



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Gabriel Inácio Barboza

**Descobrimento de Parâmetros em Modelos
Epidemiológicos Compartmentais usando Redes
Neurais Informadas pela Física**

Vitória, ES

2025

Gabriel Inácio Barboza

**Descobrimento de Parâmetros em Modelos
Epidemiológicos Compartmentais usando Redes Neurais
Informadas pela Física**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

Universidade Federal do Espírito Santo – UFES

Centro Tecnológico

Programa de Pós-Graduação em Informática

Orientador: Prof. Dr. Isaac Pinheiro dos Santos

Vitória, ES

2025

Gabriel Inácio Barboza

Descobrimto de Parâmetros em Modelos Epidemiológicos Compartimentais usando Redes Neurais Informadas pela Física/ Gabriel Inácio Barboza. – Vitória, ES, 2025-

65 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Isaac Pinheiro dos Santos

Dissertação de Mestrado – Universidade Federal do Espírito Santo – UFES
Centro Tecnológico

Programa de Pós-Graduação em Informática, 2025.

1. Palavra-chave1. 2. Palavra-chave2. I. Souza, Vítor Estêvão Silva. II. Universidade Federal do Espírito Santo. IV. Descobrimto de Parâmetros em Modelos Epidemiológicos Compartimentais usando Redes Neurais Informadas pela Física

CDU 02:141:005.7

Gabriel Inácio Barboza

Descobrimento de Parâmetros em Modelos Epidemiológicos Compartmentais usando Redes Neurais Informadas pela Física

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

Trabalho aprovado. Vitória, ES, 25 de setembro de 2014:

Prof. Dr. Isaac Pinheiro dos Santos
Orientador

Professor
Convidado 1

Professor
Convidado 2

Vitória, ES
2025

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis malesuada laoreet leo at interdum. Nullam neque eros, dignissim sed ipsum sed, sagittis laoreet nisi.

Agradecimentos

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis malesuada laoreet leo at interdum. Nullam neque eros, dignissim sed ipsum sed, sagittis laoreet nisi. Duis a pulvinar nisl. Aenean varius nisl eu magna facilisis porttitor. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut mattis tortor nisi, facilisis molestie arcu hendrerit sed. Donec placerat velit at odio dignissim luctus. Suspendisse potenti. Integer tristique mattis arcu, ut venenatis nulla tempor non. Donec at tincidunt nulla. Cras ac dignissim neque. Morbi in odio nulla. Donec posuere sem finibus, auctor nisl eu, posuere nisl. Duis sit amet neque id massa vehicula commodo dapibus eu elit. Sed nec leo eu sem viverra aliquet. Nam at nunc nec massa rutrum aliquam sed ac ante.

Vivamus nec quam iaculis, tempus ipsum eu, cursus ante. Phasellus cursus euismod auctor. Fusce luctus mauris id tortor cursus, volutpat cursus lacus ornare. Proin tristique metus sed est semper, id finibus neque efficitur. Cras venenatis augue ac venenatis mollis. Maecenas nec tellus quis libero consequat suscipit. Aliquam enim leo, pretium non elementum sit amet, vestibulum ut diam. Maecenas vitae diam ligula.

Fusce ac pretium leo, in convallis augue. Mauris pulvinar elit rhoncus velit auctor finibus. Praesent et commodo est, eu luctus arcu. Vivamus ut porta tortor, eget facilisis ex. Nunc aliquet tristique mauris id sollicitudin. Donec quis commodo metus, sit amet accumsan nibh. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus.

*“Science is a differential equation. Religion is a boundary condition.
(Alan Turing)”*

Resumo

Desde 2020, com o surto de COVID-19, uma grande quantidade de dados foram coletados sobre a dispersão desta doença. Para entender como doenças se propagam e ajudar tomadores de decisão na criação de políticas para conter o avanço e mitigar os efeitos que doenças contagiosas trazem são empregados modelos que explicam o comportamento de tais doenças. Modelos compartimentais baseados em equações diferenciais ordinárias são comumente aplicados para a modelagem de fenômenos epidemiológicos por sua simplicidade, facilidade de análise, e pela existência de métodos numéricos capazes de resolver estes sistemas. Com o avanço de áreas como aprendizado de máquina e ciência de dados, foram desenvolvidas técnicas de descobrimento de equações em grandes bases de dados. Uma dessas técnicas, são as redes neurais informadas pela física, que atuam como um ajustador de curvas regularizado por equações que regem o fenômeno ao incorporá-las na função de perda da rede. Elas podem ser aplicadas para a solução numérica de equações diferenciais, mas também para a solução de problemas inversos envolvendo equações diferenciais, utilizando dados reais e sintéticos. Neste trabalho, são aplicadas redes neurais informadas pela física para o descobrimento de parâmetros que compõem os modelos compartimentais. São realizados experimentos com dados sintéticos para averiguar a capacidade das redes neurais informadas pela física de resolver problemas inversos e testar a resiliência dos métodos a ruídos. Em seguida, são utilizados dados reais para extrair o valor de infectiosidade ao longo do tempo.

Palavras-chaves: Redes Neurais Informadas pela Física. Epidemiologia. Modelos Compartimentais. Problemas Inversos.

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis malesuada laoreet leo at interdum. Nullam neque eros, dignissim sed ipsum sed, sagittis laoreet nisi. Duis a pulvinar nisl. Aenean varius nisl eu magna facilisis porttitor. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut mattis tortor nisi, facilisis molestie arcu hendrerit sed. Donec placerat velit at odio dignissim luctus. Suspendisse potenti. Integer tristique mattis arcu, ut venenatis nulla tempor non. Donec at tincidunt nulla. Cras ac dignissim neque. Morbi in odio nulla. Donec posuere sem finibus, auctor nisl eu, posuere nisl. Duis sit amet neque id massa vehicula commodo dapibus eu elit. Sed nec leo eu sem viverra aliquet. Nam at nunc nec massa rutrum aliquam sed ac ante.

Vivamus nec quam iaculis, tempus ipsum eu, cursus ante. Phasellus cursus euismod auctor. Fusce luctus mauris id tortor cursus, volutpat cursus lacus ornare. Proin tristique metus sed est semper, id finibus neque efficitur. Cras venenatis augue ac venenatis mollis. Maecenas nec tellus quis libero consequat suscipit. Aliquam enim leo, pretium non elementum sit amet, vestibulum ut diam. Maecenas vitae diam ligula.

Fusce ac pretium leo, in convallis augue. Mauris pulvinar elit rhoncus velit auctor finibus. Praesent et commodo est, eu luctus arcu. Vivamus ut porta tortor, eget facilisis ex. Nunc aliquet tristique mauris id sollicitudin. Donec quis commodo metus, sit amet accumsan nibh. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus.

Duis elementum dictum tristique. Integer mattis libero sit amet pretium euismod. Curabitur auctor eu augue ut ornare. Integer bibendum eros ullamcorper rhoncus convallis. Pellentesque non pretium ligula, sit amet bibendum eros. Nam venenatis ex felis, quis blandit nunc auctor sit amet. Maecenas ut eros pharetra, lobortis neque id, fermentum arcu. Cras neque dui, rhoncus feugiat leo id, semper facilisis lorem. Fusce non ex turpis. Nullam venenatis sed ligula ac lacinia.

Keywords: Physics Informed Neural Networks. Epidemiology. Compartmental Models. Inverse Problems.

Lista de ilustrações

Figura 1 – Representação gráfica de um neurónio de MacCough-Pitts. Fonte: elaborada pelos autores.	19
Figura 2 – Representação gráfica das redes <i>feedforward</i> . Fonte: elaborada pelos autores.	21
Figura 3 – Representação gráfica das PINNs. Fonte: elaborada pelos autores.	23
Figura 4 – Grafo para o <i>SIR</i> . Fonte: elaborada pelos autores.	30
Figura 5 – Exemplo do SIR com $\beta = 0.8$ e $\gamma = 0.1$	31
Figura 6 – Grafo para o <i>SIRD</i> . Fonte: elaborada pelos autores.	35
Figura 7 – Grafo para o <i>SIRD</i> . Fonte: elaborada pelos autores.	35
Figura 8 – Grafo para o <i>SIS</i> . Fonte: elaborada pelos autores.	35
Figura 9 – Grafo para o <i>SIRD</i> . Fonte: elaborada pelos autores.	36
Figura 10 – Representação gráfica das redes <i>feedforward</i> . Fonte: elaborada pelos autores.	43
Figura 11 – Na primeira. Fonte: elaborada pelos autores.	44
Figura 12 – Na primeira. Fonte: elaborada pelos autores.	45
Figura 13 – Na primeira. Fonte: elaborada pelos autores.	46
Figura 14 – Na primeira. Fonte: elaborada pelos autores.	47
Figura 15 – Na primeira. Fonte: elaborada pelos autores.	47
Figura 16 – Na primeira. Fonte: elaborada pelos autores.	48
Figura 17 – Na primeira. Fonte: elaborada pelos autores.	48
Figura 18 – Na primeira. Fonte: elaborada pelos autores.	52
Figura 19 – Na primeira. Fonte: elaborada pelos autores.	52
Figura 20 – Na primeira. Fonte: elaborada pelos autores.	53
Figura 21 – Na primeira. Fonte: elaborada pelos autores.	54
Figura 22 – Na primeira. Fonte: elaborada pelos autores.	54
Figura 23 – Na primeira. Fonte: elaborada pelos autores.	54

Lista de tabelas

Tabela 1 – Valores das métricas de erro (MSE , norma \mathcal{L}_2 e norma \mathcal{L}_∞) para as soluções aproximadas pela rede neural, em comparação com as soluções analíticas.	53
Tabela 2 – Valores das métricas de erro (MSE , norma \mathcal{L}_2 e norma \mathcal{L}_∞) para as soluções aproximadas pela rede neural, em comparação com as soluções analíticas.	53

Lista de abreviaturas e siglas

PINNs	Physics Informed Neural Networks
DA	Diferenciação Automática
EDO	Equação Diferencial Ordinária
EDP	Equação Diferencial Parcial
MEF	Método de Elementos Finitos
MDF	Método de Diferenças Finitas
SIR	Susceptible-Infected-Removed
SEIR	Susceptible-Exposed-Infected-Removed
SEIRD	Susceptible-Exposed-Infected-Removed-Deceased
MSE	Mean Square Error
RMSE	Root Mean Square Error
MLP	Multi-layer Perceptron
EINNs	Epidemiology Informed Neural Networks
DINNs	Deceased Informed Neural Networks
AWGN	Additive White Gaussian Noise

Sumário

1	INTRODUÇÃO	14
2	REDES NEURAIS INFORMADAS PELA FÍSICA	19
2.1	Redes Neurais Feedforward	19
2.2	Redes Neurais Informadas pela Física	22
2.2.1	PINNs como Solucionadores de Equações Diferenciais	22
2.2.2	Formulação para Problemas Inversos	23
2.2.3	Pontos de Colocação	24
2.3	Arquiteturas Alternativas	24
3	MODELOS COMPARTIMENTAIS	29
3.1	Origens	29
3.2	O modelo SIR	29
3.3	Pontos de equilíbrio	31
3.4	Pontos críticos	32
3.5	Matriz de nova geração	33
3.6	Identificabilidade de um Modelo	33
3.7	Análise de Sensibilidade	34
3.8	Problemas Inversos	34
3.9	Adicionando Compartimentos	35
3.10	Variantes do Modelo SIR	36
3.11	Aplicação de PINNs com Modelos Compartimentais	36
4	PROPOSTA DO TRABALHO	42
4.1	Estimativa de parâmetros	42
4.2	Justificativa para a Escolha do Modelo SIR	43
4.3	Testes com Dados Sintéticos	43
4.3.1	Avaliação dos Resultados	45
4.4	Testes com Dados Sintéticos Ruidosos	46
4.5	Testes com diferentes pontos de equilíbrio	47
4.6	Testes com Base de Dados Reais	47
4.6.1	Base de dados da OMS	48
4.6.2	Bases de dados do DataSUS	48
4.6.3	Tratamento dos Dados	48
4.7	Uso da Matriz de Nova Geração	49
4.8	Arquitetura da Rede	49

4.8.1	Parâmetro Ajustável	49
4.8.2	Regularização L2	49
4.8.3	Correlação com a Temperatura	50
4.8.3.1	Teste de Correlação de Pearson	50
4.8.3.2	Teste de Correlação de Spearman	50
4.9	Implementação e Execução	50
5	AVALIAÇÃO DO TRABALHO	52
5.1	Testes com Dados Sintéticos	52
5.2	Testes com Dados Ruidosos	53
5.3	Testes com Base de Dados Reais	55
6	CONSIDERAÇÕES FINAIS	56
	REFERÊNCIAS	57
	APÊNDICES	65

1 Introdução

Equações diferenciais são equações que descrevem uma relação entre uma função e suas derivadas, são de particular interesse para as ciências naturais por sua aplicação na modelagem de fenômenos naturais e de leis físicas. A principal distinção entre elas é feita pelo número de variáveis independentes, no caso de apenas uma variável, chama-se de equação parcial ordinária (EDO), e equação diferencial parcial (EDP) para o caso mais de uma variável independente. Muitas das equações de interesse para áreas como engenharia, física, ecologia, química e epidemiologia, apenas para nomear algumas áreas, não possuem soluções analíticas conhecidas, fazendo com a busca por métodos numéricos seja uma área de pesquisa ativa da matemática aplicada. Ao longo dos século passado, foram desenvolvidos métodos para a solução de equações, como o método de diferenças finitas (MDF) e o método de elementos finitos (MEF). Entretanto é necessário uma discretização do domínio das equações, gerando uma malha de pontos em que a solução será resolvida. A criação desta malha (*mesh*) não é uma tarefa trivial, e a qualidade da solução obtida está diretamente ligada a obtenção desta malha (RAISSI; PERDIKARIS; KARNIADAKIS, 2019). Logo, há um interesse em métodos que não necessitem de malhas.

Uma rede neural artificial, ou apenas redes neurais, são um modelo de computação inspirado no funcionamento do cérebro, seu desenvolvimento remonta a trabalhos pioneiros ainda nos anos 40 como (MCCULLOCH; PITTS, 1943). O potencial que redes neurais têm como solucionadores numéricos de equações diferenciais pode ser facilmente observado, pois redes neurais são aproximadores universais, ou seja são capazes de aproximar qualquer função, inclusive uma que seja a solução de uma equação diferencial. Esta propriedade é atestada pelo teorema da aproximação universal, demonstrado primeiramente para redes com largura arbitrária e função sigmóide por (CYBENKO, 1989), e para redes com no mínimo uma camada escondida por (HORNIK; STINCHCOMBE; WHITE, 1989). A versão do teorema demonstrada nesses artigos, atesta que que redes neurais com uma largura suficientemente grande são capazes de aproximar qualquer função. Anos depois, em (GRIPENBERG, 2003), a mesma propriedade foi demonstrada para redes com profundidade arbitrária e largura fixa.

O potencial das redes neurais para a solução numérica de equações diferenciais já havia percebido nos anos 90, em trabalhos como (PSICHOGIOS; UNGAR, 1992) que incorporou uma rede neural na modelagem de um bioreator para a estimativa de parâmetros que seriam difíceis de serem estimados apenas com princípios físicos e químicos. Trabalhos seguintes focariam em propor métodos para a solução de qualquer equação diferencial através de redes neurais. Um exemplo que pode ser citado é (MEADE; FERNANDEZ, 1994), sendo um dos primeiros a criar uma forma para solucionar EDOs

arbitrárias utilizando redes neurais, entretanto o método proposto necessita que certas limitações sejam impostas às entredas, pesos e viéses da rede. Um outro exemplo de proposta para a solução de qualquer equação diferencial é encontrada em (LAGARIS; LIKAS; FOTIADIS, 1998), sendo um dos primeiros trabalhos a propor um método não apenas para a solução de EDOs, mas também para sistemas de EDOs e até mesmo EDPs. Os autores separaram o problema em duas partes, as equações diferenciais em si e as condições de fronteira, sendo a primeira parte aproximada por uma rede neural *feedforward*, e a segunda parte é obtida por meio de restrições duras (*hard constraints*) impostas ao modelo.

Uma série de métodos alternativos, ou seja, que não usam apenas redes neurais *feedforward* como base, foram propostos utilizando redes neurais para a solução de equações diferenciais nos últimos anos. Como exemplo, pode-se citar as *Neural ODE* (CHEN et al., 2018), pensadas para a resolução de EDOs, o método define a evolução de uma EDO (interpretando a variável independente como sendo o tempo) como uma rede neural, criando uma solução contínua para a EDO, em oposição às redes *feedforward*, em que representam a evolução do sistema por meio de camadas discretas. Outro exemplo é o *Deep Galerkin* (SIRIGNANO; SPILIOPOULOS, 2018) que combina o princípio do método de Galerkin de solucionar uma EDP em partes do domínio, mas utiliza uma rede neural profunda para aproximar as soluções no lugar de métodos lineares aplicados a um conjunto de funções de bases. Entre os exemplos mais recentes, é possível citar as *Finite Element Neural Networks* (MITUSCH; FUNKE; KUCHTA, 2021), que combinam redes neurais com o FEM, criando um método capaz de resolver problemas diretos e inversos com reduzido custo computacional.

Entretanto, foi apenas com (RAISSI; PERDIKARIS; KARNIADAKIS, 2019) e a introdução do conceito das *Physics Informed Neural Networks* (PINNs) que se renovou o interesse em aplicar redes neurais para a solução de problemas científicos. A grande diferença na abordagem das PINNs em relação a propostas anteriores, é tratar não apenas as equações que compõem o modelo, mas também as condições de fronteira e iniciais, como residuais a serem minimizados pela função de perda. Ou seja, tratando a solução de uma equação diferencial como um problema de minimização, sendo as equações e as condições de fronteira e iniciais tratadas como restrições leves (*soft constraints*). Outra inovação, é a incorporação de dados ao treinamento da rede neural, permitindo a descoberta de parâmetros do modelo, através da reformulação do mesmo como um problema inverso. Inclusive, os autores elaboraram diversos experimentos com dados perturbados para atestar a resiliência do método a dados ruidosos, indicando que PINNs podem ser treinadas com dados comprometidos sem afetar drasticamente a qualidade da solução obtida. Outra possibilidade que o uso de dados proporciona, é a descoberta de partes das equações que compõem o modelo,

Esta abordagem foi possível com o desenvolvimento de técnicas de diferenciação automática (DA). A importância que a DA tem para as PINNs está ligada ao fato de que incorporar as equações do modelo a função de perda faz com que o cálculo da derivada para o algoritmo de retropropagação seja uma tarefa muito complexa. A diferenciação automática permite que não seja necessário calcular uma derivada analítica, nem utilizar derivadas numéricas, que podem levar a erros de arredondamento. Apesar da técnica ter sido desenvolvida ainda nos anos 60 e 70 com os trabalhos de (WENGERT, 1964) para *forward accumulation*, e (LINNAINMAA, 1976) para *inverse accumulation*, foi com sua incorporação às bibliotecas que implementam estes algoritmos como PyTorch (PASZKE et al., 2019) e TensorFlow (ABADI et al., 2016), que facilitou o uso desta técnica, e a adoção de funções de perda mais complexas.

Desde sua concepção, PINNs vem sendo aplicadas para diversos problemas de engenharia, por exemplo, escoamento de fluxos incompressíveis (JIN et al., 2021). Solução de problemas de física, como física quântica como a equação de Schrödinger (JIN; MATTHEAKIS; PROTOPAPAS, 2022) e modelos cosmológicos (CHANTADA et al., 2023). Pode-se também citar exemplos de aplicações para problemas de ecologia e química, como o problema de formação de padrões em modelos de reação-difusão, como em (GIAMPAOLO et al., 2022). Problemas envolvendo epidemiologia, como (SHAIER; RAISSI; SESHAIYER, 2022), em que os autores utilizam PINNs para ajustar modelos compartimentais para várias doenças como dengue, rubéola e gripe. Atestando a efetividade das PINNs para problemas de epidemiologia modelados por equações diferenciais.

Em epidemiologia, modelos compartimentais baseados em equações diferenciais são modelos que separam a população em compartimentos e modelam fluxos de indivíduos entre estes compartimentos como proporcional ao tamanho dos compartimentos envolvidos multiplicados por parâmetros. Estes parâmetros são taxas que indicam a evolução da doença ao longo do tempo, como a taxa de infecção, taxa de mortalidade e taxa de imunização da população. Fazendo com que a variação total de cada compartimento ao longo do tempo, ou seja a diferenciação do mesmo em relação ao tempo, seja igual a soma destes fluxos. Eles foram introduzidos por (KERMACK; MCKENDRICK, 1927), ao proporem o modelo *Susceptible-Infected-Removed* (SIR), que se separa a população, como o nome indica, em suscetíveis, infectados e recuperados. Modelos compartimentais são normalmente empregados em epidemiologia, em detrimento de modelos mais complexos, como aqueles baseados em agentes, por sua simplicidade e capacidade de indicar tendências a curto prazo na evolução de uma pandemia. Por serem modelos formados por um sistema de equações ordinárias que, se sabidas as condições iniciais de cada variável, ou seja o número inicial de indivíduos em cada compartimento, tem-se um problema de valor inicial (PVI), que pode ser facilmente resolvido por métodos numéricos mais simples, como o método de Runge-Kutta de quarta ordem.

Com a declaração da OMS (Organização Mundial de Saúde) em dezembro de 2019 da pandemia de COVID-19, o mundo teve que adotar medidas para conter o avanço da doença, como isolamento social e uso de máscaras. Junto às medidas de contenção da pandemia, governos mobilizaram os sistemas de saúde para coletar dados e permitir aos governantes e entidades responsáveis pela saúde pública, tomar decisões acerca das medidas restrição. Essas ações geraram uma quantidade grande de dados acerca da quantidade de casos notificados, internações, tempo de internação e mortes causadas pela COVID-19. Pesquisadores de diferentes áreas trabalharam em cima destes dados tentando encaixá-los em modelos compartimentais para oferecer às autoridades projeções da evolução da pandemia. As PINNs, que haviam sido propostas há pouco tempo, foram logo encaradas como mais uma ferramenta a ser utilizada para a estimativa de parâmetros, sendo que estes podem ser usados para prover uma projeção da evolução da doença.

Pode-se citar (LONG; KHALIQ; FURATI, 2021) como um dos primeiros trabalhos a utilizar PINNs com modelos compartimentais para dados epidemiológicos. Os autores inovaram ao, não apenas utilizar PINNs para solucionar este problema, mas também ao modelar as taxas de variação entre os compartimentos como uma função que tem que ser aproximada pela rede neural. A vantagem desta abordagem é considerar que medidas de contenção do avanço da doença como isolamento social e vacinação da população influenciam em taxas como infecção e mortalidade. Vale notar que no trabalho mencionado os autores focaram nas taxas de transmissão, por períodos curtos de tempo, e os modelos foram capazes de aproximar apenas funções monotônicas para os parâmetros.

Neste trabalho é proposto uma abordagem utilizando PINNs para a indentificação de parâmetros em modelos compartimentais, considerando que há uma variabilidade na taxa de transmissão ao longo tempo. O principal objetivo é averiguar se PINNs são capazes de aproximar as taxas de infecção com precisão, mesmo que elas sigam diferentes tendências ao longo de um período considerável de tempo. Para atingir este objetivo são feitos testes com dados sintéticos para avaliar se as PINNs conseguem aproximar a taxa de infecção seguindo diferentes tipos de função. Em seguida, os mesmos experimentos são repetidos, mas adicionando ruído aos dados sintéticos, para avaliar a resiliência do método a dados ruidosos. Por fim, são realizados testes com dados reais.

O restante do trabalho está organizado da seguinte forma:

No capítulo 2 são formalizados os conceitos de redes neurais e redes neurais informadas pela física, além de uma pequena revisão da literatura sobre algumas das diferentes arquiteturas alternativas já propostas.

No capítulo 3 são apresentados os modelos epidemiológicos compartimentais e são discutidas suas propriedades mais relevantes para a compreensão do trabalho desenvolvido nesta dissertação, como pontos de equilíbrio e identificabilidade, além disto são apresentadas algumas das variantes mais comuns. Também é feita uma pequena revisão da literatura

sobre o uso de PINNs para a estimativa de parâmetros em modelos compartimentais.

No capítulo 4 são providos os detalhes do método proposto por este trabalho, como a arquitetura da rede e como o modelo compartimental escolhido é integrado à rede. São definidos os experimentos para averiguar a efetividade do método, tanto os experimentos com dados sintéticos quanto os com dados reais, além dos métodos de avaliação dos resultados.

No capítulo 5 são apresentados os resultados dos experimentos propostos no capítulo 4 e são feitas pequenas avaliações dos mesmos.

Por fim, no capítulo 6 são apresentadas as conclusões, lacunas de pesquisa encontradas, e possibilidades de trabalhos futuros.

2 Redes Neurais Informadas pela Física

Neste capítulo é feita uma revisão dos principais conceitos relacionados a redes neurais e redes neurais informadas pela física para a compreensão do trabalho desenvolvido. Também é feita uma revisão das principais variantes das redes neurais informadas pela física.

2.1 Redes Neurais Feedforward

O primeiro conceito a ser compreender são as redes neurais *feedforward*, também conhecidas como *perceptrons* de multiplas camadas, do inglês, *multitayer perceptrons* (MLP). Redes *feedforward* podem ser entendidos como um modelo que simula o funcionamento de um cérebro, em que os neurônios formam um rede de conexões, em que o processamento se dá pela passagem de informação por essa rede considerando a topologia, ou seja, as conexões sinápticas entre os neurônios e força des mesmas. Sendo que a força pode ser de ativação (positiva), ou de inibição (negativa). Nesta analogia, um neurônio é entendido como uma unidade de processamento que ao receber estímulos de outros neurônios, processa estas entradas e produz uma saída. Esta estrutura de neurônio artificial foi proposta em (MCCULLOCH; PITTS, 1943) e está esquematizada na figura 1.

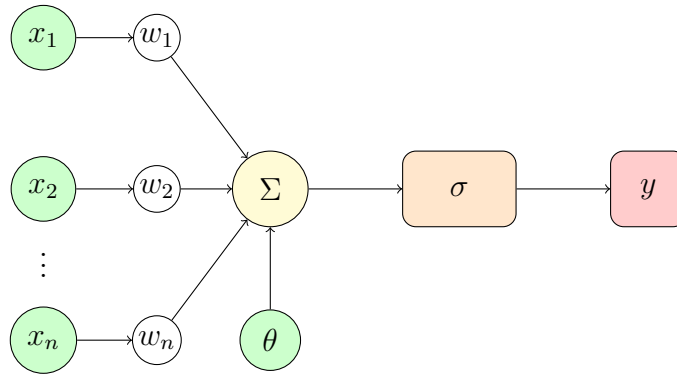


Figura 1 – Representação gráfica de um neurônio de MacCough-Pitts. Fonte: elaborada pelos autores.

Formalizando matematicamente a figura 1, sendo W um vetor de pesos $\in [-1, 1]$, \mathbf{x} um vetor de entradas de tamanho n e pertencente a R^n , θ um valor pertencente a R e σ uma função não linear. Um neurônio é uma função $f : R^n \rightarrow R$,

$$\sigma\left(\left(\sum_{i=1}^n W_i \mathbf{x}_i\right) + \theta\right) \quad (2.1)$$

O termo θ pode ser entendido como o último elemento de \mathbf{x} que está sendo sempre multiplicado pelo último elemento de W que sempre tem valor igual a 1, logo a equação 2.1 pode ser simplificada como,

$$\sigma(W\mathbf{x}) \quad (2.2)$$

Um neurônio pode ser entendido como uma transformação linear, a multiplicação das entradas pelos pesos e vieses somados, seguida por uma transformação não linear. Esta é feita pela função de ativação σ . Alguns exemplos de funções de ativação empregadas nas redes neurais são A função *sigmoid* 2.3, tangente hiperbólica (*tanh*) 2.4e *Rectified Linear Unit* (*ReLU*) 2.5.

$$\text{sigmoid}(x) = \frac{e^x}{1 + e^x} \quad (2.3)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.4)$$

$$\text{ReLU}(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \quad (2.5)$$

Um único neurônio não é capaz de aproximar qualquer função, para isso, eles são organizados em camadas que a saída de cada neurônio de uma camada forma parte da entrada de todos os neurônios da camada seguinte. A primeira camada apenas representa os dados de entrada, a última camada representa um vetor de saída de tamanho m . Logo, uma rede neural *feedforward* com L camadas escondidas pode ser entendida como uma função composta f formada por um conjunto de consecutivas transformações lineares e não lineares num vetor de dados de entrada de tamanho n $\mathbf{x} \in R^n$ que produz como saída um vetor de tamanho m $\hat{\mathbf{y}} \in R^m$, sendo as transformações intermediárias representadas pelas saídas y^l da camada escondida l , $l \in [1, \dots, L]$,

$$f(\mathbf{x}) = \sigma^{L+1} y^{L+1} \circ \sigma^L \circ y^L \dots \circ \sigma^2 \circ y^2 \circ \sigma^1 \circ y^1 \quad (2.6)$$

A figura 2 resume graficamente a definição presente na equação 2.6.

O vetor de saída $\hat{\mathbf{y}}$ é então comparado com um vetor desejado \mathbf{y} para calcular o erro entre a saída da rede e a saída desejada. Este é o papel da função de perda (*loss function*). Usualmente, para o caso de uma regressão, emprega-se a função de erro quadrático médio, do inglês, *mean root square* (*MSE*), como definida na equação 2.7.

$$\mathcal{L} = MSE = \frac{1}{N} \sum_{i=0}^N (y_i - \hat{y}_i)^2 \quad (2.7)$$

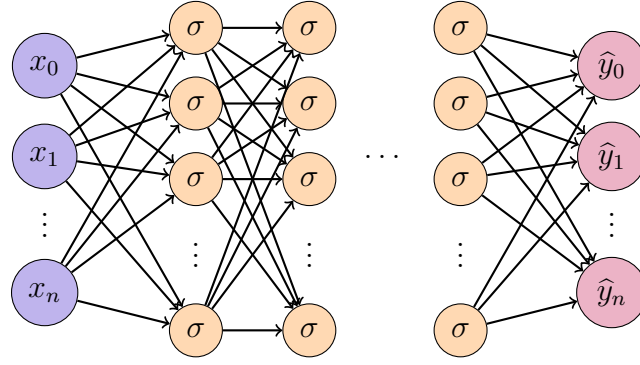


Figura 2 – Representação gráfica das redes *feedforward*. Fonte: elaborada pelos autores.

Para problemas de classificação de classificação, aplica-se a função de entropia cruzada (*cross-entropy*) para problemas de classificação binária, sendo p e q duas distribuições a função de perda é definida pela fórmula 2.8.

$$\mathcal{L} = \mathcal{H}(p, q) = - \sum_{x \in \mathcal{X}} p(x) \log q(x) \quad (2.8)$$

Já para o caso de classificação com múltiplos rótulos, aplica-se a função de *softmax*. Para um problema de classificação de múltiplos rótulos com N classes, a função de *softmax* é definida pela fórmula 2.10.

$$\rho_i = \frac{e^{z_i}}{\sum_{i=0}^N e^{z_j}} \quad (2.9)$$

$$\mathcal{L} = \text{argmax}(\rho_0, \rho_1, \dots, \rho_r) \quad (2.10)$$

A atualização dos pesos de cada camada se dá pela cálculo do gradiente de uma função de erro \mathcal{L} . O "tamanho do passo" que será dado é determinado por um parâmetro α chamado de taxa de aprendizagem. O conjunto de pesos W é então atualizado segundo a fórmula 2.11,

$$W_{t+1} = W_t + \alpha \nabla \mathcal{L} \quad (2.11)$$

A propagação dos erros se dá pelo algoritmo de retropropagação, sendo um caso de aplicação da regra da cadeia. A derivada da função de perda \mathcal{L} de uma rede com L camadas escondidas pode ser expressa então pela fórmula 2.12.

$$\frac{d\mathcal{L}}{d\sigma^L} \cdot \frac{d\sigma^L}{dy^L} \cdot \frac{dy^L}{d\sigma^{L-1}} \cdot \frac{d\sigma^{L-1}}{dy^{L-1}} \cdot \frac{dy^{L-2}}{d\sigma^{L-2}} \cdot \dots \cdot \frac{d\sigma^1}{dy^1} \cdot \frac{dy^1}{dx} \quad (2.12)$$

A atualização dos pesos se dar por meio da avaliação da expressão acima da esquerda para a direita, computando o gradiente δ em cada camada l e pesos W^l como descrito pela equação 2.13

$$\delta_l = (\sigma^l)' \circ (W^{l+1})(\sigma^{l+1})' \circ \dots \circ (W^{L-1})(\sigma^{L-1})' \circ (W^L)(\sigma^L)' \circ \nabla \mathcal{L} \quad (2.13)$$

A atualização dos pesos W da camada l utilizando os gradientes δ_l pode ser expresso pela equação 2.14.

$$\nabla_W \mathcal{L} = \delta_l(\sigma^{l-1}) \quad (2.14)$$

2.2 Redes Neurais Informadas pela Física

Apresentadas em (RAISSI; PERDIKARIS; KARNIADAKIS, 2019), PINNs podem ser entendidas como uma forma avançada de regularização, ou como um problema de otimização que transforma as condições de fronteira e iniciais em penalizações para a função custo.

2.2.1 PINNs como Solucionadores de Equações Diferenciais

Cabe aqui prover uma definição formal do que são equações diferenciais e como as PINNs resolvem este tipo de equação. PINNs são capazes de resolver problemas no seguinte formato:

$$\mathcal{D}(u(\mathbf{x}, t); \boldsymbol{\lambda}) = f(u, \mathbf{x}, t), \quad \mathbf{x} \in \Omega, t \in I, \quad (2.15)$$

$$\mathcal{B}(u(\mathbf{x}, t)) = g(\mathbf{x}, t), \quad \mathbf{x} \in \Gamma, t \in I, \quad (2.16)$$

$$\mathcal{I}(u(\mathbf{x}, t_0)) = q(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (2.17)$$

Em que $\Omega \subset \mathbb{R}^d$ é o domínio espacial limitado pela fronteira Γ ; d é a dimensão do domínio espacial; $T = [t_0, t_f]$ é o intervalo de tempo, sendo $t_0 < t_f$; $\mathbf{x} = (x_1, x_2, \dots, x_d)$ é um vetor de coordenadas espaciais; t denota o tempo; $u = u(\mathbf{x}, t)$ denota a solução desconhecida do problema; $\boldsymbol{\lambda}$ é um vetor de parâmetros das equações; \mathcal{D} é um operador diferenciais associado às equações; f é um termo fonte ou sorvedouro; \mathcal{B} and \mathcal{I} são operadores representando, respectivamente, as condições de fronteira e iniciais; por fim, g e q são funções conhecidas que definem essas condições.

A equação 2.18 define o termo da função de perda que engloba todos as equações que compõem o modelo. Trata-se de um treinamento não-supervisionado que busca minimizar o residual definido.

$$\mathcal{L}_{\text{física}}(\boldsymbol{\theta}) = \mathcal{L}_{\mathcal{D}}(\mathbf{x}, t, \boldsymbol{\theta}) + \mathcal{L}_{\mathcal{B}}(\mathbf{x}, t, \boldsymbol{\theta}) + \mathcal{L}_{\mathcal{I}}(\mathbf{x}, t_0, \boldsymbol{\theta}) \quad (2.18)$$

Sendo $\omega_{\text{domínio}}$, $\omega_{\text{fronteira}}$ e ω_{inicial} pesos atribuídos a cada um dos residuais. Caso haja dados disponíveis, é feito um treinamento supervisionado utilizando tais dados. A função de perda final da rede neural é então definida pela equação 2.19.

$$\mathcal{L}_{\text{total}}(\boldsymbol{\theta}) = \omega_{\text{física}} \mathcal{L}_{\text{física}}(\boldsymbol{\theta}) + \omega_{\text{dados}} \mathcal{L}_{\text{dados}}(\boldsymbol{\theta}) \quad (2.19)$$

Aqui vale mencionar que esta não é a única forma de distribuir os pesos da loss, a implementação da biblioteca *DeepXDE* (LU et al., 2021) permite atribuir pesos diferentes a cada condição inicial e de fronteira. Então o problema passa a ser encontrar os conjuntos $\boldsymbol{\theta}^*$ de parâmetros e vieses da rede que minimiza a função 2.19.

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \mathcal{L}_{\text{total}}(\boldsymbol{\theta}), \quad (2.20)$$

Existem muitos métodos de otimização para encontrar os argumentos da equação 2.20. Pode-se citar o método de primeira ordem *Adam* (KINGMA; BA, 2014) e o método quase-newtoniano *BFGS*, ou como comumente usado, a sua versão para ambientes de pouca memória, o *L-BFGS* (LIU; NOCEDAL, 1989).

A figura 3 mostra uma representação gráfica das PINNs.

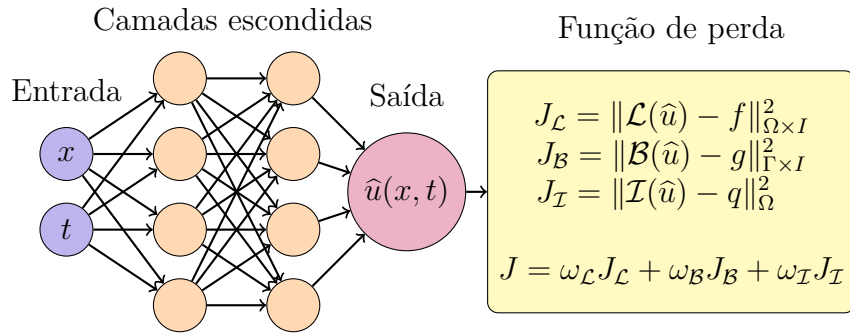


Figura 3 – Representação gráfica das PINNs. Fonte: elaborada pelos autores.

2.2.2 Formulação para Problemas Inversos

A formulação para as PINNs apresentada até então foca na resolução de problemas diretos. Em (RAISSI; PERDIKARIS; KARNIADAKIS, 2019) é apresentada uma extensão para a mesma que também soluciona problemas inversos. Solucionar um problema inverso consiste em, dado um conjunto de dados representados pela $\mathcal{L}_{\text{dados}}$, encontrar um conjunto de parâmetros $\boldsymbol{\lambda}$ que minimiza a $\mathcal{L}_{\text{total}}$. A equação 2.19 pode então ser reescrita como,

$$\mathcal{L}_{\text{total}}(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \omega_{\text{física}} \mathcal{L}_{\text{física}}(\boldsymbol{\theta}, \boldsymbol{\lambda}) + \omega_{\text{dados}} \mathcal{L}_{\text{dados}}(\boldsymbol{\theta}) \quad (2.21)$$

O problema de minimização descrito em 2.20 passa então a ser encontrar o conjunto θ^* de parâmetros da rede, como também o conjunto λ^* que minimiza a função 2.21,

$$(\theta^*, \lambda^*) = \arg \min_{\theta, \lambda} \mathcal{L}_{\text{total}}(\theta, \lambda), \quad (2.22)$$

Vale salientar que a rede ao encontrar os conjuntos θ^* e λ^* , a rede neural está solucionando tanto o problema direto (aproximar uma função que satisfaça as equações diferenciais), quanto o problema inverso (encontrar λ^*).

2.2.3 Pontos de Colocação

Pontos de colocação são um conceito importante para as PINNs, sendo análogos a criação da malha do *MEF*. Entretanto, diferente do *MEF*, não há um critério método bem definido para a obtenção destes pontos. Os pontos de colocação podem ser entendidos como uma discretização do domínio de uma equação diferencial. O único critério que deve ser atendido é o uso de pontos nas condições de fronteira e iniciais do problema. Por esta perspectiva, os pontos de colocação podem ser entendidos como uma amostra do domínio na qual a rede neural será treinada. Por este motivo, muitas técnicas de amostragem multidimensional são aplicadas às PINNs para definir os pontos de colocação, pode-se citar como exemplo o hipercubo latino.

2.3 Arquiteturas Alternativas

A definição de PINNs como uma rede neural com informação física não se aplica apenas a redes *feedforward*, qualquer arquitetura de redes neurais que incorpora as equações que descrevem os dados em que rede está sendo treinada, pode ser considerada uma rede neural informada pela física, uma PINN. Desde a publicação do artigo seminal em 2019, uma série de arquiteturas alternativas foram propostas utilizando a ideia das PINNs como princípio.

Uma classe de EDPs de especial interesse para áreas como teoria dos campos quânticos e mecânica estatística são as EDPs estocásticas. Esta classe de equações diferenciais pode ser entendida como uma equação diferencial com um termo ξ que representa um campo com ruído. O exemplo mais simples é a equação do calor estocástica 2.23.

$$\frac{\partial u}{\partial t} = \nabla \cdot u + \xi \quad (2.23)$$

Uma proposta de PINNs para EDPs estocásticas é apresentada em (ZHANG; GUO; KARNIADAKIS, 2020), em que são empregados métodos espectrais ortogonais para modelar processos que representam a estocasticidade nos dados. Estes processos são incorporados à função de perda como termos regularizadores extras. Outra arquitetura pensada para a

solução de EDPs estocásticas são as FINNs, apresentadas em (KARLBAUER et al., 2022), que combinam o método de volumes finitos com redes neurais para obter soluções para problemas de advecção-difusão como a equação de *Burgers* e de *Allen-Cahn*. A arquitetura obteve resultados superiores a redes neurais sem informação física e até mesmo de PINNs convencionais.

Um exemplo de proposta de PINNs alternativas que explora limitações das PINNs originais pode ser encontrada em (SUN; FENG, 2023). Os autores propõem as gPINNs (Gradient Enhanced PINNs) que tentam contornar as complexidades dos algoritmos de diferenciação automática existentes para equações de alta ordem. A arquitetura utiliza uma função de ativação adaptativa, uma técnica de melhoria nos gradientes para acelerar a convergência do modelo, e uma técnica chamada de *deep mixed residual method* para compensar o *overhead* gerado pela melhoria nos gradientes. A arquitetura é empregada na solução de equações de onda da luz. Outra proposta similar que também explora o uso de melhorias no gradientes pode ser encontrada em (MOHAMMADIAN; BAKER; FIORETTO, 2023), tendo como maior vantagem a necessidade de poucos dados para treinamento, acelerando o processo de treinamento. Os autores utilizam o método proposto para gerar estimativas para sistemas de controle de redes elétricas.

Técnicas como convolução e recorrência, comuns em áreas como visão computacional e processamento de linguagem natural, também encontram espaço na solução de problemas científicos. Em (REN et al., 2022) são propostas as PhyCRNet e PhyCRNet-s para resolução de EDPs sem dados de treinamento. A arquitetura emprega uma rede LSTM *encoder-decoder* com convolução para a extração de características espaciais e simulação da evolução do fenômeno no tempo. Para evitar o ajuste de hiper-parâmetros, as condições iniciais e de contorno não são impostas por meio de penalidades na função custo, mas como *hard-constraints*. Também são utilizadas conexões residuais e autoregressivas para simular a progressão do tempo. OS autores obtiveram resultados satisfatórios para problemas como a equação de *Burgers* e problemas de reação-difusão. Outro exemplo envolvendo camadas convolucionais pode ser encontrado em (SHI; ZENG; LIANG, 2024). No artigo é proposto uma arquitetura em que uma única camada convolucional é empregada. Nos testes realizados pelos autores, a rede neural proposta apresentou uma convergência melhor do que uma PINN convencional para problemas de difusão com diferentes frequências.

Pensando para lidar com problemas ligadas a solução de PDEs com dados que possuem componentes com frequências diferentes, a arquitetura proposta em (HAN; LEE, 2023) utiliza aprendizado hierárquico para lidar com esse problema. O princípio por trás do aprendizado hierárquico é utilizar camadas com redes neurais responsáveis por aproximar uma frequência, assim cada camada adicionada deve aprender a partir dos dados residuais da camada anterior. A saída dessas camadas é então alimentada em uma rede de características de Fourier embutidas. As primeiras camadas capturam os componentes de

menor frequência, enquanto que a rede de características embutidas fica responsável por impor limites a solução. A arquitetura proposta é testada com uma série de equações diferenciais parciais.

Outro exemplo de arquiteturas alternativas de redes neurais com informação física pode ser encontrada em (WU et al., 2022). Os autores propõem as *rnn-dct networks*, que combinam a arquitetura clássica de redes residuais (RNN) e DCT (*discrete cosine transformations*) com redes MLP (feedforward). O DCT é responsável por codificar a parte espacial da equação, enquanto que a RNN fica responsável por interpretar os valores temporais. O resultado dessas duas redes são então combinados e usados para alimentar a rede *feedforward*, ajudando esta a capturar as relações espaciais e temporais do problema. A arquitetura proposta apresentou resultados excelentes em problemas envolvendo as equações de Navier-Stokes, como o problema de vortices de Taylor-Green.

Em (PAKRAVAN et al., 2021) é apresentado um método focado na solução de problemas inversos envolvendo EDPs com lacunas nos dados chamado *Blended inverse-PDE networks* BiPDE. Os parâmetros da equação são descobertos utilizando não apenas conhecimentos sobre a física incorporados a função de perda, mas também utilizando conhecimentos específicos do domínio. A rede neural exerce o papel de aproximador universal dos parâmetros, que combinado com um solucionador de EDPs, livra a rede de ter que resolver o problema de forma direta, focando apenas no problema inverso, tornando o método mais eficiente, ao não necessitar de tantos dados para o treinamento. O método proposto é testado em problemas inversos envolvendo encontrar coeficientes de difusão variantes para as equações de *Burgers* e de *Poisson*, além de problemas de alta dimensionalidade.

Propostas em (JAGTAP; KHARAZMI; KARNIADAKIS, 2020), as conservative PINNs (cPINNs) utilizam leis de conservação não-lineares para a solução de problemas em domínios discretos. Para aplicar as cPINNs, o domínio é fragmentado em subdomínios nos quais as cPINNs são empregadas, a propriedade de conservação é garantida pela aplicação de continuidade ao longo de todos os subdomínios. Nas fronteiras entre os subdomínios, a solução é obtida pela média das saídas de cada cPINN. Uma vantagem desta arquitetura é permitir que em cada subdomínio sejam utilizados diferentes meta-parâmetros, como número de camadas escondidas, funções de ativação e otimizadores. A arquitetura é capaz de tirar grande proveito de processamento paralelo.

Utilizando as cPINNs como base, as *Extended Physics-informed Neural Networks* (XPINNs) são propostas em (JAGTAP; KARNIADAKIS, 2020). Assim como as cPINNs, o domínio é fragmentado em subdomínios permitindo que a otimização dos meta-parâmetros possa ser feita em cada rede. A grande novidade da arquitetura é a possibilidade de aproximar qualquer EDP. Outra novidade é a decomposição arbitrária do domínio, o que permite tirar ainda mais proveito de paralelismo durante o treinamento do modelo. A

arquitetura foi aplicada em PDE não-lineares com geometrias complexas, aproveitando a decomposição de domínio. Continuando o desenvolvimento da arquitetura, em (HU et al., 2023) são propostas as *Augmented PINNs* (APINNs), a novidade da arquitetura é tornar a fragmentação dos subdomínios treinável. A arquitetura permite que dados de treinamento sejam usados entre diferentes fragmentos, aumentando a generalização do modelo. A performance das APINNs é consideravelmente superior a das XPINNs em termos de fragmentação do domínio e generalização.

Aproveitando as possibilidades que o uso de paralelismo pode trazer em termos de velocidade no tempo de treinamento, em (DWIVEDI; PARASHAR; SRINIVASAN, 2019) são propostas as *Distributed PINNs* (DPINNs). A principal ideia por trás deste método é fragmentar o domínio do problema em partes que devem ser aproximadas por uma rede neural independente. Uma vantagem dessa abordagem é que é possível utilizar redes mais simples em cada fragmento, ao invés de aplicar uma única rede mais complexa para todo o domínio. A função de perda é calculada com base na aproximação de cada rede neural que funcionam como PINNs paralelas. O método é capaz de aproximar corretamente problemas difíceis como as equações de Navier-Stokes de forma satisfatória utilizando redes simples, obtendo resultados melhores do que aqueles com uma rede neural informada pela física centralizada.

Problemas de engenharia costumam ter parâmetros incertos que são dependentes de variáveis temporais e espaciais, estas incertezas são contornadas utilizando processos que as modelam. Métodos como o MEF são capazes de resolver tais equações, mas é necessário ter um conhecimento *a priori* sobre as distribuições que modelam o fenômeno estudado, fazendo com que métodos puramente estatísticos sejam insuficientes. É comum empregar técnicas como teoria de conjuntos difusos junto aos métodos tradicionais de solução de equações diferenciais. Aproveitando o potencial que as PINNs e conjuntos difusos tem para a soluções de equações com tais características, em (FUHG et al., 2022) são propostos dois novos modelos as *Interval Physics-Informed Neural Networks* (iPINNs) e *Fuzzy Physics-Informed Neural Networks* (fPINNs) que combinadas, são capazes de resolver problemas envolvendo equações diferenciais e conjuntos difusos sem conhecimentos prévios. Uma rede neural é responsável pela a aproximação das entradas, enquanto a outra fica responsável por aproximar a solução da PDE. O método proposto mantém as vantagens das PINNs, como não precisar de uma malha de pontos, e ser capaz de resolver problemas de forma direta e inversa sem alterações significativas nos métodos.

As equações de ordem fracionaria, também chamadas de equações diferenciais extraordinárias, são uma classe de equações diferenciais que encontram aplicações na modelagem de diversos fenômenos físicos como escoamentos subterrâneos (ATANGANA; BILDIK, 2013) e equações de conservação de massa (WHEATCRAFT; MEERSCHAERT,

2008). A equação 2.24 mostra a forma geral para tais equações.

$$\frac{d^\alpha u}{dx^\alpha} = u, \quad \alpha \in \mathbb{R} \quad (2.24)$$

Em (PANG; LU; KARNIADAKIS, 2019) é proposto uma arquitetura de redes neurais informadas pela física pensada para equações diferenciais de ordem fracionária. O principal entrave para o uso de PINNs convencionais na solução dessas equações está ligado ao fato de que algoritmos de diferenciação automática são otimizados para equações de ordem inteira. A arquitetura proposta, chamado de *fPINNs*, contorna tais limitações ao aplicar uma abordagem mista que utiliza diferenciação automática para a parte inteira das derivadas e derivadas numéricas para a parte fracionária. A arquitetura proposta foi capaz de solucionar problemas clássicos como a equação de *Burgers* utilizando derivadas parciais de ordem fracionária, obtendo resultados equiparáveis ao MEF.

Baseado nas *KANs* (*Kolmogorov-Arnold Neural Networks*) (LIU et al., 2025), redes neurais baseadas no teorema de representação universal de Kolmogorov-Arnold, foram propostas as *PIKANs* (*Physically Informed Kolmogorov-Arnold Neural Networks*) que combinando as funções de ativação baseadas em *b-splines* das *KANs* com a técnica de regressão simbólica, permite encontrar representações analíticas das soluções encontradas pela rede neural. As *KANs* não se mostraram superiores aos MLPs para todo tipo de tarefa (TOSCANO et al., 2025), mas sua superioridade com regressão simbólica demonstra o potencial da arquitetura para problemas envolvendo explicabilidade.

3 Modelos Compartimentais

Neste capítulo são apresentados os principais conceitos acerca dos modelos compartimentais para a compreensão do trabalho desenvolvido nesta dissertação. No fim do capítulo é apresentada uma pequena revisão de aplicação de PINNs com modelos compartimentais.

3.1 Origens

Apresentado no conjunto de trabalhos seminais (KERMACK; MCKENDRICK, 1927), (KERMACK; MCKENDRICK, 1932) e (KERMACK; MCKENDRICK, 1933) e consolidados em estudos como (KENDALL, 2023), modelos Compartimentais nada mais são do que modelos epidemiológicos utilizados no estudo de doenças contagiosas que separam a população em grupos. Os fluxos de indivíduos entre esses grupos, chamados de compartimentos, são modelados por meio de operações de diferenciação.

3.2 O modelo SIR

Apresentado no trabalho seminal de (KERMACK; MCKENDRICK, 1927), o modelo *SIR* (Susceptible-Infected-Recovered) é definido pelo conjunto de equações 3.1, 3.2 e 3.3.

$$\frac{dS(t)}{dt} = \frac{-\beta S(t)}{N} I(t), \quad t > t_0, \quad (3.1)$$

$$\frac{dI(t)}{dt} = \frac{\beta S(t)}{N} I(t) - \gamma I(t), \quad t > t_0, \quad (3.2)$$

$$\frac{dR(t)}{dt} = \gamma I(t), \quad t > t_0, \quad (3.3)$$

Estas três equações formam um sistema de equações diferenciais de fácil interpretação. A primeira equação modela a interação entre pessoas infectadas e pessoas suscetíveis, o parâmetro β , a taxa de infecção, diz quantos destes encontros resultaram em novos casos da doença. A terceira equação modela a recuperação de indivíduos ao longo do tempo, o parâmetro γ é taxa de indivíduos que se recuperam por unidade de tempo, sendo que a recuperação pode ser a cura da doença ou o falecimento do indivíduo, já que o *SIR* não distingue os dois casos. A segunda equação é a soma da primeira e da terceira equação, mas com seus sinais trocados. Ela modela o fluxo de indivíduos que entram e saem do compartimento de infectados.

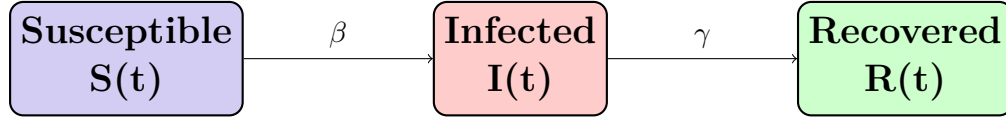


Figura 4 – Grafo para o *SIR*. Fonte: elaborada pelos autores.

Por se tratar de um sistema de equações diferenciais ordinárias com três equações, são necessárias três condições iniciais para se obter um problema de valor inicial.

$$S(0) = S_0 \quad (3.4)$$

$$I(0) = I_0 \quad (3.5)$$

$$R(0) = R_0 \quad (3.6)$$

Como o modelo não inclui mortes naturais ou por outras causas que não a doença que está sendo modelada, nem o nascimento de pessoas na população estudada, assume-se que a soma dos três compartimentos para qualquer tempo t é igual ao total N da população,

$$S(t) + I(t) + R(t) = N, \quad t > t_0, \quad (3.7)$$

O modelo pode ser entendido com um grafo em que os indivíduos fluem de um compartimento para outro a uma taxa β e γ . A figura 4 representa o *SIR* por esta perspectiva.

Certas definições são necessárias neste ponto para se compreender os modelos compartimentais. Em modelos epidemiológicos, incidência é o fluxo de novos casos por unidade de tempo, enquanto que a prevalência é a quantidade de casos na população num instante t . No caso do *SIR*, estes conceitos são representados respectivamente pelo termo $\frac{-\beta S(t)}{N}$ e pela equação $I(t)$.

O número de reprodução básico R_0 é definido como a relação entre os parâmetros β e γ . Ele é de particular interesse para o estudo da propagação de uma doença, pois se $R_0 > 1$, significa que a dispersão da doença ainda está em curso e o número de infectados tende a aumentar. Se $R_0 < 1$, significa que a dispersão já atingiu seu pico e o número de indivíduos infectados tende a diminuir. Se $R_0 = 1$ a doença está em equilíbrio e o número de infectados tende a se manter o mesmo.

$$R_0 = \frac{\beta}{\gamma} \quad (3.8)$$

Um valor relacionado ao número de reprodução básico R_0 , é o número de reprodução efetivo R_e . Enquanto que o número de reprodução básico R_0 é utilizado para medir o

dispersão de uma doença no início de uma pandemia. Seu valor é obtido pela multiplicação de R_0 por S num instante t .

$$R_e = R_0 S(t) \quad (3.9)$$

O limite \mathcal{T} define o valor máximo de indivíduos que estarão contaminados no pico da pandemia. Este limite é importante para tomadores de decisão e governos para entender os efeitos sobre os serviços de saúde pública.

$$\mathcal{T} = \frac{\gamma}{\beta} N = \frac{N}{R_0} \quad (3.10)$$

A figura 5 mostra um exemplo do SIR com $\beta = 0.8$ e $\gamma = 0.1$ e as três curvas características desse modelo.

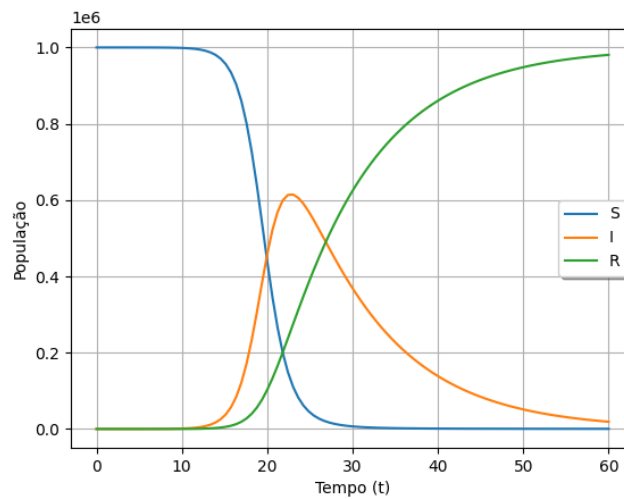


Figura 5 – Exemplo do SIR com $\beta = 0.8$ e $\gamma = 0.1$.

3.3 Pontos de equilíbrio

Os modelos compartimentais possuem dois pontos de equilíbrio chamados de ponto de equilíbrio livre de doenças (*Disease Free Equilibrium*) e ponto de equilíbrio endêmico (*Endemic Equilibrium*). O primeiro ponto é atingido quando não há mais indivíduos infectados na população, ou seja, no ponto em que I é igual a zero. O ponto de equilíbrio endêmico é atingido justamente quando o número de novos casos é compensado pelo número de indivíduos que se recuperam, mantendo o número de indivíduos no compartimento *Infected* constante, mas diferente de zero.

3.4 Pontos críticos

Utilizando o foi desenvolvido na seção anterior, pode-se facilmente encontrar os pontos críticos do modelo SIR. Considerando que os pontos de equilíbrio do modelo SIR são determinados pela solução do sistema composto pelas equações 3.11, 3.12 e 3.13.

$$-\beta SI = 0 \quad (3.11)$$

$$\beta SI - \gamma I = 0 \quad (3.12)$$

$$\gamma I = 0 \quad (3.13)$$

A última equação $\gamma I = 0$ implica que $I = 0$. Substituindo este valor na primeira e segunda equações, as equações são satisfeitas para qualquer valor de S . Portanto, o sistema tem um conjunto de pontos de equilíbrio da forma:

$$(S^*, I^*, R^*) = (S^*, 0, R^*) \quad (3.14)$$

Sendo $S^* + R^* = N$, ou seja, qualquer ponto em que não haja infectados e a soma dos suscetíveis e recuperados seja igual a N representa um estado de equilíbrio. Isso significa que o sistema pode estabilizar-se em diferentes estados dependendo da quantidade de indivíduos que permaneceram suscetíveis ou que foram recuperados. Esse é o estado de equilíbrio livre de doença, onde toda população está no compartimento *Susceptible* ou *Recovered* e não há mais infectados.

A análise de estabilidade dos pontos de equilíbrio do modelo SIR pode ser facilmente feita através da matriz jacobiana do sistema de equações diferenciais. Considere a matriz 3.15.

$$J(S, I, R) = \begin{pmatrix} \frac{\partial f_1}{\partial S} & \frac{\partial f_1}{\partial I} & \frac{\partial f_1}{\partial R} \\ \frac{\partial f_2}{\partial S} & \frac{\partial f_2}{\partial I} & \frac{\partial f_2}{\partial R} \\ \frac{\partial f_3}{\partial S} & \frac{\partial f_3}{\partial I} & \frac{\partial f_3}{\partial R} \end{pmatrix} = \begin{pmatrix} -\beta I & -\beta S & 0 \\ \beta I & \beta S - \gamma & 0 \\ 0 & \gamma & 0 \end{pmatrix} \quad (3.15)$$

Utilizando a equação 3.14 e aplicando-a na matriz 3.15, chega-se ao ponto de equilíbrio 3.16.

$$J(S^*, 0, R^*) = \begin{pmatrix} 0 & -\beta S & 0 \\ 0 & \beta S - \gamma & 0 \\ 0 & \gamma & 0 \end{pmatrix} \quad (3.16)$$

Os autovalores dessa matriz são $\gamma_1 = 0$, $\gamma_2 = 0$, $\gamma_3 = \beta S^* - \gamma$. Isso significa que para $S^* < \gamma/\beta$ tem-se $\gamma_3 < 0$ e a solução de equilíbrio é estável.

3.5 Matriz de nova geração

Uma forma de se calcular o número de reprodução básico R_0 é a aplicar a técnica de matriz de nova geração. Proposta por (DIEKMANN; HEESTERBEEK; METZ, 1990), o método consiste em

- Divida toda a população em n compartimentos de tamanho m tal que $m < n$.

Aplicando o método descrito acima para o modelo SIR obtem-se as seguintes definições:

$$\frac{dI}{dt} = \underbrace{\beta SI}_{\mathcal{F}} - \underbrace{\gamma I}_{\mathcal{V}} \quad (3.17)$$

$$F = \frac{\partial \mathcal{F}}{\partial I} = \beta S \Big|_{DFE} = \beta N, \quad V = \frac{\partial \mathcal{V}}{\partial I} = \gamma \quad (3.18)$$

$$FV^{-1} = (\beta N) \left(\frac{1}{\gamma} \right) = \frac{\beta N}{\gamma} \quad (3.19)$$

$$R_0 = \rho(FV^{-1}) = \frac{\beta N}{\gamma} \quad (3.20)$$

3.6 Identificabilidade de um Modelo

Uma questão pertinente a modelos baseados em equações diferenciais é a identificabilidade. A identificabilidade está relacionada a capacidade de estabelecer os parâmetros de um modelo de forma única com base nos dados disponíveis e independente das condições iniciais. Há dois tipos de identificabilidade:

- **Identificabilidade estrutural:** Refere-se a capacidade de estabelecer os parâmetros de um modelo utilizando dados ideais. Isto é, desprovidos de ruídos e lacunas e dispersos de forma uniforme pelo domínio do problema.
- **Identificabilidade prática:** Refere-se a capacidade de estabelecer os parâmetros de um modelo utilizando dados reais que possuem ruído, lacunas (não foram coletados em todos os dias da semana, por exemplo) ou não estão uniformemente distribuídos pelo domínio do problema (há mais dados referentes a uma semana epidemiológica do outra).

A análise dos dois tipos de identificabilidade é importante para o estudo de modelos compartimentais. No primeiro caso para o estudo teórico de modelos compartimentais,

mais especificamente, para demonstrar teoricamente ser possível estimar parâmetros como β , γ e α de novas propostas de modelos compartimentais utilizando métodos analíticos como álgebra diferencial e expansões por séries de Taylor. No segundo caso para lidar como dados epidemiológicos reais e ter garantias que os parâmetros estimados para o modelo têm algum grau de confiabilidade. As técnicas mais comuns para se obter os parâmetros a partir de um conjunto de dados são os métodos não-paramétricos como *bootstrapping* e simulações de Monte Carlo, além de outras técnicas estatísticas como matrizes de informação de *Fisher*.

Formalmente, um problema de identificabilidade pode ser expresso pelas equações 3.21 e 3.22. Sendo \mathbf{p} um conjunto de parâmetros, o problema de identificabilidade consiste em determinar se $\mathbf{p} \rightarrow \mathbf{y}$ é injetivo.

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}, \mathbf{p}, t) \quad \text{Equações de estado} \quad (3.21)$$

$$\mathbf{y} = h(\mathbf{x}, \mathbf{p}, t) \quad \text{Equações de observação} \quad (3.22)$$

Um dos fatores mais importantes para determinar a identificabilidade de um modelo compartimental são as condições iniciais. O que cria um problema para a aplicação desses modelos sob qualquer situação, pois geralmente, não se sabe as condições iniciais reais.

3.7 Análise de Sensibilidade

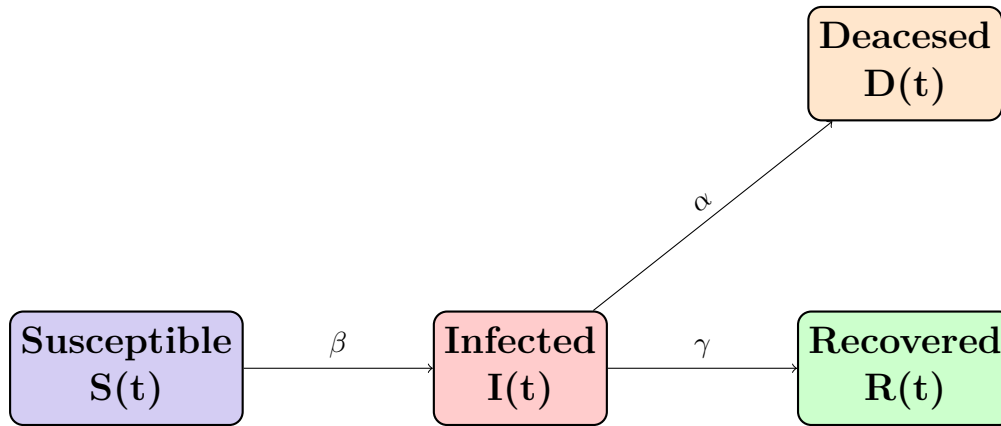
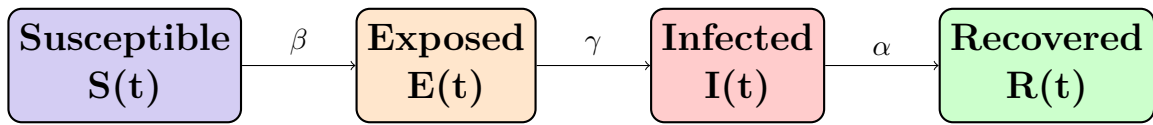
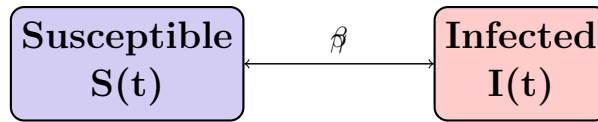
Partial Rank Correlation Coefficients (PRCC)

3.8 Problemas Inversos

Por se tratarem de problemas de valor inicial, os problemas envolvendo modelos compartimentais podem ser facilmente resolvidos por meio de métodos numéricos tradicionais, porém quando se deseja resolver um problema inverso envolvendo modelos compartimentais, ou seja, descobrir parâmetros como β e γ a partir de um conjunto de dados, deve se considerar que problemas inversos são mal-postos. Cabe aqui prover uma formalização do que são problemas bem postos e mal postos.

Um problema é considerado bem-posto se apresentar as seguintes 3 condições:

- **Existência:** Existe ao menos uma solução.
- **Unica:** Há no máximo uma solução.
- **Estável:** A solução depende continuamente nos dados.

Figura 6 – Grafo para o *SIRD*. Fonte: elaborada pelos autores.Figura 7 – Grafo para o *SIRD*. Fonte: elaborada pelos autores.Figura 8 – Grafo para o *SIS*. Fonte: elaborada pelos autores.

Enquanto que problemas diretos são considerados bem postos, problemas inversos são mal-postos. O motivo é que problemas inversos não satisfazem o segundo critério, há mais de uma solução possível para um problema inverso.

3.9 Adicionando Compartimentos

Baseados no modelo *SIR*, foram propostos outros modelos com mais compartimentos, como o *SIRD* (Susceptible-Infected-Recovered-Deceased) (GILES, 1977) que separa o compartimento de *Recovered* em dois, um para os indivíduos que se recuperaram (*R*) e outro para os que vieram a falecer devido a doença (*D*).

Outro exemplo é o *SEIR* que inclui um compartimento para indivíduos que foram expostos a doença, mas ainda não manifestaram sintomas, ele é aplicável a doenças que possuem um tempo de latência entre o contanto com patógeno e o surgimento dos sintomas e capacidade do indivíduo de transmitir a doença.

Compartimentos podem também ser removidos, como é o caso do modelo *SIS* apresentado no mesmo conjunto de trabalhos de Kermack e McKendrick (KERMACK; MCKENDRICK, 1932), é um modelo com aplicação para doenças que não geram imunidade, como gripe e resfriado. Obviamente, é possível estender esse modelo, como é o caso do *SEIS*, sendo o *SIS* com um compartimento para indivíduos expostos.

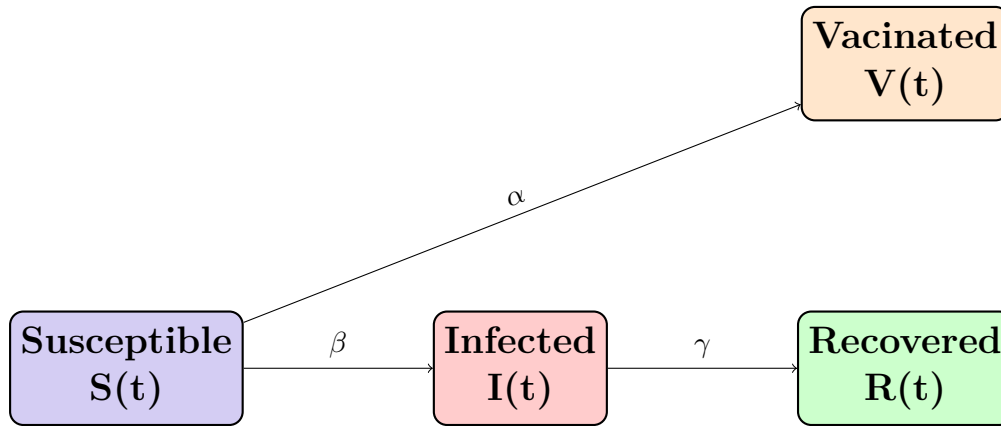


Figura 9 – Grafo para o *SIRD*. Fonte: elaborada pelos autores.

Compartimentos podem ser usados para modelar fenômenos específicos como a vacinação. É o caso de modelos como o *SIRV* (SCHLICKESER; KRÖGER, 2021) e o *SIRVD*. Indivíduos são vacinados a uma taxa α e adquirem imunidade à doença. Um parâmetro ϕ pode ser adicionado para dar conta de casos em que a vacinação não é eficaz.

Exemplos de modelos mais recentes, como o *SVIHRD*, podem ser encontrados em trabalhos como (NELSON et al., 2024). que inclui compartimentos para hospitalizados e vacinados e foi aplicado para modelar a pandemia de COVID-19. Outro exemplo é o *SEIHDR*.

3.10 Variantes do Modelo SIR

Ao longo dos anos foram propostas diferentes variantes do modelo *SIR* original, esta seção se dedica a abordar alguns exemplos brevemente. Em (NOBLE, 1974) é proposto uma variante do modelo que Em (SINGH; GUPTA, 2022) é proposto

3.11 Aplicação de PINNs com Modelos Compartimentais

Com a pandemia de COVID-19 no fim de 2019, renovou-se o interesse em modelos epidemiológicos compartimentais. As PINNs, que haviam acabado de ser propostas, foram encaradas como uma ferramenta nova que poderia ser utilizada junto com a grande quantidade de dados que estavam sendo gerados. Nesta seção são apresentados alguns trabalhos que aplicaram PINNs para solucionar problemas em epidemiologia. São destacados novos modelos compartimentais, modificações feitas nos modelos clássicos, e aplicações inovadoras das PINNs.

Em (OUYOUSSEF et al., 2024) é apresentada uma variação do *SIR* em os compartimentos de suscetíveis e infectados são divididos em dois subgrupos. Há também um subgrupo de indivíduos vacinados V , que assume constante ao longo do tempo do tempo,

sendo que apenas uma porcentagem ξ está totalmente imunizada. Os autores não estavam interessados na taxa de recuperação, logo ela é assumida como constante e a terceira equação do sistema *SIR* é removida. A PINN passa a ter que aproximar as quatro curvas de interesse do modelo (S_1, S_2, I_1, I_2) , além de descobrir, as quatro taxas $(\beta_{11}, \beta_{12}, \beta_{21}$ e $\beta_{22})$ de infecção. A aplicação destes subgrupos varia conforme a modelagem, pode se dividir a população por grupos de risco, por exemplo. O modelo passaria a capturar as iterações entre estes subgrupos. São feitos testes apenas com dados sintéticos.

Outro exemplo utilizando subgrupos dentro dos compartimentos pode ser encontrado em (ARULANDU; SESHAIYER, 2023). Mas os autores vão além, ao utilizar o modelo de meta-populações de (JACQUEZ et al., 1988) que divide a população em n subcompartimentos, transformando o parâmetro β numa matriz $n \times n$. Através de simplificações algébricas, é proposto uma variação do *SIR* com $3n$ equações e $3n + 3$ parâmetros. Os autores então empregam este modelo junto com uma PINN para estimar o número de reprodução efetivo R_i para cada população i , sendo $i \in [1, \dots, n]$. Este valor é então usado para ajudar na criação de um plano de vacinação que priorize populações mais vulneráveis, ou seja, aquelas em que R_i é maior. São realizados testes com a famosa base de dados sobre a dispersão de influenza numa escola londrina.

Modelos compartmentais utilizam sistemas de ODEs para modelar a evolução de uma pandemia, considerando apenas o tempo como variável independente e ignorando a dimensão espacial. Em (BERTAGLIA et al., 2022) é feita uma modificação do modelo *SIR*, criando o *SIR* hiperbólico. A modificação consiste em inserir uma dimensão espacial e transformar o PVI num problema de reação-difusão, numa EDP hiperbólica em função da dimensão espacial x e temporal t . A função de perda da rede é modificada, mas tomando cuidado para garantir que a rede satisfaça as propriedades de convergência assintótica. São realizados testes envolvendo problemas diretos e inversos para averiguar a efetividade do modelo.

O uso de PINNs com modelos com mais compartimentos é explorado em (NELSON et al., 2024). Os autores propõem o já mencionado *SVIHRD*, um modelo que inclui compartimentos para vacinados (V), hospitalizados (H), e separa o compartimento de recuperados em entre os que se curaram e as fatalidades (D). O compartimento de hospitalizados é usado para medir a ocupação dos serviços de saúde pública. O modelo ainda inclui uma taxa de nascimentos para os suscetíveis e uma taxa de mortes naturais para todos os outros compartimentos. É demonstrada a estabilidade do modelo e são feitos testes com dados da pandemia de COVID-19 no Japão.

Em (HAN et al., 2024) há um outro exemplo de aplicação de PINNs com modelos comportamentais e parâmetros que variam no tempo. Os autores aplicaram o modelo *SAIRD* para gerar dados sintéticos para o treinamento de uma PINN. Uma vez validado o modelo, ele é utilizado para se ajustar aos dados de COVID-19 da Alemanha. É feito um

estudo sobre o parâmetro ω_{dados} para diferentes valores e seus efeitos na convergência da rede e valores obtidos para os parâmetros do problema inverso.

Um dos primeiros usos de PINNs com parâmetros variando no tempo é proposto em (LONG; KHALIQ; FURATI, 2021). Os autores aplicam uma PINN com o modelo *SIRD* para estimar a taxa de transmissão da COVID-19 em três cidades americanas ao longo de março de 2020 a outubro de 2020. Uma vez calculada a estimativa da taxa de transmissão ao longo do tempo, o número de reprodução básico é calculado para o mesmo período. Os parâmetros estimados são alimentados numa *Long Short Term Memory* (LSTM) para prever o número de infectados no mês seguinte, outubro.

Em (SHAMSARA; KÖNIG; PFEIFER, 2025) é apresentado uma aplicação de PINNs junto ao modelo *SVIHRD* com parâmetros variando no tempo. Os autores introduzem uma técnica para ajustar o peso ω_{dados} ao longo do treinamento da rede. O modelo é então aplicado para ajustar a dados de COVID-19 dos países Itália, França e Alemanha. Uma vez estimada a taxa de transmissão da doença, ela é correlacionada com O surgimento de variantes do vírus que causa a COVID-19.

Outra extensão feita nos modelos clássicos é proposta em (NGUYEN; RAISSI; SESHAIYER, 2022). Os autores modificam o *SEIR* ao adicionar um compartimento *P* para modelar a concentração de patógenos em ambientes fechados, e a possibilidade de contaminação pela doença através da interação não apenas com o contato com indivíduos contaminados, mas também com a interação com patógenos presentes no ambiente. Os autores deduzem os pontos de equilíbrio do modelo e realizam testes com PINNs nos dados de influenza de um colégio privado londrino.

Um exemplo com modelos de ordem fracionária é apresentado em (LI et al., 2025). Os autores utilizam uma versão do *SEIHDR* (*Susceptible-Exposed-Invected-Hospitalized-Deceased-Recovered*) com derivadas fracionárias de Caputo e parâmetros variando no tempo. São feitas análises dos pontos de estabilidade do modelo, e o número de reprodução básico é calculado utilizando matrizes de próxima geração. As PINNs são usadas para ajustar o modelo aos dados de COVID-19 do Canadá.

Em (HELDMANN et al., 2023) é aplicado o modelo *SVIHRD* com uma função de custo multi-objetivo ao tratar a $\mathcal{L}_{\text{física}}$ e a $\mathcal{L}_{\text{dados}}$ como duas funções objetivos distintas. A função de perda da rede é composta por uma aproximação da curva Pareto entre as duas funções objetivos. Outra inovação do artigo, é estimar alguns parâmetros utilizando o método de diferenças finitas alternativo. As PINNs são aplicadas para os dados de COVID-19 da Alemanha. São feitos ajustes aos dados para estimar os parâmetros de transmissão e previsões de curto tempo.

A combinação de modelos compartimentais com PINNs podem ser utilizados para outros fins que não apenas estimar a taxa de transmissão da doença. Em (GHOSH et

al., 2023) é apresentado uma variação do *SEIR* com níveis de imunidade dependentes do tempo e do número de doses de vacinas ministradas, o modelo também considera a eficácia da vacina. Estas adições ao *SEIR* são modeladas na forma de parâmetros extras que podem mover indivíduos do compartimento de recuperados para o compartimento de suscetíveis. O modelo é aplicado junto a PINNs para estimar o nível de subnotificação de infectados e mortes não atribuídas a COVID-19 na China após o fim da política de COVID zero.

Outro exemplo que combina PINNs com métodos já consolidados na literatura pode ser encontrado em (MILLEVOI; PASETTO; FERRONATO, 2024). Os autores propõem uma chamada "abordagem separada" (*split approach*), a ideia principal é separar o treinamento da rede em duas fases. Na primeira fase, o compartimento de infectados é aproximado utilizando uma rede neural que se ajusta aos dados. Na segunda fase, a saída desta rede é então utilizada junto a uma PINN para para estimar o número de suscetíveis e a taxa de transmissão β do modelo *SIR*. A vantagem desta abordagem em comparação com a abordagem tradicional, ou "conjunta" (*joint approach*) é simplificar a função de perda da rede neural ao simplificar as equações do modelo *SIR* que são usadas na função de perda, fazendo com que a PINN tenha uma curva a menos para se ajustar, melhorando a convergência da mesma. Os autores testam a abordagem proposta com dados sintéticos e para dados de COVID-19 da Itália, além de testes com lacunas nos dados para testar a resiliência da proposta.

Em (OGUEDA-OLIVA et al., 2023) é feita uma modificação no modelo *SIRD* para modelar o deslocamento de pessoas duas cidades colombianas de Bogotá e Medellín. A incorporação deste movimento é feita através de uma matriz de transporte e o resultado é um modelo *SIRD* com duas populações distintas e, por consequência, dois compartimentos para cada compartimento original ($S_1, S_2, I_1, I_2, \dots$), assim como parâmetros duplicados ($\beta_1, \beta_2, \gamma_1, \gamma_2, \dots$). O modelo é utilizado para se ajustar aos dados de COVID-19 da Colômbia de 2021.

Em (NING et al., 2023) são usadas PINNs paralelas para aproximar cada compartimento do modelo *SEIRD* com parâmetros variando no tempo. Uma PINN paralela usa uma rede *feedforward* independente para aproximar cada saída da rede, ou para um conjunto de saídas. No artigo, os autores aproximam cada compartimento com uma rede independente, além de usar *skip-connections* para prevenir explosão e desaparecimento de gradientes. As funções para os parâmetros β , γ e μ são extraídos dos dados de COVID-19 da Itália em 2020 e são feitas previsões de curto prazo. Os resultados obtidos com as PINNs são comparados com valores obtidos por métodos Bayesianos.

Outro exemplo de uso das PINNs paralelas é apresentado em (YANG; WANG; LOU, 2025). Os autores propõem um modelo com compartimentos para vacinação e reinfeção, sendo que recuperados são movidos para um compartimento S_R e, caso infectados, movidos

para um compartimento I_2 . A intenção dos autores era demonstrar que a mortalidade entre reinfectados é maior do que entre os infectados. Outra novidade, e a proposição das PINNs discretas, as *dtPINNs*. A ideia por de trás destas PINNs é utilizar uma rede neural para aproximar cada um dos parâmetro que variam no tempo, um método numérico, como o Runge-Kutta de quarta ordem, é empregado para resolver o sistema de equações com os parâmetros que variam no tempo e com os parâmetros estáticos. A função de perda é calculada com base no *MSE* entre a solução obtida pelo método numérico e os dados. Os autores utilizam o método desenvolvido para estimar os parâmetros de COVID-19 do estado da Carolina do Norte, nos Estados Unidos.

Em (SHAIR; RAISSI; SESHAIYER, 2022), é feito um estudo dos efeitos de ruído nos dados e dados faltantes na qualidade das soluções obtidas com PINNs para o modelo *SIRD*. Também são realizados estudos dos efeitos de metaparâmetros como taxa de aprendizagem, número de pontos de colocação, largura das camadas escondidas e número de camadas escondidas. PINNs são aplicadas juntas a vários modelos compartimentais para uma série de doenças como COVID-19, ebola, rubéola, pneumônia, polio, dengue e zika.

Em (MADDEN et al., 2024) é utilizada uma versão diferente do *SIR* que utiliza séries temporais, o *TSIR*. A PINN utilizada tem uma entrada de alta dimensionalidade que incorpora dados da população, retardos na transmissão da doença, dados sobre as cidades vizinhas. O modelo desenvolvido é aplicado para os dados de sarampo da Inglaterra e do país de Gales entre os anos de 1944 e 1965. Uma novidade deste artigo é tentar buscar formas de inserir explicabilidade nas PINNs através de técnicas como o *SHAP* (*SHapley Additive exPlanations*).

Um exemplo de trabalho utilizando um modelo alternativo pode ser encontrado em (HU et al., 2022), os autores aplicam o modelo *SICDR* (*Susceptible-Infected-Confirmed-Recovered-Deceased*). Uma novidade apresentada nesse trabalho, é uma modificação na função de perda da PINN que mantém a estabilidade do modelo, mesmo com compartimentos faltantes. Outra abordagem inovadora é o uso de funções *wavelet* para suavizar o ruído nos dados. O modelo desenvolvido é aplicado para os dados de COVID-19 do Alabama no ano de 2020.

Em (SCHIASI et al., 2021) é proposto uma nova abordagem que combina PINNs com uma técnica de interpolação chamada Teoria de Conexões Funcionais. O algoritmo funciona aproximando funções por expressões condicionadas, uma expressão condicionada uma é a soma de funções livres e um funcional que analiticamente satisfaz as condições impostas não importando quais funções livres são escolhidas. O modelo proposto é feito *SIR*, *SEIR* e *SEIRS* (*Susceptible-Exposed-Infectious-Recovered-Susceptible*, um *SEIR* com reinfeção). São feitos testes com dados sintéticos com e sem ruído.

Um exemplo utilizando redes neurais recorrentes pode ser encontrado em (RODRI-

[GUEZ et al., 2023](#)). Os autores propõem uma *framework* que combina uma rede *feedforward* informada pela física com redes recorrentes. O intuito dos autores é desenvolver uma *framework* pensada para a previsão de tendências no desenvolvimento de uma pandemia a longo prazo incorporando dados das mais diversas fontes.

4 Proposta do Trabalho

Nesta seção é apresentada uma proposta para a estimativa dos parâmetros do modelo compartimental para dados epidemiológicos. O modelo utilizado é apresentado, assim como a arquitetura da PINN utilizada. São também detalhados os testes para averiguar a efetividade do método.

4.1 Estimativa de parâmetros

Os modelos compartimentais possuem parâmetros de transmissão e mortalidade fixos, considerando que estes modelos foram pensados apenas para dar uma projeção de como uma epidemia evoluirá. Entretanto, medidas de afastamento social são capazes de alterar o parâmetro de transmissão ao longo do tempo. Utilizando (LONG; KHALIQ; FURATI, 2021) como inspiração, é proposto obter o parâmetro β como uma função em função do tempo. A rede neural deverá se ajustar a um $\beta(t)$. O modelo *SIR* passa a ser escrito como,

$$\frac{dS(t)}{dt} = -\frac{\beta(t)S(t)}{N}I(t), \quad t > t_0, \quad (4.1)$$

$$\frac{dI(t)}{dt} = \frac{\beta(t)S(t)}{N}I(t) - \gamma I(t), \quad t > t_0, \quad (4.2)$$

$$\frac{dR(t)}{dt} = \gamma I(t), \quad t > t_0, \quad (4.3)$$

A taxa de mortalidade de um doença permanece a mesma para a maioria das doenças, logo não há a necessidade de transformá-lo numa equação em função do tempo. Considerando esse fato e seguindo o que foi aplicado em trabalhos como (MILLEVOI; PASETTO; FERRONATO, 2024) e (OUYOUSSEF et al., 2024), pode-se simplificar o sistema de equações acima, considerando que $S(t) + I(t) + R(t) = N$ para todo $t > t_0$. A função $R(t)$ passa então a ser obtida pela subtração do total de indivíduos na população pelos valores de $S(t)$ e $I(t)$.

$$R(t) = N - S(t) - I(t) \quad (4.4)$$

A rede neural passar a ter que aproximar uma função $R \rightarrow R^3$ de (S, I, β) em função de t como ilustrado na figura 10.

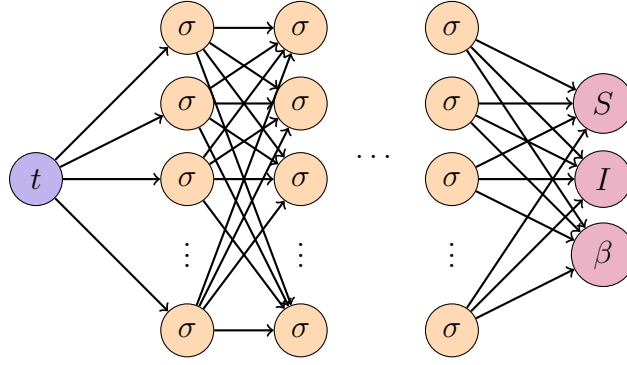


Figura 10 – Representação gráfica das redes *feedforward*. Fonte: elaborada pelos autores.

4.2 Justificativa para a Escolha do Modelo SIR

O modelo *SIR* pode não ser o melhor para os testes com dados reais de influenza, entretanto, considerando o período de tempo em que o algoritmo foi executado não há necessidade de modelar a possibilidade de reinfecção. O motivo é que, apesar da gripe não gerar imunidade permanente, ainda assim gera imunidade por tempo suficiente para poder ser desconsiderada pelo período de um ano. Logo não há necessidade de usar modelos como o *SIRS* ou *SIS*.

4.3 Testes com Dados Sintéticos

Para averiguar se o método proposto funcionará bem com dados reais, será feito primeiramente um teste com dados sintéticos obtidos a partir da solução do modelo compartimental utilizando o método de Runge-Kutta de 4º ordem (RK4) implementado na biblioteca *SciPy* (VIRTANEN et al., 2020). O método resolverá o sistema de equações diferenciais ao aproximar funções para os 3 compartimentos do modelo SIR com o parâmetro β em função do tempo. O único que será conhecido previamente, será justamente a função que descreve β . Para emular um beta que varia em função do tempo, a função 4.5 será utilizada. (ANDREU-VILARROIG; GONZÁLEZ-PARRA; VILLANUEVA, 2025) (EDLUND et al., 2011).

$$\beta(t) = \sin\left(\frac{2\pi t}{t_f - t_0}\right) + \beta_{min} \quad (4.5)$$

O modelo empregado é ligeiramente diferente do *SIR* original. Será utilizada a sua forma normalizada, bastando remover o divisor N das equações,

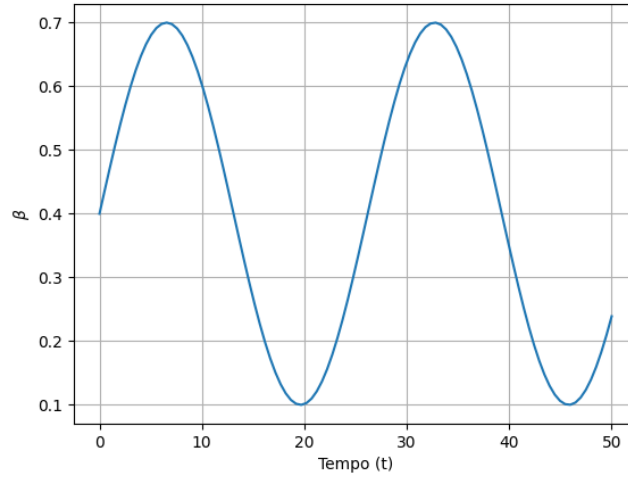


Figura 11 – Na primeira. Fonte: elaborada pelos autores.

$$\frac{dS(t)}{dt} = -\beta(t)S(t)I(t), \quad t > t_0, \quad (4.6)$$

$$\frac{dI(t)}{dt} = \beta(t)S(t)I(t) - \gamma I(t), \quad t > t_0, \quad (4.7)$$

$$\frac{dR(t)}{dt} = \gamma I(t), \quad t > t_0, \quad (4.8)$$

Os compartimentos do modelo passam a ter um outro significado, não mostrado mais o número de indivíduos em cada compartimento, mas sim a porcentagem da população que está em cada compartimento no instante t . Logo, a soma dos três compartimentos será igual a 1 para todo instante t

$$S(t) + I(t) + R(t) = 1, \quad t > t_0 \quad (4.9)$$

Levando isso em conta, as condições iniciais do problema são também expressadas como porcentagens.

$$S_0 = 0.99 \quad (4.10)$$

$$I_0 = 0.01, \quad (4.11)$$

$$R_0 = 0 \quad (4.12)$$

O problema é então resolvido para $t_0 \leq t \leq t_f$, sendo $t_0 = 0$ e $t_f = 50$. A solução numérica obtida é então utilizada para obter um conjunto de N_{data} valores para cada compartimento do modelo *SIR*. Para este experimento, foram usados 100 A figura 12

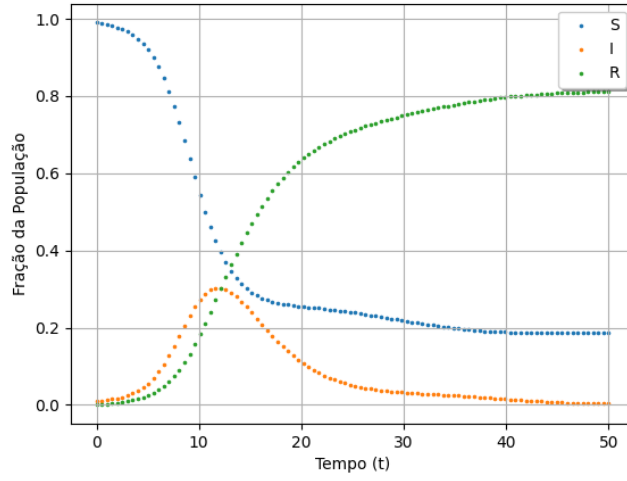


Figura 12 – Na primeira. Fonte: elaborada pelos autores.

A solução com RK4 retorna dados para os três compartimentos, mas como a versão do *SIR* utilizada neste trabalho não possui a terceira equação, os dados para o terceiro compartimento (R) não são usados. Tomando como exemplo trabalhos como (HAN et al., 2024) que utilizam apenas dados do compartimento de infectados, mas conseguem bons resultados, o modelo também será treinado apenas com dados sintéticos de incidência de infectados.

4.3.1 Avaliação dos Resultados

A qualidade da solução obtida pela rede neural é medida pelo uso das métricas de *MSE*, norma euclidiana (\mathcal{L}_2) e norma no infinito (\mathcal{L}_∞). O *MSE*, é definida pela equação 4.13

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4.13)$$

A norma \mathcal{L}_2 , é definida pela equação 4.14 e corresponde a distância euclidiana entre as superfícies. É uma grandeza escalar

$$\mathcal{L}_2(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4.14)$$

A norma \mathcal{L}_∞ , também chamada de norma máxima, é definida pela equação 4.15 e corresponde à distância máxima entre as superfícies \mathbf{x} e \mathbf{y} . A vantagem dessa métrica é mostrar qual é o maior erro pontual. Uma desvantagem desta métrica, é ser sensível a *outliers*.

$$\mathcal{L}_\infty(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_\infty = \max_i |x_i - y_i| \quad (4.15)$$

A comparação é feita com a solução numérica obtida com RK4 no caso dos compartimentos de suscetíveis e infectados. Para o parâmetro de taxa de infecção β , a comparação é feita com a equação 4.5.

4.4 Testes com Dados Sintéticos Ruidosos

Segundo (RAISSI; PERDIKARIS; KARNIADAKIS, 2019), PINNs são resilientes a dados ruidosos. Para testar se PINNs são capazes de corretamente aproximar as curvas do *SIR* mesmo com dados ruidosos, é adicionado ruído gaussiano branco aos dados obtidos na seção anterior. A equação 4.16 descreve este processo.

$$Z_t \sim \mathcal{N}(0, \sigma) \quad (4.16)$$

$$\mathcal{N}_t = \max(\mathcal{C}_t + Z_t, 0) \quad (4.17)$$

A figura 13 mostra o resultado deste processo sobre os dados de treinamento. Vale observar que para evitar usar dados negativos, pois são valores que não fazem sentido no contexto do modelo, foi utilizada a função de máximo entre o valor do dado de treinamento e zero. Com este experimento, espera-se validar o método para a aplicação com dados reais, que normalmente são perturbados.

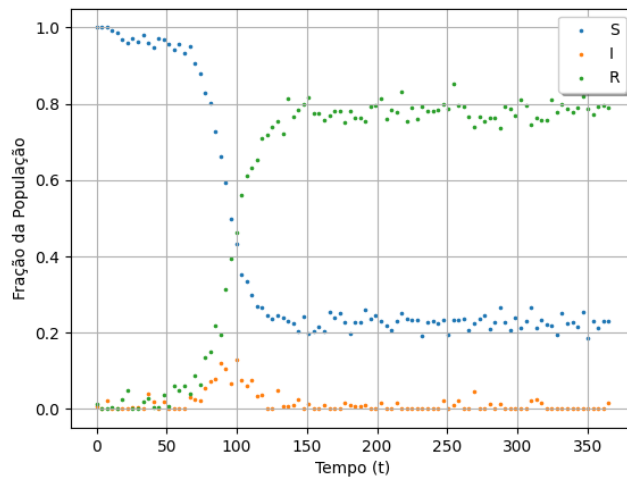


Figura 13 – Na primeira. Fonte: elaborada pelos autores.

4.5 Testes com diferentes pontos de equilíbrio

São feitos testes com diferentes pontos de equilíbrio livre de doença, pois uma vez atingindo este ponto de equilíbrio, a rede para de aproximar corretamente o valor de β . As figuras 14 e 15 mostram os dados para o segundo ponto de equilíbrio para os casos de dados com e sem ruído, respectivamente.

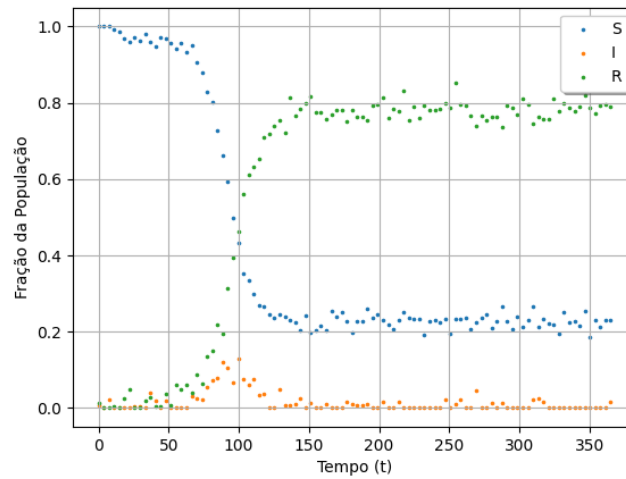


Figura 14 – Na primeira. Fonte: elaborada pelos autores.

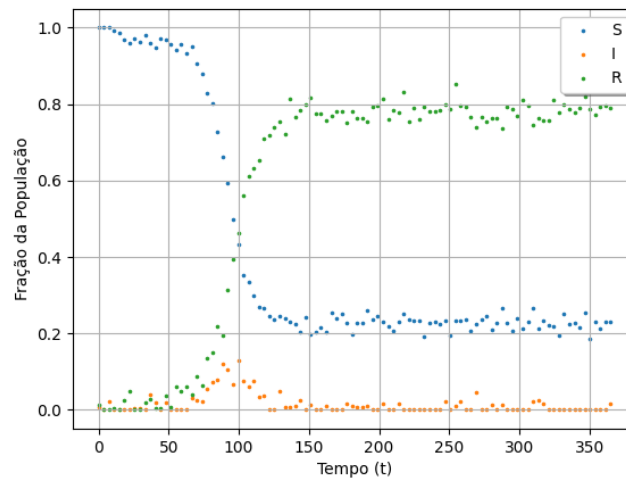


Figura 15 – Na primeira. Fonte: elaborada pelos autores.

4.6 Testes com Base de Dados Reais

São realizados testes utilizando dados reais...

4.6.1 Base de dados da OMS

Os dados foram coletados da plataforma *FluNet* ([ORGANIZATION,](#))

4.6.2 Bases de dados do DataSUS

Os dados foram coletados do plataforma *OpenDataSUS* ([SAÚDE, 2024](#))

4.6.3 Tratamento dos Dados

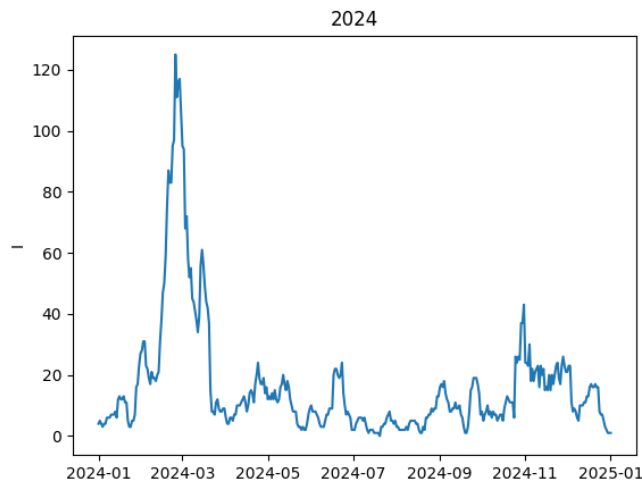


Figura 16 – Na primeira. Fonte: elaborada pelos autores.

Janela móvel de 7 dias como em ([HAN et al., 2024](#)), ([LONG; KHALIQ; FURATI, 2021](#)) e ([SHAMSARA; KÖNIG; PFEIFER, 2025](#)) para suavizar o ruído

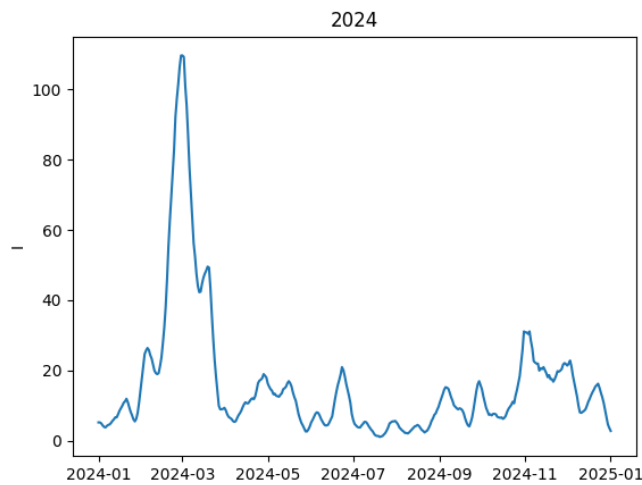


Figura 17 – Na primeira. Fonte: elaborada pelos autores.

Pesos ajustáveis entre a loss física e a loss dos dados como em ([LONG; KHALIQ; FURATI, 2021](#)) e ([SHAMSARA; KÖNIG; PFEIFER, 2025](#))

Segundo (BONFANTI et al., 2024), PINNs não generalizam bem fora do domínio de treinamento. PINNs podem estimar os parâmetros para fora do domínio de treinamento como em (MILLEVOI; PASETTO; FERRONATO, 2024).

Assim como em (GHOSH et al., 2023), fazer testes com lacunas nos dados para testar a resiliência do método.

4.7 Uso da Matriz de Nova Geração

As condições inicial para os compartimentos *Susceptible* e *Infected* podem ser facilmente determinadas pelo número total da população e número de infectados no primeiro dia. O mesmo não acontece para a taxa de transmissão β . Para estimar o valor β_0 , foi utilizado método de matriz de nova geração.

4.8 Arquitetura da Rede

Baseando-se em (SHAIER; RAISSI; SESHAIYER, 2022), o número de camadas escolhido foi...

Como em (MILLEVOI; PASETTO; FERRONATO, 2024), aplicar uma *hard-constraint* na rede neural ao utilizar nos nós de saída uma função de ativação que retorna apenas valores positivos.

4.8.1 Parâmetro Ajustável

Aplicando a ideia de (SHAMSARA; KÖNIG; PFEIFER, 2025), é adicionado uma parâmetros ω_{dados} ajustável.

$$\omega_{\text{dados}} = \frac{\max(\text{gradientes de } \mathcal{L}_{\text{física}})}{\text{média}(\text{gradientes de } \mathcal{L}_{\text{dados}})} \quad (4.18)$$

4.8.2 Regularização L2

Seguindo os trabalhos de (LONG; KHALIQ; FURATI, 2021) e (SHAMSARA; KÖNIG; PFEIFER, 2025), é aplicado uma regularização L2 na loss.

$$\mathcal{L}_2(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4.19)$$

Os benefícios ligados ao uso da regularização L2 estão ligados a evitar gradientes muito grandes, o que pode comprometer o processo de convergência da rede neural.

4.8.3 Correlação com a Temperatura

A gripe é uma doença com maior taxa de transmissão nos meses frios. Para testar se a transmissão em função do tempo aproximada pelo modelo é plausível, é feito um teste de correlação de *Pearson* entre $\beta(t)$ e a temperatura ao longo do ano.

4.8.3.1 Teste de Correlação de Pearson

Sendo β e T duas amostras de tamanho n com dados pareados (β_i, T_i) :

$$\rho = \frac{\sum_{i=1}^n (\beta_i - \bar{\beta})(T_i - \bar{T})}{\sqrt{\sum_{i=1}^n (\beta_i - \bar{\beta})^2 \sum_{i=1}^n (T_i - \bar{T})^2}} \quad (4.20)$$

Espera-se um valor de ρ acima de 0.5 para indicar uma correlação no mínimo moderada entre a taxa de transmissão β e a temperatura.

4.8.3.2 Teste de Correlação de Spearman

A correlação de Spearman é equivalente a calcular a correlação de Pearson nos *ranks* dos valores.

$$r_s = \frac{\sum_{i=1}^n (R(\beta_i) - \bar{R}_\beta)(R(T_i) - \bar{R}_T)}{\sqrt{\sum_{i=1}^n (R(\beta_i) - \bar{R}_\beta)^2 \sum_{i=1}^n (R(T_i) - \bar{R}_T)^2}} \quad (4.21)$$

Caso haja empates entre os elementos das amostras, aplica-se a fórmula de ajuste abaixo.

$$r_s = \frac{\sum_{i=1}^n (R(\beta_i) - \bar{R}_\beta)(R(T_i) - \bar{R}_y)}{\sqrt{\left[\sum_{i=1}^n (R(x_i) - \bar{R}_x)^2 - T_x\right] \left[\sum_{i=1}^n (R(y_i) - \bar{R}_y)^2 - T_y\right]}} \quad (4.22)$$

Sendo que os fatores T_x e T_y de correção de empate são calculados usando as fórmulas abaixo.

$$\mathcal{T} = \sum \frac{t^3 - t}{12} \quad (4.23)$$

4.9 Implementação e Execução

A implementação foi feita utilizando a biblioteca *DeepXDE* (LU et al., 2021), utilizando o *TensorFlow* (ABADI et al., 2016) como *backend*. Todo o código, dados utilizados nos experimentos, encontram-se disponíveis no repositório público no GitHub¹. As sementes para a geração de pseudonúmeros foram fixadas para garantir a reprodutibilidade

¹ <<https://github.com/ginbar/inverse-cm>>

dos experimentos.

Todos os experimentos foram feitos numa máquina com 4GB de ram, processador

5 Avaliação do Trabalho

Nesta seção são apresentados os resultados dos experimentos propostos no capítulo 4. Para cada experimento, são feitas pequenas discussões acerca dos resultados.

5.1 Testes com Dados Sintéticos

A figura 19 mostra o valor de beta ao longo do tempo.

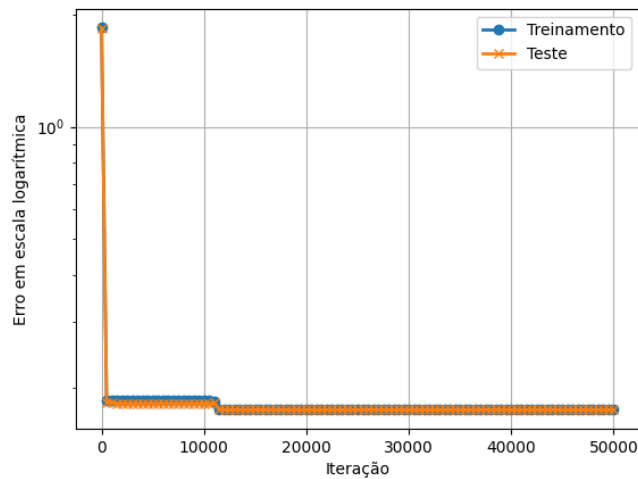


Figura 18 – Na primeira. Fonte: elaborada pelos autores.

A figura 19 mostra o valor de beta ao longo do tempo.

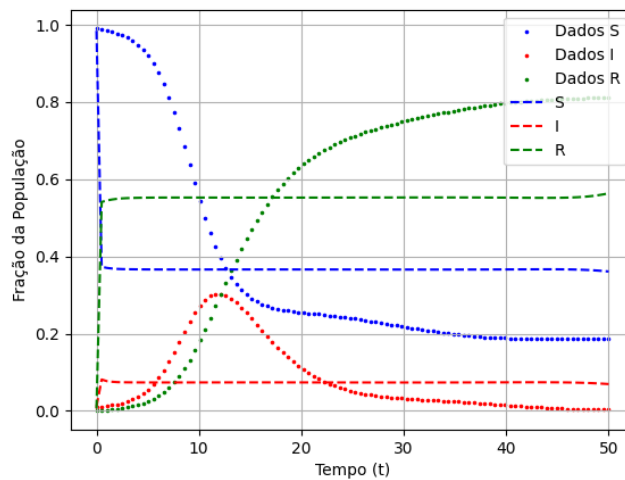


Figura 19 – Na primeira. Fonte: elaborada pelos autores.

A figura 20 mostra o valor de beta ao longo do tempo.

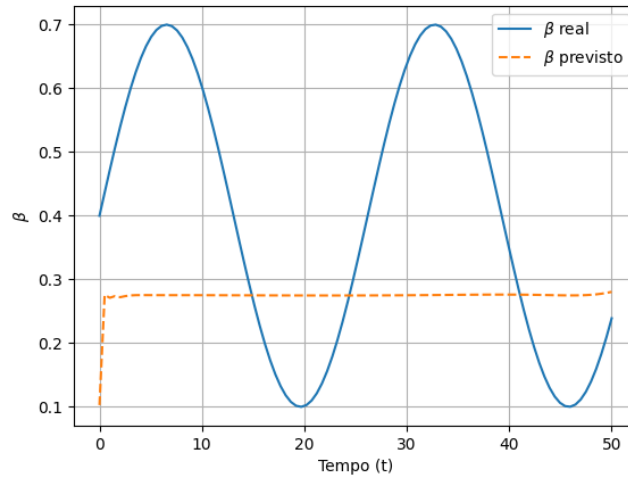


Figura 20 – Na primeira. Fonte: elaborada pelos autores.

A tabela 1 mostra os valores para

Tabela 1 – Valores das métricas de erro (MSE , norma \mathcal{L}_2 e norma \mathcal{L}_∞) para as soluções aproximadas pela rede neural, em comparação com as soluções analíticas.

Compartimento	Métricas		
	MSE	\mathcal{L}_2	\mathcal{L}_∞
S	$8,628 \times 10^{-6}$	$2,444 \times 10^{-3}$	$1,182 \times 10^{-2}$
I	$1,005 \times 10^{-6}$	$1,252 \times 10^{-2}$	$4,81 \times 10^{-3}$
R	$1,997 \times 10^{-6}$	$4,913 \times 10^{-3}$	$8,635 \times 10^{-3}$

5.2 Testes com Dados Ruidosos

A figura 19 mostra o valor de beta ao longo do tempo.

A figura 22 mostra o valor de beta ao longo do tempo.

A figura 23 mostra o valor de beta ao longo do tempo.

A tabela 2 mostra os valores para

Tabela 2 – Valores das métricas de erro (MSE , norma \mathcal{L}_2 e norma \mathcal{L}_∞) para as soluções aproximadas pela rede neural, em comparação com as soluções analíticas.

Compartimento	Métricas		
	MSE	\mathcal{L}_2	\mathcal{L}_∞
S	$8,628 \times 10^{-6}$	$2,444 \times 10^{-3}$	$1,182 \times 10^{-2}$
I	$1,005 \times 10^{-6}$	$1,252 \times 10^{-2}$	$4,81 \times 10^{-3}$
R	$1,997 \times 10^{-6}$	$4,913 \times 10^{-3}$	$8,635 \times 10^{-3}$

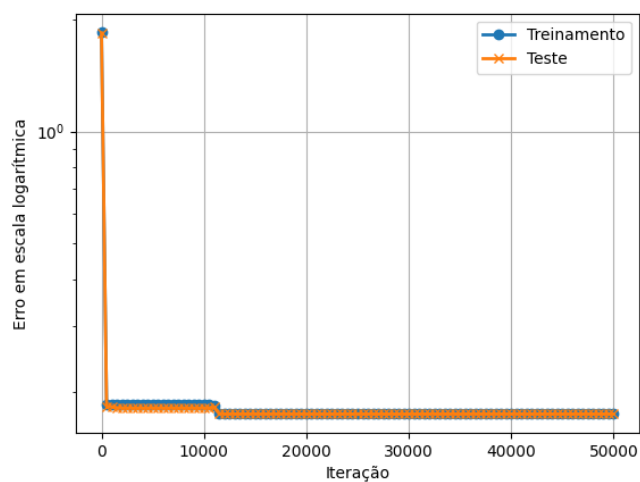


Figura 21 – Na primeira. Fonte: elaborada pelos autores.

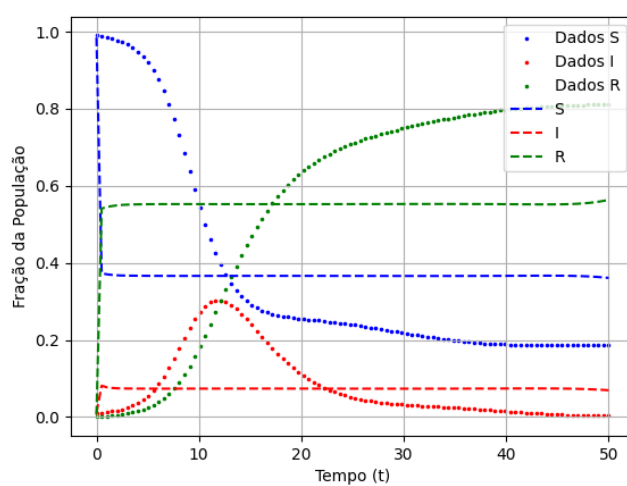


Figura 22 – Na primeira. Fonte: elaborada pelos autores.

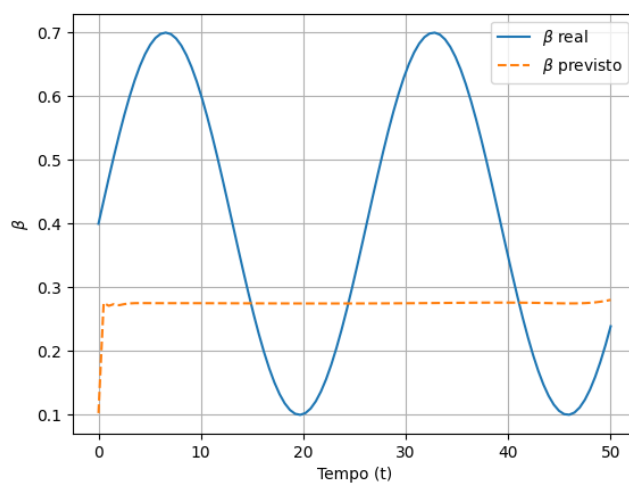


Figura 23 – Na primeira. Fonte: elaborada pelos autores.

5.3 Testes com Base de Dados Reais

6 Considerações Finais

Aplicar PINNs Bayesianas ([YANG; MENG; KARNIADAKIS, 2021](#)) para estimar o desvio padrão do $\beta(t)$.

Referências

ABADI, M. et al. Tensorflow: A system for large-scale machine learning. *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, p. 265–283, 2016. Citado 2 vezes nas páginas 16 e 50.

ANDREU-VILARROIG, C.; GONZÁLEZ-PARRA, G.; VILLANUEVA, R.-J. Mathematical modeling of influenza dynamics: Integrating seasonality and gradual waning immunity. *Bulletin of Mathematical Biology*, v. 87, n. 6, p. 75, maio 2025. ISSN 1522-9602. Disponível em: <<https://doi.org/10.1007/s11538-025-01454-w>>. Citado na página 43.

ARULANDU, A. C.; SESHAIYER, P. Physics-informed neural networks for informed vaccine distribution in meta-populations. *Journal of Machine Learning for Modeling and Computing*, v. 4, n. 3, p. 83 – 99, 2023. Cited by: 0. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-105006553822&doi=10.1615%2fJMachLearnModelComput.2023047642&partnerID=40&md5=d04e4b230160aee713a7bc1571692a31>>. Citado na página 37.

ATANGANA, A.; BILDIK, N. The use of fractional order derivative to predict the groundwater flow. *Mathematical Problems in Engineering*, v. 2013, n. 1, p. 543026, 2013. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1155/2013/543026>>. Citado na página 27.

BERTAGLIA, G. et al. Asymptotic-preserving neural networks for multiscale hyperbolic models of epidemic spread. *Mathematical Models and Methods in Applied Sciences*, v. 32, n. 10, p. 1949 – 1985, 2022. Cited by: 17; All Open Access, Green Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85140742473&doi=10.1142%2fS0218202522500452&partnerID=40&md5=170623db80cb3dca5ce350b87afb42f1>>. Citado na página 37.

BONFANTI, A. et al. On the generalization of PINNs outside the training domain and the hyperparameters influencing it. *Neural Computing and Applications*, v. 36, n. 36, p. 22677–22696, dez. 2024. ISSN 1433-3058. Disponível em: <<https://doi.org/10.1007/s00521-024-10178-2>>. Citado na página 49.

CHANTADA, A. T. et al. Cosmology-informed neural networks to solve the background dynamics of the universe. *Phys. Rev. D*, American Physical Society, v. 107, p. 063523, Mar 2023. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevD.107.063523>>. Citado na página 16.

CHEN, R. T. et al. Neural ordinary differential equations. In: . [s.n.], 2018. v. 2018-December, p. 6571 – 6583. Cited by: 3352. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85064831041&partnerID=40&md5=a0ffdeab131ce4f8c209514e5f53df1e>>. Citado na página 15.

CYBENKO, G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, v. 2, p. 303–314, 1989. Disponível em: <<https://doi.org/10.1007/BF02551274>>. Citado na página 14.

- DIEKMANN, O.; HEESTERBEEK, J. A. P.; METZ, J. A. J. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, v. 28, n. 4, p. 365–382, jun. 1990. ISSN 1432-1416. Disponível em: <<https://doi.org/10.1007/BF00178324>>. Citado na página 33.
- DWIVEDI, V.; PARASHAR, N.; SRINIVASAN, B. Distributed physics informed neural network for data-efficient solution to partial differential equations. *ArXiv*, abs/1907.08967, 2019. Disponível em: <<https://api.semanticscholar.org/CorpusID:198147990>>. Citado na página 27.
- EDLUND, S. et al. Comparing three basic models for seasonal influenza. *Epidemics*, v. 3, n. 3, p. 135–142, 2011. ISSN 1755-4365. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1755436511000259>>. Citado na página 43.
- FUHG, J. N. et al. Interval and fuzzy physics-informed neural networks for uncertain fields. *Probabilistic Engineering Mechanics*, v. 68, p. 103240, 2022. ISSN 0266-8920. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0266892022000273>>. Citado na página 27.
- GHOSH, S. et al. Understanding the implications of under-reporting, vaccine efficiency and social behavior on the post-pandemic spread using physics informed neural networks: A case study of china. *PLOS ONE*, v. 18, 11 2023. Citado 2 vezes nas páginas 39 e 49.
- GIAMPAOLO, F. et al. Physics-informed neural networks approach for 1D and 2D Gray-Scott systems. *Advanced Modeling and Simulation in Engineering Sciences*, Springer, v. 9, n. 1, p. 1–17, maio 2022. Disponível em: <<https://hal.science/hal-04456081>>. Citado na página 16.
- GILES, P. The mathematical theory of infectious diseases and its applications. *Journal of the Operational Research Society*, v. 28, n. 2, p. 479–480, jul. 1977. ISSN 1476-9360. Disponível em: <<https://doi.org/10.1057/jors.1977.92>>. Citado na página 35.
- GRIPENBERG, G. Approximation by neural networks with a bounded number of nodes at each level. *Journal of Approximation Theory*, v. 122, n. 2, p. 260–266, 2003. ISSN 0021-9045. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0021904503000789>>. Citado na página 14.
- HAN, J.; LEE, Y. Hierarchical learning to solve pdes using physics-informed neural networks. In: MIKYŠKA, J. et al. (Ed.). *Computational Science – ICCS 2023*. Cham: Springer Nature Switzerland, 2023. p. 548–562. ISBN 978-3-031-36024-4. Citado na página 25.
- HAN, S. et al. Approaching epidemiological dynamics of covid-19 with physics-informed neural networks. *Journal of the Franklin Institute*, v. 361, n. 6, 2024. Cited by: 12; All Open Access, Green Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85186494686&doi=10.1016%2fj.jfranklin.2024.106671&partnerID=40&md5=c6f62df7ebafa637a85e89ac5ee26f45>>. Citado 3 vezes nas páginas 37, 45 e 48.
- HELDMANN, F. et al. Pinn training using biobjective optimization: The trade-off between data loss and residual loss. *Journal of Computational Physics*, v. 488, p. 112211,

2023. ISSN 0021-9991. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0021999123003066>>. Citado na página 38.

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. *Neural Networks*, v. 2, n. 5, p. 359–366, 1989. ISSN 0893-6080. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0893608089900208>>. Citado na página 14.

HU, H. et al. A modified pinn approach for identifiable compartmental models in epidemiology with application to covid-19. *Viruses*, v. 14, n. 11, 2022. Cited by: 10; All Open Access, Gold Open Access, Green Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85141594087&doi=10.3390%2fv14112464&partnerID=40&md5=a4e3585cc4da89e3d73b5a5303bdee78>>. Citado na página 40.

HU, Z. et al. Augmented physics-informed neural networks (apinns): A gating network-based soft domain decomposition methodology. *Engineering Applications of Artificial Intelligence*, v. 126, p. 107183, 2023. ISSN 0952-1976. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0952197623013672>>. Citado na página 27.

JACQUEZ, J. A. et al. Modeling and analyzing hiv transmission: the effect of contact patterns. *Mathematical Biosciences*, v. 92, n. 2, p. 119–199, 1988. ISSN 0025-5564. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0025556488900314>>. Citado na página 37.

JAGTAP, A. D.; KARNIADAKIS, G. E. Extended physics-informed neural networks (xpinns): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. *Communications in Computational Physics*, v. 28, n. 5, p. 2002–2041, 2020. Citado na página 26.

JAGTAP, A. D.; KHARAZMI, E.; KARNIADAKIS, G. E. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, v. 365, p. 113028, 2020. ISSN 0045-7825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0045782520302127>>. Citado na página 26.

JIN, H.; MATTHEAKIS, M.; PROTOPAPAS, P. Physics-informed neural networks for quantum eigenvalue problems. In: IEEE. *2022 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2022. p. 1–8. Citado na página 16.

JIN, X. et al. Nsfnets (navier-stokes flow nets): Physics-informed neural networks for the incompressible navier-stokes equations. *Journal of Computational Physics*, v. 426, p. 109951, 2021. ISSN 0021-9991. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0021999120307257>>. Citado na página 16.

KARLBAUER, M. et al. *Composing Partial Differential Equations with Physics-Aware Neural Networks*. 2022. Disponível em: <<https://arxiv.org/abs/2111.11798>>. Citado na página 25.

KENDALL, D. G. Deterministic and stochastic epidemics in closed populations. In:

_____. *Volume 4 Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume IV*. Berkeley: University of California Press, 2023. p. 149–166. ISBN 9780520350717. Disponível em: <<https://doi.org/10.1525/9780520350717-011>>. Citado na página 29.

KERMACK, W. O.; MCKENDRICK, A. G. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, v. 115, n. 772, p. 700–721, 1927. Citado 2 vezes nas páginas 16 e 29.

KERMACK, W. O.; MCKENDRICK, A. G. Contributions to the mathematical theory of epidemics. ii. —the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, v. 138, n. 834, p. 55–83, 1932. Disponível em: <<https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1932.0171>>. Citado 2 vezes nas páginas 29 e 35.

KERMACK, W. O.; MCKENDRICK, A. G. Contributions to the mathematical theory of epidemics. iii.—further studies of the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, v. 141, n. 843, p. 94–122, 1933. Disponível em: <<https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1933.0106>>. Citado na página 29.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *International Conference on Learning Representations*, 2014. Citado na página 23.

LAGARIS, I. E.; LIKAS, A.; FOTIADIS, D. I. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks*, v. 9, n. 5, p. 987–1000, Sep. 1998. ISSN 1045-9227. Citado na página 15.

LI, R. et al. Dynamic analysis and data-driven inference of a fractional-order seihdr epidemic model with variable parameters. *Mathematics and Computers in Simulation*, v. 230, p. 1 – 19, 2025. Cited by: 0. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85208764841&doi=10.1016%2fj.matcom.2024.10.042&partnerID=40&md5=5df06da84bff8326ce9e8e000816f191>>. Citado na página 38.

LINNAINMAA, S. Taylor expansion of the accumulated rounding error. *BIT Numerical Mathematics*, v. 16, n. 2, p. 146–160, jun. 1976. ISSN 1572-9125. Disponível em: <<https://doi.org/10.1007/BF01931367>>. Citado na página 16.

LIU, D. C.; NOCEDAL, J. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, v. 45, p. 503–528, 1989. Disponível em: <<https://doi.org/10.1007/BF01589116>>. Citado na página 23.

LIU, Z. et al. KAN: Kolmogorov-Arnold Networks. 2025. Disponível em: <<https://arxiv.org/abs/2404.19756>>. Citado na página 28.

LONG, J.; KHALIQ, A.; FURATI, K. Identification and prediction of time-varying parameters of covid-19 model: a data-driven deep learning approach. *International Journal of Computer Mathematics*, v. 98, n. 8, p. 1617 – 1632, 2021. Cited by: 33; All Open Access, Green Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85106627387&doi=10.1080%2f10687125.2021.1911111>>.

2f00207160.2021.1929942&partnerID=40&md5=803b3e59fc4241f274b002aae280c6ef>. Citado 5 vezes nas páginas 17, 38, 42, 48 e 49.

LU, L. et al. Deepxde: A deep learning library for solving differential equations. *SIAM Review*, Society for Industrial & Applied Mathematics (SIAM), v. 63, n. 1, p. 208–228, jan. 2021. ISSN 1095-7200. Disponível em: <<http://dx.doi.org/10.1137/19M1274067>>. Citado 2 vezes nas páginas 23 e 50.

MADDEN, W. G. et al. Deep neural networks for endemic measles dynamics: Comparative analysis and integration with mechanistic models. *PLoS Computational Biology*, v. 20, n. 11, 2024. Cited by: 1. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85209953346&doi=10.1371%2fjournal.pcbi.1012616&partnerID=40&md5=062a1370cac26e07e1443b565102f756>>. Citado na página 40.

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, dez. 1943. ISSN 1522-9602. Disponível em: <<https://doi.org/10.1007/BF02478259>>. Citado 2 vezes nas páginas 14 e 19.

MEADE, A. J.; FERNANDEZ, A. A. The numerical solution of linear ordinary differential equations by feedforward neural networks. *Math. Comput. Model.*, Elsevier Science Publishers B. V., v. 19, n. 12, p. 1–25, 1994. Citado na página 14.

MILLEVOI, C.; PASETTO, D.; FERRONATO, M. A physics-informed neural network approach for compartmental epidemiological models. *PLoS Computational Biology*, v. 20, n. 9, 2024. Cited by: 2; All Open Access, Gold Open Access, Green Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85203432916&doi=10.1371%2fjournal.pcbi.1012387&partnerID=40&md5=5fc619838ca1b408adc707c6a90ac0a8>>. Citado 3 vezes nas páginas 39, 42 e 49.

MITUSCH, S. K.; FUNKE, S. W.; KUCHTA, M. Hybrid fem-nn models: Combining artificial neural networks with the finite element method. *Journal of Computational Physics*, v. 446, p. 110651, 2021. ISSN 0021-9991. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0021999121005465>>. Citado na página 15.

MOHAMMADIAN, M.; BAKER, K.; FIORETTTO, F. Gradient-enhanced physics-informed neural networks for power systems operational support. *Electric Power Systems Research*, v. 223, p. 109551, 2023. ISSN 0378-7796. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378779623004406>>. Citado na página 25.

NELSON, S. P. et al. Modeling the dynamics of covid-19 in japan: employing data-driven deep learning approach. *International Journal of Machine Learning and Cybernetics*, 2024. Cited by: 3. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85201409261&doi=10.1007%2fs13042-024-02301-5&partnerID=40&md5=fcc869d76810d112561a40fcb39068b8>>. Citado 2 vezes nas páginas 36 e 37.

NGUYEN, L.; RAISSI, M.; SESHAIYER, P. Modeling, analysis and physics informed neural network approaches for studying the dynamics of covid-19 involving human-human and human-pathogen interaction. *Computational and Mathematical Biophysics*, v. 10, n. 1, p. 1 – 17, 2022. Cited by: 22; All Open Access, Gold Open Access. Disponível em:

<<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85125789439&doi=10.1515%2fcmb-2022-0001&partnerID=40&md5=90f6956fb3cafb7a76f6498f77fef89>>. Citado na página 38.

NING, X. et al. Physics-informed neural networks integrating compartmental model for analyzing covid-19 transmission dynamics. *Viruses*, v. 15, n. 8, 2023. Cited by: 14; All Open Access, Gold Open Access, Green Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85168905930&doi=10.3390%2fv15081749&partnerID=40&md5=0d75aa500c625235ce5bda5a1fccedbc>>. Citado na página 39.

NOBLE, J. V. Geographic and temporal development of plagues. *Nature*, v. 250, n. 5469, p. 726–729, ago. 1974. ISSN 1476-4687. Disponível em: <<https://doi.org/10.1038/250726a0>>. Citado na página 36.

OGUEDA-OLIVA, A. G. et al. Machine learning for predicting the dynamics of infectious diseases during travel through physics informed neural networks. *Journal of Machine Learning for Modeling and Computing*, v. 4, n. 3, p. 17 – 35, 2023. Cited by: 1. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-105006554742&doi=10.1615%2fJMachLearnModelComput.2023047213&partnerID=40&md5=672b3d4fc327420a2609ab66cd4a62c2>>. Citado na página 39.

ORGANIZATION, W. H. *FluNet*. <<https://www.who.int/tools/flunet>>. Accessed: 2025-09-07. Disponível em: <<https://www.who.int/tools/flunet>>. Citado na página 48.

OUYOUSSEF, K. I. et al. Physics-informed neural networks for parameter estimation and simulation of a two-group epidemiological model. *International Journal of Modeling, Simulation, and Scientific Computing*, v. 15, n. 3, 2024. Cited by: 1. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85199247126&doi=10.1142%2fS1793962324500429&partnerID=40&md5=75c96ad2ee9d4ef94eb1f14d23b58425>>. Citado 2 vezes nas páginas 36 e 42.

PAKRAVAN, S. et al. Solving inverse-pde problems with physics-aware neural networks. *Journal of Computational Physics*, v. 440, p. 110414, 2021. ISSN 0021-9991. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0021999121003090>>. Citado na página 26.

PANG, G.; LU, L.; KARNIADAKIS, G. E. fpinns: Fractional physics-informed neural networks. *SIAM Journal on Scientific Computing*, v. 41, n. 4, p. A2603–A2626, 2019. Disponível em: <<https://doi.org/10.1137/18M1229845>>. Citado na página 28.

PASZKE, A. et al. Pytorch: An imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems*. [S.l.: s.n.], 2019. p. 8024–8035. Citado na página 16.

PSICHOGIOS, D. C.; UNGAR, L. H. A hybrid neural network-first principles approach to process modeling. *AIChE Journal*, v. 38, n. 10, p. 1499–1511, 1992. Citado na página 14.

RAISSI, M.; PERDIKARIS, P.; KARNIADAKIS, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, Elsevier, v. 378, p. 686–707, 2019. Citado 5 vezes nas páginas 14, 15, 22, 23 e 46.

- REN, P. et al. Phycrnet: Physics-informed convolutional-recurrent network for solving spatiotemporal pdes. *Computer Methods in Applied Mechanics and Engineering*, v. 389, p. 114399, 2022. ISSN 0045-7825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0045782521006514>>. Citado na página 25.
- RODRIGUEZ, A. et al. Einns: Epidemiologically-informed neural networks. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2023. v. 37. Citado na página 41.
- SAÚDE, M. da. *Notificações de Síndrome Gripal - 2024*. 2024. <<https://opendatasus.saude.gov.br/dataset/notificacoes-de-sindrome-gripal-leve-2024>>. Accessed: 2025-09-07. Disponível em: <<https://opendatasus.saude.gov.br/dataset/notificacoes-de-sindrome-gripal-leve-2024>>. Citado na página 48.
- SCHIASSI, E. et al. Physics-informed neural networks and functional interpolation for data-driven parameters discovery of epidemiological compartmental models. *Mathematics*, v. 9, n. 17, 2021. Cited by: 34; All Open Access, Gold Open Access, Green Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85114263779&doi=10.3390%2fmath9172069&partnerID=40&md5=6d6d5add48f2d06e3d42ad56f2d97802>>. Citado na página 40.
- SCHLICKER, R.; KRÖGER, M. Analytical modeling of the temporal evolution of epidemics outbreaks accounting for vaccinations. *Physics*, v. 3, n. 2, p. 386–426, 2021. ISSN 2624-8174. Disponível em: <<https://www.mdpi.com/2624-8174/3/2/28>>. Citado na página 36.
- SHAIER, S.; RAISSI, M.; SESHAIYER, P. Data-driven approaches for predicting spread of infectious diseases through dinns: Disease informed neural networks. *Letters in Biomathematics*, v. 9, n. 1, p. 71 – 105, 2022. Cited by: 18. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85136169389&partnerID=40&md5=2f5c407588258f3d630368da396af06a>>. Citado 3 vezes nas páginas 16, 40 e 49.
- SHAMSARA, E.; KÖNIG, F.; PFEIFER, N. An informed deep learning model of the omicron wave and the impact of vaccination. *Computers in Biology and Medicine*, v. 191, 2025. Cited by: 0. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-105002151013&doi=10.1016%2fj.compbiomed.2025.109968&partnerID=40&md5=b1244c15295694df7ec626c4394d6c0d>>. Citado 3 vezes nas páginas 38, 48 e 49.
- SHI, P.; ZENG, Z.; LIANG, T. Physics-informed convnet: Learning physical field from a shallow neural network. *Communications in Nonlinear Science and Numerical Simulation*, v. 132, p. 107911, 2024. ISSN 1007-5704. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1007570424000972>>. Citado na página 25.
- SINGH, P.; GUPTA, A. Generalized SIR (GSIR) epidemic model: An improved framework for the predictive monitoring of COVID-19 pandemic. *ISA Transactions*, v. 124, p. 31–40, maio 2022. Citado na página 36.
- SIRIGNANO, J.; SPILIOPOULOS, K. Dgm: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics*, v. 375, p. 1339–1364, 2018. ISSN 0021-9991. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0021999118305527>>. Citado na página 15.

- SUN, K.; FENG, X. A second-order network structure based on gradient-enhanced physics-informed neural networks for solving parabolic partial differential equations. *Entropy*, v. 25, n. 4, 2023. ISSN 1099-4300. Disponível em: <<https://www.mdpi.com/1099-4300/25/4/674>>. Citado na página 25.
- TOSCANO, J. D. et al. From PINNs to PIKANs: recent advances in physics-informed machine learning. *Machine Learning for Computational Science and Engineering*, v. 1, n. 1, p. 15, mar. 2025. ISSN 3005-1436. Disponível em: <<https://doi.org/10.1007/s44379-025-00015-1>>. Citado na página 28.
- VIRTANEN, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, v. 17, p. 261–272, 2020. Citado na página 43.
- WENGERT, R. E. A simple automatic derivative evaluation program. *Commun. ACM*, Association for Computing Machinery, New York, NY, USA, v. 7, n. 8, p. 463–464, ago. 1964. ISSN 0001-0782. Disponível em: <<https://doi.org/10.1145/355586.364791>>. Citado na página 16.
- WHEATCRAFT, S. W.; MEERSCHAERT, M. M. Fractional conservation of mass. *Advances in Water Resources*, v. 31, n. 10, p. 1377–1381, 2008. ISSN 0309-1708. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0309170808001188>>. Citado na página 28.
- WU, B. et al. Physics informed rnn-dct networks for time-dependent partial differential equations. In: GROEN, D. et al. (Ed.). *Computational Science – ICCS 2022*. Cham: Springer International Publishing, 2022. p. 372–379. ISBN 978-3-031-08754-7. Citado na página 26.
- YANG, J.; WANG, X.; LOU, J. Covid-19 reinfection processes: Mathematical modeling, data fitting with time-varying parameters, and optimal control studies. *Journal of Biological Systems*, v. 33, n. 2, p. 423 – 449, 2025. Cited by: 0. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85212778806&doi=10.1142%2fS021833902550010X&partnerID=40&md5=0c66b2d85a85efc24d1997bdd2e3278a>>. Citado na página 39.
- YANG, L.; MENG, X.; KARNIADAKIS, G. E. B-pinns: Bayesian physics-informed neural networks for forward and inverse pde problems with noisy data. *Journal of Computational Physics*, v. 425, p. 109913, 2021. ISSN 0021-9991. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0021999120306872>>. Citado na página 56.
- ZHANG, D.; GUO, L.; KARNIADAKIS, G. E. Learning in modal space: Solving time-dependent stochastic pdes using physics-informed neural networks. *SIAM Journal on Scientific Computing*, v. 42, n. 2, p. A639–A665, 2020. Disponível em: <<https://doi.org/10.1137/19M1260141>>. Citado na página 24.

Apêndices