara Jiménez, Noemí Merayo, Anaïs Andrés, Ramón J Durán, Juan C Aguado, Ignacio de Miguel, Patricia Fernández, Rubén M Lorenzo, and Evaristo J Abril. An auto-tuning pid control system based on genetic algorithms to provide delay guarantees in passive optical networks. *Expert systems with applications*, 42(23):9211–9220, 2015.
- [17] Lotfi A Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.
- [18] Ioannis G Damousis, KJ Satsios, DP Labridis, and PS Dokopoulos. Combined fuzzy logic and genetic algorithm techniques—application to an electromagnetic field problem. *Fuzzy sets and systems*, 129(3):371–386, 2002.
- [19] Anastasiya V Kostikova, Pavel V Tereliansky, Alexander V Shubaev, Valentina N Parakhina, and Pavel N Timoshenko. Expert fuzzy modeling of dynamic properties of complex systems. *ARPN J. Eng. Appl. Sci.*, 11(17):10601–10608, 2016.
- [20] Ebrahim H Mamdani and Sedrak Assilian. An experiment in linguistic synthesis with a fuzzy logic controller. In *Readings in Fuzzy Sets for Intelligent Systems*, pages 283–289. Elsevier, 1993.
- [21] AK Chu, KH Huang, CH Chiang, WC Tien, and FZ Lee. Strain-induced tunability of thin dwdm thin-film filters. *Electronics letters*, 48(18):1141–1142, 2012.
- [22] Katsunari Okamoto. Wavelength-division-multiplexing devices in thin soi: Advances and prospects. *IEEE Journal of Selected Topics in Quantum Electronics*, 20(4):248–257, 2014.
- [23] João Victor Marangoni, Francisco José Arnold, and Marcos Sérgio Gonçalves. Proposta de um filtro Óptico para aplicações em sistemas dwdm. *sbrt*, 2016.



# Análise do perfil de mutações de resistência em paciente HIV positivo multirresistente aos antirretrovirais

Ana Luiza Martins Karl<sup>1</sup> and Adriana de Oliveira Afonso<sup>2</sup>

<sup>1</sup> Laboratório Nacional de Computação Científica, Petrópolis/RJ, Brasil

<sup>2</sup> Universidade Católica de Petrópolis, Petrópolis/RJ, Brasil

---

## Resumo

Uma das primeiras características a se tornar notável na infecção pelo HIV foi seu alto potencial em acumular diversidade genética. Essa característica confere ao vírus a capacidade de responder rápida e eficientemente à pressões seletivas (Coffin and Swanson, 2013). A seleção de isolados virais com mutações de resistência aos antirretrovirais é um dos principais obstáculos para o tratamento do HIV, uma vez que afeta o controle da infecção. A identificação dessas mutações através dos testes de genotipagem e fenotipagem tem permitido estratégias personalizadas para o tratamento de indivíduos infectados pelo HIV-1, possibilitando a escolha do tratamento com maior chance de supressão viral (Van Laethem et al., 2015). O presente trabalho usou informações de genotipagem obtidas de um paciente multi-resistente aos antirretrovirais para análises computacionais. Os dados foram coletados do prontuário no DIP no Hospital Municipal Dr. Nelson de Sá Earp, em Petrópolis/RJ e modelos computacionais das enzimas mutantes protease e transcriptase reversa foram gerados utilizando a modelagem comparativa. Em seguida, estudos de atracamento molecular com os inibidores do esquema terapêutico do paciente foram realizados. Nossos resultados preliminares mostraram pequenas variações conformacionais dos modelos das proteínas mutantes. Os estudos de atracamento molecular indicaram diferenças nos modos de ligação dos inibidores com os modelos da enzima protease, indicando a influência da mudança da cadeia lateral dos aminoácidos do sítio ativo na interação com os inibidores.

**Palavras-chave:** HIV, antirretrovirais, multirresistência, estudo de caso, modelagem molecular.

---

## REFERÊNCIAS

- [1] Coffin, J. and Swanson, R. (2013). "Hiv pathogenesis: dynamics and genetics of viral populations and infected cells". *Cold Spring Harbor perspectives in medicine*, 3(1):a012526.
- [2] Van Laethem, K., Theys, K., and Vandamme, A.-M. (2015). "Hiv-1 genotypic drug resistance testing: digging deep, reaching wide?" *Current opinion in virology*, 14:16–23.



# Estudo comparativo de métodos de limitação de inclinação e fluxo para Leis de Conservação

Ana Caroline Francisca da Silva<sup>1</sup> and Daniel Gregorio Alfaro Vigo<sup>2</sup>

<sup>1</sup> Programa de Pós Graduação em Informática - PPGI-UFRJ ,RJ, Brazil

<sup>2</sup> Programa de Pós Graduação em Informática - Departamento Ciência da Computação, UFRJ/RJ, Brazil

---

## Abstract

Neste trabalho apresentamos um estudo aplicado a problemas modelados por leis de conservação, com variáveis limitadas por condições iniciais e contorno. A pesquisa apresenta um estudo sobre método limitador de inclinação e fluxo (Slope-limiter methods, Flux limiters), fazendo um levantamento comparativo de métodos de alta ordem em paralelo com o método limitador de inclinação. Abordaremos situações que envolve simulação com equações específicas como águas rasas e elasticidade unidimensional, analisando seu comportamento ao longo do tempo e com as condições estabelecidas.

O método escolhido é o Limitador de inclinação e fluxo são esquemas numéricos de alta resolução; que segundo Randall (2008), este aplica-se a métodos que são pelo menos de segunda ordem precisos em soluções suaves e ainda resolve o problema apropriadamente. Com base nos estudos apresentados pela literatura Cuminato et al. (2013), Maugin et al. (2008) temos que Métodos de alta ordem como TVD (Total Variation Diminishing) e Godunov tendem a corrigir o erro de fase. Sendo assim com o limitador de inclinação alcançamos o objetivo de comparação dos esquemas numéricos para aproximação da solução para sistema de leis de conservação.

**Keywords:** Leis de conservação, Limitador de inclinação, Equações de Águas Rasas, Equação Unidimensional da Elasticidade

---

## REFERENCES

- [1] Cuminato, J. A., Junior, and Meneguette, M. (2013). *Discretização de Equações Diferenciais Parciais- Técnicas de diferenças finitas.* 1 ed.
- [2] Maugin, G., Engelbrecht, J., and Berezovski, A. (2008). *Numerical Simulation of Waves and Fronts in Inhomogeneous Solid.* 1 ed.
- [3] Randall, J. L. (2008). *Numerical Methods for Conservation Laws.* 2 ed.



# Redes Neurais Artificiais: modelagem computacional da previsão de volume de uma usina hidrelétrica

Bárbara Raquel Mendonça Rezende<sup>1</sup>, Eliane da Silva Christo<sup>1</sup> e Fernando Tadeu Pereira de Medeiros<sup>1</sup>

<sup>1</sup> *Modelagem Computacional em Ciência e Tecnologia, Universidade Federal Fluminense - EEIMVR, Volta Redonda/RJ, Brasil*

---

## Abstract

A matriz energética brasileira é composta, na grande maioria, por usinas hidrelétricas. Elas produzem energia limpa e renovável, controlam as cheias dos rios, abastecem populações e geram empregos. dessa forma, é importante garantir a máxima eficiência no funcionamento das mesmas, já que a consequência direta é um impacto significativo no custo da produção, na especificação de energia e no meio ambiente. Determinar o volume de uma usina hidrelétrica é uma etapa fundamental de importância na eficiência desta operação. Um projeto de redes neurais consiste em pré-processamento, processamento e, por fim, um pós-processamento dos dados. São capazes de reconhecer regularidades e padrões de dados através do treinamento e realizar generalizações baseadas no conhecimento adquirido. Desta forma, as Redes Neurais Artificiais (RNA) podem auxiliar na prevenção de riscos e transtornos, como cheias e desabamentos, proporcionando um volume das hidrelétricas sempre dentro dos padrões ou agindo em cima das prevenções. Adotou-se essa técnica devido ao seu alto grau de desempenho computacional para resoluções de problemas relacionados com previsões de séries temporais, o que proporciona maior segurança e confiabilidade nos resultados esperados, gerando satisfação nas tomadas de decisões. Portanto, o presente estudo tem como objetivo a elaboração de um modelo de previsão, com a utilização das ferramentas de Redes Neurais Artificiais, para estimar o volume útil da usina hidrelétrica de Furnas - MG, mediante dados obtidos pela Agência Nacional de Águas, durante o período de 1 ano. Esse rede foi implementada no software MatLab devido ao alto desempenho computacional. A rede foi parametrizada com 80% dos dados e a validação foi realizada com os outros 20%. Utilizou-se o algoritmo *trainlm* para treinamento da rede e, as funções de ativação, foram logarítmica (*logsig*) e linear (*purelin*). Através desse modelo, foi possível analisar e avaliar a melhor combinação das camadas de saída. Após diversos treinamentos, a rede retornou um coeficiente de determinação maior de 0,98, provando que a qualificação da mesma é válida.

**Keywords:** Modelagem Computacional, Redes Neurais Artificiais, Energia, Previsão.

---

## REFERENCES

- [1] Ballini, R. (2000). *Análise e previsões de vazões utilizando modelos de séries temporais, redes neurais e redes neurais nebulosas.*
- [2] Braga, A. (2000). “Redes neurais artificiais: teoria e aplicações.” *Livros Técnicos e Científicos.*
- [3] Haykin, S. (2008). *Redes Neurais: Princípios e Práticas.* 2 ed.
- [4] Kovács, Z. (2006). “Redes neurais artificiais - fundamentos e aplicações”.
- [5] Matsunaga, V. (2012). *Curso de redes neurais utilizando o matlab.* Belém do Pará.
- [6] Soares Filho, S. (1987). “Planejamento da operação de sistemas hidrotérmicos”. *SBA - Controle e Automação*, pages 122–131.



# Boa Colocação e Estabilidade via Semigrupos Lineares para o Problema Termoelástico

Bruna Thais Silva Sozzo<sup>1</sup> e Jaime E. Muñoz Rivera<sup>1</sup>

<sup>1</sup> Laboratório Nacional de Computação Científica, Petrópolis/RJ, Brasil

---

## Resumo

A teoria de semigrupos de operadores lineares ou teoria exponencial dos operadores é usada principalmente para o estudo de boa colocação e estabilidade de problemas de valor inicial e de contorno (PVIC). O ponto-chave da teoria é que seus resultados são válidos para operadores lineares não limitados em espaço de dimensão infinita. Quando nos referimos a essa categoria de operadores, é comum questionarmos a respeito de suas aplicações, as quais estão muito mais próximas do que podemos imaginar. Por exemplo, quando trabalhamos com sistemas de equações diferenciais parciais e utilizamos a teoria de semigrupos, nos deparamos com espaço de funções (espaço de dimensão infinita) e operadores não limitados (operadores diferenciais), e o que fazemos é converter o PVIC em um problema de Cauchy Abstrato. Assim, a fim de mostrar a boa colocação e estabilidade para um dado problema, a teoria de semigrupos desenvolve resultados qualitativos em torno do operador associado ao sistema. Neste trabalho, utilizamos os principais resultados dessa teoria, baseados em Pazy (1992) e Liu and Zheng (1999) para mostrar a boa colocação e estabilidade do problema Termoelástico, tal sistema pode ser encontrado em Dafermos (1968) e Rivera (1992). Vale lembrar que o principal objetivo aqui é mostrar com um exemplo simples, como utilizar alguns resultados dessa teoria.

**Palavras-chave:** Estabilidade exponencial, Operadores lineares, Operadores não limitados, Espaços de dimensão infinita, Equações diferenciais parciais

---

## REFERÊNCIAS

- [1] Dafermos, C. M. (1968). “On the existence and the asymptotic stability of solutions to the equations of linear thermoelasticity”. *Archive for Rational Mechanics and Analysis*, 29(4):241–271.
- [2] Liu, Z. and Zheng, S. (1999). *Semigroups Associated with Dissipative Systems*. Taylor & Francis.
- [3] Pazy, A. (1992). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer New York.
- [4] Rivera, J. (1992). “Energy decay rates in linear thermoelasticity”. *Funkcialaj Ekvacioj. Serio Internacia*, 35.

Contato: Bruna Sozzo, bsozzo@lncc.br



# Simulação da produção de óleo por poços horizontais hidráulicamente fraturados em reservatórios de baixa permeabilidade

Camila Lima Chung<sup>1</sup>, Ralph Alves Bini da Silva Almeida<sup>1</sup>, Leonardo Figueira Werneck<sup>1</sup>, Grazione de Souza<sup>1</sup> e Helio Pedro Amaral Souto<sup>1</sup>

<sup>1</sup> Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo/RJ, Brazil

---

## Abstract

Neste trabalho, utilizou-se o Modelo de Fratura Discreta (MFD) no estudo da influência do emprego de poços horizontais hidráulicamente fraturados na determinação do campo de pressões no reservatório e da pressão no poço Karimi-Fard et al. (2004); Cavalcante and Sepehrnoori (2017). Considera-se o escoamento monofásico de um óleo leve em formações de baixa permeabilidade. Um simulador tridimensional de reservatórios de óleo, apropriado para aplicações na escala de campo, empregando o sistema de coordenadas cartesianas e uma abordagem clássica do tipo diferenças finitas, é utilizado. A equação diferencial parcial não-linear que governa o escoamento, escrita em termos da pressão, é resolvida numericamente, após discretização e linearização, usando o método iterativo dos Gradientes Conjugados. O objetivo principal é o de estudar os efeitos combinados das fraturas hidráulicas com o poço no perfil da perda de pressão no poço produtor, considerando diferentes cenários de produção de óleo leve. Os resultados observados mostram como a presença das fraturas favorece o escoamento e como o contraste entre fratura e matriz influencia na sua dinâmica. A API OpenMP Dietrich et al. (2017) foi aplicada no intuito de reduzir o esforço computacional empregado nas simulações, o qual depende, por exemplo, do grau da heterogeneidade em termos das propriedades das fraturas.

**Keywords:** Poço hidráulicamente fraturado, reservatório de óleo, método da fratura discreta, método de diferenças finitas, OpenMP.

---

## REFERENCES

- [1] Cavalcante, Filho, J. S. A. and Sepehrnoori, K. (2017). “Simulation of planar hydraulic fractures with variable conductivity using the embedded discrete fracture model”. *Journal of Petroleum Science and Engineering*, pages 212–122.
- [2] Dietrich, R., Schmitt, F., Grund, A., and Stolle, J. (2017). “Critical-blame analysis for openmp 4.0 offloading on intel xeon phi”. *Journal of Computer and System Sciences*, 125:381–388.
- [3] Karimi-Fard, M., Durlofsky, L., and Aziz, K. (2004). “An efficient discrete-fracture model applicable for general-purpose reservoir simulators”. *SPE Journal*, pages 227–236.



# Aplicação de Redes Complexas para Estudo da Correlação Geográfica de Doenças

Caroline de Oliveira Costa Souza Rosa<sup>1</sup> e Artur Ziviani<sup>1</sup>

<sup>1</sup> Laboratório Nacional de Computação Científica (LNCC), Petrópolis/RJ, Brazil

---

## Resumo

A identificação de características em comum entre doenças pode auxiliar na compreensão das enfermidades em si ou na elaboração de estratégias de saúde pública mais eficientes. Duas ou mais doenças podem estar ligadas, por exemplo, se têm um padrão geográfico de incidência semelhante (Sassolas et al., 2010; Veley and Malka, 2013) ou se são afetadas pelo mesmo gene. A relação entre doenças com base nos genes associados a elas foi obtida através da abordagem de redes complexas (Goh et al., 2007). No entanto, não foram encontrados registros da aplicação dessa metodologia ao problema de correlação entre doenças com base em sua incidência geográfica. Assim, o objetivo deste trabalho é verificar a viabilidade de se utilizar a modelagem por redes complexas para comparar a ocorrência geográfica de doenças. Para isso, serão considerados os 5570 municípios brasileiros e notificações de casos de 23 doenças, cujas ocorrências são acompanhadas e disponibilizadas pelo Sistema Único de Saúde (SUS) do Brasil. A metodologia proposta consiste em elaborar uma rede bipartite de doenças-municípios, em que arestas são criadas entre um município e as doenças que foram registradas nele. Pode-se então extrair uma projeção da rede de doenças. Duas enfermidades estarão fortemente conectadas, se tiverem um número elevado de municípios em comum. Métricas clássicas de Ciência de Redes como diâmetro, grau médio e coeficiente de clusterização são utilizadas para compreensão da estrutura dessa rede complexa. Pode-se obter também uma projeção da rede de municípios, que representa ligações entre cidades com notificações de doenças em comum e cuja estrutura também deve ser avaliada através das métricas clássicas.

**Palavras-chave:** Redes complexas, Doenças, Rede bipartite

---

**Agradecimentos:** Este trabalho teve o apoio de CAPES, CNPq e FAPERJ.

---

## REFERÊNCIAS

- [1] Goh, K.-I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., and Barabási, A.-L. (2007). “The human disease network”. *Proceedings of the National Academy of Sciences*, 104(21):8685–8690.
- [2] Sassolas, G., Hafdi-Nejjari, Z., Schott, A. M., Bournaud, C., Peix, J. L., Orgiazzi, J., Dutrieux-Berger, N., and Borson-Chazot, F. (2010). “Geographical correlation between incidence of benign disease and that of cancer of the thyroid among the population of the Rhône-Alpes région of France”. *European Journal of Endocrinology*, 162(1):127–135.
- [3] Veley, K. and Malka, E. (2013). “The Geographic Correlation Between Lyme Disease Incidence and Degenerative Neurological Disease Mortality: An Ecological Study”. *Value in Health*, 16(7):A340–A341.



# Emprego do *OpenFlow* em *Smart Grid*: Uma Abordagem Utilizando o Ambiente GENI

Cássia C. da Silva Pereira<sup>1</sup>, Raul Vitor Arantes Monteiro<sup>2</sup>, Ailton Akira Shinoda<sup>3</sup>

<sup>1,3</sup> Universidade Estadual Paulista “Júlio Mesquita Filho”, Ilha Solteira/SP, Brasil

<sup>2</sup> Universidade Federal do Mato Grosso, Cuiabá/MT, Brasil

---

## Resumo

A tecnologia da Informação é algo que está cada vez mais presente no cotidiano da sociedade atual e o uso de recursos de informática tem gerado um crescente aumento na demanda por serviços avançados de comunicação. No entanto, muitas vezes a “rigidez” da infraestrutura das redes atuais, acaba por limitar o processo de implantação de algumas tecnologias, resultando em uma busca por arquiteturas de redes inovadoras e flexíveis. Pesquisas e experimentos em plataformas de experimentação (testbeds) de diferentes países têm apontado para o protocolo *OpenFlow*, utilizados em SDNs (*Software Defined Networking*), como uma solução promissora em diversas áreas, incluindo setor elétrico. Utilizar SDN em *Smart Grid* pode ser vantajoso, uma vez que possibilita aos operadores, através de um *software* instalado em um Controlador, monitorar e controlar dinamicamente toda a rede.

O GENI (*Global Environment for Network Innovation*), um ambiente laboratorial para redes e sistemas distribuídos, possibilita experimentos com uma ampla gama de tópicos, no qual muitos podem ser investigados em escalas globais (MAMBRETTI, 2016). Dessa maneira, utilizando o ambiente GENI, este trabalho tem como principal contribuição a aplicação de uma arquitetura dinâmica de comunicação para análise da confiabilidade dos dados da *Smart Grid*, fluxo de dados e segurança da SDN.

Na primeira fase do experimento, que é dividido em 2 fases, através de uma aplicação de monitoramento os dados capturados na simulação de uma microrrede de distribuição serão analisados pelo Controlador SDN afim de verificar anomalias. Na microrrede de distribuição será fornecido o parâmetro de tensão da barra escolhida para monitoramento, esse parâmetro deve estar dentro dos limites estabelecidos por PRODIST (2012). O controlador deverá agir no sentido de restaurar os padrões do nível de tensão.

**Palavras-chave:** OpenFlow, SDN, Smart Grid.

---

## REFERENCIAS

- [1] MAMBRETTI, Joe, et al. Creating a Worldwide Network For The Global Environment fo Network Innovations (GENI) and Related Experimental Environments. InThe GENI Book Springer, Cham, 2016. P. 589-632.
- [1] PRODIST, ANEEL, Agência Nacional de Energia Elétrica Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional, Módulo 8 – Qualidade de Energia Elétrica, 2012.



# Problem Formalization of Embedding Methods for Dynamic Graphs

Claudio Daniel Tenório de Barros<sup>1</sup> and Artur Ziviani<sup>1</sup>

<sup>1</sup> National Laboratory for Scientific Computing (LNCC), Petrópolis/RJ, Brazil

---

## Abstract

Network embedding has arisen as a remarkable machine learning task aiming to represent networks in a low-dimensional vector space in which semantic, relational, and structural information about the graph can be captured (Cai et al., 2018). Applying these representations as features to a machine learning model enables performing tasks such as node classification, ranking and clustering, link prediction and graph classification, among others. Although existing work on network embedding primarily focuses on static graphs, real networks often exhibit dynamical temporal properties, including topological evolution and information diffusion. Hitherto, only few methods to advance graph embedding for dynamic graphs have appeared in the literature (Goyal et al., 2019; Trivedi et al., 2019), thus the area has not fully established and consensual definitions. Therefore, we propose a formal definition for the dynamic network embedding problem that can cover all existing methods, besides allowing the construction of new techniques based on the proposed theory. The formalization of the aforementioned problem unifies current approaches, giving to the scientific community a mathematical and computational foundation to organize previously established works and guide future research. We also propose a taxonomy based on problem settings and algorithmic paradigms (Cai et al., 2018), along with application scenarios and the discussion of the temporal characteristics captured by network embedding. At last, we discuss the possibilities of extending embedding methods to multidimensional graphs (Wehmuth et al., 2016).

**Keywords:** dynamic networks, representation learning, dynamic graph embedding, time-varying graphs

---

## REFERENCES

- [1] Cai, H., Zheng, V. W., and Chang, K. C.-C. (2018). “A Comprehensive Survey of Graph Embedding: Problems, Techniques, and Applications”. *IEEE Transactions on Knowledge and Data Engineering*, 30(9):1616–1637.
- [2] Goyal, P., Chhetri, S. R., and Canedo, A. (2019). “dyngraph2vec: Capturing Network Dynamics using Dynamic Graph Representation Learning”. *Knowledge-Based Systems*.
- [3] Trivedi, R., Farajtabar, M., Biswal, P., and Zha, H. (2019). “DyRep: Learning Representations over Dynamic Graphs”. *International Conference on Learning Representations*.
- [4] Wehmuth, K., Fleury, É., and Ziviani, A. (2016). “On Multiaspect Graphs”. *Theoretical Computer Science*, 651:50–61.

Contato: Claudio Daniel Tenório de Barros, cdtb@lncc.br



# Programação Genética para Problemas de Classificação

Gianni Ribeiro Silvério da Conceição<sup>1</sup> e Camila Silva de Magalhães<sup>1</sup>

<sup>1</sup> Núcleo Multidisciplinar de Pesquisa em Computação- UFRJ, Duque de Caxias/RJ, Brazil

---

## Abstract

A Programação Genética (PG) é um dos algoritmos da área de Computação Evolucionista, que comprehende algoritmos inspirados na teoria da evolução das espécies por seleção natural e na genética de populações biológicas. Estes algoritmos são baseados na sobrevivência do mais apto e na reprodução de indivíduos. Inicialmente, uma população aleatória de indivíduos (possíveis soluções para o problema alvo) é gerada. Para cada indivíduo é atribuído um valor de aptidão, ou *fitness*, que é uma medida de qualidade da solução. Estes indivíduos são submetidos ao processo de evolução, através da aplicação de operadores de seleção, recombinação e mutação [1]. Na PG, proposta por Koza [2], os indivíduos são representados na forma de árvores e, quando aplicada para problemas de classificação, seu objetivo é encontrar a árvore que melhor classifique um conjunto de dados de treinamento e que consiga classificar novos dados. Neste caso, o fitness é medido pela acurácia na classificação do conjunto de dados de treinamento. Os melhores modelos obtidos são, posteriormente, avaliados em um conjunto teste externo, não apresentado anteriormente ao classificador.

Neste trabalho, um algoritmo de PG foi implementado e seu desempenho avaliado para classificação de seis conjunto de dados, geralmente utilizados para teste de classificadores, obtidos do UC Irvine Machine Learning Repository (<https://archive.ics.uci.edu/ml/index.php>). Os resultados do algoritmo de PG, implementado em Python 3.7, foram comparados com resultados da literatura [3], e evidenciaram a competitividade da técnica. Adicionalmente, o desempenho do algoritmo de PG com utilização de uma técnica de *niching* também será avaliado.

**Keywords:** Computação Evolucionista, Programação Genética, Classificação de Dados

---

## REFERENCES

- [1] Eiben, A. and Smith, J. E. (2003). *Introduction to evolutionary computing*. Springer.
- [2] Koza, John R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA: The MIT Press.
- [3] Ramalingaswamy Cheruku, Damodar Reddy Edla, Venkata nareshbabu Kuppili, Ramesh Dharavath, “RST-BatMiner: A fuzzy rule miner integrating rough set feature selection and Bat optimization for detection of diabetes disease” (2018), *Applied Soft Computing* (67):764-780

Contato: Gianni Conceição, ribeirogianni1@hotmail.com



# Modelagem Computacional da Transmissão da Dengue através do Modelo Epidêmico Compartimental SIRS

Gicelle Martins Da Cruz<sup>1</sup>, João David Mangabeira Barbosa<sup>1</sup>, Magna Pereira Amorim<sup>1</sup>, Yoisell Rodríguez Núñez<sup>1</sup> e Vanessa da Silva Garcia<sup>2</sup>

<sup>1</sup> Universidade Federal Fluminense — EEIMVR, Volta Redonda/RJ, Brazil

<sup>2</sup> Universidade Federal Fluminense — ICHS, Volta Redonda/RJ, Brazil

---

## Abstract

A epidemiologia matemática fundamenta-se em hipóteses que quantificam aspectos biológicos da propagação de epidemias e propõe modelos que facilitam a adoção de políticas de controle. Os modelos matemáticos utilizam-se de características intrínsecas de determinadas doenças que envolvem parâmetros que descrevem o comportamento da epidemia ao longo do tempo, facilitando a análise do modelo e a evolução da doença (Anderson and May, 2008). A modelagem matemática determinística em Epidemiologia é descrita por equações diferenciais ordinárias (EDOs) ou sistemas de EDOs (Edelstein-Keshet, 1987; Garner and Hamilton, 2011), que permitem estabelecer uma análise detalhada a respeito da doença que estiver sendo estudada, por exemplo: Dengue, Zika, Chikungunya, Tuberculose, Ebola, AIDS, etc. O modelo explorado nesta pesquisa e descrito por um sistema de EDOs é o SIRS (Hethcote, 1989; Kermack and McKendrick, 1927), um modelo compartmentalizado com os parâmetros biológicos ajustados para a Dengue, que se tornou um dos principais problema de saúde pública no mundo, provocando um enorme ônus financeiro para os governos. Os compartimentos S, I e R representam o grupo da população suscetível de contrair a doença, a parcela da população que contraiu a doença e que pode transmiti-la a outro indivíduo, e a classe da população que recuperou-se da doença, respectivamente. Além das taxas de transição entre os diferentes compartimentos, são consideradas as taxas de natalidade e mortalidade dos indivíduos, caracterizando o estudo com dinâmica vital. As simulações permitem analisar a dinâmica do sistema de forma qualitativa e quantitativa. A importância desse estudo se dá ao fato de quanto mais se conhece a respeito da doença e o modo como ela se propaga, mais eficazes serão os métodos para controlar sua propagação, e até mesmo o estudo de ações preventivas, como por exemplo, campanhas de vacinação.

**Keywords:** Equações Diferenciais, Modelagem Computacional, Epidemiologia Matemática, Modelos Compartimentais

---

## REFERENCES

- [1] Anderson, R. M. and May, R. (2008). “Infectious Diseases of Humans: Dynamics and Control”. *Oxford University Press, 1992. Brauer, F.; van den Driessche, P.; Wu, J. Mathematical Epidemiology (Lecture Notes in Mathematics)*. Springer.
- [2] Edelstein-Keshet, L. (1987). “Mathematical Models in Biology”. *McGraw-Hill, Inc.*
- [3] Garner, M. G. and Hamilton, S. A. (2011). “Principles of epidemiological modelling”. *Rev Sci Tech.*, 30(2):407–416.
- [4] Hethcote, H. W. (1989). “Three basic epidemiological models”. In *Applied mathematical ecology*. Springer Berlin Heidelberg, pages 119–144.
- [5] Kermack, W. and McKendrick, A. (1927). “A Contribution to the Mathematical Theory of Epidemics”. *Proceedings of the Royal Society of London A*, 115:700–721.



# Does the Ataxia group have genetic similarities?

Gustavo Simões Carniali<sup>1</sup> and Sérgio Vale Aguiar Campos<sup>1</sup>

<sup>1</sup> Universidade Federal de Minas Gerais, Belo Horizonte/MG, Brazil

---

## Abstract

Spinocerebellar Ataxias, or SCAs, are a group of 37 known genetically heterogeneous clinical neurodegenerative diseases. The most common type of SCAs have an occurrence of 1 to 5 cases per 100,000 people. Commonly, they attack the nervous system generating coordination loss (Paulson, 2009). This paper analyzes the genetic similarity level between the most common Ataxias, of types 1, 2, 3 and 7, we identify genes that are associated with more than one Ataxia. Studies have shown that genes may relate to other genes, that is, the increased action of a gene can culminate in the increase or decrease of phenotype effects of another gene (Snel et al., 2000). The String tool (Snel et al., 2000) allows the creation of the genetic relationships such as the ones described above. For each of the 4 most common Ataxias, a set of 20 genes linked to the main gene responsible for Ataxia will be obtained by the String tool. It has already been shown that the Ataxias have many proteins in common (Lim et al., 2006). However, a similar study with genes has not yet been done. This study will allow us to understand and classify better these diseases, with benefits such as the reuse of treatments between similar diseases. In addition, a set of 3 neurodegenerative and genetic diseases will have their genes compared to the Ataxia genes, the diseases are: Type 1 Neuronal Lipofuscinose ceroid, Duchenne Dystrophy and Type 1 Spinal Muscular Atrophy. A total of 17 common genes were found among the SCAs, representing a total of 13% of the genes studied. No gene has been found that is in an SCA as well in one of other neurodegenerative diseases. This study shows an indicative that the SCAs have more genes in common, which may explain their similar phenotypic effects and their unique classification as Ataxia.

**Keywords:** Ataxia, Similarities, Gene, Genetic Correlation

---

## REFERENCES

- [1] Lim, J., Hao, T., Shaw, C., Patel, A. J., Szabó, G., Rual, J.-F., Fisk, C. J., Li, N., Smolyar, A., Hill, D. E., et al. (2006). “A protein–protein interaction network for human inherited ataxias and disorders of purkinje cell degeneration”. *Cell*, 125(4):801–814.
- [2] Paulson, H. L. (2009). “The spinocerebellar ataxias”. *Journal of neuro-ophthalmology: the official journal of the North American Neuro-Ophthalmology Society*, 29(3):227.
- [3] Snel, B., Lehmann, G., Bork, P., and Huynen, M. A. (2000). “String: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene”. *Nucleic acids research*, 28(18):3442–3444.



# Análise Mutacional via Implementação do Algoritmo Soma com Transporte

José dos Santos Fernandes<sup>1</sup>, Mariane Moreira de Souza<sup>1</sup> e Anderson José de Oliveira<sup>2</sup>

<sup>1</sup> Departamento de Ciência da Computação, ICEX, Universidade Federal de Alfenas, Alfenas/MG, Brasil

<sup>2</sup> Departamento de Matemática, ICEX, Universidade Federal de Alfenas, Alfenas/MG, Brasil

---

## Resumo

A Modelagem Computacional pode ser utilizada para auxiliar no desenvolvimento de áreas correlatas, como a Matemática Aplicada na modelagem do código genético. Em Oliveira e Palazzo Jr. (2012), é feita a proposta de um algoritmo de Soma com Transporte para análises mutacionais, que acarretam em mudanças durante o processo de síntese proteica, além de caracterizar os rotulamentos do código genético, propostos por Rocha e Palazzo Jr. (2010) e Faria e Palazzo Jr. (2011), apresentando as influências dos mesmos durante o processo. Os objetivos deste trabalho são: a) implementar o algoritmo para realização da Soma com Transporte, b) comprovar que o uso da ferramenta facilita os cálculos envolvidos, além de utilizar os rotulamentos A, B e C, propostos em Rocha e Palazzo Jr. (2010) e Faria e Palazzo Jr. (2011) durante a modelagem, c) analisar o comportamento das sequências, tanto fictícias quanto reais, em especial na geração dos aminoácidos e nos efeitos da mutação em um determinado organismo, d) comparar os resultados entre os rotulamentos e visualizar quais são os diferentes efeitos que uma mesma sequência gera em um organismo, tendo em vista os rotulamentos A, B e C. Os resultados obtidos possibilitam a utilização do algoritmo em problemas de bioinformática e biomedicina, gerando uma ferramenta que pode ser facilmente utilizada pelo usuário final ou pelo pesquisador.

**Palavras Chave:** Síntese Proteica, Rotulamentos do Código Genético, Mutações, Algebra, Soma com Transporte

---

## REFERÊNCIAS

- [1] Rocha, A. S. L. e Palazzo Jr., R. (2010). “Modelo de sistema de comunicações digital para o mecanismo de importação de proteínas mitocondriais através de códigos corretores de erros”. *Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas*, Campinas.
- [2] Faria, L. C. B. e Palazzo Jr., R. (2011). “Existências de Códigos Corretores de Erros e Protocolos de Comunicação em Sequências de DNA”. *Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas*, Campinas.
- [3] Oliveira, A. J. e Palazzo Jr., R. (2012). “Análise Algébrica dos Rotulamentos Associados ao Mapeamento do Código Genético”. *Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas*, Campinas.



# Modelagem Computacional do Câncer de Mama: Estudo Comparativo via Fatores de Tratamento

Josué Cleber de Azevedo Custódio<sup>1</sup>, Felipe Bezerra de Azevedo<sup>2</sup>,  
Yoisell Rodríguez Núñez<sup>2</sup> e Vanessa da Silva Garcia<sup>3</sup>

<sup>1</sup> Universidade Federal Fluminense — ICEx, Volta Redonda/RJ, Brazil

<sup>2</sup> Universidade Federal Fluminense — EEIMVR, Volta Redonda/RJ, Brazil

<sup>3</sup> Universidade Federal Fluminense — ICHS, Volta Redonda/RJ, Brazil

---

## Abstract

O câncer é um evidente problema de saúde pública mundial. Na tentativa de se obter a cura, diversos avanços no tratamento da doença têm sido alcançados através da quimioterapia e outros tratamentos, porém quantitativamente pouco se conhece sobre o efeito das drogas nas células tumorais e nas sadias. Dentre os diversos tipos de câncer, o mais temido pelas mulheres é o de mama, pois, a cada ano, mais de um milhão de novos casos são diagnosticados em todo o mundo. Segundo o Instituto Nacional de Câncer (INCA), em 2018 o câncer de mama atingiu quase 60 mil mulheres no Brasil, destacando-se pela crescente taxa de incidência e pela elevada taxa de mortalidade (INCA, 2018). Dentre as terapias convencionais, além da quimioterapia destaca-se a radioterapia, que destrói as células cancerígenas por meio de radiação ionizante. Sabendo da necessidade de um melhor entendimento quantitativo e analítico, é comum utilizar modelos matemáticos contínuos, baseados em EDOs, como ferramenta de grande valia para tomada de decisões médicas, especialmente quando o diagnóstico e o tratamento são precoces, pois nesses casos a doença apresenta bom prognóstico. Nesta pesquisa, estudamos um modelo matemático contínuo do tipo presa-predador logístico, proposto em Gatenby (1996), que descreve a dinâmica do crescimento tumoral. O objetivo fundamental do nosso estudo consiste em compreender os principais aspectos da dinâmica do crescimento tumoral mamário através dos cenários estabelecidos: para o crescimento tumoral livre de tratamento e o crescimento tumoral inibido por medicação (fator de tratamento) (Rodrigues, 2011). Com a inserção de diferentes fatores de tratamento e considerando dados específicos, já conhecidos na literatura referentes ao câncer de mama, são realizadas simulações numéricas comparativas que mostram que sob a ação dos diferentes tratamentos considerados, ocorre redução na taxa de crescimento do tumor aumentando o tempo de sobrevida do paciente.

**Keywords:** Equações Diferenciais, Modelagem Computacional, Crescimento Tumoral, Fator de Tratamento, Tempo de Sobrevida

---

## REFERENCES

- [1] Gatenby, R. A. (1996). "Application of competition theory to tumour growth: implications for tumour biology and treatment". *European Journal of Cancer*, 32A(4):722–726.
- [2] INCA (2018). "Incidência de Câncer no Brasil". Disponível em: <http://www1.inca.gov.br/estimativa/2018/sintese-de-resultados-comentarios.asp>. Acesso em: 20 de setembro de 2019.
- [3] Rodrigues, D. S. (2011). *Modelagem Matemática em Câncer: Dinâmica Angiogênica e Quimioterapia Anti-Neoplásica*. Dissertação de Mestrado, Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp), IBB.



# Numerical Simulation of Non-Isothermal Water-Oil Flow in Petroleum Reservoirs

Juan Diego dos Santos Heringer<sup>1</sup>, Grazione de Souza<sup>1</sup> e Helio P. Amaral Souto<sup>1</sup>

<sup>1</sup> Rio de Janeiro State University, Nova Friburgo/RJ, Brazil

---

## Abstract

The development and commercial deployment of modern EORs is today seen as the way to extend the oil production in recoverable reserves (Petitfrere, 2014). Enhanced oil recovery(EOR) is normally used to extract an extra amount of oil from the reservoir. One of the main methods considered in EOR are the thermal methods those of which take advantage of the heat addition and reduces the oil viscosity. In this work a three-dimensional non-isothermal two-fluid flow in an oil reservoir with one vertical well and four static heaters is studied. This is an enhancement of the previous work developed by the authors that can be seen in Heringer et al. (2019) , but in this case a two-fluid and a one-equation model is investigated. We use mathematical and physical modeling of a fluid flow in porous media in order to obtain the governing equations for this particular problem. A one equation model is developed in order to solve the two-phase flow problem. For the non-isothermal fluid flow modeling, one equation is used to the temperature variable, obtained from the principal of energy conservation. After this process, a Finite Volume Method is used to discretize the systems of partial differential equations. In order to calculate pressure, temperature and saturation IMPES Method (IMplicit for Pressure and Explicit for Saturation), an operator splitting method (Vennemo, 2016) and a linearization process (Ertekin et al., 2001) are implemented. Two set of linear equations are generated, one for pressure and the other for temperature and, a preconditioned conjugate gradient method is used to solve those set of equations. After the solution of pressure, the saturation is explicitly calculated, then, the temperature is solved.

**Keywords:** Two-phase flow, Finite Volume Methods, One-equation model, Reservoir engineering

---

## REFERENCES

- [1] Ertekin, T., Abou-Kassem, J., and King, G. (2001). *Basic Applied Reservoir Simulation*. Richardson, USA: Society of Petroleum Engineers, Texas.
- [2] Heringer, J. D. S., Debossam, J. G. S., de Souza, G., and Souto, H. P. A. (2019). “Numerical simulation of non-isothermal flow in oil reservoirs using a two-equation model”. *Coupled Systems Mechanics*, 8(2).
- [3] Petitfrere, M. (2014). *EOS based simulations of thermal and compositional flows in porous media*, An optional note. Universit de Pau et des Pays de L'Adours, France.
- [4] Vennemo, S. B. (2016). *Multiscale Simulation of Thermal Flow in Porous Media*. Norwegian University of Science and Technology, Trondheim, Norway.

Contato: Juan Diego dos Santos Heringer, jheringer21@gmail.com



# Neuroevolução aplicada ao *trading* de criptomoedas

Lincon O. Vidal<sup>1</sup>

<sup>1</sup> Laboratório Nacional de Computação Científica, Petrópolis/RJ, Brasil

---

## Abstract

O conceito de criptomoeda foi introduzido com a criação do *bitcoin*, um sistema de dinheiro eletrônico *peer-to-peer* para permitir pagamentos online diretamente de uma parte a outra, sem necessidade de intermediários, como descrito por Nakamoto (2008). Uma década depois do primeiro *bitcoin* criado, já foram registradas mais de 450 milhões de transações na rede descentralizada, embora o uso do ativo se dê em maior parte relacionado à investimento especulativo como mostra Baur et al. (2018). Independente do motivo, pessoas do mundo todo têm, diariamente, transformando dinheiro local em criptomoedas e se aproveitado das variações do mercado ainda volátil para fazer lucro nas operações de *trading*. Como reportado por Khan et al. (2011), com crescimento desse mercado de maneira geral, seja na compra e venda de ações tradicionais ou criptoativos, os investidores vêm buscando ferramentas para prever as variações de preços e automatizar as operações de compra e venda a fim de aumentar seus ganhos e minimizar o risco. Ao investir no mercado financeiro, encontrar os melhores pontos de entrada (*buy*) e saída (*sell*) de forma a obter lucro máximo tem sido um problema difícil, já que são muitos os fatores que podem impactar o preço de ativos, como mostra o trabalho de Chang and Liu (2008). A fim de se obter algum poder preditivo sobre o preço futuro do ativo, duas estratégias são geralmente aplicadas, como detalhado por Kirkpatrick II and Dahlquist (2010): *análise fundamental*, que avalia o valor futuro do ativo com base nas variáveis e métricas relacionadas ao valor intrínseco da empresa a qual o ativo está relacionado, sendo valiosa para os investidores de longo prazo; *análise técnica*, que prevê o valor futuro com base na investigação de movimentos passados, sem tentar medir o valor intrínseco do ativo, envolvendo apenas a análise da atividade do mercado – como preço e volume – para identificar padrões que possam ser usados na tomada de decisão, sendo mais adequada para negociações de curto prazo. Este trabalho apresenta uma abordagem que combina a Análise de Componentes Principais (PCA) com a Neuroevolução de Topologia Aumentada (NEAT) para gerar um sinal de *trading* a partir de indicadores de análise técnica, com objetivo de gerar lucro diário com baixo risco associado na negociação de *bitcoin*. A abordagem proposta foi avaliada com dados diários do par BTC/USDT de duas grandes corretoras de criptomoedas (Binance e Bitstamp). O treinamento dos agentes envolveu a otimização dos pesos e da topologia de redes neurais que operam na compra e venda de *bitcoins* em ambientes de simulação. A partir de dados de movimentação diária (preço, volume etc), foram gerados novos atributos calculados de indicadores de análise técnica, totalizando 121 atributos, que foram posteriormente reduzidos a 11 com a técnica de PCA descrito por Wold et al. (1987). O método de seleção de atributos apresentado no trabalho de Song et al. (2010) foi utilizado para selecionar os 10 atributos que mais contribuíam para a variância do dado dentre os 121 originais, gerando um outro conjunto de treinamento para a construção

de um segundo modelo. Os modelos desenvolvidos e treinados independentemente com cada conjunto de dados – 11 atributos transformados ou 10 melhores atributos originais – e foram comparados à estratégia mais simples de investimento '*Buy Hold*', na qual o agente investidor realiza a compra do ativo no início do período e executa a venda ao final. Os conjuntos foram utilizados no treinamento e avaliação de redes neurais com o algoritmo NEAT, definindo a função objetivo de forma a favorecer as redes com maior lucro médio diário (MDP do trabalho de Nadkarni and Neves (2018)) em simulações de *trading*. Os resultados preliminares obtidos mostram que, mesmo com PCA, o melhor agente não foi capaz de obter margem de lucro média por período maior do que a estratégia *BH*, embora a tendência de alta do mercado presente nos conjuntos de teste (em decorrência do período recente selecionado) pode estar associada à incapacidade do modelo de ser mais lucrativo que a estratégia mais simples para mercados de alta. Atributos transformados por PCA obtiveram melhor generalização no conjunto externo de teste (Binance) quando comparados aos atributos selecionados manualmente (10 melhores), o que pode indicar um caminho para redução da dimensionalidade nesse tipo de dado. Apesar da lucratividade baixa, os resultados preliminares indicam a viabilidade da evolução agentes autônomos de *trading* baseados em redes neurais.

**Keywords:**neuroevolução, bitcoin, trading, aprendizado de máquina, redes neurais

---

## REFERENCES

- [1] Baur, D. G., Hong, K., and Lee, A. D. (2018). “Bitcoin: Medium of exchange or speculative assets?” *Journal of International Financial Markets, Institutions and Money*, 54:177–189.
- [2] Chang, P.-C. and Liu, C.-H. (2008). “A tsk type fuzzy rule based system for stock price prediction”. *Expert Systems with applications*, 34(1):135–144.
- [3] Khan, Z. H., Alin, T. S., and Hussain, M. A. (2011). “Price prediction of share market using artificial neural network (ann)”. *International Journal of Computer Applications*, 22(2):42–47.
- [4] Kirkpatrick II, C. D. and Dahlquist, J. A. (2010). *Technical analysis: the complete resource for financial market technicians*. FT press.
- [5] Nadkarni, J. and Neves, R. F. (2018). “Combining neuroevolution and principal component analysis to trade in the financial markets”. *Expert Systems with Applications*, 103:184–195.
- [6] Nakamoto, S. (2008). “A peer-to-peer electronic cash system”. *Bitcoin*. URL: <https://bitcoin.org/bitcoin.pdf>.
- [7] Song, F., Guo, Z., and Mei, D. (2010). “Feature selection using principal component analysis”. In: *2010 international conference on system science, engineering design and manufacturing informatization*, pages 27–30.
- [8] Wold, S., Esbensen, K., and Geladi, P. (1987). “Principal component analysis”. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52.



# Heat Transfer Solution for Circular Channels with Axial Heat Diffusion using Integral Transforms

Lívia Mantuano Corrêa<sup>1</sup> and Daniel José Nahid Mansur Chalhub<sup>1</sup>

<sup>1</sup> Graduate Program of Mechanical Engineering (PPG-EM), Group for Studies and Environmental Simulations in Reservoirs (GESAR), Rio de Janeiro State University (UERJ), Rua Fonseca Teles, 121, 20940-903, Rio de Janeiro, RJ, Brazil

---

## Abstract

The present paper proposes a semi-analytical solution for an extended version of Graetz problem for slip-flow in a duct with axisymmetric flow. The steady-state formulation considers axial heat diffusion in a semi-infinite channel using a single domain. The duct temperature profile is determined using the Generalized Integral Transform Technique (GITT). The work also aims to analyze Nusselt number distribution for the simulated cases and the Nusselt convergence rate. The axial diffusion effect was also considered and analyzed, varying Péclet number. Through mathematical operations, a dimensionless generalized formulation is defined, and using the GITT combined with Bessel eigenvalue problem, the eigenvalues are calculated by a numerical method. Although the original problem transformation takes to a coupled ordinary differential equations system(ODE), in this work, an analytical matrix method is used, whose solution is obtained in terms of eigenvalues and eigenvectors from a matrix involving the coupling coefficients of the ODEs system. Thus, only one numerical implementation is necessary, that is the calculation of eigenvalues and eigenvectors from the coefficient matrix mentioned. Finally, the temperature profile and the Nusselt number behavior are obtained. By the results, GITT is an efficient technique according to the convergence analysis, presenting worse rates in the channel entrance; Péclet value influences the way as the flow develops thermally, as well as Biot and the duct aspect ratio, and as Péclet raises, Nusselt approaches the solution for the case with no axial diffusion, as expected.

**Keywords:** Graetz problem, Generalized Integral Transform Technique, Axial heat diffusion, Orthotropic tube, Axisymmetric flow

---



# Identificação de bolsões específicos como potenciais inibidores da proteína NS2B-NS3 dos principais Flavivírus de importância médica

Luan Antônio de Moraes<sup>1</sup>, Gabriela de Lima Menezes<sup>2</sup> e Roosevelt Alves da Silva<sup>3</sup>

<sup>1,2,3</sup> Universidade Federal de Goiás, Jataí/GO, Brazil

---

## Abstract

Arbovírus do gênero Flavivírus como os vírus da dengue, febre amarela e Zika, vêm ganhando notoriedade devido aos crescentes casos observados e, principalmente, por não ter ainda vacinas e fármacos disponíveis para profilaxia e tratamento.

A NS2B-NS3 protease, transcrita por todos os representantes desse gênero viral, é uma das proteínas responsáveis por atuar na replicação e maturação viral, se tornando atrativa para o desenvolvimento de fármacos que possam inibir a multiplicação do vírus. (Li Y., 2017)

Com este objetivo, busca-se neste trabalho identificar bolsões de ligações nas proteínas NS2B-NS3 dos principais Flavivírus que permitam a ligação de inibidores. No entanto, serão avaliados quais bolsões nestas proteínas podem ser mais interessantes para este propósito, observando quais deles não estão presentes em estruturas de proteínas humanas já determinadas.

A comparação entre estas estruturas será realizada usando o programa TM-align (Y. Zhang, 2005) e outros programas para caracterizar as similaridades entre bolsões de ligação de compostos. Os resultados deste trabalho fornecerão informações importantes a respeito da especificidade dos bolsões a fim de avaliar as reações cruzadas dos compostos já propostos na literatura, além de contribuir como uma estratégia racional para a triagem de compostos mais seletivos contra estes vírus e com menos prejuízos à saúde humana.

**Keywords:** Flavivírus, Proteínas, Docking, Ligantes, Fármacos

---

## REFERENCES

- [1] Li Y., Zhang Z., P. W. L. Y. W. W. L. S. C. M. H. A. K. T. L. D. K. C. (2017). "Structural Dynamics of Zika Virus NS2B-NS3 Protease Binding to Dipeptide Inhibitors". *Structure*, 25(8):1242–1250.
- [2] Y. Zhang, J. S. (2005). "TM-align: A protein structure alignment algorithm based on TM-score". *Nucleic Acids Research*, 33:2302–2309.

Contato: Luan Antônio de Moraes, luanmoraes1997@gmail.com



# Modelagem da Vascularização Arterial do Rim Humano

Luís Fernando Mendes Cury<sup>1</sup>, Caio Sergio Bianchi Reis Guimarães<sup>3</sup>, Mauricio Younes-Ibrahim<sup>2,3</sup> e Pablo Javier Blanco<sup>1</sup>

<sup>1</sup> Laboratório Nacional de Computação Científica, Petrópolis/RJ, Brasil

<sup>2</sup> Departamento de Medicina Interna da Faculdade de Ciências Médicas da UERJ, Rio de Janeiro/RJ, Brasil

<sup>3</sup> Disciplina de Nefrologia do Departamento de Medicina da PUC-Rio, Rio de Janeiro/RJ, Brasil

---

## Abstract

Os rins são órgãos vitais para o corpo humano responsáveis pelo balanço hidroeletrolítico e de água, pela excreção de produtos metabólicos e de substâncias bioativas, pela regulação da pressão arterial, pela regulação da produção de células vermelhas, de vitamina D e de gluconeogênese. Cada rim é suprido por uma ou mais artérias renais, que se dividem sequencialmente em artérias interlobares, artérias arqueadas, artérias interlobulares, arteríolas aferentes e eferentes, interpostas pelos capilares glomerulares. As arteríolas eferentes deixam o glomérulo e se dividem, parte em capilares peritubulares, a partir das arteríolas que saem de néfrons corticais, e, parte em *vasa recta*, a partir das arteríolas que saem de néfrons justamedulares. Os capilares peritubulares e a *vasa recta*, em sequência, dão origem às veias interlobulares, veias arqueadas, veias interlobares e veia renal. O objetivo deste trabalho é utilizar ferramentas computacionais desenvolvidas no grupo HeMoLab para gerar as árvores arteriais renais em modelos bidimensionais e tridimensionais. A modelagem terá a resolução espacial adequada às artérias interlobulares. O método usado é baseado numa adaptação do *Constrained Constructive Optimization* (CCO), que gera as árvores a partir da minimização de um funcional custo associado ao volume dos elementos vasculares.

**Keywords:** Anatomia Vascular, Rim, Hemodinâmica, Modelagem

---

## Referências

- [1] Blanco, PJ de Queiroz, R. B. a. F. R. (2013). “A computational approach to generate concurrent arterial networks in vascular territories”. *International journal for numerical methods in biomedical engineering*, 29(5):601–614.
- [2] Eaton, Douglas C. Pooler, J. P. (2009). *Vander's Renal Physiology*. McGraw-Hill Companies.
- [3] Khamanarong, K., Prachaney, P., Utraravichien, A., Tong-Un, T., and Sriparaya, K. (2004). “Anatomy of renal arterial supply”. *Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists*, 17(4):334–336.
- [4] Layton, Anita T. Edwards, A. (2014). *Mathematical Modeling in Renal Physiology*. Springer.
- [5] Mohamed, Tahagod Sequeira-Lopez, M. L. S. (2018). “Development of the renal vasculature”. Elsevier.

Contato: Luís F. M. Cury, luiscury@lncc.br



# Controle ótimo aplicado a um modelo epidemiológico compartimental de transmissão da dengue

Magna Pereira Amorim<sup>1</sup>, Yoisell Rodríguez Núñez<sup>1</sup>, Panters Rodríguez Bermúdez<sup>1</sup> e  
Vanessa da Silva Garcia<sup>2</sup>

<sup>1</sup> Universidade Federal Fluminense - EEIMVR, Volta Redonda/RJ, Brasil

<sup>2</sup> Universidade Federal Fluminense - ICHS, Volta Redonda/RJ, Brasil

---

## Abstract

Nas últimas décadas, a prevalência global da dengue aumentou consideravelmente e se tornou um dos principais problemas de saúde pública no mundo. A dengue é uma doença transmitida por vetores, e atualmente, a única maneira de controlar a doença é minimizar a população de vetores. Prevê-se que a vacina contra a dengue, para prevenção eficaz e controle de longo prazo, seja a solução. Através de modelos matemáticos epidemiológicos a modelagem matemática desempenha um papel fundamental no estudo da evolução de doenças infecciosas (Anderson et al., 1991; Hethcote, 2000; Siettos and Russo, 2013). Nesta pesquisa exploramos um modelo epidémico da dengue com controle, considerando o modelo epidemiológico compartimental SIR + ASI (Rodrigues et al., 2014; Schreppel and Chudej, 2018), definido por um sistema de equações diferenciais ordinárias (EDOs), que permite avaliar os padrões de transmissão e controle da dengue. São considerados dois tipos de populações neste modelo: hospedeiros e vetores. Os hospedeiros (humanos) são divididos em três classes complementares: suscetível,  $S_h$ , indivíduos que podem contrair a doença; infectado,  $I_h$ , indivíduos capazes de transmitir a doença a outros; e resistente,  $R_h$ , indivíduos que adquiriram imunidade. Da mesma forma, o modelo também possui três compartimentos para os vetores, ou seja, mosquitos fêmeas (mosquitos machos não são considerados porque não picam seres humanos e, consequentemente, não influenciam a dinâmica da doença):  $A_m$ , que representa a fase aquática da mosquito (incluindo ovo, pupas e larvas) e a fase adulta do mosquito, com  $S_m$  e  $I_m$ , suscetíveis e infectados, respectivamente. Nos compartimentos  $S_h$  e  $R_h$  é introduzida a variável de controle  $u$  para estudar a melhor estratégia de vacinação (estratégia ideal), considerando os custos de tratamento de indivíduos infectados e os custos de vacinação, minimizando a variável de controle via Princípio Máximo de Pontryagin (Bakare et al., 2014; Di Liddo, 2016; Fitria et al., 2017; Pontryagin et al., 1962). Além desse estudo da variável de controle, a pesquisa consiste no desenvolvimento da análise qualitativa do modelo, identificando as condições de estabilidade em torno dos pontos críticos, o plano de fase para cada ponto de equilíbrio e o número reprodutivo básico relativo ao modelo estudado, para entender sua influência na propagação da doença.

---

**Keywords:** Modelagem Computacional, Modelo Epidemiológico, Dengue, Controle.

## REFERENCES

- [1] Anderson, R. M., Anderson, B., and May, R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press.
- [2] Bakare, E. A., Nwagwo, A., and Danso-Addo, E. (2014). “Optimal control analysis of an sir epidemic model with constant recruitment”. *International Journal of Applied Mathematics Research*, 3(3):273.
- [3] Di Liddo, A. (2016). “Optimal control and treatment of infectious diseases. the case of huge treatment costs”. *Mathematics*, 4(2):21.
- [4] Fitria, I., Winarni, Pancahayani, S., and Subchan (2017). “An optimal control strategies using vaccination and fogging in dengue fever transmission model”. In: *AIP Conference Proceedings*, AIP Publishing, pages 020068.
- [5] Hethcote, H. W. (2000). “The mathematics of infectious diseases”. *SIAM review*, 42(4):599–653.
- [6] Pontryagin, L. S., Mishchenko, E. F., Boltyanskii, V. G., and Gamkrelidze, R. (1962). *The Mathematical Theory of Optimal Processes*. John Wiley & Sons, New York.
- [7] Rodrigues, H. S., Monteiro, M. T. T., and Torres, D. F. (2014). “Vaccination models and optimal control strategies to dengue”. *Mathematical Biosciences*, 247:1–12.
- [8] Schreppel, C. and Chudej, K. (2018). “Numerical optimal control applied to an epidemiological model”. *IFAC-PapersOnLine*, 51(2):1–6.
- [9] Siettos, C. I. and Russo, L. (2013). “Mathematical modeling of infectious disease dynamics”. *Virulence*, 4(4):295–306.



# Modelagem computacional de processos de contaminação de aquíferos considerando biodegradação e adsorção

Marcelo Lemos da Silva<sup>1</sup>, João Gabriel de Souza Debossam<sup>1</sup> e Grazione de Souza<sup>1</sup>

<sup>1</sup> Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo/RJ, Brazil

---

## Abstract

Os recursos hídricos são atingidos por processos de poluição, apesar de a água constituir-se em um recurso fundamental para a sobrevivência dos seres humanos e ter aproveitamentos em aplicações que envolvem irrigação, processos industriais, transporte, obtenção de energia e higiene, por exemplo. No caso dos aquíferos, formações rochosas subterrâneas onde há água armazenada, processos de contaminação podem inviabilizar a utilização. Neste contexto, a modelagem e a simulação numérica de escoamentos em meios porosos podem ser aplicadas no entendimento da dinâmica da contaminação de aquíferos e na elaboração de estratégias de contenção e/ou remediação Datta et al. (2009); Zhong et al. (2008); McLaren and Illman (2012). Neste trabalho, estudou-se via simulação numérica o processo de contaminação em um aquífero. Utilizou-se o método de diferenças finitas para a solução em termos da pressão da água e da concentração do contaminante, tendo sido estudados os efeitos da adsorção e da biodegradação combinados na dinâmica do escoamento. Os resultados aqui obtidos mostram a influência de diferentes efeitos físicos, tais como a difusão e advecção, assim como numéricos, como a difusão numérica. Neste sentido, estratégias foram avaliadas para a redução desta, em um estudo comparativo, incluindo testes para limitadores de fluxo. Cabe destacar que a equação diferencial parcial para a incógnita concentração é não linear.

**Keywords:** Aquíferos, Contaminação, Método de diferenças finitas, Adsorção, Biodegradação

---

## REFERENCES

- [1] Datta, B., Chakrabarty, D., and Dhar, A. (2009). “Simultaneous identification of unknown groundwater pollution sources and estimation of aquifer parameters”. *Journal of Hydrology*, 376:48–57.
- [2] McLaren, R.G., S. E. P. Y.-J. and Illman, W. (2012). “Numerical simulation of dnapl emissions and remediation in a fractured dolomitic aquifer”. *Journal of Contaminant Hydrology*, 136-137:56–71.
- [3] Zhong, L., Oostrom, M., Wietsma, T., and Covert, M. (2008). “Enhanced remedial amendment delivery through fluid viscosity modifications: experiments and numerical simulations”. *Journal of Contaminant Hydrology*, 101:29–41.



# Study of *Ideonella sakaiensis* PETase enzyme via Computational Molecular Biophysics

Marcio Fernandes da Silva<sup>1</sup>, Stéphano Andrade de Vasconcelos<sup>1</sup> e Marcos Serrou do Amaral<sup>1</sup>

<sup>1</sup> Laboratório de Modelagem Molecular, Instituto de Física, UFMS, Campo Grande/MS, Brazil

---

## Abstract

The problem of plastic waste on the environment is well known. The industry most commonly used plastic is polyethylene terephthalate (PET) (Kibria, 2017). Its durability and chemical inertia are commercially advantageous although very harmful for the environment (Fecker et al., 2018). Hence studying degrading plastic microorganisms had become an interesting wayout. For the referred problem, the bacteria *Ideonella sakaiensis* (*Is*) is an excellent candidate since it can break PET polymers in its former constituents via an enzyme called PETase (*IsPETase*) (Yoshida et al., 2016). Furthermore, it has its optimal thermal spot on room temperature and the bacteria feeds itself from the PET hydrolyses products (Ma et al., 2018). Just as the role of water on biological interactions (Bellissent-Funel et al., 2016) and the behavior of lipases in the presence of several solvents were studied via computational methods (Mohtashami et al., 2019), so was the case for PETase in the present work. Here Constant pH Molecular Dynamics (CpHMD) (Swails et al., 2014) were used to study functional/structural features of *IsPETase* (PDB id: 6EQE).The AMBER18 and AmberTools software packages were used for all simulations and analysis (Case et al., 2018). *ff14SB* force field (Maier et al., 2015) was employed for the atomic/molecular interactions description of the system. The *IsPETase* was diluted in water (TIP3P model) and a 12 Å cutoff distance was settled for the electrostatic interactions. Temperatures studied were 300 K and 330 K both at pH 7.0 and pH 9.0. So far, the MD has confirmed a stronger pH-dependence of the residues protonations states instead of a temperature-dependence. The RMSD values were lower at pH 9.0.

---

**Keywords:** Biophysics, Protein analysis, Computational modeling

---

## REFERENCES

- [1] Bellissent-Funel, M.-C., Hassanali, A., Havenith, M., Henchman, R., Pohl, P., Sterpone, F., van der Spoel, D., Xu, Y., and Garcia, A. E. (2016). “Water determines the structure and dynamics of proteins”. *Chemical Reviews*, 116(13):7673–7697. PMID: 27186992.
- [2] Case, D., Ben-Shalom, I., Brozell, S., Cerutti, D., Cheatham, T., III, Cruzeiro, V., Darden, T., Duke, R., Ghoreishi, D., Gilson, M., nad A.W. Goetz, H. G., Greene, D., Harris, R., Homeyer, N., Izadi, S., Kovalenko, A., Kurtzman, T., Lee, T., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Mermelstein, D., Merz, K., Miao, Y., Monard, G., Nguyen, C., Nguyen, H., Omelyan, I., Onufriev, A., Pan, F., Qi, R., Roe, D., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C., Smith, J., Salomon-Ferrer, R., Swails, J., Walker, R., Wang, J., Wei, H., Wolf, R., Wu, X., Xiao, L., York, D., and Kollman, P. (2018). “Amber 18”.
- [3] Fecker, T., Galaz-Davison, P., Engelberger, F., Narui, Y., Sotomayor, M., Parra, L. P., and Ramírez-Sarmiento, C. A. (2018). “Active site flexibility as a hallmark for efficient pet degradation by i. sakaiensis petase”. *Biophysical Journal*, 114(6):1302 – 1312.
- [4] Kibria, G. (2017). “Plastic waste, plastic pollution- a threat to all nations”.
- [5] Ma, Y., Yao, M., Li, B., Ding, M., He, B., Chen, S., Zhou, X., and Yuan, Y. (2018). “Enhanced poly(ethylene terephthalate) hydrolase activity by protein engineering”. *Engineering*, 4(6):888 – 893.
- [6] Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). “ff14sb: Improving the accuracy of protein side chain and backbone parameters from ff99sb”. *Journal of Chemical Theory and Computation*, 11(8):3696–3713. PMID: 26574453.
- [7] Mohtashami, M., Fooladi, J., Haddad-Mashadrizeh, A., Housaindokht, M. R., and Monhemi, H. (2019). “Molecular mechanism of enzyme tolerance against organic solvents: Insights from molecular dynamics simulation”. *International Journal of Biological Macromolecules*, 122:914 – 923.
- [8] Swails, J. M., York, D. M., and Roitberg, A. E. (2014). “Constant ph replica exchange molecular dynamics in explicit solvent using discrete protonation states: Implementation, testing, and validation”. *Journal of Chemical Theory and Computation*, 10(3):1341–1352. PMID: 24803862.
- [9] Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., Toyohara, K., Miyamoto, K., Kimura, Y., and Oda, K. (2016). “A bacterium that degrades and assimilates poly(ethylene terephthalate)”. *Science*, 351(6278):1196–1199.



# Modelagem Computacional do Crescimento Tumoral Mamário com Quimioterapia

Matheus Fonseca da Silva<sup>1</sup>, Paulo Raul Cardoso Faria<sup>1</sup>, Carlos Alberto Reis Vale Júnior<sup>1</sup>, Yoisell Rodríguez Núñez<sup>1</sup> e Vanessa da Silva Garcia<sup>2</sup>

<sup>1</sup> Universidade Federal Fluminense — EEIMVR, Volta Redonda/RJ, Brazil

<sup>2</sup> Universidade Federal Fluminense — ICHS, Volta Redonda/RJ, Brazil

---

## Abstract

O câncer de pulmão e o de mama são os mais diagnosticados no mundo todo, são também os principais responsáveis pelo número de mortes causados por esta doença. No âmbito mundial, o câncer mamário corresponde à principal causa de morte entre as mulheres. No Brasil o cenário não é diferente, uma vez que, segundo o Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA), esta patologia representa o tipo mais comum na população feminina (INCA, 2019). Desta forma, considerando o impacto desta patologia no âmbito da saúde mundial, faz-se necessário estudos aprofundados que possibilitem a compreensão do crescimento tumoral, bem como a resposta aos tratamentos quimioterápicos existentes. Neste trabalho são explorados dois sistemas de equações diferenciais ordinárias (EDOs) que modelam o crescimento tumoral mamário. O primeiro deles considera que não há tratamento (Gatenby, 1996), ou seja, sem inserção de um agente quimioterápico e, o segundo, aborda um tratamento quimioterápico (Rodrigues, 2011). Esses modelos de EDOs representativos do câncer com ação da quimioterapia são baseados em duas populações de células (tumorais e normais), que competem entre si por recursos limitados, como o oxigênio e espaço tecidual. A partir das simulações realizadas são gerados gráficos que mostram o comportamento das células tumorais e normais diante do aumento ou redução de massa corpórea do paciente junto à ação da quimioterapia, que pode ser administrada de forma contínua ou por ciclos. Os resultados indicam que o aumento de massa interfere na ação do agente quimioterápico prejudicando sua eficiência. E a redução de massa, por sua vez, potencializa o efeito tóxico da quimioterapia agravando o quadro geral do paciente. Com esta pesquisa esperamos contribuir para a melhoria do atendimento, rastreamento e diagnóstico do câncer de mama.

**Keywords:** Equações Diferenciais, Modelagem Computacional, Crescimento Tumoral, Métodos Numéricos

---

## REFERENCES

- [1] Gatenby, R. A. (1996). "Application of competition theory to tumour growth: implications for tumour biology and treatment". *European Journal of Cancer*, 32A(4):722–726.
- [2] INCA (2019). "A mulher e o câncer de mama no Brasil". Disponível em: <https://www.inca.gov.br/exposicoes/mulher-e-o-cancer-de-mama-no-brasil>. Acesso em: 15 de setembro de 2019.
- [3] Rodrigues, D. S. (2011). *Modelagem Matemática em Câncer: Dinâmica Angiogênica e Quimioterapia Anti-Neoplásica*. Dissertação de Mestrado, Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp), IBB.



# Engenharia Virtual da PETase termoestável da *Ideonella sakaiensis* para degradação altamente eficiente do PET

Stéphano Andrade de Vasconcelos<sup>1</sup>, Sérgio Leandro Espindola Preza<sup>2</sup>, Marcio Fernandes da Silva<sup>1</sup> e Marcos Serrou do Amaral<sup>1</sup>

<sup>1</sup> Laboratório de Modelagem Molecular, Instituto de Física, UFMS, Campo Grande/MS, Brasil

<sup>2</sup> UCDB, Campo Grande/MS, Brasil

---

## Abstract

No mundo moderno são produzidos mais de 300 milhões de toneladas de plástico anualmente incluindo garrafas, malas, bolsas, etc. Dado que metade destes itens são descartáveis, eles contribuem para o acúmulo de plástico na natureza. Estima-se que até 2050 a população do planeta vai aumentar em 40%. Já para o consumo de plástico, no mesmo período, estima-se um aumento de 300%. Isso vai contribuir para o agravamento do acúmulo de plástico (Programme, 2018). Diante disso fica clara a importância para solução deste problema. Em 2016 foi descoberta bactéria *Ideonella Sakaiensis* a qual apresentou habilidade em degradar plástico (Yoshida et al., 2016). Ela sintetiza duas enzimas capazes de degradar o PET: PETase e MHETase. Elas catalizam a hidrólise do PET, transformando-o em polímeros menores: Mono(2-hidroxietilíco) ácido tereftálico (MHET), Bis(2-hidroxietilíco) ácido tereftálico (MHET), Ácido Tereftálico (TPA) e Etileno glicol (EG) (Austin et al., 2018). Este trabalho tem por objetivo fazer mutações virtuais na PETase para obtenção de melhor estabilidade térmica, mantendo-se a atividade enzimática, através de Dinâmica Molecular (DM). A estrutura da PETase foi obtida através do *Protein Data Bank* (Código PDB: 6EQE) (Berman et al., 2000) As análises foram realizadas com o pacote AMBER 18 e as visualizações com os programas VMD e Chimera. A PETase foi submetida a diferentes temperaturas: 300 K e 326 K. Os sistemas estudados foram PETase natural e duas variantes da PETase, apresentando as seguintes mutações: a) Arginina 280 substituída por uma Alanina; e b) Serina 121 substituída por um Ácido aspartâmico, Ácido Aspartâmico 186 substituído por uma Histidina e a Arginina 280 substituída por uma Alanina. O pH foi mantido constante em 7.0 durante as simulações.

---

**Keywords:** Polietileno Tereftálico, Hidrólise, Modelagem Computacional

---

## REFERENCES

- [1] Austin, H. P., Allen, M. D., Donohoe, B. S., Rorrer, N. A., Kearns, F. L., Silveira, R. L., Pollard, B. C., Dominick, G., Duman, R., El Omari, K., Mykhaylyk, V., Wagner, A., Michener, W. E., Amore, A., Skaf, M. S., Crowley, M. F., Thorne, A. W., Johnson, C. W., Woodcock, H. L., McGeehan, J. E., and Beckham, G. T. (2018). "Characterization and engineering of a plastic-degrading aromatic polyesterase". *Proceedings of the National Academy of Sciences*, 115(19):E4350–E4357.
- [2] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). "The Protein Data Bank". *Nucleic Acids Research*, 28(1):235–242.
- [3] Programme, U. N. E. (2018). "SINGLE-USE PLASTICS:A Roadmap for Sustainability". pages 1–104.
- [4] Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., Toyohara, K., Miyamoto, K., Kimura, Y., and Oda, K. (2016). "A bacterium that degrades and assimilates poly(ethylene terephthalate)". *Science*, 351(6278):1196–1199.



# Evolução Diferencial para Classificação de Dados

Taciano dos Passos Ferreira Pinheiro<sup>1</sup> and Camila Silva de Magalhães<sup>2</sup>

<sup>1</sup> Núcleo Multidisciplinar de Pesquisa em Computação, UFRJ, Duque de Caxias/RJ, Brazil

---

## Abstract

Algoritmos evolucionistas são métodos estocásticos baseados na Teoria da Evolução por Seleção Natural, proposta por Darwin, em que indivíduos mais adaptados têm maiores chances reprodutivas e dão continuidade às espécies. Estes algoritmos têm obtido sucesso na solução de problemas complexos de otimização em várias áreas do conhecimento (Eiben and Smith, 2003). O algoritmo de Evolução Diferencial (ED) é um dos algoritmos da área de Computação Evolucionista, proposto por Storn e Price (Storn and Price, 1997), que tem ganho popularidade pela simplicidade de implementação e eficácia na solução de problemas de otimização complexos (Swagatam et al., 2016). Neste trabalho, propomos uma metodologia que integra o algoritmo de ED com uma função polinomial (ED+FP) visando a sua utilização para classificação de dados. O principal objetivo deste trabalho é, portanto, avaliar o desempenho do algoritmo ED+FP proposto, para classificação de conjunto de dados obtidos do *UC Irvine Machine Learning Repository* (<http://archive.ics.uci.edu/ml>), geralmente utilizados para teste de classificadores. Resultados preliminares com o conjunto de dados Iris apresentaram alta acurácia nos conjuntos de treinamento e teste, podendo ser um indicativo do potencial da metodologia proposta.

**Keywords:** Evolução Diferencial, Função Polinomial, Classificação de Dados

---

## REFERENCES

- [1] Eiben, A. and Smith, J. (2003). *Introduction to Evolutionary Computing*.
- [2] Storn, R. and Price, K. (1997). “Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces”. *Journal of global optimization*, 11(4):341–359.
- [3] Swagatam, D., Mullick, S., and Ponnuthurai, S. S. (2016). “Recent Advances in Differential Evolution: An Updated Survey”. *Swarm and Evolutionary Computation*, 27:1–30.



# Atenuação da Dimensionalidade de Dados para Detecção de Anomalias Através da Utilização de Mapas Auto Organizáveis

Thiago Moeda Sant'Anna<sup>1</sup> and Fabio Porto<sup>1</sup>

<sup>1</sup> LNCC-DEXL, Petrópolis/RJ, Brasil

---

## Abstract

A utilização ubíqua de sensores no monitoramento de ambientes e processos permite interpretar o comportamento de fenômenos a partir de séries temporais multi-variáveis. Neste contexto, a detecção de eventos anômalos aparece como um desafio importante, principalmente em fenômenos com análise em tempo real. O tema principal deste trabalho tem como objetivo um estudo detalhado sobre a detecção automática dos eventos onde o fluxo de dados é continuamente acumulado e reprocessado ao longo do tempo. É útil que se tenha um sistema de supervisão automática, para que se possa monitorar quaisquer observações anormais como medidas preventivas ou de emergência. Considerando a base de dados como um conjunto extenso de variáveis independentes, o estudo de *clusters* e *outliers* em função da heterogeneidade do conjunto de dados tornam-se fundamentais para a descoberta de padrões frequentes. Métodos baseados em *cluster* e em densidade indicam que regiões densas nos dados são encontradas diretamente e os *outliers* são definidos como os pontos que não residem nessas regiões, enquanto clustering tenta determinar grupos de pontos de dados similares, outliers são pontos de dados individuais remanescentes de uma certa amostra (Aggarwal and Zhai, 2012). O desafio está no processo de localização destas anomalias pelo fato de que os dados anômalos geralmente podem estar em subespaços de baixa dimensão com características irrelevantes ocultando-os (Sathe and Aggarwal, 2016). Nesse sentido, propomos uma análise do método automático de análise de dados baseado em mapas de recursos auto organizáveis (SOFMs). Esta técnica possibilita a atenuação de um determinado conjunto de dados com uma alta dimensão para dimensões mais baixas, basicamente para uma dimensão com topologia plana, porém preservando as relações métricas relevantes entre os dados onde a distância euclidiana entre a observação dos dados de teste e o centroide dos dados vizinhos é caracterizada como um indicador (Kohonen et al., 2001). Portanto, percebe-se que como se trata de uma técnica de investigação indireta em um ambiente não-linear, onde a correlação não implica em causalidade, esperamos, de certa forma, contribuir consistentemente na detecção de padrões de eventos ocasionados pelas frequências anômalas dos conjuntos de dados das séries temporais. Pretendemos avaliar as técnicas investigadas no problema de detecção de eventos anômalos associados ao processo de perfuração de poços de petróleo.

**Keywords:** Anomalias, Mapas de recursos auto organizáveis, Aprendizagem de máquina

---

## REFERENCES

- [1] Aggarwal, C. and Zhai, C. (2012). *Mining Text Data*. Springer New York.
- [2] Kohonen, T., Huang, T., and Schroeder, M. (2001). *Self-Organizing Maps*. Springer Berlin Heidelberg.
- [3] Sathe, S. and Aggarwal, C. C. (2016). “Subspace outlier detection in linear time with randomized hashing”. pages 459–468.

