



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS FLORIANÓPOLIS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE AUTOMAÇÃO E
SISTEMAS

Rodrigo Luiz da Cruz Schwartz

**Identificação de Sistemas Orientada para o Controle do Tratamento de Câncer por
Hipertermia Magnética com Redes Neurais de Estado de Eco**

Florianópolis

2023

Rodrigo Luiz da Cruz Schwartz

**Identificação de Sistemas Orientada para o Controle do Tratamento de Câncer por
Hipertermia Magnética com Redes Neurais de Estado de Eco**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Automação e Sistemas da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Engenharia de Automação e Sistemas. Orientador: Prof. Rodolfo César Costa Flesch, Dr. Eng. Coorientador: Prof. Bernardo Barancelli Schwedersky, Dr. Eng.

Florianópolis

2023

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Sehwartz, Rodrigo Luiz da Cruz

Identificação de Sistemas Orientada para o Controle do Tratamento de Câncer por Hipertermia Magnética com Redes Neurais de Estado de Eco / Rodrigo Luiz da Cruz Sehwartz ; orientador, Rodolfo César Costa Flesch, coorientador, Bernardo Barancelli Schwedersky, 2023.

94 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia de Automação e Sistemas, Florianópolis, 2023.

Inclui referências.

1. Engenharia de Automação e Sistemas. 2. Rede de estado de eco. 3. Identificação de sistemas. 4. Hipertermia magnética. 5. Sistemas não lineares. I. Flesch, Rodolfo César Costa. II. Schwedersky, Bernardo Barancelli. III. Universidade Federal de Santa Catarina. Programa de Pós Graduação em Engenharia de Automação e Sistemas. IV. Título.

Rodrigo Luiz da Cruz Schwartz

**Identificação de Sistemas Orientada para o Controle do Tratamento de Câncer por
Hipertermia Magnética com Redes Neurais de Estado de Eco**

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Gustavo Artur de Andrade, Dr. Eng.
Universidade Federal de Santa Catarina

Prof. Jefferson Luiz Brum Marques, Dr. Eng.
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** da dissertação que foi julgada adequada para obtenção do título de Mestre em Engenharia de Automação e Sistemas.

Coordenação do Programa de Pós-Graduação

Prof. Rodolfo César Costa Flesch, Dr. Eng.
Orientador

Florianópolis, 2023.

Este trabalho é dedicado aos meus pais, aos meus sobrinhos e
à minha eterna avó, Ana da Rocha Veiga.

AGRADECIMENTOS

Primeiramente, gostaria de agradecer a Deus por ter me dado forças durante os momentos difíceis e por ter me apresentado a pessoas que foram de grande ajuda no desenvolvimento deste trabalho, seja de forma direta ou indireta.

Gostaria de destacar os meus agradecimentos aos meus pais, Armando e Edna, ao meu irmão, Danilo, e aos meus sobrinhos, Davi e Daniel, pelo suporte, pelos momentos felizes e por terem me ensinado a perseverar nos momentos de maiores dificuldades.

Agradeço também às minhas tias, Maria Cristina e Maria José Malheiros, por sempre estarem presentes e pelo enorme apoio. Sem elas, muitos dos meus sonhos seriam impossíveis de serem realizados.

Sou também grato aos meus amigos, José Cláudio, Paulo Patez, Marina Sinimbu, Délis Palheta, Ian Castro e Grazielle Castilho. Mesmo eu estando longe, todos foram compreensivos, estiveram presentes nos momentos difíceis e acreditaram no meu potencial.

Destaco também o meu agradecimento ao meu amigo Eric Mochiutti por ter me apoiado nas grandes mudanças que ocorreram nos últimos dois anos deste mestrado e pela enorme ajuda durante o desenvolvimento desta dissertação.

Agradeço aos professores orientadores, Rodolfo César Costa Flesch e Bernardo Barancelli Schwedersky, por serem extremamente compreensivos com as minhas dificuldades, pela enorme atenção e pelos ensinamentos que contribuíram para o desenvolvimento deste trabalho.

Agradeço à Universidade Federal de Santa Catarina e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pela estrutura e pelo suporte financeiro que foram, sem sombra de dúvidas, um dos pilares para a conclusão deste trabalho.

Em memória da minha amada avó Ana da Rocha Veiga, que embora tenha partido desta vida, sempre me ajudou e demonstrou amor e afeto sem igual. Deixo aqui minha eterna saudade e agradeço por ter sido a melhor avó do mundo.

RESUMO

Um modelo de processo é usualmente necessário para o projeto de um controlador, especialmente quando se trabalha com controladores preditivos baseados em modelo (MPCs). No contexto da hipertermia por fluidos magnéticos (MFH), a criação de um modelo baseado em leis físicas para prever o aumento da temperatura do tumor durante o tratamento depende de equações diferenciais parciais não lineares. O maior desafio para a utilização de tais modelos para fins de predição em MPC é o tempo significativo que é requerido para obter soluções numéricas para essas equações. Como uma abordagem alternativa, este trabalho propõe um modelo de predição rápido das temperaturas na região do tumor durante um tratamento de MFH, com erros de predição aceitáveis para que seja utilizado como um modelo de predição em uma implementação de MPC. O método proposto baseia-se no uso de uma rede neural de estado de eco (ESN) para modelar e prever temperaturas em quatro pontos de medição no tumor. A ESN foi treinada usando dados gerados a partir de três modelos numéricos e considerou a frequência e a amplitude do campo magnético como entradas. Os resultados mostram que a ESN proposta obteve um erro quadrático médio normalizado inferior a 5% para todas as temperaturas avaliadas nos conjuntos de teste. Além disso, o tempo computacional necessário para estimar valores de temperatura para os quatro pontos de medição em um horizonte com mais de 1000 amostras foi de até 263 ms. Isso representa um tempo cerca de três ordens de magnitude menor que o necessário para resolver o modelo numérico para as mesmas predições.

Palavras-chave: Rede neurais de estado de eco. Hipertermia por fluidos magnéticos. Identificação de sistemas. Sistemas não lineares.

ABSTRACT

A process model is usually required for controller design, particularly when working with model predictive controllers (MPCs). In the context of magnetic fluid hyperthermia (MFH), creating a first-principles model to predict tumor temperature rise during treatment relies on nonlinear partial differential equations. The major challenge for the use of such models for prediction in MPC is the significant time required to obtain numerical solutions for these equations. As an alternative approach, this work proposes a fast and accurate prediction model for the temperatures within the tumor region during an MFH treatment which is suitable for being used as a prediction model in an MPC implementation. The proposed method is based on the use of an echo state neural network (ESN) to model and estimate temperatures at four measurement points in the tumor. The ESN was trained using data generated from three numerical models and considered the magnetic field frequency and amplitude as inputs. The results demonstrate that the proposed ESN achieved a normalized mean squared error of less than 5% for all temperatures evaluated in the test datasets. Furthermore, the computational time required to estimate temperature values for all measurement points over a horizon of more than 1000 samples was less than 263 ms. This represents a significant improvement, as it is more than three orders of magnitude faster compared to the time required to solve the numerical model for the same predictions.

Keywords: Echo state networks. Magnetic hyperthermia. System identification. Nonlinear systems.

LISTA DE FIGURAS

Figura 1 – Etapas do tratamento de hipertermia magnética.	30
Figura 2 – Esboço de um protótipo de um sistema de hipertermia local por fluidos magnéticos.	32
Figura 3 – Representação do ciclo de histerese de um material magnético.	33
Figura 4 – Mecanismo de perdas por relaxação de Brown e de Néel.	34
Figura 5 – Representação do funcionamento de um neurônio artificial.	40
Figura 6 – Representação esquemática de uma rede neural MLP.	41
Figura 7 – Representação de uma rede neural recorrente com neurônios ocultados.	42
Figura 8 – Representação geral da arquitetura de uma rede neural de estado de eco.	43
Figura 9 – Exemplos de sinais de excitação: (A) PRBS; e (B) APRBS.	51
Figura 10 – Geometrias dos modelos numéricos: (A) esférico; (B) elipsoidal oblato; e (C) elipsoidal prolato.	53
Figura 11 – Verificação do modelo de transporte intersticial do nanofluido com os resultados da literatura: (A) pressão intersticial; e (B) velocidade intersticial.	64
Figura 12 – Verificação do modelo de convecção-difusão com os resultados da literatura	65
Figura 13 – Verificação do modelo de biotransferência de calor com os resultados da literatura: (A) perfil radial de distribuição temperatura; e (B) perfil transiente de temperatura.	66
Figura 14 – Sinais APRBS: (A) frequência de operação; e (B) amplitude do campo magnético.	68
Figura 15 – Respostas do processo aos sinais de entrada: (A) modelo esférico; (B) modelo oblato; e (C) modelo prolato.	70
Figura 16 – Minimização do NMSE a partir da otimização bayesiana no caso do modelo de geometria esférica.	72
Figura 17 – Minimização do NMSE a partir da otimização bayesiana no caso do modelo de geometria elipsoidal oblata.	72
Figura 18 – Minimização do NMSE a partir da otimização bayesiana no caso do modelo de geometria elipsoidal prolata.	73
Figura 19 – Influência do tamanho do reservatório sobre a ESN: (A) NMSE médio; e (B) tempo de execução	74
Figura 20 – Predições para temperaturas na região do tumor no modelo esférico: (A) ponto I; (B) ponto II; (C) ponto III; e (D) ponto IV.	77
Figura 21 – Comparação entre o NMSE e o tempo de execução para cada modelo no conjunto de teste do modelo esférico.	78

Figura 22 – Predições para temperaturas na região do tumor no modelo elipsoidal oblato: (A) ponto I; (B) ponto II; (C) ponto III; e (D) ponto IV.	79
Figura 23 – Comparação entre o NMSE e o tempo de execução para cada modelo no conjunto de teste do modelo oblato.	81
Figura 24 – Predições para temperaturas na região do tumor no modelo elipsoidal prolato: (A) ponto I; (B) ponto II; (C) ponto III; e (D) ponto IV.	82
Figura 25 – Comparação entre o NMSE e o tempo de execução para cada modelo no conjunto de teste do modelo prolato.	83

LISTA DE TABELAS

Tabela 1	– Valores para as propriedades pertencentes à física de escoamento laminar. . .	56
Tabela 2	– Valores para as propriedades pertencentes à física de transporte intersticial de nanofluido.	58
Tabela 3	– Valores utilizados para as propriedades dos domínios durante a simulação do modelo de transporte do nanofluido.	61
Tabela 4	– Valores utilizados para as propriedades dos materiais durante a simulação do modelo de biotransferência de calor.	63
Tabela 5	– Coordenadas, em milímetros, ao longo do eixo radial e vertical de cada ponto de medição.	69
Tabela 6	– Valores do NMSE para a ESN proposta e os dois modelos de referência na tarefa de predição da temperatura nos pontos I, II, III e IV na região do tumor esférico.	78
Tabela 7	– Valores do NMSE para a ESN proposta e os dois modelos de referência na tarefa de predição da temperatura nos pontos I, II, III e IV na região do tumor elipsoidal oblato.	80
Tabela 8	– Valores do NMSE para a ESN proposta e os dois modelos de referência na tarefa de predição da temperatura nos pontos I, II, III e IV na região do tumor elipsoidal prolato.	83

LISTA DE ABREVIATURAS E SIGLAS

MFH	hipertermia por fluidos magnéticos – do inglês <i>Magnetic Fluid Hyperthermia</i>
AMF	campo magnético alternado – do inglês <i>Alternating Magnetic Field</i>
PI	controlador Proporcional-Integral
PID	controlador Proporcional-Integral-Derivativo
MPC	controle preditivo baseado em modelo – do inglês <i>Model Predictive Control</i>
ESN	rede neural de estado de eco – do inglês <i>Echo State Network</i>
MIMO	sistema com múltiplas entradas e múltiplas saídas – do inglês <i>Multiple Input and Multiple Outputs</i>
MNP	nanopartícula magnética – do inglês <i>Magnetic NanoParticle</i>
RNA	Rede Neural Artificial
SNC	Sistema Nervoso Central
MLP	perceptron de múltiplas camadas – do inglês <i>MultiLayer Perceptron</i>
RNN	rede neural recorrente – do inglês <i>Recurrent Neural Network</i>
LSTM	redes de memória longa de curto prazo – do inglês <i>Long Short-Term Memory</i>
LSM	máquina de estado líquido – do inglês <i>Liquid State Machine</i>
RC	computação em reservatório – do inglês <i>Reservoir Computing</i>
GP	processo gaussiano – do inglês <i>Gaussian Process</i>
ARD	determinação automática de relevância – do inglês <i>Automatic Relevance Determination</i>
PRBS	sequência binária pseudo-aleatória – do inglês <i>Pseudo Random Binary Sequence</i>
APRBS	sequência binária pseudo-aleatória modulada em amplitude – do inglês <i>Amplitude Modulated Random Binary Sequence</i>
NMSE	erro médio quadrático normalizado – do inglês <i>Normalized Mean Squared Error</i>

FT	Função de Transferência
HW	modelo de Hammerstein-Wiener

LISTA DE SÍMBOLOS

B_r	remanescência
H_c	campo coercitivo
μ	momento magnético
τ_N	tempo de relaxação de Néel de uma nanopartícula
τ_B	tempo de relaxação de Brown de uma nanopartícula
μ_{nf}	viscosidade dinâmica do nanofluido
V_H	volume hidrodinâmico de uma nanopartícula
k_B	constante de Boltzmann
T_i	temperatura na região i
τ_0	tempo característico do material
V_M	volume de uma nanopartícula
K_a	constante de anisotropia de uma nanopartícula
τ	tempo de relaxação efetivo de uma nanopartícula
B	densidade do fluxo magnético
U	energia interna do sistema
M	magnetização de um material
μ_0	constante de permeabilidade do vácuo
ΔU	variação da energia interna do material
$H(t)$	intensidade do campo magnético ao longo do tempo
H_0	amplitude máxima do campo magnético
ω	frequência angular
t	tempo
χ	suscetibilidade complexa do nanofluido magnético
χ'	componente em fase da suscetibilidade complexa do nanofluido magnético
χ''	componente fora de fase da suscetibilidade complexa do nanofluido magnético
χ_0	suscetibilidade de equilíbrio do nanofluido magnético

ϕ_i	fração de nanopartículas no meio i
M_d	magnetização do domínio de uma nanopartícula suspensa
P	potência dissipada por uma nanopartícula em função de um campo magnético externo
y_k	saída do k -ésimo neurônio artificial
$g(\cdot)$	função de ativação do neurônio artificial
x_k	saída da junção aditiva do k -ésimo neurônio artificial
w_{ki}	i -ésimo peso sináptico associado a uma entrada do k -ésimo neurônio artificial
u_{ki}	i -ésima entrada do k -ésimo neurônio artificial
b_k	<i>bias</i> do k -ésimo neurônio artificial
n	contador do tempo discreto
$\mathbf{x}[n]$	vetor com os estados do reservatório da ESN no instante n
N_x	número de neurônios presentes no reservatório da ESN
α	taxa de vazamento do reservatório da ESN
$\tilde{\mathbf{x}}[n]$	vetor com a atualização dos estados do reservatório da ESN no instante n
$\mathbf{u}[n]$	vetor com sinais de entrada no instante n
N_u	número de entradas da ESN
\mathbf{W}^{in}	matriz de pesos que conecta a entrada da ESN até o reservatório
\mathbf{W}^{res}	matriz dos pesos relacionada às conexões recorrentes dentro do reservatório
s^{in}	fator de escalonamento da matriz \mathbf{W}^{in}
$\hat{\mathbf{y}}[n]$	vetor com sinais de saída preditos no instante n
N_y	número de saídas da ESN
\mathbf{W}^{out}	matriz de pesos da saída
$\lambda(\mathbf{W}^{\text{res}})$	raio espectral da matriz \mathbf{W}^{res}
N_{treino}	tamanho do conjunto de dados para o treinamento de um modelo
$\mathbf{Y}^{\text{treino}}$	matriz que contém todos os sinais de saída para treinar a ESN
$\mathbf{U}^{\text{treino}}$	matriz que contém todos os sinais de entrada para treinar a ESN

$\mathbf{X}^{\text{treino}}$	matriz que contém os estados do reservatório durante o treinamento
\mathbf{M}	matriz que concatena o <i>bias</i> , os sinais de entrada e os estados do reservatório da ESN
$\mathbf{w}_i^{\text{out}}$	vetor que contém a <i>i</i> -ésima linha da matriz \mathbf{W}^{out}
$\hat{y}_i[n]$	<i>i</i> -ésima saída predita de um modelo para o instante <i>n</i>
$y_i^{\text{treino}}[n]$	<i>i</i> -ésima saída real de um sistema para o instante <i>n</i> no conjunto de treinamento
\mathbf{I}	matriz identidade
$\ \cdot\ $	norma euclidiana de um vetor
β	coeficiente de regularização
$J(\cdot)$	função objetivo a ser minimizada pela otimização bayesiana
\mathbf{z}	vetor com as entradas para a função $J(\cdot)$
$m(\cdot)$	função média do processo gaussiano
$\text{cov}(\cdot, \cdot)$	função de covariância do processo gaussiano
σ_{cov}^2	amplitude da função de covariância
$d^2(\cdot, \cdot)$	distância entre dois pontos no espaço das entradas da função $J(\cdot)$
ℓ_m	escala de comprimento da medida de distância
$\text{PI}(\cdot)$	função de aquisição da probabilidade de melhoria
$\mu_{\text{GP}}(\cdot)$	média da distribuição gaussiana em torno de um ponto
$\sigma_{\text{GP}}^2(\cdot)$	variância da distribuição gaussiana em torno de um ponto
\mathbf{z}^*	melhor ponto calculado até a presente iteração pelo processo gaussiano
ξ	hiperparâmetro das funções de aquisição de probabilidade de melhoria e de melhoria esperada
$\Phi(\cdot)$	função de distribuição acumulada normal
$\text{EI}(\cdot)$	função de aquisição da melhoria esperada
$\varphi(\cdot)$	função de distribuição de probabilidade
t_h	tempo mínimo de duração para um patamar do sinal de excitação
τ_{min}	menor constante de tempo de um sistema
\mathbf{v}_{nf}	velocidade do nanofluido ao longo da agulha

ρ_{nf}	massa específica do nanofluido
p_{nf}	pressão do nanofluido ao longo da agulha
$\phi_{m\acute{a}x}$	fração de nanopartículas no nanofluido magnético
ρ_{MNP}	massa específica da nanopartícula magnética
ρ_s	massa específica do solvente
∇	gradiente de uma variável em coordenadas cilíndricas
p_0	pressão intersticial do tumor no local de injeção
\mathbf{n}	Vetor normal à superfície
\mathbf{v}_i	velocidade intersticial do nanofluido na região i
κ_i	condutividade hidráulica da região i
p_i	pressão intersticial do nanofluido na região i
θ_V	vazão volumétrica para fora da vasculatura por unidade de volume do tecido
L_p	condutividade hidráulica da parede microvascular
S/V	área de superfície por unidade de volume para transporte no tumor
p_V	pressão vascular
σ_T	coeficiente médio de reflexão osmótica para as proteínas plasmáticas
π_V	pressão osmótica do plasma sanguíneo
π_i	pressão osmótica do fluido intersticial na região i
L_{pL}	condutividade hidráulica da parede linfática
θ_L	vazão volumétrica para dentro dos vasos linfáticos
p_L	pressão hidrostática dos vasos linfáticos
Ψ	interface entre o núcleo necrótico do tumor e sua zona viável
Υ	interface entre a zona viável do tumor e o órgão saudável
r_{agulha}	raio da agulha
q_{nf}	vazão volumétrica de injeção do nanofluido
Ξ	limite do órgão saudável
C_i	concentração de nanopartículas na região i

D_{eff}	coeficiente de difusão das nanopartículas na matriz porosa
Θ_V	vazão volumétrica de soluto entre o sistema vascular e a região analisada
Θ_L	vazão volumétrica de soluto entre o sistema linfático e a região analisada
D_0	coeficiente de difusão das nanopartículas na fase líquida
T_{abs}	Temperatura absoluta
r_{MNP}	raio da nanopartícula magnética
ε	porosidade do órgão
σ_s	coeficiente de reflexão do soluto
C_p	concentração de soluto no plasma sanguíneo
P_c	permeabilidade dos vasos capilares
Pe	número de Péclet
ε_c	porosidade da parede dos vasos capilares
γ	razão entre o valor do raio da nanopartícula e o raio dos poros dos vasos capilares
r_p	raio dos poros dos vasos capilares
L_c	espessura da parede dos vasos capilares
$C_{\text{máx}}$	concentração de nanopartículas no nanofluido magnético
M_{MNP}	massa molar da nanopartícula magnética
ρ_b	densidade do sangue
c_b	calor específico do sangue
Ω_b	taxa de perfusão sanguínea
T_b	temperatura do sangue
η	fator de correção para a potência dissipada por uma nanopartícula
Q_{met}	calor gerado em função do metabolismo
f	frequência de operação do campo magnético externo
c_{MNP}	calor específico da nanopartícula magnética
k_{MNP}	condutividade térmica da nanopartícula magnética

SUMÁRIO

1	INTRODUÇÃO	20
1.1	CONTEXTUALIZAÇÃO	20
1.2	JUSTIFICATIVA	21
1.3	OBJETIVOS	22
1.4	ESTRUTURA DO DOCUMENTO	23
2	TRABALHOS RELACIONADOS	24
2.1	HIPERTERMIA LOCAL POR FLUIDOS MAGNÉTICOS	24
2.2	CONTROLE APLICADO A SISTEMAS DE HIPERTERMIA	25
2.3	REDES DE ESTADO DE ECO COMO MODELOS DE PREDIÇÃO	27
2.4	COMENTÁRIOS FINAIS	28
3	FUNDAMENTAÇÃO TEÓRICA	29
3.1	HIPERTERMIA LOCAL POR FLUIDOS MAGNÉTICOS	29
3.1.1	Funcionamento	29
3.1.2	Mecanismos de Perdas em Nanopartículas	31
3.1.2.1	Perdas por Histerese	32
3.1.2.2	Perdas por Relaxação	33
3.1.3	Formulação Matemática de Rosensweig	35
3.1.4	Restrições Operacionais	36
3.1.5	Controle de Temperatura em Sistemas de Hipertermia Magnética	37
3.2	REDES NEURAIS ARTIFICIAIS	38
3.2.1	Neurônio Artificial	39
3.2.2	Arquiteturas de Redes Neurais	40
3.2.3	Redes Neurais de Estado de Eco	42
3.2.3.1	Modelo Básico	43
3.2.3.2	Inicialização do Modelo	44
3.2.3.3	Treinamento	45
3.2.3.4	Otimização de Hiperparâmetros	47
3.3	IDENTIFICAÇÃO DE SISTEMAS DINÂMICOS	50
3.4	COMENTÁRIOS FINAIS	51
4	MODELO NUMÉRICO DO SISTEMA	53
4.1	GEOMETRIA E IMPLEMENTAÇÃO DO MODELO	53
4.2	EQUAÇÕES GOVERNANTES	54
4.2.1	Escoamento Laminar	55
4.2.2	Transporte Intersticial de Nanofluido	56

4.2.3	Distribuição de Nanopartículas	59
4.2.4	Biotransferência de Calor	61
4.3	VERIFICAÇÃO DOS MODELOS NUMÉRICOS	63
4.4	COMENTÁRIOS FINAIS	66
5	ENSAIOS DE IDENTIFICAÇÃO DO SISTEMA	67
5.1	AQUISIÇÃO E PROCESSAMENTO DA SÉRIE DE DADOS	67
5.2	TREINAMENTO E OTIMIZAÇÃO DA REDE NEURAL	69
5.3	ANÁLISE DO CUSTO COMPUTACIONAL	73
5.4	COMENTÁRIOS FINAIS	75
6	RESULTADOS E DISCUSSÃO	76
6.1	MODELO ESFÉRICO	76
6.2	MODELO ELIPSOIDAL OBLATO	78
6.3	MODELO ELIPSOIDAL PROLATO	81
6.4	COMENTÁRIOS FINAIS	83
7	CONSIDERAÇÕES FINAIS	84
7.1	CONCLUSÕES	84
7.2	SUGESTÕES PARA TRABALHOS FUTUROS	85
	REFERÊNCIAS	86

1 INTRODUÇÃO

Neste capítulo, é apresentada uma introdução ao tema desta dissertação, assim como as justificativas para o desenvolvimento da mesma. Para isso, é feita uma contextualização do tema na Seção 1.1. Em seguida, são apresentados os desafios relacionados ao controle do tratamento de câncer por hipertermia magnética, o conceito de redes neurais de estado de eco e o seu potencial de contribuição para o controle do procedimento. Na Seção 1.3, são abordados os objetivos gerais e objetivos específicos deste trabalho. Por fim, a organização deste trabalho é descrita na Seção 1.4.

1.1 CONTEXTUALIZAÇÃO

O câncer é um grupo de mais de 100 doenças que afeta a humanidade há milhares de anos, originando-se de anormalidades genéticas frequentemente desencadeadas por diversos fatores, como exposição a substâncias carcinogênicas, medicamentos, infecções virais e radiação ultravioleta. A consequência disso é o descontrole no processo de divisão celular e a formação de tumores, definidos como uma proliferação descontrolada das células afetadas para além dos tecidos normais e com a capacidade de entrar em metástase em órgãos distantes (Stratton; Campbell; Futreal, 2009; Hausman, 2019; Jose *et al.*, 2020).

Atualmente, as estatísticas apontam que a incidência e a mortalidade por câncer estão crescendo rapidamente no cenário mundial. Por exemplo, projeções para o fim de 2023 sugerem que cerca de 609 mil pessoas irão falecer neste ano em decorrência de câncer nos Estados Unidos, equivalendo a mais de mil mortes por dia (Siegel *et al.*, 2023). No Brasil, as previsões para o triênio de 2023 a 2025 indicam uma tendência de mais de 704 mil novos registros de casos de câncer (Brasil, 2023).

Diante desse cenário alarmante, é de extrema importância promover um debate sobre os tratamentos atuais para o combate ao câncer. Por exemplo, além de não serem totalmente eficazes, tratamentos convencionais como quimioterapia, radioterapia e intervenção cirúrgica ainda possuem efeitos colaterais, como a morte de células saudáveis durante o processo e o aumento da resistência do organismo do paciente a medicamentos (Zaimy *et al.*, 2017; Schirmacher, 2019). Desse modo, à medida que existe uma busca por aumentar a eficácia e minimizar os efeitos colaterais de tais procedimentos, estudos recentes também têm buscado novas alternativas para o combate ao câncer, como a hipertermia por fluidos magnéticos (MFH, do inglês *Magnetic Fluid Hyperthermia*) (Rodrigues; Capistrano; Bakuzis, 2020; Zhang, C. *et al.*, 2022).

Idealizada a partir do estudo de Gilchrist *et al.* (1957), a hipertermia por fluidos magnéticos é uma modalidade de tratamento oncológico baseada em nanotecnologia, cuja proposta

é danificar as células tumorais por meio do aumento da temperatura local, ao mesmo tempo que as células normais são preservadas. Para isso, administra-se uma determinada dose de um nanofluido magnético na região do tumor, seguida da exposição do paciente a um campo magnético alternado (AMF, do inglês *Alternating Magnetic Field*), visando promover o aquecimento induzido da área (Rodrigues; Capistrano; Bakuzis, 2020; Tang *et al.*, 2023b).

1.2 JUSTIFICATIVA

Um dos desafios associados à execução do tratamento de hipertermia magnética é o controle da temperatura. Para o tratamento ser eficaz, a temperatura na região do tumor deve ser elevada até um intervalo entre 41 °C e 46 °C, podendo exceder até 50 °C em algumas situações. Desse modo, caso o procedimento seja conduzido inadequadamente, há o risco de causar danos aos tecidos saudáveis que cercam a região do tumor (Obaidat *et al.*, 2019; Suleman; Riaz, 2020b).

Para o controlar o aumento da temperatura local durante o tratamento de MFH, uma das possibilidades é ajustar a potência dissipada pelas nanopartículas magnéticas durante a sua exposição a um AMF. Para isso, é possível realizar um procedimento em malha aberta, no qual o operador define manualmente os parâmetros do campo magnético, a exemplo do que foi desenvolvido nos estudos de Johannsen *et al.* (2010) e Tang *et al.* (2021); ou então, utilizar estratégias de controle em malha fechada. Nesse contexto, pode-se citar o uso de controladores do tipo Proporcional-Integral (PI), investigados em Haemmerich e Webster (2005), além dos controladores baseado em lógica difusa, propostos em Wang e Tai (2017), e dos controladores adaptativos do tipo Proporcional-Integral-Derivativo (PID), descritos em Tang *et al.* (2023a).

Normalmente, há a necessidade de conhecer o modelo do processo para realizar o projeto de um controlador. No contexto do controle preditivo baseado em modelo (MPC, do inglês *Model Predictive Control*), o modelo também é usado como parte do controlador para fazer as estimativas para comportamento futuro das variáveis do processo ao longo de um horizonte de predição (Camacho; Bordons, 2007). No entanto, no caso do tratamento de hipertermia magnética, há algumas dificuldades associadas ao uso de um modelo baseado em leis físicas.

Para modelar todas as etapas do processo de hipertermia magnética, uma das abordagens mais utilizadas envolve o uso de um conjunto de equações diferenciais parciais não lineares, cujas soluções numéricas podem levar um tempo considerável para serem computadas. Por exemplo, a distribuição de temperatura na região do tumor pode ser estimada pela combinação da equação de biotransferência de calor de Pennes (1948) com a formulação de Rosensweig (2002) para produção de calor a partir de nanofluidos magnéticos. Como consequência, há uma dificuldade para realizar as predições em tempo real, uma vez que o tempo necessário para

simular 1 s do comportamento do sistema é maior que 1 s.

Nesse contexto, devido à sua capacidade de modelar fenômenos dinâmicos não lineares com tempo menor que o requerido por modelos físicos tradicionais, as redes de estado de eco (ESNs, do inglês *Echo State Networks*) apresentam potencial para serem utilizadas em aplicações de controle de malha fechada para o tratamento de hipertermia magnética. Introduzidas por Jaeger (2001), as ESNs são um modelo de rede neural recorrente que podem ser rapidamente treinadas por meio de algoritmos simples, como o de regressão linear (Lukoševičius, 2012).

Desse modo, justifica-se a relevância deste trabalho pela possibilidade de utilizar as redes neurais de estado de eco como modelo de predição de temperatura em sistemas que se utilizem de MPCs para o tratamento de hipertermia magnética, o que pode contribuir para o aprimoramento da segurança e da eficácia do procedimento.

1.3 OBJETIVOS

O objetivo geral deste trabalho é explorar a possibilidade de utilizar redes neurais de estado de eco como modelos de predição de temperatura para o tratamento de câncer com a terapia de hipertermia magnética. Isso envolve a realização de um estudo que abrange a modelagem matemática, simulações numéricas e a identificação do sistema dinâmico em um cenário com múltiplas entradas e múltiplas saídas (MIMO, do inglês *Multiple Input and Multiple Output*). A identificação é realizada com foco na aplicação como modelo de predição em MPC, porém o trabalho não abrange a avaliação do uso dos modelos para tarefas de controle.

Para o desenvolvimento desta dissertação, foram estabelecidos os seguintes objetivos específicos:

- a) revisar a literatura a fim de compreender e modelar matematicamente as etapas principais do tratamento de hipertermia magnética;
- b) propor modelos numéricos baseados em leis físicas com capacidade de representar diferentes cenários do tratamento de hipertermia magnética;
- c) propor um método que permita selecionar a arquitetura apropriada para a rede neural de estado de eco para essa aplicação;
- d) identificar a dinâmica das temperaturas em diferentes pontos do tumor durante a fase de aquecimento induzido do tratamento com emprego da rede neural escolhida;
- e) avaliar o desempenho da rede neural para a tarefa desejada e inferir sobre a possibilidade de uso em aplicações reais.

1.4 ESTRUTURA DO DOCUMENTO

Os trabalhos relacionados ao tema desta dissertação estão apresentados no Capítulo 2. Em seguida, no Capítulo 3 são explorados os conceitos que fundamentaram o desenvolvimento da dissertação, como o funcionamento da terapia de hipertermia magnética e das redes neurais artificiais, além do procedimento usual para a identificação de sistemas dinâmicos. Por sua vez, os modelos numéricos propostos neste estudo são detalhados no Capítulo 4. Em sequência, no Capítulo 5, é apresentado o procedimento proposto para a identificação dos sistemas não lineares representados pelos modelos numéricos. No Capítulo 6, são feitas as discussões a respeito dos resultados obtidos a partir dos ensaios de identificação de sistema. Por fim, as considerações finais sobre o trabalho são apresentadas no Capítulo 7.

2 TRABALHOS RELACIONADOS

Neste capítulo, são destacados estudos relevantes no contexto desta dissertação. A primeira linha de pesquisa abordada é a de hipertermia magnética, discutida na Seção 2.1. Em seguida, os trabalhos influentes na área de controle de sistemas de hipertermia são explorados na Seção 2.2. Na Seção 2.3, são apresentados trabalhos relacionados ao uso das redes neurais de estado de eco como modelos de predição em sistemas que envolvem controladores preditivos baseados em modelo. Por fim, na Seção 2.4, são apresentados comentários finais sobre os estudos listados, destacando questões relevantes para o contexto deste trabalho.

2.1 HIPERTERMIA LOCAL POR FLUIDOS MAGNÉTICOS

Um dos trabalhos pioneiros na área da hipertermia magnética foi desenvolvido por Gilchrist *et al.* (1957). Nesse estudo, foi introduzida a proposta de utilizar um processo de aquecimento induzido, envolvendo nanopartículas magnéticas, para eliminar células malignas localizadas em nódulos linfáticos.

Desde a publicação do artigo de Gilchrist *et al.* (1957), diversos trabalhos se propuseram a investigar o potencial do tratamento de hipertermia magnética. Por exemplo, em Johannsen *et al.* (2005) foram investigados os possíveis efeitos da terapia para o tratamento de cânceres de próstata. Para isso, os pesquisadores desenvolveram um estudo *in vivo* que induziu o desenvolvimento de um tumor desse tipo em ratos. Ao fim, foi relatado que a terapia levou a uma inibição do crescimento do tumor.

Em Jordan *et al.* (2006), foi promovido um estudo *in vivo* com foco no tratamento de tumores de glioblastoma¹ por meio da hipertermia magnética. Ao fim, os pesquisadores também verificaram que a terapia também pode contribuir para a redução do crescimento de tumores em caso desse tipo. Anos depois, em Maier-Hauff *et al.* (2011) foi realizado um estudo clínico para investigar os efeitos da hipertermia magnética em casos de glioblastoma. A partir disso, observaram que a terapia, quando combinada com radioterapia, pode aumentar a sobrevida dos pacientes.

A literatura também relata resultados positivos em pesquisas que envolvem a aplicação da hipertermia magnética em cenários relacionados ao câncer de pâncreas. Por exemplo, estudos *in vivo* conduzidos por Pan *et al.* (2021) e Soleymani *et al.* (2023) relataram que a hipertermia magnética possui potencial para ajudar a inibir o crescimento de tumores de fígado, além de contribuir para redução do volume dos mesmos.

¹ O glioblastoma é um tipo de tumor cerebral caracterizado por uma elevada taxa de mortalidade e notória resistência a tratamentos convencionais, incluindo quimioterapia e radioterapia, mesmo após intervenções cirúrgicas (Jordan *et al.*, 2006).

Além de estudos *in vivo*, é possível encontrar trabalhos com foco na modelagem matemática do tratamento de hipertermia magnética. Muitos dos resultados da modelagem são utilizados para planejar as etapas do tratamento, em especial para definir as propriedades dos nanofluidos e do campo magnético que resultam em um tratamento adequado. Nesse contexto, um dos artigos mais influentes é o de Baxter e Jain (1989), no qual foi apresentado um sistema de equações diferenciais parciais, baseado na lei de Darcy para meios porosos e na equação de convecção-difusão, para descrever o transporte intersticial e a distribuição de soluto ao longo de um tumor.

Para modelar o perfil de distribuição de temperatura em tumores, estudos como os de Tang *et al.* (2021) e Mahesh, Singh e Talukdar (2022) consideram o uso da equação de bio-transferência de calor de Pennes (1948). No entanto, é importante mencionar que há outras formulações que também podem ser utilizadas para a mesma finalidade. Por exemplo, o modelo de não equilíbrio térmico local, que considera os efeitos da porosidade dos tecidos biológicos, é explorado nos trabalhos de Rezvanian *et al.* (2020) e Dutta, Kundu e Yook (2021).

Para modelar os efeitos do campo magnético sobre a região do tumor durante o tratamento de hipertermia magnética, estudos como os de Wang e Tai (2017) e Rahpeima e Lin (2022) exploram as equações de Maxwell. No entanto, é possível encontrar modelagens mais simplificadas, com o intuito de reduzir a complexidade do modelo matemático. Por exemplo, em Suleman e Riaz (2020b) e Tang *et al.* (2021) é assumido que os valores de frequência e amplitude do AMF são uniformes em toda a região do tumor.

2.2 CONTROLE APLICADO A SISTEMAS DE HIPERTERMIA

No estudo de Haemmerich e Webster (2005), foi apresentado um controlador PI implementado em um programa escrito em C++ para sistemas de hipertermia por radiofrequência, modelados por meio do método dos elementos finitos. Nesse caso, o programa tinha como função verificar e garantir o seguimento de referência para a temperatura registrada na ponta do eletrodo utilizado no processo, por meio de ajustes na tensão fornecida ao sistema. No fim, apesar de relatarem dificuldade para sintonizar os parâmetros do controlador, os pesquisadores verificaram um desempenho razoável para o mesmo, identificando um sobressinal de até 11% durante os ensaios.

Outra proposta de controle de temperatura em sistemas de hipertermia foi apresentada por Xiaohua, Fei e Yuanjin (2015). Nesse trabalho, os autores desenvolveram um sistema capaz de realizar medições de temperatura de forma não invasiva por meio de imagens fotoacústicas, ao mesmo tempo, em que o processo de aquecimento é ajustado por um controlador PID. Para sintonizar os parâmetros do controlador, os pesquisadores propuseram um modelo do processo

baseado em simplificações para as equações governantes. No final, relataram um bom desempenho do PID e discutiram quais pontos poderiam ser melhorados. Em particular, abordaram o processo de sintonia dos parâmetros do PID, que poderia ser otimizado com um algoritmo de ajuste automático. Além disso, destacaram a necessidade de estender o algoritmo PID para cenários envolvendo o controle de temperatura em vários pontos de medição.

Em Wang e Tai (2017), foi apresentada uma estratégia de controle adaptativo baseada em lógica difusa, a fim de garantir o seguimento de referência para a temperatura local, rejeição de perturbações e capacidade de lidar com incertezas paramétricas e com dinâmicas não lineares. Nesse caso, o algoritmo do controlador apresentou a capacidade de aprender com o erro de seguimento de referência, para então ajustar tanto os próprios parâmetros quanto as próprias regras de controle em tempo de execução. A partir disso, os autores demonstram que o controlador proposto apresentou uma boa robustez e adaptabilidade a diferentes cenários de tratamento. Entretanto, é importante comentar que o controlador foi projetado para um sistema com um único ponto de medição de temperatura. Na prática, é recomendável que o controle seja feito a partir de vários pontos de sensoriamento, uma vez que a distribuição heterogênea de nanopartículas no local faz com que um único ponto de sensoriamento seja insuficiente para um controle seguro da terapia (Hergt; Dutz, 2007).

No trabalho de Sebeke *et al.* (2019), foi detalhado um algoritmo de controle preditivo baseado em modelo de espaço estados para regular a temperatura na região do tumor durante procedimentos de hipertermia com ultrassom focalizado de alta intensidade. Nesse caso, os autores identificaram o modelo MIMO do processo por meio de ensaios preliminares de sonificação e, posteriormente, ajustaram o algoritmo do MPC para controlar o aumento da temperatura local por meio de ajustes na potência acústica dissipada pelo equipamento de ultrassom. Como resultado, os pesquisadores observaram que o MPC produziu pequenos erros de seguimento de referência, além de garantir uma distribuição de temperatura mais homogênea que outras técnicas clássicas de controle. No entanto, uma limitação deste estudo é a não consideração das restrições do processo durante a implementação do MPC.

Outro exemplo relevante é o estudo conduzido por Rezvanian *et al.* (2020), no qual foi proposta uma estratégia de controle por realimentação de estados com o intuito de estabelecer um segmento de referência para a temperatura média do tumor, alcançando-o por meio do ajuste da potência dissipada pelas nanopartículas. Nesse caso, os pesquisadores propuseram um algoritmo de regressão que utiliza dados de medição de temperatura no centro do tumor para estimar a distribuição de temperatura na área e, assim, determinar seu valor médio. No final, os autores demonstram que a estratégia tem potencial para atingir o objetivo desejado, ao mesmo tempo, em que citam os efeitos de incertezas do processo sobre o desempenho do controlador.

Por fim, em Tang *et al.* (2023a) foi introduzido um algoritmo envolvendo um controla-

dor PID com otimização dinâmica para os próprios parâmetros, pelo algoritmo do recozimento simulado (do inglês, *simulated annealing*). A partir disso, os pesquisadores apresentaram resultados que lhes permitiram inferir que o sistema de controle proposto pode regular a temperatura local durante o tratamento, ajustando a dissipação de energia das nanopartículas magnéticas e se adaptando automaticamente a diferentes situações durante a terapia, como nos casos em que a distribuição de nanofluido na região do tumor varia segundo a estratégia de infusão. Entretanto, tal estudo poderia considerar um cenário ainda mais realista, optando por ajustar diretamente os parâmetros do campo magnético ao invés da potência dissipada pelas nanopartículas, além de considerar mais restrições para a entrada e para a saída do processo.

2.3 REDES DE ESTADO DE ECO COMO MODELOS DE PREDIÇÃO

Desde a sua introdução em Jaeger (2001), as redes neurais de estado de eco apresentam um potencial de aplicação para a identificação de sistemas dinâmicos, mesmo que não haja um conhecimento prévio sobre os mesmos. Por exemplo, em Jaeger (2002) é discutido o uso das ESNs para tal finalidade, tanto em um cenário de identificação off-line quanto on-line.

As redes de estado de eco também são comumente utilizadas na literatura como modelos de predição de sistemas que envolvem controladores preditivos baseados em modelo. Um exemplo de aplicação é o artigo de Pan e Wang (2012), que utilizou ESNs para um problema de referência que envolvia a identificação on-line de um sistema dinâmico não linear e predição de suas saídas.

Nos últimos anos, em Jordanou *et al.* (2018), foi proposto a utilização de uma ESN como modelo de predição para a pressão em poços de petróleo com elevação de gás, com foco em aplicações no contexto do MPC. Nesse caso, a proposta dos autores visava contornar as complexidades associadas à dinâmica não linear do processo e à sua difícil modelagem.

Outra proposta de uso das ESNs como modelo de predição em MPCs foi apresentada no estudo de Schwedersky *et al.* (2018). Neste caso, as ESNs foram utilizadas em experimentos práticos que envolveram um sistema não linear composto por uma bancada de compressores. No contexto deste estudo, os pesquisadores desenvolveram um algoritmo de MPC não linear que emprega as ESNs como uma ferramenta para estimar tanto a resposta livre do sistema, quanto a sua resposta forçada.

Em Armenio *et al.* (2019), foram analisadas as propriedades de estabilidade, o procedimento de treinamento e o uso das ESNs em aplicações com MPC. Para isso, os pesquisadores testaram a rede neural em um problema de referência que envolvia um processo de neutralização de pH com dinâmicas não lineares.

Finalmente, é relevante mencionar o estudo de Schwedersky, Flesch e Rovea (2022).

Nesse caso, os pesquisadores propuseram a aplicação de redes neurais de estado de eco para identificação em tempo real de sistemas MIMO não lineares, bem como sua utilização como modelo de predição no contexto de um MPC adaptativo.

2.4 COMENTÁRIOS FINAIS

Com base no estudo da literatura apresentado, é possível verificar que a hipertermia magnética é um procedimento com um grande potencial de contribuição para o tratamento do câncer, e que está em constante avanço. No entanto, pode ser notado que a maioria dos esforços para modelar matematicamente todas as etapas da terapia envolvem equações diferenciais parciais não lineares. Essas equações tipicamente precisam ser resolvidas de forma numérica com técnicas de elementos finitos, implicando em um custo computacional considerável.

Ao analisar os trabalhos apresentados para a área de controle, é observado que a literatura apresenta soluções interessantes para o controle de temperatura durante procedimentos de hipertermia, mas que as propostas têm margem para melhorias. Por exemplo, apesar do potencial de adaptação dos controladores apresentados por Wang e Tai (2017), Sebeke *et al.* (2019) e Tang *et al.* (2023b) para qualquer cenário de tratamento, uma questão que poderia ser melhor abordada são as restrições de segurança para o processo. Além disso, os controladores propostos em Haemmerich e Webster (2005) e Xiaohua, Fei e Yuanjin (2015) poderiam ser melhorados se considerassem mais pontos de medição de sensoriamento para a temperatura na região do tumor, uma vez que é difícil fazer um controle seguro do tratamento com apenas um ponto de medição.

Finalmente, é evidente que as redes de estado de eco têm contribuído para o desenvolvimento de modelos preditivos para controladores baseados em modelo, abrangendo tanto sistemas lineares quanto não lineares. Diante da ausência, até o momento, de estudos voltados para a utilização das ESNs como modelo de predição de temperatura durante o tratamento de hipertermia magnética, e considerando a complexidade da modelagem matemática para o processo em questão, pode-se destacar o potencial de contribuição desta dissertação na etapa de modelagem do problema, que pode viabilizar o desenvolvimento de novas técnicas de controle adaptativo para a terapia de hipertermia magnética, com capacidade de considerar um processo MIMO, bem como suas restrições operacionais.

3 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são abordados os principais fundamentos teóricos que basearam o desenvolvimento deste trabalho. O primeiro assunto abordado é o tratamento de hipertermia magnética, o qual é discutido na Seção 3.1. Em seguida, na Seção 3.2, são explorados os conceitos gerais referentes ao funcionamento das redes neurais artificiais, bem como é apresentado um detalhamento sobre a arquitetura da rede de estado de eco. Na Seção 3.3, é apresentada uma visão geral sobre a identificação de sistemas dinâmicos. Por fim, na Seção 3.4, são feitos os comentários finais a respeito dos fundamentos discutidos nesta seção.

3.1 HIPERTERMIA LOCAL POR FLUIDOS MAGNÉTICOS

Nesta seção, são abordados os conceitos relacionados à terapia de hipertermia magnética. Os primeiros temas abordados são seu funcionamento e suas etapas, os quais são discutidos na Seção 3.1.1. Em seguida, na Seção 3.1.2, é apresentada uma visão geral sobre os mecanismos de geração de calor por meio de nanofluidos magnéticos. Por sua vez, na Seção 3.1.3, é destacada a formulação matemática de Rosensweig (2002), a qual permite estimar a potência dissipada pelas nanopartículas magnéticas durante o tratamento. Na Seção 3.1.4, são discutidas as restrições operacionais que podem ser aplicadas durante o controle do procedimento. Por fim, na Seção 3.1.5, são abordadas as alternativas voltadas para o controle de temperatura durante o tratamento de hipertermia magnética, além de quais aspectos podem ser melhorados no controle do tratamento.

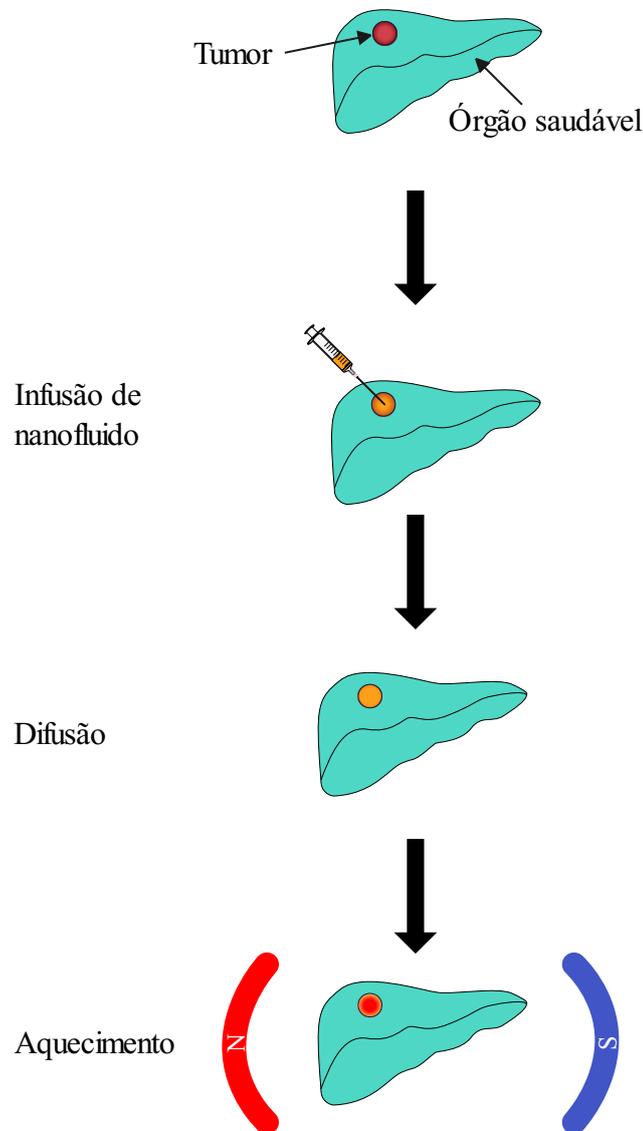
3.1.1 Funcionamento

A hipertermia magnética é uma modalidade de tratamento oncológico cuja proposta tem origens no trabalho de Gilchrist *et al.* (1957), no qual foi demonstrado que há a possibilidade de eliminar tumores em nódulos linfáticos a partir de um processo de aquecimento induzido envolvendo nanopartículas magnéticas (MNPs, do inglês *Magnetic NanoParticles*). Desde então, pode-se observar que o procedimento vem sendo alvo de diversos estudos, dado o seu potencial de aplicação (Huang; Hainfeld, 2013; Rahpeima; Lin, 2022).

O tratamento de hipertermia magnética possui três etapas básicas, conforme ilustrado na Figura 1. Primeiramente, faz-se a administração de uma determinada dose de um nanofluido magnético na região tumor. Em seguida, há um procedimento usual, no qual é aguardado um período entre o término da etapa de infusão e o início do processo de aquecimento, para garantir que haja a difusão do nanofluido no tecido, garantindo uma maior homogeneidade na concentração. Por fim, um campo magnético externo é aplicado a região do tumor, promovendo

o aquecimento induzido do local (Suleman; Riaz, 2020b; Tang *et al.*, 2022).

Figura 1 – Etapas do tratamento de hipertermia magnética.



Fonte: adaptado de Suleman e Riaz (2020b).

Para a primeira etapa, há ao menos três opções para transportar as nanopartículas magnéticas até a região do tumor. A injeção direta é considerada o método mais invasivo, mas que permite uma administração precisa das nanopartículas no tumor. No entanto, pode resultar em uma distribuição mais desigual das MNPs ao redor do local da injeção. Por outro lado, as injeções intraperitoneais e intravenosas são menos invasivas e promovem uma distribuição mais uniforme das nanopartículas ao longo da região do tumor. Porém, há o risco de tais aplicações causarem uma toxicidade sistêmica ao paciente (Pankhurst *et al.*, 2003; Huang; Hainfeld, 2013).

A segunda etapa visa permitir que as nanopartículas magnéticas se distribuam de forma mais uniforme ao longo da região do tumor por meio do processo de difusão. No entanto, é importante ressaltar que a literatura não apresenta um consenso claro sobre o tempo ideal

para essa etapa. Tal fato fica evidenciado pela variedade de intervalos considerados em estudos no âmbito computacional, o que pode ser atribuído ao entendimento de que o processo de difusão pode ser influenciado por vários fatores, incluindo o diâmetro das nanopartículas, as propriedades do nanofluido e as características do tumor (Suleman; Riaz, 2020b; Mahesh; Singh; Talukdar, 2022).

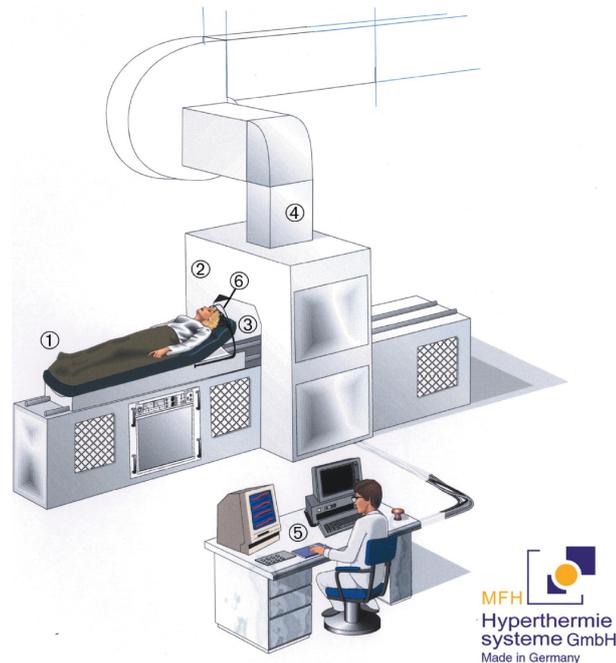
A etapa de aquecimento induzido pode ter vários objetivos, tanto o de potencializar os efeitos de tratamentos convencionais, como a radioterapia, quanto o de induzir a necrose das células malignas. De modo geral, a terceira etapa pode ser classificada quanto à sua faixa de operação da seguinte forma. A hipertermia leve é caracterizada por uma aplicação até uma faixa de temperatura entre 40 °C e 42 °C e por um período de 30 min a 60 min. Em Hannon *et al.* (2021), é relatado que essa modalidade de tratamento pode atuar como um agente radiosensibilizador, melhorando a oxigenação, o fluxo sanguíneo e a permeabilidade vascular na região do tumor, porém com menores chances de causar a morte das células malignas. Por outro lado, a hipertermia moderada ocorre em uma faixa de temperatura entre 42 °C e 45 °C, também por um período de 30 min a 60 min. Neste caso, o trabalho de Hannon *et al.* (2021) cita que esse procedimento pode induzir a apoptose das células malignas sem causar danos visíveis às células normais. Por sua vez, a ablação térmica ocorre com temperaturas acima de 50 °C e por um período de 10 min a 30 min (Suleman; Riaz, 2020b). Em Legge *et al.* (2019), é citado que essa modalidade visa induzir a morte direta das células malignas, sem se preocupar em sensibilizar o tumor a outros tratamentos.

Finalmente, do ponto de vista operacional, o tratamento de hipertermia magnética pode ser realizado utilizando um equipamento semelhante ao ilustrado na Figura 2. Nesse caso, ① representa o leito do paciente. A câmara em ② é responsável por aplicar um campo magnético perpendicular ao paciente, conforme é representado em ③. Durante o procedimento, o núcleo do equipamento é resfriado por meio de trocadores de calor, em ④. No caso de tumores cerebrais, a região de interesse para medições de temperatura, tanto no tumor quanto em pontos de referência, é indicada em ⑥. Enquanto isso, um especialista supervisiona e controla todo o processo em ⑤, fazendo ajustes na distância vertical entre a região do tumor e a fonte do campo magnético em ③, assim como nos parâmetros do AMF (Jordan *et al.*, 2001).

3.1.2 Mecanismos de Perdas em Nanopartículas

A produção de calor a partir de um material magnético em resposta à exposição a um AMF pode ser explicada por meio de pelo menos três mecanismos de perda distintos. Segundo o trabalho apresentado em Obaidat *et al.* (2019), esses mecanismos são as perdas por histerese, as perdas por relaxação de Néel e as perdas por relaxação de Brown. Com base nisso, a fim

Figura 2 – Esboço de um protótipo de um sistema de hipertermia local por fluidos magnéticos.



Fonte: Jordan *et al.* (2001).

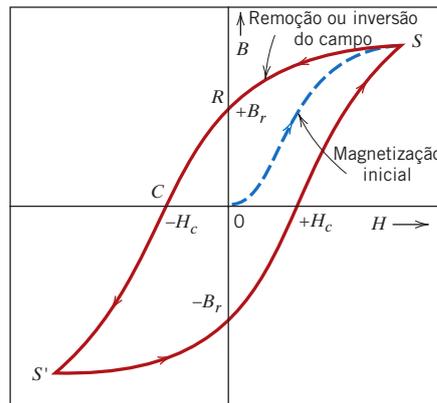
de discutir sobre como cada um desses mecanismos contribui para o processo de aquecimento durante o tratamento de hipertermia magnética, esta seção está organizada da seguinte forma. Primeiramente, na Seção 3.1.2.1, é apresentado o mecanismo de produção de calor por meio das perdas por histerese. Em seguida, na Seção 3.1.2.2, são explorados os mecanismos de perdas por relaxações de Néel e de Brown.

3.1.2.1 Perdas por Histerese

As perdas por histerese, como o nome sugere, estão relacionadas ao ciclo de histerese de um material magnético, exemplificado na Figura 3. Os domínios de um material magnético, inicialmente desordenados, tendem a se alinhar quando sujeitos a um campo magnético de amplitude H . Quando esse processo é concluído, é dito que o material atingiu sua magnetização de saturação (ponto S). Em sequência, conforme a amplitude do campo é reduzida, é observado que a curva de magnetização não segue o mesmo caminho original. Como consequência, o ponto R ilustra que mesmo na ausência de um campo magnético, o material permanece magnetizado com uma densidade de fluxo remanescente, B_r , também conhecida como remanescência (Callister; Rethwisch, 2016).

Para promover a desmagnetização da amostra (chegar ao ponto C na Figura 3), é necessário aplicar um campo coercitivo, H_c , na direção oposta ao campo original. Conforme a intensidade do campo é aumentada em módulo, a saturação no sentido oposto pode ser alcan-

Figura 3 – Representação do ciclo de histerese de um material magnético.



Fonte: Callister e Rethwisch (2016).

çada, indicada pelo ponto S' . Por fim, uma segunda inversão do campo magnético de volta ao ponto de saturação inicial (ponto S) fecha o ciclo de histerese, produzindo tanto uma remanência negativa quanto exigindo uma coercividade positiva para promover a desmagnetização da amostra (Callister; Rethwisch, 2016).

Durante o ciclo de histerese, ocorre uma dissipação da energia magnética absorvida pelo material na forma de calor, sendo esse um processo proporcional à área entre as curvas de magnetização. Contudo, a magnitude de tais perdas pode ser influenciada pela morfologia do material. Normalmente, materiais com tamanho em micro/nanoescala apresentam um estado ferro/ferrimagnético, com as perdas por histerese sendo o principal mecanismo de aquecimento. Entretanto, quando as dimensões estão abaixo de um valor crítico, o material adota um comportamento superparamagnético, reduzindo as perdas por histerese a quase zero (Yu *et al.*, 2022). Por exemplo, em Ma *et al.* (2004), esse fato foi observado para nanopartículas de magnetita com um diâmetro de até 13 nm, quando expostas a um campo magnético operando a uma frequência de 80 kHz e uma intensidade de $32,5 \text{ kA m}^{-1}$.

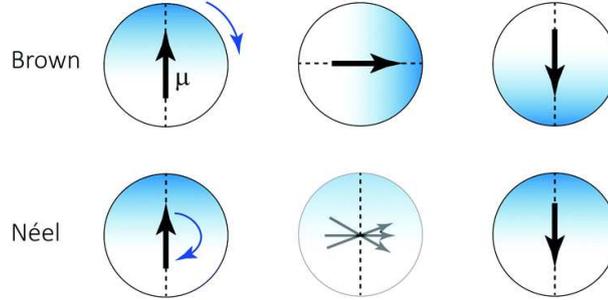
3.1.2.2 Perdas por Relaxação

Quando um campo magnético é aplicado, ocorre uma transferência de energia para as partículas, forçando os seus momentos magnéticos a se alinharem na direção do campo. Quando o campo é removido, há uma tentativa de realinhamento dos momentos magnéticos com o eixo de orientação preferencial. Como consequência, ocorre a dissipação da energia magnética recebida na forma de calor (Obaidat *et al.*, 2019).

Nesse contexto, dois mecanismos de perdas podem ser destacados: o de Néel e o de Brown, ilustrados na Figura 4. As perdas por relaxação de Brown ocorrem devido à rotação da nanopartícula, acompanhada pelo seu momento magnético, μ , em função da direção do campo

externo. Por outro lado, durante as perdas por relaxação de Néel, apenas o momento magnético da nanopartícula é rotacionado (Ilg; Kröger, 2020).

Figura 4 – Mecanismo de perdas por relaxação de Brown e de Néel.



Fonte: Ilg e Kröger (2020).

Uma maneira de avaliar a influência das perdas por relaxação de Néel e de Brown no processo de geração de calor é através da análise de seus tempos de relaxação respectivos. Para determiná-los, o seguinte par de equações pode ser empregado (Obaidat *et al.*, 2019):

$$\tau_N = \tau_0 \frac{\sqrt{\pi} \exp \Gamma}{2\sqrt{\Gamma}}, \quad (3.1)$$

$$\tau_B = \frac{3\mu_{nf}V_H}{k_B T_i}, \quad (3.2)$$

nas quais τ_N é o tempo característico das perdas por relaxação de Néel; τ_B é o tempo característico das perdas por relaxação de Brown; μ_{nf} é a viscosidade dinâmica do nanofluido magnético; V_H é o volume hidrodinâmico da nanopartícula; $k_B = 1,38 \times 10^{-23} \text{ J K}^{-1}$ é a constante de Boltzmann; T_i é a temperatura em um meio i ; τ_0 é o tempo característico do material, usualmente na ordem de $1 \times 10^{-9} \text{ s}$; e Γ é uma variável intermediária, dada por:

$$\Gamma = \frac{K_a V_M}{k_B T_i}, \quad (3.3)$$

em que V_M é o volume da nanopartícula e K_a é a constante de anisotropia da nanopartícula.

Quando as perdas por relaxação de Néel e de Brown ocorrem simultaneamente, é possível estimar um tempo de relaxação efetivo para nanopartícula, τ . Para isso, pode-se aplicar a seguinte equação (Tang; Flesch; Jin, 2017):

$$\frac{1}{\tau} = \frac{1}{\tau_B} + \frac{1}{\tau_N}. \quad (3.4)$$

O resultado da Equação (3.4) tende a ser mais influenciado pelo menor tempo de relaxação da nanopartícula magnética (τ_N ou τ_B). Em Tang, Flesch e Jin (2017), são apresentadas observações a respeito das implicações de cada tempo de relaxação sobre o processo de geração de calor por meio de nanopartículas magnéticas de magnetita (Fe_3O_4). No contexto desse trabalho, os autores citam que para situações em que $\tau \approx \tau_N$, foi verificado que a potência

dissipada pelas MNPs tende a zero, independentemente do valor da constante de anisotropia do material. Contudo, quando $\tau \approx \tau_B$, foi notado que uma nanopartícula pode dissipar mais energia magnética na forma de calor.

3.1.3 Formulação Matemática de Rosensweig

O calor produzido a partir da ação de um AMF sobre um nanofluido magnético pode ser modelado de diversas maneiras, sendo que uma das formas tipicamente utilizadas na literatura é por meio da formulação de Rosensweig (2002). Nesse modelo, é considerado que as nanopartículas se encontram em um estado superparamagnético e, portanto, as perdas por relaxação de Néel e Brown são consideradas os mecanismos predominantes de aquecimento no processo.

A partir da Primeira Lei da Termodinâmica, ao admitir um sistema de densidade constante, volume unitário e sujeito a um campo colinear, a variação infinitesimal da energia interna de um material magnético pode ser escrita como (Rosensweig, 2002):

$$dU = H \cdot dB, \quad (3.5)$$

em que B é a densidade do fluxo magnético, dada por:

$$B = \mu_0(H + M), \quad (3.6)$$

em que U é a energia interna do material magnético; M é a magnetização; e $\mu_0 = 4\pi \times 10^{-7} \text{ NA}^{-2}$ é a constante de permeabilidade no vácuo. Por sua vez, a variação da energia interna do material, ΔU , pode ser estimada a partir da integração da Equação (3.5), conforme é dado por:

$$\Delta U = -\mu_0 \oint H dM. \quad (3.7)$$

Quando a magnetização fica atrasada em relação ao campo, o termo integrador na Equação (3.7) produz um resultado positivo, indicando a conversão de trabalho magnético em energia interna. Para expressar a magnetização em função do campo magnético e da susceptibilidade complexa do nanofluido magnético, em Rosensweig (2002) foi proposto o uso do seguinte par de equações:

$$H(t) = H_0 \cos \omega t, \quad (3.8)$$

$$\chi = \chi' - j\chi'', \quad (3.9)$$

nos quais $H(t)$ representa a intensidade do campo magnético ao longo do tempo; H_0 é a amplitude máxima da intensidade do campo magnético; $\omega = 2\pi f$ é a frequência angular do AMF; t é o tempo; χ é a susceptibilidade complexa do nanofluido magnético; χ' é o componente em fase de χ , dado por:

$$\chi' = \frac{\chi_0}{1 + (\omega\tau)^2} \quad (3.10)$$

e χ'' é o componente fora de fase de χ , definido como:

$$\chi'' = \frac{\omega\tau}{1 + (\omega\tau)^2} \chi_0, \quad (3.11)$$

em que χ_0 é a suscetibilidade no equilíbrio do nanofluido magnético, dada pela seguinte equação:

$$\chi_0 = \frac{M_d \phi_i}{H} \left[\coth \left(\frac{\mu_0 M_d H V_M}{k_B T} \right) + \frac{k_B T_i}{\mu_0 M_d H V_M} \right], \quad (3.12)$$

em que ϕ_i é a fração de nanopartículas no meio i e M_d é a magnetização de domínio de uma nanopartícula suspensa. Com base nisso, a magnetização, M , pode ser descrita por:

$$M(t) = \text{Re}[\chi H_0 e^{j\omega t}]. \quad (3.13)$$

Ao substituir a Equação (3.9) na Equação (3.13), é possível expandir tal formulação em uma soma de senos e cossenos a partir da fórmula de Euler. Desse modo, tem-se que:

$$M(t) = H_0 [\chi' \cos \omega t + \chi'' \text{sen } \omega t]. \quad (3.14)$$

As definições apresentadas para o comportamento do campo magnético e para a magnetização do nanofluido magnético podem ser aplicadas para reescrever a Equação (3.7). A partir disso, em Rosensweig (2002) foi proposto que:

$$\Delta U = 2\mu_0 f H_0^2 \chi'' \int_0^{2\pi/\omega} (\text{sen } \omega t)^2 dt. \quad (3.15)$$

Por fim, a integração da Equação (3.15) e a multiplicação de seu resultado pela frequência cíclica $f = \omega/(2\pi)$ permite definir a potência dissipada pelas nanopartículas, P . Desse modo, Rosensweig (2002) concluiu que:

$$P = \mu_0 \pi \chi'' f H_0^2. \quad (3.16)$$

3.1.4 Restrições Operacionais

Ao analisar a formulação de Rosensweig (2002), fica evidente que a quantidade de calor gerada por uma nanopartícula está diretamente relacionada à frequência, f , e à amplitude máxima do campo magnético externo, H_0 . Em teoria, é possível aumentar a geração de calor ajustando tais variáveis. No entanto, há algumas limitações a serem consideradas, uma vez que existe uma grande preocupação em evitar danos às células saudáveis devido às correntes parasitas geradas por campos magnéticos operando em frequências e amplitudes elevadas (Obaidat *et al.*, 2019). Desse modo, a literatura apresenta algumas restrições operacionais para garantir a segurança do procedimento, as quais podem ser consideradas tanto para o projeto de equipamentos quanto para o controle do processo.

Para a frequência e a amplitude do AMF, em Atkinson, Brezovich e Chakraborty (1984) foi estabelecida uma restrição de que o valor do produto entre tais variáveis deve ser menor ou igual a $4,85 \times 10^8 \text{ A m}^{-1} \text{ s}^{-1}$ para tratar tumores localizados em regiões como tórax e abdômen. Porém, os pesquisadores comentam que tal restrição pode ser relaxada, dependendo do tamanho da região na qual o tratamento será realizado. Por sua vez, em Hergt *et al.* (2004) foi apresentada uma restrição de segurança para o caso do tratamento de cânceres de mama, a qual recomenda que o produto entre a frequência de operação e a amplitude do campo magnético não deve ultrapassar o valor de $3 \times 10^9 \text{ A m}^{-1} \text{ s}$. Nesse caso, os autores argumentam que tal restrição tende a ser suficiente para garantir um processo eficiente e seguro.

Por fim, é importante mencionar que também é possível aplicar restrições à temperatura na área afetada pelo campo magnético externo. Uma abordagem comum inclui a imposição de um limite superior para a temperatura na região do tumor, a fim de evitar o superaquecimento do local. Além disso, por questões de segurança, também é recomendável estabelecer um valor máximo para a temperatura em áreas distantes do tumor (Sharma *et al.*, 2023).

3.1.5 Controle de Temperatura em Sistemas de Hipertermia Magnética

Durante a terapia de hipertermia magnética, existe uma grande preocupação quanto ao controle da temperatura na região do tumor, uma vez que um superaquecimento do local tende a provocar danos indesejados aos tecidos saudáveis mais próximos. Além disso, é importante que a temperatura seja mantida na faixa desejada, a fim de garantir os efeitos esperados para a terapia, conforme discutido na Seção 3.1.1.

Para promover o controle de temperatura em sistemas de hipertermia magnética, existem diversas linhas de pesquisas. Uma delas, é dedicada ao desenvolvimento de nanopartículas com a capacidade de dissipar a maior quantidade possível de calor, mesmo sob a influência de campos magnéticos com amplitude mais baixa. Ao mesmo tempo, é de interesse que tais nanopartículas apresentem uma baixa temperatura de Curie, conferindo às MNPs a capacidade de interromper temporariamente o processo de aquecimento, caso a temperatura local exceda a sua temperatura de Curie. Por exemplo, essa capacidade foi relatada no trabalho de W. Zhang *et al.* (2019) para nanopartículas magnéticas de $\text{Zn}_{0,54}\text{Co}_{0,46}\text{Cr}_{0,6}\text{Fe}_{1,4}\text{O}_4$, as quais apresentaram uma temperatura de Curie de $48,8^\circ\text{C}$. Por outro lado, em Rao *et al.* (2019), foi observada uma temperatura de Curie de 46°C para nanopartículas de $\text{Co}_{0,37}\text{Zn}_{0,63}\text{Fe}_2\text{O}_4$.

Por sua vez, uma segunda linha de pesquisa visa propor controladores com a capacidade de regular a temperatura na área do tumor por meio de ajustes na potência dissipada pelas nanopartículas magnéticas. Paralelamente, é de interesse que tais controladores possam rejeitar perturbações, demonstrem robustez diante das variações paramétricas inerentes à aplicação do

método em diferentes cenários e considerem as restrições operacionais. No entanto, a linha em questão apresenta alguns desafios, os quais vêm sendo explorados na literatura. Um deles, está relacionado ao desenvolvimento de um modelo do processo de hipertermia, o que é normalmente requerido para o projeto e a validação de controladores. Nesse caso, a principal dificuldade está associada ao fato de que todo o tratamento é geralmente modelado por meio de equações diferenciais parciais não lineares, cujos parâmetros tendem a variar dependendo do cenário do tratamento. Além disso, dada a complexidade de tais equações, existem dificuldades para integrar todas as equações governantes em um único modelo (Rahpeima; Lin, 2022).

Perante o exposto, abordagens envolvendo controladores adaptativos vêm ganhando destaque na literatura. Por exemplo, o controlador baseado em lógica difusa, proposto em Wang e Tai (2017), visava adaptar os seus parâmetros e as suas regras de controle a partir do erro de seguimento de referência para a temperatura em um ponto de sensoriamento, enquanto o algoritmo do controlador PID explorado em Tang *et al.* (2023b) foi projetado para sintonizar os seus parâmetros com base na distribuição de temperatura ao longo do tumor.

Um aspecto que ainda pode ser melhor desenvolvido em sistemas de controle de hipertermia magnética é a capacidade de regular a temperatura em diversos pontos da região do tumor. Conforme discutido na Seção 2.1, nem todos os trabalhos consideram tal possibilidade. No entanto, existe uma perspectiva de que isso possa mudar, uma vez que diversos métodos para o sensoriamento da temperatura vêm sendo propostos nos últimos anos. Por exemplo, o estudo realizado por Hankiewicz *et al.* (2016) aborda a possibilidade de realizar o monitoramento da temperatura por meio de imagens de ressonância magnética. Além disso, em Liu *et al.* (2017) é demonstrada a viabilidade do sensoriamento utilizando técnicas de ultrassom. Por fim, vale comentar que também é possível monitorar a temperatura na região do tumor por meio de imagens fotoacústicas, conforme mencionado no estudo de Vu, Razansky e Yao (2019).

3.2 REDES NEURAIIS ARTIFICIAIS

As Redes Neurais Artificiais (RNAs) são modelos computacionais, que surgiram de uma busca por emular o funcionamento das redes neurais biológicas, de forma a produzir soluções para problemas sem a necessidade da programação direta, e sim por meio de algoritmos de aprendizado a partir de exemplos (Mitchell, 1997). Para executar essa proposta, as redes neurais artificiais adotam estruturas compostas por um conjunto de neurônios artificiais, os quais estão interconectados seguindo arquiteturas específicas. Desse modo, cada neurônio pode receber e transmitir os seus sinais para outras unidades de neurônios, formando, assim, uma estrutura similar àquela observada nos seres vivos (Mitchell, 1997).

A partir de tal organização, as redes neurais artificiais podem ser empregadas para diver-

sas atividades, como processamento de sinais, reconhecimento de padrões e predição de séries temporais (Mitchell, 1997). Porém, é importante comentar que a área de pesquisa das RNAs é extensa e que cada tarefa é tipicamente realizada por uma RNA com arquitetura adequada para tal. Por esse motivo, esta seção apresenta os conceitos referentes às RNAs que fundamentaram o desenvolvimento deste trabalho. Para isso, foi adotada a seguinte organização. Primeiramente, na Seção 3.2.1, é apresentada uma descrição geral sobre o funcionamento dos neurônios artificiais. Em seguida, na Seção 3.2.2, são detalhadas as arquiteturas das RNAs, introduzindo os conceitos de rede neural alimentada para adiante (do inglês, *feedforward*) e de rede neural recorrente. Por fim, na Seção 3.2.3 é apresentada a arquitetura das redes neurais de estado de eco.

3.2.1 Neurônio Artificial

Apesar de sua complexidade variável, o Sistema Nervoso Central (SNC) de todos os animais é composto por um elemento em comum: o neurônio. O neurônio pode ser compreendido como uma unidade do SNC cuja função é receber, processar e transmitir informações (Hales; Yudofsky; Gabbard, 2012).

Inspirados nesse princípio, McCulloch e Pitts (1943) propuseram um dispositivo binário, o qual ficou conhecido unidade de McCulloch-Pitts. Uma das principais características desse dispositivo é a sua capacidade de responder à atividade de suas sinapses, a qual reflete o estado de neurônios anteriores. Quando nenhuma sinapse inibitória estiver ativa, o neurônio soma suas entradas em um instante de tempo e , então, avalia se o somatório excedeu ou não um limiar. Se isso ocorreu, o neurônio será ativado. Se não, o neurônio permanece inativo (Anderson, 1995).

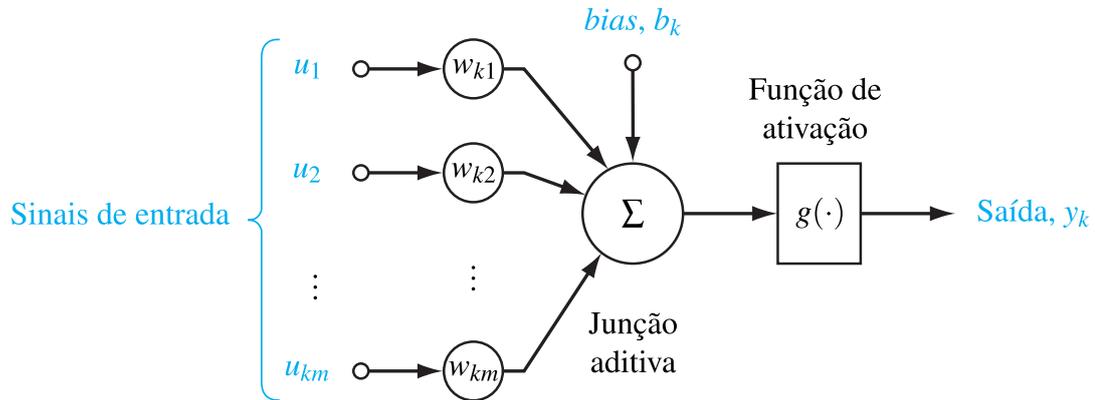
Com base no modelo de McCulloch-Pitts, em Rosenblatt (1958), foi proposta a rede *perceptron* de única camada, que consiste em um conjunto de neurônios artificiais cujos pesos sinápticos podem ser ajustados a partir de um processo de aprendizagem. Nas palavras de Haykin (2009, p. 48), pode-se classificar o *perceptron* como “a forma mais simples de uma rede neural que pode ser usada para classificação de padrões ditos linearmente separáveis (isto é, padrões localizados em lados opostos de um hiperplano)”.

Conforme ilustrado na Figura 5, um neurônio artificial opera da seguinte forma: primeiramente, o neurônio recebe um vetor de m entradas com valores reais. Em seguida, a unidade realiza uma combinação linear desses valores, considerando os pesos associados a cada conexão. No fim, essa combinação linear é processada por uma função de ativação, cujo resultado é transformado em uma saída (Haykin, 2009).

O funcionamento de um neurônio artificial k pode ser representado matematicamente por:

$$y_k = g(x_k), \quad (3.17)$$

Figura 5 – Representação do funcionamento de um neurônio artificial.



Fonte: adaptado de Haykin (2009).

em que y_k é a saída do neurônio k ; $g(\cdot)$ é a função de ativação da unidade k ; e x_k é a saída da junção aditiva, determinada por (Haykin, 2009):

$$x_k = \sum_{i=0}^m w_{ki}u_{ki} + b_k, \quad (3.18)$$

em que w_{ki} é i -ésimo peso sináptico do neurônio k ; u_{ki} é a i -ésima entrada do neurônio k ; e b_k é o viés do neurônio (do inglês, *bias*).

Na formulação original, a função de ativação usada na rede *perceptron* de única camada consiste em uma função degrau, a qual pode ser definida matematicamente como (Haykin, 2009):

$$g(x_k) = \begin{cases} 1, & x_k \geq 0, \\ 0, & x_k < 0 \end{cases}. \quad (3.19)$$

Outras funções de ativação tipicamente usadas são a função tangente hiperbólica, representada por (Schwedersky; Flesch; Danguì, 2022):

$$g(x_k) = \frac{e^{2x_k} - 1}{e^{2x_k} + 1}, \quad (3.20)$$

e a função sigmoide (Mitchell, 1997):

$$g(x_k) = \frac{1}{1 + e^{-x_k}}. \quad (3.21)$$

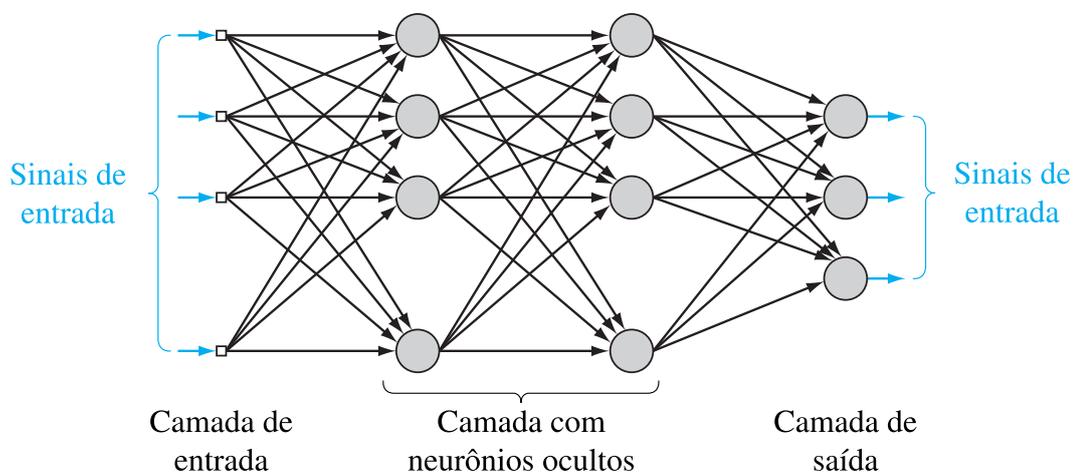
3.2.2 Arquiteturas de Redes Neurais

Atualmente, existem diversos modelos de redes neurais com a capacidade de solucionar uma variedade de problemas, incluindo tarefas de regressão, de classificação e de predição. Nesse contexto, é possível destacar as redes *feedforward*, como, por exemplo, as redes *perceptron* de

múltiplas camadas (MLPs, do inglês *MultiLayer Perceptrons*), e as redes neurais recorrentes (RNNs, do inglês *Recurrent Neural Networks*) (Haykin, 2009).

A rede neural MLP, como o nome sugere, apresenta uma arquitetura na qual os neurônios artificiais são organizados em múltiplas camadas. Em uma rede desse tipo, o sinal de entrada flui em uma única direção, percorrendo a camada de entrada, passando pelas camadas ocultas e, finalmente, chegando à camada de saída, conforme ilustrado na Figura 6. À medida que isso ocorre, cada sinal é processado em paralelo por múltiplos neurônios artificiais, similarmente ao apresentado na Seção 3.2.1 (Haykin, 2009).

Figura 6 – Representação esquemática de uma rede neural MLP.



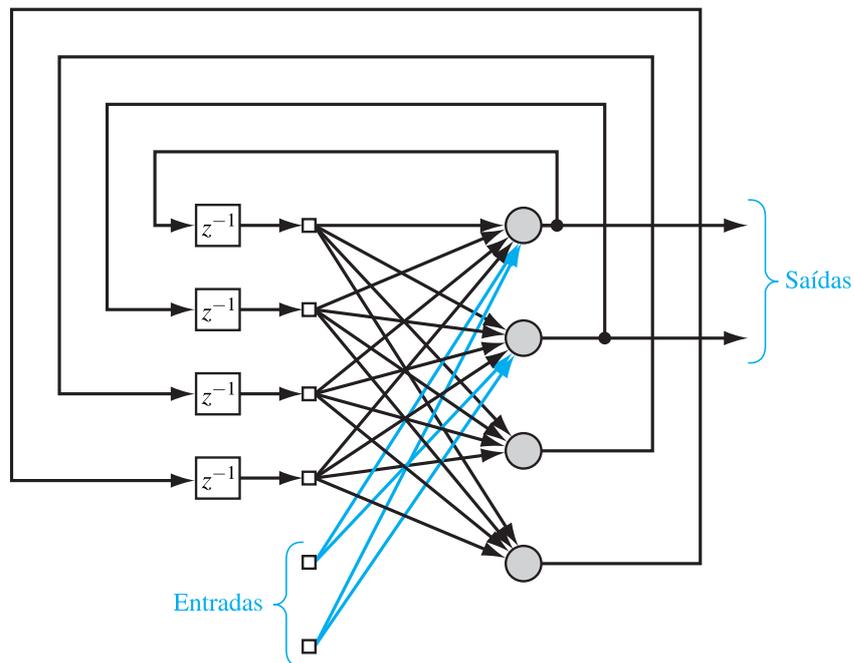
Fonte: adaptado de Haykin (2009).

A capacidade das redes MLP para executarem tarefas de regressão e classificação é um fato já reconhecido pela literatura (Mitchell, 1997). No entanto, observa-se que as redes MLP apresentam um desempenho inferior quando aplicadas a tarefas de previsão em séries temporais. Isso pode ser atribuído à arquitetura estática e desprovida de memória das MLPs. Com isso, a rede apresenta dificuldades para capturar a relação temporal entre os dados de entrada (Shoib *et al.*, 2016).

Por outro lado, a literatura reconhece a capacidade das redes neurais recorrentes para tarefas de predição envolvendo séries temporais. Uma RNN se distingue de uma rede *feedforward* por conta da existência de pelo menos uma conexão recorrente, conforme ilustrado na Figura 7. A partir dessas conexões, a rede pode aprender representações temporais, já que os neurônios artificiais são alimentados com informações atuais da entrada e também por valores passados das saídas de outras camadas (Elman, 1990; Haykin, 2009).

Por fim, é importante comentar que a área das redes neurais recorrentes é vasta, com opções de arquiteturas que vão além do escopo desta dissertação. Por exemplo, pode-se citar a rede de Hopfield (1982), a qual, embora não seja utilizada para tarefas envolvendo séries

Figura 7 – Representação de uma rede neural recorrente com neurônios ocultados.



Fonte: adaptado de Haykin (2009).

temporais, foi um dos primeiros modelos de rede neural a apresentar conexões recorrentes. Podem ser mencionadas, ainda, as redes de Jordan (1986) e Elman (1990), além das redes de memória longa de curto prazo (LSTM, do inglês *Long Short-Term Memory*), propostas em Hochreiter e Schmidhuber (1997), e a máquina de estado líquido (LSM, do inglês *Liquid State Machine*), idealizada em Maass, Natschläger e Markram (2002).

3.2.3 Redes Neurais de Estado de Eco

Para o treinamento de RNNs convencionais, é comum a utilização de algoritmos baseados nos gradientes dos erros. No entanto, trabalhos como o de Doya (1992) e Hochreiter (1998) demonstram que tais algoritmos apresentam limitações. Um dos problemas mais abordados pela literatura é o do desaparecimento dos gradientes, o qual é normalmente verificado em redes profundas e também em RNNs cujas previsões dependem dos sinais de entrada em um passado distante. Nesse caso, tais algoritmos enfrentam dificuldades para medir os efeitos dos sinais de entrada sobre a saída, prejudicando o processo de treinamento (Haykin, 2009).

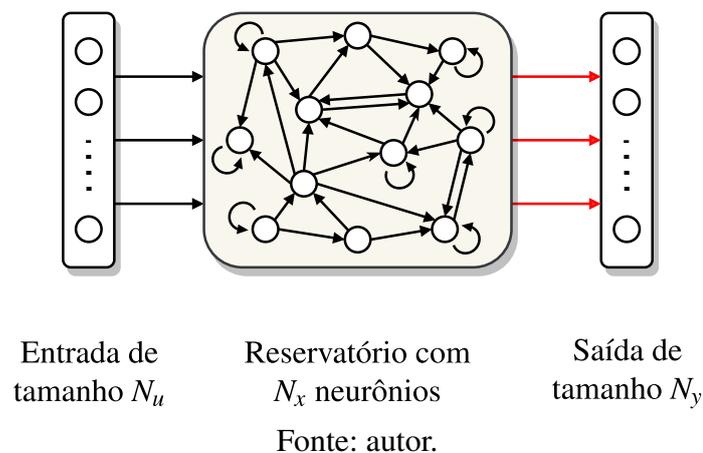
Entre as soluções propostas na literatura para contornar os desafios associados ao treinamento das RNNs por algoritmos baseados nos gradientes dos erros, pode-se destacar o paradigma da computação de reservatório (RC, do inglês *Reservoir Computing*), no qual é proposto que uma camada composta por neurônios conectados de forma recorrente, conhecido como reservatório, pode ser utilizada para mapear as entradas de uma rede neural em um espaço de alta

dimensionalidade, permitindo a realização eficiente de tarefas que dependam de um contexto temporal, sem a necessidade de utilização de algoritmos cujo treinamento sofre com o problema do desaparecimento do gradiente (Tanaka *et al.*, 2019). Nesse contexto, uma das arquiteturas que utiliza tal estratégia é a rede de estado de eco, a qual é apresentada neste documento. Por uma questão de organização, esta seção está dividida da seguinte forma. Primeiramente, na Seção 3.2.3.1, é apresentado o modelo básico de uma ESN. Em seguida, na Seção 3.2.3.3, é abordado o procedimento de treinamento das ESNs. Por sua vez, na Seção 3.2.3.2, é discutida a importância da etapa de inicialização da arquitetura de uma ESN. Por fim, na Seção 3.2.3.4, é apresentada uma visão geral sobre as alternativas para otimizar os hiperparâmetros da rede.

3.2.3.1 Modelo Básico

Proposta em Jaeger (2001), a rede neural de estado de eco é uma arquitetura de rede neural recorrente, que está inserida no paradigma de aprendizado de computação de reservatório, cuja estrutura geral é apresentada na Figura 8. As conexões em vermelho podem ser treinadas e estão relacionadas à camada de saída da rede, o que é uma característica compartilhada pelas várias redes do paradigma de RC¹. Por outro lado, as conexões em preto não são treináveis e estão associadas ao reservatório da rede, o qual consiste em uma RNN esparsamente conectada. Esse princípio permite contornar os problemas relacionados aos algoritmos baseados nos gradientes dos erros, como o desaparecimento desses gradientes. Além disso, tal abordagem reduz o custo computacional para o treinamento da rede, uma vez que as ESNs costumam empregar algoritmos mais simples para isso, como o de regressão linear (Schwedersky, 2022).

Figura 8 – Representação geral da arquitetura de uma rede neural de estado de eco.



¹ Além das redes neurais de estado de eco, a LSM e a rede *backpropagation-decorrelation* são outros exemplos de arquiteturas que compartilham o paradigma de computação de reservatório (Maass; Natschläger; Markram, 2002; Steil, 2004). Porém, as ESNs são as que mais se destacam para aplicações em engenharia (Schwedersky, 2022).

Para descrever o funcionamento de uma ESN composta por neurônios integradores com vazamento, é comum utilizar um conjunto de equações em tempo discreto, denotado por n . A primeira, está relacionada à atualização dos estados do reservatório a cada instante no tempo, podendo ser definida como (Lukoševičius, 2012):

$$\mathbf{x}[n] = (1 - \alpha)\mathbf{x}[n-1] + \alpha\tilde{\mathbf{x}}[n], \quad (3.22)$$

em que $[n]$ representa a dependência temporal do sinal em tempo discreto, ao invés da notação (t), utilizada anteriormente para sinais em tempo contínuo; $\mathbf{x}[n] \in \mathbb{R}^{N_x}$ é o vetor com os estados do reservatório no instante n ; α é a taxa de vazamento do reservatório; e $\tilde{\mathbf{x}}[n] \in \mathbb{R}^{N_x}$ é o vetor com a atualização dos estados, dado por:

$$\tilde{\mathbf{x}}[n] = g \left(s^{\text{in}} \mathbf{W}^{\text{in}} \begin{bmatrix} 1 \\ \mathbf{u}[n] \end{bmatrix} + \mathbf{W}^{\text{res}} \mathbf{x}[n-1] \right), \quad (3.23)$$

em que $\mathbf{u}[n] \in \mathbb{R}^{N_u}$ é o vetor com os sinais de entrada no instante n ; $\mathbf{W}^{\text{in}} \in \mathbb{R}^{N_x \times (1+N_u)}$ é a matriz com os pesos para as conexões que partem da entrada para o reservatório; $\mathbf{W}^{\text{res}} \in \mathbb{R}^{N_x \times N_x}$ é a matriz com os pesos para as conexões recorrentes entre os neurônios do reservatório; e s^{in} é o fator que escala os valores da matriz \mathbf{W}^{in} . Além disso, como função de ativação, $g(\cdot)$, a escolha pela função tangente hiperbólica, $\tanh(\cdot)$, é uma das mais comuns (Jordanou *et al.*, 2018).

A saída da rede ESN é obtida pela equação que define o mecanismo de leitura do reservatório, geralmente considerando uma combinação linear entre os estados do reservatório, sinais de entrada e o bias, podendo ser descrita nesse caso por (Lukoševičius, 2012):

$$\hat{\mathbf{y}}[n] = \mathbf{W}^{\text{out}} \begin{bmatrix} 1 \\ \mathbf{u}[n] \\ \mathbf{x}[n] \end{bmatrix}, \quad (3.24)$$

em que $\hat{\mathbf{y}}[n] \in \mathbb{R}^{N_y}$ é vetor contendo os valores das predições da rede no instante de tempo n ; e $\mathbf{W}^{\text{out}} \in \mathbb{R}^{N_y \times (1+N_u+N_x)}$ é a matriz com os pesos ajustáveis para a saída.

3.2.3.2 Inicialização do Modelo

A partir do que é estabelecido pelo paradigma da computação de reservatório, é necessária uma atenção especial à inicialização das matrizes de pesos do reservatório, uma vez que não é possível reajustá-las após o treinamento da ESN. Para garantir um bom desempenho para o modelo, as matrizes relacionadas à entrada de informações no reservatório, \mathbf{W}^{in} , e às conexões

recorrentes, \mathbf{W}^{res} , são geradas a partir de distribuições de probabilidade específicas, sendo a distribuição normal uma das mais escolhidas. Posteriormente, enquanto a matriz \mathbf{W}^{in} é ajustada segundo o fator de escala s^{in} , a matriz \mathbf{W}^{res} é ajustada de forma a garantir que o reservatório tenha o raio espectral, $\lambda(\mathbf{W}^{\text{res}})$, desejado (Antonelo; Camponogara; Foss, 2017; Lukoševičius, 2012).

O valor de s^{in} , juntamente com o raio espectral, tem um impacto significativo na resposta do reservatório. Por exemplo, para a escolha padrão de uma função de ativação do tipo tangente hiperbólica, um valor pequeno de s^{in} é mais recomendado para tarefas lineares, uma vez que os neurônios do reservatório irão operar em torno do ponto zero, no qual a função de ativação é aproximadamente linear. Por outro lado, valores maiores de s^{in} irão ajudar a saturar mais facilmente os neurônios, o que pode ser mais interessante para tarefas não lineares (Lukoševičius, 2012).

Por sua vez, o raio espectral, $\lambda(\mathbf{W}^{\text{res}})$, corresponde ao maior valor absoluto dos autovalores da matriz \mathbf{W}^{res} , sendo selecionado seguindo determinadas diretrizes. Nesse caso, a mais relevante está associada à escolha de um valor para $\lambda(\mathbf{W}^{\text{res}})$ que garanta a propriedade de estado de eco. Essa propriedade está diretamente relacionada à estabilidade da Equação (3.22), com um valor de raio espectral menor que um sendo suficiente para obter um reservatório estável na maioria das situações. Mais detalhes podem ser encontrados em Lukoševičius (2012).

Por fim, a taxa de vazamento $\alpha \in (0, 1]$, utilizada na Equação (3.22), regula o processo de retenção da informação dos estados anteriores do reservatório entre cada intervalo de tempo, estando relacionada à velocidade com que o reservatório é atualizado a cada iteração. Nesse caso, valores pequenos para a taxa de vazamento implicam reservatórios com maior capacidade de retenção de estímulos passados. Por outro lado, altas taxas de vazamento resultam em reservatórios com menor retenção de memória (Antonelo; Camponogara; Foss, 2017).

3.2.3.3 Treinamento

O treinamento de uma ESN consiste em calcular os valores da matriz \mathbf{W}^{out} para minimizar o erro quadrático entre as saídas reais de um sistema e as saídas previstas pela rede para um dado conjunto de treinamento de tamanho N_{treino} . Para que isso seja possível, primeiramente é necessário que o conjunto de treino seja utilizado para simular a evolução do reservatório. Em seguida, \mathbf{W}^{out} é encontrada a partir da resolução de um problema de mínimos quadrados,

representado matematicamente por (Lukoševičius, 2012):

$$\mathbf{Y}^{\text{treino}} = \mathbf{W}^{\text{out}} \begin{bmatrix} \mathbf{1} \\ \mathbf{U}^{\text{treino}} \\ \mathbf{X}^{\text{treino}} \end{bmatrix}, \quad (3.25)$$

em que $\mathbf{1} \in \mathbb{R}^{N_{\text{treino}}}$; $\mathbf{Y}^{\text{treino}} \in \mathbb{R}^{N_y \times N_{\text{treino}}}$ é matriz com as medições para a saída do sistema no conjunto de treinamento, dada por:

$$\mathbf{Y}^{\text{treino}} = \begin{bmatrix} \mathbf{y}^{\text{treino}}[n] & \mathbf{y}^{\text{treino}}[n+1] & \dots & \mathbf{y}^{\text{treino}}[n+N_{\text{treino}}-1] \end{bmatrix}; \quad (3.26)$$

$\mathbf{U}^{\text{treino}} \in \mathbb{R}^{N_u \times N_{\text{treino}}}$ é a matriz com as medições para as entradas do sistema no conjunto de treinamento, definida por:

$$\mathbf{U}^{\text{treino}} = \begin{bmatrix} \mathbf{u}^{\text{treino}}[n] & \mathbf{u}^{\text{treino}}[n+1] & \dots & \mathbf{u}^{\text{treino}}[n+N_{\text{treino}}-1] \end{bmatrix}; \quad (3.27)$$

e $\mathbf{X}^{\text{treino}} \in \mathbb{R}^{N_x \times N_{\text{treino}}}$ é a matriz que contém a evolução dos estados do reservatório durante o treinamento, representada por:

$$\mathbf{X}^{\text{treino}} = \begin{bmatrix} \mathbf{x}^{\text{treino}}[n] & \mathbf{x}^{\text{treino}}[n+1] & \dots & \mathbf{x}^{\text{treino}}[n+N_{\text{treino}}-1] \end{bmatrix}. \quad (3.28)$$

Por brevidade de notação, pode ser definida uma matriz $\mathbf{M} = [\mathbf{1} \ \mathbf{U}^{\text{treino}} \ \mathbf{X}^{\text{treino}}]^{\top}$. Com base nisso, a Equação (3.25) pode ser reescrita como:

$$\mathbf{Y}^{\text{treino}} = \mathbf{W}^{\text{out}} \mathbf{M}. \quad (3.29)$$

Para resolver o problema descrito pela Equação (3.29), há pelo menos duas opções. A primeira delas, é utilizar a seguinte solução:

$$\mathbf{W}^{\text{out}} = \mathbf{Y}^{\text{treino}} \mathbf{M}^+, \quad (3.30)$$

em que \mathbf{M}^+ é a matriz pseudo-inversa de \mathbf{M} . Entretanto, em Lukoševičius (2012) é citado que essa solução pode ser empregada quando o objetivo for treinar a ESN sem regularização. No entanto, tal abordagem eleva as chances de ocorrer um sobreajuste do modelo aos dados.

A segunda possibilidade é treinar a ESN considerando uma regressão com regularização, o que pode ser feito pelo método de Thikonov, no qual a norma L_2 é utilizada. Nesse caso, o objetivo é evitar que ocorra um sobreajuste do modelo aos dados, o que pode levar a soluções que se ajustam bem ao conjunto de treinamento, porém, não apresentam o mesmo resultado quando testadas em novos conjuntos. Para isso, é aplicada uma penalidade aos pesos com maiores valores

absolutos na matriz \mathbf{W}^{out} . Na prática, isso equivale a resolver a seguinte equação (Lukoševičius, 2012):

$$\mathbf{W}^{\text{out}} = \arg \min_{\mathbf{W}^{\text{out}}} \frac{1}{N_y} \sum_{i=1}^{N_y} \left(\sum_{n=1}^{N_{\text{treino}}} (\hat{y}_i[n] - y_i^{\text{treino}}[n])^2 + \beta \|\mathbf{w}_i^{\text{out}}\|^2 \right), \quad (3.31)$$

cuja solução analítica é dada por:

$$\mathbf{W}^{\text{out}} = \mathbf{Y}^{\text{treino}} \mathbf{M}^{\top} (\mathbf{M} \mathbf{M}^{\top} + \beta \mathbf{I})^{-1}, \quad (3.32)$$

em que $\mathbf{w}_i^{\text{out}}$ é o vetor com os pesos para i -ésima linha da matriz \mathbf{W}^{out} ; $\hat{y}_i[n]$ é a i -ésima saída predita da rede para o instante n ; $y_i^{\text{treino}}[n]$ é a i -ésima saída real do sistema amostrada no instante n do conjunto de treinamento; \mathbf{I} é uma matriz identidade de dimensões apropriadas; $\|\cdot\|$ é a norma euclidiana; e β é o coeficiente de regularização.

Por fim, um aspecto que exige atenção está relacionado ao estado transitório inicial do reservatório. Como os estados são inicializados de forma aleatória em $n = 0$ durante o treinamento, uma boa prática é descartar os primeiros pontos sempre que possível. Durante esse processo, a quantidade de pontos a serem descartados varia conforme a dinâmica de atualização do reservatório, sendo normalmente da ordem de dezenas ou centenas (Lukoševičius, 2012).

3.2.3.4 Otimização de Hiperparâmetros

Como a maioria dos algoritmos de aprendizado de máquina, as ESNs possuem vários hiperparâmetros que precisam ser cuidadosamente ajustados para obter um melhor desempenho do modelo. Para isso, a literatura menciona algumas opções, como o método de busca em grade e a otimização bayesiana. Além disso, existem diversas outras alternativas, que não são exploradas nesta dissertação, como algoritmos evolucionários e o algoritmo de otimização por enxame de partículas (Thiede; Parlitz, 2019; Yang; Shami, 2020; Racca; Magri, 2021).

O algoritmo de busca em grade, também conhecido como *grid search*, é um método que avalia os efeitos de cada combinação de hiperparâmetros sobre o desempenho de um determinado modelo em um espaço de busca. A vantagem da busca em grade é que o método pode ser facilmente implementado e executado em paralelo. No entanto, a sua desvantagem está no tempo de execução para resolver problemas cujo espaço de busca apresente alta dimensionalidade e existem diversas combinações de parâmetros a serem testadas (Yang; Shami, 2020).

Para as ESNs, o problema da dimensionalidade tende a ser frequente, uma vez que a rede possui ao menos cinco hiperparâmetros para serem otimizados, como a taxa de vazamento do reservatório, α ; o número de neurônios, N_x ; o raio espectral, $\lambda(\mathbf{W}^{\text{res}})$; o coeficiente de regularização, β ; e o fator de escalonamento para as entradas, s^{in} (Thiede; Parlitz, 2019). Desse modo, fazer o uso de algoritmos como a otimização bayesiana pode ser uma estratégia mais

adequada para ajustar a arquitetura da rede com um esforço computacional menor que o exigido por uma busca em grade.

Tanto a busca em grade quanto a otimização bayesiana visam encontrar o mínimo global de uma função objetivo, $J(\cdot)$, em um dado espaço de busca. No entanto, a principal distinção entre essas duas abordagens reside na capacidade da otimização bayesiana de usar informações acumuladas ao longo de um processo iterativo para determinar os argumentos que minimizam a função objetivo. Para isso, o algoritmo modela o problema em torno de uma função substituta (do inglês, *surrogate function*) e a função de aquisição (Snoek; Larochelle; Adams, 2012; Greenhill *et al.*, 2020).

A função substituta é responsável por construir um modelo probabilístico da função $J(\cdot)$. Normalmente, é utilizado um modelo de processo gaussiano (GP, do inglês *Gaussian Process*), o qual pode ser compreendido como uma extensão da distribuição gaussiana multivariada para um processo estocástico com dimensões que tendem ao infinito. Nesse contexto, qualquer combinação finita de dimensões resulta em uma distribuição gaussiana. Desse modo, similar à distribuição gaussiana, a qual é uma distribuição associada a uma variável aleatória, definida por sua média e covariância, um GP é uma distribuição associada a funções, sendo especificado por sua função média e por uma função de covariância. Matematicamente, isso pode ser representado por (Brochu; Cora; Freitas, 2010):

$$J(\mathbf{z}) \sim \text{GP}(m(\mathbf{z}), \text{cov}(\mathbf{z}, \mathbf{z}')), \quad (3.33)$$

em que $\mathbf{z} \in \mathbb{R}^D$ e $\mathbf{z}' \in \mathbb{R}^D$ são os vetores, de dimensão D , que contêm as entradas da função $J(\cdot)$, porém em pontos diferentes do espaço; $m(\cdot)$ é a função média, usualmente assumida como sendo igual a zero, mas que pode assumir outras formas para melhorar a interpretabilidade do modelo, conforme detalhado em Rasmussen e Williams (2005); e $\text{cov}(\cdot, \cdot)$ é a função de covariância, também conhecida como *kernel*. Normalmente, tal função é escolhida para refletir o conhecimento prévio a respeito da função de interesse, como a periodicidade ou a sua diferenciabilidade. Atualmente, existem diversas maneiras de construir as funções de covariância, sendo uma delas por meio da função exponencial quadrática com determinação automática de relevância (ARD, do inglês *Automatic Relevance Determination*, dada por (Garnett *et al.*, 2010; Brochu; Cora; Freitas, 2010):

$$\text{cov}(\mathbf{z}, \mathbf{z}') = \sigma_{\text{cov}}^2 \exp\left(-\frac{1}{2}d^2(\mathbf{z}, \mathbf{z}')\right), \quad (3.34)$$

e através da função ARD *Matérn 5/2*, representada matematicamente por (Snoek; Larochelle; Adams, 2012):

$$\text{cov}(\mathbf{z}, \mathbf{z}') = \sigma_{\text{cov}}^2 \left(1 + \sqrt{5d^2(\mathbf{z}, \mathbf{z}')}\right) \exp\left(-\sqrt{5d^2(\mathbf{z}, \mathbf{z}')}\right) \quad (3.35)$$

nas quais σ_{cov}^2 é a amplitude da função; e $d^2(\cdot, \cdot)$ é a distância entre \mathbf{z} e \mathbf{z}' , podendo ser calculada a partir da seguinte equação (Snoek; Larochelle; Adams, 2012):

$$d^2(\mathbf{z}_i, \mathbf{z}_j) = \sum_{m=1}^D \frac{(z_m - z'_m)^2}{\ell_m}, \quad (3.36)$$

em que ℓ_m é a escala de comprimento da medida de distância, cujo valor determina a influência de cada variável em \mathbf{z} sobre a função de covariância. Mais detalhes são apresentados em Rasmussen e Williams (2005), por exemplo.

Para minimizar a função objetivo, a otimização bayesiana faz o uso das informações referentes às entradas e à saída de $J(\cdot)$. É nesse contexto que as funções de aquisição exercem um papel relevante, estimando quais valores em \mathbf{z} tendem a ser o ponto ótimo de $J(\cdot)$. Para isso, a função de aquisição, baseada nas incertezas do GP, deve ser otimizada (Rasmussen; Williams, 2005; Brochu; Cora; Freitas, 2010).

Atualmente, existem diversas opções de escolha para as funções de aquisição. Uma delas, é a função de probabilidade de melhoria, $\text{PI}(\cdot)$, dada por (Snoek; Larochelle; Adams, 2012; Greenhill *et al.*, 2020):

$$\text{PI}(\mathbf{z}) = \Phi \left(\frac{\mu_{\text{GP}}(\mathbf{z}) - J(\mathbf{z}^*) - \xi}{\sigma_{\text{GP}}(\mathbf{z})} \right), \quad (3.37)$$

em que $\mu_{\text{GP}}(\cdot)$ e $\sigma_{\text{GP}}^2(\cdot)$ são a média e a variância da distribuição gaussiana no ponto \mathbf{z} ; \mathbf{z}^* é o melhor ponto calculado até a presente iteração; ξ é um hiperparâmetro que ajusta o compromisso entre o aproveitamento e exploração no processo de busca, o qual é abordado com mais detalhes em Brochu, Cora e Freitas (2010); e $\Phi(\cdot)$ é a função de distribuição acumulada normal.

Outra opção de função de aquisição é a de melhoria esperada, $\text{EI}(\cdot)$, descrita por (Brochu; Cora; Freitas, 2010):

$$\text{EI}(\mathbf{z}) = \begin{cases} (\mu_{\text{GP}}(\mathbf{z}) - J(\mathbf{z}^*))\Phi(Z) + \sigma_{\text{GP}}(\mathbf{z})\varphi(Z), & \sigma_{\text{GP}} > 0, \\ 0, & \sigma_{\text{GP}} = 0, \end{cases} \quad (3.38)$$

em que $\varphi(\cdot)$ é a função de distribuição de probabilidade; e Z é definida por:

$$Z = \frac{\mu_{\text{GP}}(\mathbf{z}) - J(\mathbf{z}^*)}{\sigma_{\text{GP}}(\mathbf{z})}. \quad (3.39)$$

Em Greenhill *et al.* (2020), são citadas as diferenças entre as Equações (3.37) e (3.38). Por exemplo, a função de probabilidade de melhoria visa, como o nome sugere, explorar regiões onde existe uma maior chance de melhoria em relação ao ponto ótimo atual. Por outro lado, a função de melhoria esperada considera não apenas a probabilidade disso, mas também a magnitude esperada dessa melhoria.

3.3 IDENTIFICAÇÃO DE SISTEMAS DINÂMICOS

Normalmente, tarefas como projeto de controladores, otimização, supervisão e detecção de falhas em uma planta estão relacionadas, direta ou indiretamente, ao conhecimento do modelo que representa o sistema. Uma forma de obter modelos para sistemas nos quais a modelagem fenomenológica é desafiadora consiste no uso de técnicas de identificação de sistemas, as quais permitem a obtenção de modelos com base nos dados de entrada e de saída adquiridos durante o funcionamento do processo (Nelles, 2001).

Para identificar um sistema dinâmico a partir de dados do processo, uma série de etapas costuma ser seguida. De forma geral, as principais etapas utilizadas são (Nelles, 2001):

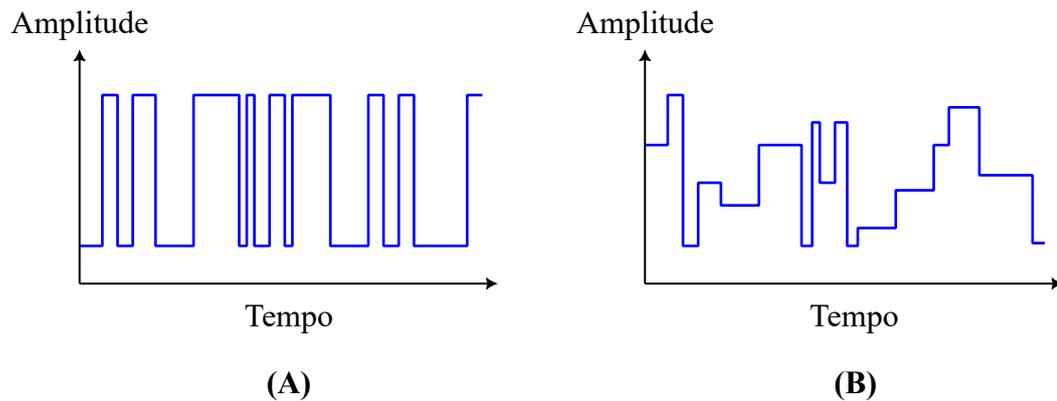
1. definição das entradas e das saídas do modelo a ser identificado;
2. escolha dos sinais de excitação;
3. escolha da estrutura do modelo;
4. treinamento do modelo;
5. validação do modelo.

A escolha das entradas e das saídas do modelo é um procedimento que deve considerar a sua finalidade. Para fins de controle, é necessário escolher como entradas e saídas as variáveis manipuladas e controladas, respectivamente (Nelles, 2001). Por exemplo, no contexto do tratamento de hipertermia magnética, é sabido que a frequência e a amplitude do campo magnético desempenham um papel significativo no processo de aquecimento. Portanto, é possível considerar a realização de um ensaio de identificação no qual essas variáveis sejam selecionadas como as entradas do modelo, enquanto a temperatura da região do tumor pode ser definida como a saída.

A determinação dos sinais de excitação é uma decisão que tende a demandar uma atenção especial do projetista. Para sistemas lineares, é comum empregar um sinal baseado em uma sequência binária pseudo-aleatória (PRBS, do inglês *Pseudo Random Binary Sequence*), devido a sua capacidade de excitar o processo em todas as frequências. Por outro lado, o uso de um PRBS para excitar sistemas não lineares é inadequado, pois esse tipo de sinal não permite capturar informações suficientes do comportamento não linear do processo ao longo de toda sua faixa de operação. Nesse caso, a recomendação é utilizar um sinal baseado em uma sequência binária pseudo-aleatória modulada em amplitude (APRBS, do inglês *Amplitude Modulated Random Binary Sequence*), o que permite excitar a planta em diversos pontos de operação (Nelles, 2001). Com base nisso, na Figura 9 são ilustrados os sinais PRBS e APRBS.

O projeto dos sinais de excitação costuma seguir algumas diretrizes. Uma delas, diz respeito ao tempo mínimo no qual o patamar de uma sequência deve ser mantida, t_h . Para uma

Figura 9 – Exemplos de sinais de excitação: (A) PRBS; e (B) APRBS.



Fonte: autor.

sequência APRBS, em Nelles (2001) é recomendado que o valor de t_h deve ser próximo da maior constante de tempo do processo. Por sua vez, para um PRBS é recomendada a seguinte heurística (Aguirre, 2007):

$$\frac{\tau_{\min}}{10} \leq t_h \leq \frac{\tau_{\min}}{3}, \quad (3.40)$$

em que τ_{\min} é a menor constante de tempo do sistema. Para processos lineares, a recomendação é que o valor de t_h seja o mais próximo de possível do limite inferior da Equação (3.40). Por outro lado, para sistemas não lineares, o valor de t_h deve ser mais próximo do limite superior apresentado na heurística.

A escolha da estrutura do modelo a ser identificado depende de diversos fatores, como a complexidade do problema e o tempo disponível para sua execução. Nesse contexto, as opções são vastas e incluem, por exemplo, estruturas baseadas em funções de transferência, modelos de Hammerstein-Wiener e redes neurais recorrentes (Ljung, 1999; Racca; Magri, 2021).

Na etapa de estimação do modelo, ocorre o ajuste de seus parâmetros a partir de uma determinada função objetivo. Após isso, uma última etapa é avaliar o desempenho do modelo treinado em um novo conjunto de dados. Se o modelo não apresentar um desempenho adequado à tarefa avaliada, é necessária revisão das etapas de identificação, podendo ser realizadas alterações em todas as etapas anteriores. Caso contrário, o modelo pode ser usado para a finalidade desejada (Nelles, 2001; Schwedersky, 2022).

3.4 COMENTÁRIOS FINAIS

Neste capítulo, foram apresentados os fundamentos que permitiram o desenvolvimento da metodologia discutida nesta dissertação. Na Seção 3.1, foram explorados os aspectos gerais do tratamento por hipertermia magnética, proporcionando uma compreensão do funcionamento dessa técnica. Isso se mostrou fundamental para a construção dos modelos numéricos apresenta-

dos no Capítulo 4. Além disso, nas Seções 3.2 e 3.3, foram discutidos os aspectos relacionados às redes neurais artificiais e ao desenvolvimento de modelos por meio da identificação de sistemas, respectivamente. Esses tópicos foram importantes no desenvolvimento discutido nos Capítulos 5 e 6.

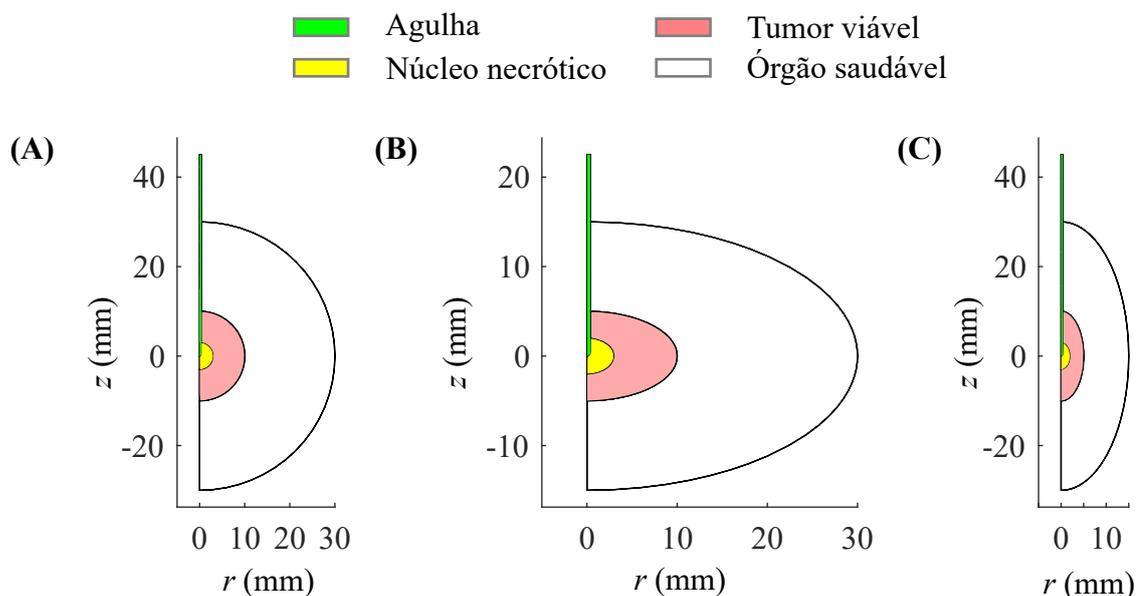
4 MODELO NUMÉRICO DO SISTEMA

Neste capítulo, são apresentados os modelos numéricos considerados para o desenvolvimento desta dissertação, assim como as equações que permitem modelar as etapas consideradas para o tratamento de hipertermia magnética. Desse modo, na Seção 4.1, são detalhadas as geometrias dos modelos numéricos e a sua implementação. Em seguida, na Seção 4.2, são apresentadas as equações diferenciais parciais que descrevem cada fenômeno incorporado aos modelos. Na Seção 4.3, é descrito o procedimento adotado para verificar a implementação das equações governantes aos modelos. Por fim, na Seção 4.4, são apresentados os comentários finais a respeito da implementação e verificação dos modelos numéricos.

4.1 GEOMETRIA E IMPLEMENTAÇÃO DO MODELO

Os tumores costumam exibir múltiplas regiões. No centro, é comum encontrar um núcleo necrótico, onde não existe um sistema vascular funcional e nem células vivas. Por outro lado, a região adjacente ao núcleo necrótico, conhecida como zona viável, é caracterizada pela presença de um sistema vascular funcional e células vivas (Baxter; Jain, 1989; Steuperaert *et al.*, 2017). Com base nisso e para avaliar o potencial das redes de estado de eco como modelos de predição em diferentes condições de tratamento, este trabalho propõe os três modelos numéricos ilustrados na Figura 10, ou seja, um modelo esférico, um modelo elipsoidal oblato e um modelo elipsoidal prolato, todos com uma geometria axissimétrica em torno do eixo z .

Figura 10 – Geometrias dos modelos numéricos: (A) esférico; (B) elipsoidal oblato; e (C) elipsoidal prolato.



Fonte: autor.

Quanto às dimensões de cada modelo, foram adotados os seguintes valores. No modelo esférico, o tumor foi representado como uma esfera com um diâmetro de 20 mm, apresentando um núcleo necrótico no centro com um diâmetro de 6 mm. Além disso, o tumor foi cercado por tecido saudável, assumido como um fígado humano, representado por uma seção esférica com um diâmetro de 60 mm, escolhido para garantir que as condições de fronteira não alterassem os resultados das simulações significativamente. Para injetar o nanofluido no centro do tumor, foi assumido o uso de uma agulha com um diâmetro de 0,8 mm, comum a todos os modelos. No modelo oblato, o eixo maior de todos os domínios conserva seus respectivos valores de diâmetro, porém as dimensões do eixo menor foram definidas como: 10 mm para o tumor, 4 mm para o núcleo necrótico e 30 mm para o fígado. A partir desses valores, foi possível construir o modelo prolato, o qual pode ser visto como uma versão do modelo oblato rotacionada em 90° .

Os modelos descritos nesta seção e as equações governantes apresentadas na Seção 4.2 foram resolvidos numericamente por meio do método dos elementos finitos, utilizando o software COMSOL Multiphysics. Por sua vez, todos os modelos foram implementados a partir de suas geometrias axissimétricas, o que permite simular um modelo tridimensional com menor custo computacional. Além disso, a opção de malha “extremamente refinada” foi selecionada com base em testes iniciais de convergência. Nessa configuração, o modelo esférico é composto por uma malha de 13792 elementos, o modelo oblato é composto por 17428 elementos e o modelo prolato utiliza uma malha de 16872 elementos. Por fim, foi escolhido um passo de tempo fixo de 10 s para realizar as simulações numéricas para a física de biotransferência de calor (detalhada na Seção 4.2.4), com o intuito de garantir que as temperaturas sejam amostradas nos instantes desejados. Por outro lado, os fenômenos relacionados à distribuição de nanopartículas foram simulados com passos de tempo automaticamente definidos pelo software, para acelerar a convergência sem prejudicar a tolerância dos resultados numéricos obtidos.

4.2 EQUAÇÕES GOVERNANTES

Nesta seção, são apresentados os modelos matemáticos considerados para descrever as três etapas básicas do tratamento de hipertermia magnética. Por isso, foi adotada a seguinte organização. Primeiramente, na Seção 4.2.1 é modelada a física do transporte de nanofluido para a região do tumor via injeções intratumorais. Em seguida, na Seção 4.2.2 são descritas as equações que permitem modelar a velocidade e a pressão intersticial de nanofluido. Além disso, na Seção 4.2.3 é apresentado o modelo matemático usado para estimar a distribuição de nanopartículas no tumor. Por fim, na Seção 4.2.4 é apresentado o modelo de biotransferência de calor, utilizado para calcular o perfil de temperatura ao longo dos tecidos durante o tratamento.

4.2.1 Escoamento Laminar

Ao considerar uma injeção intratumoral, a velocidade na qual um nanofluido escoar ao longo de uma agulha pode ser modelada usando a equação de Navier-Stokes. Desse modo, ao admitir um escoamento em regime permanente, supor um fluido incompressível e desprezar as forças do corpo, tem-se que (Suleman; Riaz, 2020a):

$$\rho_{nf}(\mathbf{v}_{nf} \cdot \nabla) \mathbf{v}_{nf} = \nabla \cdot \left[-p_{nf} \mathbf{I} + \mu_{nf} \left(\nabla \mathbf{v} + (\nabla \mathbf{v})^\top \right) \right], \quad (4.1)$$

sendo:

$$\rho_{nf} = \phi_{\text{máx}} \rho_{\text{MNP}} + (1 - \phi_{\text{máx}}) \rho_s, \quad (4.2)$$

enquanto a equação da continuidade é dada por:

$$\nabla \cdot (\rho_{nf} \mathbf{v}_{nf}) = 0, \quad (4.3)$$

em que \mathbf{v}_{nf} , ρ_{nf} , p_{nf} , μ_{nf} são a velocidade, a massa específica, a pressão e a viscosidade dinâmica do nanofluido, respectivamente; $\phi_{\text{máx}}$ é a fração de nanopartículas no nanofluido, admitida como sendo igual a 0,0177; ρ_{MNP} é a massa específica das nanopartículas magnéticas; ρ_s é a massa específica do solvente utilizado para diluir as nanopartículas; e ∇ denota o gradiente de uma variável em coordenadas cilíndricas.

Para resolver as Equações (4.1) e (4.3), foram impostas três condições de contorno. A primeira, descreve uma condição de não deslizamento nas paredes da agulha, podendo ser representada matematicamente por (Suleman; Riaz, 2020b):

$$v_{nf} = 0. \quad (4.4)$$

A segunda condição de contorno define a pressão estática na entrada da agulha, podendo ser representada por (Suleman; Riaz, 2020a):

$$\mathbf{n}^\top \left[-p_{nf} \mathbf{I} + \mu_{nf} \left(\nabla \mathbf{v} + (\nabla \mathbf{v})^\top \right) \right] \mathbf{n} = 0. \quad (4.5)$$

Por sua vez, a última condição de contorno define a pressão na saída da agulha, conforme é dado por (Suleman; Riaz, 2020a):

$$\left[-p_{nf} \mathbf{I} + \mu_{nf} \left(\nabla \mathbf{v} + (\nabla \mathbf{v})^\top \right) \right] \mathbf{n} = -p_0 \mathbf{n}, \quad (4.6)$$

em que p_0 denota, nesse caso, a pressão intersticial do tumor no local de injeção; e \mathbf{n} é o vetor normal.

Finalmente, foram assumidos os valores listados na Tabela 1 para simular o fenômeno descrito nesta seção. Como solvente, foi considerado o uso de uma solução salina, com base no trabalho de Suleman e Riaz (2020b). Além disso, foi admitida a hipótese de que as nanopartículas são feitas de magnetita (Fe_3O_4). Uma escolha comum para estudos de hipertermia magnética.

Tabela 1 – Valores para as propriedades pertencentes à física de escoamento laminar.

Propriedade	Valor
Viscosidade do nanofluido, μ_{nf} (Pa s)	$1,23 \times 10^{-3}$
Massa específica do solvente, ρ_s (kg m^{-3})	2170
Massa específica da magnetita, ρ_{MNP} (kg m^{-3})	5180

Fonte: adaptado de Suleman e Riaz (2020b) e Tang *et al.* (2021).

4.2.2 Transporte Intersticial de Nanofluido

A velocidade do nanofluido nos interstícios exerce influência sobre o perfil de distribuição das nanopartículas ao longo da região do tumor. Para modelar tal velocidade, é comum utilizar o par de equações apresentado a seguir. Nesse caso, a Equação (4.7) representa a lei de Darcy para meios porosos, enquanto a Equação (4.8) descreve a lei de Starling acoplada à equação da continuidade (Baxter; Jain, 1989; Steuperaert *et al.*, 2017):

$$\mathbf{v}_i = -\kappa_i \nabla p_i \quad (4.7)$$

$$\nabla \cdot \mathbf{v}_i = \begin{cases} 0, & i = 1, \\ \theta_V, & i = 2, \\ \theta_V - \theta_L, & i = 3, \end{cases} \quad (4.8)$$

em que \mathbf{v}_i é a velocidade intersticial do nanofluido; κ_i é a condutividade hidráulica; p_i é a pressão intersticial; θ_V é a vazão volumétrica para fora da vasculatura por unidade de volume do tecido, dada por (Baxter; Jain, 1989):

$$\theta_V = \frac{L_p S}{V} [p_V - p_i - \sigma_T (\pi_V - \pi_i)], \quad (4.9)$$

em que L_p é a condutividade hidráulica da parede microvascular; S/V é a área de superfície por unidade de volume para transporte no tumor; p_V é a pressão vascular; σ_T é o coeficiente médio de reflexão osmótica para proteínas plasmáticas; π_V é a pressão osmótica do plasma sanguíneo; π_i é a pressão osmótica do fluido intersticial; L_{pL} é a condutividade hidráulica da parede linfática; e θ_L é vazão volumétrica para dentro dos vasos linfáticos, descrita por (Baxter; Jain, 1989):

$$\theta_L = \frac{L_{pL} S_L}{V} (p_i - p_L), \quad (4.10)$$

em que L_{pL} é a condutividade hidráulica da parede linfática e p_L é a pressão hidrostática dos vasos linfáticos. Além disso, os seguintes valores do índice i são usados: $i = 1$ para denotar o núcleo necrótico, $i = 2$ para denotar o tumor viável e $i = 3$ para denotar o tecido saudável.

Para resolver as Equações (4.7) e (4.8), foram impostas quatro condições de contorno. A primeira, é responsável por modelar a continuidade da pressão e da velocidade intersticial entre todos os domínios. Desse modo, na fronteira entre o núcleo necrótico e o zona viável do tumor, têm-se que (Sefidgar; Bashooki; Shojaee, 2020):

$$-\kappa_1 \nabla p_1|_{\Psi} = -\kappa_2 \nabla p_2|_{\Psi}, \quad (4.11)$$

$$p_1|_{\Psi} = p_2|_{\Psi}, \quad (4.12)$$

enquanto a continuidade de p_i e v_i na fronteira entre o tumor e o órgão saudável é dada por (Sefidgar; Bashooki; Shojaee, 2020):

$$-\kappa_2 \nabla p_2|_{\Upsilon} = -\kappa_3 \nabla p_3|_{\Upsilon}, \quad (4.13)$$

$$p_2|_{\Upsilon} = p_3|_{\Upsilon}, \quad (4.14)$$

em que Ψ é a interface entre o núcleo necrótico e a zona viável do tumor e Υ é a interface entre a zona viável do tumor e o órgão saudável.

A segunda condição de contorno, por sua vez, define o isolamento das paredes da agulha, descrita por (Suleman; Riaz, 2020b):

$$-\mathbf{n} \cdot (\rho_{nf} \mathbf{v}_i) = 0, \quad (4.15)$$

em que \mathbf{n} é o vetor normal à superfície. Além disso, a terceira condição de contorno modela vazão mássica de nanopartículas para o centro do tumor, definida por (Suleman; Riaz, 2020b):

$$-\mathbf{n} \cdot (\rho_{nf} \mathbf{v}_i) = \frac{q_{nf} \rho_{nf}}{\pi r_{agulha}^2}. \quad (4.16)$$

em que r_{agulha} é o raio da agulha; e q_{nf} é a vazão volumétrica de nanofluido, assumida como sendo igual a $5 \mu\text{L min}^{-1}$.

Como a quarta e última condição de contorno, é imposto que a pressão nas bordas do órgão saudável, Ξ , é igual a zero, o que pode ser expresso como (Sefidgar; Bashooki; Shojaee, 2020):

$$p_i|_{\Xi} = 0. \quad (4.17)$$

Por fim, foram utilizados os valores apresentados na Tabela 2 para simular o fenômeno descrito nesta seção. Os parâmetros listados para o tumor são válidos tanto para o seu núcleo necrótico quanto para a sua zona viável. Os valores para o órgão saudável são uma aproximação para as propriedades de um fígado saudável.

Tabela 2 – Valores para as propriedades pertencentes à física de transporte intersticial de nanofluido.

Propriedade	Valor
Condutividade hidráulica da parede vascular, L_p (cm mmHg ⁻¹ s ⁻¹)	
Tumor	$2,8 \times 10^{-7}$
Órgão saudável	$0,36 \times 10^{-7}$
Condutividade hidráulica do tecido, κ_i (cm ² mmHg ⁻¹ s ⁻¹)	
Tumor	$4,13 \times 10^{-8}$
Órgão saudável	$8,53 \times 10^{-9}$
Área de superfície por unidade de volume para transporte no tumor, S/V (cm ⁻¹)	
Tumor	200
Órgão saudável	70
Coeficiente médio de reflexão osmótica para as proteínas plasmáticas, σ_T (adimensional)	
Tumor	0,82
Órgão saudável	0,91
Pressão osmótica do fluido intersticial no tecido, π_i (mmHg)	
Tumor	15
Órgão saudável	10
Pressão osmótica do plasma sanguíneo, π_V (mmHg)	
Tumor	20
Órgão saudável	20
Pressão vascular, p_V (mmHg)	
Tumor	15,6
Órgão saudável	15,6
Pressão hidrostática dos vasos linfáticos, p_L (mmHg)	
Tumor	–
Órgão saudável	0
Coeficiente de filtração linfática, $L_{pL}S_L/V$ (mmHg ⁻¹ s ⁻¹)	
Tumor	–
Órgão saudável	$1,33 \times 10^{-5}$

Fonte: adaptado de Baxter e Jain (1989).

4.2.3 Distribuição de Nanopartículas

Após simular os fenômenos descritos nas Seções 4.2.1 e 4.2.2, pode-se estimar a distribuição das nanopartículas ao longo da região do tumor por meio da equação de convecção-difusão, dada por (Baxter; Jain, 1989; Mahesh; Singh; Talukdar, 2022):

$$\frac{\partial C_i}{\partial t} = D_{\text{eff}} \nabla^2 C_i - \nabla(\mathbf{v}_i c_i) + (\Theta_V - \Theta_L), \quad (4.18)$$

em que C_i é a concentração de nanopartículas no meio; D_{eff} é o coeficiente de difusão efetivo das nanopartículas na região; Θ_V é o termo que considera a balanço de soluto entre o sistema vascular e a região analisada; e Θ_L é o termo que considera o balanço de soluto entre o sistema linfático e a região analisada.

O coeficiente de difusão das nanopartículas em cada órgão pode ser estimado a partir de um par de equações. Primeiramente, estima-se o seu valor para a fase líquida por meio da Equação (4.19). Por fim, calcula-se o seu valor efetivo por meio da Equação (4.20) (Tang *et al.*, 2022):

$$D_0 = \frac{k_B T_{\text{abs}}}{6\pi\mu_{\text{nf}}r_{\text{MNP}}}, \quad (4.19)$$

$$D_{\text{eff}} = \frac{3\varepsilon}{2-\varepsilon} D_0, \quad (4.20)$$

em que D_0 é o coeficiente de difusão das nanopartículas na fase líquida; D_{eff} é o coeficiente de difusão efetivo das nanopartículas nos órgãos; $T_{\text{abs}} = 310,15 \text{ K}$ é a temperatura absoluta do meio; r_{MNP} é o valor admitido para o raio das nanopartículas magnéticas; e ε é a porosidade de cada região. Neste trabalho, assumiu-se uma porosidade de 0,4 para o núcleo necrótico, 0,3 para a zona viável do tumor e 0,26 para o órgão saudável. Além disso, foi considerado que as nanopartículas possuem um raio de 6 nm (Tang *et al.*, 2021; Mahesh; Singh; Talukdar, 2022).

Os termos Θ_V e Θ_L podem ser calculados a partir do seguinte par de equações (Baxter; Jain, 1989; Steuperaert *et al.*, 2017):

$$\Theta_V = \begin{cases} 0, & i = 1, \\ \theta_V(1 - \sigma_s)C_p + P_c(C_p - C_i) \frac{S}{V} \frac{\text{Pe}}{e^{\text{Pe}} - 1}, & i = 2, 3, \end{cases} \quad (4.21)$$

$$\Theta_L = \begin{cases} 0, & i = 1, 2, \\ \theta_L C_i, & i = 3, \end{cases} \quad (4.22)$$

em que σ_s é o coeficiente de reflexão do soluto; C_p é a concentração de soluto no plasma sanguíneo, usualmente assumida como sendo igual a zero em casos que envolvem apenas injeções intratumorais; P_c é a permeabilidade dos vasos capilares; e Pe é o número de Péclet, dado por (Steuperaert *et al.*, 2017):

$$\text{Pe} = \frac{\theta_V(1 - \sigma_s)}{P_c(S/V)}. \quad (4.23)$$

Para estimar os valores de σ_s e P_c , podem ser empregadas as seguintes equações (Mahesh; Singh; Talukdar, 2022):

$$\sigma_f = 1 - (1 - \gamma)^2 \left[2 - (1 - \gamma)^2 \right] \left(1 - \frac{2}{3}\gamma - 0,163\gamma^3 \right), \quad (4.24)$$

$$P_c = \varepsilon_c (1 - \gamma)^2 \left(1 - 2,11\gamma + 2,09\gamma^3 - 0,95\gamma^5 \right) \frac{D_0}{L_c}, \quad (4.25)$$

em que ε_c é a porosidade da parede dos vasos capilares; $\gamma = r_{MNP}/r_p$ é a razão entre o valor do raio da nanopartícula e o raio dos poros dos vasos capilares, denotado por r_p ; e L_c é a espessura da parede dos vasos capilares.

Como condição inicial para a Equação (4.18), foi imposto que a concentração de nanopartículas é igual a zero em todos os domínios. Além disso, foram impostas duas condições de contorno. A primeira, estabelece a concentração de nanopartículas que é transportada até o centro do tumor por meio da injeção intratumoral, $C_{m\acute{a}x}$. Para isso, utilizou-se a seguinte equação (Suleman; Riaz, 2020b):

$$C_{m\acute{a}x} = \frac{\rho_{nf}\phi_{m\acute{a}x}}{M_{MNP}}, \quad (4.26)$$

em que $M_{MNP} = 231,53 \text{ g mol}^{-1}$ é a massa molar da nanopartícula de magnetita. Com isso, foi possível assumir que $C_{m\acute{a}x} \approx 170 \text{ mol m}^{-3}$.

A segunda condição de contorno estabelece a continuidade do perfil de concentração ao longo de todos os domínios. Logo, na fronteira entre o núcleo necrótico e o zona viável do tumor, Ψ , têm-se que (Sefidgar; Bashooki; Shojaee, 2020):

$$\left[D_{\text{eff}} \nabla^2 C_1 - \nabla \cdot (\mathbf{v}_1 C_1) \right]_{\Psi} = \left[D_{\text{eff}} \nabla^2 C_2 - \nabla \cdot (\mathbf{v}_2 C_2) \right]_{\Psi}, \quad (4.27)$$

$$C_1|_{\Psi} = C_2|_{\Psi}, \quad (4.28)$$

enquanto a continuidade de C_i na fronteira entre o tumor e o órgão saudável, Υ , é garantida por:

$$\left[D_{\text{eff}} \nabla^2 C_2 - \nabla \cdot (\mathbf{v}_2 C_2) \right]_{\Upsilon} = \left[D_{\text{eff}} \nabla^2 C_3 - \nabla \cdot (\mathbf{v}_3 C_3) \right]_{\Upsilon}, \quad (4.29)$$

$$C_2|_{\Upsilon} = C_3|_{\Upsilon}. \quad (4.30)$$

Finalmente, a simulação numérica do fenômeno descrito nesta seção foi realizada com base nos parâmetros listados na Tabela 3 e dividida em duas fases: uma para a infusão do nanofluido e a outra para a difusão das nanopartículas. Nesse caso, estimou-se que a administração de uma dose de $0,4 \text{ cm}^3$ de nanofluido no centro do tumor, a uma vazão de $5 \mu\text{L min}^{-1}$, necessita de 4800 s (1 h e 20 min) para ser concluída. Por outro lado, considerou-se que as nanopartículas se difundem por um período de 24 h antes do início da fase de aquecimento, conforme proposto em Mahesh, Singh e Talukdar (2022).

Tabela 3 – Valores utilizados para as propriedades dos domínios durante a simulação do modelo de transporte do nanofluido.

Propriedade	Valor
Raio dos poros dos vasos capilares, r_p (m)	
Tumor	850×10^{-9}
Órgão saudável	10×10^{-9}
Porosidade da parede dos vasos capilares, ϵ_c (adimensional)	
Tumor	0,001
Órgão saudável	0,0005
Espessura da parede dos vasos capilares L_c (m)	
Tumor	5×10^{-6}
Órgão saudável	5×10^{-6}

Fonte: adaptado de Mahesh, Singh e Talukdar (2022).

4.2.4 Biotransferência de Calor

O fenômeno da transferência de calor em tecidos biológicos pode ser modelado usando a equação de Pennes (1948). Ao combinar tal formulação com a teoria introduzida na Seção 3.1.3, é possível estimar o aumento da temperatura local por meio da seguinte equação (Tang *et al.*, 2021):

$$\rho_i c_i \frac{\partial T_i}{\partial t} = k_i \nabla^2 T_i + \rho_b c_b \Omega_b (T_b - T_i) + Q_{\text{met}} + \eta P, \quad (4.31)$$

em que ρ_b , c_b , Ω_b e T_b são a densidade, calor específico, taxa de perfusão e a temperatura do sangue, respectivamente; T_i é a temperatura local; $\eta = 0,55$ é o fator de correção para a potência dissipada; Q_{met} é o calor gerado em função do metabolismo em i ; e P é a potência dissipada pelas nanopartículas magnéticas, definida anteriormente pela Equação (3.16), mas que pode ser reescrita como:

$$P = \mu_0 \pi \chi_0 \frac{\omega \tau}{1 + (\omega \tau)^2} f H_0^2, \quad (4.32)$$

em que τ é o tempo de relaxação efetivo das nanopartículas, calculado a partir da Equação (3.4); χ_0 é a suscetibilidade no equilíbrio do nanofluido magnético, dado pela Equação (3.12); f é a frequência do campo magnético; H_0 é a amplitude do campo magnético; $\omega = 2\pi f$ é a frequência angular; e $\mu_0 = 4\pi \times 10^{-7} \text{ NA}^{-2}$ é a constante da permeabilidade no vácuo.

Para resolver a equação de biotransferência de calor, foram necessárias algumas hipóteses adicionais. Nesse caso, assumiu-se que a temperatura do sangue é constante e igual a 37°C . Além disso, considerou-se que a taxa de perfusão sanguínea é igual a zero no núcleo necrótico

($i = 1$), devido à ausência de sistema vascular. Por outro lado, definiu-se o comportamento de Ω_b para a região viável do tumor ($i = 2$) e do órgão saudável ($i = 3$) por meio das seguintes equações (Tang *et al.*, 2021; Mahesh; Singh; Talukdar, 2022):

$$\Omega_b^{(2)} = \begin{cases} 8,33 \times 10^{-4}, & T < 37 \text{ }^\circ\text{C}, \\ 8,33 \times 10^{-4} - \frac{(T - 37)^{4,8}}{5,438 \times 10^6}, & 37 \text{ }^\circ\text{C} \leq T < 42 \text{ }^\circ\text{C}, \\ 4,16 \times 10^{-4}, & T \geq 42 \text{ }^\circ\text{C}, \end{cases} \quad (4.33)$$

$$\Omega_b^{(3)} = \begin{cases} 4,5 \times 10^{-4} + 3,55 \times 10^{-3} \exp\left[-\frac{(T - 45)^2}{12}\right], & T < 45 \text{ }^\circ\text{C}, \\ 0,004, & T \geq 45 \text{ }^\circ\text{C}. \end{cases} \quad (4.34)$$

A distribuição espacial das nanopartículas magnéticas na região do tumor pode afetar as propriedades do meio. Desse modo, para estimar a densidade, a capacidade térmica e a condutividade térmica local, pode-se recorrer às seguintes equações (Tang *et al.*, 2021):

$$\rho_i = \phi_i \rho_{\text{MNP}} + (1 - \phi_i) \rho_t, \quad (4.35)$$

$$c_i = \phi_i c_{\text{MNP}} + (1 - \phi_i) c_t, \quad (4.36)$$

$$1/k_i = \phi_i/k_{\text{MNP}} + (1 - \phi_i)/k_t, \quad (4.37)$$

em que $\phi_i = C_i/\rho_{\text{MNP}}$ é a fração de nanopartículas no meio; c_{MNP} é o calor específico da nanopartícula; k_{MNP} é a condutividade térmica da nanopartícula; e o subscrito t denota as propriedades do tumor antes da infusão das MNPs.

Quanto às condições iniciais, definiu-se que a temperatura em todas as regiões é igual a $37 \text{ }^\circ\text{C}$ em $t = 0$. Além disso, foram consideradas duas condições de contorno. A primeira garante a continuidade do fluxo de calor e da temperatura em todas as interfaces entre os domínios. Desse modo, na fronteira entre o núcleo necrótico e a zona viável do tumor, Ψ , têm-se que (Tang *et al.*, 2021):

$$-k_1 \nabla T_1|_{\Psi} = -k_2 \nabla T_2|_{\Psi}, \quad (4.38)$$

$$T_1|_{\Psi} = T_2|_{\Psi}, \quad (4.39)$$

enquanto a condição da continuidade na fronteira entre a zona viável do tumor e o órgão saudável, Υ , é garantida por:

$$-k_2 \nabla T_2|_{\Upsilon} = -k_3 \nabla T_3|_{\Upsilon}, \quad (4.40)$$

$$T_2|_{\Upsilon} = T_3|_{\Upsilon}. \quad (4.41)$$

Para a segunda e última condição de contorno, definiu-se o isolamento térmico nas bordas do órgão saudável, Ξ (Lin; Liu, 2009).

$$-\mathbf{n} \cdot (k_3 \nabla T_3)|_{\Xi} = 0. \quad (4.42)$$

Finalmente, os parâmetros listados na Tabela 4 foram utilizados para solucionar as equações governantes apresentadas nesta seção. Quanto ao tempo de simulação, este foi ajustado em função dos ensaios de identificação de sistema.

Tabela 4 – Valores utilizados para as propriedades dos materiais durante a simulação do modelo de biotransferência de calor.

Propriedade	Valor
Massa específica, ρ (kg m^{-3})	
Tumor	1030
Órgão saudável	1060
Sangue	1050
Nanopartícula	5180
Condutividade térmica, k ($\text{W m}^{-1} \text{K}^{-1}$)	
Tumor	0,52
Órgão saudável	0,51
Nanopartícula	40
Calor específico, c ($\text{J kg}^{-1} \text{K}^{-1}$)	
Tumor	3540
Órgão saudável	3871
Sangue	4180
Nanopartícula	670
Calor gerado em função do metabolismo, Q_{met} (W m^{-3})	
Tumor	5790
Órgão saudável	684
Constante de anisotropia da nanopartícula, K_a (kJ m^{-3})	
	20,3
Magnetização do domínio da nanopartícula, M_d (kA m^{-1})	
	446
Tempo característico da nanopartícula, τ_0 (s)	
	1×10^{-9}

Fonte: adaptado de Rosensweig (2002) e Tang *et al.* (2021).

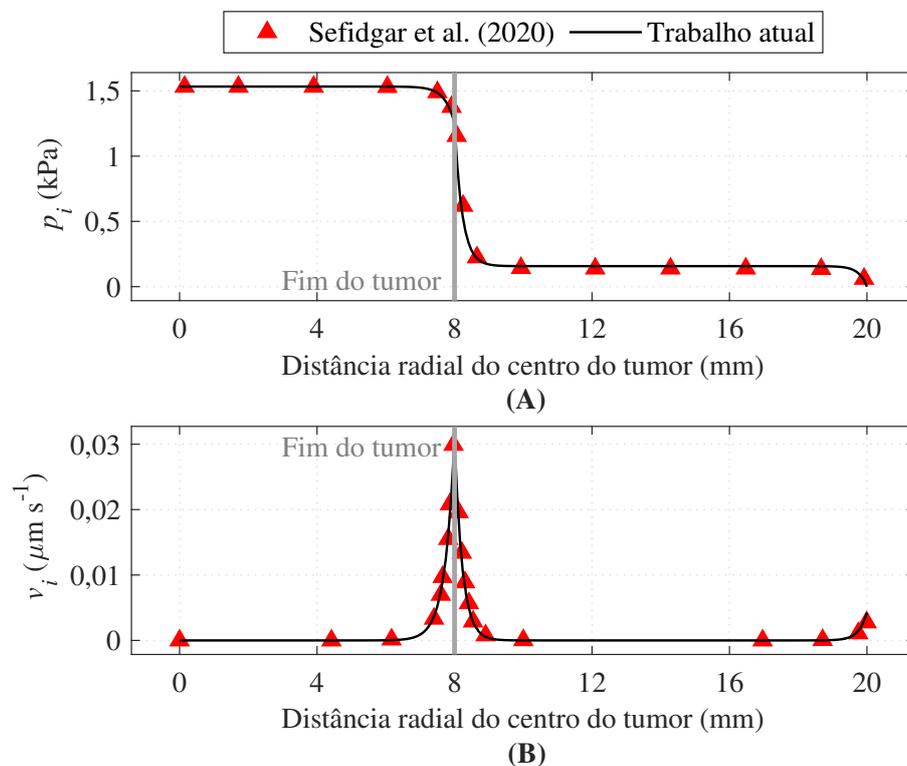
4.3 VERIFICAÇÃO DOS MODELOS NUMÉRICOS

Antes de avançar com as simulações para obter os dados necessários à identificação dos sistemas, foram realizadas verificações para garantir a correta implementação dos modelos numéricos. Para isso, selecionaram-se os estudos conduzidos por Sefidgar, Bashooki e Shojaee

(2020), Steuperaert *et al.* (2017) e Lin e Liu (2009) para fins de comparação. Em todos os casos, o modelo geométrico foi implementado de acordo com o descrito na referência empregada, como forma de poder comparar os dados. Os modelos geométricos descritos na Seção 4.1 foram empregados em etapa posterior do trabalho, uma vez verificada a funcionalidade das equações implementadas.

O estudo de Sefidgar, Bashooki e Shojaee (2020) foi utilizado para verificar os aspectos relacionados à velocidade e à pressão intersticial do nanofluido nos tecidos. Nesse estudo, assume-se um tumor esférico com um raio de 8 mm, com um núcleo necrótico de 4 mm de raio e com um órgão saudável de 40 mm de comprimento ao seu redor. Ao considerar tais hipóteses, a Figura 11(A), demonstra que a pressão na região do tumor pode chegar até 1,5 kPa, diminuindo à medida que a distância radial cresce. Por sua vez, a Figura 11(B) revela que a velocidade intersticial tende a zero no centro do tumor e aumenta na região de fronteira entre o tumor e o órgão saudável, ultrapassando o valor de $0,03 \mu\text{m s}^{-1}$. Desse modo, é verificada uma coerência entre os valores obtidos com o modelo proposto neste trabalho e os valores reportados na literatura, com um desvio médio relativo de 3,8% para os resultados referentes à pressão intersticial e um desvio médio relativo de 8,4% para a velocidade intersticial.

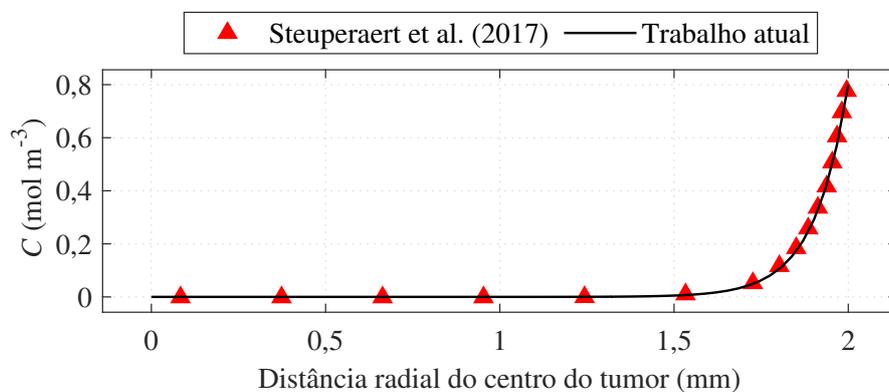
Figura 11 – Verificação do modelo de transporte intersticial do nanofluido com os resultados da literatura: (A) pressão intersticial; e (B) velocidade intersticial.



Fonte: autor.

Assim como no estudo conduzido por Mahesh, Singh e Talukdar (2022), esta dissertação também adotou a pesquisa de Steuperaert *et al.* (2017) para verificar a implementação do modelo de convecção-difusão. Nesse contexto, um dos modelos de referência considerados pelos autores mencionados envolve um tumor isolado com um raio de 2 mm, contendo um núcleo necrótico de 1 mm de raio. Além disso, tal modelo ainda pressupõe uma concentração de $0,8 \text{ mol m}^{-3}$ de um determinado fármaco nas bordas do tumor, a qual se difunde por toda a região em um intervalo de até 1 h. Com base nessas hipóteses e sem perda de generalidade, a Figura 12 evidencia que o modelo implementado pode retornar resultados consistentes com a literatura, com um desvio médio relativo de 3,0%, qualificando o seu uso para o desenvolvimento deste estudo.

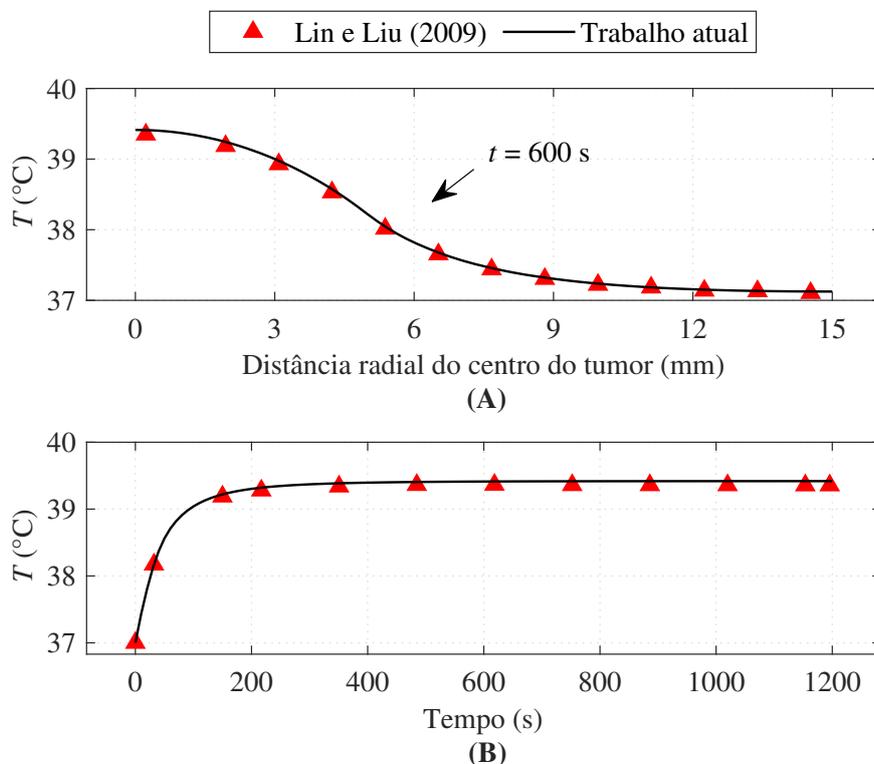
Figura 12 – Verificação do modelo de convecção-difusão com os resultados da literatura



Fonte: autor.

Finalmente, o estudo de Lin e Liu (2009) foi adotado para verificar a implementação da física de biotransferência de calor. Nesse contexto, o modelo proposto pelos pesquisadores considera um cenário de tratamento de hipertermia magnética, no qual nanopartículas de magnetita dissipam uma potência de $1,95 \times 10^5 \text{ W m}^{-3}$ em um tumor esférico com 5 mm de raio, envolto por um tecido saudável com 15 mm de raio. Com base nessas considerações e realizando as adaptações necessárias na geometria e nos parâmetros dos modelos propostos nesta dissertação, a Figura 13(A) apresenta a distribuição de temperatura em $t = 600 \text{ s}$ ao longo de todos os tecidos, permitindo observar o seu decaimento à medida que a distância radial aumenta. Por sua vez, a Figura 13(B) destaca o aumento da temperatura no centro do tumor em um intervalo de 1200 s. Além disso, ao verificar a diferença entre os resultados desta simulação e os reportados no trabalho de referência, foi observado um desvio médio relativo de 0,07% e 0,11% para os perfis de distribuição de temperatura radial e transiente, respectivamente. Com base nisso, é possível inferir que a implementação da física de biotransferência de calor está corretamente implementada.

Figura 13 – Verificação do modelo de biotransferência de calor com os resultados da literatura: (A) perfil radial de distribuição temperatura; e (B) perfil transiente de temperatura.



Fonte: autor.

4.4 COMENTÁRIOS FINAIS

A partir do que foi abordado na Seção 4.1, foi possível construir os três modelos numéricos propostos nesta dissertação. Além disso, na Seção 4.2 foram apresentadas as equações governantes relacionadas às etapas básicas do tratamento de hipertermia magnética, o que também permitiu observar a complexidade dos modelos matemáticos. Finalmente, o desenvolvimento da Seção 4.3 permitiu avaliar se a física de cada etapa foi adequadamente incorporada nos modelos numéricos, sendo uma etapa importante antes da realização dos ensaios de identificação de sistema descritos no Capítulo 5.

5 ENSAIOS DE IDENTIFICAÇÃO DO SISTEMA

Neste capítulo, é descrito o método de identificação empregado para representar a relação dinâmica entre as características do campo magnético e as temperaturas em quatro pontos do tumor, para cada um dos modelos geométricos considerados neste trabalho. Tal identificação foi feita empregando as ESNs, utilizando os dados obtidos por meio de simulações numéricas envolvendo os modelos apresentados no Capítulo 4. Desse modo, o capítulo está organizado da seguinte forma. Primeiramente, na Seção 5.1 é detalhado o procedimento para aquisição e processamento dos dados empregados para a identificação dos sistemas. Em seguida, na Seção 5.2 é abordada a estratégia de otimização dos hiperparâmetros da ESN. Na Seção 5.3, é feita uma análise a respeito do custo computacional para a execução das ESNs como modelo de predição. Por fim, na Seção 5.4 são apresentados os comentários finais a respeito de todo processo de identificação e otimização com as ESNs.

5.1 AQUISIÇÃO E PROCESSAMENTO DA SÉRIE DE DADOS

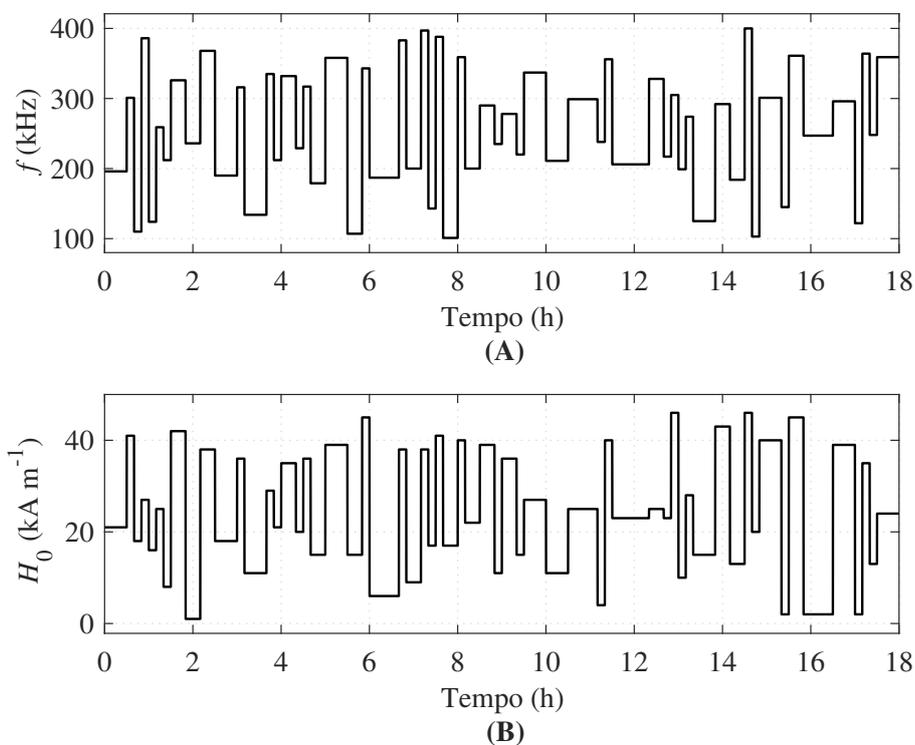
Para controlar o aumento da temperatura local durante um tratamento de hipertermia magnética, deve-se ajustar a potência dissipada pelas nanopartículas magnéticas. Tal controle pode ser obtido alterando-se a frequência ou a amplitude do campo magnético externo, ou uma combinação de ambos os parâmetros, conforme evidenciado pela formulação de Rosensweig (2002). Com base nisso, esta dissertação propõe o uso de tais variáveis como as entradas de um sistema a ser identificado.

Para excitar os modelos numéricos durante os ensaios de identificação, a escolha foi pelo uso de sinais do tipo APRBS, uma vez que os sistemas de hipertermia magnética tendem a apresentar uma dinâmica não linear. Cada sinal foi projetado para apresentar um tempo de amostragem de 10 s, o qual corresponde ao passo no tempo empregado para solucionar o modelo de biotransferência de calor. Além disso, os limites superiores e inferiores de cada sinal de entrada foram estabelecidos com base no estudo de Fernández *et al.* (2022), no qual foi utilizado um sistema capaz de gerar um campo magnético com uma frequência entre 100 kHz e 400 kHz, e uma amplitude entre 0 kA m^{-1} e 47 kA m^{-1} para avaliar o potencial de aplicação de MNPs de ferrita de manganês no contexto do tratamento de hipertermia magnética. Para aplicações clínicas, a ideia é que os limites dos sinais sejam ajustados em função das restrições operacionais e das especificações do equipamento disponível.

Para determinar a duração mínima de cada patamar nos APRBSs, foram realizados testes preliminares nos quais diferentes sinais do tipo degrau foram aplicados aos sistemas. A partir disso, foi verificado que os processos possuem uma dinâmica lenta, atingindo o regime per-

manente após 60 min, ao utilizar campos magnéticos de baixa amplitude e frequência, e entre 30 min e 60 min com AMFs mais intensos. Logo, uma vez que o tratamento de hipertermia magnética não costuma exceder 60 min, optou-se por seguir a heurística descrita na Equação (3.40) para sistemas não lineares, o que levou a definir que cada patamar nos APRBSs devem ser mantidos por, no mínimo, 10 min. Com base nessas escolhas, foi gerado um conjunto de sinais com 6481 pontos de dados para ambas as entradas do sistema, resultando em um tempo de simulação de 18 h, conforme ilustrado na Figura 14.

Figura 14 – Sinais APRBS: (A) frequência de operação; e (B) amplitude do campo magnético.



Fonte: autor.

Para monitorar o aumento de temperatura na região do tumor em resposta aos sinais de entrada e em função da simetria dos modelos numéricos, foram considerados quatro pontos de medição (I, II, III e IV). Cada ponto foi posicionado nas coordenadas listadas na Tabela 5. Além disso, um ruído branco com um desvio padrão de $0,2^\circ\text{C}$ foi adicionado para representar possíveis ruídos de medição em cada ponto de sensoriamento.

As respostas do processo aos sinais de excitação estão apresentadas na Figura 15. A fim de utilizar tais informações no treinamento do modelo de aprendizado de máquina, todos os sinais envolvidos no ensaio de identificação foram normalizados para um intervalo entre 0 e 1. A normalização dos sinais de entrada foi realizada considerando seus respectivos limites superiores

Tabela 5 – Coordenadas, em milímetros, ao longo do eixo radial e vertical de cada ponto de medição.

Modelo	I	II	III	IV
Esférico	(0, 0)	(3, 0)	(6, 0)	(9, 0)
Oblato	(0, 0)	(3, 0)	(6, 0)	(9, 0)
Prolato	(0, 0)	(1,5, 0)	(3, 0)	(4,5, 0)

Fonte: autor.

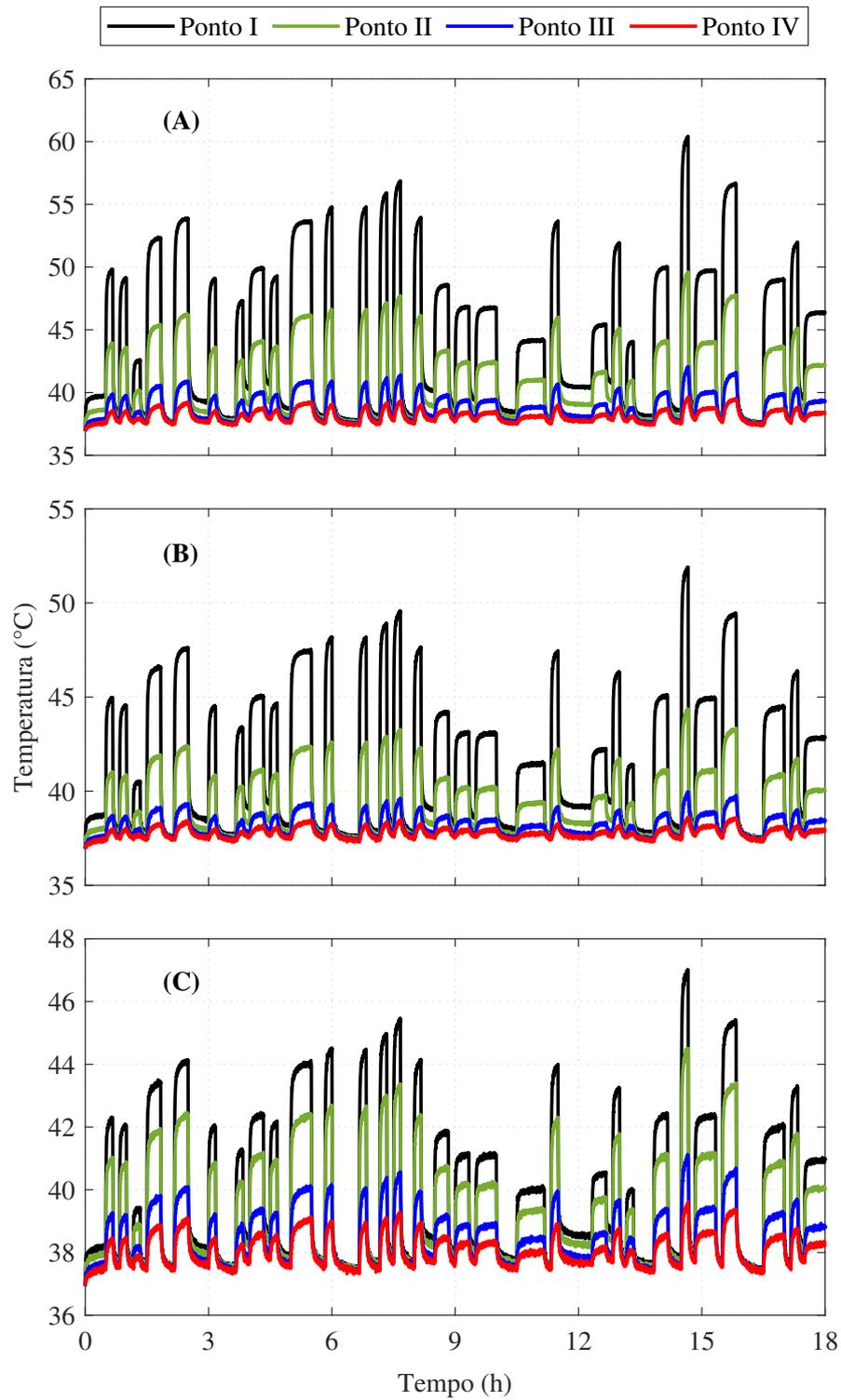
e inferiores. Para normalizar as saídas, adotou-se um limite inferior de 37°C , correspondente à temperatura normal do corpo humano, e um limite superior de 46°C , temperatura na qual a apoptose do tumor pode ocorrer sem prejudicar o tecido saudável ao redor (Tang *et al.*, 2021). Além disso, vale destacar dois aspectos. Primeiramente, caso a temperatura ultrapasse o limiar de 46°C , o processo de identificação com as ESNs não é afetado, uma vez que as saídas da rede não podem saturar os neurônios do reservatório. Em segundo lugar, na prática, haverá cuidado para que a temperatura local não exceda os limites de segurança, então o fato de otimizar o comportamento da rede até 46°C tende a levá-la a modelar da melhor forma possível os sinais que serão empregados na prática.

5.2 TREINAMENTO E OTIMIZAÇÃO DA REDE NEURAL

Após a etapa de pré-processamento, os dados foram divididos em conjuntos distintos destinados ao treinamento, validação e teste da rede neural. A porção de treinamento representa 60% do conjunto de dados e abrange 3889 pontos, dos quais os primeiros 50 pontos foram selecionados para remover os efeitos do transitório inicial nos estados do reservatório da ESN. Por sua vez, a porção de validação corresponde a 20% do conjunto de dados e apresenta 1296 pontos, sendo empregados para otimizar os hiperparâmetros do modelo. Por fim, os 20% restantes dos dados foram reservados para avaliar o desempenho da rede.

Para otimizar o desempenho da rede neural, diversos hiperparâmetros passaram por ajustes. Em relação ao fator de escala da matriz \mathbf{W}^{in} , um intervalo de busca foi definido, variando de 0,1 a 1,5, a fim de investigar se as unidades do reservatório devem operar próximas ao ponto no qual a função de ativação se torna praticamente linear ou se devem operar próximas dos limites de saturação. Devido ao custo computacional, o intervalo de busca para o tamanho do reservatório foi limitado entre 100 e 1000 unidades. Para garantir a propriedade de estado de eco, o maior valor testado para o raio espectral foi estabelecido como sendo igual a 1. Por sua vez, a taxa de vazamento do reservatório foi ajustada no intervalo entre 0 e 1. Por último, o coeficiente

Figura 15 – Respostas do processo aos sinais de entrada: (A) modelo esférico; (B) modelo oblato; e (C) modelo prolato.



Fonte: autor.

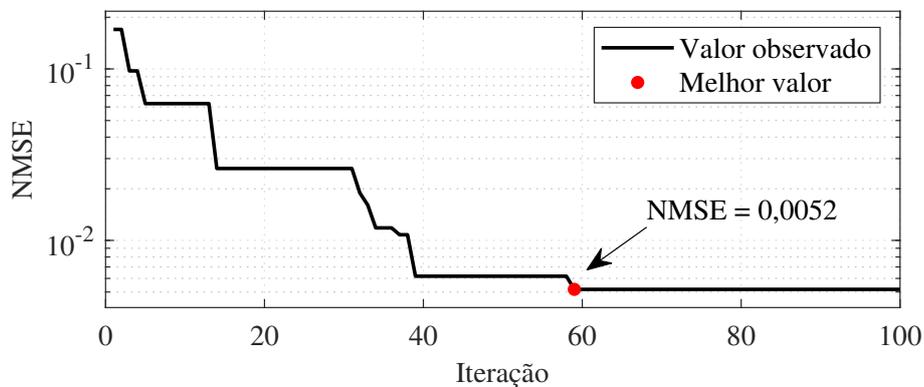
de regularização foi ajustado no intervalo entre 1×10^{-4} e 10 (Schwedersky; Flesch; Dangui, 2022).

Em função da quantidade de hiperparâmetros a serem otimizados, optou-se por ajustar o desempenho das ESNs através de um algoritmo de otimização bayesiana, implementado em MATLAB. Para isso, foram adotados os seguintes procedimentos: na primeira iteração, o algoritmo escolhe aleatoriamente uma dada combinação de hiperparâmetros para iniciar o treinamento da rede. Em seguida, o desempenho da ESN no conjunto de validação é avaliado a partir da média aritmética do erro médio quadrático normalizado (NMSE, do inglês *Normalized Mean Squared Error*) entre todas as saídas consideradas. Com base nos valores encontrados, o algoritmo de otimização modela o NMSE médio como uma função caixa preta, cujo modelo probabilístico pode ser construído a partir de um modelo de processo gaussiano. Por sua vez, para estimar os hiperparâmetros com maior potencial de minimizar o NMSE, a otimização bayesiana recorre à função de aquisição de melhoria esperada. Desse modo, ao longo deste processo iterativo, o algoritmo verifica quais hiperparâmetros podem ter maior impacto sobre o desempenho da ESN e os otimiza em até 100 iterações, sendo este um valor definido para encontrar o menor NMSE possível. Todo esse processo foi desenvolvido com base no trabalho de Yperman e Becker (2016), o qual também utilizou a otimização bayesiana para o mesmo propósito.

Na Figura 16, são apresentados os resultados relacionados ao ajuste dos hiperparâmetros da ESN através do algoritmo de otimização bayesiana para o sistema MIMO representado pelo modelo esférico. Desse modo, é verificado que a abordagem considerada permitiu tal ajuste em menos de 100 iterações. Ao fim, foi constatado que o menor valor para o NMSE foi 0,0052 para o conjunto de validação, o qual foi obtido a partir da seguinte combinação de hiperparâmetros: um fator de escalonamento de 0,3013 para a matriz \mathbf{W}^{in} ; um raio espectral de 0,9839; uma taxa de vazamento de 0,1359 para o reservatório; um reservatório com 966 unidades; e um coeficiente de regularização de 0,2805.

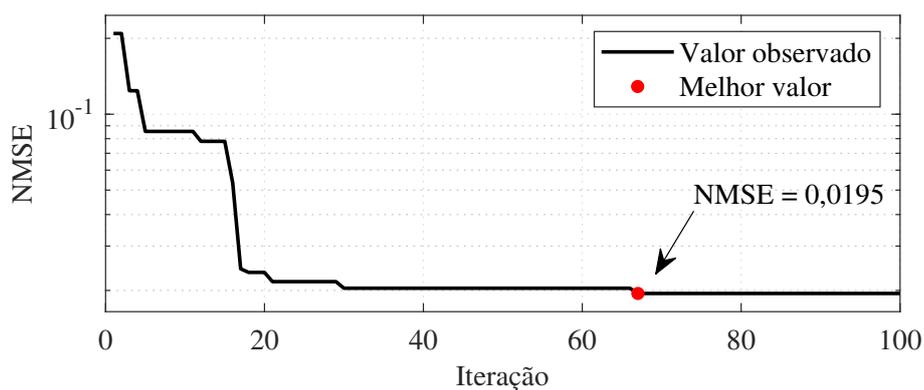
A partir dos valores encontrados para os hiperparâmetros no cenário representado pelo modelo esférico, é possível realizar uma série de inferências. A primeira, é sobre a linearidade do processo. Nesse caso, os valores para o fator de escalonamento da matriz \mathbf{W}^{in} e o raio espectral ratificam a hipótese de que o sistema proposto apresenta uma dinâmica não linear, uma vez que há a necessidade de que as unidades do reservatório não atuem tão próximas do ponto no qual a função de ativação é aproximadamente linear. Além disso, o valor do raio espectral, em conjunto com a taxa de vazamento do reservatório, permitem constatar que a ESN precisa de uma alta retenção de memória para modelar as saídas do sistema. Um das possíveis explicações para esse fato é que o sistema proposto apresenta dinâmica lenta, muito influenciada pelos efeitos da taxa de perfusão sanguínea nos tecidos (Liangruksa; Ganguly; Puri, 2011; Lukoševičius, 2012).

Figura 16 – Minimização do NMSE a partir da otimização bayesiana no caso do modelo de geometria esférica.



No contexto do modelo elipsoidal oblato, a Figura 17 ilustra o processo de ajuste dos hiperparâmetros da rede neural de estado de eco por meio da otimização bayesiana. Mais uma vez, em até 100 iterações, foi possível minimizar o NMSE entre as saídas da rede e as saídas do modelo numérico. O menor valor registrado para o NMSE foi de 0,0195, obtido a partir da seguinte combinação de hiperparâmetros: um fator de escalonamento de 1,0682 para a matriz \mathbf{W}^{in} ; um raio espectral de 0,9882; uma taxa de vazamento de 0,0729 para o reservatório; um reservatório com 273 unidades; e um coeficiente de regularização de 0,2805.

Figura 17 – Minimização do NMSE a partir da otimização bayesiana no caso do modelo de geometria elipsoidal oblata.

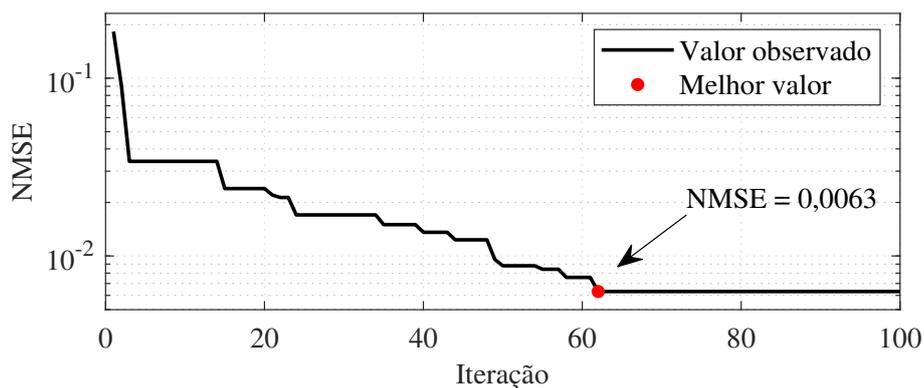


Os hiperparâmetros encontrados para o caso do modelo oblato sugerem que a geometria do modelo numérico introduz uma maior não linearidade ao processo. Nesse contexto, os valores do fator de escalonamento da matriz \mathbf{W}^{in} , em conjunto com o raio espectral, indicam que as unidades do reservatório devem operar o mais próximo possível da região de saturação da função

de ativação para modelar as saídas do sistema com o menor erro possível. Além disso, assim como no modelo esférico, a ESN continua a demandar um reservatório com uma alta retenção de memória para desempenhar a tarefa desejada.

Finalmente, na Figura 18 são apresentados os resultados referentes à otimização dos hiperparâmetros da ESN, considerando o cenário de tratamento representado pelo modelo prolato. Ao longo desse procedimento, o menor valor observado para o NMSE foi de 0,0063, o qual foi obtido a partir do uso dos seguintes hiperparâmetros: um fator de escalonamento de 0,3013 para a matriz \mathbf{W}^{in} ; um raio espectral de 0,9858; uma taxa de vazamento de 0,1359 para o reservatório; um reservatório com 966 unidades; e um coeficiente de regularização de 0,2806. Tais valores são similares aos reportados no contexto do modelo esférico e, portanto, permitem fazer as mesmas inferências sobre a não linearidade do processo e o comportamento do reservatório da ESN.

Figura 18 – Minimização do NMSE a partir da otimização bayesiana no caso do modelo de geometria elipsoidal prolata.



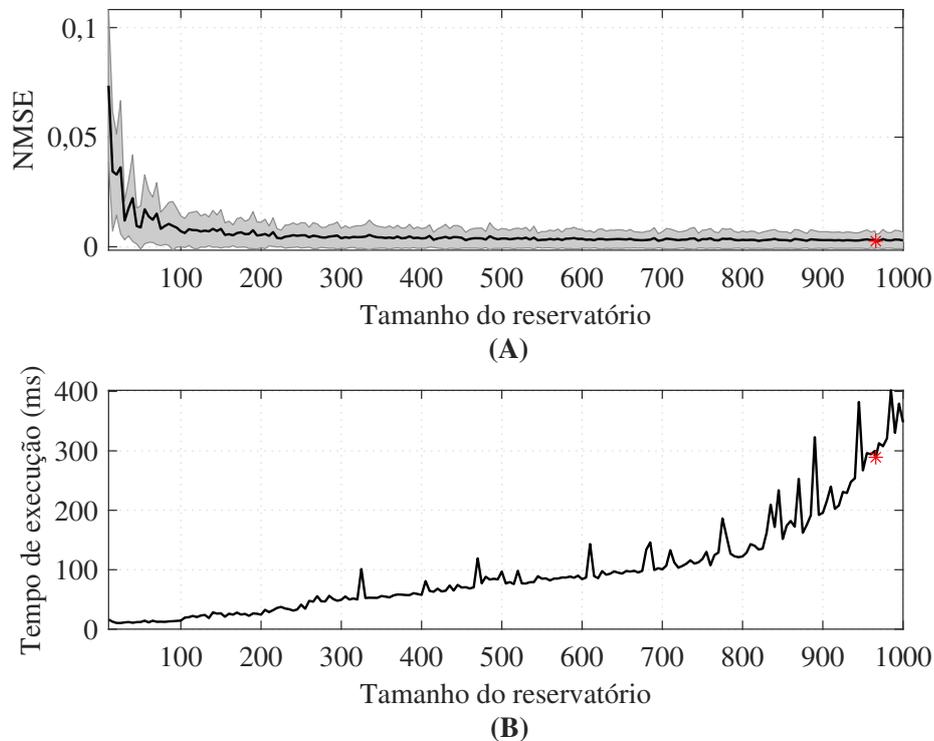
Fonte: autor.

5.3 ANÁLISE DO CUSTO COMPUTACIONAL

No contexto do modelo numérico de geometria esférica, um aspecto que se destacou após otimizar a ESN foi o tamanho do reservatório, N_x . Nesse caso, o valor desse hiperparâmetro sugere que a rede neural necessita de um reservatório com uma quantidade considerável de neurônios para alcançar o menor erro de predição para o conjunto de validação. Essa configuração, apesar de apresentar o menor erro, faz com que o modelo necessite de 289 ms para computar a predição para as mais de 1000 amostras nesse conjunto.

Haja vista que, em aplicações práticas, é importante haver um equilíbrio entre o desempenho do modelo e o seu tempo de execução, realizou-se um experimento semelhante ao descrito em Schwedersky, Flesch e Danguì (2022). Nesse caso, o objetivo foi avaliar como o tamanho do reservatório impacta tanto o NMSE quanto o tempo de execução da ESN. Para isso,

Figura 19 – Influência do tamanho do reservatório sobre a ESN: (A) NMSE médio; e (B) tempo de execução



o modelo foi novamente treinado, utilizando reservatórios contendo entre 10 e 1000 unidades de neurônios, enquanto os demais hiperparâmetros foram fixados em seus valores otimizados. Note que não necessariamente o ajuste dos demais parâmetros será ótimo para reservatórios menores, pois seus valores foram definidos em conjunto no processo de otimização bayesiana. Entretanto, essa análise permite encontrar um limitante superior para o desempenho do modelo com reservatórios com menos neurônios. Ao final, a análise do impacto de N_x sobre os aspectos mencionados foi conduzida no conjunto de validação, conforme ilustrado na Figura 19. As linhas contínuas representam os valores observados, enquanto as áreas sombreadas em cinza demonstram os limites superiores e inferiores desses valores, estabelecidos por dois desvios-padrões em torno da média calculada. Além disso, o asterisco destaca o valor escolhido para tamanho do reservatório no modelo final.

Com base nas informações apresentadas, a Figura 19(A) revela que o número de neurônios no reservatório não exerce uma influência significativa no desempenho do modelo, desde que o número seja maior que, aproximadamente, 100 neurônios. No entanto, também é possível observar que um reservatório contendo menos de 100 neurônios pode prejudicar o desempenho. Por outro lado, a Figura 19(B) ilustra que o tempo de execução dos modelos cresce de forma diretamente proporcional ao tamanho do reservatório, apresentando algumas oscilações em fun-

ção da execução de processos, em paralelo, na máquina de teste. Nesse caso, é observado que um reservatório contendo até, aproximadamente, 700 neurônios pode computar as previsões em menos de 100 ms, enquanto um reservatório com menos de 300 neurônios pode fazer isso em até 50 ms e sem elevar de forma significativa o NMSE das saídas previstas.

Para os demais cenários, representados pelos modelos elipsoidais, foi observada uma situação similar ao que foi retratado na Figura 19. Isso leva a inferir que a otimização bayesiana que se preocupe apenas em minimizar o erro de previsão pode resultar em um modelo complexo e que tenha apenas um ganho pequeno em termos de erro de previsão em relação a um modelo mais simples. Assim, para situações nas quais o desempenho computacional seja relevante, é importante considerar a complexidade do modelo diretamente na função a ser minimizada ou realizar um ajuste final do modelo, a fim de garantir um melhor equilíbrio entre o desempenho do modelo de previsão e o seu tempo de execução. Esse ajuste pode ser especialmente útil no contexto de MPC, visto que uma previsão de múltiplos instantes à frente é necessária a cada período de amostragem.

5.4 COMENTÁRIOS FINAIS

A partir do que foi discutido na Seção 5.1, foi possível adquirir os dados necessários para a identificação dos sistemas MIMO que caracterizam as relações dinâmicas entre as propriedades do campo magnético e as temperaturas em alguns pontos do tumor. Para aplicações práticas, a metodologia precisaria de ajustes adicionais, considerando as restrições operacionais e os limites do equipamento encarregado de gerar o campo magnético externo. Na Seção 5.2, foi descrito o método empregado para otimizar o desempenho da ESN com base nos conjuntos de treinamento e validação. Por sua vez, foi verificado que o método descrito em Yperman e Becker (2016) permitiu encontrar as combinações de hiperparâmetros mais adequadas para cada caso, em até 100 iterações. Além disso, vale destacar que essa estratégia se mostrou vantajosa, uma vez que abordou a alta dimensionalidade do problema de otimização, contornando uma limitação que surgiria se a busca em grade tivesse sido adotada. Ao fim, a Seção 5.3 revelou que, mesmo após a otimização, a estrutura da ESN pode ser ajustada para garantir um melhor equilíbrio entre o seu desempenho e o custo computacional para a sua execução.

6 RESULTADOS E DISCUSSÃO

Neste capítulo, são apresentados os resultados obtidos para a identificação dos sistemas por meio das redes neurais de estado de eco. Para isso, adotou-se a seguinte organização. Primeiramente, na Seção 6.1 são explorados os resultados obtidos para o modelo numérico com geometria esférica. Em seguida, na Seção 6.2 são detalhados os resultados para a identificação do modelo elipsoidal oblato. Na Seção 6.3, são abordados os resultados para o modelo elipsoidal prolato. Por fim, na Seção 6.4 são feitas as discussões finais sobre os resultados apresentados.

6.1 MODELO ESFÉRICO

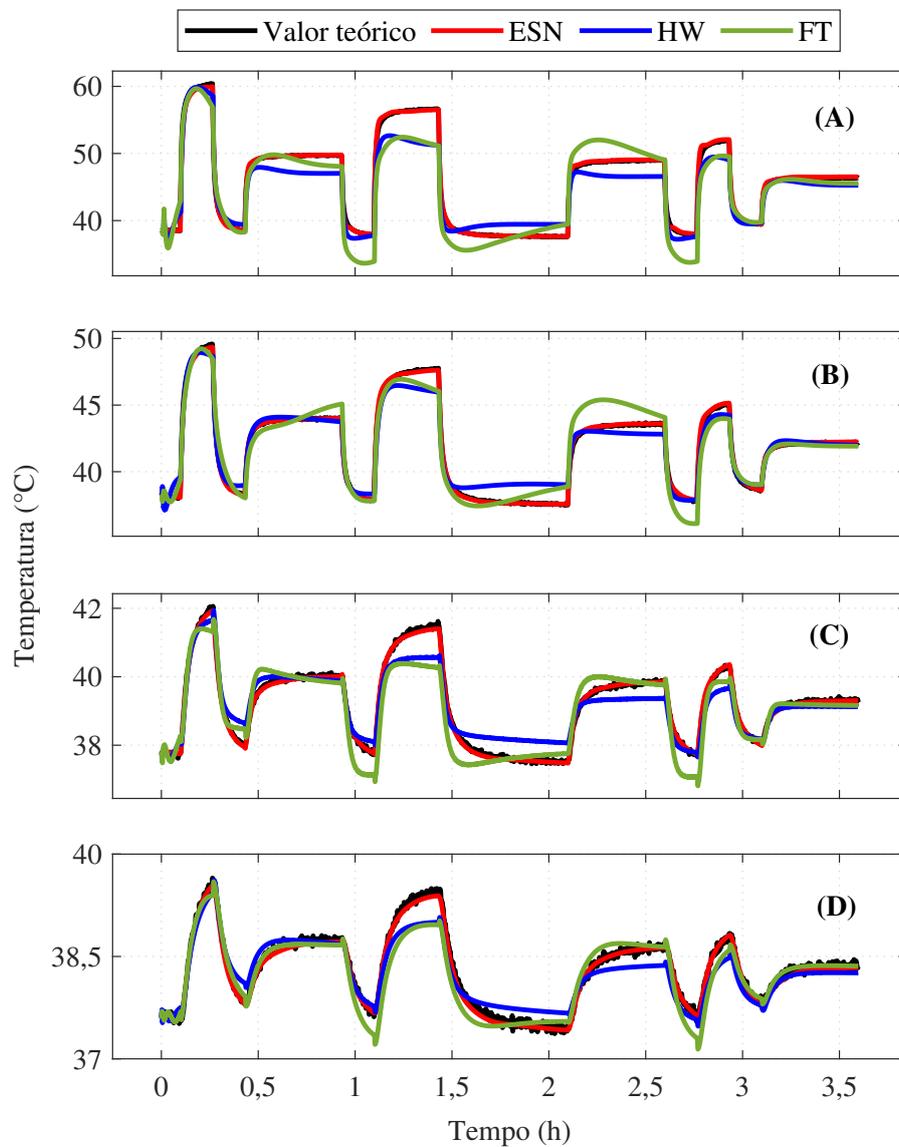
Após a definição da arquitetura mais apropriada para a ESN, empregou-se o modelo resultante para estimar a temperatura na região do tumor com base no conjunto de dados de teste. Além disso, para fins de comparações, optou-se por incorporar tanto um modelo linear quanto um modelo não linear para a mesma tarefa.

A identificação de cada modelo de referência foi feita a partir do conjunto de treinamento, enquanto a ordem de cada um foi escolhida a partir de testes envolvendo o conjunto de validação. Desse modo, optou-se por um modelo linear baseado em uma Função de Transferência (FT) de quinta ordem. Além disso, o modelo não linear adotado é um modelo do tipo Hammerstein-Wiener (HW), o qual também utiliza uma função de transferência de quinta ordem como modelo dinâmico. Nesse contexto, o modelo incorpora um polinômio de grau 2 para as saídas, um polinômio de grau 4 para o sinal correspondente à frequência do campo magnético e um polinômio de grau 3 para o sinal que representa a amplitude do campo magnético.

Na Figura 20, é possível analisar as predições de cada modelo para o sistema em questão, considerando o conjunto de teste. A partir disso, é observado que a função de transferência apresenta dificuldades para estimar a temperatura em todos os pontos considerados. Por outro lado, o modelo de Hammerstein-Wiener demonstra um desempenho superior em relação ao modelo linear, reforçando a hipótese de que o processo possui uma dinâmica não linear. No entanto, o desempenho de ambos os modelos ainda fica abaixo do que é observado para a rede neural de estado de eco. Neste caso, constata-se que as predições da ESN estão mais próximas das saídas reais.

Os fatos expostos na Figura 20 ficam ainda mais evidentes a partir da análise dos dados apresentados na Tabela 6. Como pode ser observado, a ESN demonstra maior potencial para estimar a dinâmica do aumento da temperatura na região do tumor, uma vez que todos os NMSEs para cada saída são menores do que os observados para os modelos de referência. Em todos os casos, a ESN apresentou pequenos erros de predição para os pontos de medição mais próximos

Figura 20 – Predições para temperaturas na região do tumor no modelo esférico: (A) ponto I; (B) ponto II; (C) ponto III; e (D) ponto IV.



Fonte: autor.

da borda do tumor (pontos III e IV), os quais tendem a ser os mais críticos para o tratamento de hipertermia, uma vez que estão mais próximos dos tecidos saudáveis.

Por fim, foram registrados os tempos necessários para cada modelo estimar as saídas para as 1296 amostras de cada ponto de medição no conjunto de teste. Este experimento foi conduzido em uma CPU Intel(R) Core(TM) i5-9300H, com processador de 2,60 GHz, 4 núcleos e 8 *threads*. Nesse cenário, foi verificado que a função de transferência completou essa tarefa em 25 ms com um NMSE médio de 12,36%. Por outro lado, o modelo de Hammerstein-Wiener levou 71 ms e apresentou um NMSE médio de 9,41%, enquanto a ESN computou as predições

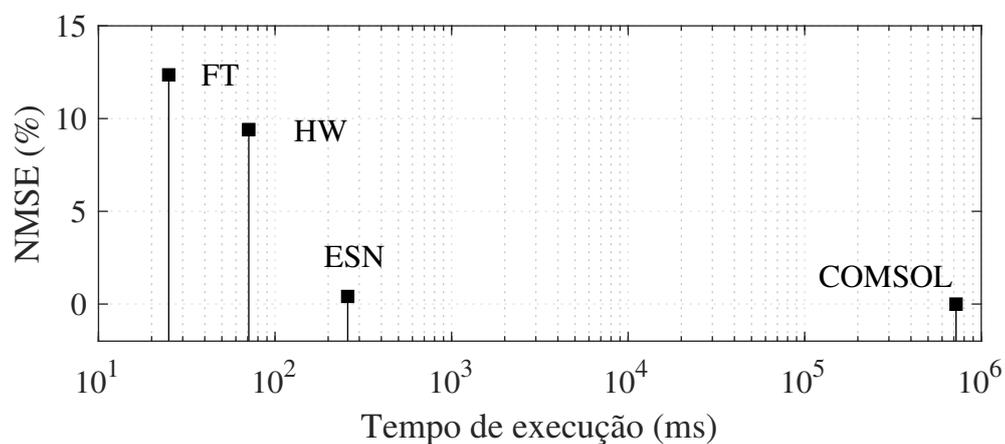
Tabela 6 – Valores do NMSE para a ESN proposta e os dois modelos de referência na tarefa de predição da temperatura nos pontos I, II, III e IV na região do tumor esférico.

Ponto	Rede de estado de eco	Modelo de Hammerstein-Wiener	Função de transferência
I	0,0021	0,0968	0,1314
II	0,0012	0,0483	0,0738
III	0,0025	0,1097	0,1533
IV	0,0103	0,1216	0,1359

Fonte: autor.

em 258 ms e apresentou um NMSE de 0,41%. Para fins de comparação, é relevante mencionar que foram necessários 12 min para obter todas amostras do conjunto de testes a partir das simulações numéricas, representando uma diferença de aproximadamente três ordens de magnitude em comparação ao tempo consumido pela ESN para a mesma tarefa, conforme ilustrado na Figura 21. Tais resultados destacam ainda mais o potencial das ESNs como modelos de predição de temperatura no contexto de tratamentos de hipertermia magnética.

Figura 21 – Comparação entre o NMSE e o tempo de execução para cada modelo no conjunto de teste do modelo esférico.



Fonte: autor.

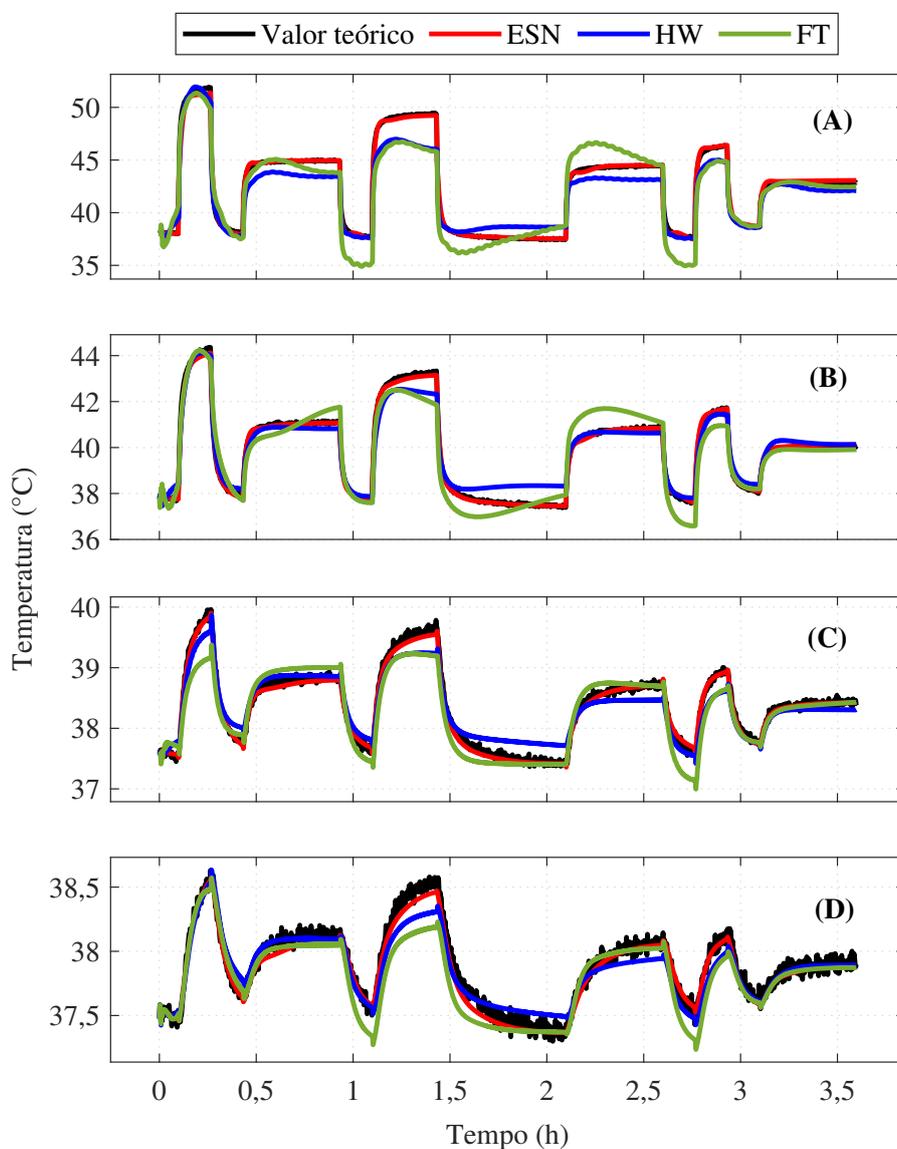
6.2 MODELO ELIPSOIDAL OBLATO

Analogamente ao que foi desenvolvido na Seção 6.1, dois modelos de referência também foram propostos para a tarefa de predição de temperatura na região do tumor oblato. Nesse caso, adotou-se um modelo linear baseado em uma função de transferência de sexta ordem, escolhida a partir de testes preliminares. Por sua vez, o modelo não linear consiste em um modelo de

Hammerstein-Wiener com a seguinte estrutura: uma função transferência de sexta ordem como modelo dinâmico, um polinômio de grau 2 para as saídas, um polinômio de grau 4 para a frequência do campo magnético e um polinômio de grau 3 para a amplitude do campo.

A Figura 22 apresenta as previsões de cada modelo proposto nesta seção, considerando o conjunto de teste. Ao observá-la, é possível realizar as mesmas inferências desenvolvidas para o modelo esférico. Os modelos de referência apresentam dificuldades para estimar corretamente os valores de temperatura em cada ponto de sensoriamento, enquanto a ESN demonstra um desempenho superior para essa tarefa.

Figura 22 – Previsões para temperaturas na região do tumor no modelo elipsoidal oblato: (A) ponto I; (B) ponto II; (C) ponto III; e (D) ponto IV.



Fonte: autor.

Além do desempenho dos modelos de predição, a Figura 22 permite observar uma menor temperatura nas bordas do tumor, sugerindo que a geometria do modelo exerce uma influência considerável sobre valores medidos nos em cada ponto de sensoriamento. No centro do tumor, onde as nanopartículas magnéticas foram injetadas, as temperaturas são mais elevadas, enquanto no ponto IV a temperatura local varia menos de 2 °C. Isso mostra que, apesar da semelhança nas coordenadas dos pontos de sensoriamento entre os tumores oblato e esférico, não se pode garantir que um único modelo seja capaz de estimar a temperatura local em todos os cenários de tratamento. Portanto, é possível enfatizar a necessidade de que os sistemas de hipertermia magnética adaptem suas estratégias de controle conforme as necessidades de cada paciente.

A Tabela 7 permite uma melhor compreensão a respeito do desempenho de cada modelo de predição no conjunto de teste. Para a ESN, verifica-se um MSE menor que 5% para todas as quatro saídas. Por sua vez, nota-se um MSE de até 11% para o modelo não linear. Para o modelo linear, verifica-se um MSE de até 24%, o que é mais um indício da forte não linearidade do processo nas condições representadas pelo tumor oblato.

Tabela 7 – Valores do NMSE para a ESN proposta e os dois modelos de referência na tarefa de predição da temperatura nos pontos I, II, III e IV na região do tumor elipsoidal oblato.

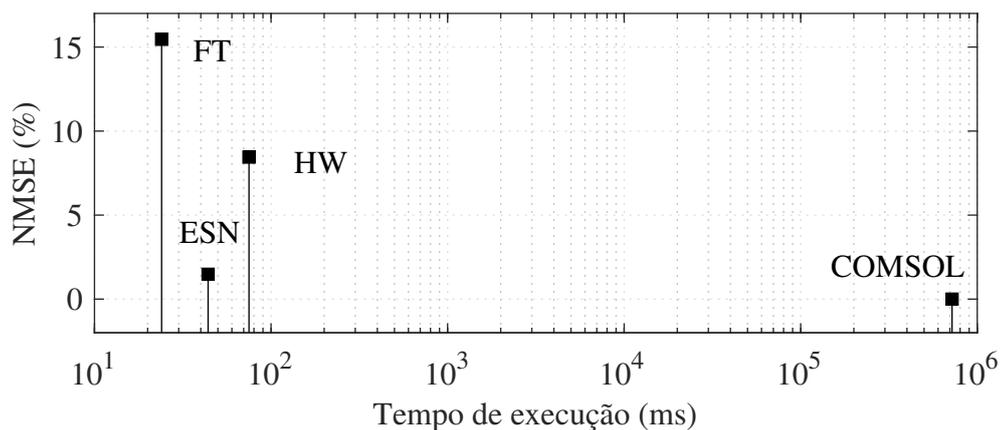
Ponto	Rede de estado de eco	Modelo de Hammerstein-Wiener	Função de transferência
I	0,0047	0,0877	0,1442
II	0,0030	0,0473	0,0930
III	0,0107	0,0954	0,1428
IV	0,0408	0,1075	0,2382

Fonte: autor.

Por fim, na Figura 23 são apresentados os tempos de execução e o NMSE médio para cada modelo no conjunto de teste. Desse modo, é possível verificar que os modelos de referência apresentam um tempo de execução similar aos relatados na Seção 6.1. Nesse caso, o menor tempo observado foi de 24 ms para o modelo linear, o qual apresentou um NMSE de 15,46%. Em contrapartida, o modelo não linear apresentou um tempo de execução de aproximadamente 75 ms e um NMSE médio de 8,45%. Por sua vez, a ESN completou todos os cálculos em 44 ms, com um NMSE médio de 1,48%. Ao comparar os tempos de execução entre a ESNs propostas nesta seção e na Seção 6.1, verifica-se uma redução de duas ordens de magnitude para o mesmo. Isso está diretamente relacionado à diferença no tamanho dos reservatórios entre os modelos. Portanto, conforme discutido na Seção 5.3, ainda há margem para reduzir o custo computacional da execução das redes de estado de eco. Para atingir esse objetivo, é possível buscar um equilíbrio

mais adequado entre o tamanho do reservatório e o desempenho da ESN, ou realizar ajustes mais significativos na esparsidade da matriz de pesos das conexões recorrentes, \mathbf{W}^{res} , conforme detalhado em Lukoševičius (2012).

Figura 23 – Comparação entre o NMSE e o tempo de execução para cada modelo no conjunto de teste do modelo oblato.



Fonte: autor.

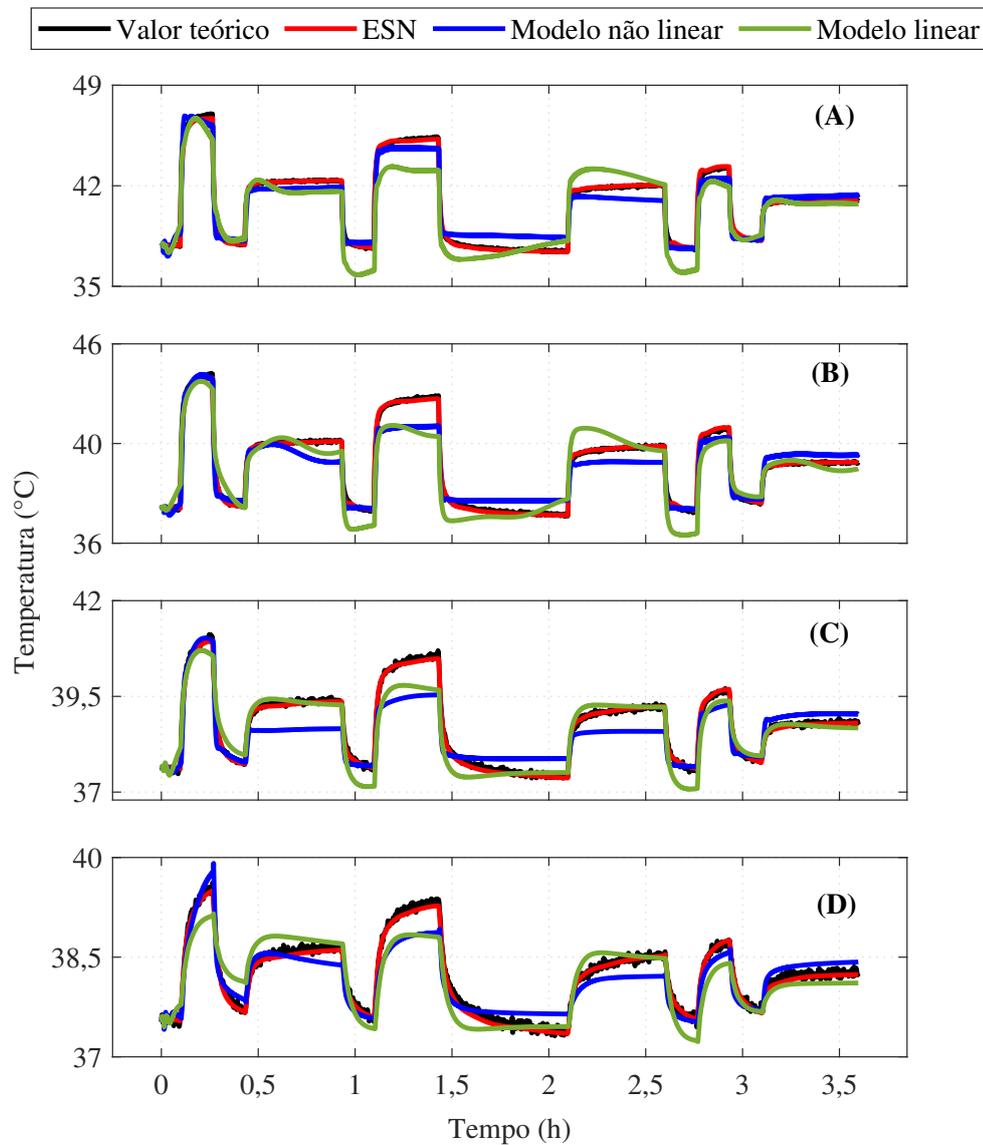
6.3 MODELO ELIPSOIDAL PROLATO

Para fins de comparação, foram propostos um modelo linear, baseado em uma função de transferência de sexta ordem, e um modelo não linear de Hammerstein-Wiener. Este último também adota uma mesma função de transferência de sexta ordem como modelo dinâmico e utiliza polinômios de grau 4 para cada uma das entradas, juntamente com um polinômio de grau 2 para as saídas.

A Figura 24 permite observar duas questões. A primeira, confirma a adaptabilidade das ESNs a diferentes cenários de tratamento. Mesmo com mudanças na geometria do tumor, a rede neural consegue modelar o perfil de temperatura local e estimar o seu valor para vários instantes à frente. Em contrapartida, os modelos de referência seguem apresentando limitações para desempenhar a mesma tarefa. A segunda questão que pode ser constatada é a proximidade dos valores de temperatura em cada ponto de medição, o que ocorre em função das coordenadas escolhidas para os mesmos.

A Tabela 8 demonstra o desempenho de cada modelo de predição no conjunto de teste. Ao observá-la, nota-se que a ESN demonstrou um NMSE médio de até 1% para todas as saídas, em contraste com o modelo de Hammerstein-Wiener, que apresentou um NMSE de até 25% no conjunto de teste. Por outro lado, embora o modelo linear tenha atingido um NMSE de até 20%, é evidente que seu desempenho é inferior em relação a todos os modelos propostos.

Figura 24 – Predições para temperaturas na região do tumor no modelo elipsoidal prolato: (A) ponto I; (B) ponto II; (C) ponto III; e (D) ponto IV.



Fonte: autor.

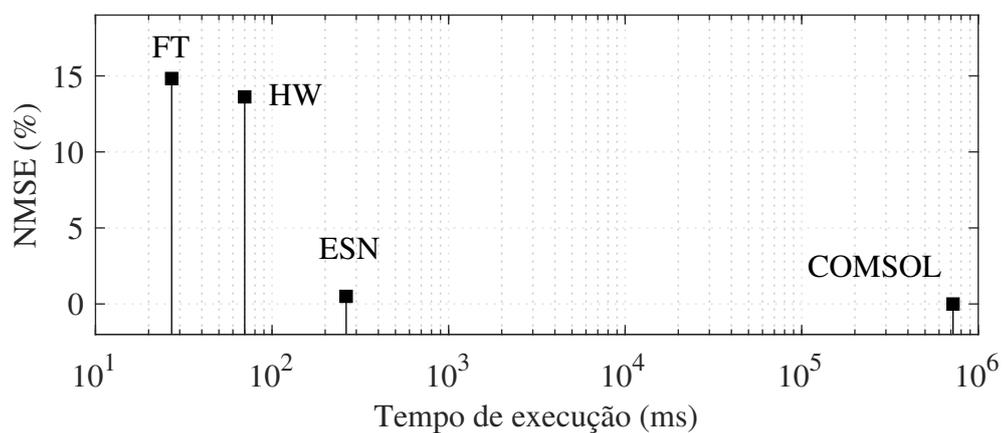
Finalmente, registrou-se o tempo requerido por cada modelo para realizara as predições para o conjunto de teste, considerando o erro de predição para cada modelo proposto. Conforme apresentado na Figura 25, o modelo linear necessitou de 27 ms para essa tarefa e apresentou um NMSE médio de 14,83%. Por outro lado, o modelo de Hammerstein-Wiener demandou 70 ms para realizar todos os cálculos. Além disso, o modelo em questão apresentou um NMSE médio de 13,62%. Quanto a ESN, esta completou a tarefa em aproximadamente 263 ms, com um NMSE de 0,5%.

Tabela 8 – Valores do NMSE para a ESN proposta e os dois modelos de referência na tarefa de predição da temperatura nos pontos I, II, III e IV na região do tumor elipsoidal prolato.

Ponto	Rede de estado de eco	Modelo de Hammerstein-Wiener	Função de transferência
I	0,0031	0,0538	0,1471
II	0,0027	0,1028	0,1434
III	0,0040	0,2450	0,1105
IV	0,0103	0,1429	0,1919

Fonte: autor.

Figura 25 – Comparação entre o NMSE e o tempo de execução para cada modelo no conjunto de teste do modelo prolato.



Fonte: autor.

6.4 COMENTÁRIOS FINAIS

Com base nos resultados apresentados neste capítulo, pode-se inferir que os modelos baseados nas ESNs são promissores para a utilização como modelos de predição em aplicações que envolvem controles preditivos baseados em modelo no contexto do tratamento de hipertermia magnética. Em todas as seções, foi observado que as ESNs demonstraram um pequeno erro de predição, o que não foi o caso para os modelos de referência. Além disso, vale destacar a capacidade de adaptação das ESNs a diferentes cenários de tratamento. Outro aspecto que pode ser ressaltado é o tempo de execução desses modelos. Enquanto a aplicação dos modelos numéricos baseados nas equações governantes exigiu aproximadamente 12 min para calcular as predições da temperatura em toda a região do tumor no conjunto de teste, as ESNs realizaram a mesma tarefa em menos de 1 s.

7 CONSIDERAÇÕES FINAIS

Neste capítulo, são apresentadas as conclusões para esta dissertação, assim como sugestões para trabalhos futuros. A Seção 7.1 discute as contribuições deste trabalho para a literatura e as conclusões gerais a respeito do problema explorado. Em seguida, na Seção 7.2 são comentadas as perspectivas para o prosseguimento da pesquisa e as sugestões para trabalhos futuros.

7.1 CONCLUSÕES

Nesta dissertação, em primeiro lugar, foi apresentado um modelo abrangente que simula todas as etapas de um tratamento de hipertermia magnética. A partir do método proposto, foi possível modelar as etapas básicas do tratamento, que incluem a infusão e difusão do nanofluido na região do tumor, bem como a fase de aquecimento por indução. Tais implementações foram, então, verificadas por meio de trabalhos disponíveis na literatura. Com base nisso, foram desenvolvidos três modelos numéricos compostos por tumores hepáticos com simetria axial, para simular diferentes cenários de tratamento.

Com base nos resultados dos estudos de caso, fica evidente que a simulação numérica de um cenário de tratamento de hipertermia magnética por meio das equações diferenciais parciais que descrevem a física do procedimento exige um custo computacional considerável. Portanto, pode-se concluir que a adoção dessas equações como um modelo de processo para o projeto de controladores, especialmente os controladores preditivos baseados em modelos, pode apresentar desafios significativos em termos de eficiência computacional e viabilidade prática.

Para lidar com essas circunstâncias, este estudo propôs o uso de uma rede neural de estado de eco para identificar a dinâmica de sistemas multivariáveis, tendo a amplitude e a frequência do campo magnético alternado como entradas e quatro pontos de medição de temperatura na região dos tumores hepáticos como saídas. Os resultados foram comparados com dois modelos de referência: um modelo linear e um modelo de Hammerstein-Wiener. Considerando que a dinâmica do aumento de temperatura na região do tumor apresenta evidências de comportamento não linear, observou-se um erro de predição considerável para funções de transferência. Por outro lado, os modelos de Hammerstein-Wiener apresentaram um desempenho superior aos modelos lineares, mas ainda abaixo daquele apresentado pelas redes neurais de estado de eco. Durante os testes, as ESNs apresentaram um NMSE de até 5% em todos os estudos de caso, enquanto os modelos de referência apresentaram um NMSE de até 25%.

Em termos de custo computacional, todos os modelos apresentaram tempos de execução da ordem de milissegundos, mesmo em horizontes de predição com mais de 1000 amostras. Esses intervalos de tempo são aproximadamente três ordens de magnitude menores do que o

tempo necessário para realizar a mesma tarefa utilizando um modelo numérico baseado em equações governantes. Portanto, o modelo que utiliza a rede neural de estado de eco demonstrou ser uma escolha vantajosa, oferecendo um equilíbrio entre a precisão das predições e o custo computacional, tornando-o um bom candidato para aplicações de controle preditivo baseado em modelo.

Considerando as conclusões apresentadas, é evidente que todos os objetivos propostos neste trabalho foram alcançados. A revisão da literatura proporcionou uma compreensão das etapas básicas do tratamento de hipertermia magnética. Com base nesse conhecimento, foram desenvolvidos diversos modelos numéricos, baseados em princípios físicos e capazes de representar diferentes cenários de tratamento. Além disso, foi possível identificar cada sistema MIMO, representado pelos modelos numéricos, por meio das ESNs. Por sua vez, as ESNs demonstraram ser uma ferramenta promissora para a predição de temperatura em tumores durante o tratamento, com perspectivas para uso em aplicações envolvendo controle preditivo baseado em modelo.

7.2 SUGESTÕES PARA TRABALHOS FUTUROS

Diversas sugestões podem ser feitas para trabalhos futuros. Primeiramente, é possível considerar um cenário mais realista do tratamento de hipertermia magnética. Para isso, é possível adicionar mais equações governantes aos modelos numéricos, como as equações de Maxwell e as equações que permitem descrever a contribuição das correntes parasitas para o processo de aquecimento. Em segundo lugar, o algoritmo para o treinamento e validação da ESN pode ser melhorado a partir da implementação de uma validação cruzada por k -partições, a exemplo do que foi desenvolvido em Racca e Magri (2021). Nesse caso, a ideia é obter um modelo ainda mais confiável e com capacidade de ser ajustado a partir de uma quantidade menor de dados (Nelles, 2001).

Outra sugestão é a implementação de um processo de identificação on-line para sistemas de hipertermia magnética. Nesse caso, pode-se propor que os modelos baseados na ESN sejam ajustados a partir de dados obtidos em tempo real, considerando as devidas restrições de segurança para o processo. Tal abordagem será investigada em trabalhos futuros, uma vez que pode conferir uma maior robustez aos controladores preditivos baseados em modelo, frente às incertezas do processo.

Além das sugestões apresentadas, a partir das discussões feitas no trabalho de Cai *et al.* (2021), é possível propor a realização de estudos voltados para a utilização de redes neurais baseadas em princípios físicos. Estes estudos também poderiam ter os mesmos objetivos desta dissertação, assim como a utilização de tais redes para o desenvolvimento de MPCs, considerando um cenário em que há poucos dados disponíveis.

REFERÊNCIAS

- AGUIRRE, L. A. **Introdução à identificação de sistemas: técnicas lineares e não lineares aplicadas a sistemas reais**. 3. ed. Minas Gerais: Editora da UFMG, 2007.
- ANDERSON, J. A. **An introduction to neural networks**. MIT Press, 1995.
- ANTONELO, E. A.; CAMPONOGARA, E.; FOSS, B. Echo state networks for data-driven downhole pressure estimation in gas-lift oil wells. **Neural Networks**, v. 85, p. 106–117, 2017.
- ARMENIO, L. B.; TERZI, E.; FARINA, M.; SCATTOLINI, R. Model predictive control design for dynamical systems learned by echo state networks. **IEEE Control Systems Letters**, v. 3, n. 4, p. 1044–1049, 2019.
- ATKINSON, W. J.; BREZOVICH, I. A.; CHAKRABORTY, D. P. Usable frequencies in hyperthermia with thermal seeds. **IEEE Transactions on Biomedical Engineering**, BME-31, n. 1, p. 70–75, 1984.
- BAXTER, L. T.; JAIN, R. K. Transport of fluid and macromolecules in tumors. I. Role of interstitial pressure and convection. **Microvascular Research**, v. 37, n. 1, p. 77–104, 1989.
- BRASIL. **Estimativa 2023: incidência de câncer no Brasil**. Rio de Janeiro: Instituto Nacional de Combate ao Câncer, 2023.
- BROCHU, E.; CORA, V. M.; FREITAS, N. de. **A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning**. 2010. arXiv: 1012.2599 [cs.LG]. Disponível em: <http://arxiv.org/abs/1012.2599>.
- CAI, S. *et al.* Physics-informed neural networks for heat transfer problems. **Journal of Heat Transfer**, v. 143, n. 6, 2021.
- CALLISTER, W. D.; RETHWISCH, D. G. **Ciência e engenharia de materiais: uma introdução**. 9. ed. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora Ltda, 2016.
- CAMACHO, E. F.; BORDONS, C. **Model predictive control**. Londres: Springer, 2007.
- DOYA, K. Bifurcations in the learning of recurrent neural networks. *In: IEEE International Symposium on Circuits and Systems*. San Diego, Estados Unidos: IEEE, 1992. v. 6, p. 2777–2780.

DUTTA, J.; KUNDU, B.; YOON, S.-J. Three-dimensional thermal assessment in cancerous tumors based on local thermal non-equilibrium approach for hyperthermia treatment.

International Journal of Thermal Sciences, v. 159, p. 106591, 2021.

ELMAN, J. Finding structure in time. **Cognitive Science**, v. 14, n. 2, p. 179–211, 1990.

FERNÁNDEZ, S. D. S. *et al.* Tunable control of the structural features and related physical properties of $Mn_xFe_{3-x}O_4$ nanoparticles: implication on their heating performance by magnetic hyperthermia. **The Journal of Physical Chemistry C**, v. 126, n. 24, p. 10110–10128, 2022.

GARNETT, R. *et al.* Sequential bayesian prediction in the presence of changepoints and faults. **The Computer Journal**, v. 53, n. 9, p. 1430–1446, 2010.

GILCHRIST, R. K. *et al.* Selective inductive heating of lymph nodes. **Annals of Surgery**, v. 146, n. 4, p. 596–606, 1957.

GREENHILL, S. *et al.* Bayesian optimization for adaptive experimental design: a review. **IEEE Access**, v. 8, p. 13937–13948, 2020.

HAEMMERICH, D.; WEBSTER, J. G. Automatic control of finite element models for temperature-controlled radiofrequency ablation. **BioMedical Engineering Online**, v. 4, n. 1, p. 1–8, 2005.

HALES, R. E.; YUDOFKY, S. C.; GABBARD, G. O. **Tratado de psiquiatria clínica**. 5. ed. Porto Alegre: Artmed, 2012.

HANKIEWICZ, J. H. *et al.* Ferromagnetic particles as magnetic resonance imaging temperature sensors. **Nature Communications**, v. 7, n. 1, 2016.

HANNON, G.; TANSI, F. L.; HILGER, I.; PRINA-MELLO, A. The effects of localized heat on the hallmarks of cancer. **Advanced Therapeutics**, v. 4, n. 7, p. 2000267, 2021.

HAUSMAN, D. M. What is cancer? **Perspectives in Biology and Medicine**, v. 62, n. 4, p. 778–784, 2019.

HAYKIN, S. **Neural networks and learning machines**. 3. ed. Nova Jersey: Prentice Hall, 2009.

HERGT, R. *et al.* Maghemite nanoparticles with very high AC-losses for application in RF-magnetic hyperthermia. **Journal of Magnetism and Magnetic Materials**, v. 270, n. 3, p. 345–357, 2004.

HERGT, R.; DUTZ, S. Magnetic particle hyperthermia – biophysical limitations of a visionary tumour therapy. **Journal of Magnetism and Magnetic Materials**, v. 311, n. 1, p. 187–192, 2007.

HOCHREITER, S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. **International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems**, v. 06, n. 02, p. 107–116, 1998.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1997.

HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. **Proceedings of the National Academy of Sciences**, v. 79, n. 8, p. 2554–2558, 1982.

HUANG, H. S.; HAINFELD, J. F. Intravenous magnetic nanoparticle cancer hyperthermia. **International Journal of Nanomedicine**, v. 8, p. 2521, 2013.

ILG, P.; KRÖGER, M. Dynamics of interacting magnetic nanoparticles: effective behavior from competition between Brownian and Néel relaxation. **Physical Chemistry Chemical Physics**, v. 22, n. 39, p. 22244–22259, 2020.

JAEGER, H. Adaptive nonlinear system identification with echo state networks. *In*: BECKER, S.; THRUN, S.; OBERMAYE, K. (Ed.). **Advances in Neural Information Processing Systems 15**. MIT Press, 2002.

JAEGER, H. **The “echo state” approach to analysing and training recurrent neural networks**. Bonn, Alemanha: Fraunhofer Institute for Autonomous Intelligent Systems, 2001.

JOHANNSEN, M.; THIESEN, B.; WUST, P.; JORDAN, A. Magnetic nanoparticle hyperthermia for prostate cancer. **International Journal of Hyperthermia**, v. 26, n. 8, p. 790–795, 2010.

JOHANNSEN, M. *et al.* Magnetic fluid hyperthermia (MFH) reduces prostate cancer growth in the orthotopic dunning R3327 rat model. **The Prostate**, v. 64, n. 3, p. 283–292, 2005.

JORDAN, A. *et al.* Presentation of a new magnetic field therapy system for the treatment of human solid tumors with magnetic fluid hyperthermia. **Journal of Magnetism and Magnetic Materials**, v. 225, n. 1-2, p. 118–126, 2001.

JORDAN, A. *et al.* The effect of thermotherapy using magnetic nanoparticles on rat malignant glioma. **Journal of Neuro-Oncology**, v. 78, n. 1, p. 7–14, 2006.

JORDAN, M. I. **Serial order**: a parallel distributed processing approach. San Diego, 1986. P. 46.

JORDANOU, J. P.; CAMPONOGARA, E.; ANTONELLO, E. A.; AGUIAR, M. A. S. Nonlinear model predictive control of an oil well with echo state networks. *In*: 3rd IFAC Workshop on Automatic Control in Offshore Oil and Gas Production OOGP 2018. Esbjerg, Dinamarca: IFAC, 2018. v. 51, p. 13–18.

JOSE, J. *et al.* Magnetic nanoparticles for hyperthermia in cancer treatment: an emerging tool. **Environmental Science and Pollution Research International**, v. 27, n. 16, p. 19214–19225, 2020.

LEGGÉ, C. J.; COLLEY, H. E.; LAWSON, M. A.; RAWLINGS, A. E. Targeted magnetic nanoparticle hyperthermia for the treatment of oral cancer. **Journal of Oral Pathology & Medicine**, v. 48, n. 9, p. 803–809, 2019.

LIANGRUKSA, M.; GANGULY, R.; PURI, I. K. Parametric investigation of heating due to magnetic fluid hyperthermia in a tumor with blood perfusion. **Journal of Magnetism and Magnetic Materials**, v. 323, n. 6, p. 708–716, 2011.

LIN, C. T.; LIU, K. C. Estimation for the heating effect of magnetic nanoparticles in perfused tissues. **International Communications in Heat and Mass Transfer**, v. 36, n. 3, p. 241–244, 2009.

LIU, Y.-D. *et al.* Adaptive ultrasound temperature imaging for monitoring radiofrequency ablation. **PLOS ONE**, v. 12, n. 8, e0182457, 2017.

LJUNG, L. **System identification**: theory for the user. 2. ed. Nova Jersey: Prentice Hall, 1999.

LUKOŠEVIČIUS, M. A practical guide to applying echo state networks. *In*: MONTAVON, G.; ORR, G. B.; MÜLLER, K.-R. (Ed.). **Neural Networks: Tricks of the Trade**. 2. ed. Berlin, Heidelberg: Springer Verlag, 2012. v. 7700. p. 659–686.

MA, M. *et al.* Size dependence of specific power absorption of Fe₃O₄ particles in AC magnetic field. **Journal of Magnetism and Magnetic Materials**, v. 268, n. 1-2, p. 33–39, 2004.

MAASS, W.; NATSCHLÄGER, T.; MARKRAM, H. Real-time computing without stable states: a new framework for neural computation based on perturbations. **Neural Computation**, v. 14, n. 11, p. 2531–2560, 2002.

MAHESH, N.; SINGH, N.; TALUKDAR, P. A mathematical model for understanding nanoparticle biodistribution after intratumoral injection in cancer tumors. **Journal of Drug Delivery Science and Technology**, v. 68, p. 103048, 2022.

MAIER-HAUFF, K. *et al.* Efficacy and safety of intratumoral thermotherapy using magnetic iron-oxide nanoparticles combined with external beam radiotherapy on patients with recurrent glioblastoma multiforme. **Journal of Neuro-Oncology**, v. 103, n. 2, p. 317–324, 2011.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, v. 5, n. 4, p. 115–133, 1943.

MITCHELL, T. M. **Machine learning**. Nova York, Estados Unidos: McGraw Hill, 1997.

NELLES, O. **Nonlinear system identification**. Berlin, Heidelberg: Springer, 2001.

OBAIDAT, I. M. *et al.* Principles of Magnetic Hyperthermia: A Focus on Using Multifunctional Hybrid Magnetic Nanoparticles. **Magnetochemistry**, v. 5, n. 4, p. 67, 2019.

PAN, J. *et al.* Mild magnetic hyperthermia-activated innate immunity for liver cancer therapy. **Journal of the American Chemical Society**, v. 143, n. 21, p. 8116–8128, 2021.

PAN, Y.; WANG, J. Model predictive control of unknown nonlinear dynamical systems based on recurrent neural networks. **IEEE Transactions on Industrial Electronics**, v. 59, n. 8, p. 3089–3101, 2012.

PANKHURST, Q. A.; CONNOLLY, J.; JONES, S. K.; DOBSON, J. Applications of magnetic nanoparticles in biomedicine. **Journal of Physics D: Applied Physics**, v. 36, n. 13, r167–r181, 2003.

PENNES, H. H. Analysis of tissue and arterial blood temperatures in the resting human forearm. **Journal of Applied Physiology**, v. 1, n. 2, p. 93–122, 1948.

RACCA, A.; MAGRI, L. Robust optimization and validation of echo state networks for learning chaotic dynamics. **Neural Networks**, v. 142, p. 252–268, 2021.

RAHPEIMA, R.; LIN, C. A. Numerical study of magnetic hyperthermia ablation of breast tumor on an anatomically realistic breast phantom. **PLOS ONE**, v. 17, n. 9, e0274801, 2022.

RAO, P. A. *et al.* A systematic study of cobalt-zinc ferrite nanoparticles for self-regulated magnetic hyperthermia. **Journal of Alloys and Compounds**, v. 794, p. 60–67, 2019.

RASMUSSEN, C. E.; WILLIAMS, C. K. I. **Gaussian processes for machine learning**. The MIT Press, 2005.

REZVANIAN, A.; BEIGZADEH, B.; MARKAZI, A. H. D.; HALABIAN, M. Feedback control of temperature in specific geometry of porous media: application to hyperthermia. **Journal of Thermal Analysis and Calorimetry**, v. 141, n. 5, p. 1559–1568, 2020.

RODRIGUES, H. F.; CAPISTRANO, G.; BAKUZIS, A. F. *In vivo* magnetic nanoparticle hyperthermia: a review on preclinical studies, low-field nano-heaters, noninvasive thermometry and computer simulations for treatment planning. **International Journal of Hyperthermia**, v. 37, n. 3, p. 76–99, 2020.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological Review**, v. 65, n. 6, p. 386–408, 1958.

ROSENSWEIG, R. Heating magnetic fluid with alternating magnetic field. **Journal of Magnetism and Magnetic Materials**, v. 252, p. 370–374, 2002.

SCHIRRMACHER, V. From chemotherapy to biological therapy: a review of novel concepts to reduce the side effects of systemic cancer treatment (review). **International Journal of Oncology**, v. 54, n. 2, p. 407–419, 2019.

SCHWEDERSKY, B. B.; FLESCH, R. C.; ROVEA, S. B. Adaptive practical nonlinear model predictive control for echo state network models. **IEEE Transactions on Neural Networks and Learning Systems**, v. 33, n. 6, p. 2605–2614, 2022.

SCHWEDERSKY, B. B. **Contribuições para o controle preditivo não linear prático com modelos identificados por redes de estado de eco**. 2022. 148 f. Tese (Doutorado) – Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia de Automação e Sistemas, Florianópolis, 2022.

SCHWEDERSKY, B. B.; FLESCH, R. C. C.; DANGUI, H. A. S. Nonlinear MIMO system identification with echo-state networks. **Journal of Control, Automation and Electrical Systems**, v. 33, n. 3, p. 743–754, 2022.

SCHWEDERSKY, B. B.; FLESCH, R. C. C.; DANGUI, H. A. S.; IERVOLINO, L. A. Practical nonlinear model predictive control using an echo state network model. *In*: 2018 International Joint Conference on Neural Networks (IJCNN). Rio de Janeiro: IEEE, 2018. p. 1–8.

SEBEKE, L. *et al.* Model predictive control for MR-HIFU-mediated, uniform hyperthermia. **International Journal of Hyperthermia**, v. 36, n. 1, p. 1039–1049, 2019.

SEFIDGAR, M.; BASHOOKI, E.; SHOJAEE, P. Numerical simulation of the effect of necrosis area in systemic delivery of magnetic nanoparticles in hyperthermia cancer treatment. **Journal of Thermal Biology**, v. 94, p. 102742, 2020.

SHARMA, A. *et al.* Design of a temperature-feedback controlled automated magnetic hyperthermia therapy device. **Frontiers in Thermal Engineering**, v. 3, p. 1131262, 2023.

SHOAIB, M.; SHAMSELDIN, A. Y.; MELVILLE, B. W.; KHAN, M. M. A comparison between wavelet based static and dynamic neural network approaches for runoff prediction. **Journal of Hydrology**, v. 535, p. 211–225, 2016.

SIEGEL, R. L.; MILLER, K. D.; WAGLE, N. S.; JEMAL, A. Cancer statistics, 2023. **CA: A Cancer Journal for Clinicians**, v. 73, n. 1, p. 17–48, 2023.

SNOEK, J.; LAROCHELLE, H.; ADAMS, R. P. Practical bayesian optimization of machine learning algorithms. *In*: PEREIRA, F.; BURGESS, C. J.; BOTTOU, L.; WEINBERGER, K. Q. (Ed.). **Advances in Neural Information Processing Systems**. Curran Associates, Inc., 2012. v. 25.

SOLEYMANI, M. *et al.* Folic acid-conjugated dextran-coated $Zn_{0.6}Mn_{0.4}Fe_2O_4$ nanoparticles as systemically delivered nano heaters with self-regulating temperature for magnetic hyperthermia therapy of liver tumors. **Scientific Reports**, v. 13, n. 1, 2023.

STEIL, J. Backpropagation-decorrelation: online recurrent learning with $O(N)$ complexity. *In*: IEEE International Joint Conference on Neural Networks. Budapeste: IEEE, 2004. v. 2, p. 843–848.

STEUPERAERT, M. *et al.* Mathematical modeling of intraperitoneal drug delivery: simulation of drug distribution in a single tumor nodule. **Drug Delivery**, v. 24, n. 1, p. 491–501, 2017.

STRATTON, M. R.; CAMPBELL, P. J.; FUTREAL, P. A. The cancer genome. **Nature**, v. 458, n. 7239, p. 719–24, 2009.

SULEMAN, M.; RIAZ, S. 3D in silico study of magnetic fluid hyperthermia of breast tumor using Fe_3O_4 magnetic nanoparticles. **Journal of Thermal Biology**, v. 91, p. 102635, 2020.

SULEMAN, M.; RIAZ, S. In silico study of hyperthermia treatment of liver cancer using core-shell $CoFe_2O_4@MnFe_2O_4$ magnetic nanoparticles. **Journal of Magnetism and Magnetic Materials**, v. 498, p. 166143, 2020.

TANAKA, G. *et al.* Recent advances in physical reservoir computing: a review. **Neural Networks**, v. 115, p. 100–123, 2019.

TANG, Y.; FLESCH, R. C. C.; JIN, T. Numerical analysis of temperature field improvement with nanoparticles designed to achieve critical power dissipation in magnetic hyperthermia. **Journal of Applied Physics**, v. 122, n. 3, p. 34702, 2017.

TANG, Y.; FLESCHE, R. C. C.; JIN, T.; HE, M. Computational evaluation of malignant tissue apoptosis in magnetic hyperthermia considering intratumoral injection strategy. **International Journal of Heat and Mass Transfer**, v. 178, p. 121609, 2021.

TANG, Y.; SU, H.; JIN, T.; FLESCHE, R. C. C. Adaptive PID control approach considering simulated annealing algorithm for thermal damage of brain tumor during magnetic hyperthermia. **IEEE Transactions on Instrumentation and Measurement**, v. 72, 2023.

TANG, Y.; ZOU, J.; FLESCHE, R. C.; JIN, T. Effect of injection strategy for nanofluid transport on thermal damage behavior inside biological tissue during magnetic hyperthermia. **International Communications in Heat and Mass Transfer**, v. 133, p. 105979, 2022.

TANG, Y.; ZOU, J.; FLESCHE, R. C. C.; JIN, T. Backflow modeling in nanofluid infusion and analysis of its effects on heat induced damage during magnetic hyperthermia. **Applied Mathematical Modelling**, v. 114, p. 583–600, 2023.

THIEDE, L. A.; PARLITZ, U. Gradient based hyperparameter optimization in Echo State Networks. **Neural Networks**, v. 115, p. 23–29, 2019.

VU, T.; RAZANSKY, D.; YAO, J. Listening to tissues with new light: recent technological advances in photoacoustic imaging. **Journal of Optics**, v. 21, n. 10, p. 103001, 2019.

WANG, W.-C.; TAI, C.-C. Magnetic induction of hyperthermia by a modified self-learning fuzzy temperature controller. **The Review of Scientific Instruments**, v. 88, n. 7, p. 074702, 2017.

XIAOHUA, F.; FEI, G.; YUANJIN, Z. Photoacoustic-based-close-loop temperature control for nanoparticle hyperthermia. **IEEE Transactions on Biomedical Engineering**, v. 62, n. 7, p. 1728–1737, 2015.

YANG, L.; SHAMI, A. On hyperparameter optimization of machine learning algorithms: Theory and practice. **Neurocomputing**, v. 415, p. 295–316, 2020.

YPERMAN, J.; BECKER, T. **Bayesian optimization of hyper-parameters in reservoir computing**. arXiv, 2016. Disponível em: <https://arxiv.org/abs/1611.05193>.

YU, X. *et al.* The heating efficiency of magnetic nanoparticles under an alternating magnetic field. **Scientific Reports**, v. 12, n. 1, p. 16055, 2022.

ZAIMY, M. A. *et al.* New methods in the diagnosis of cancer and gene therapy of cancer based on nanoparticles. **Cancer Gene Therapy**, v. 24, n. 6, p. 233–243, 2017.

ZHANG, C. *et al.* Ferroptosis in cancer therapy: a novel approach to reversing drug resistance. **Molecular Cancer**, v. 21, n. 1, p. 47, 2022.

ZHANG, W. *et al.* Magnetic nanoparticles with low Curie temperature and high heating efficiency for self-regulating temperature hyperthermia. **Journal of Magnetism and Magnetic Materials**, v. 489, p. 165382, 2019.