III SIMPÓSIO DE AUTOMAÇÃO E SISTEMAS



▪ ANAIS ▪

III Simpósio de Automação e Sistemas

Pesquisa, inovação e aplicações em automação

LAUT / DCA / CT / UFRN

NATAL – RN

2026

Organizadores: LAUT/DCA/UFRN

Anais do III Simpósio de Automação e Sistemas

III SAS

Organização
Laboratório de Automação (LAUT)
Departamento de Engenharia de Computação e Automação
Universidade Federal do Rio Grande do Norte

Natal – RN
2026

Apresentação

O III Simpósio de Automação e Sistemas (III SAS) é um evento promovido pelo Laboratório de Automação (LAUT), vinculado ao Departamento de Engenharia de Computação e Automação da Universidade Federal do Rio Grande do Norte (DCA/UFRN). Realizado nos dias 12 e 13 de novembro de 2026, em Natal/RN, o simpósio reúne atividades de discussão, divulgação científica e apresentação de trabalhos relacionados à automação e a sistemas inteligentes.

Esta edição também integra as comemorações dos 30 anos do curso de Engenharia de Computação da UFRN. A programação contempla palestras, sessões plenárias e apresentações de artigos, abordando temas como inteligência artificial, desenvolvimento de software, automação e aplicações nos setores produtivos.

Organização e comissões

Organização

Laboratório de Automação (LAUT) – Departamento de Engenharia de Computação e Automação (DCA/UFRN).

Comissão Técnica

- Agostinho de Medeiros Brito Junior (UFRN)
- Anderson Cavalcanti (UFRN)
- André Laurindo Maitelli (UFRN)
- Carlos Eduardo Trabuco Dorea (UFRN)
- Fábio Meneghetti Ugulino de Araujo (UFRN)

Comissão de Apoio

- Ana Sophia Diógenes Soares (UFRN)
- Aurea da Silva Cabral (UFRN)
- Dominic N. Rodrigues (UFRN)
- Fernando Henrique do Couto (UFRN)
- Guilherme C. M. Nascimento (UFRN)
- João Paulo Ferreira (UFRN)
- Laíze Suélia da Silva Oliveira (UFRN)
- Marina Lucena Araujo (UFRN)

Programação

12 de novembro de 2026 – Auditório da Reitoria

08h00 Credenciamento.

09h00 Abertura – Prof. Anderson Cavalcanti (Chefe do DCA/UFRN) e Profa. Carla Maitelli (Diretora do Centro de Tecnologia/UFRN).

09h30 Palestra: *30 anos do curso de Engenharia de Computação* – Prof. Samuel Xavier (DCA/UFRN).

10h20 Coffee-break.

10h50 Palestra: *A disrupção no desenvolvimento de software: onde estamos e o que nos espera* – Eng. Gleydson Lima (CEO, eSIG Group).

11h40 Palestra: *O curso de Engenharia de Computação na UFRN – Evolução* – Prof. Adelardo Medeiros (DCA/UFRN).

12h30 Almoço.

12 de novembro de 2026 – Auditório do CTEC

14h00 Plenária: *Da automação à autonomia: como a inteligência artificial está redefinindo a engenharia* – Prof. Ivanovitch Silva (DCA/UFRN).

15h00 Coffee-break.

15h30 Apresentação de artigos (15 minutos de apresentação e 5 minutos para perguntas) – 6 trabalhos.

13 de novembro de 2026 – Auditório do CTEC

09h00 Plenária: *Automação aplicada ao setor de óleo e gás* – Mohamad Sadeque Abou Ali (Idus Inteligência Industrial) e Prof. André Maitelli (DCA/UFRN).

10h00 Coffee-break.

10h30 Apresentação de artigos (15 minutos de apresentação e 5 minutos para perguntas) – 6 trabalhos.

12h30 Almoço.

14h00 Apresentação de artigos (15 minutos de apresentação e 5 minutos para perguntas) – 6 trabalhos.

16h00 Coffee-break.

16h30 Encerramento.

Sumário

Apresentação **ii**

Organização e comissões **iii**

Programação **iv**

Trabalhos publicados **1**

Estudos preliminares para o desenvolvimento de uma arquitetura modular baseada em IA para interação com sistemas de gestão de malhas de controle 2

Sistema de inspeção visual automatizada para controle de qualidade em peças fabricadas 9

Análise teórica e experimental da resposta de um gerador síncrono à inserção de um degrau de carga 14

Análise da influência de variações de atraso em controladores GPC aplicados ao modelo de slate para pressão arterial 21

Gêmeos digitais na indústria 4.0: aplicação em tanques acoplados 26

Uma revisão de metodologias inteligentes de classificação de curtos circuitos em sistemas elétricos de potência 34

Interface baseada em agentes de IA para exploração e controle de qualidade de dados sísmicos 42

Identificação de padrões e anomalias em reclamações de telecomunicações no Nordeste brasileiro usando aprendizado não supervisionado 49

Geração automática de equações de estado a partir de Bond Graphs: protótipo e validação numérica 57

Biblioteca para cálculo de conjuntos invariantes com Julia 64

Desenvolvimento de uma ferramenta computacional multicritério para avaliação de desempenho de controladores PID 72

Framework paralelo de ajuste de campos de força atomísticos e coarse-grain de grafeno através de otimização por enxame de partículas PSO 77

Plataforma Linux embarcada para integração de simulador de bombeio mecânico com controlador WEG ALC11 85

Ressintonia automática de controladores PID baseada em estimativa de robustez . 91

Indústria sem telas: perspectivas para o uso de dispositivos vestíveis assistidos por modelos de linguagem de pequeno porte 99

Arquitetura edge-cloud para monitoramento inteligente de veículos conectados com auditabilidade via blockchain 107

Desenvolvimento de uma ferramenta computacional em aplicação WEB para simular o comportamento dinâmico do bombeio por cavidades progressivas (BCP) e estudar o fenômeno de reversão do fluxo 115

Índice de autores **122**

Expediente **123**

Trabalhos publicados

Estudos Preliminares para o Desenvolvimento de uma Arquitetura Modular Baseada em IA para Interação com Sistemas de Gestão de Malhas de Controle

***Ana Sophia Diógenes Soares*[1]**
***Fernando Eduardo Dias*[1]**
***Fábio Meneghetti Ugulino de Araujo*[1]**

[1] *Universidade Federal do Rio Grande do Norte*

Palavras-chave: Chatbot; Inteligência Artificial; Copiloto Operacional; Avaliação e Sintonia de Malhas de Controle; Automação de Processos Industriais.

Resumo

Este trabalho apresenta o desenvolvimento de uma arquitetura modular baseada em Inteligência Artificial para apoiar a interação de usuários com sistemas de gestão de malhas de controle. A proposta consiste na utilização de um *chatbot* como camada de interação entre o operador e módulos especializados responsáveis por funções como avaliação e ressintonia de malhas. A arquitetura foi concebida de forma a separar as responsabilidades entre a interface conversacional, os componentes de comunicação e os módulos responsáveis pelo processamento técnico, evitando que o *chatbot* execute diretamente algoritmos de controle ou tenha acesso à planta industrial. A integração entre os componentes é realizada por meio de APIs, seguindo uma arquitetura cliente-servidor, na qual agentes específicos são responsáveis pelo envio dos parâmetros necessários aos serviços e pelo recebimento dos resultados produzidos. Para a implementação da comunicação entre os módulos é utilizado o *FastAPI*, enquanto o *Streamlit* é empregado na construção da interface de interação com o usuário. A abordagem proposta busca estabelecer uma estrutura flexível e modular, capaz de incorporar, em etapas futuras, recursos de processamento de linguagem natural, aprendizado de máquina, contextualização e armazenamento de informações.

1 Introdução

1.1 Contextualização

As malhas de controle desempenham um papel fundamental na automação de processos industriais, sendo responsáveis por manter variáveis de processo próximas aos valores desejados. Entre as estratégias de controle amplamente utilizadas na indústria encontra-se o controlador proporcional-integral-derivativo (PID), cuja combinação das ações proporcional, integral e derivativa permite atender a diferentes requisitos de regulação e rastreamento. O desempenho de uma malha está relacionado às características do processo e aos parâmetros utilizados pelo controlador, de modo que alterações nas condições de operação podem tornar necessária a avaliação de seu comportamento e o reajuste desses parâmetros.

A integração de métodos de Inteligência Artificial (IA), especialmente às relacionadas a incorporação de técnicas de Processamento de Linguagem Natural, do inglês *Natural Language Processing* (NLP), possibilita o desenvolvimento de sistemas conversacionais projetados para interagir com usuários por meio de linguagem natural. Entre essas aplicações, destacam-se os *chatbots*, que podem atuar como interfaces para consulta e apresentação de informações em domínios específicos. Diante disso, uma inter-

face conversacional pode facilitar o acesso a informações técnicas produzidas por diferentes módulos computacionais, possibilitando que usuários consultem e interpretem esses resultados por meio de uma interação mais próxima da linguagem natural, funcionando como um sistema de copiloto para o auxilio de operadores de plantas industriais.

1.2 Lacuna

A aplicação de funções como sintonia e avaliação podem envolver diferentes ferramentas, parâmetros de configuração e formas de apresentação dos resultados. Em um ambiente no qual os procedimentos técnicos estão distribuídos entre aplicações especializadas, o operador precisa interpretar as informações disponibilizadas por cada módulo e selecionar a operação correspondente à tarefa desejada. Esse cenário instiga a busca por uma camada de interação capaz de apresentar uma forma mais unificada de acesso às funcionalidades sem transferir ao sistema conversacional a responsabilidade pela execução dos algoritmos de controle.

Sistemas conversacionais constituem uma alternativa para essa camada de interação. A literatura caracteriza esses sistemas como aplicações capazes de processar linguagem natural para estabelecer uma interação orientada a objetivos, utilizando desde estratégias predefinidas até técnicas de aprendizado de máquina.

1.3 Objetivo

Sob o prisma da pesquisa de gestão inteligente de malhas de controle, propõe-se o desenvolvimento de uma ferramenta baseada em IA, na forma de um agente conversacional, destinada a apoior a interação do operador com informações provenientes de módulos especializados de avaliação e ressintonia de malhas. A arquitetura proposta estabelece uma separação entre a camada de interação com o usuário e os módulos responsáveis pelo processamento técnico das malhas. Dessa forma, é possível garantir que o *chatbot* não execute os algortimos de avaliação ou ressintonia, sequer possua acesso direto à planta industrial ou aos mecanismos responsáveis pelo controle do processo.

Sendo assim, este trabalho tem como objetivo desenvolver uma ferramenta baseada em IA, na forma de um *chatbot*, para apoiar a interação de usuários com informações provenientes de módulos especializados de avaliação e ressintonia de malhas de controle. Para isso, é proposta uma arquitetura de integração baseada em uma interface de comunicação e em uma API, permitindo que o *chatbot* atue como uma camada de interação e apoio à decisão, mantendo separadas as responsabilidades da inteligência artificial, dos módulos especializados e da planta industrial.

2 Fundamentação Téorica

O desenvolvimento de sistemas conversacionais envolve diferentes conceitos e técnicas de IA, que permitem estabelecer a interação entre usuários e aplicações computacionais. Nesse contexto, são abordados os principais fundamentos relacionados ao funcionamento e à organização desses sistemas.

2.1 Chatbots

Chatbots são programas que simulam uma conversa humana, comunicando-se por texto ou voz. Trata-se de um modelo de Interação Humano-Máquina que utiliza NLP e análise de sentimento para se comunicar em linguagem humana com pessoas ou outros sistemas automatizados.

O conceito existe desde o surgimento do Teste de Turing, em 1950, que questionou se uma máquina seria capaz de se comportar como um ser humano. No entanto, o primeiro *chatbot* amplamente reconhecido foi o ELIZA, desenvolvido por Joseph Weizenbaum em 1966. O sistema simulava um psicoterapeuta por meio de regras de correspondência de padrões (*pattern matching*), reformulando as frases do usuário em perguntas. Apesar de sua simplicidade e da incapacidade de compreender o contexto da conversa, o ELIZA demonstrou o potencial da interação em linguagem natural e serviu de base para sistemas posteriores, como PARRY, ALICE e *SmarterChild*, que introduziram melhorias relacionadas ao contexto, à personalidade e ao acesso a bases de conhecimento (Adamopoulou; Moussiades, 2020).

A partir da década de 2010, a evolução das técnicas de NLP e Aprendizado de Máquina, do inglês Machine Learning (ML) possibilitou o surgimento de assistentes virtuais mais sofisticados, como *Siri*, *Google Assistant*, *Cortana*, *Alexa* e *IBM Watson*. Diferentemente dos sistemas baseados exclusivamente em regras, esses assistentes utilizam modelos estatísticos e redes neurais para interpretar intenções, extrair informações relevantes e produzir respostas mais naturais, tornando possível sua aplicação em diferentes áreas, como educação, saúde, comércio eletrônico e automação industrial.

Os *chatbots* podem ser classificados por meio de diferentes critérios. Quanto ao domínio de conhecimento, podem ser de domínio específico, quando

atuam em uma única área, ou de domínio aberto, capazes de responder a perguntas sobre diversos assuntos. Em relação ao objetivo, destacam-se os *chatbots* informativos, voltados ao fornecimento de informações; os conversacionais, cujo foco é manter um diálogo natural; e os orientados a tarefas, destinados à execução de ações como agendamentos, consultas e automação de processos.

Outra classificação importante considera a forma de geração das respostas. Os sistemas baseados em regras utilizam padrões previamente definidos para selecionar respostas, oferecendo maior previsibilidade e facilidade de implementação. Já os modelos baseados em recuperação (*retrieval-based*) selecionam a resposta mais adequada a partir de um conjunto de exemplos, enquanto os modelos generativos empregam técnicas de aprendizado profundo para produzir novas respostas de forma dinâmica. Atualmente, soluções híbridas que combinam essas abordagens têm sido amplamente adotadas, buscando conciliar precisão, flexibilidade e naturalidade nas interações.

2.2 Processamento de Linguagem Natural

Por definição, o NLP corresponde ao conjunto de técnicas utilizadas para permitir que sistemas computacionais processem, interpretem e produzam informações expressas em linguagem natural. Em sistemas conversacionais, o NLP constitui um dos principais componentes responsáveis pela interpretação das mensagens fornecidas pelos usuários e pela construção de respostas adequadas ao contexto da interação.

A utilização de NLP permite que a comunicação com o sistema seja feita por meio de linguagem natural, reduzindo a necessidade de que o usuário conheça comandos específicos ou a estrutura interna da aplicação. Em um sistema de domínio especializado, essas técnicas podem ser utilizados para identificar a intenção do usuário, extrair informações relevantes da solicitação e determinar quais recursos do sistemas devem ser acionados.

2.3 Inteligência Artificial e Aprendizado de Máquina

A Inteligência Artificial pode ser definida como o estudo de sistemas que agem de uma forma que, para qualquer observador, pareça inteligente, ou ainda, o uso de métodos baseados no comportamento inteligente humano e animais para a resolução de problemas complexos. Conceitualmente, a IA está na linha de frente na transformação da análise de dados e da melhoria da tomada de decisões por meio do raciocínio, da aprendizagem e da resolução de problemas. Além disso, dentre as técnicas que se destacam dentro do campo da IA, destacam-se os métodos de ML, nos quais modelos computacionais podem identificar padrões a partir de dados e utilizá-los na realização de determinadas tarefas.

No contexto dos sistemas conversacionais, técnicas de IA e ML podem ser utilizadas em diferentes etapas do processamento, incluindo a interpretação da linguagem, classificação das solicitações, recuperação de informações e geração de respostas. A escolha das técnicas utilizadas depende das características do domínio, dos dados e dos requisitos do sistema.

2.4 Arquiteturas de Chatbots

A arquitetura de um *chatbot* corresponde à organização dos componentes responsáveis pela interação com o operador, processamento das solicitações e geração das respostas. A complexidade dessa arquitetura pode variar de acordo com o domínio da aplicação e com as técnicas utilizadas para processar as mensagens.

Em sistemas mais básicos, a interação pode ser baseada em regras predefinidas, nas quais determinadas entradas são associadas a respostas específicas. Abordagens baseadas em recuperação podem utilizar conjuntos de informações previamente disponíveis para selecionar respostas relacionadas à solicitação do usuário. Já abordagens generativas utilizam modelos capazes de produzir respostas a partir da entrada recebida e do contexto disponível.

Além do mecanismo de geração das respostas, arquiteturas de *chatbots* podem ser compostas por diferentes módulos, como interface de usuário, processamento da linguagem, gerenciamento de contexto, memória e integração com serviços externos.

2.5 Domínio Específico

Chatbots de domínio específico são desenvolvidos para atuar em contextos nos quais as interações estão relacionadas a um conjunto delimitado de conhecimentos, tarefas ou informações. Diferentemente de sistemas conversacionais destinados a atender a uma ampla variedade de assuntos, esses sistemas podem ser projetados considerando as características e necessidades particulares do domínio de aplicação.

A especialização do domínio permite definir quais informações são relevantes para as interações e quais operações podem ser realizadas pelo sistema. Essa característica é particularmente relevante em aplicações técnicas, nas quais as respostas precisam estar

relacionadas a informações específicas do contexto de utilização.

No presente trabalho, o domínio de aplicação está relacionado à gestão inteligente de malhas de controle. Dessa forma, o *chatbot* é concebido para interagir com informações produzidas por módulos especializados desse domínio, mantendo delimitadas suas responsabilidades em relação ao processamento técnico das malhas.

3 Aplicação

A abordagem adotada busca estabelecer a integração entre a interface conversacional e os módulos especializados de avaliação e sintonia de malhas, mantendo a independência entre os diferentes serviços.

3.1 Visão Geral

A ferramenta proposta consiste em um sistema conversacional baseado em Inteligência Artificial desenvolvido para facilitar a interação de usuários com informações relacionadas às malhas de controle de instalações industriais. A solução foi concebida como uma camada intermediária entre o usuário e módulos especializados responsáveis pelo processamento das informações técnicas das malhas.

A arquitetura adotada busca manter separadas as responsabilidades de cada componente. Enquanto os módulos especializados permanecem responsáveis pelos procedimentos de avaliação e sintonia das malhas, o *chatbot* atua na interpretação das solicitações do usuário na comunicação com os serviços disponíveis e na apresentação dos resultados obtidos.

Desse modo, a ferramenta não possui como finalidade substituir os algoritmos especializados ou realizar diretamente o controle do processo. Sua função consiste em fornecer uma interface de interação em linguagem natural para facilitar o acesso às informações produzidas pelos demais componentes, funcionando como um auxiliar, um copiloto, para o operador do sistema.

3.2 Estratégia de modulação

A arquitetura da ferramenta vista na Figura 1 foi definida de forma modular, estabelecendo uma separação entre a interface de interação, o *chatbot*, a camada de comunicação e os módulos especializados da pesquisa. Essa organização permite que cada componente possua uma responsabilidade específica e que a comunicação entre eles seja realizada por interfaces bem definidas.

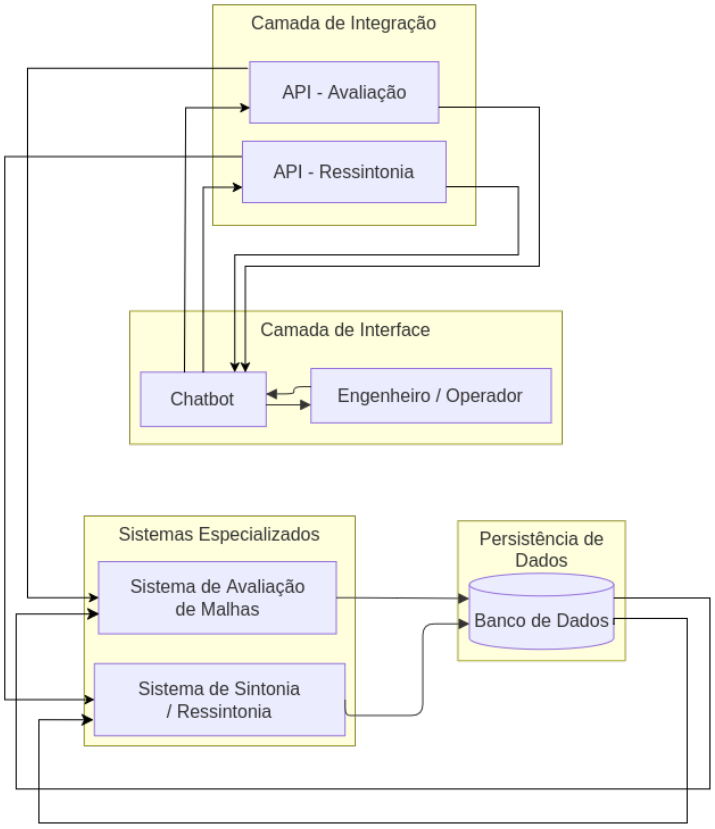


Figura 1: *Estrategia de integração dos sistemas.*

3.3 Integração entre módulos

A comunicação entre o *chatbot* e os módulos responsáveis pela avaliação e ressintonia das malhas será implementada por meio do *FastAPI*, um *framework* moderno da linguagem Python destinado ao desenvolvimento de APIs de alto desempenho. Essa abordagem possibilita a troca eficiente de informações entre os componentes do sistema, mantendo a independência entre a interface conversacional e os módulos de processamento, como demonstrado na Figura 2.

A arquitetura utilizada parte da ideia cliente-servidor, em que o assistente virtual servirá como cliente enquanto o servidor será a aplicação de avaliação/sintonia.

3.4 Arquitetura de Ferramenta

Para a implementação, foi estabelecida uma separação entre a interface de interação, o *chatbot*, a camada de comunicação e os módulos especializados da pesquisa. Essa organização permite que cada componente possua uma responsabilidade específica e que a comunicação entre eles seja realizada por interfaces bem definidas.

A Figura 3 apresenta a organização dos principais componentes da solução. O operador interage com



Figura 2: *Página do Modelo de Sintonia de Malhas*

a aplicação por meio da interface conversacional, enquanto o *chatbot* realiza o processamento da solicitação e, quando necessário, consulta os módulos especializados por meio da API.

3.5 Interface

A interface responsável pela interação entre o usuário e as aplicações será desenvolvida utilizando o *Streamlit*, um *framework* da linguagem Python destinado à construção de aplicações web interativas. Essa abordagem permitirá integrar, em uma única interface, o *chatbot* e os módulos responsáveis pela avaliação e ressintonia de malhas.

3.6 Agente

Neste trabalho, denomina-se agente o componente de software responsável por intermediar a comunicação entre o *chatbot* e os módulos especializados de avaliação e ressintonia. Sua função consiste em realizar as chamadas às respectivas APIs, fornecendo os parâmetros necessários para a execução de cada serviço e recebendo os resultados produzidos pelos módulos externos. Dessa forma, o agente atua como uma camada de comunicação, sem executar diretamente os algoritmos responsáveis pelo processamento técnico das malhas.

Para preservar a modularidade da arquitetura proposta, cada funcionalidade possui um agente específico, responsável pela comunicação com o respectivo serviço. Em sua implementação, são definidos o endereço do *endpoint* da API, os parâmetros necessários à requisição e a estrutura dos dados enviados. As informações são organizadas no formato JSON e encaminhadas ao serviço correspondente. Após o processamento, a resposta retornada pela API é recebida pelo agente e disponibilizada aos demais componentes da aplicação para posterior interpretação e apresentação ao usuário.

4 Direcionamentos futuros

Para dar continuidade ao desenvolvimento do *chatbot*, será necessário integrar módulos voltados ao processamento de linguagem natural, bem como mecanismos de contextualização específicos para o domínio no qual a ferramenta está inserida. Essa etapa é fundamental para adaptar o sistema à sua finalidade de aplicação, permitindo que ele compreenda não apenas a linguagem utilizada pelos usuários, mas também as particularidades e o contexto do ambiente em que será empregado. Nesse sentido, torna-se necessário incorporar conhecimentos específicos da área de controle, possibilitando ao *chatbot* interpretar adequadamente os jargões e a terminologia técnica característica desse domínio.

De maneira análoga à capacidade humana de interpretar informações e formular respostas, a ferra-

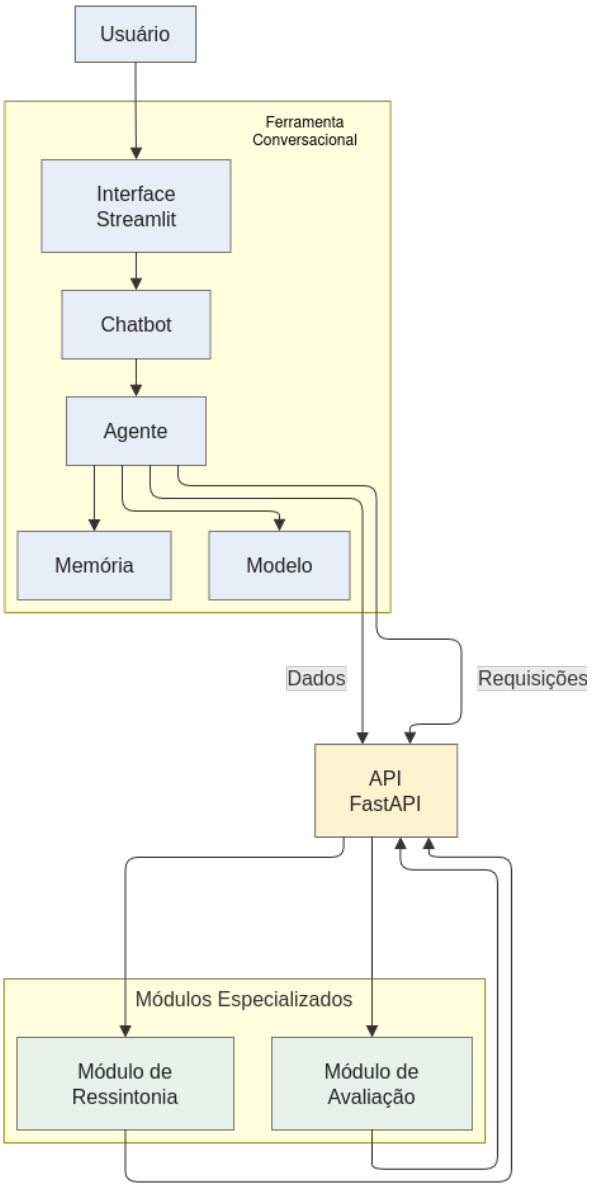


Figura 3: *Arquitetura proposta para a ferramenta conversacional.*

menta deverá dispor de módulos capazes de processar os dados recebidos, relacioná-los ao contexto e gerar respostas coerentes e compatíveis com as necessidades do usuário. Para viabilizar essas funcionalidades, *frameworks* de aprendizado de máquina (ML) deverão ser integrados à arquitetura do sistema, contribuindo para o processamento e a interpretação das informações.

Além disso, considerando que a utilização da ferramenta resultará na geração e no armazenamento de um histórico de interações, torna-se necessária a implementação de um banco de dados. Esse recurso permitirá tanto o armazenamento e posterior processamento das informações quanto a recuperação de dados relevantes em situações nas quais seja necessário consultar interações ou conhecimentos previamente registrados. Dessa forma, a integração entre os módulos de processamento de linguagem natural, aprendizado de máquina e armazenamento de dados constitui uma etapa importante para a evolução das funcionalidades e da capacidade de contextualização do *chatbot*.

5 Conclusões

Este trabalho apresentou a proposta de uma arquitetura modular baseada em Inteligência Artificial para facilitar a interação entre usuários e sistemas especializados de gestão de malhas de controle. A solução foi estruturada de maneira a estabelecer uma separação clara entre a camada conversacional e os módulos responsáveis pelo processamento técnico, permitindo que o *chatbot* desempenhe o papel de interface de comunicação sem assumir diretamente funções de avaliação e sintonia, que são desempenhadas por módulos por módulos específicos, desenvolvidos por outros membros da equipe do projeto em questão, ou de controle da planta que continuará sob responsabilidade do sistema de automação já presente na mesma.

A adoção de uma arquitetura baseada em APIs possibilita que os diferentes componentes da solução mantenham suas responsabilidades independentes, ao mesmo tempo em que estabelece uma interface padronizada para a troca de informações. Nesse contexto, os agentes de comunicação são responsáveis por encaminhar os parâmetros necessários aos módulos especializados e receber os resultados produzidos, preservando a modularidade proposta para o sistema. Essa organização também favorece a inclusão ou substituição de funcionalidades sem exigir alterações significativas nos demais componentes da aplicação.

A utilização do *FastAPI* para a comunicação en-

tre os módulos e do *Streamlit* para a construção da interface fornece uma base para a implementação e evolução da ferramenta. A estrutura desenvolvida permite concentrar, em uma única interface, o acesso às funcionalidades disponibilizadas pelos diferentes módulos, ao mesmo tempo em que mantém separados o processamento técnico das malhas e a interação com o usuário.

Como continuidade deste trabalho, deverão ser incorporados mecanismos de Processamento de Linguagem Natural e Aprendizado de Máquina capazes de aprimorar a interpretação das solicitações realizadas em linguagem natural e considerar as particularidades do domínio de controle. Também está prevista a implementação de mecanismos de contextualização e armazenamento das interações, permitindo que informações previamente registradas possam contribuir para o processamento de novas solicitações. Dessa forma, a arquitetura apresentada constitui uma base para o desenvolvimento progressivo de uma ferramenta conversacional voltada ao apoio à interação com sistemas de gestão inteligente de malhas de controle.

Referências

ADAMOPOULOU, Eleni; MOUSSIADES, Lefteris. Chatbots: history, technology, and applications. Machine Learning with Applications, v. 2, artigo 100006, 2020. DOI: https://doi.org/10.1016/j.mlwa.2020.100006.

Sistema de Inspeção Visual Automatizada para Controle de Qualidade em Peças Fabricadas

***Israel Lucas*[1]**

[1]*Escola de Ciências e Tecnologia, UFRN,*
israel.lucas.126@ufrn.edu.br

Resumo

Este trabalho apresenta o desenvolvimento de um sistema de inspeção visual automatizada para detecção de defeitos em peças metálicas utilizando técnicas de visão computacional e aprendizagem profunda. O experimento foi desenvolvido no Google Colab com a biblioteca Ultralytics e o modelo YOLOv8n. O conjunto de dados utilizado, denominado Detection_Defects_in_Metal, possuía anotações originalmente no formato Pascal VOC (XML), que foram convertidas para o formato YOLO (TXT), com normalização das coordenadas das caixas delimitadoras. O conjunto foi dividido em 80% para treinamento e 20% para validação, resultando em 1.824 imagens de treinamento e 456 imagens de validação. O modelo foi treinado durante 10 épocas, com imagens de 640 × 640 pixels e batch size 16. Na validação final, foram obtidos precisão de 0,674, recall de 0,470, mAP@50 de 0,497 e mAP@50–95 de 0,247. Os resultados demonstram a viabilidade da aplicação do YOLOv8 na identificação automatizada de defeitos, embora também indiquem limitações relacionadas à distribuição desigual das classes e ao número reduzido de exemplos de algumas categorias.

Abstract

This work presents the development of an automated visual inspection system for detecting defects in metallic parts using computer vision and deep learning techniques. The experiment was developed in Google Colab using the Ultralytics library and the YOLOv8n model. The dataset, named Detection_Defects_in_Metal, originally contained Pascal VOC (XML) annotations, which were converted to the YOLO (TXT) format, with normalized bounding-box coordinates. The dataset was split into 80% for training and 20% for validation, resulting in 1,824 training images and 456 validation images. The model was trained for 10 epochs using 640 × 640 pixel images and a batch size of 16. In the final validation, the model achieved a precision of 0.674, recall of 0.470, mAP@50 of 0.497, and mAP@50–95 of 0.247. The results demonstrate the feasibility of applying YOLOv8 to automated defect identification, while also indicating limitations associated with class imbalance and the small number of samples in some categories.

Keywords: computer vision; visual inspection; object detection; YOLOv8; metallic parts.

1 Introdução

A Visão Computacional é uma área da ciência da computação que permite aos sistemas computacionais interpretar e extrair informações de imagens digitais, possibilitando a identificação de objetos, padrões e

características visuais. Segundo Gonzalez e Woods (2010), o processamento digital de imagens utiliza técnicas como filtragem, segmentação e reconhecimento de padrões para obter informações relevantes. No contexto industrial, essas técnicas são aplicadas à inspeção visual automatizada, contribuindo para o controle de qualidade e a identificação de irregularidades em produtos fabricados.

Em peças metálicas, a detecção automatizada permite identificar diferentes tipos de imperfeições superficiais por meio da análise de imagens. Conforme Xie (2008), técnicas de análise de textura apresentam aplicações importantes na detecção de defeitos em superfícies industriais. Dessa forma, a utilização de sistemas computacionais pode auxiliar na padronização das inspeções e reduzir a dependência de avaliações exclusivamente manuais.

Com os avanços da Inteligência Artificial, as Redes Neurais Convolucionais (*CNNs*) passaram a desempenhar um papel importante na análise de imagens, permitindo aprender características como bordas, formas e texturas diretamente dos dados de treinamento (LECUN; BENGIO; HINTON, 2015). Entre os algoritmos de detecção de objetos destaca-se o YOLO (*You Only Look Once*), apresentado por Redmon et al. (2016), que realiza simultaneamente a classificação e a localização dos objetos em uma imagem. Neste trabalho, foi utilizado o modelo YOLOv8n, da biblioteca Ultralytics, para identificar e localizar defeitos em peças metálicas por meio de caixas delimitadoras (*bounding boxes*).

O treinamento desses modelos depende de imagens devidamente anotadas, contendo a identificação das classes e as coordenadas dos objetos. Para isso, as anotações podem ser armazenadas em diferentes formatos, como Pascal VOC (XML) e YOLO (TXT). O desempenho do modelo pode ser avaliado por métricas como precisão (*Precision*), revocação (*Recall*) e precisão média (mAP), que permitem analisar a capacidade de identificação e localização dos defeitos.

No cenário industrial, a inspeção de produtos é uma etapa importante para garantir a qualidade e a conformidade das peças fabricadas. Em processos com grande quantidade de produtos, a inspeção exclusivamente manual pode apresentar limitações relacionadas à repetibilidade, ao tempo de execução e à possibilidade de variações na avaliação realizada pelo operador. Nesse contexto, sistemas de inspeção visual automatizada podem auxiliar na identificação de características e defeitos presentes nas peças.

A visão computacional possibilita transformar imagens em informações úteis para sistemas automatizados. Entre as aplicações dessa área encontra-se a detecção de objetos, na qual o modelo deve localizar objetos ou regiões de interesse e atribuir uma classe a cada detecção. Para a inspeção de peças metálicas, essa abordagem pode ser utilizada para identificar diferentes tipos de imperfeições a partir de imagens.

2 Metodologia

A metodologia adotada neste trabalho consistiu no desenvolvimento de um sistema de inspeção visual automatizada para a detecção de defeitos em peças metálicas, utilizando técnicas de visão computacional e aprendizagem profunda. Para isso, foi utilizado um *dataset* encontrado no site do *Kaggle*. Nele é possível encontrar diversos compartilhamentos de *datasets*, bem como a possibilidade de criar um do zero. O experimento foi realizado no ambiente Google Colab, com a biblioteca Ultralytics e o modelo de detecção de objetos YOLOv8n.O procedimento foi dividido em três etapas principais: preparação do conjunto de dados, treinamento do modelo e avaliação dos resultados obtidos.

2.1 Conjunto de dados e preparação

O conjunto de dados utilizado no experimento é denominado Detection_Defects_in_Metal. As imagens estavam armazenadas em uma pasta de imagens e as respectivas anotações em arquivos XML. As anotações seguem a estrutura Pascal VOC, contendo, entre outras informações, as dimensões da imagem, o nome da classe e as coordenadas da caixa delimitadora (xmin, xmax, ymin e ymax).

Para permitir o treinamento com YOLO, os arquivos XML foram convertidos para arquivos TXT. As coordenadas das caixas foram transformadas para o formato YOLO, representado por classe, centro x, centro y, largura e altura, com os valores normalizados pelas dimensões da imagem. A divisão dos dados foi realizada com test_size=0,2 e random_state=42, resultando em 80% dos arquivos para treinamento e 20% para validação.

O processamento realizado no notebook gerou 1.824 imagens no conjunto de treinamento e 456 imagens no conjunto de validação. O treinamento utilizou 12 classes de defeitos presentes nas anotações.

2.2 Modelo e treinamento

Foi utilizado o modelo YOLOv8n pré-treinado, carregado por meio da biblioteca Ultralytics. O treinamento foi configurado para 10 épocas, tamanho de imagem de 640 pixels e batch size igual a 16. O treinamento foi executado no ambiente Google Colab e,

conforme o registro do notebook, utilizou CPU. O processo completo de treinamento teve duração aproximada de 4,769 horas.

Durante o treinamento, o modelo apresentou evolução das perdas de caixa (box loss), classificação (classification loss) e distribuição focal (DFL loss). Na décima época, os valores registrados foram 1,694 para box loss, 1,861 para classification loss e 1,411 para DFL loss. Na validação final, realizada com os pesos armazenados em best.pt, foram obtidos os resultados apresentados na **Tabela 1.**

2.3 Métricas de avaliação

A avaliação foi realizada utilizando precisão (Precision), revocação (Recall), mAP@50 e mAP@50–95. A precisão representa a proporção de detecções classificadas corretamente entre as detecções realizadas, enquanto o recall indica a capacidade do modelo de recuperar as ocorrências presentes nas imagens. O mAP é uma medida agregada de desempenho da detecção, sendo calculado para diferentes critérios de interseção sobre união (IoU).

Conj.	Imag.	Instân.	Precision	Recall
Validação	456	683	0,674	0,470
Validação	456	683	mAP@50 = 0,497	mAP@50–95 = 0,247

Tabela 1: Resultados globais obtidos na validação final do YOLOv8n.

3 Resultados e Discussões

A validação final apresentou precisão de 0,674 e recall de 0,470. O valor de mAP@50 foi 0,497, enquanto o mAP@50–95 foi 0,247. Esses valores mostram que o modelo conseguiu realizar detecções de diferentes classes de defeitos, mas ainda apresenta margem significativa para melhoria, principalmente na localização mais precisa das caixas e na recuperação de todas as ocorrências.

O desempenho variou consideravelmente entre as classes. A classe 3_yueyawan apresentou mAP@50 de 0,952 e recall igual a 1,000, enquanto a classe 1_chongkong apresentou mAP@50 de 0,917 e recall de 0,939. Por outro lado, classes com poucos exemplos apresentaram desempenho inferior. A classe 10_yaozhe possuía apenas 4 imagens e 6 instâncias na validação e apresentou recall igual a zero e mAP@50 de 0,0109. A classe 9_zhehen, com 11 imagens e 19 instâncias, também apresentou recall igual a zero. Esses resultados indicam a influência da quantidade e da distribuição das amostras sobre o treinamento.

A classe 7_yiwu também apresentou desempenho reduzido, com precisão de 0,423, *recall* de 0,129, mAP@50 de 0,268 e mAP@50–95 de 0,0896. Em contraste, classes com maior quantidade de exemplos, como 6_siban, 2_hanfeng e 1_chongkong, apresentaram resultados mais consistentes. Portanto, os resultados sugerem que o aumento e o balanceamento das amostras das classes minoritárias podem contribuir para melhorar a capacidade de generalização do modelo.

Classe	Imag.	Inst.	P	R	mAP50	mAP50–95
6_siban	148	175	0.709	0.432	0.517	0.241
4_shuiban	52	63	0.688	0.571	0.603	0.352
5_youban	34	68	0.547	0.544	0.520	0.261
7_yiwu	49	85	0.423	0.129	0.268	0.090
1_chongkong	66	66	0.888	0.939	0.917	0.462
2_hanfeng	97	98	0.703	0.867	0.783	0.347
10_yaozhe	4	6	1.000	0.000	0.011	0.004
3_yueyawan	50	50	0.880	1.000	0.952	0.619
9_zhehen	11	19	1.000	0.000	0.362	0.120
10_yaozhed	34	34	0.282	0.529	0.362	0.125
8_yahen	9	19	0.296	0.158	0.175	0.094

Tabela 2: Métricas por classe obtidas na validação final.

O recall global de 0,470 indica que mais da metade dos defeitos não foi detectada. Em inspeção industrial, isso equivale a falsos negativos, ou seja, peças defeituosas que seriam aprovadas, o que é mais crítico do que falsos positivos. As principais causas prováveis são o treinamento curto (10 épocas, ainda sem convergência, pois as perdas continuavam altas), o desbalanceamento de classes (*10_yaozhe* e *9_zhehen* com recall zero), o uso do modelo de menor capacidade (YOLOv8n) e defeitos pequenos ou de baixo contraste. Seriam necessárias mais épocas, aumento de dados e ajuste do limiar de confiança para priorizar o recall.

Como o conjunto de validação foi usado para avaliar e para selecionar o melhor modelo (best.pt), as métricas podem estar otimistas. Não houve um conjunto de teste independente (por exemplo, 70/15/15), o que limita a avaliação da generalização. A distribuição das

classes pode ser observada na **Imagem 1** e **Imagem 2**, com as Matrizes de Confusão.

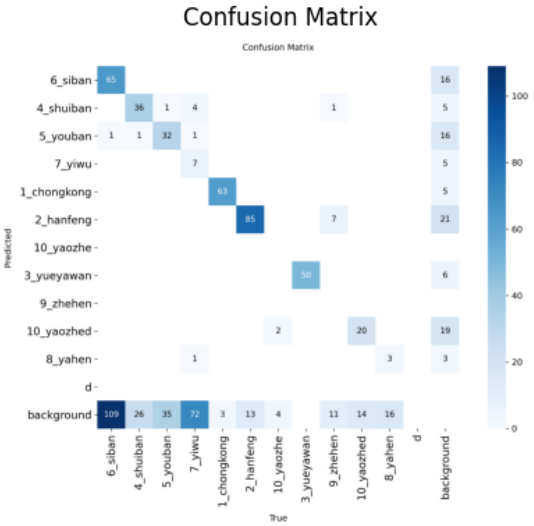


Imagem 1 – *Matriz de Confusão*

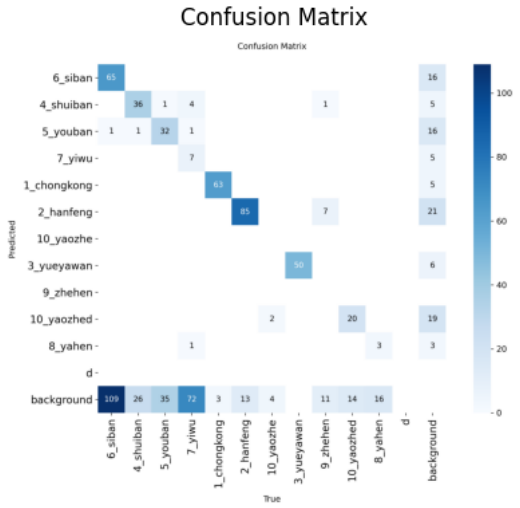


Imagem 2 – *Matriz de Confusão Normalizada*

3.1 Teste de Inferência

Após o treinamento, o arquivo best.pt foi utilizado para realizar inferência em uma imagem externa obtida por URL. O resultado registrado no notebook indicou uma detecção da classe 2_hanfeng. A inferência foi executada em uma imagem redimensionada para 288 × 640 durante o processamento, com tempo aproximado de 196,0 ms para a inferência, além de 7,1 ms de pré-processamento e 30,4 ms de pós-processamento.

O teste também confirmou a existência de uma bounding box, com coordenadas aproximadas [59,23; 158,50; 392,12; 182,63] e confiança de 0,5949. O ID de classe retornado foi 5, correspondente à classe 2_hanfeng na configuração gerada pelo dataset. Não foram identificadas máscaras de segmentação, pois o experimento foi configurado para detecção por caixas delimitadoras.

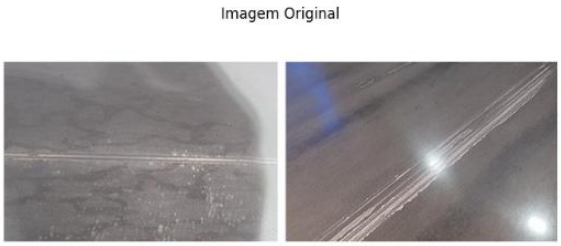


Imagem 3: *Imagem Original.*



Imagem 4: *Imagem com Detecção Yolov8.*

3.2 Conjunto de dados (12 vs. 11 classes)

O conjunto de dados utilizado neste trabalho apresenta 12 classes de defeitos em peças metálicas. Entretanto, durante a etapa de validação do modelo YOLO, foram identificadas apenas 11 classes. Essa diferença indica que uma das categorias não apresentou ocorrências no subconjunto de validação ou que pode existir alguma inconsistência na identificação e no mapeamento das classes durante o processamento das anotações.

Essa diferença entre o número de classes disponíveis no conjunto original e aquelas presentes na validação é importante, pois pode influenciar a avaliação do desempenho do modelo. A ausência de exemplos de determinada categoria na validação impossibilita verificar

diretamente sua capacidade de reconhecimento para essa classe específica.

Além disso, a distribuição das imagens e das instâncias entre treinamento e validação deve ser considerada, uma vez que o equilíbrio entre as categorias contribui para uma avaliação mais representativa do modelo. Portanto, a diferença observada entre as 12 classes originais e as 11 classes identificadas na validação deve ser considerada na interpretação dos resultados experimentais.

4 Conclusões

Este trabalho constitui uma prova de conceito da aplicação do YOLOv8n à inspeção de peças metálicas. Os resultados (mAP@50 de 0,497 e recall de 0,470), obtidos apenas em validação, com treinamento curto e classes desbalanceadas, ainda não permitem afirmar que a solução seja robusta o bastante para uso industrial. São necessários mais dados nas classes minoritárias, treinamento mais longo, modelos maiores e avaliação em um conjunto de teste independente.

O experimento demonstrou a possibilidade de utilizar o YOLOv8n para a inspeção visual automatizada de peças metálicas. A metodologia contemplou a conversão das anotações XML para o formato YOLO, a divisão do dataset em treinamento e validação, o treinamento do modelo e a realização de inferência em uma imagem externa. Assim, o experimento indica que o sistema é funcional como prova de conceito, mas ainda requer aprimoramentos antes de ser considerado uma solução robusta para uso industrial.

Como trabalhos futuros, recomenda-se ampliar o número de imagens das classes minoritárias, revisar a qualidade e consistência das anotações, aumentar o número de épocas, avaliar modelos YOLO de maior capacidade e comparar diferentes configurações de treinamento. Também é recomendável realizar uma avaliação com um conjunto de teste independente, permitindo medir de forma mais rigorosa a capacidade de generalização do sistema.

Referências

GONZALEZ, Rafael C.; WOODS, Richard E. Processamento Digital de Imagens. 3. ed. São Paulo: Pearson Prentice Hall, 2010.

MAIA, Helton. Visão Computacional. Disponível em: https://heltonmaia.com/computervision/intro.html. Acesso em: 22 jul. 2025.

ULTRALYTICS. Ultralytics YOLO. Documentação e biblioteca para detecção de objetos. Disponível em: https://docs.ultralytics.com/. Acesso em: 3 out. 2026.

XIE, X. A review of recent advances in surface defect detection using texture analysis techniques. Image and Vision Computing, v. 26, n. 6, p. 850–863, 2008.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. Nature, v. 521, p. 436–444, 2015.

REDMON, Joseph et al. You Only Look Once: Unified, Real-Time Object Detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, p. 779–788.

Análise Teórica e Experimental da Resposta de um Gerador Síncrono à Inserção de um Degrau de Carga

***Levi da Silva Lopes*[1]**
***José Higor Rodrigues Soares*[1]**
***Rodrigo Prado de Medeiros*[1]**
***Thiago de Oliveira Alves Rocha*[1]**
***Ricardo Lúcio de Araújo Ribeiro*[1]**
***José Patrocínio da Silva*[1]**

[1] *Departamento de Engenharia Elétrica, UFRN, <<levi.lopes.026@ufrn.edu.br>>*
[1] *Departamento de Engenharia Elétrica, UFRN, <<jose.higor.soares.018@ufrn.edu.br>>*
[1] *Departamento de Engenharia Elétrica, UFRN, <<rodrigo.prado@ufrn.br>>*
[1] *Departamento de Engenharia Elétrica, UFRN, <<thiago.rocha@ufrn.br>>*
[1] *Departamento de Engenharia Elétrica, UFRN, <<rlucio@dee.ufrn.br>>*
[1] *Departamento de Engenharia Elétrica, UFRN, <<jose.patrocinio@ufrn.br>>*

Resumo

A estimação e o monitoramento de grandezas elétricas fundamentais, como tensão, corrente, potência e frequência, são essenciais para as funções de controle e supervisão de sistemas elétricos. Em regimes de operação isolada, manobras abruptas como o chaveamento de cargas provocam transitórios caracterizados por elevação de corrente e afundamentos de tensão. Este trabalho apresenta uma análise circuital simplificada e uma avaliação experimental da resposta de um gerador síncrono à inserção de carga, considerando as variações de tensão, corrente, potência ativa e frequência estimada no Ponto de Acoplamento Comum (PAC) de uma bancada experimental composta por um gerador síncrono conectado a uma carga passiva, no laboratório PTX da UFRN. A análise considera duas condições operacionais: antes e após o fechamento da chave responsável pela conexão da carga adicional. Os ensaios foram conduzidos na bancada física dotada de um conjunto motor-gerador constituído por um motor de corrente contínua (CC) acoplado a um gerador síncrono, com aquisição digital dos sinais no PAC via plataforma de prototipagem rápida dSPACE (DS1103) e correlação com a velocidade medida via encoder óptico incremental. Os resultados experimentais apresentam tendências qualitativamente e quantitativamente compatíveis com a formulação circuital desenvolvida, acompanhadas da quantificação das variações transitórias de potência e afundamento de tensão.

Abstract

Accurate estimation and monitoring of fundamental electrical quantities such as voltage, current, power, and frequency form the basis for control, dispatch, and stability in modern electrical grids. During islanded operation, sudden load switching may produce current transients and terminal-voltage deviations. This paper presents a simplified circuit-based analysis and an experimental investigation of a synchronous generator response to load insertion, considering voltage, current, active power, and estimated frequency variations at the Point of Common Coupling (PCC) of an experimental testbed located at the PTX laboratory at UFRN. The analysis considers two operating

conditions: before and after the connection of an additional load. Tests were conducted on a physical setup featuring a DC motor mechanically coupled to a salient-pole synchronous generator, with digital signal acquisition at the PCC via a dSPACE (DS1103) rapid prototyping system and electromechanical correlation via incremental encoder. The experimental results exhibit qualitative and quantitative trends consistent with the simplified circuit analysis, accompanied by the quantification of transient power variations and voltage sags.

1 Introdução

As transformações recentes observadas na matriz elétrica mundial têm impulsionado a integração de unidades de geração distribuída, com ênfase no aproveitamento de fontes renováveis de energia operando em estruturas de microrredes (KUNDUR; BALU; LAUBY, 1994; BEVRANI; WATANABE; MITANI, 2014). Sob essa arquitetura de operação, particularmente durante o modo ilhado ou isolado da rede principal, a análise e o controle da dinâmica do sistema dependem da medição e da estimação adequadas de grandezas elétricas, como tensão, corrente e frequência (MILANO, 2010).

Em sistemas de potência reais, variações abruptas da demanda, decorrentes da conexão ou desconexão de cargas conectadas em paralelo no PAC, impõem perturbações eletromagnéticas e transitórios imediatos. Esses distúrbios refletem-se simultaneamente em: (i) sobrecorrentes transitórias e de regime permanente decorrentes da redução da impedância equivalente de carga; (ii) afundamento imediato na amplitude da tensão senoidal terminal, governado pela queda fasorial sobre a impedância interna do gerador; e (iii) desvios dinâmicos de rotação e frequência provocados pelo desbalanço temporário entre a demanda elétrica solicitada e a potência mecânica do acionamento primário (GRAINGER; STEVENSON, 1994).

No âmbito experimental, a reprodução física fidedigna desses regimes operacionais exige uma plataforma de bancada robusta e flexível. O laboratório PTX da Universidade Federal do Rio Grande do Norte (UFRN) dispõe de uma microrrede constituída por um grupo eletromecânico que conjuga um motor de corrente contínua (CC), atuando como máquina primária motriz, mecanicamente acoplado a um gerador síncrono (GS) trifásico de polos salientes (FITZGERALD; KINGSLEY; UMANS, 2003).

Neste experimento, a aquisição digital e o processamento digital de sinais (PDS) das grandezas elétricas e das malhas de sincronização em tempo real (BLAABJERG; TEODORESCU, 2006) são executados por uma plataforma de prototipagem rápida dSPACE (DS1103). Módulos de condicionamento com transdutores de efeito Hall (sensores de tensão LV-25P e de corrente LAH-25-NP) e encoder óptico incremental acoplado ao rotor permitem adquirir as tensões terminais, as correntes de linha e monitorar a velocidade mecânica angular do eixo. Neste trabalho, são analisados os sinais elétricos e a frequência estimada pelo SRF-PLL. O objetivo deste trabalho é analisar a resposta de um gerador síncrono à inserção de carga, confrontando uma representação circuital simplificada com medições reais de tensão, corrente, potência ativa e a frequência estimada via SRF-PLL. Para isso, o ensaio de degrau de carga é realizado por meio de uma manobra controlada através de contatores rápidos, conectando repentinamente um bloco de carga passiva em derivação no PAC da bancada.

Embora a resposta de máquinas síncronas a degraus de carga seja um tópico clássico em sistemas de potência, a contribuição deste trabalho reside na integração didático-experimental entre a modelagem analítica circuital e a instrumentação digital em tempo real em uma microrrede física real. Investiga-se aqui o comportamento das grandezas elétricas (tensão e corrente) no PAC e a dinâmica de frequência estimada via SRF-PLL. As previsões do modelo teórico são confrontadas com a resposta experimental obtida, o que inclui o monitoramento eletromecânico do rotor via encoder, delimitando a aplicabilidade do modelo circuital linear diante da resposta de uma máquina síncrona sem regulação automática de tensão (AVR).

2 Modelagem Analítica do Circuito sob Degrau de Carga

Para fundamentar a resposta transitória e estacionária das variáveis elétricas com base na teoria clássica de circuitos elétricos lineares (ALEXANDER; SADIKU, 2013; NILSSON; RIEDEL, 2015), modela-se a microrrede através do circuito equivalente ilustrado na Figura 1.

Para a análise circuital simplificada, o gerador síncrono é representado por uma fonte senoidal interna associada a uma impedância série equivalente $\mathbf{Z}_g = R_g + jX_g$. Essa representação não contempla explicitamente as dinâmicas subtransitórias e eletromecânicas da máquina durante a perturbação. O arranjo é alimentado pela máquina síncrona cuja força eletromotriz interna induzida $e_g(t)$ é senoidal pura, expressa no domínio do tempo e em notação fasorial

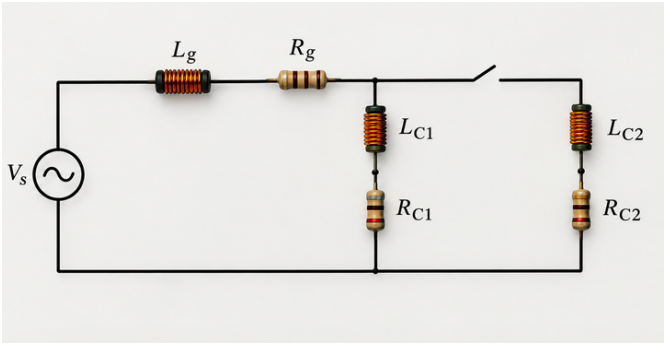


Figura 1: *Circuito equivalente elétrico monofásico da microrrede sob comutação de carga no PAC.*

por:

$$e_g(t) = \sqrt{2}E_g \sin(\omega t + \theta_0) \iff \dot{E}_g = E_g \angle \theta_0 \quad (1)$$

onde E_g é o valor eficaz da fem, $\omega = 2\pi f_0$ é a frequência angular nominal ($f_0 = 60$ Hz) e θ_0 é o ângulo de fase inicial.

A impedância interna série do gerador é modelada por:

$$\mathbf{Z}_g = R_g + j\omega L_g = R_g + jX_g \quad (2)$$

conectada diretamente ao PAC, cuja tensão instantânea terminal é designada por $v_{\text{PAC}}(t)$. O barramento alimenta uma resistência de carga permanente inicial dada por R_{c1} e dispõe de um ramo em derivação com carga resistiva comutável R_{c2}, acionado por uma chave de manobra rápida S (contator acionado pela plataforma dSPACE).

2.1 Condição com a Chave Aberta ($t < 0$): Carga Inicial

No intervalo pré-perturbação ($t < 0$), a chave S permanece aberta e o ramo contendo R_{c2} encontra-se eletricamente desconectado. O circuito opera como uma malha série única, cuja impedância total equivalente vista pela fonte é:

$$\mathbf{Z}_{\text{tot1}} = \mathbf{Z}_g + R_{c1} = (R_g + R_{c1}) + jX_g \quad (3)$$

Em coordenadas polares, expressa-se:

$$\mathbf{Z}_{\text{tot1}} = |\mathbf{Z}_{\text{tot1}}| \angle \phi_1 \quad (4)$$

onde o módulo e o ângulo de defasagem de impedância são:

$$|\mathbf{Z}_{\text{tot1}}| = \sqrt{(R_g + R_{c1})^2 + X_g^2} \quad (5)$$

$$\phi_1 = \arctan\left(\frac{X_g}{R_g + R_{c1}}\right) \quad (6)$$

Aplicando a Lei de Ohm fasorial, a corrente de linha fornecida pelo gerador no regime estacionário inicial resulta em:

$$\dot{I}_{g1} = \frac{\dot{E}_g}{\mathbf{Z}_{\text{tot1}}} = \frac{E_g \angle \theta_0}{|\mathbf{Z}_{\text{tot1}}| \angle \phi_1} = I_{g1} \angle (\theta_0 - \phi_1) \quad (7)$$

sendo $I_{g1} = E_g / |\mathbf{Z}_{\text{tot1}}|$ o valor eficaz. No domínio do tempo:

$$i_{g1}(t) = \sqrt{2} I_{g1} \sin(\omega t + \theta_0 - \phi_1) \quad (8)$$

A tensão terminal no barramento (v_{PAC1}) é estabelecida pelo divisor de tensão sobre a resistência R_{c1}:

$$\dot{V}_{\text{PAC1}} = \dot{I}_{g1} \cdot R_{c1} = \left(\frac{R_{c1}}{\mathbf{Z}_g + R_{c1}}\right) \dot{E}_g \quad (9)$$

A equação senoidal da tensão no PAC para $t < 0$ assume a forma:

$$v_{\text{PAC1}}(t) = \sqrt{2} V_{\text{PAC1}} \sin(\omega t + \theta_0 - \phi_1) \quad (10)$$

onde $V_{\text{PAC1}} = I_{g1} R_{c1}$. Em virtude da relação $R_{c1} \gg |\mathbf{Z}_g|$, a corrente de carga I_{g1} é modesta, acarretando queda de tensão residual reduzida na impedância interna do gerador e assegurando que $V_{\text{PAC1}} \approx E_g$.

A potência complexa consumida no barramento em $t < 0$ é descrita por:

$$\mathbf{S}_1 = \dot{V}_{\text{PAC1}} \cdot \dot{I}_{g1}^* = P_{e1} + jQ_{e1} \quad (11)$$

onde $P_{e1} = I_{g1}^2 R_{c1}$ e $Q_{e1} \approx 0$ em virtude da característica predominantemente resistiva da carga.

2.2 Condição com a Chave Fechada ($t \geq 0$): Carga em Paralelo

No instante $t = 0$, a chave S fecha bruscamente, conectando a resistência comutável R_{c2} em paralelo com a carga inicial R_{c1} nos terminais do PAC. A nova resistência equivalente de carga vista pelo barramento torna-se:

$$R_{\text{eq}} = \frac{R_{c1} R_{c2}}{R_{c1} + R_{c2}} \quad (12)$$

onde a conexão dos ramos em paralelo resulta na redução da resistência equivalente de carga:

$$R_{\text{eq}} < R_{c1} \quad (13)$$

A impedância total resultante do circuito pós-chaveamento é descrita por:

$$\mathbf{Z}_{\text{tot2}} = \mathbf{Z}_g + R_{\text{eq}} = (R_g + R_{\text{eq}}) + jX_g \quad (14)$$

com módulo e fase definidos por:

$$|\mathbf{Z}_{\text{tot2}}| = \sqrt{(R_g + R_{\text{eq}})^2 + X_g^2} \tag{15}$$

$$\phi_2 = \arctan\left(\frac{X_g}{R_g + R_{\text{eq}}}\right) \tag{16}$$

onde, em decorrência direta de (13), resulta $|\mathbf{Z}_{\text{tot2}}| < |\mathbf{Z}_{\text{tot1}}|$.

Embora a rede completa envolva mais de uma variável de estado em decorrência dos enrolamentos internos da máquina, a aproximação por uma equação diferencial de primeira ordem para a malha equivalente de Thévenin é plenamente aceitável quando o sistema opera com excitação constante de campo. Sob essa agregação paramétrica, a corrente total de linha fornecida pelo gerador pode ser aproximada por:

$$(R_g + R_{\text{eq}})i_{g2}(t) + L_g\frac{di_{g2}(t)}{dt} = \sqrt{2}E_g\sin(\omega t + \theta_0) \tag{17}$$

A solução geral é a soma linear da resposta natural homogênea $i_{\text{trans}}(t)$ com a resposta forçada particular $i_{\text{perm}}(t)$:

$$i_{g2}(t) = i_{\text{trans}}(t) + i_{\text{perm}}(t) \tag{18}$$

sendo:

$$i_{\text{perm}}(t) = \sqrt{2}I_{g2}\sin(\omega t + \theta_0 - \phi_2) \tag{19}$$

com o novo valor eficaz de regime permanente expresso por:

$$I_{g2} = \frac{E_g}{|\mathbf{Z}_{\text{tot2}}|} > I_{g1} \tag{20}$$

A componente transitória associada à constante de tempo indutiva equivalente $\tau = L_g/(R_g+R_{\text{eq}})$ possui formato exponencial:

$$i_{\text{trans}}(t) = A \cdot e^{-t/\tau} \tag{21}$$

A constante A é obtida a partir da continuidade da corrente nos enrolamentos indutivos: no instante imediatamente anterior e posterior à manobra, tem-se $i_{g2}(0^+) = i_{g1}(0^-)$. Aplicando essa condição inicial:

$$A = \sqrt{2}I_{g1}\sin(\theta_0 - \phi_1) - \sqrt{2}I_{g2}\sin(\theta_0 - \phi_2) \tag{22}$$

Substituindo (22) em (18), obtém-se a expressão da corrente:

$$i_{g2}(t) = \left[\sqrt{2}I_{g1}\sin(\theta_0 - \phi_1) - \sqrt{2}I_{g2}\sin(\theta_0 - \phi_2)\right]e^{-t/\tau} + \sqrt{2}I_{g2}\sin(\omega t + \theta_0 - \phi_2) \tag{23}$$

Para a tensão no PAC em $t \geq 0$, a expressão fasorial de regime estacionário pós-transitório torna-se:

$$\dot{V}_{\text{PAC2}} = \left(\frac{R_{\text{eq}}}{\mathbf{Z}_g + R_{\text{eq}}}\right)\dot{E}_g = V_{\text{PAC2}}\angle(\theta_0 - \phi_2) \tag{24}$$

cuja envoltória instantânea é dada por:

$$v_{\text{PAC2}}(t) = \sqrt{2}V_{\text{PAC2}}\sin(\omega t + \theta_0 - \phi_2) \tag{25}$$

Como a parcela $\mathbf{Z}_g$ ganha relevância percentual diante da resistência de carga reduzida R_{eq}, a variação da queda de tensão interna sobre a máquina síncrona é dada por:

$$\Delta\dot{V}_g = \mathbf{Z}_g(\dot{I}_{g2} - \dot{I}_{g1}) \tag{26}$$

resultando na variação da tensão terminal no PAC ($\Delta\dot{V}_{\text{PAC}} = -\mathbf{Z}_g\Delta\dot{I}_g$), provocando a redução na magnitude da tensão eficaz do barramento ($V_{\text{PAC2}} < V_{\text{PAC1}}$).

A potência ativa total solicitada pelas cargas em paralelo salta para:

$$P_{e2} = I_{g2}^2 R_{\text{eq}} > P_{e1} \tag{27}$$

gerando o degrau líquido de potência ativa $\Delta P = P_{e2} - P_{e1}$.

2.3 Comparativo Analítico das Condições do Circuito

Na Tabela 1, são apresentadas as expressões utilizadas para caracterizar as condições anteriores e posteriores à conexão da carga adicional.

Tabela 1: *Resumo Analítico da Resposta Circuital sob Manobra*

Grandeza	**Chave Aberta** ($t < 0$)	**Chave Fechada** ($t \geq 0$)
Carga	R_{c1}	$R_{\text{eq}} = R_{c1} \parallel R_{c2}$
Total	$\mathbf{Z}_g + R_{c1}$	$\mathbf{Z}_g + R_{\text{eq}}$
Corrente	$I_{g1} = E_g/\lvert\mathbf{Z}_{\text{tot1}}\rvert$	$I_{g2} = E_g/\lvert\mathbf{Z}_{\text{tot2}}\rvert > I_{g1}$
Tensão PAC	$V_{\text{PAC1}} = I_{g1}R_{c1}$	$V_{\text{PAC2}} = I_{g2}R_{\text{eq}} < V_{\text{PAC1}}$
Potência	$P_{e1} = I_{g1}^2 R_{c1}$	$P_{e2} = I_{g2}^2 R_{\text{eq}} > P_{e1}$

3 Metodologia Experimental

3.1 Infraestrutura Experimental do PTX Lab

Os ensaios foram desenvolvidos na microrrede experimental do laboratório PTX da UFRN, conforme ilustrado na Figura 2. O aparato físico utilizado é composto pelos seguintes elementos:

1. Conjunto motor-gerador constituído por um motor de corrente contínua (CC) de 3 kW,

220 V, atuando como máquina primária motriz operando sob controle de velocidade de rotação síncrona, mecanicamente acoplado a um gerador síncrono trifásico de polos salientes de 5 kVA, 220 V, 60 Hz, 4 polos, mantido com corrente de excitação de campo fixa (sem AVR);

2. Banco de cargas comutável acionado por contatores para comando do degrau no PAC, cujos parâmetros nominais de ensaio por fase são constituídos por uma carga base permanente de $R_{c1} = 300\,\Omega$ associada a um bloco comutável em derivação de $R_{c2} \approx 98{,}2\,\Omega$, resultando em uma resistência equivalente de regime pós-degrau de $R_{\text{eq}} \approx 74\,\Omega$;

3. Plataforma de prototipagem rápida dSPACE (DS1103) para aquisição digital dos sinais de tensão e corrente medidos no PAC, utilizando frequência de amostragem de 10 kHz. Sensores de corrente (LAH-25-NP) e de tensão (LV-25P) baseados no efeito Hall foram utilizados para medir as correntes e tensões, respectivamente, em conjunto com um encoder óptico incremental acoplado ao eixo rotórico para a monitorização da velocidade angular mecânica.

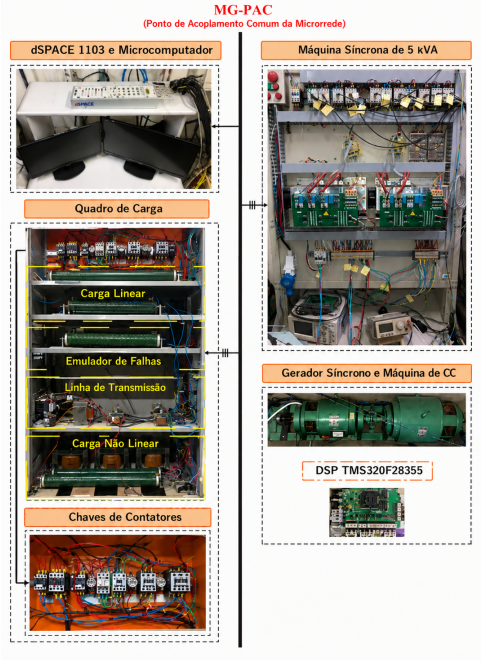


Figura 2: *Plataforma experimental utilizada para validação dos resultados.*

3.2 Rotinas Digitais de Estimação da Tensão, da Corrente, da Potência e da Frequência

Os sinais de tensão e corrente adquiridos pela plataforma dSPACE (DS1103) foram processados por meio das seguintes rotinas computacionais:

- Estimação das tensões e correntes eficazes via janela deslizante considerando o período fundamental de um ciclo ($T = 1/60$ s):

$$V_{\text{RMS}} = \sqrt{\frac{1}{T}\int_{t-T}^{t} v_{\text{PAC}}^2(\tau)\,d\tau} \tag{28}$$

$$I_{\text{RMS}} = \sqrt{\frac{1}{T}\int_{t-T}^{t} i_g^2(\tau)\,d\tau} \tag{29}$$

- Estimação da potência ativa via transformação de Park no referencial síncrono (dq), conforme fundamentado por MEDEIROS (2025) considerando a relação $p(t) = v_d i_d + v_q i_q$, em que v_d, i_d, v_q e i_q correspondem, respectivamente, às variáveis: tensão de eixo direto, corrente de eixo direto, tensão de eixo de quadratura e corrente de eixo de quadratura.

- Estimação da frequência instantânea via malha de sincronização de fase no referencial síncrono (SRF-PLL, do inglês *Synchronous Reference Frame Phase-Locked Loop*), equipada com filtro passa-baixas digital de segunda ordem para atenuação de ruídos de medição e harmônicos de chaveamento (BLAABJERG; TEODORESCU, 2006). A dinâmica da frequência estimada no PAC foi sincronizada com a plataforma de aquisição e confrontada com os sinais de rotação obtidos pelo encoder óptico incremental acoplado ao rotor.

4 Resultados e Discussão

A análise experimental das deduções analíticas apresentadas na Seção 2 foi realizada a partir da manobra de chaveamento do bloco de carga adicional R_{c2} em paralelo com a carga inicial R_{c1}.

Na Figura 3, são apresentadas as respostas experimentais de tensão, corrente, potência ativa e frequência estimada durante a inserção da carga adicional. O gerador começa fornecendo aproximadamente 180 W à carga linear. Em $t \approx 6{,}42$ s, as chaves de contatores conectam o bloco de carga em paralelo, elevando a potência de saída para aproximadamente 500 W a partir de $t > 6{,}6$ s. Nas Figs. 3(a) e 3(b), são apresentados, respectivamente, os sinais de corrente

e tensão da fase A, incluindo seus valores instantâneos e eficazes.

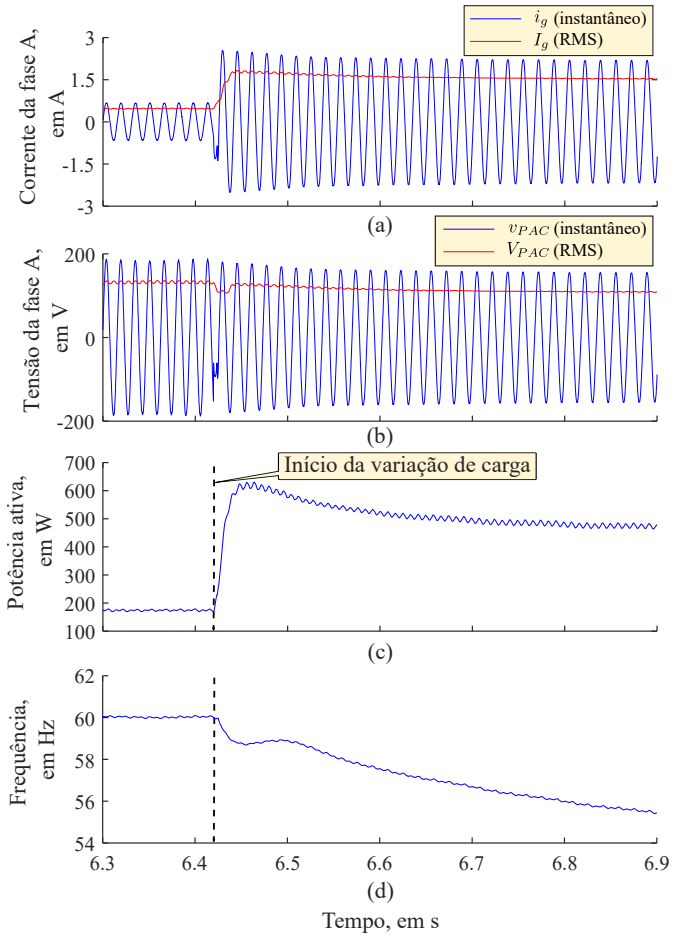


Figura 3: *Comportamento das variáveis elétricas, medidas no PAC, mediante a aplicação do degrau de carga em ambiente experimental: (a) Corrente elétrica da fase A (valores instantâneo e RMS), em A; (b) Tensão da fase A (valores instantâneo e RMS), em V; (c) Potência ativa, em W; (d) Frequência, em Hz.*

4.1 Dinâmica Transitória da Potência Ativa

De acordo com os resultados apresentados na Figura 3(c), a conexão da carga adicional foi acompanhada por um aumento da potência ativa medida, correspondendo a um incremento de $\Delta P \approx 320$ W (aumento relativo de aproximadamente 177,8% em relação à condição inicial). Do ponto de vista circuital, a conexão de ramos em paralelo reduz a resistência equivalente vista pela máquina ($R_{\text{eq}} < R_{c1}$), exigindo o aumento imediato da corrente elétrica fornecida pelo gerador síncrono.

A potência ativa medida apresenta uma transição gradual, em vez de uma variação instantânea entre os valores anteriores e posteriores ao chaveamento. Essa característica decorre das propriedades físicas do circuito indutivo: a oposição natural da indutância síncrona a variações bruscas de corrente impõe uma constante de tempo $\tau = L_g/(R_g + R_{\text{eq}})$ ao transitório, combinada ao tempo de acomodação inerente à janela de integração móvel do algoritmo de potência RMS processado pela plataforma dSPACE.

Como o sistema opera sem regulação automática de tensão (sem AVR), a força eletromotriz interna E_g permanece essencialmente constante durante o transitório. Dessa forma, a maior circulação de corrente provoca uma elevação na queda fasorial sobre a impedância interna do gerador ($\Delta \dot{V}_g = \mathbf{Z}_g \Delta \dot{I}_g$), resultando na redução da tensão terminal observada experimentalmente.

Na Tabela 2, apresenta-se o confronto quantitativo entre os valores obtidos pela formulação circuital e os valores medidos experimentalmente na bancada do PTX, antes e após a inserção do degrau de carga, considerando a relação de carga no PAC com $R_{c1} = 300\,\Omega$ e a resistência equivalente pós-chaveamento $R_{\text{eq}} \approx 74\,\Omega$.

Tabela 2: *Comparativo Quantitativo entre Resultados Teóricos e Experimentais sob Degrau de Carga*

	Pré-Degrau ($t < 6{,}42$ s)			**Pós-Degrau** ($t > 6{,}60$ s)		
Grandeza	**Teórico**	**Exp.**	**Erro (%)**	**Teórico**	**Exp.**	**Erro (%)**
V_{RMS} (V)	135,0	≈ 135,0	0,0%	112,0	≈ 112,0	0,0%
I_{RMS} (A)	0,45	≈ 0,52	13,5%	1,51	≈ 1,58	4,4%
P_e (W)	182,3	≈ 180,0	1,3%	508,0	≈ 500,0	1,6%
f (Hz)	60,0	60,0	0,0%	—	≈ 55,5	—

4.2 Comportamento Dinâmico da Frequência e Resposta Eletromecânica

O degrau de potência ativa consumida pelas cargas repercute de forma imediata no equilíbrio eletromecânico do grupo gerador (MACHOWSKI; BIALEK; BUMBY, 2020). Pela equação de oscilação da máquina síncrona (*swing equation*), expressa em unidades por centro (*pu*) como $2H\frac{d\omega_{pu}}{dt} = P_{m_pu} - P_{e_pu} - D_{pu}(\omega_{pu} - 1)$, o balanço entre o torque mecânico motriz fornecido pela máquina primária e o torque eletromagnético resistente solicitado pela armadura governa a aceleração angular do rotor. A modelagem dinâmica de transitórios e a estimação em tempo real das constantes de inércia e amortecimento sob perturbações fundamenta-se nos desenvolvimentos reportados em (MEDEIROS, 2025).

No instante imediatamente subsequente ao fechamento da chave, a potência motriz fornecida pelo motor CC não se altera instantaneamente devido à inércia mecânica e aos atrasos de transporte do acionamento primário ($P_m \approx P_{e1}$). Em contrapartida, a demanda elétrica salta bruscamente para o novo pa-

tamar (P_{e2}), gerando um déficit imediato de potência ativa ($P_m < P_e$). Para suprir esse desequilíbrio energético, o sistema extrai energia cinética armazenada nas massas rotantes do rotor, induzindo uma desaceleração mecânica do eixo com taxa de variação negativa da frequência (RoCoF).

Na Figura 3(d), observa-se a redução da frequência estimada pelo SRF-PLL após a inserção da carga. Durante o intervalo analisado, a frequência estimada apresenta tendência decrescente. Não é possível identificar, a partir dos resultados apresentados, um novo regime permanente. A confrontação entre a resposta estimada pela malha de sincronização SRF-PLL e o monitoramento da rotação mecânica pelo encoder confirma que a trajetória decrescente de 60 Hz até aproximadamente 55,5 Hz reflete a desaceleração física real do eixo rotórico diante da sobrecarga imposta, e não um artefato numérico transitório de alinhamento de fase.

5 Conclusões

Neste trabalho, realizou-se uma análise teórica simplificada e uma investigação experimental da resposta de um gerador síncrono a um degrau de carga.

A formulação circuital permitiu equacionar os transitórios de corrente, tensão e potência ativa, corroborados quantitativamente pelo salto medido de $\Delta P \approx 320$ W e pelo afundamento proporcional de tensão terminal. As rotinas implementadas na plataforma dSPACE em conjunto com o encoder óptico permitiram registrar com fidelidade as grandezas elétricas e a resposta dinâmica da frequência estimada no PAC da microrrede experimental.

Referências

ALEXANDER, C. K.; SADIKU, M. N. O. *Fundamentos de Circuitos Elétricos*. 5. ed. Porto Alegre: McGraw-Hill, 2013.

BEVRANI, H.; WATANABE, M.; MITANI, Y. *Microgrid Dynamics and Control: One Modern Synthesis*. Hoboken: John Wiley & Sons, 2014.

BLAABJERG, F.; TEODORESCU, R. Overview of control and grid synchronization for distributed power generation systems. *IEEE Transactions on Industrial Electronics*, IEEE, v. 53, n. 5, p. 1398–1409, 2006.

FITZGERALD, A. E.; KINGSLEY, C.; UMANS, S. D. *Electric Machinery*. 6. ed. New York: McGraw-Hill, 2003.

GRAINGER, J. J.; STEVENSON, W. D. *Power System Analysis*. New York: McGraw-Hill, 1994.

KUNDUR, P.; BALU, N. J.; LAUBY, M. G. *Power System Stability and Control*. New York: McGraw-Hill, 1994.

MACHOWSKI, J.; BIALEK, J. W.; BUMBY, J. R. *Power System Dynamics: Stability and Control*. 3. ed. Hoboken: John Wiley & Sons, 2020.

MEDEIROS, R. P. d. Real-time estimation of power system inertia and damping constant using forgetting factor rls. In: *2025 Brazilian Power Electronics Conference (COBEP)*. [S.l.: s.n.], 2025. p. 1–5.

MILANO, F. *Power System Modelling and Scripting*. London: Springer Science & Business Media, 2010.

NILSSON, J. W.; RIEDEL, S. A. *Circuitos Elétricos*. 10. ed. São Paulo: Pearson, 2015.

Análise da Influência de Variações de Atraso em Controladores GPC Aplicados ao Modelo de Slate para Pressão Arterial

***Fernando Henrique de Couto*[1]**
***Aurea da Silva Cabral*[2]**
***André Laurindo Maitelli*[3]**

[1] *Departamento de Engenharia de Computação e Automação, UFRN, <maitelli@dca.ufrn.br>*

Resumo

O controle automático da pressão arterial média (PAM) utilizando fármacos vasodilatadores, como o nitroprussiato de sódio (NPS), é um desafio clássico na engenharia de controle biomédico. A principal dificuldade reside na variação contínua da dinâmica do paciente, especialmente no atraso de transporte (tempo morto) do sistema circulatório. Este artigo analisa o impacto da variação estocástica do atraso na malha de controle da PAM utilizando um Controlador Preditivo Generalizado (GPC). O comportamento dinâmico do paciente é representado pelo modelo farmacocinético/farmacodinâmico de Slate. Simulações demonstram que flutuações no tempo morto geram oscilações severas (*chattering*) no sinal de controle nominal. Para mitigar esse efeito, investiga-se a incorporação do polinômio de filtragem de robustez $T(z^{-1})$ na formulação do GPC, avaliando seu impacto na estabilidade e no esforço de controle.

Palavras Chaves: Controle Preditivo Generalizado (GPC), Pressão Arterial Média (PAM), Atraso de Transporte (ou Tempo Morto), Filtro de Robustez e Modelo de Slate.
.

Abstract

The automatic control of mean arterial pressure (MAP) using vasodilator drugs, such as sodium nitroprusside (SNP), is a challenging problem in biomedical control enginneering. A major difficulty lies in the continuous variation of the patient's dynamics, particularly in the transport delay of the circulatory system. This work analyses the impact of stochastic transport-delay variations on the MAP control loop using a Generalized Predictive Controller (GPC). The patient's dynamic behavior is represented by the Slate pharmacokinetic/pharmacodynamic model. Simulations show that variations in the transport delay generate severe oscillations (chattering) in the nominal control signal. To mitigate this effect, the incorporation of the robustness polynomical $T(z^{-1})$ into the GPC formulation is investigated, evaluating its impact on control stability and effort.

Keywords: Generalized Predictive Control (GPC); Mean Arterial Pressure (MAP); Transport Delay; Robustness Filter; Slate Model.

1 Introdução

A regulação automatizada da Pressão Arterial Média (PAM) em pacientes submetidos a cirurgias críticas ou em unidades de terapia intensiva é frequentemente realizada pela infusão intravenosa de nitroprussiato de sódio (NPS). Devido às respostas altamente não lineares e variantes no tempo do corpo humano, o controle manual da infusão exige atenção constante do corpo médico clínico.

O modelo desenvolvido por Slate (SLATE, 1980), frequentemente adaptado na literatura de controle, é utilizado para representar a relação entre a taxa de infusão de NPS e a variação da PAM. Neste trabalho, adota-se a formulação apresentada em *Modern Control Systems* (DORF; BISHOP, 2017), que caracteriza o sistema biológico como uma dinâmica de segunda ordem com atraso de transporte puro, associado à dinâmica integradora da bomba de infusão. Na prática clínica, o atraso de transporte não é um parâmetro fixo; ele varia continuamente devido a flutuações no débito cardíaco, alterações no volume sanguíneo e respostas termodinâmicas do paciente.

Controladores tradicionais e estratégias preditivas assumem, em sua formulação nominal, que o atraso do sistema é constante. Quando a planta real apresenta variações nesse tempo morto, surge um erro de predição interno que leva a ações de controle excessivamente agressivas, resultando em *chattering* no atuador (bomba de infusão).

O Controle Preditivo Generalizado (GPC) é uma estratégia natural para lidar com sistemas com atraso. Este trabalho avalia a degradação de desempenho do GPC frente a atrasos variantes e propõe a análise da utilização do polinômio de robustez $T(z^{-1})$ associado ao modelo CARIMA para restaurar a estabilidade do sinal de controle.

2 Revisão Bibliográfica

A literatura de engenharia de controle aplicada à biomedicina tem se dedicado extensivamente ao desafio da regulação automatizada da Pressão Arterial Média (PAM) utilizando fármacos vasodilatadores ou vasopressores. O trabalho fundacional de Slate (SLATE, 1980) estabeleceu as bases matemáticas rigorosas para compreender a dinâmica farmacocinética e farmacodinâmica da infusão de nitroprussiato de sódio (NPS). Na literatura de sistemas de controle, como apresentado por Dorf e Bishop (DORF; BISHOP, 2017), este modelo biológico é frequentemente adaptado para uma função de transferência de segunda ordem associada a um tempo morto puro e ao comportamento integrador da bomba de infusão.

O grande obstáculo no controle hemodinâmico reside no fato de que o atraso de transporte (tempo morto) degrada o desempenho do sistema em malha fechada, diminuindo a margem de fase e podendo induzir oscilações severas ou instabilidade. Além disso, devido à variabilidade inter e intra-paciente, este atraso não é um parâmetro constante; ele varia continuamente com as alterações nas condições fisiológicas, como flutuações no débito cardíaco e no volume sanguíneo.

Para compensar o efeito prejudicial deste atraso, diversas estratégias de controle têm sido propostas. Martin *et al.* (MARTIN; SCHNEIDER; SMITH, 1987) propuseram um Controle Adaptativo de Múltiplos Modelos (MMAC) que incorpora um Preditor de Smith clássico para remover eficazmente o tempo de atraso da infusão do NPS, permitindo ao controlador identificar e se adaptar a diferentes perfis de resposta dos pacientes. Mais recentemente, Khodadadi *et al.* (KHODADADI; GHADIRI; DEHGHANI, 2026) desenvolveram uma arquitetura robusta que integra um controlador PID de ordem fracionária com autossintonia *fuzzy* e um Preditor de Smith Modificado (MSP). Esta estrutura foi projetada para lidar com a regulação simultânea da PAM e do débito cardíaco, demonstrando que a compensação explícita de atrasos variantes no tempo é crucial para manter a estabilidade diante das incertezas dinâmicas dos pacientes.

Abordagens no domínio da frequência também têm sido exploradas para contornar a variabilidade dos parâmetros fisiológicos e os grandes atrasos variantes. Tasoujian *et al.* (TASOUJIAN et al., 2019) investigaram estratégias de Controle por Modelo Interno (IMC-PID) e de *loop-shaping* com parâmetros variantes para a regulação da PAM através da infusão de fenilefrina. Recorrendo à aproximação de Padé, os autores aproximaram a dinâmica de dimensão infinita inerente ao atraso por um sistema de fase não mínima, aplicando o teorema do pequeno ganho para extrair condições necessárias e suficientes de estabilidade robusta diante das incertezas.

Em suma, a literatura recente e clássica confirma que, seja utilizando Preditores de Smith, estruturas IMC ou Estratégias Preditivas (como o GPC), a inclusão de mecanismos de robustez para lidar com erros de modelagem do atraso — à semelhança do filtro $T(z^{-1})$ investigado neste artigo — é fundamental. Estas abordagens são essenciais para mitigar o erro de predição gerado pela dinâmica variante do tempo morto, evitando o *chattering* no atuador e garantindo a segurança clínica.

3 Modelagem do Sistema

3.1 Modelo de Slate e Atuador

A resposta da PAM à infusão de NPS pode ser descrita como a variação da pressão $\Delta P(s)$ em relação à taxa de infusão do fármaco $U(s)$. O modelo fisiológico do paciente adotado é uma função de transferência de primeira ordem com tempo morto. O atuador (bomba de infusão) é modelado como um integrador

puro $(1/s)$, o que torna o sistema global em malha aberta de segunda ordem. A função de transferência da planta (paciente) é dada por:

$$G(s) = \frac{\Delta P(s)}{U(s)} = \frac{Ke^{-T_i s}(1 + \alpha e^{-T_c s})}{\tau s^2 + s} \quad (1)$$

onde τ representa a constante de tempo da resposta fisiológica do paciente, T_i é o atraso de transporte inicial (tempo morto) e T_c o atraso de recirculação.

3.2 Variação Dinâmica do Atraso

Para emular o cenário clínico realista, o atraso T_d não é mantido constante. Em vez disso, é modelado como uma variável variante no tempo $T_d(t)$, sujeita a perturbações contínuas modeladas como um passeio aleatório (*random walk*) filtrado em baixa frequência:

$$T_d(t) = T_{nominal} + \delta(t) \quad (2)$$

onde $\delta(t)$ representa a variação de atraso restrita entre limites físicos máximos e mínimos.

4 Controlador Preditivo Generalizado (GPC)

O algoritmo GPC baseia-se no modelo CARIMA (*Controlled Auto-Regressive Integrated Moving Average*) para prever as saídas futuras da planta. A dinâmica do sistema sob essa representação é descrita pela seguinte equação polinomial:

$$A(z^{-1})y(t) = B(z^{-1})u(t-1) + \frac{e(t)}{\Delta} \quad (3)$$

onde $y(t)$ é a saída do sistema, $u(t)$ é a entrada de controle, $e(t)$ é uma sequência de ruído branco de média nula que representa as perturbações, $\Delta = 1 - z^{-1}$ é o operador de diferença, e $A(z^{-1})$ e $B(z^{-1})$ são polinômios no operador de atraso discreto z^{-1}.

A partir das predições geradas por este modelo, a função de custo a ser minimizada a cada instante de amostragem é dada por:

$$J = \sum_{j=N_1}^{N_2} [y(t+j) - w(t+j)]^2 + \lambda \sum_{j=1}^{N_u} [\Delta u(t+j-1)]^2 \quad (4)$$

onde N_1 e N_2 definem o horizonte de predição, N_u o horizonte de controle, λ o fator de ponderação do esforço de controle, y a saída predita, w a referência de pressão e Δu o incremento de controle.

4.1 Filtro de Robustez T

A presença de distúrbios não modelados, como a variação estrutural de $T_d(t)$, compromete o estimador de saída livre do GPC. Para mitigar esse efeito, o polinômio observador $T(z^{-1})$ é introduzido diretamente na equação diofantina utilizada para o cálculo das predições futuras, modificando a identidade clássica para a seguinte forma:

$$T(z^{-1}) = E_j(z^{-1})A(z^{-1})\Delta + z^{-j}F_j(z^{-1}) \quad (5)$$

onde $E_j(z^{-1})$ e $F_j(z^{-1})$ são os polinômios de ordem dependente do passo de predição j. Ao ser incorporado na equação de predição, o polinômio $T(z^{-1})$ atua intrinsecamente como um filtro passa-baixa sobre os dados passados da planta, atenuando a sensibilidade do controlador ao erro de modelagem em alta frequência sem alterar os polos de malha fechada do sistema nominal.

5 Metodologia

A malha de controle foi implementada em ambiente de simulação discreto com tempo de amostragem definido. O sistema operou sob as seguintes restrições reais de bomba de infusão clínica:

- Saturação de atuador: $u_{min} \leq u(t) \leq u_{max}$
- Taxa de variação (Slew Rate): $\Delta u_{min} \leq \Delta u(t) \leq \Delta u_{max}$

Foram avaliados três cenários distintos:

1. **Cenário Ideal**: Atraso da planta fixo, idêntico ao modelo interno do GPC.
2. **Cenário com Atraso Variante**: Variação contínua e aleatória do atraso de transporte, com a formulação padrão do GPC ($T(z^{-1}) = 1$).
3. **Cenário de Robustez**: Mesmo perfil de atraso variante, ativando o filtro polinomial de primeira ordem $T(z^{-1}) = 1 - \alpha z^{-1}$.

6 Resultados e Discussões

Os testes computacionais avaliaram o desempenho da malha de controle de pressão arterial sob a influência de diferentes níveis de sensibilidade do paciente ao medicamento, representados pela variação do ganho da planta (K). Para ilustrar o impacto dessa variação fisiológica na dinâmica de controle, os resultados foram divididos em cenários de baixa, média

(nominal) e alta sensibilidade.No cenário de baixa sensibilidade ($K = -0.5$), a variável de saída controlada (y_1) apresenta grande lentidão na resposta ao degrau. Ao longo dos 1800 segundos de simulação, a Pressão Arterial Média (PAM) decai de forma excessivamente suave de 140 mmHg para aproximadamente 105 mmHg, falhando em alcançar a referência estipulada de 100 mmHg dentro da janela de tempo analisada. Consequentemente, o sinal de infusão u_1 cresce de maneira contínua e sem atingir o estado estacionário, ultrapassando 50 ml/h ao final do período. A análise do incremento de controle (du_1) mostra flutuações ruidosas, porém de magnitude muito baixa (inferiores a 0.06), evidenciando a limitação do controlador em forçar uma resposta mais rápida de uma planta pouco reativa.

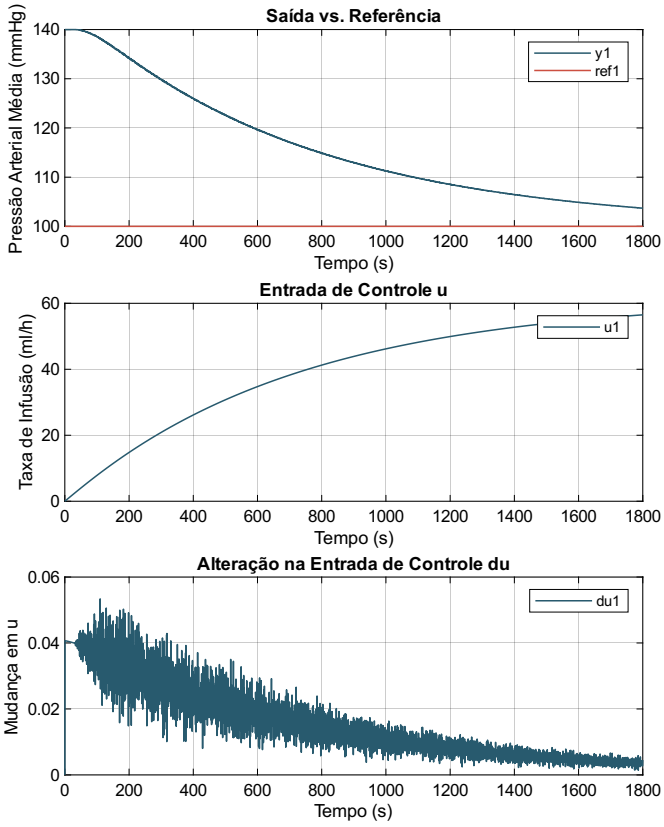


Figura 1: *Resultados do controle para paciente com baixa sensibilidade ($K = -0.5$): resposta lenta e dificuldade em alcançar a referência.*

Para o paciente com sensibilidade nominal ($K = -1$), o desempenho do sistema demonstra um rastreamento eficaz, embora com uma fase transiente pronunciada. A pressão arterial atinge a referência de 100 mmHg com um sub-sinal (*undershoot*), onde a pressão cai até um valor próximo a 88 mmHg em torno de 400 segundos, antes de se recuperar e estabilizar-se completamente em 1000 segundos. O sinal de atuação u_1 eleva-se até um pico de aproximadamente 42 ml/h durante o período de maior erro, para então acomodar-se de forma clinicamente segura em torno de 31 ml/h. O esforço incremental (du_1) concentra picos de variação (entre -5 e 5) estritamente na fase transiente (até 300 segundos), decaindo para zero logo em seguida e garantindo a ausência de *chattering* em regime permanente.

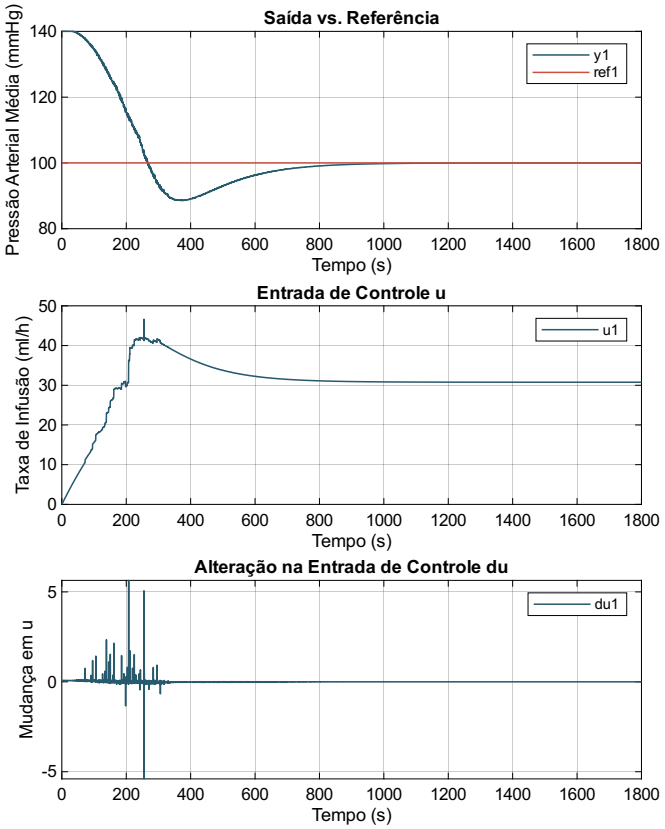


Figura 2: *Resultados do controle para paciente com sensibilidade nominal ($K = -1$): rastreamento eficaz com sub-amortecimento inicial.*

Por fim, no cenário de alta sensibilidade ($K = -8$), o sistema reage de forma muito mais abrupta e rápida à infusão de nitroprussiato de sódio. A PAM apresenta oscilações iniciais visíveis entre 110 e 120 mmHg, mas consegue atingir e estabilizar-se na referência de 100 mmHg de forma muito mais rápida, em pouco mais de 300 segundos. Devido à alta reatividade do paciente, o sinal de controle u_1 apresenta uma ação inicial oscilatória e agressiva, chegando a registrar valores negativos antes de se estabilizar rapidamente em uma taxa de infusão muito baixa, próxima a 4 ml/h. A variação da entrada de controle (du_1) atesta a ocorrência de *chattering* inicial expressivo, que é eficientemente neutralizado assim que a pressão se estabiliza no valor desejado.

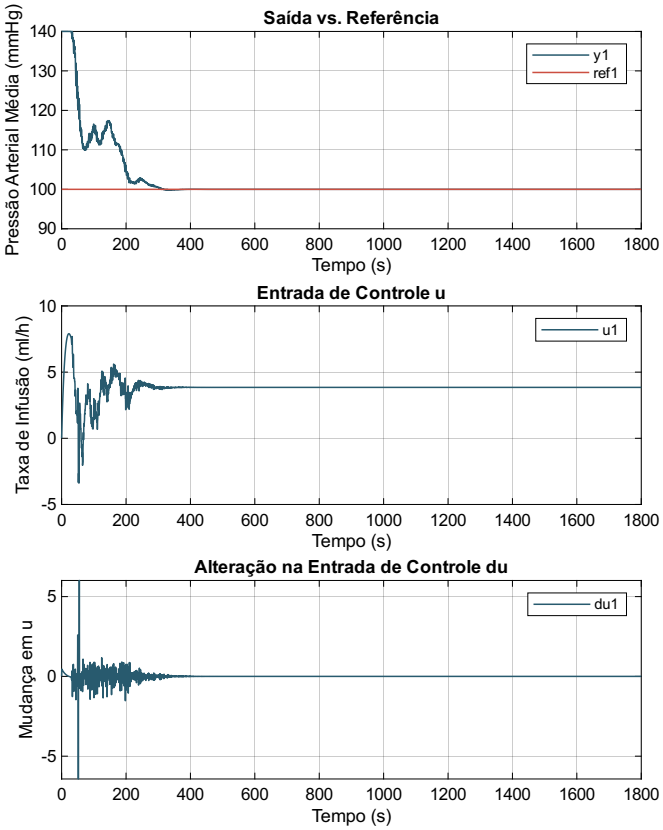


Figura 3: *Resultados do controle para paciente com alta sensibilidade ($K = -8$): resposta rápida com oscilações e chattering inicial na atuação.*

7 Conclusão

Neste trabalho, demonstrou-se que a regulação automatizada da Pressão Arterial Média (PAM) por meio da infusão de nitroprussiato de sódio (NPS) enfrenta desafios críticos inerentes à dinâmica variante dos pacientes. A análise do modelo de Slate evidenciou que as flutuações contínuas no atraso de transporte (tempo morto) causam uma severa degradação no desempenho do Controlador Preditivo Generalizado (GPC) em sua formulação padrão. Sem compensação, o erro de predição interno induz o sistema a comportamentos oscilatórios agressivos e inaceitáveis, caracterizados por um intenso *chattering* na bomba de infusão e instabilidade no rastreamento da pressão.

Para superar essa vulnerabilidade, a incorporação do filtro polinomial de robustez $T(z^{-1})$ provou-se uma estratégia altamente eficaz. Os testes computacionais confirmaram que a sintonia adequada deste filtro (como demonstrado com $\alpha = 0.9$) reconfigura a dinâmica de predição do GPC, permitindo que o controlador atue como um filtro passa-baixa sobre os distúrbios de tempo morto e rejeite o ruído de modelagem.

Conclui-se que esta abordagem não apenas restaura a precisão e a estabilidade da malha de controle, mas também garante uma ação de controle notavelmente suave. Dessa forma, a estratégia proposta atende aos rigorosos requisitos da prática clínica, assegurando a integridade mecânica dos atuadores físicos e garantindo a segurança e estabilidade hemodinâmica de pacientes críticos submetidos a terapias preditivas.

Referências

DORF, R. C.; BISHOP, R. H. *Modern Control Systems*. [S.l.]: Pearson Education, 2017.

KHODADADI, H.; GHADIRI, H.; DEHGHANI, A. Modified smith predictor-based fuzzy self-tuning fopid controller for mean arterial pressure regulation. *Scientific Reports*, 2026.

MARTIN, J. F.; SCHNEIDER, A. M.; SMITH, N. T. Multiple-model adaptive control of blood pressure using sodium nitroprusside. *IEEE Transactions on Biomedical Engineering*, BME-34, n. 8, p. 603–611, 1987.

SLATE, J. B. *Model-based design of a controller for infusing sodium nitroprusside during postsurgical hypertension*. Tese (Doutorado) — Dept. Elect. Eng., Univ. Wisconsin-Madison, Madison, WI, 1980.

TASOUJIAN, S. et al. Robust imc-pid and parameter-varying control strategies for automated blood pressure regulation. *International Journal of Control, Automation and Systems*, v. 17, n. 7, p. 1803–1813, 2019.

Gêmeos Digitais na Indústria 4.0: Aplicação em Tanques Acoplados

***Kaillane Araújo da Silva*[1]**
***Fabio Meneghetti Ugulino de Araujo*[1]**

[1] *Departamento de Engenharia de Computação e Automação, UFRN, <kaillane.silva.707@ufrn.edu.br>*

[1] *Departamento de Engenharia de Computação e Automação, UFRN, <meneghet@dca.ufrn.br>*

Resumo

A Quarta Revolução Industrial, ou Indústria 4.0, estabelece-se por meio da convergência entre tecnologias digitais, automação avançada, *Big Data* e ciência de dados. Nesse cenário, surge o conceito de Gêmeos Digitais (*Digital Twins*), que se apresenta como uma ferramenta fundamental, sendo uma réplica virtual de um objeto ou sistema do mundo real, possibilitando o monitoramento de sistemas físicos em tempo real. Este trabalho propõe o estudo da planta *Coupled Tanks*, da Quanser, voltada ao aprendizado de processos industriais de controle de nível. A metodologia adotada fundamenta-se na Engenharia de Sistemas Baseada em Modelos (MBSE) e na exploração de soluções comerciais sólidas, com o intuito de integrar algoritmos de Inteligência Artificial e sensoriamento remoto. O objetivo do projeto é a capacitação técnica no domínio dos Gêmeos Digitais, culminando no desenvolvimento de uma interface computacional capaz de realizar o prognóstico de falhas e a otimização de estratégias de controle. Desse modo, busca-se a mitigação de riscos operacionais e a redução dos custos de manutenção dos sistemas.

Palavras-chave: Gêmeos Digitais. Indústria 4.0. Tanques Acoplados. Controle em Cascata. Node-RED.

Abstract

The Fourth Industrial Revolution, or Industry 4.0, is established through the convergence of digital technologies, advanced automation, Big Data, and data science. In this context, the concept known as Digital Twins emerges as a fundamental tool, as it is an exact virtual replica of a real-world object or system, enabling the real-time monitoring of physical systems. This work proposes the study of the Quanser Coupled Tanks, a plant designed for learning industrial level control processes. The adopted methodology is based on Model-Based Systems Engineering (MBSE) and the exploration of robust commercial solutions, with the aim of integrating Artificial Intelligence algorithms and remote sensing. The objective of the project is to provide technical training in the field of Digital Twins, culminating in the development of a computational interface capable of performing fault prognosis and optimizing control strategies. Thus, the project seeks to mitigate operational risks and reduce system maintenance costs.

Keywords: Digital Twins. Industry 4.0. Coupled Tanks. Cascade Control. Node-RED.

1 Introdução

A partir de 2011, surge na literatura o termo Quarta Revolução Industrial, ou Indústria 4.0. O conceito foi apresentado na feira comercial Hannover Messe, na Alemanha, onde um grupo de trabalho voltado à Indústria 4.0 foi patrocinado pelo governo federal alemão (LIN et al., 2018). A Quarta Revolução Industrial surge com o objetivo de possibilitar a personalização de produtos de acordo com as necessidades dos clientes (FRONTONI et al., 2018). A Indústria 4.0 é fomentada por tecnologias fundamentais, como

Big Data, sistemas integrados, computação em nuvem, Internet das Coisas (IoT), segurança da informação, simulação e manufatura aditiva.

Mais recentemente, surge o conceito de Indústria 5.0, que amplia a perspectiva da transformação industrial ao incorporar aspectos relacionados à centralidade humana, à sustentabilidade e à resiliência dos sistemas produtivos. Diferentemente de uma substituição da Indústria 4.0, essa nova abordagem busca complementar seus avanços tecnológicos, direcionando-os também ao bem-estar humano e à responsabilidade ambiental (BREQUE; NUL; PETRIDIS, 2021).

No contexto da Indústria 4.0, identificam-se tecnologias como os Gêmeos Digitais (*Digital Twins*), considerados ferramentas importantes para a integração entre os ambientes físico e virtual. As atuais tecnologias de informação e comunicação são utilizadas para representar e conectar esses dois ambientes. De acordo com Schleich et al. (2017), o Gêmeo Digital estabelece uma relação bidirecional entre um artefato físico e um conjunto de modelos virtuais. Essa relação permite utilizar a representação virtual durante diferentes etapas do ciclo de vida do produto, incluindo projeto, fabricação e manutenção ou serviço. Glaessgen e Stargel (2012) ampliam essa perspectiva ao apresentar o Gêmeo Digital como uma representação virtual capaz de integrar modelos de alta fidelidade, dados de sensores, histórico de operação e informações relacionadas à condição do sistema, contribuindo para o monitoramento, a previsão e o gerenciamento do ciclo de vida.

No âmbito desta pesquisa de iniciação científica, utilizou-se o sistema de *Coupled Tanks* da Quanser, uma planta didática composta por tanques interligados e sensores de nível, que mimetiza as dinâmicas de indústrias de saneamento e refino de petróleo. O objetivo central deste trabalho consiste no desenvolvimento de um protótipo funcional voltado ao monitoramento contínuo do comportamento do processo, integrando fundamentos teóricos e aplicações ao ciclo de vida do Gêmeo Digital.

2 Gêmeos Digitais e Tecnologias da Indústria 4.0

Nesta seção são apresentados os principais conceitos e fundamentos necessários para a compreensão do desenvolvimento deste trabalho. Inicialmente, são abordados os conceitos relacionados à Indústria 4.0 e às tecnologias que contribuem para a integração entre os ambientes físico e digital. Em seguida, são discutidos os Gêmeos Digitais, considerando sua evolução conceitual, características e aplicações em sistemas industriais. Por fim, são apresentados os fundamentos relacionados à modelagem, simulação e controle de processos, estabelecendo a base teórica necessária para a aplicação do conceito de Gêmeo Digital ao sistema estudado.

2.1 Indústria 4.0

Segundo Pfeiffer (2017), o termo Indústria 4.0 ganhou destaque na Alemanha em 2011, passando a representar uma visão de transformação da produção industrial. Nesse cenário, o conceito passou a ocupar espaço tanto no debate tecnológico quanto nas discussões econômicas e sociais relacionadas ao futuro da produção. Zhong et al. (2017) destacam que a Indústria 4.0 busca promover maior flexibilidade nos processos de manufatura, personalização em massa, melhoria da qualidade e aumento da produtividade. Para isso, diferentes tecnologias são empregadas na construção de sistemas de manufatura inteligentes, entre elas a Internet das Coisas (IoT), os Sistemas Ciberfísicos (CPS), a computação em nuvem, a análise de *Big Data* e as tecnologias de informação e comunicação.

2.2 Gêmeos Digitais

Os Gêmeos Digitais possuem antecedentes no programa Apollo, desenvolvido pela NASA, especialmente a missão Apollo 11, realizada em 1969. Durante as missões, modelos e sistemas de simulação eram utilizados em solo para reproduzir condições de operação das espaçonaves, permitindo aos engenheiros analisar e mitigar erros. Embora essas aplicações ainda não constituíssem Gêmeos Digitais em seu sentido atual, contribuíram para o desenvolvimento do conceito.

A partir de 2002, Michael Grieves (GRIEVES, 2002) apresentou, no contexto do *Product Lifecycle Management* (PLM), uma estrutura conceitual formada por três elementos principais: o espaço físico, o espaço virtual e um mecanismo de conexão responsável pelo fluxo de informações entre esses dois ambientes. Essa proposta estabelecia a ideia de uma representação virtual associada ao produto físico durante seu ciclo de vida. Posteriormente, o modelo passou a ser denominado *Information Mirroring Model*, mantendo como princípio a conexão entre os ambientes físico e virtual.

Posteriormente, Glaessgen e Stargel (2012) ampliaram essa concepção ao integrar modelos de simulação, dados de sensores e informações sobre o estado do sistema ao longo de seu ciclo de vida. Em

2017, Schleich et al. (2017) destacaram a relação entre o artefato físico e seus modelos virtuais, considerando características como interoperabilidade, escalabilidade, expansibilidade e fidelidade, contribuindo para a consolidação do conceito que conhecemos atualmente.

2.3 Sistema de Tanques Acoplados da Quanser

O sistema de tanques acoplados da Quanser (*Coupled Tanks*) é uma planta experimental utilizada para o estudo e desenvolvimento de estratégias de controle de nível em processos industriais. O sistema é composto por dois tanques transparentes de acrílico, uma bomba e sensores responsáveis pela medição da pressão e do nível do fluido (Quanser, 2012). Sua arquitetura permite a implementação de diferentes configurações de controle, possibilitando o estudo de sistemas SISO, em cascata e MIMO.

No contexto deste trabalho, o sistema é utilizado como planta física para o desenvolvimento de um Gêmeo Digital. A partir da integração entre os dados obtidos pelos sensores e o modelo virtual do processo, busca-se representar o comportamento do sistema físico no ambiente computacional, possibilitando a simulação, o monitoramento e a análise de seu funcionamento.

2.4 Estratégias de controle

As estratégias de controle são utilizadas para regular o comportamento de sistemas dinâmicos, buscando manter determinadas variáveis próximas aos valores de referência. Entre as estratégias utilizadas neste trabalho estão o controle PID, o controle baseado em lógica fuzzy e o controle em cascata.

O controlador Proporcional-Integral-Derivativo (PID) é um dos algoritmos de controle em malha fechada mais utilizados na indústria, devido à sua estrutura relativamente simples e à possibilidade de ajustar seus parâmetros para atender aos requisitos de desempenho de diferentes processos (OGATA, 2010). Seu sinal de controle é determinado pela combinação de três ações que atuam sobre o erro, definido como a diferença entre o valor de referência e a variável medida. A ação proporcional (P) produz uma resposta proporcional ao erro instantâneo; a ação integral (I) acumula o erro ao longo do tempo, contribuindo para sua eliminação em regime permanente; e a ação derivativa (D) considera a taxa de variação do erro, podendo melhorar o amortecimento e reduzir oscilações, conforme a dinâmica do sistema e a sintonia adotada (NISE, 2012). Embora sua estrutura seja simples, o desempenho do controlador depende da escolha adequada dos parâmetros de sintonia e das características dinâmicas do processo controlado.

Diferentemente das abordagens de controle baseadas em modelos matemáticos explícitos da planta, a Lógica Fuzzy (ou lógica nebulosa) permite desenvolver estratégias de controle a partir de regras linguísticas, sem exigir necessariamente uma representação matemática precisa da dinâmica do sistema. Fundamentada na teoria dos conjuntos nebulosos, introduzida por Zadeh (1965), essa abordagem utiliza funções de pertinência para representar diferentes graus de associação de uma variável a conjuntos linguísticos. Um controlador fuzzy convencional compreende três etapas principais: a fuzzificação, que converte as entradas numéricas em representações nebulosas; a inferência, responsável por avaliar um conjunto de regras do tipo “Se–Então”; e a defuzzificação, que transforma a saída nebulosa em um sinal numérico de controle (PASSINO; YURKOVICH, 1998). Essa abordagem pode ser particularmente útil em sistemas com comportamento não linear, incertezas e dificuldade de modelagem precisa, embora seu desempenho e sua robustez dependam da estrutura das regras, das funções de pertinência e dos parâmetros adotados.

O controle em cascata é uma estratégia de controle em malha fechada caracterizada pela utilização de duas malhas de realimentação interligadas, nas quais um controlador primário define o valor de referência de um controlador secundário. A malha primária atua sobre a variável principal do processo, enquanto a malha secundária regula uma variável intermediária, permitindo que perturbações detectadas nessa malha sejam corrigidas antes de afetarem significativamente a variável principal. Para que essa estratégia apresente resultados satisfatórios, a dinâmica da malha secundária deve ser suficientemente mais rápida que a da malha primária, favorecendo uma resposta mais ágil às perturbações (SEBORG et al., 2016). Dessa forma, o controle em cascata pode melhorar o desempenho do sistema e a rejeição de perturbações, desde que as variáveis controladas sejam adequadamente selecionadas e os controladores sejam devidamente sintonizados.

3 Desenvolvimento e Implementação do Sistema

A primeira etapa deste projeto consistiu na realização de uma pesquisa bibliográfica exploratória e no mapeamento de tecnologias do mercado. Essa

fase teve como objetivo principal consolidar o embasamento teórico sobre Gêmeos Digitais, estratégias de controle e a arquitetura da Indústria 4.0. Além disso, foi conduzido um levantamento técnico de soluções comerciais voltadas para a indústria de processos, analisando produtos de empresas como Sensia, Sorba.ai, Vidyatec, Seeq, Visco e SLB (OptiSite). Essa análise permitiu compreender a estrutura funcional de sistemas reais, garantindo que as melhores práticas da indústria para a qualificação e validação de Gêmeos Digitais orientassem o desenvolvimento e a arquitetura do protótipo laboratorial.

3.1 MATLAB/Simulink

Para a etapa de modelagem, simulação e validação do sistema de controle, utilizou-se o ambiente computacional MATLAB integrado ao *Simulink*. Inicialmente, o modelo matemático das equações diferenciais que controlam a dinâmica do sistema foi implementado e parametrizado na plataforma. Utilizando a interface gráfica baseada em diagramas de blocos do *Simulink*, construiu-se a malha de controle em circuito fechado, adicionando os controladores. Esse ambiente virtual permitiu o desenvolvimento e a validação prévios das estratégias a serem aplicadas na planta.

A comunicação entre o modelo computacional e a planta do sistema de Tanques Acoplados da Quanser é realizada por meio de blocos HIL. Conforme apresentado na Figura 1, o projeto utiliza o bloco *HIL Read Analog* para realizar a leitura dos sinais provenientes dos sensores de pressão da planta física, que representam os níveis dos tanques. Já o bloco *HIL Write Analog* é utilizado para enviar os sinais de tensão responsáveis pelo acionamento da bomba de alimentação.

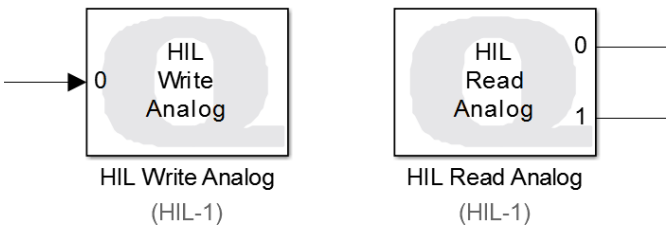


Figura 1: *Blocos HIL. Fonte: Autoria própria (2026).*

Utiliza-se também o bloco *Stream Client*, responsável por estabelecer a comunicação por meio do protocolo TCP/IP, possibilitando o monitoramento remoto e a integração com sistemas externos. Essa comunicação permite que os dados operacionais da planta, como os níveis dos tanques, os sinais de tensão e o estado do sistema físico, sejam transmitidos para a interface gráfica. Neste projeto, essa arquitetura de comunicação foi configurada com o objetivo de exportar os dados provenientes do MATLAB/*Simulink*, que serão utilizados, na próxima etapa da pesquisa, para alimentar a interface gráfica do Gêmeo Digital e possibilitar a representação do sistema em tempo real.

3.2 Sistema de Controle

Para lidar com a dinâmica interligada dos tanques e otimizar o desempenho global, optou-se pela implementação de uma malha de controle em cascata. Essa arquitetura é justificada pela natureza física do sistema *Coupled Tanks*, que apresenta dinâmicas distintas em cada estágio do processo de escoamento. O Tanque 1, que recebe o fluxo direto da bomba hidráulica, comporta-se de maneira aproximada a um sistema de primeira ordem. Por possuir uma resposta mais rápida e dinâmica menos complexa, a malha interna da cascata foi projetada utilizando um controlador clássico PID. O objetivo dessa malha interna é corrigir rapidamente os distúrbios locais e garantir que o nível do Tanque 1 atinja seu valor de referência de forma ágil.

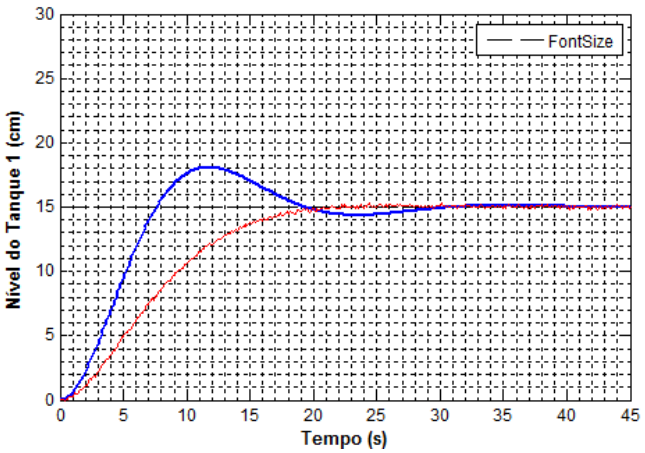


Figura 2: *Gráfico de comparação do controle PID e modelo matemático. Fonte: Autoria própria (2026).*

Como ilustrado na Figura 2, a avaliação do controle PID no Tanque 1 permite comparar o comportamento dinâmico da bancada real (linha vermelha) com o do Gêmeo Digital (linha azul). Embora o objetivo de controle seja alcançado em ambos os casos, observa-se uma diferença entre as respostas transitórias: o modelo virtual apresenta resposta mais rápida e sobressinal mais acentuado, enquanto a bancada real atinge o *setpoint* de forma mais lenta e com

sobressinal reduzido. Essa discrepância evidencia a necessidade de aprimoramentos na modelagem matemática, de modo a representar com maior fidelidade o comportamento dinâmico da planta física.

Por outro lado, o fluxo gravitacional oriundo do Tanque 1 atua como entrada do Tanque 2, conferindo a este segundo estágio um comportamento dinâmico de segunda ordem, mais lento e acentuado por não linearidades do escoamento. Para governar o nível do Tanque 2, a malha externa (mestre) da arquitetura em cascata foi projetada utilizando um controlador baseado em Lógica *Fuzzy*. O núcleo *Fuzzy* avalia o erro de rastreamento (diferença entre o nível desejado e o medido) em conjunto com a sua taxa de variação, extraída por um filtro derivativo. Através de regras de inferência linguísticas, o controlador gera uma ação suave e adaptativa.

A característica central dessa topologia em cascata é que a saída do controlador *Fuzzy* não atua diretamente na bomba. Em vez disso, ela estabelece dinamicamente o valor de referência exigido para a malha interna (o PID do Tanque 1). Adicionalmente, um mecanismo seletor (*Switch*) foi implementado para permitir a comutação entre a ação do mestre *Fuzzy* e uma malha alternativa direta, conferindo flexibilidade aos ensaios experimentais.

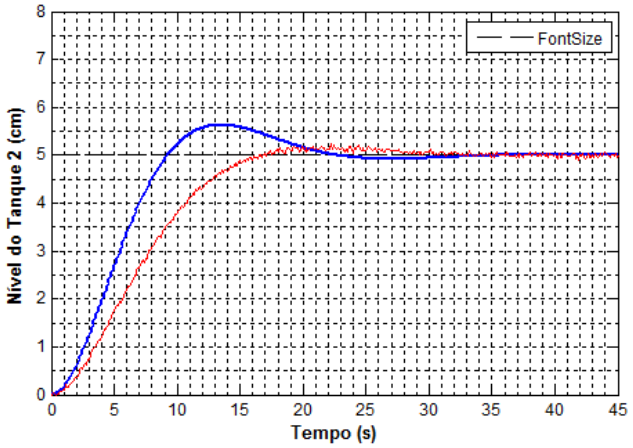


Figura 3: *Comparação entre o controlador Fuzzy e o modelo matemático. Fonte: Autoria própria (2026).*

Conforme representado na Figura 3, a ação do controlador *Fuzzy* demonstra eficácia na mitigação do efeito não linear do sistema acoplado. A abordagem em cascata permite que o PID absorva os distúrbios iniciais rapidamente, enquanto o *Fuzzy* garante um rastreamento de trajetória suave, com um tempo de acomodação de 21 segundos e estabilização robusta no segundo tanque. Apesar da estabilização do sistema, a comparação entre a bancada física e o modelo matemático também evidencia diferenças no comportamento transitório, indicando que a representação virtual ainda apresenta limitações na reprodução da dinâmica real do processo.

3.3 Sistema de Tanques Acoplados da Quanser

O sistema experimental utilizado neste trabalho baseia-se na bancada de tanques acoplados da Quanser, apresentada na Figura 4, (*Coupled Tanks*), uma plataforma amplamente empregada na validação de estratégias de controle avançado e na implementação de Gêmeos Digitais. A planta física é composta essencialmente pelos seguintes elementos:

- Bomba Hidráulica (Atuador): Responsável por bombear a água a partir de um reservatório inferior e fornecer a vazão de entrada diretamente para o sistema, operando através de um sinal de tensão elétrica condicionado.

- Tanque 1 (Malha Interna): Posicionado no estágio superior, recebe o fluxo direto da bomba. O escoamento deste tanque ocorre por gravidade através de um orifício calibrado em sua base, funcionando como afluente direto para o estágio inferior.

- Tanque 2 (Malha Externa): Posicionado abaixo do primeiro, recebe a água proveniente do escoamento do Tanque 1. O escoamento final ocorre também por gravidade em direção ao reservatório base, fechando o circuito hidráulico.

- Instrumentação e Sensores: Ambos os tanques estão equipados com sensores de nível que convertem a altura da coluna de água em sinais de tensão elétrica, permitindo a aquisição de dados em tempo real por meio da interface de hardware (*HIL*).

3.4 Interface Gráfica

Para consolidar o conceito de Gêmeo Digital e permitir a interação em tempo real com a planta física, desenvolveu-se um sistema supervisório baseado em tecnologias web, utilizando a plataforma Node-RED. A integração entre o modelo matemático processado no MATLAB/*Simulink* e a interface gráfica é viabilizada pelo protocolo de comunicação TCP/IP, através do bloco *Stream Client*, garantindo a transmissão contínua dos dados operacionais e de controle.

O ambiente do supervisório foi estruturado para fornecer tanto o monitoramento quanto o comando

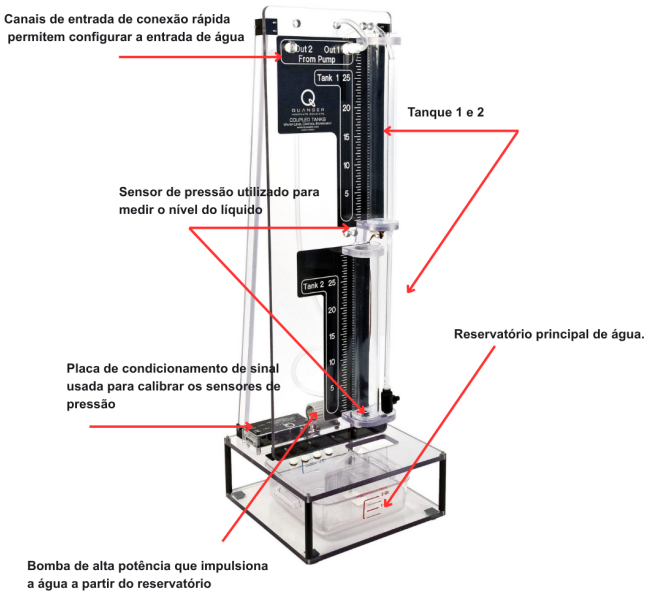


Figura 4: *Descrição do sistema de Tanques Acoplados da Quanser. Fonte: Autoria própria (2026).*

remoto do processo. Como visto na Figura 5, o Painel de Operação centraliza as ações do usuário, disponibilizando chaves de comutação para habilitar o controle via web e o modo de simulação. Por meio deste painel, o operador pode selecionar o tanque alvo da ação de controle e ajustar o valor de referência (*setpoint*) utilizando uma barra de rolagem deslizante que define a altura desejada em centímetros.

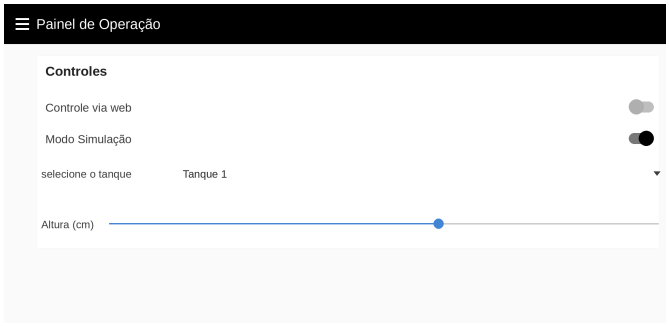


Figura 5: *Painel de controle virtual. Fonte: Autoria própria (2026).*

Além dos controles, a interface exibe, através de um gráfico, os valores numéricos instantâneos dos níveis do Tanque 1 e do Tanque 2, a tensão aplicada à bomba hidráulica e um indicador de status que confirma a normalidade do sistema, como apresentado na Figura 6.

Como elemento central do Gêmeo Digital, o painel conta também com uma representação virtual bidimensional do sistema físico, visto na Figura 7.



Figura 6: *Representação gráfica do sistema físico. Fonte: Autoria própria (2026).*

Essa interface ilustra graficamente os dois reservatórios em cascata. Animada dinamicamente pelos dados recebidos via TCP/IP, a representação mimetiza o comportamento da planta real, atualizando o volume da coluna de água virtual de forma sincronizada com as variações físicas. Essa topologia de supervisório não apenas facilita a análise visual do escoamento, mas também comprova a viabilidade da aplicação de tecnologias da Indústria 4.0 para o monitoramento e a teleoperação de processos industriais de controle de nível.

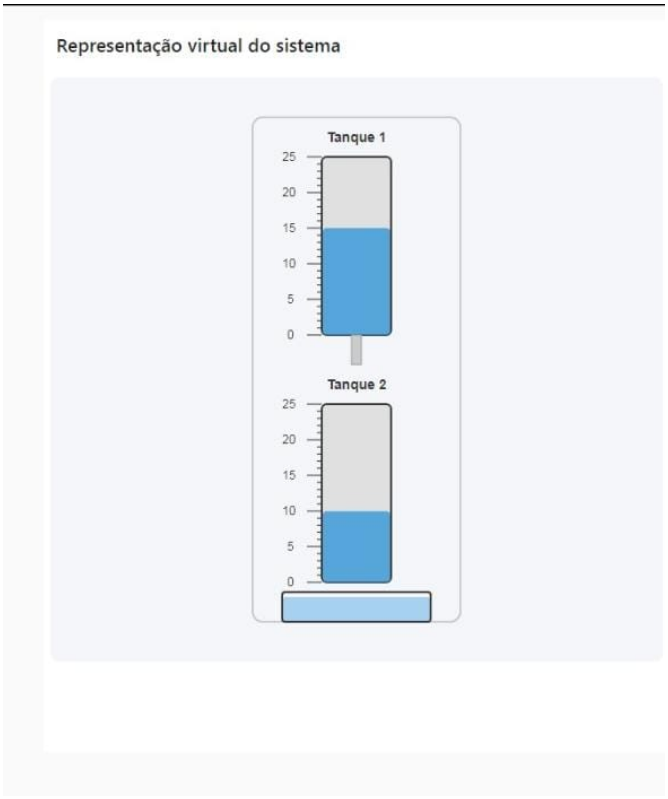


Figura 7: *Painel de Operação da interface gráfica do Gêmeo Digital. Fonte: Autoria própria (2026).*

Cabe ressaltar que a atual infraestrutura de supervisório compõe a etapa de validação da interface e comunicação do projeto. Como trabalhos futuros, a pesquisa dará continuidade à consolidação definitiva do Gêmeo Digital, visando a sua operação autônoma em ambiente *web* e em nuvem. O objetivo é embarcar os modelos matemáticos e algoritmos de controle diretamente na plataforma online, permitindo que o sistema inteligente funcione de forma independente, descentralizada e ininterrupta, sem a necessidade de execução local simultânea do MATLAB/*Simulink* para o monitoramento e atuação sobre a planta física.

4 Discussão da Proposta em Relação às Soluções Existentes

O levantamento de soluções comerciais permitiu identificar diferentes abordagens para a aplicação de Gêmeos Digitais em ambientes industriais. Entre elas, a Sensia integra tecnologias de automação, conectividade e análise de dados para apoiar a digitalização e a otimização de operações no setor de petróleo e gás (SENSIA, 2026). A SLB, por meio da solução OptiSite, combina Gêmeos Digitais e inteligência artificial para fornecer informações preditivas e apoiar a otimização de instalações, equipamentos e dutos (SLB, 2025). A SORBA.ai, por sua vez, disponibiliza uma plataforma de inteligência artificial industrial que integra ferramentas de análise de dados, aprendizado de máquina e Gêmeos Digitais. A solução permite desenvolver modelos para monitoramento de ativos, identificação de anomalias, manutenção preditiva e otimização de processos industriais, oferecendo recursos de simulação e controle avançado (SORBA.AI, 2026).Outras soluções apresentam enfoques distintos, como a Vidya, que utiliza representações digitais para o gerenciamento da integridade de ativos industriais (VIDYATEC, 2025); a Seeq, voltada à análise avançada de dados operacionais e à identificação de oportunidades de melhoria em processos industriais (SEEQ, 2026); e a Visco, que oferece um ambiente de visualização integrado a modelos tridimensionais, dados operacionais e ferramentas de análise de riscos (VISCO, 2026). Essas aplicações demonstram a diversidade de possibilidades associadas aos Gêmeos Digitais, que podem atender desde necessidades de monitoramento e análise de dados até a gestão da integridade e a otimização de processos.

Em relação a essas soluções, o trabalho desenvolvido apresenta uma abordagem experimental, direcionada à integração entre uma planta física de tanques acoplados, sua representação matemática e uma interface de supervisão remota. A arquitetura proposta combina o MATLAB/Simulink, utilizado na modelagem e implementação das estratégias de controle, com o Node-RED, responsável pela visualização dos dados e pela interação com o processo por meio de uma interface web. Além disso, a aplicação de controladores PID e Fuzzy em uma estrutura de controle em cascata permite explorar, em um mesmo ambiente, a integração entre modelagem, controle e monitoramento. Dessa forma, a contribuição do trabalho está no desenvolvimento de um protótipo experimental que permite investigar aspectos práticos da implementação de Gêmeos Digitais em sistemas de controle de processos, constituindo uma base para futuras ampliações, como a incorporação de recursos de análise de dados, diagnóstico de falhas e otimização operacional.

5 Conclusões

Este trabalho apresentou o desenvolvimento e a validação de um protótipo de Gêmeo Digital aplicado ao sistema experimental de Tanques Acoplados da Quanser. Fundamentado nos preceitos da Indústria 4.0, o projeto integrou modelagem matemática, estratégias de controle híbridas e a criação de uma interface de supervisão remota, consolidando uma arquitetura ciberfísica funcional.

A avaliação da malha de controle em cascata no ambiente MATLAB/*Simulink* evidenciou a eficácia da abordagem proposta. A alocação de um controlador PID na malha interna permitiu uma resposta rápida na mitigação de distúrbios e na estabilização do Tanque 1. Simultaneamente, o controlador baseado em Lógica *Fuzzy*, atuando na malha externa, demonstrou elevada capacidade adaptativa ao lidar com as dinâmicas não lineares e de segunda ordem do Tanque 2. A análise comparativa confirmou que o Gêmeo Digital mimetiza de forma satisfatória o comportamento da planta real.

Adicionalmente, a implementação do painel de operação via plataforma Node-RED cumpriu o objetivo de estabelecer uma comunicação de dados em tempo real via protocolo TCP/IP. Este sistema supervisório viabilizou a operação via *web* e o monitoramento visual do escoamento, servindo como a interface definitiva entre o operador e o sistema interligado.

Como perspectivas para trabalhos futuros, almeja-se a transição desta infraestrutura para um ambiente de computação em nuvem. O objetivo é permitir que os algoritmos de controle e o Gêmeo Digital operem

de forma autônoma e descentralizada na *web*, dispensando a execução local. Em suma, o projeto atingiu os seus objetivos iniciais, fornecendo uma base tecnológica robusta e reforçando a capacitação técnica, analítica e de desenvolvimento de sistemas essenciais no domínio da Engenharia de Computação.

Referências

BREQUE, M.; NUL, L. D.; PETRIDIS, A. *Industry 5.0: Towards a sustainable, human-centric and resilient European industry*. Luxembourg, 2021.

FRONTONI, E. et al. Industry 4.0: A bibliometric analysis and systematic review. *Journal of Manufacturing Systems*, 2018.

GLAESSGEN, E.; STARGEL, D. The digital twin paradigm for future nasa and us air force vehicles. In: *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*. [S.l.: s.n.], 2012. p. 1818.

GRIEVES, M. *Complex systems engineering theory*. 2002. Apresentação no Product Lifecycle Management (PLM) Center of Excellence, University of Michigan.

LIN, D. et al. Strategic response to industry 4.0: An empirical investigation on the chinese automotive industry. *International Journal of Production Economics*, v. 200, p. 589–605, 2018.

NISE, N. S. *Engenharia de Sistemas de Controle*. 6. ed. Rio de Janeiro: LTC, 2012.

OGATA, K. *Engenharia de Controle Moderno*. 5. ed. São Paulo: Pearson Prentice Hall, 2010.

PASSINO, K. M.; YURKOVICH, S. *Fuzzy Control*. Menlo Park, CA: Addison-Wesley, 1998.

PFEIFFER, S. The vision of "Industrie 4.0" in the making—a case of future told, tamed, and traded. *Nanoethics*, v. 11, n. 1, p. 107–121, 2017.

Quanser. *Coupled Tanks User Manual*. Markham, ON, Canada, 2012.

SCHLEICH, B. et al. Shaping the digital twin for design and production engineering. *CIRP Annals*, v. 66, n. 1, p. 141–144, 2017.

SEBORG, D. E. et al. *Process Dynamics and Control*. 4. ed. Hoboken, NJ: John Wiley & Sons, 2016.

SEEQ. *Seeq: Plataforma de Análise de Dados Industriais*. 2026. Disponível em: <https://www.seeq.com>. Acesso em: 6 nov. 2026.

SENSIA. *Automation and Digitalization*. 2026. Disponível em: <https://www.sensiaglobal.com>. Acesso em: 5 nov. 2026.

SLB. *SLB Launches Industry-First AI Solution for Entire Production Networks*. 2025. Disponível em: <https://www.slb.com>. Acesso em: 5 nov. 2026.

VIDYATEC. *Suíte Digital de Gestão de Integridade de Ativos*. 2025. Disponível em: <https://vidyatec.com>. Acesso em: 6 nov. 2026.

VISCO. *VCOG Digital Twin Solution*. 2026. Disponível em: <https://viscocg.com>. Acesso em: 5 nov. 2026.

ZADEH, L. A. Fuzzy sets. *Information and Control*, v. 8, n. 3, p. 338–353, 1965.

ZHONG, R. Y. et al. Intelligent manufacturing in the context of industry 4.0: A review. *Engineering*, v. 3, n. 5, p. 616–630, 2017.

Uma Revisão de Metodologias Inteligentes de Classificação de Curtos-circuitos em Sistemas Elétricos de Potência

Guilherme Nascimento da Silva [1]
Fabio Meneghetti Ugulino de Araujo [2]
Rodrigo Prado de Medeiros [3]

[1] *Programa de Pós-Graduação em Engenharia Elétrica e Computação, UFRN, <guilherme.gnds@gmail.com>*
[2] *Departamento de Engenharia de Computação e Automação, UFRN, <meneghet@dca.ufrn.br>*
[3] *Departamento de Engenharia Elétrica, UFRN, <rodrigo.prado@ufrn.br>*

Resumo

A confiabilidade dos sistemas elétricos de potência (SEP) é fortemente desafiada por curtos-circuitos, cujos impactos operacionais demandam esquemas de proteção céleres e seletivos. Historicamente, os critérios convencionais baseados em limiares determinísticos e restrição harmônica enfrentam limitações para discriminar faltas reais de transitórios operacionais benignos, como correntes de magnetização (*inrush*) e manobras de chaveamento. Diante disso, a integração entre processamento digital de sinais (PDS) e inteligência artificial (IA) consolidou-se como um paradigma indispensável para a superação desses limites. Este trabalho apresenta uma revisão bibliográfica sistemática sobre o avanço das metodologias inteligentes de detecção e classificação de curtos-circuitos na rede elétrica. Adotando o transformador de potência como ativo paradigmático de severidade transitória e estendendo a investigação a linhas de transmissão, bancos de capacitores, microrredes e dispositivos FACTS, analisam-se abordagens baseadas na transformada *wavelet*, filtragem de Fourier e componentes simétricas associadas a redes neurais artificiais, lógica *fuzzy* e aprendizado de máquina. A análise comparativa demonstra que modelos híbridos atingem expressivas taxas de acerto (superiores a 98%) e tempos de resposta subcíclicos (< 10 ms), evidenciando *trade-offs* fundamentais entre acurácia, latência, interpretabilidade e resiliência à perda de medições (*missing data*). Por fim, delineiam-se os desafios abertos para a integração desses algoritmos em dispositivos eletrônicos inteligentes (IEDs) comerciais.

Palavras-chave: Classificação de faltas; Processamento digital de sinais; Inteligência artificial; Proteção de sistemas elétricos; Transformadores de potência.

Abstract

The reliability of modern power systems is constantly challenged by short-circuit faults, whose impacts demand fast, sensitive, and selective protection schemes. Historically, conventional criteria based on deterministic thresholds and harmonic restraint have faced operational constraints in discriminating internal faults from benign transient events, such as magnetizing inrush currents and switching maneuvers. Consequently, the combination of digital signal processing (DSP) and artificial intelligence (AI) has emerged as an essential paradigm to overcome these limitations. This paper presents a systematic review of intelligent methodologies for fault detection and classification in electrical grids. Adopting the power transformer as the benchmark asset due to its high transient severity and expanding the investigation to transmission lines, shunt capacitor banks, microgrids, and FACTS devices, we review approaches based on wavelet transforms, Fourier filtering, and symmetrical sequence components combined with artificial neural networks, fuzzy logic, and machine le-

arning. The comparative analysis demonstrates that hybrid models achieve high accuracy rates (exceeding 98%) and sub-cycle operating times (< 10 ms), highlighting critical trade-offs regarding accuracy, latency, interpretability, and resilience under missing data conditions. Finally, open challenges and technical barriers for embedding these algorithms into commercial intelligent electronic devices (IEDs) are outlined.

Keywords: *Fault classification; Digital signal processing; Artificial intelligence; Power system protection; Power transformers.*

1 Introdução

O contínuo aprimoramento na confiabilidade dos sistemas elétricos de potência (SEP) está diretamente ligado aos avanços tecnológicos que sustentam a crescente demanda por uma alimentação de energia segura e estável. Para acompanhar a alta dependência da sociedade contemporânea pela energia elétrica, as concessionárias e órgãos responsáveis buscam otimizar o fornecimento, implementando formas cada vez mais sofisticadas de identificar e contornar intercorrências ao longo do sistema.

Nesse cenário, curtos-circuitos figuram entre as perturbações de maior impacto, podendo causar desde quedas de tensão até apagões sustentados que afetam milhares de consumidores. Quando essas faltas ocorrem em áreas com elevado fluxo de carga, os transtornos econômicos e operacionais são imensos (MAMEDE FILHO; MAMEDE, 2011). Além da perda do suprimento de energia, curtos-circuitos causam danos térmicos e mecânicos irreversíveis aos equipamentos que compõem o sistema (MARTINS, 2021).

Dentre os dispositivos essenciais presentes em todas as etapas da matriz energética, da geração à distribuição, os transformadores de potência exigem atenção primordial. Eles são responsáveis por adequar os níveis de tensão e corrente para o transporte e consumo de energia (KITANO, 2019). Dada a relevância estratégica e o elevado custo desses equipamentos, falhas não mitigadas rapidamente podem resultar em catástrofes operacionais. Estudos indicam que grande parte das faltas de energia ocorre em transformadores, e a maioria delas é oriunda de curtos-circuitos entre espiras (KITANO, 2019).

Para evitar a evolução de falhas internas e proteger esses ativos, as diretrizes da Agência Nacional de Energia Elétrica (ANEEL) recomendam a aplicação de sistemas de proteção intrínseca e principal (ANEEL, 2020). A proteção diferencial atua como a linha de frente da proteção principal em transformadores de médio e grande porte, sendo sua função principal a detecção célere de defeitos internos. Diante de sua importância, as abordagens para análise de curtos-circuitos nesses equipamentos têm evoluído consideravelmente ao longo dos anos.

Com o avanço acelerado da literatura voltada à aplicação de inteligência computacional e processamento de sinais no setor elétrico, surge a possibilidade de mapear os diferentes *trade-offs* práticos, tais como: tempo de resposta a eventos de distúrbio, sensibilidade a faltas incipientes e viabilidade de implementação em relés comerciais. Este trabalho propõe, portanto, uma revisão bibliográfica sistemática sobre a evolução das estratégias de proteção, com ênfase no avanço dos métodos de detecção e classificação de faltas em transformadores de potência, traçando uma linha evolutiva desde as abordagens analógicas e digitais convencionais até os modelos híbridos contemporâneos.

2 Metodologia da Pesquisa

A condução deste levantamento bibliográfico fundamentou-se em um protocolo estruturado de busca, triagem e síntese crítica, organizado em três fases de investigação:

- Fase 1 - Mapeamento dos Fundamentos Convencionais: Resgate histórico dos princípios canônicos de proteção diferencial, distância e sobrecorrente, analisando as contribuições seminais da literatura técnica e as limitações das formulações analógicas e digitais baseadas em restrição harmônica.

- Fase 2 - O Transformador de Potência como Ativo Paradigmático: Mapeamento da inserção de técnicas de processamento digital de sinais (PDS) e modelos de inteligência artificial (IA) no equipamento que historicamente concentra os transitórios eletromagnéticos mais severos do sistema, identificando quais arquiteturas de extração de atributos e quais paradigmas de inferência computacional foram desenvolvidos para superar os limites dos critérios determinísticos.

- Fase 3 - Expansão e Validação Cruzada no SEP: Ampliação do escopo investigativo para avaliar como as técnicas híbridas de PDS e inteligência computacional se comportam na discriminação e classificação de faltas em outros equipamentos e arranjos críticos da rede elétrica, incluindo linhas de transmissão, bancos de capa-

citores, microrredes e compensadores estáticos (STATCOM).

As buscas foram efetuadas nas bases indexadoras *IEEE Xplore*, *ScienceDirect* e *SpringerLink*, além de diretrizes normativas do setor elétrico nacional (ANEEL). Empregaram-se combinações booleanas estruturadas a partir de três eixos de descritores: (i) fenômenos e equipamentos (*"power transformer"*, *"transmission line"*, *"capacitor bank"*, *"inrush current"*, *"internal fault"*); (ii) processamento de sinais e extração de atributos (*"signal processing"*, *"time-frequency analysis"*, *"spectral features"*, *"wavelet"*); e (iii) inteligência computacional (*"artificial intelligence"*, *"machine learning"*, *"pattern recognition"*, *"fault classification"*).

Foram adotados como critérios de inclusão artigos revisados por pares que apresentassem modelagem transitória e métricas quantitativas de desempenho (acurácia, sensibilidade, tempo de resposta ou resiliência a perda de dados). Excluíram-se trabalhos focados exclusivamente em monitoramento de variáveis térmicas de regime permanente e propostas conceituais desprovidas de simulação ou ensaios experimentais.

3 Evolução dos Critérios Convencionais de Detecção e Discriminação de Faltas

Historicamente, os esquemas de proteção no SEP baseiam-se na monitoração de desvios nas grandezas de corrente e tensão para identificar anomalias operacionais. Seja por meio de funções de sobrecorrente, distância ou princípio diferencial, o desafio fundamental sempre consistiu em segregar faltas reais de fenômenos transitórios benignos, tais como energização de bancos de capacitores, chaveamentos em linhas de transmissão e, de modo especialmente crítico, a corrente de magnetização (*inrush*) em transformadores de potência. As primeiras implementações industriais, introduzidas na primeira metade do século XX por Cordray (1931), utilizavam relés eletromecânicos simples, cuja atuação era frequentemente comprometida por transitórios que mimetizavam correntes de curto-circuito interno.

Para contornar tais atuações indevidas, a engenharia de proteção desenvolveu métodos de restrição baseados em filtragem espectral, aproveitando a presença marcante de componentes harmônicas durante regimes de saturação e magnetização transitória (KENNEDY; HAYWARD, 1938; HAYWARD, 1941; SHARP; GLASSBURN, 1958). Paralelamente, esquemas de compensação de magnitude e defasagem angular foram consolidados (MCKENNA, 1950; EINVALL; LINDERS, 1975), permitindo que, na década de 1970, a transição para relés digitais viabilizasse a aplicação de filtros computacionais para cálculo fasorial e bloqueio por harmônicos (SYKES; MORRISON, 1972). Contudo, à medida que a topologia do SEP se tornou mais interconectada, incorporando linhas longas, compensações em série e níveis elevados de curto-circuito, as lógicas tradicionais com limiares fixos de restrição harmônica começaram a apresentar desarmes incorretos ou retardos indesejados (LIU et al., 1992).

No início do século XXI, a necessidade de aprimorar a sensibilidade para faltas de difícil detecção (a exemplo de curtos-circuitos envolvendo poucas espiras em enrolamentos ou faltas de alta impedância em alimentadores) motivou o surgimento de novos critérios analíticos. Destacam-se o reconhecimento dinâmico de formas de onda (GUZMAN et al., 2002) e a aplicação de grandezas de sequência negativa para diagnóstico de assimetrias incipientes (GAJIĆ et al., 2005; BABIY; GOKARAJU; GARCIA, 2011). Não obstante a eficácia dessas abordagens, o comportamento não linear dos equipamentos do SEP e a dependência de ajustes paramétricos rígidos evidenciaram as limitações dos critérios determinísticos convencionais, impulsionando a pesquisa em direção a métodos de aprendizado de máquina e processamento tempo-frequência.

4 Inteligência Artificial e Processamento de Sinais na Proteção

Embora as abordagens baseadas em lógica diferencial digital e componentes de sequência negativa apresentem avanços significativos, a busca por uma proteção imune a perturbações transitórias e com capacidade de detecção de faltas incipientes impulsionou o emprego de técnicas de inteligência artificial (IA) no setor elétrico (ANEEL, 2025). A flexibilidade dos modelos inteligentes possibilita lidar com o comportamento altamente não linear do núcleo magnético e com a diversidade de regimes operacionais sem depender de modelos analíticos rígidos.

Dentre os paradigmas de aprendizado e inferência aplicados ao diagnóstico de faltas no SEP, destacam-se:

- **Redes neurais artificiais (RNA):** Modelos conexionistas cujo aprendizado por ajuste de pesos sinápticos, consolidado a partir do al-

goritmo de retropropagação do erro (RUMELHART; HINTON; WILLIAMS, 1986), confere elevada capacidade de reconhecimento de padrões e generalização frente a assinaturas de faltas complexas;

- **Sistemas baseados em lógica *fuzzy* (LF):** Estruturas de inferência fundamentadas em variáveis linguísticas e graus de pertinência (TAKAGI; SUGENO, 1985; ELZALIK et al., 2025), adequadas ao tratamento de incertezas operacionais em sistemas sem um modelo linear equivalente, saturações parciais e limites de transição suaves;
- **Métodos de aprendizado de máquina supervisionado (AdM):** Classificadores estatísticos e baseados em margem, como máquinas de vetores de suporte e abordagens correlatas, amplamente investigados na segregação de distúrbios operacionais (GUPTA; SUHAVAN; VUPPULURI, 2025).

Apesar do potencial desses algoritmos, a aplicação direta de modelos inteligentes sobre amostras brutas de corrente frequentemente resulta em perdas de generalização e baixa robustez frente a ruídos. Essa limitação é corroborada pelo estudo comparativo de Gupta, Suhavan e Vuppuluri (2025), no qual, ao se avaliarem onze métodos de aprendizado de máquina aplicados a registros de faltas sem tratamento prévio, apenas um algoritmo superou 90% de acurácia. Conforme sintetizado por Panigrahi et al. (2025), o processamento digital de sinais (PDS) estabeleceu-se como uma etapa indispensável de pré-processamento e extração de características no domínio tempo-frequência.

Em particular, a transformada *wavelet* discreta (TWD) e variações como pacotes *wavelet* (WP) e transformada *wavelet* discreta redundante (TWDR) tornaram-se referências na literatura, em virtude de sua sensibilidade na detecção de descontinuidades e transitórios rápidos de alta frequência característicos da corrente *inrush* e do início dos curtos-circuitos (QIU, 2022; DANTAS et al., 2026). De forma complementar, ferramentas como a transformada de Fourier discreta (TFD) para filtragem harmônica (MOHANTY; MOHAPATRA; RAMANA, 2024) e a análise por componentes de sequência simétrica (CS) (KATYARA et al., 2018) também são empregadas na composição de vetores de atributos robustos.

No domínio específico dos transformadores de potência, as primeiras abordagens baseadas em lógica *fuzzy* buscaram modelar as incertezas nas fronteiras operacionais da proteção diferencial (WISZNIEWSKI; KASZTENNY, 1993). Posteriormente, a fusão entre a análise multirresolução da transformada *wavelet* e a inferência *fuzzy* demonstrou elevada eficiência na discriminação entre energização e curto-circuito interno (ELDIN, 2003; ANANWATTANAPORN et al., 2017).

De forma simultânea, a combinação de TWD com redes neurais e outros modelos de aprendizado de máquina alcançou patamares expressivos de acurácia na proteção desse ativo. Em simulações no *software* ATP, Fernandes, Costa e Medeiros (2016) obtiveram 100% de eficácia na distinção entre faltas internas e externas e 99,14% na correta classificação do tipo de defeito entre espiras, corroborando resultados pioneiros como os de Khorashadi-Zadeh e Sanaye-Pasand (2004). Avanços mais recentes exploram a resiliência dos classificadores inteligentes sob condições de contingência operacional severa: no estudo de Dantas et al. (2026), o emprego de TWDR e diferentes técnicas de aprendizado de máquina atingiu até 95% de acurácia mesmo sob perda temporária de canais de dados de medição (*missing data*), alinhando-se às demandas da Indústria 4.0. Em outra frente complementar, Alkhardawi et al. (2025) demonstraram que a integração de RNAs a parâmetros físico-químicos de análise de gases dissolvidos (DGA) atinge 93,3% de acurácia na identificação de anomalias estruturais, evidenciando o potencial da fusão sensorial no diagnóstico do equipamento.

5 Análise Comparativa das Metodologias

A necessidade de discriminar fenômenos transitórios e classificar faltas com rapidez e confiabilidade não se restringe aos transformadores, estendendo-se por todo o SEP. Conforme sintetizado na Tabela 1, o acoplamento entre ferramentas de PDS e inteligência computacional tem sido amplamente investigado em diversos equipamentos e topologias de rede.

No âmbito das linhas de transmissão, Katyara et al. (2018) associaram lógica *fuzzy* digitalizada à teoria das componentes simétricas em sistemas de 400 kV, reportando tempos de atuação inferiores a 10 ms para todas as classes de falta. Para acomodar a alta dinamicidade e a variabilidade de curto-circuito em microrredes AC, Jayasinghe et al. (2024) implementaram um classificador baseado em TWD e RNA com taxa de acerto de 99,39%. Em aplicações voltadas a parques eólicos *offshore* com compensação por STATCOM, Mohanty, Mohapatra e Ramana

Tabela 1: *Comparativo de metodologias de classificação de faltas no SEP.*

Referência	Metodologia	Equipamento Alvo	Desempenho / Métrica de Destaque
Qiu (2022)	WP + LF	Linha de Transmissão	Classificação de 10 tipos de faltas sob variações de R_f
Katyara et al. (2018)	CS + LF	Linha de Transmissão	Tempo de resposta < 10 ms para 400 kV
Gupta, Suhavan e Vuppuluri (2025)	AdM	Linha de Transmissão	Apenas 1 de 11 modelos $> 90\%$ de acurácia sem PDS prévio
Jayasinghe et al. (2024)	TWD + RNA	Microrrede	Acurácia de 99,39% em rede dinâmica
Mohanty, Mohapatra e Ramana (2024)	TFD + LF	STATCOM	99,99% de acerto global na discriminação
Songsukthawan et al. (2024)	TWD + LF	Banco de Capacitores	Acurácia entre 70% e 100% em transitórios
Patcharoen e Ngaopitakkul (2020)	TWD + RNP	Banco de Capacitores	Acurácia de 98% (falta vs. *inrush*)
Patcharoen et al. (2024)	TWD + LF	Banco de Capacitores	Precisão de 88% ($> 90\%$ em *back-to-back*)
Fernandes, Costa e Medeiros (2016)	TWD + RNA	Transf. de Potência	100% (interna/externa) e 99,14% (tipo de falta)
Alkhardawi et al. (2025)	DGA + RNA	Transf. de Potência	Acurácia de 93,3% na detecção com DGA
Dantas et al. (2026)	TWDR + AdM	Transf. de Potência	Acurácia de até 95% sob *missing data*

Legenda: CS: componentes de sequência; **LF:** lógica *fuzzy*; **RNP:** rede neural probabilística. As demais siglas de PDS e inteligência computacional estão detalhadas na Seção 4.

(2024) obtiveram 99,99% de sucesso na discriminação de eventos via TFD aliada a sistemas de inferência *fuzzy*, ressalvando, contudo, a maior sobrecarga computacional demandada pela decomposição espectral frente a modelos *wavelet*.

Outro segmento fortemente contemplado por essas metodologias são os bancos de capacitores de alta tensão, componentes submetidos a severos regimes de manobra. Songsukthawan et al. (2024) e Patcharoen e Ngaopitakkul (2020) demonstraram que arquiteturas híbridas baseadas em TWD combinadas a sistemas nebulosos e redes neurais probabilísticas garantem índices de acurácia que alcançaram até 100% na segregação entre correntes de falta e transitórias de energização. Para subestações de 230 kV, Patcharoen et al. (2024) reportaram precisão superior a 90% na classificação específica de transitórios em chaveamentos do tipo *back-to-back*, atestando a robustez desses esquemas frente a distúrbios de chaveamento de alta severidade.

A análise comparativa dos estudos compilados permite identificar importantes *trade-offs* de projeto entre as diferentes abordagens:

1. Acurácia e Generalização vs. Dependência de Dados: Modelos baseados em RNAs e classificadores de aprendizado de máquina supervisionado proporcionam elevadas taxas de acerto, sendo eficazes na separação de curtos entre espiras incipientes e regimes de sobreexcitação. Contudo, seu desempenho depende criticamente de uma base abrangente de casos de treinamento, tornando-os sensíveis a desvios operacionais não contemplados na calibração do modelo.

2. Interpretabilidade e Rastreabilidade Operacional: Sistemas fundamentados em lógica *fuzzy* oferecem transparência no processo de decisão por meio de regras compreensíveis e auditáveis por engenheiros de proteção, atenuando a resistência prática quanto à natureza de "caixa-preta" inerente às redes neurais profundas. Não obstante, o delineamento das funções de pertinência e o crescimento exponencial da base de regras em sistemas com múltiplos atributos configuram desafios de parametrização.

3. Complexidade Computacional e Operação em Tempo Real: A viabilidade de integração desses algoritmos em Dispositivos Eletrônicos Inteligentes (IEDs) comerciais é condicionada pela latência conjunta do pré-processamento de sinais e da inferência computacional. Enquanto a transformada *wavelet* entrega excelente resolução temporal para janelas de amostragem reduzidas, a sobrecarga de processamento associada à decomposição em múltiplos níveis deve ser compatibilizada com os rigorosos limites de tempo de atuação da proteção diferencial.

Assim, a literatura técnica converge para o con-

senso de que a escolha da metodologia ótima está vinculada aos requisitos de sensibilidade, capacidade de processamento embarcado e disponibilidade de registros históricos de faltas, sobressaindo-se as soluções híbridas que unem a extração tempo-frequência à inteligência computacional.

6 Conclusões

Este trabalho apresentou uma revisão bibliográfica sistemática sobre o avanço das metodologias inteligentes de detecção e classificação de curtos-circuitos no SEP. A análise histórica e comparativa consolidou que a superação das limitações dos critérios convencionais depende essencialmente da sinergia entre processamento digital de sinais e inteligência computacional. O emprego de ferramentas como a transformada *wavelet*, filtragem fasorial e componentes simétricas provou-se indispensável para viabilizar taxas de assertividade superiores a 98% e tempos de resposta subcíclicos (< 10 ms), assegurando a correta discriminação entre faltas reais e perturbações operacionais severas em transformadores de potência, linhas de transmissão, bancos de capacitores e microrredes.

A compilação dos estudos evidenciou que a escolha da arquitetura ótima decorre do balanceamento prático entre interpretabilidade operacional (favorecida pela lógica *fuzzy*), capacidade de generalização (destacada em RNAs e aprendizado supervisionado) e resiliência a contingências de medição (*missing data*). Como diretrizes para pesquisas futuras, sobressai-se a necessidade de transpor as abordagens simuladas para ensaios em tempo real via *hardware-in-the-loop* (HIL), avaliando a latência de processamento em IEDs comerciais, além da incorporação de inteligência artificial explicável (XAI) e da integração com a norma IEC 61850 em subestações digitais.

Referências

ALKHARDAWI, W. et al. Classification of power transformer faults using ann. *2025 7th Asia Energy and Electrical Engineering Symposium (AEEES)*, mar. 2025.

ANANWATTANAPORN, S. et al. Internal fault classification algorithm in power transformer based on discrete wavelet transform and fuzzy logic. *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, jul. 2017.

ANEEL. *Resolução Normativa nº 903, de 8 de dezembro de 2020: Aprova a reestruturação e a revisão dos Procedimentos de Rede e estabelece procedimentos e critérios para alterações*. Brasília, DF: Agência Nacional de Energia Elétrica, 2020. Disponível em: <https://www2.aneel.gov.br/cedoc/ren2020903.pdf>.

ANEEL. *Bibliografia Temática: Inteligência Artificial (IA) no Setor Elétrico*. Brasília, DF: Agência Nacional de Energia Elétrica, 2025. 30 p. p. (Bibliografia Temática, 3). Disponível em: <https://biblioteca.aneel.gov.br/Busca/Download?codigoArquivo=180406&tipoMidia=0>.

BABIY, M.; GOKARAJU, R.; GARCIA, J. C. Turn-to-turn fault detection in transformers using negative sequence currents. *2011 IEEE Electrical Power and Energy Conference*, out. 2011.

CORDRAY, R. E. Percentage-differential transformer protection. *Electrical Engineering*, v. 50, p. 361–363, maio 1931.

DANTAS, I. T. A. et al. Power transformer fault classification using industry 4.0-compliant machine learning under data missing conditions. *IEEE Access*, v. 14, p. 98741 – 98751, jun. 2026.

EINVALL, C. H.; LINDERS, J. R. A three-phase differential relay for transformer protection. *IEEE Transactions on Power Apparatus and Systems*, v. 94, p. 1971–1980, nov. 1975.

ELDIN, E. T. A new algorithm for the classification of different transient phenomena in power transformers combining wavelet transforms and fuzzy logic. *2003 46th Midwest Symposium on Circuits and Systems*, dez. 2003.

ELZALIK, M. et al. Enhancing frequency response using artificial intelligence techniques in the power system with renewable energy resources. *Sustainable Futures*, v. 10, p. 100848, 2025.

FERNANDES, J. F.; COSTA, F. B.; MEDEIROS, R. P. de. Power transformer disturbance classification based on the wavelet transform and artificial neural networks. *2016 International Joint Conference on Neural Networks (IJCNN)*, nov. 2016.

GAJIĆ, Z. et al. Sensitive turn-to-turn fault protection for power transformers. *Joint Meetings of Cigré Study Committee B5 and IEEE Power System Relaying Committee, with the Study Committee B5 Colloquium*, v. 1, set. 2005.

GUPTA, R.; SUHAVAN, D.; VUPPULURI, P. P. Effective machine learning models for fault detection and classification in power transmission lines. *2025 7th International Conference on Energy, Power and Environment (ICEPE)*, maio 2025.

GUZMAN, A. et al. A current-based solution for transformer differential protection. ii. relay description and evaluation. *IEEE Transactions on Power Delivery*, v. 13, p. 886–893, out. 2002.

HAYWARD, C. D. Harmonic-current-restrained relays for transformer differential protection. *Transactions of the American Institute of Electrical Engineers*, v. 60, p. 377–382, jun. 1941.

JAYASINGHE, J. A. R. R. et al. Classification and localization of faults in ac microgrids through discrete wavelet transform and artificial neural networks. *IEEE Open Access Journal of Power and Energy*, v. 11, p. 303–313, out. 2024.

KATYARA, S. et al. Adaptive fault classification approach using digitized fuzzy logic (dfl) based on sequence components. *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, jun. 2018.

KENNEDY, L. F.; HAYWARD, C. D. Harmonic-current-restrained relays for differential protection. *Transactions of the American Institute of Electrical Engineers*, v. 57, p. 262–271, maio 1938.

KHORASHADI-ZADEH, H.; SANAYE-PASAND, M. Power transformer differential protection scheme based on wavelet transform and artificial neural network algorithms. *39th International Universities Power Engineering Conference, 2004. UPEC 2004.*, set. 2004.

KITANO, F. T. *Avaliação das Práticas Adotadas na Proteção De Transformadores de Potência Conectados ao Sistema Interligado Nacional*. Dissertação (Mestrado) — Departamento de Engenharia Elétrica, UNB, Brasília, DF, dez. 2019.

LIU, P. et al. Improved operation of differential protection of power transformers for internal faults. *IEEE Transactions on Power Delivery*, v. 7, p. 1912–1919, out. 1992.

MAMEDE FILHO, J.; MAMEDE, D. R. *Proteção de Sistemas Elétricos de Potência*. Rio de Janeiro, RJ: LTC Editora, 2011.

MARTINS, D. de B. *Análise de Curto-Circuito Probabilístico em Sistemas de Distribuição de Energia Elétrica com Geração Distribuída*. Dissertação (Mestrado) — Escola de Engenharia de São Carlos, USP, São Carlos, SP, fev. 2021.

MCKENNA, G. W. Theory and application of transformer differential protection. *Transactions of the American Institute of Electrical Engineers*, v. 69, p. 1197–1202, jan. 1950.

MOHANTY, S. K.; MOHAPATRA, S.; RAMANA, M. Protection scheme for statcom compensated line using fuzzy inference system based approach. *2024 IEEE 4th International Conference on Sustainable Energy and Future Electric Transportation (SEFET)*, jul. 2024.

PANIGRAHI, B. K. et al. An extensive review of techniques for transmission line fault classification. *2025 International Conference in Advances in Power, Signal, and Information Technology (APSIT)*, maio 2025.

PATCHAROEN, T.; NGAOPITAKKUL, A. Combination of wavelet transforms and artificial intelligence in hv shunt capacitor banks: Fault and inrush currents classification. *15th International Conference on Developments in Power System Protection (DPSP 2020)*, mar. 2020.

PATCHAROEN, T. et al. Classification of capacitor bank switching using fuzzy interference systems in 230 kv substation. *2024 16th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, jul. 2024.

QIU, G. Application of wavelet packet and fuzzy algorithm in power system short circuit fault classification. *Mathematical Problems in Engineering*, v. 2022, p. 1–12, jun. 2022.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, v. 323, p. 533–536, out. 1986.

SHARP, R. L.; GLASSBURN, W. E. A transformer differential relay with second-harmonic restraint. *Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems*, v. 60, p. 913–918, abr. 1958.

SONGSUKTHAWAN, P. et al. Combining fuzzy logic and discrete wavelet transform for accurate fault and inrush current classification in high voltage capacitor banks. *2024 11th International*

Conference on Power and Energy Systems Engineering (CPESE), set. 2024.

SYKES, J. A.; MORRISON, I. F. A proposed method of harmonic restraint differential protecting of transformers by digital computer. *IEEE Transactions on Power Apparatus and Systems*, PAS-91, p. 1266–1272, maio 1972.

TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-15, p. 116–132, fev. 1985.

WISZNIEWSKI, A.; KASZTENNY, B. Fuzzy set approach to transformer differential relay. *1993 Fifth International Conference on Developments in Power System Protection*, mar. 1993.

Interface Baseada em Agentes de IA para Exploração e Controle de Qualidade de Dados Sísmicos

***José A. A. Oliveira*[1]**
***Luiz F. S. Silva*[1]**
***Ernane Ferreira Rocha Junior*[1]**
***Victor Gomes*[1]**
***Ignacio Sánchez-Gendriz*[2]**
***Ivanovitch Silva*[1]**
***Luiz Affonso Guedes*[1]**
***Tiago Barros*[1]**

[1] Departamento de Engenharia de Computação e Automação (DCA), Universidade Federal do Rio Grande do Norte (UFRN), Natal–RN, Brasil. <augusto.oliveira.099@ufrn.edu.br>
[2] Centro de Tecnologias Estratégicas do Nordeste (CETENE/MCTI), Recife–PE, Brasil.

Resumo

A inspeção inicial de dados sísmicos envolve formatos especializados, metadados de aquisição e operações de processamento que, em geral, dependem de software específico ou programação. Este trabalho apresenta uma interface baseada em agentes de inteligência artificial para interação em linguagem natural com dados sísmicos no formato SEG-Y. A arquitetura emprega dois agentes com responsabilidades complementares: um Assistente Geral, responsável pela interpretação da solicitação e pela coordenação do fluxo, e um Agente de Visualização, especializado na geração de representações gráficas. A leitura, preparação e análise dos dados permanecem implementadas em ferramentas determinísticas de domínio, separando o raciocínio em linguagem natural do processamento sísmico. A avaliação considera seis tarefas representativas, dois modelos de linguagem, três formulações semanticamente equivalentes por tarefa e cinco repetições por formulação. Os resultados mostram que a interface executa operações de inspeção de cabeçalhos, análise de fold, seleção e visualização de traços e transformações exploratórias, com diferenças entre os modelos quanto à latência e à produção da saída esperada na primeira interação. Esses resultados indicam o potencial de interfaces agênticas para tornar a exploração inicial e o controle de qualidade de dados sísmicos mais acessíveis e interativos.

Abstract

Initial seismic-data inspection involves specialized file formats, acquisition metadata, and processing operations that typically depend on dedicated software or programming. This paper presents an artificial-intelligence agent-based interface for natural-language interaction with SEG-Y seismic data. The architecture employs two agents with complementary responsibilities: a General Assistant, responsible for interpreting requests and coordinating the workflow, and a Visualization Agent, specialized in generating graphical representations. Data reading, preparation, and analysis remain implemented as deterministic domain tools, separating natural-language reasoning from seismic processing. The evaluation considers six representative tasks, two language models, three semantically equivalent prompt formulations per task, and five repetitions per formulation. Results show

that the interface can perform header inspection, fold analysis, trace selection and visualization, and exploratory transformations, while the evaluated models differ in latency and in their ability to produce the expected output in the first interaction. These findings indicate the potential of agentic interfaces to make early-stage seismic exploration and quality control more accessible and interactive.

1 Introdução

Dados sísmicos são fundamentais para estudos geofísicos e para atividades de exploração do subsolo, mas sua análise inicial envolve arquivos de grande volume, estruturas multidimensionais e metadados de aquisição que precisam ser interpretados em conjunto com os sinais registrados (KHOSRO ANJOM; VACCARINO; SOCCO, 2024; DAVIS; HUNT, 2024). O formato SEG-Y é amplamente empregado para intercâmbio desses dados e armazena, além dos traços sísmicos, cabeçalhos textuais, binários e de traço contendo informações relevantes para a geometria da aquisição, organização dos registros e procedimentos de controle de qualidade (SEG TECHNICAL STANDARDS COMMITTEE, 2002).

Antes de etapas avançadas de processamento e interpretação, é comum realizar tarefas como inspeção de cabeçalhos, verificação de cobertura, seleção de traços, análise de fold e visualização de seções. Essas operações normalmente dependem de software especializado ou de rotinas de programação. Nesse contexto, modelos de linguagem de grande porte (LLMs) e agentes capazes de acionar ferramentas externas permitem utilizar linguagem natural como camada de interação com sistemas computacionais (WANG et al., 2024; CHENG et al., 2024). O LLM pode interpretar a intenção do usuário e coordenar a execução, enquanto operações numéricas e acesso aos dados permanecem sob responsabilidade de funções especializadas e auditáveis.

O uso de LLMs como interface para dados científicos vem sendo explorado em diferentes domínios. Zhang et al. (2025) apresentam um fluxo orientado por LLM para simplificar etapas de análise de dados em geociências. Em aplicações de engenharia, arquiteturas com múltiplos atores têm sido empregadas para integrar análise e visualização de sistemas de controle (XAVIER et al., 2024), enquanto estratégias de orquestração com múltiplos LLMs também foram investigadas para simulação ultrassônica baseada em texto (KIM; YU; SEO, 2025). Esses trabalhos reforçam uma tendência em direção a interfaces que combinam interpretação em linguagem natural com ferramentas computacionais especializadas. Para dados sísmicos, entretanto, essa integração requer atenção particular às estruturas SEG-Y, aos metadados de aquisição e à necessidade de preservar a execução determinística das operações de domínio.

Neste trabalho, propõe-se uma interface baseada em agentes de Inteligência Artificial (IA) para exploração inicial e controle de qualidade de dados sísmicos. A contribuição central está na integração de três elementos: interação em linguagem natural, coordenação entre agentes com responsabilidades distintas e um conjunto de ferramentas determinísticas para leitura, preparação, visualização e transformação dos dados. A arquitetura é avaliada em seis tarefas representativas, permitindo analisar tanto a execução funcional quanto aspectos de responsividade da interação. O objetivo não é substituir sistemas completos de processamento sísmico, mas oferecer uma camada de interação capaz de reduzir a necessidade de programação durante as etapas iniciais de inspeção e exploração dos dados.

2 Metodologia

A metodologia é organizada em três componentes: preparação dos dados sísmicos, arquitetura baseada em agentes e protocolo experimental. A Figura 1 apresenta o fluxo funcional da interface e explicita a separação entre interpretação em linguagem natural, coordenação dos agentes e execução determinística das operações sísmicas.

2.1 Dados sísmicos e preparação

A implementação foi desenvolvida para arquivos sísmicos marinhos no formato SEG-Y fornecidos pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP). Para a avaliação quantitativa apresentada neste artigo foi utilizado um conjunto pré-empilhado de aproximadamente 50 MB, representativo das operações de inspeção de metadados, seleção de traços e visualização consideradas.

Na etapa inicial, os cabeçalhos binários e de traço são extraídos e organizados em uma representação tabular. As bibliotecas Segyio e Segy-SAK são utilizadas para leitura das estruturas SEG-Y e acesso aos metadados (EQUINOR, 2026; HALLAM, 2025). Essa organização permite realizar sucessivas operações de inspeção, filtragem e ordenação sem exigir que todo o conteúdo de amostras dos traços seja mantido continuamente em memória. Quando uma solicitação demanda informação do sinal, os traços correspondentes são recuperados e encaminhados às ferramentas específicas.

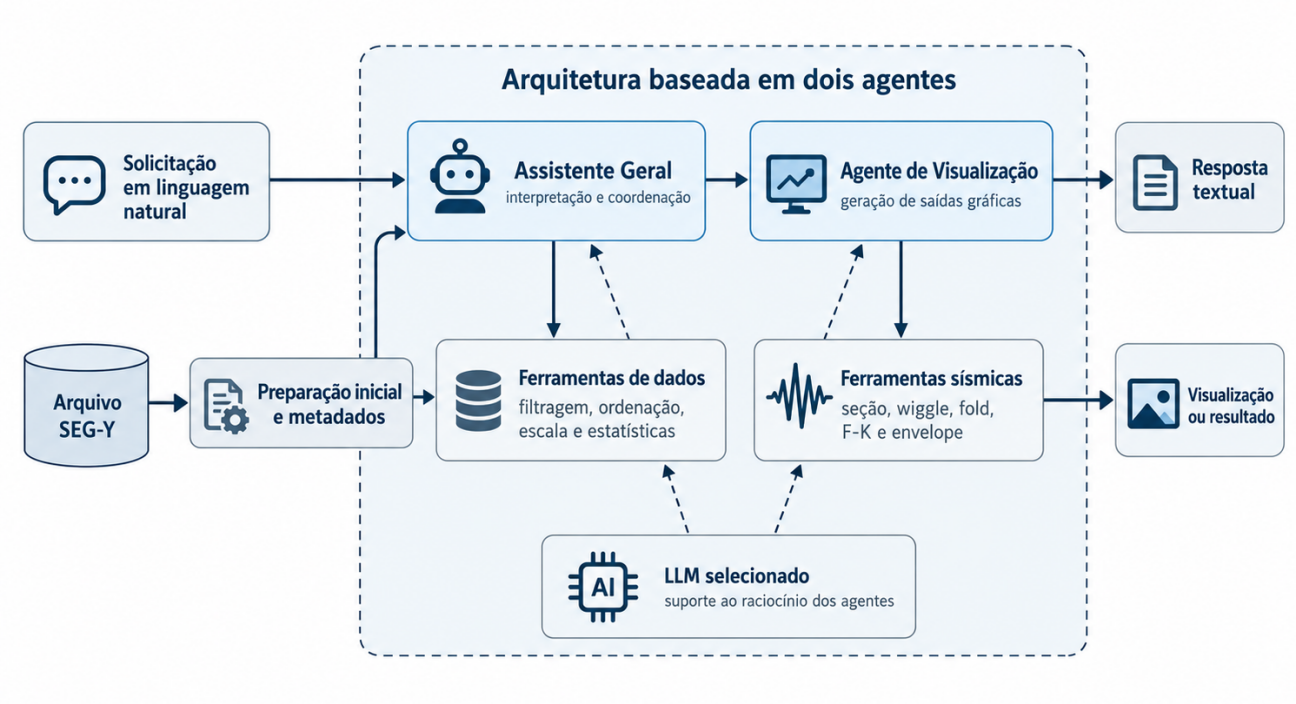


Figura 1: *Arquitetura da interface baseada em dois agentes. A solicitação em linguagem natural é interpretada pelo Assistente Geral, que coordena a interação, aciona ferramentas de preparação de dados quando necessário e pode delegar tarefas gráficas ao Agente de Visualização. Em paralelo, o arquivo SEG-Y é processado para disponibilizar metadados e estruturas necessárias às operações de domínio. As ferramentas de dados realizam filtragem, ordenação, escala e estatísticas, enquanto as ferramentas sísmicas executam visualizações e transformações específicas. O LLM fornece suporte ao raciocínio dos agentes, mas o acesso aos dados e os cálculos permanecem implementados em funções determinísticas. O fluxo pode produzir respostas textuais, visualizações ou outros resultados derivados das ferramentas.*

Essa separação entre metadados e amostras é relevante para a interação com o sistema. Solicitações simples, como consulta de cabeçalhos ou análise de cobertura, podem ser resolvidas a partir da representação tabular, enquanto operações como seção sísmica, gráfico *wiggle*, F–K ou envelope requerem acesso aos sinais. Assim, o fluxo de dados é condicionado pela tarefa solicitada, evitando carregar informação de maior volume quando ela não é necessária.

2.2 Arquitetura baseada em agentes

A arquitetura utiliza dois agentes com responsabilidades complementares, implementados a partir de componentes do LangChain e do LangGraph (CHASE, 2026; LANGCHAIN INC., 2026). O LangChain fornece as abstrações utilizadas para integrar o modelo de linguagem aos agentes e às funções externas disponíveis para execução. Essas funções, usualmente denominadas *tools* no contexto de sistemas agênticos, correspondem a rotinas de software que podem ser selecionadas e invocadas pelos agentes durante a execução de uma solicitação. O LangGraph, por sua vez, organiza o fluxo de interação como um grafo com estado, permitindo controlar transições entre etapas, preservar o contexto da interação e implementar decisões condicionais, como responder diretamente, executar uma ferramenta ou delegar a tarefa a outro agente.

Como mostrado na Figura 1, o *Assistente Geral* constitui o primeiro ponto de decisão após a entrada da solicitação em linguagem natural. O agente utiliza o LLM selecionado para interpretar a intenção do usuário em conjunto com o contexto associado ao arquivo SEG-Y carregado. A partir dessa interpretação, pode produzir uma resposta textual diretamente, selecionar uma *tool* de preparação ou análise de dados, ou encaminhar a solicitação ao *Agente de Visualização* quando a tarefa requer uma saída gráfica.

As *tools* associadas ao fluxo de dados realizam operações como filtragem, ordenação, aplicação de fatores de escala e cálculo de estatísticas. Nessas operações, o LLM não realiza diretamente o processamento numérico. Seu papel é selecionar a função apropriada, interpretar os argumentos fornecidos pelo usuário e definir os parâmetros necessários para a chamada. A execução propriamente dita é realizada por código determinístico. Dessa forma, uma vez definida a função e seus parâmetros de entrada, o mesmo conjunto de dados produz o mesmo resultado computacional.

Essa separação é particularmente importante em tarefas especializadas, pois reduz a dependência da geração textual do LLM para operações cujo resultado deve ser obtido diretamente a partir dos cabeçalhos ou das amostras sísmicas. A possível variabilidade introduzida pelo modelo fica, portanto, concentrada principalmente na interpretação da solicitação, na seleção da *tool*, na definição dos parâmetros e na sequência de ações, enquanto o processamento sísmico permanece sob responsabilidade de funções explicitamente implementadas.

O *Agente de Visualização* recebe a solicitação já contextualizada pelo fluxo de orquestração e dispõe de um conjunto específico de *tools* voltadas à representação e à análise exploratória dos dados. Entre as operações consideradas estão o cálculo de fold por CDP, a geração de seções sísmicas e gráficos *wiggle*, a transformação F–K e o cálculo do envelope sísmico. Por exemplo, uma solicitação para visualizar uma seção correspondente a determinado intervalo de CDPs pode exigir a seleção do subconjunto de traços, sua ordenação e, posteriormente, a invocação da ferramenta responsável pela geração da seção. Nesse caso, o agente coordena a sequência de operações, enquanto cada etapa numérica é executada pelas respectivas funções de domínio.

O mesmo LLM selecionado é utilizado pelos dois agentes como mecanismo de interpretação, tomada de decisão e coordenação das ações. Assim, o modelo atua como camada de raciocínio entre a solicitação em linguagem natural e as funcionalidades disponíveis, enquanto as *tools* realizam o acesso aos dados e as operações determinísticas. O LLM, portanto, não substitui os algoritmos de processamento sísmico, mas determina quando e como eles devem ser utilizados. A resposta final pode assumir a forma de texto, de um resultado numérico ou de uma visualização produzida pelas ferramentas de domínio.

Essa separação entre camada de raciocínio e camada de execução também favorece a modularidade da arquitetura. O modelo de linguagem pode ser substituído sem alterar a implementação das operações sísmicas, assim como novas *tools* podem ser incorporadas sem modificar a estrutura fundamental de interação entre os agentes.

2.3 Protocolo experimental

Foram selecionadas seis tarefas representativas da exploração inicial: (1) exibição de cabeçalhos; (2) cálculo e visualização do fold por CDP; (3) geração de uma seção sísmica após seleção e ordenação de traços; (4) geração de gráfico *wiggle*; (5) cálculo da transformação F–K; e (6) geração da seção de envelope. As tarefas abrangem desde consultas baseadas exclusivamente em metadados até operações que exigem seleção de traços e transformação do sinal.

Cada tarefa foi expressa por três formulações semanticamente equivalentes e repetida cinco vezes, totalizando 15 execuções por tarefa e por modelo. A utilização de formulações alternativas permite observar se pequenas variações linguísticas alteram o fluxo de execução, enquanto as repetições fornecem uma indicação da consistência operacional da interação.

Foram avaliados os modelos GPT-5-mini e o4-mini, ambos acessados por API (OPENAI, 2026a, 2026b), com esforço de raciocínio padrão (*medium*). São reportadas duas medidas obtidas nas execuções: o tempo médio de resposta e a taxa de geração, na primeira resposta, do tipo de saída esperada. A segunda medida verifica se o fluxo alcança a saída solicitada sem interação adicional; portanto, representa um indicador de execução do fluxo e não uma validação numérica integral do conteúdo produzido.

3 Resultados e Discussão

A Tabela 1 resume os resultados obtidos para o conjunto de dados utilizado. O GPT-5-mini apresentou maior regularidade na geração da saída esperada, com média de 96,67% nas seis tarefas. O o4-mini apresentou média de 88,89%, porém com menor tempo de resposta: 16,85 s em média, contra 54,49 s para o GPT-5-mini.

As tarefas de inspeção de cabeçalhos e fold foram executadas com alta regularidade pelos dois modelos. Nesses casos, a solicitação exige um caminho relativamente direto entre interpretação, consulta aos metadados e apresentação da informação. Diferenças mais evidentes aparecem nas tarefas que combinam seleção ou ordenação de traços com a geração de uma representação do sinal. Para o GPT-5-mini, as menores taxas ocorreram na transformação F–K e na seção de envelope; no o4-mini, a redução nessas duas tarefas foi mais acentuada.

A análise das execuções em que a visualização não foi produzida na primeira interação indica dois comportamentos recorrentes: solicitação de confirmação antes da geração do gráfico e encerramento da resposta após etapas intermediárias de preparação, sem acionamento da ferramenta gráfica final. Esses casos mostram que uma solicitação pode ser corretamente compreendida sem que toda a cadeia de ações seja concluída em uma única interação. Dessa forma, a métrica utilizada nesta avaliação captura simultaneamente a interpretação da intenção e a continuidade do fluxo de orquestração.

Tabela 1: *Resultados para seis tarefas no conjunto pré-empilhado de aproximadamente 50 MB. Cada valor resume 15 execuções por tarefa e modelo.*

Tarefa	GPT-5-mini		o4-mini	
	Tempo (s)	Saída esperada (%)	Tempo (s)	Saída esperada (%)
Exibição de cabeçalhos	11,02	100,00	11,30	100,00
Fold por CDP	29,10	100,00	11,69	100,00
Seção sísmica	59,87	100,00	17,39	93,33
Gráfico wiggle	58,77	100,00	21,45	93,33
Transformação F–K	84,58	86,67	21,20	73,33
Seção de envelope	83,59	93,33	18,04	73,33
Média	**54,49**	**96,67**	**16,85**	**88,89**

O o4-mini concentrou 10 das 13 execuções sem a saída esperada, oito delas nas tarefas de F–K e envelope, que exigem transformação do sinal após a seleção dos traços. Em ambos os comportamentos descritos, o turno é encerrado antes do acionamento da ferramenta final, o que sugere menor persistência do modelo em fluxos com várias etapas encadeadas. Essa diferença é coerente com a documentação da OpenAI, segundo a qual a família GPT-5 foi desenvolvida com ênfase no uso de ferramentas e é, por padrão, minuciosa ao reunir contexto em ambientes agênticos (OPENAI, 2025a). No τ^2-bench, que avalia agentes conversacionais com uso de ferramentas, a OpenAI também reporta, com esforço de raciocínio alto, desempenho superior do GPT-5-mini em dois de três domínios, com a maior diferença no de telecomunicações (74,1% contra 40,5%) (OPENAI, 2025b).

Os tempos de execução também evidenciam diferenças importantes para uma interface interativa. Na exibição de cabeçalhos, que envolve poucas decisões, os dois modelos levaram cerca de 11 s. À medida que a tarefa exige mais etapas, o tempo do GPT-5-mini chega a cerca de 85 s, enquanto o do o4-mini permanece entre 11 e 22 s; a razão entre os tempos passa de 2,5 no fold para 4,0 e 4,6 na F–K e no envelope. Como as ferramentas executadas são as mesmas para os dois modelos, a diferença decorre essencialmente da camada de raciocínio: cada interpretação, seleção de *tool* ou delegação entre agentes corresponde a uma chamada ao LLM, e a hipótese mais provável é que o GPT-5-mini dedique mais raciocínio a cada decisão, custo que se acumula nos fluxos mais longos. Os dados sugerem, portanto, um compromisso entre responsividade e regularidade do fluxo que deve ser considerado na seleção do modelo para aplicações interativas.

É importante distinguir a responsabilidade do agente daquela das ferramentas. Uma vez acionada a função correta com os parâmetros apropriados, operações como filtragem, ordenação, cálculo de fold, F–K ou envelope são executadas deterministicamente. Os erros observados nesta avaliação estão, portanto, associados principalmente à decisão de fluxo — seleção, sequência ou acionamento das ferramentas — e não a uma geração livre dos valores sísmicos pelo LLM. Essa característica é particularmente relevante em aplicações científicas, pois preserva uma fronteira clara entre interpretação linguística e processamento numérico.

A avaliação possui limitações que definem o alcance dos resultados. O experimento utiliza um único arquivo para a análise quantitativa condensada e uma métrica operacional de saída esperada, sem comparar numericamente cada elemento das visualizações com uma referência independente. Além disso, os resultados se restringem às seis tarefas implementadas e aos dois modelos avaliados. Ainda assim, a diversidade das operações — desde leitura de cabeçalhos até transformações do sinal — demonstra que uma mesma interface pode coordenar tarefas de diferentes níveis de complexidade sem exigir que o usuário programe explicitamente cada etapa.

4 Conclusões

Foi apresentada uma interface baseada em agentes de IA para exploração e controle de qualidade de dados sísmicos em formato SEG-Y. A arquitetura combina um Assistente Geral, responsável pela interpretação e coordenação das solicitações, com um Agente de Visualização especializado em saídas gráficas. Ambos utilizam um conjunto de ferramentas determinísticas de domínio, mantendo separado o papel do LLM no raciocínio e na coordenação do processamento sísmico propriamente dito.

Em seis tarefas representativas, a interface foi capaz de transformar solicitações em linguagem natural

em operações de inspeção de metadados, seleção e ordenação de traços, visualização e transformação exploratória. Os dois modelos avaliados apresentaram comportamentos distintos: o GPT-5-mini obteve maior taxa média de geração da saída esperada na primeira interação, enquanto o o4-mini apresentou menor latência média. A análise sugere que o GPT-5-mini dedica mais raciocínio a cada decisão, o que eleva a latência em tarefas com várias etapas, mas conclui a cadeia de ferramentas com mais frequência; já o o4-mini tende a encerrar a interação antes da ferramenta de visualização final. Esses resultados evidenciam que a escolha do modelo afeta não apenas o tempo de resposta, mas também a continuidade do fluxo agentivo.

Para trabalhos futuros, torna-se importante ampliar a avaliação para outros arquivos SEG-Y e classes de tarefas, incorporar critérios quantitativos de validação do conteúdo gerado e investigar formas de tornar a coordenação mais robusta diante de solicitações compostas. A manutenção das operações sísmicas em ferramentas determinísticas permanece como princípio central para que a interface possa evoluir sem perder rastreabilidade e controle sobre o processamento realizado.

Agradecimentos

Os autores agradecem o apoio da Petrobras por meio do projeto “Cognitive assistant for geophysical data” e reconhecem a importância estratégica do apoio da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), por meio da regulamentação de P&D. Os autores também agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio às atividades de pesquisa associadas a este trabalho, e ao Departamento de Engenharia de Computação e Automação (DCA) da Universidade Federal do Rio Grande do Norte (UFRN), pela disponibilização do espaço e da infraestrutura do Laboratório Model2AI.

Referências

CHASE, H. *LangChain*. 2026. Disponível em: <https://github.com/langchain-ai/langchain>. Acesso em: nov. 2026.

CHENG, Y. et al. *Exploring Large Language Model based Intelligent Agents: Definitions, Methods, and Prospects*. 2024. Disponível em: <https://arxiv.org/abs/2401.03428>.

DAVIS, W.; HUNT, C. R. Knowledge graphs for seismic data and metadata. *Applied Computing and Geosciences*, Elsevier BV, v. 21, p. 100151, mar. 2024. ISSN 2590-1974.

EQUINOR. *Segyio: Fast Python library for SEG-Y files*. 2026. Disponível em: <https://segyio.readthedocs.io/en/stable/>. Acesso em: nov. 2026.

HALLAM, A. *SEGY-SAK*. 2025. Disponível em: <https://trhallam.github.io/segysak/>. Acesso em: nov. 2026.

KHOSRO ANJOM, F.; VACCARINO, F.; SOCCO, L. V. Machine learning for seismic exploration: Where are we and how far are we from the holy grail? *GEOPHYSICS*, Society of Exploration Geophysicists, v. 89, n. 1, p. WA157–WA178, jan. 2024. ISSN 1942-2156.

KIM, S.; YU, Y.; SEO, H. Artificial intelligence orchestration for text-based ultrasonic simulation via self-review by multi-large language model agents. *Scientific Reports*, Springer Science and Business Media LLC, v. 15, n. 1, p. 12474, abr. 2025. ISSN 2045-2322.

LANGCHAIN INC. *LangGraph*. 2026. Disponível em: <https://github.com/langchain-ai/langgraph>. Acesso em: nov. 2026.

OPENAI. *GPT-5 prompting guide*. 2025. Disponível em: <https://developers.openai.com/cookbook/examples/gpt-5/gpt-5_prompting_guide>. Acesso em: nov. 2026.

OPENAI. *Introducing GPT-5 for developers*. 2025. Disponível em: <https://openai.com/index/introducing-gpt-5-for-developers/>. Acesso em: nov. 2026.

OPENAI. *GPT-5-mini Model Documentation*. 2026. Disponível em: <https://platform.openai.com/docs/models/gpt-5-mini>. Acesso em: nov. 2026.

OPENAI. *o4-mini Model Documentation*. 2026. Disponível em: <https://platform.openai.com/docs/models/o4-mini>. Acesso em: nov. 2026.

SEG TECHNICAL STANDARDS COMMITTEE. *SEG Y rev 1 Data Exchange Format*. [S.l.], 2002.

WANG, L. et al. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, Springer Science and Business Media LLC, v. 18, n. 6, p. 186345, mar. 2024. ISSN 2095-2236.

XAVIER, M. et al. LLM-powered multi-actor system for intelligent analysis and visualization of IEC 61499 control systems. In: *IECON 2024 - 50th Annual Conference of the IEEE Industrial Electronics Society*. [S.l.]: IEEE, 2024. p. 1–8.

ZHANG, J. et al. Streamlining geoscience data analysis with an LLM-driven workflow. *Applied Computing and Geosciences*, Elsevier BV, v. 25, p. 100218, fev. 2025. ISSN 2590-1974.

Identificação de padrões e anomalias em reclamações de telecomunicações no Nordeste brasileiro usando aprendizado não supervisionado

Sayonara Cristina de Oliveira Magalhães [1]
Emelyn Clementino Freire [1,2]
Arthur Phelipe de Oliveira Queiroz [2]
Marcelo Jorge Oliveira de Albuquerque Melo [2]

[1] *Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PPgEEC), Centro de Tecnologia, Universidade Federal do Rio Grande do Norte (UFRN), Natal, RN, Brasil*
[2] *Departamento de Engenharia de Computação e Automação (DCA), UFRN, Natal, RN, Brasil*
<sayonara.magalhaes.068@ufrn.edu.br>,
<emelyn.freire.116@ufrn.edu.br>,
<arthur.queiroz.115@ufrn.edu.br>,
<marcelo.melo.703@ufrn.edu.br>

Resumo

Reclamações de consumidores de telecomunicações dos nove estados do Nordeste brasileiro, registradas entre janeiro de 2020 e agosto de 2026, foram agregadas por estado, município, mês e serviço, o que resultou em 212.626 registros e 2.003.561 solicitações. Para cada um dos 1.794 municípios, foram calculados 11 atributos de volume, variabilidade, serviços e tipos de problema. Os atributos padronizados foram submetidos ao K-Means e, após redução por análise de componentes principais (PCA), ao DBSCAN, com avaliação pelo Silhouette Score e pelo índice de Davies-Bouldin. O K-Means separou os municípios em dois grupos, de 698 e 1.096 municípios, com separação fraca (Silhouette de 0,1870). O DBSCAN (`eps` = 1,0; `min_samples` = 10) formou um grupo principal de 1.282 municípios e dois grupos menores, de 17 e 6, e rotulou 489 municípios (27,26%) como ruído, entre eles grandes capitais e 282 municípios com menos de 100 solicitações. Para ordenar prioridades, calculou-se a taxa de reclamações por 100 mil habitantes nos 12 meses mais recentes e sua variação. Os 20 municípios de maior taxa estão fora do grupo principal do DBSCAN, e Salvador lidera, com taxa 4,0 vezes a regional. Os resultados indicam onde investigar e não equivalem a uma medida de qualidade do serviço.

Palavras-chave: aprendizado não supervisionado; K-Means; DBSCAN; telecomunicações; detecção de anomalias.

Abstract

Consumer complaints about telecommunications services in the nine states of the Brazilian Northeast, recorded from January 2020 to August 2026, were aggregated by state, municipality, month and service, yielding 212,626 records and 2,003,561 requests. For each of the 1,794 municipalities, we computed 11 attributes describing volume, variability, services and problem types. The standardized attributes were clustered with K-Means and, after principal component analysis (PCA), with DBSCAN; both were assessed with the Silhouette Score and the Davies-Bouldin index. K-Means split the municipalities into two groups of 698 and 1,096, with weak separation (Silhouette of 0.1870). DBSCAN (`eps` = 1.0; `min_samples` = 10) formed a main cluster of 1,282 municipalities and two smaller ones, of 17 and 6, and labelled 489 municipalities (27.26%) as noise, among

them large capitals and 282 municipalities with fewer than 100 requests. To rank priorities, we computed the complaint rate per 100,000 inhabitants over the most recent 12 months and its change. The 20 municipalities with the highest rates lie outside the main DBSCAN cluster, and Salvador ranks first, at 4.0 times the regional rate. The results indicate where to investigate and are not a measure of service quality.

Keywords: unsupervised learning; K-Means; DBSCAN; telecommunications; anomaly detection.

1 Introdução

Os serviços de telecomunicações sustentam a comunicação, o acesso à informação e boa parte da atividade econômica. As reclamações registradas pelos consumidores constituem um registro contínuo da qualidade percebida desses serviços e podem indicar localidades que merecem investigação. Somente em 2020, a plataforma Anatel Consumidor recebeu 2,96 milhões de reclamações no país (FREITAS; ANDREÃO, 2021).

Esse volume inviabiliza o exame manual, e os registros não trazem rótulo que indique quais municípios apresentam comportamento fora do esperado. O aprendizado não supervisionado é adequado a esse cenário, pois agrupa observações semelhantes e destaca as que se afastam do comportamento predominante sem depender de classes definidas previamente (JAIN, 2010; CHANDOLA; BANERJEE; KUMAR, 2009).

O objetivo deste trabalho é identificar perfis municipais e comportamentos atípicos nas reclamações de telecomunicações dos nove estados do Nordeste brasileiro. Para isso, descreve-se um fluxo reproduzível que parte da agregação mensal dos registros, constrói atributos por município e compara dois algoritmos de natureza distinta: o K-Means, baseado em centroides, e o DBSCAN, baseado em densidade e capaz de rotular observações como ruído.

2 Trabalhos relacionados

Entre os métodos de agrupamento, o K-Means é um dos mais difundidos. Ele particiona as observações em k grupos de modo a minimizar a variância interna de cada grupo (MACQUEEN, 1967). Jain (2010) revisa cinco décadas de uso do algoritmo e aponta suas restrições conhecidas: o número de grupos precisa ser informado, o resultado depende da inicialização e grupos de tamanho ou formato muito desiguais tendem a ser mal representados.

O DBSCAN define grupos como regiões de alta densidade, delimitadas por um raio de vizinhança (`eps`) e por um número mínimo de vizinhos (`min_samples`), e rotula como ruído os pontos que não pertencem a nenhuma dessas regiões (ESTER et al., 1996). Schubert et al. (2017) mostram que o desempenho e o resultado do algoritmo dependem fortemente da escolha desses dois parâmetros.

Na taxonomia de Chandola, Banerjee e Kumar (2009), uma família de técnicas de detecção de anomalias baseadas em agrupamento assume que observações normais pertencem a algum grupo e anomalias, a nenhum. O rótulo de ruído do DBSCAN corresponde a essa definição. Para avaliar partições sem rótulos de referência, são usuais o Silhouette Score (ROUSSEEUW, 1987) e o índice de Davies-Bouldin (DAVIES; BOULDIN, 1979).

No domínio das reclamações de telecomunicações, os estudos localizados tratam a reclamação individual como unidade de análise. Freitas e Andreão (2021) propõem uma rotina de limpeza do texto de respostas a reclamações da plataforma Anatel Consumidor, voltada à construção de classificadores supervisionados. Cardoso (2024) aplica o K-Means a 794 reclamações sobre chamadas não solicitadas. Ubeda et al. (2025) classificam a avaliação atribuída pelo consumidor à resposta da operadora em mais de 900 mil reclamações do Consumidor.gov.br. Oliveira et al. (2025) descrevem 2.061.022 reclamações de telecomunicações da mesma plataforma, registradas de 2015 a 2024, com estatística descritiva e teste qui-quadrado. Assis et al. (2017) relacionam reclamações coletadas no Twitter, localizadas por cidade, ao número de estações rádio base da operadora. Fora do contexto brasileiro, Faed et al. (2016) combinam PCA e análise envoltória de dados no tratamento de reclamações de clientes. Este trabalho muda a unidade de análise para o município: cada um é descrito por atributos calculados sobre todo o período e agrupado com o K-Means e o DBSCAN.

3 Metodologia

3.1 Dados e preparação mensal

Os dados provêm do Anatel Consumidor, sistema da Agência Nacional de Telecomunicações (Anatel) destinado ao registro de reclamações de consumidores de serviços de telecomunicações, e são disponibilizados publicamente pela agência (AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES, 2026b). Cada registro informa o mês de referência, a unidade da federação, o município (nome e código), o serviço, a descrição do problema e a quantidade de solicitações.

Os arquivos obtidos foram consolidados em uma única base, que cobre o período de janeiro de 2020 a agosto de 2026. Foram mantidos os registros dos nove estados da região Nordeste (Alagoas, Bahia, Ceará, Maranhão, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe).

Os registros foram agregados por estado, município, mês e serviço, o que produziu 212.626 linhas. Nove municípios apareciam com duas grafias do nome e foram unificados pelo código do IBGE, de modo que a base reúne 1.794 municípios. Os serviços presentes são telefonia móvel pré-paga (SMP_Pré) e pós-paga (SMP_Pós), banda larga fixa (SCM), telefonia fixa (STFC), TV por assinatura (SeAC) e uma categoria Outros, não detalhada na base. Após a agregação, verificaram-se colunas, duplicidades, cobertura de períodos e estados, valores nulos e valores negativos.

O campo de problema é textual. A partir dele, após conversão para minúsculas, três indicadores foram criados por busca de expressões: lentidão ("lentidão" ou "velocidade reduzida"), queda de conexão ("queda de conexão", "linha muda" ou "sem sinal") e interrupção do serviço ("interrupção do serviço" ou "sem conexão"). Cada indicador soma as solicitações dos registros cuja descrição contém alguma das expressões. Uma mesma descrição pode satisfazer mais de um padrão textual; por isso, os indicadores são analisados separadamente e não formam categorias mutuamente exclusivas.

3.2 Atributos municipais e padronização

Cada município foi descrito por 11 atributos: o logaritmo do total de solicitações no período, na forma $\log(1 + x)$; o coeficiente de variação das solicitações entre os registros de mês e serviço do município; as taxas de lentidão, queda e interrupção, dadas pela razão entre as ocorrências de cada indicador e o total de solicitações do município; e a participação de cada um dos seis serviços no total do município. A transformação logarítmica reduz a influência dos municípios de maior volume. Em municípios com um único registro, o coeficiente de variação foi fixado em zero.

Como os atributos têm escalas diferentes e tanto o K-Means quanto o DBSCAN dependem de distâncias, todas as variáveis foram padronizadas para média zero e desvio-padrão unitário.

3.3 Redução de dimensionalidade

A análise de componentes principais (PCA) projeta os dados em direções ortogonais de máxima variância (JOLLIFFE; CADIMA, 2016). Foram retidos cinco componentes, que explicam 71,67% da variância total (25,82%, 13,27%, 12,00%, 11,18% e 9,40%, na ordem). Esse espaço reduzido foi a entrada do DBSCAN, cuja noção de densidade se degrada à medida que a dimensionalidade cresce. Uma segunda projeção, com dois componentes (39,10% da variância), foi usada apenas para visualização (Figura 2).

3.4 K-Means

O K-Means foi aplicado diretamente aos 11 atributos padronizados, sem PCA, para k de 2 a 8 com inicialização k-means++, 20 reinicializações e semente fixa (42). Selecionou-se o valor de k com maior Silhouette Score.

3.5 DBSCAN

O DBSCAN foi avaliado em uma grade com `eps` $\in \{0{,}4; 0{,}6; 0{,}8; 1{,}0; 1{,}2; 1{,}5; 2{,}0\}$ e `min_samples` $\in \{5; 10\}$, num total de 14 combinações. Selecionou-se a de maior Silhouette Score entre as que produziram ao menos dois grupos e no máximo 30% de ruído. Para o DBSCAN, o Silhouette e o Davies-Bouldin foram calculados apenas sobre os municípios atribuídos a algum grupo. Os pontos de ruído não formam um grupo, mas o conjunto dos municípios que não pertencem a nenhum; tratá-los como um grupo adicional reuniria observações sem semelhança entre si e distorceria a métrica. Nesse caso, a silhueta média desses municípios seria de $-0{,}263$, e a do conjunto cairia para 0,1334. Por isso, a qualidade dos grupos e a fração de ruído são relatadas separadamente. Os municípios com rótulo -1 foram tratados como observações de comportamento atípico em relação aos grupos encontrados. O rótulo descreve apenas a posição do município no espaço de atributos: ele não mede gravidade nem qualidade do serviço.

3.6 Métricas de avaliação

Os agrupamentos foram avaliados por duas métricas que dispensam rótulos de referência. O Silhouette Score (ROUSSEEUW, 1987) compara, para cada observação, a distância média aos membros do próprio grupo com a distância média ao grupo vizinho mais próximo; sua média varia de -1 a 1, e valores maiores indicam grupos mais coesos e separados. O índice de Davies-Bouldin (DAVIES; BOULDIN, 1979) relaciona a dispersão interna dos grupos à distância entre seus centroides, e valores menores indicam melhor separação. As análises foram implementadas em Python com a biblioteca scikit-learn (PEDREGOSA et al., 2011).

3.7 Taxas por habitante e priorização

Os grupos descrevem perfis, mas não ordenam os municípios. Para indicar onde concentrar a investigação, calculou-se a taxa de reclamações por 100 mil habitantes nos 12 meses mais recentes da base (setembro de 2025 a agosto de 2026) e sua variação em relação aos 12 meses anteriores. A população municipal foi recuperada do indicador de densidade de TV por assinatura publicado pela Anatel, pela razão entre acessos e densidade (AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES, 2026a); os valores de 2022 são compatíveis com os do Censo Demográfico. Para evitar taxas instáveis, a ordenação considera apenas os 592 municípios com 20 mil habitantes ou mais. Na banda larga fixa e na telefonia móvel, que somam cerca de 80% das solicitações, calculou-se também a taxa por mil acessos: as solicitações do serviço em 12 meses divididas pela média mensal de acessos do município no mesmo período (AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES, 2026a). Os acessos móveis estavam disponíveis até junho de 2026, e a média considera os meses disponíveis.

4 Resultados

4.1 Análise exploratória

A base agregada soma 2.003.561 solicitações entre janeiro de 2020 e agosto de 2026. A Bahia concentra 733.882 delas (36,63%), seguida por Pernambuco (18,60%) e Ceará (15,90%); os outros seis estados têm participações entre 3,32% e 6,43%. Entre os serviços, o SMP_Pós é o mais frequente, com 682.663 solicitações (34,07%), seguido por SMP_Pré (23,72%), SCM (21,73%), STFC (10,22%), SeAC (5,38%) e Outros (4,87%).

A concentração repete-se entre os municípios. Salvador, Fortaleza e Recife lideram o volume, e as nove capitais estão entre os dez municípios com mais solicitações, ao lado de Feira de Santana.

O volume mensal diminuiu ao longo da maior parte do período. O máximo ocorreu em maio de 2020, com 44.281 solicitações, e o mínimo em novembro de 2024, com 16.583. A média mensal caiu de 38.691 solicitações em 2020 para 18.233 em 2024 e subiu para 21.525 nos oito meses observados de 2026 (Figura 1).

Os três indicadores textuais somam 226.226 ocorrências: 107.940 de interrupção (47,71%), 84.335 de queda de conexão (37,28%) e 33.951 de lentidão (15,01%). Esses percentuais têm como base a soma das ocorrências, que equivale a, no máximo, 11,29% do total de solicitações, já que uma descrição pode ser contada em mais de um indicador. A banda larga fixa (SCM) concentra 73,80% das ocorrências de lentidão e 53,70% das de interrupção, enquanto a telefonia fixa (STFC) responde por 48,78% das de queda.

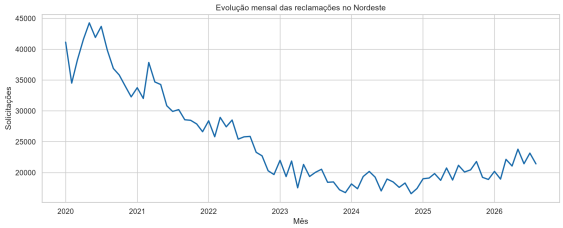


Figura 1: *Evolução mensal do total de solicitações no Nordeste (jan. 2020 a ago. 2026).*

4.2 Comparação dos modelos

No K-Means, o maior Silhouette foi obtido com $k = 2$ (0,1870), seguido por $k = 4$ (0,1797), enquanto o Davies-Bouldin diminui com o número de grupos. Em todas as configurações, o Silhouette é baixo, o que indica grupos pouco separados (Tabela 1).

No DBSCAN, a fração de ruído variou de 2,29% a 100% entre as 14 combinações, e oito delas produziram no máximo um grupo. Entre as que respeitam o limite de 30% de ruído, `eps` = 1,0 e `min_samples` = 10 obteve o maior Silhouette (0,2907), com três grupos e um conjunto de ruído.

Tabela 1: *Métricas de avaliação por configuração (selecionadas em negrito). No K-Means, o parâmetro é k; no DBSCAN, `eps` e `min_samples` (MS), apenas nas configurações com dois ou mais grupos. Silh.: Silhouette; DB: Davies-Bouldin.*

Modelo	Parâmetro	Grupos	Ruído (%)	Silh.	DB
K-Means	$\boldsymbol{k = 2}$	**2**	—	**0,1870**	**2,0104**
K-Means	$k = 3$	3	—	0,1778	2,2374
K-Means	$k = 4$	4	—	0,1797	1,8006
K-Means	$k = 5$	5	—	0,1562	1,6836
K-Means	$k = 6$	6	—	0,1293	1,7330
K-Means	$k = 7$	7	—	0,1283	1,6741
K-Means	$k = 8$	8	—	0,1310	1,6023
DBSCAN	0,4; MS 5	8	97,44	0,4654	0,7698
DBSCAN	0,6; MS 5	11	61,37	−0,1077	0,8245
DBSCAN	0,8; MS 5	6	35,01	0,0743	0,7067
DBSCAN	1,0; MS 5	3	19,12	0,2661	0,6463
DBSCAN	**1,0; MS 10**	**3**	**27,26**	**0,2907**	**0,7967**
DBSCAN	1,2; MS 5	3	11,87	0,2899	0,6956

A leitura da Tabela 1 exige cautela. O K-Means atribui todos os 1.794 municípios a algum grupo e opera sobre os atributos padronizados. O DBSCAN opera sobre cinco componentes principais e deixa 489 municípios fora dos grupos, e esses municípios não entram no cálculo de suas métricas. As métricas,

portanto, descrevem partições de conjuntos e espaços diferentes e não ordenam diretamente os dois algoritmos.

4.3 Perfis dos grupos do K-Means

O Grupo A reúne 698 municípios (38,91%), com média de 2.695,61 solicitações por município. O Grupo B reúne 1.096 municípios (61,09%), com média de 111,34. As taxas médias de lentidão, queda e interrupção do Grupo A são de 1,9 a 2,2 vezes as do Grupo B, e seu coeficiente de variação médio também é maior (Tabela 2).

A composição dos serviços difere entre os grupos. No Grupo B, o SMP_Pré responde, em média, por 51,22% das solicitações do município; no Grupo A, por 34,96%. O Grupo A tem participações médias maiores de SMP_Pós, SCM e STFC (8,43% contra 3,17%).

Na projeção sobre os dois primeiros componentes principais (Figura 2, à esquerda), os dois grupos ocupam regiões contíguas, separadas ao longo do primeiro componente, sem lacuna visível entre elas.

4.4 Grupos e municípios atípicos no DBSCAN

O DBSCAN identificou um grupo principal com 1.282 municípios (71,46%), dois grupos menores, com 17 e 6 municípios, e 489 municípios (27,26%) rotulados como ruído (Figura 2, à direita). O grupo 1 tem a maior média de solicitações (23.596,35) e a maior participação média de banda larga fixa: 26,30%, contra 5,33% no grupo principal (Tabela 2). Ele reúne cinco capitais (João Pessoa, Aracaju, Maceió, Natal e Teresina) e outros doze municípios: Campina Grande, Vitória da Conquista, Caruaru, Itabuna, Parnamirim, Juazeiro do Norte, Caucaia, Mossoró, Nossa Senhora do Socorro, São José de Ribamar, Barreiras e Crato. Seus totais somam 401.138 solicitações (20,02% do total). O grupo 2 (Parnaíba, Timon, Aquiraz, Quixadá, Cascavel e Ceará-Mirim) tem volume bem menor, de 1.129 a 3.534 solicitações por município, e também se distingue pela banda larga fixa (22,52%). Os 23 municípios dos dois grupos pertencem ao Grupo A do K-Means.

Na Figura 2, o grupo principal concentra-se na região central da projeção, e os grupos 1 e 2 aparecem em valores elevados do primeiro componente. Os pontos de ruído predominam na periferia, mas também se sobrepõem ao grupo principal. Essa sobreposição é esperada, pois o DBSCAN foi ajustado em cinco componentes e a figura exibe apenas dois, que retêm 39,10% da variância.

Os municípios rotulados como ruído não foram tratados como problemáticos. O conjunto é heterogêneo e reúne os dois extremos de porte. De um lado, a mediana é de 68 solicitações, e 282 dos 489 municípios (57,67%) registraram menos de 100 no período. De outro, estão nele Salvador, Fortaleza, Recife, São Luís e Feira de Santana, e o conjunto soma 1.210.580 solicitações (60,42% do total). Em relação ao K-Means, 262 desses municípios pertencem ao Grupo A e 227 ao Grupo B.

4.5 Taxas por habitante e prioridades

Nos 12 meses mais recentes, o Nordeste registrou 252.564 solicitações, ou 462,2 por 100 mil habitantes, 11,0% acima dos 12 meses anteriores (416,3). A Bahia tem a maior taxa (592,0) e o Maranhão, a menor (241,1); as maiores altas ocorreram no Piauí (25,3%) e no Rio Grande do Norte (24,6%). Entre os serviços, cresceram o SMP_Pós (18,2%), o SCM (9,4%) e a categoria Outros (64,8%), enquanto recuaram o STFC (−24,6%), o SeAC (−20,0%) e o SMP_Pré (−3,6%).

Entre os 592 municípios com 20 mil habitantes ou mais, a taxa mediana é de 176,8, e nove superam o dobro da taxa regional. A Tabela 3 lista os dez primeiros. A divisão pela população não desfaz a concentração nos grandes centros: Salvador permanece em primeiro lugar, com taxa 4,0 vezes a regional, seis dos dez municípios são capitais, e outros dois, Lauro de Freitas e Olinda, são vizinhos de capitais. A correlação de postos entre população e taxa (0,45) é, ainda assim, menor que entre população e volume (0,79).

Dois casos ilustram o uso conjunto de taxa e variação. Cocal (PI), com 28 mil habitantes, é o único município de pequeno porte na lista: suas solicitações mais que quadruplicaram, e 69,8% delas referem-se à TV por assinatura, o que sugere um problema localizado em um único serviço. Natal combina taxa elevada (2,2 vezes a regional) com a maior alta entre as capitais (27,5%).

Na banda larga fixa, os acessos permitem um denominador mais adequado. A região registrou 6,44 solicitações por mil acessos, praticamente o mesmo valor do período anterior (6,41): o aumento de 9,4% nas solicitações do serviço acompanha o crescimento de 8,8% da base de acessos. A ordem dos municípios pouco se altera (Tabela 4). Salvador mantém a primeira posição, com taxa 4,3 vezes a regional, embora tenha menos acessos por 100 habitantes (21,7) do que João Pessoa (35,8) ou Natal (34,2); sua posição, portanto, não decorre de maior difusão do serviço. Quatro municípios da região metropolitana do

Tabela 2: *Perfil médio dos grupos do K-Means e do DBSCAN (taxas e participações em % das solicitações do município). CV: coeficiente de variação.*

Grupo	Municípios	Média de solic.	CV	Lentidão	Queda	Interrupção	SMP_Pré	SMP_Pós	SCM
K-Means A	698	2.695,61	0,69	1,56	5,41	5,63	34,96	31,81	9,32
K-Means B	1.096	111,34	0,49	0,82	2,43	3,03	51,22	21,23	4,88
DBSCAN principal	1.282	295,61	0,58	0,81	2,95	3,50	47,85	25,54	5,33
DBSCAN grupo 1	17	23.596,35	0,87	2,07	2,86	4,17	22,37	35,20	26,30
DBSCAN grupo 2	6	2.145,67	0,77	2,38	3,11	5,03	32,12	24,72	22,52
DBSCAN ruído	489	2.475,62	0,53	1,83	5,29	5,44	38,09	24,51	9,09

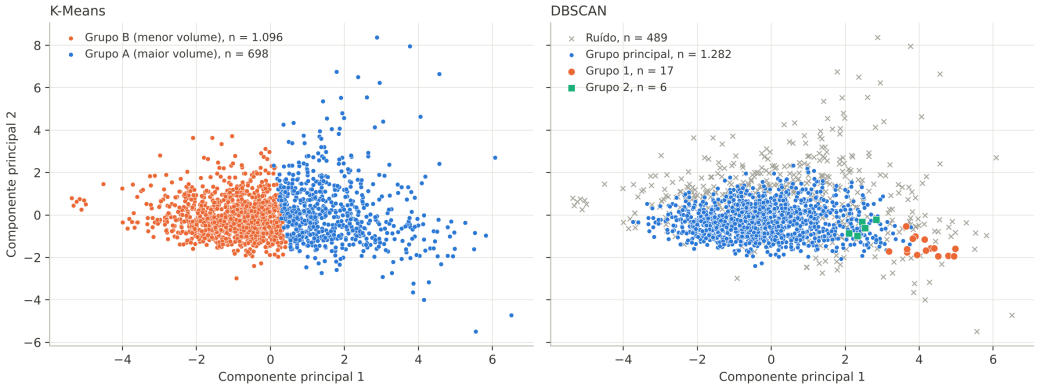


Figura 2: *Municípios no plano dos dois primeiros componentes principais, com os grupos do K-Means (esquerda) e do DBSCAN (direita).*

Recife estão entre os sete primeiros. As maiores altas entre as capitais ocorreram em Natal (54,1%) e Teresina (43,7%).

Na telefonia móvel, o quadro é outro. A taxa regional foi de 2,54 solicitações por mil acessos, 7,4% acima do período anterior: as solicitações cresceram 9,2% e a base, 1,7%, de modo que a alta não se explica pelo crescimento da base. O pós-pago concentra o problema, com 3,45 solicitações por mil acessos, o dobro do pré-pago (1,73), e alta de 7,6%. Entre os municípios, a taxa é menos associada ao porte (correlação de postos de 0,24 com a população). Salvador tem a segunda maior taxa (6,24, ou 2,5 vezes a regional), e seis dos dez primeiros municípios são baianos; seis dos dez, porém, têm menos de cem solicitações no período e devem ser lidos com reserva.

Os rótulos do DBSCAN acompanham a ordenação por habitante. Nenhum dos 20 municípios de maior taxa por habitante pertence ao grupo principal: 11 estão no ruído e 9 no grupo 1. Entre os municípios com 20 mil habitantes ou mais, a taxa mediana é de 165,4 no grupo principal, 264,4 no ruído, 247,8 no grupo 2 e 629,9 no grupo 1.

5 Discussão

A partição do K-Means é dominada pelo volume: a média de solicitações do Grupo A é cerca de 24 vezes a do Grupo B.

A solução com dois grupos é simples de comunicar, mas o Silhouette de 0,1870 indica separação fraca, coerente com a Figura 2, em que os grupos formam uma nuvem contínua. Os perfis municipais parecem variar de modo gradual, e a partição funciona mais como um corte ao longo de um gradiente de volume do que como identificação de grupos naturais.

O DBSCAN complementa essa leitura por não forçar todo município a pertencer a um grupo. Seu Silhouette (0,2907) deve ser lido com reserva, pois, calculado sem os pontos de ruído, refere-se a uma partição com um grupo de 1.282 municípios e dois grupos pequenos. A fração de 27,26% de ruído é elevada para ser lida como conjunto de anomalias em sentido estrito e depende dos parâmetros: entre as configurações com dois ou mais grupos, variou de 11,87% a 97,44%, o que está de acordo com a sensibilidade descrita por Schubert et al. (2017).

A composição do ruído reforça a interpretação adotada. Ele reúne capitais como Salvador, Fortaleza e Recife e, em maior número, municípios com poucas dezenas de solicitações. O que esses casos têm em comum é a distância em relação ao perfil mais frequente, e não um nível de qualidade de serviço. Os grupos 1 e 2, por sua vez, sugerem um perfil comum a parte das capitais e a municípios urbanos de porte intermediário, marcado pela maior partici-

Tabela 3: *Dez municípios com maior taxa de reclamações por 100 mil habitantes entre set. 2025 e ago. 2026 (municípios com 20 mil habitantes ou mais). Variação em relação aos 12 meses anteriores; serviço com maior número de solicitações no período e sua participação.*

	Município (UF)	População	Solicitações	Taxa	Variação (%)	Serviço principal	DBSCAN
1	Salvador (BA)	2.418.005	45.006	1.861,3	+0,5	SMP_Pós (37,9%)	ruído
2	Lauro de Freitas (BA)	203.334	3.125	1.536,9	−2,2	SMP_Pós (40,7%)	ruído
3	Recife (PE)	1.488.920	20.379	1.368,7	+10,1	SMP_Pós (37,1%)	ruído
4	Cocal (PI)	28.212	354	1.254,8	+359,7	SeAC (69,8%)	ruído
5	Aracaju (SE)	602.757	7.408	1.229,0	+17,0	SMP_Pós (39,9%)	grupo 1
6	João Pessoa (PB)	833.932	9.619	1.153,5	+19,1	SMP_Pós (39,3%)	grupo 1
7	Fortaleza (CE)	2.428.678	27.687	1.140,0	+13,8	SCM (38,4%)	ruído
8	Natal (RN)	751.300	7.521	1.001,1	+27,5	SMP_Pós (36,5%)	grupo 1
9	Olinda (PE)	349.976	3.372	963,5	+12,8	SMP_Pós (34,5%)	ruído
10	Feira de Santana (BA)	616.279	5.378	872,7	+5,0	SMP_Pós (48,3%)	ruído

Tabela 4: *Municípios com maior taxa de solicitações de banda larga fixa por mil acessos entre set. 2025 e ago. 2026 (20 mil habitantes ou mais). Acessos: média mensal, em milhares.*

Município (UF)	Acessos	Taxa	Variação (%)
Salvador (BA)	525,2	27,9	−6,3
Olinda (PE)	44,9	24,4	+5,6
Recife (PE)	286,3	23,4	+6,9
Lauro de Freitas (BA)	46,8	22,0	−38,7
Fortaleza (CE)	718,3	14,8	+15,9
Paulista (PE)	40,9	12,5	+6,6
Jaboatão dos Guararapes (PE)	72,5	12,4	+3,7
Eunápolis (BA)	16,8	10,8	+10,9

pação da banda larga fixa.

As taxas por habitante dão a esses rótulos uma leitura prática. O agrupamento, feito sem informação de população, separou do grupo principal os municípios que depois se mostraram de maior taxa, de modo que o ruído e os grupos menores delimitam onde procurar. A ordenação por taxa e variação indica por onde começar: municípios de taxa alta e estável, como Salvador, apontam problemas persistentes, e os de alta recente, como Cocal e Natal, apontam mudanças a examinar. Essa leitura requer cautela, pois a taxa por habitante também reflete a difusão dos serviços e a propensão a recorrer à Anatel, maiores nos grandes centros. Na banda larga fixa, a taxa por acessos manteve a ordem, o que sugere que a difusão não explica o resultado nesse serviço. Na telefonia móvel, a concentração nas capitais diminui, embora Salvador e Lauro de Freitas permaneçam entre os primeiros.

6 Limitações

Os tipos de problema foram identificados por palavras-chave, de modo que descrições equivalentes redigidas com outros termos podem não ter sido captadas. Os acessos foram usados como denominador na banda larga fixa e na telefonia móvel; a taxa geral e a dos demais serviços usam a população, o que impede separar a incidência de problemas da penetração dos serviços. A população foi obtida de forma indireta, a partir de um indicador da Anatel. A categoria Outros depende de documentação da fonte para ser interpretada.

No agrupamento, a agregação de todo o período em um único vetor por município não distingue mudanças de comportamento ao longo dos anos. O número de municípios rotulados como ruído depende diretamente de `eps` e `min_samples`, e a estabilidade dos grupos frente a variações desses parâmetros e da inicialização do K-Means não foi avaliada. Os limites de 30% de ruído, na seleção do DBSCAN, e de 20 mil habitantes, na ordenação, são arbitrários. Por fim, os dois algoritmos foram aplicados a espaços distintos, o que restringe a comparação entre suas métricas.

7 Conclusão

A agregação mensal das reclamações e a construção de atributos por município permitiram descrever 1.794 municípios do Nordeste com dois algoritmos não supervisionados. O K-Means separou os municípios em dois grupos, diferenciados sobretudo pelo volume de solicitações, com separação fraca (Silhouette de 0,1870). O DBSCAN identificou um grupo principal, dois grupos menores, de 17 e 6 municípios, e 489 municípios fora do padrão dominante.

A taxa por 100 mil habitantes e sua variação recente acrescentaram uma ordem de prioridade. Os 20 municípios de maior taxa estão fora do grupo principal do DBSCAN, e a lista é encabeçada por Salvador, Lauro de Freitas e Recife; Cocal destaca-se pela alta abrupta, concentrada na TV por assinatura. Na banda larga fixa, a taxa por mil acessos mantém Salvador na primeira posição; na telefonia

móvel, a taxa regional subiu 7,4%, puxada pelo pós-pago. Os resultados são preliminares: indicam onde investigar e não permitem afirmar, por si, onde o serviço é pior. Como trabalhos futuros, propõe-se estender o cálculo por acessos à telefonia fixa e à TV por assinatura, usar as categorias de problema da própria Anatel, avaliar a estabilidade dos grupos e construir um painel de acompanhamento mensal.

Referências

AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES. *Painéis de dados: acessos*. Brasília: Anatel, 2026. Dados brutos de banda larga fixa, telefonia móvel e TV por assinatura, em arquivos ZIP. Disponível em: <https://informacoes.anatel.gov.br/paineis/acessos>. Acesso em: 9 out. 2026.

AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES. *Painéis de dados: reclamações*. Brasília: Anatel, 2026. Dados brutos em arquivo ZIP. Disponível em: <https://informacoes.anatel.gov.br/paineis/consumidor/reclamacoes>. Acesso em: 9 set. 2026.

ASSIS, E. K. et al. Information system for the improvement of telecommunication service quality. In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO (SBSI), 13., 2017, Lavras. *Anais [...]*. Porto Alegre: Sociedade Brasileira de Computação, 2017. p. 452–455.

CARDOSO, I. L. *Modelo integrado para melhoria da qualidade em serviços de telecomunicações*. 2024. Projeto de Graduação (Engenharia de Produção) – Universidade de Brasília, Brasília, 2024. Disponível em: <https://bdm.unb.br/handle/10483/41311>. Acesso em: 6 out. 2026.

CHANDOLA, V.; BANERJEE, A.; KUMAR, V. Anomaly detection: a survey. *ACM Computing Surveys*, v. 41, n. 3, p. 1–58, 2009.

DAVIES, D. L.; BOULDIN, D. W. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1, n. 2, p. 224–227, 1979.

ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 2., 1996, Portland. *Proceedings [...]*. Menlo Park: AAAI Press, 1996. p. 226–231.

FAED, A. et al. Intelligent customer complaint handling utilising principal component and data envelopment analysis (PDA). *Applied Soft Computing*, v. 47, p. 614–630, 2016.

FREITAS, M. S.; ANDREÃO, R. V. Automatização do processamento do texto bruto oriundo de um serviço de atendimento de reclamações. In: ESCOLA REGIONAL DE INFORMÁTICA DO RIO DE JANEIRO (ERI-RJ), 4., 2021, Online. *Anais [...]*. Porto Alegre: Sociedade Brasileira de Computação, 2021. p. 72–79.

JAIN, A. K. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, v. 31, n. 8, p. 651–666, 2010.

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A*, v. 374, n. 2065, p. 20150202, 2016.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY, 5., 1967, Berkeley. *Proceedings [...]*. Berkeley: University of California Press, 1967. v. 1, p. 281–297.

OLIVEIRA, L. D. M. et al. A voz do consumidor no setor de telecomunicações: um estudo das reclamações no consumidor.gov.br. *Revista de Gestão e Secretariado*, v. 16, n. 2, p. 1–29, 2025.

PEDREGOSA, F. et al. Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.

ROUSSEEUW, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, v. 20, p. 53–65, 1987.

SCHUBERT, E. et al. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems*, v. 42, n. 3, p. 1–21, 2017.

UBEDA, G. N. et al. Avaliação automática da qualidade das respostas para reclamações no setor de telefonia no Brasil. In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO (SBSI), 21., 2025, Recife. *Anais estendidos [...]*. Porto Alegre: Sociedade Brasileira de Computação, 2025. p. 147–151.

Geração automática de equações de estado a partir de Bond Graphs: protótipo e validação numérica

***Emelyn Clementino Freire*[1,2]**
***Arthur Phelipe de Oliveira Queiroz*[2]**
***Jaysa Keylla Siqueira Barbosa*[3]**
***Guilherme Souza de Farias*[1,2]**
***Sayonara Cristina de O. Magalhães*[2]**
***André Laurindo Maitelli*[2]**

[1] Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PPgEEC), Centro de Tecnologia, Universidade Federal do Rio Grande do Norte (UFRN), Natal, RN, Brasil
[2] Departamento de Engenharia de Computação e Automação (DCA), UFRN, Natal, RN, Brasil
[3] Programa de Pós-Graduação em Sistemas e Computação (PPgSC), UFRN, Natal, RN, Brasil
<emelyn.freire.116@ufrn.edu.br>, <arthur.queiroz.115@ufrn.edu.br>, <jaysa.barbosa.108@ufrn.edu.br>, <guilhermefariias1995@gmail.com>, <sayonara.magalhaes.068@ufrn.edu.br>, <maitelli@dca.ufrn.br>

Resumo

A modelagem por Bond Graph representa sistemas dinâmicos de diferentes domínios físicos com um único conjunto de elementos, mas a passagem do diagrama às equações de estado costuma ser feita à mão ou em ferramentas comerciais. Este trabalho apresenta um protótipo em Python que recebe a descrição de um sistema (em JSON ou em texto simples em português), monta o Bond Graph, obtém simbolicamente as equações de estado, integra-as numericamente e verifica o resultado por autovalores e balanço de energia. As equações são obtidas resolvendo, de forma simbólica, o sistema algébrico formado pelas relações constitutivas dos elementos e pelas restrições das junções, o que também permite detectar causalidade derivativa. O protótipo foi avaliado em quatro sistemas lineares com solução analítica conhecida (circuitos RC e RLC, sistema massa-mola-amortecedor e motor de corrente contínua). O erro máximo relativo ficou entre $3{,}1\times10^{-12}$ e $4{,}3\times10^{-8}$; as ordens de convergência observadas para os métodos de Euler (1,00 a 1,08) e Runge-Kutta de quarta ordem (4,00 a 4,06) coincidiram com as teóricas; e seis integradores adaptativos do SciPy mantiveram o erro abaixo de 10^{-6}. Os resultados ainda não abrangem sistemas não lineares nem modelos de maior porte. Os experimentos são reproduzíveis em contêiner Docker com versões fixadas.

Abstract

Bond graph modelling represents dynamic systems from different physical domains with a single set of elements, but the step from the diagram to the state equations is usually done by hand or in commercial tools. This paper presents a Python prototype that takes a system description (JSON or plain Portuguese text), builds the bond graph, derives the state equations symbolically, integrates them numerically and checks the result through eigenvalues and an energy balance. The equations are obtained by symbolically solving the algebraic system formed by the element constitutive laws and the junction constraints, which also allows derivative causality to be detected. The prototype was evaluated on four linear systems with known analytical solutions (RC

and RLC circuits, a mass-spring-damper and a DC motor). The maximum relative error ranged from 3.1×10^{-12} to 4.3×10^{-8}; the observed convergence orders of the Euler (1.00 to 1.08) and fourth-order Runge-Kutta (4.00 to 4.06) methods matched the theoretical ones; and six adaptive SciPy integrators kept the error below 10^{-6}. The results do not yet cover nonlinear systems or larger models. The experiments can be reproduced in a Docker container with pinned versions.
Palavras-chave: Bond Graph; modelagem de sistemas dinâmicos; simulação numérica; equações de estado; Python.

1 Introdução

O Bond Graph, proposto por Paynter (1961), descreve um sistema físico pela troca de potência entre seus componentes. Cada ligação (*bond*) carrega um par de variáveis, esforço e fluxo, cujo produto é a potência, e um mesmo conjunto de elementos representa resistores, capacitores, massas, molas ou amortecedores. Por isso a técnica se aplica a sistemas que combinam domínios, como os eletromecânicos, sem trocar de formalismo (KARNOPP; MARGOLIS; ROSENBERG, 2012; BORUTZKY, 2010; GAWTHROP; BEVAN, 2007).

Depois de desenhado o diagrama, obter as equações de estado exige atribuir causalidade e combinar as relações de cada elemento, etapa sujeita a erros quando feita manualmente. Ferramentas como o 20-sim (BROENINK, 1999) automatizam esse processo em ambiente comercial. Entre as alternativas abertas, a biblioteca BondGraphTools, em Python, constrói modelos por composição simbólica com aplicações descritas sobretudo em biologia de sistemas (CUDMORE et al., 2019, 2021).

O protótipo descrito aqui não se propõe a substituir essas ferramentas. O 20-sim integra a entrada e o processamento de modelos hierárquicos, que combinam Bond Graphs e diagramas de blocos, verifica a consistência do modelo durante sua edição e dispõe de simulador próprio (BROENINK, 1999). O BondGraphTools é voltado à modelagem por script de sistemas multifísicos complexos e foi concebido para uso em conjunto com as demais bibliotecas científicas do Python (CUDMORE et al., 2019). O foco do protótipo está em três aspectos. Na entrada, o sistema pode ser descrito em alto nível, por uma árvore de ligações em série e em paralelo ou por um texto curto em português, e o Bond Graph é montado a partir dessa descrição. Na obtenção das equações, o sistema algébrico completo é resolvido simbolicamente, sem etapa explícita de atribuição de causalidade. Na verificação, um diagnóstico automático, por autovalores e balanço de energia, examina o resultado de cada simulação, e não apenas o modelo, e o conjunto é avaliado por um protocolo reproduzível contra soluções analíticas.

Este trabalho descreve um protótipo de simulador, também em Python, que cobre a cadeia completa: descrição do sistema, montagem do Bond Graph, geração simbólica das equações de estado, integração numérica e verificação automática do resultado. Nesta etapa, o objetivo é verificar, com métricas objetivas, se essa cadeia produz resultados corretos. As contribuições são: (i) a formulação da geração de equações como solução simbólica de um sistema algébrico linear, que dispensa a atribuição explícita de causalidade e sinaliza causalidade derivativa; (ii) um diagnóstico automático baseado em autovalores e em balanço de energia; e (iii) um protocolo experimental reproduzível, com comparação contra soluções analíticas e estudo de convergência.

2 Fundamentos

Em um Bond Graph, cada bond i tem esforço e_i e fluxo f_i, com potência $P_i = e_i f_i$. A meia seta indica o sentido em que a potência é considerada positiva. No domínio elétrico, esforço e fluxo correspondem a tensão e corrente; no mecânico translacional, a força e velocidade.

Os elementos de uma porta são as fontes de esforço (S_e) e de fluxo (S_f), o dissipador (R) e os armazenadores capacitivo (C) e inercial (I). Os armazenadores definem as variáveis de estado: o deslocamento generalizado q em cada C (carga, deslocamento) e o momento generalizado p em cada I (fluxo magnético, quantidade de movimento). Os elementos de duas portas, transformador (TF) e girador (GY), acoplam bonds sem armazenar nem dissipar energia. As junções impõem as leis de conservação: na junção 1, todos os bonds têm o mesmo fluxo e a soma algébrica dos esforços é nula (ligação em série); na junção 0, todos têm o mesmo esforço e a soma algébrica dos fluxos é nula (ligação em paralelo). A Tabela 1 resume as relações usadas pelo protótipo.

3 Protótipo

3.1 Arquitetura

O protótipo foi implementado em Python com NetworkX (HAGBERG; SCHULT; SWART, 2008), SymPy (MEURER et al., 2017) e SciPy (VIRTA-

Tabela 1: *Relações constitutivas implementadas.*

Elemento	Relação	Exemplos
S_e	$e = u(t)$	fonte de tensão; força
S_f	$f = u(t)$	fonte de corrente
R	$e = R\,f$	resistor; amortecedor
C	$e = q/C,\ \dot{q} = f$	capacitor; mola ($C = 1/k$)
I	$f = p/I,\ \dot{p} = e$	indutor; massa
TF	$e_1 = n\,e_2,\ f_2 = n\,f_1$	transformador
GY	$e_1 = r\,f_2,\ e_2 = r\,f_1$	motor CC
0	e comum, $\sum \pm f = 0$	ligação em paralelo
1	f comum, $\sum \pm e = 0$	ligação em série

NEN et al., 2020), organizado nos módulos da Figura 1. A entrada pode ser um JSON com elementos e bonds explícitos, um JSON em forma de árvore série/paralelo ou um texto curto em português, como "um resistor de 100 ohms em série com um capacitor de 1 µF conectado a uma fonte de tensão de 10 V". O texto é interpretado por regras (palavras-chave, valores com prefixos de unidade e as expressões "em série" e "em paralelo"), e o programa exibe o que entendeu para conferência. Essa interpretação cobre um vocabulário restrito (resistor, capacitor, indutor, diodo, fontes de tensão e de corrente e, no domínio mecânico, massa, mola, amortecedor e força aplicados a um mesmo corpo) e topologias formadas apenas por ligações em série e em paralelo; descrições fora desse padrão podem ser interpretadas de forma incorreta. A taxa de acerto da interpretação não foi avaliada neste trabalho, e todos os experimentos relatados a seguir usaram descrições em JSON. Na conversão de árvores, ligações em série viram junções 1 e ligações em paralelo, junções 0. O Bond Graph é armazenado como grafo direcionado, em que o sentido de cada aresta é o sentido positivo da potência.

3.2 Geração das equações de estado

Para um Bond Graph com N bonds há $2N$ incógnitas, os esforços e fluxos de todos os bonds. Cada extremidade de bond contribui com exatamente uma equação: um elemento de uma porta fornece sua relação constitutiva, uma junção com k bonds fornece $k-1$ igualdades e uma soma, e TF e GY fornecem duas relações. O sistema resultante tem, portanto, $2N$ equações. Nas junções, o sinal de cada termo das somas é positivo quando o bond entra na junção.

Com os estados $\mathbf{x}$ (os q e p) e as entradas $\mathbf{u}(t)$ tratados como conhecidos, as equações formam um sistema linear $\mathbf{A}\,\mathbf{z} = \mathbf{b}(\mathbf{x}, \mathbf{u})$ nas incógnitas $\mathbf{z}$, cujos coeficientes dependem dos parâmetros. O protótipo resolve esse sistema simbolicamente com o SymPy e substitui a solução em $\dot{q} = f$ e $\dot{p} = e$, obtendo $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})$ com os parâmetros $\boldsymbol{\theta}$ ainda simbólicos. Assim, não é preciso atribuir causalidade explicitamente; em contrapartida, o custo cresce com o tamanho do sistema simbólico, o que não foi avaliado para modelos grandes. Quando o sistema não tem solução única, há estados dependentes, situação que no Bond Graph corresponde à causalidade derivativa (por exemplo, fonte de tensão ideal em paralelo com capacitor). Nesse caso a simulação é interrompida com uma mensagem que aponta as causas típicas.

O motor de corrente contínua usado nos experimentos é controlado pela armadura (Figura 2a): com a corrente de campo i_f constante, a tensão de armadura v_a (denotada V no modelo) alimenta o circuito formado por L_a e R_a, e o rotor, de momento de inércia J, está sujeito a atrito viscoso de coeficiente b. A força contraeletromotriz e o torque são proporcionais, respectivamente, à velocidade angular e à corrente de armadura, com a mesma constante K_m (MEDEIROS; MAITELLI, 2010). No Bond Graph da Figura 2b, essa relação é representada pelo girador GY, que acopla o domínio elétrico ao mecânico. As equações geradas pelo protótipo para esse modelo foram

$$\dot{p}_{L_a} = V - \frac{R_a}{L_a}\,p_{L_a} - \frac{K_m}{J}\,p_J, \tag{1}$$

$$\dot{p}_J = \frac{K_m}{L_a}\,p_{L_a} - \frac{b}{J}\,p_J, \tag{2}$$

em que p_{L_a}/L_a é a corrente de armadura e p_J/J a velocidade angular, as mesmas equações obtidas pelas leis de Kirchhoff e de Newton.

Resistores não lineares também podem ser declarados, como o diodo descrito pelo modelo de Shockley (1949). Nesse caso a parte linear do sistema é resolvida simbolicamente, e as poucas incógnitas restantes são calculadas numericamente a cada avaliação de $\mathbf{F}$. Essa funcionalidade não foi validada quantitativamente neste trabalho.

3.3 Integração e diagnóstico

As equações são convertidas em funções numéricas e integradas pelos métodos adaptativos do SciPy, entre eles o Runge-Kutta de Dormand-Prince (RK45) (DORMAND; PRINCE, 1980), os implícitos Radau e BDF (HAIRER; WANNER, 1996) e o LSODA, que alterna entre métodos rígidos e não rígidos (PETZOLD, 1983). Para estudo de convergência, foram implementados ainda os métodos de passo fixo de Euler explícito e Runge-Kutta clássico de quarta ordem (RK4) (HAIRER; NØRSETT; WANNER, 1993).

Após cada simulação, o protótipo calcula os autovalores da jacobiana (por diferenças finitas), dos

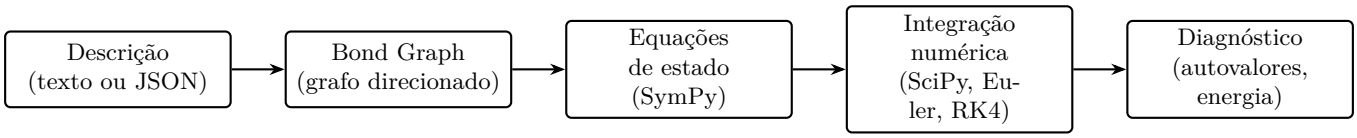


Figura 1: *Etapas do protótipo, da descrição do sistema ao diagnóstico do resultado.*

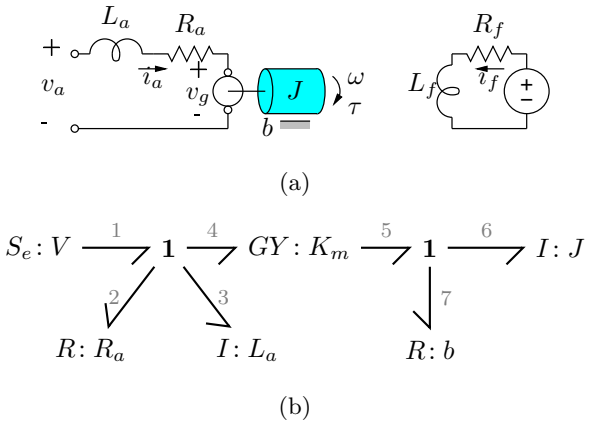


Figura 2: *Motor CC controlado pela armadura: (a) diagrama esquemático, adaptado de Medeiros e Maitelli (2010); (b) Bond Graph correspondente, com a parte elétrica à esquerda do girador e a mecânica à direita. Os números identificam os bonds.*

quais estima constantes de tempo, frequências naturais, fatores de amortecimento e a razão de rigidez, e emite avisos quando o tempo final ou o passo parecem inadequados. Calcula também o balanço de energia, comparando a energia fornecida pelas fontes, E_f, com a dissipada nos elementos R, E_d, e a variação da energia armazenada, $\Delta E_a = \Delta(\sum_C q^2/2C + \sum_I p^2/2I)$:

$$\varepsilon_E = \frac{|E_f - E_d - \Delta E_a|}{\max(|E_f|, |E_d|, |\Delta E_a|)}. \tag{3}$$

As energias E_f e E_d são integrais da potência no tempo, calculadas pela regra dos trapézios sobre os pontos de saída.

4 Metodologia experimental

Foram usados quatro sistemas lineares com solução analítica conhecida e condições iniciais nulas (Tabela 2). As soluções de referência foram escritas a partir dos parâmetros físicos, sem usar as equações geradas pelo protótipo: resposta ao degrau de primeira ordem para o RC, de segunda ordem subamortecida para o RLC e o massa-mola-amortecedor, e decomposição espectral da matriz do modelo do motor, montada a partir das leis de Kirchhoff e de Newton.

Tabela 2: *Sistemas avaliados, parâmetros, grandeza comparada e tempo final T.*

Sistema	Parâmetros	T
RC série (v_C)	$V = 10$ V, $R = 100\ \Omega$, $C = 1$ µF	0,5 ms
RLC série (v_C)	$V = 5$ V, $R = 10\ \Omega$, $L = 10$ mH, $C = 100$ µF	20 ms
Massa-mola-amortecedor (x)	$F = 10$ N, $m = 2$ kg, $k = 50$ N/m, $b = 3$ N·s/m	8 s
Motor CC (ω)	$V = 12$ V, $R_a = 1\ \Omega$, $L_a = 5$ mH, $K_m = 0{,}05$ N·m/A, $J = 2 \cdot 10^{-4}$ kg·m^2, $b = 10^{-4}$ N·m·s	0,4 s

Sendo y_k o valor simulado e y_k^* o analítico nos instantes de saída, adotaram-se o erro máximo relativo, $\varepsilon_{\max} = \max_k |y_k - y_k^*| / \max_k |y_k^*|$, e o NRMSE, raiz do erro quadrático médio dividida pela amplitude $\max_k y_k^* - \min_k y_k^*$. Três experimentos foram realizados:

E1 – precisão. Integração com RK45 (tolerâncias relativa 10^{-7} e absoluta 10^{-13}, passo máximo $T/200$) e 1000 pontos de saída; cálculo de $\varepsilon_{\max}$, NRMSE e ε_E.

E2 – convergência. Euler e RK4 com $N \in \{50, 100, \ldots, 3200\}$ passos ($h = T/N$). A ordem observada é a inclinação do ajuste por mínimos quadrados de $\log \varepsilon_{\max}$ contra $\log h$ nos quatro passos mais finos com $10^{-11} < \varepsilon_{\max} < 1$, o que exclui o piso de arredondamento e os casos instáveis.

E3 – métodos adaptativos. RK45, RK23, DOP853, Radau, BDF e LSODA com as mesmas tolerâncias; registro de $\varepsilon_{\max}$ e do número de avaliações de $\mathbf{F}$, incluindo as usadas na aproximação da jacobiana pelos métodos implícitos e no cálculo inicial dos autovalores.

Os resultados apresentados foram obtidos com Python 3.12, NumPy 2.5.3, SciPy 1.18.1 e SymPy 1.14.0. Os experimentos são executados por um único comando em contêiner Docker, sem acesso à rede, com uma CPU, versões fixadas e parâmetros lidos de um arquivo de configuração; cada execução registra o ambiente e uma assinatura SHA-256

Tabela 3: *E1: erro em relação à solução analítica e erro do balanço de energia (RK45).*

Sistema	$\varepsilon_{\max}$	NRMSE	ε_E
RC série	$3{,}13 \cdot 10^{-12}$	$5{,}82 \cdot 10^{-13}$	$2{,}12 \cdot 10^{-6}$
RLC série	$1{,}65 \cdot 10^{-9}$	$5{,}30 \cdot 10^{-10}$	$3{,}34 \cdot 10^{-5}$
Massa-mola-amort.	$4{,}27 \cdot 10^{-8}$	$1{,}72 \cdot 10^{-8}$	$1{,}34 \cdot 10^{-4}$
Motor CC	$1{,}48 \cdot 10^{-9}$	$2{,}10 \cdot 10^{-10}$	$3{,}00 \cdot 10^{-5}$

Tabela 4: *E2: ordem de convergência observada (teórica: 1 para Euler e 4 para RK4).*

Sistema	Euler	RK4
RC série	1,00	4,03
RLC série	1,01	4,01
Massa-mola-amortecedor	1,08	4,00
Motor CC	1,00	4,06

dos resultados numéricos, que permite verificar se duas execuções produziram os mesmos valores. O código-fonte será disponibilizado em repositório público após a avaliação do trabalho.

5 Resultados e discussão

5.1 Precisão

A Tabela 3 mostra que o erro máximo relativo ficou entre $3{,}1 \cdot 10^{-12}$ e $4{,}3 \cdot 10^{-8}$, valores compatíveis com a tolerância relativa de 10^{-7} do integrador. O maior erro ocorreu no massa-mola-amortecedor, possivelmente porque é o único caso em que a simulação cobre várias oscilações pouco amortecidas (fator de amortecimento 0,15). Como os quatro sistemas foram descritos com o mesmo conjunto de elementos, inclusive o motor, que combina dois domínios, o resultado indica que, nesses casos, as equações geradas automaticamente coincidem com as obtidas pelas leis físicas de cada domínio.

O erro do balanço de energia, entre $2 \cdot 10^{-6}$ e $1{,}3 \cdot 10^{-4}$, é várias ordens de grandeza maior que o erro dos estados. A diferença vem da quadratura: ao repetir as simulações com 10 000 pontos de saída em vez de 1000, ε_E caiu cerca de cem vezes nos casos RC, massa-mola-amortecedor e motor, como se espera da regra dos trapézios, cujo erro decresce com o quadrado do intervalo de amostragem. O balanço serve, assim, como verificação de consistência física, mas sua sensibilidade depende do número de pontos de saída.

5.2 Convergência

A Figura 3 e a Tabela 4 mostram que as ordens observadas coincidem com as teóricas: de 1,00 a 1,08 para Euler e de 4,00 a 4,06 para RK4. No massa-mola-amortecedor, Euler diverge nos dois maiores passos ($h = 0{,}16$ s e 0,08 s). Para os autovalores $\lambda = -0{,}75 \pm 4{,}94j$ desse sistema, $|1 + h\lambda|$ vale 1,18 e 1,02 nesses passos, fora da região de estabilidade do método, e 0,99 para $h = 0{,}04$ s. Esses pontos ficaram fora do ajuste da ordem. Para RK4 no circuito RC, o erro no passo mais fino ($2 \cdot 10^{-14}$) já se aproxima do piso de arredondamento.

Esse resultado verifica a implementação dos integradores de passo fixo e, de forma independente do E1, a coerência entre as equações geradas e as soluções analíticas: se houvesse erro nas equações, o erro numérico convergiria para um valor não nulo, e não para zero na taxa prevista.

5.3 Métodos adaptativos

Na Tabela 5, todos os métodos mantiveram $\varepsilon_{\max}$ abaixo de 10^{-6} (o maior valor, $6{,}3 \cdot 10^{-7}$, ocorreu com BDF no massa-mola-amortecedor). O DOP853 obteve os menores erros, de $5 \cdot 10^{-16}$ a $3 \cdot 10^{-11}$, com cerca de 3000 avaliações em todos os casos, e o LSODA usou o menor número de avaliações (366 a 1124). Nenhum dos quatro sistemas é rígido (a maior razão entre autovalores, no motor, é de cerca de 13). Nessas condições, o Radau exigiu mais avaliações que o RK45 em todos os casos, e o BDF usou menos avaliações, porém com erros maiores; o RK23, de ordem mais baixa, chegou a 8495 avaliações no massa-mola-amortecedor. O número de avaliações não depende da máquina e por isso foi preferido ao tempo de execução como medida de custo.

6 Limitações e trabalhos futuros

Os resultados se restringem a quatro sistemas lineares com solução analítica conhecida, com no máximo dois estados e sete bonds. Eles não validam, portanto, o tratamento de sistemas não lineares nem de modelos de maior complexidade. Ficam para trabalhos futuros: a comparação de casos não lineares, como o retificador com diodo, com ferramentas de referência; a avaliação do custo da solução simbólica em modelos com muitos bonds; a medição da taxa de acerto da entrada em texto sobre um conjunto de descrições; e o tratamento da causalidade derivativa, hoje apenas detectada. O protótipo inclui ainda módulos preliminares de ajuste de parâmetros a partir de dados medidos e de representação de elementos não lineares por redes neurais, testados apenas com dados sintéticos, que motivam uma linha de pesquisa

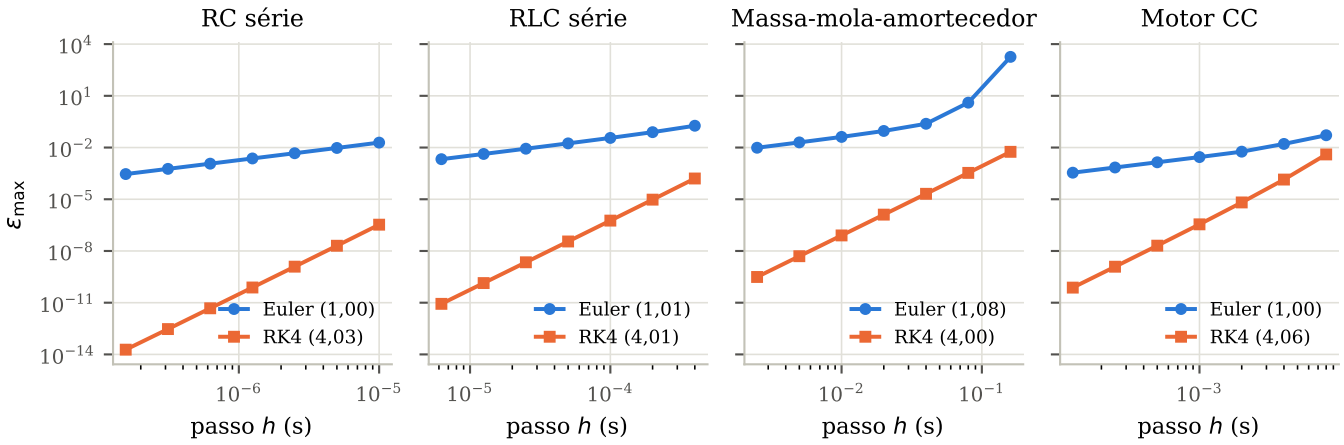


Figura 3: *E2: erro máximo relativo em função do passo para Euler e RK4. Entre parênteses, a ordem observada (inclinação ajustada nos quatro passos mais finos).*

Tabela 5: *E3: erro máximo relativo $\varepsilon_{\max}$ e número de avaliações de* **F** *(entre parênteses) para os métodos adaptativos do SciPy.*

Método	RC série	RLC série	Massa-mola-amortecedor	Motor CC
RK45	$3{,}13 \cdot 10^{-12}$ (1204)	$1{,}65 \cdot 10^{-9}$ (1235)	$4{,}27 \cdot 10^{-8}$ (1679)	$1{,}48 \cdot 10^{-9}$ (1307)
RK23	$9{,}11 \cdot 10^{-8}$ (730)	$7{,}52 \cdot 10^{-8}$ (3494)	$2{,}88 \cdot 10^{-7}$ (8495)	$4{,}65 \cdot 10^{-8}$ (1640)
DOP853	$5{,}37 \cdot 10^{-16}$ (3004)	$1{,}01 \cdot 10^{-14}$ (3020)	$8{,}91 \cdot 10^{-13}$ (3047)	$2{,}66 \cdot 10^{-11}$ (3035)
Radau	$4{,}11 \cdot 10^{-10}$ (1409)	$7{,}73 \cdot 10^{-9}$ (2380)	$1{,}16 \cdot 10^{-8}$ (5566)	$8{,}01 \cdot 10^{-10}$ (1902)
BDF	$3{,}85 \cdot 10^{-8}$ (445)	$1{,}40 \cdot 10^{-7}$ (686)	$6{,}30 \cdot 10^{-7}$ (1407)	$1{,}07 \cdot 10^{-8}$ (703)
LSODA	$1{,}35 \cdot 10^{-8}$ (404)	$5{,}74 \cdot 10^{-9}$ (477)	$1{,}11 \cdot 10^{-7}$ (1124)	$3{,}69 \cdot 10^{-8}$ (366)

sobre a integração de técnicas de aprendizado de máquina à modelagem por Bond Graph. Um editor gráfico e a avaliação do uso da ferramenta no ensino também estão previstos.

7 Conclusões

Este trabalho apresentou um protótipo que gera simbolicamente as equações de estado a partir de um Bond Graph e as integra numericamente. Nos quatro sistemas lineares avaliados, que incluem um modelo eletromecânico, o erro em relação às soluções analíticas ficou no nível da tolerância do integrador, as ordens de convergência de Euler e RK4 coincidiram com as teóricas e o balanço de energia fechou dentro do erro da quadratura. Esses resultados verificam a cadeia de geração e simulação apenas para sistemas lineares de pequeno porte. A validação com modelos não lineares e de maior complexidade é o passo seguinte, entre as extensões listadas na seção anterior.

Declaração de uso de inteligência artificial

Os autores utilizaram o assistente de inteligência artificial Claude (Anthropic) como apoio na implementação do protótipo, na preparação dos scripts dos experimentos e na redação deste texto. Todo o conteúdo gerado foi revisado e editado pelos autores, que assumem total responsabilidade pelo trabalho.

Referências

BORUTZKY, W. *Bond Graph Methodology: Development and Analysis of Multidisciplinary Dynamic System Models*. London: Springer, 2010.

BROENINK, J. F. 20-sim software for hierarchical bond-graph/block-diagram models. *Simulation Practice and Theory*, v. 7, n. 5–6, p. 481–492, 1999.

CUDMORE, P. et al. *Computer-aided modelling of complex physical systems with BondGraphTools*. 2019. ArXiv:1906.10799.

CUDMORE, P. et al. Analysing and simulating energy-based models in biology using BondGraphTools. *The European Physical Journal E*, v. 44, p. 148, 2021.

DORMAND, J. R.; PRINCE, P. J. A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, v. 6, n. 1, p. 19–26, 1980.

GAWTHROP, P. J.; BEVAN, G. P. Bond-graph modeling: a tutorial introduction for control engineers. *IEEE Control Systems Magazine*, v. 27, n. 2, p. 24–45, 2007.

HAGBERG, A. A.; SCHULT, D. A.; SWART, P. J. Exploring network structure, dynamics, and function using NetworkX. In: *Proceedings of the 7th Python in Science Conference (SciPy 2008)*. Pasadena, CA: [s.n.], 2008. p. 11–15.

HAIRER, E.; NØRSETT, S. P.; WANNER, G. *Solving Ordinary Differential Equations I: Nonstiff Problems*. 2. ed. Berlin: Springer, 1993.

HAIRER, E.; WANNER, G. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. 2. ed. Berlin: Springer, 1996.

KARNOPP, D. C.; MARGOLIS, D. L.; ROSENBERG, R. C. *System Dynamics: Modeling, Simulation, and Control of Mechatronic Systems*. 5. ed. Hoboken, NJ: John Wiley & Sons, 2012.

MEDEIROS, A. A. D.; MAITELLI, A. L. *Modelagem e Análise de Sistemas Dinâmicos: material didático*. Natal: Universidade Federal do Rio Grande do Norte, Departamento de Engenharia de Computação e Automação, 2010.

MEURER, A. et al. SymPy: symbolic computing in Python. *PeerJ Computer Science*, v. 3, p. e103, 2017.

PAYNTER, H. M. *Analysis and Design of Engineering Systems*. Cambridge, MA: MIT Press, 1961.

PETZOLD, L. Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. *SIAM Journal on Scientific and Statistical Computing*, v. 4, n. 1, p. 136–148, 1983.

SHOCKLEY, W. The theory of p-n junctions in semiconductors and p-n junction transistors. *Bell System Technical Journal*, v. 28, n. 3, p. 435–489, 1949.

VIRTANEN, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, v. 17, p. 261–272, 2020.

Biblioteca para Cálculo de Conjuntos Invariantes com *Julia*

Nicholas W. Lopes[1]
Felipe R. Silva[2]
Carlos E. T. Dórea[3]

[1] *Departamento de Engenharia de Computação, UFRN, <wallace.lopes.162@ufrn.edu.br>*

[2] *Departamento de Engenharia de Computação, UFRN, <felipe.rocha.702@ufrn.edu.br>*

[3] *Departamento de Engenharia de Computação, UFRN, <cetdorea@dca.ufrn.br>*

Resumo

Este artigo apresenta o desenvolvimento de uma biblioteca em *Julia* para o cálculo e a análise de conjuntos invariantes de sistemas dinâmicos lineares discretos. A ferramenta permite definir sistemas a partir de modelos em espaço de estados, simular trajetórias, verificar a invariância de conjuntos poliédricos considerando restrições e buscar conjuntos que satisfaçam as propriedades de invariância e contração. A implementação utiliza o Lema de Farkas para formular condições matemáticas e integra as bibliotecas JuMP, Polyhedra e CDDlib para a modelagem de problemas de otimização e a manipulação de poliedros. A proposta busca centralizar, em um único ambiente computacional, etapas anteriormente realizadas com diferentes ferramentas. Por meio de exemplos, são apresentadas as funcionalidades implementadas e sua aplicação a problemas de invariância, incluindo sistemas com atraso na entrada de controle.

Abstract

This paper presents the development of a *Julia* library for the computation and analysis of invariant sets for discrete-time linear dynamical systems. The library allows users to define systems using state-space models, simulate trajectories, verify the invariance of polyhedral sets under constraints, and search for sets satisfying invariance and contraction properties. The implementation uses Farkas' lemma to formulate mathematical conditions and integrates the JuMP, Polyhedra, and CDDlib libraries for optimization modeling and polyhedral manipulation. The proposed approach aims to consolidate, within a single computational environment, tasks that were previously performed using different tools. Examples illustrate the implemented functionalities and their application to invariant-set problems, including systems with input delays.

1 Introdução

O estudo da estabilidade de sistemas dinâmicos é fundamental para a teoria de controle, uma vez que permite analisar o comportamento das trajetórias do sistema ao longo do tempo e estabelecer condições para sua operação segura e previsível. Nesse contexto, a abordagem de conjuntos invariantes permite caracterizar regiões do espaço de estados nas quais as trajetórias permanecem confinadas sob a dinâmica do sistema. Essas regiões são definidas por hiperplanos de conjuntos convexos, em particular, esse trabalho enfatiza os conjuntos hiperpoliédricos (*hyperpolyhedral sets*), que despertam interesse devido

à sua capacidade de representar restrições sobre as variáveis de estado e de entrada. O cálculo dessas regiões, entretanto, pode exigir procedimentos computacionais específicos, motivando o desenvolvimento de ferramentas para sua determinação.

Na literatura, diferentes formulações de problemas de otimização têm sido propostas para a determinação de conjuntos poliédricos invariantes. Em trabalhos anteriores (DóREA; OLARU; NICULESCU, 2023), a modelagem dos problemas era realizada por meio da linguagem AMPL (*A Mathematical Programming Language*) (FOURER; GAY; KERNIGHAN, 1990), e sua solução era obtida com o auxílio de solvers de otimização como o Knitro (BYRD; NOCEDAL; WALTZ, 2006). Posteriormente, os resultados eram processados no MATLAB (The MathWorks, Inc.,), utilizando ferramentas específicas para manipulação de poliedros. Embora essa abordagem permita combinar recursos especializados, ela distribui as etapas do procedimento entre diferentes ambientes computacionais, podendo aumentar a complexidade de implementação e manutenção. Nesse contexto, a utilização de um único ambiente de programação surge como uma alternativa para integrar a modelagem, a resolução dos problemas de otimização e as simulações com os poliedros resultantes.

Com esse objetivo, este trabalho utiliza *Julia* (BEZANSON et al., 2017), uma linguagem de programação de código aberto que apresenta recursos particularmente adequados à computação científica e à otimização matemática. Seu ecossistema dispõe de bibliotecas que atendem às diferentes etapas do cálculo de conjuntos invariantes. Entre elas, destaca-se JuMP (LUBIN et al., 2023), utilizada para a modelagem de problemas de otimização, e Polyhedra (LEGAT, 2023), que fornece uma interface para a representação e a manipulação de poliedros. A integração dessas bibliotecas permite explorar recursos já disponíveis no ecossistema *Julia* e reduzir a necessidade de implementar, do zero, funcionalidades matemáticas consolidadas.

Diante disso, este trabalho apresenta o desenvolvimento de um pacote (*package*) em *Julia* para auxiliar o cálculo de conjuntos invariantes, integrando bibliotecas existentes a funções desenvolvidas especificamente para o problema considerado. A implementação busca centralizar as etapas computacionais envolvidas no procedimento e oferecer uma estrutura modular, de código aberto, que facilite a manutenção, a extensão das funcionalidades e a contribuição de outros desenvolvedores. Dessa forma, pretende-se disponibilizar uma ferramenta que apoie estudos sobre conjuntos invariantes e favoreça o desenvolvimento de novas funcionalidades nesse contexto.

2 Metodologia

2.1 Fundamentação Teórica

Inicialmente, definimos o escopo e a estrutura matemática adotados neste trabalho. Consideram-se sistemas dinâmicos discretos no tempo, descritos na representação em espaço de estados por:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k), \\ y(k) &= Cx(k) \end{aligned} \tag{1}$$

em que $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}$ e $y(k) \in \mathbb{R}$ representam, respectivamente, o vetor de estados, a entrada de controle e a saída do sistema. As matrizes A, B e C descrevem a dinâmica dos estados, a influência da entrada de controle e a relação entre os estados e a saída.

Um conjunto $S \subseteq \mathbb{R}^n$ é dito positivamente invariante em relação ao sistema definido em 1 se, para qualquer condição inicial $x(0) \in S$, a trajetória do sistema permanecer em S para todo instante $k \geq 0$, isto é,

$$x(0) \in S \Rightarrow x(k) \in S, \quad \forall k \geq 0.$$

Essa propriedade garante que, uma vez que o estado do sistema pertença a S, ele permanecerá nessa região ao longo de toda a evolução temporal. Uma propriedade mais restritiva pode ser estabelecida por meio da definição de conjuntos λ-contrativos positivamente invariantes, para os quais existe um fator de contração $0 < \lambda < 1$ tal que

$$x(k) \in S \Rightarrow x(k+1) \in \lambda S. \tag{2}$$

Essa condição garante que o $x(k)$ converge para a origem quando $k \to \infty$, desde que S seja limitado e fechado. Assim, além de assegurar a permanência da trajetória em uma região, a propriedade de contração permite estabelecer resultados de convergência assintótica.

No contexto de problemas de controle com restrições, é comum que os limites impostos às variáveis de estado e de entrada sejam expressos por meio de desigualdades lineares. Essas restrições podem ser representadas geometricamente por conjuntos poliédricos convexos. Em particular, um conjunto de restrições no espaço de estados pode ser descrito na forma

$$\Omega = \{x \in \mathbb{R}^n : Fx \leq w\}, \quad F \in \mathbb{R}^{f \times n}, \quad w \in \mathbb{R}^f.$$

A representação poliédrica é especialmente interessante porque permite transformar determinadas propriedades geométricas em condições algébricas verificáveis. Em particular, o Lema de Farkas Estendido fornece ferramentas para estabelecer condições de inclusão entre poliedros. Essa possibilidade é fundamental para o estudo de conjuntos invariantes, pois permite formular condições suficientes para garantir que a evolução do sistema permaneça dentro das restrições impostas.

Dessa forma, a combinação da representação poliédrica com ferramentas de verificação algébrica fornece uma base sistemática para o estudo de sistemas de controle sujeitos a restrições. Mais a diante abordaremos como essas definições são utilizadas em projetos de controladores baseados em problemas de otimização com restrições.

Para verificar inclusão de poliedros, utilizamos o Lema de Farkas Estendido. a implicação $F_2x \leq w_2 \Rightarrow F_1x \leq w_1$ é válida se existir uma matriz D não negativa tal que:

$$\begin{aligned} DF_2 &= DF_1, \\ Dw_2 &\leq Dw_1. \end{aligned}$$

Ou seja, para $\Omega_1 = \{x : F_1x \leq w_1\}$ e $\Omega_2 = \{x : F_2x \leq w_2\}$, a existência de D garante $\Omega_2 \subset \Omega_1$. Isso é especialmente útil quando há restrições sobre os estados, pois, garantindo que o poliedro invariante esteja contido nelas, o sistema nunca violará as restrições para um estado dentro desse poliedro.

Mais adiante, o Lema de Farkas também é utilizado para garantir a invariância de poliedros, como já mencionado anteriormente. Para um sistema com dinâmica $x(k+1) = Ax(k)$, por exemplo, essa propriedade pode ser escrita como $Fx(k) \leq w \Rightarrow FAx(k) \leq w$. Essa expressão se assemelha bastante à inclusão de poliedros descrita acima, de forma que, para que o poliedro convexo Ω seja λ-contrativo e positivamente invariante, deve existir uma matriz H não negativa tal que:

$$\begin{aligned} HF &= FA, \\ Hw &\leq \lambda w. \end{aligned}$$

Essas condições são necessárias e suficientes para garantir a propriedade de invariância e podem ser adaptadas para outras dinâmicas de $x(k+1)$.

2.2 Design da Biblioteca

O paradigma de despacho múltiplo (*multiple dispatch*) é uma das principais características de *Julia*. Nesse paradigma, a seleção do método a ser executado depende dos tipos de todos os argumentos da chamada da função. Essa abordagem difere do mecanismo tradicional de despacho simples, comum em linguagens orientadas a objetos, no qual a seleção do método depende principalmente do tipo do objeto receptor. Dessa forma, *Julia* permite implementar padrões de programação orientada a objetos, além de oferecer outras possibilidades de organização e extensão do código. Mais informações sobre os benefícios do *multiple dispatch* podem ser encontradas na discussão apresentada em (The Julia Programming Language, 2020).

Diferentemente das linguagens orientadas a objetos tradicionais, *Julia* organiza seus dados por meio de tipos, incluindo tipos abstratos e concretos, que podem estabelecer relações hierárquicas de subtipagem. Nesse contexto, utilizamos estruturas definidas com a palavra-chave `struct` para estabelecer os tipos concretos que compõem a biblioteca, conforme representado na Figura 1. Essa organização busca tornar a arquitetura clara, modular e escalável, facilitando a manutenção do código e a incorporação de futuras contribuições.

O principal tipo definido é `InvariantPolyhedron`, que representa o objeto de estudo da biblioteca: um conjunto poliédrico λ-contrativo positivamente invariante, conforme definido na equação (2). A ideia é permitir que as funções relacionadas a esses conjuntos recebam objetos desse tipo. Para isso, seus campos incluem as matrizes F e w, que caracterizam o poliedro, o escalar λ, que representa o fator de contração, e o campo `system`, que identifica o sistema em relação ao qual o conjunto é invariante.

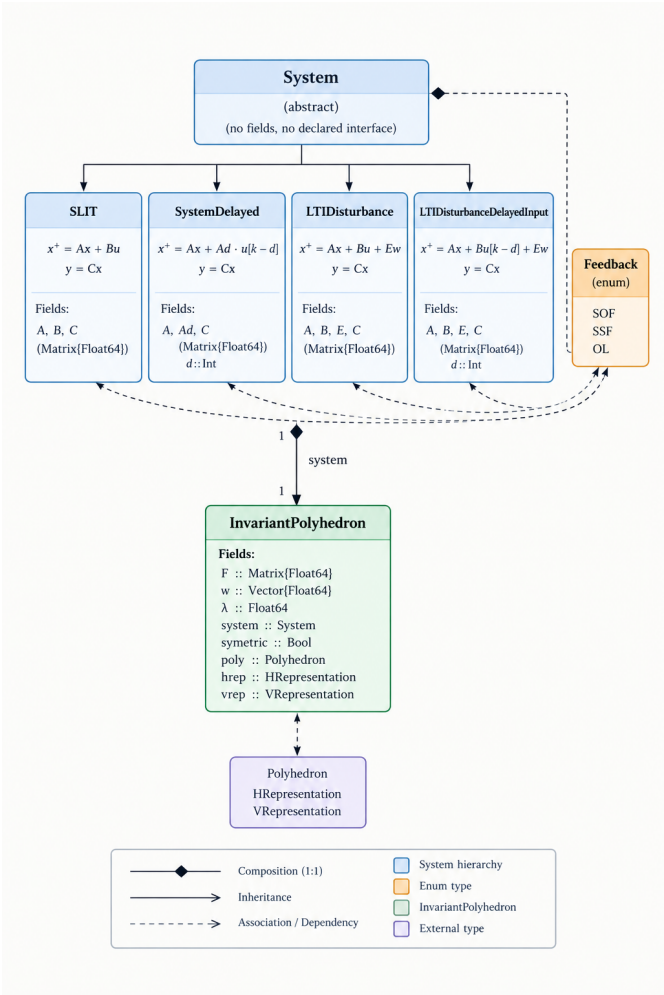


Figura 1: *Diagrama UML da estrutura de tipos da biblioteca*

Além desses campos, a estrutura armazena representações do poliedro fornecidas por uma biblioteca específica para computação poliédrica, que utiliza algoritmos da CDDlib (FUKUDA, 2022). Essa biblioteca permite trabalhar com representações por desigualdades lineares e por vértices e raios. Como cada representação pode ser mais adequada a determinadas operações, armazenamos ambas para evitar conversões recorrentes durante os cálculos.

Para representar os sistemas dinâmicos, definimos o tipo abstrato `System`, do qual derivam os tipos concretos que representam os sistemas suportados pela biblioteca. Além das matrizes da representação em espaço de estados, cada sistema possui uma matriz de ganhos G e um campo `feedback`, definido por um tipo enumerado (`@enum`), cujos valores correspondem às leis de controle descritas na Tabela 1.

Em geral, os ganhos do controlador não são conhecidos no momento da definição do sistema. Por esse motivo, disponibilizamos um construtor que recebe apenas as matrizes do modelo e inicializa o ganho com um valor unitário. Esse valor pode ser posteriormente atualizado por algoritmos responsáveis pelo cálculo dos ganhos do controlador.

Tipo	Lei de controle	Dinâmica em malha fechada
SOF	$u(k) = Gy(k) = GCx(k)$	$x(k+1) = (A + BGC)\,x(k)$
SSF	$u(k) = Gx(k)$	$x(k+1) = (A + BG)\,x(k)$

Tabela 1: *Tipos de realimentação e suas dinâmicas em malha fechada*

Para encontrar os ganhos do controlado um problema de otimização foi formulado por (BRIÃO et al., 2021), com as restrições dadas pelas condições do Lema de Farkas e como função objetivo o fator de contração λ a ser minimizado. Esse problema foi traduzido para o JuMP (LUBIN et al., 2023) por meio de uma função para cada tipo de sistema, adaptando as condições de invariância às dinâmicas específicas de cada um. Com isso, obteve-se uma função que recebe como entrada um conjunto poliédrico, um sistema de interesse e as restrições e que retorna, caso o conjunto seja invariante, o fator de contração associado. Caso não seja invariante, a função indica isso ao usuário.

Inicialmente, foi implementado um sistema linear invariante no tempo, descrito na equação (1). A partir desse modelo, foram consideradas extensões que contemplam um atraso na entrada de controle e a presença de uma perturbação. A mesma estrutura pode ser estendida para incorporar outros modelos dinâmicos, permitindo ampliar as funcionalidades da biblioteca de forma modular.

Vale ressaltar que os sistemas implementados até então foram sendo adicionados conforme necessário, pois essa biblioteca estava sendo desenvolvida para atacar o problema do segmento de referência para sistemas com atraso. Mais adiante abordaremos sobre funções específicas para esse propósito.

2.3 Verificação de Invariância Positiva

2.4 Determinação de Poliedros

E se não houver, de antemão, um poliedro candidato para testar a sua invariância? Nesse caso, pode-se utilizar o problema de otimização (BRIÃO et al.,

2021), mas agora considerando as matrizes que caracterizam o poliedro como variáveis do problema. Isso implica que o problema de otimização que antes era convexo passa a ser não convexo, tornando mais difícil a busca por um ótimo global.

Outra implicação é que passa a fazer sentido incorporar à função objetivo algum termo relativo ao tamanho do poliedro, pois, como mostrado por (DóREA; OLARU, 2022), há um compromisso entre o tamanho do poliedro e o fator de contração. Para tentar contornar isso, adiciona-se um fator de ponderação π à função objetivo, de forma que $obj = \pi\lambda - (1-\pi)\,\text{enl}(\cdot)$, em que $\text{enl}(\cdot)$ é uma função para maximizar o tamanho do poliedro.

Na primeira versão da biblioteca, foram implementadas duas funções $\text{enl}(\cdot)$. Na primeira, definem-se n vetores igualmente espaçados, cada um associado a um escalar γ_k, de forma a maximizar a soma desses escalares, como descreve a equação abaixo:

$$enl(.) = \frac{\sum_{k=1}^{n} \gamma_k}{n} \tag{3}$$

Além disso foi preciso adicionar mais n restrições ao problema para garantir que todos os vetores estejam contidos em Ω.

$$L * \gamma_k \mathbf{v}_k \leq \phi, k = 1, 2, ..., n \tag{4}$$

Para sistemas com apenas dois estados, a interpretação geométrica da estratégia pode ser ilustrada pela Figura 2.

Para a outra implementação da função $\text{enl}(\cdot)$, criamos um poliedro com o mesmo formato das restrições de estado e, com auxílio do *Lemma de Farkas*, adicionamos uma restrição no problema de otimização para que ele estivesse contido em Ω que queremos encontrar com a seguinte expressão.

$$\begin{aligned} Q\mathcal{S} &= F, \\ Q\mathbf{1} &\leq \gamma\mathbf{1} \end{aligned}$$

Dessa forma, para $\text{enl} = \gamma$, quando minimizamos γ estamos expandindo o poliedro no interior de Ω e consequentemente expandindo Ω também. Como pode ser visto na figura (3).

figure poliedros contidos

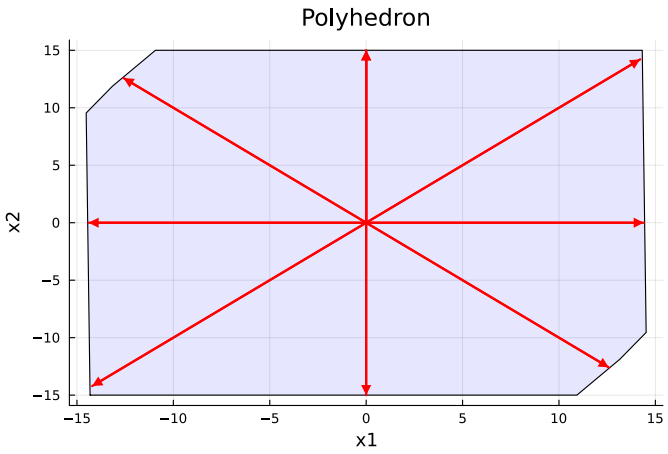


Figura 2: *Poliedro com os vetores utilizados para maximizar o seu tamanho*

2.5 Programas específicos para atraso

Considere o sistema a seguir:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k-d), \\ y(k) &= Cx(k) \end{aligned} \tag{5}$$

A entrada de controle u está atrasada em d amostras. O foco do trabalho principal que motivou a criação dessa biblioteca é sobre invariância com atraso, dessa forma foram desenvolvidos alguns programas auxiliares que lidam com problemas que surgem quando consideramos o atraso.

Utilizamos o conceito de $\mathcal{D}$-invariância *delay-dependent* (DóREA; OLARU; NICULESCU, 2023) que é uma abordagem menos conservadora que leva em conta o tamanho do atraso para calcular o conjunto invariante. Por outro lado, nessa abordagem, a invariância só é válida para um conjunto de condições iniciais, que pode ser calculado a partir do poliedro invariante e das matrizes do sistema original. Para tanto nós trabalhamos com o espaço expandido em que cada estado é uma amostra do sistema original. A ideia é encontrar vetores de condições iniciais $\mathbf{x}^T = [x(0), x(-1), ..., x(-d+1), x(-d)]$ para os quais

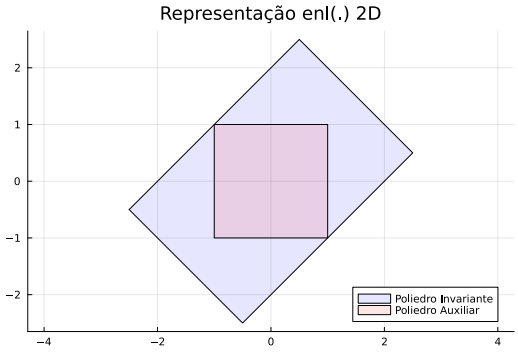


Figura 3: *Poliedro auxiliar contido no poliedro invariante*

a invariância é valida para o sistema original.

$$X(k+1) = \mathcal{A}X(k) = \begin{bmatrix} A & 0 & \cdots & 0 & A_d \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ x(k-1) \\ \vdots \\ x(k-d+1) \\ x(k-d) \end{bmatrix}$$

Pode ser demonstrado que para a condição inicial ser válida ela deve fazer parte do poliedro Ω_{adm} definido por

$$\begin{bmatrix} \mathcal{F} \\ \mathcal{F}\mathcal{A} \\ \mathcal{F}\mathcal{A}^2 \\ \vdots \\ \mathcal{F}\mathcal{A}^d \end{bmatrix} X(k) \leq \begin{bmatrix} \mathcal{W} \\ \mathcal{W} \\ \mathcal{W} \\ \vdots \\ \mathcal{W} \end{bmatrix}, \quad \mathcal{F} = \begin{bmatrix} F & 0 & \cdots & 0 & 0 \\ 0 & F & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & F & 0 \\ 0 & 0 & \cdots & 0 & F \end{bmatrix}.$$

O que o programa desenvolvido fez então foi montar essas matrizes estendidas, com base nas matrizes do sistema e no poliedro. Note que, quanto maior o atraso, maior a dimensão dessas matrizes, o que acaba tornando a computação mais custosa. Algo que fizemos para mitigar esses efeitos foi ir eliminando as redundâncias dessas matrizes no momento da criação delas. Dessa forma, conseguimos definir o conjunto de condições iniciais admissíveis de forma mais eficiente.

Uma vez que temos Ω_{adm} podemos partir para as simulações. Se o usuário já tiver uma condição inicial em mente ele pode testar se essas condições estão contidas em Ω_{adm} pela verificação da desigualdade. Por outro lado, se não houver nenhuma condição candidata, podemos escolher um dos vértices do poliedro de condições iniciais admissíveis, plotados na figura 4, como ponto de partida. Essa segunda abordagem é interessante pois não precisamos de um chute inicial para encontrar uma condição e em geral conseguimos pontos interessantes para simular, que se encontram perto dos limites do poliedro invariante encontrado.

.

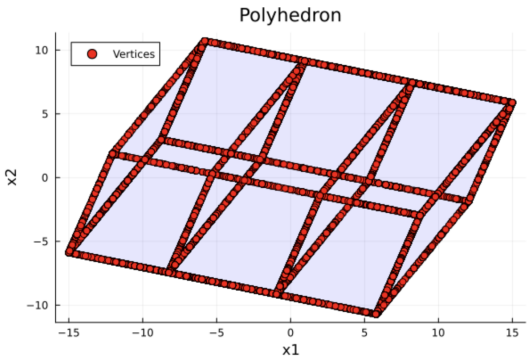


Figura 4: *x(0) de todos os vértices do poliedro de condições iniciais admissíveis, com o poliedro invariante de fundo*

3 Exemplos

Para mostrar um pouco mais sobre o fluxo de trabalho com essa biblioteca e como todo o processo que antes era dividido entre vários ambientes foi todo sintetizado em um só lugar, separamos um exemplo numérico. Mais demonstrações do uso juntamente com a documentação podem ser vistas em https://github.com/Nicholas-Wallace/Polyhedron .

Com o *Julia* instalado na máquina o usuário tem a opção de usar o REPL nativo da linguagem dentro do próprio terminal, ou, uma alternativa mais interativa e visualmente agradável é o notebook Pluto (PLAS et al.,).

Nesse exemplo vamos estar trabalhando com um sistema descrito por:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 \\ 1 \end{bmatrix} .C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

Definimos então o nosso sistema com essas matrizes e com realimentação estado (SSF).

Considerando também as seguintes matrizes que serão utilizadas para representar conjuntos de restrições:

```
sys = SLIT(
        A = 2×2 Matrix{Float64}:
             1.0  1.0
             0.0  1.0
        B = 2×1 Matrix{Float64}:
             2.0
             1.0
        C = 1×2 Matrix{Float64}:
             1.0  0.0
        feedback = SSF::Feedback = 1
        G = 1×2 Matrix{Float64}:
             1.0  1.0
    )
sys = SLIT(A, B, C, SSF)
```

Figura 5: *Definição do sistema de interesse no ambiente Pluto*

$$X^T = \begin{bmatrix} 0.8 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}, \qquad U = \begin{bmatrix} 1.2 \\ -1.5 \end{bmatrix}.$$

Nos quais $Xx \leq \mathbf{1}$ representa as retrições de estado e $Uu \leq \mathbf{1}$ as restrições na entrada.

Podemos então verificar se o poliedro das restrições de estado é invariante com respeito ao sistema que anteriormente definido utilizando a função *is_pinvariant*.

```
ip = is_pinvariant(sys, U, X)

Job 23640258 dispatched
password: mWyKrvHT

---------- Begin Solver Output -----------
Condor submit: 'neos.submit'
Condor submit: 'watchdog.submit'
Job submitted to NEOS HTCondor pool.

termination_status(model) = MathOptInterface.LOCALLY_SOLVED
```

Figura 6: *Resultado Locally Solved do problema de otimização*

Esse exemplo foi escolhido propositalmente, pois o mesmo poliedro de restrições não é invariante se o sistema fosse definido com realimentação de saída. Entretanto podemos utilizar a função que procura um poliedro dentro das restrições que seja invariante. Na figura (7) podemos ver os dois poliedros sobrepostos.

É perceptível que o poliedro não preenche toda o poliedro de restrições dos estados, mas cobre quase toda área, rejeitando apenas uma pequena região do espaço.

3.1 Exemplo: sistema com atraso

Para um sistema com atraso (5) para o qual:

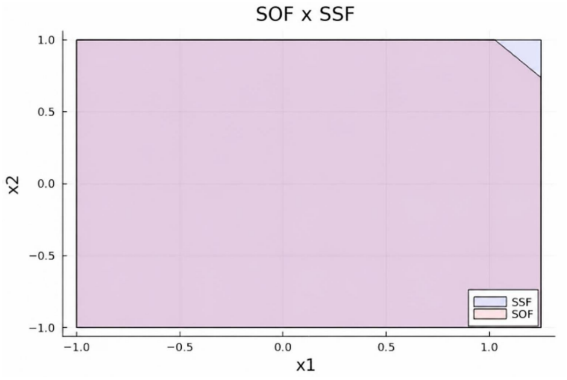


Figura 7: *Poliedro invariante com realimentação de saída*

$$A = \begin{bmatrix} 1.2 & 0.2 \\ -0.4 & 0.6 \end{bmatrix}, \qquad A_d = \begin{bmatrix} -0.3 & -0.2 \\ 0.4 & 0.2 \end{bmatrix},$$

e um conjunto poliedrico $\Omega = \{x \in \mathbb{R}^n : Fx \leq w\}$ com

$$F^T = \begin{bmatrix} -1 & 2 & 1 & -2 \\ -1 & 1 & 1 & -1 \end{bmatrix}, \qquad w^T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}.$$

Utilizando os programas desenvolvidos verificamos que, para d = 1, o conjunto é invariante, e plotamos na figura (8) um exemplo de trajetória inciando em um dos vértices. Podemos observar que a trajetória vai para origem e não saí dos limites do poliedro.

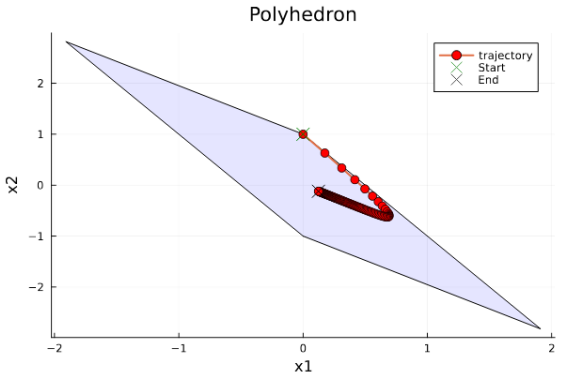


Figura 8: *trajetória dos estados do sistema*

4 Conclusões

Este trabalho apresentou uma biblioteca em *Julia* para o cálculo e a análise de conjuntos invariantes, integrando ferramentas de otimização matemática e

computação poliédrica em um único ambiente. As funcionalidades implementadas permitem definir sistemas dinâmicos, verificar a invariância de conjuntos sob restrições e buscar conjuntos invariantes para vários tipos de sistema, considerando diferentes configurações de realimentação. Os exemplos demonstram a aplicação da biblioteca aos problemas considerados e sua utilidade no estudo de sistemas com atraso na entrada de controle. Como trabalhos futuros, pretende-se ampliar as funcionalidades para outros tipos de conjuntos invariantes e integrar com outras bibliotecas do ecossistema *Julia*.

Referências

BEZANSON, J. et al. Julia: A fresh approach to numerical computing. *SIAM Review*, v. 59, n. 1, p. 65–98, 2017.

BRIÃO, S. L. et al. Output feedback design for discrete-time constrained systems subject to persistent disturbances via bilinear programming. *Journal of the Franklin Institute*, Elsevier, v. 358, n. 18, p. 9741–9770, 2021.

BYRD, R. H.; NOCEDAL, J.; WALTZ, R. A. Knitro: An integrated package for nonlinear optimization. In: PILLO, G. D.; ROMA, M. (Ed.). *Large-Scale Nonlinear Optimization*. Boston: Springer, 2006.

DóREA, C.; OLARU, S. Computation of low complexity invariant sets for time-delay systems: An optimization-based approach. In: *Proc. 18th IFAC Workshop on Control Applications of Optimization (accepted)*. [S.l.: s.n.], 2022. p. 1–6.

DóREA, C. E.; OLARU, S.; NICULESCU, S.-I. Delay-dependent invariance of polyhedral sets for discrete-time linear systems. *IFAC-PapersOnLine*, v. 56, n. 2, p. 3259–3264, 2023. ISSN 2405-8963. 22nd IFAC World Congress.

FOURER, R.; GAY, D. M.; KERNIGHAN, B. W. A modeling language for mathematical programming. *Management Science*, v. 36, n. 5, p. 519–554, 1990.

FUKUDA, K. *cddlib: A C Implementation of the Double Description Method*. [S.l.], 2022. Disponível em: <https://github.com/cddlib/cddlib>.

LEGAT, B. Polyhedral computation. In: *JuliaCon*. [s.n.], 2023. Disponível em: <https://pretalx.com/juliacon2023/talk/JP3SPX/>.

LUBIN, M. et al. JuMP 1.0: Recent improvements to a modeling language for mathematical optimization. *Mathematical Programming Computation*, v. 15, p. 581–589, 2023.

PLAS, F. van der et al. *Pluto.jl: Simple reactive notebooks for Julia*. Disponível em: <https://github.com/JuliaPluto/Pluto.jl>.

The Julia Programming Language. *Functions and types: Multiple dispatch*. 2020. MIT 18.S191, Week 3, Fall 2020. Vídeo publicado em 15 set. 2020. Disponível em: <https://www.youtube.com/watch?v=DXtbaSP_LFI>.

The MathWorks, Inc. *MATLAB*. The MathWorks, Inc. Acesso em: 16 out. 2026. Disponível em: <https://www.mathworks.com/products/matlab.html>.

Desenvolvimento de Ferramenta Computacional Multicritério para Avaliação de Desempenho de Controladores PID

***Thyfanne Dayane Cosme do Nascimento*[1]**
***Dominic Nascimento Rodrigues*[2]**
***André Laurindo Maitelli*[3]**

[1] *Centro de Tecnologia, UFRN, thyfanne.cosme.116@ufrn.edu.br*
[2] *Escola de Ciência e Tecnologia, UFRN, dominic.rodrigues.183@ufrn.edu.br*
[3] *Centro de Tecnologia, UFRN, maitelli@dca.ufrn.br*

Resumo

A avaliação de desempenho de malhas de controle requer a análise conjunta do comportamento dinâmico e das restrições físicas do processo. Este artigo apresenta o desenvolvimento de uma ferramenta computacional multicritério, implementada em Python, capaz de automatizar a simulação, a análise e o ranqueamento de sintonias de controladores Proporcionais-Integrais-Derivativos, conhecidos pela sigla PID. A aplicação preliminar do *software* foi conduzida por meio de um estudo de caso aplicado ao modelo dinâmico com atraso da coluna de destilação de *Wood-Berry*. Os resultados indicam a capacidade da ferramenta em reprovar objetivamente configurações agressivas e selecionar, de forma automática e estruturada, o controlador que oferece melhor compromisso entre rapidez de resposta e atendimento às restrições operacionais adotadas.

1 Introdução

Os controladores Proporcionais Integrais Derivativos, amplamente conhecidos como PIDs, permanecem entre as estruturas de controle mais utilizadas na automação industrial, devido à sua simplicidade estrutural e facilidade de implementação [1,3,9,10]. No entanto, para preservar esse rendimento ao longo do ciclo de vida operacional, é preciso lidar com desafios frequentes, como sintonias mal ajustadas, desgastes físicos de componentes e variações na dinâmica do processo. Assim, o monitoramento contínuo e a avaliação sistemática do comportamento dinâmico tornaram-se fundamentais para assegurar a eficiência operacional e a qualidade do produto final [14,15].

Na prática industrial, o desempenho da malha é tradicionalmente avaliado por métricas de erros integrais. Destacam-se a Integral do Erro Absoluto, referenciada como IAE, e a Integral do Erro Quadrático, a ISE, que medem a magnitude acumulada dos desvios, além de variações com ponderação temporal, como o ITAE e o ITSE [2,4], os quais penalizam erros persistentes no regime permanente.

Contudo, avaliar o controlador fundamentando-se em métricas isoladas revela-se insuficiente, podendo conduzir a sintonias agressivas que provocam o desgaste prematuro de atuadores ou violações operacionais. Trabalhos recentes em avaliação de desempenho de malhas, *retuning* de controladores PID e sintonia multiobjetivo reforçam a importância de considerar simultaneamente critérios de erro, sobressinal, esforço de controle e restrições operacionais [11–13, 16–18]. Para superar esta limitação, o presente trabalho adota uma abordagem multicritério, na qual múltiplos indicadores de erro, resposta transitória e esforço de controle são normalizados e combinados num índice composto global J, denominado neste trabalho como *Score*. Por meio de travas de segurança e pesos configuráveis, este sistema computacional engloba as principais métricas de avaliação, torna explícita a

influência de cada critério de ponderação e permite ranquear diversas configurações de controladores PID.

2 Metodologia

O desenvolvimento do sistema baseia-se em uma arquitetura modular orientada a objetos e funções, com separação entre os componentes de modelagem, simulação e avaliação. Essa organização preserva a independência das fontes de dados em relação ao núcleo analítico. A estrutura contempla módulos dedicados aos modelos matemáticos, controladores, simulação em malha fechada, cálculo de métricas e armazenamento de resultados. Essa modularidade permite a incorporação de novos modelos sem alterações substanciais no núcleo das avaliações.

2.1 Modelagem e discretização dos sistemas

No modo de simulação, o sistema emprega funções de transferência contínuas para representar as plantas. A dinâmica é convertida para o domínio discreto de acordo com o período de amostragem T_s, utilizando o método de Retenção de Ordem Zero, termo derivado do inglês *Zero-Order Hold* cuja sigla é ZOH [5,6], que considera o sinal de controle constante durante cada intervalo.

Para modelos com atraso de transporte inteiro em relação a T_s, o atraso é representado diretamente no domínio discreto por um deslocamento de $d = T/T_s$ amostras, correspondendo ao operador z^{-d}. Essa abordagem evita a aproximação contínua por funções de Padé e permite implementação por estados de memória. Os subsistemas interconectados são representados em Espaço de Estados, contemplando a planta, o atraso discreto e os estados internos, permitindo tratar de forma uniforme a malha fechada e reduzindo a manipulação de polinômios complexos.

Paralelamente, os controladores PID são definidos a partir da sua implementação paralela, permitindo a utilização independente dos parâmetros de sintonia K_p, K_i e K_d. Para execução digital, são discretizados pela transformação bilinear de Tustin [5]. A resposta temporal é obtida a partir de um vetor de tempo com incrementos correspondentes ao período de amostragem, utilizando-se a resposta forçada do modelo.

2.2 Execução dos ensaios e estímulos

O comportamento dinâmico é avaliado por ensaios de variação controlada do *setpoint*. A aplicação do estímulo é precedida por um intervalo de pré-estabilização t_{sp}, estabelecendo uma janela observável anterior à aplicação do degrau, permitindo separar eventuais efeitos de inicialização.

2.3 Avaliação do desempenho

A avaliação é estruturada em múltiplos grupos de métricas. Os índices integrais IAE, ISE, ITAE e ITSE quantificam o erro acumulado. Para ITAE e ITSE, o tempo de ponderação é contado a partir de cada alteração de referência, evitando que o período de pré-estabilização introduza uma penalização injusta.

A resposta transitória é avaliada por métricas temporais. O sobressinal, também referido como *overshoot*, é calculado sobre a amplitude efetiva do degrau aplicado. O tempo de acomodação baseia-se na entrada e permanência do sinal dentro de uma banda de tolerância em torno da referência. O tempo de subida é o intervalo necessário para que a variável percorra de 10% a 90% da variação, definição amplamente presente na literatura clássica da área [2, 3]. O módulo de oscilação avalia ainda a dispersão das variáveis e a quantidade de cruzamentos. O erro em regime permanente é estimado pelo desvio absoluto entre a referência final e a média da cauda da resposta, considerando os últimos 5% do trecho avaliado. A contagem de cruzamentos do *setpoint* é obtida pelas mudanças de sinal da diferença entre a variável controlada e a referência, sendo usada como indicador complementar de oscilação em torno do valor desejado.

2.4 Restrições e consolidação dos indicadores

Sintonias que induzem instabilidade, excedem sobressinal limite ou demandam valores de variável manipulada, referenciada pela sigla MV, fora da faixa operacional admissível, são rejeitadas. Neste estudo, a saturação física do atuador não é modelada como uma não linearidade interna da malha; os limites de MV são empregados como critério de triagem aplicado ao comando calculado. A esses controladores é atribuída a penalidade máxima de valor 1,0, bloqueando-os para a operação. Para os aprovados, aplica-se a normalização *Min-Max*, transformando grandezas físicas heterogêneas em escalas adimensionais contínuas entre 0 e 1, onde valores menores refletem desempenhos superiores.

O índice global é obtido por uma matriz de ponderação gerada via questionário submetido ao engenheiro de processo, abrangendo cinco dimensões fundamentais. O sistema avalia a tolerância ao sobressinal,

definindo o rigor da restrição a ultrapassagens da referência, e o comportamento dinâmico, que mensura a preferência entre uma atuação conservadora ou agressiva. Considera-se também o desgaste do atuador, relacionado à suscetibilidade do elemento final à fadiga mecânica. Por fim, o questionário afere a rapidez transitória, traduzindo a urgência operacional na rejeição de distúrbios, e a precisão em regime permanente, que estabelece a tolerância a erros residuais de acomodação.

Os níveis definidos no questionário não correspondem diretamente a uma única métrica isolada; cada dimensão operacional redistribui prioridades para um subconjunto de indicadores da matriz de ponderação, permitindo traduzir requisitos qualitativos em pesos quantitativos.

O índice global J, denominado *Score*, é calculado conforme definido na Equação 1:

$$J_i = \sum_{j=1}^{m} w_j \overline{M}_{ij} \tag{1}$$

Em que J_i representa o índice global do controlador i, w_j é o peso associado à métrica j, $\overline{M}_{ij}$ é o valor normalizado da métrica j para o controlador i, e m é o número total de métricas avaliadas. Como a normalização foi definida de modo que valores menores indiquem melhor desempenho, o controlador com menor J_i é considerado superior dentro do conjunto analisado.

3 Estudo de Caso e Aplicação

O *software* foi avaliado inicialmente utilizando o clássico modelo da coluna de destilação de *Wood e Berry* [7]. A Tabela 1 exibe a parametrização adotada na simulação digital.

Tabela 1: *Parâmetros do processo e configurações do ensaio.*

Parâmetro	**Símbolo**	**Valor**
Planta		
Ganho	K	$-19{,}4$
Constante de tempo	τ	864 s
Tempo morto	T	180 s
Simulação e Restrições		
Período de amostragem	T_s	10 s
Atraso discreto	d	18 amostras
Pré-estabilização	t_{sp}	300 s
Amplitude da variação	ΔSP	0,1
Sobressinal máximo	OV_{max}	15%
Restrição do atuador	$MV_{min,max}$	$[-0{,}5; 0{,}5]$

O modelo FOPDT selecionado é a função de transferência $G_{22}(s)$ da matriz de *Wood-Berry*, relacionando a vazão de vapor no refervedor, que atua como variável manipulada, à composição do fundo da coluna, que serve como variável controlada.

O modelo original de *Wood-Berry* descreve uma coluna de destilação binária de metanol e água, com duas entradas (vazões de refluxo e de vapor no refervedor) e duas saídas (composições dos produtos de topo e de fundo). A sigla FOPDT corresponde ao inglês *First-Order Plus Dead Time*, ou modelo de primeira ordem com tempo morto. Para a malha $G_{22}(s)$ adotada neste estudo, a função de transferência, com os parâmetros de tempo expressos em segundos, é dada por:

$$G_{22}(s) = \frac{-19{,}4\, e^{-180s}}{864s + 1}. \tag{2}$$

O ganho estático negativo indica uma relação inversa, requerendo ação direta do controlador para garantir estabilidade. A relação entre a constante de tempo e o tempo morto permite avaliar a influência relativa do atraso de transporte na dinâmica do processo. Para o caso analisado, essa razão é dada por $\tau/T = 864/180 = 4{,}8$, indicando que a inércia do processo é superior ao atraso, embora este ainda exerça influência significativa na resposta. Além disso, o período de amostragem ($T_s = 10$ s) fornece aproximadamente $\tau/T_s = 86{,}4$ amostras por constante de tempo.

Essa escolha de amostragem também estabelece uma relação inteira com o tempo morto do processo, resultando no atraso discreto de $d = T/T_s = 180/10 = 18$ amostras, gerando associadamente o termo z^{-18}. A simulação utilizou variáveis de desvio, com reconstrução do sinal absoluto por $PV(t) = PV_0 + \Delta PV(t)$, adotando $PV_0 = 0{,}5$ e $MV_0 = 0$. Foram aplicadas variações de referência com amplitude de 0,1 nos dois sentidos de marcha, com duração máxima de 8000 s.

Foram estabelecidos limites rigorosos, como o *overshoot* máximo de 15% e a restrição da variável manipulada na faixa $MV \in [-0{,}5; 0{,}5]$. Cinco sintonias PID com níveis de agressividade gradativa, variando desde um perfil muito conservador até a estrutura agressiva de Ziegler-Nichols [8], foram testadas. Após serem filtrados pelas travas de segurança, os controladores aprovados receberam a normalização *Min-Max*. O questionário de ponderação priorizou o comportamento dinâmico, configurado com peso de nível 3, mantendo as demais exigências no limiar basal correspondente ao peso de nível 1.

4 Resultados e Discussão

A avaliação dos controladores perante a coluna de *Wood-Berry* resultou no ranqueamento da Tabela 2. Duas janelas completas compuseram o *Score*; o trecho parcial final foi excluído desse cálculo, mas considerado nas verificações de segurança. O comportamento transitório correspondente é ilustrado na Figura 1.

Tabela 2: *Ranqueamento final e estado dos controladores avaliados.*

Pos.	ID	K_p	K_i	K_d	Score	Status
1º	#2	−0,080	−0,00009	0,000	0,2079	Aprovado
2º	#3	−0,124	−0,00014	0,000	0,3234	Aprovado
3º	#1	−0,043	−0,00005	0,000	0,6766	Aprovado
–	#4	−0,200	−0,00040	−10,00	1,0000	Bloq. (OV)
–	#5	−0,297	−0,00083	−26,73	1,0000	Bloq. (OV)

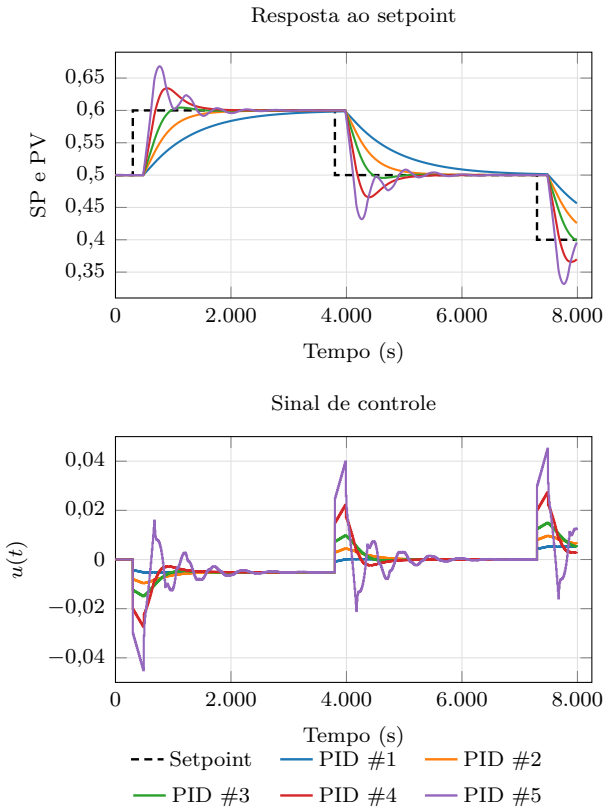


Figura 1: *Resposta da variável controlada às mudanças de referência (acima) e sinais de controle dos cinco controladores PID (abaixo).*

O contraste visual na Figura 1 corrobora o ranqueamento. O gráfico de resposta na parte superior revela que as sintonias #4 e #5 apresentam sobressinais máximos de 34,21% e 68,19%, respectivamente, superiores ao limite de 15%. As sintonias #4 e #5 foram classificadas pelo *software* como bloqueadas por violação do critério de sobressinal, recebendo penalidade máxima no *Score*, conforme evidenciado na Tabela 2. O gráfico de esforço, alocado na parte inferior, reforça esta inadequação pela maior variação relativa demandada à variável manipulada. A curva do PID #1 é demasiadamente amortecida, resultando numa aproximação excessivamente lenta com grande erro acumulado. O PID #3 reage rapidamente, mas apresenta sobressinal máximo de 4,35%, inferior ao limite de 15%, além de cruzamentos da referência e maior variância do comando. Já o PID #2 atenua a agressividade sem prejudicar criticamente a rapidez de subida, apresentando um *Score* equilibrado.

A aplicação da restrição de sobressinal permitiu identificar e rejeitar as sintonias incompatíveis com o limite operacional estabelecido. Para os classificados, a Tabela 3 detalha as métricas e a ponderação. Os pesos percentuais exibidos na Tabela 3 foram arredondados para uma casa decimal apenas para fins de apresentação. O cálculo do *Score* foi realizado com os pesos normalizados em precisão completa no *software*.

Tabela 3: *Métricas normalizadas e matriz de pesos percentuais.*

Métrica	Peso	PID #1	PID #2	PID #3	PID #4	PID #5
IAE	8,2%	1,0000	0,2837	0,0000	1,0000	1,0000
ISE	8,2%	1,0000	0,2785	0,0000	1,0000	1,0000
ITAE	8,9%	1,0000	0,1854	0,0000	1,0000	1,0000
ITSE	9,3%	1,0000	0,1855	0,0000	1,0000	1,0000
Sobressinal	10,8%	0,0000	0,0000	1,0000	1,0000	1,0000
Tempo de Acom.	10,8%	1,0000	0,2823	0,0000	1,0000	1,0000
Tempo de Subida	4,8%	1,0000	0,2857	0,0000	1,0000	1,0000
Erro em Regime	7,1%	1,0000	0,0485	0,0000	1,0000	1,0000
Var. da Saída	10,4%	1,0000	0,6003	0,0000	1,0000	1,0000
Var. do Controle	7,8%	0,0000	0,2301	1,0000	1,0000	1,0000
Cruzamentos SP	13,8%	0,0000	0,0000	1,0000	1,0000	1,0000

Entre os aprovados, o PID #1, possuindo o ganho $K_p = -0{,}043$, apresentou o maior *Score* de 0,6766, conforme a Tabela 2. Esse resultado indica que a baixa agressividade proporcional penalizou os índices integrais, reduzindo sua classificação global. O PID #3 gerou um *Score* de 0,3234, sendo penalizado pelo sobressinal, pelos cruzamentos da referência e pela maior variância do controle. Para os pesos adotados neste estudo, o PID #2 apresentou o menor *Score*, igual a 0,2079, indicando o melhor compromisso multicritério entre erro, rapidez de resposta e esforço de controle. A diferença de 0,1155 entre os *Scores* dos controladores #2 e #3 evidencia o efeito da ponderação dos diferentes critérios sobre a classificação.

5 Conclusão

A ferramenta computacional apresentada automatiza a avaliação multicritério de controladores PID. Sua arquitetura modular separa o motor de simulação da estrutura analítica de avaliação de desempenho, permitindo calcular índices integrais de erro, métricas transitórias e indicadores associados ao esforço de controle. A utilização da normalização *Min-Max*, associada a travas algorítmicas de segurança, reduz a dependência de avaliações baseadas em um único índice quantitativo e torna explícita a influência dos critérios de ponderação adotados.

No estudo de caso analisado, a ferramenta demonstrou capacidade de rejeitar sintonias incompatíveis com as restrições operacionais adotadas e de selecionar a configuração com melhor compromisso multicritério entre erro, rapidez de resposta e esforço de controle. Os resultados indicam a viabilidade da abordagem proposta para apoiar a avaliação comparativa de sintonias PID sob critérios múltiplos. Como limitação, destaca-se que o ranqueamento obtido depende dos pesos atribuídos às métricas e do conjunto de controladores avaliados, uma vez que a normalização *Min-Max* é relativa à amostra analisada.

Como continuidade imediata, pretende-se ampliar a avaliação da ferramenta por meio de análises de sensibilidade dos pesos, variações paramétricas da planta e inclusão de perturbações de carga. Esses estudos permitirão verificar a estabilidade do ranqueamento diante de diferentes cenários operacionais e configurações de prioridade. Em continuidade ao escopo de desenvolvimento previsto para o *software*, etapas posteriores deverão contemplar integração com protocolos industriais, importação de dados históricos, diagnóstico de fricção de atuadores, recomendação automatizada de *retuning* e estruturação futura da ferramenta como apoio à construção de gêmeos digitais de malhas de controle.

Referências

[1] M. C. M. M. Campos, H. C. G. Teixeira: Controles Típicos de Equipamentos e Processos Industriais. Editora Edgard Blücher, 2006.

[2] R. C. Dorf, R. H. Bishop: Sistemas de Controle Modernos. 12ª Edição. LTC Livros Técnicos e Científicos, 2013.

[3] K. Ogata: Engenharia de Controle Moderno. 5ª Edição. Prentice-Hall do Brasil, 2011.

[4] G. F. Franklin, J. D. Powell, A. E. Naeini: Sistemas de Controle para Engenharia. 6ª Edição. Bookman Editora LTDA, 2013.

[5] K. Ogata: Discrete-Time Control Systems. Prentice-Hall, 1987.

[6] K. J. Åström, B. Wittenmark: Computer-Controlled Systems: Theory and Design. 3ª Edição. Prentice-Hall, 1997.

[7] R. K. Wood, M. W. Berry: Terminal composition control of a binary distillation column. Chemical Engineering Science **28** (1973) 1707–1717.

[8] J. G. Ziegler, N. B. Nichols: Optimum settings for automatic controllers. Transactions of the ASME **64** (1942) 759–768.

[9] R. P. Borase, D. K. Maghade, S. Y. Sondkar, S. N. Pawar: A review of PID control, tuning methods and applications. International Journal of Dynamics and Control **9** (2021) 818–827.

[10] O. A. Somefun, K. Akingbade, F. M. Dahunsi: The dilemma of PID tuning. Annual Reviews in Control **52** (2021) 65–74.

[11] X. Gao, F. Yang, C. Shang, D. Huang: A novel data-driven method for simultaneous performance assessment and retuning of PID controllers. Industrial & Engineering Chemistry Research **56** (2017) 2127–2139.

[12] S. Yu, X. Li: Proportional-Integral-Derivative controller performance assessment and retuning based on general process response data. ACS Omega **6**(15) (2021) 10207–10223.

[13] W. Zhang, H. Dong, Y. Xu, D. Cao, X. Li: Multiobjective tuning and performance assessment of PID using Teaching-Learning-Based Optimization. ACS Omega **6** (2021) 31765–31774.

[14] S. X. Ding, L. Li: Control performance monitoring and degradation recovery in automatic control systems: A review, some new results, and future perspectives. Control Engineering Practice **111** (2021) 104790.

[15] W. Bounoua, M. F. Aftab, C. W. P. Omlin: Controller performance monitoring: A survey of problems and a review of approaches from a data-driven perspective with a focus on oscillations detection and diagnosis. Industrial & Engineering Chemistry Research **61**(49) (2022) 17735–17765.

[16] C. J. Munaro, M. R. Pimentel, R. Bacci di Capaci, L. Campestrini: Data driven performance monitoring and retuning using PID controllers. Computers & Chemical Engineering **178** (2023) 108360.

[17] T. Marques, G. Reynoso-Meza: Applications of multi-objective optimisation for PID-like controller tuning: a 2015–2019 review and analysis. IFAC-PapersOnLine **53**(2) (2020) 7933–7940.

[18] A. Rodríguez-Molina, E. Mezura-Montes, M. G. Villarreal-Cervantes, M. Aldape-Pérez: Multi-objective meta-heuristic optimization in intelligent control: A survey on the controller tuning problem. Applied Soft Computing **93** (2020) 106342.

Framework paralelo de ajuste de campos de força atomísticos e *coarse-grain* de grafeno através de otimização por enxame de partículas (PSO)

***Iwri Simplicio Felipe*[1]**
***Elton José Figueiredo de Carvalho*[2]**

[1] *Departamento de Engenharia de Computação e Automação, UFRN, <iwri.felipe.707@ufrn.edu.br>*
[2] *Escola de Ciências e Tecnologia, UFRN, <elton.carvalho@ufrn.br>*

Resumo

A determinação acurada de parâmetros para campos de força em simulações de dinâmica molecular tradicionalmente depende de calibrações empíricas sujeitas a erros e à exaustiva carga de trabalho manual. Este trabalho apresenta e avalia um ecossistema computacional em malha fechada para a parametrização automatizada e concorrente dos parâmetros de dissociação e decaimento do potencial de Morse (D_0 e α) e da constante elástica de flexão angular (k_θ) para monocamadas de grafeno. A metodologia é versátil, abrangendo desde modelos atomísticos até formulações de granulação grossa (*coarse-grain*), nas quais cada partícula efetiva condensa quatro átomos de carbono. O fluxo integra o algoritmo de Otimização por Enxame de Partículas (PSO), uma camada de orquestração distribuída via MPI e o simulador LAMMPS (THOMPSON et al., 2022). Amostras virtuais com quiralidade *zigzag* foram submetidas a ensaios concorrentes de tração uniaxial e cisalhamento puro. No módulo de análise, rotinas de truncamento autônomo no limite de integridade estrutural alimentam um modelo polinomial de segunda ordem com intercepto nulo ($\sigma = E_2\varepsilon + E_3\varepsilon^2$), onde o cômputo da aptidão baseado na diferença quadrática entre os coeficientes ajustados da partícula e os valores-alvo garante uma superfície de custo suave e bem-condicionada. Configurado sob as diretrizes cinéticas de inércia e estabilidade propostas por Shi e Clerc, o algoritmo demonstrou convergência estruturada e consistente atingindo o limiar físico de estagnação em poucas dezenas de gerações, mitigando o custo típico de milhares de passos em direção aos critérios de parada estabelecidos. Os parâmetros ótimos identificados reproduziram com fidelidade as propriedades elásticas canônicas e não lineares de referência da literatura, comprovando a robustez, precisão e repetibilidade da metodologia proposta para a parametrização sistemática de nanoestruturas bidimensionais.

Abstract

The accurate determination of force field parameters in molecular dynamics simulations traditionally relies on empirical calibrations that are error-prone and labor-intensive. This work presents a closed-loop computational ecosystem for the automated and concurrent parameterization of the Morse potential dissociation and decay parameters (D_0 and α) and the angular bending elastic constant (k_θ) for graphene monolayers. The methodology is versatile, encompassing both atomistic models and coarse-grained formulations, in which each effective particle condenses four carbon atoms. The workflow integrates the Particle Swarm Optimization (PSO) algorithm, a distributed orchestration layer via MPI, and the LAMMPS simulator (THOMPSON et al., 2022). Virtual samples with *zigzag* chirality were subjec-

ted to concurrent uniaxial tensile and pure shear tests. In the analysis module, autonomous truncation routines at the structural integrity limit feed a second-order polynomial model with a zero intercept ($\sigma = E_2\varepsilon + E_3\varepsilon^2$), where the fitness computation based on the squared difference between the particle's fitted coefficients and the target values ensures a smooth and well-conditioned cost surface. Configured under the kinetic guidelines of inertia and stability proposed by Shi and Clerc, the algorithm demonstrated a structured and consistent convergence reaching the physical stagnation threshold in a few dozen generations, thereby mitigating the typical computational cost of thousands of steps toward the established stopping criteria. The identified optimal parameters faithfully reproduced the canonical and non-linear elastic properties from literature references, proving the robustness, precision, and repeatability of the proposed methodology for the systematic parameterization of two-dimensional nanostructures.

1 Introdução

A calibração automatizada de modelos físicos a partir de dados empíricos ou de simulações de alta fidelidade é um desafio central em engenharia de sistemas e computação científica. No contexto da nanotecnologia computacional, a modelagem mecânica de nanomateriais bidimensionais (2D), cujo expoente canônico é o grafeno, servindo de base para estruturas de maior complexidade atômica como os dicalcogenetos de metais de transição (TMDs) (MANESCO; PULKIN, 2022), demanda modelos contínuos ou mesoscópicos (*coarse-grained*) capazes de estender a escala de análise para a ordem de centenas de nanômetros, dimensões típicas de dispositivos experimentais (WILLIAMS; LÍSAL, 2020; ZHAI; SANDLER, 2019). Nesses modelos, a resposta mecânica macroscópica decorre da ação combinada de potenciais intra e intermoleculares, englobando interações de ligação axial, flexão angular no plano, torção fora do plano, bem como contribuições não-ligadas de van der Waals e eletrostáticas. Contudo, a identificação dos parâmetros ótimos desses potenciais raramente admite solução analítica direta devido ao forte acoplamento entre os modos de deformação sob solicitações mecânicas distintas, como tração e cisalhamento (RUIZ et al., 2015).

Tradicionalmente, a parametrização de tais campos de força apoia-se em abordagens manuais iterativas ou varreduras pontuais em malha aberta, procedimentos suscetíveis a armadilhas de mínimos locais e ineficientes computacionalmente. A transição para estratégias de calibração em malha fechada baseadas em algoritmos heurísticos, como a Otimização por Enxame de Partículas (PSO), permite explorar sistematicamente espaços de busca multivariáveis. Entretanto, o fechamento desse laço de controle introduz uma dependência estrita da integridade numérica do módulo de avaliação: a cada iteração, o simulador de dinâmica molecular gera séries temporais de tensão e deformação que precisam ser pós-processadas e convertidas de forma autônoma em um escalar de aptidão (*fitness*).

É justamente na interface entre a simulação física e a avaliação algorítmica que residem os principais gargalos computacionais. Ensaios virtuais executados em ferramentas de dinâmica molecular como o LAMMPS (THOMPSON et al., 2022) exibem não-linearidades geométricas, a exemplo do enrugamento fora do plano (*buckling*) induzido por cisalhamento, e dispersões térmicas inerentes às respostas dinâmicas. Se a rotina de regressão matemática não for formulada com restrições físicas adequadas e baixa sensibilidade a covariâncias espúrias, a função-objetivo torna-se mal-condicionada, gerando degenerescências numéricas ou falhas de execução que paralisam a busca estocástica.

Com o intuito de mitigar esses entraves, este trabalho propõe um sistema computacional integrado para a identificação autônoma de parâmetros de campo de força para o grafeno e outros materiais bidimensionais com simetria hexagonal, como TMDs e nitreto de boro. O ecossistema orquestra ensaios mecânicos de tração uniaxial e cisalhamento paralelizados via MPI, acoplados a rotinas de regressão polinomial tensão-deformação com intercepto nulo na origem para a extração dos módulos elásticos. A investigação foca na estabilidade do processamento numérico das curvas mecânicas, na sensibilidade da função de custo frente a variações da energia de ligação (D_0) e do parâmetro de rigidez (α) do potencial de Morse e da constante angular (k_θ), bem como na eficácia do algoritmo de Otimização por Enxame de Partículas (PSO – *Particle Swarm Optimization*)(SHI; EBERHART, 1998) na varredura do espaço de busca em ambientes de computação de alto desempenho.

2 Metodologia

2.1 Modelo Mecânico Molecular e Potenciais Harmônicos

A representação discreta da monocamada de grafeno baseia-se em uma rede bidimensional que preserva a simetria hexagonal característica da hibridização

sp^2, com distância interatômica de equilíbrio $r_0 = 1{,}42\,\text{Å}$. A energia potencial total do sistema, U_{total}, é descrita por contribuições intramoleculares ligadas e interações não-ligadas de van der Waals:

$$U_{\text{total}} = U_{\text{lig}} + U_{\text{ang}} + U_{\text{diedro}} + U_{\text{vdW}} \tag{1}$$

Para a resposta mecânica no plano, as contribuições dominantes decorrem do estiramento axial das ligações covalentes (U_{lig}) e da flexão angular interligações (U_{ang}). A fim de capturar adequadamente o comportamento não linear e a energia de dissociação da malha sob maiores deformações, adota-se o potencial de Morse para as ligações, mantendo-se a formulação harmônica clássica para os ângulos:

$$U_{\text{lig}}(r) = D_0 \left[1 - e^{-\alpha(r-r_0)}\right]^2 \tag{2}$$

$$U_{\text{ang}}(\theta) = k_\theta(\theta - \theta_0)^2 \tag{3}$$

em que r e θ denotam, respectivamente, o comprimento instantâneo da ligação e o ângulo formado por trincas de átomos adjacentes, com $\theta_0 = 120°$ representando o ângulo de equilíbrio da rede hexagonal plana. Os parâmetros do potencial de Morse D_0 [kcal/mol] e α [1/Å], juntamente com a constante elástica k_θ [kcal/(mol · rad^2)], compõem as variáveis de decisão a serem identificadas pelo sistema de otimização. Mantemos fixo o comprimento de equilíbrio $r_0 = 1{,}42\,\text{Å}$ pois os ensaios não são sensíveis a essa grandeza e o modelo *coarse-grained* é obtido alterando esse valor e a massa das partículas, então é um parâmetro do nível de refinamento geométrico do modelo. As interações fora do plano (diédricas) e os termos não-ligados de Lennard-Jones (U_{vdW}) são mantidos fixos nos valores do campo de força OPLS-AA (KAMINSKI et al., 2001) para assegurar a integridade conformacional da rede.

2.2 Cinemática dos Ensaios Mecânicos Virtuais

A extração das propriedades constitutivas é realizada por meio de simulações de dinâmica molecular em regime quase-estático implementadas no simulador LAMMPS (THOMPSON et al., 2022), sob condições de contorno periódicas no plano e controle termodinâmico a $T = 300\,\text{K}$. Dois experimentos mecânicos complementares são configurados:

Tração Uniaxial (Estiramento Longitudinal): A dimensão longitudinal da caixa de simulação (L_x) é expandida monotonicamente no tempo segundo uma taxa de deformação constante $\dot{\varepsilon}_x$:

$$L_x(t) = L_{x,0}(1 + \dot{\varepsilon}_x t), \tag{4}$$

permitindo que a dimensão ortogonal relaxe via barostato para reproduzir o efeito de Poisson transversal livre de tensões artificiais.

Cisalhamento Puro (Angular): A caixa de simulação ortogonal sofre uma distorção angular afim através da variação da componente de inclinação (*tilt factor*) xy:

$$xy(t) = xy_0 + \dot{\gamma}_{xy} L_{y,0} t, \tag{5}$$

induzindo um estado de cisalhamento uniforme no plano da monocamada.

Em ambos os ensaios, a deformação máxima aplicada é delimitada para cobrir integralmente a zona elástica linear sem exceder o limiar de fratura estrutural ou regimes governados puramente por instabilidades termomecânicas tardias.

2.3 Arquitetura de Orquestração Paralela em Malha Fechada

O fluxo computacional de identificação paramétrica é estruturado como um sistema em malha fechada implementado em Python, integrando o orquestrador ao algoritmo PSO através da biblioteca `FFFit` (CARVALHO; SOUZA, 2021). A execução distribuída é escalonada em clusters de alto desempenho operados pelo agendador Slurm. A Figura 1 esquematiza a arquitetura do ecossistema.

O pipeline opera segundo as seguintes etapas de processamento:

Geração Topológica e Inicialização O nó orquestrador (*Rank 0*) constrói a topologia molecular inicial, indexando ligações, ângulos e diedros, e transmite a malha consistente a todos os processos por meio de primitivas de difusão (`MPI_Split`, `MPI_Bcast`).

Despacho Paralelo de Candidatos A cada geração j, o enxame propõe N partículas no espaço $\mathbb{R}^3$. Através da interface `mpi4py`, cada processo avaliador recebe um vetor de parâmetros candidato $\mathbf{x}_i = [D_i, \alpha_i, k_{\theta i}]$ e cria um diretório isolado de trabalho (`step_j/particle_i`).

Simulação Concorrente Em cada processo avaliador, os arquivos de configuração do LAMMPS são reparametrizados dinamicamente com base nas constantes de $\mathbf{x}_i$, executando os ensaios de tração longitudinal e cisalhamento transversal de forma concorrente e independente.

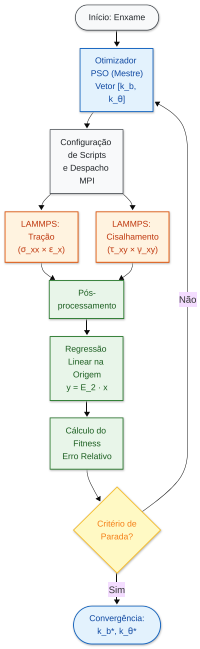


Figura 1: *Fluxograma da arquitetura computacional em malha fechada. O otimizador PSO distribui os parâmetros $(D,\ \alpha,\ k_\theta)$ via MPI para ensaios concorrentes de tração e cisalhamento no LAMMPS. Os perfis de tensão-deformação são tratados no módulo de análise por regressão quadrática forçando a origem, retroalimentando o escalar de aptidão até a convergência global.*

Pós-processamento e Cálculo do Fitness Os registros temporais do tensor de tensão virial, correspondentes à tração axial (σ_{xx}) e ao cisalhamento (τ_{xy}), são extraídos para o ajuste linear das relações constitutivas pelo módulo `analysis.py`, retornando a métrica escalar de aptidão ao nó mestre para a atualização cinemática do enxame.

A orquestração distribuída do enxame adotou uma topologia hierárquica em dois níveis de paralelismo para avaliar simultaneamente a população de $N = 32$ partículas. O comunicador global (`MPI_COMM_WORLD`) foi particionado dinamicamente por meio da primitiva `MPI_Comm_split`, gerando subcomunicadores independentes dedicados a cada candidato do enxame. Cada subconjunto alocou 4 núcleos de CPU para a execução concorrente de cada instância do LAMMPS via decomposição espacial de domínio, otimizando o balanço de carga e mitigando a sobrecarga de comunicação no cluster.

2.4 Processamento de Curvas, Modelagem Constitutiva e Função de Custo

Para cada candidato paramétrico avaliado, o simulador LAMMPS registra a evolução temporal das componentes do tensor de tensão virial ao longo do incremento das deformações aplicadas. O módulo de pós-processamento automatizado (`analysis.py`) realiza a leitura dessas séries temporais e executa a extração das propriedades mecânicas por meio de três etapas consecutivas: delimitação do regime elástico, regressão constitutiva e cômputo da divergência escalar.

Na primeira etapa, os perfis de tração e cisalhamento são truncados no limite de resposta mecânica estável da monocamada. Dado que o potencial de Morse admite dissociação química de ligações, o ponto de truncamento é identificado de forma autônoma pelo índice de máxima tensão sustentada na malha, delimitando o regime de interesse anterior à perda de linearidade e a instabilidades estruturais:

$$i_{\text{rup}} = \arg\max_i \left(\sigma_i\right) \tag{6}$$

restringindo os vetores de deformação ($\boldsymbol{\varepsilon}_{\text{rup}}$) e tensão ($\boldsymbol{\sigma}_{\text{rup}}$) ao intervalo pré-fratura. No ensaio de cisalhamento, a estrutura exibe uma transição geométrica em que a monocamada perde a planaridade basal e desenvolve ondulações transversais (*buckling*), alterando a inclinação da curva antes da ruptura catastrófica. O truncamento automatizado assegura que regiões de instabilidade pós-ruptura não contaminem os cálculos constitutivos.

Na segunda etapa, adota-se um modelo elástico linear puro com a restrição física elementar de tensão nula sob deformação nula, aplicado via ajuste de mínimos quadrados:

$$\sigma_{xx}(\varepsilon_x) = E_3^{\text{trac}}\varepsilon_x^2 + E_2^{\text{trac}}\varepsilon_x \tag{7}$$

$$\tau_{xy}(\gamma_{xy}) = E_3^{\text{cis}}\gamma_{xy}^2 + E_2^{\text{cis}}\gamma_{xy} \tag{8}$$

em que E_2^{trac} e E_2^{cis} correspondem, respectivamente, ao módulo de Young (E) e ao módulo de cisalhamento (G) da folha. E_3^{trac} e E_3^{cis} correspondem a correções de segunda ordem para reproduzir a concavidade das curvas observadas nos dados, compatíveis com trabalhos anteriores (RUIZ et al., 2015).

Por fim, a função de custo (*fitness*) é formulada como uma soma ponderada dos erros quadráticos relativos normalizados em relação aos alvos de referência da literatura:

$$\begin{aligned} f(D_0, \alpha, k_\theta) = \ & w_{\text{trac}}^{(2)}\Delta(E_2^{\text{trac}})^2 + w_{\text{trac}}^{(3)}\Delta(E_3^{\text{trac}})^2 \\ & + w_{\text{cis}}^{(2)}\Delta(E_2^{\text{cis}})^2 + w_{\text{cis}}^{(3)}\Delta(E_3^{\text{cis}})^2, \end{aligned} \tag{9}$$

onde

$$\Delta(E_2^{\text{trac}}) = \frac{E_2^{\text{trac}} - E_{\text{alvo}}}{E_{\text{alvo}}} \quad \Delta(E_3^{\text{trac}}) = \frac{E_3^{\text{trac}} - E_{3,\text{alvo}}}{E_{3,\text{alvo}}}$$
$$\Delta(E_2^{\text{cis}}) = \frac{E_2^{\text{cis}} - G_{\text{alvo}}}{G_{\text{alvo}}} \quad \Delta(E_3^{\text{cis}}) = \frac{E_3^{\text{cis}} - G_{3,\text{alvo}}}{G_{3,\text{alvo}}} \tag{10}$$

sendo w_{trac} e w_{cis} os pesos adimensionais atribuídos a cada modo de solicitação. A formulação expressa pela Equação (9) retorna uma função de aptidão estritamente não-negativa, cujo mínimo global é atingido no valor zero quando as propriedades simuladas coincidem exatamente com os valores-alvo, orientando as partículas do algoritmo de enxame em direção à região de parametrização ótima.

Para estabelecer os valores-alvo de rigidez que guiam a função de custo do algoritmo de otimização, realizaram-se simulações preliminares de tração uniaxial e cisalhamento utilizando o potencial reativo AIREBO (STUART; TUTEIN; HARRISON, 2000). Devido à sua precisão na descrição de interações covalentes em nanoestruturas de carbono, o AIREBO reproduz os efeitos não lineares sob grandes deformações. Consequentemente, as respostas tensão-deformação obtidas exibiram perda gradual de linearidade até o ponto de fratura (em $\varepsilon_x \approx 0{,}20$ e $\gamma_{xy} \approx 0{,}31$).

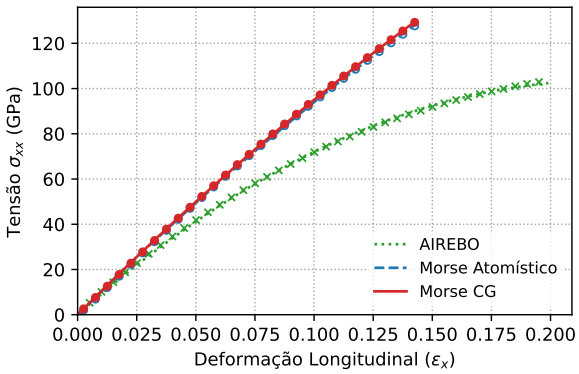
Para isolar o comportamento elástico linear, regime de interesse para a parametrização do modelo harmônico, aplicou-se uma regressão polinomial de segunda ordem ($\sigma = E_2\varepsilon + E_3\varepsilon^2$). A partir deste ajuste, extraíram-se os coeficientes lineares, estabelecendo como alvos paramétricos o módulo elástico bidimensional de tração $E_{2,\text{tracao}} \approx 926{,}6\,\text{GPa}$ e o módulo de cisalhamento $E_{2,\text{cis}} \approx 253{,}9\,\text{GPa}$. Esses módulos contínuos equivalentes serviram como referência (*ground truth*) para o cálculo do erro relativo avaliado pelas partículas do enxame.

3 Resultados e Discussão

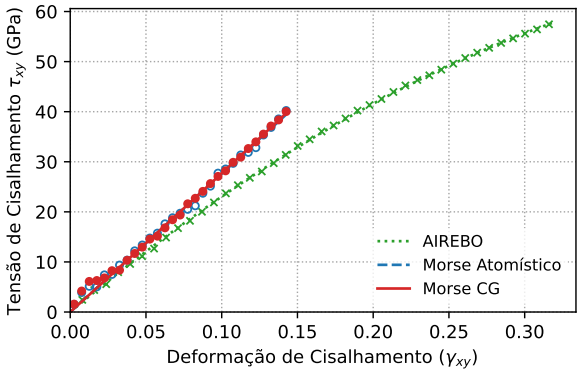
3.1 Comportamento Mecânico sob Tração e Cisalhamento

A resposta mecânica da monocamada de grafeno foi avaliada sob ensaios computacionais de tração uniaxial e cisalhamento simples no ambiente LAMMPS, adotando espessura efetiva de camada de 3,35 Å e quiralidade estrutural *zigzag* alinhada ao eixo x. Os perfis resultantes são ilustrados na Figura 2.

No ensaio de tração uniaxial (Figura 2a), observa-se um comportamento elástico inicial com regime linear bem definido entre o tensor de tensão virial e a deformação imposta. Para deformações maiores, tanto o modelo atomístico quanto o *coarse-grained* apresentam concavidade para baixo, compatível com o observado nas simulações de referência com o campo de forças AIREBO. Embora a curvatura observada nos campos ajustados seja menos intensa que aquela obtida na simulação de referência, está compatível com os valores de (RUIZ et al., 2015), que leva em conta referências experimentais.



(a) *Tração uniaxial ($\sigma_{xx} \times \varepsilon_x$)*



(b) *Cisalhamento puro ($\tau_{xy} \times \gamma_{xy}$)*

Figura 2: *Curvas de resposta mecânica tensão-deformação da monocamada de grafeno sob carregamentos no plano. Os gráficos contrastam o comportamento mecânico do potencial reativo AIREBO (referência) com os modelos fundamentados no campo de forças de Morse (Atomístico e* Coarse-Grained*), cujos parâmetros foram otimizados via PSO.*

De modo análogo, a resposta sob cisalhamento puro é apresentada na Figura 2b. Durante a deformação angular, a estrutura bidimensional preserva sua rigidez elástica até sofrer uma transição geométrica com perda de planaridade da folha, manifestada na forma de enrugamento fora do plano (*buckling*). A restrição do ajuste aos primeiros incrementos de deformação angular ($\gamma_{xy} \leq 0{,}05$) viabiliza o isolamento estrito do regime linear elástico, garantindo a extração estável e precisa do módulo de cisalhamento (E_2^{cis}).

3.2 Dinâmica de Convergência do PSO e Parametrização Ótima

A etapa de calibração paramétrica foi conduzida por meio do algoritmo de Otimização por Enxame de Partículas (PSO), configurado para minimizar a divergência mecânica entre o modelo computacional e as referências do potencial AIREBO. Para avaliar o impacto do nível de refinamento geométrico, a otimização foi dividida em duas frentes de modelagem com o campo de forças de Morse: uma atomística (todos os átomos de carbono explicitamente representados) e uma *coarse-grained* (CG), que aglutina átomos para reduzir o custo computacional. As propriedades inerciais e geométricas fixas, bem como os domínios contínuos de busca para as variáveis de decisão (D_0, α e k_θ), estão sumarizados na Tabela 1.

O enxame de partículas foi inicializado sob dispersão estocástica uniforme, estritamente contida nos limites de busca definidos. A cinética de otimização fundamentou-se nas diretrizes de seleção paramétrica propostas por Shi e Eberhart (SHI; EBERHART, 1998), adotando o fator de inércia $w = 0{,}728\,994$ (CLERC; KENNEDY, 2002) e os coeficientes de aceleração $c_1 = 0{,}5$ e $c_2 = 2{,}5$, orientando a convergência para uma tolerância residual de 1×10^{-4}. A dinâmica evolutiva do processo de minimização é apresentada na Figura 3.

Tabela 1: *Parâmetros físicos e limites do espaço de busca da otimização PSO.*

Parâmetro	Atomístico	CG
Massa (g/mol)	12,011	48,044
r_0 (Å)	1,420	2,824
Limites de Otimização		
D_0 (kJ/mol)	$[50, 20000]$	$[1000, 20000]$
α (1/Å)	$[0{,}1, 2{,}0]$	$[0{,}1, 2{,}0]$
k_θ (kJ/(mol · rad^2))	$[10, 5000]$	$[10, 5000]$

Tabela 2: *Resultados: Parâmetros do campo de força com valor mínimo de fitness resultantes do PSO*

Parâmetro	Atomístico	CG
D_0 [kJ/mol]	4106,28	20000,00
α [1/Å]	0,6563	0,2974
k_θ [kJ/(mol · rad)]	1014,07	3707,50
Fitness Final (f)	0,0872	0,0893

A Tabela 3 consolida os parâmetros elásticos de primeira ordem (E e G) e as constantes não lineares de terceira ordem (E_3 e G_3) extraídos dos ajustes polinomiais das curvas de tensão-deformação. A análise revela que os modelos fenomenológicos otimizados apresentaram um Módulo de Young na ordem de 1,04 TPa. Este resultado é consistente com o valor teórico clássico aceito para a monocamada de grafeno (~ 1 TPa), corrigindo a subestimativa do modelo de Min e mantendo a coerência com as predições do AIREBO e de Ruiz.

Tabela 3: *Parâmetros elásticos e constantes de terceira ordem extraídos dos ajustes de curva com o campo de força otimizado apresentado na Tabel 2.*

	Tração Uniaxial		Cisalhamento	
Modelo	E (GPa)	E_3 (GPa)	G (GPa)	G_3 (GPa)
AIREBO (Referência)	926,6	−2071,4	252,9	−224,5
Ruiz (RUIZ et al., 2015)	900,0	−975,0	445,0	—
Min (MIN; ALURU, 2011)	—	—	370 ± 10	—
PSO Morse Atomístico	1036,5	−968,0	275,2	9,7
PSO Morse CG	1046,8	−972,3	273,7	33,8

No regime de cisalhamento, os módulos transversais (G) dos modelos de Morse alinharam-se de forma satisfatória com o referencial do AIREBO (≈ 253 GPa), divergindo das estimativas mais elevadas da literatura, o que reflete a priorização da função de custo em relação ao alvo estabelecido. Contudo, um achado notável da otimização reside na predição do comportamento não linear sob tração: embora o algoritmo tenha sido guiado pela minimização do erro em relação ao AIREBO, as constantes de terceira ordem (E_3) identificadas pelo PSO (−968,0 GPa e −972,3 GPa) convergiram quase perfeitamente para o valor de referência independente reportado por Ruiz (−975,0 GPa). Este alinhamento não intencional corrobora fortemente a adequação física da formulação de Morse para descrever a anarmonicidade e o amolecimento das ligações carbono-carbono sob grandes deformações.

Por fim, a extrema similaridade paramétrica entre as respostas do modelo de Morse Atomístico e do modelo *Coarse-Grained* em todas as métricas avaliadas valida categoricamente a estratégia de aglutinação espacial. Os resultados comprovam que o modelo CG preserva a fidelidade mecânica macroscópica da malha completa de átomos, viabilizando a execução de simulações em escalas de comprimento e tempo significativamente maiores com uma expressiva redução do custo computacional.

A dispersão inicial das partículas viabilizou uma ampla exploração do domínio de busca, sucedida por uma descida abrupta do erro na fase exploratória, conforme ilustrado na Figura 3. Contudo, a evolução do enxame estagnou em patamares superiores à tole-

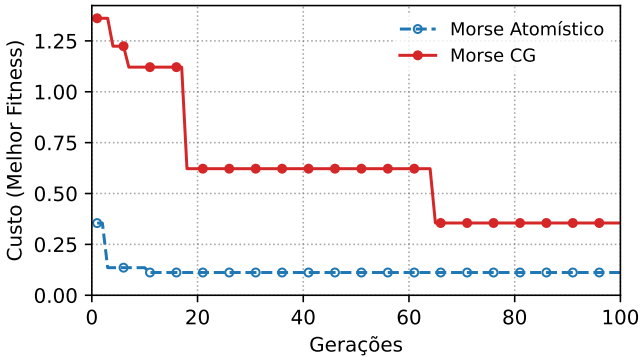


Figura 3: *Evolução da função de aptidão (*fitness*) em função das iterações para um enxame com $N = 32$ partículas. As curvas descrevem a descida monotônica do melhor valor global acumulado (f_{best}) para a parametrização do modelo Morse Atomístico (azul, tracejada) e Morse* Coarse-Grained *(vermelho, linha contínua). Observa-se a rápida estagnação do erro em patamares superiores à tolerância residual (1×10^{-4}), limitados pelas restrições de correlação física da formulação adotada.*

rância residual estrita de 1×10^{-4}, encerrando a execução ao atingir o limite estipulado de 500 gerações. É fundamental ressaltar que essa saturação não configura uma falha do otimizador, mas reflete uma limitação da formulação fenomenológica adotada. Como os modos microscópicos de estiramento de ligação e distorção angular estão fortemente correlacionados em ambos os ensaios macroscópicos (tração e cisalhamento) e as funções de Morse e harmônica no ângulo são aproximações das superfícies de energia potencial do sistema experimental (e também do campo de forças AIREBO), o erro residual atinge um limite inferior físico inevitável (na ordem de $8{,}9 \times 10^{-2}$), restando ao algoritmo apenas identificar o mínimo global da aptidão, em outras palavras, a projeção do modelo ideal no espaço de parâmetros disponível. A incorporação de graus de liberdade adicionais, a exemplo das interações fora do plano (termos diédricos), assim como a inclusão de outras medidas, como tensão e deformação no ponto de rutura, configuram perspectivas promissora para aprimorar a aderência da calibração em trabalhos futuros.

A intensificação deliberada da atração social ($c_2 = 2{,}5$) em detrimento da individual ($c_1 = 0{,}5$) conferiu forte estabilidade direcional ao enxame, mitigando a dependência de flutuações estocásticas da inicialização e assegurando que o espaço de parâmetros fosse explorado de forma estruturada. No entanto, a estagnação acima de zero sugere que o uso de um limiar de aptidão absoluto não é o critério de paragem mais adequado para este tipo de calibração paramétrica. Recomenda-se, para modelagens subsequentes, a adoção de um critério secundário baseado na estagnação relativa da bacia de convergência, interrompendo a otimização após um número fixo de iterações sem evolução significativa.

Apesar do erro residual na função de custo, os parâmetros ótimos identificados reproduziram com elevada fidelidade a resposta mecânica alvo (Tabela 3). Esse nível de acurácia física confirma que a regressão quadrática tensão–deformação com intercepto nulo na origem assegura uma superfície de custo suave e localmente convexa na vizinhança do mínimo. Simultaneamente, a arquitetura distribuída via MPI viabilizou o despacho concorrente das instâncias do LAMMPS com alta eficiência de escalonamento. Dessa forma, consolida-se um fluxo computacional fechado e reprodutível, eliminando por completo a dependência de calibrações empíricas manuais.

4 Conclusão

Neste trabalho, foi desenvolvido e validado um ecossistema computacional em malha fechada para a parametrização automatizada de campos de força moleculares para o grafeno. A integração entre o algoritmo PSO, o orquestrador distribuído via MPI e o simulador LAMMPS viabilizou a identificação autônoma e concorrente dos parâmetros de ligação axial (D_0, α) e constante elástica de flexão angular (k_θ).

A formulação constitutiva quadrática com restrição na origem ($y = E_3x^2 + E_2x$), combinada ao truncamento autônomo nos limites de integridade mecânica sob tração e cisalhamento, proveu uma superfície de custo suave e bem-condicionada para a busca heurística. Guiado pelas diretrizes cinéticas de Shi e Eberhart, o enxame de partículas demonstrou convergência rápida e consistente em direção aos critérios de parada estabelecidos, reproduzindo com fidelidade os módulos elásticos canônicos de referência da literatura.

A metodologia estabelecida comprova a eficácia de sistemas de automação na identificação de propriedades mecânicas fundamentais e estabelece uma base algorítmica extensível para a calibração de modelos mesoscópicos de materiais bidimensionais com geometrias atômicas mais complexas, tais como os dicalcogenetos de metais de transição (TMDs) e nitreto de boro (BN).

Iterações futuras deste trabalho envolverão uma *meta-otimização* do algoritmo PSO, avaliando o efeito de: condições de contorno do espaço de busca; parâmetros cognitivo e social, c_1 e c_2; fator de

[inércia w e critérios de parada. Está prevista a avaliação de outras heurísticas de busca, como Simplex de *Nelder–Mead*(NELDER; MEAD, 1965), algoritmo genético (GA) ou Evolução Diferencial (DE)(WAHAB; NEFTI-MEZIANI; ATYABI, 2015).

Referências

CARVALHO, E. J. F.; SOUZA, J. H. A. *FFFIT: Ajuste heurístico de parâmetros de campo de força para dinâmica molecular*. Natal, 2021. Registro de Programa de Computador no INPI BR512021000174-5.

CLERC, M.; KENNEDY, J. The particle swarm—explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, v. 6, n. 1, p. 58–73, 2002.

KAMINSKI, G. A. et al. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides†. *The Journal of Physical Chemistry B*, American Chemical Society (ACS), v. 105, n. 28, p. 6474–6487, jul 2001. Disponível em: <https://doi.org/10.1021\%2Fjp003919d>.

MANESCO, A. L. R.; PULKIN, A. Flat bands in twisted bilayer transition metal dichalcogenides. *Physical Review B*, v. 106, n. 12, p. 125421, 2022.

MIN, K.; ALURU, N. R. Mechanical properties of graphene under shear deformation. *Applied Physics Letters*, AIP Publishing, v. 98, n. 1, jan. 2011. ISSN 1077-3118. Disponível em: <http://dx.doi.org/10.1063/1.3534787>.

NELDER, J. A.; MEAD, R. A simplex method for function minimization. *The Computer Journal*, Oxford University Press (OUP), v. 7, n. 4, p. 308–313, jan. 1965. ISSN 1460-2067. Disponível em: <http://dx.doi.org/10.1093/comjnl/7.4.308>.

RUIZ, L. et al. A coarse-grained model for the mechanical behavior of multi-layer graphene. *Carbon*, v. 82, p. 103–115, 2015.

SHI, Y.; EBERHART, R. C. Parameter selection in particle swarm optimization. In: *Evolutionary Programming VII*. [S.l.]: Springer, 1998. (Lecture Notes in Computer Science, v. 1447), p. 591–600.

STUART, S. J.; TUTEIN, A. B.; HARRISON, J. A. A reactive potential for hydrocarbons with intermolecular interactions. *The Journal of Chemical Physics*, AIP Publishing, v. 112, n. 14, p. 6472–6486, 2000.

THOMPSON, A. P. et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications*, v. 271, p. 108171, 2022.

WAHAB, M. N. A.; NEFTI-MEZIANI, S.; ATYABI, A. A comprehensive review of swarm optimization algorithms. *PLOS ONE*, Public Library of Science (PLoS), v. 10, n. 5, p. e0122827, may 2015. Disponível em: <https://doi.org/10.1371\%2Fjournal.pone.0122827>.

WILLIAMS, C. D.; LÍSAL, M. Coarse grained models of graphene and graphene oxide for use in aqueous solution. *2D Materials*, v. 7, n. 2, p. 025025, 2020.

ZHAI, D.; SANDLER, N. Electron dynamics in strained graphene. *Modern Physics Letters B*, v. 33, n. 35, p. 1950429, 2019.

Plataforma Linux Embarcada para Integração de Simulador de Bombeio Mecânico com Controlador WEG ALC11

***Diego Moura*[1]**

***Andres Ortiz*[1]**

***André Maitelli*[1]**

***Rodrigo Dacol*[2]**

[1]*Departamento de Engenharia de Computação, UFRN,*
diegomoura@dca.ufrn.br
andres@dca.ufrn.br
maitelli@dca.ufrn.br
[2]*WEG*
rdacol@weg.net

Resumo

Este trabalho apresenta uma plataforma Linux embarcada para integrar um simulador de bombeio mecânico (BM), implementado em Python, ao controlador ALC11 da WEG. O sistema utiliza um módulo *MYIR* com imagem Linux customizada pelo *Yocto Project*. O simulador de BM que está sendo executado na placa *MYIR* envia informações de carga na haste polida e posição ao controlador, de modo que ele interaja com o simulador como um poço físico. A informação de velocidade em ciclos por minuto retorna ao simulador via *Modbus* TCP, que atualiza a execução dos ciclos simulados. Resultados obtidos a partir da comparação entre cartas dinamométricas geradas pelo simulador e aquelas plotadas pelo supervisório validam a proposta da plataforma embarcada.

Abstract

This paper presents an embedded Linux platform designed to integrate a sucker-rod pumping (SRP) simulator, implemented in Python, with the WEG ALC11 controller. The system utilizes a MYIR module running a custom Linux image generated by the Yocto Project. The SRP simulator executing on the MYIR board sends polished rod load and position data to the controller, so that it interacts with the simulator as a physical well. The speed information in strokes per minute is sent back to the simulator via Modbus TCP, which updates the execution of the simulated cycles. Results obtained from the comparison between the dynamometer cards generated by the simulator and those plotted by the supervisory system validate the feasibility of the proposed embedded platform.

Palavras-chave

Yocto Project, bombeio mecânico, Linux embarcado

1 Introdução

O bombeio mecânico (BM) é um dos métodos de elevação artificial utilizados na indústria de petróleo. O monitoramento e o controle desses sistemas envolvem a aquisição de grandezas operacionais, como a carga e a posição da haste polida, utilizadas para avaliar as condições de funcionamento do poço. O desenvolvimento e a validação de controladores de BM podem exigir ensaios em diferentes condições operacionais, cuja realização em instalações reais pode apresentar limitações relacionadas a custos, disponibilidade de equipamentos e segurança. Nesse contexto, plataformas capazes de reproduzir sinais representativos de um sistema de bombeio mecânico constituem uma alternativa para a realização de testes em laboratório. [1].

Em 2022, foi iniciado um projeto para o desenvolvimento de um simulador de bombeio mecânico, implementado em *Python*, capaz de fornecer informações de carga e posição da haste polida.

Para permitir a utilização dessas informações por controladores físicos, foi proposta sua execução em uma plataforma embarcada, possibilitando a conversão dos dados simulados em sinais elétricos compatíveis com as entradas desses equipamentos. Para essa finalidade, foi selecionado o módulo MYC-Y6ULY2-V2-4E512D-50-I, da MYIR, utilizando um sistema operacional Linux customizado por meio do *Yocto Project*, de modo a disponibilizar os recursos necessários à aplicação.

Para integrar o simulador ao controlador físico WEG ALC11, foi desenvolvida uma interface em *Python* responsável pela geração dos sinais elétricos a partir dos dados simulados. A carga na haste polida é reproduzida por meio de um sinal PWM, posteriormente filtrado e condicionado para produzir uma tensão analógica compatível com a entrada do controlador. A referência de posição, por sua vez, é fornecida por um pulso digital que simula o acionamento de uma chave *reed switch*, indicando a passagem da cabeça da unidade de bombeio pelo ponto mais baixo do ciclo. Embora o simulador também disponibilize dados de posição que permitem sua representação por um sinal analógico, essa alternativa não foi implementada nos ensaios apresentados.

A plataforma também recebe do controlador, por meio do protocolo *Modbus*, a informação de velocidade de bombeio, expressa em ciclos por minuto (CPM), permitindo que essa variável seja utilizada pelo simulador. Dessa forma, estabelece-se uma comunicação bidirecional entre a plataforma embarcada e o controlador físico.

Este trabalho apresenta o desenvolvimento de uma plataforma Linux embarcada para integração de um simulador computacional a controladores de bombeio mecânico. As principais contribuições compreendem a customização do sistema operacional utilizando o *Yocto Project*, o desenvolvimento de uma interface de software para gerenciamento dos sinais de entrada e saída e a implementação de circuitos de conversão PWM-analógico, visando à realização de ensaios laboratoriais com controladores de BM.

2 Arquitetura da Plataforma Embarcada

A plataforma desenvolvida tem como elemento central o módulo embarcado MYC-Y6ULY2-V2-4E512D-50-I, da *MYIR*, responsável pela execução do simulador de bombeio mecânico e pelo gerenciamento das interfaces de comunicação com o controlador WEG ALC11. O módulo utiliza uma distribuição Linux customizada por meio do *Yocto Project*, configurada para disponibilizar os recursos de software e os periféricos necessários à aplicação. [2, 3].

A arquitetura proposta integra três componentes principais: o ambiente computacional embarcado, a interface de geração e condicionamento de sinais elétricos e a comunicação com o controlador físico. No ambiente embarcado, uma aplicação *Python* realiza a integração entre os dados fornecidos pelo simulador e os periféricos da placa. A informação de carga na haste polida é convertida em um sinal PWM, posteriormente filtrado e condicionado para produzir um sinal analógico compatível com a entrada do ALC11. A referência de posição é reproduzida por meio de um pulso digital que simula o acionamento de uma chave *reed switch*.

A comunicação no sentido inverso é realizada por meio do protocolo *Modbus* TCP, utilizando a interface Ethernet. Por esse canal, a plataforma recebe do controlador ALC11 a velocidade de bombeio, expressa em ciclos por minuto (CPM), que é encaminhada ao simulador para atualização das condições de operação. Essa organização permite integrar a execução computacional do simulador à aquisição e geração de sinais físicos, estabelecendo uma interface bidirecional com o controlador de bombeio mecânico. A Figura 1 apresenta o diagrama de blocos da arquitetura desenvolvida.

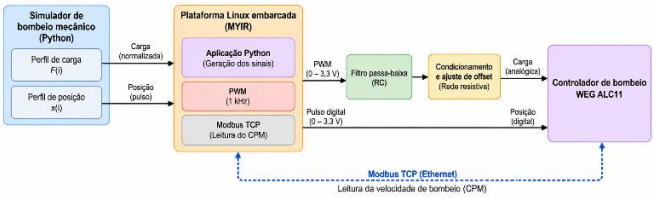


Figura 1: *Representação conceitual da arquitetura da plataforma embarcada e das interfaces com o controlador de bombeio mecânico.*

3 Desenvolvimento e Implementação

3.1 Customização do Sistema Operacional Linux

Para a implementação da plataforma embarcada, foi desenvolvida uma distribuição Linux customizada utilizando o *Yocto Project*, com o objetivo de disponibilizar os recursos necessários à execução da aplicação e à comunicação com os dispositivos externos. Durante esse processo, foram incorporados componentes de software, incluindo *Python*, *mbpoll*, suporte a *WebSocket* e recursos para servidor web, além das dependências necessárias ao funcionamento desses componentes. [4].

Uma das principais modificações realizadas envolveu a configuração dos pinos do barramento de vídeo do módulo *MYIR*, originalmente definidos como entradas e saídas digitais de propósito geral (GPIO). Para isso, foi criado o arquivo *myd-y6ull-emmc-ufrn.dts*, a partir do arquivo original *myd-y6ull-emmc.dts*, fornecido pelo fabricante. O novo arquivo *Device Tree Source* (DTS) foi modificado para habilitar a funcionalidade PWM nos pinos 10, 11, 12 e 13, permitindo sua utilização na geração dos sinais necessários à interface com o controlador de bombeio mecânico. [3, 5].

Após essas modificações, o processo de construção foi executado novamente utilizando a ferramenta *BitBake*,

responsável pelo gerenciamento das tarefas de compilação e geração da imagem Linux. Nesse processo, as alterações realizadas no DTS foram incorporadas ao arquivo binário *Device Tree Blob* (DTB), disponibilizando as novas configurações dos periféricos ao sistema operacional. Com isso, os canais PWM passaram a estar acessíveis às aplicações executadas na plataforma embarcada.

3.2 Desenvolvimento da Interface de Entrada e Saída

A interface de entrada e saída foi desenvolvida para reproduzir eletricamente os sinais de carga e posição da haste polida, a partir dos dados fornecidos pelo simulador de bombeio mecânico. Inicialmente, foi considerada a utilização de um conversor digital-analógico (DAC) externo, porém, devido às limitações encontradas durante sua implementação, optou-se pela utilização dos periféricos PWM disponíveis na plataforma MYIR, acessados por meio da biblioteca *python-periphery*. [6].

Para a geração do sinal de carga, cada amostra da carta dinamométrica é normalizada em relação à capacidade nominal de 30.000 lb da célula de carga considerada no projeto. O valor normalizado é utilizado para definir o ciclo de trabalho (*duty cycle*) do PWM, conforme a Eq. (1):

$$D_i = \frac{trunc\left(100\frac{Fi}{F_{max}}\right)}{100} \qquad (1)$$

em que (D_i) representa o ciclo de trabalho, (F_i) é a carga simulada na amostra (i) e (F_{max} = 30.000) lb é a carga máxima de referência. O truncamento em duas casas decimais estabelece incrementos de 1% no ciclo de trabalho, equivalentes a uma resolução de comando de 300 lb na faixa considerada. Para cargas superiores ao valor de referência, o ciclo de trabalho é limitado a 100%.

O sinal PWM foi configurado com frequência de 10 kHz e amplitude nominal de 3,3 V. Para obter uma tensão analógica proporcional ao ciclo de trabalho, foi implementado um filtro passa-baixa RC, composto por um resistor de 1 kΩ e um capacitor de 1 µF, resultando em uma frequência de corte aproximada de 159 Hz. A tensão filtrada é posteriormente aplicada a uma rede resistiva de atenuação, constituída por um resistor de 330 kΩ e três resistores de 1 kΩ, responsável pela adequação da amplitude do sinal à entrada diferencial de carga do controlador WEG ALC11.

Durante os ensaios iniciais, foram identificadas diferenças entre os níveis de carga produzidos pelo simulador e aqueles apresentados pelo supervisório ALCTOOLS. Para adequar os potenciais elétricos dos sinais aplicados à entrada diferencial, foi desenvolvido um circuito de ajuste de *offset*, formado por um divisor resistivo de 1,15 kΩ e 1 kΩ, alimentado pelos terminais de excitação V+ e V− do próprio ALC11. A tensão obtida no ponto intermediário do divisor é utilizada como referência do circuito de atenuação, deslocando simultaneamente os potenciais dos sinais aplicados aos terminais S+ e S−. Dessa forma, o circuito modifica a tensão de modo comum sem alterar, idealmente, a diferença de potencial que representa a carga simulada. Após essa adequação, observou-se maior correspondência entre os níveis de carga das cartas dinamométricas geradas pelo simulador e aquelas registradas pelo ALCTOOLS.

Para representar a posição da haste polida, foi adotada a geração de um pulso digital correspondente à passagem da cabeça da unidade de bombeio pelo ponto inferior do ciclo. O sinal gerado pela plataforma MYIR é aplicado a um circuito de interface composto por um optoacoplador PC817A e um estágio transistorizado NPN, responsável pela adequação do pulso à entrada digital do controlador. O circuito utiliza a alimentação de 24 V disponibilizada pelo próprio ALC11, mantendo o isolamento galvânico entre a referência da plataforma embarcada e a referência do circuito digital do controlador. Essa separação evita a conexão direta entre os retornos elétricos, contribuindo para preservar as condições de referência da medição diferencial de carga.

Os circuitos de filtragem, atenuação, ajuste de *offset* e condicionamento do pulso digital (Figura 2) foram implementados em uma placa de interface, responsável pela conexão elétrica entre a plataforma MYIR e o controlador ALC11. Embora o simulador também disponibilize dados contínuos de posição, permitindo sua representação por um sinal analógico, essa alternativa não foi implementada nos ensaios apresentados.

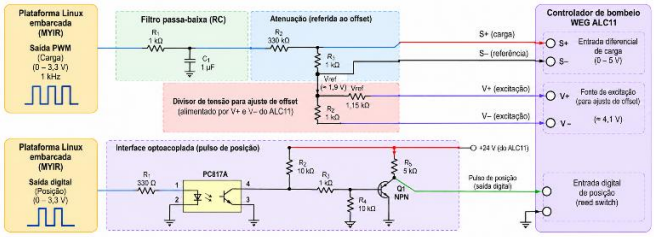


Figura 2: *Circuitos de condicionamento dos sinais de carga e posição: filtragem e ajuste de offset da carga e interface optoacoplada para o pulso digital de posição.*

3.3 Comunicação com o Controlador WEG ALC11

A comunicação entre a plataforma embarcada e o controlador WEG ALC11 foi implementada utilizando o protocolo *Modbus* TCP, por meio da interface Ethernet. Essa comunicação permite obter a velocidade de bombeio disponibilizada pelo controlador, expressa em ciclos por minuto (CPM), e utilizá-la na atualização

da velocidade de execução das cartas dinamométricas geradas pelo simulador. [7, 8].

Para realizar a leitura dos dados, foi utilizado o utilitário *mbpoll*, executado por meio de comandos invocados pela aplicação *Python*. A consulta é direcionada a dois registradores consecutivos do controlador, a partir do endereço 13815, cujos conteúdos de 16 bits são combinados para reconstruir o valor da velocidade em formato de ponto flutuante de 32 bits (*IEEE 754*). Após a conversão, o valor obtido é encaminhado ao simulador.

A integração entre a geração dos sinais e a comunicação *Modbus* TCP foi realizada de forma sequencial. Durante cada ciclo de bombeio, a aplicação atualiza o sinal PWM correspondente à carga e gera o pulso digital de referência de posição. Após a transmissão dos pontos da carta dinamométrica, é realizada uma nova leitura da velocidade no ALC11, cujo valor é utilizado para ajustar a execução dos ciclos seguintes. Esse procedimento foi empregado nos ensaios experimentais, permitindo que a velocidade de reprodução dos sinais acompanhasse as alterações do parâmetro CPM disponibilizado pelo controlador.

A estrutura de comunicação utilizada nos ensaios também contemplou dois computadores auxiliares. O primeiro foi utilizado para acesso remoto à plataforma embarcada por meio do protocolo SSH, enquanto o segundo executou o software ALCTools, da WEG, empregado na configuração e no monitoramento dos parâmetros do controlador ALC11. Todos os equipamentos foram interligados por uma rede Ethernet, conforme apresentado na Figura 3.

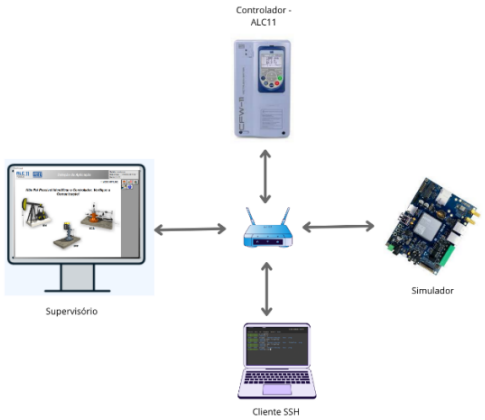


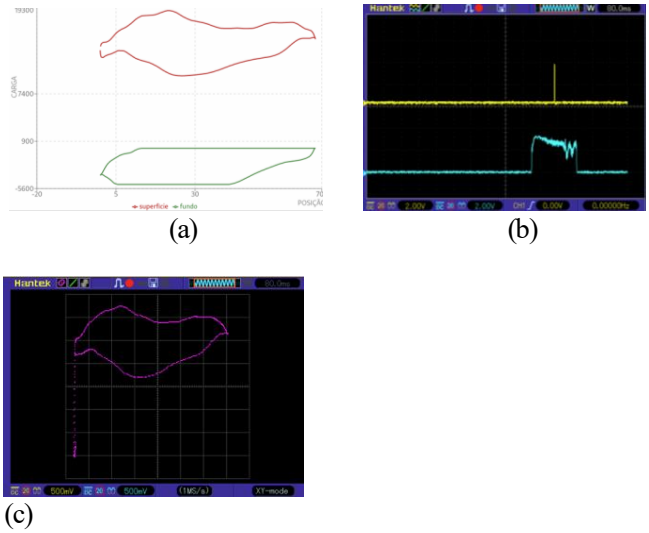
Figura 3: *Estrutura de comunicação Ethernet utilizada nos ensaios, incluindo a plataforma embarcada MYIR, o controlador WEG ALC11 e os computadores destinados ao acesso remoto e ao monitoramento pelo ALCTOOLS.*

4 Resultados Experimentais

A plataforma desenvolvida foi avaliada experimentalmente por meio de sua integração com o controlador de bombeio mecânico WEG ALC11. Os ensaios tiveram como objetivo verificar a geração e o condicionamento dos sinais elétricos, a reprodução das cartas dinamométricas e a comunicação *Modbus* TCP entre o controlador e a aplicação executada no Linux embarcado.

4.1 Validação dos Sinais Elétricos e das Cartas Dinamométricas

A geração dos sinais foi inicialmente avaliada por meio de medições com osciloscópio, considerando a tensão correspondente à carga na haste polida e o pulso digital utilizado como referência de posição. As formas de onda obtidas evidenciaram a reprodução do perfil de carga e a geração do pulso de referência em um ponto específico do ciclo de bombeio, conforme Figura 4.



(a) (b)

(c)

Figura 4: *(a)Carta dinamométrica de referência; obtida diretamente na saída da plataforma MYIR, utilizando o osciloscópio (b)em modo XT e (c)em modo XY.*

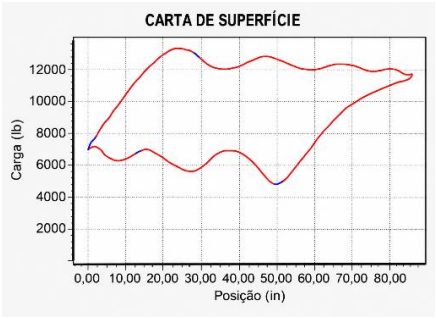
A carta dinamométrica também foi observada no osciloscópio em modo XY, utilizando diretamente os sinais produzidos pela plataforma MYIR, antes de sua aplicação ao controlador ALC11. O perfil obtido apresentou correspondência qualitativa com a carta produzida pelo simulador, permitindo verificar a reprodução dos sinais elétricos pela interface embarcada.

Nos testes iniciais de integração com o ALC11, foram identificadas diferenças entre os valores máximos e mínimos das cartas dinamométricas geradas pelo simulador e aqueles apresentados pelo supervisório ALCTools. Para avaliar a correspondência entre as cartas dinamométricas produzidas pelo simulador e aquelas registradas pelo controlador ALC11, foi utilizado o erro relativo percentual dos valores extremos de carga, calculado conforme a Eq. (2):

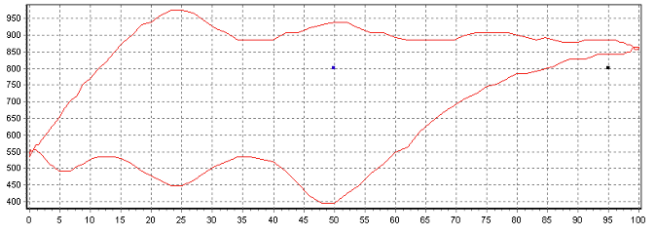
$$E_r = \frac{|F_{ALC11} - F_{sim}|}{|F_{sim}|} x100 \qquad (2)$$

em que (F_{sim}) representa a carga fornecida pelo simulador e (F_{ALC11}) corresponde à carga registrada pelo controlador, ambas expressas em libras (lb).

A Figura 5 apresenta as cartas dinamométricas obtidas antes da implementação do circuito de ajuste de *offset*. Nessa condição, foram observadas diferenças entre os níveis de carga fornecidos pelo simulador e aqueles registrados pelo ALC11. Considerando o fator de escala de 10 aplicado aos valores apresentados pelo ALCTOOLS, os erros relativos estimados foram de 26,69% para a carga máxima e 18,75% para a carga mínima. Essas diferenças motivaram a implementação do circuito de ajuste de *offset*, com o objetivo de adequar os níveis de tensão de modo comum aplicados à entrada diferencial do controlador.



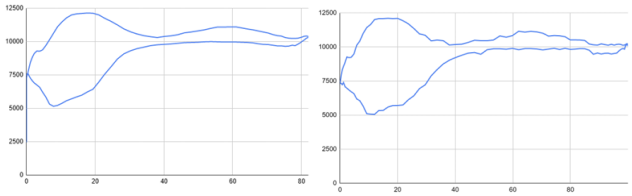
(a)



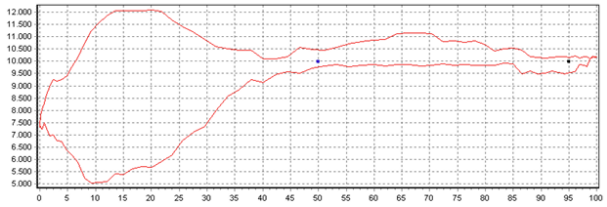
(b)

Figura 5: *Cartas dinamométricas (a) gerada pelo simulador e (b) apresentada pelo ALCTools.*

Após a implementação do circuito, foram realizados novos ensaios utilizando outra carta dinamométrica, apresentada na Figura 6. Nesse caso, a comparação entre os valores extremos da carta simulada e da carta registrada pelo ALC11 apresentou erros relativos estimados próximos de 0% para a carga máxima e de 3,85% para a carga mínima. Os resultados mostram uma maior correspondência dos níveis de carga nessa configuração, além de demonstrar a capacidade da plataforma de reproduzir diferentes perfis de cartas dinamométricas.



(a) (b)



(c)

Figura 6: *Comparação das cartas dinamométricas após o ajuste de offset: (a) simulador; (b) pontos registrados pelo ALCTOOLS; (c) supervisório ALCTOOLS.*

A Tabela 1 reúne os valores extremos e os respectivos erros relativos obtidos nas duas condições. Ressalta-se que os valores foram estimados a partir das escalas gráficas e que os ensaios utilizaram cartas diferentes. Dessa forma, os resultados permitem comparar o desempenho observado em cada configuração, mas não quantificar isoladamente o efeito do circuito de *offset* sobre o erro de reprodução da carga.

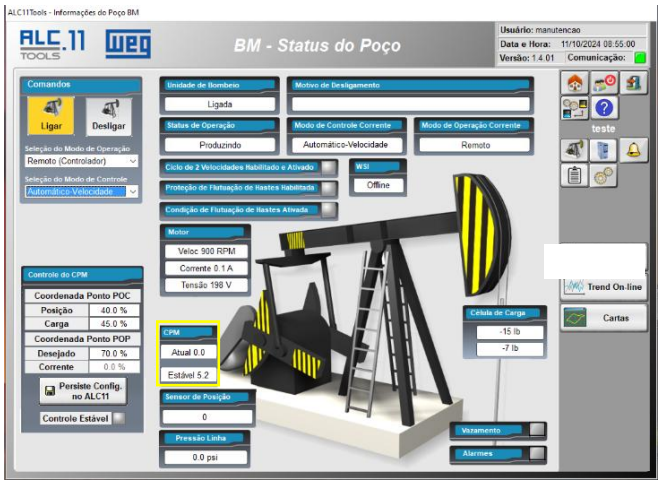
Tabela 1 – Comparação dos valores extremos de carga antes e após o ajuste de *offset*.

Condição	Parâmetro	Simulador (lb)	ALC11 (lb)	Erro relativo
Antes do offset	Carga máxima	13.300	9.750	26,69%
Antes do offset	Carga mínima	4.800	3.900	18,75%
Após o offset	Carga máxima	12.000	12.000	≈0%
Após o offset	Carga mínima	5.200	5.000	≈3,85%

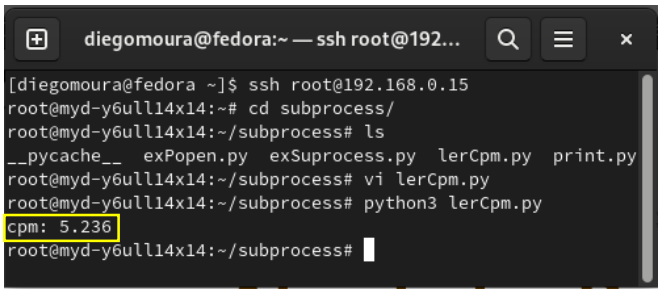
4.2 Validação da Comunicação *Modbus* TCP

A comunicação *Modbus* TCP foi avaliada por meio da leitura dos registradores associados à velocidade de bombeio disponibilizada pelo controlador. Inicialmente, dois registradores de 16 bits foram consultados utilizando o utilitário *mbpoll*. Os valores obtidos foram convertidos para representação binária e

combinados para formar uma palavra de 32 bits, posteriormente interpretada conforme o padrão IEEE 754. [7, 8].
A conversão manual resultou em aproximadamente 5,236 CPM, valor consistente com os 5,2 CPM apresentados pelo software ALCTools, considerando o arredondamento utilizado na interface de supervisão (Figura 7). Em seguida, a conversão foi implementada em *Python*, sendo obtido o valor de 5,236 CPM, novamente correspondente à indicação do controlador.



(a)



(b)

Figura 7: *Validação da comunicação Modbus TCP entre a plataforma Linux embarcada e o controlador WEG ALC11, comparando (a) o valor apresentado no supervisório ALCTOOLS com (b) a velocidade de bombeio obtida pela aplicação Python.*

Durante os ensaios, o valor de CPM recebido pela plataforma embarcada foi utilizado para atualizar a velocidade de execução das cartas dinamométricas nos ciclos seguintes. Esses resultados demonstram o funcionamento da comunicação *Modbus* TCP e sua integração à aplicação de simulação, permitindo a troca de informações entre a plataforma Linux embarcada e o controlador físico.

5 Conclusões

Este trabalho apresentou o desenvolvimento de uma plataforma Linux embarcada para a integração de um simulador de bombeio mecânico ao controlador físico WEG ALC11. A solução foi implementada utilizando um módulo MYIR com sistema operacional customizado por meio do *Yocto Project*, incluindo modificações no *Device Tree* para habilitação dos periféricos PWM e a incorporação dos recursos de software necessários à aplicação.
A interface desenvolvida em *Python* permitiu reproduzir os sinais elétricos de carga na haste polida e de referência de posição, utilizando os canais PWM da plataforma. Para adequação à entrada diferencial do controlador, foram implementados circuitos de filtragem, atenuação e ajuste de *offset*. Os resultados experimentais demonstraram a reprodução das cartas dinamométricas e uma melhor correspondência entre os níveis de carga apresentados pelo simulador e pelo ALC11 após o condicionamento dos sinais.
A comunicação *Modbus* TCP possibilitou a leitura da velocidade de bombeio disponibilizada pelo controlador e sua utilização na atualização dos ciclos de simulação. A comparação entre os valores obtidos pela aplicação *Python* e aqueles apresentados pelo supervisório ALCTOOLS confirmou o funcionamento da leitura e interpretação dos dados.
Os resultados demonstram a viabilidade da plataforma desenvolvida como ferramenta de integração entre simulação computacional e equipamentos físicos de controle de bombeio mecânico. Como continuidade, propõe-se avaliar quantitativamente a fidelidade dos sinais reproduzidos, os efeitos da temporização do sistema embarcado e o desempenho da comunicação em diferentes condições de operação.

Referências

[1] NASCIMENTO, J. M. A. Simulador Computacional para Poços de Petróleo com Método de Elevação Artificial por Bombeio Mecânico. Dissertação (Mestrado), UFRN, Natal, 2005.
[2] MYIR. MYC-Y6ULX Product Manual. Versão 1.3, 2018.
[3] MYIR. MYD-Y6ULX-V2 Development Board Overview. Versão 1.3, 2018.
[4] YOCTO PROJECT. Yocto Project Documentation. Disponível em: https://docs.yoctoproject.org/. Acesso em: 8 nov. 2026.
[5] NXP SEMICONDUCTORS. i.MX 6ULL Applications Processor Reference Manual. 2017.
[6] PYTHON-PERIPHERY. PWM — Python-periphery Documentation. Disponível em: https://python-periphery.readthedocs.io/en/latest/pwm.html. Acesso em: 8 nov. 2026.
[7] MODBUS ORGANIZATION. MODBUS Application Protocol Specification V1.1b3. 2012. Disponível em: https://modbus.org/docs/Modbus_Application_Protocol_V1_1b3.pdf.
[8] MODBUS ORGANIZATION. MODBUS Messaging on TCP/IP Implementation Guide. Versão 1.0b, 2006. Disponível em: https://modbus.org/docs/Modbus_Messaging_Implementation_Guide_V1_0b.pdf.
[9] WEG. Manual do Usuário ALC-11 — Bombeio por Cavidades Progressivas. Série CFW-11, documento 10003747005/02, maio 2022. Disponível em: https://static.weg.net/medias/downloadcenter/h11/h93/WEG-CFW11-ALC11-BCP-manual-de-usuario-10003747005-pt.pdf. Nota: manual da variante BCP; não utilizado como comprovação das especificações elétricas do ALC11 para BM.

Ressintonia automática de controladores PID baseada em estimativa de robustez

Marina L. Araújo [1]
Carlos E. T. Dórea [2]

[1] *Graduanda em Engenharia Mecatrônica, UFRN, <marina.lucena.099@ufrn.edu.br>*
[2] *Departamento de Engenharia de Computação e Automação, UFRN, <cetdorea@dca.ufrn.br>*

Resumo

Neste trabalho, é proposto um método de ressintonia automática de controladores PID baseado na estimativa da robustez do sistema, visando atender a um valor de máxima sensibilidade especificado pelo usuário. A partir de um controlador PID em funcionamento, a abordagem consiste em determinar pontos da resposta em frequência necessários ao cálculo de fatores utilizados para modificar as constantes do controlador. Esses parâmetros são então ajustados de modo que os pontos analisados da resposta em frequência permaneçam fora do círculo de máxima sensibilidade, garantindo a margem de robustez desejada. Por fim, são realizadas simulações e apresentados exemplos numéricos para ilustrar a aplicação do método e verificar o comportamento dos sistemas antes e após a ressintonia.

Palavras-chave: Controladores PID, Ressintonia, Robustez, Máxima Sensibilidade, Resposta em Frequência

Abstract

An automatic PID controller retuning method is proposed based on system robustness estimation, aiming to meet a maximum sensitivity value specified by the user. Starting from an operating PID controller, the method consists of determining frequency-response points required to calculate factors used to modify the controller parameters. These parameters are then adjusted in order to keep the analyzed frequency-response points outside the maximum sensitivity circle, thus ensuring a desired robustness margin. Finally, simulations are performed and numerical examples are presented to illustrate the application of the method and assess system behavior before and after the retuning procedures.

Keywords: PID Controllers, Retuning, Robustness, Maximum Sensitivity, Frequency Response

1 Introdução

Desde a década de 1940, os controladores Proporcional-Integral-Derivativo (PID) consolidaram-se como uma das principais estratégias de controle em sistemas industriais, contribuindo para o aumento da confiabilidade, da acurácia e do desempenho dos processos por meio do ajuste da ação de controle em função do erro (PATIL; JADHAV; PATIL, 2024). Em razão de sua estrutura relativamente simples, facilidade de implementação e manutenção, esses controladores permanecem amplamente empregados em diferentes áreas, como controle de movimento e de processos, eletrônica de potência, manufatura e sistemas hidráulicos e pneumáticos (BORASE et al., 2021).

Estima-se que controladores PID estejam presentes em aproximadamente 90-95% das malhas de controle industriais (BORASE et al., 2021), evidenciando sua ampla adoção e relevância tanto em aplicações convencionais quanto em sistemas tecnológicos mais recentes. Entre esses últimos, destacam-se veículos autônomos, veículos aéreos não tripulados (UAVs) e robôs autônomos (PATIL; JADHAV; PATIL, 2024), nos quais o controle preciso e a capacidade de adaptação às variações do sistema são fundamentais para garantir um desempenho adequado.

Em contrapartida, embora exista uma ampla diversidade de métodos para determinar os parâmetros dos controladores de acordo com objetivos específicos de desempenho, ainda persiste uma importante lacuna no processo de sintonia e ressintonia dessas malhas de controle, pois apenas 1/3 dos processos industriais apresentam desempenho satisfatório (NEVES, 2009).

Nesse contexto, sabe-se que uma parcela significativa das malhas industriais é sintonizada com base nas configurações predefinidas pelos fabricantes dos controladores ou manualmente pelos operadores (ASTRöM; HäGGLUND, 1995). Essa prática pode tornar o processo de sintonia trabalhoso e demorado, além de estar sujeito a inconsistências (HANG; ÅSTRöM; WANG, 2002). Ademais, em muitos casos práticos, os controladores PID são mal sintonizados devido à falta de tempo e à falta de habilidade do operador (VERONESI; VISIOLI, 2009). Portanto, se a performance do controlador PID não for satisfatória, surge a necessidade de ressintonizá-lo (VERONESI; VISIOLI, 2009), muitas vezes por métodos iterativos e, consequentemente, de difícil implementação em CLP (Controlador Lógico-Programável).

Diante desse cenário, REGO et al. (2017) propôs um método para o ajuste dos parâmetros de um controlador PI baseado na estimativa de robustez, com implementação em CLP, baixo custo computacional e utilização de expressões matemáticas simples. Posteriormente, PONTAROLO et al. (2019) estendeu essa abordagem para controladores PID por meio da inclusão do termo derivativo, permitindo a conversão de um controlador PI em um controlador PID, mais robusto e com melhor desempenho da resposta temporal do sistema (PONTAROLO et al., 2019).

Apesar dessa adaptação, as abordagens de REGO et al. (2017) e PONTAROLO et al. (2019) mantêm uma estrutura metodológica semelhante, baseada na estimativa do comportamento do processo a partir do experimento do relé e no cálculo de parâmetros capazes de retirar a curva de Nyquist da malha aberta de uma área circular que representa uma margem de segurança em relação ao ponto crítico do plano complexo (PONTAROLO et al., 2019).

Finalmente, o presente trabalho propõe uma extensão do método de PONTAROLO et al. (2019) para a ressintonia de controladores PID, considerando a existência de um termo derivativo preexistente. Nesse contexto, mais um parâmetro é introduzido para ajustar especificamente o termo derivativo, sendo calculados três fatores de ajuste para modificar as constantes T_d, T_i e K_p do controlador e deslocar a curva de Nyquist para uma região compatível com a máxima sensibilidade especificada no projeto.

2 Medidas de Robustez

2.1 Critério de Estabilidade de Nyquist

O critério de estabilidade de Nyquist permite avaliar a estabilidade em malha fechada a partir da resposta em frequência da função de transferência de malha aberta $L(s)$. Nesse caso, o sistema será estável em malha fechada se, e somente se, o número de voltas que o diagrama de Nyquist de $L(s)$ dá em torno do ponto crítico $(-1, 0)$ for igual ao número de polos de $G(s)$ no semi-plano direito (SKOGESTAD; POSTLETHWAITE, 2005). Portanto, o ideal para um sistema sem polos no semi-plano direito é evitar a proximidade de sua malha aberta do ponto crítico (PONTAROLO et al., 2019).

2.2 Margens de ganho e de fase

A margem de ganho MG de um sistema corresponde ao fator máximo pelo qual o ganho da malha aberta pode ser aumentado antes que a curva de Nyquist atinja o ponto crítico, levando o sistema ao limite de estabilidade. É definida por

$$MG = \frac{1}{|L(j\omega_u)|}, \tag{1}$$

em que ω_u é chamada frequência de cruzamento de fase, definida por $\angle L(j\omega_u) = -\pi$.

Por outro lado, a margem de fase MF corresponde à defasagem adicional que pode ser introduzida na malha aberta, na frequência de cruzamento de ganho, antes que o sistema atinja o limite de estabilidade. É calculada por

$$MF = \pi + \angle L(j\omega_c), \tag{2}$$

em que ω_c é chamada frequência de cruzamento de ganho, definida por $|L(j\omega_c)| = 1$.

2.3 Máxima Sensibilidade

A função de sensibilidade $S(s)$ de um sistema é definida pela razão entre as variações relativas das funções de transferência da malha fechada $T(s)$ e da planta $G(s)$ (SKOGESTAD; POSTLETHWAITE, 2005), sendo calculada como

$$S(s) = \frac{\dfrac{\Delta T(s)}{T(s)}}{\dfrac{\Delta G(s)}{G(s)}} = \frac{1}{1 + L(s)}, \tag{3}$$

em que $L(s)$ é a função de transferência de malha aberta. A função $S(s)$ pode ser utilizada para a veri-

ficação da atenuação de distúrbios, o que só ocorrerá caso $|S(j\omega)| < 1$ (ASTRöM; HäGGLUND, 2006). Assim, a sensibilidade máxima M_s corresponde à frequência em que há maior sensibilidade do sistema a perturbações, isto é:

$$M_s = \max_{\omega} \left| \frac{1}{1 + L(j\omega)} \right| \tag{4}$$

Esse parâmetro é representado no diagrama de Nyquist pelo chamado círculo M_s, uma circunferência centrada no ponto crítico $(-1, 0)$ de raio $\frac{1}{M_s}$. Ao projetar um controlador que garanta que a curva de Nyquist permaneça fora do círculo M_s para qualquer valor de frequência ω, assegura-se a robustez do sistema, implicando em valores de margens de ganho MG e de fase MF que estabilizam o sistema e minimizam o efeito das perturbações (PONTAROLO et al., 2019).

Åström e Hägglund (2006) definem como valores aceitáveis M_s entre 1.4 e 2. Logo, como uma especificação de M_s implica diretamente no valor das margens de estabilidade, é possível relacioná-las para estabelecer limites inferiores para essas margens, como segue:

$$MG \geq \frac{M_s}{M_s - 1} \tag{5}$$

$$MF \geq 2 \arcsin\left(\frac{1}{2M_s}\right) \tag{6}$$

3 Método de Ressintonia

3.1 Ajuste do controlador PID

Considera-se que o controlador PID de um sistema realimentado foi previamente parametrizado por alguma regra de sintonia. A proposta de ressintonia consiste em modificar seus parâmetros de modo a reposicionar a curva de Nyquist da malha aberta de acordo com uma especificação de robustez definida pela máxima sensibilidade M_s. Para isso, parte-se da forma do controlador PID no domínio da frequência, dada por:

$$C_{PID}(j\omega) = K_p \left(1 - \frac{j}{\omega T_i} + j\omega T_d\right) \tag{7}$$

em que K_p, T_i e T_d correspondem, respectivamente, ao ganho proporcional, ao tempo integral e ao tempo derivativo do controlador.

Para o ajuste, são introduzidos os parâmetros α, β e γ, que permitem modificar, respectivamente, a contribuição do termo integral, o ganho do controlador e a contribuição do termo derivativo preexistente. Assim, o controlador PID ressintonizado $C_{\text{PID},r}(j\omega)$ é expresso no domínio da frequência por:

$$C_{\text{PID},r}(j\omega) = \frac{K_p}{\beta} \left(1 - \frac{j}{\omega T_i \alpha} + j\omega T_d \gamma\right), \tag{8}$$

em que $\alpha > 0$, $\beta > 0$ e $\gamma > 0$. Nesse contexto, β modifica o ganho do controlador e, consequentemente, o módulo da curva de Nyquist, enquanto α e γ alteram as contribuições de fase de seus respectivos termos do controlador, além do comportamento da curva.

Nas subseções seguintes, são apresentados os procedimentos para determinação dos parâmetros α e β e, posteriormente, do parâmetro γ, a partir da máxima sensibilidade, possibilitando o reposicionamento da curva de Nyquist conforme a especificação de robustez adotada.

3.2 Cálculo de α e β

Estendendo as definições do método de REGO et al. (2017) para os controladores PID, o parâmetro α permite ajustar a fase dos pontos da curva de Nyquist de malha aberta, enquanto β move esses pontos radialmente, expandindo ou contraindo a curva.

Com a finalidade de manter o sistema seguro em relação à robustez e à instabilidade, os cálculos de α e β levam em conta uma margem de estabilidade reduzida. Isto é, nos procedimentos dessa seção, deve-se utilizar o valor da máxima sensibilidade como o valor do M_s desejado adicionado de 0.2, incremento determinado empiricamente por Pontarolo (2019). Em suma, entende-se que

$$M_s' = M_s + 0,2 \tag{9}$$

em que M_s corresponde à real especificação do projeto.

Com isso, o parâmetro α pode ser calculado de forma simples, com vistas a deslocar o ponto $L(j\omega_i)$ para interceptar o círculo M_s' (REGO et al., 2017). Nesse procedimento, considera-se $L(j\omega_i)$ como um ponto da malha aberta entre $L(j\omega_u)$ e $L(j\omega_c)$. É definido como o ponto da curva de Nyquist de frequência intermediária ω_i, dada por

$$\omega_i = \sqrt{\omega_u \omega_c} \tag{10}$$

Dessa forma, como contribuição desta pesquisa, o cálculo do parâmetro α é modificado para considerar a fase de um controlador PID e passa a ser definido por

$$\alpha = \frac{1}{(\omega_i T_i (\omega_i T_d - \tan(\angle C_1(j\omega_i)))} \tag{11}$$

em que $\angle C_1(j\omega)$ corresponde à contribuição de fase do controlador após a inclusão de α e pode ser calculada pelos seguintes procedimentos:

$$\Re(L_1(j\omega_i)) = \frac{\left(\frac{1}{M_s'}\right)^2 - (|L(j\omega_i)|)^2 - 1}{2} \tag{12}$$

$$\angle L_1(j\omega_i) = \arctan\left(\sqrt{\left(\frac{|L(j\omega_i)|}{\Re(L_1(j\omega_i))}\right)^2 - 1}\right) - \pi \tag{13}$$

$$\Delta(\angle L(j\omega_i)) = \angle L_1(j\omega_i) - \angle L(j\omega_i) \tag{14}$$

$$\angle C_1(j\omega_i) = \Delta(\angle L(j\omega_i)) + \arctan\left(T_d\omega_i - \frac{1}{T_i\omega_i}\right) \tag{15}$$

É importante notar que o valor de α calculado nesta etapa só poderá ser utilizado se o ângulo $\Delta(\angle L(j\omega_i))$ necessário ao ajuste puder ser provido por um controlador PID, cuja fase está sujeita aos limites:

$$-\frac{\pi}{2} \leq \angle C_{\text{PID}}(j\omega) \leq \frac{\pi}{2} \tag{16}$$

Caso contrário, o mesmo procedimento é feito com a frequência ω_c, verificando novamente a condição em (16) para $\Delta(\angle L(j\omega_c))$. Em último caso, se o ajuste de fase calculado ainda não for viável, toma-se $\alpha = 1$ e mantém-se o controlador original. Adicionalmente, embora seja recomendado verificar que $\alpha > 0$, considera-se que para um controlador inicial razoavelmente sintonizado, o ajuste de fase exigido é moderado. Nesse caso, prevê-se que a fase total desejada não superará o avanço fornecido pelo termo derivativo, mantendo α naturalmente positivo na prática.

Com isso, acrescentando a contribuição de α na função de $C(j\omega)$, é possível obter $C_1(j\omega)$ e $L_1(j\omega)$, as novas funções de transferência do controlador e de malha aberta. Consequentemente, pode-se calcular as posições atualizadas dos três pontos: $L_1(j\omega_u)$, $L_1(j\omega_c)$ e $L_1(j\omega_i)$. Se pelo menos um desses novos pontos ainda estiver dentro do círculo M_s', é necessário determinar o valor de β. Caso contrário, adota-se $\beta = 1$, diretamente (REGO et al., 2017).

Assim, o valor de β é dado por

$$\beta = \frac{|L_1(j\omega_u)|}{|L_2(j\omega_u)|} \tag{17}$$

em que $|L_2(j\omega_u)| =$

$$\left|\frac{-1 + \sqrt{-\left(\frac{\Im(L_1(j\omega_u))}{\Re(L_1(j\omega_u))}\right)^2 + \left(\frac{1}{M_s'}\right)^2 + \left(\frac{\Im(L_1(j\omega_u))}{M_s'\Re(L_1(j\omega_u))}\right)^2}}{\sqrt{\left(\frac{\Im(L_1(j\omega_u))}{\Re(L_1(j\omega_u))}\right)^2 + 1}}\right| \tag{18}$$

Esse mesmo procedimento é feito para os outros pontos $L_1(j\omega_c)$ e $L_1(j\omega_i)$, desconsiderando aqueles que já estiverem fora do círculo. Em seguida, toma-se β como o maior valor dentre os três calculados, assegurando a remoção dos três pontos de dentro do círculo M_s'. Feito isso, é realizada a mesma etapa de atualização do controlador, da função de transferência de malha aberta e dos pontos da curva, considerando agora a inclusão do parâmetro β.

3.3 Cálculo de γ

Como resultado deste trabalho, o método de ressintonia de PONTAROLO et al. (2019), que inclui um termo derivativo a um controlador inicial PI, foi adaptado para ressintonizar um controlador inicial PID em um outro PID mais robusto. Dessa forma, é evidente que os cálculos realizados devem considerar o valor inicial da constante de tempo derivativa.

Para isso, é introduzido mais um parâmetro de ajuste das constantes do controlador: γ, que passa a multiplicar a constante de tempo derivativa. Assim, de maneira semelhante ao ajuste da constante de tempo integral e do ganho proporcional por α e β, respectivamente, γ atua diretamente sobre T_d, modificando a contribuição de fase da ação derivativa do controlador.

Para $\gamma = 1$, mantém-se o valor original de T_d, não havendo alteração na ação derivativa. Para $0 < \gamma < 1$, a contribuição da ação derivativa e a fase positiva fornecida pelo controlador são reduzidas, o que tende a deslocar a curva de Nyquist no sentido de reduzir sua fase. Por outro lado, para $\gamma > 1$, ocorre um aumento da ação derivativa e da sua contribuição de fase, o que provoca um deslocamento da curva no sentido de aumentar sua fase. Dessa forma, o ajuste de γ permite modificar a posição da curva de Nyquist por meio do controle da contribuição de fase da ação derivativa, possibilitando aproximar o ponto de maior sensibilidade dos limites do círculo M_s.

Nesse sentido, a fase de um controlador PID em uma determinada frequência ω pode ser calculada por

$$\angle C_{\text{PID}}(j\omega) = \arctan\left(T_d\omega - \frac{1}{T_i\omega}\right) \tag{19}$$

Então, até esta etapa do processo de ressintonia, a fase do controlador com a contribuição de α corres-

ponde a

$$\angle C_2(j\omega) = \arctan\left(T_d\omega - \frac{1}{T_i\alpha\omega}\right) \tag{20}$$

Note que, na equação (20), T_d corresponde ao valor da constante de tempo derivativa do controlador inicial, até agora não alterada pelo método, enquanto a constante de tempo integral já está ajustada por α.

Após o cálculo dos parâmetros α e β, constata-se que a curva de Nyquist pode permanecer no interior do círculo M_s para uma determinada faixa de frequência. Portanto, é necessário determinar o ponto mais sensível da curva de Nyquist, isto é, aquele mais próximo do ponto crítico (-1,0). Para isso, são calculados mais dois pontos, a partir de duas frequências intermediárias.

A frequência ω_{i2} dada por

$$\omega_{i2} = \sqrt{\omega_u\omega_i} \tag{21}$$

que denota o ponto $L(j\omega_{i2})$, entre $L(j\omega_u)$ e $L(j\omega_i)$.

E a frequência ω_{i3} dada por

$$\omega_{i3} = \sqrt{\omega_i\omega_c} \tag{22}$$

que denota o ponto $L(j\omega_{i3})$, entre $L(j\omega_i)$ e $L(j\omega_c)$.

Com isso, é possível determinar o ponto mais sensível entre cinco candidatos possíveis: $L_2(j\omega_u)$, $L_2(j\omega_i)$ e $L_2(j\omega_c)$, determinados após o cálculo de α e β para o círculo M_s reduzido, além de $L_2(j\omega_{i2})$ e $L_2(j\omega_{i3})$, calculados nesta etapa. Conhecendo-se esse ponto mais sensível entre os cinco, denominado agora de $L_2(j\omega_s)$, considera-se ω_s como a frequência associada a ele, que será usada no cálculo de γ para o ajuste de T_d.

Os procedimentos a seguir consideram que $L_2(j\omega_s)$ é o ponto escolhido da curva de Nyquist após o ajuste por α e β, enquanto $L_3(j\omega_s)$ é o correspondente que resultará da inclusão de γ. É importante ressaltar que, neste processo, o valor de M_s considerado deve retornar ao original, visando deslocar $L_2(j\omega_s)$ para fora do círculo de objetivo.

Inicialmente, calcula-se a parte real de $L_3(j\omega_s)$ por:

$$\Re(L_3(j\omega_s)) = \frac{\left(\frac{1}{M_s}\right)^2 - (|L_2(j\omega_s)|)^2 - 1}{2} \tag{23}$$

e a sua fase por:

$$\angle(L_3(j\omega_s)) = \arctan\left(\sqrt{\left(\frac{|L_2(j\omega_s)|}{\Re(L_3(j\omega_s))}\right)^2 - 1}\right) - \pi \tag{24}$$

Em seguida, é mensurada a contribuição de fase dada pelo controlador após o ajuste de T_d, necessária para mover $L_2(j\omega_s)$ para os limites do círculo M_s:

$$\angle C_3(j\omega_s) = \Delta(\angle L(j\omega_s)) + \angle C_2(j\omega_s) \tag{25}$$

em que

$$\Delta(\angle L(j\omega_s)) = \angle L_3(j\omega_s) - \angle L_2(j\omega_s) \tag{26}$$

sendo $\angle L_3(j\omega_s)$ e $\angle L_2(j\omega_s)$ as fases do ponto $L(j\omega_s)$, respectivamente, depois e antes do ajuste por γ. Ademais, $\angle C_2(j\omega_s)$ é dado segundo a equação (20). Logo, obtém-se a nova fase do controlador ressintonizado:

$$\angle C_3(j\omega_s) = \arctan\left(\gamma\ T_d\omega - \frac{1}{T_i\alpha\omega}\right) \tag{27}$$

que pode ser reescrita como

$$\tan(\angle C_3(j\omega_s)) = \gamma\ T_d\omega - \frac{1}{T_i\alpha\omega} \tag{28}$$

Portanto,

$$\gamma = \frac{\tan(\angle C_3(j\omega_s)) + \left(\frac{1}{T_i\alpha\omega_s}\right)}{\omega_s\ T_d} \tag{29}$$

em que $\angle C_3(j\omega_s)$ é obtida pela equação (27).

Por fim, é necessário verificar se a variação de fase $\Delta(\angle L(j\omega_s))$ para a qual γ foi calculado satisfaz à equação (16), tal que a frequência ω_s selecionada seja adequada para calcular esse parâmetro. Caso contrário, a primeira frequência escolhida é descartada e considera-se, sucessivamente, o próximo ponto de maior sensibilidade que atenda à restrição de fase. Repete-se esse procedimento até encontrar uma frequência admissível, tomando-se $\gamma = 1$ caso contrário.

3.4 Resumo do método

1. Incrementar M_s segundo a Equação (9)
2. Calcular a frequência intermediária com a Equação (10)
3. Calcular o parâmetro α pela Equação (11)

4. Incorporar a contribuição de α ao controlador, obtendo $C_1(j\omega)$ e $L_1(j\omega)$

5. Resolver a Equação (17) de β para os pontos $L_1(j\omega_u)$, $L_1(j\omega_c)$ e $L_1(j\omega_i)$, adotando o maior valor obtido

6. Incorporar a contribuição de β ao controlador, obtendo $C_2(j\omega)$ e $L_2(j\omega)$

7. Calcular as frequências ω_{i2} e ω_{i3}, e seus pontos $L_2(j\omega_{i2})$ e $L_2(j\omega_{i3})$ com as Equações (21) e (22)

8. Identificar o ponto de máxima sensibilidade $L_2(j\omega_s)$ entre os cinco conhecidos

9. Calcular γ utilizando M_s original e o ponto $L_2(j\omega_s)$, pela Equação (29)

10. Incorporar a contribuição de γ ao controlador, obtendo o PID final ressintonizado $C_3(j\omega)$ e $L_3(j\omega)$

4 Implementação

Inicialmente, o método de ressintonia de controladores PID foi desenvolvido analiticamente, a partir das relações e condições apresentadas anteriormente. Em seguida, foi realizada sua implementação computacional em linguagem *Python*, com o objetivo de verificar os cálculos e avaliar o comportamento do método em diferentes condições de operação.

Para isso, utilizou-se a biblioteca *Python Control Systems Library* (FULLER et al., 2021), que fornece recursos para a representação e análise de sistemas de controle, incluindo funções de transferência, construção das malhas de controle e geração de diagramas de Nyquist.Logo, foi possível simular os sistemas controlados e analisar características como o rastreamento de referência e a rejeição de perturbações.

Os códigos foram desenvolvidos recebendo como parâmetros de entrada apenas a função de transferência da planta $G(s)$, os parâmetros do controlador inicial K_p, T_i e T_d, e a máxima sensibilidade desejada após o ajuste M_s. As rotinas permitem obter as funções de transferência finais do controlador $C_3(s)$ e da malha aberta $L_3(s)$, além dos pontos da resposta em frequência, possibilitando visualizar o afastamento da curva de Nyquist dos limites da margem de estabilidade, conforme intencionado. Por fim, os resultados obtidos computacionalmente são apresentados na sequência, a fim de verificar o funcionamento dos procedimentos propostos.

5 Resultados

Considerando uma planta com função de transferência dada por

$$G(s) = \frac{1}{(s^2 + 7s + 4.5)} 6e^{-0.2s},$$

foi inicialmente empregado um controlador PID com parâmetros $K_p = 3$, $T_i = 1.6$ e $T_d = 0.0688$ (PONTAROLO et al., 2019), resultando em

$$C(s) = 3\left(1 + \frac{1}{1.6s} + 0.0688s\right).$$

Em seguida, o método proposto neste trabalho foi aplicado à malha aberta definida por $L_1(s) = G(s)C(s)$, adotando-se como especificação de robustez o valor $M_s = 1.5$. Assim, foram calculados $\alpha \approx 0.4534$, $\beta = 1.34106$ e $\gamma \approx 2.04666$ e, como resultado, os novos parâmetros do controlador $K_p \approx 2.237$, $T_i \approx 0.725$ e $T_d \approx 0.141$, com o maior ajuste sobre a ação derivativa. O controlador resultante da ressintonia é, portanto, dado por

$$C_3(s) = 2.237\left(1 + \frac{1}{0.725s} + 0.141s\right).$$

O efeito da ressintonia sobre o sistema pode ser observado nos diagramas de Nyquist apresentados na Figura 1, cujas curvas correspondem à malha inicial e à malha após a aplicação do método proposto. Nota-se que, após a ressintonia, a curva de Nyquist se modificou de modo a atender à especificação de robustez estabelecida por $M_s = 1.5$, representada pelo círculo de raio $\frac{1}{M_s} \approx 0.667$ no plano complexo.

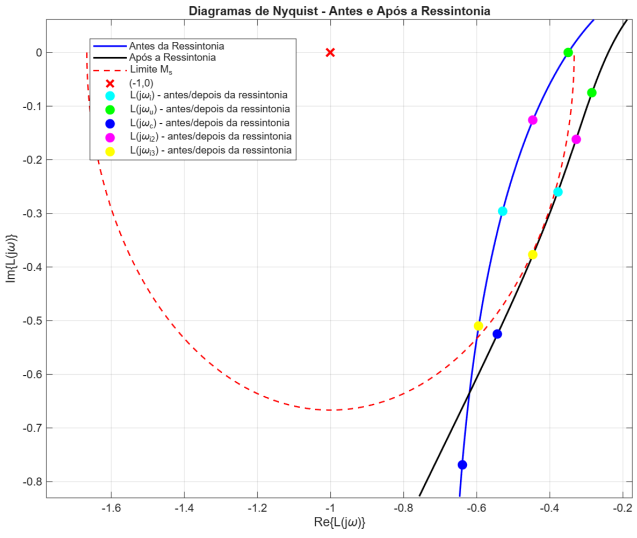


Figura 1: *Comparativo das curvas de Nyquist antes e depois da ressintonia (Caso 1)*

Para além da resposta em frequência, o compor-

tamento transitório do sistema também foi avaliado para uma entrada do tipo degrau, antes e depois da ressintonia. A Figura 2 apresenta as respostas do sistema com os controladores inicial e ressintonizado, considerando uma perturbação de amplitude 0.1 aplicada em $t = 5\,\text{s}$. A comparação entre as curvas evidencia uma melhoria na rejeição da perturbação, acompanhada, no entanto, por um maior sobressinal no seguimento de referência após a ressintonia.

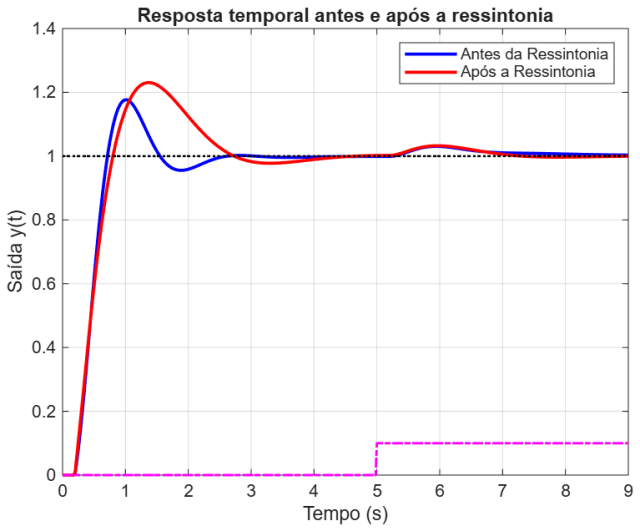


Figura 2: *Comparativo da resposta transitória antes e depois da ressintonia (Caso 1)*

Em segunda análise, para uma planta com atraso de transporte, descrita pela função de transferência

$$G(s) = \frac{1}{(s+1)^3}e^{-s},$$

foi considerado um controlador inicial com $K_p = 0.8$, $T_i = 1.6$ e $T_d = 1.03$, isto é

$$C(s) = 0.8\left(1 + \frac{1}{1.6s} + 1.03s\right).$$

(VERONESI; VISIOLI, 2009).

Nesse caso, o método foi aplicado visando obter $M_s = 1.5$ para o sistema ressintonizado. Como resultado, obteve-se $\alpha \approx 0.91$, $\beta = 1$ e $\gamma \approx 1.555$, implicando em $K_p = 0.8$, $T_i \approx 1.456$ e $T_d \approx 1.602$, com o maior ajuste sobre a ação derivativa. O controlador resultante da ressintonia é, portanto, dado por

$$C_3(s) = 0.8\left(1 + \frac{1}{1.456s} + 1.602s\right).$$

O resultado desse procedimento é ilustrado pelas Figuras 3 e 4, que apresentam, respectivamente, os diagramas de Nyquist antes e após a ressintonia, bem como a resposta temporal de regime transitório para o sistema em ambas as condições. Nesse caso, a mesma perturbação é aplicada em $t = 30\,\text{s}$ e observa-se a redução do sobressinal da resposta do sistema após a ressintonia.

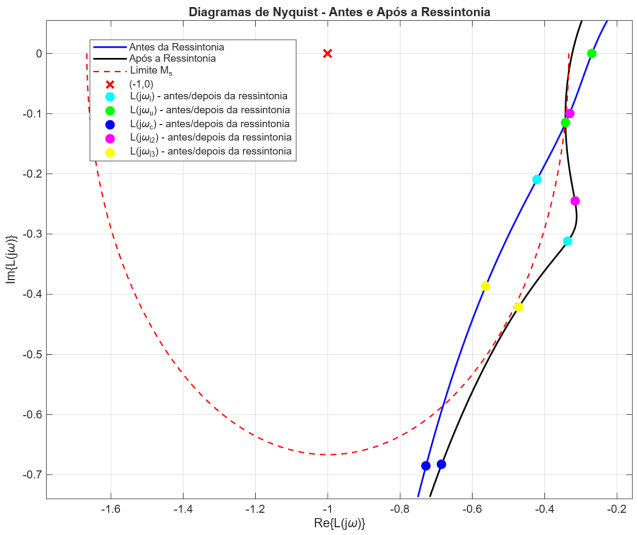


Figura 3: *Comparativo das curvas de Nyquist antes e depois da ressintonia (Caso 2)*

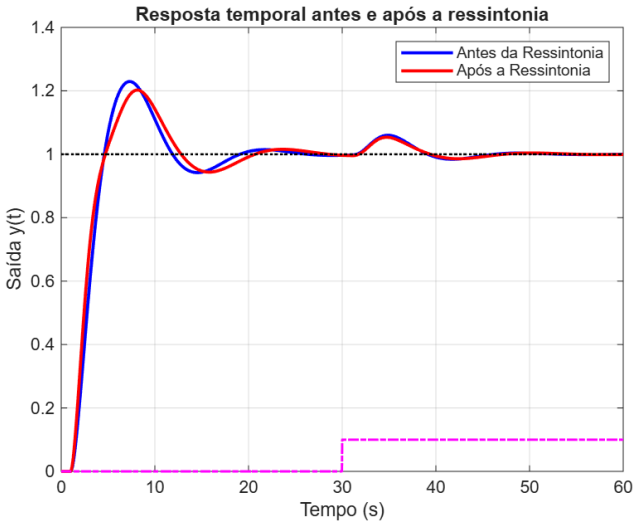


Figura 4: *Comparativo da resposta transitória antes e depois da ressintonia (Caso 2)*

6 Conclusão

Neste trabalho foi proposto um método de ressintonia de controladores PID baseado na estimativa de robustez do sistema pela máxima sensibilidade M_s, a partir da extensão do método de PONTAROLO et al. (2019) para controladores que já possuem um termo derivativo. Os resultados da implementação em Python permitiram verificar os cálculos propostos e avaliar o comportamento do sistema antes e após a ressintonia, indicando a viabilidade do procedimento para o ajuste de controladores PID previamente parametrizados.

Como continuidade desta pesquisa, pretende-se estender a abordagem para seguir uma especificação de robustez por margem de disco, capaz de considerar conjuntamente as variações de ganho e fase. Diferentemente das margens clássicas de ganho e de fase, que avaliam separadamente variações de ganho ou de fase, a margem de disco considera perturbações simultâneas dessas duas grandezas, proporcionando uma avaliação mais abrangente da robustez do sistema (SEILER; PACKARD; GAHINET, 2020).

Agradecimentos

O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), por meio dos processos nº 180754/2026-2 e 313368/2026-0.

Referências

ASTRöM, K. J.; HäGGLUND, T. *PID Controllers: Theory, Design, and Tuning*. [S.l.]: ISA – Instrumentation, Systems, and Automation Society, 1995.

ASTRöM, K. J.; HäGGLUND, T. *Advanced PID Control*. Research Triangle Park, North Carolina: ISA – Instrumentation, Systems and Automation Society, 2006.

BORASE, R. P. et al. A review of pid control, tuning methods and applications. *International Journal of Dynamics and Control*, v. 9, p. 818–827, 2021.

FULLER, S. et al. The python control systems library (python-control). In: IEEE. *60th IEEE Conference on Decision and Control (CDC)*. [S.l.], 2021. p. 4875–4881.

HANG, C. C.; ÅSTRöM, K. J.; WANG, Q. G. Relay feedback auto-tuning of process controllers: tutorial review. *Journal of Process Control*, v. 12, n. 2, p. 143–162, 2002.

NEVES, M. G. S. *Auto-tuning de Controladores PID pelo Método Relay*. Dissertação (Dissertação de Mestrado) — Universidade Técnica de Lisboa, Lisboa, Portugal, 2009.

PATIL, R. S.; JADHAV, S. P.; PATIL, M. D. Review of intelligent and nature-inspired algorithms-based methods for tuning pid controllers in industrial applications. *Journal of Robotics and Control (JRC)*, v. 5, n. 2, p. 336–358, 2024.

PONTAROLO, L. A. et al. Ajuste de controlador pid por método de autossintonia baseado em estimativa de robustez. In: *Congresso Rio Automação 2019*. Rio de Janeiro, RJ, Brasil: [s.n.], 2019.

REGO, E. J. C. et al. Ressintonia automática de controladores pi embarcados em clp baseada em estimativa de robustez. In: *Anais do XIII Simpósio Brasileiro de Automação Inteligente*. Porto Alegre, RS, Brasil: [s.n.], 2017. p. 1733–1737. ISSN 2175-8905.

SEILER, P.; PACKARD, A.; GAHINET, P. An introduction to disk margins. *IEEE Control Systems*, v. 40, n. 5, p. 78–95, 2020.

SKOGESTAD, S.; POSTLETHWAITE, I. *Multivariable Feedback Control*. [S.l.]: John Wiley & Sons, 2005.

VERONESI, M.; VISIOLI, A. Performance assessment and retuning of PID controllers. *Industrial & Engineering Chemistry Research*, v. 48, n. 5, p. 2616–2623, 2009.

Indústria sem telas: perspectivas para o uso de dispositivos vestíveis assistidos por modelos de linguagem de pequeno porte

Mateus Araujo[1], Leonardo Cavalcanti[1], Alice Freire[1], Morsinaldo Medeiros[1], Hilton Machado [1], Erick Justino [1], Marianne Silva[2], Carlos M. D. Viegas[1], Ivanovitch Silva[1]

[1] *Dep. Eng. de Computação e Automação (DCA), UFRN*
<mateus.araujo.117@ufrn.edu.br>,
<leonardo.cavalcanti.136@ufrn.edu.br>,
<alice.victorino.061@ufrn.edu.br>,
<morsinaldo.medeiros.075@ufrn.edu.br>,
<hilton.machado.088@ufrn.edu.br>,
<erick.justino.701@ufrn.edu.br>, <carlos.viegas@ufrn.br>,
<ivanovitch.silva@ufrn.br>
[2] *Curso de Sistemas de Informação, UFAL*
<marianne.silva@penedo.ufal.br>

Resumo

Sistemas industriais são tradicionalmente operados por meio de interfaces baseadas em telas, como interfaces homem-máquina (IHMs), sistemas supervisórios, planilhas e painéis. Este trabalho investiga uma perspectiva alternativa de interação: a *indústria sem telas*, apoiada por dispositivos vestíveis, interação por voz e modelos de linguagem de pequeno porte (*Small Language Models* - SLMs). Essa perspectiva se alinha à Indústria 5.0, especialmente ao princípio da centralidade no ser humano, ao buscar formas de interação mais naturais e contextuais, sem necessidade de recorrer a interfaces visuais convencionais. As possibilidades de aplicação são discutidas nos domínios de manutenção, produção, operação e assistência, considerando quatro óculos inteligentes comerciais, sendo um deles acompanhado de um anel de controle. Propõe-se também uma arquitetura de referência que integra dispositivos vestíveis, processamento na borda e sistemas de dados industriais. Como estudo de caso, apresenta-se uma implementação automotiva, adotada como ambiente análogo de validação por suas restrições visuais e de mãos ocupadas, que permite consultar, por comandos de voz predefinidos, indicadores obtidos da telemetria OBD-II durante a condução. Os resultados indicam a viabilidade técnica da abordagem com dispositivos de consumo, com respostas rápidas (TTFA mediano inferior a 500 ms) e baixa sobrecarga de comunicação e de processamento. O estudo estabelece, assim, uma base experimental de referência para a posterior integração de SLMs na borda e para a extensão dos testes a ambientes fabris reais.

Abstract

Industrial systems traditionally rely on screen-based interfaces, including human-machine interfaces (HMIs), supervisory systems, spreadsheets, and dashboards. This work explores the screenless industry, combining wearable devices, voice interaction, and Small Language Models (SLMs) in line with the human-centric principle of Industry 5.0. Applications in maintenance, production, operation, and assistance are discussed, along with four commercial smart glasses, one paired with a control ring. A reference architecture integrating wearable devices, edge computing, and industrial data systems is proposed. As a preliminary case study, an automotive implementation uses voice commands to query OBD-II telemetry while driving, a setting with hands-free and visual-attention constraints analogous to industrial tasks. The prototype relies on predefined, rule-based commands; SLM integration remains future work. Results indicate technical feasibility with consumer-grade devices, median TTFA below 500 ms, and low communication and processing overhead. The study provides an experimental baseline for future edge-

SLM integration and validation in industrial environments.

1 Introdução

Historicamente, a interação entre profissionais e sistemas industriais tem sido mediada por interfaces visuais. Nos sistemas de automação tradicionais, informações e comandos são concentrados em painéis físicos compostos por instrumentos, indicadores, chaves, botoeiras e diagramas. Com a digitalização da automação industrial, grande parte dessas funções migrou para estações de operação baseadas em interfaces homem-máquina (IHMs) e sistemas supervisórios, que passaram a concentrar em telas, a supervisão e o controle dos processos. Mais recentemente, planilhas, sistemas corporativos e, com a Indústria 4.0, *dashboards* ampliaram essa logística ao integrar e resumir dados provenientes de diferentes sistemas da planta (XU et al., 2021). Assistentes baseados em modelos de linguagem de grande porte (LLMs) acrescentam uma nova camada a essas interfaces, permitindo consultar dados e gerar análises em linguagem natural.

Apesar dessa evolução, permanece uma característica comum: na maior parte dessas interfaces, o profissional ainda precisa ir até a informação. Consultar um painel, monitor, *tablet* ou *smartphone* exige desviar a atenção da tarefa e, em muitos casos, ocupar as mãos para interagir com o dispositivo. Essa característica é particularmente relevante no trabalho de campo, em atividades de manutenção, inspeção e operação, nas quais a atenção visual e a disponibilidade das mãos fazem parte da própria execução da tarefa.

A Indústria 5.0 reposiciona o trabalhador no centro dos sistemas produtivos e propõe uma indústria centrada no ser humano, resiliente e sustentável (MARTÍNEZ et al., 2026; GRAZIANI et al., 2026), na qual a tecnologia amplia as capacidades do operador em vez de disputar sua atenção (ROMERO et al., 2016; GIUGLIANO et al., 2026). Décadas antes, Weiser (1991) já defendia que as tecnologias mais profundas são aquelas que se integram de forma tão natural ao cotidiano que deixam de demandar atenção consciente. A convergência dessas perspectivas convida a repensar não apenas o que a tecnologia pode fazer pelo trabalhador, mas também a forma como ele precisa interagir com ela. Surge, assim, a questão que orienta este trabalho: e se a próxima interface industrial fosse sem telas?

A Figura 1 resume essa trajetória. Neste trabalho, uma *indústria sem telas* é entendida como aquela em que o acesso à informação deixa de depender de monitores, painéis ou dispositivos que o profissional precise segurar. A informação passa a acompanhar o trabalhador no contexto da tarefa, podendo ser solicitada por voz e apresentada por áudio ou, quando necessário, por elementos visuais no campo de visão.

Nesse cenário, a interação entre o trabalhador e os sistemas industriais pode ser mediada por dispositivos vestíveis que ofereçam diferentes modalidades de entrada e saída, como voz, toque, áudio e apresentação de informações no campo de visão. Entre esses dispositivos, os óculos inteligentes se destacam por poderem reunir essas modalidades em uma única armação, enquanto outros vestíveis, como anéis, podem atuar como interfaces complementares de controle e sensoriamento. O desenvolvimento comercial dos óculos inteligentes passou por diferentes fases, inicialmente com propostas voltadas ao consumidor final e, posteriormente, com maior direcionamento ao mercado corporativo e industrial. Nesse processo, um dos primeiros dispositivos a ganhar ampla visibilidade foi o Google Glass, apresentado ao público em 2013. Após sua baixa adoção como produto de consumo, o dispositivo foi relançado em 2017 em uma edição voltada ao mercado corporativo (LEVY, 2017). Fabricantes como Epson e Microsoft também lançaram seus próprios óculos, aplicados sobretudo em contextos profissionais, como saúde, manutenção e treinamento (KIM; CHOI, 2021). Mais recentemente, a integração de assistentes de inteligência artificial renovou o interesse por esses dispositivos no mercado de consumo, incluindo modelos com *display* disponíveis no varejo (META, 2026), ampliando sua disponibilidade para a prototipagem de aplicações industriais.

O dispositivo vestível, entretanto, representa apenas a interface com o trabalhador. Para que a interação seja contextual, é necessário interpretar o pedido e consultar os sistemas da planta relacionados à tarefa em andamento, papel que pode ser desempenhado por modelos de linguagem capazes de traduzir a fala em consultas e ações sobre esses sistemas. Os LLMs, porém, costumam ser executados na nuvem, o que acrescenta latência e dependência de conectividade a cada interação (FRIHA et al., 2024). Modelos de linguagem de pequeno porte (*Small Language Models*, SLMs) já podem desempenhar essa função localmente, em *smartphones* ou em computadores de borda instalados na própria planta (LAMAAKAL et al., 2025; BELCAK et al., 2025). Com dispositivos vestíveis comercialmente disponíveis e modelos capazes de operar localmente, a ideia de uma indústria sem telas torna-se tecnicamente plausível.

Para explorar essa perspectiva, este artigo apre-

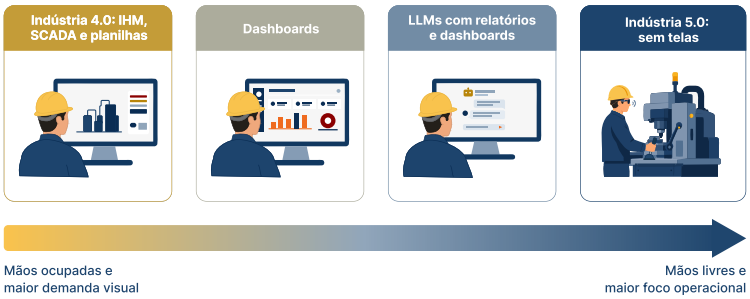


Figura 1: *Evolução das interfaces industriais: da supervisão por telas à interação inteligente e não intrusiva no ambiente operacional.*

senta: (i) a visão de uma indústria sem telas, com aplicações organizadas em quatro domínios (manutenção, produção, operação e assistência); (ii) a análise de quatro óculos inteligentes comerciais, um deles acompanhado de um anel de controle, quanto ao custo, às formas de interação e às possibilidades de uso em ambientes industriais; (iii) uma arquitetura de referência que integra dispositivos vestíveis, processamento na borda com SLMs e sistemas de dados; e (iv) um estudo de caso no setor automotivo, adotado como análogo operacional de cenários industriais. Esse ambiente reproduz condições severas de sobrecarga atencional, mãos ocupadas e necessidade de leitura contínua de sensores. Durante a condução real, o motorista consulta indicadores do veículo por meio da voz. Nesta etapa, o protótipo utiliza comandos predefinidos por regras para avaliar a infraestrutura física de comunicação. Com isso, estabelece-se uma linha de base de desempenho que orientará, em trabalhos subsequentes, a incorporação de SLMs na borda e a transposição do sistema para células fabris.

2 Trabalhos Relacionados

Na indústria, óculos inteligentes e dispositivos de realidade aumentada vêm sendo avaliados como interfaces vestíveis de apoio ao operador, com revisões recentes que comparam os modelos disponíveis por suas características técnicas e ergonômicas e recomendam critérios de escolha conforme a aplicação (SOLOMASHENKO et al., 2025). No contexto da manutenção de máquinas industriais, Angelopoulos, Manettas e Alexopoulos (2025) combinaram realidade aumentada e um LLM para sugerir ações ao técnico diante de falhas inesperadas, guiando a execução por sobreposições visuais, e observaram que um modelo pequeno ajustado ao domínio produziu respostas mais específicas que um modelo de propósito geral. Já Nussbaum, Legler e Ruskowski (2025) propuseram um assistente para montagem manual, baseado em um LLM hospedado localmente e em geração aumentada por recuperação, com interação exclusivamente por voz e áudio.

Para consultar sistemas industriais, um modelo de linguagem precisa saber quando acionar ferramentas externas (SCHICK et al., 2023). A viabilidade dessa capacidade em modelos compactos executados localmente é o foco do estudo de Kavathekar et al. (2025), que avalia a geração de chamadas de função por SLMs, com medidas de latência e memória em um dispositivo de borda e testes de robustez a *prompt injection*, e da investigação de Lu et al. (2025), que analisa a capacidade, o aprendizado em contexto e as oportunidades de otimização de mais de 60 SLMs voltados a *smartphones* e dispositivos de borda.

No domínio veicular, a telemetria OBD-II sustenta diagnóstico, análise do comportamento de condução e estimativa de emissões (MEDEIROS; SILVA; SILVA, 2025; ANDRADE et al., 2024). Do lado do motorista, telas sensíveis ao toque tendem a prejudicar a segurança, enquanto interfaces por voz são, em geral, consideradas mais seguras (KRSTAČIĆ; ŽUŽIĆ; OREHOVAČKI, 2024), embora um estilo de conversa informal tenda a aumentar a distração cognitiva (DITTRICH; KO; RIEGER, 2025). Este trabalho defende que essas frentes sejam pensadas em conjunto e ilustra essa visão com uma instância da arquitetura proposta, avaliada em condução real.

3 Óculos inteligentes na indústria

O uso de óculos inteligentes na indústria não implica necessariamente substituir as interfaces convencionais, mas oferecer formas complementares de acesso à informação durante atividades que exigem atenção ao ambiente e liberdade das mãos. Sua adoção, contudo, depende tanto de seus custos quanto das funcionalidades e limitações dos dispositivos. Nesse contexto, esta seção analisa alternativas comerciais e discute suas possibilidades de aplicação industrial.

3.1 Custo de aquisição

O uso de dispositivos vestíveis de cabeça, durante muito tempo, esteve associado a equipamentos especializados de alto custo. Modelos como HoloLens 2 (2019), Magic Leap 2 (2022) e RealWear Navigator 520 (2023) exemplificam essa realidade, com preços no mercado americano entre US$ 2.900 e US$ 3.500[1].

Esse cenário, no entanto, vem mudando com a chegada de alternativas mais acessíveis. Óculos como X3 Pro e Air 4 Pro da RayNeo (RAYNEO, 2026), Ray-Ban Meta Gen 2 (META, 2026) e Even G2 (EVEN REALITIES, 2026), lançados entre 2025 e 2026, oferecem, em diferentes combinações, recursos de comunicação por voz e de exibição de informações no campo de visão. Esses modelos chegaram ao mercado por valores entre US$ 299 e US$ 1.299, uma fração do custo das soluções especializadas, o que reduz a barreira econômica para sua adoção por equipes de manutenção.

3.2 Óculos avaliados

O laboratório do grupo Conect2AI da UFRN adquiriu quatro modelos de óculos inteligentes de consumo para prototipagem e experimentos, apresentados na Tabela 1. Esses dispositivos combinam, de diferentes maneiras, recursos como *display*, câmera, microfone e alto-falante, representando abordagens distintas para aplicações industriais.

O X3 Pro é o único entre os modelos avaliados que reúne processamento embarcado, sistema baseado em Android, câmera e *display* colorido, permitindo sobrepor informações ao ambiente real. Essas características o tornam particularmente adequado à apresentação de instruções passo a passo e a tarefas de visão computacional. Diferentemente do X3 Pro, o Air 4 Pro depende de um dispositivo externo, como um *smartphone* ou computador, ao qual se conecta por USB-C, exibindo os conteúdos diretamente no campo de visão do usuário. Essa abordagem permite aproveitar sistemas supervisórios e aplicações já existentes, podendo dispensar o desenvolvimento de um *software* específico para os óculos.

Por outro lado, o Ray-Ban Meta Gen 2 não possui *display*, mas reúne câmera, microfones e alto-falantes em uma armação semelhante à de óculos convencionais. Seu kit de desenvolvimento permite que aplicativos de terceiros acessem a câmera, enquanto o acesso aos microfones e aos alto-falantes ocorre por meio dos perfis Bluetooth do sistema.

[1] https://vr-expert.com/en-us e https://www.vr52.com/ar

Tabela 1: *Recursos dos óculos inteligentes disponíveis no laboratório Conect2AI.*

	RayNeo X3 Pro	**RayNeo Air 4 Pro**	**Ray-Ban Meta Gen 2**	**Even G2**
Display	✓	✓	✗	✓
Câmera	✓	✗	✓	✗
Microfone	✓	✗	✓	✓
Alto-falante	✓	✓	✓	✗
Executa aplicativos	✓	✗	✗	✗
Kit de desenvolvimento	✓	✗	✓	✓

Por fim, o Even G2 prioriza a discrição e possui *display* monocromático, mas não dispõe de câmera nem de alto-falantes. O anel R1 complementa suas funcionalidades, permitindo interagir sem comandos de voz, o que pode ser útil em ambientes ruidosos.

Essas diferenças mostram que a escolha do dispositivo depende menos de uma superioridade tecnológica em geral e mais dos requisitos de interação, processamento e integração de cada aplicação.

3.3 Potencialidades

As potencialidades da indústria sem telas vão além da substituição de monitores e terminais convencionais. O uso de dispositivos vestíveis permite que informações, instruções e alertas sejam apresentados de forma contextual, mantendo as mãos do trabalhador livres e reduzindo a necessidade de interromper a tarefa para consultar uma interface externa. Nesse cenário, aplicações de manutenção, produção, operação e assistência podem se beneficiar de interações baseadas em voz, imagem e informações apresentadas diretamente no campo de visão.

Na manutenção, o dispositivo poderia funcionar como uma etapa adicional de verificação após a execução do serviço. Depois de realizar o cabeamento ou o reparo de um painel elétrico, por exemplo, o técnico poderia solicitar uma inspeção do conjunto, permitindo que o sistema analise visualmente o painel e identifique possíveis inconsistências ou problemas que tenham passado despercebidos durante a realização da manutenção.

Na produção, a aplicação serviria como um mecanismo de apoio à verificação de etapas críticas de montagem. Em um processo no qual a instalação segura de um componente, como um *airbag*, depende da ocorrência de uma sequência específica de cliques de travamento, o sistema poderia monitorar continuamente os sinais sonoros gerados durante o procedimento. Ao término da montagem, poderia então confirmar se todos os eventos esperados foram identificados com base no áudio captado.

Na operação, os óculos poderiam estabelecer um elo entre o trabalhador em campo e um especialista remoto, permitindo que situações que exijam conhecimento técnico específico sejam acompanhadas e solucionadas à distância. Por meio da câmera e do áudio integrados, o especialista poderia visualizar a situação a partir da perspectiva do operador, orientá-lo na realização de procedimentos e esclarecer dúvidas em tempo real.

Na assistência, a tecnologia ofereceria acesso a informações relevantes para a tarefa ou situação vivenciada pelo usuário. Em atividades técnicas, isso pode envolver a consulta a especificações, parâmetros de operação ou procedimentos associados a determinado equipamento. No contexto veicular, pode auxiliar o motorista na avaliação de seu comportamento de condução ou na interpretação de alertas e indicadores presentes no painel do carro. Dessa forma, o dispositivo funcionaria como um assistente capaz de fornecer informações relevantes de maneira contextualizada e relacionadas à situação vivenciada pelo usuário.

4 Arquitetura Proposta

A arquitetura proposta organiza a interação em três camadas complementares, sem depender de telas convencionais, conforme ilustra a Figura 2. A camada vestível permite ao usuário expressar suas intenções e receber respostas. A camada de borda interpreta essas intenções e determina como atender às solicitações, recorrendo, quando necessário, às informações disponíveis na camada de dados. Essa organização prioriza o processamento próximo ao usuário para reduzir a latência. Também separa o sensoriamento contínuo das interações sob demanda e utiliza a linguagem natural como interface de acesso às funcionalidades dos sistemas. Além disso, busca minimizar a sobrecarga computacional e operacional associada ao uso dos dispositivos vestíveis.

Os dispositivos apresentados na Tabela 1 oferecem diferentes combinações de recursos. Na camada vestível, a arquitetura utiliza esses recursos para adaptar a interação às características de cada dispositivo e às condições do ambiente. Assim, uma resposta pode ser apresentada por áudio ou diretamente no campo de visão do usuário, considerando, por exemplo, o nível de ruído e a disponibilidade de um *display*.

A camada de borda é executada em um *smartphone* ou *gateway* industrial. Essa camada concentra o processamento responsável por interpretar as solicitações e produzir respostas. Nas interações por voz, o fluxo começa com uma palavra de ativação, seguida pelo reconhecimento automático de fala (ASR), que converte o áudio em texto. A interpretação pode utilizar regras, no caso de comandos com vocabulário limitado, ou um SLM executado localmente. Essa segunda abordagem permite compreender solicitações mais variadas e gerar chamadas de função (BELCAK et al., 2025; KAVATHEKAR et al., 2025).

A intenção identificada é associada a uma *skill*, definida como uma função com parâmetros tipados que consulta as informações na camada de dados. Inspirada no uso de ferramentas por modelos de linguagem (SCHICK et al., 2023), essa abordagem limita as consultas a funções previamente definidas e auditáveis. Isso reduz o risco de apresentar valores críticos não verificados, como o torque de aperto durante uma manutenção. A confiabilidade das respostas também depende da integridade e da atualização dos dados consultados. Os resultados são convertidos em linguagem natural por meio de *templates* ou do próprio SLM. As respostas são apresentadas por síntese de voz (TTS) ou no *display*. Um mecanismo de contexto e salvaguardas considera o usuário, o equipamento envolvido e as condições de uso. Com isso, é possível aplicar restrições de segurança, como permitir apenas consultas de leitura durante a condução.

Por fim, a camada de dados reúne as fontes de informações já disponíveis no ambiente industrial, incluindo a telemetria de veículos e máquinas, dados de controladores industriais, sistemas de supervisão e gestão, além de manuais e procedimentos técnicos. Quando necessário, a arquitetura também permite recorrer a um especialista remoto. Em vez de substituir esses sistemas, a arquitetura amplia o acesso a eles por meio da linguagem natural.

5 Estudo de caso no setor automotivo

O setor automotivo é um exemplo de ambiente em que a interação sem telas pode trazer benefícios reais. Durante a condução, manter a atenção na via e as mãos no volante é essencial para a segurança. Nesse contexto, a necessidade de consultar informações em telas no painel pode levar o motorista a desviar o olhar da estrada, contribuindo para a distração e possíveis acidentes de trânsito (KRSTAČIĆ; ŽUŽIĆ; OREHOVAČKI, 2024). Por outro lado, os veículos modernos monitoram continuamente diversos parâmetros de funcionamento e condução, produzindo dados de telemetria que permitem compreender o comportamento do veículo e do motorista.

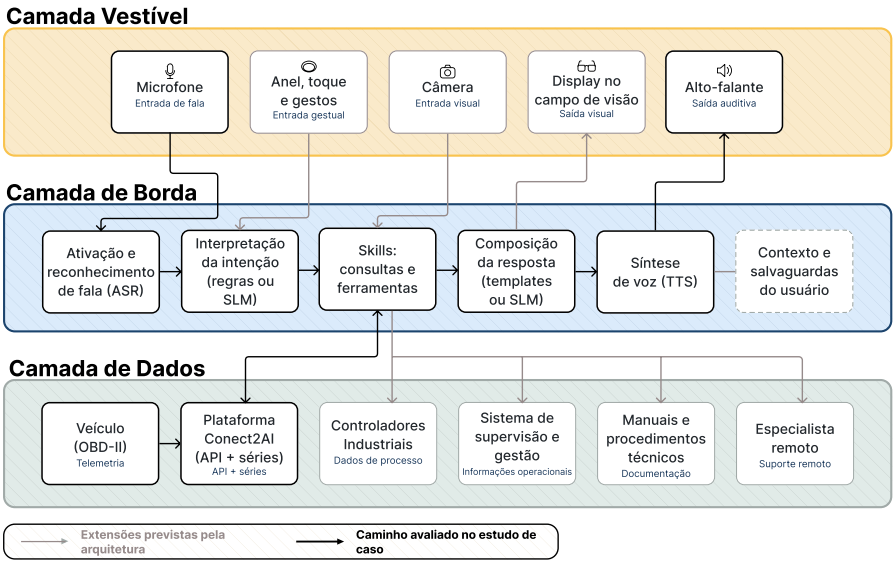


Figura 2: *Arquitetura de referência da indústria sem telas com a especificação das três camadas: vestível, de borda e de dados.*

Parte dessas informações pode ser acessada por meio da interface OBD-II e utilizada para oferecer assistência contextualizada. Assim, uma pergunta simples: "como estou dirigindo?", pode ser respondida por voz, com base nos dados do próprio veículo, sem exigir que o motorista desvie a atenção da via.

5.1 Instância da arquitetura

Para explorar as possibilidades de interação sem telas no ambiente automotivo, os autores desenvolveram o *Glasses2Car*, um aplicativo que permite ao motorista consultar informações sobre sua condução e sua viagem por meio de comandos de voz. A proposta é aproveitar os dados de telemetria do veículo para oferecer informações contextualizadas, utilizando óculos inteligentes como interface de interação e um *smartphone* como plataforma de processamento e integração. A implementação avaliada segue o caminho destacado na Figura 2, organizando-se em dois fluxos que compartilham o *smartphone* como plataforma de borda. No fluxo de sensoriamento, o aplicativo recebe os parâmetros OBD-II obtidos por intermédio de um adaptador ELM327 e encaminha a telemetria à plataforma Conect2AI, responsável por estimar as emissões de CO_2 e armazenar os dados (MACHADO et al., 2025). No fluxo de interação, os óculos Ray-Ban Meta Gen 2 capturam a voz do motorista e reproduzem a resposta, enquanto o aplicativo *Glasses2Car*, desenvolvido com o *Meta Wearables Device Access Toolkit*[2], gerencia a interação no *smartphone*.

[2]https://developers.meta.com/wearables

Na implementação desenvolvida, a interação por voz é iniciada pela frase de ativação "*Hey Conecta*", que permite ao usuário realizar três consultas relacionadas à viagem. A consulta *"How am I driving?"* retorna a avaliação do comportamento recente de condução como cauteloso, normal ou agressivo (MEDEIROS; SILVA; SILVA, 2025); *"How am I emitting?"* informa as emissões acumuladas de CO_2 ao longo da viagem (ANDRADE et al., 2024); e *"Trip summary"* apresenta, ao final do percurso, um resumo que reúne esses indicadores, a distância percorrida e a duração da viagem. Nesta versão, a interpretação é baseada em regras predefinidas, enquanto as respostas são construídas a partir de *templates*, com ambas as etapas executadas localmente no *smartphone*. Essa implementação estabelece uma linha de base funcional para a evolução da arquitetura, especialmente para a futura incorporação de um SLM na interpretação das solicitações.

5.2 Experimentos e Resultados

Foram realizadas duas sessões experimentais em um percurso urbano de aproximadamente 6 km, em Natal (RN), utilizando o veículo Volkswagen Polo 1.0 Turbo 2019 automático. O aplicativo foi executado em um iPhone 16 Pro, operando com sua própria bateria e conexão móvel 4G/5G. Em cada sessão, foram realizadas 11 interações em inglês, sendo seis consultas sobre o comportamento de condução, quatro sobre as emissões de CO_2 e uma solicitação de resumo ao final da viagem, totalizando 22 interações. A principal métrica de desempenho foi o *Time to First Audio* (TTFA), definido como o intervalo en-

tre o reconhecimento do comando e o início da resposta falada, desconsiderando o tempo de fala do motorista. Também foram monitorados o tráfego de rede do *Glasses2Car*, o uso de CPU, o consumo de memória, o estado térmico e o nível de bateria do *smartphone*.

Nas 22 interações realizadas, todos os comandos foram reconhecidos e respondidos, enquanto a telemetria era coletada e transmitida continuamente. Entretanto, a robustez do reconhecimento de voz diante do ruído da cabine não foi avaliada especificamente. Nas 20 consultas realizadas durante a condução, o TTFA mediano foi de 440 ms para o comportamento de condução e de 390 ms para as emissões de CO_2, valores inferiores ao limite de 2 s utilizado como referência para a resposta de agentes conversacionais automotivos (FUNK et al., 2020). A consulta à API respondeu pela maior parte desse tempo, enquanto a formatação das respostas e o início da síntese de voz contribuíram com apenas alguns milissegundos. Já as duas consultas de resumo da viagem apresentaram latências de 6,5 e 8,3 s, devido à necessidade de consolidar os dados de toda a sessão e finalizar seu registro no servidor, diferentemente das demais consultas, que acessam apenas dados recentes. Esses resultados indicam que a latência observada está associada principalmente ao processamento realizado pelos serviços consultados, variando conforme a complexidade das operações solicitadas.

A sobrecarga introduzida pela interface vestível foi pequena. Nas 22 interações, o *Glasses2Car* trocou cerca de 24,1 KiB com a plataforma em decorrência das consultas realizadas pelo vestível, volume equivalente a aproximadamente 5% dos 475 KiB de telemetria veicular transmitidos no mesmo período. No *smartphone*, o aplicativo apresentou uso mediano de CPU de 0,7% de um núcleo e consumo de cerca de 29 MB de memória. O aparelho permaneceu no estado térmico nominal, com variação do nível de bateria inferior a 1% nas duas sessões.

A interação sem depender de telas foi demonstrada com um dispositivo de consumo de US$ 379 e um *smartphone* de uso geral. Nas condições avaliadas, a principal limitação funcional do protótipo não esteve no *hardware*, mas no repertório restrito a três comandos em inglês, que exige do motorista conhecer previamente as consultas disponíveis. A incorporação de um SLM na etapa de interpretação permitiria formular perguntas de maneira mais flexível, inclusive em português, mantendo sua associação ao mesmo conjunto auditável de *skills*. Essa evolução, entretanto, acrescentaria o tempo de inferência do modelo ao TTFA. A diferença de aproximadamente 1,5 s entre as medianas observadas e a referência de 2 s sugere uma margem para incorporar esse processamento adicional, hipótese que será avaliada na próxima etapa do trabalho.

5.3 Limitações e riscos

A avaliação envolveu um único veículo, motorista, percurso, modelo de óculos e *smartphone*, sem comparação direta com interfaces convencionais ou experimentos em ambiente fabril. Assim, os resultados demonstram a viabilidade técnica nas condições avaliadas, mas não permitem generalizações sobre o desempenho em outros cenários nem conclusões sobre segurança viária. A robustez do reconhecimento de fala em ambientes ruidosos, a carga cognitiva e a usabilidade ainda precisam ser investigadas em estudos com usuários (DITTRICH; KO; RIEGER, 2025; REIMER et al., 2026). Além disso, a presença de câmeras em dispositivos vestíveis levanta questões de privacidade, enquanto mecanismos de autenticação inadequados podem expor os sistemas a acessos não autorizados (ZHANG et al., 2024). A palavra de ativação "*Hey Conecta*" também não permite verificar a identidade de quem realiza a solicitação. Nesse contexto, restringir as *skills* a operações de leitura em situações críticas, conforme previsto no mecanismo de salvaguardas, constitui uma medida inicial para reduzir os riscos associados ao controle indevido dos sistemas.

6 Conclusão

Este artigo propôs uma perspectiva para os próximos anos da automação: uma indústria em que a informação acompanha o trabalhador, em vez de permanecer em uma tela. Essa visão foi explorada em quatro domínios de aplicação, na comparação de quatro modelos comerciais de óculos inteligentes, mais acessíveis que as soluções especializadas, e em uma arquitetura que integra vestíveis, processamento na borda com SLMs e sistemas de dados existentes. O estudo de caso automotivo demonstrou seu potencial com dispositivos de consumo: as consultas durante a condução apresentaram tempos medianos de resposta inferiores a meio segundo, com baixa sobrecarga de comunicação e processamento.

Como próximos passos, pretende-se incorporar um SLM à camada de borda e avaliar seu impacto na latência, além de explorar os recursos de *display* e câmera dos demais óculos para ampliar as possibilidades de interação, incluindo o reconhecimento visual de elementos do ambiente, como indicadores luminosos em painéis. A arquitetura também deverá ser estendida a atividades de manutenção e montagem

industrial, com avaliações envolvendo usuários, ruído ambiental, usabilidade e carga cognitiva.

Agradecimentos

Este trabalho foi apoiado pelo CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), processo n.º 444677/2024-0, e pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), Código de Financiamento 001.

Referências

ANDRADE, M. et al. On the use of biofuels for cleaner cities: Assessing vehicular pollution through digital twins and machine learning algorithms. *Sustainability*, MDPI, v. 16, n. 2, p. 708, 2024.

ANGELOPOULOS, J.; MANETTAS, C.; ALEXOPOULOS, K. Industrial maintenance optimization based on the integration of large language models (LLM) and augmented reality (AR). In: *Advances in Artificial Intelligence in Manufacturing II*. Cham: Springer, 2025. p. 197–205.

BELCAK, P. et al. *Small Language Models are the Future of Agentic AI*. 2025. ArXiv:2506.02153.

DITTRICH, M.; KO, S.; RIEGER, T. Impact of voice assistants' conversational style on cognitive driver distraction. In: *Proceedings of the 25th ACM International Conference on Intelligent Virtual Agents*. New York, NY, USA: ACM, 2025.

EVEN REALITIES. *Even G2 smart glasses*. 2026. Acesso em: 2 out. 2026. Disponível em: <https://www.evenrealities.com/smart-glasses>.

FRIHA, O. et al. LLM-based edge intelligence: A comprehensive survey on architectures, applications, security and trustworthiness. *IEEE Open Journal of the Communications Society*, v. 5, p. 5799–5856, 2024.

FUNK, M. et al. Usable and acceptable response delays of conversational agents in automotive user interfaces. In: *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. New York, NY, USA: ACM, 2020. p. 262–269.

GIUGLIANO, G. et al. Smart wearable devices and technologies for human-augmentation industry 5.0: A systematic review. *Applied Sciences*, MDPI, v. 16, n. 3, p. 1481, 2026.

GRAZIANI, A. R. et al. A shift toward Industry 5.0: A practical assessment framework for human-centric, sustainable, and resilient industry. *Sustainability*, v. 18, n. 12, p. 6330, 2026. ISSN 2071-1050.

KAVATHEKAR, I. et al. Small models, big tasks: An exploratory empirical study on small language models for function calling. In: *Proceedings of the 29th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2025. p. 1117–1126.

KIM, D.; CHOI, Y. Applications of smart glasses in applied sciences: A systematic review. *Applied Sciences*, v. 11, n. 11, p. 4956, 2021.

KRSTAČIĆ, R.; ŽUŽIĆ, A.; OREHOVAČKI, T. Safety aspects of in-vehicle infotainment systems: A systematic literature review from 2012 to 2023. *Electronics*, v. 13, n. 13, 2024.

LAMAAKAL, I. et al. Tiny language models for automation and control: Overview, potential applications, and future research directions. *Sensors*, MDPI, v. 25, n. 5, p. 1318, 2025.

LEVY, S. *Google Glass 2.0 Is a Startling Second Act*. WIRED, 2017. Disponível em: <https://www.wired.com/story/google-glass-2-is-here>. Acesso em: 2 out. 2026.

LU, Z. et al. Demystifying small language models for edge deployment. In: *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics*. Vienna, Austria: ACL, 2025. p. 14747–14764.

MACHADO, H. et al. Timing challenges in vehicular tinyml: Characterizing end-to-end latency. In: *2025 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0 & IoT)*. Castelldefels, Spain: IEEE, 2025. p. 29–33.

MARTÍNEZ, M. A. D. et al. A systematic review of Industry 5.0 and sustainability in manufacturing: clustering of concepts, characteristics, and technological enablers. *Discover Sustainability*, v. 7, n. 1, p. 175, 2026. ISSN 2662-9984.

MEDEIROS, M.; SILVA, M.; SILVA, I. Evolving online clustering for real-time driver behavior analysis in iot-enabled vehicles. *Evolving Systems*, Springer, v. 16, n. 4, p. 118, 2025.

META. *Óculos com IA*. 2026. Disponível em: <https://www.meta.com/br/ai-glasses/>. Acesso em: 2 out. 2026.

NUSSBAUM, J.; LEGLER, T.; RUSKOWSKI, M. CPPM Copilot: Proposing an AI-based assistant for manual assembly tasks in a flexible production. In: ALEXOPOULOS, K.; MAKRIS, S.; STAVROPOULOS, P. (Ed.). *Advances in Artificial Intelligence in Manufacturing II*. Cham: Springer, 2025. p. 80–88.

RAYNEO. *RayNeo AR Smart Glasses*. 2026. Disponível em: <https://www.rayneo.com/products>. Acesso em: 2 out. 2026.

REIMER, B. et al. From distraction to support: evolving perspectives on driver attention support and the case for a holistic, context-aware framework. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 121, p. 103658, 2026.

ROMERO, D. et al. Towards an operator 4.0 typology: A human-centric perspective on the fourth industrial revolution technologies. In: *Proceedings of the 46th International Conference on Computers and Industrial Engineering*. Tianjin, China: CIE, 2016. p. 608–618.

SCHICK, T. et al. Toolformer: Language models can teach themselves to use tools. *Advances in neural information processing systems*, v. 36, p. 68539–68551, 2023.

SOLOMASHENKO, A. B. et al. Industrial applications of AR headsets: a review of the devices and experience. *Light: Advanced Manufacturing*, v. 6, n. 2, p. 358–387, 2025.

WEISER, M. The computer for the 21st century. *Scientific American*, v. 265, n. 3, p. 94–104, 1991.

XU, X. et al. Industry 4.0 and industry 5.0–inception, conception and perception. *Journal of Manufacturing Systems*, v. 61, p. 530–535, 2021.

ZHANG, X. et al. From virtual touch to tesla command: Unlocking unauthenticated control chains from smart glasses for vehicle takeover. In: *2024 IEEE symposium on security and privacy*. San Francisco, CA, USA: IEEE, 2024. p. 2366–2384.

Arquitetura *Edge–Cloud* para Monitoramento Inteligente de Veículos Conectados com Auditabilidade via *Blockchain*

Miguel Euripedes [1], Morsinaldo Medeiros [1], Thais Medeiros [1], Hilton Machado [1], Erick Justino [1], Alice Freire [1], Marianne Silva[2], Carlos M. D. Viegas[1], Ivanovitch Silva[1]

[1] Dep. Eng. de Computação e Automação (DCA), UFRN
<miguel.amaral.111@ufrn.edu.br>,
<morsinaldo.medeiros.075@ufrn.edu.br>,
<thais.araujo.707@ufrn.edu.br>,
<hilton.machado.088@ufrn.edu.br>,
<erick.justino.701@ufrn.edu.br>,
<alice.victorino.061@ufrn.edu.br>, <carlos.viegas@ufrn.br>,
<ivanovitch.silva@ufrn.br>
[2] Curso Sistemas de Informação - Campus Arapiraca, UFAL
<marianne.silva@penedo.ufal.br>

Resumo

O monitoramento de sistemas veiculares conectados requer a integração entre aquisição de sinais, processamento inteligente e infraestrutura computacional capaz de lidar com fluxos contínuos de dados. Neste trabalho, propõe-se e avalia-se uma arquitetura modular para monitoramento veicular que integra aquisição de dados veiculares via OBD-II, processamento na borda (*edge*), métodos de inteligência artificial, *soft-sensors*, serviços em nuvem e registro distribuído em *blockchain*. Como estudo de caso, a arquitetura foi aplicada à análise do comportamento de condução e à estimativa de emissões de CO_2. Nos ensaios de infraestrutura, foram processados 627.515 registros a uma taxa média de 9,86 registros/s, com latência predominantemente entre 0 e 2 s no trecho instrumentado. Em condições reais de operação, os dados veiculares foram integrados às saídas dos modelos TEDA e MMCloud e às estimativas de CO_2. A integração com uma rede permissionada Hyperledger Besu também foi avaliada, com registro e posterior recuperação das informações consolidadas de uma viagem e tempo médio de aproximadamente 4,40 s entre o envio da requisição e a confirmação da transação. Os resultados fornecem evidências da viabilidade funcional da arquitetura nas condições avaliadas, integrando monitoramento, análise inteligente, infraestrutura de dados e registro distribuído e auditável.

Abstract

Monitoring connected vehicular systems requires integrating signal acquisition, intelligent processing, and infrastructure capable of handling continuous data streams. This work proposes and evaluates a modular architecture combining OBD-II data acquisition, edge processing, artificial intelligence methods, soft sensors, cloud services, and blockchain-based record-keeping. As a case study, the architecture was applied to driving behavior analysis and CO_2 emission estimation. Infrastructure experiments processed 627,515 records at an average rate of 9.86 records/s, with latency predominantly between 0 and 2 s in the instrumented segment. Under real-world operating conditions, vehicular data were integrated with TEDA and MMCloud outputs and CO_2 estimates. Integration with a permissioned Hyperledger Besu network was evaluated by recording and subsequently retrieving consolidated trip information, with an average time of approximately 4.40 s between request submission and transaction confirmation. The results support the architecture's functional feasibility under the evaluated conditions.

1 Introdução

O monitoramento de sistemas físicos depende da transformação de dados de sensores em informações sobre suas condições de operação (ZENDRON et al., 2025). No domínio veicular, parâmetros como velocidade, rotação do motor, carga do motor e posição do acelerador podem ser obtidos por meio de PIDs (*Parameter IDs*) disponibilizados pela interface *On-Board Diagnostics* (OBD-II) do veículo, constituindo uma fonte de dados para aplicações de monitoramento e supervisão (MICHAILIDIS; PANAGIOTOPOULOU; PAPADAKIS, 2025).

A combinação de computação em borda com serviços em nuvem e inteligência artificial amplia as possibilidades de processamento desses dados (KAGAWADE et al., 2026). Enquanto a borda permite executar análises próximas à fonte, a nuvem possibilita persistir e disponibilizar registros para diferentes aplicações (NILSSON et al., 2025). Métodos de inteligência artificial e *soft-sensors* complementam esse fluxo ao identificar padrões operacionais e inferir variáveis não diretamente medidas (LU; WANG, 2026). Entretanto, a integração dessas tecnologias exige coordenar aquisição, processamento, comunicação, armazenamento e disponibilização dos dados (ZHANG et al., 2026).

No contexto veicular, essa integração permite combinar a análise do comportamento de condução com a estimativa de grandezas como emissões de CO_2 a partir dos parâmetros operacionais (RYKAŁA; GRZELAK; BORUCKA, 2024). Além do processamento e da disponibilização dessas informações, aplicações que dependem de registros verificáveis requerem mecanismos capazes de preservar sua integridade e auditabilidade. Nesse contexto, tecnologias *blockchain* podem complementar a infraestrutura de monitoramento por meio do registro distribuído e auditável das informações produzidas pela plataforma (TYAGI; TIWARI; NAITHANI, 2024).

Diante desse cenário, este trabalho propõe e avalia uma arquitetura modular que integra aquisição de dados via interface OBD-II, processamento na borda, análise inteligente, infraestrutura em nuvem e registro distribuído em *blockchain*. Um estudo de caso é utilizado para avaliar a aquisição e disponibilidade dos dados, a execução dos métodos de análise comportamental, a estimativa de CO_2, o desempenho da infraestrutura e a integração com uma rede permissionada Hyperledger Besu (BESU, 2026).

As principais contribuições são: (i) a especificação de uma arquitetura modular que integra computação na borda, nuvem e *blockchain* para o monitoramento de dados veiculares; (ii) a integração de métodos de inteligência artificial e *soft-sensors* ao fluxo de processamento; e (iii) a implementação e avaliação dos componentes da arquitetura, incluindo aquisição de dados via OBD-II, infraestrutura de dados e registro distribuído das informações consolidadas das viagens.

O restante do artigo está organizado em cinco seções: a Seção 2 apresenta os trabalhos relacionados; a Seção 3 descreve a arquitetura proposta; a Seção 4 apresenta o estudo de caso; a Seção 5 apresenta e discute os resultados; e a Seção 6 apresenta as conclusões.

2 Trabalhos Relacionados

Estudos recentes têm explorado dados veiculares, inteligência artificial e processamento distribuído para monitoramento de veículos conectados. Michailidis, Panagiotopoulou e Papadakis (2025) revisam aplicações de aprendizado de máquina com dados OBD-II para análise de consumo, emissões, comportamento e diagnóstico. Andrade et al. (2022) apresentam um *soft-sensor* para estimativa de CO_2 a partir de sinais OBD-II, enquanto Medeiros et al. (2024) combinam *soft-sensors*, TEDA e agrupamento incremental para análise do comportamento do condutor na borda.

No nível arquitetural, Nilsson et al. (2025) apresentam a AutoSPADA para processamento distribuído entre veículos e nuvem, enquanto Ku et al. (2025) propõem o V2X-Car Edge Cloud para gerenciamento de serviços distribuídos em ambientes veiculares.

No monitoramento de emissões, Farrag et al. (2025) combinam sensores de gases, comunicação sem fio, nuvem e aprendizado profundo, enquanto Biswas, Chowdhury e Chandra (2026) empregam sensoriamento, processamento em borda e aprendizado de máquina. Diferentemente dessas abordagens, neste trabalho o CO_2 é estimado a partir de parâmetros OBD-II por meio de um *soft-sensor*, sem sensores dedicados de gases.

A tecnologia *blockchain* também tem sido aplicada à rastreabilidade de emissões. Liu et al. (2024) propõem um mecanismo de autenticação para ambientes de contabilização de emissões veiculares, utilizando *blockchain* para o registro dessas informações, enquanto Lu et al. (2022) apresentam o STRICTs, baseado em Hyperledger Fabric para auditoria e negociação de permissões de emissão.

Deste modo, este trabalho integra elementos tratados separadamente na literatura em uma arquitetura modular que articula aquisição OBD-II, processamento inteligente na borda, infraestrutura em nuvem, supervisão e registro auditável em *blockchain*.

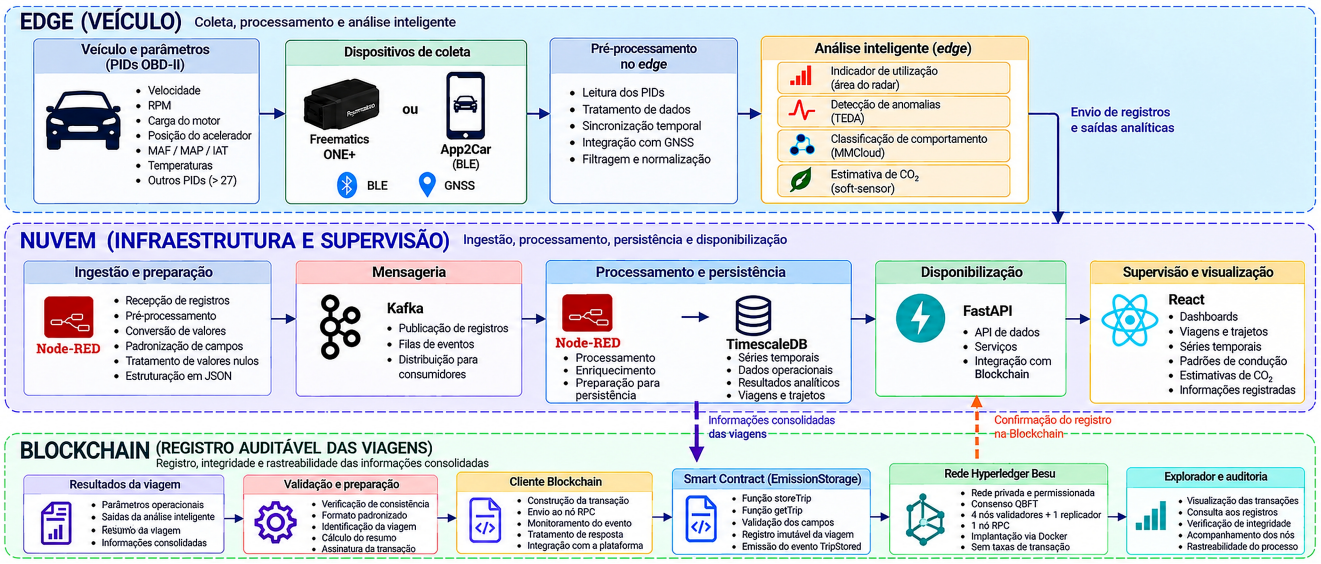


Figura 1: *Visão geral da arquitetura modular proposta.*

A contribuição está na integração e avaliação desse fluxo, incluindo o registro das informações consolidadas das viagens em uma rede permissionada Hyperledger Besu.

3 Arquitetura Proposta

A arquitetura proposta, ilustrada pela Figura 1, organiza o monitoramento veicular em um fluxo modular que integra leitura dos parâmetros veiculares, coleta de dados, processamento inteligente, infraestrutura de dados, supervisão e registro distribuído.

O fluxo inicia-se na borda com a leitura dos parâmetros veiculares, identificados por PIDs (*Parameter IDs*), por meio da interface OBD-II conectada ao veículo. Os dados são coletados por dispositivos de borda e utilizados pelos métodos de análise inteligente para geração de indicadores, detecção de anomalias, classificação do comportamento de condução e estimação de emissões de CO_2. Os registros e as saídas analíticas são então encaminhados à infraestrutura em nuvem, onde são preparados, distribuídos, processados e armazenados como séries temporais. Serviços de acesso e visualização disponibilizam essas informações para aplicações de monitoramento e supervisão. Complementarmente, informações consolidadas podem ser encaminhadas ao módulo *blockchain* para registro verificável e auditável, fornecendo informações rastreáveis que podem subsidiar mecanismos de créditos de carbono.

3.1 Coleta e análise inteligente na borda

A coleta de dados veiculares é realizada por meio da interface OBD-II, abrangendo mais de 27 PIDs disponibilizados pelo veículo. Esses parâmetros fornecem informações sobre diferentes condições de operação, incluindo velocidade, rotação do motor, carga do motor, posição do acelerador, temperaturas e variáveis relacionadas à admissão de ar, entre outras. Para as análises consideradas neste trabalho, são utilizados subconjuntos desses parâmetros de acordo com a informação a ser inferida.

Na análise comportamental, são empregados principalmente velocidade, rotação do motor, carga do motor e posição do acelerador. Para a estimativa de CO_2, utiliza-se o fluxo de massa de ar (MAF), quando disponível, ou parâmetros necessários à sua estimação, como pressão absoluta no coletor de admissão (MAP), temperatura do ar de admissão (IAT) e rotação do motor. Informações de localização obtidas por GNSS (*Global Navigation Satellite System*) complementam os registros e permitem associar os dados operacionais aos trajetos percorridos.

Para isso, a arquitetura admite duas opções de coleta. A primeira utiliza o dispositivo Freematics ONE+, baseado no ESP32 com interface OBD-II e receptor GNSS, que possibilita a aquisição e o processamento local dos sensores (ANDRADE et al., 2022). A segunda utiliza o aplicativo *App2Car*, desenvolvido pelo laboratório Conect2AI, que se comunica com o veículo por meio de um adaptador OBD-II sem fio, baseado na tecnologia *Bluetooth Low Energy* (BLE) e emprega o dispositivo móvel como *gateway*. Ambas as modalidades alimentam a mesma infraestrutura de dados, permitindo desacoplar a fonte de aquisição das etapas posteriores de processamento e armazenamento.

A camada de análise inteligente combina dois fluxos complementares. O primeiro caracteriza padrões

de condução a partir de indicadores derivados dos sensores veiculares. Para isso, os sensores selecionados são normalizados e representados em um gráfico de radar, no qual cada eixo corresponde a uma variável operacional. A área do polígono formado por esses sinais é utilizada para produzir um indicador que sintetiza conjuntamente as condições de utilização do veículo.

Conforme ilustra a Figura 2, o polígono do radar pode ser decomposto em triângulos formados pelo centro da representação e por dois vértices consecutivos.

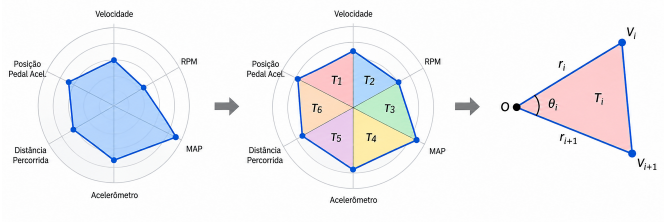


Figura 2: *Decomposição do gráfico de radar para o cálculo da área do indicador de utilização.*

A área total é obtida pela soma das áreas desses triângulos. Considerando n eixos igualmente espaçados e $r_i(t)$ o valor normalizado do sinal associado ao eixo i no instante t, a área do polígono é dada por:

$$A(t) = \frac{1}{2}\sin\left(\frac{2\pi}{n}\right)\sum_{i=1}^{n} r_i(t)r_{i+1}(t), \qquad r_{n+1}(t) = r_1(t) \tag{1}$$

A partir de uma configuração de referência de área B, define-se o indicador normalizado:

$$I(t) = \frac{A(t)}{B} \tag{2}$$

O indicador $I(t)$ sintetiza a magnitude conjunta dos sinais operacionais e alimenta os métodos de análise comportamental. O *framework* TEDA (*Typicality and Eccentricity Data Analytics*) analisa cada nova observação a partir de medidas de tipicidade e excentricidade, que representam, respectivamente, sua similaridade e seu afastamento em relação ao comportamento observado no fluxo de dados (MEDEIROS; SILVA; SILVA, 2025). Essas medidas são atualizadas recursivamente, permitindo sinalizar observações atípicas sem armazenar todo o histórico.

Para caracterizar os padrões de condução, o MMCloud utiliza o indicador de utilização dos recursos do veículo e a carga do motor como entradas de um processo de agrupamento incremental. A cada nova observação, sua distância em relação aos centroides existentes é avaliada e as estatísticas dos grupos são atualizadas, permitindo adaptar os agrupamentos à evolução do fluxo de dados (MEDEIROS; SILVA; SILVA, 2025). Na configuração utilizada, são considerados até três grupos, associados aos perfis *cauteloso*, *normal* e *agressivo* conforme a ordenação das normas de seus centroides.

O segundo fluxo analítico utiliza um *soft-sensor* para estimar as emissões de CO_2 a partir do fluxo de massa de ar admitido pelo motor. Quando o sinal MAF não está disponível, esse fluxo pode ser estimado pelo método *Speed Density*, de acordo com os parâmetros disponíveis no veículo. Seja $\dot{m}_{ar}(t)$ o fluxo de massa de ar, AFR_f a razão mássica ar–combustível, ρ_f a densidade do combustível e γ_f o fator de emissão correspondente. O fluxo estimado de CO_2 é calculado por:

$$\hat{m}_{CO_2}(t) = \frac{\gamma_f}{\rho_f AFR_f}\dot{m}_{ar}(t) \tag{3}$$

A Equação 3 fornece a taxa estimada de emissão em g/s. Para uma viagem, a massa acumulada de CO_2 é obtida pela integração temporal dessa taxa, aproximada no domínio discreto por:

$$\widehat{M}_{CO_2} \approx \sum_{k=1}^{N} \hat{m}_{CO_2}(t_{k-1})\Delta t_k, \qquad \Delta t_k = t_k - t_{k-1} \tag{4}$$

O uso de Δt_k permite considerar eventuais variações nos intervalos de amostragem durante a integração. Os parâmetros AFR_f, ρ_f e γ_f devem ser definidos de acordo com o combustível utilizado. Dessa forma, a camada analítica combina informações diretamente adquiridas do veículo com variáveis inferidas, produzindo indicadores de comportamento e estimativas de emissão que são encaminhados aos demais componentes da arquitetura.

3.2 Infraestrutura de dados e supervisão

Os dados provenientes da borda são encaminhados à infraestrutura de dados, onde fluxos implementados em Node-RED (NODE-RED, 2026) realizam operações de preparação, incluindo conversão de valores, padronização dos nomes dos campos e tratamento de valores nulos. Os registros são estruturados em JSON e publicados no Apache Kafka (APACHE, 2026), que atua como intermediário entre a produção e o consumo das mensagens. Esse desacoplamento permite que aquisição e persistência operem em ritmos distintos, de acordo com a capacidade e a configuração dos serviços.

No estágio seguinte, outro fluxo Node-RED consome as mensagens e prepara as operações de persistência no TimescaleDB. O armazenamento orientado a séries temporais permite preservar a evolução

dos sinais ao longo das viagens e relacioná-los aos resultados produzidos pela camada analítica. Essa organização também possibilita avaliar propriedades da infraestrutura, como taxa de processamento e latência no trecho instrumentado.

A disponibilização dos dados é realizada por uma API desenvolvida com FastAPI, utilizada tanto por aplicações externas quanto pela interface de supervisão. Uma aplicação desenvolvida em React permite consultar viagens, trajetos, séries temporais, padrões de condução e estimativas instantâneas e acumuladas de CO_2. A separação entre persistência, acesso programático e apresentação permite que os mesmos registros sejam utilizados por diferentes aplicações sem acoplá-las diretamente aos mecanismos de aquisição.

3.3 Registro em *blockchain*

Após o processamento dos dados de uma viagem, as informações consolidadas são encaminhadas ao módulo *blockchain*. Um cliente intermediário é responsável por construir, assinar e submeter a transação ao nó RPC (*Remote Procedure Call*) de uma rede permissionada Hyperledger Besu. A rede utiliza o protocolo de consenso QBFT (*Quorum Byzantine Fault Tolerance*) e opera sem taxas de transação (*free gas*).

O registro é realizado pelo contrato inteligente *EmissionStorage*, implementado em linguagem *Solidity*. A função `storeTrip` recebe as informações consolidadas da viagem, verifica a consistência dos campos e a existência prévia do registro e, após a persistência, emite o evento `TripStored`. Os dados registrados podem posteriormente ser recuperados pela função `getTrip`, permitindo verificar a correspondência entre as informações processadas pela plataforma e aquelas persistidas na rede.

A infraestrutura utilizada é composta por quatro nós validadores, um nó replicador e um nó RPC. Dessa forma, o módulo acrescenta ao fluxo da arquitetura um mecanismo de registro distribuído e auditável das informações produzidas a partir dos dados veiculares.

4 Estudo de Caso

A arquitetura proposta foi avaliada por meio de experimentos complementares, realizados em condições reais de operação e em ambiente controlado. A avaliação possui caráter exploratório e descritivo, abrangendo quatro aspectos: aquisição e disponibilidade dos dados, execução das análises inteligentes, desempenho da infraestrutura e integração com o módulo *blockchain*.

4.1 Cenários

A aquisição em cenário real foi realizada por meio do aplicativo *App2Car*, utilizando duas configurações compostas por adaptadores OBD-II e dispositivos móveis. Para a análise apresentada neste trabalho, foi selecionada uma das viagens registradas, contendo parâmetros operacionais do veículo e as saídas dos métodos TEDA e MMCloud. Esse conjunto foi utilizado para avaliar a regularidade temporal da aquisição, a disponibilidade dos parâmetros e a execução integrada dos métodos de análise comportamental.

A avaliação do *soft-sensor* de CO_2 considerou séries obtidas em viagens realizadas com os veículos VW Polo e Mitsubishi Eclipse. Essas séries foram utilizadas para analisar as estimativas instantâneas de emissão e sua acumulação ao longo do tempo.

Complementarmente, o desempenho da infraestrutura foi caracterizado por meio de um ensaio controlado com dados veiculares simulados. O ensaio teve duração aproximada de 17 h e 40 min e produziu 627.515 registros, permitindo avaliar a taxa de processamento, a latência e a utilização de recursos computacionais.

A integração com o módulo *blockchain* foi avaliada em um percurso experimental, no qual os dados adquiridos foram processados pela plataforma e um registro consolidado da viagem foi submetido a uma rede permissionada Hyperledger Besu. Esse experimento foi utilizado para verificar a integração entre o processamento dos dados veiculares e o registro distribuído associado ao mecanismo de créditos de carbono.

4.2 Procedimentos e métricas de avaliação

Na viagem selecionada do *App2Car*, os registros foram organizados temporalmente e examinados quanto à presença dos parâmetros operacionais e das saídas analíticas. A regularidade da aquisição foi caracterizada pelos intervalos entre observações consecutivas, enquanto a disponibilidade foi determinada pela presença das variáveis nos registros analisados.

Para a análise inteligente, foram consideradas as observações sinalizadas pelo TEDA e a distribuição das classificações produzidas pelo MMCloud, relacionando essas saídas temporalmente aos parâmetros operacionais disponíveis. Para o *soft-sensor*, foram analisadas as estimativas instantâneas de CO_2 e a massa acumulada obtida por integração temporal, conforme descrito na Equação 4.

No ensaio de infraestrutura, foram considerados o número de registros processados, a taxa média de

processamento, a utilização de CPU e memória e a latência. A taxa média foi calculada pela razão entre o número de registros e a duração do ensaio. A latência corresponde à diferença entre o instante de envio do registro e sua chegada ao estágio instrumentado de pré-processamento no Node-RED, não representando, portanto, a latência de ponta a ponta da arquitetura.

Por fim, a integração com a *blockchain* foi avaliada a partir das informações consolidadas da viagem, verificando-se o fluxo entre processamento, submissão da transação, execução do contrato inteligente e confirmação do registro na rede permissionada Hyperledger Besu. Foram considerados a emissão do evento `TripStored`, a recuperação posterior dos dados por meio da função `getTrip` e o tempo entre o envio da requisição pela API e a confirmação da transação. Dessa forma, a avaliação caracteriza a integração funcional e a rastreabilidade das informações utilizadas pelo mecanismo de créditos de carbono, não constituindo um ensaio de capacidade máxima da rede.

5 Resultados

Os resultados são apresentados de acordo com os componentes avaliados da arquitetura, o ensaio controlado de infraestrutura, as séries de emissão e o experimento de integração com *blockchain*.

5.1 Aquisição, plataforma e desempenho da infraestrutura

A viagem selecionada resultou em 684 observações analisadas ao longo de aproximadamente 683 s. O intervalo mediano entre observações consecutivas foi de 1,000 s, com valores entre 0,994 e 1,005 s, caracterizando uma cadência de registro próxima de 1 Hz. Velocidade, rotação do motor, carga do motor, área do radar e as saídas TEDA e MMCloud estavam disponíveis nos registros analisados, permitindo relacionar temporalmente os parâmetros operacionais às saídas da camada de análise inteligente.

Os dados processados são disponibilizados pela plataforma de supervisão, que integra informações das viagens, parâmetros veiculares e resultados produzidos pelos módulos analíticos. A Figura 3 apresenta a interface implementada para acompanhamento das informações processadas pela arquitetura.

No ensaio de infraestrutura, foram processados 627.515 registros durante 63.653 s, correspondendo a uma taxa média de 9,86 registros/s. A latência no trecho instrumentado permaneceu predominantemente entre 0 e 2 s, com picos de até 4 s. A utilização

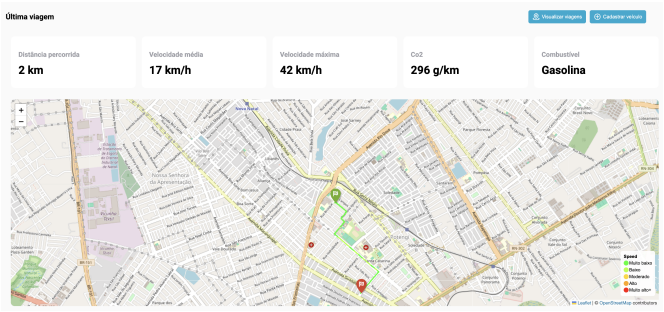


Figura 3: *Interface da plataforma para supervisão dos dados veiculares e dos resultados produzidos pela arquitetura.*

de CPU atingiu 11,3% durante o ensaio, enquanto o uso de memória permaneceu em 1,9%. A Tabela 1 destaca os principais resultados quantitativos obtidos nas diferentes etapas da avaliação.

Tabela 1: *Resultados da avaliação da arquitetura.*

Indicador	Resultado obtido
Observações analisadas	684
Intervalo mediano	1,000 s
Faixa dos intervalos	0,994 a 1,005 s
Registros processados	627.515
Taxa média	9,86 registros/s
Latência predominante	0 a 2 s (pico 4 s)
Uso de CPU e Memória RAM	11,3% e 1,9%
TEDA	57 (8,33%)
MMCloud – normal	514 (75,15%)
MMCloud – cauteloso	170 (24,85%)

Os resultados demonstram a disponibilidade conjunta dos parâmetros veiculares e das saídas analíticas, bem como o funcionamento da infraestrutura sob a carga aplicada. O estudo de desempenho, entretanto, caracteriza as condições avaliadas e não determina a capacidade máxima da arquitetura.

5.2 Análise inteligente e estimativa de CO_2

Na viagem analisada, o TEDA sinalizou 57 observações, correspondendo a 8,33% dos registros. Essas ocorrências representam observações identificadas como atípicas em relação ao comportamento do fluxo de dados, permitindo relacioná-las temporalmente às condições operacionais do veículo.

O MMCloud produziu dois códigos de comportamento no conjunto analisado. O comportamento *normal* foi observado em 514 registros (75,15%) e o comportamento *cauteloso* em 170 registros (24,85%). Esses resultados demonstram a execução integrada dos métodos de análise comportamental sobre os dados adquiridos pelo *App2Car*. As proporções observadas não representam, entretanto, uma medida de

acurácia, uma vez que não foi utilizada uma referência independente para as condições de condução.

Além da análise comportamental, o *soft-sensor* foi utilizado para estimar as emissões de CO_2. A Figura 4 apresenta a evolução temporal das emissões acumuladas estimadas para os veículos VW Polo e Mitsubishi Eclipse.

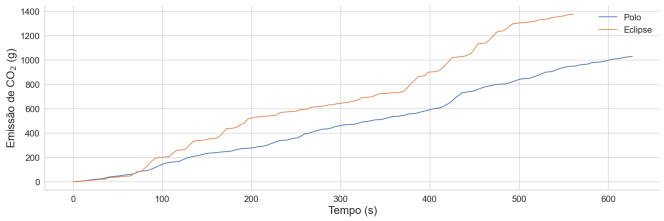


Figura 4: *Evolução temporal das estimativas acumuladas de CO_2 (em gramas) para os veículos VW Polo e Mitsubishi Eclipse.*

As estimativas instantâneas apresentam variações ao longo das viagens, enquanto sua integração temporal resulta no crescimento da massa acumulada de CO_2. Ao final dos respectivos períodos analisados, a emissão acumulada estimada para o veículo Eclipse foi superior à do veículo Polo. Essa diferença, entretanto, não permite comparar diretamente a eficiência dos veículos, pois fatores como duração, distância percorrida e condições de operação também influenciam a emissão acumulada.

Na ausência de uma referência independente de emissão, esses resultados devem ser interpretados como uma demonstração do funcionamento e da integração do *soft-sensor* à arquitetura, e não como uma validação de sua exatidão metrológica.

5.3 Integração com *blockchain*

Na avaliação, o fluxo entre aquisição, processamento e registro distribuído foi executado até a confirmação da transação na rede Hyperledger Besu. Ao término da viagem, as informações consolidadas foram encaminhadas ao cliente *blockchain*, que invocou o contrato inteligente e submeteu a transação à rede. A Figura 5 apresenta a transação correspondente visualizada no explorador da rede. O registro foi confirmado no bloco 1514 e posteriormente recuperado por meio da função `getTrip`.

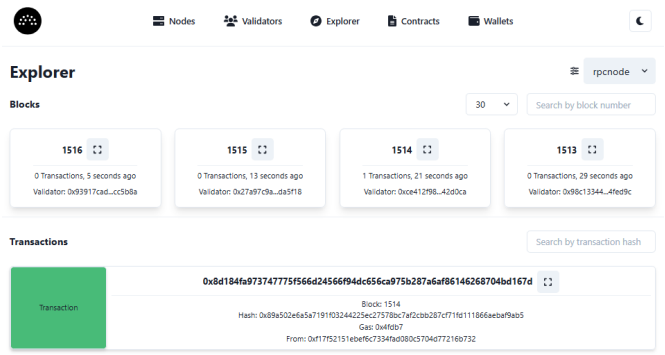


Figura 5: *Transação da viagem registrada no bloco 1514 da rede permissionada Hyperledger Besu.*

O registro recuperado continha as informações consolidadas da viagem, incluindo a estimativa de emissão de 154 g de CO_2. O tempo médio entre o envio da requisição pela API e a confirmação na *blockchain* foi de aproximadamente 4,40 s. O evento `TripStored` foi detectado pelo cliente e os dados recuperados foram validados sem inconsistências.

A confirmação da transação e a posterior recuperação dos dados demonstram a execução do ciclo de escrita, confirmação e consulta das informações da viagem na rede permissionada. Esse mecanismo acrescenta rastreabilidade ao fluxo de dados da arquitetura e fornece uma base auditável para as informações utilizadas pelo mecanismo de créditos de carbono. O experimento caracteriza a integração funcional do módulo nas condições avaliadas, não permitindo inferir a capacidade máxima ou a escalabilidade da rede sob cargas crescentes.

6 Conclusão

Este trabalho apresentou e avaliou uma arquitetura modular para monitoramento e análise inteligente de dados veiculares, integrando coleta via OBD-II, processamento na borda, análise comportamental, estimação de CO_2, infraestrutura em nuvem e registro em *blockchain*. Os resultados demonstraram, nas condições avaliadas, o processamento de 627.515 registros a uma taxa média de 9,86 registros/s, a integração das saídas TEDA e MMCloud aos dados veiculares e a geração de estimativas instantâneas e acumuladas de CO_2. O percurso experimental também demonstrou a integração entre aquisição, processamento e registro consolidado da viagem em uma rede permissionada Hyperledger Besu.

Os resultados fornecem evidências da viabilidade funcional da arquitetura, embora não determinem a capacidade máxima da infraestrutura, a acurácia das análises ou o desempenho da *blockchain* sob cargas

crescentes. Como trabalhos futuros, pretende-se realizar avaliações integradas de ponta a ponta, ampliar os cenários experimentais e avaliar o mecanismo de geração de créditos de carbono.

Agradecimentos

Os autores agradecem à FUNDEP pelo apoio ao desenvolvimento deste trabalho, bem como aos parceiros Volkswagen Truck & Bus, Volkswagen do Brasil, Stellantis e Embeddo pela colaboração no âmbito do projeto.

Referências

ANDRADE, P. et al. A tinyml soft-sensor approach for low-cost detection and monitoring of vehicular emissions. *Sensors*, MDPI, v. 22, n. 10, p. 3838, 2022.

APACHE. *Apache Kafka Documentation*. 2026. Acesso em: 6 out. 2026. Disponível em: <https://kafka.apache.org/documentation>.

BESU. *Besu Documentation*. 2026. Acesso em: 5 out. 2026. Disponível em: <https://docs.besu-eth.org>.

BISWAS, S.; CHOWDHURY, S.; CHANDRA, S. A cloud-edge based multivariate, real time in-vehicle pollution monitoring system to support a greener environment. *Remote Sensing Applications: Society and Environment*, Elsevier, p. 102035, 2026.

FARRAG, O. et al. Ivemps: Iot-based vehicle emission monitoring and prediction system. *IEEE Access*, IEEE, v. 13, p. 95628–95646, 2025.

KAGAWADE, S. S. et al. Harnessing edge ai for real-time telematics: A framework for low-latency vehicle data analytics. In: *Artificial Intelligence for Transportation*. Boca Raton: CRC Press, 2026. p. 169–188.

KU, D. et al. Vehicle-to-everything-car edge cloud management with development, security, and operations automation framework. *Electronics*, MDPI, v. 14, n. 3, p. 478, 2025.

LIU, X. et al. Lcaa: Lightweight conditional anonymous authentication scheme for vehicular carbon emissions environment. *IEEE Internet of Things Journal*, IEEE, v. 11, n. 19, p. 32025–32040, 2024.

LU, Y. et al. Stricts: A blockchain-enabled smart emission cap restrictive and carbon permit trading system. *Applied Energy*, Elsevier, v. 313, p. 118787, 2022.

LU, Z.; WANG, J. Review on soft sensing of complex industrial processes: Advances, opportunities, and challenges. *Measurement*, Elsevier, p. 123088, 2026.

MEDEIROS, M. et al. A multi-layered methodology for driver behavior analysis using tinyml and edge computing. In: *2024 IEEE International Conference on Evolving and Adaptive Intelligent Systems*. Madrid, Spain: IEEE, 2024. p. 1–8.

MEDEIROS, M.; SILVA, M.; SILVA, I. Evolving online clustering for real-time driver behavior analysis in iot-enabled vehicles. *Evolving Systems*, Springer, v. 16, n. 4, p. 118, 2025.

MICHAILIDIS, E. T.; PANAGIOTOPOULOU, A.; PAPADAKIS, A. A review of obd-ii-based machine learning applications for sustainable, efficient, secure, and safe vehicle driving. *Sensors*, MDPI, v. 25, n. 13, p. 4057, 2025.

NILSSON, A. et al. The autospada platform: User-friendly edge computing for distributed learning and data analytics in connected vehicles. *Internet of Things*, Elsevier, v. 30, p. 101480, 2025.

NODE-RED. *Node-RED Documentation*. 2026. Acesso em: 6 out. 2026. Disponível em: <https://nodered.org/docs/>.

RYKAŁA, M.; GRZELAK, M.; BORUCKA, A. Analysis of the impact of selected dynamic parameters of a motor vehicle on co2 emissions using logistic regression. *Applied Sciences*, MDPI, v. 14, n. 22, p. 10349, 2024.

TYAGI, A. K.; TIWARI, S.; NAITHANI, K. Blockchain-enabled internet of things (iots) platforms for vehicle sensing and transportation monitoring. *Digital Twin and Blockchain for Smart Cities*, Wiley Online Library, p. 257–276, 2024.

ZENDRON, L. A. S. et al. Enhancing human activity recognition with machine learning: insights from smartphone accelerometer and magnetometer data. *PeerJ Computer Science*, PeerJ Inc., v. 11, p. e3137, 2025.

ZHANG, C. et al. Integrating cfd-based digital twins into an iot platform for enhanced greenhouse temperature spatial distribution monitoring. *Internet of Things*, Elsevier, p. 102064, 2026.

Desenvolvimento de uma ferramenta computacional em aplicação *WEB* para simular o comportamento dinâmico do Bombeio por Cavidades Progressivas (BCP) e estudar o fenômeno de reversão de hastes (*Backspin*)

Guilherme da Costa Medeiros Nascimento [1]
Prof. Dr. André Laurindo Maitelli [2]
Camila Raquel Sena de Almeida [3]
Pedro Manuel Pereira Chamorro [4]
Prof. Dr. Rutácio de Oliveira Costa [5]

[1] *Departamento de Engenharia Elétrica, UFRN, <guilherme.costa.071@ufrn.edu.br>*
[2] *Departamento de Engenharia de Computação e Automação, UFRN, <maitelli@dca.ufrn.br>*
[3] *Departamento de Engenharia Elétrica, UFRN, <camila.sena.016@ufrn.edu.br>*
[4] *Departamento de Engenharia de Computação e Automação, UFRN, <pedro.chamorro.117@ufrn.edu.br>*
[5] *Departamento de Engenharia de Petróleo, UFRN, <rutacio.costa@ufrn.br>*

Resumo

O método de elevação artificial por Bombeio por Cavidades Progressivas (BCP) é amplamente empregado na indústria petrolífera devido à sua elevada eficiência no processamento de óleos pesados e viscosos. Contudo, o desligamento do motor provoca o fenómeno do *backspin* (reversão da coluna de hastes), o qual pode ocasionar danos mecânicos severos e riscos operacionais. Este trabalho apresenta a modelagem matemática via solução numérica implícita da Equação Diferencial Parcial de torção na coluna de hastes discretizada por diferenças finitas, a solução numérica e a arquitetura de um software simulador dedicado para simular o comportamento dinâmico. Os resultados das simulações numéricas demonstram coerência com o modelo e a eficácia dos mecanismos de frenagem na mitigação do*backspin*.

Abstract

The Progressing Cavity Pump (PCP) artificial lift method is widely utilized in the petroleum industry due to its high efficiency in processing heavy and viscous crude oils. However, motor shutdown triggers the backspin phenomenon (rod string reversal), which can cause severe mechanical damage and operational hazards. This paper presents the transient mathematical modeling of the rod string via a torsional Partial Differential Equation discretized by implicit finite differences, the numerical solution, and the architecture of a dedicated simulator software designed to simulate the dynamic behavior. The numerical simulation results demonstrate consistency with the model and the effectiveness of braking mechanisms in mitigating backspin.

1 Introdução, Justificativa e Motivação

O método de elevação artificial por Bombeio por Cavidades Progressivas (BCP) — conhecido internacionalmente como *Progressing Cavity Pump* (PCP) — estabeleceu-se como uma das tecnologias mais eficientes e versáteis da indústria petrolífera, sobretudo no processamento de óleos pesados, altamente viscosos ou com elevada presença de sólidos e areia (ASSMANN, 2008), (CHOLET, 1997). O sistema baseia-se na interação entre um rotor metálico helicoidal e um estator revestido por elastômero. A rotação transmitida da superfície até ao fundo do poço através de uma coluna de hastes provoca o deslocamento axial contínuo de cavidades vedadas, elevando os fluidos até à superfície.

O *backspin* ocorre quando a alimentação elétrica do motor de superfície é interrompida de forma repentina ou programada. A energia potencial de torção acumulada ao longo da coluna de hastes e a energia potencial hidráulica da coluna de fluido no interior do tubo de produção são subitamente libertadas. Este movimento faz com que parte do conjunto de superfície e a coluna de hastes girem violentamente no sentido reverso, atingindo por vezes velocidades superiores a 1500 rpm, enquanto em uma operação normal não passa de 500 RPM.

A reversão descontrolada pode provocar a desconexão ou ruptura das hastes por fadiga/torção, danos graves nos equipamentos de cabeça de poço, superaquecimento e riscos severos à integridade física dos operadores. Além disso, a frenagem prolongada e o tempo de espera até à estabilização do poço geram perdas expressivas de produção. Estudos recentes demonstram que a implementação de sistemas de controle e frenagem pode mitigar significativamente esses riscos (ASSMANN, 2008; WEG Equipamentos Elétricos, 2026).

Para mitigar estes riscos e otimizar a operação, torna-se indispensável o desenvolvimento de modelos matemáticos rigorosos e de uma plataforma de software especializada. Em relação aos trabalhos anteriores, a principal contribuição científica desta pesquisa é a integração dos modelos em uma interface WEB interativa contendo exportação de simulações, algoritmos de redução, e um módulo de geração automática de relatórios em PDF, eliminando a transcrição manual de dados. O objetivo principal é simular o comportamento transiente do sistema BCP, prever os torques e rotações críticas durante o *backspin*, e dimensionar sistemas de frenagem capazes de dissipar a energia acumulada em intervalos de tempo seguros. Neste contexto, o projeto em questão também propõe o desenvolvimento de uma bancada experimental para validar os modelos e testar estratégias de frenagem em um ambiente controlado (WEG Equipamentos Elétricos, 2026).

2 Estado da Arte

2.1 Modelagem Computacional do Sistema BCP

A aplicação da simulação computacional ao sistema BCP iniciou-se com o intuito de prever o comportamento dinâmico interligado dos seus subsistemas. Vidal (VIDAL, 2005) desenvolveu um simulador computacional baseado em Programação Orientada a Objetos (POO) que integrava os subsistemas de motor de indução, coluna de hastes, bomba BCP, espaço anular e escoamento multifásico. O trabalho demonstrou a importância da modulação transiente para avaliar esforços axiais, torques na haste polida e oscilações de pressão de sucção e recalque. Posteriormente, Assmann (ASSMANN, 2008) expandiu o modelo para incluir um estudo aprofundado de estratégias de controle e otimização, além de um sistema especialista para dimensionamento.

2.2 Análise da Reversão e Sistemas de Travagem

Araújo (ARAÚJO, 2022) aprofundou a análise dinâmica do *backspin*, categorizando o fenómeno em dois cenários operacionais distintos: Rotor Preso e Parada Normal (Rotor Livre).

- **Rotor Preso:** Ocorre quando o rotor fica bloqueado no estator (ex.: por areia ou bloqueio mecânico), acumulando um torque de torção extremo antes do desligamento do motor. Toda a energia libertada é de origem elástica/torcional.

- **Parada Normal (Rotor Livre):** Ocorre no desligamento convencional da bomba. A energia libertada resulta da combinação entre a energia torcional da coluna de hastes e a energia potencial da coluna de fluido drenada de volta para o anular através da bomba.

Araújo (ARAÚJO, 2022) demonstrou que a implementação de um sistema de travagem proporcional à velocidade angular reversa reduz drasticamente os picos de rotação e elimina oscilações perigosas na coluna de hastes.

2.3 Soluções Industriais de Controlo de *Backspin*

Ernst et al. (ERNST et al., 2006) documentaram a aplicação de variadores de velocidade (VSD) com resistores de frenagem dinâmica e controladores dedicados para poços no campo de Boscán (Venezuela). O estudo revelou que a utilização de um *Backspin Controller* integrado ao VSD reduz o tempo de reversão/paragem de até 6 horas para cerca de 1 a 1,5 minutos, preservando a produção diária e prevenindo danos mecânicos.

Cholet (CHOLET, 1997) e as normas internacionais como a ISO 15136-2 (International Organization for Standardization, 2009) estabelecem requisitos fundamentais para o dimensionamento dos mecanismos de dissipação de energia de superfície (freios centrífugos, hidráulicos, discos de fricção ou motores PMM).

3 Modelagem Matemática do Sistema BCP e do *Backspin*

3.1 Dedução da Equação Diferencial Parcial (EDP) da Coluna de Hastes

A coluna de hastes é modelada como uma barra cilíndrica submetida à torção elástica (ASSMANN, 2008). Considerando a coordenada axial z a partir da superfície ($z = 0$) até à profundidade da bomba ($z = L$) e o ângulo de rotação $\theta(z, t)$, o torque de torção transmitido na secção transversal é expresso por:

$$T(z,t) = -GJ_p \frac{\partial \theta}{\partial z} \tag{1}$$

onde G é o módulo de cisalhamento do material do aço (≈ 77 GPa) e J_p é o momento polar de inércia da secção transversal da haste ($J_p = \frac{\pi D_h^4}{32}$).

O atrito viscoso por unidade de comprimento entre o fluido e a coluna de hastes gera um torque resistivo de amortecimento:

$$T_{\text{res}} = -B \frac{\partial \theta}{\partial t} \Delta z \tag{2}$$

onde B é o coeficiente de amortecimento viscoso (ASSMANN, 2008):

$$B = 4\pi\mu \frac{D_h^2 D_t^2}{D_t^2 - D_h^2} \tag{3}$$

Aplicando a segunda lei de Newton a um elemento infinitesimal de massa ($\rho_h J_p \Delta z$), obtém-se a Equação Diferencial Parcial do movimento de torção:

$$\frac{\partial^2 \theta}{\partial t^2} = v^2 \frac{\partial^2 \theta}{\partial z^2} - c \frac{\partial \theta}{\partial t} \tag{4}$$

onde $v = \sqrt{\frac{G}{\rho_h}}$ representa a velocidade de propagação da onda de torção na haste (≈ 3130 m/s) e $c = \frac{B}{\rho_h J_p}$ é o coeficiente de atenuação viscosa.

3.2 Discretização Numérica Implícita e Solução Matricial

Para garantir estabilidade numérica sem restrições severas foi utilizada a equação numérica implícita no passo de tempo (Δt), a EDP é discretizada no espaço ($\Delta z = L/M$) e no tempo (Δt) utilizando o método das diferenças finitas com um esquema implícito.

É importante ressaltar que a precisão dos resultados e a convergência da simulação dependem intimamente da escolha dos passos de discretização espacial (Δz) e temporal (Δt). Embora o esquema implícito forneça estabilidade numérica incondicional, passos muito longos podem introduzir erros de truncamento significativos, amortecendo artificialmente a dinâmica do sistema transiente.

A equação discretizada para o nó interno i no tempo $k+1$ assume a forma:

$$\begin{aligned} &-\frac{v^2}{2(\Delta z)^2}\theta_{i-1}^{k+1} + \left(\tfrac{1}{(\Delta t)^2} + \tfrac{v^2}{(\Delta z)^2} + \tfrac{c}{2\Delta t}\right)\theta_i^{k+1} \\ &-\frac{v^2}{2(\Delta z)^2}\theta_{i+1}^{k+1} = d_i^k \end{aligned} \tag{5}$$

Esta formulação resulta num Sistema Linear Tridiagonal da forma $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, onde a matriz de coeficientes possui estrutura tridiagonal simétrica.

3.3 Condições de Contorno

3.3.1 Operação Normal (Partida e Regime Permanente)

Na superfície ($z = 0$), impõe-se a velocidade angular nominal da bomba N (rpm):

$$\theta_0^{k+1} = \omega \cdot (k+1)\Delta t, \quad \text{onde } \omega = \frac{2\pi N}{60} \tag{6}$$

No fundo ($z = L$), o torque no rotor (T_{rot}) é composto pelo torque hidráulico e pelo torque de fricção estator-rotor ($T_{\text{rot}} = T_{\text{rf}} + c_B \Delta p$). O deslocamento angular da coluna de hastes e é obtida por:

$$\theta_M^{k+1} = \frac{1}{3}\left(-\frac{2\Delta z}{GJ_p} T_{\text{rot}}^{k+1} + 4\theta_{M-1}^{k+1} - \theta_{M-2}^{k+1}\right) \tag{7}$$

3.3.2 Reversão (*Backspin*) Sem Freio

Após o desligamento do motor, a aceleração das massas rotativas da superfície (J_{sup}) passa a ser governada pela reação do torque elástico da haste polida:

$$\theta_0^{k+1} = 2\theta_0^k - \theta_0^{k-1} + \frac{GJ_p(\Delta t)^2}{2J_{\text{sup}}\Delta z}\left(-3\theta_0^k + 4\theta_1^k - \theta_2^k\right) \tag{8}$$

3.3.3 Reversão (*Backspin*) Com Sistema de Travagem

Com a atuação de um sistema de travagem que aplica um torque de retenção T_{hpf} oposto ao movimento reversor:

$$\theta_0^{k+1} = 2\theta_0^k - \theta_0^{k-1} + \frac{GJ_p(\Delta t)^2}{2J_{\text{sup}}\Delta z}\left(-3\theta_0^k + 4\theta_1^k - \theta_2^k\right) \pm T_{\text{hpf}}\frac{(\Delta t)^2}{J_{\text{sup}}} \tag{9}$$

4 Desenvolvimento do Software e Arquitetura do Simulador

4.1 Requisitos e Funcionalidades do Sistema

O simulador contempla parâmetros de entrada configuráveis (Coluna de Produção de 2 7/8″ e 3 1/2″, hastes sólidas e ocas, modelos comerciais de bombas PCP NETZSCH e Weatherford, propriedades dos fluidos e curvas IPR do reservatório), suporte à execução de múltiplos cenários transientes e rotinas otimizadas de pós-processamento (WEG Equipamentos Elétricos, 2026).

4.2 Novas Funcionalidades Implementadas

4.2.1 Redutor da Curva Tempo x Torque via Visvalingam-Whyatt

Para tratar volumes extensos de dados temporais sem perda de precisão, foi implementado o Algoritmo de Visvalingam-Whyatt. O algoritmo elimina progressivamente vértices que formam triângulos com áreas inferiores a um determinado limiar. Uma deteção híbrida de regimes preserva na totalidade os pontos da fase transitória enquanto reduz drasticamente os pontos do regime permanente (WEG Equipamentos Elétricos, 2026).

4.2.2 Módulo de Geração de Relatório em PDF

O software inclui um gerador interno de relatórios executivos em formato PDF, compilando parâmetros do poço, gráficos de alto rendimento (Torque, RPM, Níveis Dinâmicos, Potência Dissipada) e um resumo com indicadores chave do *backspin* (Rotação Máxima Reversa, Tempo Total de Parada e Energia Dissipada) (WEG Equipamentos Elétricos, 2026).

4.2.3 Gestão de Histórico e Guia do Utilizador

O sistema possui armazenamento de simulações com metadados e botões para exportação (CSV) e geração de relatórios comparativos, além de um manual técnico e tutorial integrados na interface (WEG Equipamentos Elétricos, 2026).

5 Análise de Resultados das Simulações Numéricas

5.1 Simulação da Partida em Operação Normal

Na partida, o torque na haste polida cresce até ao preenchimento completo da coluna. O nível dinâmico no anular desce até atingir o ponto de equilíbrio onde $Q_{\text{bomba}} \approx Q_{\text{res}}$ (ASSMANN, 2008).

5.2 Simulação da Reversão (*Backspin*)

A Tabela 1 apresenta o comparativo entre os dois cenários de reversão analisados para um poço de 1500 m.

Tabela 1: *Comparativo do Backspin: Livre vs. Com Frenagem*

Parâmetro	*Backspin* Livre	*Backspin* Com Freio
Rotação Reversa Máxima	$>$ 1500 rpm	$<$ 250 rpm
Perfil de Velocidade	Altamente Oscilatório	Desaceleração Suave
Tempo Total de Parada	Elevado (Minutos/Horas)	Reduzido
Risco de Danos Mecânicos	Muito Elevado	Praticamente Nulo

5.3 Avaliação da influência do modelo de bomba, diâmetro de coluna de produção e torque de frenagem

Foi realizada uma bateria de 16 testes adicionais contemplando dois modelos de bombas, dois diâmetros de coluna de produção e distintos valores de torque de frenagem. A bomba NTZ 400-200-ST-50 fornece 50 m^3/d para cada 100 rpm, enquanto a bomba NTZ 350-200-ST-20 está limitada a 20 $m3/d$ para cada 100 rpm. Observam-se tempos de parada significativamente maiores para a bomba de menor vazão, o

que, de fato, seria esperado, porque o fenômeno se prolonga enquanto há um desnível de fluido suficiente para vencer o torque de atrito entre o rotor e o estator.

O torque requerido para frear instantaneamente a bomba NTZ 350-200-ST-20 é de cerca de 200 Nm, enquanto que são requeridos 800 Nm para parar instantaneamente a bomba NTZ 400-200-ST-50. Portanto, a bomba de maior vazão é a que requer mais torque de frenagem para fazê-la parar instantaneamente. Na ausência de frenagem, verifica-se que a bomba de maior vazão para em menos tempo. Para uma dada bomba, quanto maior o diâmetro da coluna de produção, menor o tempo de parada. Percebe-se que a energia dissipada na frenagem é maior para a bomba de maior vazão, e o efeito do aumento do diâmetro da coluna de produção é aumentar a energia dissipada.

Tabela 2: *Resultados da Bateria de Simulações de Backspin (Poço de 1500 m)*

#	Tubagem	Modelo da Bomba	Freio (N·m)	Tempo Parada (s)	Energia Dissipada (kJ)
1	2 7/8″	NTZ 400-200-ST-50	0	468,3	0,0
2	2 7/8″	NTZ 400-200-ST-50	400	391,4	-8088,3
3	2 7/8″	NTZ 400-200-ST-50	600	267,5	-3374,0
4	2 7/8″	NTZ 400-200-ST-50	800	9,1	-8,1
5	3 1/2″	NTZ 400-200-ST-50	400	523,0	-11422,4
6	3 1/2″	NTZ 350-200-ST-20	100	2362,2	-3620,6
7	3 1/2″	NTZ 350-200-ST-20	200	13,4	-0,8

5.4 Bateria de Testes para analisar influência da Viscosidade na dinâmica do BCP

Com objetivo de avaliar a dinâmica do Bombeio por Cavidades Progressivas, uma bateria de testes foi realizada dentro da versão mais atualizada do Simulador BCP-WEG. A principal característica observada foi a variação dos valores de viscosidade do fluído, e foi avaliada sua influência em relação aos resultados de Nível, RPM e Torque.

5.4.1 Ensaio 1: Torque de Frenagem Nulo ($T_{freio} = 0$ N · m)

O primeiro teste considera nulo o torque de frenagem. Para o fluido com viscosidade de 100 cp, os níveis no anular e na coluna de produção apresentam uma tendência única de aumento (coluna) ou redução (anular), sendo a taxa de variação com o tempo cada vez menor, como visto na Figura 1. Já para a viscosidade de 1 cp observa-se que os níveis apresentam altas taxas de variação, especialmente no início do fenômeno, mas invertem a tendência de aumento ou diminuição, o que sugere um comportamento oscilatório amortecido. Após algumas inversões de tendência, o nível na coluna tende a uma estabilidade e o nível no anular tem uma pequena tendência de crescimento, devido a alimentação de fluido do reservatório. Em ambos os casos, devido ao atrito rotor/estator, não ocorre um equilíbrio de níveis à semelhança do que ocorreria em um tubo em U.

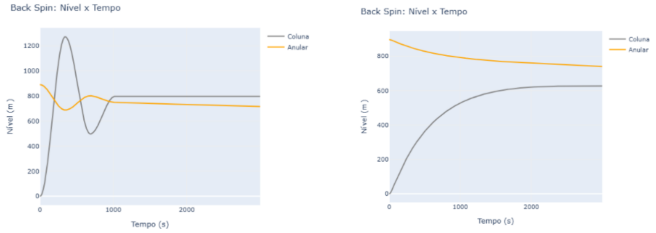


Figura 1: *Nível vs Tempo com 1cp (esq.) e 100cp (dir.), respectivamente*

Em relação ao RPM, Figura 2, gerada com viscosidade de 100 cp, a coluna de hastes rapidamente inverte o sentido de rotação, permanecendo em rotação reversa até parar pela ação do atrito haste/fluido e do atrito rotor/estator. Com viscosidade do fluido de 1 cp, observa-se que o back spin começa com a rotação reversa e atinge rapidamente rotações maiores, mas inverte o sentido de rotação e passa a operar no sentido normal, posteriormente invertendo novamente o sentido de rotação, passando a operar em rotação reversa até parar. Portanto, um observador externo descreveria de forma diferente o fenômeno nos dois casos em que a única diferença é a viscosidade do fluido.

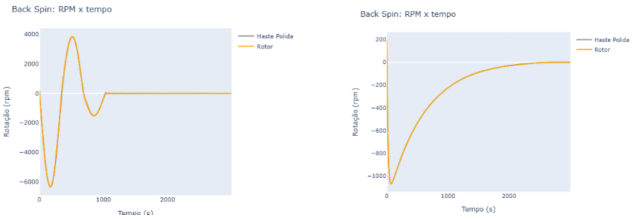


Figura 2: *RPM vs Tempo com 1cp (esq.) e 100cp (dir.), respectivamente*

A dinâmica de torque, apresentada na Figura 3 apresenta comportamento oscilatório ou de estabilização diretamente atrelado às respostas do nível e RPM evidenciadas.

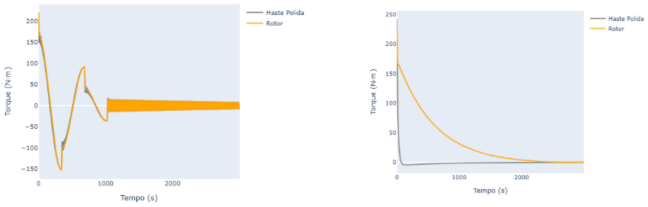


Figura 3: *Torque vs Tempo com 1cp (esq.) e 100cp (dir.), respectivamente*

5.4.2 Ensaio 2: Torque de Frenagem de 50 N · m

A mesma variação de viscosidade (1 cP e 100 cP) foi aplicada em um segundo cenário, considerando um torque de frenagem constante de 50 N · m.

Na avaliação dos níveis dinâmicos sob frenagem de 50 N · m, mantém-se a tendência geral de elevação na coluna e redução no anular.

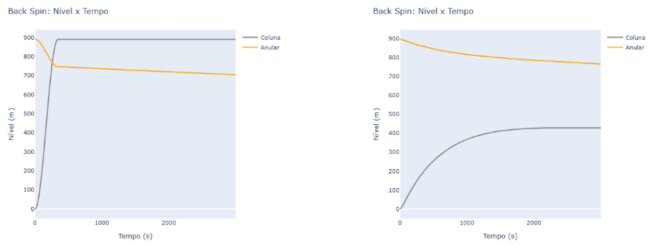


Figura 4: *Nível vs Tempo com 1cp (esq.) e 100cp (dir.), respectivamente*

No comportamento de rotação (RPM), a coluna de hastes inverte o sentido de rotação rapidamente e permanece em rotação reversa até cessar o movimento.

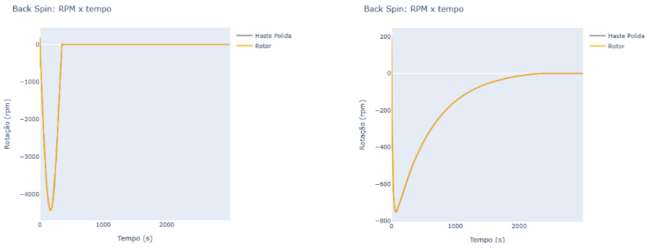


Figura 5: *RPM vs Tempo com 1cp (esq.) e 100cp (dir.), respectivamente*

Em relação à evolução do torque sob frenagem, observa-se a adaptação do sistema até alcançar o patamar de estabilização correspondente aos limites operacionais de frenagem.

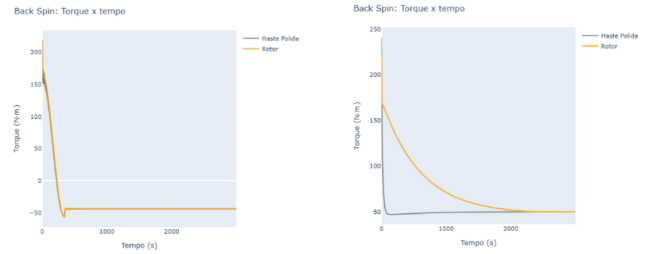


Figura 6: *Torque vs Tempo com 1cp (esq.) e 100cp (dir.), respectivamente*

6 Conclusões

A modelagem matemática aliada à simulação computacional demonstrou ser eficiente para simular com sucesso o comportamento dinâmico de bombeio de uma unidade BCP, bem como o fenômeno de reversão (*Backspin*). Foi possível simular poços em operação normal e com alta viscosidade, variando o valor de torque de frenagem para simular um mecanismo de frenagem. Dessa forma, o Simulador pode ser utilizado com sucesso para calcular a energia de frenagem para diferentes casos e observar o tempo de frenagem baseado no valor selecionado para o torque.

Como recomendação futura e visando a validação física do modelo matemático desenvolvido, os resultados numéricos do simulador devem exaustivamente comparados com dados experimentais, ainda não disponíveis.

Referências

ARAÚJO, R. F. F. *Análise do Fenômeno da Reversão em Poços de Petróleo Equipados com o Sistema de Bombeio por Cavidades Progressivas.* Dissertação (Mestrado) — Universidade Federal do Rio Grande do Norte, Natal, RN, 2022.

ASSMANN, B. W. *Estudo de Estratégias de Otimização para Poços de Petróleo com Elevação por Bombeio de Cavidades Progressivas.* Tese (Doutorado) — Universidade Federal do Rio Grande do Norte, Natal, RN, 2008.

CHOLET, H. *Progressing Cavity Pumps.* Paris: Editions Technip, 1997.

ERNST, T. et al. Back spin control in progressive cavity pump for oil well. *IEEE/PES Transmission & Distribution Conference and Exposition: Latin America*, 2006.

International Organization for Standardization. *Petroleum and natural gas industries – Progressing cavity pump systems for artificial lift.* Geneva, Switzerland, 2009.

VIDAL, F. J. T. *Desenvolvimento de um simulador de bombeio por cavidades progressivas.* Dissertação (Mestrado) — Universidade Federal do Rio Grande do Norte, Natal, RN, 2005.

WEG Equipamentos Elétricos. *Controle de Reversão de Hastes no Método de Elevação Artificial por Bombeio por Cavidades Progressivas (BCP) - Quarto Relatório Parcial.* Natal, RN, 2026.

Índice de autores

Alice Freire, 99, 107
Ana Sophia Diógenes Soares, 2
Andres Ortiz, 85
André Laurindo Maitelli, 21, 57, 72, 85, 115
Araújo, Marina L., 91
Arthur Phelipe de Oliveira Queiroz, 49, 57
Aurea da Silva Cabral, 21

Camila Raquel Sena de Almeida, 115
Carlos E. T. Dórea, 64
Carlos M. D. Viegas, 99, 107

Diego Moura, 85
Dominic Nascimento Rodrigues, 72
Dórea, C. E. Trabuco, 91

Elton José Figueiredo de Carvalho, 77
Emelyn Clementino Freire, 49, 57
Erick Justino, 99, 107
Ernane Ferreira Rocha Junior, 42

Fabio Meneghetti Ugulino de Araujo, 26, 34
Felipe R. Silva, 64
Fernando Eduardo Dias, 2
Fernando Henrique de Couto, 21
Fábio Meneghetti Ugulino de Araujo, 2

Guilherme da Costa Medeiros Nascimento, 115
Guilherme Nascimento da Silva, 34
Guilherme Souza de Farias, 57

Hilton Machado, 99, 107

Ignacio Sánchez-Gendriz, 42
Israel Lucas, 9
Ivanovitch Silva, 42, 99, 107
Iwri Simplicio Felipe, 77

Jaysa Keylla Siqueira Barbosa, 57
José A. A. Oliveira, 42
José Higor Rodrigues Soares, 14
José Patrocínio da Silva, 14

Kaillane Araújo da Silva, 26

Leonardo Cavalcanti, 99
Levi da Silva Lopes, 14
Luiz Affonso Guedes, 42
Luiz F. S. Silva, 42

Marcelo Jorge Oliveira de Albuquerque Melo, 49
Marianne Silva, 99, 107
Mateus Araujo, 99
Miguel Euripedes, 107
Morsinaldo Medeiros, 99, 107

Nicholas W. Lopes, 64

Pedro Manuel Pereira Chamorro, 115

Ricardo Lúcio de Araújo Ribeiro, 14
Rodrigo Dacol, 85
Rodrigo Prado de Medeiros, 14, 34
Rutácio de Oliveira Costa, 115

Sayonara Cristina de O. Magalhães, 57
Sayonara Cristina de Oliveira Magalhães, 49

Thais Medeiros, 107
Thiago de Oliveira Alves Rocha, 14
Thyfanne Dayane Cosme do Nascimento, 72
Tiago Barros, 42

Victor Gomes, 42

Expediente

III Simpósio de Automação e Sistemas – III SAS
12 e 13 de novembro de 2026 – Natal/RN
Organização: Laboratório de Automação (LAUT), DCA/UFRN.

Anais organizados pela Comissão Técnica do III SAS.
Apoio: FUNPEC, Idus Inteligência, eSIG Group, Centro de Tecnologia, DCA, LAUT.