



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

ARTHUR ARAÚJO BRUM

**Modelos de Crescimento Aplicados à Pandemia de Covid-19 e Versão Entrópica
Heterotípica da Teoria H**

Recife

2024

ARTHUR ARAÚJO BRUM

Modelos de Crescimento Aplicados à Pandemia de Covid-19 e Versão Entrópica Heterotípica da Teoria H

Tese apresentada ao Programa de Pós-Graduação em Física da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Doutor em Física.

Área de Concentração: Física Teórica e Computacional

Orientador: Antônio Murilo Santos Macêdo

Recife

2024

Catálogo na fonte
Bibliotecária: Luiza de Oliveira/CRB1316

B893m Brum, Arthur Araújo.
Modelos de crescimento aplicados à pandemia de Covid-19 e versão entrópica heterotípica da Teoria H /Arthur Araújo Brum.– 2024.
99 f.: il.

Orientador: Antônio Murilo Santos Macêdo.
Tese (Doutorado) – Universidade Federal de Pernambuco. Centro de Ciências Exatas e da Natureza. Programa de Pós-graduação em Física, Recife, 2024.
Inclui referências.

1 Mecânica estatística. 2. Sistemas complexos. 3. Entropia. 4. Modelos de crescimento. 5. Covid-19 I. Macêdo, Antônio Murilo Santos. II. Título

530.1

CDD (23. ed.)

UFPE - CCEN 2024 – 61

ARTHUR ARAÚJO BRUM

**MODELOS DE CRESCIMENTO APLICADOS À PANDEMIA DE COVID-19 E
VERSÃO ENTRÓPICA HETEROTÍPICA DA TEORIA H**

Tese apresentada ao Programa de Pós-Graduação em Física da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Física.

Aprovado em: 28/03/2024.

BANCA EXAMINADORA

Prof. Dr. Antônio Murilo Santos Macêdo
Orientador
Universidade Federal de Pernambuco

Prof. Dr. Andre Luiz Alves Lima
Examinador Interno
Universidade Federal de Pernambuco

Prof. Dr. Ernesto Carneiro Pessoa Raposo
Examinador Interno
Universidade Federal de Pernambuco

Prof. Dr. Americo Barbosa da Cunha Junior
Examinador Externo
Universidade do Estado do Rio de Janeiro

Prof. Dr. Marcus Werner Beims
Examinador Externo
Universidade Federal do Paraná

AGRADECIMENTOS

Agradeço em primeiro lugar aos meus pais, por me possibilitarem chegar até aqui; em especial à minha mãe, que me deu todo o suporte necessário para percorrer essa maratona que foi o Doutorado.

Aos meus mestres, Antônio Murilo e Giovani Vasconcelos, que me possibilitaram participar de um grupo de pesquisa extremamente dinâmico, ativo e com discussões e estudos de alto nível.

Aos amigos que fiz e mantive durante esse tempo.

Às pessoas que passaram pela minha vida e deixaram sua contribuição positiva.

Ao tempo, que me fez amadurecer e aprender, encarando os problemas e dificuldades como oportunidades para crescer.

RESUMO

Esta Tese apresenta estudos em duas áreas distintas da Física. No primeiro, analisamos um sistema hierárquico e computamos a entropia heterotípica da teoria H para as duas classes de universalidade resultantes da teoria; neste âmbito, mostramos que partindo de uma densidade de probabilidade conhecida, podemos encontrar a entropia para o sistema em questão. Estas entropias são denominadas heterotípicas, pois diferem da tradicional entropia de Boltzmann-Gibbs e são representadas de forma analítica em termos de funções G de Meijer. Ainda neste tema, discutimos a termodinâmica estocástica e mostramos sua conexão com a formulação entrópica heterotípica. No segundo, aplicamos modelos de crescimento generalizados às curvas epidêmicas de Covid-19 e realizamos ajustes numéricos, nos quais obtivemos excelentes resultados. Ainda neste tema, propusemos um modelo de crescimento com parâmetros dependentes do tempo para modelar curvas epidêmicas da Covid-19 com múltiplas ondas de infecções (ou mortes) e também neste caso mostramos através de excelentes ajustes numéricos que o modelo proposto descreve muito bem as curvas em questão. Finalmente, desenvolvemos um aplicativo web que aplica os modelos de crescimentos às curvas epidêmicas da Covid-19 para Países e Estados e Cidades no Brasil e nos Estados Unidos, através do qual o usuário pode monitorar a evolução da pandemia em uma dada localidade, obter informações relevantes sobre sua dinâmica e fazer previsões de curto prazo baseadas nos modelos de crescimento implementados.

Palavras-chave: mecânica estatística; sistemas complexos; entropia; modelos de crescimento; Covid-19.

ABSTRACT

This Thesis presents studies in two distinct areas of Physics. In the first one, we analyze a hierarchical system and compute the heterotypic entropy of the H-theory for the two universality classes resulting from this theory; in this scope, we show that starting from a known probability density, one can find the expression for the entropy of the system at hand. These entropies are called heterotypical, because they differ from the standard Boltzmann-Gibbs entropy and are represented analytically in terms of the Meijer-G functions. Still in this theme, we discuss the stochastic thermodynamics and show its connection with the entropic heterotypical formulation. In the second one, we apply generalized growth models to the epidemic curves of COVID-19 and perform numerical fits, in which we obtained excellent results. Still in that scope, we proposed a growth model with time dependent parameters to model COVID-19 epidemic curves with multiple waves of infections (or deaths) and also in that case we showed through great numerical fits that the model proposed describes very well the curves. We also studied the growth of works related to COVID-19 deposited in preprint repositories and showed that it can be described by a growth model previously applied to the epidemic curves; we also showed that the rank-frequency distribution of the works in the preprint repositories follows a Zipf-Law. Finally, we developed a web application that applies the growth models to the epidemic curves of COVID-19 to Countries, States and Cities of Brazil and the United States. through which the user can monitor the evolution of the pandemic in a given location, obtain relevant information about its dynamics and make short-term predictions based on the growth models implemented.

Keywords: statistical mechanics; complex systems; entropy; growth models; COVID-19.

LISTA DE FIGURAS

- Figura 1 – Comparativo entre o total de mortes causadas pelas diferentes epidemias e pandemias de doenças infecciosas ao longo da história ao redor do mundo em relação à população global estimada na época. 15
- Figura 2 – Ajuste numérico feito por Kermack e McKendrick utilizando um modelo SIR para descrever o número de mortes em decorrência de uma epidemia de Peste Bubônica na ilha de Bombay, na Índia, entre dezembro de 1905 e julho de 1906 em função do número de semanas desde a primeira morte. . . 16
- Figura 3 – Curvas de mortes acumuladas atribuídas à Covid-19 até 8 de maio de 2020 para (a) Itália, (b) Espanha, (c) França, (d) Alemanha, (e) Irã e (f) Coreia do Sul. Nas figuras, os círculos em vermelho representam o número acumulado de mortes no respectivo dia e a curva sólida preta representa o ajuste numérico com o modelo de Richards, cujos parâmetros do melhor fit estão indicados no *inset* das figuras, juntamente com seus erros. 25
- Figura 4 – Gráfico comparativo ilustrando o comportamento das curvas acumuladas geradas pelos diferentes modelos de crescimento no mesmo intervalo de tempo e com o mesmo tamanho final da epidemia, $K = 100$ 27
- Figura 5 – Gráfico comparativo ilustrando o comportamento das curvas diárias geradas pelos diferentes modelos de crescimento no mesmo intervalo de tempo e com o mesmo tamanho final da epidemia, $K = 100$ 28
- Figura 6 – Coluna da esquerda: Número acumulado de mortes (círculos vermelhos) em função do número de dias desde o primeiro registro de morte até a data de 31 de julho de 2020 para (a) Itália e (c) França e até a data de 07 de julho de 2020 para (e) Estados Unidos. Nesta coluna, as curvas pretas sólidas representam o melhor ajuste numérico utilizando o modelo logístico beta. Coluna da direita: Dados diários de mortes (círculos vermelhos conectados pelas linhas vermelhas) para os mesmos países e até as mesmas datas, respectivamente. As curvas pretas neste caso representam a derivada da curva acumulada resultante do ajuste numérico com o modelo logístico beta; no *inset* de cada figura é exibido um plot em escala log-log para ressaltar o comportamento assintótico de lei de potência. 30

Figura 7 – Evolução temporal de um parâmetro genérico do modelo $\zeta(t)$ entre dois valores ζ_1 e ζ_2 com tempo de transição t_1 e taxa de transição ρ_1 . A curva tracejada representa uma aproximação linear da curva sólida.	33
Figura 8 – Variação do formato da curva $\zeta(t)$ para valores fixos de t_1 , ζ_1 e ζ_2 e diferentes valores de ρ_1 , ilustrando transições mais ou menos suaves a depender do valor do parâmetro.	33
Figura 9 – Evolução temporal de $\zeta(t)$ para o caso ilustrativo de 4 ondas epidêmicas com os seguintes parâmetros: $\zeta_1 = 0,4$, $\zeta_2 = 0,8$, $\zeta_3 = 0,6$, $\zeta_4 = 0,3$, $t_1 = 100$, $t_2 = 250$, $t_3 = 400$, $\rho_1 = 0,12$, $\rho_2 = 0,1$ e $\rho_3 = 0,15$	34
Figura 10 – Coluna da esquerda: número acumulado de mortes (círculos vermelhos) atribuídos à Covid-19 para (a) Canadá e (c) África do Sul até 3 de abril de 2021. A linha preta representa o melhor ajuste numérico com o modelo logístico beta de duas ondas e o ponto preto na curva representa o tempo t_1 em que ocorre a transição entre as duas ondas; ao lado de cada gráfico são exibidos os parâmetros encontrados pelo ajuste. Coluna da direita: número diário de mortes para os mesmos países (círculos vermelhos ligados por segmentos de reta). A curva preta representa a derivada temporal do modelo logístico beta com os parâmetros encontrados pelo melhor ajuste numérico.	36
Figura 11 – Coluna da esquerda: número acumulado de mortes (círculos vermelhos) atribuídos à Covid-19 para (a) Japão e (c) Estados Unidos até 3 de abril de 2021. A linha preta representa o melhor ajuste numérico com o modelo logístico beta de duas ondas e o ponto preto na curva representa o tempo t_1 em que ocorre a transição entre as duas ondas; ao lado de cada gráfico são exibidos os parâmetros encontrados pelo ajuste. Coluna da direita: número diário de mortes para os mesmos países (círculos vermelhos ligados por segmentos de reta). A curva preta representa a derivada temporal do modelo logístico beta com os parâmetros encontrados pelo melhor ajuste numérico.	37
Figura 12 – Comparação das curvas epidêmicas de três integrações numéricas para o modelo SIR com taxa de transmissão $\beta = 0,14$, $\beta = 0,155$ e $\beta = 0,17$ mantendo o parâmetro $\gamma = 0,1$ fixo.	42

Figura 13 – Comparação das curvas epidêmicas de três integrações numéricas para o modelo SIRD com taxa de transmissão $\beta = 0,15$ e taxa de recuperação $\gamma = 0,06$ mantidas fixas e taxa de mortalidade variando nos valores $\delta = 0,04$, $\delta = 0,02$ e $\delta = 0,01$	45
Figura 14 – O gráfico da esquerda ilustra a curva de mortes acumuladas $D(t)$ de uma realização do modelo SIRD (curva contínua em vermelho) juntamente com a curva de mortes acumuladas $C(t)$ do modelo de Richards (curva tracejada em preto) com parâmetros obtidos a partir do mapa entre os dois modelos. No gráfico da direita é apresentado um plot da quantidade $\eta(t)$ para este mapa.	48
Figura 15 – Comportamento dos parâmetros do modelo de Richards r (curva contínua) e α (curva tracejada), obtidos através do mapa, e do taxa básica de reprodução R_0 (curva pontilhada) em função do parâmetro β do modelo SIRD com γ e δ fixos.	49
Figura 16 – Curva de mortes acumuladas atribuídas à Covid-19 (círculos vermelhos) para a Suécia até 30 de Julho de 2020 juntamente com a curva de mortes acumuladas do modelo SIRD padrão com parâmetros obtidos através do melhor ajuste numéricos, cujas estimativas e erros são exibidos no inset da figura.	50
Figura 17 – Número acumulado de mortes (círculos vermelhos) atribuídos à Covid-19 até 30 de Julho de 2020 para (a) Holanda, (b) Itália, (c) Alemanha, (d) Cuba, (e) Suécia, e (f) Japão. As curvas contínuas em preto representam o melhor ajuste numérico pelo modelo SIRD modificado com parâmetro β dependente do tempo e cujos valores estimados dos parâmetros estão descritos na legenda da figura. As curvas tracejadas em verde representam a curva teórica obtida pelo modelo de Richards com os parâmetros computados numericamente através do mapa dado pelas equações (2.42) e (2.43) e os parâmetros do modelo SIRD encontrados pelo ajuste numérico.	51
Figura 18 – Figura esquemática de uma curva epidêmica apresentando duas ondas de infecções. As linhas verticais coloridas representam os estágios epidêmicos; o ponto e a linha vertical em preto representam o início da segunda onda epidêmica.	57

Figura 19 – Figura esquemática da velocidade (primeira derivada) da curva acumulada mostrada na figura 18.	58
Figura 20 – Figura esquemática da aceleração (segunda derivada) da curva acumulada mostrada na figura 18.	59
Figura 21 – Figura esquemática descrevendo o funcionamento do aplicativo ModInterv desde sua inicialização até o ajuste numérico.	60
Figura 22 – Ilustrações esquemáticas de quatro dos cinco <i>widgets</i> do módulo <i>iPywidgets</i> usados no aplicativo.	61
Figura 23 – Curva acumulada de mortes (círculos vermelhos) atribuídas à Covid-19 para a Holanda até o dia 09/03/2023.	69
Figura 24 – Ajuste numérico para a Holanda utilizando todos os dados disponíveis na base de dados, o que equivale à data 09/03/2023. (a) Curva epidêmica acumulada de óbitos (círculos vermelhos) sobreposta com a curva acumulada teórica (linha contínua na cor preta). (b) Curva epidêmica da óbitos diários (círculos vermelhos) sobreposta com a curva diária teórica (linha contínua na cor preta).	70
Figura 25 – Ajuste numérico para a Holanda utilizando parte dos dados disponíveis na base de dados, com o último ponto da curva correspondendo à data 19/06/2021. (a) Curva epidêmica acumulada de óbitos (círculos vermelhos) sobreposta com a curva acumulada teórica (linha contínua na cor preta). (b) Curva epidêmica da óbitos diários (círculos vermelhos) sobreposta com a curva diária teórica (linha contínua na cor preta).	72
Figura 26 – Ajuste numérico para o estado de <i>Wyoming</i> nos Estados Unidos utilizando todos os dados disponíveis na base de dados, com o último ponto da curva correspondendo à data 09/03/2023. (a) Curva epidêmica acumulada de óbitos (círculos vermelhos) sobreposta com a curva acumulada teórica (linha contínua na cor preta). (b) Curva epidêmica da óbitos diários (círculos vermelhos) sobreposta com a curva diária teórica (linha contínua na cor preta).	73

Figura 27 – Ajuste numérico para o estado de Pernambuco no Brasil utilizando todos os dados disponíveis na base de dados, com o último ponto da curva correspondendo à data 18/03/2023. (a) Curva epidêmica acumulada de óbitos (círculos vermelhos) sobreposta com a curva acumulada teórica (linha contínua na cor preta). (b) Curva epidêmica da óbitos diários (círculos vermelhos) sobreposta com a curva diária teórica (linha contínua na cor preta).	74
Figura 28 – Plot paramétrico de $s(p)$ para N variando de 1 a 5 e valores de parâmetros $a = 1$ e ω e ε_0 foram escolhidos para cada caso de forma que o quociente $\omega/2\varepsilon_0 = 1$	83
Figura 29 – Plot paramétrico de $s(p)$ para N variando de 1 a 5 e valores de parâmetros $a = 1$ e ω e ε_0 foram escolhidos para cada caso de forma que o produto $2\omega\varepsilon_0 = 1$	85
Figura 30 – Comparação entre as curvas paramétricas $s(p)$ para as classes de universalidade Exponencial Esticada e Lei de Potência com N variando de 1 a 5, $a = 1$ e parâmetros ω e ε_0 escolhidos de tal forma que o termo que acompanha a variável E nas equações (4.29) e (4.41) sejam sempre iguais a unidade.	86
Figura 31 – Plot paramétrico de $\Delta(p)$ para as diferentes classes de universalidade com N variando de 1 a 5.	90

SUMÁRIO

1	INTRODUÇÃO	14
2	MODELAGEM MATEMÁTICA DA PANDEMIA DE COVID-19	20
2.1	MODELOS DE CRESCIMENTO LOGÍSTICO	21
2.1.1	Modelo de Richards	21
2.1.1.1	<i>Dados Analisados</i>	22
2.1.1.2	<i>Ajustes numéricos</i>	23
2.1.2	Modelo Logístico Beta	24
2.1.2.1	<i>Ajustes Numéricos</i>	29
2.2	MODELOS DE CRESCIMENTO COM MÚLTIPLAS ONDAS	32
2.2.1	Ajustes Numéricos	35
2.3	MODELOS COMPARTIMENTAIS	39
2.3.1	Modelo SIR	40
2.3.2	A taxa básica de reprodução R_0: determinando se há ou não uma epidemia	43
2.3.3	Modelo SIRD	44
2.3.4	Mapa entre os modelos SIRD e Richards	46
2.3.5	Modelo SIRD com parâmetros dependentes do tempo	48
3	MODINTERV: UMA FERRAMENTA ONLINE PARA MONITORAMENTO DA COVID-19	54
3.1	ESTÁGIOS DE UMA CURVA EPIDÊMICA	56
3.2	ESTRUTURA DE <i>BACKEND</i> DO APLICATIVO MODINTERV	60
3.2.1	Interface do usuário	60
3.2.2	Aquisição de dados	61
3.2.3	Detecção de ondas epidêmicas	62
3.2.4	Ajustes numéricos	63
3.2.5	Classificação do estágio epidêmico	65
3.2.6	Interatividade do aplicativo e hospedagem	66
3.3	ESTRUTURA DE INTERFACE DO USUÁRIO DO APLICATIVO MODINTERV	66
3.3.1	Países	67

3.3.2	Cidades e Estados do Brasil e dos Estados Unidos	68
3.3.3	Exemplos	69
4	VERSÃO ENTRÓPICA HETEROTÍPICA DA TEORIA H	75
4.1	UMA BREVE INTRODUÇÃO À TEORIA-H	76
4.2	FUNCIONAL ENTRÓPICO A PARTIR DE UMA DISTRIBUIÇÃO DE PRO- BABILIDADE	77
4.2.1	Classe de Universalidade Exponencial Esticada	78
4.2.1.1	<i>O Ensemble de Boltzmann-Gibbs ($N=0$)</i>	79
4.2.1.2	<i>O Ensemble K ($N=1$)</i>	80
4.2.1.3	<i>Caso geral ($N>1$)</i>	81
4.2.2	Classe de Universalidade Lei de Potência	82
4.2.2.1	<i>Ensemble Gaussiano Extendido ($N=1$)</i>	83
4.2.3	Caso geral ($N>1$)	84
4.3	TERMODINÂMICA ESTOCÁSTICA	86
5	CONCLUSÕES E PERSPECTIVAS	91
5.1	CONCLUSÕES	91
5.2	PERSPECTIVAS	92
	REFERÊNCIAS	94

1 INTRODUÇÃO

Desde os primórdios da civilização, a humanidade sofre com doenças dos mais variados tipos e transmitidas por uma variedade de patógenos, como a peste bubônica que assolou a Europa em meados do século XIV e estima-se que tenha dizimado entre 50% e 60% da população daquela região e causou uma onda de terror generalizada durante os seus 6 anos de duração (BENEDICTOW, 2006); as duas pandemias de cólera que afetaram diversos continentes e ao todo vitimaram aproximadamente 150 mil pessoas nos Estados Unidos e 200 mil pessoas no México, ou ainda a pandemia de gripe espanhola, que surgiu no rastro de destruição da primeira guerra mundial e matou por volta de 50 milhões de pessoas no total ao longo de toda sua duração (JOHNSON; MUELLER, 2002). A figura 1 mostra uma lista com as maiores pandemias que o mundo já enfrentou e o respectivo número de mortes causadas por cada uma delas em relação à população mundial daquela época. Como é possível notar, a Peste Negra (*Black Death* na figura) foi de longe a maior pandemia presenciada pela humanidade até agora.

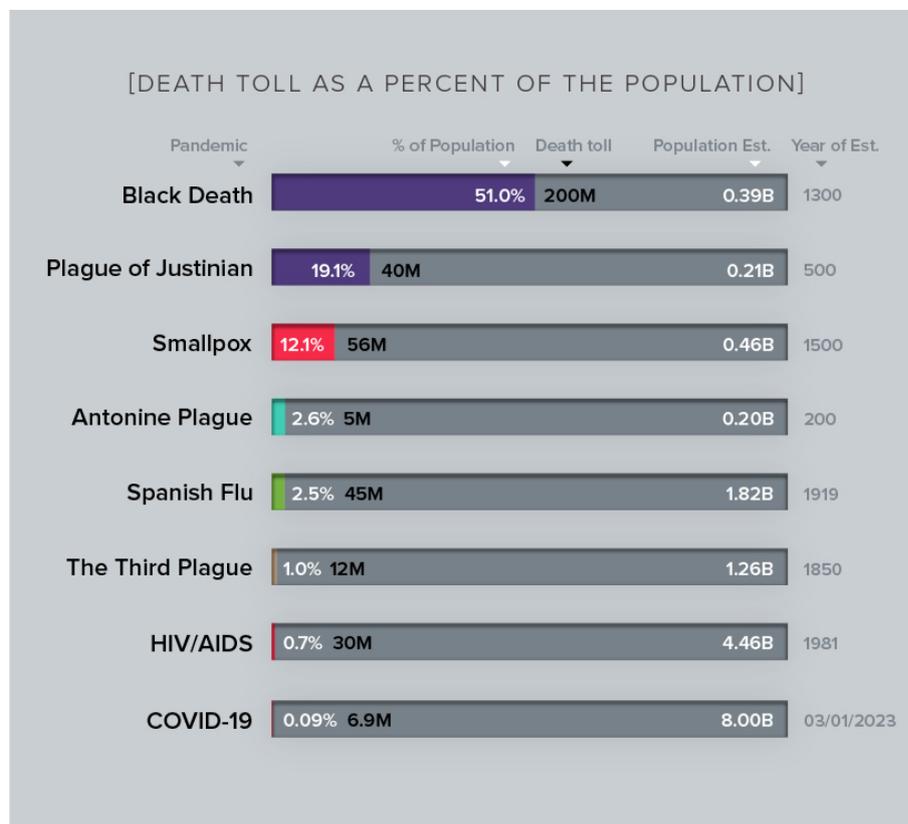
Por conta dessas e de tantas outras epidemias vivenciadas pela humanidade surgiu a necessidade de descrever e modelar as epidemias a fim de fazer previsões e suprimir os danos causados por seu espalhamento desenfreado. O conhecimento da dinâmica assumida pela disseminação de uma doença é peça fundamental na implementação de medidas preventivas ou de mitigação, como vacinação, fechamento de escolas, isolamento social, uso de máscaras e sanitizante para mãos.

O trabalho de Daniel Bernoulli (BERNOULLI, 1760; DIETZ; HEESTERBEEK, 2002), publicado em 1760 e sucedido por uma versão expandida e mais detalhada publicada em 1766 (BERNOULLI, 1766), é considerado o primeiro modelo de epidemiologia matemática. Em seu trabalho, o matemático suíço analisou a eficácia da inoculação de pequenas quantidades do vírus para o aumento da expectativa de vida da população. O próximo grande passo na epidemiologia matemática foi dado com o trabalho do médico londrino e ganhador do prêmio Nobel de medicina do ano de 1902, Sir Ronald Ross (ROSS, 1911). Em seu trabalho Ross mostrou que, ao contrário do que se pensava na época, seria possível eliminar a malária de uma dada população sem a necessidade de eliminar por completo os mosquitos, mas sim reduzindo seu número abaixo de um certo limiar. Para isto, Ross utilizou um modelo compartimental que incluía humanos e mosquitos e analisou a dinâmica da transmissão de tal doença para a po-

pulação; ainda nesse trabalho Ross introduziu o conceito de taxa básica de reprodução, que tem ocupado papel central na epidemiologia matemática desde então.

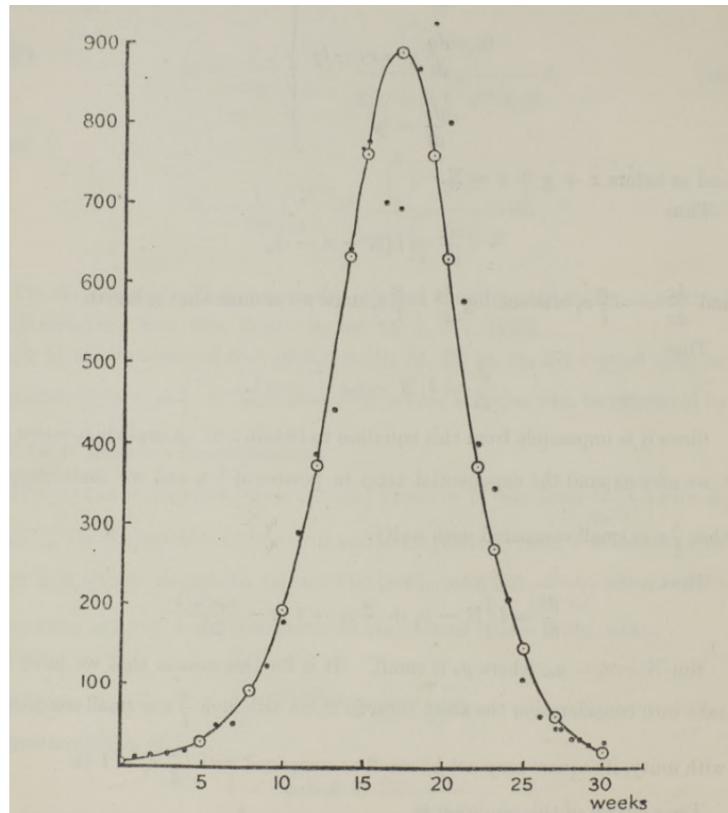
Os modelos compartimentais chegaram na forma que são utilizados na atualidade com o extenso trabalho de Kermack e McKendrick publicado ao longo de 6 anos em um total de 3 artigos (KERMACK; MCKENDRICK, 1927; KERMACK; MCKENDRICK, 1932; KERMACK; MCKENDRICK, 1933). Em seu estudo eles apresentaram um modelo que contava com três compartimentos, a saber: Suscetíveis, Infecciosos e Removidos (que incluía mortos e recuperados da doença); esta versão do modelo SIR tornou-se bastante popular e teve uma massificada aplicação para as mais diversas doenças e ganhou, ao longo dos anos, vários desdobramentos com a inclusão de novos compartimentos, gerando modelos altamente personalizáveis e complexos. A figura 2 mostra um ajuste numérico feito por Kermack e McKendrick utilizando seu modelo compartimental para descrever o número de mortes causadas pela Peste Bubônica em uma epidemia na ilha de Bombay na Índia entre dezembro de 1905 e julho de 1906 em função do número de semanas desde a primeira morte.

Figura 1 – Comparativo entre o total de mortes causadas pelas diferentes epidemias e pandemias de doenças infecciosas ao longo da história ao redor do mundo em relação à população global estimada na época.



Fonte:(LEPAN, 2020).

Figura 2 – Ajuste numérico feito por Kermack e McKendrick utilizando um modelo SIR para descrever o número de mortes em decorrência de uma epidemia de Peste Bubônica na ilha de Bombay, na Índia, entre dezembro de 1905 e julho de 1906 em função do número de semanas desde a primeira morte.



Fonte:(KERMACK; MCKENDRICK, 1927).

A epidemiologia matemática foi mais uma vez posta a prova quando no início de 2020 o mundo se deparou com o que viria a ser uma das maiores crises sanitárias de todos os tempos. A doença causada pelo vírus SARS-CoV-2 surpreendeu a humanidade com seu poder destrutivo e fez, ao longo de mais de 3 anos, quase 7 milhões de vítimas segundo a Organização Mundial da Saúde (OMS). Inicialmente detectada em Wuhan, na China, no final de 2019, a Covid-19 apresentou-se como uma Síndrome Respiratória Aguda Grave (SRAG) na maioria dos pacientes infectados. Dentre os sintomas mais comuns estavam tosse, febre, dores no corpo e dispneia (dificuldade para respirar). No entanto, alguns pacientes evoluíram para quadros de saúde mais sérios, apresentando falta de ar e baixos níveis de saturação de oxigênio, o que fez com que muitos pacientes necessitassem de acompanhamento médico em Unidades de Tratamento Intensivo (UTIs) e intubação para realização de respiração por meio de aparelhos.

Logo nos primeiros meses a comunidade científica engajou-se nas mais diversas áreas, atuando no sequenciamento genético do vírus e de suas variantes, no teste de possíveis medicamentos para prevenir e amenizar os sintomas causados pela doença, no desenvolvimento de

vacinas, ou ainda na descrição da dinâmica e evolução da epidemia nas diversas localidades, considerando uma variedade de fatores coadjuvantes como clima, localização geográfica, índice de desenvolvimento humano, entre tantos outros.

Ainda nesta última área, foram utilizados os mais diversos modelos, como os modelos baseados em agentes (KERR et al., 2021) descrevendo a transmissão do vírus entre indivíduos, os modelos compartimentais tradicionais como SIR e SIRD (MACÊDO et al., 2021) e também versões modificadas e estendidas destes, e os modelos fenomenológicos, como os modelos de crescimento logístico (VASCONCELOS et al., 2020) que descrevem a evolução da epidemia em termos de parâmetros simples, porém sem uma clara interpretação epidemiológica.

A presente tese aborda duas das classes de modelos mencionadas anteriormente, a saber: modelos compartimentais e modelos fenomenológicos. Os modelos compartimentais possuem uma clara interpretação epidemiológica e têm vasta aplicação e abordagem na literatura da epidemiologia matemática desde sua consolidação e popularização com os trabalhos de Kermack e McKendrick. No entanto, sua aplicação encontra alguns desafios pois necessita de dados confiáveis e bem estimados e coletados para cada um de seus compartimentos para o bom funcionamento do modelo. A pandemia de Covid-19 se mostrou especialmente desafiadora para tais modelos devido à subnotificação de casos confirmados e de pacientes recuperados da doença, restando apenas os dados de mortes em decorrência da infecção como fonte confiável de informação, o que gera consideráveis problemas na realização de ajustes numéricos com o modelo.

Os modelos de crescimento logístico, de certa maneira desprovidos de clara interpretação epidemiológica para seus parâmetros, trazem em contrapartida a simplicidade computacional e matemática, muitas vezes podendo ser descritos em termos de expressões analíticas. Além da simplicidade, os modelos fenomenológicos não dependem dos diversos compartimentos epidemiológicos como os modelos compartimentais, tornando possível sua aplicação com excelentes resultados a curva epidêmicas das mais diversas localidades e com diferentes graus de complexidade. Notamos ainda que é possível construir um mapa entre os modelos compartimentais e os modelos de crescimento logístico, trazendo significado epidemiológico para os parâmetros destes últimos.

Utilizando os modelos fenomenológicos discutidos ao longo deste trabalho, foi implementado um software online chamado ModInterv (BRUM et al., 2023; BRUM et al., 2022), que disponibiliza para pesquisadores, autoridades de saúde pública e o público em geral uma ferramenta de fácil acesso para o acompanhamento e realização de algumas previsões de curto

prazo sobre a evolução das curvas epidêmicas em uma dada localidade escolhida pelo usuário. O software está disponível para acesso no endereço <<https://fisica.ufpr.br/modinterv>> e através dele podem ser consultadas as curvas epidêmicas (curvas acumuladas de casos confirmados ou mortes) para todos os países, todas as cidades e estados do Brasil e dos Estados Unidos. Além disso os modelos de crescimento logístico podem ser aplicados a estas curvas para analisar e determinar o atual estágio epidêmico em que se encontra tal localidade.

Ainda nesta tese é apresentado um trabalho desenvolvido na área de mecânica estatística que aborda um tema central tanto na física quanto na nossa vida cotidiana: a entropia. Esta grandeza que é considerada como um dos conceitos mais difíceis de se compreender é uma velha conhecida da humanidade e está presente em todos os aspectos da vida, desde a sensação de frio ou de calor, fome ou saciedade, até a expansão de um gás em um ambiente sem barreiras. Apesar de só ter sido formalmente definida em 1865 por Clausius (CLAUSIUS, 1865), a entropia já estava presente nas discussões de Carnot e na sua versão inicial das leis da termodinâmica em 1824.

A entropia encontrou aplicação nas mais diversas áreas, passando pela teoria dos jogos, economia e até mesmo cosmologia, sempre guiando a dinâmica dos processos, servindo como uma "seta do tempo" (BLUM, 1968) para os sistemas. Nesta tese abordamos o aspecto heterotípico da entropia para sistemas hierárquicos através da Teoria-H, uma teoria que é aplicada na descrição de sistemas hierárquicos com múltiplas escalas (de comprimento, tempo, etc) em que há fluxo de energia entre elas, mostrando como é possível obter um funcional entrópico partindo de uma distribuição de probabilidade.

A sequência da presente tese está dividida da seguinte forma: no capítulo 2 são discutidos em detalhes os modelos compartimentais e fenomenológicos. Para cada uma destas classes de modelos é apresentada uma introdução histórica mostrando o desenvolvimento de cada um deles desde sua aparição na literatura até as versões mais modernas e utilizadas nos estudos discutidos. Além disso, são exibidos ajustes numéricos que mostram a excelente concordância entre os modelos e as curvas epidêmicas empíricas.

Em seguida, no capítulo 3, é discutida a implementação do software ModInterv, um aplicativo que implementa os modelos de crescimento logístico discutidos no capítulo 2 e que disponibiliza amplamente e de forma acessível para qualquer público uma forma de monitorar a evolução da pandemia de Covid-19 em todos os países, nas cidades e estados do Brasil e dos Estados Unidos.

No capítulo 4 é elaborada a versão entrópica heterotípica da Teoria-H. Neste capítulo é

mostrado como se obter um funcional entrópico a partir de uma distribuição de probabilidade e este procedimento é aplicado para as duas classes de universalidade previstas pela Teoria-H para um número arbitrário N de escalas. Além disso, é feita a conexão com a teoria da Termodinâmica estocástica, ressaltando-se o papel fundamental da entropia na dinâmica dos sistemas fora do equilíbrio.

A presente tese é encerrada com a apresentação das conclusões e discussão sobre as perspectivas no capítulo 5.

2 MODELAGEM MATEMÁTICA DA PANDEMIA DE COVID-19

A pandemia de Covid-19 se consolidou como uma das crises sanitárias mais graves da história, tendo vitimado desde sua primeira fatalidade um total de quase 7 milhões de pessoas ao redor do mundo. Tendo seu fim decretado oficialmente pela Organização das Nações Unidas em 5 de maio de 2023, a emergência global de saúde causada pela Covid-19 mostrou a importância de entender a evolução de uma doença na população para tomar decisões acerca de medidas de controle como *lockdown*, fechamento de escolas, uso obrigatório de máscaras.

Durante a referida pandemia, as medidas de controle tiveram papel fundamental na diminuição da velocidade de espalhamento da doença, evitando assim centenas de milhares (ou até milhões) de mortes. Entretanto, a adoção destas medidas deve ser feita baseada em dados científicos que consigam capturar o panorama geral e dar indicativos sobre a situação tende a evoluir. Isto é obtido através de modelos matemáticos bem estabelecidos na literatura científica e quando aplicados aos dados empíricos, conseguem capturar características importantes da evolução da pandemia em uma dada localidade através dos seus parâmetros.

Através da aplicação destes modelos aos dados empíricos de uma determinada localidade é possível fazer previsões e tomar decisões bem embasadas. Tais medidas ajudam a diminuir o espalhamento da doença, a manter o sistema de saúde funcionando em níveis aceitáveis e a salvar milhares de vidas (DAVIES et al., 2020; DAVIES et al., 2021).

Neste capítulo vamos apresentar uma breve revisão dos modelos matemáticos para a modelagem de epidemias. Estes modelos foram utilizados de forma bem sucedida para a modelagem da pandemia da Covid-19 em vários países, estados e cidades.

Nas seções 2.1 e 2.2 apresentaremos uma coleção de modelos de crescimento logístico com diferentes níveis de complexidade, desde o modelo logístico simples com apenas 2 parâmetros até o mais complexo com 5 parâmetros e sua extrapolação com parâmetros dependentes do tempo para a acomodação das várias ondas epidêmicas.

Na seção 2.3 revisitamos os modelos compartimentais, considerando desde o modelo mais simples, com apenas 2 compartimentos, até o modelo com 4 compartimentos e parâmetros dependentes do tempo para novamente acomodar diferentes ondas epidêmicas.

2.1 MODELOS DE CRESCIMENTO LOGÍSTICO

Os modelos de crescimento logístico têm sido usados para modelagem de problemas e sistemas há quase 200 anos, desde quando o matemático belga Pierre François Verhulst publicou uma série de 3 artigos entre 1838 e 1847 introduzindo o modelo logístico como ferramenta matemática para modelar o crescimento populacional (VERHULST, 1838; VERHULST, 1845; VERHULST, 1847). Este modelo sofreu várias generalizações ao longo dos anos, aumentando cada vez mais a sua complexidade e expandindo a gama de fenômenos que podem ser descritos por ele através da adição de parâmetros que modificam seu comportamento assintótico. Estas modificações e generalizações serão abordadas em detalhes, juntamente com aplicações a modelagem da pandemia de Covid-19, ao longo desta seção.

2.1.1 Modelo de Richards

Uma das primeiras versões modificadas do modelo logístico original foi introduzida em 1959 pelo botânico inglês Francis John Richards (RICHARDS, 1959) como forma de auxiliar os ajustes numéricos em suas análises experimentais. O modelo ou curva de Richards, como ficou conhecido, é definido pela equação diferencial ordinária. A equação diferencial ordinária que descreve tal modelo é dada por (WANG; WU; YANG, 2012):

$$\frac{dC}{dt} = rC(t) \left[1 - \left(\frac{C(t)}{K} \right)^\alpha \right], \quad (2.1)$$

onde $C(t)$ é o número de casos ou mortes acumuladas em uma dada localidade em um tempo t , contado em dias a partir do primeiro caso ou da primeira morte, K é o tamanho final da epidemia, r é a taxa de crescimento na fase inicial da evolução e α é um parâmetro que controla a assimetria da curva em relação ao formato de S simétrico inerente ao modelo logístico padrão, caso que é recuperado quando $\alpha = 1$. A equação anterior deve ser suplementada com uma condição de contorno. Levando em consideração a forma empírica das curvas de mortes acumuladas, que apresentam um ponto de inflexão e um platô de saturação, foi escolhida a condição de contorno:

$$d^2C(t = t_c)/dt^2 = 0, \quad (2.2)$$

para um dado instante t_c que marca o ponto de inflexão da curva acumulada. Integrando diretamente a eq. (2.1) com a condição (2.2), temos:

$$C(t) = \frac{K}{\{1 + \alpha \exp[-\alpha r (t - t_c)]\}^{1/\alpha}}. \quad (2.3)$$

Para os ajustes numéricos realizados nos trabalhos apresentados nesta tese, escolhemos fazer a identificação $C(0) = C_0$, onde C_0 é o número de mortes registradas no primeiro dia em que uma morte foi reportada, e pode ser explicitamente expressa como

$$C_0 = \frac{K}{[1 + \alpha \exp(\alpha r t_c)]^{1/\alpha}} \quad (2.4)$$

e podemos eliminar t_c em detrimento dos outros parâmetros, de forma que ficamos apenas com 3 parâmetros livres, (α, r, K) , a serem determinados numericamente.

Antes de prosseguirmos para a aplicação do Modelo de Richards aos dados reais da pandemia de Covid-19, vale a pena ressaltar que este modelo não é totalmente confiável para curvas que estão muito abaixo do ponto de inflexão, t_c . Nestes casos, a pandemia está em um regime de crescimento exponencial e é mais confiável usar o modelo de crescimento generalizado (WU et al., 2020; CHOWELL, 2017), dado pela equação

$$C(t) = A^{1/(1-q)} \exp_q(rt/A), \quad (2.5)$$

onde $0 < q < 1$ é um parâmetro que interpola o regime de crescimento entre exponencial, quando $q = 1$, e subexponencial, quando $q = 0$. A função $\exp_q(x) = [1 + (1 - q)x]^{1/(1-q)}$ é conhecida na literatura como a função q-exponencial (GASPER; RAHMAN, 2004). Finalmente, o parâmetro A pode ser determinado a partir da condição inicial $C(0) = A^{1/(1-q)}$.

Nesta subseção, nos ateremos apenas à aplicação do Modelo de Richards aos dados da pandemia de Covid-19, como discutido em (VASCONCELOS et al., 2020), deixando de lado os casos em que o modelo de crescimento generalizado se faz necessário.

2.1.1.1 Dados Analisados

Desde o início da pandemia, um fato tornava as estatísticas de casos de infecções pelo vírus SARS-CoV-2, e consequente desenvolvimento da Covid-19, pouco confiáveis: a subnotificação. Este problema aconteceu durante toda a pandemia devido a grande parte da população, tanto no início da pandemia, quanto durante sua ampla disseminação, não ter sido submetida a testes, o que faz com que a diferença entre o número de casos confirmados e o número real de

infecções na população ser subjetiva em relação à política de testagem de cada localidade. Um outro fator que corroborou com a subnotificação foi o fato de que muitas pessoas desenvolviam sintomas muito brandos, tais como o de uma gripe ou resfriado, ou sequer apresentavam sintomas, e isto fazia com que uma grande parcela da população optasse por não se submeter aos testes e muitas infecções não fossem detectadas (LI et al., 2020).

Ao contrário do panorama de casos confirmados, os dados de mortes causadas pela Covid-19 é uma estimativa mais confiável do avanço da pandemia. Esta classe de dados ainda carrega consigo alguma incerteza devido às políticas e protocolos de notificação de mortes, atrasos nas notificações e na confirmação da causa de morte por Covid-19, entre outros. Ainda assim, os dados de associados a mortes são mais confiáveis para estimar a evolução da pandemia em comparação aos de casos confirmados e, portanto, ao longo de toda a seção 2.1 os modelos serão aplicados às curvas de mortes acumuladas.

Todos os dados utilizados nas análises realizadas em (VASCONCELOS et al., 2020) e aqui reproduzidas foram obtidos no repositório do Centro de Recursos do Coronavírus da Universidade John Hopkins (JHU, 2020), que disponibiliza os dados acumulados de casos confirmados, mortes e recuperados de forma automática e atualizada diariamente.

2.1.1.2 Ajustes numéricos

Os ajustes numéricos com o modelo de Richards consistiram em resolver o problema de otimização usando o algoritmo de Levenberg-Marquardt (MORÉ, 1978) para resolver o problema de mínimos quadrados não linear com os devidos parâmetros do modelo. O modelo de Richards, bem como suas variantes, é conhecido por ser suscetível ao problema de *overfitting* e seu ajuste numérico pode portanto apresentar parâmetros fora dos limites aceitáveis. Para evitar este tipo de problema, nos ajustes numéricos os parâmetros r e α foram limitados ao intervalo $(0, 1)$.

Como já foi comentado anteriormente, o modelo de Richards não é totalmente confiável quando as curvas em questão estão muito abaixo do ponto de inflexão, t_c , e portanto, foram considerados apenas os casos em que o parâmetro t_c estimado pelo ajuste numérico foi menor que a data do último dado empírico para a curva em questão.

No trabalho em questão, foram selecionados 6 países, a saber: Itália, Espanha, França, Alemanha, Irã e Coreia do Sul. Em todos os casos foram considerados dados de mortes acumuladas até 8 de maio de 2020. Nos gráficos apresentados na 3, os círculos vermelhos representam

o número de mortes acumuladas até o respectivo dia, contado a partir do dia em que houve o primeiro registro de mortes associadas à Covid-19, a curva preta sólida representa o ajuste numérico feito com o modelo de Richards e no *inset* de cada gráfico estão indicados os valores dos parâmetros encontrados no melhor ajuste, juntamente com seus erros numéricos, calculados através da matriz de covariância.

Como podemos ver, através dos ajustes numéricos, o modelo de Richards consegue capturar e descrever muito bem o comportamento da curva epidêmica em questão para todos os 6 países. No entanto, no caso da Itália, Espanha e Irã, a porção da curva que se aproxima ao platô de saturação (tamanho final da epidemia) fica um pouco fora dos dados empíricos. Isto se dá ao fato de que o modelo de Richards prevê uma aproximação exponencial ao platô, enquanto nos dados este fenômeno acontece de forma mais lenta. Como será discutido posteriormente, a aproximação exponencial ao platô de saturação está diretamente relacionada com a implementação de medidas de controle, como o uso obrigatório de máscaras, fechamento de escolas, quarentena compulsória, *lockdown*, entre outras.

2.1.2 Modelo Logístico Beta

Como visto na seção 2.1.1.2, o modelo de Richards desempenha um ótimo trabalho em descrever as curvas epidêmicas que já passaram do ponto de inflexão t_c . No entanto, na fase inicial da epidemia, ou seja, quando $C(t) \ll K$, a Equação Diferencial Ordinária que define o modelo de Richards fica

$$\frac{dC}{dt} = rC(t), \quad (2.6)$$

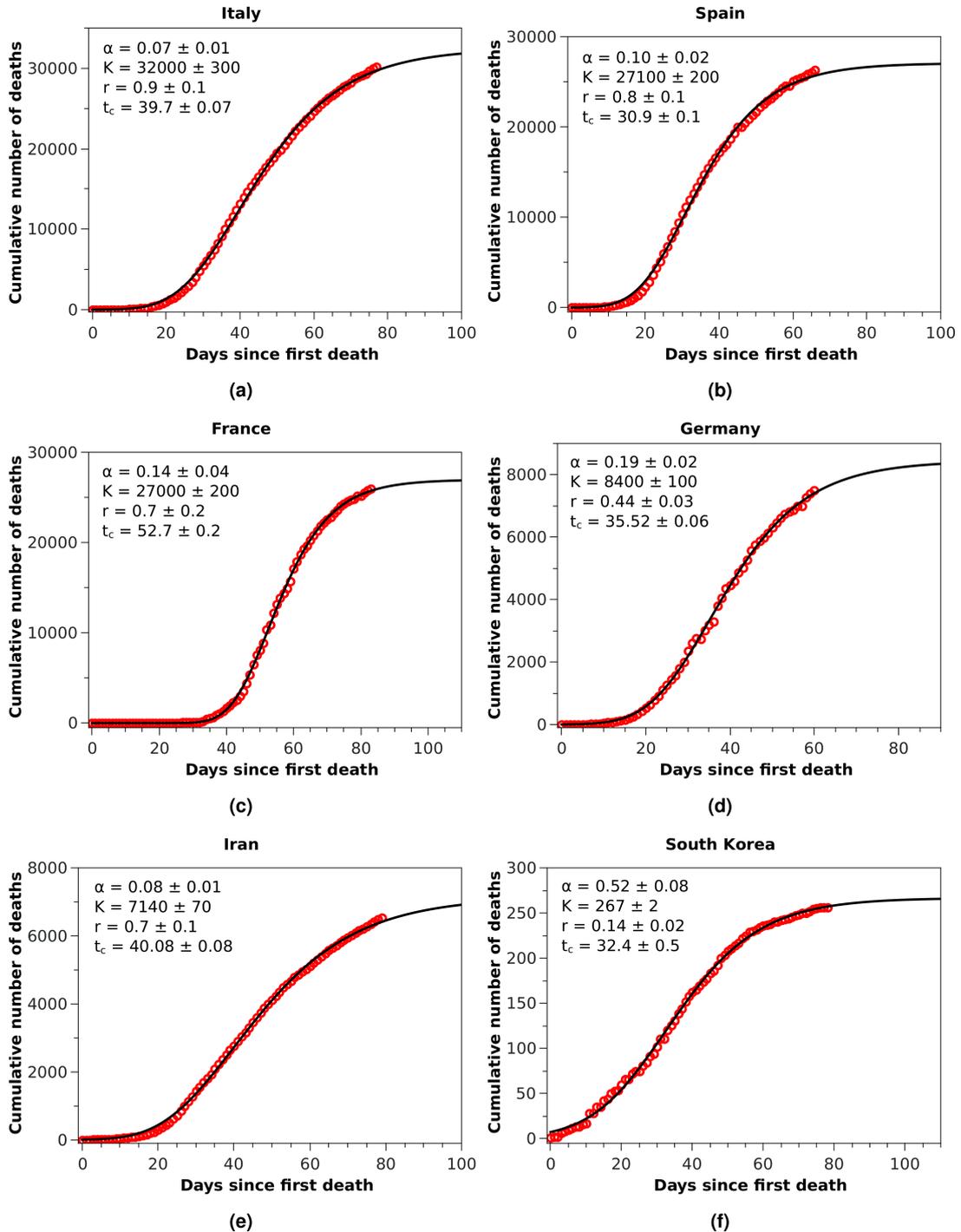
que após integração direta resulta em uma curva exponencial. Portanto, o modelo de Richards prevê sempre um crescimento exponencial na fase inicial da epidemia. Além disso, o modelo de Richards também apresenta, como falado anteriormente, uma aproximação exponencialmente rápida ao platô de saturação, como pode ser mostrado ao introduzirmos em 2.1 no regime assintótico de grandes tempos — caso em que $C(t) \rightarrow K$ — a variável $\varepsilon(t) = K - C(t)$, tal que $\varepsilon \ll 1$. Neste caso, a equação 2.1 se torna

$$-\frac{d\varepsilon}{dt} = r(K - \varepsilon(t)) \left[1 - \left(\frac{K - \varepsilon(t)}{K} \right)^\alpha \right], \quad (2.7)$$

que com alguma manipulação algébrica fica

$$\frac{d\varepsilon}{dt} = -rK [1 - (\varepsilon(t)/K)] \left[1 - \left(1 - \frac{\varepsilon(t)}{K} \right)^\alpha \right]. \quad (2.8)$$

Figura 3 – Curvas de mortes acumuladas atribuídas à Covid-19 até 8 de maio de 2020 para (a) Itália, (b) Espanha, (c) França, (d) Alemanha, (e) Irã e (f) Coreia do Sul. Nas figuras, os círculos em vermelho representam o número acumulado de mortes no respectivo dia e a curva sólida preta representa o ajuste numérico com o modelo de Richards, cujos parâmetros do melhor fit estão indicados no *inset* das figuras, juntamente com seus erros.



Fonte: (VASCONCELOS et al., 2020).

Finalmente, tomando o limite $\varepsilon(t)/K \rightarrow 0$ e usando a expansão $(1-x)^\alpha \approx 1 - \alpha x$ em torno de $x = 0$, chegamos em

$$\frac{d\varepsilon}{dt} = -r\alpha\varepsilon, \quad (2.9)$$

que novamente mediante integração direta nos fornece um decrescimento exponencialmente rápido para a variável $\varepsilon(t)$ e conseqüentemente, uma aproximação exponencialmente rápida ao platô de saturação. Estas características assintóticas — tanto no crescimento inicial da curva epidêmica, quanto na sua aproximação ao platô — podem fazer com que o modelo de Richards falhe em descrever as curvas de mortes ou casos acumulados para algumas localidades.

Esta falha do modelo de Richards em descrever as curvas epidêmicas de algumas localidades pode ser explicada pelo fato de que a literatura das áreas de epidemiologia e modelagem matemática de epidemias mostrou em diversas oportunidades que a fase inicial de uma epidemia não necessariamente possui um crescimento exponencialmente rápido, podendo ter um comportamento polinomial ou até mesmo linear (CHOWELL et al., 2016). De forma análoga, recentemente foi mostrado (VASCONCELOS et al., 2021) que a aproximação ao platô de saturação pode acontecer de forma subexponencial (lei de potência) e que este fato está ligado com a implementação de medidas de controle da epidemia em uma dada localidade, acelerando ou retardando a aproximação ao tamanho final da epidemia.

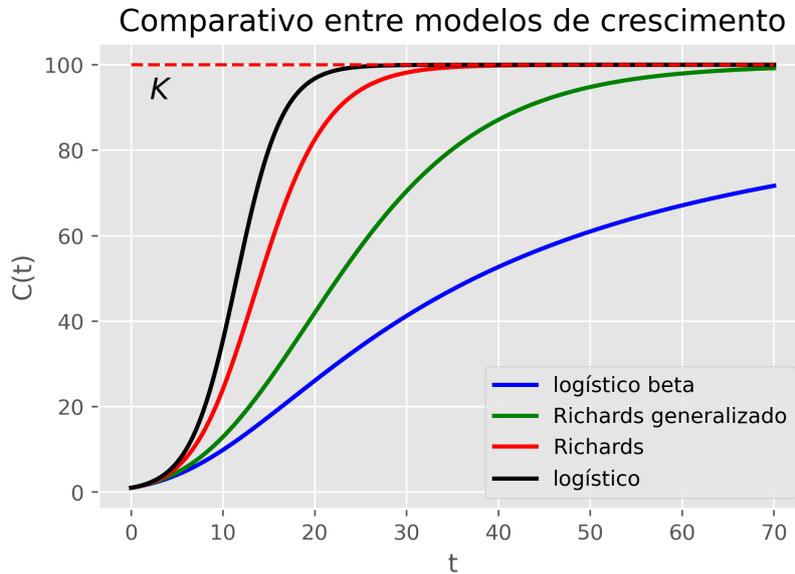
Uma forma de adaptar o modelo de Richards para abranger tais regimes assintóticos nas fases inicial e final da curva epidêmica é através da introdução de dois novos parâmetros na EDO que define o modelo. Este modelo é conhecido na literatura como modelo logístico generalizado (TSOULARIS; WALLACE, 2002) como mostrado a seguir:

$$\frac{dC}{dt} = r [C(t)]^q \left[1 - \left(\frac{C(t)}{K} \right)^{\alpha} \right]^p, \quad (2.10)$$

onde agora o parâmetro q controla o regime de crescimento inicial da curva epidêmica, interpolando entre crescimento linear, $q = 0$, crescimento subexponencial, $0 < q < 1$, e crescimento exponencial, $q = 1$. O parâmetro p controla a forma como a curva se aproxima ao platô de saturação; quando $p > 1$, a aproximação ao platô acontece com uma taxa polinomial, enquanto o caso $p = 1$ corresponde à aproximação exponencial ao platô de saturação. Os demais parâmetros, r , α e K continuam com as mesmas funções já discutidas na equação 2.1. Este modelo foi denominado modelo logístico beta (MLB) pelos autores (VASCONCELOS et al., 2021) pelo fato de o lado direito da equação 2.10 se tratar de uma distribuição beta generalizada do primeiro tipo (MCDONALD, 1984).

Notamos que ao tomarmos alguns limites nos parâmetros da equação 2.10, recuperamos vários casos conhecidos de modelos fenomenológicos de crescimento. Ao fazermos $p = 1$, temos o modelo generalizado de Richards (WU et al., 2020) que apresenta, além das características do modelo de Richards, a possibilidade de crescimento subexponencial e linear na fase inicial

Figura 4 – Gráfico comparativo ilustrando o comportamento das curvas acumuladas geradas pelos diferentes modelos de crescimento no mesmo intervalo de tempo e com o mesmo tamanho final da epidemia, $K = 100$.



Fonte: elaborada pelo autor (2023).

da curva epidêmica. Ao fazermos $p = q = 1$, recuperamos o modelo de Richards, já discutido anteriormente nesta tese. Finalmente, ao fazermos $p = q = \alpha = 1$, recuperamos a curva em formato de S simétrica do modelo logístico tradicional.

Reproduzindo mais uma vez o procedimento realizado entre as equações 2.7 e 2.9, obtemos o resultado

$$\frac{d\varepsilon}{dt} = -\gamma\varepsilon^p, \quad (2.11)$$

onde $\gamma = r\alpha^p/K^{p-q}$. Através de integração direta, obtemos a seguinte expressão para $\varepsilon(t)$:

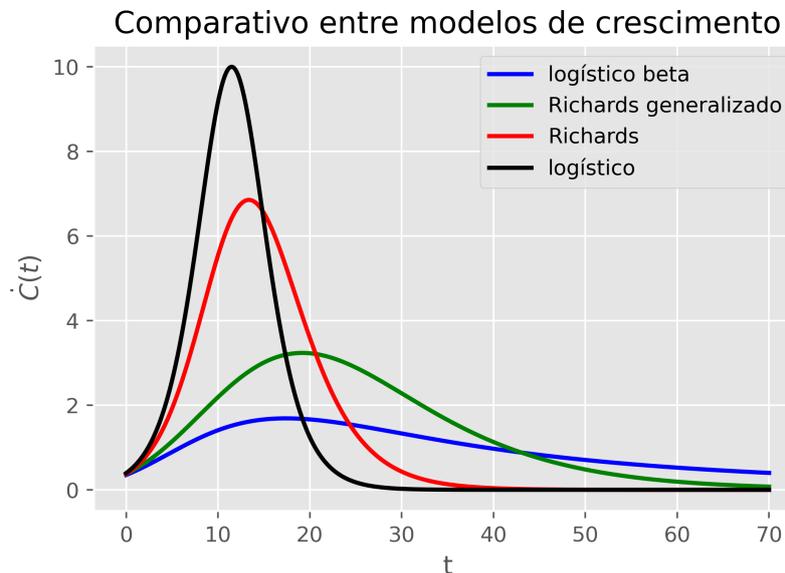
$$\varepsilon(t) = D^{1/(1-p)} \exp_p\left(-\frac{\gamma t}{D}\right), \quad (2.12)$$

onde novamente a função $\exp_p(x)$ é a função q-exponencial. Fica claro então que no modelo MLB, a aproximação ao platô pode acontecer desde forma linear até a forma exponencial.

A figura 4 ilustra os diferentes modelos fenomenológicos de crescimento para um mesmo tamanho final de epidemia, $K = 100$. A curva em azul é referente ao modelo logístico beta com parâmetros $r = 0,4$, $\alpha = 0,6$, $q = 0,8$ e $p = 2$; a curva em verde se refere ao modelo generalizado de Richards com parâmetros $r = 0,4$, $\alpha = 0,6$ e $q = 0,8$; a curva em vermelho representa o modelo de Richards com parâmetros $r = 0,4$ e $\alpha = 0,6$; a curva em preto é referente ao modelo logístico tradicional com parâmetro $r = 0,4$.

Pela figura podemos ver os diferentes comportamentos dos modelos, tanto na fase inicial do crescimento quanto na fase final da curva. O modelo logístico, juntamente com o modelo

Figura 5 – Gráfico comparativo ilustrando o comportamento das curvas diárias geradas pelos diferentes modelos de crescimento no mesmo intervalo de tempo e com o mesmo tamanho final da epidemia, $K = 100$.



de Richards, apresentam crescimento exponencial na fase inicial e portanto atingem valores mais altos em um intervalo de tempo menor se comparados com os outros dois modelos; além disso, o fato de os dois modelos apresentarem aproximação exponencialmente rápida ao platô faz com que a curva "dobre" (termo usado para descrever uma mudança considerada rápida no comportamento de um dado gráfico), saindo do regime de crescimento linear e entrando no regime final de saturação. É possível notar também que dos quatro modelos apresentados, o beta logístico é o que mais demora a atingir o platô de saturação devido tanto ao seu crescimento inicial subexponencial, quanto ao seu crescimento final polinomial.

A figura 5 ilustra o comportamento das curvas diárias para os diferentes modelos de crescimento discutidos nesta seção. As chamadas curvas diárias são a derivada com relação ao tempo das curvas acumuladas, ou seja, $\dot{C}(t) = dC/dt$. Elas representam o número de casos confirmados ou mortes em um certo dia, contado a partir do primeiro dia em que houve casos confirmados ou mortes; tais curvas apresentam um pico que corresponde ao ponto de inflexão na curva acumulada. Vale a pena notar na figura que a curva em preto, que representa o modelo logístico, apresenta o pico mais pronunciado e este acontece mais cedo que nas curvas de todos os outros modelos. Em completa oposição, o modelo logístico beta apresenta uma curva mais distribuída ao longo do tempo, com um pico bem menos pronunciado. Os outros dois modelos interpolam estas duas curvas discutidas.

Os diferentes regimes, tanto de crescimento inicial da curva epidêmica, como de sua saturação surgem como resultado de uma coleção de fatores, como a forma de transmissão da doença, o vetor que a transmite (caso haja), as medidas de controle impostas (ou a falta delas) pelas autoridades responsáveis e a adesão dessas medidas por parte da população. Estudos recentes mostram que os valores do parâmetro q , que controla o crescimento inicial da curva epidêmica, podem variar bastante de acordo com a doença e com a localidade.

Quando analisados os dados de uma pandemia de influenza em São Francisco no ano de 1918 e de um surto de varíola em Khulna, Bangladesh, no ano de 1972, foram obtidos parâmetros q acima de 0.85, indicando um perfil de crescimento rápido e que se aproximava de exponencial (CHOWELL et al., 2016). No outro lado do espectro, quando modelados os dados de uma epidemia de AIDS no Japão entre os anos de 1985 e 2012, o parâmetro q encontrado foi próximo de 0.5, indicando um crescimento aproximadamente linear (VIBOUD; SIMONSEN; CHOWELL, 2016). Ainda no mesmo estudo, no caso mais extremo, foram encontrados parâmetros q muito próximos de 1 para surtos de Ebola nos distritos de Margibi, na Libéria, e nos distritos de Bo e Bombali em Serra Leoa, indicando um crescimento exponencial na fase inicial da epidemia.

Assim como a Equação Diferencial Ordinária 2.1, a EDO 2.10 pode ser integrada de forma exata e é possível obter uma solução analítica na forma implícita $t = t(C)$ (VASCONCELOS et al., 2021):

$$t = \frac{C^{1-q}}{r(1-q)} {}_2F_1 \left(p, \frac{1-q}{\alpha}; 1 + \frac{1-q}{\alpha}; \left(\frac{C}{K} \right)^\alpha \right) - t_i, \quad (2.13)$$

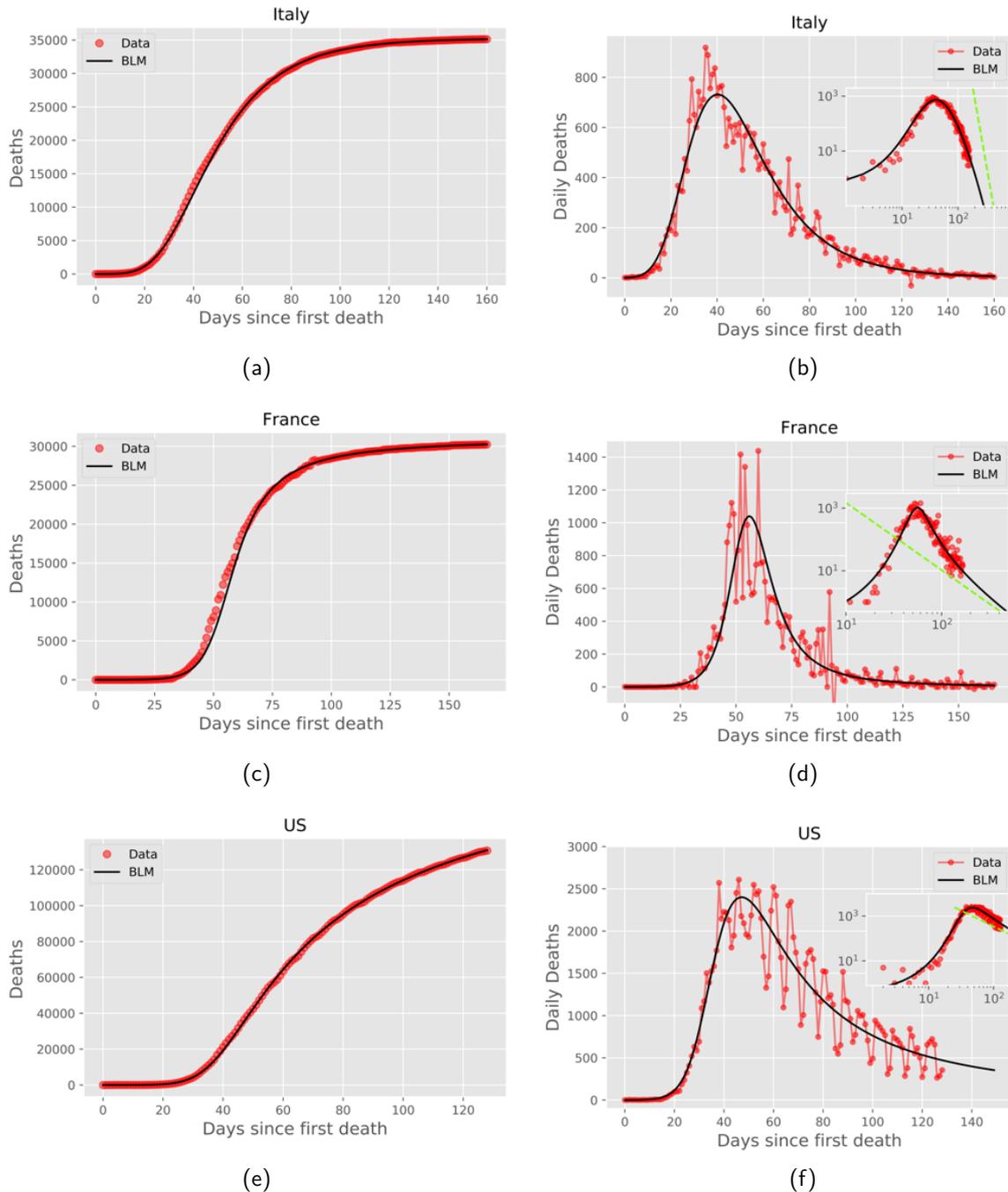
onde ${}_2F_1(a, b; c; x)$ é a função hipergeométrica de Gauss e

$$t_i = \frac{C_0^{1-q}}{r(1-q)} {}_2F_1 \left(p, \frac{1-q}{\alpha}; 1 + \frac{1-q}{\alpha}; \left(\frac{C_0}{K} \right)^\alpha \right). \quad (2.14)$$

2.1.2.1 Ajustes Numéricos

Assim como nos ajustes numéricos discutidos no caso do modelo de Richards, para o beta logístico os ajustes também consistiram em resolver o problema de otimização usando o algoritmo de Levenberg-Marquardt, neste caso utilizando o módulo `Imfit` (NEVVILLE et al., 2014) da linguagem de programação Python. Neste estudo foram analisados 8 países cujas curvas epidêmicas acumuladas já haviam se distanciado do ponto de inflexão e estavam no platô de saturação (VASCONCELOS et al., 2021). Os países escolhidos foram: Canadá, Itália, Reino Unido, Holanda, Alemanha, Bélgica, França e Estados Unidos. Foram analisados para

Figura 6 – Coluna da esquerda: Número acumulado de mortes (círculos vermelhos) em função do número de dias desde o primeiro registro de morte até a data de 31 de julho de 2020 para (a) Itália e (c) França e até a data de 07 de julho de 2020 para (e) Estados Unidos. Nesta coluna, as curvas pretas sólidas representam o melhor ajuste numérico utilizando o modelo logístico beta. Coluna da direita: Dados diários de mortes (círculos vermelhos conectados pelas linhas vermelhas) para os mesmos países e até as mesmas datas, respectivamente. As curvas pretas neste caso representam a derivada da curva acumulada resultante do ajuste numérico com o modelo logístico beta; no *inset* de cada figura é exibido um plot em escala log-log para ressaltar o comportamento assintótico de lei de potência.



Fonte: (VASCONCELOS et al., 2021).

estes países os dados de mortes acumuladas; estes dados foram retirados do repositório da Universidade John Hopkins (JHU, 2020) com datas finais variadas de acordo com cada país.

Nos ajustes numéricos com o modelo logístico beta, com o propósito de evitar o *overfitting*, os parâmetros livres (r, α, q, p) foram limitados ao intervalo $(0, 1)$, evitando assim que fossem obtidos parâmetros fora do limite aceitável. Na figura 6 são exibidos os ajustes numéricos realizados utilizando as curvas acumuladas de mortes para Itália, França e Estados Unidos. Na coluna da esquerda temos o número acumulado de mortes (círculos vermelhos) como função do número de dias contados a partir do primeiro dia em que foi contabilizada uma morte para a dada localidade. As curvas cheias em preto representam o modelo logístico beta utilizando os parâmetros obtidos a partir do melhor ajuste numérico com os dados empíricos. No caso da Itália e da França, figuras 6a e 6c respectivamente, os dados utilizados vão até o dia 31 de julho de 2020; para os Estados Unidos, figura 6e, os dados vão até o dia 07 de julho de 2020. Na coluna da direita são mostradas as mortes diárias (círculos vermelhos ligados por linhas vermelhas) para os mesmos países. A curva cheia em preto representa a curva diária teórica, obtida a partir da derivada com relação ao tempo do modelo logístico beta com os parâmetros obtidos pelo melhor ajuste numérico. No *inset* das figuras da coluna da direita, são exibidos ainda gráficos em escala log-log para destacar o comportamento assintótico de lei de potência dos dados e capturado pelo modelo.

Como podemos ver, os ajustes com o modelo logístico beta têm uma ótima concordância com os dados de mortes acumuladas para os três países exibidos na figura. Notamos ainda que para a Itália e França, a aproximação ao platô se dá de forma muito mais rápida e acentuada que no caso dos Estados Unidos, em que a aproximação ao platô se dá de forma muito mais lenta e suave. Esta característica é capturada de forma explícita no modelo através do valor do parâmetro p para cada país; O melhor ajuste para cada país resultou em $p = 1,09$ para a Itália, $p = 1,87$ para a França e $p = 3,43$ para os Estados Unidos.

Por fim, notamos ainda a ótima concordância entre os dados de mortes diárias e a curva diária teórica, obtida através da derivada temporal da curva acumulada com os parâmetros encontrados no ajuste numérico do modelo. Ressaltamos que mesmo o ajuste sendo feito através da curva acumulada, quando calculada a curva diária teórica, todas as características dos dados empíricos são capturadas de forma excelente, o que reforça a ótima consistência do modelo.

2.2 MODELOS DE CRESCIMENTO COM MÚLTIPLAS ONDAS

Os modelos de crescimento logístico apresentados na seção 2.1 se provaram uma excelente ferramenta para descrever as curvas epidêmicas da pandemia de Covid-19 para várias localidades e durante um bom período de tempo. No entanto, à medida em que segundas ondas de infecções e conseqüentemente de mortes foram surgindo, o modelo logístico beta, que descrevia de forma excepcional a evolução da epidemia, passou a não ser mais bem sucedido em tal tarefa. Para solucionar este problema e fornecer uma resposta rápida no enfrentamento à doença, um modelo de crescimento logístico generalizado com parâmetros dependentes do tempo foi proposto (VASCONCELOS et al., 2021).

Esta nova versão do modelo é descrita pela Equação Diferencial Ordinária 2.10, no entanto, agora assumimos que os parâmetros desta EDO são dependentes do tempo, ou seja, $r = r(t)$, $\alpha = \alpha(t)$, $K = K(t)$, $q = q(t)$ e $p = p(t)$ e as múltiplas ondas de infecções e mortes são capturadas através da variação destes parâmetros no tempo, permitindo o modelo possuir um número arbitrário de ondas. Começaremos nossa descrição com o caso de uma curva epidêmica com apenas duas ondas e em seguida faremos a generalização para um número N arbitrário.

Seja $\zeta(t)$ a representação de um parâmetro qualquer do modelo logístico beta, consideramos a dependência temporal de $\zeta(t)$ descrita pela seguinte equação:

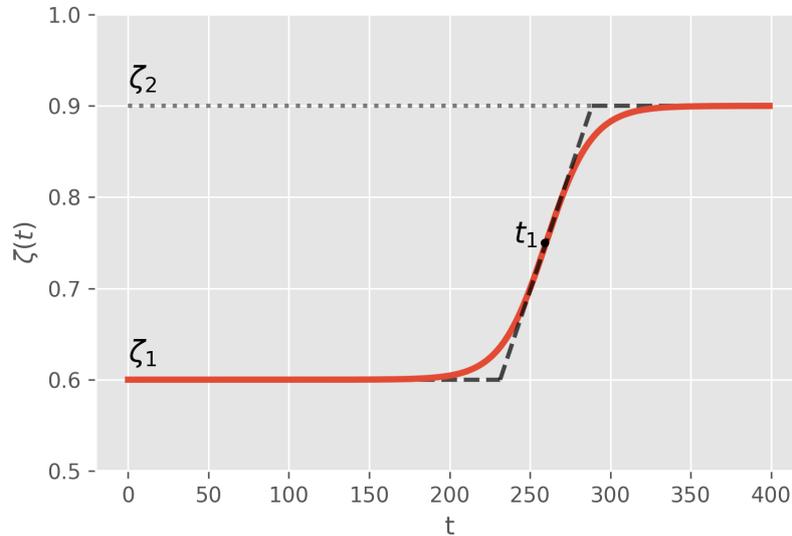
$$\zeta(t) = \zeta_1 + \frac{(\zeta_2 - \zeta_1)}{2} \left[1 + \tanh \left(\frac{\rho_1(t - t_1)}{2} \right) \right], \quad (2.15)$$

onde ζ_1 e ζ_2 representam os valores dos parâmetros na primeira e segunda onda, respectivamente. O parâmetro t_1 indica o tempo em que ocorre a transição, enquanto ρ_1 caracteriza a rapidez da transição entre os dois parâmetros em torno do tempo t_1 ; quanto maior ρ_1 , mais rápida é a transição.

Note que tanto a escala de tempo t_1 , quanto a taxa de transição ρ_1 são as mesmas para todos os parâmetros do modelo; isto é consistente com o que é observado em uma epidemia, visto que uma nova onda — que vem acompanhada de uma mudança geral nas características da dinâmica — deve alterar todos os parâmetros do modelo simultaneamente.

Na figura 7, podemos ver o comportamento do parâmetro genérico $\zeta(t)$ variando entre os valores ζ_1 e ζ_2 representado pela curva sólida; a linha tracejada representa a aproximação linear de $\zeta(t)$. Ressaltamos que, apesar de o tempo de transição ser determinado pelo parâmetro t_1 , os efeitos de uma nova onda podem ser sentidos algum tempo antes deste instante, visto que a velocidade com a qual a transição ocorre depende também do parâmetro ρ_1 .

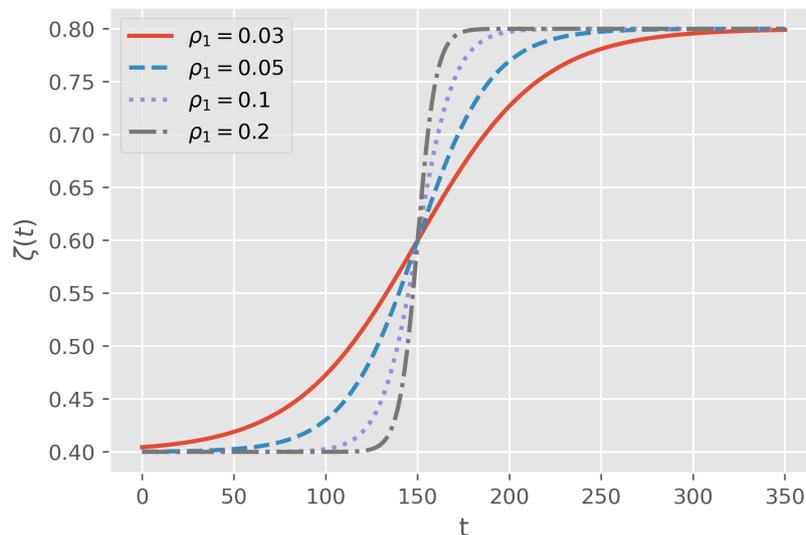
Figura 7 – Evolução temporal de um parâmetro genérico do modelo $\zeta(t)$ entre dois valores ζ_1 e ζ_2 com tempo de transição t_1 e taxa de transição ρ_1 . A curva tracejada representa uma aproximação linear da curva sólida.



Fonte: elaborada pelo autor e publicada em (VASCONCELOS et al., 2021).

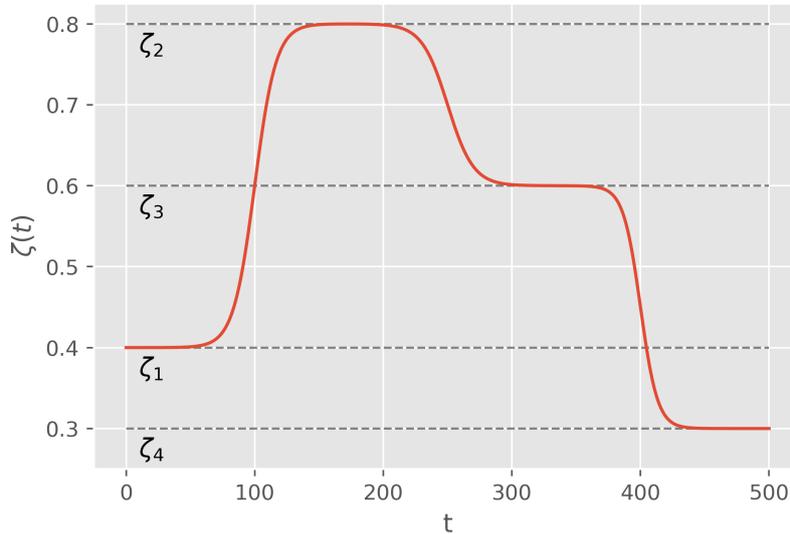
Na figura 8 temos diferentes formatos da curva $\zeta(t)$ para valores fixos de $t_1 = 250$, $\zeta_1 = 0,4$ e $\zeta_2 = 0,8$ e diferentes valores de ρ_1 . Note que o valor do parâmetro ρ_1 muda significativamente a velocidade da transição; para $\rho_1 = 0,03$, ζ em $t = 200$ é próximo de 0,73, enquanto para $\rho_1 = 0,2$ no mesmo instante de tempo, ζ já chegou no seu valor final, 0,8.

Figura 8 – Variação do formato da curva $\zeta(t)$ para valores fixos de t_1 , ζ_1 e ζ_2 e diferentes valores de ρ_1 , ilustrando transições mais ou menos suaves a depender do valor do parâmetro.



Fonte: elaborada pelo autor (2023).

Figura 9 – Evolução temporal de $\zeta(t)$ para o caso ilustrativo de 4 ondas epidêmicas com os seguintes parâmetros: $\zeta_1 = 0,4$, $\zeta_2 = 0,8$, $\zeta_3 = 0,6$, $\zeta_4 = 0,3$, $t_1 = 100$, $t_2 = 250$, $t_3 = 400$, $\rho_1 = 0,12$, $\rho_2 = 0,1$ e $\rho_3 = 0,15$.



Fonte: elaborada pelo autor (2023).

Agora que já visualizamos bem como se dá a variação de um parâmetro genérico do modelo, $\zeta(t)$, para o caso de duas ondas epidêmicas, podemos generalizar a equação 2.15 para o caso de um número arbitrário de ondas:

$$\zeta(t) = \zeta_1 + \sum_{i=1}^{N-1} \frac{(\zeta_{i+1} - \zeta_i)}{2} \left[1 + \tanh \left(\frac{\rho_i(t - t_i)}{2} \right) \right], \quad (2.16)$$

onde N é o número de ondas epidêmicas exibidas pela curva, ρ_i e t_i são a taxa e o tempo de transição entre ondas subsequentes. O modelo logístico beta para múltiplas ondas tem um total de $5N + 2(N - 1) = 7N - 2$ parâmetros, sendo $5N$ referentes a cada uma das ondas epidêmicas do modelo logístico beta e $2(N - 1)$ referentes aos parâmetros de transição entre as ondas subsequentes. Na figura 9 está ilustrada a evolução temporal de $\zeta(t)$ para o caso de 4 ondas epidêmicas com parâmetros utilizados indicados na legenda da figura.

Com a substituição dos parâmetros estáticos por parâmetros dependentes do tempo na equação 2.10, a integração direta não é mais possível e se faz necessário recorrer à integração numérica para calcular a curva acumulada a partir da EDO, o que não impõe qualquer restrição ou dificuldade na realização dos ajustes numéricos com o modelo logístico beta de múltiplas ondas, como será mostrado em seguida.

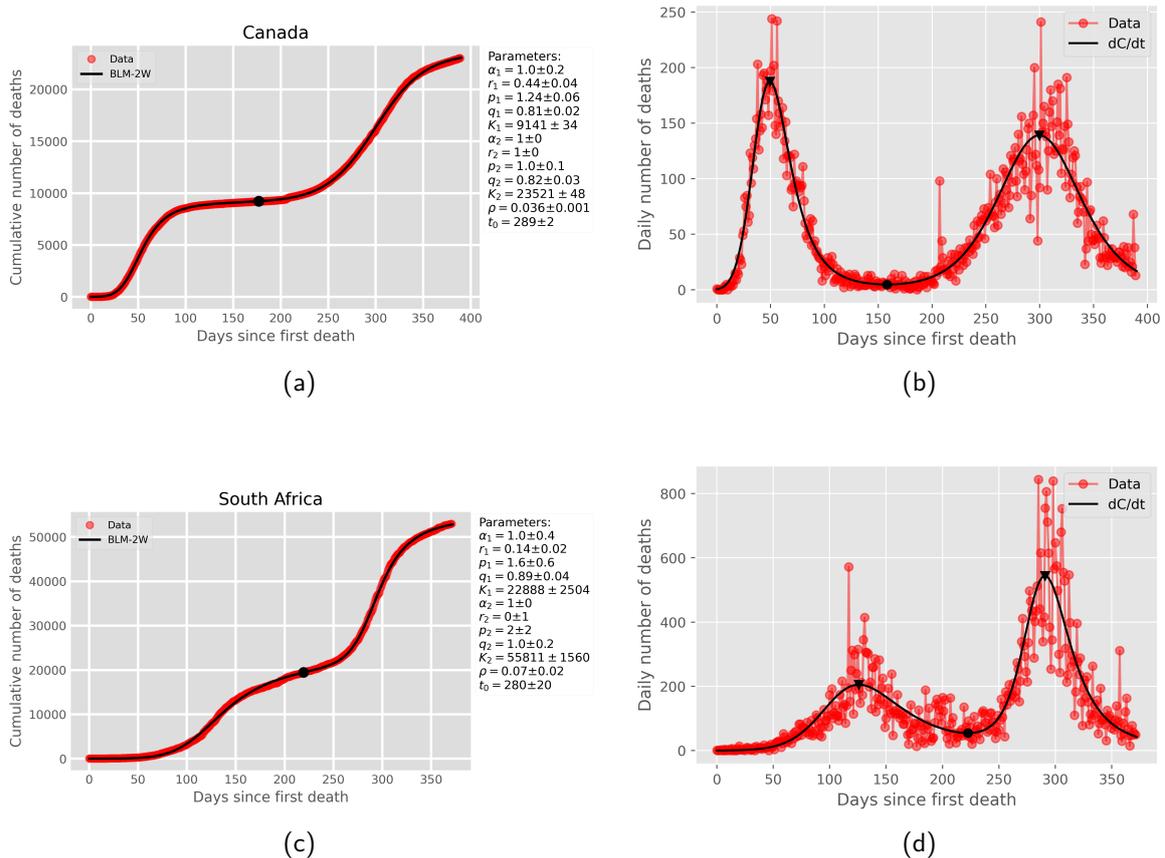
2.2.1 Ajustes Numéricos

Como discutido recentemente, o modelo logístico beta para múltiplas ondas tem $7N - 2$ parâmetros, ou seja, para o caso de uma curva epidêmica com duas ondas, teríamos 12 parâmetros livres para serem determinados através do ajuste numérico. Esta elevada quantidade pode levar à ocorrência de problemas com *overfitting*, gerando valores espúrios ou com grandes erros associados para os parâmetros do modelo. Antes de analisarmos o desempenho do modelo logístico beta na descrição de curvas epidêmicas com múltiplas ondas, vamos descrever o procedimento que foi utilizado para minimizar a ocorrência destes problemas.

Em todos os ajustes numéricos utilizando o modelo logístico beta com parâmetros dependentes do tempo para descrever curvas epidêmicas com múltiplas ondas, foi aplicado um método com múltiplas etapas. Na primeira fase deste método, é feita uma estimativa do tempo em que ocorre a primeira transição, t_1 , e a curva acumulada é truncada neste ponto; em seguida é feito um ajuste numérico com o modelo logístico beta de uma onda utilizando sua solução analítica dada pela equação 2.13 utilizando chutes iniciais para os parâmetros do modelo. Os parâmetros obtidos como resultado do ajuste são então utilizados como chutes iniciais para um ajuste subsequente usando o modelo logístico beta para o caso de 2 ondas, enquanto para t_1 é assumido como chute inicial o valor estimado na primeira etapa do procedimento e ρ_1 é escolhido aleatoriamente. Vale notar que se a curva epidêmica apresentar mais de duas ondas epidêmicas, é novamente estimado o tempo de transição t_2 entre a segunda e terceira ondas e a curva é truncada neste ponto; caso contrário, a curva inteira é utilizada no ajuste numérico. Este primeiro ajuste numérico serve como uma calibração dos chutes iniciais dos parâmetros referentes à primeira onda epidêmica no modelo, o que mitiga o problema de *overfitting* do modelo devido ao grande número de parâmetros (12 já no caso de 2 ondas).

No caso de uma curva acumulada que apresente 3 ou mais ondas, o procedimento continua nos mesmos moldes discutidos. Neste caso, os 12 parâmetros obtidos através do ajuste com o modelo de 2 ondas epidêmicas servirão agora como chute inicial para os parâmetros do modelo logístico beta com 3 ondas epidêmicas, mais uma vez mitigando o problema de *overfitting* em um ajuste que agora possui 19 parâmetros a serem determinados. Este procedimento pode ser aplicado de forma sequenciada para uma curva acumulada com um número arbitrário N de ondas, no entanto, o número de parâmetros $7N - 2$ escala de forma muito rápida com N , aumentando de forma considerável o problema com *overfitting*; para contornar esta situação, um procedimento automatizado e com um algoritmo para detecção e filtragem de

Figura 10 – Coluna da esquerda: número acumulado de mortes (círculos vermelhos) atribuídos à Covid-19 para (a) Canadá e (c) África do Sul até 3 de abril de 2021. A linha preta representa o melhor ajuste numérico com o modelo logístico beta de duas ondas e o ponto preto na curva representa o tempo t_1 em que ocorre a transição entre as duas ondas; ao lado de cada gráfico são exibidos os parâmetros encontrados pelo ajuste. Coluna da direita: número diário de mortes para os mesmos países (círculos vermelhos ligados por segmentos de reta). A curva preta representa a derivada temporal do modelo logístico beta com os parâmetros encontrados pelo melhor ajuste numérico.



Fonte: elaborada pelo autor e publicada em (VASCONCELOS et al., 2021).

ondas epidêmicas foi implementado e será discutido no capítulo 3.

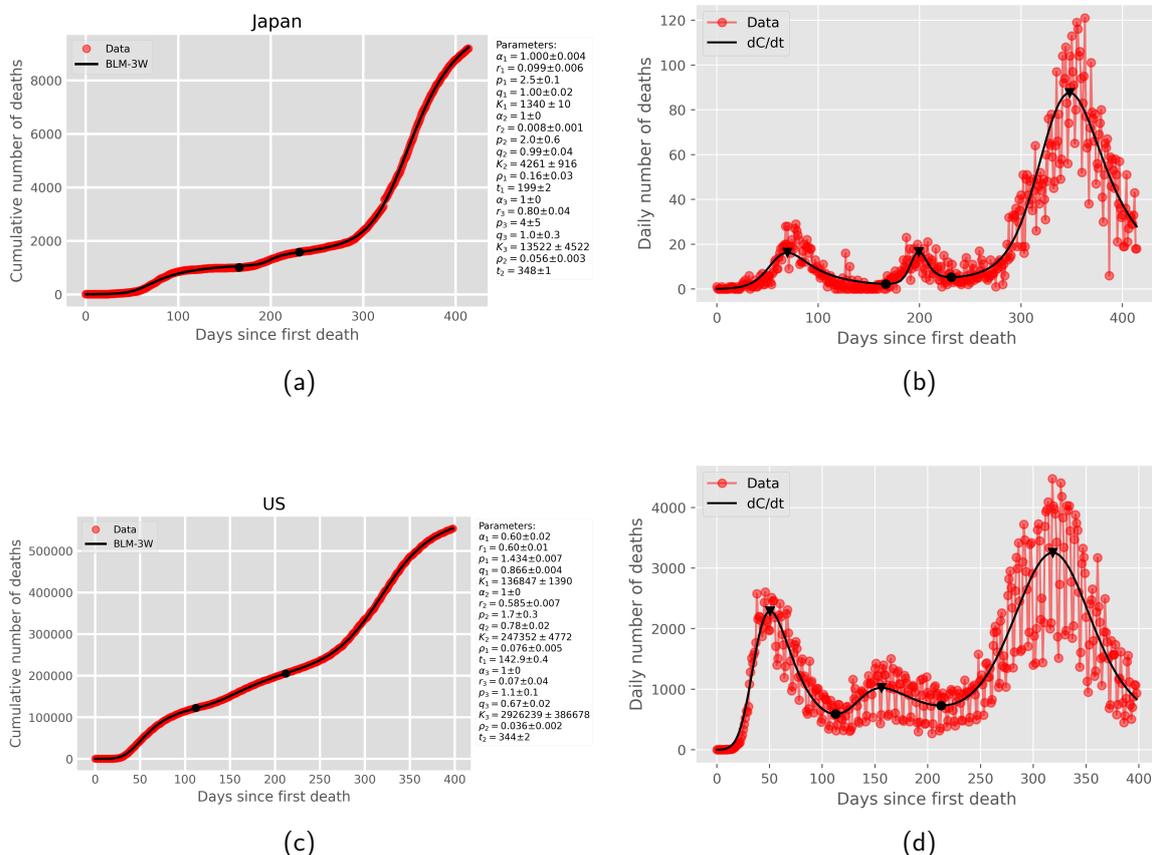
Os ajustes numéricos apresentados em (VASCONCELOS et al., 2021) e aqui revisitados foram realizados resolvendo o problema de otimização através do algoritmo de Levenberg-Marquardt implementado no módulo `lmfit` (NEWVILLE et al., 2014) da linguagem Python. Foram considerados os dados de mortes atribuídas à Covid-19 até o dia 3 de abril de 2021 para países com até 3 ondas epidêmicas, a saber: Brasil, Canadá, Alemanha, México, África do Sul e Suécia apresentando duas ondas epidêmicas e Irã, Itália, Japão e Estados Unidos apresentando três ondas. Nesta tese, por concisão, revisitaremos apenas dois países que apresentavam na referida data duas e três ondas epidêmicas, respectivamente.

Na figura 10 são apresentados os ajustes numéricos das curvas de mortes acumuladas para Canadá e África do Sul com o modelo logístico beta de duas ondas. Na coluna da esquerda,

representados por círculos vermelhos, temos os dados de mortes acumuladas para cada um dos países e a curva preta indica o modelo logístico beta com os parâmetros encontrados através do melhor ajuste numérico, que estão indicados (juntamente com seus erros associados) no lado direito do gráfico. Na coluna da direita, indicados por círculos vermelhos, temos os dados de mortes diárias para os referidos países, a curva preta representa a derivada temporal da curva acumulada apresentada na coluna da esquerda, os triângulos invertidos em preto indicam os pontos de máximo (picos) da curva diária teórica, enquanto o círculo preto indica o ponto em que ocorre o mínimo desta curva.

Como pode ser visto, a concordância entre a curva teórica e os dados de mortes acumuladas é excelente; além disso, apesar de o ajuste numérico ser feito com base nos dados de mortes acumuladas, a concordância entre a derivada da curva teórica e os dados de mortes diárias

Figura 11 – Coluna da esquerda: número acumulado de mortes (círculos vermelhos) atribuídos à Covid-19 para (a) Japão e (c) Estados Unidos até 3 de abril de 2021. A linha preta representa o melhor ajuste numérico com o modelo logístico beta de duas ondas e o ponto preto na curva representa o tempo t_1 em que ocorre a transição entre as duas ondas; ao lado de cada gráfico são exibidos os parâmetros encontrados pelo ajuste. Coluna da direita: número diário de mortes para os mesmos países (círculos vermelhos ligados por segmentos de reta). A curva preta representa a derivada temporal do modelo logístico beta com os parâmetros encontrados pelo melhor ajuste numérico.



Fonte: elaborada pelo autor e publicada em (VASCONCELOS et al., 2021).

reforça ainda mais a eficácia do modelo em descrever a evolução da pandemia de Covid-19 nas dadas localidades. Notamos ainda que a transição entre duas ondas epidêmicas subsequentes pode acontecer de duas formas distintas: na primeira delas, o número de mortes diárias atinge um valor muito baixo e a curva acumulada atinge um platô de saturação bem definido; este caso foi denominado de transição "padrão" e é ilustrado através das curvas epidêmicas do Canadá nas figuras 10a e 10b. Na segunda delas, o número de mortes diárias diminui significativamente com relação ao pico da onda anterior, mas não chega a números tão baixos como no caso da transição "padrão"; neste caso, a curva acumulada não chega a atingir um platô de saturação e antes que os efeitos da primeira onda epidêmica desapareçam, a segunda onda já entra em cena e faz com que o número diários de mortes volte a aumentar, gerando assim o caso que foi denominado transição "anômala" e que é ilustrado com o caso da África do Sul nas figuras 10c e 10d.

Na figura 11 são apresentadas as curvas de mortes acumuladas do Japão e dos Estados Unidos. Nesta figura é seguida a mesma convenção de que os pontos empíricos são representados por círculos vermelhos e a curva teórica (ou sua derivada, no caso da curva diária) é representada por uma linha contínua na cor preta. Na coluna da esquerda são apresentadas as curvas de mortes acumuladas para cada um dos países, juntamente com o melhor ajuste numérico utilizando o modelo logístico beta com três ondas; os parâmetros obtidos estão indicados no lado direito de cada gráfico. Na coluna da direita são exibidas as curvas diárias para os mesmos países juntamente com a curva diária teórica; nos gráficos das curvas diárias temos ainda círculos em preto indicando os mínimos da curva diária e triângulos invertidos também em preto indicando os picos da mesma curva.

Aqui novamente é possível ver a excelente concordância entre os dados de mortes acumuladas e o ajuste numérico com o modelo logístico beta de três ondas, bem como entre os dados de mortes diárias e curva diária teórica. Note que no caso de três ondas epidêmicas, o modelo tem um total de 19 parâmetros livres para serem determinados através do ajuste numérico; ainda assim, os erros relativos nos parâmetros encontrados pelo ajuste são todos menores que 100%, como pode ser visto nas figuras 11a e 11c. Assim como no caso anterior, podemos ver que o Japão, 11a, apresenta três ondas na sua curva de mortes acumuladas e que a transição entre cada uma das três ondas se dá da forma denominada "padrão", ou seja, temos a presença de platôs de saturação bem definidos; isto pode ser reforçado ainda mais através da curva de mortes diárias 11b, na qual é possível ver mínimos bastante pronunciados aproximando-se do zero. Em oposição ao Japão, os Estados Unidos, 11c, apresentam entre

suas ondas a transição do tipo "anômala", como pode ser visto pela ausência de platôs bem pronunciados e mínimos mais distantes do zero e relativamente próximos dos picos na curva de mortes diárias.

Com isto encerramos nossa análise dos modelos de crescimento aplicados à pandemia de Covid-19, e podemos concluir que tais modelos fenomenológicos têm excelente desempenho na modelagem das curvas epidêmicas para a pandemia causada pelo vírus SARS-Cov-2, como mostrado nas diversas aplicações abordadas nesta seção através de ajustes numéricos com os modelos logísticos. Notamos ainda que o modelo logístico beta consegue capturar padrões de crescimento extremamente complexos com a adição de parâmetros dependentes do tempo, sendo capaz de modelar curvas com um elevado número de ondas epidêmicas.

2.3 MODELOS COMPARTIMENTAIS

Nesta seção vamos fazer uma breve revisão dos modelos compartimentais, começando pelo mais simples, com 3 compartimentos, até um modelo com 4 compartimentos e parâmetros dependentes do tempo, capaz de modelar o comportamento de curvas epidêmicas que apresentam mais de uma onda de casos ou mortes. Esta classe de modelos matemáticos tem sido amplamente utilizada para descrever epidemias dos mais diversos tipos de forma muito bem sucedida. Os modelos compartimentais surgiram a partir da pesquisa de um físico e um bioquímico, ambos escoceses, que resultaram em uma série de três artigos seminais na área de epidemiologia (KERMACK; MCKENDRICK, 1927; KERMACK; MCKENDRICK, 1932; KERMACK; MCKENDRICK, 1933).

Como foi mostrado na seção anterior, os modelos de crescimento são bastante versáteis e conseguem capturar e descrever muito bem o comportamento de curvas epidêmicas tanto para casos simples quanto para os casos mais complexos, nos quais há a presença de múltiplas ondas. No entanto, os parâmetros dos modelos de crescimento não têm uma interpretação tão direta em termos de conceitos epidemiológicos (WANG; WU; YANG, 2012).

Considere por exemplo o parâmetro que representa a taxa de transmissão β do modelo SIR (KERMACK; MCKENDRICK, 1927). Podemos ver este parâmetro como sendo a probabilidade de um contato entre um indivíduo infectado e um indivíduo suscetível ocasione a transmissão do agente patogênico e gere uma nova infecção — dando origem assim a dois indivíduos infectados — multiplicada pelo número de contatos que ocorrem em um dia. Apesar de não ser possível medir este parâmetro de forma empírica e sem base em um modelo, a interpretação dele é

bastante simples de entender. É possível ainda fazer uma analogia entre a taxa de transmissão β e a velocidade com a qual uma reação ocorre na lei de ação das massas.

Esta classe de modelos consiste em dividir uma dada população em um conjunto de compartimentos mutuamente acoplados, entre os quais há um fluxo de indivíduos da população determinado pelos parâmetros do modelo e das condições iniciais do problema. Dentre os compartimentos utilizados para dividir a população total, alguns dos mais comuns são:

- Suscetível: indivíduo que não desenvolveu a doença e que, em contato com um indivíduo infectado, tem a chance de adquirir a doença;
- Exposto: indivíduo que teve contato com um infectado, mas ainda não desenvolveu a infecção e ainda não é capaz de infectar outros. Este compartimento é especialmente útil quando se quer modelar ou estudar a evolução de doenças causadas por patógenos que possuem período de incubação;
- Infectado: indivíduo que desenvolveu a doença em questão e é capaz de infectar outros;
- Recuperado: indivíduo que desenvolveu a doença e se curou dela. Nos modelos que estudaremos, não é considerada a reinfeção de um indivíduo que já se recuperou da doença, portanto, este é considerado imune;
- Morto: indivíduo que desenvolveu a doença e veio a óbito em decorrência dela.

Note que além dos compartimentos mencionados nesta lista, diversos outros podem ser utilizados para modelar problema mais complexos e que abrangem mais hipóteses, como por exemplo medidas de mitigação e supressão da epidemia (através de quarentena e vacinação), ciclo de vida de hospedeiros e vetores da doença em questão, demografia, questões socioculturais e ambientais, entre outras.

2.3.1 Modelo SIR

Nesta variante do modelo usaremos apenas três dos compartimentos discutidos acima, a saber: Suscetíveis (S), Infectados (I) e Recuperados (R). Agora que já discutimos e escolhemos os compartimentos para o modelo a ser abordado, precisamos entender como ele funciona. Os modelos compartimentais são baseados em um sistema de Equações Diferenciais Ordinárias

acopladas, cujos parâmetros de acoplamento geram a dinâmica do modelo, fazendo os indivíduos passarem de um compartimento para outro à medida em que o tempo passa. Retomando a discussão anterior, sabemos que a taxa de transmissão se relaciona com a probabilidade de haver transmissão do patógeno em um contato entre um indivíduo suscetível e um indivíduo infectado multiplicado pelo número de contatos por dia.

A probabilidade de um dado indivíduo escolhido ao acaso de um total de N , onde $N = S + I + R$ é o número total e fixo de integrantes do sistema, ser suscetível é dada por $s = S/N$ e analogamente para os infectados, $i = I/N$. Desta forma, a probabilidade de um indivíduo suscetível ser acometido pela patologia através de um contato com um infectado é dada por βsi , o que gera uma variação em s de

$$\frac{ds}{dt} = -\beta si. \quad (2.17)$$

Note que do lado direito da equação acima temos um sinal negativo, o que está de acordo com o fato de que quando ocorre a infecção de um indivíduo suscetível, o número total de suscetíveis diminui. Conseqüentemente, os integrantes que passam do compartimento dos suscetíveis para o dos infectados geram uma variação em i de βsi com o sinal positivo, ou seja, aumentando esta quantidade.

Além disso, uma vez que um integrante do sistema entra no estado infectado, ele leva um certo tempo até se recuperar da infecção. O período de infecção de um indivíduo tem uma distribuição exponencial com parâmetro γ , ou seja, a probabilidade de um indivíduo estar infectado depois de τ unidades de tempo é $p(\tau) = \exp^{-\gamma\tau}$, o que gera uma variação temporal em i de $-\gamma i$. Finalmente, os integrantes do sistema que estavam infectados e se recuperam geram uma variação de γi em r . Sumarizando todas estas variações em forma de um sistema de equações diferenciais ordinárias acopladas, temos:

$$\frac{ds}{dt} = -\beta si, \quad (2.18)$$

$$\frac{di}{dt} = \beta si - \gamma i, \quad (2.19)$$

$$\frac{dr}{dt} = \gamma i, \quad (2.20)$$

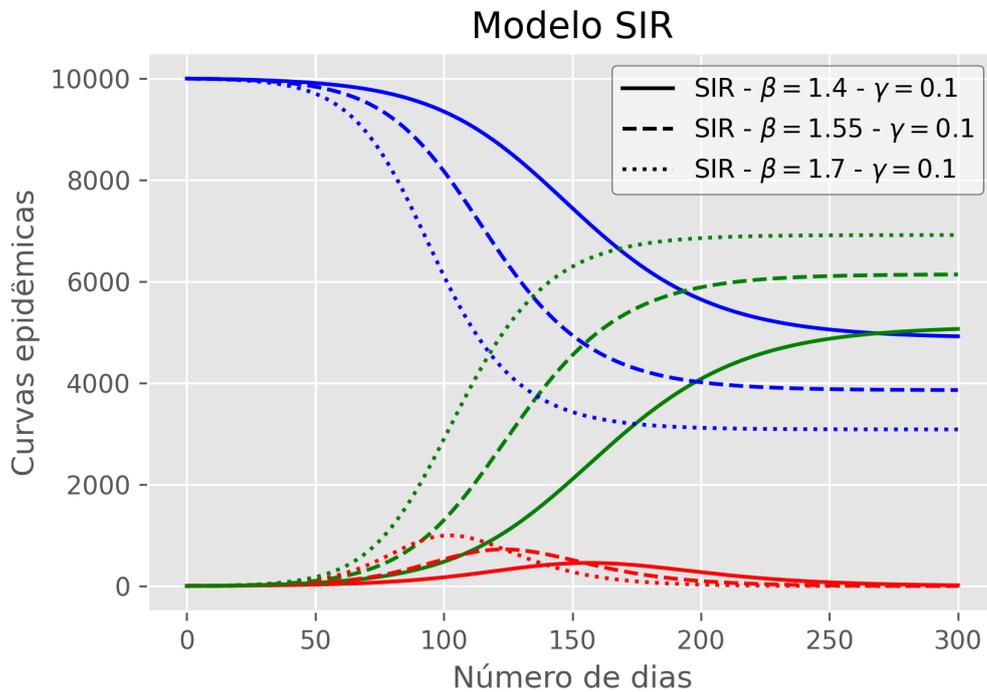
ou em sua versão mais comum, usando o número de indivíduos em cada compartimento,

$$\frac{dS}{dt} = -\frac{\beta SI}{N}, \quad (2.21)$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I, \quad (2.22)$$

$$\frac{dR}{dt} = \gamma I. \quad (2.23)$$

Figura 12 – Comparação das curvas epidêmicas de três integrações numéricas para o modelo SIR com taxa de transmissão $\beta = 0,14$, $\beta = 0,155$ e $\beta = 0,17$ mantendo o parâmetro $\gamma = 0,1$ fixo.



Fonte: elaborada pelo autor.

Note ainda que

$$\frac{dN}{dt} = \frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0, \quad (2.24)$$

o que está de acordo com a condição estabelecida inicialmente de população total fixa.

Na figura 12, podemos ver uma comparação das curvas epidêmicas do modelo SIR; na cor azul temos as curvas epidêmicas dos suscetíveis, na cor vermelha temos as curvas dos infectados e na cor verde temos as curvas dos recuperados. Foram feitas três integrações numéricas do modelo com população total de $N = 10000$ indivíduos, com parâmetro $\gamma = 0,1$ fixo e a taxa de transmissão $\beta = 0,14$ representada no gráfico pelas curvas sólidas, $\beta = 0,155$ representada pelas curvas tracejadas e $\beta = 0,17$ representada pelas curvas pontilhadas. Perceba que a mudança na taxa de transmissão β mantendo-se γ fixo, faz com que a epidemia se desenvolva de forma mais rápida e agressiva, como pode ser verificado pela curva de infectados no caso de $\beta = 0,17$ (curva pontilhada em vermelho), na qual vemos um pico mais pronunciado que os demais e também mais estreito, o que significa que neste caso mais indivíduos ficam infectados ao mesmo tempo. No cenário oposto, $\beta = 0,14$, vemos que o pico da curva de infectados é menos pronunciado, mais largo e ocorre em um instante de tempo posterior ao do caso anterior, no qual a taxa de transmissão era mais elevada.

2.3.2 A taxa básica de reprodução R_0 : determinando se há ou não uma epidemia

Dado um sistema de EDOs como nas equações 2.21-2.23 e seus respectivos parâmetros, como podemos saber se há ou não uma epidemia em curso? Ou seja, como saber se haverá um estado da dinâmica em que a doença se espalha pelos indivíduos de forma consistente, gerando um pico de infecções com posterior decaimento? Podemos responder estas perguntas através de uma quantidade bastante conhecida e usada na área de epidemiologia: a chamada taxa básica de reprodução, R_0 .

A taxa básica de reprodução corresponde ao número esperado de infecções secundárias geradas por uma infecção primária em uma população em que a doença ainda não estava presente (também conhecida como virgem). Esta quantidade tem um limiar bem definido que distingue dois cenários distintos, $R_0 = 1$. Desta forma, se um patógeno para o qual $R_0 > 1$ é introduzido em uma população virgem, a epidemia se desenvolve e há um cenário de transmissão sustentada da doença entre a população, gerando assim uma epidemia. Em contrapartida, se o patógeno em questão tiver $R_0 < 1$, então o número de infecções secundárias é menor que o número de infecções primárias, portanto o número de infectados logo decai e o surto da referida doença não chega a se tornar uma epidemia.

Tal quantidade pode ser calculada para um modelo compartimental através do método de *next generation matrix* (DIEKMANN; HEESTERBEEK; METZ, 1990; DIEKMANN; HEESTERBEEK; ROBERTS, 2009), mas neste caso simples, podemos encontrar R_0 através da análise direta das equações diferenciais do modelo. Considere a equação que descreve a evolução temporal de $I(t)$, eq. 2.22:

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I.$$

Logo que um patógeno é introduzido em uma população virgem, o número de indivíduos suscetíveis é extremamente próximo da população total, N . Neste caso podemos fazer a aproximação $S/N \approx 1$, portanto a equação acima se torna

$$\frac{dI}{dt} = \beta I - \gamma I = (\beta - \gamma)I. \quad (2.25)$$

Nesta aproximação vemos que há um limiar bem definido pela relação entre os parâmetros β e γ ; se $\beta = \gamma$, então $\frac{dI}{dt} = 0$ e o número de infectados se mantém constante (basicamente o número de indivíduos que são infectados com a doença e o de indivíduos que se curam é igual), até que S se distancia de N e a nossa aproximação deixa de valer e o número de infectados começa a decair até chegar em zero. No caso em que $\beta < \gamma$, então $\frac{dI}{dt} < 0$ e o número de

infectados rapidamente cai para zero. Apenas quando $\beta > \gamma$, $\frac{dI}{dt} > 0$ havendo uma transmissão sustentada da doença fazendo com que o número de infectados aumente consistentemente, atingindo um pico em algum momento como em uma curva epidêmica de infecções padrão. A partir desta análise podemos concluir que para o modelo SIR, a taxa básica de reprodução, R_0 é

$$R_0 = \frac{\beta}{\gamma} \quad (2.26)$$

2.3.3 Modelo SIRD

Esta classe de modelos compartimentais diverge do modelo SIR abordado na seção 2.3.1 pois no presente caso, a taxa de mortalidade da doença agora é levada em conta e temos além dos indivíduos recuperados, os indivíduos mortos; este é o modelo compartimental mais simples capaz de descrever a evolução de uma epidemia que tem capacidade de tirar a vida de indivíduos infectados pelo patógeno. Portanto, um indivíduo contaminado pela doença pode tanto se recuperar depois de um certo período de $1/\gamma$ dias ou falecer em decorrência da doença com taxa de fatalidade δ . Para esta classe de modelos, o sistema de EDOs se torna

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \quad (2.27)$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I - \delta I \quad (2.28)$$

$$\frac{dR}{dt} = \gamma I \quad (2.29)$$

$$\frac{dD}{dt} = \delta I. \quad (2.30)$$

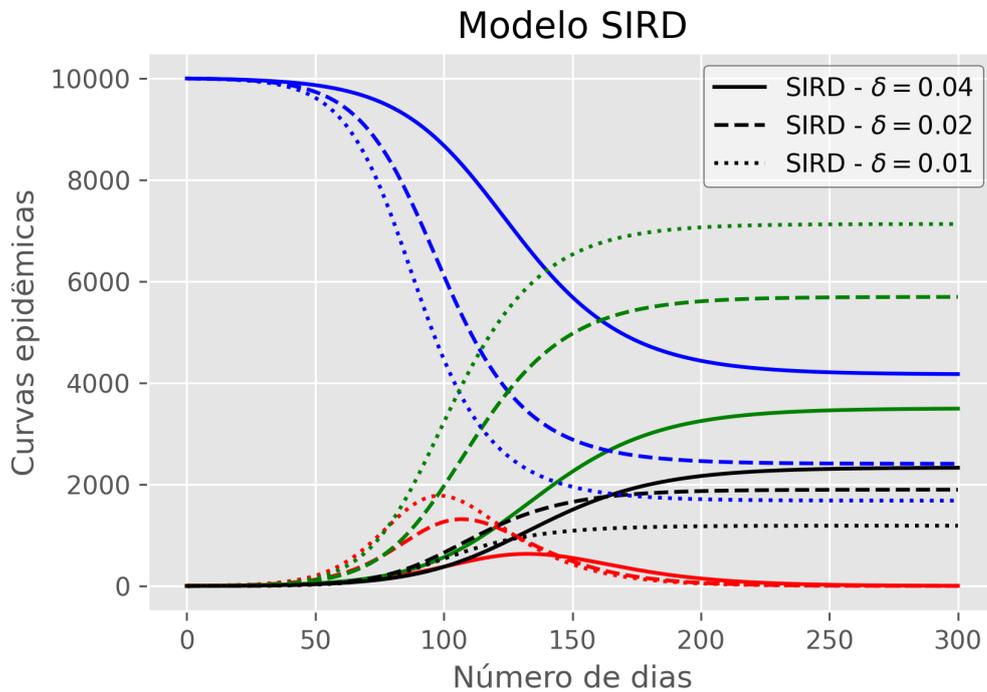
Para o modelo SIRD, repetindo um procedimento análogo ao feito para o modelo SIR na equação 2.28, podemos mostrar que a taxa básica de reprodução agora depende tanto da taxa de mortalidade δ quanto da taxa de recuperação γ da seguinte forma:

$$R_0 = \frac{\beta}{\gamma + \delta}, \quad (2.31)$$

o que faz bastante sentido se analisarmos que uma doença que tenha uma taxa de recuperação ou mortalidade muito elevada faz com que os indivíduos infectados não gerem uma transmissão continuada, pois os indivíduos se curam rapidamente ou falecem em decorrência da patologia, e a epidemia tende a perder força e acabar rapidamente.

Na figura 13 é mostrada uma comparação das curvas epidêmicas do modelo SIRD; na cor azul temos as curvas epidêmicas para os suscetíveis, em vermelho temos as curvas de

Figura 13 – Comparação das curvas epidêmicas de três integrações numéricas para o modelo SIRD com taxa de transmissão $\beta = 0,15$ e taxa de recuperação $\gamma = 0,06$ mantidas fixas e taxa de mortalidade variando nos valores $\delta = 0,04$, $\delta = 0,02$ e $\delta = 0,01$.



infectados, em verde são mostradas as curvas para o compartimento dos recuperados e em preto a curva referente ao compartimento dos mortos. Foram feitas três integrações numéricas das EDOs do modelo com população total $N = 10000$ indivíduos, com parâmetros $\beta = 0,15$ e $\gamma = 0,06$ fixos, enquanto a taxa de mortalidade foi variada nos valores $\delta = 0,04$, representado pelas curvas sólidas para cada compartimento, $\delta = 0,02$, representado pelas curvas tracejadas em cada compartimento e $\delta = 0,01$, representado pelas curvas pontilhadas. A variação da taxa de mortalidade muda de forma significativa o formato das curvas epidêmicas; quando $\delta = 0,04$ temos uma quantidade expressiva de indivíduos morrendo em decorrência da infecção e, conseqüentemente, a curva de infectados tem um pico largo e pouco pronunciado. No cenário oposto, quando $\delta = 0,01$, a taxa de mortalidade é relativamente baixa e mais pessoas ficam infectadas com a doença simultaneamente, o que pode ser visto através da curva de infectados como um pico estreito e bem pronunciado.

2.3.4 Mapa entre os modelos SIRD e Richards

Vamos considerar agora um modelo SIRD modificado, seguindo a referência (WANG; WU; YANG, 2012), em que nas equações 2.27 e 2.28 o número total de indivíduos do sistema é substituído pela população parcial dos compartimentos S e I combinados. Isto leva em consideração o fato de que os indivíduos mortos e os recuperados (assumindo que se tornem imunes) não contribuem mais para a dinâmica da transmissão da doença. Portanto, o novo sistema de equações diferenciais do modelo se torna:

$$\frac{dS}{dt} = -\frac{\beta SI}{S+I} \quad (2.32)$$

$$\frac{dI}{dt} = \frac{\beta SI}{S+I} - \gamma I - \delta I \quad (2.33)$$

$$\frac{dR}{dt} = \gamma I \quad (2.34)$$

$$\frac{dD}{dt} = \delta I. \quad (2.35)$$

A taxa básica de reprodução R_0 pode ser calculada da mesma forma para esta versão modificada do modelo SIRD. Desta forma, temos novamente

$$R_0 = \frac{\beta}{\gamma + \delta}. \quad (2.36)$$

O próximo passo no nosso desenvolvimento é definir a variável $y(t) = S(t) + I(t)$ e dividir a equação (2.33) pela equação (2.32), obtendo

$$\frac{dI}{dS} = -1 + \frac{(\gamma + \delta)(S + I)}{\beta S}. \quad (2.37)$$

Reconhecendo o termo R_0 na equação acima e fazendo algumas manipulações algébricas, ficamos com

$$\frac{dy}{y} = \frac{1}{R_0} \frac{dS}{S}. \quad (2.38)$$

Integrando ambos os lados da equação (2.38), coletando as constantes de integração em um único termo e inserindo em (2.32), chegamos a uma equação de crescimento do tipo Richards

$$\frac{dS}{dt} = -\beta S \left[1 - \left(\frac{S}{L} \right)^\alpha \right], \quad (2.39)$$

onde $\alpha = 1 - 1/R_0$ e $L = [S(0) + I(0)]^{1/\alpha} S(0)^{1-1/\alpha}$. Agora que temos uma equação de crescimento do tipo Richards para $S(t)$ e conseqüentemente para $D(t)$, queremos aproximar esta curva ao próprio modelo de Richards $C(t)$, dado pela equação (2.1) e encontrar uma

relação entre os parâmetros dos dois modelos. Para isto, começamos impondo as condições de contorno $K = D(\infty)$ e $t_c = t_i$, onde t_i é o ponto de inflexão da curva $D(t)$. Pela definição de ponto de inflexão, $\ddot{D}(t_i) = 0$, e por consequência da equação (2.35), $\dot{I}(t_i) = 0$.

Além das condições de contorno mencionadas acima, fazemos a imposição de que no ponto $t = t_i$, tanto $C(t)$ quanto sua derivada temporal, $\dot{C}(t)$, coincidam respectivamente com $D(t)$ e $\dot{D}(t)$, ou seja

$$C(t_c) = D(t_i); \quad (2.40)$$

$$\dot{C}(t_c) = \dot{D}(t_i). \quad (2.41)$$

O instante de tempo t_i pode ser calculado explicitamente a partir da solução da equação (2.39), mas por brevidade este cálculo será omitido no texto desta tese.

Através das equações (2.40) e (2.41), juntamente com (2.3) e sua derivada temporal, podemos encontrar a conexão entre os parâmetros (r, α) do modelo de Richards e os parâmetros (β, δ, γ) do modelo compartimental SIRD

$$\frac{1}{(1 + \alpha)^{1/\alpha}} = \frac{D(t_i)}{D(\infty)}, \quad (2.42)$$

$$\frac{r\alpha}{1 + \alpha} = \delta \frac{I(t_i)}{D(t_i)}, \quad (2.43)$$

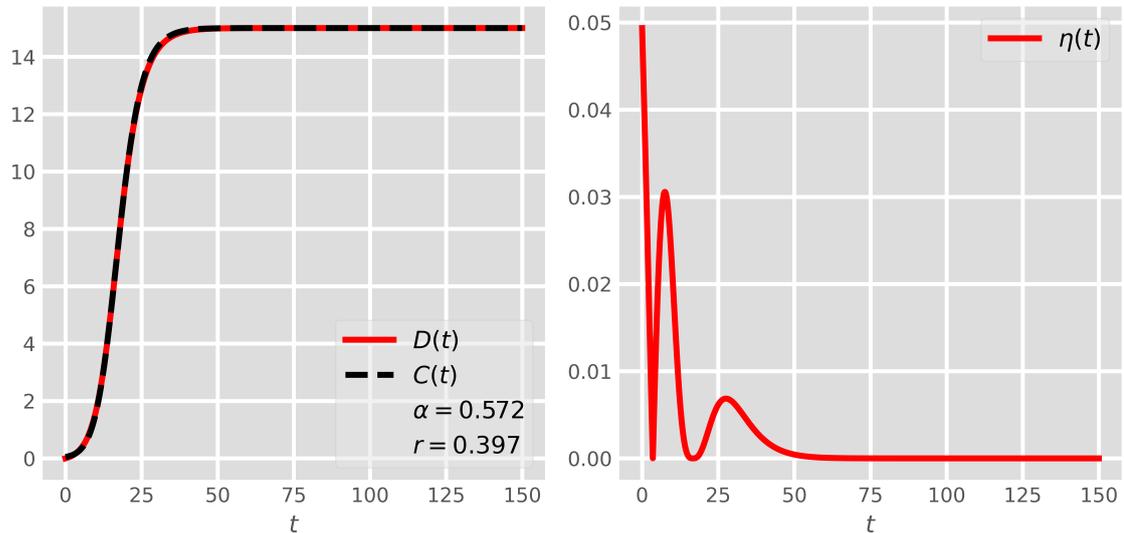
onde usamos a condição de contorno $K = D(\infty)$ e usamos na segunda equação acima o resultado obtido na primeira para simplificação. Estas equações compõem um mapa entre os parâmetros dos dois modelos e uma vez tendo os parâmetros do modelo compartimental SIRD em mãos juntamente com os pontos t_i e $D(\infty)$, podemos encontrar univocamente os parâmetros r e α do modelo de Richards.

Agora que temos uma relação entre os parâmetros dos dois modelos, podemos estimar a precisão do referido mapa através da função erro relativo, dada por:

$$\eta(t) = \frac{|C(t) - D(t)|}{D(t)}. \quad (2.44)$$

Na figura 14, o gráfico do lado esquerdo exibe em a curva de mortes acumuladas $D(t)$ de uma realização do modelo SIRD (curva contínua em vermelho) sobreposta pela curva de mortes acumuladas $C(t)$ do modelo de Richards (curva tracejada em preto) com parâmetros (r, α) obtidos pelo mapa contruído entre os dois modelos. Note que a concordância entre as duas curvas é excelente, o que é corroborado pelo plot da função erro relativo $\eta(t)$ para este mesmo caso, exibido no gráfico do lado direito. A função erro relativo tem um valor máximo

Figura 14 – O gráfico da esquerda ilustra a curva de mortes acumuladas $D(t)$ de uma realização do modelo SIRD (curva contínua em vermelho) juntamente com a curva de mortes acumuladas $C(t)$ do modelo de Richards (curva tracejada em preto) com parâmetros obtidos a partir do mapa entre os dois modelos. No gráfico da direita é apresentado um plot da quantidade $\eta(t)$ para este mapa.



Fonte: elaborada pelo autor e publicada em (MACÊDO et al., 2021).

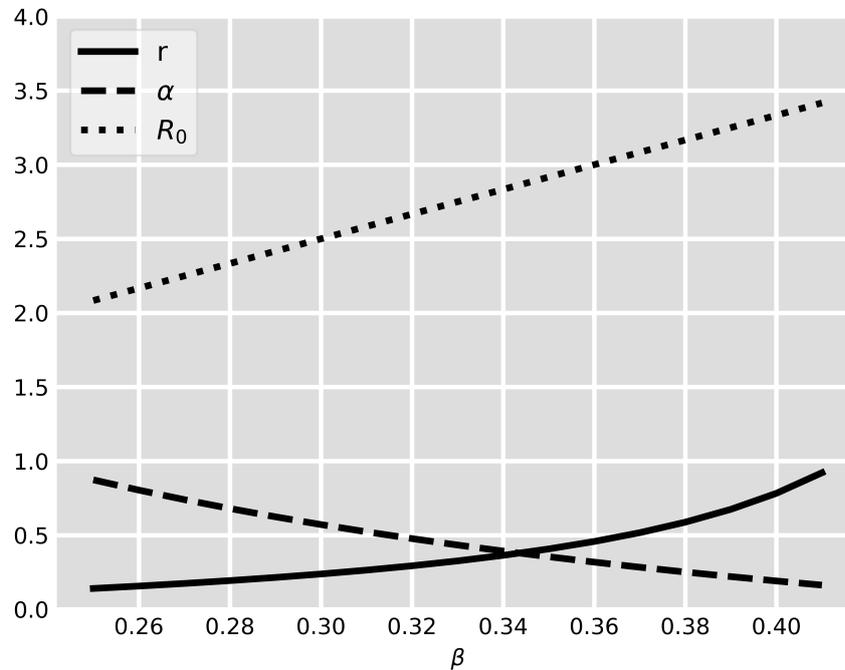
de 5%, logo nos primeiros pontos em que é feita a comparação, mas rapidamente decai para muito próximo de zero nos pontos seguintes. O segundo e o terceiro picos no gráfico do lado direito da figura 14 coincidem com os "joelhos" da curva acumulada, que correspondem ao crescimento exponencial e à aproximação exponencialmente rápida ao platô de saturação.

A figura 15 mostra o comportamento dos parâmetros r (representado na figura pela curva contínua) e α (curva tracejada no gráfico) do modelo de Richards e da taxa básica de reprodução R_0 (curva pontilhada no gráfico) em função do parâmetro β do modelo SIRD mantendo-se fixos os parâmetros γ e δ . O valor do parâmetro β for variado de forma a manter os parâmetros r e α dentro do seu limite biologicamente aceitável, ou seja, entre 0 e 1. Notamos ainda o respectivo crescimento e decréscimo monotônico dos parâmetros com a variação de β .

2.3.5 Modelo SIRD com parâmetros dependentes do tempo

O modelo SIRD tradicional com parâmetros constantes, dado pelas equações (2.27)-(2.30), é extremamente comum na área de epidemiologia para a modelagem das mais diversas doenças. No entanto, o modelo SIRD tradicional exibe um crescimento muito rápido depois de um certo ponto característico e também tem uma aproximação exponencialmente rápida ao platô de saturação. Estas características inerentes do modelo SIRD tradicional são incompatíveis com

Figura 15 – Comportamento dos parâmetros do modelo de Richards r (curva contínua) e α (curva tracejada), obtidos através do mapa, e do taxa básica de reprodução R_0 (curva pontilhada) em função do parâmetro β do modelo SIRD com γ e δ fixos.



Fonte: elaborada pelo autor e publicada em (MACÊDO et al., 2021).

as curvas de casos e mortes acumuladas na pandemia de Covid-19, nas quais uma plethora de ações envolvendo a obediência ou desobediência de intervenções não farmacológicas (como uso de máscaras, fechamento de escolas e locais públicos como lojas e centros de compras ou ainda lockdown), faz com que a curva tenha um crescimento sub-exponencial (polinomial) e uma lenta aproximação ao platô de saturação (lei de potência), como mostrado em (VASCONCELOS et al., 2021).

Esta incompatibilidade entre o modelo SIRD tradicional com parâmetros constantes e as curvas epidêmicas da Covid-19 pode ser vista na figura 16. A curva de mortes acumuladas atribuídas à Covid-19 para a Suécia até a data de 30 de Julho de 2020 é exibida em círculos vermelhos, enquanto a curva de mortes acumuladas $D(t)$ do modelo SIRD padrão com parâmetros constantes e obtidos através do melhor ajuste numérico, exibidos no inset da figura juntamente com os erros da estimativa.

Uma solução simples e que, como veremos mais a frente, se provou extremamente eficiente para a modelagem das curvas epidêmicas da Covid-19 foi introduzir uma dependência temporal

no parâmetro epidemiológico β de acordo com a seguinte equação diferencial (CACCAVO, 2020):

$$\frac{d\beta}{dt} = \frac{1}{\tau_1}[\beta_0\beta_1 - \beta(t)], \quad t \geq \tau_0, \quad (2.45)$$

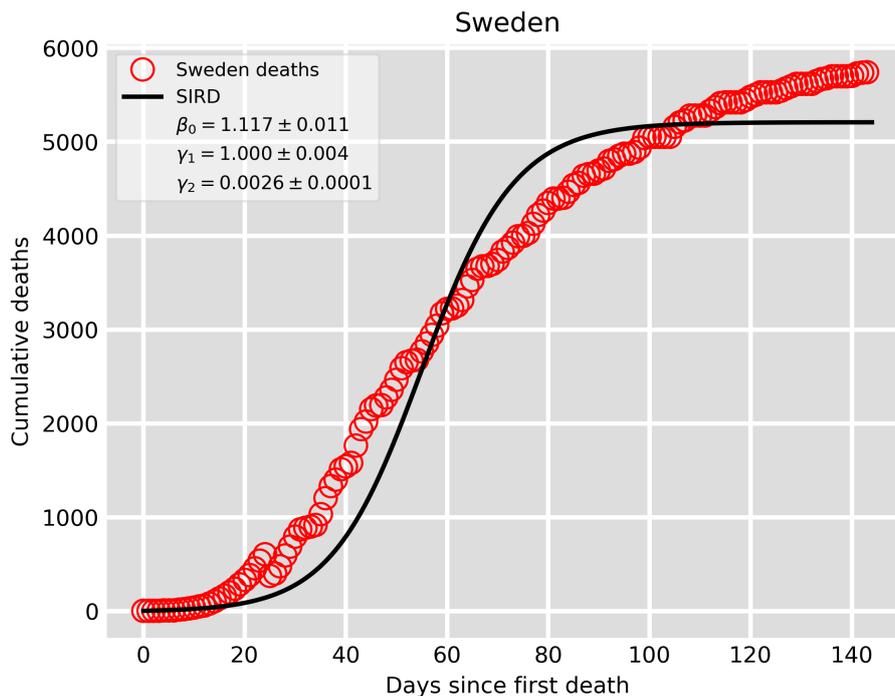
onde τ_0 é o tempo em que se iniciam as intervenções, τ_1 é o tempo médio de duração das intervenções, β_0 é o valor inicial da taxa de transmissão antes das intervenções e o produto $\beta_0\beta_1$ representa seu valor final no final da pandemia. Integrando a equação diferencial (2.45) com a condição inicial $\beta(t) = \beta_0$, ficamos com

$$\beta(t) = \begin{cases} \beta_0, & t < \tau_0; \\ \beta_0 \left(\beta_1 + (1 - \beta_1)e^{-\frac{(t-\tau_0)}{\tau_1}} \right), & t \geq \tau_0. \end{cases} \quad (2.46)$$

Mesmo com esta mudança na taxa de transmissão β , as equações centrais do mapa (2.42) e (2.43), que nos permitem calcular os parâmetros r e α do modelo de Richards a partir da curva $D(t)$, continuam válidas. No entanto, o instante de tempo t_i deve agora ser determinado a partir do pico da curva $I(t)$ ou do ponto de inflexão da curva de mortes acumuladas $D(t)$.

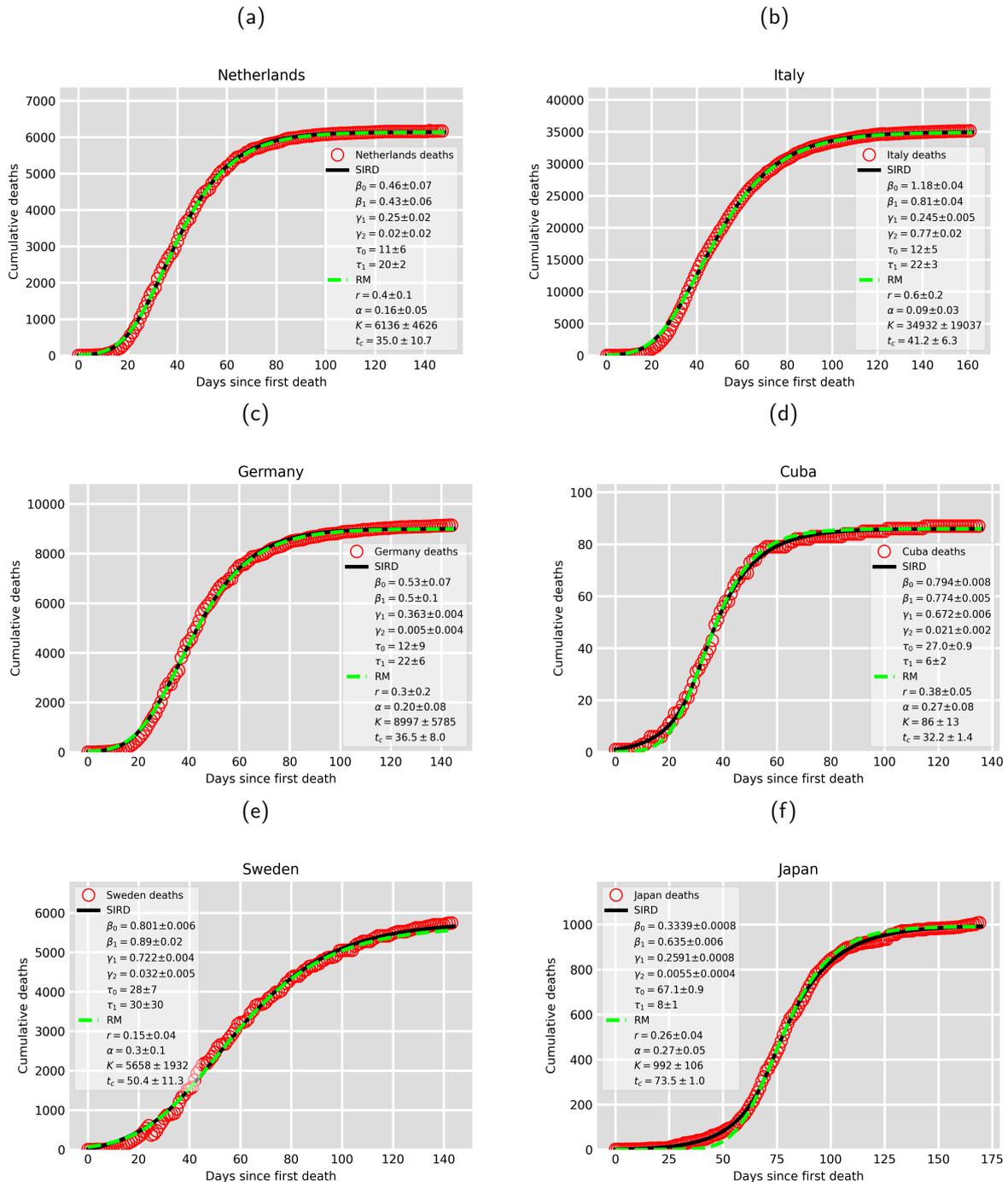
A figura 17 mostra a aplicação do mapa entre o modelo SIRD modificado com parâmetro β dependente do tempo e o modelo de Richards para Holanda, Itália, Alemanha, Cuba, Suécia

Figura 16 – Curva de mortes acumuladas atribuídas à Covid-19 (círculos vermelhos) para a Suécia até 30 de Julho de 2020 juntamente com a curva de mortes acumuladas do modelo SIRD padrão com parâmetros obtidos através do melhor ajuste numéricos, cujas estimativas e erros são exibidos no inset da figura.



Fonte: elaborada pelo autor e publicada em (MACÊDO et al., 2021).

Figura 17 – Número acumulado de mortes (círculos vermelhos) atribuídos à Covid-19 até 30 de Julho de 2020 para (a) Holanda, (b) Itália, (c) Alemanha, (d) Cuba, (e) Suécia, e (f) Japão. As curvas contínuas em preto representam o melhor ajuste numérico pelo modelo SIRD modificado com parâmetro β dependente do tempo e cujos valores estimados dos parâmetros estão descritos na legenda da figura. As curvas tracejadas em verde representam a curva teórica obtida pelo modelo de Richards com os parâmetros computados numericamente através do mapa dado pelas equações (2.42) e (2.43) e os parâmetros do modelo SIRD encontrados pelo ajuste numérico.



Fonte: elaborada pelo autor e publicada em (MACÊDO et al., 2021).

e Japão até o dia 30 de Julho de 2020. Em todos os casos, o número acumulado de mortes atribuídas à Covid-19 em função do número de dias desde a primeira morte para cada país é denotado através dos círculos vermelhos. A curva contínua em preto a curva acumulada de

mortes do modelo SIRD modificado com β dependente do tempo obtida através do melhor ajuste numérico com os dados empíricos; a curva tracejada em verde representa a curva acumulada gerada pelo modelo de Richards com parâmetros obtidos através do mapa entre os dois modelos.

A partir desta figura, podemos ver que a correspondência entre estes dois modelos através do mapa discutido anteriormente é muito boa para todos os casos apresentados. Isto nos leva a concluir que, apesar de os parâmetros do modelo de Richards não terem uma interpretação epidemiológica direta como no caso do modelo compartimental SIRD, vemos que estes ainda carregam informações importantes sobre as características epidemiológicas da doença em questão, visto que possuem uma relação (ainda que não linear) com os parâmetros do modelo SIRD.

Perceba ainda que o modelo de Richards, que possui apenas dois parâmetros independentes do tempo, consegue descrever as curvas epidêmicas em questão tão bem quanto o modelo SIRD modificado com parâmetro β dependente do tempo, o que corresponde efetivamente a um total de 6 parâmetros ($\beta_0, \beta_1, \gamma, \delta, \tau_0, \tau_1$). Portanto, ao passo que perde-se um pouco da interpretação epidemiológica ao se fazer uso do modelo de Richards, ganha-se em flexibilidade e simplicidade computacional.

Todos os ajustes numéricos apresentados na figura 17 foram feitos utilizando o algoritmo de Levenberg-Marquardt (MORÉ, 1978) implementado no módulo `lmfit` (NEVVILLE et al., 2014) da linguagem de programação Python, em que também foram produzidos todos os gráficos desta seção. Para cada ajuste numérico, o módulo `lmfit` foi utilizado para solucionar o problema de mínimos quadrados não linear a partir dos dados empíricos e o modelo SIRD modificado com a finalidade de achar o melhor conjunto de parâmetros para modelar os dados juntamente com os erros associados a cada parâmetro, que são calculados através do método da matriz de covariância.

Para um determinado conjunto de parâmetros do modelo SIRD modificado, os erros associados aos parâmetros do modelo de Richards obtidos através do mapa constituído pelas equações (2.42) e (2.43) são encontrados através de um procedimento de *bootstrap*. Tal procedimento consiste em considerar os valores extremos dos parâmetros obtidos através do ajuste numérico dos dados empíricos com o modelo SIRD modificado, $\chi \pm \delta\chi$, onde χ é o valor do parâmetro encontrado pelo ajuste e $\delta\chi$ seu erro associado. Em seguida é feita uma integração do modelo SIRD modificado para cada uma das combinações possíveis dos valores extremos – um total de 64 – e em cada integração é aplicado o mapa para calcular os parâ-

metros do modelo de Richards (r, α) , que são armazenados se estiverem no intervalo aceitável $(0 < r, \alpha \leq 1)$. Os erros associados a cada parâmetro são calculados através da média do valor absoluto da diferença entre os parâmetros encontrados em cada uma das iterações discutidas e os valores encontrados no mapa feito inicialmente.

3 MODINTERV: UMA FERRAMENTA ONLINE PARA MONITORAMENTO DA COVID-19

Nos últimos dias do ano de 2019 o mundo se deparou com as notícias de uma nova doença que começara a se disseminar na China e que pouco tempo depois tomaria conta de todo o mundo, vitimando milhões de pessoas e mudando para sempre a vida, e até certo ponto, o funcionamento da sociedade. Um surto de proporções globais como a pandemia de Covid-19 não era visto desde o início do século XX, quando o mundo testemunhou a letalidade da gripe espanhola, causada pelo vírus influenza, durante os quase três anos de pandemia.

O espalhamento de uma doença infectocontagiosa é extremamente rápido e a dinâmica por trás da disseminação é bastante complexa e envolve diversas escalas, como a transmissão nos núcleos familiares, nos bairros, nas cidades, a nível estadual e assim por diante. Uma pandemia pode ser dividida em diversas fases, cada uma com suas respectivas possíveis respostas e intervenções por parte do poder público e da população em geral; no início acontece a detecção dos primeiros casos e mortes, em seguida são analisados a transmissibilidade e a letalidade da doença, que sucede o monitoramento do espalhamento da doença juntamente com a implementação de intervenções não-farmacológicas (distanciamento social, uso obrigatório de máscaras, fechamento de escolas e locais públicos, lockdown) e farmacológicas (uso de medicamentos eficazes contra a doença, vacinação), se disponíveis.

Durante o monitoramento da pandemia de Covid-19, vários recursos modernos foram utilizados para auxiliar esta tarefa, a exemplo de dados vinculados a localizações geográficas como estatísticas a níveis de estado, município e até bairros; o uso de smartphones para informar novos casos, teleatendimento médico no caso de sintomas pertinentes à Covid-19 e dados de geolocalização dos indivíduos a fim de estimar a parcela da população que estava seguindo as orientações de distanciamento social e lockdown. Uma outra vertente crucial no monitoramento do espalhamento da doença foi a análise de dados (CORI et al., 2017) juntamente com a modelagem da epidemia utilizando-se diversos modelos (compartimentais, logísticos, de agentes) para entender a dinâmica específica de cada local para assim traçar medidas eficazes no combate da pandemia.

Uma das classes de dados mais significativas e que foi amplamente explorada pela mídia internacional são as curvas epidêmicas, constituídas pelo número diário ou acumulado de casos ou mortes atribuídas à Covid-19 em função do tempo (em sua maioria, do número de dias desde o início da pandemia ou do primeiro caso ou morte naquela localidade) (KAMVAR et al.,

2019). As curvas epidêmicas são uma ótima ferramenta para monitorar um surto de qualquer que seja a doença pois serve como uma ilustração visual da dinâmica da epidemia, evidenciando o crescimento ou decréscimo do número diário de casos ou mortes e o resultado das intervenções implementadas pelas autoridades de saúde pública; além disso, elas são a base para a aplicação de diversos tipos de modelos epidemiológicos, como os modelos compartimentais (KERMACK; MCKENDRICK, 1927; MATHEMATICAL... , 2020), modelos de agentes (KERR et al., 2021) e modelos fenomenológicos (VASCONCELOS et al., 2020; VASCONCELOS et al., 2021; MACÊDO et al., 2021).

O aplicativo ModInterv foi desenvolvido pela Rede Cooperativa de Pesquisa em Modelagem da Epidemia de Covid-19 e Intervenções Não-Farmacológicas (ModInterv-Covid19) (MODINTERV-COVID19, 2020) com o intuito de prover um aplicativo amplamente disponível ao público, autoridades de saúde pública e outros pesquisadores, de fácil uso e grande abrangência. O aplicativo permite ao usuário monitorar as curvas epidêmicas de casos e mortes atribuídas à Covid-19 para vários países ao redor do mundo e para estados e cidades no Brasil e nos Estados Unidos. O ModInterv traz em sua implementação uma classe generalizada de modelos logísticos de crescimento com parâmetros dependentes do tempo que são capazes de modelar curvas epidêmicas com um número arbitrário de ondas de infecções ou mortes através de ajustes numéricos realizados automaticamente pelo software.

Uma vez que o usuário acessa o aplicativo, ele deve escolher o tipo de curva epidêmica que deseja analisar, casos ou mortes; em seguida deverá ser escolhido o país, estado ou cidade do Brasil ou dos Estados Unidos que servirá de base para o modelo. O software então prossegue para a determinação do número de ondas epidêmicas apresentadas na curva epidêmica. Este procedimento de determinação do número de ondas é feito através de um algoritmo que começa com a suavização dos dados através da aplicação de uma regressão do tipo LOWESS (CLEVELAND, 1981). Com os dados suavizados, um módulo de processamento de sinais nativo da linguagem python é utilizado para detectar máximos e mínimos locais na curva diária suavizada, o que servirá como indicativo de possíveis ondas na curva epidêmica. Em seguida, um filtro implementado com base no algoritmo (HARVEY et al., 2023) é utilizado para detectar e eliminar ondas espúrias (ondas com duração muito curta, ondas muito "rasas" e também falsas ondas detectadas erroneamente por máximos ou mínimos locais causados por grandes flutuações nos dados diários).

Depois que as ondas espúrias são eliminadas e o número N de ondas "reais" é determinado, o modelo usa os dados empíricos juntamente com os intervalos de cada onda determinados

pelo filtro para fazer o ajuste numérico com o o modelo logístico de crescimento para múltiplas ondas com parâmetros dependentes do tempo. Sendo o ajuste bem sucedido, as curvas epidêmicas diária e acumulada para os dados escolhidos são exibidas juntamente com as curvas teóricas obtidas através do modelo com os parâmetros encontrados pelo ajuste numérico. Nos gráficos são indicados com símbolos as datas de início e pico de cada onda epidêmica (e seus respectivos valores indicados na legenda da figura); neles também são traçadas linhas verticais de diferentes cores para indicar visualmente o regime de aceleração que a epidemia apresentava naquele dado momento. Comparando a posição do último ponto da curva com os pontos que marcam estes regimes de aceleração, o aplicativo ModInterv pode determinar o regime de aceleração atual para a região escolhida pelo usuário, permitindo inferir se a epidemia naquele local tende a continuar acelerando ou se está em regime de desaceleração.

Os detalhes de todas as características de implementação do aplicativo, como construção da interface do usuário, hospedagem em servidores, obtenção dos dados brutos, ajuste numérico, processamento e análise dos dados serão discutidos cuidadosamente ao longo deste capítulo. Inicialmente serão elaborados as definições e cálculos para a determinação dos estágios das curvas epidêmicas e em seguida serão discutidos os tópicos voltados à parte computacional da implementação do aplicativo.

3.1 ESTÁGIOS DE UMA CURVA EPIDÊMICA

As curvas epidêmicas, sejam elas diárias ou acumuladas, de casos ou de mortes, são uma coleção de dados empíricos que servem como um ótimo indicativo visual da evolução da epidemia em um dado local. Conhecer as características de tais curva é assunto de tremenda importância, pois pode auxiliar as autoridades de saúde a entender o comportamento que a curva tende a ter nos dias que se sucedem, embasando a implementação, fortalecimento ou relaxamento de medidas de enfrentamento à pandemia. O aplicativo ModInterv utiliza modelos fenomenológicos generalizados de crescimento com parâmetros dependentes do tempo (discutidos previamente na seção 2.2), que são capazes de modelar curvas epidêmicas com múltiplas ondas de infecções ou mortes de forma excelente, como mostrado na seção 2.2.1.

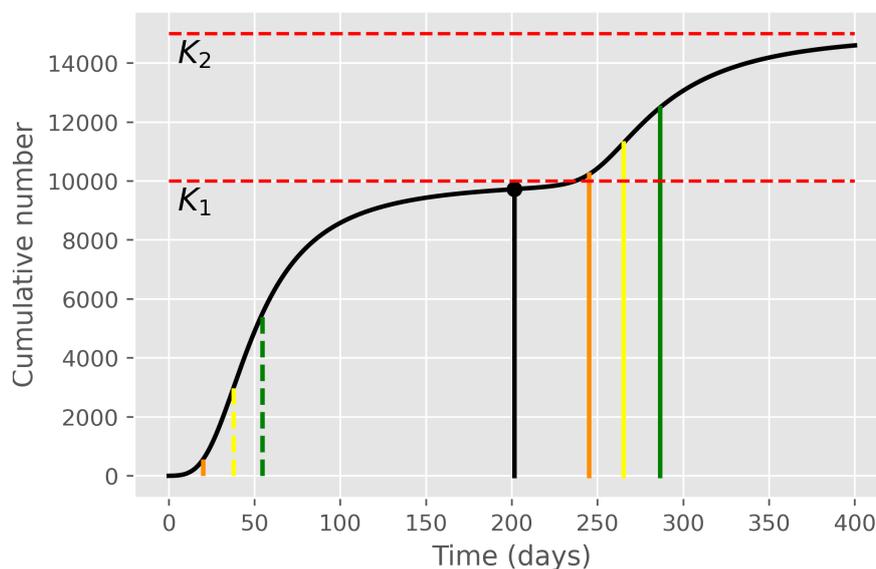
Através da análise das curvas epidêmicas com os modelos de crescimento, o aplicativo ModInterv fornece informações que não poderiam ser obtidas através de médias móveis ou outros tipos de suavização de curvas, como será mostrado ao longo desta seção. De forma geral, as curvas epidêmicas exibem três regimes de crescimento: uma fase inicial de crescimento rápido

e acelerado; uma fase intermediária com crescimento aproximadamente linear e finalmente uma fase de crescimento em que a curva se afasta do crescimento linear e tende a um platô de saturação, aproximando-se do valor final de casos ou mortes. É possível fazer uma divisão mais detalhada das fases de crescimento através de considerações sobre a aceleração da curva epidêmica.

Para manter a generalidade, será considerada inicialmente uma curva epidêmica com $N \geq 1$ ondas. As figuras esquemáticas 18, 19 e 20 ilustram a curva acumulada, a sua velocidade (primeira derivada) e a sua aceleração (segunda derivada), respectivamente, para uma curva com duas ondas de infecção ($N = 2$). O início de cada onda é marcado por um regime de *aceleração crescente*, quando a aceleração vai de um valor próximo de zero até o seu máximo, instante denotado por $t_1^{(k)}$, onde $k = 1, \dots, N$ se refere a onda em questão, e são representados para as diferentes ondas na figura por linhas verticais na cor laranja. Em seguida, a curva epidêmica entra em uma fase intermediária que é composta por dois estágios epidêmicos com regimes de aceleração distintos.

O primeiro deles apresenta um regime de *aceleração decrescente*, no qual a aceleração diminui do seu máximo em $t_1^{(k)}$ até chegar em zero no instante $t_2^{(k)}$, indicados para cada onda na figura 18 pelas linhas verticais amarelas. Em seguida a curva entra em um regime de desaceleração crescente até atingir seu pico no instante $t_3^{(k)}$, indicado para cada onda na figura

Figura 18 – Figura esquemática de uma curva epidêmica apresentando duas ondas de infecções. As linhas verticais coloridas representam os estágios epidêmicos; o ponto e a linha vertical em preto representam o início da segunda onda epidêmica.

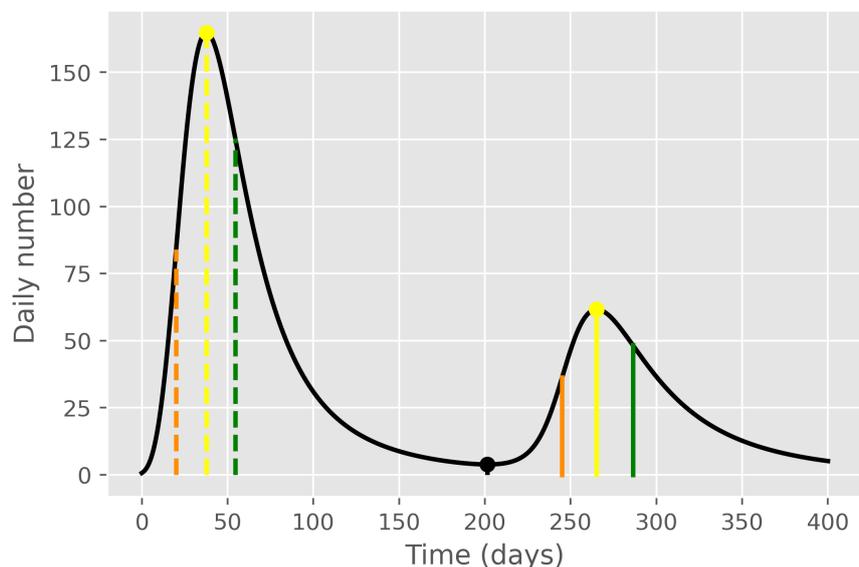


Fonte: elaborada pelo autor e publicada em (BRUM et al., 2023).

pelas linhas verticais na cor verde. Depois de passar do instante $t_3^{(k)}$ a curva entra em sua última fase de crescimento, apresentando um regime de desaceleração decrescente e se aproxima de zero já no fim de cada onda epidêmica. Como já comentado anteriormente, a forma com que a curva epidêmica se aproxima do platô e conseqüentemente a característica da desaceleração, assumindo comportamento de lei de potência ou exponencial, depende de diversos fatores como as medidas de restrição imposta pelas autoridades de saúde e a adesão da população a estas medidas, índice de vacinação, entre outros. O final de uma onda epidêmica acontece quando a velocidade chega a seu mínimo e a aceleração é nula; este instante é denotado nas figuras por um ponto na cor preta. Finalmente, ainda na figura da curva acumulada, temos a indicação do platô que cada onda epidêmica atingiria, caso não houvesse uma onda subsequente. Estes platôs são indicados pelas linhas horizontais tracejadas em vermelho e denotados pelos símbolos K_k . No caso em tela, são apresentados os platôs das duas ondas, K_1 e K_2 .

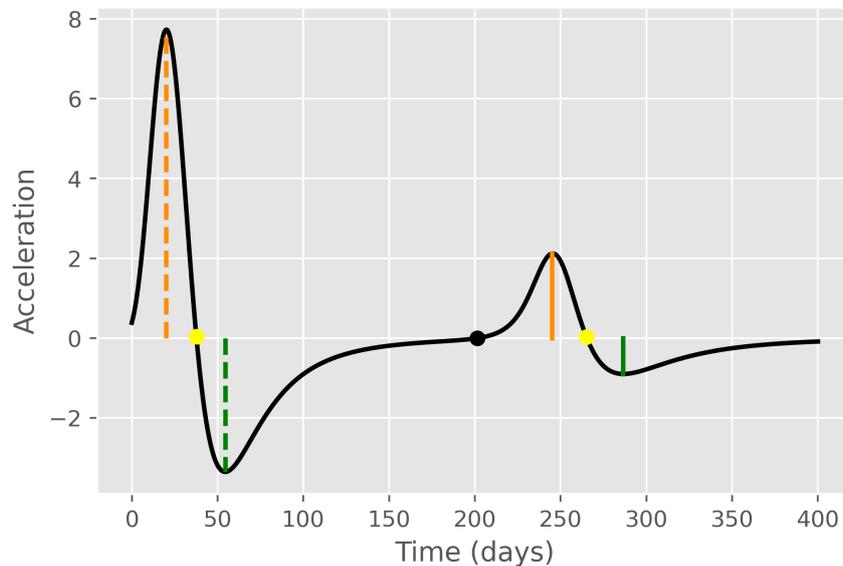
O conhecimento de tais instantes de tempo $t_i^{(k)}$ nos permite determinar o estágio atual da epidemia em uma determinada localidade ao compararmos a posição do último ponto da curva empírica, denominado como tempo atual e representado por t_f , com os pontos característicos calculados pelo modelo. Para determinar estes pontos característicos, o aplicativo realiza um ajuste numérico com o modelo matemático apropriado escolhido automaticamente para ajustar ondas com um número arbitrário de ondas. Depois de realizar o ajuste e calcular os pontos

Figura 19 – Figura esquemática da velocidade (primeira derivada) da curva acumulada mostrada na figura 18.



Fonte: elaborada pelo autor e publicada em (BRUM et al., 2023).

Figura 20 – Figura esquemática da aceleração (segunda derivada) da curva acumulada mostrada na figura 18.



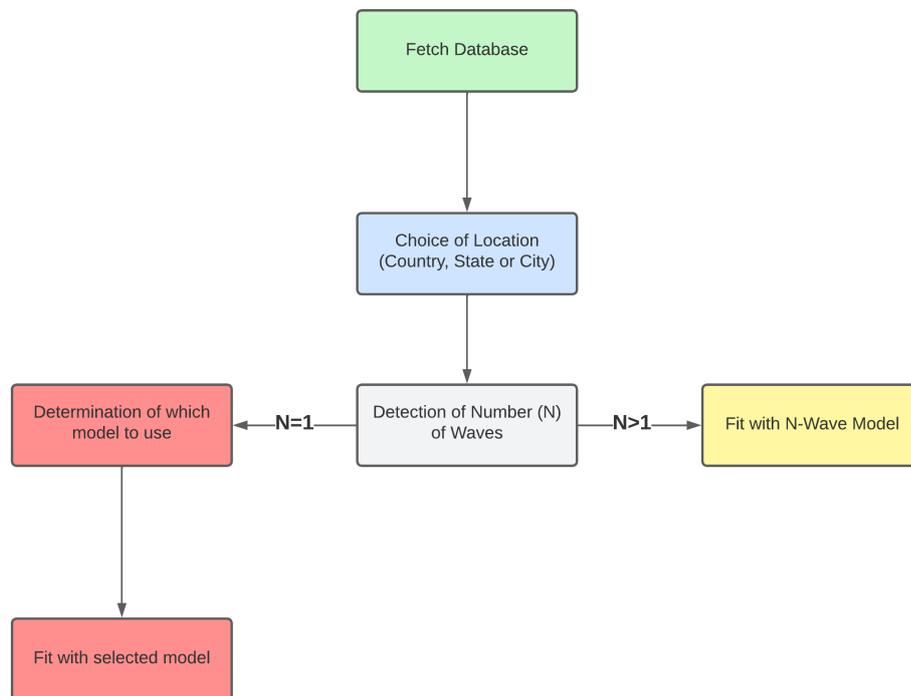
Fonte: elaborada pelo autor e publicada em (BRUM et al., 2023).

característicos (procedimento que será discutido em detalhe nas seções que seguem) o aplicativo determina o estágio atual da curva epidêmica analisada. Por fim é gerado um gráfico exibindo a curva empírica, a curva teórica dada pelo modelo com os parâmetros obtidos pelo melhor ajuste numérico, as linhas verticais representando os instantes descritos acima e na legenda é indicado o atual estágio da epidemia na localidade em questão, que recebe uma das seguintes denominações:

1. Aceleração crescente: $t_f < t_1^{(k)}$;
2. Aceleração decrescente: $t_1^{(k)} < t_f < t_2^{(k)}$;
3. Desaceleração crescente: $t_2^{(k)} < t_f < t_3^{(k)}$;
4. Desaceleração decrescente: $t_f > t_3^{(k)}$.

O aplicativo ModInterv implementa um modelo de crescimento logístico capaz de descrever curvas que apresentam um número arbitrário N de ondas epidêmicas de forma completamente automatizada. Este modelo foi discutido em detalhe na seção 2.2. O funcionamento do ajuste numérico utilizando tal modelo será discutida em detalhes nas seções que seguem.

Figura 21 – Figura esquemática descrevendo o funcionamento do aplicativo ModInterv desde sua inicialização até o ajuste numérico.



Fonte: elaborada pelo autor e publicada em (BRUM et al., 2023).

3.2 ESTRUTURA DE *BACKEND* DO APLICATIVO MODINTERV

Nesta seção serão discutidas a implementação, funcionamento e ferramentas utilizadas na implementação do aplicativo ModInterv. Dentre estes temas serão discutidos como os dados são obtidos, como os ajustes numéricos são realizados com o modelo implementado, a construção da interface do usuário e a estrutura geral do aplicativo. A figura 21 ilustra esquematicamente o processo de funcionamento do aplicativo em questão, desde sua inicialização até chegar no ajuste numérico.

3.2.1 Interface do usuário

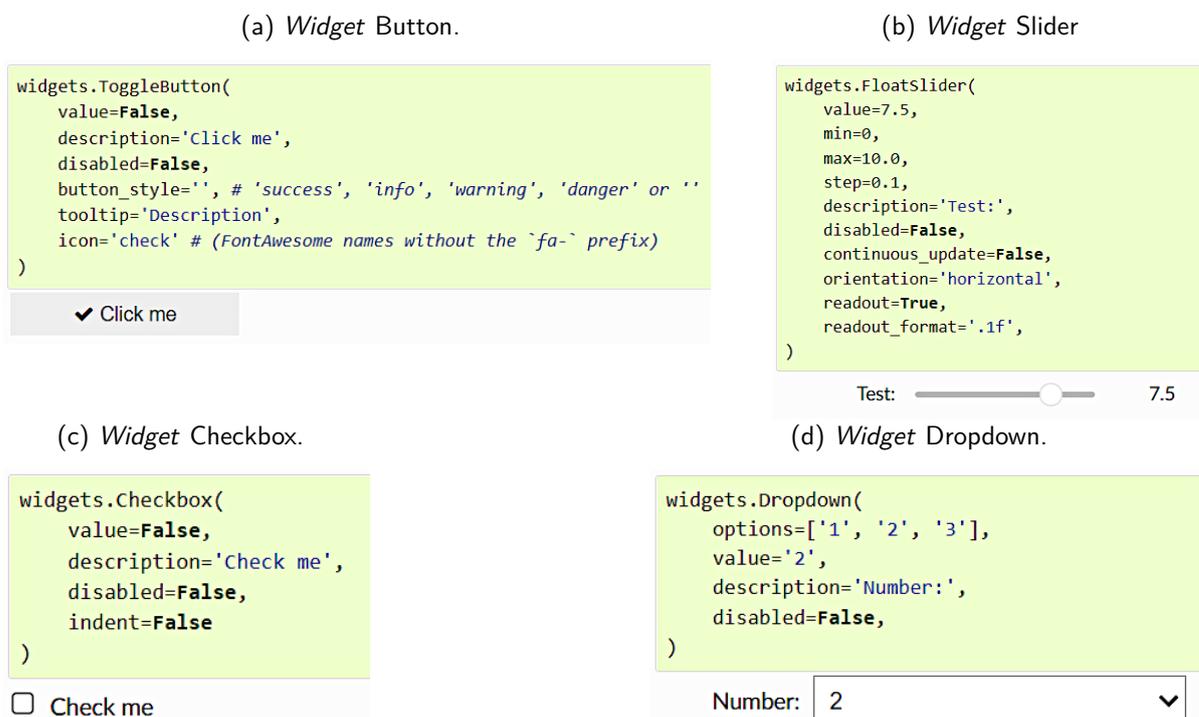
A interface do usuário do aplicativo ModInterv foi desenvolvida com o intuito de fornecer uma experiência acessível para o público em geral sem deixar de contar com uma riqueza de opções de personalizações dos resultados gerados pelo aplicativo. A língua padrão do aplicativo é o inglês, no entanto, o usuário poderá optar pelo português clicando em um botão no topo do site. Desta forma, todo o texto do aplicativo e dos resultados gerados será exibido em

português do Brasil. A interface do usuário do aplicativo ModInterv foi construída usando objetos chamados *widgets*, já implementados no módulo *iPywidgets* (COMMUNITY, 2015) da linguagem Python. Dentre os múltiplos *widgets* disponíveis neste módulo, foram utilizados apenas cinco tipos, a saber: HTML, Button, IntSlider, Dropdown e Checkbox. O primeiro deles, HTML, foi usado para exibir elementos de texto dinâmicos sem a necessidade de recarregar a página (quando o usuário mudar a língua do aplicativo, por exemplo). Os outros 4 *widgets* foram utilizados para receber informações vindas do usuário, como o tipo de curva epidêmica a ser analisada (Casos ou Mortes), o país, estado ou cidade escolhido, o número de dias a ser considerado no ajuste numérico, os gráficos a serem mostrados, dentre outras. A figura 22 ilustra quatro dos cinco tipos de *widget* usados na implementação do ModInterv.

3.2.2 Aquisição de dados

Os dados utilizados para analisar as curvas epidêmicas dos países e dos estados e cidades nos Estados Unidos são obtidos da base de dados disponibilizada ao público pela Johns Hopkins University (JHU, 2020); esta base de dados fornece as séries temporais de casos confirmados e mortes para as localidades mencionadas. Os dados para estados e cidades do Brasil são

Figura 22 – Ilustrações esquemáticas de quatro dos cinco *widgets* do módulo *iPywidgets* usados no aplicativo.



Fonte: elaborada pelo autor e publicada em (BRUM et al., 2023).

obtidos da base de uma base de dados no GitHub mantida por Wesley Cota, da Universidade Federal de Viçosa. Ambas as bases de dados foram atualizadas diariamente até meados de Março, quando a pandemia de Covid-19 foi oficialmente declarada extinta pela Organização Mundial da Saúde. Toda vez que o aplicativo é inicializado, as bases de dados são acessadas automaticamente e os respectivos dados ficam disponíveis para serem analisados pelo usuário.

Nos dois conjuntos de dados, as séries temporais são fornecidas em formato .csv e lidas e armazenadas em *cache* pelo aplicativo sem qualquer tipo de suavização ou pré processamento, uma vez escolhido o tipo de curva epidêmica a ser analisada e a localidade desejada. As bases de dados mencionadas não fornecem dados diários, portanto, para obter as curvas epidêmicas diárias, são subtraídos valores consecutivos da série temporal acumulada. Deve-se ainda notar que eventualmente as autoridades de saúde de certos locais podem fazer correções nas curva epidêmicas, adicionando ou removendo de uma só vez um grande número de casos confirmados ou mortes que antes não haviam sido contabilizadas ou erroneamente atribuídas à doença em questão, o que pode gerar valores negativos ou saltos na curva diária. No entanto, os efeitos de tais correções para a análises realizadas pelo aplicativos são diminutas, visto que os ajustes numéricos são feitos com base nas curvas acumuladas.

3.2.3 Detecção de ondas epidêmicas

A detecção de forma automática do número de ondas presentes em uma curva epidêmica mostra-se um dos pontos mais desafiadores para ser implementado utilizando apenas técnicas matemáticas, estatísticas e computacionais sem o auxílio de inteligência artificial. O que para um humano seria tarefa simples, para um algoritmo que utiliza apenas as ferramentas mencionadas acima é um verdadeiro desafio devido à presença de flutuações de natureza rápida (ruído) e variações de natureza lenta que caracterizam uma onda epidêmica; realizar tal distinção a partir de argumentos matemáticos e estatísticos não é tarefa fácil.

Para resolver tal problema, foi implementado no ModInterv um algoritmo baseado no procedimento de suavização de LOWESS (*Locally Weighted Scatterplot Smoothing*) (CLEVELAND, 1981) juntamente com ferramentas previamente implementadas em um módulo de análise de sinal padrão da linguagem Python para detecção de máximos e mínimos locais, que são utilizados para estimar o número de ondas epidêmicas presentes na curva em questão. Uma vez selecionada a localidade e o tipo de curva desejado pelo usuário (Casos confirmados ou Mortes), a curva diária é obtida através da manipulação discutida anteriormente a partir

da curva acumulada obtida das bases de dados. O procedimento de suavização de LOWESS é então aplicado à curva diária e em seguida o módulo de análise de sinais é utilizado para detectar máximos e mínimos locais na curva diária já suavizada. Os máximos e mínimos locais detectados na curva são de crucial importância, pois representam potenciais picos ou separações entre ondas subsequentes, respectivamente. Uma vez encontrados tais pontos, é importante filtrar e descartar ondas espúrias; isto é feito utilizando um algoritmo proposto em (HARVEY et al., 2023) que descarta as ondas espúrias ao comparar a proximidade entre os máximos e mínimos locais e a altura relativa dos picos. O número N de ondas epidêmicas presentes na curva é então determinado como o número de mínimos locais restantes depois da filtragem acrescido de uma unidade.

É válido notar que a determinação do número de ondas poderia ser deixado alternativamente a cargo do usuário, no entanto esta opção vai de contra o intuito com o qual o aplicativo ModInterv foi idealizado: fornecer uma ferramenta completamente automatizada e de fácil utilização pelo público em geral. Além disso, o algoritmo de detecção de ondas epidêmicas adicionalmente ao número N de ondas presentes na curva escolhida, também fornece chutes iniciais para os tempos de transição t_i entre ondas sucessivas. Desta forma, este trabalho também teria de ser delegado ao usuário, tornando o aplicativo mais complexo de ser usado e passível de gerar ajustes numéricos de baixa qualidade ou que não convirjam, visto que o elevado número de parâmetros do modelo o deixam bastante sensível à escolha dos chutes iniciais quando na ocasião de um ajuste.

3.2.4 Ajustes numéricos

Todos os ajustes numéricos realizados pelo aplicativo ModInterv utilizam o algoritmo de Levenberg-Marquardt, implementado no módulo `lmfit` (NEWVILLE et al., 2014), para resolver o problema de otimização não linear de mínimos quadrados. A implementação de um algoritmo que realiza de forma automatizada ajustes numéricos para curvas epidêmicas com um número arbitrário de ondas é bastante desafiadora, especialmente para um número N grande, devido ao fato que para cada onda epidêmica o modelo possui 5 parâmetros e mais $2(N - 1)$ parâmetros responsáveis pela transição entre os parâmetros de cada onda, ou seja, os tempos de transição t_i e as taxas de transição ρ_i , totalizando $7N - 2$ parâmetros para serem otimizados pelo ajuste numérico. O elevado número de parâmetros livres do modelo matemático utilizado traz consigo o problema de *overfitting* que demanda bastante cuidado na hora de selecionar os

chutes iniciais e torna o ajuste bastante sensível a tal escolha. Para mitigar este problema, foi implementado no aplicativo ModInterv uma rotina de ajuste numérico recursivo com N passos, como será descrito em seguida.

O procedimento é iniciado utilizando um chute inicial para o tempo de transição t_1 entre a primeira e a segunda onda fornecido pela rotina que detecta o número de ondas epidêmicas presentes na curva em questão. Então é feito um ajuste numérico com a curva acumulada truncada até o tempo t_1 com o modelo de uma onda utilizando chutes iniciais arbitrários dentro dos seus limites de definição para os parâmetros $\{r_1, \alpha_1, q_1, K_1, p_1\}$.

Em seguida, um chute também fornecido pela rotina de detecção de ondas epidêmicas para o tempo de transição t_2 entre a segunda e terceira ondas é utilizado para truncar a curva acumulada em análise e um novo ajuste numérico é realizado até tal data. Os parâmetros obtidos na primeira etapa do procedimento descrito são usados como chutes iniciais para $\{r_1, \alpha_1, q_1, K_1, p_1\}$ no ajuste da curva com duas ondas, enquanto os chutes iniciais para $\{r_2, \alpha_2, q_2, K_2, p_2\}$ são novamente escolhidos arbitrariamente dentro do seu intervalo de definição.

Então realiza-se um novo ajuste numérico a partir da curva acumulada truncada até a data t_3 (estimada pela rotina que filtra as ondas epidêmicas) utilizando os valores obtidos para $\{r_i, \alpha_i, q_i, K_i, p_i\}$ com $i = 1, 2$ e t_1 e p_1 como chutes iniciais para tais parâmetros no novo ajuste, enquanto para $\{r_3, \alpha_3, q_3, K_3, p_3\}$ estes são escolhidos arbitrariamente dentro do seu intervalo de definição. Este processo é repetido até que a curva epidêmica acumulada seja utilizada integralmente para o ajuste numérico. O ajuste recursivo ajuda a diminuir o problema com *overfitting* visto que os chutes iniciais para os parâmetros livres no ajuste final (considerando-se a curva epidêmica em sua totalidade) são calibrados nas etapas intermediárias, o que por sua vez diminui a sensibilidade do modelo a tais quantidades. Nota-se ainda que ao passo que o procedimento de ajustes recursivos diminui o *overfitting* ele também aumenta o tempo computacional necessário para obter resultados, sobretudo quando o número N de ondas epidêmicas presentes na curva da localidade escolhida é grande.

Além do procedimento descrito acima para o ajuste recursivo das curva epidêmicas, outras medidas foram tomadas para ajudar a minimizar o *overfitting*. Assim como discutido na seção 2.1.2.1, foi observado que a restrição dos parâmetros livres do Modelo Logístico Beta, $\{r, \alpha, q\}$, ao limite $(0, 1)$ e $p \geq 1$ diminui *overfitting* no caso de curvas com apenas uma onda, o que é extremamente útil no primeiro passo da rotina de ajuste recursivo ou no caso de curvas com apenas uma onda epidêmica. Para os parâmetros referentes às demais ondas epidêmicas, ou

seja, $\{r_i, \alpha_i, q_i, p_i\}$ com $i = 2, \dots, N$, as mesmas condições são impostas exceto para α_k , que é mantido fixo em 1 pois foi observado que deixá-lo variar gerava um aumento significativo no *overfitting* pois este parâmetro está ligado à assimetria da curva diária em torno do pico de uma dada onda, como foi discutido em detalhe em (VASCONCELOS et al., 2021). Finalmente, o parâmetro p_N para a última onda epidêmica é mantido fixo em 1, visto que tal parâmetro controla o quão rápido a curva diária decai depois do pico (ou alternativamente, o quão rápido a curva acumulada se aproxima do platô de saturação). No entanto, para ter uma boa estimativa deste parâmetro a curva precisa ter uma cauda bem formada, o que não necessariamente é o caso para curvas epidêmicas referentes a locais nos quais a pandemia ainda está evoluindo, consequentemente esta escolha ajuda a diminuir o *overfitting* com relação a este parâmetro.

3.2.5 Classificação do estágio epidêmico

Como foi comentado anteriormente, o aplicativo ModInterv após analisar a curva epidêmica pra uma determinada localidade, consegue determinar em que estágio epidêmico esta curva está. Tal determinação pode ajudar as autoridades de saúde pública a implementar ou relaxar medidas para conter o avanço da doença.

Uma vez que o ajuste numérico usando o modelo de múltiplas ondas é realizado e seus parâmetros livres são determinados, o aplicativo segue para computar os pontos $\{t_i^k\}$, $i = 1, \dots, 3$; $k = 1, \dots, N$, que determinam os diferentes regimes de aceleração da curva epidêmica. Estes pontos são computados numericamente através do procedimento que segue: a curva epidêmica acumulada é reconstruída a partir do modelo implementado no aplicativo utilizando os parâmetros determinados pelo ajuste numérico.

Esta curva é então interpolada por um *spline* cúbico e sua primeira e segunda derivadas, curva diária ou velocidade e aceleração, respectivamente, são computadas e em seguida são encontrados o ponto em que ocorre o máximo da aceleração, t_1^k , o ponto onde ocorre o zero da aceleração, t_2^k , e o ponto onde ocorre o mínimo da aceleração, t_3^k . Comparando o tempo final, t_f , do último ponto da curva epidêmica com os pontos t_i^N calculados para a última onda, o aplicativo ModInterv determina o estágio epidêmico para a localidade em questão como um dos três estágios possíveis: *aceleração crescente*, *aceleração decrescente*, *desaceleração crescente*, *desaceleração decrescente*. O estágio determinado pelo aplicativo é informado na legenda do gráfico gerado após o ajuste numérico, o qual também exibe a curva empírica, a curva obtida através do modelo matemático com os parâmetros determinados pelo ajuste

numérico e mais algumas informações relevantes que serão discutidas em detalhe nas próximas seções.

3.2.6 Interatividade do aplicativo e hospedagem

O aplicativo ModInterv foi escrito utilizando iPython *notebooks* (PÉREZ; GRANGER, 2007), uma ferramenta que pode ser utilizada na nuvem e permite o usuário desenvolver códigos sem precisar instalar as dependências em sua máquina e sem consumir seu poder de processamento. Apesar de muito versáteis, os *notebooks* não são apropriados para serem compartilhados com o público em geral por não possuírem uma interface do usuário bem definida e para que os usuários pudessem utilizar o aplicativos, todos deveriam necessariamente ter permissão de editar e compilar o código; o que poderia gerar resultados indesejados.

Tendo em vista os possíveis problemas de distribuir ao público geral o código fonte do aplicativo e ao mesmo tempo com o intuito de gerar uma página da internet facilmente acessível e com uma interface do usuário simples e interativa, foi utilizado Voilà, um módulo Python que cria páginas HTML a partir dos *notebooks* e permite que os *widgets* permaneçam interativos nesse processo, ou seja, a partir do *notebook* é gerado uma página HTML com *frontend* e *backend* funcionando separadamente. O aplicativo atualmente está hospedado em um servidor localizado na Universidade Federal do Paraná (UFPR) e sua página web pode ser acessado através do link <http://fisica.ufpr.br/modinterv>.

3.3 ESTRUTURA DE INTERFACE DO USUÁRIO DO APLICATIVO MODINTERV

Nesta seção as características da interface do usuário serão discutidas em detalhes, comentando como o usuário pode navegar pela página do aplicativo, acessar as diferentes seções designadas para cada tipo de localidade (países e estados e cidades do Brasil e dos Estados Unidos) disponível no aplicativo, customizar os gráficos e resultados gerados depois do ajuste numérico.

Para tornar a página do aplicativo mais simples de prática de usar, ela foi dividida em duas seções: Países e Estados e Cidades do Brasil e dos Estados Unidos. Em cada uma destas seções o usuário encontrará basicamente a mesma interface e com as mesmas opções de customização, como será discutido a seguir.

3.3.1 Países

A primeira seção do aplicativo é dedicada à análise dos países e para fazer ajustes numéricos com os dados de um país dentre os disponíveis, primeiramente o usuário deve escolher no menu Data o tipo de curva que deseja analisar (Casos ou Mortes) e em seguida selecionar no menu *dropdown* Country o país de interesse. Depois de selecionar o tipo de curva epidêmica e o país, um gráfico é gerado automaticamente mostrando a curva epidêmica acumulada (círculos vermelhos) em função do tempo medido em dias desde o primeiro caso/morte e na legenda é exibido o nome do país juntamente com a data do último ponto exibido no gráfico.

Em seguida, para realizar o ajuste numérico, o usuário deve clicar no botão Perform fit; uma caixa de texto irá aparecer com o texto "*Computing...*" e logo depois um novo gráfico com o ajuste numérico (curva sólida na cor preta) sobreposto à curva curva empírica (círculos vermelhos) será exibido abaixo do gráfico anterior. No gráfico, o nome do país escolhido pelo usuário é mostrado no título, juntamente com a data do último ponto considerado na análise; na legenda é indicado o modelo que melhor ajustou a curva epidêmica em análise; a data do primeiro caso ou morte; a data de início de cada onda epidêmica e o estágio epidêmico do país considerado segundo a classificação discutida anteriormente.

Além das informações já mostradas no gráfico do ajuste numérico da curva acumulada, o usuário pode ter acesso a informações adicionais através da interface do aplicativo ModInterv. Ao marcar a caixa Check to display/hide the daily curve, um novo gráfico será exibido, desta vez com a curva diária empírica (círculos vermelhos) calculada a partir da diferença entre pontos consecutivos da curva acumulada, juntamente com a curva diária teórica (linha contínua na cor preta); esta última sendo calculada através da derivada temporal da curva acumulada obtida pelo ajuste numérico do modelo escolhido. Logo abaixo da caixa mencionada anteriormente, o usuário pode encontrar uma outra caixa com o texto Check to display the parameters of the fit, que quando marcada gera uma tabela com os valores dos parâmetros encontrados pelo ajuste numérico do modelo, juntamente com o erro estimados dos parâmetros. O usuário pode ainda deslizar o cursor Time range para obter uma previsão de curto prazo, variando de 7 a 28 dias (com incrementos de 7 dias) depois do último ponto da curva empírica. Esta previsão é obtida através da curva teórica gerada pelo modelo matemático com parâmetros estimados pelo ajuste numérico; o número acumulado de mortes ou casos no período escolhido para a previsão é mostrado na legenda do gráfico.

Para permitir uma análise mais refinada das curvas acumulada e diária, o usuário pode

ainda determinar a escala de sua preferência para cada um dos eixos ao selecionar nos menus *dropdown* x-Axis e y-Axis entre as escalas linear e logarítmica. Finalmente, o usuário pode fazer o download de todos os gráficos produzidos ao clicar no botão *Generate figure* e em seguida clicar nos respectivos links gerados pelo aplicativo logo abaixo do botão. Os formatos png e eps são suportados e podem também ser escolhidos pelo usuário no momento de gerar os gráficos para download.

3.3.2 Cidades e Estados do Brasil e dos Estados Unidos

Na segunda seção do aplicativo, o usuário encontrará um layout similar ao anterior mas agora ele poderá analisar cidades e estados do Brasil e condados e estados dos Estados Unidos. Primeiramente deve ser selecionado o país de interesse, Brasil ou Estados Unidos, e em seguida o tipo de dado a ser analisado deve ser escolhido (casos ou mortes). Então o usuário deverá escolher o tipo de região a ser analisada (cidades ou estados) e em seguida o estado ou cidade (ou condado no caso dos Estados Unidos) desejada; assim como na seção anterior, um gráfico de pré visualização será mostrado apresentando a curva epidêmica a ser analisada e a data do último ponto desta curva.

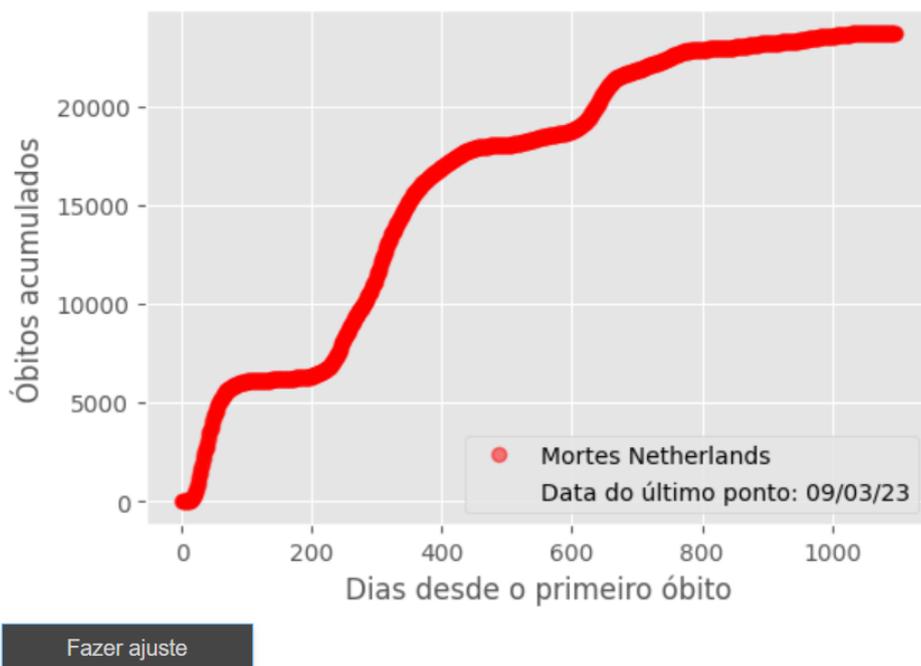
Devemos notar que para o caso das cidades do Brasil, o tempo de carregamento dos dados é relativamente longo se comparado ao caso dos condados dos Estados Unidos. Esta diferença se dá devido ao fato de existirem mais de 5500 cidades no Brasil contra pouco mais de 3000 condados nos Estados Unidos, ou seja, quase o dobro. Além disso, para cada cidade ou condado existem mais de 1300 pontos empíricos, o que no caso do Brasil culmina em um arquivo com quase 2GB de dados para as cidades; demandando um tempo computacional bastante elevado para a leitura de tal arquivo. Por este motivo, foi implementado no aplicativo *ModInterv* uma caixa de texto exibindo "Loading..." a fim de informar o usuário que os dados para as cidades do Brasil ou condados dos Estados Unidos ainda estão sendo lidos e processados pelo aplicativo. As demais funções discutidas e presentes na seção dos países, como a possibilidade de exibir a curva diária, os parâmetros do ajuste, mudança da escala dos eixos e download das figuras também estão presentes na seção de cidades e estados do Brasil e dos Estados Unidos de forma idêntica.

3.3.3 Exemplos

Nesta seção serão apresentados alguns exemplos de ajustes numéricos realizados para locais variados (países, estados e cidades do Brasil e dos Estados Unidos) usando a curva epidêmica em sua totalidade, e também usando a curva até uma certa data escolhida pelo usuário, com a finalidade de demonstrar todas as funcionalidades do aplicativo ModInterv, bem como sua capacidade de produzir ajustes numéricos de qualidade para curvas epidêmicas com múltiplas ondas.

Uma vez que o usuário acessa o aplicativo ModInterv, na seção pertinente aos países, ele deve escolher o tipo de curva a ser analisada (Casos ou Mortes) e em seguida o país desejado. Um gráfico de pré-visualização da curva epidêmica da localidade escolhida será exibido, como na figura 23 para o caso da Holanda. Note que todo o texto do gráfico está em português, exceto o nome do país escolhido; isto pode ser configurado através de dois botões localizados no cabeçalho do site para que o usuário possa optar entre português e inglês. Por padrão, o aplicativo exibe e considera todos os pontos disponíveis até aquele momento e o usuário pode, caso deseje, alterar a quantidade de pontos (e conseqüentemente a data final referente à curva epidêmica) que devem ser analisados ao deslizar o cursor ao lado do texto Time fit.

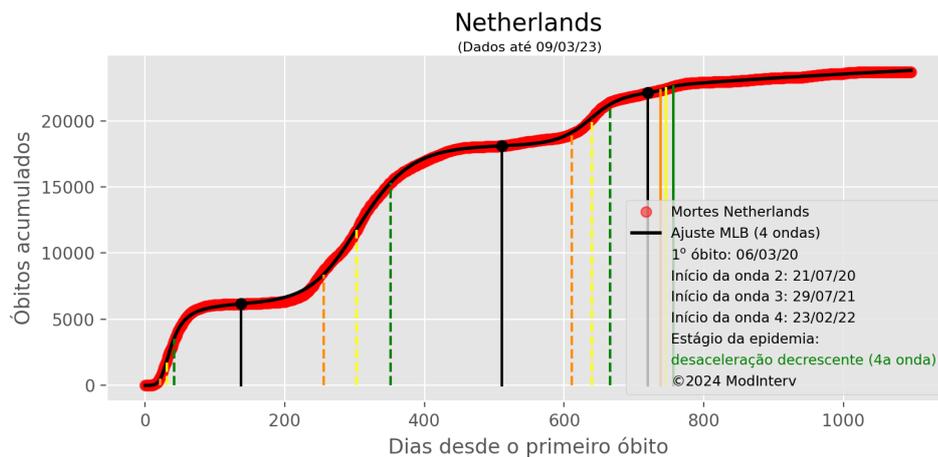
Figura 23 – Curva acumulada de mortes (círculos vermelhos) atribuídas à Covid-19 para a Holanda até o dia 09/03/2023.



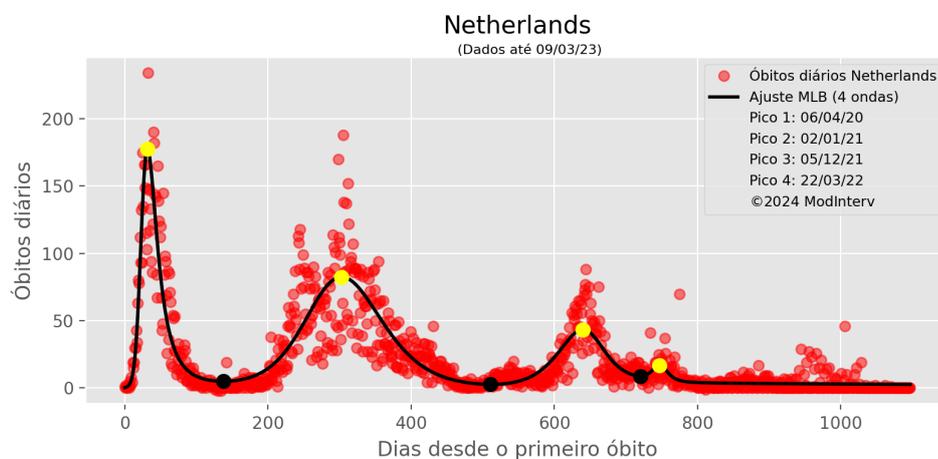
Fonte: elaborada pelo autor.

No caso em questão, será exibido o ajuste numérico para a Holanda utilizando todos os dados disponíveis na base de dados. Isto pode ser realizado pelo usuário ao clicar no botão Perform fit ou Fazer ajuste e o resultado do ajuste numérico para esta curva epidêmica é mostrado na figura 24. Na figura 24a é apresentada a curva epidêmica de mortes (círculos vermelhos) com a curva acumulada teórica (linha contínua na cor preta) sobreposta; as linhas verticais coloridas indicam o estágio epidêmico de cada onda; as linhas verticais na cor preta indicam o início de uma nova onda epidêmica. Na legenda do gráfico são disponibilizadas informações como a data do primeiro óbito, as datas de início de cada onda subsequente e o estágio da epidemia para aquele determinado país. Na figura 24b é apresentada a curva

Figura 24 – Ajuste numérico para a Holanda utilizando todos os dados disponíveis na base de dados, o que equivale à data 09/03/2023. (a) Curva epidêmica acumulada de óbitos (círculos vermelhos) sobreposta com a curva acumulada teórica (linha contínua na cor preta). (b) Curva epidêmica da óbitos diários (círculos vermelhos) sobreposta com a curva diária teórica (linha contínua na cor preta).



(a)



(b)

Fonte: elaborada pelo autor.

epidêmica de óbitos diários (círculos vermelhos) com a curva diária (linha contínua na cor preta) sobreposta. Neste gráfico são destacados em círculos na cor amarela os picos de cada onda epidêmica e em círculos na cor preta o início de uma nova onda epidêmica. Na legenda do gráfico são mostradas as datas dos picos correspondentes aos círculos em amarelo.

Além do ajuste numérico mostrado anteriormente, usando todos os pontos disponíveis, o usuário poderia ter optado por fazer o ajuste numérico até uma data anterior ao deslocar o cursor ao lado do texto *Time fit*, como mostrado na figura 25 para a Holanda até o dia 19/06/2021. Na figura 25a é mostrada a curva de óbitos acumulados (círculos vermelhos) sobreposta pela curva teórica produzida pelo modelo escolhido pelo algoritmo com os parâmetros determinados pelo ajuste numérico (curva contínua na cor preta). As mesmas informações discutidas anteriormente são mostradas na legenda do gráfico e as linhas verticais coloridas novamente representam as datas limites de cada estágio epidêmico da curva.

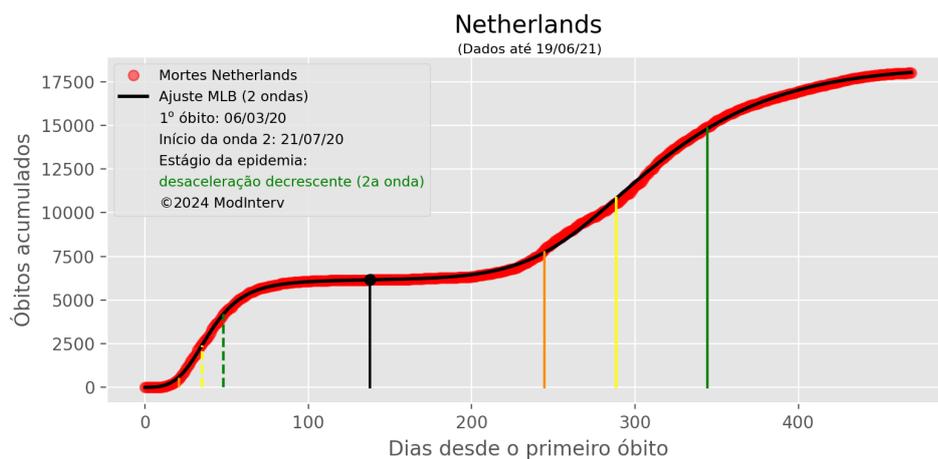
Na figura 25b é mostrada a curva empírica de mortes diárias (círculos vermelhos) sobreposta pela curva teórica diária (curva contínua na cor preta) e neste caso também são marcados com círculos amarelos e pretos os picos e o início de cada onda epidêmica, respectivamente. Neste gráfico também são exibidas as datas de cada pico correspondente aos círculos amarelos.

Na seção dos estados e cidades do Brasil e dos Estados Unidos, o usuário pode escolher entre os dois países mencionados, em seguida escolher o tipo de curva epidêmica que deseja analisar (casos ou óbitos) para então escolher o tipo de localidade (estados ou cidades). Caso selecione cidades, o usuário deve primeiramente escolher o estado ao qual a cidade de interesse pertence e em seguida, no campo *City*, selecionar a cidade. Considerando o caso de um usuário que escolheu os Estados Unidos como país, óbitos como tipo de curva epidêmica, Estados como tipo de localidade e *Wyoming* como estado para ser analisado, a figura 26 mostra o resultado do ajuste numérico, com a figura 26a exibindo a curva acumulada de óbitos (círculos vermelhos) sobreposta pela curva teórica resultante do ajuste numérico (curva contínua na cor preta) e a figura 26b apresentando a curva empírica de óbitos diários (círculos vermelhos) sobreposta pela curva diária teórica (curva contínua na cor preta). Em ambas as figuras, as mesmas informações fornecidas para a seção de países, como data do primeiro óbito, data de início de cada onda, data do pico de cada onda epidêmica e o estágio epidêmico da localidade.

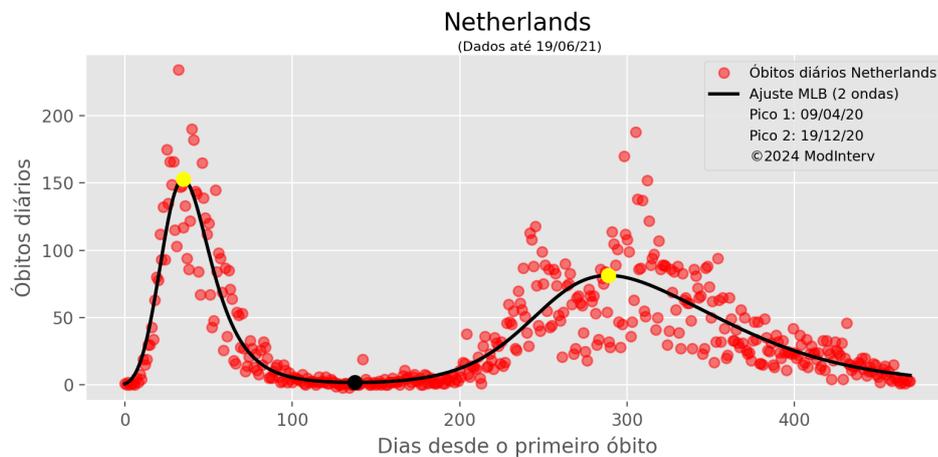
Para finalizar esta seção, será apresentado o ajuste numérico para o estado de Pernambuco, no Brasil, utilizando todos os dados disponíveis na base de dados, o que corresponde à data de 18/03/2023. A figura 27a mostra a curva de óbitos acumulados (círculos vermelhos) sobreposta

pela curva teórica obtida através do ajuste numérico com o modelo escolhido pelo algoritmo (curva contínua na cor preta). Na figura 27b é exibida a curva de óbitos diários (círculos vermelhos) sobreposta pela curva diária teórica (curva contínua na cor preta), obtida pela derivada temporal da curva teórica acumulada. Novamente, todas as informações relevantes disponibilizadas nos outros exemplos, como data do primeiro óbito, data de início das ondas epidêmicas, datas dos picos de cada onda e estágio epidêmico da localidade, estão presentes nos gráficos apresentados.

Figura 25 – Ajuste numérico para a Holanda utilizando parte dos dados disponíveis na base de dados, com o último ponto da curva correspondendo à data 19/06/2021. (a) Curva epidêmica acumulada de óbitos (círculos vermelhos) sobreposta com a curva acumulada teórica (linha contínua na cor preta). (b) Curva epidêmica da óbitos diários (círculos vermelhos) sobreposta com a curva diária teórica (linha contínua na cor preta).



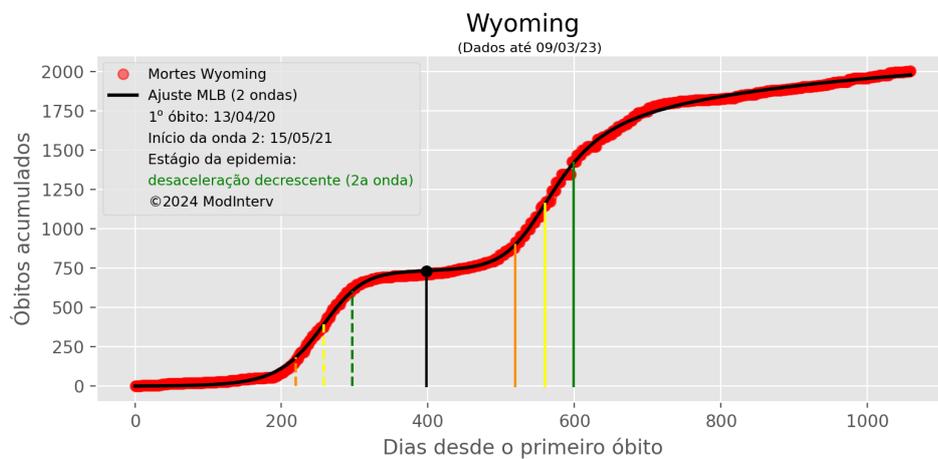
(a)



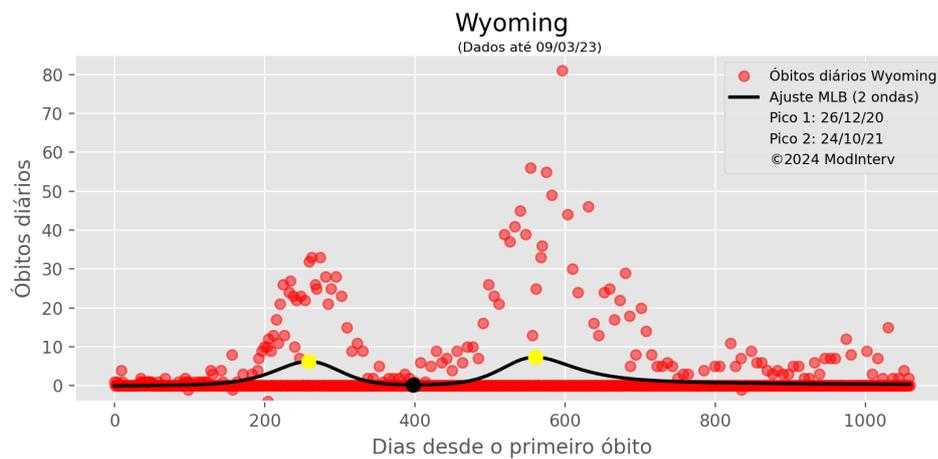
(b)

Fonte: Elaborada pelo autor.

Figura 26 – Ajuste numérico para o estado de *Wyoming* nos Estados Unidos utilizando todos os dados disponíveis na base de dados, com o último ponto da curva correspondendo à data 09/03/2023. (a) Curva epidêmica acumulada de óbitos (círculos vermelhos) sobreposta com a curva acumulada teórica (linha contínua na cor preta). (b) Curva epidêmica da óbitos diários (círculos vermelhos) sobreposta com a curva diária teórica (linha contínua na cor preta).



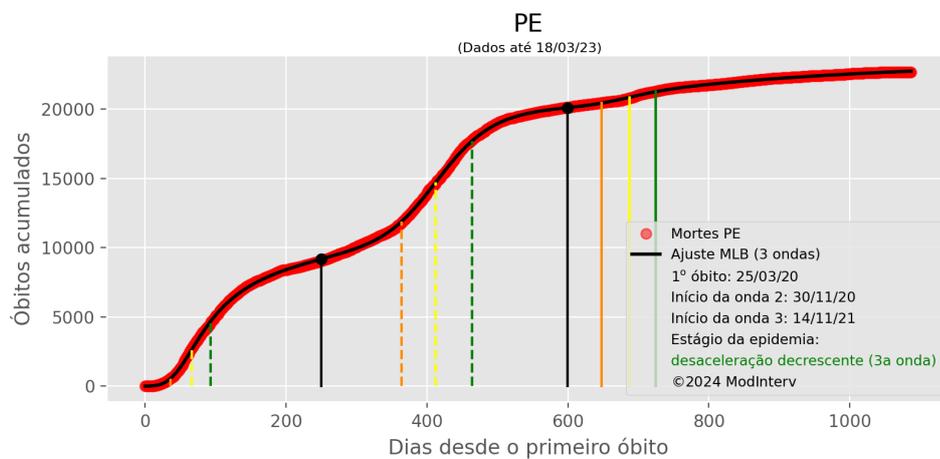
(a)



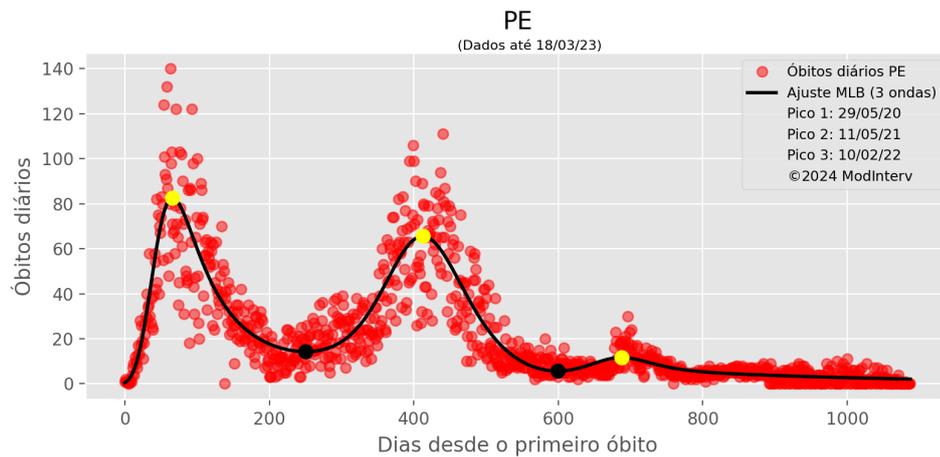
(b)

Fonte: elaborada pelo autor.

Figura 27 – Ajuste numérico para o estado de Pernambuco no Brasil utilizando todos os dados disponíveis na base de dados, com o último ponto da curva correspondendo à data 18/03/2023. (a) Curva epidêmica acumulada de óbitos (círculos vermelhos) sobreposta com a curva acumulada teórica (linha contínua na cor preta). (b) Curva epidêmica da óbitos diários (círculos vermelhos) sobreposta com a curva diária teórica (linha contínua na cor preta).



(a)



(b)

Fonte: elaborada pelo autor.

4 VERSÃO ENTRÓPICA HETEROTÍPICA DA TEORIA H

A entropia se faz presente no cotidiano de toda a humanidade, desde a sensação de frio ou calor, cansaço ou relaxamento, fome ou saciedade; todas estas manifestações são facetas da entropia ou de seu fluxo no corpo humano. O seu conceito só foi formalmente definido por volta de 1865 no trabalho seminal de Clausius (CLAUSIUS, 1865), no entanto, Carnot em seu trabalho "*Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance*", publicado em 1824 (CARNOT, 1824), já discutia tal conceito sem a nomenclatura *entropia*.

Encontrando aplicações nas mais diversas áreas, como Cosmologia (PRIGOGINE; GÉHÉNIU, 1986), Economia (GEORGESCU-ROEGER, 1986) e até Teoria dos Jogos (GRÜNWALD; DAWID, 2004), a entropia passou por diversas formulações para se encaixar a cada uma delas, como foi o caso das amplamente conhecidas entropia de Boltzmann-Gibbs, Shannon, Renyi e Tsallis. No entanto, nem todas elas carregam consigo uma interpretação física; as entropias formuladas com base na Teoria da Informação, também conhecidas como *Infoentropias* (LUZZI; VASCONCELLOS; RAMOS, 2003), são na verdade funcionais geradores de distribuições de probabilidade, enquanto entropias como a de Boltzmann-Gibbs são funcionais geradores de funções de estado termodinâmico e têm uma interpretação física bastante palpável.

Neste capítulo será calculada em detalhe a entropia relacionada à Teoria-H (MACÊDO et al., 2017a), usada para descrever sistemas com múltiplas escalas, como tempo ou comprimento, nas quais a energia flui entre elas. Dentre as aplicações naturais da Teoria-H estão a turbulência, como de fato foi mostrado em (GONZÁLEZ et al., 2017) para descrever a presença de hierarquia de turbulência em um laser aleatório de fibra, e também a coexistência de comportamento do tipo turbulento e do tipo vidro em um sistema de fótons (GONZÁLEZ et al., 2018).

Apesar de aparentemente desconexos, os dois temas principais apresentados nesta tese mantêm uma conexão significativa. As pandemias são fenômenos inerentemente complexos, estocásticos e dotados de múltiplas escalas. A dinâmica da pandemia quando olhada a nível de um país como um todo difere da dinâmica da mesma pandemia na escala estadual, municipal, distrital, até chegarmos à menor "escala", as residências.

No entanto, foge do objetivo do presente trabalho analisar a pandemia de Covid-19 do ponto de vista estocástico e considerando-se múltiplas escalas e suas respectivas dinâmicas.

Os modelos apresentados nos capítulos anteriores abordam a referida pandemia do ponto de vista determinístico, desprezando as flutuações e as múltiplas escalas de interação.

4.1 UMA BREVE INTRODUÇÃO À TEORIA-H

A Teoria-H é resultado de uma abordagem unificada para descrever a dinâmica de sistemas complexos hierárquicos. Nesta formulação é assumido que o sistema é dividido em um número N arbitrário de escalas, cujos tempos ou comprimentos característicos são muito diferentes entre escalas consecutivas e também que há equilíbrio local em cada uma delas.

Partimos de um sistema cuja distribuição de probabilidade da quantidade de interesse x depende da escala em que é medida. Realizando medidas em uma dada escala de tempo τ_0 a partir da qual as flutuações observadas não são mais correlacionadas, é assumido que o sistema é caracterizado por uma distribuição Gaussiana conhecida $P(x|\varepsilon_0)$, em que ε_0 é um parâmetro que caracteriza o equilíbrio global do sistema. Desta forma, assumindo ainda que a quantidade x varia muito mais rapidamente que o seu ambiente local, o sistema apresenta a mesma distribuição $P(x|\varepsilon)$, onde agora ε , denominada como variável de "background" representa o equilíbrio local, cuja distribuição de probabilidade é dada por:

$$P(x|\varepsilon) = \frac{1}{\sqrt{2\pi\varepsilon}} \exp\left(-\frac{x^2}{2\varepsilon}\right). \quad (4.1)$$

Fazendo amostragens do sistema em curtos intervalos de tempo durante um período prolongado, a distribuição marginal da quantidade de interesse, também denominada como "sinal" assume a forma:

$$P(x) = \int_0^\infty P(x|\varepsilon)f(\varepsilon)d\varepsilon, \quad (4.2)$$

onde $f(\varepsilon)$ é a distribuição de probabilidade da variável de *background*.

É possível determinar a forma de $f(\varepsilon)$ através de argumentos físicos ao se construir uma Equação Diferencial Estocástica com certas condições limitadoras. Assumindo que o sistema é composto por um número N de escalas de tempo bem separadas τ_i , $i = 1, \dots, N$ em que para cada escala subsequente tem-se $\tau_i \ll \tau_{i-1}$. Portanto, para uma dada escala i , a variável ε_i é dada pela EDE:

$$d\varepsilon_i = F(\varepsilon_i, \varepsilon_{i-1}) dt + G(\varepsilon_i, \varepsilon_{i-1}) dW_i, \quad (4.3)$$

em que $F(\varepsilon_i, \varepsilon_{i-1})$ é o termo determinístico de deriva, $G(\varepsilon_i, \varepsilon_{i-1})$ é a amplitude do ruído e dW_i são processos de Wiener independentes.

Os funcionais F e G devem obedecer a três condições físicas: condição de equilíbrio ($\langle \varepsilon_i(t) \rangle = \varepsilon_0$ quando $t \rightarrow \infty$), invariância por mudança de escala e positividade de ε_i . A EDE mais geral que satisfaz estas condições é:

$$d\varepsilon_i = -\gamma_i (\varepsilon_i - \varepsilon_{i-1}) dt + \kappa_i \varepsilon_i^s \varepsilon_{i-1}^{1-s} dW_i, \quad i = 1, \dots, N, \quad (4.4)$$

em que γ_i e κ_i são constantes reais e $0 \leq s \leq 1$. No entanto, se for imposta a condição de que a equação de Fokker-Planck correspondente deve possuir coeficientes analíticos, os valores de s se restringem a $s = 1/2$ e $s = 1$. O caso $s = 1/2$ dá origem à classe de universalidade Exponencial Esticada, enquanto o caso $s = 1$ dá origem à classe de universalidade Lei de Potência, que serão abordadas ao longo do capítulo.

4.2 FUNCIONAL ENTRÓPICO A PARTIR DE UMA DISTRIBUIÇÃO DE PROBABILIDADE

Nesta seção será apresentado o procedimento utilizado para se obter um funcional entrópico a partir de uma distribuição de probabilidade dada, como demonstrado em detalhes em (TSALLIS; SOUZA, 2003). Partindo de um funcional

$$\Phi[p] \equiv S[p] - \alpha \sum_i p_i - \beta \frac{\sum_i u(p_i) E_i}{\sum_i u(p_i)}, \quad (4.5)$$

onde $S[p] = \sum_i s(p_i)$, α e β são parâmetros de Lagrange, $\{E_i\}$ é o conjunto de autovalores do Hamiltoniano do sistema, $p_i = p(E_i)$ são probabilidades e a função $u(x)$ é em geral uma função monotonicamente crescente e definida tal que $u(0) = 0$ e $u(1) = 1$. É possível notar que a quantidade multiplicando β é a energia interna do sistema e portanto, para o caso particular do ensemble de Boltzmann-Gibbs, deve-se ter $u(x) = x$, enquanto para a estatística de Tsallis, $u(x) = x^q$. Ao extremizar a equação 4.5 sob a condição de energia não limitada, obtém-se $u(p) = p$ e ainda

$$s(p) = \alpha p + \int_0^p dp' E(p'), \quad (4.6)$$

onde $s(p)$ é a entropia e a função $E(p')$ é a inversa de $p(E)$. A constante α pode ser determinada a partir da condição $s(1) = 0$, o que leva a

$$\alpha = - \int_0^1 dp' E(p'). \quad (4.7)$$

Como mencionado acima, $E(p')$ é a função inversa de $p(E)$ mas no geral não é possível avaliar esta integral analiticamente conhecendo apenas $p(E)$, excetuando-se os casos em que

tal função é facilmente invertível. Para avaliar esta integral em um caso genérico, é necessário utilizar a propriedade da integral do produto, cujo resultado é dado por

$$\int_0^1 dp' E(p') = pE(p) \Big|_0^1 - \int_{E(p=0)}^{E(p=1)} dE' p(E'). \quad (4.8)$$

Em geral, $E(p = 1) = 0$ e $E(p = 0) \rightarrow \infty$, portanto,

$$\alpha = - \int_0^\infty dE' p(E'). \quad (4.9)$$

Nos sistemas complexos hierárquicos, existem duas classes de universalidade, cada uma com sua distribuição de probabilidade, que surgem naturalmente a partir do formalismo desenvolvido para descrevê-los, são elas (MACÊDO et al., 2017a): a classe Exponencial Esticada – uma generalização da distribuição K – e a classe de Lei de Potência. Nesta seção serão discutidas ambas as classes de universalidade no âmbito do cálculo da entropia heterotípica.

4.2.1 Classe de Universalidade Exponencial Esticada

A distribuição de probabilidade da classe Exponencial Esticada é dada por

$$\tilde{p}(E) = P_N(\sqrt{E}) = \frac{\omega^{1/2}}{\sqrt{2\pi\varepsilon_0}\Gamma(\vec{\beta})} G_{0,N+1}^{N+1,0} \left(\begin{matrix} - \\ \vec{\beta} - 1/2, 0 \end{matrix} \middle| \frac{\omega E}{2\varepsilon_0} \right), \quad (4.10)$$

onde $P_N(E)$ é a distribuição de probabilidade marginal associada à classe de universalidade exponencial esticada, como discutido em detalhes em (MACÊDO et al., 2017a), $G_{p,q}^{m,n} \left(\begin{matrix} \vec{a}_p \\ \vec{b}_q \end{matrix} \middle| x \right)$ é a função G de Meijer (BATEMAN, 1953), $\Gamma(\vec{a}) \equiv \prod_{j=1}^N \Gamma(a_j)$, $\Gamma(x)$ representa a função Gama e β é o vetor $\vec{\beta} \equiv (\beta_1, \dots, \beta_N)$, com β_i sendo constantes positivas. Em geral, como mostrado em (MACÊDO et al., 2017a), tem-se $\beta_i = \beta$ e conseqüentemente, $\omega = \beta^N$. A quantidade ε , que aparece no desenvolvimento das distribuições de probabilidade associadas a cada classe de universalidade é uma variável que flutua mas que varia de forma mais devagar no tempo ou no espaço que a quantidade E . A interpretação física desta variável seria por exemplo o fluxo local de energia para o caso de um escoamento turbulento (SALAZAR; VASCONCELOS, 2010) ou a temperatura local em um sistema de múltiplas escalas em equilíbrio térmico (SALAZAR; VASCONCELOS, 2012). Neste sentido, a quantidade ε_0 é o seu equivalente na escala integral, ou seja, o fluxo de energia ou a temperatura da escala integral. A versão normalizada de $\tilde{p}(E)$

é dada por

$$p(E) = \frac{\omega}{2\varepsilon_0 \Gamma(\vec{a} + \vec{1})} G_{0, N+1}^{N+1, 0} \left(\begin{array}{c} - \\ \vec{a}, 0 \end{array} \middle| \begin{array}{c} \omega E \\ 2\varepsilon_0 \end{array} \right), \quad (4.11)$$

onde agora $a = \beta - 1/2$ e $\omega = (a + 1/2)^N$.

Esta classe de distribuição é uma generalização da distribuição K, que é recuperada pela equação (4.11) no caso $N = 1$, e tem aplicações em espalhamento em meios aleatórios (JAKEMAN; PUSEY, 1978) e também em turbulência (ANDREWS et al., 1989). Em seguida será realizado o procedimento adotado para obter o funcional entrópico a partir da distribuição de probabilidade para os dois casos mais simples, $N = 0$ e $N = 1$, e depois para o caso geral de um número de escalas N qualquer. É importante notar que esta classe de universalidade leva o referido nome pois no limite assintótico de argumentos grandes, a distribuição exponencial esticada é recuperada.

4.2.1.1 O Ensemble de Boltzmann-Gibbs ($N=0$)

Fazendo $N = 0$ na distribuição de probabilidade da classe de universalidade exponencial esticada, obtém-se $p(E) = \frac{1}{2\varepsilon_0} e^{-\frac{E}{2\varepsilon_0}}$ (notando-se que pela definição de ω , quando $N = 0$ então $\omega = 1$ para qualquer valor de β ou a e, neste caso, torna-se fácil encontrar a função inversa $E(p) = -2\varepsilon_0 \ln 2\varepsilon_0 p$. Da condição apresentada na equação (4.7), é possível calcular o valor de α para o caso $N = 0$.

$$\alpha = - \int_0^1 -2\varepsilon_0 \ln 2\varepsilon_0 p' dp' = 2\varepsilon_0 \ln 2\varepsilon_0 - 2\varepsilon_0, \quad (4.12)$$

portanto, da equação (4.6) obtém-se

$$s(p) = (2\varepsilon_0 \ln 2\varepsilon_0 - 2\varepsilon_0)p - 2\varepsilon_0 p \ln 2\varepsilon_0 p + 2\varepsilon_0 p \quad (4.13)$$

$$= -2\varepsilon_0 p \ln p, \quad (4.14)$$

que tem exatamente a mesma forma funcional da entropia de Shannon, e esta é recuperada quando $2\varepsilon_0 = 1$. Utilizando a definição $s(E) \equiv s(p(E))$ e substituindo a expressão para $p(E)$ em $s(p)$, tem-se

$$s(E) = \frac{E}{2\varepsilon_0} e^{-\frac{E}{2\varepsilon_0}} - \ln 2\varepsilon_0. \quad (4.15)$$

Note que, como mostrado acima, o caso $N = 0$ recupera a tão conhecida entropia de Shannon e este cenário equivale a Teoria-H levando em conta todas as escalas de comprimento

ou tempo de uma só vez. Considere por exemplo o caso de um jato turbulento como discutido em detalhes em (MACÊDO et al., 2017b); ao se medir a distribuição de probabilidade dos incrementos de velocidade na escala integral, esta vai ter um formato Gaussiano, ou seja, $p(x^2)$ com $N = 0$ (onde x é o incremento de velocidade). Por outro lado, ao se medir a distribuição de probabilidade dos incrementos de velocidade nas menores escalas (intervalo inercial), o comportamento hierárquico se mostra bastante presente e distribuição de probabilidade se afasta de um formato Gaussiano para assumir o formato de exponencial esticada, ou seja, $p(x^2)$ com $N > 0$.

4.2.1.2 O Ensemble K ($N=1$)

Partindo da equação (4.2.1) com $N = 1$, recupera-se a distribuição K:

$$p(E) = \frac{\omega}{2\varepsilon_0\Gamma(a+1)} G_{0,2}^{2,0} \left(- \left| \frac{\omega E}{2\varepsilon_0} \right. \right)_{a,0} = \frac{2\omega(\omega E/2\varepsilon_0)^{a/2}}{2\varepsilon_0\Gamma(a+1)} K_a \left(2\sqrt{\frac{\omega E}{2\varepsilon_0}} \right). \quad (4.16)$$

O procedimento que será seguido neste caso para obter o funcional entrópico a partir da distribuição de probabilidade é o mesmo adotado no caso anterior, para $N = 0$. No entanto, como para $N = 1$ a distribuição não é invertível, serão necessários vários passos intermediários para chegar no resultado desejado. Este procedimento é iniciado pelo cálculo de α fazendo uso da equação (4.9) com $p(E)$ assumindo a forma dada na equação (4.16). Para se computar esta integral é necessário utilizar a definição da função G de Meijer, que é construída a partir de uma transformada inversa de Mellin. Aplicando uma transformação de Mellin à definição da função G de Meijer fornecida em (BATEMAN, 1953), obtém-se

$$\int_0^\infty x^{s-1} G_{p,q}^{m,n} \left(\begin{matrix} \vec{a}_p \\ \vec{b}_q \end{matrix} \middle| \eta x \right) dx = \frac{\eta^{-s} \prod_{j=1}^m \Gamma(b_j + s) \prod_{j=1}^n \Gamma(1 - a_j - s)}{\prod_{j=m+1}^q \Gamma(1 - b_j - s) \prod_{j=n+1}^p \Gamma(a_j + s)}. \quad (4.17)$$

Utilizando este resultado na equação (4.9) com $p(E)$ dado pela equação (4.16) tem-se

$$\alpha = - \int_0^\infty dE \frac{\omega}{2\varepsilon_0\Gamma(a+1)} G_{0,2}^{2,0} \left(- \left| \frac{\omega E}{2\varepsilon_0} \right. \right)_{a,0}. \quad (4.18)$$

Comparando a integral acima com a equação (4.17) é simples notar que $s = 1$ e $\eta = \omega/2\varepsilon_0$.

Portanto,

$$\alpha = - \frac{\omega}{2\varepsilon_0\Gamma(a+1)} \frac{2\varepsilon_0}{\omega} \Gamma(a+1)\Gamma(1) = -1. \quad (4.19)$$

Definindo $s(E) \equiv s(p(E))$ e derivando com respeito a E em ambos os lados a equação (4.6), obtém-se

$$s'(E) = (E + \alpha)p'(E) = (E - 1)p'(E). \quad (4.20)$$

Integrando ambos os lados da equação acima é possível encontrar a expressão para $s(E)$:

$$s(E) = \int dE' (E' - a)p'(E'), \quad (4.21)$$

que fazendo uso da regra do produto para a integração, fica

$$s(E) = (E - 1)p(E) - \int p(E)dE. \quad (4.22)$$

Utilizando a propriedade de integração da função G de Meijer (BATEMAN, 1953), é possível chegar na expressão final para o funcional entrópico $s(E)$

$$s(E) = (E - 1)p(E) + \frac{\omega}{2\pi\varepsilon_0\Gamma(a + 1)} G_{0,2}^{2,0} \left(\begin{matrix} - \\ a + 1, 0 \end{matrix} \middle| \frac{\omega E}{2\varepsilon_0} \right). \quad (4.23)$$

4.2.1.3 Caso geral ($N > 1$)

Nesta seção será calculado o funcional entrópico $s(E)$ a partir da distribuição de probabilidade $p(E)$ para o caso geral com um número N de escalas hierárquicas. A sequência lógica seguida será exatamente a mesma adotada nos casos anteriores e tem início ao se calcular o valor de α através da definição (4.9) utilizando para $p(E)$ a expressão geral (4.11). Desta forma,

$$\alpha = - \int_0^\infty dE \frac{\omega}{2\varepsilon_0\Gamma(\vec{a} + \vec{1})} G_{0,N+1}^{N+1,0} \left(\begin{matrix} - \\ \vec{a}, 0 \end{matrix} \middle| \frac{\omega E}{2\varepsilon_0} \right). \quad (4.24)$$

Usando mais uma vez a transformação de Mellin para avaliar esta integral, obtém-se

$$\alpha = - \frac{1}{\Gamma(\vec{a} + \vec{1})} \underbrace{\Gamma(a + 1) \dots \Gamma(a + 1)}_{N \text{ vezes}} \Gamma(1) = -1, \quad (4.25)$$

onde na última igualdade utilizamos a definição $\Gamma(\vec{a} + \vec{1}) = \prod_{j=1}^N \Gamma(a + 1) = \Gamma(a + 1)^N$.

Novamente, definindo $s(E) \equiv s(p(E))$ e derivando ambos os lados em relação a E juntamente com a equação (4.6),

$$s'(E) = -p'(E) + Ep'(E) = (E - 1)p'(E). \quad (4.26)$$

Integrando ambos os lados da equação acima e utilizando a propriedade da integral do produto, obtém-se

$$s(E) = (E - 1)p(E) - \int p(E)dE. \quad (4.27)$$

Utilizando as propriedades da função G de Meijer é possível avaliar a integral na equação acima:

$$\int \frac{\omega}{2\varepsilon_0\Gamma(\vec{a} + \vec{1})} G_{0,N+1}^{N+1,0} \left(\begin{matrix} - \\ \vec{a}, 0 \end{matrix} \middle| \frac{\omega E}{2\varepsilon_0} \right) dE = -\frac{\omega E}{2\varepsilon_0\Gamma(\vec{a} + \vec{1})} G_{0,N+1}^{N+1,0} \left(\begin{matrix} - \\ \vec{a}, -1 \end{matrix} \middle| \frac{\omega E}{2\varepsilon_0} \right), \quad (4.28)$$

e finalmente chega-se na expressão final para o funcional entrópico obtido a partir da distribuição de probabilidade $p(E)$ dada,

$$s(E) = (E - 1)p(E) + \frac{\omega}{2\varepsilon_0\Gamma(\vec{a} + \vec{1})} G_{0,N+1}^{N+1,0} \left(\begin{matrix} - \\ \vec{a} + \vec{1}, 0 \end{matrix} \middle| \frac{\omega E}{2\varepsilon_0} \right). \quad (4.29)$$

A figura 28 mostra um plot paramétrico de $s(p)$ para N variando de 1 a 5 para a classe de universalidade Exponencial Esticada. Para estas curvas foram usados os valores $a = 1$ e valores de ω e ε_0 tais que o quociente $\omega/2\varepsilon_0 = 1$.

4.2.2 Classe de Universalidade Lei de Potência

Nesta seção será aplicado o mesmo procedimento adotado anteriormente para a obtenção do funcional entrópico $s(E)$ a partir da distribuição de probabilidade $p(E)$, que neste caso é dada por

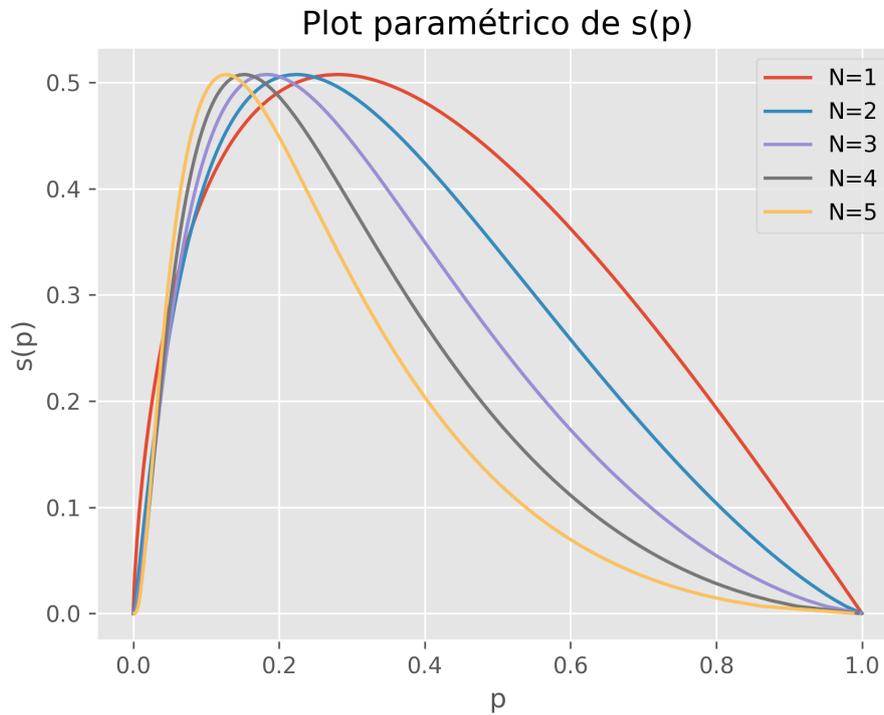
$$\tilde{p}(E) = P_N(\sqrt{E}) = \frac{1}{\sqrt{2\pi\omega\varepsilon_0}\Gamma(\vec{\beta} + \vec{1})} G_{N,1}^{1,N} \left(\begin{matrix} -\vec{\beta} - \vec{1}/2 \\ 0 \end{matrix} \middle| \frac{E}{2\omega\varepsilon_0} \right), \quad (4.30)$$

onde as definições expostas para a classe de universalidade Exponencial Esticada continuam válidas e cuja versão normalizada é dada por

$$p(E) = \frac{1}{2\omega\varepsilon_0\Gamma(\vec{a})} G_{N,1}^{1,N} \left(\begin{matrix} -\vec{a} \\ 0 \end{matrix} \middle| \frac{E}{2\omega\varepsilon_0} \right), \quad (4.31)$$

onde foi utilizada a substituição $a = \beta + 1/2$. Note que neste caso também se tomou $N = 0$, é recuperado o ensemble de Boltzmann-Gibbs com uma pequena modificação no termo que multiplica E ; portanto, não se faz necessária a abordagem deste caso particular para a classe

Figura 28 – Plot paramétrico de $s(p)$ para N variando de 1 a 5 e valores de parâmetros $a = 1$ e ω e ε_0 foram escolhidos para cada caso de forma que o quociente $\omega/2\varepsilon_0 = 1$.



Fonte: elaborada pelo autor.

de Lei de Potência. Novamente, esta classe de universalidade ganhou o referido nome pois uma distribuição na forma de lei de potência é recuperada no limite assintótico de argumentos grandes para $\tilde{p}(E)$.

4.2.2.1 Ensemble Gaussiano Extendido ($N=1$)

Nesta seção será obtido o funcional entrópico $s(E)$ a partir da distribuição de probabilidade associada à classe de universalidade de Lei de Potência para o caso específico $N = 1$. Da equação (4.31) com $N = 1$, tem-se

$$p(E) = \frac{1}{2\omega\varepsilon_0\Gamma(a)} G_{1,1}^{1,1} \left(\begin{matrix} -a \\ 0 \end{matrix} \middle| \frac{E}{2\omega\varepsilon_0} \right). \quad (4.32)$$

Nesta forma assumida pela distribuição de probabilidade, a função G de Meijer assume uma forma funcional bastante simples, dada por

$$G_{1,1}^{1,1} \left(\begin{matrix} a \\ b \end{matrix} \middle| x \right) = \Gamma(1 - a + b) x^b (1 + x)^{a-b-1}. \quad (4.33)$$

Portanto, $p(E)$ pode ser escrita como

$$p(E) = \frac{1}{2\omega\varepsilon_0\Gamma(a)}\Gamma(1+a)\left(1+\frac{E}{2\omega\varepsilon_0}\right)^{-a-1} = \frac{a}{2\omega\varepsilon_0}\left(1+\frac{E}{2\omega\varepsilon_0}\right)^{-(a+1)}. \quad (4.34)$$

A equação acima pode ser escrita em termos de uma distribuição q-Gaussiana como mostrado abaixo:

$$p(E) = C_q f\left(\sqrt{\frac{E}{2\omega\varepsilon_0(1-q)}}\right), \quad (4.35)$$

onde $f(x)$ é a distribuição q-Gaussiana com parâmetro $\beta = 1$, C_q é uma constante de normalização e o parâmetro q pode ser determinado em termos do parâmetro a através da seguinte relação

$$q = \frac{a+2}{a+1}. \quad (4.36)$$

O parâmetro q deve sempre ser menor que 3, do contrário a distribuição q-Gaussiana não é normalizável; isto pode ser verificado como válido para o caso em tela, pois $q < 3$ implica em um valor no parâmetro a a condição $a > -1/2$ e o parâmetro a é sempre positivo, por definição. Por fim, note que apesar de ser possível representar a distribuição de probabilidade da classe de universalidade da Lei de Potência como uma distribuição q-Gaussiana, não é possível recuperar a entropia de Tsallis utilizando o procedimento descrito devido à condição de energia ilimitada escolhida para extremizar o funcional (4.5).

4.2.3 Caso geral ($N > 1$)

Agora será calculado o funcional entrópico $s(E)$ a partir da distribuição de probabilidade $p(E)$ associada à classe de universalidade Lei de Potência para um caso genérico com $N > 1$.

Novamente o processo se inicia com o cálculo de α

$$\alpha = - \int_0^\infty dE \frac{1}{2\omega\varepsilon_0\Gamma(\vec{a})} G_{N,1}^{1,N} \left(\begin{matrix} -\vec{a} \\ 0 \end{matrix} \middle| \frac{E}{2\omega\varepsilon_0} \right). \quad (4.37)$$

Utilizando o resultado da equação (4.17) é possível encontrar o valor de α

$$\alpha = - \frac{1}{2\omega\varepsilon_0\Gamma(\vec{a})} 2\omega\varepsilon_0\Gamma(1)\Gamma(\vec{a}) = -1. \quad (4.38)$$

Definindo $s(E) = s(p(E))$ e derivando ambos os lados da equação com relação a E e representando o lado direito da equação pela (4.9)

$$s'(E) = s'(p(E))p'(E) = (E-1)p'(E). \quad (4.39)$$

Integrando ambos os lados da equação acima, chegamos a uma expressão para o funcional entrópico

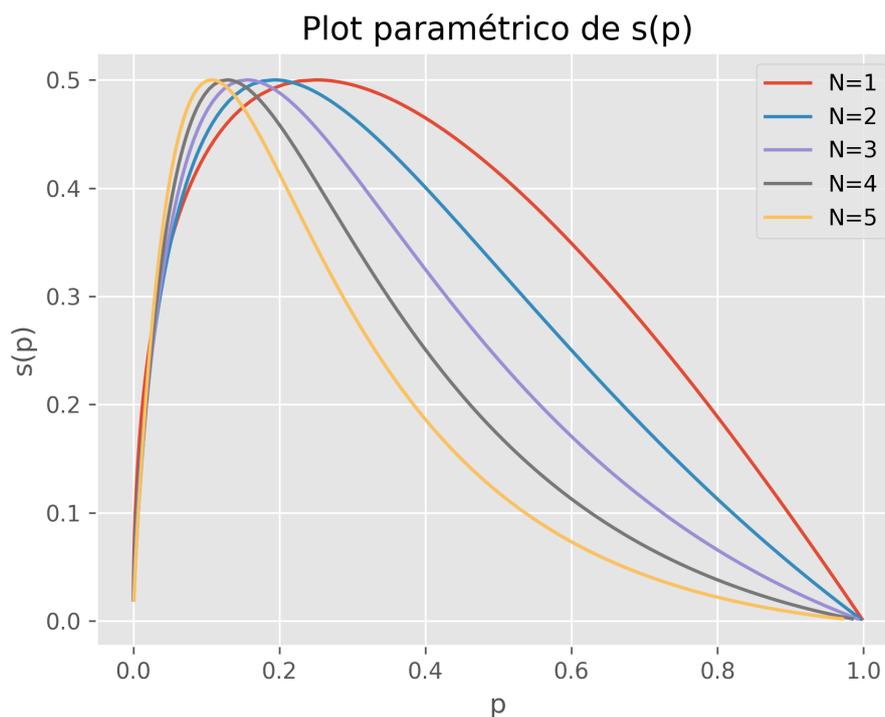
$$s(E) = (E - 1)p(E) - \int \frac{1}{2\omega\varepsilon_0\Gamma(\vec{a})} G_{N,1}^{1,N} \left(\begin{matrix} -\vec{a} \\ 0 \end{matrix} \middle| \frac{E}{2\omega\varepsilon_0} \right). \quad (4.40)$$

Utilizando as propriedades da função G de Meijer é possível computar a integral no lado direito da igualdade da equação acima para chegar na versão final do funcional entrópico

$$s(E) = (E - 1)p(E) + \frac{1}{2\omega\varepsilon_0\Gamma(\vec{a})} G_{N,1}^{1,N} \left(\begin{matrix} -\vec{a} + \vec{1} \\ 0 \end{matrix} \middle| \frac{E}{2\omega\varepsilon_0} \right). \quad (4.41)$$

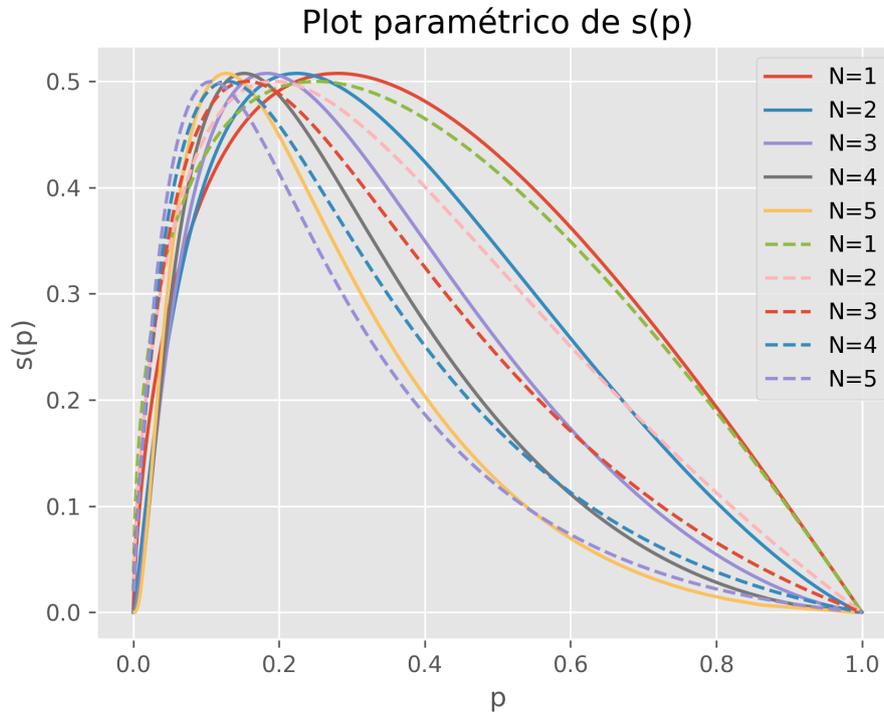
A figura 29 mostra um plot paramétrico de $s(p)$ para N variando de 1 a 5 para a classe de universalidade Lei de Potência. Para estas curvas foram usados os valores $a = 1$ e valores de ω e ε_0 tais que o produto $2\omega\varepsilon_0 = 1$. A figura 30 mostra uma comparação entre as curvas paramétricas $s(p)$ para as classes de universalidade Exponencial Esticada (linhas contínuas no gráfico) e Lei de Potência (linhas tracejadas no gráfico) com N variando de 1 a 5, $a = 1$ e parâmetros escolhidos de tal forma que o termo que acompanha a variável E nas equações (4.29) e (4.41) sejam sempre iguais a unidade para critério de comparação.

Figura 29 – Plot paramétrico de $s(p)$ para N variando de 1 a 5 e valores de parâmetros $a = 1$ e ω e ε_0 foram escolhidos para cada caso de forma que o produto $2\omega\varepsilon_0 = 1$.



Fonte: elaborada pelo autor.

Figura 30 – Comparação entre as curvas paramétricas $s(p)$ para as classes de universalidade Exponencial Esticada e Lei de Potência com N variando de 1 a 5, $a = 1$ e parâmetros ω e ε_0 escolhidos de tal forma que o termo que acompanha a variável E nas equações (4.29) e (4.41) sejam sempre iguais a unidade.



Fonte: elaborada pelo autor.

4.3 TERMODINÂMICA ESTOCÁSTICA

A teoria-H é inerentemente estocástica, desta forma, os fenômenos por ela descritos são governados pela Termodinâmica Estocástica. Nesta seção será feita a conexão entre a teoria da Termodinâmica Estocástica e as descobertas apresentadas anteriormente.

A partir de uma distribuição de probabilidade dependente do tempo $p(x, t)$ que satisfaz a seguinte equação de continuidade:

$$\partial_t p(x, t) = -\vec{\nabla} \cdot \vec{J}(x, t), \quad (4.42)$$

onde $J(x, t)$ é a densidade de corrente de probabilidade associada à distribuição de probabilidade $p(x, t)$. Esta densidade de corrente de probabilidade é dada por

$$\vec{J}(x, t) = \vec{u}(x, t)R(p(x, t)) - D\vec{\nabla}p(x, t), \quad (4.43)$$

onde $u(x, t)$ é um campo potencial arbitrário, D é o coeficiente de difusão e $R(p(x, t))$ é o "raio de curvatura" no espaço de probabilidades

$$R(p(x, t)) = -\frac{1}{s''(p)} \geq 0. \quad (4.44)$$

A entropia dinâmica é

$$S(t) = \int dxs(p(x, t)). \quad (4.45)$$

Tomando da derivada com relação ao tempo da equação acima, obtém-se

$$\dot{S}(t) = \int dxs'(p(x, t)) \frac{\partial p}{\partial t}. \quad (4.46)$$

Substituindo o lado direito da equação de continuidade (4.42) na última equação:

$$\dot{S}(t) = - \int dxs'(p(x, t)) \vec{\nabla} \cdot \vec{J}(x, t). \quad (4.47)$$

Para avaliar a integral presente no lado direito da igualdade acima, é necessário usar a identidade do cálculo vetorial

$$\vec{\nabla} \cdot (f\vec{a}) = f\vec{\nabla} \cdot \vec{a} + \vec{\nabla} f \cdot \vec{a}, \quad (4.48)$$

onde f é uma função escalar \vec{a} é um campo vetorial. Se na equação (4.48), for substituído \vec{a} por $\vec{u}(x, t)$ e f por $S(p(x, t))$, após rearranjar os termos e integrar os dois lados da equação,

$$\int dxs'(p(x, t)) \vec{\nabla} \cdot \vec{J}(x, t) = \int dx \vec{\nabla} \cdot (s(p(x, t)) \vec{J}(x, t)) - \int dx \vec{\nabla} s'(p(x, t)) \cdot \vec{J}(x, t). \quad (4.49)$$

Aplicando o teorema de Stokes no primeiro termo do lado direito da equação acima e lembrando que a densidade de corrente de probabilidade desaparece no infinito, é possível eliminar este termo. Além disso, usando a regra da cadeia para $\vec{\nabla} s'(p(x, t))$, a equação (4.47) vira

$$\dot{S}(t) = \int dxs''(p(x, t)) \vec{\nabla} p(x, t) \cdot \vec{J}(x, t) = - \int dx \frac{\vec{\nabla} p(x, t) \cdot \vec{J}(x, t)}{R(p(x, t))}. \quad (4.50)$$

Na última igualdade da equação anterior foi utilizada a definição de $R(p(x, t))$. Da equação (4.43), é possível reescrever $\vec{\nabla} p(x, t)$ como

$$\vec{\nabla} p(x, t) = - \frac{\vec{J}(x, t)}{D} + \frac{\vec{u}(x, t) R(p(x, t))}{D}. \quad (4.51)$$

Substituindo a expressão por $\vec{\nabla} p(x, t)$ na equação (4.50), tem-se

$$\dot{S}(t) = \Pi(t) - \Phi(t), \quad (4.52)$$

onde

$$\Pi(t) = \int dx \frac{J(x, t)^2}{DR(p(x, t))} \geq 0 \quad (4.53)$$

é a taxa de produção de entropia e

$$\Phi(t) = \int dx \frac{\vec{J}(x, t) \cdot \vec{u}(x, t)}{D} \quad (4.54)$$

é o fluxo de entropia. É possível ver que assim como no caso da termodinâmica, a entropia também desempenha um papel central na dinâmica dos sistemas longe do equilíbrio. Enquanto o equilíbrio termodinâmico é obtido no estado de máxima entropia, na termodinâmica estocástica, os estados estacionários são selecionados pelo princípio de máxima produção de entropia, como discutido detalhadamente em (ENDRES, 2017). Portanto, é a entropia a responsável pela dinâmica dos sistemas longe do equilíbrio.

Agora será introduzida a condição de balanço detalhado

$$\frac{\vec{u}(x, t)}{D} = -\beta \vec{\nabla} \varepsilon(x, t), \quad (4.55)$$

onde $\beta = 1/K_B T$ e $\varepsilon(x, t)$ é uma densidade de energia dependente do tempo com comportamento assintótico dado por

$$\lim_{t \rightarrow \infty} \varepsilon(x, t) = \varepsilon_{eq}(x). \quad (4.56)$$

No equilíbrio, tem-se $J(x, t) = 0$ e portanto, da equação (4.43),

$$\frac{\vec{u}(x, t)}{D} = \frac{\vec{\nabla} p(x, t)}{R(p(x, t))} = -s''(p) \vec{\nabla} p. \quad (4.57)$$

Substituindo a última igualdade da equação anterior na condição de balanço detalhado, obtém-se

$$\beta \vec{\nabla} \varepsilon_{eq}(x) = s''(p) \vec{\nabla} p \quad (4.58)$$

Da definição de $s(p)$, e da equação (4.6),

$$s''(p) = E'(p), \quad (4.59)$$

portanto,

$$\beta \vec{\nabla} \varepsilon_{eq}(x) = E'(p) \vec{\nabla} p = \vec{\nabla} E(p), \quad (4.60)$$

e

$$E(p) = \beta \varepsilon_{eq}(x). \quad (4.61)$$

Finalmente, como $p(E)$ é a inversa de $E(p)$, tem-se

$$p_{eq}(x) = \frac{1}{Z(\beta) \Gamma(\vec{a})} G_{0, N+1}^{N+1, 0} \left(\begin{matrix} - \\ \vec{a}, 0 \end{matrix} \middle| \beta \varepsilon_{eq}(x) \right), \quad (4.62)$$

onde

$$Z(\beta) = \frac{1}{\Gamma(\vec{a})} \int dx G_{0, N+1}^{N+1, 0} \left(\begin{matrix} - \\ \vec{a}, 0 \end{matrix} \middle| \beta \varepsilon_{eq}(x) \right). \quad (4.63)$$

É exatamente esta a conexão entre a formulação entrópica discutida na seção 4.2 e a Termodinâmica Estocástica que acabou de ser abordada nesta seção. Definindo a energia interna como

$$U(t) = \int dx \varepsilon(x, t) p(x, t) \quad (4.64)$$

e computando a derivada temporal da energia interna, tem-se

$$\dot{U}(t) = \int dx \varepsilon(x, t) \frac{\partial p(x, t)}{\partial t} + \int dx p(x, t) \frac{\partial \varepsilon(x, t)}{\partial t}. \quad (4.65)$$

Substituindo a equação da continuidade no primeiro termo do lado direito da equação anterior,

$$\dot{U}(t) = - \int dx \varepsilon(x, t) \vec{\nabla} \cdot \vec{J}(x, t) + \int dx p(x, t) \frac{\partial \varepsilon(x, t)}{\partial t}. \quad (4.66)$$

Realizando a integração por partes o primeiro termo no lado direito da equação anterior e lembrando que a densidade de corrente de probabilidade desaparece no infinito,

$$\dot{U}(t) = \int dx \vec{\nabla} \varepsilon(x, t) \cdot \vec{J}(x, t) + \int dx p(x, t) \frac{\partial \varepsilon(x, t)}{\partial t}. \quad (4.67)$$

Substituindo a condição de balanço detalhado por $\vec{\nabla} \varepsilon(x, t)$, obtém-se a versão estocástica da primeira lei da termodinâmica

$$\dot{U}(t) = \dot{Q}(t) - \dot{W}(t), \quad (4.68)$$

onde

$$\dot{Q}(t) = -K_B T \dot{\Phi}(t) \quad (4.69)$$

é a taxa de troca de calor e

$$\dot{W}(t) = - \int dx p(x, t) \frac{\partial \varepsilon(x, t)}{\partial t} \quad (4.70)$$

é a taxa de trabalho.

Finalmente, da definição de $R(p)$ juntamente com a equação (4.6),

$$R(p) = - \frac{1}{E'(p)} = - \frac{dp}{dE}. \quad (4.71)$$

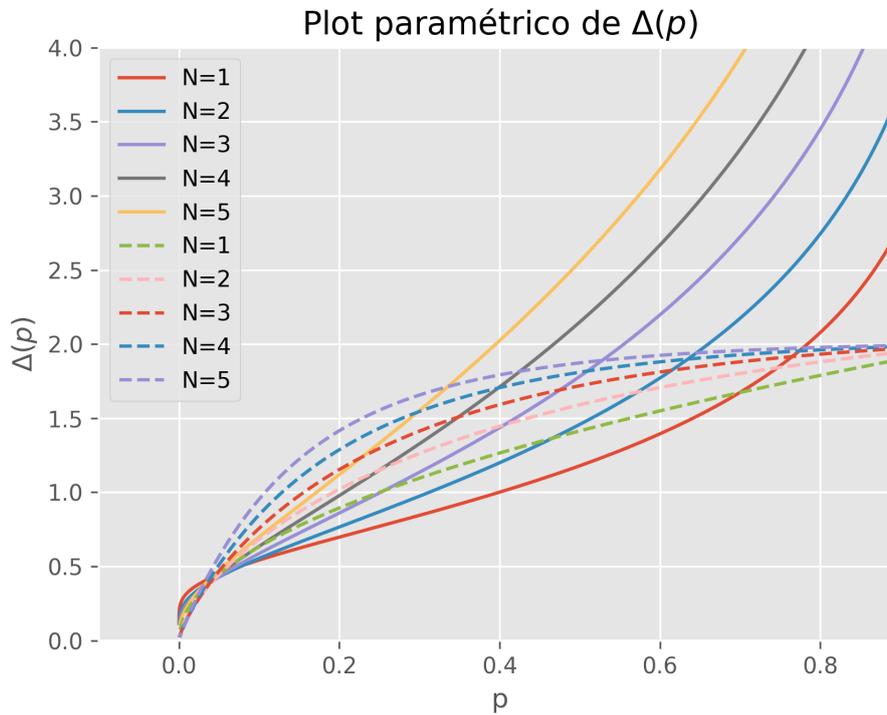
Desta forma é possível definir a função

$$R(E) = -p'(E). \quad (4.72)$$

Tendo em vista que para o ensemble de Boltzmann-Gibbs $R(p) = p$, é instrutivo introduzir a função

$$\Delta(p) = \frac{R(p)}{p} \quad (4.73)$$

Figura 31 – Plot paramétrico de $\Delta(p)$ para as diferentes classes de universalidade com N variando de 1 a 5.



Fonte: elaborada pelo autor.

para medir o desvio com relação à entropia de Boltzmann-Gibbs de um funcional entrópico dado. De forma equivalente, é possível definir a mesma função em termo da energia

$$\Delta(E) = \frac{R(E)}{p(E)}. \quad (4.74)$$

Um plot paramétrico de $\Delta(p)$ é mostrado na figura 31. Note que no gráfico a entropia se desvia fortemente da linearidade quando $p \rightarrow 0$, o que indica uma forte tendência para eventos raros, o que explica as caudas pesadas inerentes do ensemble K.

5 CONCLUSÕES E PERSPECTIVAS

5.1 CONCLUSÕES

Nesta tese foram apresentados dois temas distintos e seus desdobramentos: a modelagem matemática da pandemia de Covid-19 e versão entrópica heterotípica da Teoria-H. Começando pelo capítulo 2, foi feita uma revisão histórica do desenvolvimento dos modelos epidemiológicos, desde seu surgimento com o estudo de Daniel Bernoulli, passando pelo trabalho seminal de Kermack e McKendrick, até os modelos atuais.

Dentro do tema de modelagem matemática da pandemia de Covid-19 foram utilizadas duas classes de modelos para estudar as curvas epidêmicas: os modelos compartimentais e os modelos de crescimento logístico. A primeira classe de modelos tem vasta literatura na área e é amplamente utilizada por sua completeza na descrição de uma epidemia através de seus parâmetros, que têm uma clara interpretação epidemiológica e podem inclusive incluir a dinâmica de vetores de infecção como mosquitos ou roedores, períodos de incubação da doença ou até mesmo a demografia da população.

Apesar das vantagens de citadas, foi mostrado ao longo desta tese que os modelos compartimentais tradicionais não são capazes de descrever bem as curvas epidêmicas da Covid-19 devido a sua complexidade, trazida pelo comportamento da população e da efetividade (ou falta dela) das medidas de prevenção e mitigação implementadas pelas autoridades de saúde pública, como uso obrigatório de máscara em locais públicos, distanciamento social, fechamento de escolas ou até mesmo *lockdown*.

Os modelos de crescimento logístico, que compõem a segunda classe de modelos epidemiológicos estudados nesta tese, trazem consigo a simplicidade dos modelos fenomenológicos, descrevendo as curvas epidêmicas acumuladas em termos de parâmetros genéricos e, *a priori*, sem uma clara interpretação epidemiológica. Além disso, essa classe de modelos muitas vezes tem uma formulação matemática muito simples, o que faz com que possam ser descritos de forma explícita, contrapondo-se aos modelos compartimentais, que são formulados em termos de equações diferenciais ordinárias e necessitam de integração numérica para produzirem seus resultados.

Com esta classe de modelos foi possível descrever as curvas epidêmicas com altíssimo grau de compatibilidade através dos ajustes numéricos realizados, até mesmo nos casos mais complexos. Ainda utilizando os modelos de crescimento logístico mostramos que, utilizando

parâmetros dependentes do tempo, é possível modelar curvas epidêmicas que apresentam um número arbitrário N de ondas.

Em seguida foi apresentado o software ModInterv, que implementa todos os modelos de crescimento logístico discutidos ao longo da tese. Este software foi desenvolvido com uma interface bastante simples e intuitiva para possibilitar seu uso por vários tipos de usuários, desde pesquisadores a autoridades de saúde pública ou mesmo o público em geral. Através do ModInterv é possível acompanhar a evolução da epidemia de Covid-19 em todos os países e nas cidades e estados do Brasil e dos Estados Unidos, possibilitando analisar as curvas epidêmicas destas localidades para fazer previsões de curto prazo e determinar o estágio atual da epidemia para a curva em questão.

O segundo tema abordado nesta tese foi a formulação entrópica heterotípica da Teoria-H, na qual foi descrito o procedimento geral para encontrar um funcional entrópico a partir de uma distribuição de probabilidade. Este cálculo foi conduzido para as duas classes de universalidade previstas pela Teoria-H para um número arbitrário N de escalas hierárquicas. A conexão entre as discussões sobre o tema referido e a teoria da Termodinâmica Estocástica foi mostrada em detalhes e o papel da entropia como peça fundamental na dinâmica dos sistemas fora do equilíbrio foi ressaltada.

5.2 PERSPECTIVAS

A versatilidade dos modelos de crescimento logístico nos permitiu analisar e entender melhor a dinâmica da epidemia de Covid-19 ao redor do mundo, como a rapidez com que a curva epidêmica acumulada se aproxima do platô de saturação a depender das medidas de mitigação (ou falta delas) impostas pelas autoridades de saúde pública. Estes modelos podem ser utilizados para estudar e lançar luz sobre outras epidemias que já assolaram o mundo anteriormente.

No âmbito do software ModInterv, com sua facilidade de uso e versatilidade da aplicação de modelos matemáticos sofisticados de uma forma completamente automatizada o tornam uma ferramenta de grande valor para a classe científica e para as autoridades de saúde pública. Nosso grupo de pesquisa planeja expandir a aplicabilidade do software para receber os dados de outras pandemias que já passaram pelo mundo a fim de fornecer uma ferramenta unificada para estudo de epidemias dos mais variados tipos.

Para a formulação entrópica heterotípica da Teoria-H, planejamos incrementar seu estudo

com análises mais detalhadas em alguns pontos e com algumas aplicações teóricas e computacionais para ilustrar o potencial de sua utilização em vários fenômenos.

REFERÊNCIAS

- ANDREWS, L. C.; PHILLIPS, R. L.; SHIVAMOGGI, B. K.; BECK, J. K.; JOSHI, M. L. A statistical theory for the distribution of energy dissipation in intermittent turbulence. *Physics of Fluids A: Fluid Dynamics*, v. 1, n. 6, p. 999–1006, 1989. Disponível em: <<https://doi.org/10.1063/1.857412>>.
- BATEMAN, H. *Higher transcendental functions [volumes i-iii]*. [S.l.]: McGraw-Hill Book Company, 1953. v. 1.
- BENEDICTOW, O. J. *The black death 1346-1353: The complete history*. Woodbridge, England: Boydell Press, 2006.
- BERNOULLI, D. Réflexions sur les avantages de l'inoculation. *Mém. Paris*, p. 439–482, 1760.
- BERNOULLI, D. Essai d'une nouvelle analyse de la mortalité cause par la petite vérole et des avantages de l'inoculation pour la prévenir. *histoire de l'académie royale des sciences avec les mémoires de mathématique et de physique tirés des registres de cette académie. paris 1766 (année 1760)*. *History of Actuarial Science*, v. 8, p. 1766, 1766.
- BLUM, H. F. *Time's arrow and evolution*. Princeton, NJ: Princeton University Press, 1968.
- BRUM, A. A.; DUARTE-FILHO, G. C.; OSPINA, R.; ALMEIDA, F. A.; MACÊDO, A. M.; VASCONCELOS, G. L. Modinterv: An automated online software for modeling epidemics. *Software Impacts*, Elsevier BV, v. 14, p. 100409, dez. 2022. ISSN 2665-9638. Disponível em: <<http://dx.doi.org/10.1016/j.simpa.2022.100409>>.
- BRUM, A. A.; VASCONCELOS, G. L.; DUARTE-FILHO, G. C.; OSPINA, R.; ALMEIDA, F. A.; MACÊDO, A. M. Modinterv covid-19: An online platform to monitor the evolution of epidemic curves. *Applied Soft Computing*, Elsevier BV, v. 137, p. 110159, abr. 2023. ISSN 1568-4946. Disponível em: <<http://dx.doi.org/10.1016/j.asoc.2023.110159>>.
- CACCAVO, D. Chinese and italian COVID-19 outbreaks can be correctly described by a modified SIRD model. Cold Spring Harbor Laboratory, mar. 2020. Disponível em: <<https://doi.org/10.1101/2020.03.19.20039388>>.
- CARNOT, S. *Réflexions sur la Puissance Motrice du Feu et sur les Machines Propres à Développer cette Puissance*. Paris: Bachelier, 1824.
- CHOWELL, G. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infectious Disease Modelling*, Elsevier BV, v. 2, n. 3, p. 379–398, ago. 2017. Disponível em: <<https://doi.org/10.1016/j.idm.2017.08.001>>.
- CHOWELL, G.; SATTENSPIEL, L.; BANSAL, S.; VIBOUD, C. Mathematical models to characterize early epidemic growth: A review. *Physics of Life Reviews*, v. 18, p. 66–97, 2016. ISSN 1571-0645. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1571064516300641>>.
- CLAUSIUS, R. Ueber verschiedene für die anwendung bequeme formen der hauptgleichungen der mechanischen wärmetheorie. *Annalen der Physik*, v. 201, n. 7, p. 353–400, 1865. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.18652010702>>.

CLEVELAND, W. S. Lowess: A program for smoothing scatterplots by robust locally weighted regression. *The American Statistician*, JSTOR, v. 35, n. 1, p. 54, fev. 1981. ISSN 0003-1305. Disponível em: <<http://dx.doi.org/10.2307/2683591>>.

COMMUNITY, J. widgets. ipywidgets, a github repository. Retrieved from <https://github.com/jupyter-widgets/ipywidgets>. 2015.

CORI, A.; DONNELLY, C. A.; DORIGATTI, I.; FERGUSON, N. M.; FRASER, C.; GARSKE, T.; JOMBART, T.; NEDJATI-GILANI, G.; NOUVELLET, P.; RILEY, S.; KERKHOVE, M. D. V.; MILLS, H. L.; BLAKE, I. M. Key data for outbreak evaluation: building on the ebola experience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, The Royal Society, v. 372, n. 1721, p. 20160371, abr. 2017. ISSN 1471-2970. Disponível em: <<http://dx.doi.org/10.1098/rstb.2016.0371>>.

DAVIES, N. G.; BARNARD, R. C.; JARVIS, C. I.; RUSSELL, T. W.; SEMPLE, M. G.; JIT, M.; EDMUNDS, W. J. Association of tiered restrictions and a second lockdown with covid-19 deaths and hospital admissions in england: a modelling study. *The Lancet Infectious Diseases*, Elsevier BV, v. 21, n. 4, p. 482–492, abr. 2021. ISSN 1473-3099. Disponível em: <[http://dx.doi.org/10.1016/S1473-3099\(20\)30984-1](http://dx.doi.org/10.1016/S1473-3099(20)30984-1)>.

DAVIES, N. G.; KUCHARSKI, A. J.; EGGO, R. M.; GIMMA, A.; EDMUNDS, W. J.; JOMBART, T.; O'REILLY, K.; ENDO, A.; HELLEWELL, J.; NIGHTINGALE, E. S.; QUILTY, B. J.; JARVIS, C. I.; RUSSELL, T. W.; KLEPAC, P.; BOSSE, N. I.; FUNK, S.; ABBOTT, S.; MEDLEY, G. F.; GIBBS, H.; PEARSON, C. A. B.; FLASCHE, S.; JIT, M.; CLIFFORD, S.; PREM, K.; DIAMOND, C.; EMERY, J.; DEOL, A. K.; PROCTER, S. R.; ZANDVOORT, K. van; SUN, Y. F.; MUNDAY, J. D.; ROSELLO, A.; AUZENBERGS, M.; KNIGHT, G.; HOUBEN, R. M. G. J.; LIU, Y. Effects of non-pharmaceutical interventions on covid-19 cases, deaths, and demand for hospital services in the uk: a modelling study. *The Lancet Public Health*, Elsevier BV, v. 5, n. 7, p. e375–e385, jul. 2020. ISSN 2468-2667. Disponível em: <[http://dx.doi.org/10.1016/S2468-2667\(20\)30133-X](http://dx.doi.org/10.1016/S2468-2667(20)30133-X)>.

DIEKMANN, O.; HEESTERBEEK, J.; METZ, 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*, Springer Science and Business Media LLC, v. 28, n. 4, jun. 1990. Disponível em: <<https://doi.org/10.1007/bf00178324>>.

DIEKMANN, O.; HEESTERBEEK, J. A. P.; ROBERTS, M. G. The construction of next-generation matrices for compartmental epidemic models. *Journal of The Royal Society Interface*, The Royal Society, v. 7, n. 47, p. 873–885, nov. 2009. Disponível em: <<https://doi.org/10.1098/rsif.2009.0386>>.

DIETZ, K.; HEESTERBEEK, J. Daniel bernoulli's epidemiological model revisited. *Mathematical biosciences*, Elsevier, v. 180, n. 1-2, p. 1–21, 2002.

ENDRES, R. G. Entropy production selects nonequilibrium states in multistable systems. *Scientific Reports*, Springer Science and Business Media LLC, v. 7, n. 1, out. 2017. Disponível em: <<https://doi.org/10.1038/s41598-017-14485-8>>.

GASPER, G.; RAHMAN, M. *Basic Hypergeometric Series*. Cambridge University Press, 2004. ISBN 9780511526251. Disponível em: <<http://dx.doi.org/10.1017/CBO9780511526251>>.

GEORGESCU-ROEGEN, N. The entropy law and the economic process in retrospect. *Eastern Economic Journal*, Palgrave Macmillan Journals, v. 12, n. 1, p. 3–25, 1986. ISSN 00945056, 19394632. Disponível em: <<http://www.jstor.org/stable/40357380>>.

GONZÁLEZ, I. R. R.; LIMA, B. C.; PINCHEIRA, P. I. R.; BRUM, A. A.; MACÊDO, A. M. S.; VASCONCELOS, G. L.; MENEZES, L. de S.; RAPOSO, E. P.; GOMES, A. S. L.; KASHYAP, R. Turbulence hierarchy in a random fibre laser. *Nature Communications*, Springer Science and Business Media LLC, v. 8, n. 1, maio 2017. ISSN 2041-1723. Disponível em: <<http://dx.doi.org/10.1038/ncomms15731>>.

GONZÁLEZ, I. R. R.; RAPOSO, E. P.; MACÊDO, A. M. S.; MENEZES, L. de S.; GOMES, A. S. L. Coexistence of turbulence-like and glassy behaviours in a photonic system. *Scientific Reports*, Springer Science and Business Media LLC, v. 8, n. 1, nov. 2018. ISSN 2045-2322. Disponível em: <<http://dx.doi.org/10.1038/s41598-018-35434-z>>.

GRÜNWARD, P. D.; DAWID, A. P. Game theory, maximum entropy, minimum discrepancy and robust bayesian decision theory. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 32, n. 4, p. 1367–1433, 2004. ISSN 00905364. Disponível em: <<http://www.jstor.org/stable/3448538>>.

HARVEY, J.; CHAN, B.; SRIVASTAVA, T.; ZAREBSKI, A. E.; DŁOTKO, P.; BŁASZCZYK, P.; PARKINSON, R. H.; WHITE, L. J.; AGUAS, R.; MAHDI, A. Epidemiological waves - types, drivers and modulators in the covid-19 pandemic. *Heliyon*, Elsevier BV, v. 9, n. 5, p. e16015, maio 2023. ISSN 2405-8440. Disponível em: <<http://dx.doi.org/10.1016/j.heliyon.2023.e16015>>.

JAKEMAN, E.; PUSEY, P. N. Significance of k distributions in scattering experiments. *Phys. Rev. Lett.*, American Physical Society, v. 40, p. 546–550, Feb 1978. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.40.546>>.

JHU. *Coronavirus COVID-19 global cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)*. 2020. <<https://coronavirus.jhu.edu/map.html>>.

JOHNSON, N. P. A. S.; MUELLER, J. Updating the accounts: Global mortality of the 1918-1920 “spanish” influenza pandemic. *Bulletin of the History of Medicine*, Project MUSE, v. 76, n. 1, p. 105–115, mar. 2002. ISSN 1086-3176. Disponível em: <<http://dx.doi.org/10.1353/bhm.2002.0022>>.

KAMVAR, Z. N.; CAI, J.; PULLIAM, J. R.; SCHUMACHER, J.; JOMBART, T. Epidemic curves made easy using the r package incidence. *F1000Research*, F1000 Research Ltd, v. 8, p. 139, jan. 2019. ISSN 2046-1402. Disponível em: <<http://dx.doi.org/10.12688/f1000research.18002.1>>.

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*, The Royal Society, v. 115, n. 772, p. 700–721, ago. 1927. Disponível em: <<https://doi.org/10.1098/rspa.1927.0118>>.

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*, The Royal Society, v. 138, n. 834, p. 55–83, out. 1932. Disponível em: <<https://doi.org/10.1098/rspa.1932.0171>>.

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*, The Royal Society, v. 141, n. 843, p. 94–122, jul. 1933. Disponível em: <<https://doi.org/10.1098/rspa.1933.0106>>.

KERR, C. C.; STUART, R. M.; MISTRY, D.; ABEYSURIYA, R. G.; ROSENFELD, K.; HART, G. R.; NÚÑEZ, R. C.; COHEN, J. A.; SELVARAJ, P.; HAGEDORN, B.; GEORGE, L.; JASTRZĘBSKI, M.; IZZO, A. S.; FOWLER, G.; PALMER, A.; DELPORT, D.; SCOTT, N.; KELLY, S. L.; BENNETTE, C. S.; WAGNER, B. G.; CHANG, S. T.; ORON, A. P.; WENGER, E. A.; PANOVSKA-GRIFFITHS, J.; FAMULARE, M.; KLEIN, D. J. Covasim: An agent-based model of covid-19 dynamics and interventions. *PLoS Computational Biology*, Public Library of Science (PLoS), v. 17, n. 7, p. e1009149, jul. 2021. ISSN 1553-7358. Disponível em: <<http://dx.doi.org/10.1371/journal.pcbi.1009149>>.

LEPAN, N. *Visualizing the History of Pandemics*. 2020. Disponível em: <<https://www.visualcapitalist.com/history-of-pandemics-deadliest/>>.

LI, R.; PEI, S.; CHEN, B.; SONG, Y.; ZHANG, T.; YANG, W.; SHAMAN, J. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, American Association for the Advancement of Science (AAAS), v. 368, n. 6490, p. 489–493, maio 2020. Disponível em: <<https://doi.org/10.1126/science.abb3221>>.

LUZZI, R.; VASCONCELLOS, A. R.; RAMOS, J. G. Entropy: Mystery and controversy. plethora of informational-entropies and unconventional statistics. *arXiv: Statistical Mechanics*, 2003.

MACÊDO, A. M. S.; BRUM, A. A.; DUARTE-FILHO, G. C.; ALMEIDA, F. A. G.; OSPINA, R.; VASCONCELOS, G. L. A comparative analysis between a SIRD compartmental model and the richards growth model. *Trends in Computational and Applied Mathematics*, Brazilian Society for Computational and Applied Mathematics (SBMAC), v. 22, n. 4, p. 545–557, out. 2021. Disponível em: <<https://doi.org/10.5540/tcam.2021.022.04.00545>>.

MACÊDO, A. M. S.; GONZÁLEZ, I. R. R.; SALAZAR, D. S. P.; VASCONCELOS, G. L. Universality classes of fluctuation dynamics in hierarchical complex systems. *Phys. Rev. E*, American Physical Society, v. 95, p. 032315, Mar 2017. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.95.032315>>.

MACÊDO, A. M. S.; GONZÁLEZ, I. R. R.; SALAZAR, D. S. P.; VASCONCELOS, G. L. Universality classes of fluctuation dynamics in hierarchical complex systems. *Phys. Rev. E*, American Physical Society, v. 95, p. 032315, Mar 2017. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.95.032315>>.

MATHEMATICAL Modelling and Analysis of Infectious Diseases. Springer International Publishing, 2020. ISSN 2198-4190. ISBN 9783030498962. Disponível em: <<http://dx.doi.org/10.1007/978-3-030-49896-2>>.

MCDONALD, J. B. Some generalized functions for the size distribution of income. *Econometrica*, JSTOR, v. 52, n. 3, p. 647, maio 1984. Disponível em: <<https://doi.org/10.2307/1913469>>.

- MODINTERV-COVID19. *Cooperative Research Network on Modelling the COVID-19 Epidemic and Non-Pharmacological Interventions*. 2020. <<https://fisica.ufpr.br/redecovid19>> [Acessado: 02/11/2023].
- MORÉ, J. J. The levenberg-marquardt algorithm: Implementation and theory. In: WATSON, G. A. (Ed.). *Numerical Analysis*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1978. p. 105–116. ISBN 978-3-540-35972-2.
- NEVVILLE, M.; STENSITZKI, T.; ALLEN, D. B.; INGARGIOLA, A. *LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python*. Zenodo, 2014. Disponível em: <<https://zenodo.org/record/11813>>.
- PÉREZ, F.; GRANGER, B. E. IPython: a system for interactive scientific computing. *Computing in Science and Engineering*, IEEE Computer Society, v. 9, n. 3, p. 21–29, maio 2007. ISSN 1521-9615. Disponível em: <<https://ipython.org>>.
- PRIGOGINE, I.; GÉHÉNIAT, J. Entropy, matter, and cosmology. *Proceedings of the National Academy of Sciences*, Proceedings of the National Academy of Sciences, v. 83, n. 17, p. 6245–6249, set. 1986. Disponível em: <<https://doi.org/10.1073/pnas.83.17.6245>>.
- RICHARDS, F. J. A Flexible Growth Function for Empirical Use. *Journal of Experimental Botany*, v. 10, n. 2, p. 290–301, 06 1959. ISSN 0022-0957. Disponível em: <<https://doi.org/10.1093/jxb/10.2.290>>.
- ROSS, R. *The prevention of malaria*. [S.l.]: John Murray, 1911.
- SALAZAR, D. S. P.; VASCONCELOS, G. L. Stochastic dynamical model of intermittency in fully developed turbulence. *Physical Review E*, American Physical Society (APS), v. 82, n. 4, out. 2010. ISSN 1550-2376. Disponível em: <<http://dx.doi.org/10.1103/PhysRevE.82.047301>>.
- SALAZAR, D. S. P.; VASCONCELOS, G. L. Multicanonical distribution: Statistical equilibrium of multiscale systems. *Physical Review E*, American Physical Society (APS), v. 86, n. 5, nov. 2012. ISSN 1550-2376. Disponível em: <<http://dx.doi.org/10.1103/PhysRevE.86.050103>>.
- TSALLIS, C.; SOUZA, A. M. C. Constructing a statistical mechanics for beck-cohen superstatistics. *Phys. Rev. E*, American Physical Society, v. 67, p. 026106, Feb 2003. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.67.026106>>.
- TSOULARIS, A.; WALLACE, J. Analysis of logistic growth models. *Mathematical Biosciences*, Elsevier BV, v. 179, n. 1, p. 21–55, jul. 2002. Disponível em: <[https://doi.org/10.1016/s0025-5564\(02\)00096-2](https://doi.org/10.1016/s0025-5564(02)00096-2)>.
- VASCONCELOS, G. L.; BRUM, A. A.; ALMEIDA, F. A. G.; MACÊDO, A. M. S.; DUARTE-FILHO, G. C.; OSPINA, R. Standard and anomalous waves of COVID-19: A multiple-wave growth model for epidemics. *Brazilian Journal of Physics*, Springer Science and Business Media LLC, v. 51, n. 6, p. 1867–1883, out. 2021. Disponível em: <<https://doi.org/10.1007/s13538-021-00996-3>>.
- VASCONCELOS, G. L.; MACÊDO, A. M.; OSPINA, R.; ALMEIDA, F. A.; DUARTE-FILHO, G. C.; BRUM, A. A.; SOUZA, I. C. Modelling fatality curves of COVID-19 and the effectiveness of intervention strategies. *PeerJ*, PeerJ, v. 8, p. e9421, jun. 2020. Disponível em: <<https://doi.org/10.7717/peerj.9421>>.

VASCONCELOS, G. L.; MACÊDO, A. M. S.; DUARTE-FILHO, G. C.; BRUM, A. A.; OSPINA, R.; ALMEIDA, F. A. G. Power law behaviour in the saturation regime of fatality curves of the COVID-19 pandemic. *Scientific Reports*, Springer Science and Business Media LLC, v. 11, n. 1, fev. 2021. Disponível em: <<https://doi.org/10.1038/s41598-021-84165-1>>.

VERHULST, P.-F. Notice sur la loi que la population poursuit dans son accroissement. *Correspondance Mathématique et Physique*, v. 10, p. 113–121, 1838.

VERHULST, P.-F. Recherches mathématiques sur la loi d'accroissement de la population. *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, v. 18, p. 1–41, 1845.

VERHULST, P.-F. Deuxième mémoire sur la loi d'accroissement de la population. *Bulletins de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique*, v. 13, p. 114–123, 1847.

VIBOUD, C.; SIMONSEN, L.; CHOWELL, G. A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks. *Epidemics*, Elsevier BV, v. 15, p. 27–37, jun. 2016. Disponível em: <<https://doi.org/10.1016/j.epidem.2016.01.002>>.

WANG, X.-S.; WU, J.; YANG, Y. Richards model revisited: Validation by and application to infection dynamics. *Journal of Theoretical Biology*, v. 313, p. 12–19, 2012. ISSN 0022-5193. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0022519312003761>>.

WU, K.; DARCET, D.; WANG, Q.; SORNETTE, D. Generalized logistic growth modeling of the COVID-19 outbreak: comparing the dynamics in the 29 provinces in china and in the rest of the world. *Nonlinear Dynamics*, Springer Science and Business Media LLC, v. 101, n. 3, p. 1561–1581, ago. 2020. Disponível em: <<https://doi.org/10.1007/s11071-020-05862-6>>.