

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
CAMPUS DE SÃO JOÃO DA BOA VISTA

ANA JÚLIA NORA FRANCISCO

**Modelo inverso de amplificador à fibra dopada com érbio para o controle
de potências de saída no regime de alta modulação de ganho cruzado**

São João da Boa Vista

2024

Ana Júlia Nora Francisco

**Modelo inverso de amplificador à fibra dopada com érbio para o controle
de potências de saída no regime de alta modulação de ganho cruzado**

Trabalho de Graduação apresentado ao Conselho de Curso de Graduação em Engenharia Eletrônica e de Telecomunicações do Campus de São João da Boa Vista, Universidade Estadual Paulista, como parte dos requisitos para obtenção do diploma de Graduação em Engenharia Eletrônica e de Telecomunicações .

Orientador: Profº Dr. Ivan Aritz Aldaya Garde

São João da Boa Vista

2024

F819m	Francisco, Ana Júlia Nora Modelo inverso de amplificador à fibra dopada com érbio para o controle de potências de saída no regime de alta modulação de ganho cruzado / Ana Júlia Nora Francisco. -- São João da Boa Vista, 2024 43 p. Trabalho de conclusão de curso (Bacharelado - Engenharia de Telecomunicações) - Universidade Estadual Paulista (UNESP), Faculdade de Engenharia, São João da Boa Vista Orientador: Ivan Aritz Aldaya Garde 1. Inteligência artificial. 2. Fibras ópticas. 3. Amplificadores óticos. 4. Érbio. I. Título.
-------	--

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), Faculdade de Engenharia, São João da Boa Vista. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE ENGENHARIA - CÂMPUS DE SÃO JOÃO DA BOA VISTA
GRADUAÇÃO EM ENGENHARIA ELETRÔNICA E DE TELECOMUNICAÇÕES**

TRABALHO DE CONCLUSÃO DE CURSO

**MODELO INVERSO DE AMPLIFICADOR À FIBRA DOPADA COM ÉRBIO PARA
A CONTROLE DE POTÊNCIAS DE SAÍDA NO REGIME DE ALTA MODULAÇÃO
DE GANHO CRUZADO**

Aluna: Ana Júlia Nora Francisco

Orientador: Prof. Dr. Ivan Aritz Aldaya Garde

Banca Examinadora:

- Ivan Aritz Aldaya Garde (Orientador)
- Leandra Isabel de Abreu (Examinadora)
- Rafael Abrantes Penchel (Examinador)

A ata da defesa com as respectivas assinaturas dos membros encontra-se no prontuário do aluno (Processo nº 282/2023)

São João da Boa Vista, 26 de junho de 2024

A todos professores que, ao longo da minha vida, guiaram-me com sabedoria.

AGRADECIMENTOS

A Deus, fonte inesgotável de força e sabedoria, por guiar meus passos durante este caminho, sustentando minhas forças e me amando incondicionalmente.

Aos meus pais, Eldin Aldrin Francisco e Luciana Aparecida Nora Francisco, pelo incentivo ao longo de toda a minha graduação. Eles são meu alicerce, fonte de minha alegria e motivação para superar todos os desafios da vida.

Ao meu namorado, Lucas Ferreira Janizello Reggio, pela compreensão nos momentos de ausência e pelo suporte emocional.

Às minhas amigas, Elígia Simionato e Marcela Thaiane Pires de Souza, por cada conselho e gesto de apoio. Sem elas, minha jornada acadêmica seria muito mais difícil.

À Profa. Dra. Cintya Wink de Oliveira Benedito, que me motivou a iniciar minha jornada na iniciação científica e que constantemente estimula mais mulheres a se dedicarem à ciência.

Finalmente, ao Prof. Dr. Ivan Aritz Aldaya Garde, que me acompanhou como orientador de iniciação científica por três anos e também me orientou neste trabalho, desempenhando este papel com muita dedicação e paciência.

Este trabalho contou com o apoio das seguintes entidades:

Processo nº 2023/00591-0, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)

“As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade do(s) autor(es) e não necessariamente refletem a visão da FAPESP.”

Processos nº 2715 e 6163 - Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

*“Meu filho, aceita a instrução desde teus jovens anos;
ganharás uma sabedoria que durará até a velhice.
Vai ao encontro dela, como aquele que lavra e semeia;
espera pacientemente seus excelentes frutos.”
(Eclesiástico 6:18-19)*

RESUMO

A maior capacidade de transmissão de dados na Internet é uma necessidade crescente devido a atual implementação de novas tecnologias, como o 5G, que sobrecarregam não apenas os sistemas de comunicação móveis, mas também as conexões entre radiobases e a conexão dos sistemas celulares à Internet. As redes de comunicações ópticas são capazes de atender esta demanda, mas ainda dependem de amplificadores ópticos. Estes dispositivos utilizam fibras ópticas que permitem aumentar o produto comprimento de enlace por banda e são equipamentos indispensáveis para alcançar as distâncias impostas ao sistema. Entre os mais usados, estão os amplificadores à fibra dopada com érbio que são fundamentais para amplificar os sinais ópticos na janela de comprimento de onda de 1550 nm. Como o próprio nome expressa, o núcleo das fibras desses amplificadores é dopado com o elemento érbio durante o processo de fabricação. Já na etapa de uso, as fibras são bombeadas opticamente, excitando os seus íons e resultando no aumento do ganho óptico. No entanto, nos sistemas multiplexados por divisão de comprimento de onda existe o efeito do ganho cruzado, o qual é resultante da competição entre os feixes do sinal pelos íons excitados e faz com que o ganho experimentado por cada canal não seja dependente apenas da própria intensidade, mas também é influenciado pelas intensidades dos outros canais. Isso dificulta o controle das potências dos canais na saída dos amplificadores, fazendo com que determinar as potências de saída do amplificador a partir das potências de entrada se torne uma tarefa difícil. Existem modelos matemáticos que representam este efeito, mas necessitam do conhecimento de diversos parâmetros. Uma alternativa a eles, são os diversos modelos baseados em redes neurais artificiais encontrados em pesquisas da área, que são capazes de levar em conta a interação entre os distintos canais do sistema. Além disso, em alguns problemas práticos, não é necessário encontrar as potências de saída do amplificador, mas é preciso encontrar as potências de entrada que fornecem determinadas potências de saída desejadas. Este é o objetivo deste trabalho. Para atendê-lo, dois métodos foram explorados: (I) um modelo inverso de EDFA baseado em redes neurais artificiais e (II) um método baseado em otimização substituta. O método (I) foi empregado de forma inversa para manter potências constantes nas saídas do EDFA, partindo das potências de entrada do amplificador. Já no método (II), foram combinados um modelo de EDFA baseado em redes neurais artificiais com um algoritmo de otimização heurística, visando o mesmo objetivo. Os resultados obtidos por essas abordagens foram comparados para avaliação e a otimização substituta obteve os melhores resultados.

PALAVRAS-CHAVE: inteligência artificial; fibras ópticas; amplificadores ópticos; érbio.

ABSTRACT

The increasing demand for higher data transmission capacity on the Internet is driven by the current implementation of new technologies, such as 5G, which not only overload mobile communication systems but also the connections between base stations and the cellular systems' connection to the Internet. Optical communication networks are capable of meeting this demand, but still rely on optical amplifiers. These devices use optical fibers to increase the length-bandwidth product of the link and are indispensable equipment for achieving the distances imposed on the system. Among the most commonly used are erbium-doped fiber amplifiers, which are essential for amplifying optical signals in the 1550 nm wavelength window. As the name suggests, the cores of the fibers in these amplifiers are doped with erbium during the manufacturing process. In the operational stage, the fibers are optically pumped, exciting their ions and resulting in increased optical gain. However, in wavelength-division multiplexed systems, there is the phenomenon of cross-gain, which results from the competition between signal beams for the excited ions and causes the gain experienced by each channel to depend not only on its own intensity but also on the intensities of the other channels. This complicates the control of channel powers at the amplifier output, making it difficult to determine the amplifier output powers from the input powers. There are mathematical models that represent this effect, but they require knowledge of various parameters. An alternative to them is the various models based on artificial neural networks found in research in the field, which are capable of taking into account the interaction between the different channels of the system. Furthermore, in some practical problems, it is not necessary to find the output powers of the amplifier, but it is necessary to find the input powers that provide certain desired output powers. This is the objective of this work. To address it, two methods were explored: (I) an inverse model of EDFA based on artificial neural networks and (II) a method based on surrogate optimization. Method (I) was employed inversely to maintain constant powers at the EDFA outputs, starting from the input powers of the amplifier. In method (II), an EDFA model based on artificial neural networks was combined with a heuristic optimization algorithm, aiming for the same objective. The results obtained by these approaches were compared for evaluation, and surrogate optimization achieved the best results.

KEYWORDS: artificial intelligence; optical fibers; optical amplifiers; erbium.

LISTA DE ILUSTRAÇÕES

Figura 1	Representação gráfica do (a) mapeamento de elementos do espaço de soluções no espaço imagem e (b) mapeamento inverso do espaço imagem.	17
Figura 2	Representações dos neurônios (a) natural e (b) artificial.	19
Figura 3	ANN <i>feedforward</i> e ANN recorrente (<i>recurrent</i>).	20
Figura 4	ANN <i>shallow</i> (rasa) e ANN <i>deep</i> (profunda).	21
Figura 5	ANN <i>Fully connected/dense</i> (densamente conectada) e ANN <i>sparse</i> (esparsa/dispersa).	21
Figura 6	Rede neural artificial com arquitetura MLP.	22
Figura 7	Mecanismo clássico de evolução diferencial.	23
Figura 8	Diagrama de blocos do EDFA simulado em VPI Transmission Maker	25
Figura 9	Relação entre as potências de entrada e as potências de saída de cada canal do sistema WDM simulado.	26
Figura 10	Relação entre a potência total de entrada e a potência total de saída obtidas por meio da soma, em unidades lineares, das potências dos canais.	27
Figura 11	Gráfico da matriz de correlação cruzada entre as potências de entrada e de saída obtidas através das simulações.	27
Figura 12	Esquema do EDFA modelado e o modelo de EDFA baseado em ANN.	29
Figura 13	$\log_{10}(1 - score)$ em termos do número de neurônios para 1, 2 e 3 camadas ocultas.	30
Figura 14	Funções de ativação testadas: Identidade, Logística, Tangente Hiperbólica e Unidade Linear Retificada.	31
Figura 15	<i>Score</i> da validação cruzada do modelo baseado em ANN conforme o número de neurônios foi variado para cada uma das funções de ativação testadas.	31
Figura 16	Comparação entre as potências de saída desejadas e as potências de saída previstas pelo modelo baseado em ANN utilizando uma única camada de 140 neurônios para cada função de ativação analisada.	32
Figura 17	Gráfico da relação entre as potências de saída ideais e das potências de saída previstas pelo modelo inicial baseado em ANN para cada canal.	32
Figura 18	Histograma de distribuição de erros de treinamento e predição do modelo direto otimizado, onde o valor médio é indicado em linha descontínua e a predição de erro em termos do número de amostras.	33
Figura 19	Representação do modelo baseado em ANN e o modelo inverso baseado em ANN.	34
Figura 20	Combinação do modelo inverso de EDFA e do modelo de EDFA para fins de análise dos resultados obtidos.	35
Figura 21	Integração da ANN com o algoritmo DE.	35
Figura 22	Resultados do desempenho do modelo inverso baseado em ANN.	36
Figura 23	Resultados obtidos utilizando a otimização substituta.	37

LISTA DE TABELAS

Tabela 1 – Exemplo de configuração de potências do EDFA normalizadas.	36
---	----

LISTA DE ABREVIATURAS E SIGLAS

ADAM	<i>Adaptive Moment Estimation</i> (Estimador de Momento Adaptativo)
ANN	<i>Artificial Neural Network</i> (Rede Neural Artificial)
DE	<i>Differential Evolution</i> (Evolução Diferencial)
DEMUX	<i>Demultiplexer</i> (Demultiplexador)
EDF	<i>Erbium Doped Fiber</i> (Fibra Dopada com Érbio)
EDFA	<i>Erbium-Doped Fiber Amplifiers</i> (Amplificadores à Fibra Dopada com Érbio)
MLP	<i>Multilayer Perceptron</i> (Perceptron Multicamada)
MUX	<i>Multiplexer</i> (Multiplexador)
MSE	<i>Mean Squared Error</i> (Erro Quadrático Médio)
OBPF	<i>Optical Band Pass Filter</i> (Filtro Óptico Passa Faixa)
PLD	PLD, <i>Pump laser diode</i> (Diodo Laser de Bombeio)
RNN	<i>Recurrent Neural Network</i> (Rede Neural Recorrente)
ReLu	<i>Rectified Linear Unit</i> (Unidade Linear Retificada)
WDM	<i>Wavelength Division Multiplexing</i> (Multiplexação por Divisão de Comprimento de Onda)

LISTA DE SÍMBOLOS

Er^{3+}	Íon de érbio trivalente
k	Número de canais do sistema WDM
r	Coordenada radial
ϕ	Coordenada angular
z	Coordenada axial
P_k	Potência do k-ésimo canal
n_1	Densidade de íons não excitados na fibra
n_2	Densidade de íons excitados na fibra
σ_{ak}	Seção transversal de absorção da fibra
σ_{ek}	Seção transversal de emissão da fibra
τ	Tempo de vida do fóton
E_p	Energia do fóton na frequência do canal k
h	Constante de Planck
ν_k	Frequência do fóton no canal k
α_k	Espectro de absorção
g^*_k	Espectro de ganho
l_k	Perda extra devido ao pico de absorção da água
m	Número de modos suportado pela fibra
$\Delta\nu_k$	Largura de banda de frequência dos feixes ópticos
I_k	Intensidade de luz do canal k
i_k	Intensidade óptica normalizada
n_t	Densidade de íons de érbio local
Γ_k	Integral de sobreposição entre o modo dopante e óptico
t	Número de entradas de um perceptron
$S_{in\ j}$	J-ésima entrada de um perceptron

w_j	J-ésimo peso sináptico de um perceptron
S_{out}	Saída de um perceptron
w_0	Constante proveniente de viés adicional
S_0	Viés adicional
$\mathbf{w}$	Vetor de pesos
$\mathbf{S}_{in}$	Vetor de entradas
f	Função de ativação
N	Número de observações feitas
Y_i	Valores reais do conjunto de dados
$\hat{Y}_i$	Valores previstos pelo modelo
P_{in}	Potência de entrada do EDFA
P_{out}	Potência de saída do EDFA
P'_{out}	Potência de saída do EDFA desejada
P''_{out}	Potência de saída do EDFA encontrada pela rede Tandem
Pop^0	* População inicial gerada no estágio de inicialização do DE ¹
Pop^1	* População gerada pela evolução da população inicial
Pop^2	* População gerada pela evolução da população inicial
Pop^n	* População gerada na n-ésima evolução
Pop^{n+1}	* População gerada pela evolução da população Pop^n
N_p	* Número de indivíduos da população inicial
$p^{0,i}$	* Indivíduos da população inicial, com $1 \leq i \leq N_p$
$p^{n,i}$	* I-ésimo indivíduo da população n
$v^{n+1,i}$	* Indivíduo mutante gerado na etapa de mutação diferencial
$x^{n+1,v,i}$	* Vetor mutante de $v^{n+1,i}$
F_y	* Intensidade de mutação para a y-ésima diferença de vetor $x^{n,p1_y} - x^{n,p2_y}$

¹ Devido a quantidade de símbolos utilizada neste trabalho, houve a necessidade de reutilizar alguns caracteres presentes em símbolos de seções anteriores. Assim, todos os símbolos indicados com * são utilizados na Seção 2.5. Eles foram marcados para facilitar a identificação.

$x^{n,p_{1y}}$	* Vetor de diferença do indivíduo $p^{n,p_{1y}}$ da população P^n
$x^{n,p_{2y}}$	* Vetor de diferença do indivíduo $p^{n,p_{2y}}$ da população P^n
$x^{n+1,b,i}$	* Vetor de parâmetros de otimização da base de mutação diferencial
$b^{n,i}$	* Base de mutação diferencial
$c^{n+1,i}$	* Indivíduo gerado do cruzamento entre $v^{n+1,i}$ e $p^{n,i}$ na etapa de <i>crossover</i>

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Amplificadores ópticos à fibra dopada com érbio	14
1.2	Modelagem de amplificadores à fibra dopada com érbio	15
1.3	Objetivos	18
1.4	Estrutura do documento	18
2	REDES NEURAIS PARA REGRESSÃO SUPERVISIONADA	19
2.1	Neurônios artificiais	19
2.2	Arquiteturas das redes neurais artificiais	20
3	EVOLUÇÃO DIFERENCIAL	23
4	MÉTODOS	25
4.1	Simulação do sistema e análise estatística inicial	25
4.2	Modelo de EDFA baseado em ANN	28
4.3	Modelo inverso de EDFA	34
4.4	Otimização substituta	35
5	RESULTADOS	36
5.1	Rede inversa	36
5.2	Otimização substituta	37
6	CONCLUSÕES	38
6.1	Sugestões de trabalhos futuros	38
	REFERÊNCIAS	39

1 INTRODUÇÃO

1.1 AMPLIFICADORES ÓPTICOS À FIBRA DOPADA COM ÉRBIO

O desenvolvimento dos sistemas 5G contribui cada vez mais para um aumento no fluxo de dados transmitidos na Internet. De acordo com o relatório sobre mobilidade da Ericsson publicado em novembro de 2023, as medições apontam que o tráfego de envio de dados do usuário para a rede (*uplink*) representa 54% do volume total nos serviços de armazenamento em nuvem enquanto o tráfego de *downlink* corresponde a 46% (ERICSSON, 2023). Além disso, o relatório diz que o consumo global de tráfego de dados móveis poderá atingir 56 GB por *smartphone* a cada mês até o final de 2029 (ERICSSON, 2023). Todos estes dados apontam para um futuro onde haverá uma demanda por maior capacidade de transmissão de dados na Internet, sobrecarregando não apenas os sistemas de comunicação móveis, mas também a conexão entre os diferentes nós das redes. As redes de comunicações ópticas são soluções capazes de atender a essa crescente necessidade de banda, já que são o meio de transmissão guiado com o produto alcance por banda mais significativo (LATHI B. P E DING, 2012).

Apesar da maturidade das tecnologias de construção das fibras ópticas, que reduzem as perdas de transmissão, os sistemas de comunicação de longa distância ainda dependem de amplificadores ópticos para operar com melhor desempenho. Durante os anos 1980, várias tecnologias de amplificadores foram desenvolvidas. Foi a partir da década de 1990, com o desenvolvimento dos amplificadores à fibra dopada com érbio (*Erbium-Doped Fiber Amplifiers*, EDFA), que estes dispositivos passaram a ser amplamente empregados nos sistemas de transmissão ópticos. Os EDFAs desempenham um papel central em sistemas de comunicação óptica, operando principalmente na faixa de 1550 nm (AGRAWAL, 2016). Dentre os diversos tipos de amplificadores existentes na atualidade, destacam-se os amplificadores semicondutores, os amplificadores Raman e os EDFAs. A partir de 1995, os EDFAs permitiram a implementação eficiente dos sistemas multiplexados por divisão de comprimento de onda (*Wavelength Division Multiplexing*, WDM), possibilitando a amplificação simultânea de múltiplos canais com diferentes formatos de modulação (KEISER, 2014). Recentemente, os EDFAs começaram a ser combinados com amplificação distribuída baseada em espalhamento estimulado de Raman.

Independentemente da configuração específica, os EDFAs compartilham um princípio de funcionamento comum: fazem uso de uma fibra dopada com íon érbio (Er^{3+}) como meio de ganho. Contudo, durante o processo de fabricação, o núcleo da fibra do amplificador não é dopado exclusivamente com érbio, mas com combinações de elementos de terras raras, como o itérbio ou túlio (KEISER, 2014). Quando as fibras dopadas com esses elementos são bombeadas opticamente, os íons são excitados a níveis energéticos superiores. Ao atingir uma quantidade suficientemente alta de íons excitados, ocorre o processo denominado inversão de população. Quando um fóton de sinal, com o comprimento de onda apropriado, encontra um desses íons excitados, acontece a emissão estimulada de um novo fóton que compartilha as mesmas características do fóton original. O ganho óptico é alcançado quando a densidade de fótons atinge um nível elevado, aumentando a probabilidade de ocorrência de eventos de emissão estimulada (AGRAWAL, 2021). Para entender a dinâmica desse processo de ganho, é necessá-

rio considerar diversos parâmetros como as seções de captura e emissão, tempos de vida de decaimento e a distribuição transversal e longitudinal da intensidade dos feixes de bombeio e sinal (DESURVIRE; ZERVAS, 1995). Além disso, o ganho do EDFA possui uma seletividade frequencial característica, uma vez que o perfil modal e as seções transversais de captura e emissão dependem do comprimento de onda. Nos sistemas WDM, os feixes de sinal dos diversos canais competem pelos íons excitados. Essa competição resulta em um ganho que não depende apenas da intensidade do feixe em questão, mas também das intensidades dos demais canais. Esse fenômeno é chamado de modulação de ganho cruzado e dificulta o dimensionamento da rede (MENIF et al., 2001; FREEMAN; CONRADI, 1993). A modelagem de um EDFA é ainda mais desafiadora devido à possibilidade do dispositivo ser composto por múltiplos estágios, cada um com diferentes configurações de bombeamento e comprimentos de fibra. Esses estágios são necessários para equalizar o espectro de ganho e estender a largura de banda de operação, mas introduzem complexidade adicional no processo de modelagem (AGRAWAL, 2016).

1.2 MODELAGEM DE AMPLIFICADORES À FIBRA DOPADA COM ÉRBITO

Na literatura, podem ser encontrados diferentes modelos de EDFA. Um desses modelos é o modelo de Giles-Desurvire que é amplamente empregado para descrever os EDFAs com precisão. Esse modelo pode descrever amplificadores de vários estágios. As equações diferenciais acopladas que descrevem este modelo levam em consideração a seção transversal da fibra e as potências de cada canal para diferentes frequências, constituindo uma modelagem espaço-temporal (GILES; DESURVIRE, 1991). São elas:

$$\frac{dn_2}{dt} = \sum_k \frac{P_k i_k \sigma_{ak}}{h\nu_k} n_1(r, \phi, z) - \sum_k \frac{P_k i_k \sigma_{ek}}{h\nu_k} n_2(r, \phi, z) - \frac{n_2(r, \phi, z)}{\tau}, \quad (1)$$

e

$$\begin{aligned} \frac{dP_k}{dz} = & u_k \sigma_{ek} \int_0^{2\pi} \int_0^\infty i_k(r, \phi) n_2(r, \phi, z) r dr d\phi (P_k(z) + m h \nu_k \Delta \nu_k) \\ & - u_k \sigma_{ak} \int_0^{2\pi} \int_0^\infty i_k(r, \phi) \cdot n_1(r, \phi, z) r dr d\phi (P_k(z)). \end{aligned} \quad (2)$$

Nelas, r é a coordenada radial, ϕ é a coordenada angular, z é a coordenada axial e k é o número de canais do sistema WDM. Além disso, têm-se que a densidade de íons não excitados é n_1 , a densidade de íons excitados é n_2 , as seções transversais de absorção e emissão são, respectivamente, σ_{ak} e σ_{ek} , h é a constante de Planck, ν_k é a frequência do fóton no canal k e τ é o tempo de vida do fóton. O número de modos m é geralmente igual a 2 e $\Delta \nu_k$ é a largura de banda de frequência dos feixes ópticos. O termo u_k indica se o feixe está viajando na direção propagante ($u_k = 1$) ou contra-propagante ($u_k = -1$) (GILES; DESURVIRE, 1991). A potência do k -ésimo canal, P_k , é a potência total de um feixe k sobre as coordenadas radial e azimutal e pode ser obtida integrando a intensidade de luz I_k , ou seja,

$$P_k(z) = \int_0^{2\pi} \int_0^\infty I_k(r, \phi, z) r dr d\phi. \quad (3)$$

Outro parâmetro presente nas equações diferenciais acopladas desse modelo é a intensidade óptica normalizada. Ela é independente de z e encontrada através de

$$i_k(r, \phi) = \frac{I_k(r, \phi, z)}{P_k(z)}. \quad (4)$$

Por um lado, a Equação 1 descreve a evolução temporal de n_2 , a densidade de íons no estado excitado, que é responsável pela modulação do ganho cruzado. Por outro lado, a Equação 2 descreve a propagação dos feixes através da fibra. Assim, o número de equações obtidas por meio da Equação 2 é igual ao número de feixes que se propagam pela fibra, incluindo os feixes de sinal e de bombeio. O primeiro termo da parte direita da Equação 1 descreve os efeitos de absorção, o segundo é responsável por descrever a emissão estimulada e o terceiro a emissão espontânea. Já na Equação 2, o primeiro termo da parte direita está relacionado com o ganho e o segundo, com as perdas. Tanto a Equação 1, quanto a Equação 2 consideram um sistema WDM de k canais com um único bombeio e dois níveis. A solução no regime estacionário que pode ser obtida pelas equações anteriores é

$$n_2(r, \phi, z) = n_t \left(\frac{\sum_k \frac{\tau \sigma_{ak}}{h\nu_k} P_k i_k}{1 + \sum_k \tau \frac{\sigma_{ak} + \sigma_{ek}}{h\nu_k} P_k i_k} \right), \quad (5)$$

sendo que n_t é a densidade de íons de érbio local descrita por

$$n_t(r, \phi, z) = n_1(r, \phi, z) + n_2(r, \phi, z), \quad (6)$$

ou seja, é a soma da densidade de íons no estado excitado e da densidade de íons no estado não excitado.

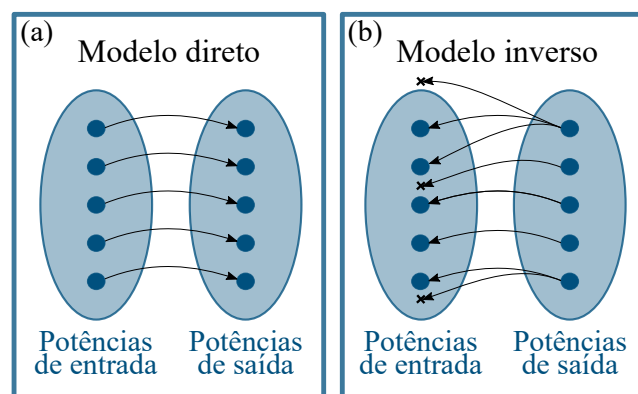
Em conjunto com a disponibilização de novas faixas de frequência para transmissão e o uso da multiplexação por divisão espacial, estão sendo conduzidas pesquisas voltadas para aprimorar o uso eficiente dos recursos. Modelos precisos dos subsistemas de transmissão, como as fibras ópticas e amplificadores, desempenham um papel fundamental na busca pela otimização completa de ponta a ponta (ROS; MOURA; YANKOV, 2020). O modelo Giles-Desurvire considera adequadamente a modulação de ganho cruzado entre os canais WDM e, por isso, é possível utilizá-lo para dimensionar redes de comunicação ópticas. Ele possibilita a determinação das potências de saída do amplificador com base nas potências de entrada. No entanto, ele requer a caracterização prévia de muitos parâmetros que podem ser difíceis de caracterizar e o custo computacional associado a esse modelo é considerável.

Por outro lado, a área de aprendizado de máquina tem emergido como uma ferramenta importante para abordar problemas de alta complexidade sujeitos a limitações estritas de custo computacional. Nesse campo, estão incluídas as redes neurais artificiais (*Artificial Neural Network*, ANNs), que são estruturas matemáticas e computacionais baseadas no cérebro humano e têm o potencial de funcionar como aproximadores universais (ALPAYDIN, 2010). Elas são candidatas promissoras para servir como modelos de sistemas e podem ser treinadas com base em dados experimentais. Portanto, o uso de um modelo de EDFA baseado em ANN que possa descrever o ganho cruzado entre os diferentes

canais do sistema WDM emerge como uma alternativa à utilização do modelo Giles-Desurvire para prever as potências de saída do amplificador em relação às potências de entrada. Existem diferentes trabalhos que realizam esta análise com abordagens semelhantes. Em [Ros, Moura e Yankov \(2020\)](#), foram obtidos dados experimentais a partir de um *setup* que possuía três EDFAs de mesma construção em cascata. A ANN escolhida para a implementação do modelo foi elaborada em PyTorch e treinada usando o Estimador de Momento Adaptativo (*Adaptive Moment Estimation*, ADAM). Uma análise de ganho alternativa também foi proposta em [Wang, Kilper e Chen \(2023\)](#), onde foi utilizada uma rede neural artificial profunda e o aprendizado de transferência, que permite a construção de um novo modelo com base em um modelo pré-treinado usando poucas amostras. Já em [Lin, Lin e Jiang \(2023\)](#), foi utilizada uma rede perceptron multicamada com duas camadas ocultas, mas não foram analisados os efeitos da utilização de diferentes funções de ativação, número de neurônios ou número de camadas possíveis. Por sua vez, em [Meseguer et al. \(2021\)](#) foi desenvolvido um modelo híbrido de ganho de EDFA baseado em técnicas de aprendizado de máquina e medida agnóstica, que previa o ganho de um EDFA com 0,05 dB de erro quadrático médio quando uma entrada WDM não plana de banda C (1530 nm-1565 nm) era fornecida. Em [Bastos-Filho, Barboza e Martins-Filho \(2017\)](#) foi proposto o uso de uma ANN para estimar o ganho do canal e a figura de ruído, ou seja, as saídas da ANN não são os valores de potências de saída do EDFA. E outras pesquisas com análises semelhantes são consideradas em [Barboza et al. \(2021\)](#), [Filho et al. \(2022\)](#) e em [Zhu et al. \(2018\)](#).

No entanto, em alguns problemas de dimensionamento de rede, não são conhecidas as potências de entrada do EDFA, mas sim as potências de saída desejadas. Neste caso, o modelo de Giles-Desurvire mencionado não permite encontrar as potências de entrada que são necessárias para alcançar essas potências de saída desejadas, pois opera de maneira unidirecional e as relações entre entradas e saídas podem não ser estritamente bijetivas. Isso significa que o modelo não possui um mapeamento bidirecional e único entre o domínio e o espaço imagem. Logo, cada configuração de saída é atingida somente por uma configuração de entrada e o inverso não pode ser considerado, já que o modelo pode encontrar diferentes soluções (esquerda) e imagem (direita) que podem ou não ser válidas. Isso é representado na [Figura 1](#), onde as elipses representam os conjuntos de soluções, os círculos menores em azul escuro são as entradas e as saídas do modelo e os símbolos *X* representam soluções inválidas.

Figura 1 – Representação gráfica do (a) mapeamento de elementos do espaço de soluções no espaço imagem e (b) mapeamento inverso do espaço imagem.



fonte: Produção da própria autora.

Portanto, a inversão de um modelo de EDFA para encontrar as potências de entrada que resultem em potências de saída específicas, pode ser desafiadora devido à natureza não linear do modelo matemático e do sistema físico do EDFA. Para enfrentar este desafio, uma abordagem eficaz para atender às necessidades específicas do sistema é a utilização de modelos baseados em ANN e algoritmos de otimização heurística. Neste trabalho, foram desenvolvidos modelos inversos de EDFA, utilizando ANNs e um algoritmo de evolução diferencial (*Differential Evolution*, DE), que consideraram o efeito da modulação de ganho entre os diferentes canais WDM.

1.3 OBJETIVOS

O objetivo principal deste trabalho é inverter um modelo de EDFA, ou seja, encontrar a configuração de potências de entrada para atingir determinadas potências de saída iguais. Para isso, foram implementadas duas abordagens. A primeira abordagem baseia-se num modelo inverso de EDFA baseado em uma rede neural artificial. Já na segunda, utilizou-se um modelo direto de EDFA baseado em uma rede neural artificial em conjunto com um algoritmo de otimização heurística, método conhecido como otimização substituta. Ambas abordagens serão melhor apresentadas nos próximos capítulos deste trabalho. Além disso, os resultados obtidos foram comparados a fim de verificar qual método obteve o melhor desempenho.

1.4 ESTRUTURA DO DOCUMENTO

O restante deste documento está organizado em cinco capítulos principais: Arquiteturas das Redes Neurais Artificiais, Evolução Diferencial, Métodos, Resultados e Conclusões. No Capítulo 2, são apresentados alguns conceitos e arquiteturas básicas de redes neurais artificiais. Já no Capítulo 3, é descrito o mecanismo clássico da Evolução Diferencial, algoritmo de otimização heurística empregado neste trabalho. Por outro lado, no Capítulo 4, é descrita a metodologia de coleta de dados e os procedimentos adotados na pesquisa, bem como a explicação das duas abordagens implementadas para a solução do problema considerado. No Capítulo 5, são apresentados os gráficos e análises dos resultados obtidos. Por fim, o Capítulo 6 engloba as principais observações do trabalho e oferece sugestões para possíveis estudos posteriores.

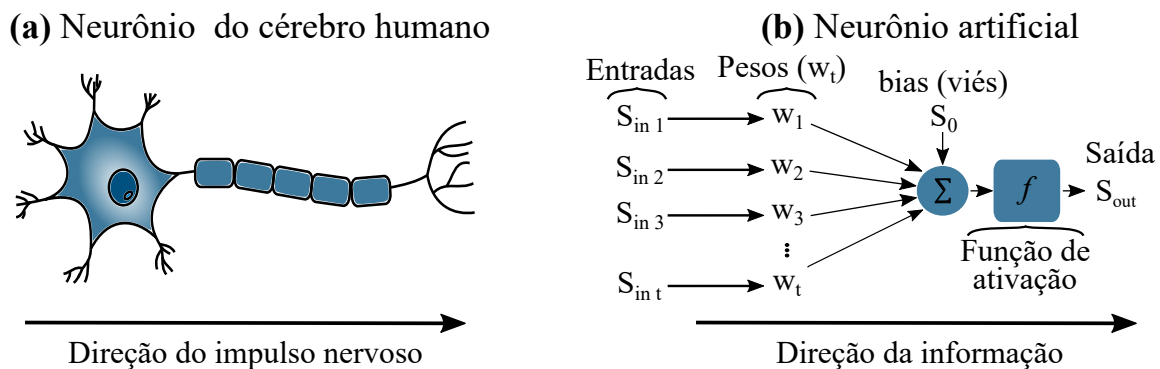
2 REDES NEURAIS PARA REGRESSÃO SUPERVISIONADA

As redes neurais artificiais foram criadas com base no cérebro humano (AGATONOVIC-KUSTRIN; BERESFORD, 2000). Elas são formadas por neurônios artificiais que atuam como as células do sistema nervoso, transmitindo e processando a informação (ALPAYDIN, 2010). Esses neurônios possuem uma função de ativação e estão conectados uns aos outros, formando camadas. Utilizando uma base de dados, é possível treinar essas estruturas para que elas aprendam o comportamento de um sistema por meio do reconhecimento de padrões e relação entre os dados (AGATONOVIC-KUSTRIN; BERESFORD, 2000). Assim, as ANNs podem atuar como modelos que são capazes de prever os valores de saída de um sistema a partir dos valores de entrada.

2.1 NEURÔNIOS ARTIFICIAIS

Na Figura 2 é possível observar a representação de uma célula do cérebro humano e a estrutura de um neurônio artificial. Os neurônios naturais são células do sistema nervoso que têm a capacidade de estabelecer conexões entre si ao receber estímulos do ambiente externo ou do próprio organismo e transmiti-los ao cérebro (VARELLA, 2024). Num cérebro humano, cada neurônio faz essas conexões, também conhecidas como sinapses, com aproximadamente 10^4 outros neurônios, que operam em paralelo, e isso resulta no grande poder computacional que o cérebro possui (ALPAYDIN, 2010; AGATONOVIC-KUSTRIN; BERESFORD, 2000; GERSHENSON, 2003).

Figura 2 – Representações dos neurônios (a) natural e (b) artificial.



fonte: Produção da própria autora.

Assim como no cérebro, o neurônio artificial é a unidade básica da rede. Existem diferentes modelos de neurônios artificiais. Entre eles, o modelo *perceptron* é o mais utilizado na maioria das redes neurais e é o neurônio representado à direita na Figura 2. Este modelo possui entradas que podem ser os dados de entrada do sistema que deseja-se modelar ou saídas de outros *perceptrons*. Considerando que esta unidade básica possui n entradas, cada entrada $S_{in j} \in \mathbb{R}$, com $j = 1, 2, \dots, n$, é multiplicada por um peso sináptico, também chamado de peso de conexão, $w_j \in \mathbb{R}$. A saída S_{out} , no caso mais simples, é a soma ponderada das entradas, mas geralmente é acrescentada uma constante w_0 que torna o modelo mais geral e que é proveniente de uma unidade de viés (*bias*) adicional S_0 . Assim,

a saída do *perceptron*, dada por

$$S_{out} = \sum_{j=1}^n w_j S_{in,j} + w_0, \quad (1)$$

é um produto escalar que também pode ser escrito como um produto ponto na forma

$$S_{out} = \mathbf{w}^T \mathbf{S}_{in}, \quad (2)$$

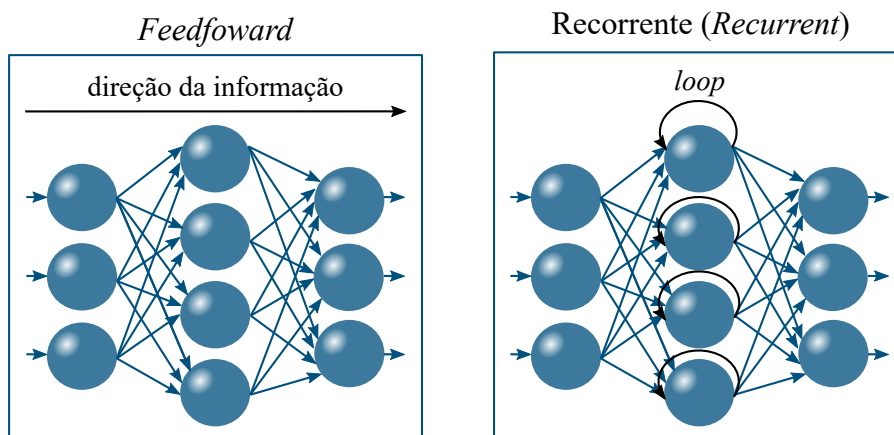
onde $\mathbf{w} = [w_0, w_1, w_2, \dots, w_n]^T$ e $\mathbf{S}_{in} = [1, S_{in,1}, S_{in,2}, \dots, S_{in,n}]^T$ são vetores aumentados para incluir o peso de viés e a entrada (ALPAYDIN, 2010).

A saída dessa combinação linear passa então por uma função de ativação f . Existem diferentes funções de ativação que podem ser utilizadas e, em sua maioria, são não lineares. Entre as mais comumente empregadas estão a função unidade linear retificada (*Rectified Linear Unit*, ReLu), que é muito utilizada em redes neurais convolucionais profundas, a função tangente hiperbólica, que possui saída limitada ao intervalo $[-1, 1]$, e a função logística, que tem saída limitada em $[0, 1]$ (RASAMOELINA; ADJAILIA; SINCAK, 2020; QIU; XU; CAI, 2018). Em geral as redes neurais são organizadas em camadas que contêm vários neurônios artificiais interligados e que podem também estar conectados às unidades de outra camada. Dessa maneira, a saída das redes é contínua e uma ANN configurada para operar em modo de regressão pode ser vista como uma caixa preta flexível capaz de realizar mapeamentos complicados entre variáveis (YU et al., 2021).

2.2 ARQUITETURAS DAS REDES NEURAIS ARTIFICIAIS

Podemos identificar diferentes tipos de ANNs dependendo do número de neurônios. Estes tipos não variam apenas de acordo com a quantidade de neurônios, mas com a quantidade de camadas, a forma como os neurônios estão conectados ou como a informação é transmitida. Na Figura 3, por exemplo, são representadas as redes do tipo *feedforward* e as redes recorrentes. As redes neurais do tipo *feedforward* são aquelas em que a informação é passada apenas em uma direção, sem retornar para neurônios anteriores. Por isso, elas são ANNs sem memória, ou seja, a saída depende apenas da entrada atual (FINE, 2006).

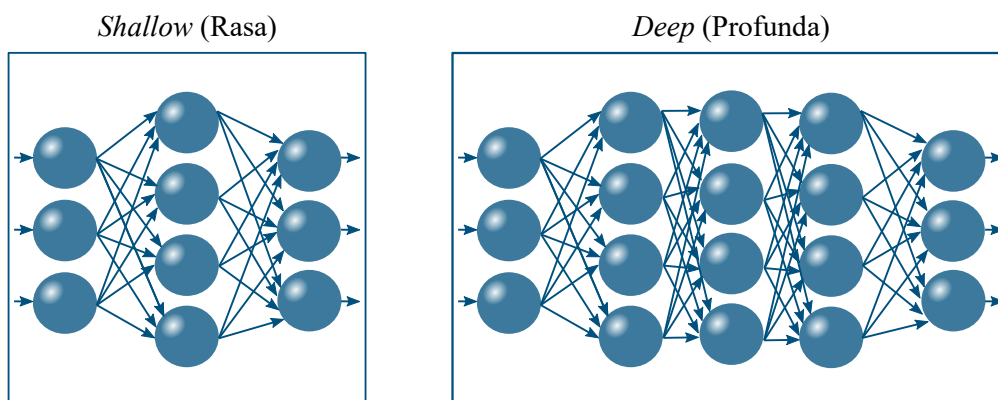
Figura 3 – ANN *feedforward* e ANN recorrente (*recurrent*).



fonte: Produção da própria autora.

Por outro lado, as redes do tipo recorrentes (*Recurrent Neural Network*, RNN) possuem *loops* que são responsáveis pela realimentação da ANN, fazendo com que a informação não flua apenas em um sentido, mas fique retida na rede, caracterizando esta arquitetura como uma rede com memória. Essa categoria de ANN pode apresentar problemas de instabilidade durante a etapa de treinamento, já que a informação fica retida na rede durante certo tempo antes de se perder (SHARKAWY, 2020). Outras duas categorias são apresentadas na Figura 4. A *shallow* (rasa), é o tipo de rede com uma ou apenas algumas camadas ocultas, e a *deep* (profunda), aquela que possui várias camadas ocultas.

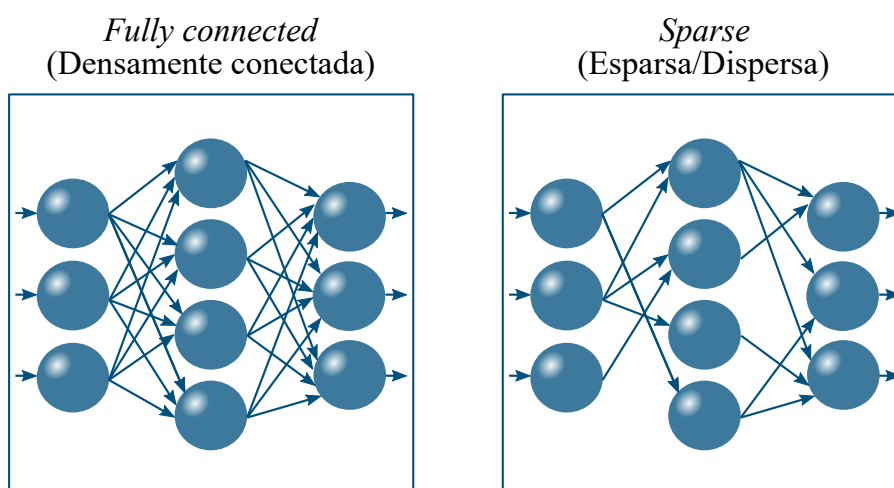
Figura 4 – ANN *shallow* (rasa) e ANN *deep* (profunda).



fonte: Produção da própria autora.

Já na Figura 5, os tipos *Fully connected/dense* (densamente conectada) e *sparse* (esparsa/dispersa) são exibidos. Enquanto nas redes neurais densamente conectadas todos os neurônios de uma camada estão conectados com todos os neurônios da outra camada, nas redes do tipo *sparse* nem todos estão conectados. Isso resulta na necessidade de otimizar um menor número de parâmetros, diminuindo a probabilidade de ocorrer *overfitting*.

Figura 5 – ANN *Fully connected/dense* (densamente conectada) e ANN *sparse* (esparsa/dispersa).

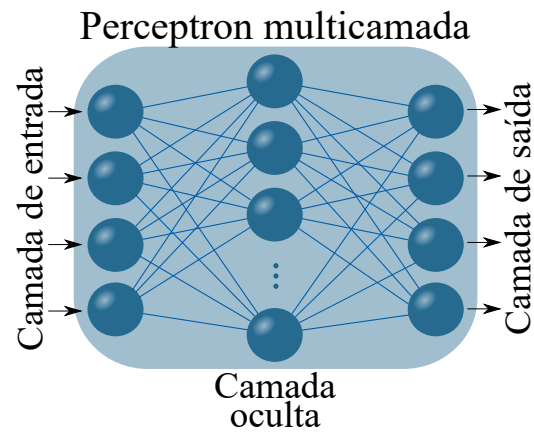


fonte: Produção da própria autora.

Uma arquitetura possível para as redes é a *perceptron* multicamada (*Multilayer Perceptron*, MLP). Essas redes são formadas por no mínimo três camadas: a camada de entrada, uma ou mais camadas ocultas e uma camada de saída, como pode ser observado na Figura 6. Esta arquitetura pode ser utilizada

tanto para resolução de problemas de classificação, quanto para problemas de regressão (Scikit-learn developers, 2024).

Figura 6 – Rede neural artificial com arquitetura MLP.

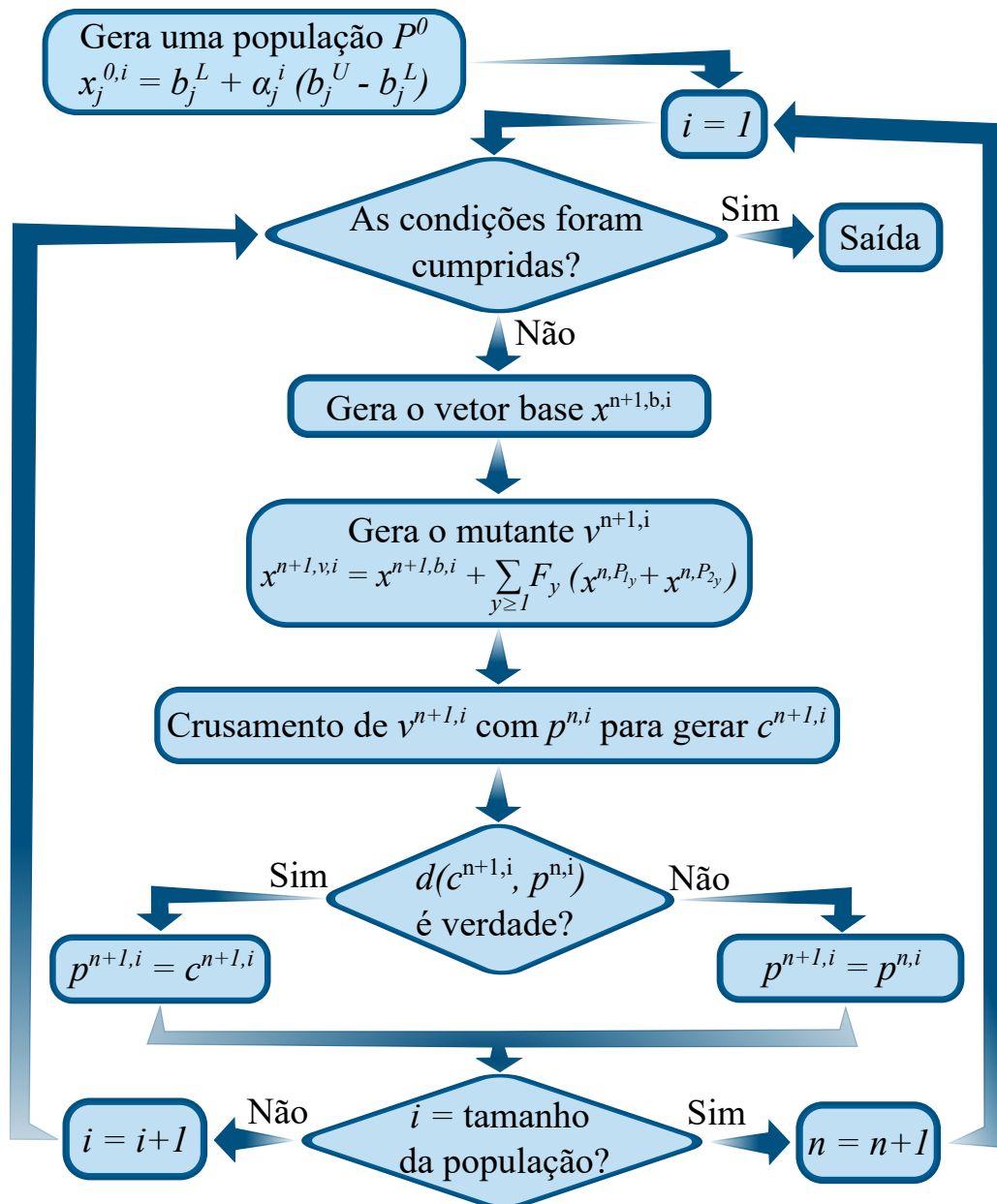


fonte: Produção da própria autora.

3 EVOLUÇÃO DIFERENCIAL

O DE é um algoritmo derivado do algoritmo genético que foi proposto por Kenneth V. Price e R. Storn, em 1995. Enquanto tentavam resolver o problema de ajuste polinomial de Chebyshev, eles acabaram descobrindo que o DE era eficiente para a otimização de funções e poderia ser utilizado em diferentes áreas. O DE emula a evolução das espécies, mas se diferencia do algoritmo genético pelo processo de mutação, já que no DE são utilizados três soluções tentativas anteriores ao invés de duas, o que acelera a convergência do algoritmo (QING, 2009). No fluxograma apresentado na Figura 7, está descrito o mecanismo da DE clássica que pode ser dividido nos estágios de inicialização e de evolução, a qual é composta pelas etapas de seleção, *crossover* (ou mistura) e, por fim, mutação.

Figura 7 – Mecanismo clássico de evolução diferencial.



fonte: Produção da própria autora.

Na inicialização, uma população inicial Pop^0 é gerada de forma aleatória com parâmetros uniformemente distribuídos em intervalos determinados. Essa população Pop^0 evolui para uma população Pop^1 , a qual evolui a uma população Pop^2 e assim iterativamente, até que na evolução n , ou seja, a evolução da população Pop^n para Pop^{n+1} , as condições de restrição estabelecidas sejam cumpridas. A população inicial Pop^0 é composta por N_p indivíduos $p^{0,i}$ com $1 \leq i \leq N_p$. O algoritmo é encerrado se os indivíduos da população satisfazem as condições desejadas. Se isso não ocorre, dá-se início à etapa de mutação diferencial. Nessa etapa, para cada indivíduo $p^{n,i}$ na população P^n , é gerado um indivíduo mutante $v^{n+1,i}$ com o vetor mutante $x^{n+1,v,i}$ obtido por meio de

$$x^{n+1,v,i} = x^{n+1,b,i} + \sum_{y \geq 1} F_y (x^{n,p_{1y}} - x^{n,p_{2y}}), \quad (1)$$

em que p_{1y} e p_{2y} são valores inteiros aleatórios, a constante F_y representa a intensidade de mutação para a y -ésima diferença de vetor $x^{n,p_{1y}} - x^{n,p_{2y}}$ e seu valor está limitado no intervalo $[0, 1]$. Por sua vez, $x^{n,p_{1y}}$ e $x^{n,p_{2y}}$ são os vetores de diferença dos indivíduos $p^{n,p_{1y}}$ e $p^{n,p_{2y}}$, respectivamente, da população P^n . O vetor $x^{n+1,b,i}$ é denominado vetor de parâmetros de otimização da base de mutação diferencial $b^{n,i}$. Na etapa de *crossover*, os indivíduos $v^{n+1,i}$ e $p^{n,i}$ são cruzados. A partir desse cruzamento é gerado o indivíduo $c^{n+1,i}$. Após isso, é realizada a etapa de seleção. Na seleção, se o indivíduo $c^{n+1,i}$ é mais dominante, ou seja, apresenta uma menor função de custo que o indivíduo $p^{n,i}$, então ele é acrescentado a nova população. Se não, $p^{n,i}$ é o indivíduo mantido na população. Porém, a fim de definir qual é o melhor indivíduo, é necessário avaliar as funções de custo do antecessor e do novo indivíduo, o que demanda um custo computacional.

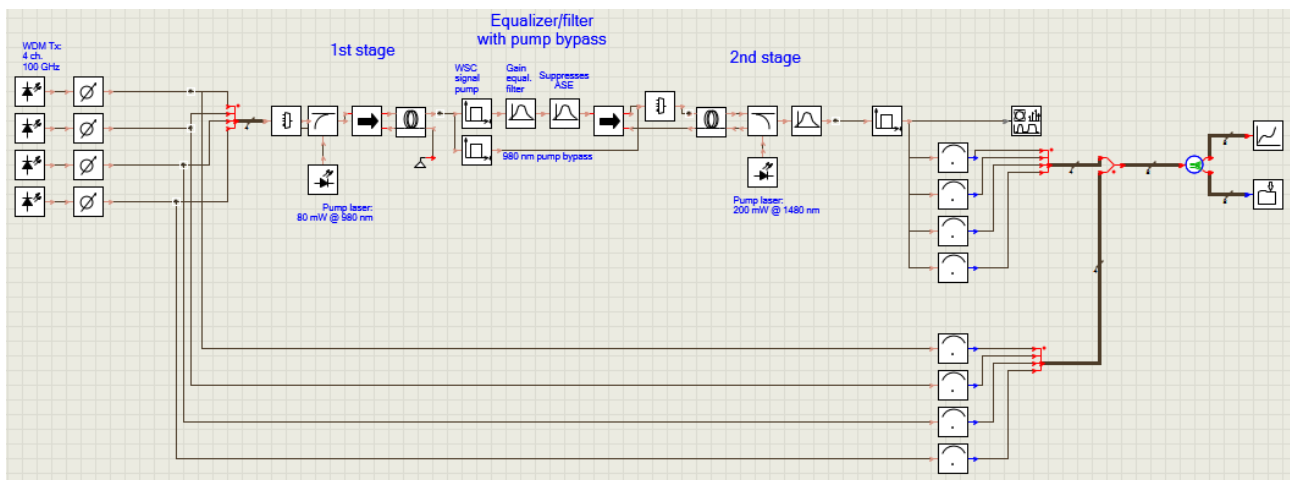
4 MÉTODOS

Neste capítulo, será apresentado o método e as ferramentas empregadas na coleta dos dados utilizados no trabalho. Também há uma breve descrição sobre as redes neurais e é apresentado o modelo baseado em uma rede neural artificial que serviu como base para as duas abordagens escolhidas. Por fim, o modelo inverso e a otimização substituta também são descritos neste capítulo.

4.1 SIMULAÇÃO DO SISTEMA E ANÁLISE ESTATÍSTICA INICIAL

O ponto de partida deste estudo consistiu na simulação computacional de um EDFA de dois estágios, apresentado na [Figura 8](#), empregando o *software* VPI Transmission Maker. O modelo das fibras utilizadas em ambos os estágios é baseado nas equações de Giles-Desurvire, o que possibilita uma simulação detalhada dos segmentos de fibra que constituem cada estágio do amplificador, por isso ele foi escolhido para a coleta dos dados que seriam utilizados neste trabalho. Cada estágio representado na figura incluiu um laser de bombeio, cuja finalidade era a excitação dos íons de érbio, em conjunto com um segmento de fibra e um isolador. O isolador foi usado com o propósito de bloquear o feixe retro-propagado causado pelo fenômeno de espalhamento de Brillouin estimulado, que poderia afetar o desempenho do amplificador ([AGRAWAL, 2021](#)).

Figura 8 – Diagrama de blocos do EDFA simulado em VPI Transmission Maker

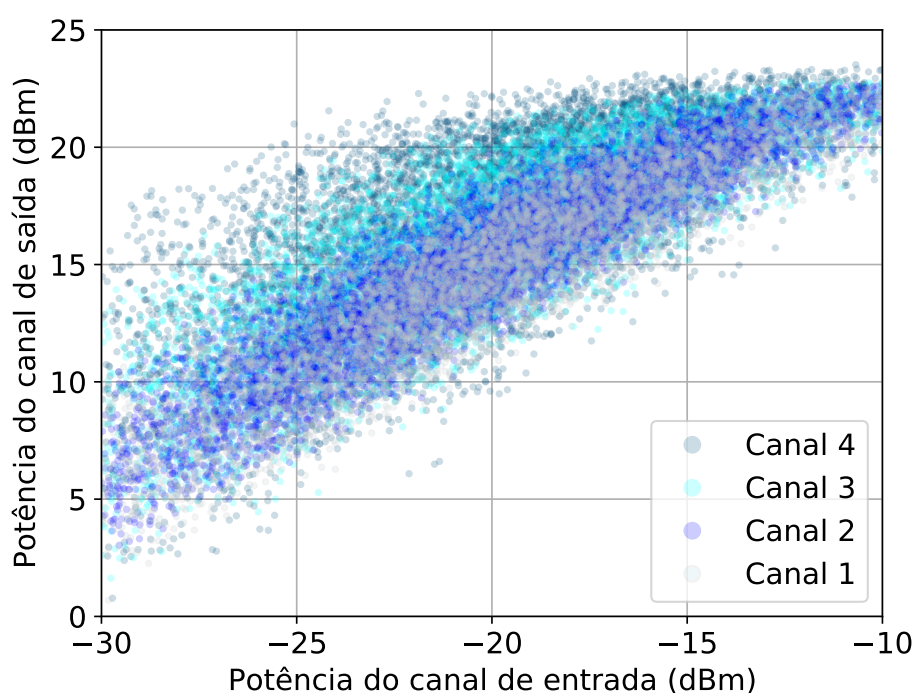


fonte: Produção da própria autora.

O primeiro estágio do EDFA simulado atuou como um pré-amplificador, garantindo uma figura de ruído baixa. Essa etapa era composta por uma fibra dopada com érbio com extensão de 30 m, submetida a um bombeamento na direção co-propagante a 980 nm e com potência de 80 mW. Por sua vez, o segundo estágio comportava-se como um amplificador de potência, possuía bombeio contra-propagante e era composto por uma fibra dopada de 30 m bombeada a 1480 nm à uma potência de 200 mW. Além disso, foram considerados apenas 4 canais, o que é um número relativamente baixo. Decidiu-se utilizar este número de canais, pois representa o pior cenário em termos do ganho cruzado. Os quatro canais foram frequencialmente separados em 100 GHz na banda C (1530 nm-1565 nm) e as potências

de entrada dos canais foram configuradas considerando uma distribuição uniforme entre -30 dBm e -10 dBm. Foram efetuadas 7412 simulações e foram obtidos os dados de potências de entrada e saída do EDFA simulado. Esses dados foram organizados em um arquivo com 8 colunas, sendo elas as potências de entrada e de saída dos 4 canais do EDFA, e 7412 linhas, o número de ensaios coletados. Na [Figura 9](#), é apresentada a relação entre a potência de entrada e a potência de saída de cada canal do sistema. Pela dispersão dos pontos nesta figura, é possível notar que a potência de saída de um canal não é apenas dependente de sua potência de entrada, indicando que existe uma dependência entre os canais. Também é possível notar que a dispersão dos pontos é maior para alguns canais do que para outros. Isso significa que estes canais sofrem mais com o efeito da modulação cruzada do que os outros.

Figura 9 – Relação entre as potências de entrada e as potências de saída de cada canal do sistema WDM simulado.

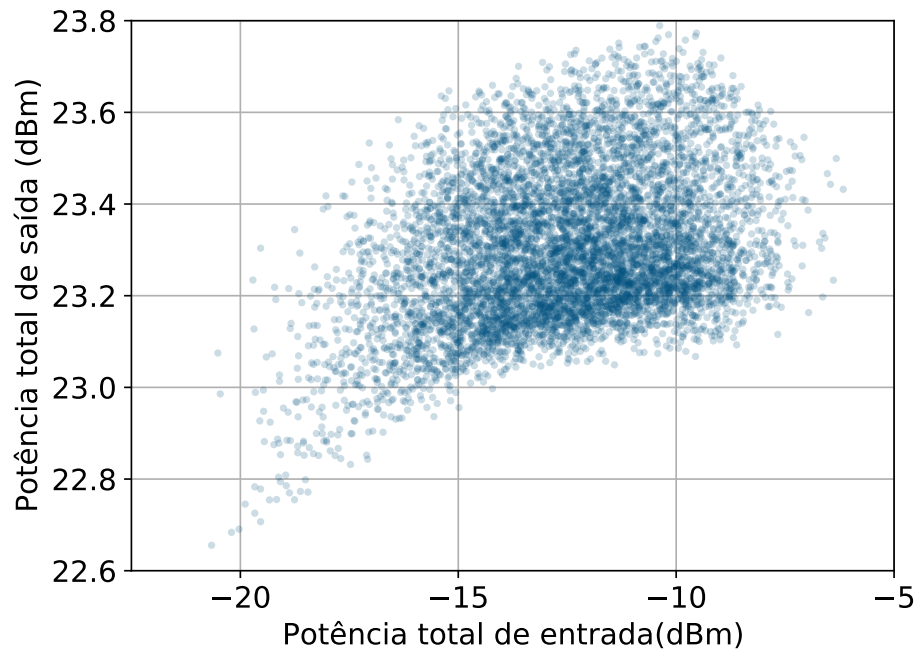


fonte: Produção da própria autora.

A partir desta primeira análise dos dados, poder-se-ia pensar que mesmo se a potência de entrada e a potência de saída de cada canal não estão completamente relacionadas, a potência total de entrada e de saída poderiam estar. Assim, após encontrar a potência total de entrada e a potência total de saída por meio da adição em unidades lineares, foi gerada a [Figura 10](#). Porém, nela também é possível observar que a potência total de saída não é apenas dependente da potência total de entrada, já que existe uma grande variabilidade nos pontos da figura. Dessa forma, a fim de compreender melhor a dependência entre as potências de entrada e saída dos canais devido à modulação do ganho cruzado, foi realizada uma análise de correlação entre os canais. No gráfico exibido na [Figura 11](#) apresenta-se a matriz de correlação entre as potências de entrada e saída de todos os canais. A escala à direita indica o coeficiente de correlação, índice que informa o quanto os dados estão relacionados. Quanto mais esse índice se aproxima dos valores 1 ou -1 , maior a dependência entre os dados, ou seja, valores absolutos representam total correlação. Isso pode ser observado na diagonal do gráfico, onde é exibida a relação

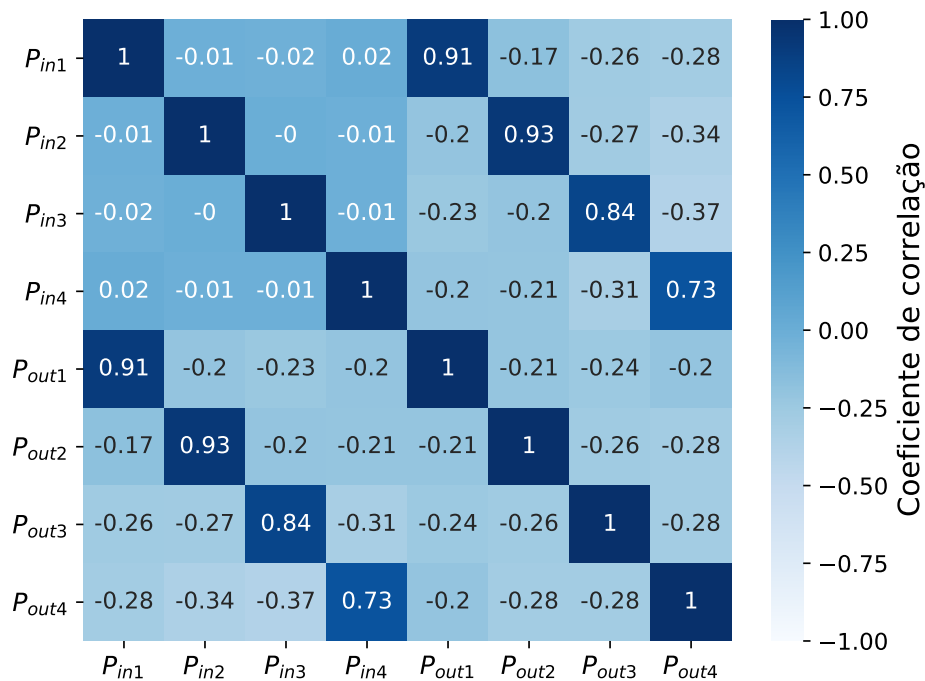
entre cada potência de entrada e de saída com ela mesma e, por isso, o coeficiente de correlação é igual a 1.

Figura 10 – Relação entre a potência total de entrada e a potência total de saída obtidas por meio da soma, em unidades lineares, das potências dos canais.



fonte: Produção da própria autora.

Figura 11 – Gráfico da matriz de correlação cruzada entre as potências de entrada e de saída obtidas através das simulações.



fonte: Produção da própria autora.

Para facilitar a análise da [Figura 11](#), é possível dividir o gráfico em quatro quadrantes. Considerando que cada quadrante é composto por 16 quadrados menores, nota-se que a relação entre todas as potências

de entrada (P_{in1} , P_{in2} , P_{in3} , P_{in4}) pode ser observada no quadrante superior esquerdo. Analisando as cores dos quadrados e os valores do coeficiente de correlação, verifica-se que a correlação das potências de entrada entre diferentes canais é aproximadamente nula. Isso já era esperado, pois os níveis de potência de entrada foram obtidos usando variáveis aleatórias independentes durante a simulação. Analisando a relação entre as potências de saída (P_{out1} , P_{out2} , P_{out3} , P_{out4}) por meio do quadrante inferior direito, nota-se que, diferentemente das potências de entrada, elas não estão igualmente relacionadas e apresentam uma pequena dependência entre si. Examinando as diagonais dos outros dois quadrantes restantes é perceptível que as potências de entrada de cada canal estão muito relacionadas com as potências de saída. É possível identificar que os canais 1 e 2 apresentam uma correlação superior a 0,9, enquanto os canais 3 e 4 tem uma correlação significativamente menor, indicando que estes últimos são mais comprometidos pela modulação de ganho cruzado. Por fim, observando os quadrados fora dessas diagonais, nota-se que a correlação entre a potência de saída de um determinado canal e a de entrada dos outros canais apresenta valores negativos, o que ocorre devido à competição dos sinais pelos íons de érbio excitados na fibra. Assim, ao aumentar a potência de um canal, mais eventos de emissão estimulada são causados, diminuindo a quantidade de íons excitados e, conseqüentemente, o ganho experimentado pelos outros canais. Isso significa que, ao utilizar inteligência artificial para modelar com precisão a potência de saída de um canal, é necessário levar em consideração as quatro potências de entrada dos canais simultaneamente.

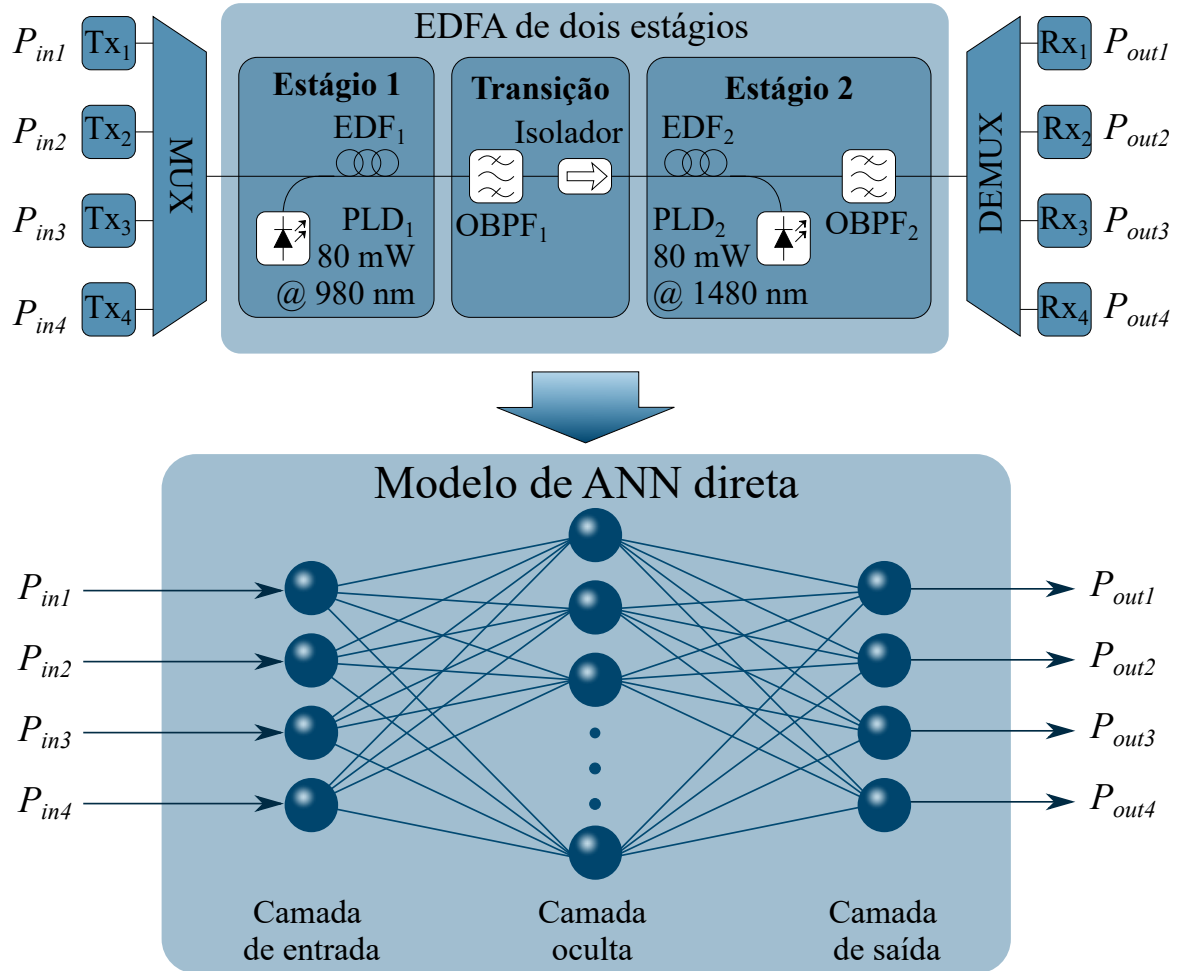
Os modelos de EDFA, que serão mais detalhados na próxima seção deste trabalho, foram implementados em linguagem de alto nível Python, já que ela possui grandes capacidades para aprendizado de máquinas. Para isso, foi utilizado o aplicativo *web open-source* Jupyter Notebook. Além disso, todas as ANNs foram treinadas com ADAM, que é o solucionador definido nas configurações padrão do modelo MLP fornecido pela biblioteca *Scikit-learn*. O solucionador ADAM, proposto por Diederik P. Kingma e Jimmy Lei Ba, é um algoritmo baseado em gradiente estocástico capaz de suportar milhares de amostras de treinamento e ter bons resultados de pontuação de validação (KINGMA; BA, 2015). "O ADAM utiliza a ideia de momento inspirado no conceito físico para definir e adaptar as taxas de aprendizado de cada um dos parâmetros ou pesos durante o processo de otimização." (PEREIRA; CUSTODIO, 2021).

4.2 MODELO DE EDFA BASEADO EM ANN

Assim como já mencionado em seções anteriores, o primeiro método utilizado para encontrar a configuração de potências de entrada em função das potências de saída do amplificador foi treinar e testar um modelo inverso de EDFA baseado em uma ANN. Para a elaboração deste modelo, foi necessário treinar um modelo de EDFA, que serviu como alicerce, e analisar o seu comportamento. As duas abordagens propostas neste trabalho utilizam redes do tipo *feedforward*, *shallow*, *fully connected* e de arquitetura MLP. Esta configuração é simples, mas foi escolhida devido a arquitetura MLP ter um baixo custo computacional. A modelagem inicial foi realizada usando a biblioteca *scikit-learn* (ou *sklearn*), específica para aplicações de *machine learning* (Scikit-learn developers, 2023). Também foram utilizadas as bibliotecas *pandas*, *numpy* e *matplotlib* (Pandas, 2023; NumPy Developers, 2022; The Matplotlib development team, 2022). Este modelo de EDFA baseado em ANN atuou como uma

caixa preta, modelando o EDFA completo composto pelos dois estágios. Dessa forma, as entradas da ANN eram as potências dos canais de entrada do amplificador (P_{in1} , P_{in2} , P_{in3} , P_{in4}), enquanto que as saídas eram as potências de saída (P_{out1} , P_{out2} , P_{out3} , P_{out4}), como pode ser observado na **Figura 12**.

Figura 12 – Esquema do EDFA modelado e o modelo de EDFA baseado em ANN.



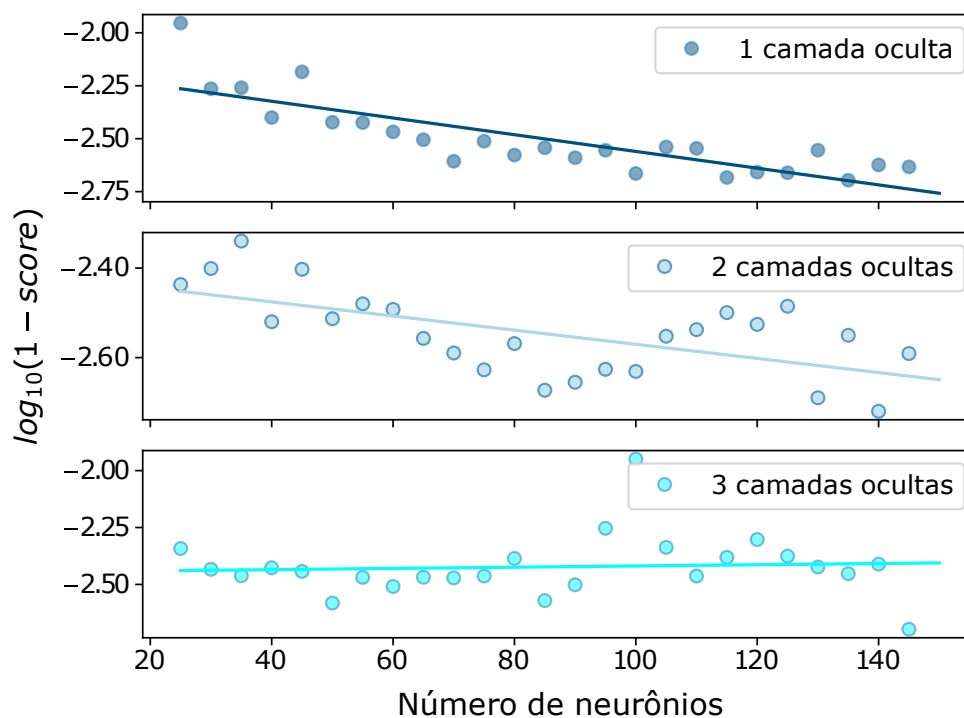
fonte: Produção da própria autora.

Na **Figura 12**, também é possível observar os dois estágios do EDFA modelado compostos pelos trechos de fibras dopadas com érbio (EDF, *Erbium Doped Fiber*), os lasers utilizados para bombeio (PLD, *Pump laser diode*), o Isolador e os filtros ópticos passa faixa (OBPF, *Optical Band Pass Filter*), bem como o multiplexador (MUX, *Multiplexer*) e o demultiplexador (DEMUX, *Demultiplexer*). A ANN foi treinada e testada a partir das amostras de dados de potências de entrada e saída obtidas por meio das simulações citadas na primeira subseção deste capítulo. Logo, para realizar a regressão, foram usadas aproximadamente 5000 configurações no treinamento e 2500 no teste com o objetivo de estimar os níveis de potências de saída do EDFA. Primeiramente, o MLP utilizado tinha configuração padrão com uma única camada oculta com 100 neurônios, que utilizavam a função identidade como função de ativação. Com essa configuração inicial, foram realizadas no máximo 500 iterações utilizando treinamento ADAM e, com a finalidade de obter um modelo otimizado, foi realizada uma análise de validação cruzada.

Para verificar o efeito do número de neurônios na camada oculta, o tamanho da camada foi variado de 25 a 150 neurônios em um passo de 5 neurônios. O número de camadas ocultas também foi variado

considerando 1, 2 e 3 camadas com o intuito de verificar as consequências na otimização da rede. Os resultados desta análise são apresentados na [Figura 13](#), onde é possível observar a pontuação (*score*) da variação do número de neurônios para 1, 2 e 3 camadas ocultas. O *score* é a pontuação média obtida para cada execução da validação cruzada e indica o quão próximos são os resultados preditos dos esperados. Se o *score* é próximo de 1, então a predição realizada foi muito boa.

Figura 13 – $\log_{10}(1 - \text{score})$ em termos do número de neurônios para 1, 2 e 3 camadas ocultas.



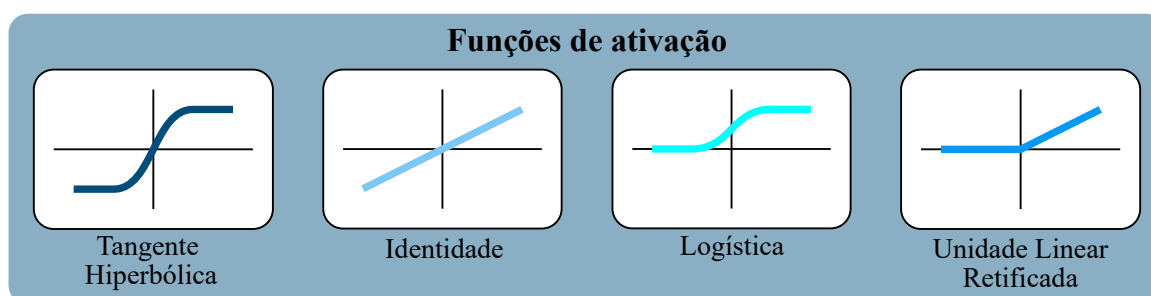
fonte: Produção da própria autora.

Na [Figura 13](#), a escala vertical da figura foi dada por $\log_{10}(1 - \text{score})$ para facilitar a visualização. Assim, os melhores resultados são os que se aproximam do valor $-2,8$, ou seja, possuem um *score* de cerca de 0,998. Também é importante ressaltar que foram traçadas retas de tendência para melhor indicar o comportamento dos resultados. Com apenas uma 1 camada oculta, a medida que o número de neurônios aumentava, melhores as previsões se tornavam, o que não ocorreu com os testes de 2 e 3 camadas. Com 3 camadas, conforme o número de neurônios aumentava, os resultados obtidos pelo modelo pioravam. Além disso, os valores preditos se aproximavam mais lentamente da reta de tendência para 2 camadas do que para 1 camada. Logo, as previsões foram mais precisas por volta de 145 neurônios, considerando 1 camada oculta.

Após este estudo, foi constatado que uma única camada oculta com 145 neurônios era suficiente para modelar o amplificador com precisão. A partir disto, também foi testada a influência das funções de ativação Identidade, Logística, Tangente hiperbólica e Unidade linear retificada, representadas na [Figura 14](#). Assim, o modelo foi testado novamente, mas agora com apenas uma camada e o número de neurônios foi variado de 10 a 150 neurônios para cada uma das funções de ativação testadas. Isso foi necessário para verificar se, ao mudar a função de ativação, a quantidade de 145 neurônios na camada oculta continuava sendo a melhor configuração para o modelo. Os resultados dessa análise são apresentados na [Figura 15](#), onde é possível verificar o *score* obtido pelo modelo treinado com apenas uma camada oculta, mas realizando a variação da quantidade de neurônios e das diferentes funções de

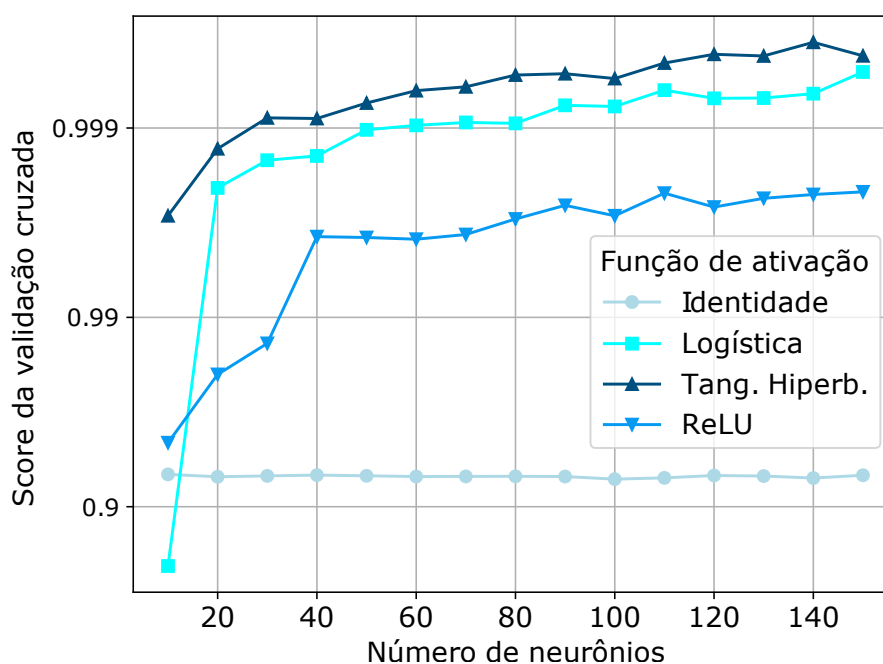
ativação. Comparando as diferentes pontuações obtidas, nota-se claramente que, exceto para a função Identidade que se manteve praticamente sem alteração, quanto maior o número de neurônios, melhor foi a pontuação do modelo. Logo, a partir do estudo realizado, constatou-se que os melhores resultados são realmente obtidos quando o modelo utiliza apenas uma camada oculta com aproximadamente 145 neurônios. Também é possível perceber que a maior pontuação, foi obtida pelo modelo com uma camada oculta de 140 neurônios utilizando a função Tangente Hiperbólica.

Figura 14 – Funções de ativação testadas: Identidade, Logística, Tangente Hiperbólica e Unidade Linear Retificada.



fonte: Produção da própria autora.

Figura 15 – Score da validação cruzada do modelo baseado em ANN conforme o número de neurônios foi variado para cada uma das funções de ativação testadas.

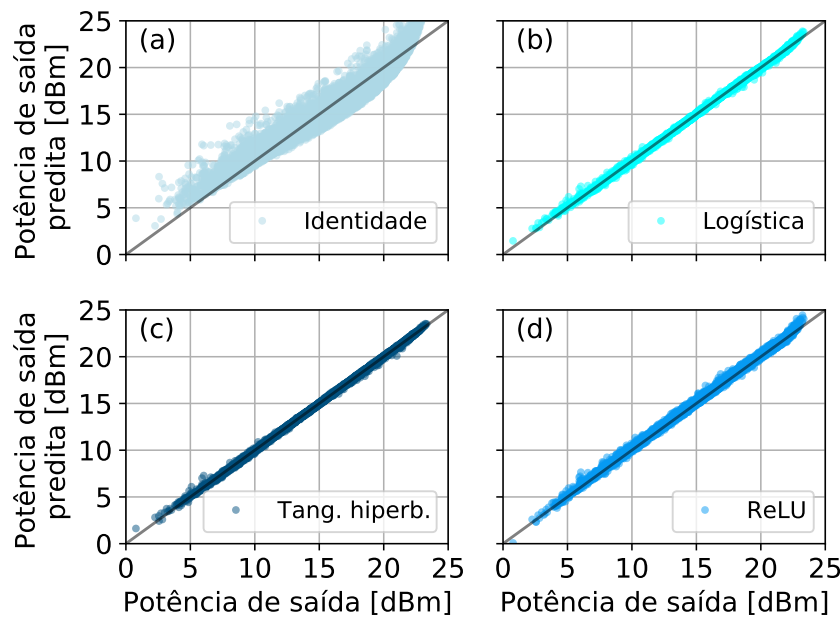


fonte: Produção da própria autora.

Para melhor visualizar e comparar os resultados de um modelo com uma única camada de 140 neurônios, a relação entre as potências de saída esperadas e as potências de saída previstas obtidas para cada uma das funções de ativação utilizadas foi apresentada na [Figura 16](#). No gráfico (a) da [Figura 16](#), pode-se observar que, ao utilizar a função Identidade, houve um grande espalhamento dos dados em relação à reta ideal. É importante notar que esse espalhamento ocorreu também acima da reta e não apenas nas extremidades. Quando os resultados da [Figura 16](#) (d) são analisados, é possível notar que quando foi empregada a função de ativação *Relu*, que é a função padrão do MLP, o desempenho

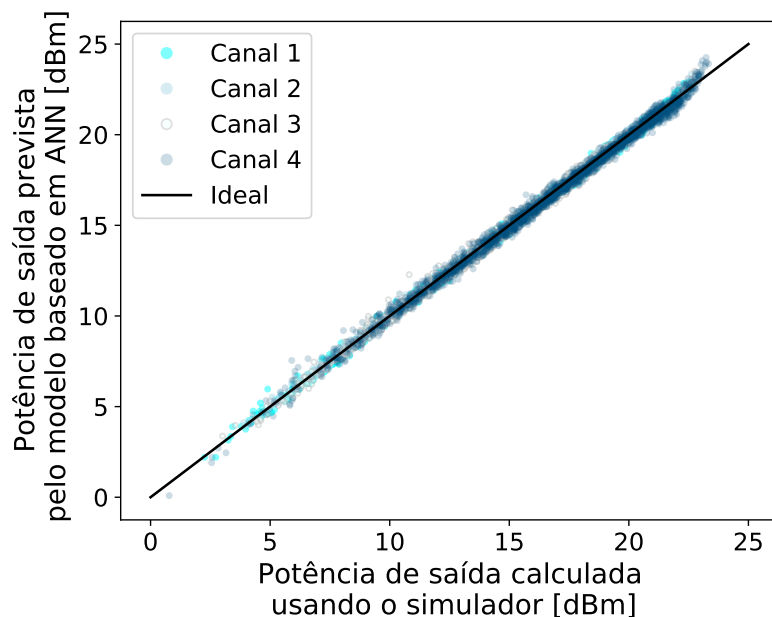
obtido é melhor em relação a função Identidade, mas os resultados ainda apresentam um notável espalhamento nas extremidades dos dados. Por sua vez, na [Figura 16](#) (b) são comparadas as potências após a aplicação da função logística e na [Figura 16](#) (c) após a aplicação da função tangente hiperbólica. Ambas apresentaram resultados muito próximos e melhores que os alcançados pela utilização das funções Identidade e *ReLU*. Porém, ocorreram menos dispersões nas extremidades dos dados quando a função tangente hiperbólica foi utilizada, indicando que esta função obteve os melhores resultados. Por isto, na [Figura 17](#), os resultados da tangente hiperbólica foram destacados.

Figura 16 – Comparação entre as potências de saída desejadas e as potências de saída previstas pelo modelo baseado em ANN utilizando uma única camada de 140 neurônios para cada função de ativação analisada.



fonte: Produção da própria autora.

Figura 17 – Gráfico da relação entre as potências de saída ideais e das potências de saída previstas pelo modelo inicial baseado em ANN para cada canal.



fonte: Produção da própria autora.

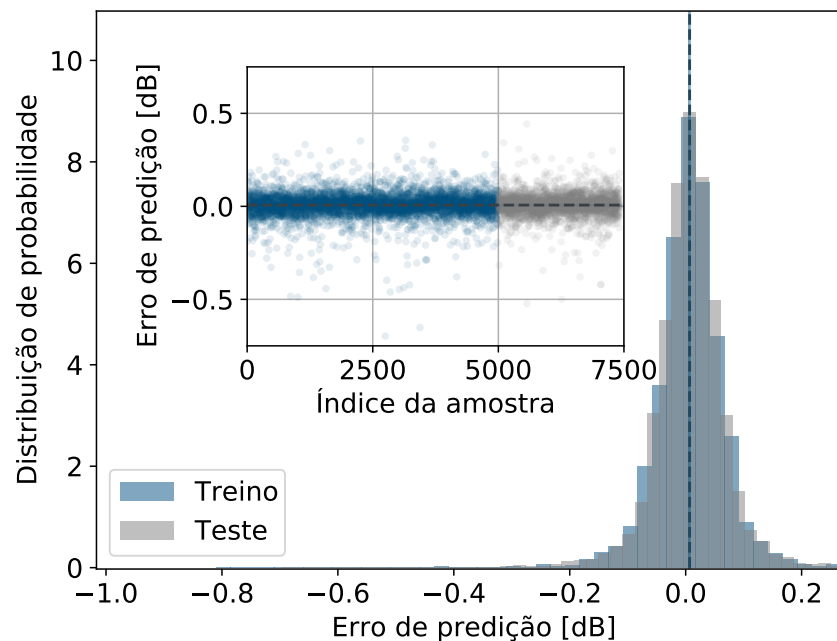
Na [Figura 17](#), é possível notar que os resultados de todos os canais estão muito próximos aos esperados, indicados pela reta ideal. Isso significa que o desempenho do modelo foi bom, porém existe um pequeno deslocamento (*bias*) dos pontos abaixo da reta. Nota-se também que os pontos presentes nas extremidades dos resultados, ou seja, as menores e as maiores potências, apresentam maiores erros. Esse comportamento é comum na regressão, já que os extremos possuem uma menor quantidade de valores e uma dependência maior dos valores dos parâmetros estimados.

Na [Figura 18](#) são apresentados os melhores resultados alcançados pelo modelo com uma única camada de 140 neurônios e função de ativação tangente hiperbólica. A função de custo utilizada para avaliar estes resultados é o Erro Quadrático Médio (*Mean Squared Error*, MSE) dado por

$$MSE = \frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2, \quad (1)$$

onde N é o número de configurações consideradas, Y_i são os valores obtidos por simulação e $\hat{Y}_i$ são os valores previstos pelo modelo. Observando os histogramas de distribuição de erro de predição dos conjuntos de treino (azul) e de teste (cinza), nota-se uma extremidade um pouco mais longa para os valores negativos. Esta assimetria causa uma pequena mudança na média, sendo de 0,032 para treinamento e 0,034 para teste. Além disso é evidente que a maioria dos erros estão concentrados no intervalo entre $-0,2$ dB e $+0,2$ dB.

Figura 18 – Histograma de distribuição de erros de treinamento e predição do modelo direto otimizado, onde o valor médio é indicado em linha descontinua e a predição de erro em termos do número de amostras.



fonte: Produção da própria autora.

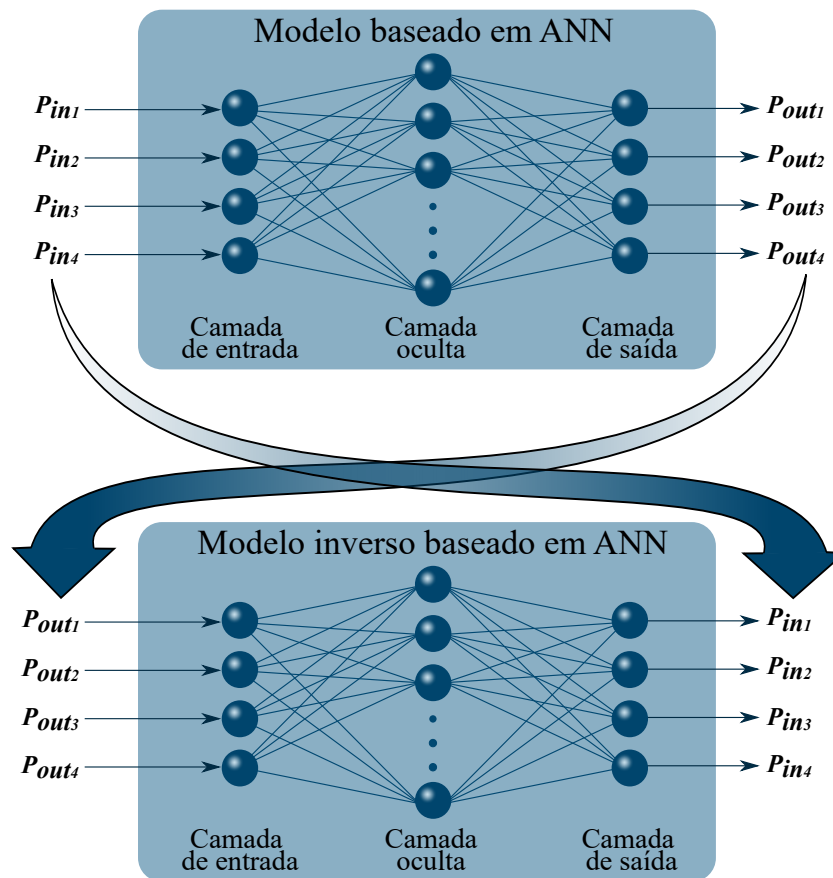
Na [Figura 18](#), o erro também é apresentado em termos do identificador da amostra. Como os histogramas de predição de treino e de teste são muito próximos, conclui-se que o modelo foi treinado corretamente e não apresentou *overfitting*. Além disso, os dados não possuíam ruído, o que contribuiu para a distribuição de erro semelhante dos conjuntos. Estes resultados confirmam que o modelo direto

baseado em ANN é capaz de prever as potências de saída em função das potências de entrada do amplificador, o que também foi constatado em trabalhos de outros autores (YOU; JIANG; JANZ, 2018; YU et al., 2021; HUANG et al., 2017). A partir destes resultados, o próximo passo foi realizar a inversão deste modelo otimizado a fim de encontrar as potências de entrada que proporcionassem as potências de saídas iguais desejadas, que é o objetivo deste trabalho.

4.3 MODELO INVERSO DE EDFA

O modelo inverso de EDFA baseado em ANN, proposto neste trabalho, foi treinado e testado utilizando a configuração de modelo de EDFA que obteve o melhor desempenho, apresentado na Figura 18: arquitetura MLP *feedforward* completamente conectada, com uma única camada oculta com 140 neurônios e função de ativação tangente hiperbólica. Também foram utilizados os mesmos conjuntos de dados de treino e teste obtidos anteriormente. Assim como o modelo baseado em ANN, o modelo inverso também considera o EDFA em sua totalidade, ou seja, os dois estágios. Porém, a abordagem de treino e teste foi invertida, já que agora as potências de saída do EDFA foram empregadas como entradas da ANN e as potências de entrada como saídas. Na Figura 19, há uma comparação entre os modelos.

Figura 19 – Representação do modelo baseado em ANN e o modelo inverso baseado em ANN.

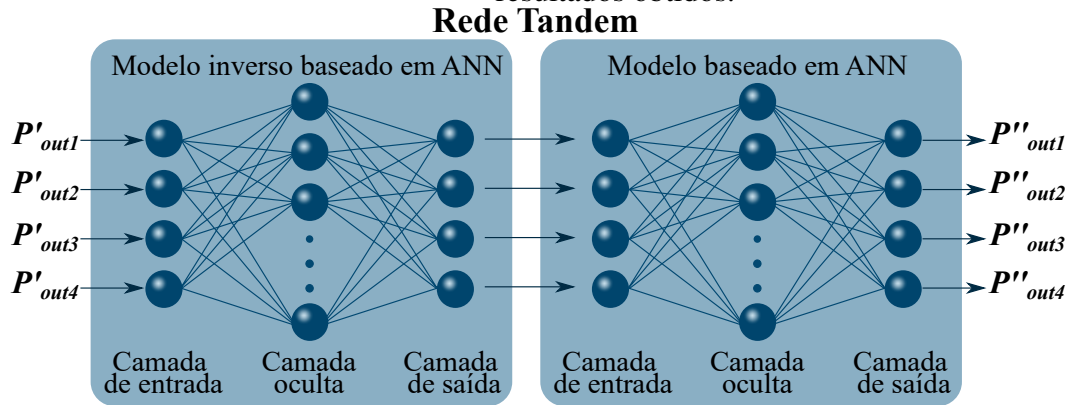


fonte: Produção da própria autora.

Para avaliar o desempenho do modelo inverso, foi necessário unir o modelo de EDFA e o modelo inverso de EDFA, já treinados, formando uma Rede Tandem. A palavra "Tandem" foi utilizada,

pois descrever um conjunto de duas unidades colocadas uma atrás da outra como, por exemplo, uma carruagem que é puxada por dois cavalos, um atrás do outro, ou até mesmo bicicletas de dois lugares (Michaelis, 2024). Como apresentado na Figura 20, P'_{out1} , P'_{out2} , P'_{out3} e P'_{out4} são as potências de saída desejadas e P''_{out1} , P''_{out2} , P''_{out3} e P''_{out4} são as potências de saída encontradas pela Rede Tandem. Os resultados obtidos por meio desta configuração são apresentados no próximo capítulo deste trabalho.

Figura 20 – Combinação do modelo inverso de EDFA e do modelo de EDFA para fins de análise dos resultados obtidos.

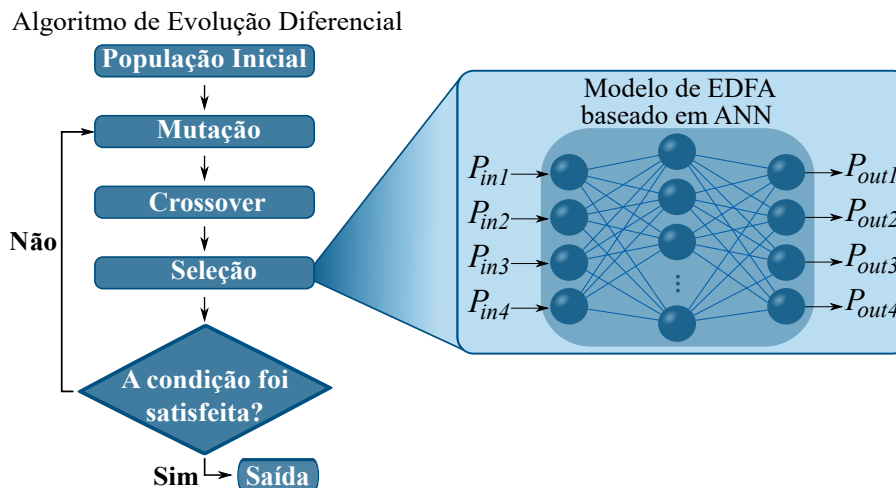


fonte: Produção da própria autora.

4.4 OTIMIZAÇÃO SUBSTITUTA

O segundo método proposto neste trabalho é uma otimização substituta baseada na combinação do modelo de EDFA direto e de um algoritmo multi-objetivo heurístico de otimização com a finalidade de obter as potências de entrada que fornecessem as potências de saída do EDFA desejadas. Entre os diferentes algoritmos heurísticos, optou-se pelo algoritmo de *Differential Evolution* por ser de fácil implementação e também não exigir um grande custo computacional. Neste trabalho, na etapa de seleção, a simulação do EDFA foi substituída pelo modelo baseado em ANN apresentado em seções anteriores. Essa substituição contribui para a redução do custo computacional das avaliações das funções de custo dos indivíduos. A integração da ANN com o DE é apresentada na Figura 21.

Figura 21 – Integração da ANN com o algoritmo DE.



fonte: Produção da própria autora.

5 RESULTADOS

Neste capítulo, serão apresentados e comparados os resultados obtidos pelos dois métodos utilizados para a obtenção das potências de entrada do EDFA que proporcionassem as potências de saída iguais desejadas: a rede neural artificial inversa e a otimização substituta.

5.1 REDE INVERSA

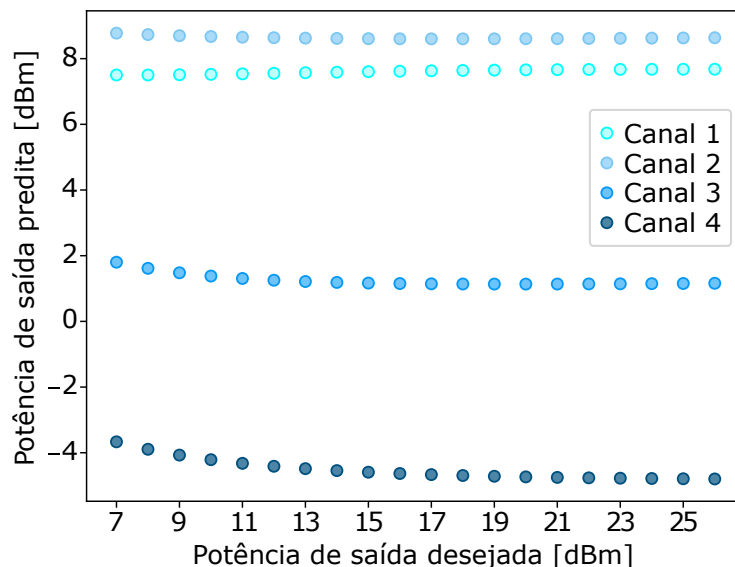
Esta seção do trabalho é dedicada à apresentação e análise dos resultados obtidos pelo primeiro método utilizado a fim de encontrar as potências de entrada do EDFA que fornecessem as potências de saída iguais desejadas. A primeira tentativa de treino e teste foi realizada considerando potências de saída desejadas de 7 dBm a 26 dBm, porém não foram obtidos bons resultados, o que levou à necessidade de realizar uma normalização dos dados. Logo, a função *preprocessing.MinMaxScaler_* da biblioteca *sklearn* foi utilizada para normalizar os dados no intervalo entre [0,1] (Scikit-learn developers, 2023). Um exemplo desta normalização está apresentado na Tabela 1, onde são exibidos os valores de um conjunto de potências antes e depois da normalização. A normalização de dados de entrada e saída pode levar a melhores resultados dependendo da estrutura do modelo. Assim, o modelo inverso pôde ser treinado e testado novamente. Os resultados do modelo inverso de EDFA baseado em ANN, obtidos por meio da Rede Tandem apresentada no capítulo anterior, podem ser observados na Figura 22.

Tabela 1 – Exemplo de configuração de potências do EDFA normalizadas.

	P_{in1}	P_{in2}	P_{in3}	P_{in4}	P_{out1}	P_{out2}	P_{out3}	P_{out4}
Antes da normalização (dBm)	-22,21	-26,26	-15,34	-10,46	10,06	6,68	17,43	22,08
Depois da normalização	0,39	0,19	0,73	0,98	0,43	0,18	0,74	0,94

fonte: Produção da própria autora.

Figura 22 – Resultados do desempenho do modelo inverso baseado em ANN.



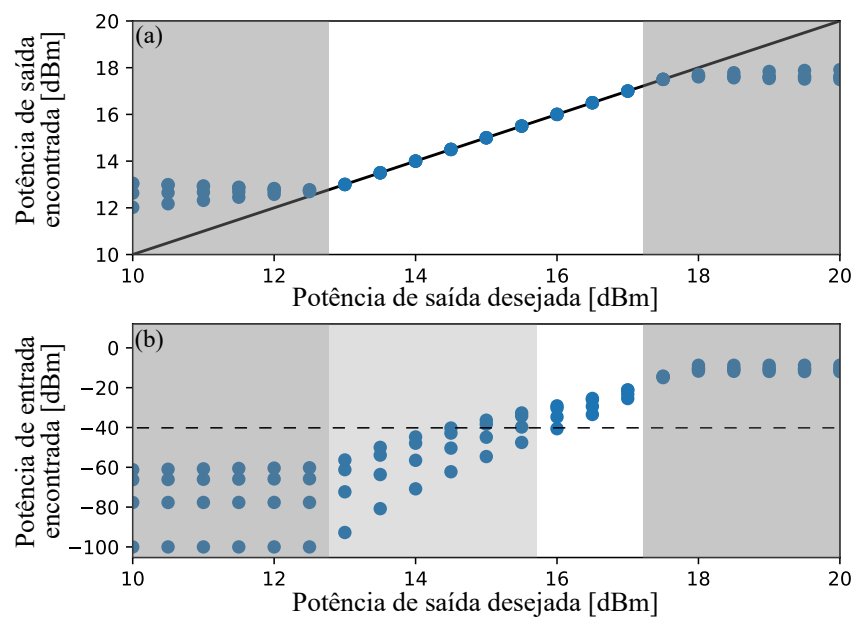
fonte: Produção da própria autora.

Na [Figura 22](#), é apresentada a relação entre potências de saída desejadas e as potências de saída encontradas para cada um dos quatro canais do amplificador. Pode-se concluir que o modelo inverso não apresentou resultados satisfatórios, o que faz sentido devido à falta de bijetividade. Todas as potências encontradas por este modelo eram diferentes das potências desejadas, comportamento que ocorreu em todos os canais.

5.2 OTIMIZAÇÃO SUBSTITUTA

Após a apresentação dos resultados do modelo inverso, nesta seção são avaliados os resultados obtidos pelo segundo método empregado neste trabalho: a otimização substituta. Na [Figura 23](#) (a), é possível observar a relação entre as potências de saída desejadas e as potências de saída encontradas pelo modelo. Os valores que podem ser considerados satisfatórios são aqueles em que os resultados estão sobrepostos à reta ideal (reta em preto). Assim, a combinação entre o DE e o modelo baseado em ANN atingiu os resultados desejados para a faixa de potências de saída desejadas entre 13 dBm e 17 dBm. Nas regiões em cinza escuro da figura, estão os valores que não foram alcançados pela otimização substituta devido ao limite de -100 dBm a 100 dBm imposto às potências de entrada. Já na [Figura 23](#) (b), é apresentada a relação entre as potências de saída desejadas e as potências de entrada encontradas. Nela, os resultados que atendem ao objetivo deste trabalho estão limitados a uma faixa mais estreita, compreendida entre 15,5 dBm e 17 dBm. Nesta faixa estão localizados os resultados de potências de entrada acima da linha tracejada, ou seja, maiores que -40 dBm, valor de potência que é considerado como mínimo para um canal WDM. A área em cinza claro indica os valores de potências de saída desejadas que foram alcançadas pelo algoritmo, mas precisaram de potências de entrada muito baixas para o sistema e as regiões em cinza escuro representam novamente os valores que não foram alcançados pela otimização substituta.

Figura 23 – Resultados obtidos utilizando a otimização substituta.



fonte: Produção da própria autora.

6 CONCLUSÕES

O objetivo deste trabalho era encontrar as potências de entrada de um EDFA, operando no regime de alta modulação de ganho cruzado, que proporcionassem um nível de potências de saída desejadas. Estas potências de entrada deveriam fornecer potências de saída desejadas com o mesmo valor em cada canal, ou seja, os canais deveriam possuir uma potência de 16 dBm cada, por exemplo. Durante o trabalho, os dois métodos propostos inicialmente foram realizados. Em ambos os casos, o primeiro passo foi treinar, testar e otimizar um modelo de EDFA baseado em ANN que serviria de base tanto para o modelo inverso baseado em ANN, quanto para a otimização substituta. O modelo inverso não apresentou bons resultados. Isso se deve ao fato de este modelo ter sido treinado com os dados obtidos por meio do *software* VPI Transmission Maker, o qual é baseado no modelo de Giles-Desurvire. Como o modelo de Giles-Desurvire opera de maneira unidirecional e as relações entre entradas e saídas podem não ser estritamente bijetivas, o modelo inverso baseado em ANN também adotou este comportamento. Por outro lado, a otimização substituta, desempenhou muito bem o seu papel, já que alcançar este objetivo era difícil devido à modulação do ganho cruzado, que também pode ser a responsável pela limitação dos resultados obtidos por este segundo método em uma determinada faixa de potências. Desta forma, este trabalho mostrou que é possível obter um modelo capaz de representar adequadamente um EDFA de dois estágios utilizando a combinação entre uma ANN e o algoritmo de evolução diferencial.

6.1 SUGESTÕES DE TRABALHOS FUTUROS

Já que no atual trabalho foi considerado um EDFA com apenas quatro canais WDM, em um possível trabalho futuro, seria também interessante realizar as mesmas análises levando em conta um EDFA com mais canais. Além disso, neste trabalho foram consideradas todas as saídas desejadas com o mesmo valor. Logo, também seria possível testar a combinação a fim de encontrar as potências de entrada que fornecessem as potências de saída desejadas de valores diferentes de potência.

REFERÊNCIAS

- AGATONOVIC-KUSTRIN, S.; BERESFORD, R. Basic concepts of artificial neural network (ann) modeling and its application in pharmaceutical research. **Journal of Pharmaceutical and Biomedical Analysis**, v. 22, n. 5, p. 717–727, 2000. ISSN 0731-7085. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0731708599002721>.
- AGRAWAL, G. **Fiber-Optic Communication Systems**. Wiley, 2021. ISBN 9781119737360. Disponível em: <https://books.google.com.br/books?id=RkwvEAAQBAJ>.
- AGRAWAL, G. P. Optical communication: its history and recent progress. In: **Optics in our time**. [S.l.]: Springer, Cham, 2016. p. 177–199.
- ALPAYDIN, E. **Introduction to Machine Learning**. [S.l.: s.n.], 2010.
- BARBOZA, E. d. A. et al. Optical amplifier response estimation considering non-flat input signals characterization based on artificial neural networks. **Journal of Lightwave Technology**, v. 39, n. 1, p. 208–215, Jan. 2021.
- BASTOS-FILHO, C. J. A.; BARBOZA, E. d. A.; MARTINS-FILHO, J. F. Estimating the spectral gain and the noise figure of EDFA using artificial neural networks. In: **19th International Conference on Transparent Optical Networks (ICTON)**. Girona, Spain: [s.n.], 2017. p. 1–4.
- DESURVIRE, E.; ZERVAS, M. N. Erbium-doped fiber amplifiers: principles and applications. **Physics Today**, v. 48, n. 2, p. 56, 1995.
- ERICSSON. **Ericsson Mobility Report November**. 2023. Disponível em: <https://www.ericsson.com/4ae12c/assets/local/reports-papers/mobility-report/documents/2023/ericsson-mobility-report-november-2023.pdf>.
- FILHO, J. C. P. et al. Estimating amplifier cascade output signal using an artificial neural network and considering tilted signals. In: **SBFoton International Optics and Photonics Conference (SBFoton IOPC)**. Recife, Brazil: [s.n.], 2022. p. 1–5.
- FINE, T. **Feedforward Neural Network Methodology**. Springer New York, 2006. (Information Science and Statistics). ISBN 9780387226491. Disponível em: <https://books.google.com.br/books?id=s-PIBwAAQBAJ>.
- FREEMAN, J.; CONRADI, J. Gain modulation response of erbium-doped fiber amplifiers. **IEEE Photonics Technology Letters**, IEEE, v. 5, n. 2, p. 224–226, 1993.
- GERSHENSON, C. Artificial neural networks for beginners. **CoRR**, cs.NE/0308031, 2003. Disponível em: <http://arxiv.org/abs/cs/0308031>.
- GILES, C. R.; DESURVIRE, E. Modeling erbium-doped fiber amplifiers. **Journal of Lightwave Technology**, IEEE, v. 9, n. 2, p. 271–283, 1991.
- HUANG, Y. et al. Dynamic mitigation of EDFA power excursions with machine learning. **Optics Express**, Optica Publishing Group, v. 25, n. 3, p. 2245–2258, Feb 2017. Disponível em: <https://opg.optica.org/oe/abstract.cfm?URI=oe-25-3-2245>.
- KEISER, G. **Comunicações por Fibras Ópticas**. [S.l.]: Grupo A, 2014.

- KINGMA, D. P.; BA, J. L. Adam: A method for stochastic optimization. In: **3rd International Conference for Learning Representations**. [s.n.], 2015. Disponível em: <https://doi.org/10.48550/arXiv.1412.6980>.
- LATHI B. P. E DING, Z. **Sistemas de Comunicações Analógicos e Digitais Modernos**. [S.l.]: Grupo GEN, 2012.
- LIN, J.; LIN, X.; JIANG, Z. Auxiliary neural network assisted machine learning EDFA gain model. In: **Optical Fiber Communications Conference and Exhibition (OFC)**. San Diego, CA, USA: [s.n.], 2023. p. 1–3.
- MENIF, M. et al. Cross-gain modulation effect on the behaviour of packetized cascaded EDFAs. **Journal of Optics A: Pure and Applied Optics**, IOP Publishing, v. 3, n. 3, p. 210, 2001.
- MESEGUER, A. C. et al. Highly accurate measurement-based gain model for constant-pump EDFA for non-flat WDM inputs. In: **Optical Fiber Communications Conference and Exhibition (OFC)**. San Francisco, CA, USA: [s.n.], 2021. p. 1–3.
- Michaelis. **Tandem**. 2024. Acessado em: 2 jul. 2024. Disponível em: <https://michaelis.uol.com.br/busca?r=0&f=0&t=0&palavra=Tandem+>.
- NumPy Developers. **NumPy documentation**. 2022. Disponível em: <https://numpy.org/doc/stable/>.
- Pandas. **About pandas**. 2023. Disponível em: <https://pandas.pydata.org/about/>.
- PEREIRA, G. C.; CUSTODIO, R. Redes neurais para regressão uni- e multivariada. **Revista Chemkeys**, v. 3, Ago. 2021.
- QING, A. **Differential evolution: fundamentals and applications in electrical engineering**. [S.l.]: John Wiley & Sons, 2009.
- QIU, S.; XU, X.; CAI, B. Frelu: Flexible rectified linear units for improving convolutional neural networks. In: **24th International Conference on Pattern Recognition (ICPR)**. [S.l.: s.n.], 2018. p. 1223–1228.
- RASAMOELINA, A. D.; ADJAILIA, F.; SINCAK, P. A review of activation function for artificial neural network. In: **IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)**. [S.l.: s.n.], 2020. p. 281–286.
- ROS, F. da; MOURA, U. C. de; YANKOV, M. P. Machine learning-based EDFA gain model generalizable to multiple physical devices. In: **European Conference on Optical Communications (ECOC)**. Brussels, Belgium: [s.n.], 2020. p. 1–4.
- Scikit-learn developers. **User guide**. 2023. Disponível em: https://scikit-learn.org/stable/user_guide.html.
- Scikit-learn developers. **Neural network models (supervised)**. 2024. Disponível em: https://scikit-learn.org/stable/modules/neural_networks_supervised.html.
- SHARKAWY, A.-N. Principle of neural network and its main types. **Journal of Advances in Applied & Computational Mathematics**, v. 7, p. 8–19, 2020.
- The Matplotlib development team. **matplotlib.pyplot**. 2022. Disponível em: https://matplotlib.org/3.5.3/api/_as_gen/matplotlib.pyplot.html.
- VARELLA, M. **Neurônio**. 2024. Disponível em: <https://drauziovarella.uol.com.br/corpo-humano/neuronio/>.

WANG, Z.; KILPER, D.; CHEN, T. Transfer learning-based roadm EDFA wavelength dependent gain prediction using minimized data collection. In: **Optical Fiber Communications Conference and Exhibition (OFC)**. San Diego, CA, USA: [s.n.], 2023. p. 1–3.

YOU, Y.; JIANG, Z.; JANZ, C. Machine learning-based EDFA gain model. In: **European Conference on Optical Communication (ECOC)**. [S.l.: s.n.], 2018. p. 1–3.

YU, J. et al. Machine-learning-based EDFA gain estimation [invited]. **Journal of Optical Communications and Networking**, v. 13, n. 4, p. B83–B91, 2021.

ZHU, S. et al. Machine learning based prediction of erbium-doped fiber WDM line amplifier gain spectra. In: **European Conference on Optical Communication (ECOC)**. Rome, Italy: [s.n.], 2018. p. 1–3.