



UNIVERSIDADE
FEDERAL DE
ALAGOAS



UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Dissertação de mestrado

**Modelo substituto para controle adaptativo do ponto de
operação do amplificador óptico baseado em Aprendizado
de Máquina**

Leandro Martins de Freitas

lmfreitas@ic.ufal.br

Orientador:

Erick de Andrade Barboza

MACEIÓ, SETEMBRO DE 2021

Leandro Martins de Freitas

**Modelo substituto para controle adaptativo do ponto de
operação do amplificador óptico baseado em Aprendizado
de Máquina**

Orientador: Prof.º Drº Erick de Andrade Baboza

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal de Alagoas como requisito parcial para obtenção do grau de Mestre em Informática.

**MACEIÓ-AL
2021**

Catálogo na Fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico

Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4 - 1767

F866m Freitas, Leandro Martins de.

Modelo substituto para controle adaptativo do ponto de operação do amplificador óptico baseado em aprendizado de máquina / Leandro Martins de Freitas. – 2021.

41 f. : il.

Orientador: Erick de Andrade Barboza.

Dissertação (mestrado em Informática) - Universidade Federal de Alagoas. Instituto de Computação. Maceió, 2021.

Bibliografia: f. 34-36.

Apêndices: f. 36-41.

1. Amplificadores ópticos. 2. Aprendizagem de máquina. 3. Redes dinâmicas. 4. Modelos substitutos. 5. Modelos lineares. I. Título.

CDU: 004.81:159.953.5



UNIVERSIDADE FEDERAL DE ALAGOAS/UFAL
Programa de Pós-Graduação em Informática – PPGI
Instituto de Computação/UFAL
Campus A. C. Simões BR 104-Norte Km 14 BL 12 Tabuleiro do Martins
Maceió/AL - Brasil CEP: 57.072-970 | Telefone: (082) 3214-1401



Folha de Aprovação

LEANDRO MARTINS DE FREITAS

MODELO SUBSTITUTO PARA CONTROLE ADAPTATIVO DO PONTO DE OPERAÇÃO DO AMPLIFICADOR ÓPTICO BASEADO EM APRENDIZADO DE MÁQUINA

Dissertação submetida ao corpo docente do Programa de Pós-Graduação em Informática da Universidade Federal de Alagoas e aprovada em 28 de setembro de 2021.

Banca Examinadora:

Prof. Dr. ERICK DE ANDRADE BARBOZA
UFAL – Instituto de Computação
Orientador

Prof. Dr. ICARO BEZERRA QUEIROZ DE ARAUJO
UFAL – Centro de Ciências Agrárias
Examinador Interno

Prof. Dr. CARMELO JOSE ALBANEZ BASTOS FILHO
UPE – Universidade de Pernambuco
Examinador Externo

Agradecimentos

Agradeço,

Primeiramente a Deus, que com sua misericórdia, apesar de minhas imperfeições, me ajudou a chegar aqui. Agradeço pela sabedoria e guia que Ele me fornece a todos os momentos.

Ao meu pai, José Ivan Oliveira de Freitas, pelo incentivo de buscar sempre mais conhecimento, mas manter-se humilde acima de tudo. À minha mãe, Maria Célia Martins de Freitas, pelo carinho e sábios conselhos, sempre nos momentos mais incertos da minha vida. E à minha esposa, Ágata dos Santos Paranhos Ferro Freitas, que me suporta e auxilia a todo momento, sendo um porto seguro em tempos tempestuosos.

Aos meus amigos de graduação e trabalho que continuam ao meu lado e sempre dispostos a me ajudar e me animar, não importando a situação. Aos professores do Programa de Pós Graduação em Informática da UFAL e aos colegas de curso pelo apoio.

Ao professor e orientador Erick de Andrade Barboza pela paciência e guia no processo de desenvolvimento e escrita desse trabalho, bem como pelo compartilhamento de seu precioso conhecimento.

Resumo

O controle adaptativo do ponto de operação do amplificador óptico (ACOP) é um dos desafios para a operação dinâmica em redes e comunicações ópticas. As abordagens ACOP visam definir os ganhos dos amplificadores ópticos de forma dinâmica para aumentar a qualidade de transmissão após uma cascata de amplificadores. Uma das abordagens ACOP mais recente usa um algoritmo de otimização evolutiva multiobjetivo para definir os ganhos dos amplificadores para maximizar a relação sinal-ruído óptica (OSNR) e minimizar a ondulação da OSNR. Apesar dos resultados promissores em relação à Qualidade de Transmissão, contar com um algoritmo evolutivo para tomar decisões em tempo real não é desejável, pois seu método iterativo geralmente implica em alto tempo de execução. Portanto, este trabalho propõe um modelo substituto que pode obter soluções tão boas quanto o algoritmo evolutivo multiobjetivo, mas em menos tempo. Foram consideradas cinco técnicas de regressão de aprendizado de máquina (ML), treinadas com as soluções do algoritmo de otimização. Os resultados mostraram que, para todos os casos, o erro mediano da regressão é menor que 1,15 dB e que um regressor pode ser usado para definir os ganhos dos amplificadores e as perdas dos atenuadores ópticos variáveis de um link óptico inteiro. Ele também mostrou que o regressor mais simples é 28.000 vezes mais rápido do que a abordagem de otimização evolutiva.

Palavras-chave: Amplificador Óptico, Rede dinâmica, Otimização de ganho, Modelo substituto, Regressão Linear, Aprendizado de máquina.

Abstract

The adaptive control of optical amplifier operating point (ACOP) is one of the problems presented in the challenge of Dynamic operation in optical communication and networks. The ACOP approaches aim to define the gains of the optical amplifiers dynamically to increase the transmission quality after a cascade of amplifiers. One of the most recent ACOP approach uses a multi-objective evolutionary optimization algorithm to define the gains of the amplifiers to maximize the optical signal to noise ratio (OSNR) and minimize OSNR ripple. Despite the promising results regarding Quality of Transmission, relying on an evolutionary algorithm to make decisions in real-time is not desirable because its iterative method usually implies high execution time. Therefore, this work proposes a surrogate model that can obtain solutions as good as the multi-objective evolutionary algorithm but in less time. We considered five machine learning (ML) regression techniques, trained with the optimization algorithm solutions. Results showed that for all cases, the regression median error is less than 1.15 dB and that one regressor can be used to define amplifiers' gains and variable optical attenuators' losses of an entire optical link. It also showed that the most straightforward regressor is 28,000 times faster than the evolutionary optimization approach.

Keywords: Optical amplifier, Dynamic network, Gain optimization, Surrogate model, Linear Regression, Machine Learning.

Lista de ilustrações

Figura 1 – Máscara de potência associada à Planicidade do Ganho (a) e Figura de Ruído (b). Fonte: (OLIVEIRA et al., 2013)	7
Figura 2 – Exemplo de espaço de busca para potência de entrada de -15 dBm. Fonte: (OLIVEIRA et al., 2013)	7
Figura 3 – Partições do espaço de busca (à esquerda) e representação gráfica de uma árvore de classificação (à direita). Fonte: (LOH, 2011)	13
Figura 4 – Etapas de desenvolvimento de um modelo substituto	16
Figura 5 – Um problema com duas funções objetivo: <i>Ripple</i> e OSNR. A frente de Pareto está destacada pela linha que une os pontos entre A e E. Fonte: (BARBOZA, 2017)	17
Figura 6 – Ilustração da interação entre o NSGA-II e o simulador do sistema óptico para a realização da otimização. Fonte: adaptado de (BARBOZA, 2017)	18
Figura 7 – Ativação dos canais do amplificador óptico caracterizado.	19
Figura 8 – Ilustração das três etapas de extração de dados: 1) usar MOO para definir ganhos de amplificadores e perdas de VOAs, 2) selecionar 30% das melhores soluções e 3) adicionar soluções selecionadas ao base de dados.	20
Figura 9 – Ilustração das três estratégias de alocação de regressores para um enlace com 2 amplificadores.	22
Figura 10 – Distribuição do erro médio absoluto (MAE) da fase de teste para ganho do amplificador e estimativa de perda do VOA considerando (a, d, g) um regressor por variável (\mathbf{R}/\mathbf{V}), (b, e, h) um regressor por amplificador (\mathbf{R}/\mathbf{A}) e (c, f, i) um regressor por enlace (\mathbf{R}/\mathbf{E}) para 2, 5 e 8 amplificadores. Os cinco regressores considerados são: Bayesian Ridge Regressor (BR), Random Forest Regressor (RF), Decision Tree Regressor (DT), Lasso CV (LCV) e SimpleMeanRegressor (SM).	26

Figura 11 – Distribuição do erro médio absoluto (MAE) da fase de teste para ganho do amplificador e estimativa de perda do VOA considerando 2 a 8 amplificadores. Os cinco regressores considerados são: Bayesian Ridge Regressor (BR), Random Forest Regressor (RF), Decision Tree Regressor (DT), Lasso CV (LCV) e SimpleMeanRegressor (SM).	28
Figura 12 – Distribuição dos valores de ganho do amplificador e perda do VOA para cada base de dados obtida através da execução do MOO, considerando diferentes números de amplificadores. A_i representa o amplificador i	29
Figura 13 – Ripple OSNR e OSNR mínima retornadas pelo algoritmo MOO proposto em (BARBOZA; BASTOS-FILHO; FILHO, 2019) e pelas técnicas de regressão utilizadas neste trabalho considerando 10, 20 e 40 canais. Os cinco regressores considerados são: Bayesian Ridge Regressor (BR), Random Forest Regressor (RF), Decision Tree Regressor (DT), Lasso CV (LCV) e SimpleMeanRegressor (SM). Apenas um regressor é utilizado para o enlace inteiro.	30
Figura 21 – Frentes de Pareto para enlace de 2 amplificadores	39
Figura 22 – Frentes de Pareto para enlace de 3 amplificadores	40
Figura 23 – Frentes de Pareto para enlace de 4 amplificadores	40
Figura 24 – Frentes de Pareto para enlace de 5 amplificadores	40
Figura 25 – Frentes de Pareto para enlace de 6 amplificadores	41
Figura 26 – Frentes de Pareto para enlace de 7 amplificadores	41
Figura 27 – Frentes de Pareto para enlace de 8 amplificadores	41

Lista de tabelas

Tabela 1 – Organização das variáveis do problema em um indivíduo do NSGA-II. As primeiras N posições contém os valores dos ganhos dos amplificadores e as posições restantes contém os valores das perdas dos VOAs.	18
Tabela 2 – Demonstração do funcionamento do modelo <i>SimpleMeanRegressor</i>	23
Tabela 3 – Instâncias retornadas pelo MOO para enlaces de 2, 3, 4, 5, 6, 7 e 8 amplificadores	25
Tabela 4 – Mediana do erro médio absoluto (MAE) da validação cruzada de 5 <i> folds</i> para a estratégia <i>SimpleMeanRegressor</i>	27
Tabela 5 – Tempo médio de execução para cada a técnica após dez execuções independentes.	31

Lista de Abreviaturas e Siglas

ACOP	<i>Controle Adaptativo do Ponto de Operação do Amplificador Óptico;</i>
ASE	<i>Emissão Espontânea Amplificada;</i>
BR	<i>Bayesian Ridge;</i>
DT	<i>Decision Tree;</i>
GF	<i>Planicidade de Ganho;</i>
LCV	<i>Lasso CV;</i>
MAE	<i>Erro Médio Absoluto;</i>
ML	<i>Aprendizado de Máquina;</i>
MOO	<i>Otimização Multiobjetivo;</i>
NF	<i>Figura de Ruído;</i>
NLI	<i>Interferência Não Linear;</i>
OA	<i>Amplificador Óptico;</i>
OSNR	<i>Razão Sinal-Ruído Óptica;</i>
PF	<i>Frente de Pareto;</i>
QoT	<i>Qualidade de Transmissão;</i>
R/A	<i>Regressor por Amplificador;</i>
R/E	<i>Regressor por Enlace;</i>
R/V	<i>Regressor por Variável;</i>
RF	<i>Random Forest;</i>
SM	<i>Simple Mean Regressor;</i>
VOA	<i>Atenuador Óptico Variável;</i>

Sumário

Lista de ilustrações	iv
Lista de tabelas	vi
Sumário	viii
1	Introdução 1
1.1	Contexto e Problema 1
1.2	Objetivos e Relevância 2
1.2.1	Objetivo Geral 2
1.2.2	Objetivos Específicos 2
1.2.3	Relevância da Proposta 3
1.3	Contribuições 3
1.4	Estrutura do trabalho 3
2	Trabalhos relacionados 5
2.1	Aprendizado de máquina em comunicações ópticas 5
2.2	Modelos substitutos 5
2.3	Amplificadores Ópticos Autônomos 6
3	Fundamentação Teórica 10
3.1	Comunicações Ópticas 10
3.2	Regressão 11
3.2.1	Least Absolute Shrinkage and Selection Operator (LASSO) e Bayesian Ridge 12
3.2.2	Decision Tree 13
3.2.3	Random Forest 14
3.3	Modelos Substitutos 14
3.4	Conceitos básicos de Otimização Multi-Objetivo 16
3.5	Otimização multiobjetivo no problema de ACOP 17
4	Metodologia 18
4.1	Extração de Dados 18
4.2	ACOP como um problema de regressão 21

4.3	Definição e treinamento dos modelos	22
4.4	Comunicação entre simulador óptico e modelos de regressão	23
5	Resultados e Discussão	25
6	Conclusão	32
Referências		34
APÊNDICE A Erro médio absoluto para fase de teste em enlaces de 2 a		
	8 amplificadores	36
APÊNDICE B Desempenho dos regressores em comparação ao MOO		
	em enlaces de 2 a 8 amplificadores	39

1 Introdução

1.1 Contexto e Problema

A massiva utilização de serviços e aplicações que precisam de muita banda de transmissão tem aumentado o tráfego de dados nas redes de telecomunicações, principalmente na Internet. O aparecimento de serviços como vídeo por demanda, IP-TV, voz sobre IP, entre outros, contribuiu para que o tráfego global da rede IP aumentasse mais de 17 vezes entre os anos de 2012 e 2017. Além disso, a perspectiva é de que este tráfego ainda irá aumentar cerca de sete vezes até 2022 (FORECAST et al., 2019).

Comunicações ópticas possuem vantagens que as tornam o principal meio de transmissão utilizado para suprir a atual demanda por dados. No princípio, os sistemas de comunicações ópticas eram utilizados apenas para comunicações ponto-a-ponto. Porém, atualmente estes sistemas são utilizados para formar redes que interligam vários dispositivos. Devido às suas vantagens, as redes ópticas são utilizadas como *backbone* de várias outras redes (*ex*: redes móveis/celulares, redes IP), o que faz com que, na maioria das transmissões, os dados trafeguem em algum momento através de uma rede óptica (TANENBAUM; WETHERALL, 2011).

Devido às novas demandas de serviços de comunicação digital as redes ópticas precisam se tornar dinâmicas. Redes ópticas dinâmicas desafiam os dispositivos a se tornarem autoadaptativos (autônomos) (European Technology Platform Photonics21, 2021).

Amplificadores ópticos são dispositivos cruciais para o bom funcionamento de sistemas de comunicação óptica e têm um alto impacto na Qualidade de Transmissão (QoT). Assim, é necessário criar mecanismos que permitam que os amplificadores ópticos funcionem de forma autônoma. Um passo essencial para esse objetivo é definir autonomamente o ganho do amplificador sempre que ocorrer uma mudança no sistema. Este problema foi definido em (BARBOZA et al., 2017) como Controle Adaptativo do Ponto de Operação do Amplificador Óptico (ACOP - *Adaptive Control of optical amplifier Operating Point*) -, e algumas abordagens para resolver este problema já foram propostas (BARBOZA et al., 2017; BARBOZA; BASTOS-FILHO; FILHO, 2019; MOURA et al., 2016; OLIVEIRA et al., 2013).

Em (BARBOZA; BASTOS-FILHO; FILHO, 2019) os autores propuseram o uso de um Atenuador Óptico Variável (VOA) na saída do amplificador para obter melhores soluções

ACOP. A abordagem usou um algoritmo evolutivo de otimização multiobjetivo (MOO - *multi-objective optimization*) para definir o ganho do amplificador e a perda de VOA. Os resultados mostraram que o VOA aumenta o número de soluções com bom QoT. O uso do MOO também se mostrou mais eficiente do que as técnicas locais no fornecimento de melhores soluções para o problema.

Apesar dos resultados promissores, confiar em um algoritmo evolutivo para tomar decisões em tempo real não é desejável. Por ser uma abordagem iterativa, essa classe de algoritmo precisa de um tempo considerável para retornar soluções aceitáveis. Esta categoria de abordagem é muito adequada para projetar um link óptico (planejamento de rede), mas não para a operação dinâmica da rede óptica. Além disso, o resultado da abordagem MOO é um conjunto de soluções, o que torna necessária a definição de uma política de escolha para atuar após o processo de otimização, que pode não ser simples devido a objetivos conflitantes.

Neste trabalho, foi investigada a criação de um modelo substituto para a obtenção de uma solução para o ACOP que seja tão boa quanto o algoritmo multiobjetivo, mas que seja obtida em menos tempo. O objetivo é avaliar como criar um modelo que possa substituir, com um bom nível de precisão, o algoritmo evolutivo MOO na tarefa de definir os ganhos dos amplificadores e as perdas dos VOAs em um enlace óptico. Este modelo pode aprender os padrões de ganhos e perdas presentes nas soluções retornadas pelo algoritmo MOO e reproduzir suas escolhas sem um processo iterativo.

1.2 Objetivos e Relevância

1.2.1 Objetivo Geral

Desenvolver um modelo substituto que retorne boas soluções para o problema de ACOP de forma rápida e eficiente.

1.2.2 Objetivos Específicos

- Desenvolver uma base de dados composta de soluções retornadas pelo algoritmo ACOP MOO proposto em (BARBOZA; BASTOS-FILHO; FILHO, 2019).
- Desenvolver múltiplos modelos substitutos, utilizando técnicas de regressão diferentes.

- Avaliar os modelos regressivos de forma a apontar qual deles retorna as melhores soluções para o problema de ACOP em menos tempo.

1.2.3 Relevância da Proposta

Considerando que em um ambiente dinâmico o estado da rede óptica mudará com frequência, e considerando que a quantidade de dados transportados em um link desta rede é da ordem de Terabytes por segundo, é desejável que mudanças que visem o melhor funcionamento da rede sejam feitas em um intervalo de tempo muito pequeno. Em problemas de proteção de tráfego em redes ópticas, por exemplo, o tempo máximo para que a rede se recupere de uma determinada falha é de 50 ms (NISHITANI et al., 2017). Apesar de não existir um requisito de tempo mínimo para a definição dos ganhos dos amplificadores, o exemplo anterior mostra que é desejável ter mudanças na casa dos milissegundos.

Diante disso, a implantação do modelo proposto possibilitará que um amplificador óptico adapte seu ponto de operação de acordo com a mudança no enlace em que está inserido de uma forma mais rápida, assim, otimizando o funcionamento da rede em um cenário dinâmico de forma satisfatória.

1.3 Contribuições

- Criação da base de dados utilizando múltiplas execuções do algoritmo multiobjetivo;
- Sugestão de modelo de regressão para solução do problema;
- Obtenção de soluções para o problema de ACOP em menor tempo em relação à abordagem multiobjetivo;
- Análise de eficiência dos modelos propostos considerando cenários com diferentes números de amplificadores.

1.4 Estrutura do trabalho

Este trabalho está estruturado da seguinte forma: na Seção 2 são apresentados trabalhos sobre otimização de redes ópticas e utilização de modelos substitutos. Na Seção 3 será fornecido um breve contexto sobre comunicações ópticas, problemas de regressão e modelos

substitutos. Na Seção 4 serão descritos os métodos de extração de dados e definição e treinamento dos modelos substitutos. Os resultados obtidos são demonstrados e analisados na Seção 5 e, por fim, a conclusão do trabalho é apresentada na Seção 6.

2 Trabalhos relacionados

2.1 Aprendizado de máquina em comunicações ópticas

Há na literatura diversos trabalhos cujo objetivo é utilizar técnicas de aprendizado de máquina para otimizar ou prever o desempenho de comunicações ópticas. Em (ROTTONDI et al., 2018), por exemplo, um classificador binário baseado em Random Forest é utilizado para prever se a Taxa de Erro de Bits (BER - Bit Error Rate) de rotas ainda não estabelecidas excede um certo limite. O estudo mostra que apesar de uma alta acurácia nas predições, um grande número de rotas auxiliares deve ser adicionado à rede para que seja possível avaliar a BER.

Em (FERNÁNDEZ et al., 2013) e (FERNÁNDEZ et al., 2015) o objetivo foi o desenvolvimento de um modelo de predição de tráfego para reconfiguração da topologia da rede. Foi proposto um método Auto-Regressivo Integrado de Médias Móveis (ARIMA - *Auto Regressive Integrated Moving Average*), cuja entrada é uma base de dados em formato de série temporal. A partir daí um modelo de decisão é utilizado para realizar a reconfiguração da rede.

Quanto à detecção de falhas, é possível citar o trabalho realizado em (VELA et al., 2017), o qual propõe dois algoritmos. No primeiro algoritmo, um conjunto de dados históricos da rede (tais como BER máximo, limite de BER desejado e BER atual) é utilizado para detectar mudanças abruptas na taxa de erro, característica essa que pode indicar falhas. A saída desse algoritmo, em adição a outros parâmetros, é analisada pelo segundo modelo proposto, cujo objetivo é prever o tipo de falha que pode ter gerado essa anomalia.

2.2 Modelos substitutos

A aplicação de modelos substitutos também é um estudo emergente, tendo várias aplicações para redes ópticas. Em (ARAÚJO; BASTOS-FILHO; MARTINS-FILHO, 2015), por exemplo, foi proposto um algoritmo evolutivo multiobjetivo para projetar redes ópticas. Se fez uso de modelos substitutos para acelerar a avaliação da aptidão, proporcionando um tempo de execução até 84% e 88% inferior ao tempo necessário por abordagens anteriores para tráfego uniforme e não uniforme, respectivamente.

A aplicação de modelos substitutos também é um estudo emergente, sendo utilizada em várias áreas. Em (WESTERMANN; WELZEL; EVINS, 2020), por exemplo, um modelo substituto baseado em redes neurais é utilizado para estimar a demanda de aquecimento ou resfriamento de casas utilizando variáveis de clima.

Um modelo substituto conjunto para identificação de fontes de contaminação em lençóis freáticos foi apresentado em (XING et al., 2019). Neste trabalho três estratégias de regressão são utilizadas para formar o modelo final, sendo elas Kriging, RBF e LSSVM. O modelo final supera as desvantagens individuais das abordagens, porém necessita de mais tempo de processamento.

Em (DONG; QIN; WANG, 2019) um modelo substituto multiobjetivo baseado em Redes Neurais é utilizado para definir o design de antenas de forma a otimizar seu ganho, tamanho, eficiência, etc. Os resultados são promissores, mostrando melhor desempenho e boa acurácia com relação ao uso de simulações e outras estratégias de otimização de antenas.

No trabalho (SONG et al., 2018), um algoritmo de otimização multiobjetivo usando modelo substituto é apresentado. Seu objetivo é a otimização do bombeamento de água em lençóis freáticos de grande escala, minimizando o aumento da massa de água salgada e maximizando a taxa de bombeamento em aquíferos costeiros. Em comparação à simulação, o algoritmo proposto pode levar a 94% de economia de tempo de CPU, e ainda assim garantir a precisão das soluções.

2.3 Amplificadores Ópticos Autônomos

No trabalho desenvolvido em (OLIVEIRA et al., 2013) objetivou-se automatizar amplificadores ópticos utilizando informações de suas máscaras de potência. As máscaras de potência foram obtidas através da medição da Planicidade de Ganho (GF - *Gain Flatness*) e Figura de Ruído (NF - *Noise Figure*) para cada combinação possível de ganho do amplificador e potência de entrada. Um exemplo das máscaras de GF e NF pode ser encontrado nas Figuras 1a e 1b, respectivamente.

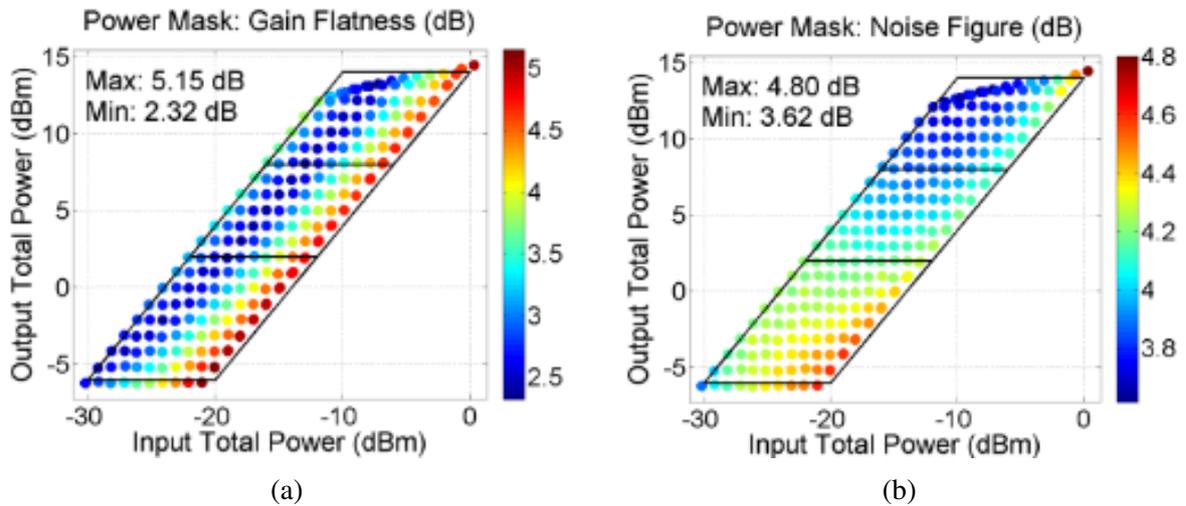


Figura 1 – Máscara de potência associada à Planicidade do Ganho (a) e Figura de Ruído (b).
Fonte: (OLIVEIRA et al., 2013)

Para cada possível potência de entrada, foi associado um conjunto de pontos no espaço de busca do mecanismo de automação, cujas coordenadas consistem do GF e NF das máscaras de potência, demonstrado na Figura 2. O ponto com menor distância euclidiana à origem é aquele que impacta mais positivamente na qualidade de transmissão do enlace.

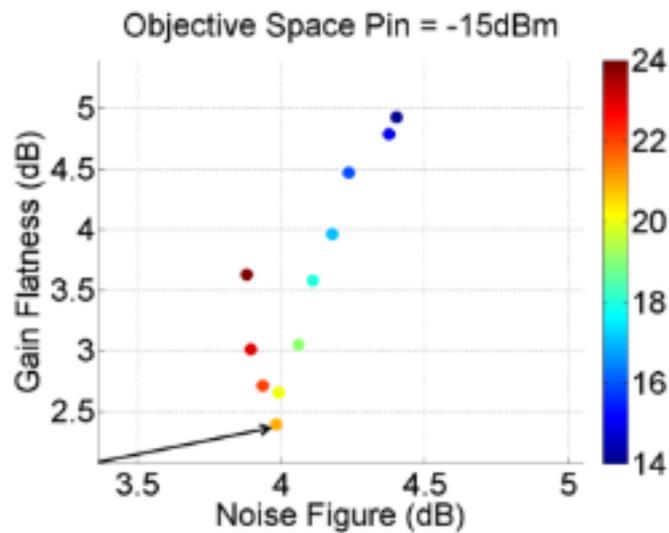


Figura 2 – Exemplo de espaço de busca para potência de entrada de -15 dBm. Fonte: (OLIVEIRA et al., 2013)

Em (MOURA et al., 2016), foi proposta uma abordagem de Raciocínio Baseado em Casos (CBR - *Case-based Reasoning*) para ajuste da ganho do amplificador óptico. Uma estratégia de CBR consiste em resolver um problema reutilizando informação e conhecimento de um problema anterior similar já resolvido. Os resultados confirmaram a capacidade de aprendizagem da metodologia e sua adequação em relação à topologia da rede, tamanho e modelos de amplificadores.

O trabalho realizado em (BARBOZA et al., 2017) define o problema de ACOP e propõe abordagens locais e globais para resolvê-lo. Foram propostos dois novos métodos de abordagem local e um global, que são comparados a outros encontrados na literatura. Os métodos de abordagem local propostos superaram os demais em desempenho e tempo de execução. A abordagem global proposta obteve os melhores resultados quanto à Figura de Ruído e Planicidade do Ganho da cascata de amplificadores. No entanto, o tempo de execução pode ser muito longo para enlace com mais de três amplificadores, visto que utiliza um método exaustivo.

Em (BARBOZA; BASTOS-FILHO; FILHO, 2019) foi apresentada uma nova estratégia para minimizar o impacto da interferência não linear no ACOP, considerando um VOA agregado à saída de cada amplificador presente no enlace. Também foi proposta uma abordagem global baseada em otimização multiobjetivo (MOO), a qual define tanto ganhos dos amplificadores quanto perdas dos VOAs, visando maximizar o OSNR e minimizar sua ondulação. Os resultados mostraram que a consideração de VOAs aumenta o número de soluções com bom QoT. Além disso, a abordagem global baseada em MOO retorna melhores soluções do que as abordagens ACOP locais disponíveis na literatura.

O artigo (BASTOS-FILHO et al., 2020) consiste de uma investigação do uso de modelos de aprendizado de máquina para substituição do algoritmo multiobjetivo apresentado em (BARBOZA; BASTOS-FILHO; FILHO, 2019) como solução do problema de ACOP. Foi constatado que nenhum modelo conseguiu aproximar-se à QoT oferecida pela abordagem multiobjetivo, porém a execução dos modelos leva milissegundos, enquanto que o MOO pode levar minutos até fornecer soluções. Também é necessário apontar que nesse artigo, houve a tentativa de usar um modelo regressivo aplicado a enlaces de diferentes tamanhos, na prática porém, o tamanho do enlace raramente muda, e se mudar, levaria muito tempo para redefinir os ganhos.

Neste trabalho, a experiência acumulada de todos os estudos referentes ao ACOP será estendida e aprimorada. Os artigos (BARBOZA; BASTOS-FILHO; FILHO, 2019) e (BASTOS-FILHO et al., 2020) serão base desse estudo, no sentido de que serão utilizados o algoritmo MOO para criação da base de dados e modelos substitutos para resolver o problema. Neste trabalho, porém, é proposta uma nova estratégia para construção da base de dados, bem como novas abordagens para alocação e treinamento dos modelos.

3 Fundamentação Teórica

3.1 Comunicações Ópticas

Sistemas de comunicação óptica inicialmente tinham como objetivo conexão de dispositivos ponto-a-ponto. Atualmente, porém, redes ópticas são utilizadas como *backbones* de diversas outras redes, portanto há grandes chances de qualquer transmissão de dados pela rede utilizar vias ópticas.

O emprego desse tipo de transmissão se faz bastante interessante devido às vantagens providas pelo uso de fibra óptica: além da baixa suscetibilidade a interferências e perda de dados, a capacidade de transmissão é bastante grande, chegando a dezenas de gigabits ou terabits por segundo (RAMASWAMI; SIVARAJAN; SASAKI, 2009).

Um Amplificador Óptico (OA - *Optical Amplifier*) tem como objetivo compensar a perda de potência sofrida durante a transmissão, de forma a manter a potência do sinal em um nível aceitável para recepção. Apesar disso, os OAs podem alterar o formato do sinal e inserir ruído, prejudicando a qualidade do sinal.

Para transmissões multi-canais, OAs podem fornecer ganhos individuais para cada canal. Ao considerar a mesma potência de entrada para os diferentes canais de um enlace, pode-se definir o *Ripple* como a diferença entre as potências dos canais com maior e menor potência. Além disso, define-se a razão sinal ruído óptica (OSNR - *Optical Signal to Noise Ratio*) como a razão entre a potência do sinal e a potência do ruído em determinado ponto do sistema. Tais métricas são muito importantes para a medição de desempenho do amplificador e, conseqüentemente, qualidade da transmissão.

Cada amplificador adiciona ruído ao sinal. A proporção de ruído adicionado é dada pela figura de ruído do amplificador (em decibel). Em (BANEY; GALLION; TUCKER, 2000), Baney e colaboradores mostraram que o fator de ruído de uma cascata de amplificadores pode ser calculado pela Equação (1).

$$F_{cascata} = F_1 + \frac{F_2}{G_1 L_1} \dots + \frac{F_N}{G_1 L_1 \dots G_{N-1} L_{N-1}} + \dots + \frac{1}{G_1 L_1 \dots G_{N-1} L_{N-1} G_N}, \quad (1)$$

na qual F_i e G_i ($i = 1, 2, \dots, N$) são respectivamente o fator de ruído e o ganho do amplificador na posição i da cascata com N amplificadores, enquanto que L_i ($i = 1, 2, \dots, N - 1$)

é a perda do enlace i da cascata. A figura de ruído da cascata ($NF_{cascata}$) é simplesmente o fator de ruído da cascata em escala logarítmica, como mostra a Equação (2)

$$NF_{cascata} = 10 \cdot \log(F_{cascata}). \quad (2)$$

Diante disso, é possível perceber como a definição dos ganhos (pontos de operação) dos amplificadores impactam na quantidade de ruído adicionado ao sinal óptico que se propaga por um enlace óptico com mais de um amplificador.

3.2 Regressão

Regressão é uma técnica estatística cujo objetivo é analisar e inferir a relação entre m variáveis dependentes (variáveis de resposta) e n variáveis independentes (variáveis exploratórias) (CHATTERJEE; HADI, 2015).

Sejam Y_1, Y_2, \dots, Y_m as variáveis dependentes e X_1, X_2, \dots, X_n as variáveis independentes, a relação entre elas pode ser dada por:

$$(Y_1, Y_2, \dots, Y_m) = f(X_1, X_2, \dots, X_n) + (\epsilon_1, \epsilon_2, \dots, \epsilon_m)$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é a função de aproximação e ϵ_i representa o erro de aproximação para cada variável dependente Y_i . Portanto, quanto menor é a somatória de ϵ_i , mais precisa é a aproximação.

O valor final de Y_i pode ser definido como a combinação linear:

$$Y_i = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n + \epsilon_i$$

onde α_j é o coeficiente linear atribuído pela função f a cada variável independente, indicando sua influência na predição de Y_i .

Modelos de regressão possuem diversas aplicações no mundo real. No trabalho apresentado em (GOPALAKRISHNAN; CHOUDHARY; PRASAD, 2018), por exemplo, um modelo de regressão linear foi utilizado para avaliar as vendas de uma grande loja, de forma a maximizar seus lucros. Seu objetivo foi retornar os lucros previstos para o ano de 2014 utilizando como base os dados reais dos anos 2011 a 2013. Para avaliação do modelo, dados reais do ano de 2014 foram comparados com as predições, revelando uma acurácia de 84%.

Outra aplicação de técnicas de regressão é o trabalho (ROOPA; ASHA, 2019), o qual utiliza a base de dados de diabetes dos Índios Pima - grupo indígena dos Estados Unidos famoso pelos altos índices de diabetes. Os dados da base de dados são escolhidos via Análise de Componentes Principais (PCA - *Principal Component Analysis*) e o resultado é utilizado no treino de um modelo regressivo. Este modelo retorna a probabilidade da instância analisada ser de um paciente diabético. A abordagem chegou à marca 82.1% de acurácia, proporção maior do que estudos anteriores, que alcançaram 81.9% de acurácia.

As seções a seguir descrevem as diferentes estratégias de regressão utilizadas neste trabalho.

3.2.1 Least Absolute Shrinkage and Selection Operator (LASSO) e Bayesian Ridge

Modelos de regressão linear tendem a sofrer com o problema de *overfitting* (sobreajuste), no qual o modelo possui alta precisão na predição das instâncias utilizadas no teste, porém apresenta baixo desempenho para novas instâncias desconhecidas, pois o modelo aprende as peculiaridades dos dados de teste, em vez de encontrar uma regra preditiva geral (DIETRICH, 1995). O processo de regularização pode minimizar este problema, pois consiste da criação de penalidades nos coeficientes de regressão do modelo de forma "desencorajar" o desenvolvimento de um modelo que se adapte demais aos dados, facilitando a definição de uma regra geral.

Duas estratégias comuns de regularização são a Ridge e Lasso (MCDONALD, 2009). Na regularização Ridge os coeficientes são alterados podendo chegar a valores muito próximos a 0, porém nunca iguais a 0, dessa forma variáveis que têm maior influência no *overfitting* podem ter coeficientes diminuídos, diminuindo o sobreajuste.

A estratégia Lasso é bastante parecida, porém os coeficientes podem chegar a 0, indicando que a sua variável relacionada é excluída da regra preditiva. Este tipo de regularização, portanto, é bastante utilizada para retirar do problema variáveis que não importam - ou têm influência mínima - na geração da função de predição, o que chamamos de redução de dimensionalidade.

Neste trabalho utilizamos dois modelos de regressão linear que utilizam as estratégias de regularização descritas anteriormente:

- Lasso CV: modelo linear baseado em regularização Lasso com validação cruzada embutida.
- Bayesian Ridge: modelo linear baseado em regularização Ridge cuja regra preditiva é baseada em Inferência Bayesiana (BOX; TIAO, 2011).

3.2.2 Decision Tree

Numa Árvore de Decisão (*Decision Tree*) os dados são continuamente divididos em subconjuntos menores e mais específicos de forma a estabelecer regras para regressão ou classificação de dados. As árvores são compostas por folhas, que representam as decisões ou resultados finais, e nós, em que os dados são divididos.

À esquerda da Figura 3 é demonstrado o particionamento do espaço de busca de um problema de classificação e à direita sua representação gráfica resultante para um problema com duas variáveis independentes - X_1 e X_2 - e a variável de saída, cujos possíveis valores são 1, 2 e 3, representados pelas cores azul, vermelho e verde, respetivamente.

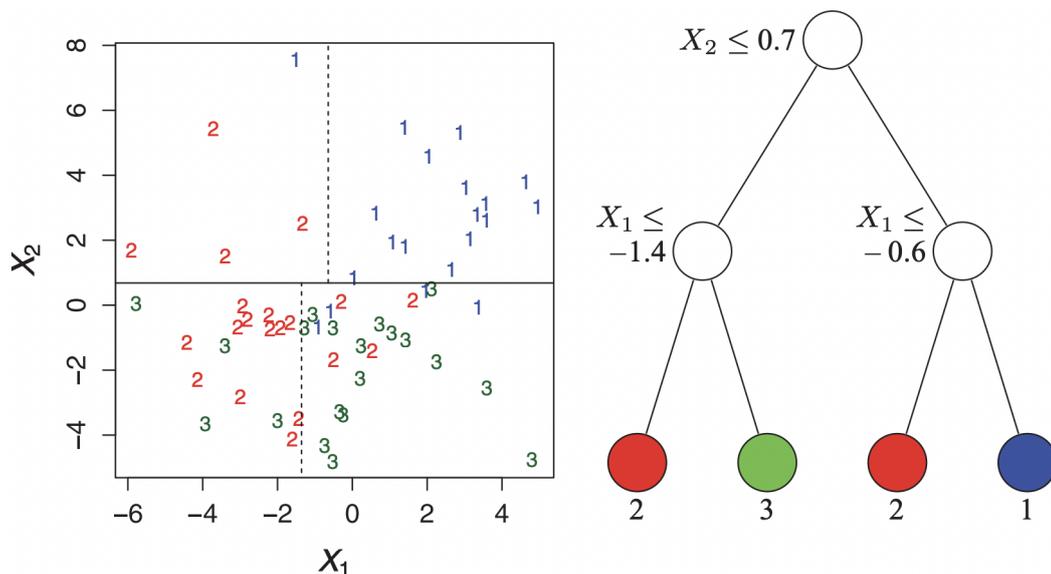


Figura 3 – Partições do espaço de busca (à esquerda) e representação gráfica de uma árvore de classificação (à direita). Fonte: (LOH, 2011)

O processo de classificação de uma instância segue a estrutura da árvore de forma iterativa: a cada nó, temos uma regra simples com duas possíveis saídas. Caso a regra seja atendida, o fluxo de verificação segue para o nó filho mais à esquerda do nó analisado, caso contrário o fluxo segue para o filho mais a direita. Esse processo ocorre até que uma folha seja alcançada, retornando predição do modelo. Uma árvore de regressão é semelhante a uma de classificação, exceto que a variável dependente assume valores ordenados e um modelo de regressão é ajustado a cada nó para fornecer seus valores previstos.

3.2.3 Random Forest

Random Forest é uma técnica de ML baseada em conjuntos: sua implementação inicializa diversas árvores de decisão ajustadas usando subconjuntos aleatórios do conjunto de dados original, provendo diferentes estruturas e regras para cada uma. Ao analisar uma instância, todas essas árvores retornam sua predição e o estimador principal (*Random Forest*) utiliza o voto majoritário (em problemas de classificação), ou a média das respostas retornadas pelas sub-árvores (no caso de problemas de regressão) para formar a previsão final.

O objetivo dessa aleatoriedade é diminuir a variância do estimador, visto que árvores de decisão individuais geralmente possuem alta variância e tendem ao *overfitting*. O modelo alcança uma variância reduzida combinando diversas árvores e diminui o problema de *overfitting*, bastante presente em árvores de decisão comuns.

3.3 Modelos Substitutos

A maioria dos problemas de engenharia exigem experimentos e/ou simulações para avaliação das soluções. Porém, para muitos problemas do mundo real, uma única simulação pode levar muitos minutos, horas ou mesmo dias para ser concluída. Há situações em que a avaliação de um experimento exige muito esforço computacional, o que pode inviabilizar o projeto dadas restrições de tempo, de recursos computacionais ou financeiras.

Uma alternativa às simulações é o uso de modelos substitutos (também chamados meta-modelos) que tentam imitar o comportamento da simulação o máximo possível e, ainda assim, ter um custo computacional menor. Modelos substitutos são construídos usando uma abordagem baseada em dados, portanto não é necessário que o funcionamento interno do

código de simulação seja conhecido (ou mesmo entendido), apenas o comportamento de entrada-saída é importante (GORISSEN et al., 2010).

É importante notar que não há garantia de que o modelo substituto implementado retorne com exatidão os resultados que seriam obtidos por meio de uma simulação; seu objetivo é minimizar o consumo de recursos utilizados na avaliação do problema e prever a saída que seria retornada pela simulação.

Quanto ao desenvolvimento do modelo, diversas técnicas podem ser citadas, tais como: Kriging, SVM, Random Forest, Redes Neurais, entre outras. É interessante evidenciar que não há uma técnica cujo desempenho é sempre melhor, pois a depender do problema, certa abordagem pode ser mais eficiente que outra.

Ao desenvolver um modelo substituto, tem-se como objetivo maximizar sua precisão utilizando o menor número de simulações possível. Este processo, como descrito em (VIANA et al., 2008), consiste de três passos principais:

1. **Escolha das amostras e realização de simulações:** nesta etapa são definidos os cenários que serão testados na simulação de forma a abranger uma parte representativa das situações possíveis, aumentando o ganho de informação;
2. **Definição do modelo substituto:** aqui o modelo substituto é definido e os dados obtidos na etapa anterior são utilizados para realizar seu treinamento;
3. **Avaliação do modelo (ou validação):** o modelo gerado a partir das etapas anteriores é avaliado através de diversas métricas para que seja possível a comparação entre os seus resultados e os obtidos através de simulações.

Na fase de avaliação é verificado se os resultados são satisfatórios, ou seja, se a precisão das predições é alta o suficiente. Caso o modelo não atinja o desempenho desejado, as etapas anteriores podem ser revisitadas.

Na Figura 4 é representado o ciclo do processo de desenvolvimento de um modelo substituto.

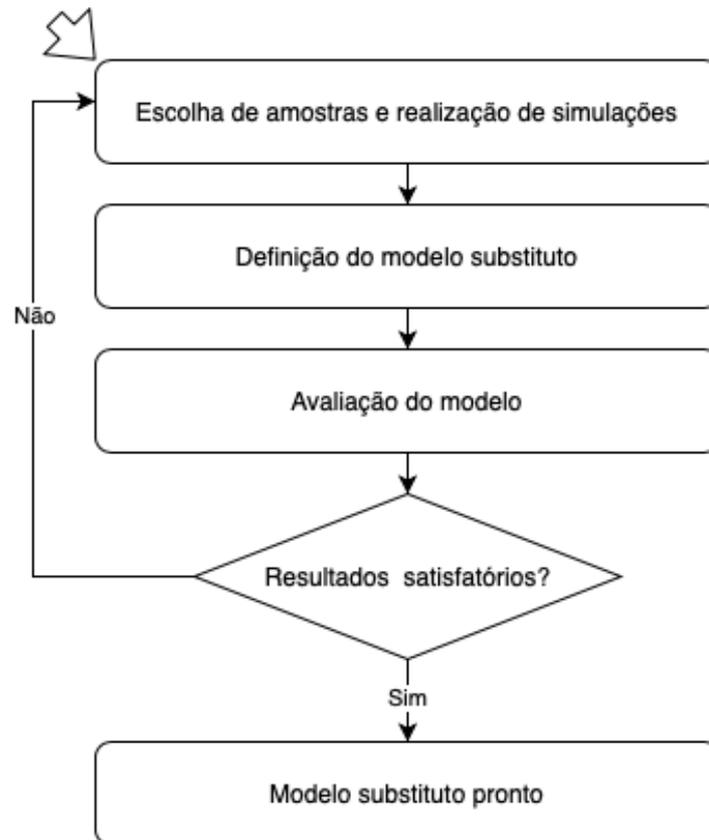


Figura 4 – Etapas de desenvolvimento de um modelo substituto

3.4 Conceitos básicos de Otimização Multi-Objetivo

Em um problema de otimização com objetivo único, tem-se um conjunto (vetor) de variáveis de decisão $\hat{x} = \{x_1, \dots, x_n\}$ que pertence a um espaço de decisão Ω e que deverá ter seus valores definidos de forma a minimizar (maximizar) o valor de uma variável independente y , mais conhecida como função objetivo $f(\hat{x})$. Em uma otimização com objetivo único apenas uma solução é retornada, pois existe apenas um $f(\hat{x})$ ótimo.

Contudo, muitos problemas de otimização do mundo real lidam com a otimização de objetivos conflitantes, como, por exemplo, a minimização do *Ripple* e a maximização da OSNR no problema ACOP. Neste contexto, o retorno do algoritmo não é uma solução mas sim um conjunto de soluções não-dominadas conhecido como frente de Pareto. Na Figura 5, as soluções A, B, C, D e E representam o conjunto de soluções da frente de Pareto, e fica claro que as soluções F e G não são boas escolhas quando comparadas às soluções na frente de Pareto. Contudo, percebe-se que ao se comparar as soluções entre A e E não é possível definir qual delas é a melhor. Por exemplo, a solução E possui menor *Ripple* se

comparada à solução A, porém a solução A possui maior OSNR. Na terminologia de Pareto (COELLO et al., 2007) as soluções B, C e D dominam a solução F ou $\{B, C, D\} \preceq F$, as soluções entre A e E dominam a solução G ou $\{A, B, C, D, E\} \preceq G$, e as soluções entre A e E são incomparáveis entre si.

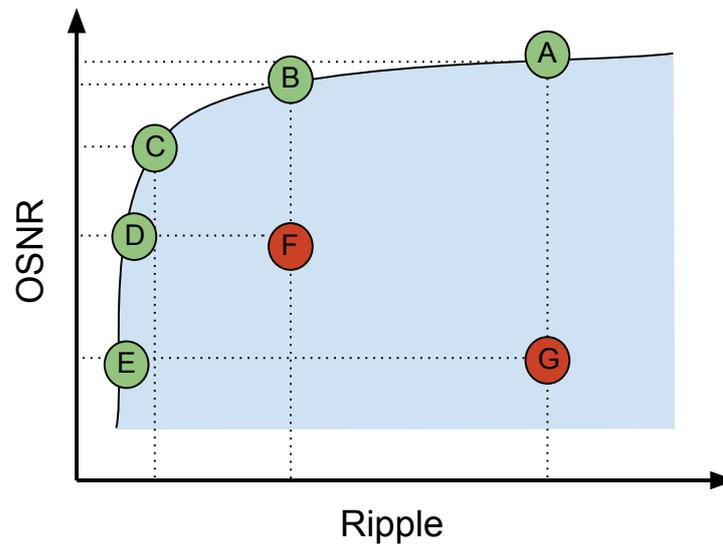


Figura 5 – Um problema com duas funções objetivo: *Ripple* e OSNR. A frente de Pareto está destacada pela linha que une os pontos entre A e E. Fonte: (BARBOZA, 2017)

3.5 Otimização multiobjetivo no problema de ACOP

Em (BARBOZA; BASTOS-FILHO; FILHO, 2019), o problema de ACOP foi estendido para realizar, em conjunto, a seleção do ganho dos amplificadores e perdas dos VOAs de forma a aumentar a QoT do enlace. Assim, foi definido o ACOP como um problema MOO (*Multi-Objective Optimization*) em que o vetor da variável de decisão contém os ganhos do amplificador e as perdas VOA a serem definidas para o enlace. Os objetivos da otimização são maximizar a $OSNR_{ASE+NLI}$ mínima e minimizar sua ondulação. A métrica $OSNR_{ASE+NLI}$ agrega o impacto de interferências lineares (ASE - *Amplified Spontaneous Emission*) e não lineares (NLI - *Non-Linear Interference*) no ruído do sistema.

O algoritmo de otimização MOO usado foi o Non-Dominated Sorting Genetic Algorithm II (NSGA-II) (DEB et al., 2002). O otimizador se comunica com um simulador óptico Java para criar um enlace óptico no qual os ganhos do amplificador e as perdas de VOAs são compostos de valores contidos em um cromossomo. Em seguida, o simulador calcula os valores

para os objetivos a serem otimizados (Figura 6). O trabalho feito em (BARBOZA; BASTOS-FILHO; FILHO, 2019) faz comparações entre o MOO e o método exaustivo - no qual todas as combinações de ganho e perda são geradas. A abordagem MOO retorna soluções viáveis em segundos, enquanto o método exaustivo pode levar horas, até dias, dependendo do tamanho do enlace.

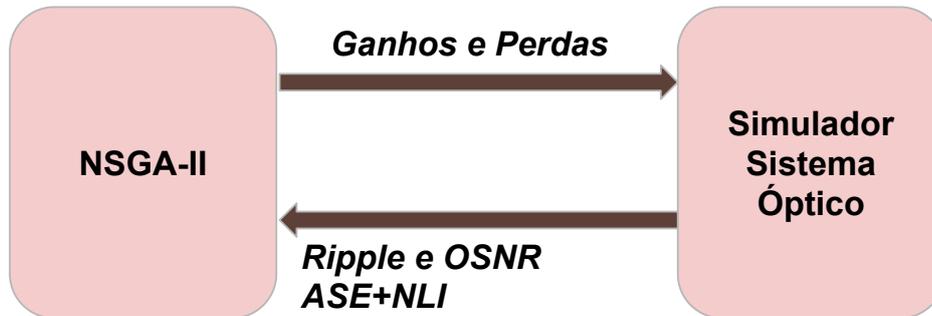


Figura 6 – Ilustração da interação entre o NSGA-II e o simulador do sistema óptico para a realização da otimização. Fonte: adaptado de (BARBOZA, 2017)

A Tabela 1 mostra a organização das variáveis do problema no indivíduo do NSGA-II. O tamanho do indivíduo será o dobro ($2N$) da quantidade de amplificadores da cascata (N). As primeiras N posições irão conter os valores dos ganhos dos amplificadores, e as posições restantes irão conter as perdas dos VOAs de cada amplificador.

Tabela 1 – Organização das variáveis do problema em um indivíduo do NSGA-II. As primeiras N posições contêm os valores dos ganhos dos amplificadores e as posições restantes contêm os valores das perdas dos VOAs.

G_1	G_2	...	G_N	L_1	L_2	...	L_N
-------	-------	-----	-------	-------	-------	-----	-------

4 Metodologia

4.1 Extração de Dados

Como apresentado na Seção 3.3, inicialmente são necessárias bases de dados que representem a resposta das simulações, isto é, execuções do algoritmo MOO, em vários cenários. Então, é possível usá-la no treinamento dos modelos de aprendizado de máquina para que

aprendam os padrões de dados que levam a uma boa solução. A criação desta base de dados é descrita nesta seção.

As simulações foram realizadas considerando enlaces construídos no simulador óptico descrito em (BARBOZA; BASTOS-FILHO; FILHO, 2019). Foi utilizado um amplificador óptico comercial, caracterizado em laboratório com sinais ópticos distribuídos em 40 canais de frequência, variando de 192,1 THz a 196,0 THz em passos de 100 GHz. Portanto, foi considerado um sinal óptico com 40 canais no mesmo espectro do amplificador caracterizado. Cada sinal foi codificado em 40 bits representando cada um dos canais, de forma que o valor 1 indica que o canal está em uso e o valor 0 representa sua inatividade. Foram considerados 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, e 40 canais ativos, ativados da menor à maior frequência, como exemplificado na Figura 7. A potência de entrada do enlace (potência de entrada no primeiro amplificador) é de -18 dBm por canal e a distância entre cada amplificador é de 90 km. Esses parâmetros foram definidos de acordo com (BARBOZA; BASTOS-FILHO; FILHO, 2019).

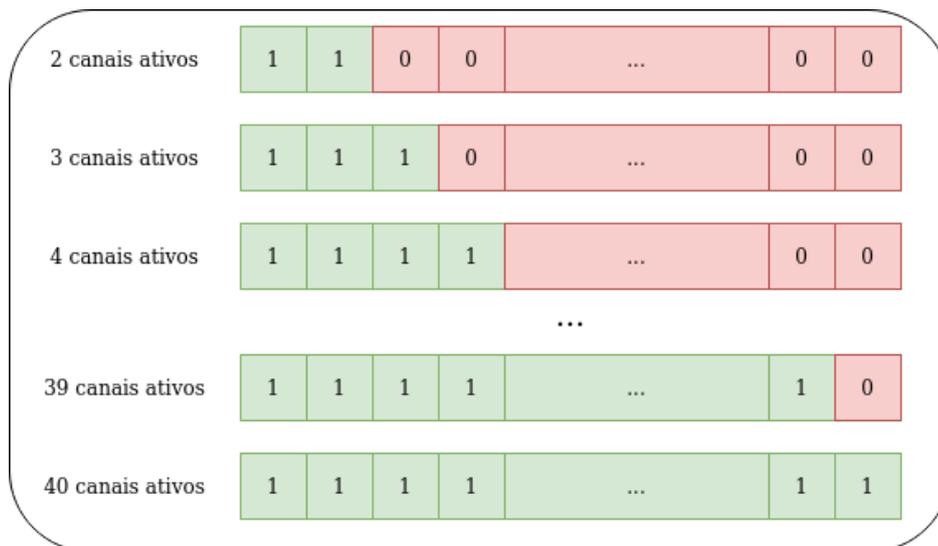


Figura 7 – Ativação dos canais do amplificador óptico caracterizado.

Considerando 2, 3, 4, 5, 6, 7, e 8 amplificadores, no final das simulações, tem-se como resultado sete bases de dados distintas. As entradas dessas bases de dados são compostas de soluções retornadas pelo MOO. Cada execução deste algoritmo retorna uma Frente de Pareto (PF), ou seja, o conjunto de soluções incomparáveis. Porém, neste trabalho, foi considerada a PF composta pelas soluções incomparáveis de dez execuções independentes do algoritmo.

Isso significa que, para cada combinação de amplificadores no enlace e canais ativos no sinal, o algoritmo MOO foi executado dez vezes. Em seguida, considerando as dez PFs gerados, todas as soluções obtidas foram agregadas e submetidas a um filtro de dominância, criando uma PF com as soluções incomparáveis. Essas múltiplas execuções foram realizadas para armazenar na base de dados um conjunto mais diversificado de soluções.

A extração de dados e estrutura da base são exemplificadas na Figura 8. Na **etapa 1**, as simulações são realizadas para enlaces com n amplificadores. Após a execução do MOO, um conjunto de soluções é retornado e, na **etapa 2**, 30% das soluções com os OSNRs mínimos mais altos são selecionadas para serem adicionadas à base de dados. Esta estratégia foi adotada pois considerar todas as soluções do MOO pode gerar uma base de dados com instâncias que são muito diferentes entre si, o que poderia prejudicar o desempenho dos modelos de regressão ao tentar encontrar padrões entre as instâncias. Então, as soluções com melhor OSNR foram privilegiadas. Na **etapa 3**, a configuração de canais simulada - representada por C_i , tal que $i \in \mathbb{Z}$, $i = [1, 40]$ -, bem como os valores de ganho e perda selecionados - G_i e L_i , respectivamente, com $G_i, V_i \in \mathbb{Z}$, $G_i = [14, 24]$, e $V_i = [0, 15]$ - são adicionados como uma nova entrada na base de dados. Uma instância para uma simulação de n amplificadores tem, portanto, 40 variáveis independentes e $2n$ variáveis dependentes.

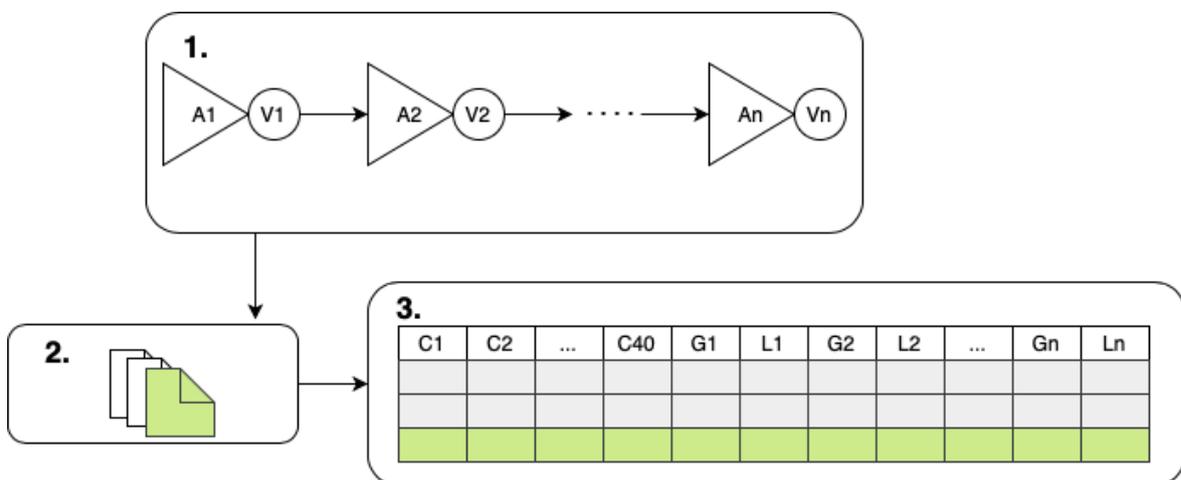


Figura 8 – Ilustração das três etapas de extração de dados: 1) usar MOO para definir ganhos de amplificadores e perdas de VOAs, 2) selecionar 30% das melhores soluções e 3) adicionar soluções selecionadas ao base de dados.

4.2 ACOP como um problema de regressão

Considerando as bases de dados com as soluções não dominadas para o problema ACOP, pode-se modelar o ACOP como um problema de regressão. Uma instância na base de dados para regressão tem 40 variáveis independentes ($C_i, i = [1, 40]$) e $(2n)$ variáveis dependentes. As variáveis independentes serão as entradas usadas para treinar os modelos de regressão e as variáveis dependentes a saída esperada.

Apesar da entrada de cada regressor ser a configuração dos canais codificada em 40 bits, é possível abordar as seguintes estratégias para alocar modelos de regressão para um enlace com n amplificadores, visto que o problema tem várias saídas:

- **Um regressor por variável (R/V):** para cada amplificador, são criados um regressor associado ao amplificador e outro associado ao VOA. Portanto, são instanciados $2n$ regressores. A saída de cada modelo representa um único valor. Um valor para o ganho, se o regressor estiver associado a um amplificador, ou um valor para a perda, se associado a um VOA;
- **Um regressor por amplificador (R/A):** um regressor é criado para cada agregação de amplificador e VOA no enlace. A saída do regressor representa um conjunto com o ganho e a perda para seu amplificador e VOA, respectivamente. Neste caso, n regressores foram instanciados;
- **Um regressor por enlace (R/E):** um único regressor é instanciado para prever o ganho e a perda de todos os amplificadores e VOAs no enlace. Assim, um regressor expressa o comportamento combinado de n valores de ganho e n valores de perda.

Na Figura 9 as três estratégias de alocação são exemplificadas para um enlace com 2 amplificadores.

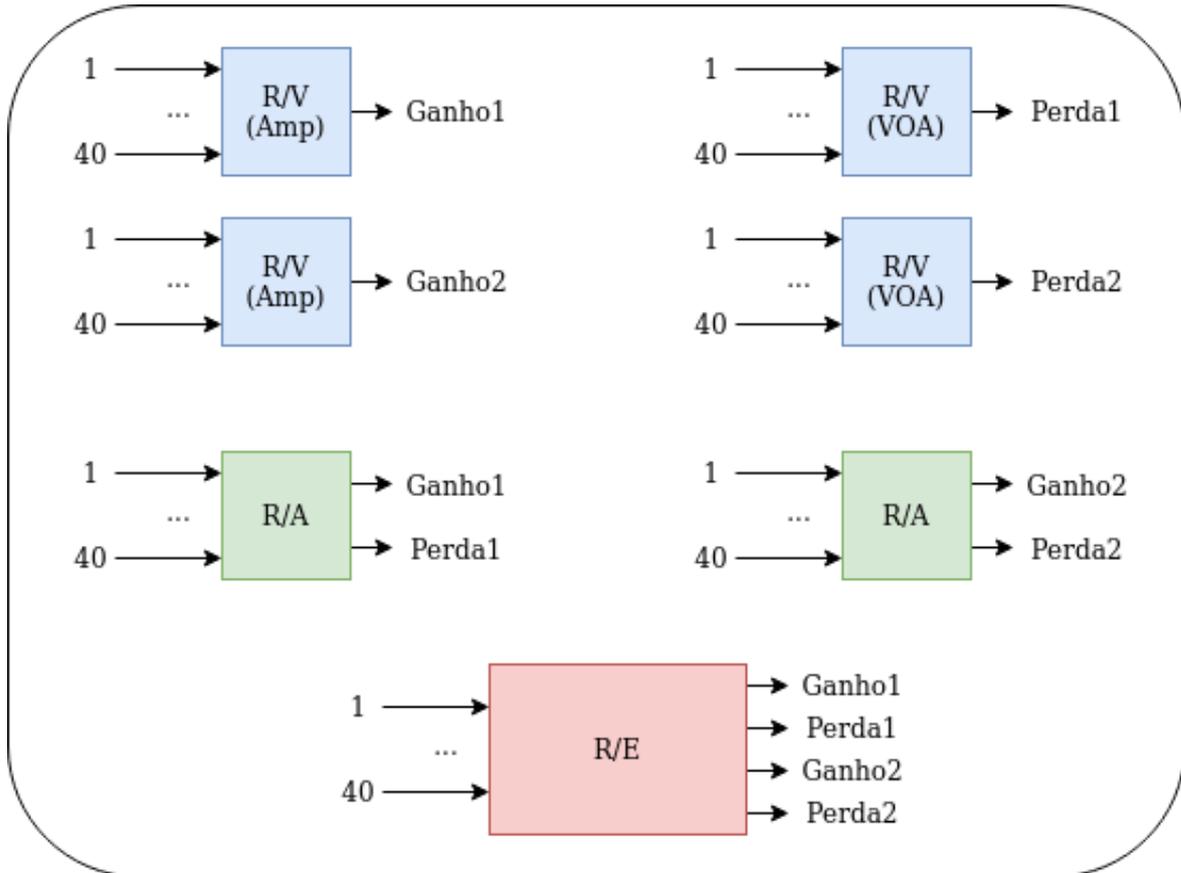


Figura 9 – Ilustração das três estratégias de alocação de regressores para um enlace com 2 amplificadores.

4.3 Definição e treinamento dos modelos

Independente da estratégia de alocação de regressores, estes devem passar pelas mesmas etapas de treinamento e teste. Foi usado um processo de validação cruzada de 5 *folds* para obter o erro médio absoluto (MAE). Essa métrica é usada para definir o desempenho de cada modelo de regressão.

Depois de concluir a extração de dados descrita em 4.1, foram investigadas algumas técnicas de aprendizado de máquina para regressão que podem servir como base para o modelo substituto. Utilizou-se a linguagem Python devido à variedade de bibliotecas de ML disponíveis e ao suporte da comunidade em relação às suas aplicações. Todos os modelos usados nos experimentos vêm da biblioteca Scikit-learn (PEDREGOSA et al., 2011). Os parâmetros de cada modelo foram mantidos como padrão para que seja possível avaliar o desempenho das instâncias mais genéricas dos modelos.

Antes de definir os regressores, é necessário um modelo para servir como referência mínima de desempenho para os outros. Para isso, foi escolhido o *SimpleMeanRegressor* (*SM*), que retorna a média das saídas alvo do base de dados, como exemplificado na tabela 2. Tal modelo, apesar de muito simples, nos indica se o uso de modelos mais complexos, de fato, vale a pena.

Amp 1	VOA 1	Amp 2	VOA 2	Amp 3	VOA 3
16	0	21	3	24	0
16	0	21	2	23	1
18	1	21	2	24	3
17	0	21	3	24	0
17	0	20	1	25	2
SimpleMeanRegressor					
16,8	0,2	20,8	2,2	24	1,2

Tabela 2 – Demonstração do funcionamento do modelo SimpleMeanRegressor

Buscando variedade nas estratégias de regressão, além de *SimpleMeanRegressor*, foram escolhidos outros quatro regressores bem conhecidos baseados em diferentes estratégias de regressão. Os regressores foram escolhidos entre outros após simulações anteriores, pois, apesar de simples, mostraram-se eficientes quanto a adaptação aos dados do problema. Os regressores utilizados neste trabalho e o motivo de sua escolha estão descritos na Seção 3.2.

4.4 Comunicação entre simulador óptico e modelos de regressão

Para comparar o desempenho dos regressores como modelos substitutos do algoritmo MOO, foi necessário conectá-los ao simulador óptico da mesma forma que o algoritmo MOO (conforme descrito na Seção 3.5). Para estabelecer essa conexão, foi utilizado um servidor Flask (GRINBERG, 2018), o qual define operações de criação, treinamento e predição dos modelos Python, permitindo sua comunicação com o simulador Java. Para esse comparativo, foram realizadas simulações com enlaces de 2 a 8 amplificadores, considerando três configurações de canais diferentes: 10, 20 e 40 canais.

O MOO e os regressores mencionados acima foram executados. A comparação foi feita com base na Frente de Pareto retornada pelo MOO e as soluções únicas retornadas pelos modelos substitutos. Para esta comparação o MOO foi executado apenas uma vez, ao contrário das dez execuções demonstradas na Seção 4.1, visando comparar a resposta do MOO

num cenário prático. O algoritmo de otimização foi configurado com os mesmos parâmetros definidos em (BARBOZA; BASTOS-FILHO; FILHO, 2019).

5 Resultados e Discussão

Após a execução das simulações para criação das bases de dados, foram obtidas 7 bases distintas, cujos números de instâncias são descritos na Tabela 3. Como apontado em (LIMA et al., 2020), o aumento de amplificadores no enlace aumenta o espaço de busca do MOO, dificultando o retorno de boas soluções para o problema.

	Número de amplificadores no enlace						
	2	3	4	5	6	7	8
Instâncias na base dados	1150	3229	4079	4041	2182	982	205

Tabela 3 – Instâncias retornadas pelo MOO para enlaces de 2, 3, 4, 5, 6, 7 e 8 amplificadores

Com as bases de dados disponíveis, foram comparados os desempenhos de cada regressor considerando as três estratégias de alocação visando definir qual delas é a mais eficiente. Para cada estratégia, o erro de teste dos regressores é agregado de forma a que, no final das simulações, todos os erros calculados sejam utilizados na construção dos gráficos. Num enlace com dois amplificadores e um regressor alocado para cada variável, por exemplo, o erro de ganho dos regressores associados aos amplificadores 1 e 2 são agregados para confecção dos gráficos de ganho, enquanto que o erro de perda dos regressores associados aos VOAs 1 e 2 são utilizados na confecção dos gráficos de perda.

Na Figura 10 é apresentado o desempenho de teste dos regressores em cenários com um regressor por variável, um regressor por amplificador e um regressor por enlace, para enlaces de 2, 5 e 8 amplificadores. Estes enlaces foram escolhidos devido por serem o menor, médio e maior enlace, considerando todos os cenários analisados. Os demais gráficos, para enlaces de 3, 4, 6 e 7 amplificadores estão disponíveis no Apêndice A.

Independente da estratégia de alocação adotada, para o mesmo enlace, não há diferença notável entre os erros medianos dos regressores, como exemplificado na Tabela 4, que mostra a mediana do MAE *SimpleMeanRegressor* para cada enlace analisado. É notável, inclusive, que o erro para as estratégias de um regressor por amplificador e um regressor por enlace é o mesmo. Tal comportamento indica que o desempenho dos modelos independe da estratégia de alocação, portanto é mais eficiente utilizar um único regressor, dado que o tempo de treinamento e processamento é minimizado.

Tendo em vista que a estratégia de um regressor por enlace é mais eficiente, na Figura 11

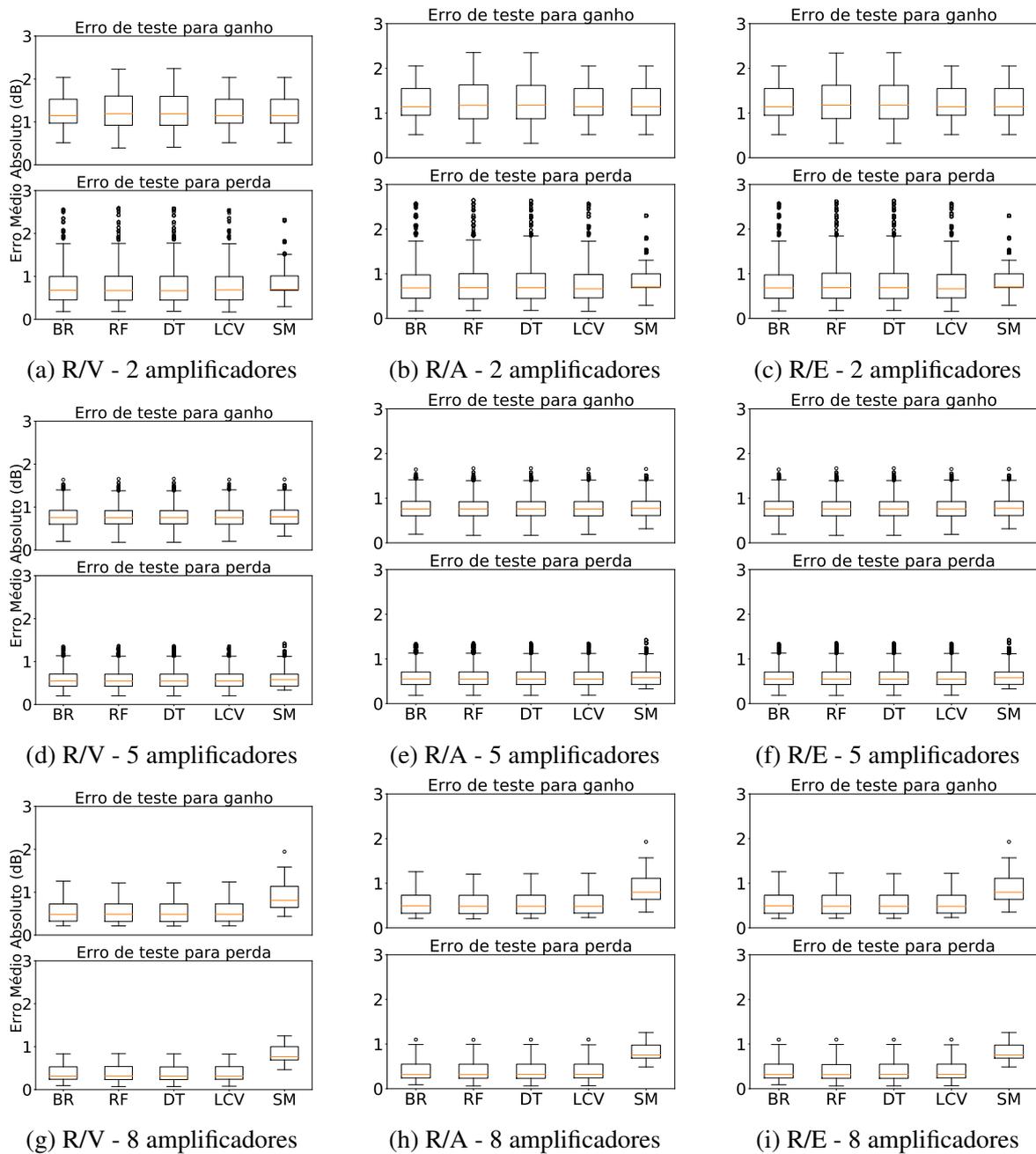


Figura 10 – Distribuição do erro médio absoluto (MAE) da fase de teste para ganho do amplificador e estimativa de perda do VOA considerando (a, d, g) um regressor por variável (**R/V**), (b, e, h) um regressor por amplificador (**R/A**) e (c, f, i) um regressor por enlace (**R/E**) para 2, 5 e 8 amplificadores. Os cinco regressores considerados são: Bayesian Ridge Regressor (BR), Random Forest Regressor (RF), Decision Tree Regressor (DT), Lasso CV (LCV) e SimpleMeanRegressor (SM).

Amplificadores no enlace	Erro Médio Absoluto (dB)					
	R / V		R / A		R / E	
	Ganho	Perda	Ganho	Perda	Ganho	Perda
2	1,147	0,697	1,140	0,706	1,140	0,706
5	0,774	0,578	0,771	0,577	0,771	0,577
8	0,811	0,767	0,799	0,755	0,799	0,755

Tabela 4 – Mediana do erro médio absoluto (MAE) da validação cruzada de 5 *folds* para a estratégia *SimpleMeanRegressor*.

foram comparados os resultados para todos os cenários considerados, isto é, 2, 3, 4, 5, 6, 7 e 8 amplificadores, considerando apenas um regressor por enlace.

É aparente que em todos os cenários - com exceção do enlace com dois amplificadores -, tem-se que a mediana do erro médio absoluto é menor que 0,9 dB para ganho e menor que 0,7 dB para perda. Além disso, percebe-se que o erro tende a diminuir ao se aumentar o número de amplificadores no enlace. Para o *SimpleMeanRegressor*, por exemplo, a mediana do erro de estimativa do ganho é 1,14 dB para dois amplificadores e chega a 0,79 dB para oito amplificadores. Este comportamento pode ser reflexo da baixa variação dos valores de ganho e perda retornados pelo MOO para enlaces com muitos amplificadores, como demonstrado na Figura 12.

Embora o erro seja pequeno, os resultados indicam que nenhuma técnica de ML pôde simular perfeitamente o MOO considerando a representação de conhecimento adotada. Para avaliar o impacto dessa imprecisão em um sistema de comunicação óptica, foi simulada a execução do ACOP MOO e dos demais modelos para enlaces de 2 a 8 amplificadores.

A Figura 13 representa as Frentes de Pareto e as previsões dos modelos ML para configurações com 10, 20 e 40 canais para enlaces de 2, 5 e 8 amplificadores. A OSNR é calculada usando o modelo GNLI (POGGIOLINI, 2012), que considera ruído linear e não linear. Ripple OSNR é a diferença entre o OSNR máximo e o OSNR mínimo entre todos os canais.

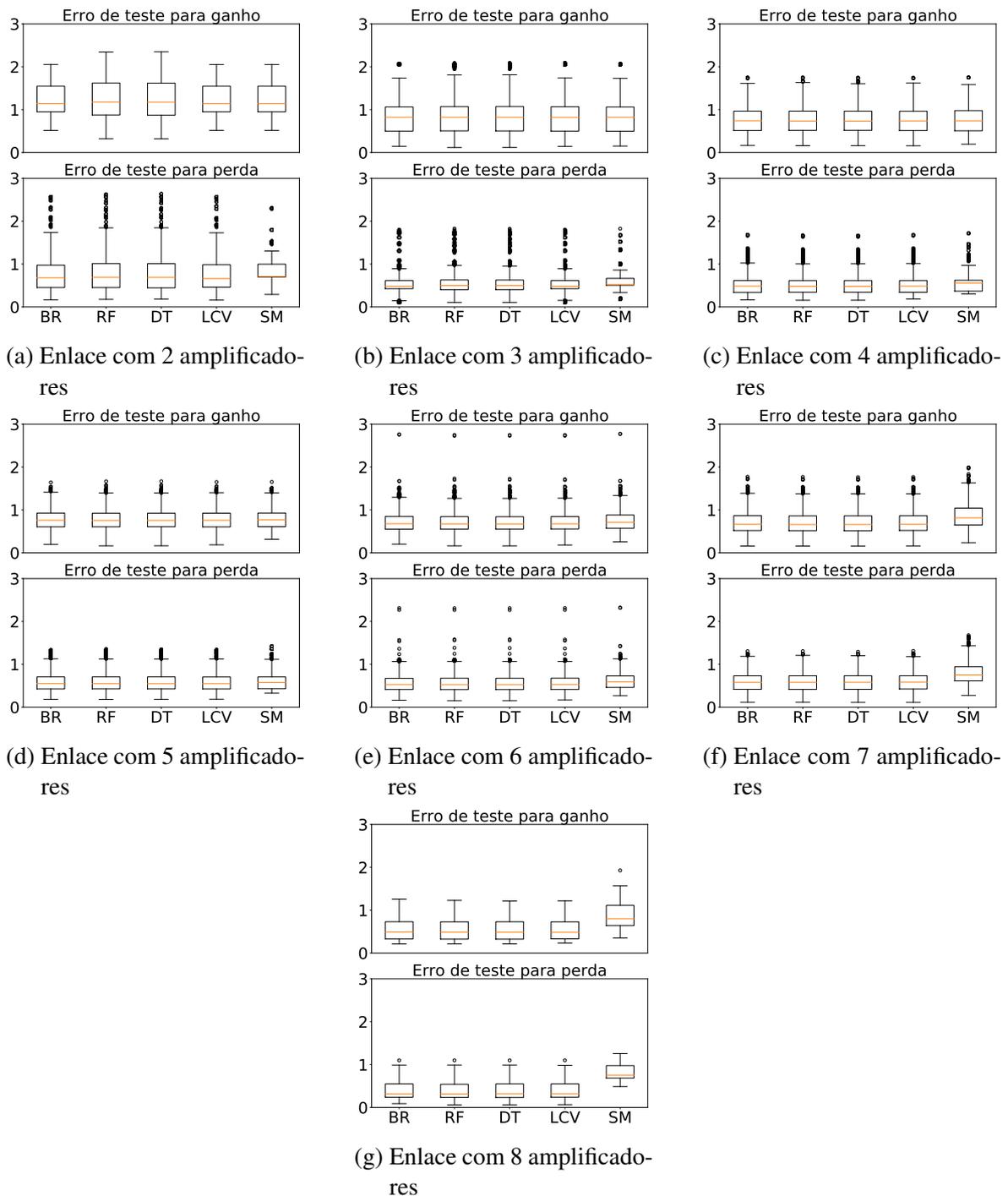


Figura 11 – Distribuição do erro médio absoluto (MAE) da fase de teste para ganho do amplificador e estimativa de perda do VOA considerando 2 a 8 amplificadores. Os cinco regressores considerados são: Bayesian Ridge Regressor (BR), Random Forest Regressor (RF), Decision Tree Regressor (DT), Lasso CV (LCV) e SimpleMeanRegressor (SM).

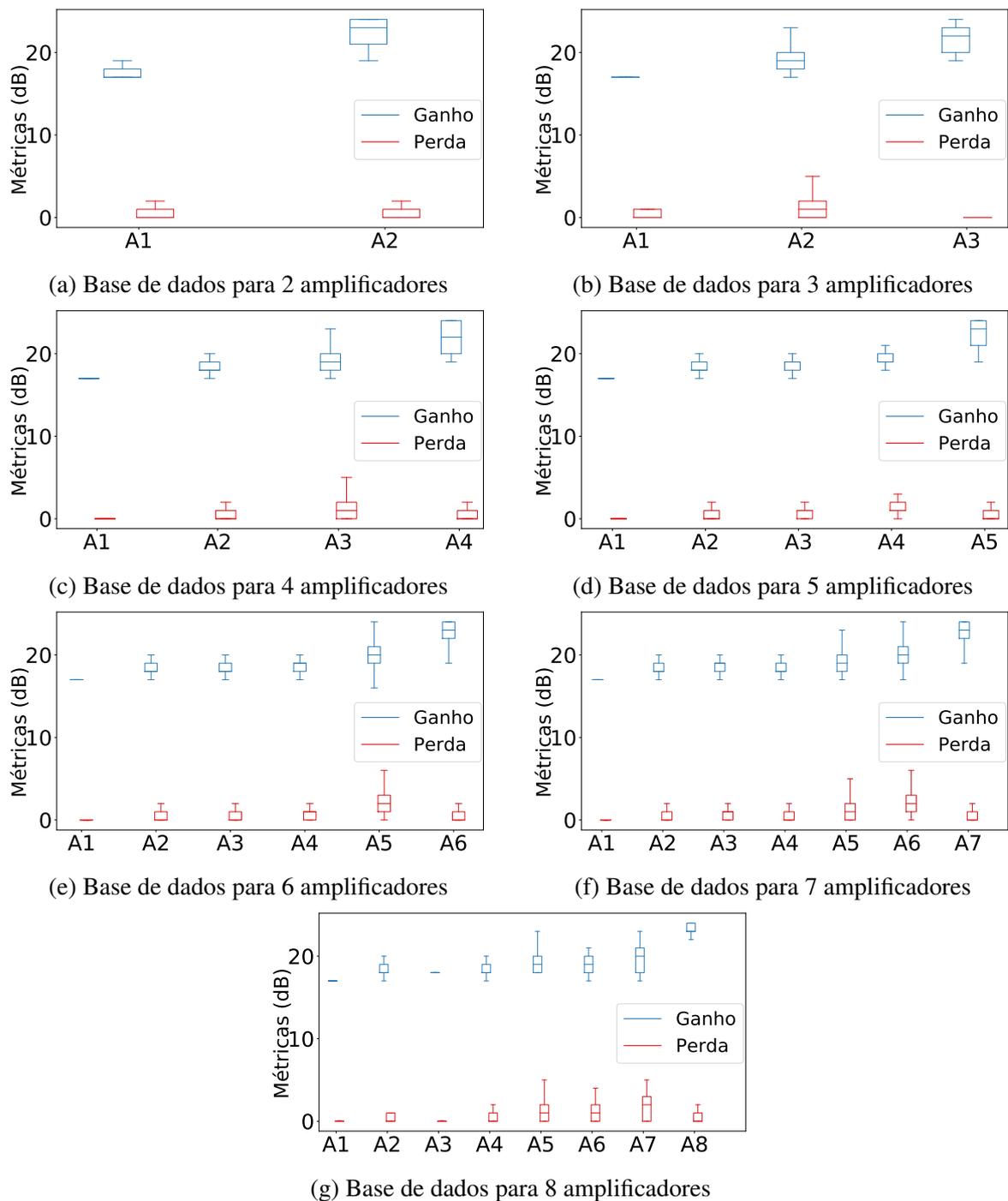
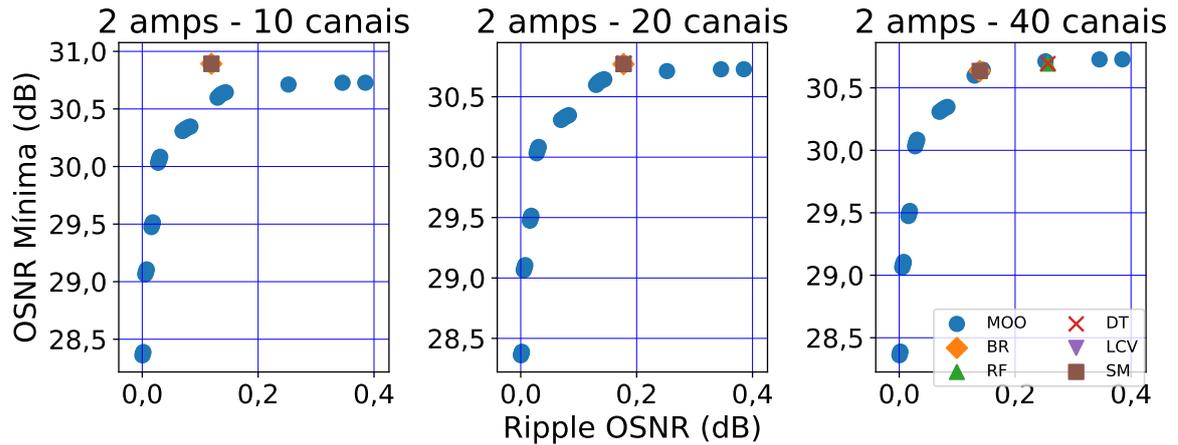
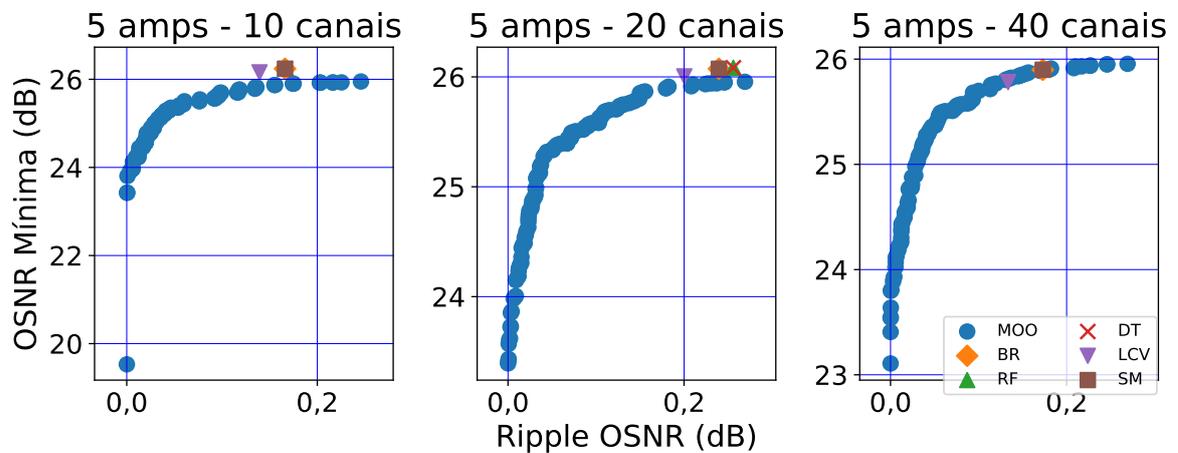


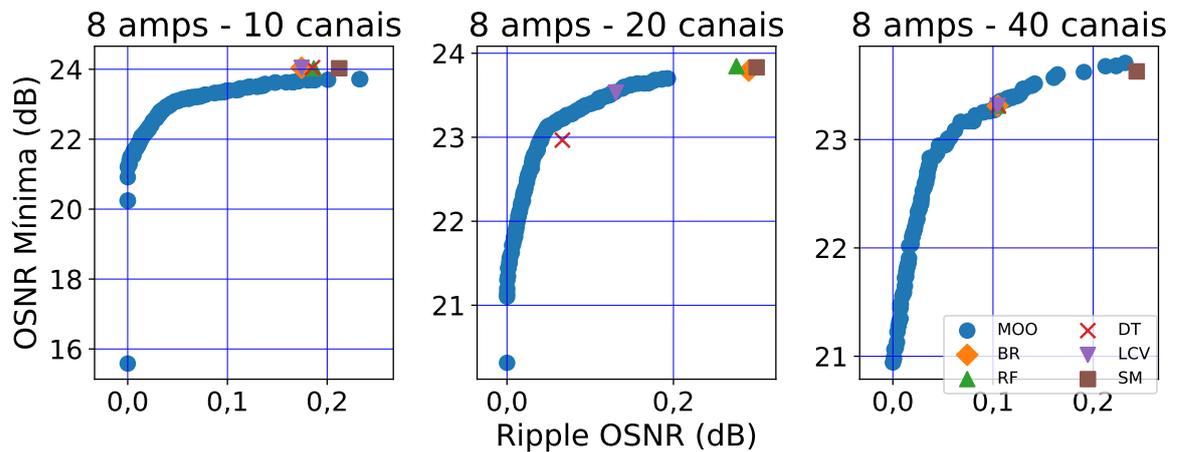
Figura 12 – Distribuição dos valores de ganho do amplificador e perda do VOA para cada base de dados obtida através da execução do MOO, considerando diferentes números de amplificadores. A_i representa o amplificador i .



(a) Frentes de Pareto para enlace de 2 amplificadores



(b) Frentes de Pareto para enlace de 5 amplificadores



(c) Frentes de Pareto para enlace de 8 amplificadores

Figura 13 – Ripple OSNR e OSNR mínima retornadas pelo algoritmo MOO proposto em (BARBOZA; BASTOS-FILHO; FILHO, 2019) e pelas técnicas de regressão utilizadas neste trabalho considerando 10, 20 e 40 canais. Os cinco regressores considerados são: Bayesian Ridge Regressor (BR), Random Forest Regressor (RF), Decision Tree Regressor (DT), Lasso CV (LCV) e SimpleMeanRegressor (SM). Apenas um regressor é utilizado para o enlace inteiro.

Percebe-se que todos os modelos obtiveram resultados iguais ou superiores ao MOO, indicando que o erro obtido na etapa anterior não impacta significativamente na solução do problema. A superioridade dos modelos substitutos em alguns casos pode ser explicada pelo fato da base de dados utilizada para o treinamento ser gerada a partir de dez execuções do MOO, permitindo a cobertura de uma parte mais representativa das saídas do MOO. Outra possível explicação para tal comportamento é que os modelos são treinados para resolver cenários com todas as configurações de canais, e não apenas 10, 20 e 40. Assim, os regressores tendem a retornar uma solução semelhante ao caso mais comum retornado pela otimização MOO.

A Tabela 5 mostra o tempo médio para cada uma das técnicas após dez execuções independentes de treinamento. É notável que o MOO requer alguns segundos para retornar soluções, enquanto outras técnicas de ML retornam resultados em milissegundos. Dentre as técnicas de ML, a que obteve os piores tempos foi a de *RandomForest*, pois, como descrito em 3.2.3, esta técnica funciona a partir de agregação de predições de diversas árvores diferentes, portanto é natural que seja uma abordagem mais custosa. Por fim percebe-se que no cenário de 8 amplificadores, por exemplo, o MOO leva cerca de 43 segundos para finalizar sua execução, enquanto que o *SimpleMeanRegressor* é finalizado em apenas 0,823 milissegundos, sendo 28.000 vezes mais rápido que o MOO.

Técnica	Tempo médio de execução (ms)						
	Amplificadores no enlace						
	2	3	4	5	6	7	8
MOO	5290	9613	12652	13108	13794	27315	42927
BayesianRidge	0,798	1,272	1,303	1,573	1,689	1,885	2,417
RandomForestRegressor	20,40	35,57	41,98	51,53	59,54	77,66	78,35
DecisionTreeRegressor	0,905	1,328	1,587	1,748	1,982	2,582	2,667
LassoCV	0,804	1,167	1,480	1,4999	1,603	2,265	2,299
SimpleMeanRegressor	0,594	0,659	0,796	0,823	0,937	1,343	1,521

Tabela 5 – Tempo médio de execução para cada a técnica após dez execuções independentes.

6 Conclusão

Neste trabalho, estudou-se a criação de um modelo substituto para o problema ACOP capaz de substituir a abordagem anterior baseada em otimização multiobjetivo que, apesar de retornar as melhores soluções relatadas na literatura, não possui um tempo de execução desejável para aplicações de tempo real. Para isso, foram utilizados diferentes modelos baseados em técnicas de aprendizado de máquina e diferentes estratégias de alocação que impactam na quantidade de regressores utilizados.

Para a análise dos modelos regressivos foram considerados enlaces ópticos com 2 a 8 amplificadores, além de sinais ópticos com diferentes números de canais. Notamos que usar apenas um regressor para o enlace completo é suficiente para obter soluções razoáveis com menor custo computacional para sua fase de treinamento. Notou-se também que não há diferença significativa na precisão entre os modelos de regressão escolhidos, indicando que o modelo mais simples, ou seja, *SimpleMeanRegressor* - o qual calcula o valor médio do ganho dos amplificadores e das perdas dos VOA em toda a base de dados - pode ser a melhor escolha para resolver o problema. Além disso, usando este modelo como um modelo substituto, foi possível reduzir - em comparação com a abordagem de otimização multiobjetivo proposta anteriormente - o tempo de execução no mínimo 8900 vezes e no máximo 28200 vezes, considerando, respectivamente, 2 e 8 amplificadores por enlace.

Não consta neste trabalho o impacto da otimização dos parâmetros dos regressores. Como especificado anteriormente, todos os modelos utilizados são definidos com os parâmetros padrão estabelecidos pela biblioteca Scikit-learn. A otimização desses parâmetros pode tornar os modelos ainda mais eficientes, seja em sua precisão, ou no tempo de execução. Em trabalhos futuros, pretende-se abordar a otimização dos regressores e se, de fato, esta alteração é refletida nos resultados finais.

Foi utilizado um amplificador óptico com sinais ópticos distribuídos em 40 canais de frequência, variando de 192,1 THz a 196,0 THz, ou seja da menor à maior frequência. Não foi considerada neste trabalho a influência da estratégia de inicialização desses canais na qualidade da base de dados. Portanto, em próximos estudos pretende-se levar em conta outras estratégias de inicialização de canais, como da maior à menor frequência, ou inicialização aleatória.

A base de dados utilizada neste trabalho consiste de 30% das soluções com os OSNRs mínimos mais altos retornadas pelo algoritmo multiobjetivo. Deseja-se em trabalhos futuros comparar a representação de informação entre essa e outras estratégias, como bases de dados com todas as soluções retornadas pelo MOO.

O trabalho aqui apresentado é um dos primeiros que abordam a utilização de modelos substitutos para solução do problema de ACOP. Verificamos que é possível utilizar esta abordagem e obter resultados satisfatórios. Em estudos futuros, pretende-se estender esse trabalho para analisar enlaces com diferentes parâmetros. Desta forma, poderá ser verificada a viabilidade e qualidade da solução em redes reais.

Referências

- ARAÚJO, D. R. D.; BASTOS-FILHO, C. J.; MARTINS-FILHO, J. F. An evolutionary approach with surrogate models and network science concepts to design optical networks. *Engineering Applications of Artificial Intelligence*, Elsevier, v. 43, p. 67–80, 2015.
- BANEY, D. M.; GALLION, P.; TUCKER, R. S. Theory and measurement techniques for the noise figure of optical amplifiers. *Optical fiber technology*, Elsevier, v. 6, n. 2, p. 122–154, 2000.
- BARBOZA, E. d. A. Amplificadores ópticos autônomos: desenvolvimento e análise de técnicas. Universidade Federal de Pernambuco, 2017.
- BARBOZA, E. d. A. et al. Local and global approaches for the adaptive control of a cascade of amplifiers. *Photonic Network Communications*, Springer, v. 33, n. 2, p. 194–207, 2017.
- BARBOZA, E. de A.; BASTOS-FILHO, C. J.; FILHO, J. F. M. Adaptive control of optical amplifier operating point using voa and multi-objective optimization. *Journal of Lightwave Technology*, IEEE, v. 37, n. 16, p. 3994–4000, 2019.
- BASTOS-FILHO, C. J. et al. Investigating the creation of a surrogate model for adaptive control of amplifier operating point using machine learning. In: IEEE. *2020 22nd International Conference on Transparent Optical Networks (ICTON)*. [S.l.], 2020. p. 1–4.
- BOX, G. E.; TIAO, G. C. *Bayesian inference in statistical analysis*. [S.l.]: John Wiley & Sons, 2011. v. 40.
- CHATTERJEE, S.; HADI, A. S. *Regression analysis by example*. [S.l.]: John Wiley & Sons, 2015.
- COELLO, C. A. C. et al. *Evolutionary algorithms for solving multi-objective problems*. [S.l.]: Springer, 2007. v. 5.
- DEB, K. et al. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, IEEE, v. 6, n. 2, p. 182–197, 2002.
- DIETTERICH, T. Overfitting and undercomputing in machine learning. *ACM computing surveys (CSUR)*, ACM New York, NY, USA, v. 27, n. 3, p. 326–327, 1995.
- DONG, J.; QIN, W.; WANG, M. Fast multi-objective optimization of multi-parameter antenna structures based on improved bpnn surrogate model. *IEEE Access*, IEEE, v. 7, p. 77692–77701, 2019.
- European Technology Platform Photonics21. *Europe’s age of light! How photonics will power growth and innovation: Strategic Roadmap 2021–2027*. 2021. <https://mcusercontent.com/2bd4f3c0e83d0ffc9c3c06a75/files/b61b8d95-76f6-42d3-e826-154c2c3fbb2e/Europes_Age_of_Light_Photonics_Roadmap21.pdf>. Acessado: 2021-08-19.

FERNÁNDEZ, N. et al. Virtual topology reconfiguration in optical networks by means of cognition: Evaluation and experimental validation. *Journal of Optical Communications and Networking*, Optical Society of America, v. 7, n. 1, p. A162–A173, 2015.

FERNÁNDEZ, N. et al. Virtual topology design and reconfiguration using cognition: Performance evaluation in case of failure. In: IEEE. *2013 5th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. [S.l.], 2013. p. 132–139.

FORECAST, G. et al. Cisco visual networking index: global mobile data traffic forecast update, 2017–2022. *Update*, v. 2017, p. 2022, 2019.

GOPALAKRISHNAN, T.; CHOUDHARY, R.; PRASAD, S. Prediction of sales value in online shopping using linear regression. In: IEEE. *2018 4th International Conference on Computing Communication and Automation (ICCCA)*. [S.l.], 2018. p. 1–6.

GORISSEN, D. et al. A surrogate modeling and adaptive sampling toolbox for computer based design. *The Journal of Machine Learning Research*, JMLR. org, v. 11, p. 2051–2055, 2010.

GRINBERG, M. *Flask web development: developing web applications with python*. [S.l.]: "O'Reilly Media, Inc.", 2018.

LIMA, F. C. N. O. et al. Maximizing the transmission rate in optical systems using swarm intelligence. In: *2020 IEEE Latin-American Conference on Communications (LATINCOM)*. [S.l.: s.n.], 2020. p. 1–6.

LOH, W.-Y. Classification and regression trees. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, Wiley Online Library, v. 1, n. 1, p. 14–23, 2011.

MCDONALD, G. C. Ridge regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, Wiley Online Library, v. 1, n. 1, p. 93–100, 2009.

MOURA, U. et al. Cognitive methodology for optical amplifier gain adjustment in dynamic dwdm networks. *Journal of Lightwave Technology*, IEEE, v. 34, n. 8, p. 1971–1979, 2016.

NISHITANI, T. et al. Protection systems for optical access networks. *Journal of Lightwave Technology*, IEEE, v. 35, n. 6, p. 1197–1203, 2017.

OLIVEIRA, J. et al. Demonstration of edfa cognitive gain control via gmpls for mixed modulation formats in heterogeneous optical networks. In: OPTICAL SOCIETY OF AMERICA. *Optical Fiber Communication Conference*. [S.l.], 2013. p. OW1H–2.

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

POGGIOLINI, P. The gn model of non-linear propagation in uncompensated coherent optical systems. *Journal of Lightwave Technology*, IEEE, v. 30, n. 24, p. 3857–3879, 2012.

RAMASWAMI, R.; SIVARAJAN, K.; SASAKI, G. *Optical networks: a practical perspective*. [S.l.]: Morgan Kaufmann, 2009.

ROOPA, H.; ASHA, T. A linear model based on principal component analysis for disease prediction. *IEEE Access*, IEEE, v. 7, p. 105314–105318, 2019.

ROTTONDI, C. et al. Machine-learning method for quality of transmission prediction of unestablished lightpaths. *Journal of Optical Communications and Networking*, Optical Society of America, v. 10, n. 2, p. A286–A297, 2018.

SONG, J. et al. Adaptive surrogate model based multiobjective optimization for coastal aquifer management. *Journal of hydrology*, Elsevier, v. 561, p. 98–111, 2018.

TANENBAUM, A. S.; WETHERALL, D. J. *Computer networks*. 5th. ed. Boston: Pearson Education, 2011. 99-105 p.

VELA, A. P. et al. Ber degradation detection and failure identification in elastic optical networks. *Journal of Lightwave Technology*, IEEE, v. 35, n. 21, p. 4595–4604, 2017.

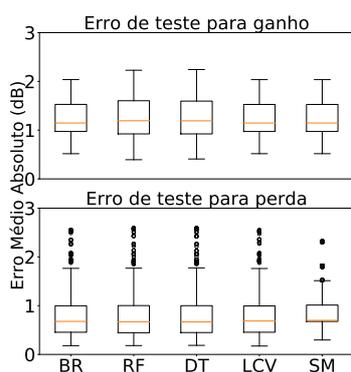
VIANA, F. A. C. et al. Surrogate modeling techniques and heuristic optimization methods applied to design and identification problems. Universidade Federal de Uberlândia, 2008.

WESTERMANN, P.; WELZEL, M.; EVINS, R. Using a deep temporal convolutional network as a building energy surrogate model that spans multiple climate zones. *Applied Energy*, Elsevier, v. 278, p. 115563, 2020.

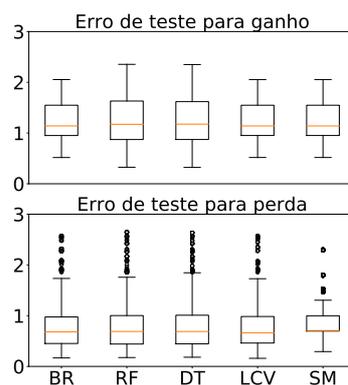
XING, Z. et al. Identifying the release history of a groundwater contaminant source based on an ensemble surrogate model. *Journal of Hydrology*, Elsevier, v. 572, p. 501–516, 2019.

APÊNDICE A – Erro médio absoluto para fase de teste em enlaces de 2 a 8 amplificadores

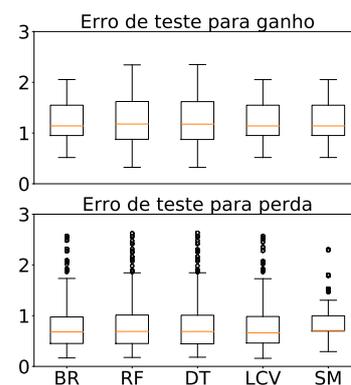
As seguintes figuras demonstram a distribuição do erro médio absoluto (MAE) da fase de teste para ganho do amplificador e estimativa de perda do VOA considerando (a) um regressor por variável (**R/V**), (b) um regressor por amplificador (**R/N**) e (c) um regressor por enlace (**R/E**) para enlaces de 2 a 8 amplificadores. Os cinco regressores considerados são: Bayesian Ridge Regressor (BR), Random Forest Regressor (RF), Decision Tree Regressor (DT), Lasso CV (LCV) e SimpleMeanRegressor (SM).



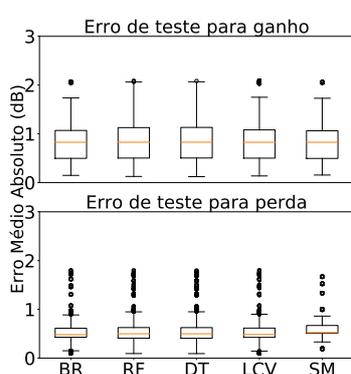
(a) R/V - 2 amplificadores



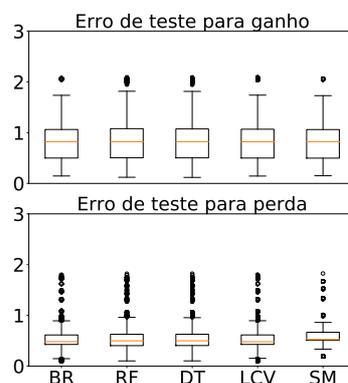
(b) R/N - 2 amplificadores



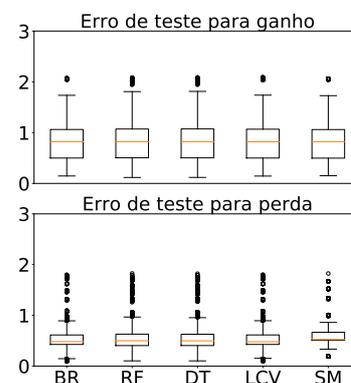
(c) R/E - 2 amplificadores



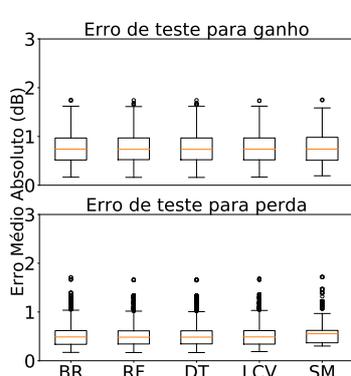
(a) R/V - 3 amplificadores



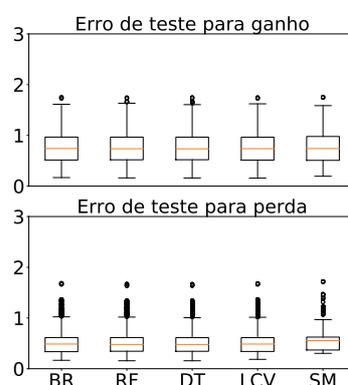
(b) R/N - 3 amplificadores



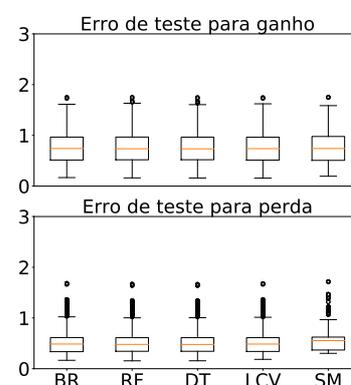
(c) R/E - 3 amplificadores



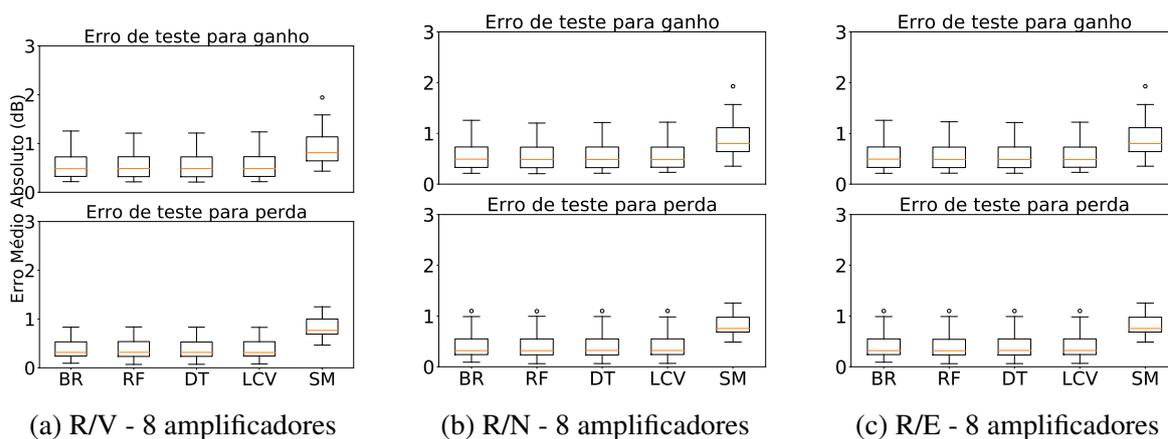
(a) R/V - 4 amplificadores



(b) R/N - 4 amplificadores



(c) R/E - 4 amplificadores



APÊNDICE B – Desempenho dos regressores em comparação ao MOO em enlaces de 2 a 8 amplificadores

Ripple OSNR e OSNR mínima retornadas pelo algoritmo MOO proposto em (BARBOZA; BASTOS-FILHO; FILHO, 2019) e pelas técnicas de regressão utilizadas neste trabalho considerando 10, 20 e 40 canais. Os cinco regressores considerados são: Bayesian Ridge Regressor (BR), Random Forest Regressor (RF), Decision Tree Regressor (DT), Lasso CV (LCV) e SimpleMeanRegressor (SM). Apenas um regressor é utilizado para o enlace inteiro.

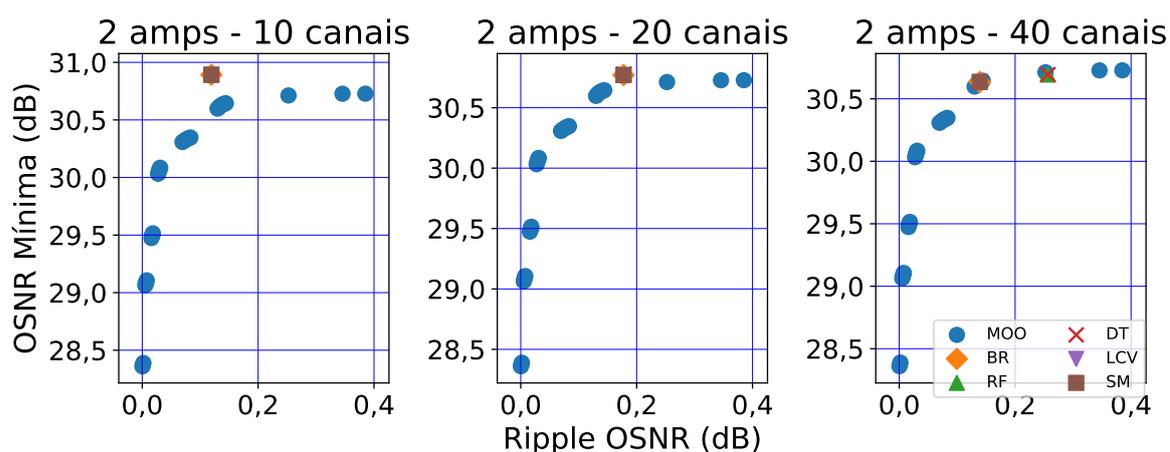


Figura 21 – Frentes de Pareto para enlace de 2 amplificadores

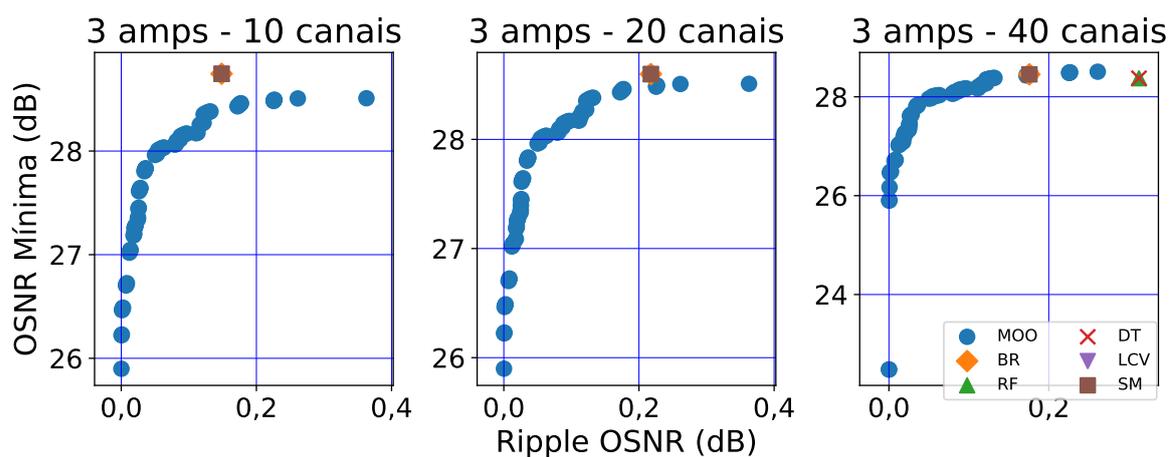


Figura 22 – Frentes de Pareto para enlace de 3 amplificadores

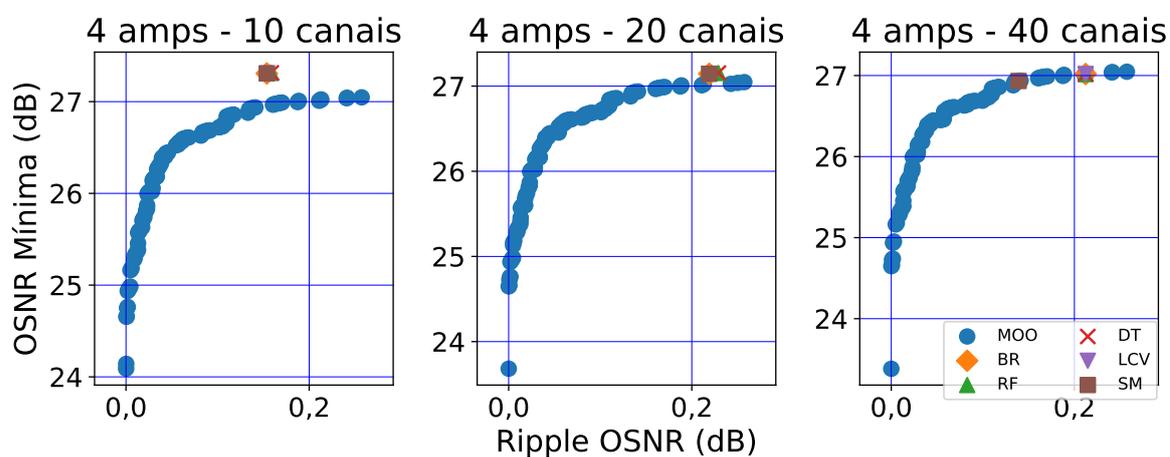


Figura 23 – Frentes de Pareto para enlace de 4 amplificadores

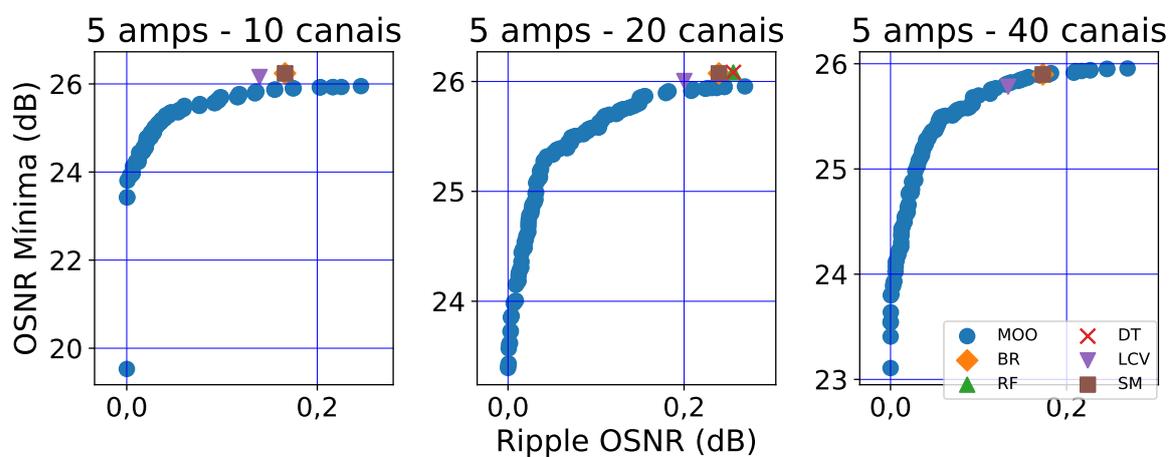


Figura 24 – Frentes de Pareto para enlace de 5 amplificadores

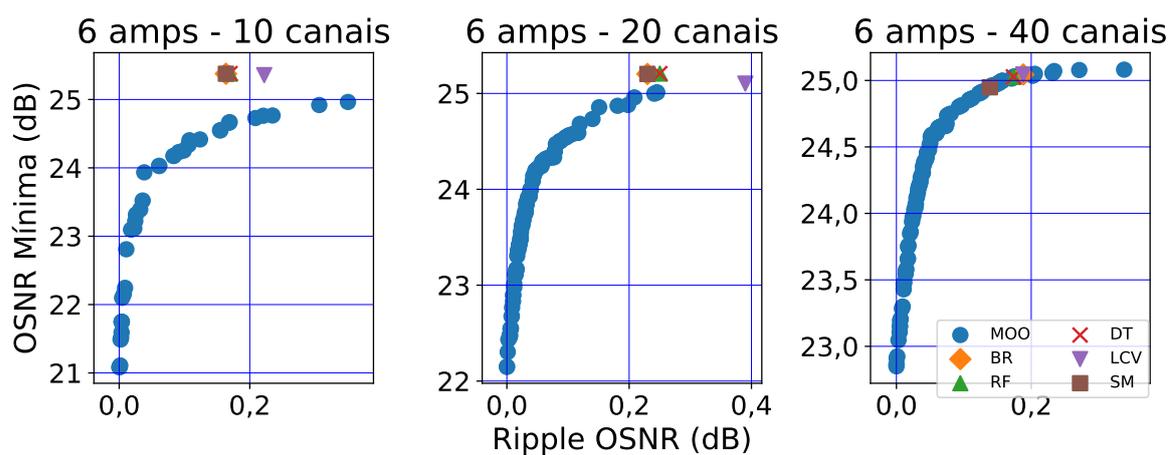


Figura 25 – Frentes de Pareto para enlace de 6 amplificadores

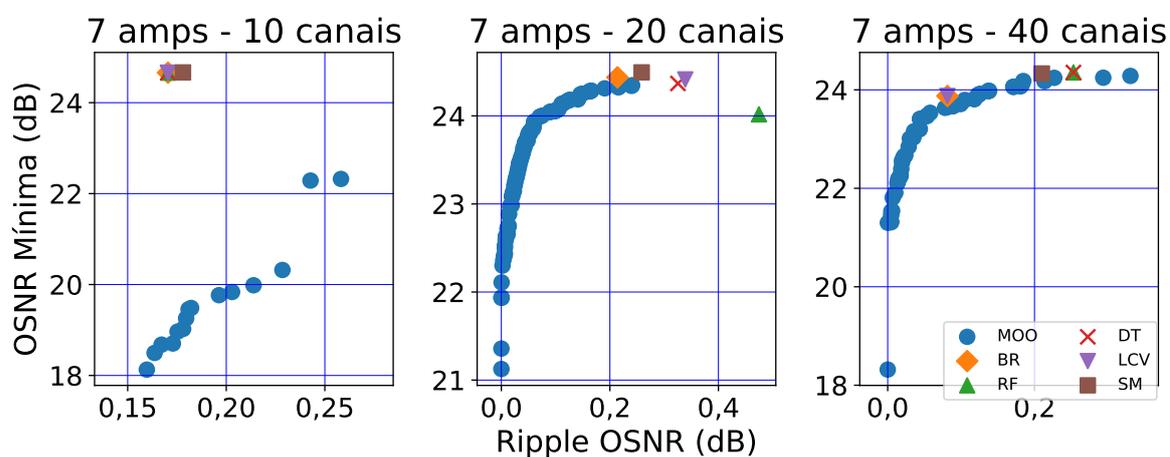


Figura 26 – Frentes de Pareto para enlace de 7 amplificadores

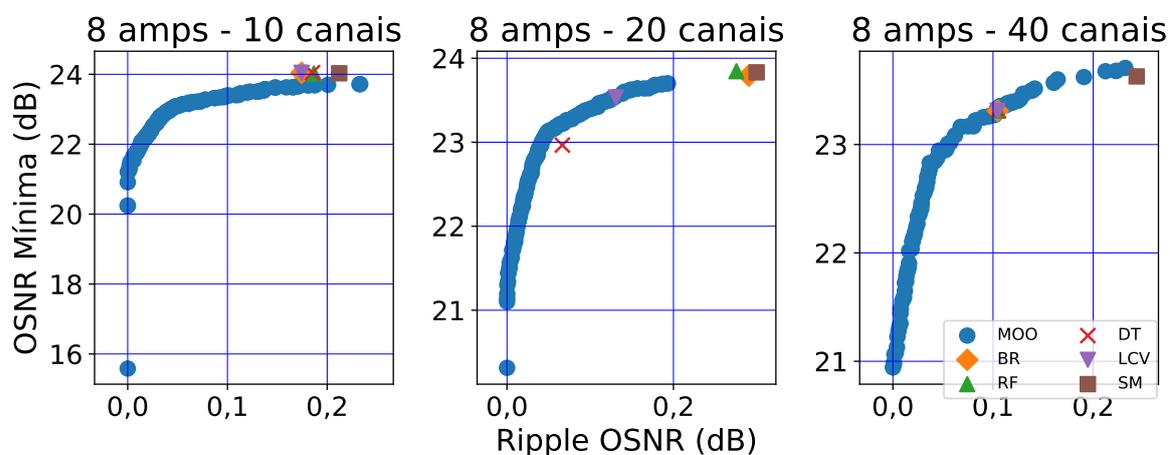


Figura 27 – Frentes de Pareto para enlace de 8 amplificadores