



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS - CEFETMG
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL - PPGMMC

**MODELOS SOCIOFÍSICOS INSPIRADOS EM
MODELAGEM EPIDEMIOLÓGICA APLICADOS EM
FENÔMENOS SOCIAIS:
UMA ABORDAGEM BASEADA EM DADOS DO GOOGLE TRENDS**

JÚLIO CÉSAR DE JESUS ONOFRE

Orientador: Prof. Dr. José Luiz Acebal Fernandes
Centro Federal de Educação Tecnológica de Minas Gerais - CEFETMG

BELO HORIZONTE
2025

JÚLIO CÉSAR DE JESUS ONOFRE

**MODELOS SOCIOFÍSICOS INSPIRADOS EM
MODELAGEM EPIDEMIOLÓGICA APLICADOS EM
FENÔMENOS SOCIAIS:
UMA ABORDAGEM BASEADA EM DADOS DO GOOGLE TRENDS**

Tese apresentada ao Programa de Pós-graduação em Modelagem Matemática e Computacional do Centro Federal de Educação Tecnológica de Minas Gerais, como requisito parcial para a obtenção do título de Doutor em Modelagem Matemática e Computacional.

Orientador: Prof. Dr. José Luiz Acebal Fernandes

BELO HORIZONTE

2025

O58m Onofre, Júlio César de Jesus
Modelos sociofísicos inspirados em modelagem epidemiológica aplicados em fenômenos sociais: uma abordagem baseada em dados do Google Trends / Júlio César de Jesus Onofre. – 2025.
108 f.

Tese (doutorado) apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.
Orientador: José Luiz Acebal Fernandes.
Tese (doutorado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Epidemiologia – Modelos matemáticos – Teses. 2. Epidemias – Serviços de informação – Teses. 3. Tecnologia da informação – Aspectos políticos – Teses. 4. Polarização (Ciências sociais) – Aspectos políticos – Teses. 5. Mídias sociais – Teses. 6. Notícias falsas – Teses. I. Fernandes, José Luiz Acebal. II. Centro Federal de Educação Tecnológica de Minas Gerais. III. Título.

CDD 511.8



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
COORDENAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

**“MODELOS SOCIOFÍSICOS INSPIRADOS EM MODELAGEM
EPIDEMIOLÓGICA APLICADOS EM FENÔMENOS SOCIAIS: UMA
ABORDAGEM BASEADA EM DADOS DO GOOGLE TRENDS”.**

Tese de Doutorado apresentada por **Júlio César de Jesus Onofre**, em 18 de dezembro de 2025, ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do CEFET-MG, e aprovada pela banca examinadora constituída pelos professores:

Prof. Dr. José Luiz Acebal Fernandes (Orientador)
Centro Federal de Educação Tecnológica de Minas Gerais

Documento assinado digitalmente
 **JOSE LUIZ ACEBAL FERNANDES**
Data: 23/12/2025 16:38:24-0300
Verifique em <https://validar.iti.gov.br>

Prof^a. Dr^a. Divane Marcon
Universidade Federal da Fronteira Sul

Documento assinado digitalmente
 **DIVANE MARCON**
Data: 05/01/2026 08:20:02-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Felipe Bacani
Universidade Federal de Ouro Preto

Documento assinado digitalmente
 **FELIPO BACANI**
Data: 05/01/2026 13:17:22-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Arthur Rodrigo Bosco de Magalhães
Centro Federal de Educação Tecnológica de Minas Gerais

Documento assinado digitalmente
 **ARTHUR RODRIGO BOSCO DE MAGALHAES**
Data: 23/12/2025 16:51:23-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Dênis Emanuel da Costa Vargas
Centro Federal de Educação Tecnológica de Minas Gerais

Documento assinado digitalmente
 **DENIS EMANUEL DA COSTA VARGAS**
Data: 23/12/2025 17:27:04-0300
Verifique em <https://validar.iti.gov.br>

Visto e permitida à impressão,

Prof. Dr. Alisson Marques da Silva
Coordenador do Programa de Pós-Graduação Stricto Sensu em
Modelagem Matemática e Computacional

Dedico este trabalho à minha família, pelo amor,
apoio e paciência em cada passo desta jornada.
Sem vocês, nada disso seria possível!

Agradecimentos

Agradeço, primeiramente, a Jesus Cristo, cuja presença constante ilumina meus caminhos, dando-me força, discernimento, sabedoria e serenidade sempre.

À minha família, presente de Deus em minha vida, que me apoiou incondicionalmente, especialmente à minha esposa Ivanilda, que sempre, diante dos desafios, chegava com palavras motivadoras e incentivadoras.

Ao meu orientador, expresso profunda gratidão pela sensibilidade e compreensão demonstradas, sobretudo em um momento pessoal delicado, bem como pela orientação e direcionamento que me permitiram perseverar, mesmo quando a vontade de desistir persistia.

Ao CEFETMG e aos professores do Programa de Pós-graduação em Modelagem Matemática e Computacional - PPGMMC, registro meu reconhecimento pelo conhecimento compartilhado e pela inspiração que cada ensinamento trouxe ao longo destes anos. Cada gesto, cada palavra e cada ensinamento, direta ou indiretamente, foi relevante e proveitoso na construção desta jornada.

A todos que, de alguma forma, contribuíram para a construção e conclusão desta tese, agradeço pelo apoio, pela paciência e pela colaboração. Mesmo diante de contratempos e desafios, cada contribuição tornou possível a finalização deste trabalho.

*“Na realidade do evangelho, a realidade das
coisas, costumam ganhar novas cores.”*
Marco Telles

Resumo

O rápido avanço tecnológico possibilitou amplo acesso à internet, o que trouxe muitos benefícios sociais, mas também muitas preocupações. De acordo com a Organização Mundial da Saúde (OMS), o fenômeno social contemporâneo chamado de infodemia é caracterizado pela disseminação de informações em excesso, sem critérios, e muitas delas caracterizadas por boatos e notícias falsas que dificultam o acesso às fontes confiáveis.

A denominação “infodemia” refere-se ao fato de ser possível identificar diversos elementos da fenomenologia das epidemias usuais no fenômeno moderno de disseminação de informações, principalmente em temas de saúde pública ou políticos. Nesta linha, trabalhos recentes abordaram o engajamento em redes sociais específicas como Gab, Reddit, YouTube, Instagram e Twitter, através de modelos epidemiológicos inspirados no modelo SIR.

Neste trabalho, exploramos esta analogia entre fenômenos epidemiológicos e fenômenos sociais, desenvolvendo modelos epidemiológicos para perfis de engajamento observados na série temporal fornecida pelo Google Trends. Tomando como hipótese a analogia entre suscetibilidade, infecção e recuperação com exposição à informação, engajamento e desengajamento, bem como reprodução de comportamentos sociais, foi possível produzir uma explicação para a dinâmica do engajamento social em temas polêmicos.

Como produto desta abordagem foi possível, por meio dos estudos e, consequentemente, dos modelos propostos, explorar quantidades da epidemiologia capazes de medir o potencial de disseminação de uma mensagem. Ou seja, a dinâmica de construção da tese em questão resultou positivamente e satisfatoriamente, considerando as descrições dos parâmetros envolvidos, do número de reprodução basal (R_0), do modelo teórico envolvendo descritores importantes da polarização social e política e da criação, por meio destes estudos, de um ambiente artificial que permite a exposição de uma forma, por contágio, de como pode ser propagada notícias verdadeiras e notícias falsas.

Palavras-chave: SIR. Infodemia. Google Trends. Polarização Política. Mídias Sociais. Notícias Falsas.

Abstract

The rapid advancement of technology has enabled widespread access to the internet, yielding numerous social benefits while also raising significant concerns. According to the World Health Organization (WHO), the contemporary social phenomenon known as infodemia is characterized by the excessive and indiscriminate dissemination of information, much of which consists of rumors and misinformation, thereby hindering access to reliable and trustworthy sources.

The term “infodemic” refers to the possibility of identifying several elements of the phenomenology of traditional epidemics within the modern phenomenon of information dissemination, particularly in topics related to public health or politics. In this context, recent studies have examined engagement on specific social media platforms such as Gab, Reddit, YouTube, Instagram, and Twitter using epidemiological models inspired by the SIR model.

In this study, we explore the analogy between epidemiological phenomena and social phenomena by developing epidemiological models for engagement profiles observed in the time series provided by Google Trends. By hypothesizing an analogy between susceptibility, infection, and recovery and, respectively, exposure to information, engagement, and disengagement as well as the reproduction of social behaviors it was possible to provide an explanation for the dynamics of social engagement with controversial topics.

As a product of this approach, it was possible through the studies conducted and, consequently, the proposed models to explore epidemiological quantities capable of measuring the potential for message dissemination. That is, the dynamics underlying the construction of the thesis in question yielded positive and satisfactory results, considering the descriptions of the parameters involved, the basic reproduction number ($\mathcal{R}_0$), the theoretical model encompassing important descriptors of social and political polarization, and the creation through these studies of an artificial environment that allows the exposure, in a contagion-like manner, of how true and false news can be propagated.

Keywords: SIR. Infodemic. Google Trends. Political Polarization. Social Media. Fake News.

Lista de Figuras

Figura 1 – Diagrama esquemático do modelo SIR mostrando o fluxo de indivíduos entre populações, segundo os parâmetros do modelo e também de suas populações (Eq. 1).	22
Figura 2 – Trajetórias e Campos de direções - Modelo predador-presa	27
Figura 3 – Planilha de dados do Google Trends e Gráfico do Google Trends referente as buscas do termo “Vacinação Infantil”.	30
Figura 4 – Gráfico gerado por meio dos dados do Google Trends referente as buscas do termo “Vacinação Infantil” efetuadas no Google.	30
Figura 5 – Representação unidimensional do operador de cruzamento real polarizado. .	33
Figura 6 – Representação unidimensional da operação de reflexão quando $x \notin [x_L, x_U]$. .	33
Figura 7 – Retrato de Fase Modelo SIR sem dinâmica vital Eq. (3)	41
Figura 8 – (a) Dados de frequência de buscas relativas realizadas no Google no período analisado e curva de ajuste gerada considerando os parâmetros otimizados do modelo SIR; (b) Gráficos dados GT $\times$ dados GT (bissetriz), dados GT $\times$ dados simulados e correlação de Pearson. GT = Google Trends.	42
Figura 9 – (a) Dados de frequência de buscas relativas realizadas no Google no período analisado e curva de ajuste gerada considerando os parâmetros otimizados do modelo SIR; (b) Gráficos dados GT $\times$ dados GT (bissetriz), dados GT $\times$ dados simulados via AGRP e correlação de Pearson.	43
Figura 10 – Diagrama do modelo 24	45
Figura 11 – Exemplo do Espaço de Fase Modelo 24	47
Figura 12 – (a) Frequência de buscas relativas realizadas no Google no período analisado e curva de ajuste considerando os parâmetros do modelo 24 via AGRP; (b) Gráficos dados GT $\times$ dados GT (bissetriz), dados GT $\times$ dados simulados e correlação de Pearson.	48
Figura 13 – Diagrama do modelo 28	49
Figura 14 – Exemplo do Espaço de Fase Modelo 28	51
Figura 15 – (a) Frequência de buscas relativas realizadas no Google no período analisado e curva fit considerando os parâmetros do modelo 28 via AGRP; (b) Gráfico de dispersão entre os dados obtidos pelo modelo 28 via AGRP e os dados do Google Trends.	52
Figura 16 – Relação entre Legitimidade λ e Ressonância ρ	55
Figura 17 – Diagrama referente as poluições S , I_+ , I_- e R do modelo 32	56
Figura 18 – Descrição de regiões de Estabilidade/Instabilidade do modelo 32	60
Figura 19 – Exemplo - Variações populacionais - $\lambda = 0$, com populações I_+ e I_- superpostas. .	60

Figura 20 – Resultado da evolução do modelo SIR com duas populações reagindo a uma mensagem política caracterizada pelo valor de $\lambda = 0$. A trajetória das populações I_+ e I_- estão representadas na projeção plana do espaço de estados.	61
Figura 21 – Resultado da evolução do modelo SIR com duas populações reagindo a uma mensagem política caracterizada pelo valor de $\lambda = 0$. A trajetória das populações I_+ e I_- estão representadas na projeção plana do espaço de estados.	61
Figura 22 – Evoluções das populações I_+ e I_- , considerando variação de γ_+ e de γ_- de 0.01 a 0.99 e $\beta_+ = 0.3$ e $\beta_- = 0.7$. Evolução mais acentuada da população I_- em comparação com I_+ induzida do fato de que taxa de adesão de β_- é superior a taxa de adesão de β_+ .	62
Figura 23 – Evoluções das populações I_+ e I_- , considerando variação de β_+ e de β_- de 0.0 a 1.0 e $\gamma_+ = 0.3$ e $\gamma_- = 0.7$. Evolução mais acentuada da população I_+ visto que a taxa de recuperação γ_- é superior a γ_+ , ou seja, a população I_- se recupera mais rapidamente.	62
Figura 24 – Evoluções das populações I_+ e I_- , considerando γ_+ , γ_- , variando de 0.01 a 0.99 e β_+ e β_- variando de 0.0 a 1.0.	63
Figura 25 – Gráficos dos parâmetros: Desengajamento, Legitimidade e Ressonância com evidenciamento da curva central caso $\mu = 0$.	64
Figura 26 – Curva Legitimidade descrita via nova função λ	65
Figura 27 – Dados Google Trends - Vacinação Infantil - 03/04/2022 a 19/02/2023 (semanas)	66
Figura 28 – Dados do Google Trends (Gráfico 28(a)) com potencial resposta do modelo (Ilustração 28(b)), descrevendo picos descendentes com ondas de engajamento/desengajamento.	69
Figura 29 – Curvas de Convergência referente ao modelo 23	102
Figura 30 – Boxplots referentes ao modelo 23	103
Figura 31 – Curvas de Convergência referente ao modelo 24	104
Figura 32 – Boxplots referentes ao modelo 24	105
Figura 33 – Curvas de Convergência referente ao modelo 28	106
Figura 34 – Boxplots referentes ao modelo 28	107

Lista de Tabelas

Tabela 1	– Descrição dos subintervalos particionados de forma empírica, no período de estudo em termos de dias de cada subintervalo, estágio de engajamento e desengajamento, e seus respectivos picos, conforme Figura 4.	29
Tabela 2	– Valores dos parâmetros N , β e γ otimizados via método <i>Trust Region Reflective</i> (sem limitante superior) e valor aproximado de $\mathcal{R}_0$ em cada período p_i e etapas do processo referente ao modelo (23) ajustado aos dados do Google Trends.	42
Tabela 3	– Valores dos parâmetros N , β e γ otimizados e valor aproximado de $\mathcal{R}_0$ em cada período p_i e etapas do processo referente ao modelo 23 ajustado aos dados do Google Trends por meio do AGRP	43
Tabela 4	– Tabela de Routh-Hurwitz referente ao ponto de equilíbrio P_1 e polinômio $\tilde{P}(\lambda) = -p(\lambda)$ - Equação 27	46
Tabela 5	– Valores dos parâmetros β , γ , μ_1 , μ_2 , $N(0)$, ξ e Λ otimizados e cálculo aproximado do valor de $\mathcal{R}_0$ em cada período p_i , referente ao modelo 24 . . .	48
Tabela 6	– Tabela de Routh-Hurwitz referente ao ponto de equilíbrio P_1 e polinômio $\tilde{P}(\lambda) = -p(\lambda)$ - Equação 31	50
Tabela 7	– Valores dos parâmetros β , μ_1 , μ_2 , $N(0)$, ξ e Λ otimizados e cálculo aproximado do valor de $\mathcal{R}_0$ em cada período p_i referente ao modelo 28.	51
Tabela 8	– Valores dos números básicos de reprodução $\mathcal{R}_0$ em cada período referente ao Modelo 23, Modelo 24 e Modelo 28, ajustado aos dados do Google Trends por meio do AGRP	52
Tabela 9	– Valores dos Coeficientes de correlação de Pearson (r), AIC (19) e AICc (20) ao Modelo 23, Modelo 24 e Modelo 28, ajustado aos dados do Google Trends por meio do AGRP	52
Tabela 10	– Tabela de Routh-Hurwitz - Modelo 32	59
Tabela 11	– Média, Desvio Padrão e Coeficiente de variação do valor ótimo da função objetivo encontrado, por meio do AGRP, em cada período p_i referente ao Modelo 23	108
Tabela 12	– Média, Desvio Padrão e Coeficiente de variação do valor ótimo da função objetivo encontrado, por meio do AGRP, em cada período p_i referente ao Modelo 24	108
Tabela 13	– Média, Desvio Padrão e Coeficiente de variação do valor ótimo da função objetivo encontrado, por meio do AGRP, em cada período p_i referente ao Modelo 28	108

Sumário

1 – Introdução	12
1.1 Contexto e Motivação	12
1.2 Objetivos	16
1.2.1 Objetivos Específicos	16
1.3 Organização da Tese	16
2 – Fundamentação Teórica	18
2.1 Infodemia e Redes Sociais	18
2.2 Polarização Social e Política	19
2.3 O Modelo SIR	20
2.3.1 A taxa de infecção β	24
2.3.2 A taxa de recuperação γ	24
2.4 O Modelo Predador-Presa (Lotka-Volterra)	25
2.4.1 Abordagem dinâmica do Lotka-Volterra	25
3 – Materiais e Métodos	28
3.1 Dados e Fontes - Descrição dos datasets	28
3.2 Descrição do uso e dos Métodos de Otimização	31
3.3 Função de Verossimilhança, Critério de Informação Akaike (AIC) e Critério de Informação Akaike Corrigido (AICc)	34
3.4 Matriz da Próxima Geração	36
4 – Modelagem do engajamento na Infodemia	39
4.1 Modelo tipo SIR - Clássico	39
4.1.1 Órbita e retrato de fase	40
4.1.2 Resultados Modelo SIR clássico	41
4.2 Para além do modelo SIR - Modelo tipo SIR - Aberto com interação IR	44
4.2.1 Pontos de Equilíbrio - Modelo 24	45
4.2.2 Estudo da estabilidade - Modelo 24	46
4.2.3 Resultados Algoritmo Genético	47
4.3 Para além do modelo SIR - Modelo tipo SIR - Aberto sem interação Infectado - Recuperado	48
4.3.1 Estudo da estabilidade - Modelo 28	49
4.3.2 Resultados Algoritmo Genético	51
5 – Modelagem da Polarização Social e Política	53
5.1 Contextualizações sociais e exemplos.	53

5.1.1	Associações e Relações	54
5.2	Modelo Sociofísico Baseado em SIR para Polarização	56
5.2.1	Pontos de Equilíbrio do Modelo de Polarização e Análise de estabilidade	57
5.2.2	Estudo da Estabilidade	58
5.2.3	Simulações referentes as populações	60
5.2.4	Desengajamento, Legitimidade e Ressonância	63
5.2.5	Conclusões sobre a experimentação referente ao modelo 32	64
5.2.6	Uma nova definição para a função Legitimidade	65
5.3	Modelo de Polarização com competição de notícias baseado no SIR e no Lotka-Volterra (LVSIR)	65
5.3.1	Pontos de equilíbrio do Modelo baseado no SIR e no Lotka-Volterra (LVSIR)	67
5.3.1.1	Número de Reprodução Básico $\mathcal{R}_0$ para o modelo LVSIR . . .	68
5.3.2	Exemplo de simulação para o modelo LVSIR	69
6	– Discussão	70
7	– Conclusão	73
	Referências	75
	Apêndices	79
	APÊNDICE A – Artigo publicado	80
	APÊNDICE B – Algoritmos	99
	APÊNDICE C – Curvas de Convergência e Boxplots	102
	APÊNDICE D – Parâmetros Estatísticos dos AGRP	108

Capítulo 1

Introdução

1.1 Contexto e Motivação

O evento da Covid-19, teve início em 30 de dezembro de 2019, quando o mercado de Wuhan, na China, foi fechado após um alerta epidemiológico. O vírus SARS-Cov-2, agente etiológico da Covid-19, espalhou-se rapidamente, evoluindo de casos isolados para uma pandemia, uma vez que a epidemia atingiu proporções mundiais, afetando um grande número de pessoas. Em 24 de janeiro de 2020, registrou-se o primeiro caso confirmado na Europa, na França. Em 30 de janeiro do mesmo ano, a Organização Mundial da Saúde (OMS) declarou estado de pandemia. No Brasil, em 20 de março de 2020, o Ministério da Saúde reconheceu a transmissão comunitária (THEY, 2020). A pandemia causou enormes prejuízos, tanto sociais quanto econômicos, à população global. Diante disso, pesquisadores de todo o mundo dedicaram esforços para mitigar e, possivelmente, impedir a disseminação do vírus. No período crítico entre o surgimento da epidemia, o desenvolvimento das vacinas e sua distribuição razoável entre os países, os temas relacionados à pandemia ganharam grande engajamento nas redes sociais.

Em meio a discussões sobre possíveis tratamentos para a Covid-19, as primeiras vacinas surgiram em menos de um ano e, rapidamente, a eficácia dos tratamentos e vacinas, tornou-se objeto de intensos debates marcados por desinformação e politização. Houve um processo de intensa disseminação de notícias falsas, tendenciosas, notícias em excesso, sem critérios e sensacionalistas nas redes sociais e em alguns veículos de mídia tradicionais.

A sociedade atual vive e vivenciou uma época marcada pela expansão da internet e dos fenômenos comunicacionais em massa. A disseminação acelerada de informações e o acesso quase ilimitado ao conhecimento transformaram a forma como interagimos, especialmente no que diz respeito a saúde pública e a política. Embora essas mudanças tenham trazido benefícios significativos, também geraram impactos negativos. O excesso de informações, a baixa qualidade e a exploração mal intencionada deste ecossistema, culmina em um adoecimento social de natureza informacional que visa a afetação do conceito de realidade na sociedade.

Segundo Buckland (1991 apud. (SIQUEIRA, 2011)), temos três significados distintos para o termo informação: informação como processo, informação como conhecimento e informação como coisa. No sentido aqui abordado, consideraremos informação ora como conhecimento e coisa tornando-se assim um objeto, ora como processo de contágio social ou de polarização social em relação ao tema em que a informação está inserida. Neste contexto, o processo de informação refere-se à transmissão do objeto informação. Notícias falsas, que envolvem produção

intencional de conteúdo enganoso (BAPTISTA; GRADIM, 2022) exacerbam a polarização política (REKKER, 2021) e fortalecem bolhas informacionais. Esse ciclo prejudica tanto a confiança em instituições jornalísticas quanto a capacidade de debate político, demandando novas competências críticas para filtrar a sobrecarga informacional contemporânea.

Na área da saúde pública o fenômeno de “infodemia”, tornou-se um desafio crescente. Esta disseminação rápida e descontrolada de informações, especialmente durante crises de saúde como pandemias, trouxe comportamentos de risco, desconfiança em relação às autoridades de saúde e adesão reduzida a medidas preventivas. A Infodemia expõe a sociedade ao risco do aumento de surtos de doenças e uma desorganização dos sistemas de saúde, com elevação do número de pessoas vulneráveis a doenças por vezes evitáveis.

A instrumentalização da mídia para difundir ideologias via desinformação não é nova, mas seu modus operacional sim: as redes sociais hoje centralizam esse processo. Sociedade e política migraram dos embates físicos para conflitos mediados digitalmente, onde atores criam cenários sensacionalistas sob o véu de um ‘progresso’. Como mostra (KOOPMANS, 2004), a extrema-direita alemã nos anos de 1990 já refinava discursos para promover agendas xenófobas-padrão que se replicam hoje em novas plataformas.

Atualmente, na esfera política, a internet desempenha um papel significativo, permitindo maior interação entre governantes e governados. Neste contexto, as redes sociais tornaram-se espaços para o debate político e mobilização popular. A disseminação acelerada de informações politicamente enviesadas associada ao fenômeno das “bolhas de informação” contribui para a polarização e radicalização política.

A polarização política mina a confiança nas instituições democráticas, gerando uma crise de representatividade e um adoecimento social caracterizado pelo enfraquecimento do tecido social e pelo agravamento de tensões sociais. Nesse contexto, as notícias falsas tornaram-se um fenômeno diário e recorrente, com estudos demonstrando que elas se disseminam mais rapidamente e amplamente do que as informações verdadeiras (ZHAO et al., 2020; VOSOUGHI; ROY; ARAL, 2018). Esse padrão é particularmente crítico quando envolve saúde pública, onde incertezas, boatos e desinformação criam ciclos de retroalimentação que intensificam debates sociais.

A complexidade desse fenômeno reside em três desafios interligados: (1) fontes científicas frequentemente possuem alcance limitado ou são subcompreendidas; (2) informações precisas tendem a ser mais complexas, enquanto versões falsas ou imprecisas simplificam indevidamente fenômenos complexos, tornando-se mais atraentes pelo apelo sensacionalista; e (3) o engajamento gerado por esses conteúdos amplifica desproporcionalmente seus impactos sociais. O caso da vacinação infantil durante a pandemia de COVID-19 exemplifica claramente essa dinâmica, quando discussões sobre imunização tornaram-se vetores tanto de desinformação quanto de polarização acentuada.

Um fenômeno comunicacional contemporâneo de relevância é a polarização política, processo no qual membros de um grupo social passam a adotar posições cada vez mais

homogêneas, percebendo como antagonistas todos aqueles que sustentam visões divergentes (SORJ et al., 2018).

(KOOPMANS, 2004), em sua análise sobre política e comunicação, propõe três conceitos fundamentais para compreender esse fenômeno:

Legitimidade: entendida conforme (WEBER, 1978) como a condição em que uma ordem social se sustenta quando as ações coletivas se orientam por máximas ou regras compartilhadas. (LIPPITT; POLANSKY; ROSEN, 1952) apud. (SOUSA, 1995), complementam essa noção, destacando a capacidade de liderança para influenciar comportamentos grupais, independentemente de méritos individuais ou benefícios concretos, tanto em relações interpessoais quanto sociais.

Segundo (FRENCH; RAVEN et al., 1959) apud. (SOUSA, 1995) e (RAVEN, 1958) apud. (SOUSA, 1995) temos a seguinte simplificação e descrição deste conceito: X tem o poder, o direito legítimo de influenciar Y, e, este tem a obrigação legítima de aceitar a influência. Temos então uma visão central de que a legitimação ocorre por meio de uma construção coletiva e influenciada da realidade social na qual os elementos de uma ordem social são vistos como consonantes com normas, valores e crenças que os indivíduos presumem ser amplamente compartilhadas, quer as compartilhem pessoalmente ou não (JOHNSON; DOWD; RIDGEWAY, 2006). Com isso, em um processo social legítimo, podemos entender que há um certo aumento de consenso e vice-versa.

Ressonância: em seu contexto físico, qualifica o ressonar, o desdobrar, como que provocando uma afetação que replicará o primeiro estímulo, um quase eco (ROCHA; SILVA, 2015). Podemos assim definir como o grau ou o potencial que uma mensagem tem de gerar interesse e discussão a seu respeito. Estas características se expressam em termos de engajamento social ao debate a respeito dessa mensagem.

Visibilidade: é segundo (PACKER; MENECHINI, 2006) apud. (FERREIRA; CAREGNATO, 2014) a capacidade de exposição que uma fonte ou fluxo de informação possui de, por um lado, influenciar seu público-alvo e, por outro, ser acessada em resposta a uma demanda de informação. Ou seja, podemos vê-la como sendo o alcance que uma mensagem é coberta pela mídia de massa.

Em síntese, por meio da Ressonância, haverá uma geração de Visibilidade impulsionada por questões ou situações, a priori legítimas, referente a um tema e para as pessoas envolvidas. Em contextos de polarização política, aumenta significativamente a probabilidade de disseminação de informações tendenciosas que sustentem visões extremadas. Esses conteúdos, frequentemente assimilados por segmentos expressivos da sociedade, cumprem uma dupla função: reforçar crenças pré-existentes e consolidar identidades grupais. O resultado é a criação de um ecossistema informacional particularmente receptivo à desinformação, onde falsas narrativas não apenas são aceitas, mas também sistematicamente, amplificadas.

A semelhança entre os fenômenos comunicacionais descritos e seu potencial de disseminação com os fenômenos biológicos, em particular com a epidemiologia, sugere a utilização da modelagem matemática para entender e antecipar estes processos. Portanto, o tratamento teórico destes fenômenos comunicacionais com o uso de modelos epidemiológicos, além de ser relevante

dada a interdisciplinaridade entre essas áreas, possibilita a associação de métricas de análise das epidemias aos fenômenos comunicacionais sociais em rede, auxiliando nas implementações de políticas públicas educacionais emergentes. Alguns trabalhos recentes foram dedicados ao estudo e modelagem do comportamento nas redes sociais do ponto de vista do fenômeno Infodemia (CINELLI et al., 2020) e da disseminação de *fakenews* (SKAZA; BLAIS, 2017).

Para analisar essas dinâmicas informacionais, optou-se pela utilização do Google Trends, ferramenta integrante do ecossistema de serviços do Google. Esta plataforma de acesso aberto possibilita: (1) o monitoramento de padrões de buscas sobre temas específicos, e (2) a identificação de flutuações temporais no interesse público. Conforme demonstra (NUTI et al., 2014), tal recurso oferece valiosos *insights* sobre comportamentos coletivos, particularmente em questões de saúde pública, categoria especialmente sensível aos fenômenos de desinformação anteriormente discutidos.

A ferramenta mostra a frequência relativa ao longo do tempo como uma razão da frequência máxima alcançada (fixada em 100 pontos) no período de interesse para uma região geográfica que pode ser previamente selecionada. Portanto, o Google Trends fornece séries temporais de frequências de pesquisa relativas a um termo escolhido onde pode-se definir o intervalo de tempo e a região geográfica das ocorrências. Por meio dessa série temporal, é possível evidenciar pulsos de engajamento e desengajamento desencadeados por fatos nas redes sociais e, na sequência, evidenciados por campanhas de informação. Consequentemente, a ferramenta fornece um método para avaliar o grau de engajamento relativo no fenômeno infodêmico. Nesta pesquisa, em um primeiro momento, propomos estudar por meio de dados obtidos do Google Trends ao tema "Vacinação Infantil" em certo período, aplicados a modelos epidemiológicos inspirados no modelo SIR, o comportamento social do engajamento/desengajamento da população brasileira em torno deste assunto.

Como segundo eixo investigativo, propõe-se um modelo teórico inspirado na estrutura da epidemiologia, adaptado para analisar dinâmicas de polarização ideológica. Nesta analogia, as "populações infectadas" representam grupos antagônicos que adotaram posições políticas diametralmente opostas. A inovação metodológica reside na articulação de modelo matemático com as três variáveis das ciências sociais: Legitimidade, Visibilidade e Ressonância. Como evolução de modelos, um segundo modelo híbrido entre a estrutura básica do modelo SIR e a dinâmica presa-predador de Lotka-Volterra, descreve os processos de engajamento mediados pela competição entre notícias verdadeiras e falsas. A análise preliminar revela que o comportamento oscilatório das populações 'infectadas' (derivado do componente SIR) exhibe padrões pseudo-periódicos qualitativamente similares às curvas de engajamento observadas empiricamente no Google Trends. Essa correspondência sugere que os mecanismos de competição informacional podem ser modelados como sistemas dinâmicos não-lineares, onde fases de intenso engajamento alternam-se com períodos de relativo desinteresse público.

1.2 Objetivos

Explorar a analogia entre fenômenos epidemiológicos e fenômenos sociais para desenvolver modelos epidemiológicos para descrever a dinâmica dos perfis de engajamento observados na série temporal fornecida pelo Google Trends. Além disso, gerar uma base matemática dinâmica para trabalhos sociológicos sobre polarização política. Com os modelos dinâmicos, propor descritores típicos da epidemiologia como o número de reprodução basal $\mathcal{R}_0$ para mensurar o potencial de engajamento e desengajamento social a assuntos polêmicos.

1.2.1 Objetivos Específicos

- Construir por meio do modelo SIR e suas variações já consolidadas, modelos teóricos inspirados na epidemiologia por meio da correspondência entre estes e elementos da fenomenologia comportamental social;
- Empregar os dados obtidos pelo Google Trends para modelar e estudar a evolução do tema “Vacinação Infantil” no período de 01/12/2021 a 27/02/2022;
- Otimizar os parâmetros do modelo SIR aplicados aos dados obtidos pelo Google Trends por meio da otimização baseada no gradiente através do método Trust Region Reflective TRR;
- Otimizar os parâmetros do modelo SIR aplicados aos dados obtidos pelo Google Trends por meio de algoritmos genéticos com cruzamento polarizado AGRP;
- Propor e executar adaptações no modelo SIR com o intuito de evoluir para modelos mais complexos cuja dinâmica entre as populações sejam mais realistas e/ou mais abrangentes utilizando a inclusão de outras dinâmicas populacionais;
- Extrair quantidades da epidemiologia que possam servir como métricas para observação e ações em políticas públicas ao fenômeno social Infodemia;
- Propor um modelo sociofísico baseado no modelo SIR, que descreva o fenômeno de polarização política considerando fatores que induzem o comportamento da sociedade em torno deste tema que são a legitimidade, ressonância e visibilidade.
- Propor um modelo híbrido (Lotka-Volterra + SIR - LVSIR) para notícias falsas/verdadeiras.

1.3 Organização da Tese

No [Capítulo 2](#) são apresentados alguns dos principais conceitos necessários que fundamentam o desenvolvimento deste trabalho. No [Capítulo 3](#) são apresentados os procedimentos, métodos e técnicas que foram considerados e adotados nesta pesquisa de forma detalhada e

com o objetivo de obtenção dos resultados, apresentados também nesta seção. Este de forma fragmentada se constrói, em algumas seções onde cada uma delas apresenta um modelo como proposta, sua construção e abordagem, simulações e execuções e possíveis aplicabilidades no contexto social. No [Capítulo 4](#) é apresentada a abordagem e simulações pertinentes ao tema Infodemia, considerando um banco de dados obtidos do Google Trends e também são apresentadas variações aplicadas ao modelo SIR original e suas simulações. No [Capítulo 5](#) é apresentado um modelo que descreve a polarização social política, construído partindo como base o modelo SIR básico sem abordagem demográfica. Na [Seção 5.3](#) é apresentado o modelo LVSIR, que baseia-se no modelo Presa-Predador (Lotka-Volterra) juntamente com o modelo SIR e que descreve o comportamento de engajamento/desengajamento por meio de curvas pseudoperiódicas. No [Capítulo 6](#) são apresentados os resultados finais descrevendo impressões e pontuações pertinentes aos modelos mais abordados e estudados durante os períodos desta pesquisa. Por fim, no [Capítulo 7](#) são apresentadas as conclusões, bem como as perspectivas de trabalhos futuros.

Capítulo 2

Fundamentação Teórica

2.1 Infodemia e Redes Sociais

Segundo a OMS (Organização Mundial da Saúde), o fenômeno social caracterizado pelo excesso de informações precisas ou imprecisas que dificulta o acesso a fontes e orientações confiáveis quando necessárias é chamado de infodemia ((WHO) et al., 2020). Este fenômeno Infodemia, recebe esse nome devido à semelhança da dinâmica da disseminação de informações e a disseminação de uma doença em uma epidemia (ZAROCOSTAS, 2020). A intensa disponibilização de informações por meio de canais de distribuição de notícias e redes sociais é fortemente alimentada por polêmicas, debates e também pela busca legítima e necessária de fontes confiáveis. Tal ebulição resulta em grande engajamento social sobre um determinado tema, normalmente polêmico e como consequência desse processo, espera-se um aumento considerável nas buscas por informações nas redes.

Muitos pesquisadores procuraram estudar a dinâmica deste tipo de fenômeno onde destacamos os trabalhos de (VOSOUGHI; ROY; ARAL, 2018; ZHAO et al., 2020) e (CINELLI et al., 2020) que abordam a disseminação de informações falsas em comparação com informações verdadeiras em algumas categorias onde conclui-se que estas se espalham de forma diferente. Em (VOSOUGHI; ROY; ARAL, 2018) especificamente é concluído que o efeito dessa propagação é particularmente acentuado quando se trata de notícias de políticas falsas e que os humanos têm uma maior propensão a compartilhar informações falsas, o que contribui para sua propagação mais ampla. De uma forma mais específica este texto destaca a prevalência das notícias falsas em comparação com as verdadeiras em diferentes áreas, explorando possíveis razões para esse fenômeno e enfatizando o papel dos seres humanos na disseminação dessas informações enganosas. Já em (CINELLI et al., 2020), temos a descrição de um estudo sobre a disseminação de informações relacionadas ao COVID-19 em diferentes plataformas de mídia social, como Twitter, Instagram, YouTube, Reddit e Gab, onde o importante Número de Reprodução Básico $\mathcal{R}_0$ é obtido e onde uma relação entre o valor deste e o surgimento de um processo infodêmico é estabelecido.

Neste estudo os pesquisadores observaram que havia diferentes volumes de desinformação em cada plataforma, ou seja, que algumas plataformas continham uma quantidade maior de informações incorretas ou enganosas sendo compartilhadas. Uma descoberta interessante e surpreendente foi que as informações provenientes de fontes confiáveis e questionáveis não apresentaram padrões de disseminação diferentes. Isto significa que as informações falsas não

se espalharam necessariamente com mais facilidade do que as informações verdadeiras nas redes sociais estudadas. Ou seja, esta dinâmica está mais relacionada a importância do assunto agregado a (des)informação compartilhada do que o valor lógico da informação em si (verdadeira ou falsa).

Sabe-se que a utilização da plataforma Google para buscas de páginas e serviços é universalmente difundida. Esse fato justifica a suposição de que o aumento no volume de buscas pode ser considerado um indicador de engajamento com um determinado tema e, consequentemente, um potencial indicador do processo infodêmico quando este está associado ao tema em questão.

Diante disso, adotamos como métrica de engajamento a série temporal obtida por meio do Google Trends (GT). Essa ferramenta foi utilizada em conjunto com o modelo SIR e duas de suas derivações para descrever curvas de ajuste, soluções dos respectivos modelos epidemiológicos aplicados aos dados do GT. Dessa forma, foi possível obter os parâmetros característicos desses modelos.

2.2 Polarização Social e Política

Um fenômeno que está bem em evidência é a polarização social derivada de uma polarização política. Trata-se do resultado de um complexo sistema de incentivos, tecnologias, identidades e instituições políticas, não sendo necessariamente ruim, visto que a pluralidade de visões e ideias enriquece uma cultura e faz com que as autoridades se preocupem em adotar políticas públicas abrangentes ((KLEIN, 2020) apud. (NUNES; TRAUMANN, 2023)). O que não é saudável é quando a polarização torna-se extremista, impedindo o debate de ideias de forma saudável, justa, verdadeira, sem ideologias cegas e incisivas, fato que justifica a atenção para este fenômeno.

O sociólogo Ruud Koopmans, pesquisou política e comunicação e propôs três conceitos que permeiam a Polarização, à saber, *legitimidade*, *ressonância* e *visibilidade* (KOOPMANS, 2004). Em seus estudos, ele procura mostrar que “mensagens polarizadoras sempre têm legitimidade questionável”(FORTE, 2021).

Sobre *Legitimidade*, Weber argumentou que uma ordem social é legítima somente se a ação for aproximadamente ou em média orientada por certas máximas ou regras (JOHNSON; DOWD; RIDGEWAY, 2006)¹. E, por esta ótica, ele oferece uma visão central de que a legitimação ocorre por meio de uma construção coletiva da realidade social na qual os elementos de uma ordem social são vistos como consonantes com normas, valores e crenças que os indivíduos presumem ser amplamente compartilhadas (JOHNSON; DOWD; RIDGEWAY, 2006). Sendo assim, em um processo social legítimo, podemos entender que há um certo aumento de consenso e vice-versa. O que está diretamente ligado com a ideia central de Koopmans.

¹ apud.(WEBER, 1978)

Desta forma, neste estudo, adotaremos uma referência social para a legitimidade do seguinte modo: uma ideia ou declaração será mais legítima quanto maior for o consenso social (positivo) em torno dela e será menos legítima quanto menor for o consenso social (negativo) em relação a esta ideia ou declaração.

Para Koopmans a *ressonância* é a capacidade que uma mensagem tem de gerar discussão a respeito de si mesma ou que está associada ao engajamento social ao debate de uma ideia. Atualmente, uma das ferramentas que estão sendo usadas para gerar ressonância são as mídias sociais.

Apesar de sua importante aplicabilidade na difusão de informações, nas construções de conhecimento, nas conscientizações relevantes em torno de uma ideia ou mensagem ou simplesmente pela sua capacidade de permitir um tipo de proximidade entre grupos ou pessoas, as mídias sociais têm sido utilizadas também para trazer dúvidas, transtornos, medo e caos para a população. Por exemplo, quando estas são utilizadas para difundir uma determinada *fakenews*, pode gerar desconfortos cujas desconstruções sejam demoradas ou até mesmo impossíveis de serem executadas.

A Visibilidade, por sua vez, seria o alcance que uma mensagem é coberta pela mídia de massa. Ou seja, é o alcance que uma mensagem possui utilizando os meios de comunicação. Para Koopmans, ressonância gera visibilidade e questões polêmicas, ou seja, sem legitimidade, sempre têm vantagem na disputa por espaço na mídia, isso possivelmente por sua atratividade intrínseca.

Mensagens que ressoam, seja de forma negativa ou positiva, tornam-se aos olhos de jornalistas e editores, mais relevantes e os atores por trás delas mais proeminentes, aumentando assim a chance de alcançar um alto nível de visibilidade para mensagens semelhantes no futuro.(KOOPMANS, 2004)²

Koopmans relaciona esse fenômeno por meio destes conceitos que considera ser três mecanismos de seleções e que ele rotula como sendo “oportunidades discursivas”, aborda a mudança na interação entre movimentos sociais e autoridades políticas e que esta interação não trata-se mais de um confronto direto e sim de encontros mediados pela mídia, correspondendo a um processo evolutivo, onde apenas algumas mensagens polêmicas conseguem espaço na mídia cuja disseminação é fortalecida por meio destes três mecanismos. No texto é apresentado exemplos da extrema direita alemã nos anos de 1990 para ilustrar como suas estratégias evoluíram devido às reações na mídia (KOOPMANS, 2004).

2.3 O Modelo SIR

O primeiro modelo em epidemiologia matemática foi o da propagação da varíola proposto por Daniel Bernoulli por volta do ano 1760 para estudar a eficácia da inoculação controlada da varíola como método de vacinação. Em 1906, W. H. Hamer sugeriu que a propagação

² Messages that resonate, be it negatively or positively, become in the eyes of journalists and editors more relevant and the actors behind them more prominent, thus increasing the speaker's chances to achieve a high level of visibility for similar messages in the future.(KOOPMANS, 2004)

de uma doença infecciosa dependa do número de indivíduos suscetíveis e de infectados. Em 1927, Kermack e McKendrick consolidaram essas ideias propondo o modelo SIR, que envolvem equações diferenciais para descrever a disseminação de doenças infecciosas entre três subpopulações de indivíduos, à saber, os indivíduos suscetíveis, os infectados e os indivíduos recuperados (BRAUER, 2017; COLOMBO; DIAMANTI, 2015). Desde então, o modelo SIR tem sido utilizado como modelo básico e com sucesso, para prever o efeito de epidemias. O modelo SIR fornece ferramentas para estudar por meio de simulações computacionais, a eficácia de ações de controle tais como a vacinação e o distanciamento social.

Desta forma, o modelo SIR constitui-se de uma descrição matemática que descreve a evolução de epidemias e que faz parte de uma classe de modelos chamada de modelos compartimentais ou populacionais. Por permitir trabalhar em classes populacionais e por auxiliar em inferências sobre epidemias e métodos de controle, o modelo e suas variações se mostraram úteis por décadas por agregarem simplicidade e correção na descrição da dinâmica de disseminação de doenças em populações.

No modelo SIR, a população é dividida em três subpopulações (compartimentos) $S(t)$ (suscetíveis), $I(t)$ (infectados ou infecciosos) e $R(t)$ (recuperados), que variam com o tempo, ou seja a população de suscetíveis a adquirir a doença, a população infectada pela doença e que pode transmiti-la e a população recuperada da doença. Neste modelo específico, os indivíduos recuperados uma vez recuperados não voltam a ser suscetíveis e por este motivo pode-se considerar também que os mortos fazem parte desta classe. No caso este sistema é fechado, o que significa que a soma das três populações em cada instante t é igual a uma constante N , ou seja, $S(t) + I(t) + R(t) = N$. A evolução das populações do modelo SIR no tempo é governada por regras dinâmicas que podem ser escritas em tempo contínuo ($t \in \mathbb{R}$), usando equações diferenciais, ou em tempo discreto, ($t \in \mathbb{Z}_+$), usando equações discretas.

A figura 1 mostra um diagrama do modelo SIR em tempo discreto. As transições de indivíduos entre populações em um intervalo de tempo correspondem a frações dessas populações. Assim, no tempo t , ocorre a interação não linear de infecção entre os indivíduos das populações $S(t)$ e $I(t)$. Os valores dessas populações compõem a taxa $\frac{\beta S(t)I(t)}{N}$ de indivíduos suscetíveis que se infectam no intervalo de tempo e mudam para a população infectada. O parâmetro β é proporcional à taxa de contato social vezes a fração de contatos que realmente produzem contágio e está definida na seção 2.3.1. Da mesma forma, a população $I(t)$ cede uma fração de seus indivíduos a uma taxa $\gamma I(t)$ à população de recuperados. O parâmetro γ é a taxa de recuperação e que está definida na seção 2.3.2.

A dinâmica do modelo SIR de tempo discreto pode ser representada por meio de um

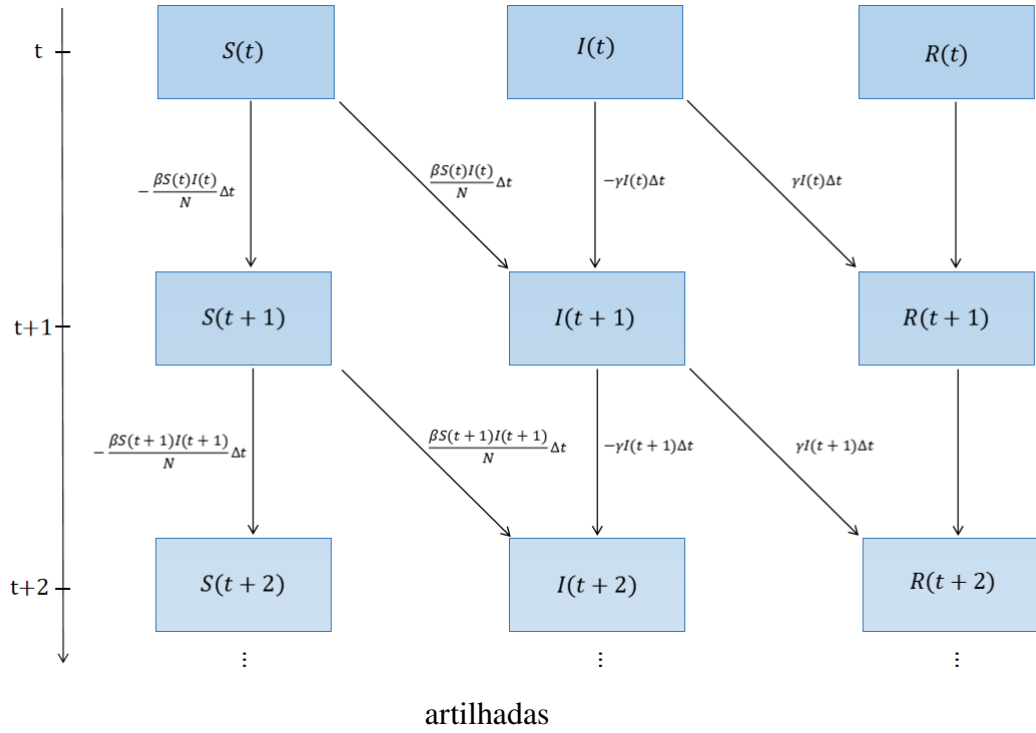


Figura 1 – Diagrama esquemático do modelo SIR mostrando o fluxo de indivíduos entre populações, segundo os parâmetros do modelo e também de suas populações (Eq. 1).

sistema de equações discretizadas:

$$\begin{cases} S(t+1) = S(t) - \left[\frac{\beta S(t)I(t)}{N} \right] \Delta t \\ I(t+1) = I(t) + \left[\frac{\beta S(t)I(t)}{N} - \gamma I(t) \right] \Delta t \\ R(t+1) = R(t) + [\gamma I(t)] \Delta t \end{cases} \quad t \in \mathbb{Z}_+ \quad (1)$$

E a dinâmica do modelo SIR em tempo contínuo é descrito pelo conjunto de equações diferenciais

$$\begin{cases} \frac{dS}{dt} = -\beta S \frac{I}{N} \\ \frac{dI}{dt} = \beta S \frac{I}{N} - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases} \quad t \in \mathbb{R}_+ \quad (2)$$

Podemos ainda normalizar as funções S , I e R , utilizando as seguintes definições, $s(t) = \frac{S(t)}{N}$, $i(t) = \frac{I(t)}{N}$ e $r(t) = \frac{R(t)}{N}$, desta forma o sistema de EDO 2 terá a forma

$$\begin{cases} \frac{ds}{dt} = -\beta si \\ \frac{di}{dt} = \beta si - \gamma i \\ \frac{dr}{dt} = \gamma i \end{cases} \quad (3)$$

A condição $S(t) + I(t) + R(t) = N$ na forma normalizada ficará $s(t) + i(t) + r(t) = 1$.

Além dos parâmetros do modelo, para determinar a intensidade do processo infeccioso utiliza-se o *Número Básico de Reprodução*, que por simplicidade definiremos aqui por: $\mathcal{R}_0 = \frac{I(1)}{I(0)}$, representa o número médio inicial de infecções secundárias produzidas por um único indivíduo infectado introduzido em uma população completamente suscetível durante o período infeccioso (CONTRERAS et al., 2020). Assim, considerando as aproximações do início de um processo infeccioso isto é que $S(0) \approx N$, $I(0) \approx 1$ (pelo menos um indivíduo infectado) e $R(0) = 0$ (não há indivíduos recuperados inicialmente) e considerando a equação para $I(t)$ no (1) podemos chegar a seguinte descrição de $\mathcal{R}_0$

$$\mathcal{R}_0 = \frac{I(\Delta t)}{I(0)} = \frac{I(0) + \left. \frac{dI}{dt} \right|_{t=0} \Delta t}{I(0)} = 1 + \left[\frac{\beta S(0)}{N} - \gamma \right] \Delta t = 1 + \left[\frac{\beta S(0)}{\gamma N} - 1 \right] \gamma \Delta t \approx 1 + \left[\frac{\beta}{\gamma} - 1 \right] \gamma \Delta t, \quad (4)$$

Fixando a escala de tempo característica $\Delta t = \frac{1}{\gamma}$, o que implicará em

$$\mathcal{R}_0 = \frac{\beta}{\gamma}. \quad (5)$$

Pode-se obter condições sobre este parâmetro através do modelo em tempo contínuo procurando estabelecer a condição para que $\frac{dI}{dt} > 0$, ou seja, verificando a condição para que a população de infectados seja crescente no início do processo, ou seja, $S(0) \approx N$ temos $\beta - \gamma > 0$ o que implica $\frac{\beta}{\gamma} > 1$, ou seja $\mathcal{R}_0 > 1$.

Portanto, esse número é expresso em termos dos parâmetros do modelo. Embora existam definições mais gerais para o número de reprodução como $\mathcal{R}_t = \frac{I(t+1)}{I(t)}$ ou $\mathcal{R}_e(t)$ que expressa uma proporção que varia com a evolução do processo, este é chamado de número efetivo de reprodução no instante t . Podemos efetuar a seguinte construção para encontrar $\mathcal{R}_t$. Considerando que $\mathcal{R}_t = \frac{I(t+1)}{I(t)}$ e $\gamma \Delta t = 1$ temos:

$$\mathcal{R}_t = \frac{I(t+\Delta t)}{I(t)} = \frac{I(t) + \left[\beta \frac{S(t)I(t)}{N} - \gamma I(t) \right] \Delta t}{I(t)} = 1 + \left[\beta \frac{S(t)}{N} - \gamma \right] \Delta t = 1 + \left[\frac{\beta}{\gamma} \frac{S(t)}{N} - 1 \right] \gamma \Delta t = 1 + \frac{\beta}{\gamma} \frac{S(t)}{N} - 1 = \frac{\beta}{\gamma} \frac{S(t)}{N} = \mathcal{R}_0 \cdot \frac{S(t)}{N}. \quad (6)$$

Neste sentido $\mathcal{R}_t(t)$ representa o número médio de infecção secundárias causadas por um único indivíduo infectado em um tempo t dada a realidade instantânea da população, que não necessariamente é totalmente suscetível, isto é, onde possivelmente temos apenas uma fração da população total suscetível.

De qualquer forma, a priori, nos restringiremos a $\mathcal{R}_0$, buscando enxergar o que desencadeou a onda de engajamento e desengajamento no início do processo tendo por premissa que este valor é compatível com uma epidemia que começou no passado com $\mathcal{R}_0$ dado pela equação (5). Observe que para $t = 0$ e considerando a condição de homogeneidade que $\mathcal{R}_t(t = 0) = \mathcal{R}_0$.

Vale ressaltar que existem outras formas mais precisas, abrangentes e matematicamente coerentes de se obter o número básico de reprodução e um destes métodos e amplamente utilizado é o método da matriz da próxima geração cuja abordagem está na seção 3.4.

2.3.1 A taxa de infecção β

A taxa de infecção, coeficiente de transmissão ou taxa efetiva de transmissão β é um parâmetro positivo que controla a transferência entre os compartimentos S e I no modelo SIR, esta significa a taxa média de contatos por infectados em dada unidade de tempo. Ela, de uma forma direta é descrita pelo produto entre o número médio de contatos c por unidade de infectados por unidade de tempo e uma probabilidade de transmissão por contato (p_t). Ou seja,

$$\beta = c \cdot p_t \quad (7)$$

Por exemplo, imagine que tenhamos 2 pessoas infectadas em um mesmo ambiente compartilhado com 14 pessoas, desta forma em média cada infectado encontra 7 pessoas ao dia ou seja, 7 contatos/dia e p_t representa a probabilidade de transmissão por contato, isto é, a chance do patógeno ser transmitido em cada encontro, por exemplo $p_t = 0,1$ significa que a chance de transmissão é de 10%. Para este caso, $\beta = 0.7$. Em outras palavras, temos a seguinte interpretação equivalente, para β , como sendo a multiplicação entre os dois fatores $\frac{\text{nº de pessoas contatadas}}{\text{nº de pessoas infectadas} \times \text{dia}}$ e probabilidade de transmissão (p_t), ou seja:

$$\beta = \frac{\text{pessoas contatadas}}{\text{pessoa infectada} \times \text{dia}} \times p_t \quad (8)$$

2.3.2 A taxa de recuperação γ

Para chegarmos numa descrição mais precisa de γ , suponhamos que toda a população esteja infectada em $t = 0$, ou seja, que as populações $S(0) = 0$ e $R(0) = 0$ e que $I(0) = N$ e que não haja reinfecções. Desta forma temos que $\frac{dI}{dt} = -\gamma I$ com $I(0) = N$. Resolvendo esta EDO temos que: $I(t) = N e^{-\gamma t}$, que é uma função decrescente. Ou seja, a medida que o tempo vai passando o número de pessoas doentes vai decrescendo.

Derivando esta função $I(t)$ concluímos que $\frac{dI}{dt} = -\gamma N e^{-\gamma t}$ o que implica $-\frac{dI}{dt} = \gamma e^{-\gamma t} > 0$. Com a evolução do tempo, todos se curam e

$$\int_0^\infty -\frac{dI}{dt} dt = \int_0^\infty \gamma e^{-\gamma t} dt = 1 \quad (9)$$

o que nos leva a concluir que a função $-\frac{dI}{N}$ é uma distribuição exponencial de parâmetro γ no intervalo $0 \leq t < \infty$. Vale ressaltar, para a conclusão anterior, que $\int_0^\infty -\frac{dI}{N} dt$ representará a variação total de infectados ao longo do tempo.

Como as pessoas que ficaram curadas em um determinado instante estiveram doentes por um período de tempo t , o tempo médio no qual as pessoas estão doentes será dado pelo valor médio de t no intervalo de tempo $0 \leq t < \infty$, ou seja,

$$\int_0^\infty -t \frac{dI}{N} dt = \gamma \int_0^\infty t e^{-\gamma t} dt = \frac{1}{\gamma} \quad (10)$$

Desta forma $\frac{1}{\gamma}$ é o tempo médio para o qual as pessoas ficam doentes. Temos que se $\frac{1}{\gamma} = \Delta t \rightarrow \left[\frac{1}{\gamma} \right] = [\Delta t] \rightarrow [\gamma] = \frac{1}{\text{unidade de tempo}}$. Se o sistema estiver sendo analisado considerando t dado em dias, por exemplo, a unidade de γ será dias⁻¹.

2.4 O Modelo Predador-Presa (Lotka-Volterra)

O modelo de Lotka-Volterra (modelo predador-presa) é um sistema de equações diferenciais utilizado para descrever a dinâmica de interações entre duas espécies em um ecossistema. Foi proposto por Alfred J. Lotka e Vito Volterra na década de 1920. O modelo trata da interação entre duas espécies, onde uma delas (presa) dispõe de alimentos em abundância e a segunda espécie (predador) tem como suprimento alimentar exclusivamente a população de presas (BASSANEZI, 2010). O modelo em sua forma mais básica consiste em duas equações diferenciais não lineares:

$$\begin{cases} \frac{dF}{dt} = \alpha F - \beta FN \\ \frac{dN}{dt} = -\gamma N + \delta FN \end{cases} \quad (11)$$

Neste sistema F , N representam respectivamente a população de presas e de predadores, α é a taxa de crescimento das presas, γ é a taxa de mortalidade dos predadores, β é a taxa de decréscimo das presas por predação e δ é a taxa de crescimento dos predadores por predação.

2.4.1 Abordagem dinâmica do Lotka-Volterra

Considerando as equações deste sistema, ou seja, as equações do modelo predador-presa temos que este possui dois pontos de equilíbrio a saber, os pontos $(0,0)$ e $\left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$, sendo respectivamente, um ponto de sela e portanto instável e o segundo um ponto de centro estável, onde as soluções do sistema descrevem uma trajetória cíclica em seu entorno.

Podemos estudar este sistema predador-presa em torno do ponto de equilíbrio $\left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$ por meio de sua linearização. Desta forma considere $u = F - \frac{\gamma}{\delta}$ e $v = N - \frac{\alpha}{\beta}$, desta forma o sistema inicial, desprezando os termos não-lineares será dado por:

$$\begin{cases} \frac{du}{dt} = -\frac{\beta\gamma c}{\delta}v \\ \frac{dv}{dt} = \frac{\alpha\delta}{\beta}u \end{cases} \quad (12)$$

Desta forma as trajetórias deste sistema linear serão dadas por:

$$\frac{dv}{du} = -\frac{\frac{\alpha\delta}{\beta}u}{-\frac{\beta\gamma c}{\delta}v} \rightarrow \alpha\delta^2 u du + \beta^2 \gamma v dv = 0 \rightarrow \alpha\delta^2 u^2 + \beta^2 \gamma v^2 = k^2,$$

o que mostra que para valores próximos do ponto crítico $(u,v) = (0,0)$ e consequentemente $(F,N) = \left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$, as trajetórias para o sistema linearizado são elípticas. Podemos ainda, de mão desta linearização, estudar o sistema considerando a seguinte transformação $U = \sqrt{\alpha\delta}u$ e $V = \sqrt{\gamma}\beta v$. Desta forma usando coordenadas polares temos que $U = k \cos(\theta)$ $V = k \sin(\theta)$, com $U^2 + V^2 = k^2$ e $\frac{V}{U} = \tan \theta$.

Por meio desta mudança de coordenadas temos o sistema

$$\begin{cases} \frac{dU}{dt} = -\sqrt{\alpha\gamma} V \\ \frac{dV}{dt} = \sqrt{\alpha\gamma} U \end{cases} \quad (13)$$

Assim chegamos que $k = C$ e $\theta = \sqrt{\alpha\gamma}t + \phi$, onde C e ϕ são constantes, desta forma temos as soluções nas coordenadas (u,v) :

$$U = C \cos(\sqrt{\alpha\gamma}t + \phi)$$

e

$$V = C \sin(\sqrt{\alpha\gamma}t + \phi)$$

Portanto $u = \frac{C}{\sqrt{\alpha\delta}} \cos(\sqrt{\alpha\gamma}t + \phi)$ e $v = \frac{C}{\sqrt{\gamma}\beta} \sin(\sqrt{\alpha\gamma}t + \phi)$ o que nos leva, considerando $C = K\gamma\sqrt{\alpha}$ as soluções nas coordenadas (F,N) :

$$F = \frac{\gamma}{\delta} + \frac{K\gamma}{\delta} \cos(\sqrt{\alpha\gamma}t + \phi)$$

e

$$N = \frac{\alpha}{\beta} + \frac{K\alpha}{\beta} \sqrt{\frac{\gamma}{\alpha}} \sin(\sqrt{\alpha\gamma}t + \phi)$$

onde podemos evidenciar um certo controle do período e da amplitude das populações F e N , levando em consideração a linearização do sistema.

Outra estratégia que pode ser usada para estudar a evolução das populações é por meio da resolução da equação $\frac{dF}{dN} = \frac{\alpha F - \beta FN}{-\gamma N + \delta FN}$. Ao considerar essa equação diferencial é possível encontrar a equação das soluções implícitas ou as trajetórias das populações soluções do sistema, a saber:

$$\gamma \ln F - \delta F + \alpha \ln N - \beta N = C$$

Um dos modelos construídos e que apresentamos aqui, descreve utilizando este modelo predador-presa um sistema envolvendo notícias falsas F e notícias verdadeiras N , por meio da

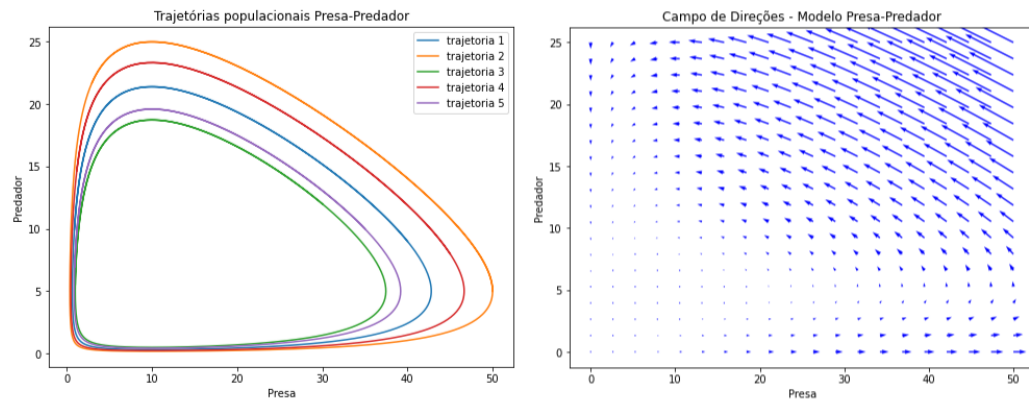


Figura 2 – Trajetórias e Campos de direções - Modelo predador-presa

respectiva relação, F com a população de presas e N com a população de predadores. Existe um sentido, no âmbito das mídias sociais e considerando algumas hipóteses, para estas descrições, haja vista que a relação entre estes dois tipos de objetos, se demonstram semelhantemente como uma espécie que se alimenta de outra.

Capítulo 3

Materiais e Métodos

3.1 Dados e Fontes - Descrição dos datasets

Um estudo de caso foi realizado para avaliar o processo de infodemia associado ao tema “vacinação infantil”, utilizando a coleta de dados do Google Trends com um filtro geográfico definido para o Brasil, no período de 01/12/2021 a 27/02/2022. A escolha deste período decorreu do crescente interesse por informações sobre a vacinação infantil, impulsionado por dois eventos marcantes: a abertura da consulta pública para vacinação de crianças de 5 a 11 anos pelo Ministério da Saúde, ocorrida entre 23 de dezembro de 2021 e 2 de janeiro de 2022 ([GOV.BR, 2021](#)) e o início (de forma simbólica) da campanha de vacinação infantil em 14 de janeiro de 2022 ([GRANCHI, 2022](#)). Estes marcos representaram passos relevantes no combate à pandemia de COVID-19 no Brasil, gerando um ambiente de busca constante por informações sobre a segurança e eficácia das vacinas para essa faixa etária. Consequentemente, este ambiente mostrou-se um ambiente de vários rumores e teorias em torno deste tema, principalmente pela urgência, necessidade e forma que tudo foi posto. Através de consulta pública foi obtida uma resposta à autorização da Anvisa para a utilização da vacina Pfizer nesse grupo. O processo de vacinação, composto por duas doses com um intervalo de três semanas entre elas, suscitou debates acalorados e enfrentou resistência de grupos contrários à imunização. No entanto, a consulta pública proporcionou um espaço para a participação ativa da sociedade civil, permitindo contribuições embasadas e reflexões pertinentes ao tema. Com isso, houve manifestações contrárias à vacinação de forma extremista, incluindo relatos de e-mails intimidatórios e, até mesmo, ameaças de morte direcionadas a diretores e servidores da Anvisa. O cronograma da consulta pública foi iniciado em 23 de dezembro de 2021, ressaltando a importância de um diálogo aberto e fundamentado em evidências científicas para o avanço da saúde pública e o bem-estar da população infantojuvenil. Considerando o gráfico de séries temporais da frequência diária relativa das pesquisas no Google sobre esse tema e a planilha com os dados correspondentes, consideramos, por meio da frequência relativa de buscas, o número de buscas como indicador de engajamento ao tema. Esta quantidade foi associada à população de infectados $I(t)$ no modelo SIR. Ou seja, considera-se $I(t)$ como indicador de engajamento ao tema. O sistema considerado é aberto no sentido de que os indivíduos podem tomar conhecimento do tema e se desinteressarem, entrando e saindo das populações de interesse do modelo SIR. Além disso, a rigor, o sistema é não autônomo, em geral, já que as taxas de contágio, recuperação e outras estão suscetíveis a notícias da mídia, campanhas de conscientização, mobilização nas redes, etc. Por este motivo, particionamos a série temporal

em intervalos, supondo que em seu interior, as taxas do modelo permaneçam aproximadamente constantes.

Dada a característica da série temporal utilizada, a dividimos em intervalos considerando os processos de engajamento (aumento) e de desengajamento (diminuição) em cada período. Mais especificamente, os dados foram divididos em oito subintervalos (ondas acentuadas de tendências), cada um deles sendo subdividido em dois períodos consecutivos, conforme apresentado na Tabela 1. Esses períodos consistem no período de engajamento(crescimento), seguido pelo período de desengajamento(decrescimento). O período de engajamento representa a sequência ascendente dos dados, que presume ser resultado de um evento midiático impulsionando o engajamento e um aumento na infecção. Já o período de desengajamento desses subintervalos corresponde ao processo de desengajamento causado por ações tomadas para informar e conscientizar o público em geral, como medidas de controle da epidemia.

Para cada período particionado, baseando-se nestas tendências de crescimento e de decrescimento, uma curva da população infectada do modelo SIR é ajustada aos dados obtendo assim, três parâmetros do modelo N , β e γ por meio da otimização do modelo SIR implementado em Python.

Tabela 1 – Descrição dos subintervalos particionados de forma empírica, no período de estudo em termos de dias de cada subintervalo, estágio de engajamento e desengajamento, e seus respectivos picos, conforme Figura 4.

Período	Dias	Pico	Período de
p_1	22º ao 24º	24/12/21 (24º dia)	Engajamento
p_2	24º ao 26º	24/12/21 (24º dia)	Desengajamento
p_3	26º ao 28º	28/12/21 (28º dia)	Engajamento
p_4	28º ao 31º	28/12/21 (28º dia)	Desengajamento
p_5	31º ao 37º	06/01/22 (37º dia)	Engajamento
p_6	37º ao 40º	06/01/22 (37º dia)	Desengajamento
p_7	40º ao 51º	20/01/22 (51º dia)	Engajamento
p_8	51º ao 89º	20/01/22 (51º dia)	Desengajamento

Após a determinação dos parâmetros, o número básico de reprodução $\mathcal{R}_0$ de cada caso é calculado para caracterizar a intensidade com a qual o processo de infodemia se desenvolveu em cada período analisado.

Inicialmente optamos por resolver numericamente o sistema de equações diferenciais ordinárias (EDOs) por meio da função *odeint* do pacote *scipy.integrate* com o método de otimização *Trust Region Reflective* disponível no pacote *scipy.optimize*, onde a função objetivo considerada foi baseada no *RMSE - Root Mean Square Error*. (Seção 3.2).

Após a determinação dos parâmetros, o número básico de reprodução $\mathcal{R}_0$, descrito na seção 2.3, foi calculado para cada período com o intuito de caracterizar a intensidade com a qual o processo de infodemia se desenvolveu.

Posteriormente, mantemos a mesma ideia inicial, mudando o método de otimização para Algoritmos Genéticos, mais precisamente para o Algoritmo Genético Real Polarizado (AGRP).

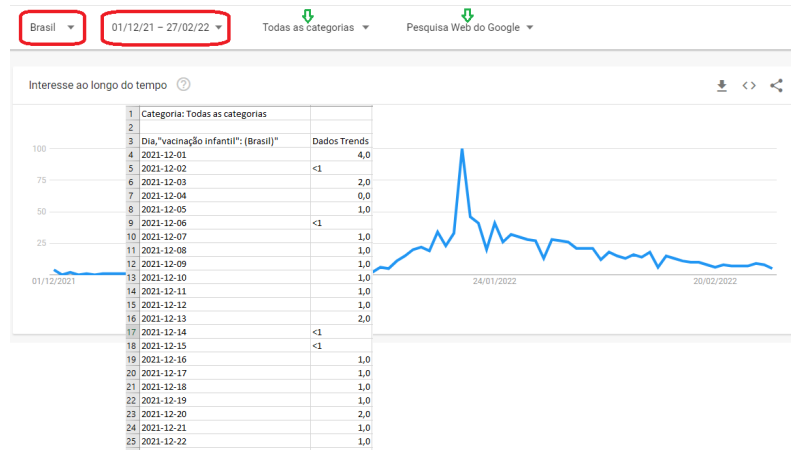


Figura 3 – Planilha de dados do Google Trends e Gráfico do Google Trends referente as buscas do termo “Vacinação Infantil”.

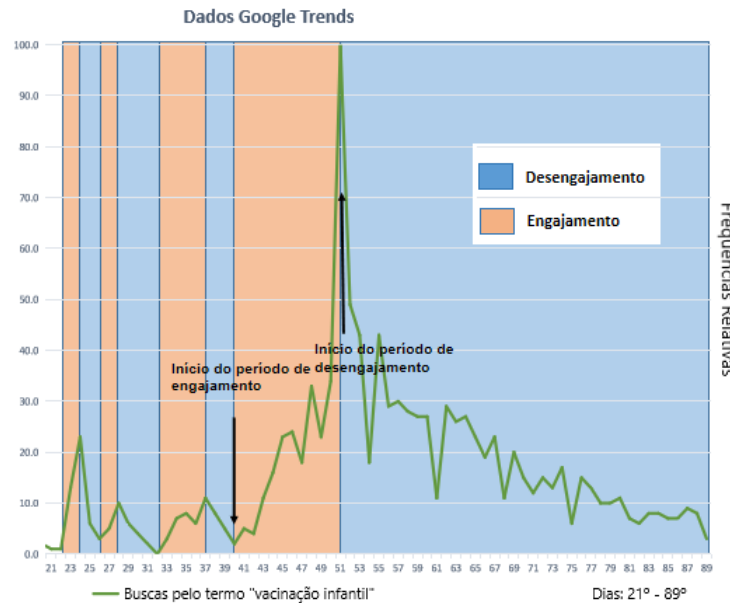


Figura 4 – Gráfico gerado por meio dos dados do Google Trends referente as buscas do termo “Vacinação Infantil” efetuadas no Google.

Desta forma, obtivemos N , β e γ por meio da otimização do modelo SIR e consequentemente o $\mathcal{R}_0$. Refinamentos no modelo SIR também foram sendo introduzidos (modelo 24 e modelo 28) procurando descrever por meio da curva de infectados $I(t)$ a curva obtida pelo Google Trends obtendo por meio de métodos de otimização os parâmetros inseridos.

Com o intuito de buscar realismo em relação ao fenômeno, novos termos foram propostos ao modelo SIR, construindo assim outros modelos sempre baseados no modelo epidemiológico SIR e que permitam descrever outros fenômenos sociais. A abordagem referente ao modelo de polarização política (modelo 32) foi construída de uma forma um pouco diferente pois o objetivo principal foi de descrever o fenômeno da polarização política por meio de três variáveis:

legitimidade, visibilidade e ressonância. Ainda sem ter posse de um conjunto de dados, nossa abordagem se deu de forma qualitativa, descrevendo o fenômeno e dando a ele uma base dinâmica. Desta forma, foram definidas estas variáveis considerando como base também o modelo SIR onde simplificações nos parâmetros do modelo foram feitas. Foi construído um modelo híbrido por meio da junção do modelo SIR e do modelo populacional Lotka-Volterra, que permite visualizar oscilações entre duas classes de notícias, notícias falsas e notícias verdadeiras, extrapolando a ideia de que uma classe é a presa e a outra é o predador, cuja dinâmica é razoável em ambientes comunicacionais.

3.2 Descrição do uso e dos Métodos de Otimização

O sistema de equações diferenciais ordinárias (EDOs) foi resolvido utilizando a função *odeint* do pacote *scipy.integrate* e, inicialmente, o método *Trust Region Reflective* disponível no pacote *scipy.optimize*¹ foi escolhido para resolver o problema de *mínimos quadrados*, mais especificamente a minimização da função objetivo $J_{p_k}^{(1)}(I_k)$ baseada, a menos de uma constante, ao *RMSE - Root Mean Square Error (Raiz Quadrada do Erro Quadrático Médio)*. Mais especificamente:

$$J_{p_k}^{(1)}(I_k) = \frac{1}{\#p_k} \sqrt{\sum_{i=1}^{\#p_k} (I_k^i - d_k^i)^2}; \quad k \in \{1, \dots, 8\} \quad 1 \leq i \leq \#p_k \quad (14)$$

Nesta descrição p_k refere-se ao k -ésimo período dos dados d_k , I_k a população do período p_k e $\#p_k$ a cardinalidade de p_k .

A técnica utilizada inicialmente para a minimização da função objetivo é também conhecida como Método da Região de Confiança Reflexiva, e trata-se de uma técnica usada em problemas de otimização matemática. A ideia básica central é a minimização da função objetivo f utilizando a sua aproximação quadrática sobre uma região de confiança, ambas construídas de forma iterativa. Mais especificamente, considera-se um ponto inicial x_0 e um tamanho inicial para a região de confiança Δ_0 . Desta forma, toma-se o gradiente $\nabla f(x)$ e a matriz Hessiana $\nabla^2 f(x)$ no ponto x_k atual. E assim utiliza-se então a aproximação quadrática (Polinômio de Taylor de segunda ordem):

$$F_k(p) = f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T \nabla^2 f(x_k) p$$

Onde p representa o vetor de passo. Portanto, é avaliada a função objetivo por meio desta aproximação e sobre a região gerada considerando este passo. Se a função objetivo melhora (ou seja, obtém-se um valor menor para f), mantém-se o passo e atualiza-se o ponto, fazendo $x_{k+1} = x_k + p$; caso contrário, mantém-se o ponto e atualiza-se o passo p .

Com relação as nossas simulações, e mais especificamente quando optou-se pelo método Algoritmo Genético Real Polarizado (AGRP), foi considerada uma função objetivo $J_{p_k}^{(2)}(I_k)$

¹ As descrições sobre ambos pacotes *scipy.integrate* e *scipy.optimize* estão disponíveis em <https://scipy.org/>.

baseada no RMSE (Root mean-square error), descrita em 15 e que foi minimizada por partes (em cada um dos 8 períodos) com termo de suavização nas descontinuidades permitindo assim uma descrição mais adequada na curva de ajuste a ser obtida por meio da obtenção do conjunto de parâmetros ótimos para cada período p_k dada por

$$J_{p_k}^{(2)}(I_k) = \frac{\sqrt{\sum_{i=1}^{\#p_k} (I_k^i - d_k^i)^2 + w_k^2 (d_k^0 - vr_k)^2}}{\#p_k + w_k}; \quad k \in \{1, \dots, 8\} \quad 1 \leq i \leq \#p_k \quad (15)$$

Em que também p_k refere-se ao k -ésimo período dos dados d_k , d_k^0 é o primeiro dado do período p_k , I_k a população do período p_k , $\#p_k$ a cardinalidade de p_k , $w_k = \sqrt{\#p_k}$ e o valor de referência $vr_k = \mu_k \cdot I_{k-1}^{final} + (1 - \mu_k) \cdot d_k^0$, onde μ_k é o fator de suavização escolhido com a condição $\mu_k \in [0, 1]$, de forma que vr_k pertença ao segmento $\overline{I_{k-1}^{final} d_k^0}$, portanto uma combinação convexa.

Conforme descrito em (SILVA et al., 2019), na etapa de cruzamento Real Polarizado, a cada geração a população corrente é dividida em duas metades, e os indivíduos são combinados em pares para reprodução. Para cada par de indivíduos (pais) selecionados, x_1 e x_2 , verifica-se a aplicação da operação de cruzamento com probabilidade de 90%. Caso o cruzamento seja aplicado, são gerados dois novos indivíduos (filhos), segundo a combinação linear

$$x_g = \mu x_1 + (1 - \mu) x_2$$

com $\mu \in (-0.1, 1.1)$ permitindo extrapolação além do segmento definido pelos pais.

Para cada geração de x_g , verifica-se ainda se o cruzamento será polarizado ou não polarizado, sendo atribuída uma probabilidade de 30% para a ocorrência do cruzamento polarizado. No caso polarizado, para um dos novos indivíduos (filhos) define-se o parâmetro

$$\mu = 1,4\beta_1\beta_2 - 0,2$$

onde β_1 e β_2 são variáveis aleatórias independentes, com distribuição uniforme no intervalo $[0, 1]$. A utilização do produto $\beta_1\beta_2$, dada a sua forma aqui posta, além de raramente assumir valores próximos de 0 e 1, induz uma maior concentração de valores de μ em regiões específicas do intervalo, aumentando a probabilidade de que o filho gerado esteja mais próximo do pai de melhor aptidão (melhor pai), previamente identificado com base na função de avaliação. O segundo indivíduo (filho) gerado a partir do mesmo par é sempre obtido sem polarização, de modo a preservar a diversidade genética na população.

Caso o cruzamento não seja polarizado, o parâmetro μ é amostrado segundo uma distribuição uniforme no intervalo $(-0.1, 1.1)$, resultando em indivíduos x_g distribuídos uniformemente no intervalo de valores possíveis definido pelos pais.



Figura 5 – Representação unidimensional do operador de cruzamento real polarizado.

Essa operação de cruzamento polarizado proposta por (TAKAHASHI, 2007) imita uma busca seguindo uma direção (ao longo de um segmento de reta) de uma tendência (com uma informação semelhante à fornecida pelo gradiente). No caso de pais localmente próximos um do outro, é realizado um passo semelhante ao do algoritmo gradiente ($dx \approx 0$) e isso acelera a convergência local para o ótimo. No caso de indivíduos que se encontram distantes (possivelmente em bacias de atração diferentes), a operação pode ser interpretada como a utilização de uma informação de tendência de longa distância.

Devido a extrapolação que envolve o processo de cruzamento, os indivíduos gerados podem estar situados fora das faixas admissíveis $[x_L, x_U]$ de parâmetros. Se isso ocorrer, uma operação de reflexão é ainda realizada para corrigir essa situação, trazendo o indivíduo de volta para o interior da região admissível. Essa operação é definida como:

$$x_r = x_L + |x - x_L|$$

Para a reflexão no limite inferior, sendo x o indivíduo que violava a restrição, x_r o novo indivíduo gerado pela reflexão e x_L o vetor de limites inferiores.

Para a reflexão no limite superior, a operação é definida por:

$$x_r = x_U - |x - x_U|$$

Para x_U o vetor de limites superiores e as demais variáveis com mesmo significado que anteriormente.



(a) Representação unidimensional do processo de reflexão quando ocorrido à esquerda do limite inferior x_L .

(b) Representação unidimensional do processo de reflexão quando ocorrido à direita do limite superior x_U .

Figura 6 – Representação unidimensional da operação de reflexão quando $x \notin [x_L, x_U]$.

Nas execuções do AGRP, consideramos para os experimentos do Capítulo 4, o tamanho da população $N = 100$, o número de gerações $G = 100$ e o número de execuções $E = 30$. A probabilidade de mutação considerada foi de 5%, a probabilidade de polarização de 30%. As outras etapas, como Função de avaliação (aptidão) foi considerada a própria função objetivo escolhendo os indivíduos mais aptos (com menores valores), para a etapa de Seleção foi utilizando

o método de torneio e inserimos também uma função de Elitismo e o critério de parada como sendo o número máximo de gerações escolhido a priori. De forma mais precisa a implementação em Python foi baseada no que foi descrito em (TAKAHASHI, 2007) juntamente com a biblioteca DEAP (Distributed Evolutionary Algorithms in Python) (FORTIN et al., 2012)

Ainda sobre a etapa de Mutação para cada indivíduo é avaliado com probabilidade de 5% se este sofrerá mutação ou não. Caso vá ocorrer a mutação é somado (ou subtraído) a cada gene do indivíduo(vetor) x uma perturbação δ_i cuja definição é dada por

$$\delta_i = 0,05\beta_i(x_R)_i$$

Com β_i sendo um número aleatório com distribuição gaussiana, média zero e variância 1 e x_R sendo o vetor de diferença entre os máximos e mínimos dos parâmetros:

$$x_R = (x_U - x_L)$$

Considerando agora as etapas Avaliação, Função de Ajuste, Seleção e Elitismo temos:

Avaliação: Cada indivíduo é avaliado na função objetivo.

Função de Ajuste: Ao utilizar o DEAP inserir a função de ajuste se torna desnecessária visto que a aptidão já está considerando o problema de minimização da função objetivo.

Seleção: Dentre os vários tipos de operações de seleção o método escolhido foi por torneio, visto que o nosso problema é mono-objetivo (minimizar a função objetivo) e pela eficiência computacional deste método.

Elitização: A implementação da elitização foi da seguinte forma: caso o melhor individuo não tenha sido selecionado na etapa de seleção para compor uma nova população, ele é introduzido para a próxima geração com exclusão do pior elemento daquela geração que ele pertence.

3.3 Função de Verossimilhança, Critério de Informação Akaike (AIC) e Critério de Informação Akaike Corrigido (AICc)

Um dos métodos de seleção e comparação entre modelos candidatos a serem utilizados na modelagem de um sistema é o Critério de Akaike (Akaike Information Criterion - AIC) que se trata de uma forma de avaliar a adequação destes, levando em consideração o equilíbrio entre a quantidade de dados disponíveis e o número de parâmetros. Sua descrição envolve a Função de Verossimilhança (Likelihood Function) (MYUNG, 2003), que é uma forma de medir, quão razoável é obter os dados observados por meio de um modelo.

Suponhamos que em um determinado modelo, tenhamos um conjunto de n dados observados $X = (x_1, x_2, \dots, x_n)$ e um conjunto de k parâmetros $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ de um modelo a serem obtidos. Ou seja, uma função do tipo $f(X, \theta)$ e após esta obtenção um outro conjunto de dados $Y = (y_1, y_2, \dots, y_n)$ simulados por meio deste modelo que, em seu teor, utiliza-se a minimização dos resíduos (erros) $\epsilon_i = y_i - f(x_i, \theta)$ entre os dados obtidos e

os dados observados seja obtido. Nestas condições a função de verossimilhança é dada por $L(\theta, X) = \prod_{i=1}^n P(X|\theta)$ onde $P(X|\theta)$ é a probabilidade (ou densidade de probabilidade) dos resíduos entre estes dois conjuntos de dados.

Se considerarmos que o erro ϵ seja distribuído normalmente com média zero e variância σ^2 constante, o que é razoável supor para problemas de minimizações de resíduos, a probabilidade de observar um ponto (x_i, y_i) é dada pela função de densidade $P(y_i|x_i, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i - f(x_i, \theta))^2}{2\sigma^2}}$. Assim a função de verossimilhança poderá ser definida da seguinte forma:

$$L(\theta, X) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i - f(x_i, \theta))^2}{2\sigma^2}}$$

Uma transformação em L é considerada, obtendo assim a função Log-Verossimilhança dada por:

$$\ln(L) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \sum_{i=1}^n \frac{(y_i - f(x_i, \theta))^2}{2\sigma^2}$$

Como dito anteriormente presume-se para esta construção que a distribuição dos resíduos seja normal, mas o que normalmente ocorre é que neste tipo de modelagem envolvendo relativa variabilidade, os resíduos apresentam heterocedasticidade (variância não constante) o que reflete que os resíduos com variância maior tendem a influenciar excessivamente o ajuste do modelo enquanto resíduos com variância menor apresentam menos impacto. Desta forma, optamos por utilizar, para a obtenção da função de verossimilhança, o método *Weighted Least Squares (WLS)*, que atribui pesos $w_i = \frac{1}{\text{Var}(\epsilon_i)} = \frac{1}{\sigma_i^2}$ inversamente proporcionais à variância dos resíduos (ϵ_i). Esta alternativa é adequada pois reflete que resíduos com alta variância recebam menos pesos e resíduos com baixa variância recebam mais pesos. Em outras palavras, o método WLS é uma abordagem adequada quando os erros não são uniformes nem homocedásticos, pois corrige o efeito da variância desigual e permite que o modelo reflita corretamente a informação contida nos dados observados.

Desta forma a função a ser usada considerando estes pesos é:

$$L_w(\theta, X) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{w_i(y_i - f(x_i, \theta))^2}{2\sigma_i^2}} \quad (16)$$

E consequentemente,

$$\ln(L_w) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma_i^2) - \sum_{i=1}^n \frac{w_i(y_i - f(x_i, \theta))^2}{2\sigma_i^2} \quad (17)$$

O AIC (Critério de Informação Akaike) é definido por

$$AIC = 2k - 2\ln(L) \quad (18)$$

Ou o equivalente para o caso utilizando WLS

$$AIC = 2k - 2\ln(L_w) \quad (19)$$

Essa fórmula reflete a diferença entre o número de parâmetros k do modelo e a Log-verossimilhança. Quanto maior o número de parâmetros maior o AIC e maior a penalização no critério. Quanto maior a soma dos quadrados dos resíduos, menor a Log-verossimilhança e maior o segundo termo, maior o AIC e maior a penalização. Portanto, o melhor modelo é aquele que controlar o menor número possível de parâmetros com a menor soma dos quadrados dos resíduos, formando o menor valor de AIC.

Quando a fração da quantidade de dados n pelo número de parâmetros k é menor que 40, ou seja, $\frac{n}{k} < 40$ é aconselhável utilizar o Critério de Informação de Akaike Corrigido - AICc (BURNHAM; ANDERSON, 2004) que trata-se de uma versão ajustada para os casos em que o tamanho da amostra é pequeno em relação ao número de parâmetros. Sua fórmula tende a penalizar um tanto mais a complexidade do modelo. Nos casos onde o tamanho da amostra é bem maior que o número de parâmetros, a diferença entre estes se torna desprezível. Por definição temos que:

$$AICc = AIC + \frac{2k^2 + 2k}{n - k - 1} \quad (20)$$

Os valores obtidos pelo critério de Akaike não são usados de forma absoluta, mas de forma relativa em comparações entre os modelos. Os melhores, do ponto de vista deste critério, são os que possuem os menores AIC nas comparações, ou no caso de tamanho de amostras pequeno, os menores AICc, como dito anteriormente.

3.4 Matriz da Próxima Geração

O método da matriz da próxima geração para obtenção do número básico de reprodução baseia-se no fato que em um modelo epidemiológico existem duas categorias populacionais, uma categoria formada pelos compartimentos contendo populações de indivíduos infectados (infecciosos ou infectantes) e uma outra categoria formada pelos outros tipos de compartimentos populacionais: suscetíveis e recuperados, no caso do modelo SIR (DRIESSCHE, 2017). Podemos defini-las da seguinte forma:

- i) $x = \{x_1, x_2, \dots, x_n\} :=$ Categoria relacionada aos compartimentos contendo indivíduos infectados(infecciosos ou infectantes); Neste caso aqui representados por n compartimentos;

- ii) $y = \{y_1, y_2, \dots, y_m\} :=$ Categoria relacionada aos outros compartimentos; Neste caso aqui representados m compartimentos.

Dessa forma estes modelos podem ser descritos como

$$\begin{cases} \frac{dx_i}{dt} = F_i(x,y) - V_i(x,y) & i = 1,2,3,\dots,n \\ \frac{dy_j}{dt} = G_j(x,y) & j = 1,2,3, \dots, m \end{cases} \quad (21)$$

Onde $F_i(x,y)$ representa a parte das taxas que envolvem a produção de novos infectados no compartimento i , $V_i(x,y)$ representa a diferença entre a taxa de transferência de indivíduos que saem do compartimento i , denotada por $V_i^-(x,y)$ e a taxa de transferência dos indivíduos que entram no compartimento i , denotada por $V_i^+(x,y)$, ou seja, $V_i(x,y) = V_i^-(x,y) - V_i^+(x,y)$ e $G_j(x,y)$ representam os outros compartimentos.

Considerando isto, o método consiste em determinar $\mathcal{R}_0$ (Número Básico de Reprodução) por meio da análise dos autovalores da matriz FV^{-1} onde $F = \frac{\partial(F_1, F_2, \dots, F_n)}{\partial(x_1, x_2, \dots, x_n)}$ e $V = \frac{\partial(V_1, V_2, \dots, V_n)}{\partial(x_1, x_2, \dots, x_n)}$, ou seja, por meio destes jacobianos. Assim, o método estabelece a identificação do máximo dos valores absolutos dos autovalores da matriz FV^{-1} , isto é, do raio espectral da matriz FV^{-1} como o número básico de reprodução $\mathcal{R}_0$.

Por exemplo, considerando o modelo SIR com dinâmica demográfica por meio da taxa de nascimentos e mortes igual a μ dado abaixo:

$$\begin{cases} \frac{dS}{dt} = \mu N - \beta \frac{SI}{N} - \mu S \\ \frac{dI}{dt} = \beta \frac{SI}{N} - (\gamma + \mu)I \\ \frac{dR}{dt} = \gamma I - \mu R \end{cases} \quad (22)$$

Desta forma temos: $x = (I)$, $y = (S, R)$, $F_1(x,y) = F_1(I, S, R) = \beta \frac{SI}{N}$, $V_1^-(x,y) = V_1^-(I, S, R) = (\gamma + \mu)I$, $V_1^+(x,y) = V_1^+(I, S, R) = 0$, $V_1(x,y) = V_1(I, S, R) = (\gamma + \mu)I$, $G_1(x,y) = G_1(I, S, R) = \mu N - \beta \frac{SI}{N} - \mu S$, $G_2(x,y) = G_2(I, S, R) = \gamma I - \mu R$.

Desta forma temos que: $F = \left[\frac{\partial F_1}{\partial I} \right] = \left[\beta \frac{S}{N} \right]$, $V = \left[\frac{\partial V_1}{\partial I} \right] = [\gamma + \mu]$. Considerando $S \approx N$ (condição de homogeneidade, toda população suscetível), $FV^{-1} = \left[\frac{\beta}{\gamma + \mu} \right]$ o que acarreta no autovalor desta matriz igual a $\frac{\beta}{\gamma + \mu}$ que é não negativo e que corresponde ao $\mathcal{R}_0$ deste modelo. Observe que através da escolha de $\mu = 0$ temos o modelo SIR básico 3 e consequentemente $\mathcal{R}_0 = \frac{\beta}{\gamma}$.

Os materiais, técnicas e procedimentos metodológicos descritos neste capítulo fundamentam o desenvolvimento das aplicações apresentadas nos capítulos seguintes. A partir dessas

definições e destes métodos, tornou-se possível a implementação e a análise das aplicações propostas, conforme serão detalhadas ao longo deste trabalho.

Capítulo 4

Modelagem do engajamento na Infodemia

Com base nos materiais e métodos apresentados no capítulo anterior, neste capítulo e no próximo, temos por objetivo apresentar as aplicações desenvolvidas neste trabalho, por meio da implementação prática da metodologia proposta e sua utilização no contexto de modelos que tratam-se do engajamento social em torno da “Infodemia” e também no contexto de Polarização Social e Política.

Efetuamos aqui e assim, um estudo do caso do grau de engajamento ao tema “vacinação infantil”, utilizando para isso a coleta de dados do Google Trends com um filtro geográfico definido para o Brasil, no período de 01/12/2021 a 27/02/2022. Foram obtidos, no dia 02/03/2022, o gráfico de séries temporais da frequência das pesquisas diárias sobre este tema e a planilha com os dados correspondentes (Figura 3). A análise inicial dos dados revelou variações inesperadas e picos (certa complexidade), sem identificar inicialmente um processo epidêmico típico (curva I do modelo SIR), mas observou-se nos dados uma série de eventos com certa complexidade caracterizados pela sobreposição de várias ondas ou picos pulsantes (Figura 4).

A observação dos dados mostrou períodos com intervalos de rápido engajamento, seguidos de intervalos de desengajamento. Não há, portanto, uma curva típica autocontrolada no sentido da referência com a população de infectados do modelo SIR (forma de sino). De fato, como o sistema é aberto, é esperado que haja a influência de diversos atores como influencers, meios de comunicação e campanhas de informação. Assim, é esperado que os parâmetros do modelo SIR, β , γ e N variem no tempo.

Esse efeito de variação dos parâmetros de modelos epidemiológicos ao longo do tempo é esperado em epidemias reais, uma vez que mudanças ambientais e comportamentais são comumente empregadas sempre que possível, tais como: ações de controle, distanciamento social, medidas sanitárias e vacinas.

A partir deste ponto de vista, a série temporal de dados de busca, associados ao compartimento de infectados do modelo foi dividida em intervalos cujas sequências lógicas estão relacionadas a processos de engajamento (aumento relativo) e desengajamento (diminuição relativa) em cada período (ver Figura 4).

4.1 Modelo tipo SIR - Clássico

Nesta seção discutiremos e apresentaremos os resultados referentes ao modelo básico SIR sem dinâmica vital considerando que este foi o primeiro modelo (Eq. 2) que utilizamos nesta

abordagem:

$$\begin{cases} \frac{dS}{dt} = -\beta S \frac{I}{N} \\ \frac{dI}{dt} = \beta S \frac{I}{N} - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases} \quad (23)$$

4.1.1 Órbita e retrato de fase

Uma forma de compreendermos a dinâmica populacional de um modelo dinâmico e consequentemente do modelo SIR é através da construção de algumas de suas órbitas e do retrato de fase. Considerando o modelo SIR normalizado (Modelo 3), procuremos inicialmente os pontos de equilíbrio do sistema. Como temos $s + i + r = 1$, teremos $r = 1 - s - i$, o que implica que podemos utilizar as primeiras duas equações populacionais para fazermos este estudo, já que elas não possuem em sua composição a população r . O espaço de fase deste sistema, considerando esta simplificação será o conjunto $D = \{(s, i) : s \geq 0, i \geq 0 \text{ e } s + i \leq 1\}$.

As equações no equilíbrio a serem consideradas são:

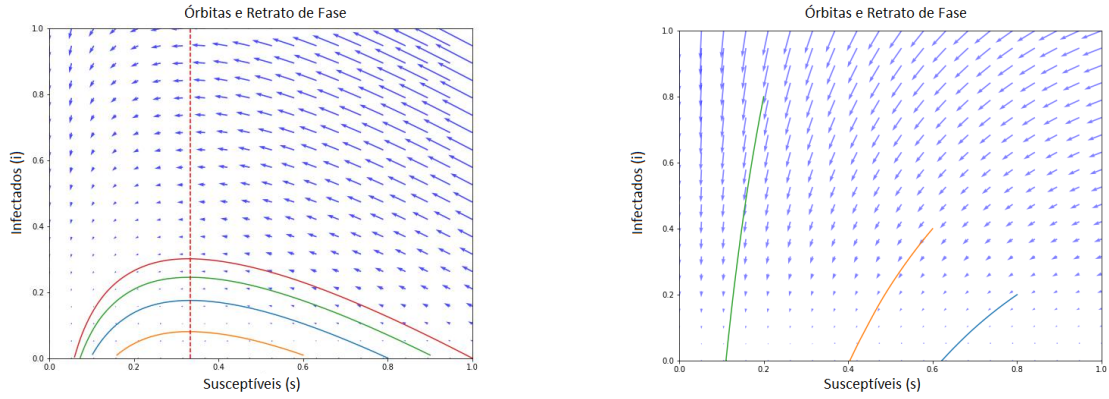
$$\begin{cases} \frac{ds}{dt} = -\beta si = 0 \\ \frac{di}{dt} = \beta si - \gamma i = 0. \end{cases}$$

Desta forma todo ponto $(s, 0)$ será em termos matemáticos um ponto de equilíbrio, já que anulam ambas equações diferenciais.

A equação $\frac{ds}{dt} = -\beta si$ para $i \neq 0$ sobre D , será sempre negativa o que remete no espaço $s \times i$ que a projeção do campo de vetores em um ponto de D apontará para a direção decrescente de s , mais especificamente para a esquerda. Já considerando a segunda equação $\frac{di}{dt} = \beta si - \gamma i = (\beta s - \gamma)i$ temos duas possibilidades para a reta nuliclinal $s = \frac{\gamma}{\beta} = \frac{1}{\mathcal{R}_0}$ excetuando o caso $(1, 0)$: o primeiro de interceptar a região D que são $\frac{\gamma}{\beta} < 1$ e o segundo de não interceptar a região D , ou seja, $\frac{\gamma}{\beta} > 1$. Considerando o primeiro caso $\frac{\gamma}{\beta} < 1$, ou seja, quando $\mathcal{R}_0 > 1$ chegamos as condições derivadas da segunda equação: para $s^* < \frac{\gamma}{\beta}$, $\frac{di}{dt} < 0$ o que leva a doença a extinguir-se e $s^* > \frac{\gamma}{\beta}$ ou seja $\frac{di}{dt} > 0$, o que leva a evolução da doença. No primeiro caso, os vetores do campo apontam para a esquerda e para baixo e no segundo caso para a esquerda e para cima (Fig. 7(a)).

Por outro lado, o caso onde $\frac{\gamma}{\beta} > 1$ temos que a reta (nuliclinal) $s = \frac{\gamma}{\beta}$ não intercepta a região D , o que implica em $\mathcal{R}_0 < 1$ e assim $s^* < \frac{\gamma}{\beta}$ o que implica como dito anteriormente $\frac{di}{dt} < 0$ ou seja leva a extinção da doença.

Temos, em termos epidemiológicos, dois casos para $\mathcal{R}_0 > 1$, representados na Figura 7(a) considerando uma perturbação $i^* > 0$ para o ponto de equilíbrio $(s^*, 0)$:



(a) Retrato de fase e órbitas, caso $\frac{\gamma}{\beta} < 1$, ou seja, $\mathcal{R}_0 > 1$.

(b) Retrato de fase e órbitas, caso $\frac{\gamma}{\beta} > 1$, ou seja, $\mathcal{R}_0 < 1$.

Figura 7 – Retrato de Fase Modelo SIR sem dinâmica vital Eq. (3)

- Sendo $s^* \leq \frac{\gamma}{\beta}$ (à esquerda da reta nuliclinial) o número de infectados decrescerá até chegar a zero;
- Sendo $s^* > \frac{\gamma}{\beta}$ (à direita da reta nuliclinial) teremos uma epidemia, ou seja, o número de infectados crescerá até atingir a um máximo em I , quando $s(t) = \frac{\gamma}{\beta}$, então a população i decrescerá até chegar a zero.

No caso onde $\mathcal{R}_0 < 1$, Figura 7(b), dadas as condições iniciais (s^*, i^*) , a tendência é o número de infectados $I(t)$ diminuir até chegar a zero.

E no caso onde $\mathcal{R}_0 = 1$, temos que o número de infecções secundárias produzidas por um único indivíduo infectado é igual a 1 o que indica que cada indivíduo infectado transmite a doença em média para um individuo apenas o que reflete que a população de infectados se mantém constante, não cresce e nem desaparece rapidamente, ou seja, a doença permanece na população de forma endêmica.

Podemos também observar para o caso $\mathcal{R}_0 > 1$ e através da figura 7(a) que para cada ponto de equilíbrio da forma (s^*, i^*) com $\frac{\gamma}{\beta} < s^* \leq 1$ temos um ponto de equilíbrio $(\tilde{s}, \tilde{i})$ com $0 \leq \tilde{s} < \frac{\gamma}{\beta}$. Este fato reforça a dinâmica referente a população dos infectados.

Considere agora o ponto de equilíbrio com a população de indivíduos suscetíveis sendo toda a população, ou seja, $s^* = 1$, $i^* = 0$ e $r^* = 0$. Por meio da análise da órbita deste caso temos que este ponto é instável se $\frac{\gamma}{\beta} < 1$ (quando $\mathcal{R}_0 > 1$) e é estável para os casos $\frac{\gamma}{\beta} > 1$ (quando $\mathcal{R}_0 < 1$).

4.1.2 Resultados Modelo SIR clássico

Para o modelo SIR clássico (23), os valores obtidos por meio da otimização usando o método *Trust Region Reflective* (Seção 3.2) estão descritos na Tabela 2 e as curvas da população infectada se ajustaram aos dados, com uma correlação forte, no sentido de uma relação linear

intensa, como podemos detectar em (Figura 8) . Por fim, a Tabela 2 apresenta também os valores de $\mathcal{R}_0$ em cada estágio de engajamento e desengajamento de cada subintervalo. Em cada período, o comportamento da curva é definido como resultado de duas tendências: de engajamento e de desengajamento. Estas tendências são caracterizadas pelas respectivas taxas de infecção β e recuperação da infecção γ , assim como o número N de indivíduos. O número de indivíduos N corresponde ao número de pessoas capazes de se interessar pelo tema, S é a parte da população que tomou contato ou está consciente de um assunto, porém não vivenciou ainda uma dúvida a respeito estando portanto, suscetível a “sofrer o contágio” ou seja, a efetuar buscas, I é a parte da população que sofreu o contágio pela dúvida e que efetivamente realizou pesquisa no Google pelo tema e R é a parte da população que, após realizar a pesquisa no Google, se sentiu suficientemente informado e superou suas dúvidas.

Tabela 2 – Valores dos parâmetros N , β e γ otimizados via método *Trust Region Reflective* (sem limitante superior) e valor aproximado de $\mathcal{R}_0$ em cada período p_i e etapas do processo referente ao modelo (23) ajustado aos dados do Google Trends.

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
N	4.3×10^1	4.463×10^3	2.9×10^1	1.9×10^1	2.2×10^1	1.1×10^1	2.02×10^2	3.67×10^4
β	2.311	9.787×10^{-4}	7.987×10^{-1}	0.8819	0.6429	0.6537	3.399×10^{-1}	1.148×10^{-5}
γ	0.4521	8.2945×10^{-1}	3.7566×10^{-26}	0.7359	0.145	0.659	4.5478×10^{-24}	7.6987×10^{-2}
$\mathcal{R}_0$	5.113	0.0012	2.126×10^{25}	1.198	4.435	0.992	7.474×10^{22}	0.0001

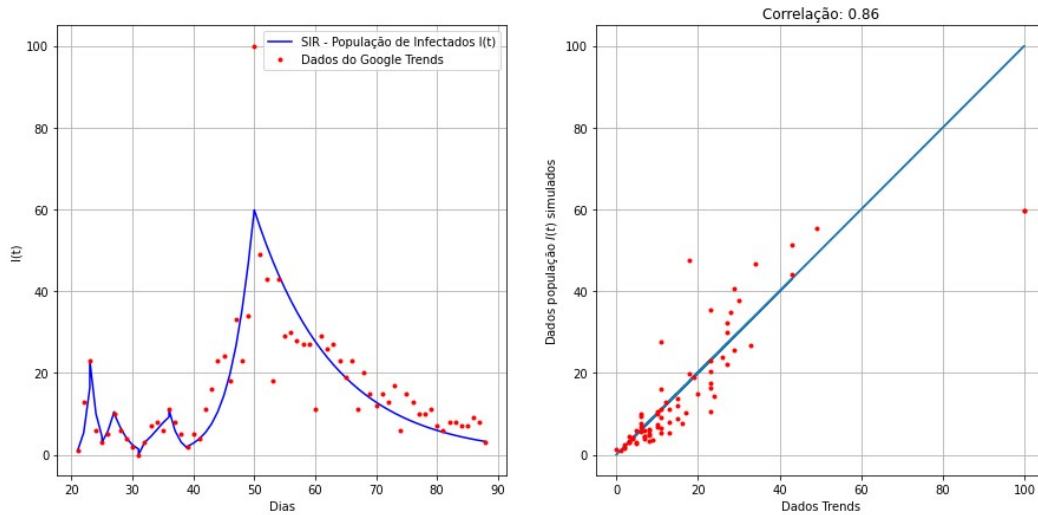


Figura 8 – (a) Dados de frequência de buscas relativas realizadas no Google no período analisado e curva de ajuste gerada considerando os parâmetros otimizados do modelo SIR; (b) Gráficos dados GT $\times$ dados GT (bissetriz), dados GT $\times$ dados simulados e correlação de Pearson. GT = Google Trends.

A possibilidade de surgimento de valores incomuns de $\mathcal{R}_0$ para epidemias biológicas (períodos p_3 e p_7 - Tabela 2) deu-se pela escolha intuitiva do intervalo de busca dos parâmetros ótimos, que neste caso inicial foi imposto apenas ser limitado inferiormente por 0, logo a

descrição dos valores dos parâmetros obtidos deve ser vista como uma forma mais ilustrativa ($\mathcal{R}_0 \gg 1$). Este fato, foi um indicador da necessidade de estendermos a busca para outras técnicas e abordagens. Opta-se assim, pelo refinamento do método de obtenção dos parâmetros via construção de um Algoritmo Genético Real Polarizado - AGRP (Seção 3) baseado nos trabalhos (TAKAHASHI, 2007), (SILVA et al., 2019), (CASTRO et al., 2004), (POZO et al., 2005) e (FORTIN et al., 2012).

Desta forma, obtemos por meio do AGRP os parâmetros descritos na Tabela 3, bem como o gráfico de ajuste e o Correlograma desta simulação (Figura 9).

Tabela 3 – Valores dos parâmetros N , β e γ otimizados e valor aproximado de $\mathcal{R}_0$ em cada período p_i e etapas do processo referente ao modelo 23 ajustado aos dados do Google Trends por meio do AGRP

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
β	0.9997	0.0110	0.2973	0.5169	0.5712	0.7472	0.3003	0.0039
$N(0)$	299.9921	253.2885	158.6266	243.8323	103.1741	239.6413	175.4934	276.5948
γ	0.0501	0.6365	0.0621	0.9999	0.9963	0.9996	0.0502	0.0913
$\mathcal{R}_0$	19.9464	0.0173	4.7878	0.5169	0.5734	0.7475	5.9822	0.0428
Min(FO)	2.3533	2.8368	0.2096	0.1773	0.7075	0.2169	3.0135	4.7201

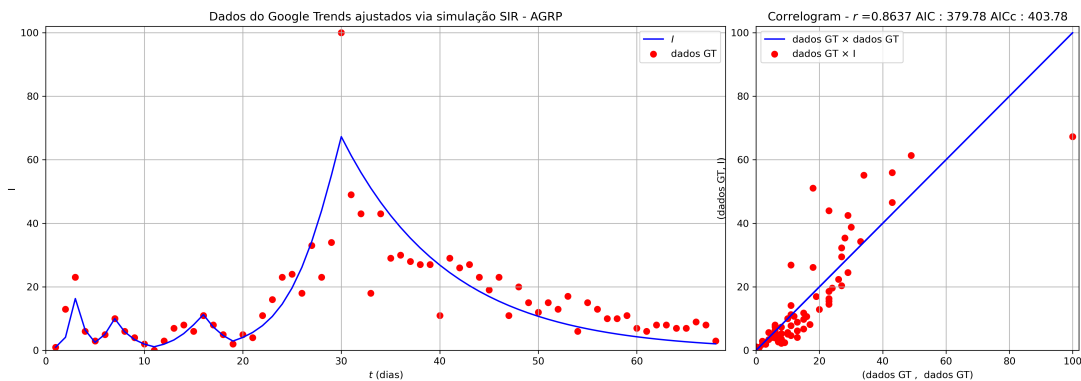


Figura 9 – (a) Dados de frequência de buscas relativas realizadas no Google no período analisado e curva de ajuste gerada considerando os parâmetros otimizados do modelo SIR; (b) Gráficos dados GT $\times$ dados GT (bissetriz), dados GT $\times$ dados simulados via AGRP e correlação de Pearson.

Em comparação com os valores de $\mathcal{R}_0$ obtidos na tabela anterior (Tabela 2) estes valores (Tabela 3) são mais epidemiologicamente comuns, visto que os espaços de buscas foram, mesmo que intuitivamente, limitados inferiormente e superiormente.

4.2 Para além do modelo SIR - Modelo tipo SIR - Aberto com interação IR

Como o objetivo é de, por meio do modelo SIR clássico na Seção 2.3, inserir novos termos baseando-nos que no processo infodêmico juntamente com o contágio social e considerando a limitação natural e dimensional aqui posta, considera-se que processos como estes carregam intrinsecamente outros desdobramentos e fenômenos. Observando o fenômeno de engajamento, nota-se que as populações que tomam conhecimento de um assunto e que efetivamente fazem buscas é variável. Assim, podemos considerar que o sistema é aberto. Isto é, as pessoas podem se desinteressar e entrar e sair de algumas subpopulações do modelo. Além disso, no caso do fenômeno informacional as transições mais gerais entre compartimentos do modelo podem ser mais gerais que os termos de infecção do modelo SIR. Desta forma, incorporamos e estudamos outros efeitos no modelo a fim de buscar resultados, mais realistas. A obtenção de parâmetros mais realistas, entretanto, fica sujeita à relação entre número de parâmetros e quantidade de dados. Outra limitação reside no fato de lidarmos apenas com a população de infectados e otimizarmos as condições iniciais das outras populações.

Um modelo com outras dinâmicas e consequentemente um pouco mais completo (Modelo 24) baseado na literatura, pode ser escrito da seguinte forma:

$$\begin{cases} \frac{dS}{dt} = -\beta \frac{SI}{N} - \mu_1 S + \Lambda \\ \frac{dI}{dt} = \beta \frac{SI}{N} - \frac{\gamma IR}{N} - \xi I - \mu_2 I \\ \frac{dR}{dt} = \frac{\gamma IR}{N} + \xi I \end{cases} \quad (24)$$

Para este modelo 24 temos sua representação dada pelo diagrama apresentado na Figura 10. As descrições para os parâmetros aplicadas no contexto de contágio social são:

- β : Taxa que indivíduos suscetíveis são infectados através da interação com indivíduos infecciosos, em síntese refere-se ao coeficiente de transmissão do contágio pela polêmica ou assunto potencialmente ressoante (t^{-1});
- μ_1 : Análogo à taxa de mortalidade. Representa a fração de indivíduos suscetíveis que não sofrem mais influências externas ou que perdem interesse ou que saem do sistema (t^{-1});
- μ_2 : Também análogo à taxa de mortalidade. Representa a fração de indivíduos infectados que não sofrem mais influências externas ou que perdem interesse ou que saem do sistema (t^{-1});
- Λ : Taxa de entrada constante responsável pelo aumento na população de suscetíveis. A mesma indica uma intensificação ocorrendo na população S , uma fonte exógena fixa representando um processo externo (t^{-1});

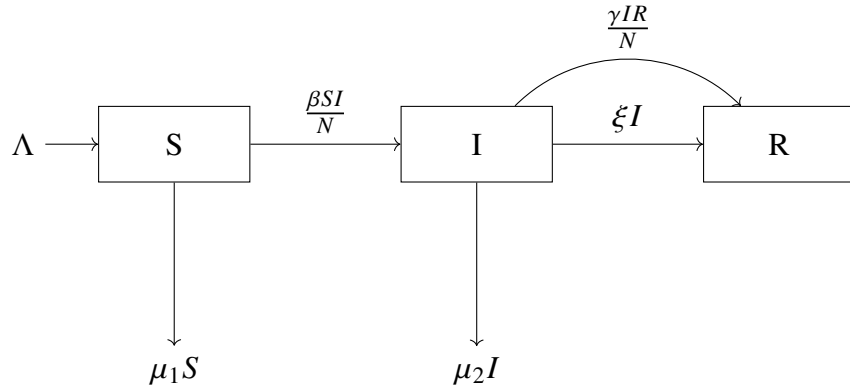


Figura 10 – Diagrama do modelo 24

- γ : Taxa por unidade de indivíduo que representa o efeito dos infectados que se tornam recuperados por influência da população recuperada. Pode-se pensar em uma taxa de ajuda por influência derivada dos próprios indivíduos recuperados (t^{-1});
- ξ : Taxa de recuperação dos indivíduos infectados (resolução da dúvida) (t^{-1}).

Ressalta-se que o modelo 24 normalizado por N , teria pelo fato de N não ser constante, uma forma mais complexa e que não agregaria no escopo deste trabalho, nenhum resultado diferente. Apresentamos apenas para título de abordagem extra o sistema normalizado por N não constante, fazendo $s = \frac{S}{N}$, $i = \frac{I}{N}$ e $r = \frac{R}{N}$:

$$\begin{cases} \frac{ds}{dt} = (1-s) \frac{\Lambda}{N} + (\mu_2 - \beta) si + \mu_1 s^2 - \mu_1 s \\ \frac{di}{dt} = (\beta + \mu_1) si - \gamma ir - \left(\frac{\Lambda}{N} + \xi + \mu_2 \right) i + \mu_2 i^2 \\ \frac{dr}{dt} = (\gamma + \mu_2) ir + \xi i + \mu_1 rs - \frac{\Lambda}{N} r \end{cases} \quad (25)$$

4.2.1 Pontos de Equilíbrio - Modelo 24

Inicialmente buscamos impor a condição $S(t) + I(t) + R(t) = N$, com N constante, ou seja, que a população total permanecesse constante em cada período e que a condição de homogeneidade natural e necessária destes modelos, fosse também levada em consideração. Impondo que $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$ em (24), implicaria na condição adicional que $\Lambda = \mu_1 S + \mu_2 I$, embora isso seja 'justificável', do ponto de vista dinâmico referente a modelos desta natureza, quando as taxas de mortalidade são iguais a taxa de natalidade por exemplo é simplificada pela consideração que a fração da população que entra no sistema é igual a fração da população que sai, tal argumentação aqui, implicaria em um mal comportamento, visto que no espaço de soluções de Λ teríamos uma função bidimensional, o que geraria uma incompatibilidade matemática, já que Λ é uma constante. Sendo assim descarta-se essa condição e tentativa de imposição.

Considerado o caso não normalizado e o fato que $S \geq 0$, $I \geq 0$ e $R \geq 0$, um único ponto de equilíbrio¹ que é o caso livre de doença, foi obtido, à saber, $P_1 = \left(\frac{\Lambda}{\mu_1}, 0, 0\right)$. Observe que escolhemos $R = 0$ por se tratar do caso livre de doença ou seja, $I = 0$ e que remeteria não termos um $R \neq 0$. Por outro lado, na busca de um ponto de equilíbrio endêmico ($I \neq 0$) teremos uma população $R = -\frac{\xi N}{\gamma}$ negativa, derivada da 3ª equação referente ao compartimento populacional R o que nos leva descartar por razões óbvias de incompatibilidade com as descrições populacionais da literatura, assim, o sistema com as condições aqui colocadas possuirá um único ponto factível de equilíbrio $P_1 = \left(\frac{\Lambda}{\mu_1}, 0, 0\right)$. Entretanto, algebricamente é possível obter de forma explícita e em função de N , o ponto de equilíbrio P_2 , sua descrição é $P_2 = \left(\frac{\mu_2 N}{\beta}, \frac{\Lambda}{\mu_2} - \frac{\mu_1 N}{\beta}, -\frac{\xi N}{\gamma}\right)$. Mas como não é um ponto de equilíbrio representativo dentro da dinâmica populacional aqui abordado sua análise não é necessária, visto que $R = -\frac{\xi N}{\gamma} < 0$.

4.2.2 Estudo da estabilidade - Modelo 24

Para P_1 , temos que sua matriz Jacobiana será dada por:

$$J(P_1) = \begin{bmatrix} -\mu_1 & -\frac{\beta\Lambda}{N_0\mu_1} & 0 \\ 0 & \frac{\beta\Lambda}{N_0\mu_1} - (\xi + \mu_2) & 0 \\ 0 & \xi & 0 \end{bmatrix} = \begin{bmatrix} -\mu_1 & -(\mu_2 + \xi) R_e(0) & 0 \\ 0 & (\mu_2 + \xi) (R_e(0) - 1) & 0 \\ 0 & \xi & 0 \end{bmatrix} \quad (26)$$

O número de reprodução efetivo $R_e(t)$ que representa o número médio de novos infectados gerados na situação atual da população (no instante t), isto é, considerando a parte já exposta, imune e removida é dada por $R_e(t) = \mathcal{R}_0 \cdot \frac{S(t)}{N(t)}$. Este parâmetro é tão importante quanto o $\mathcal{R}_0$ visto que ele fornece informações instantâneas e a medida que o número de indivíduos suscetíveis diminui. Desta forma o polinômio característico deste sistema em P_1 será dado por

$$p(\lambda) = -\lambda^3 + ((\xi + \mu_2) (R_e(0) - 1) - \mu_1) \lambda^2 + \mu_1 (\mu_2 + \xi) (R_e(0) - 1) \lambda \quad (27)$$

Aplicando o critério de Routh-Hurwitz sobre o polinômio $\tilde{P}(\lambda) = -p(\lambda)$:

λ^3	1	$\mu_1 (\xi + \mu_2) (1 - R_e(0))$
λ^2	$\mu_1 - (\xi + \mu_2) (R_e(0) - 1)$	0
λ^1	$\mu_1 (\xi + \mu_2) (1 - R_e(0))$	0
λ^0	$\mu_1 (\xi + \mu_2) (1 - R_e(0))$	

Tabela 4 – Tabela de Routh-Hurwitz referente ao ponto de equilíbrio P_1 e polinômio $\tilde{P}(\lambda) = -p(\lambda)$ - Equação 27

¹ Ressalta-se que para essa afirmação este fato está sendo considerado, já que, $\forall R^* \geq 0$ os pontos $P^* = \left(\frac{\Lambda}{\mu_1}, 0, R^*\right)$ anulariam o sistema.

Observação: Foi utilizada na Tabela 4 na última linha o polinômio auxiliar visto que a última linha da tabela de Routh resultou-se numa linha nula (caso especial).

Se $1 - R_e(0) > 0$ temos que $\mu_1 - (\xi + \mu_2)(R_e(0) - 1) > 0$ logo o sistema no ponto P_1 é estável e se $1 - R_e(0) < 0$ temos que $\mu_1(\xi + \mu_2)(1 - R_e(0)) < 0$ o que implica que o sistema em P_1 neste caso é instável, em outras palavras temos a seguinte conclusão sobre o ponto de equilíbrio P_1 :

- Se $R_e(0) < 1$ o sistema é estável;
- Se $R_e(0) > 1$ o sistema é instável;
- Se $R_e(0) = 1$ temos uma bifurcação transcritical, ou seja, onde há troca de estabilidade.

Vale a pena ressaltar que no instante $t = 0$ temos $R_e(0) = \mathcal{R}_0$. O que implica que as condições anteriores podem ser reescritas da seguinte forma:

- Se $\mathcal{R}_0 < 1$ o sistema é estável;
- Se $\mathcal{R}_0 > 1$ o sistema é instável;
- Se $\mathcal{R}_0 = 1$ temos uma bifurcação transcritical, ou seja, onde há troca de estabilidade.

Podemos observar estas conclusões acerca do ponto de equilíbrio P_1 através do retrato de fase e órbitas deste modelo, considerando o caso $R_e(0) < 1$ e o caso $R_e(0) > 1$, conforme Figura 11.

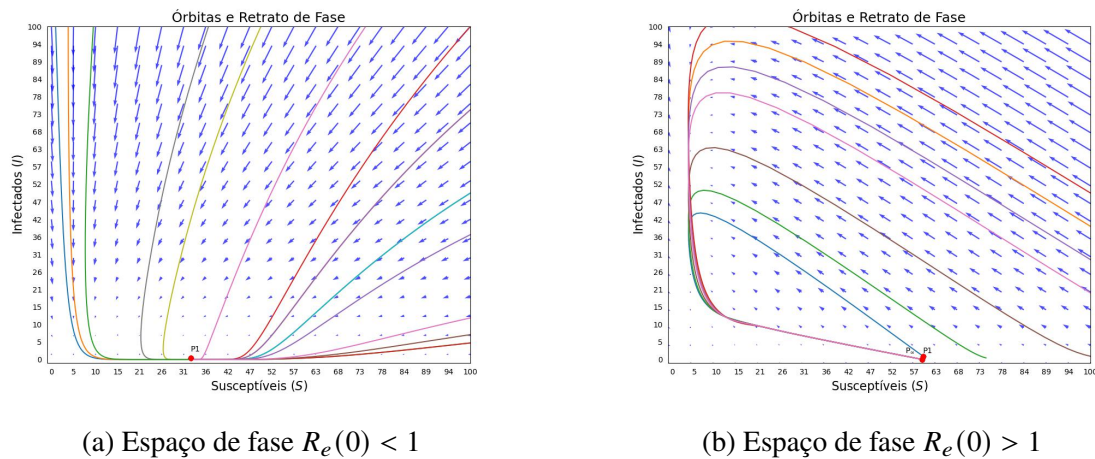


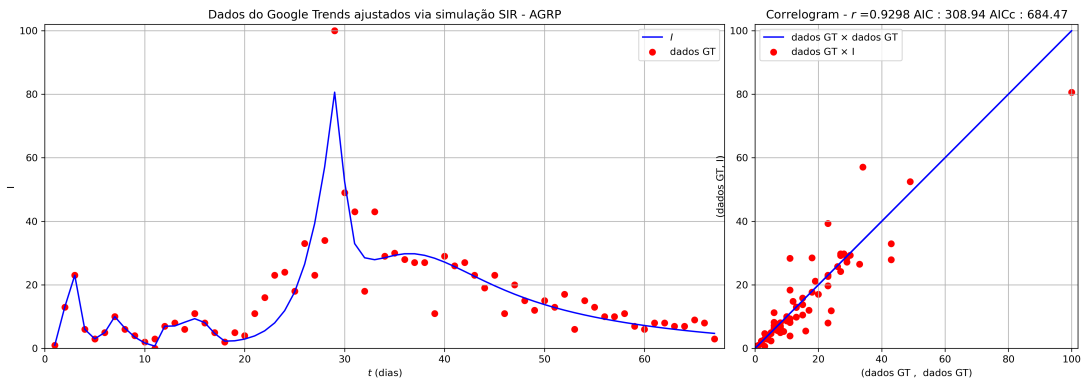
Figura 11 – Exemplo do Espaço de Fase Modelo 24

4.2.3 Resultados Algoritmo Genético

Para este modelo 24 os resultados dos parâmetros obtidos via AGRP e o resultado gráfico estão apresentados respectivamente na Tabela 5 e na figura 12, à seguir:

Tabela 5 – Valores dos parâmetros β , γ , μ_1 , μ_2 , $N(0)$, ξ e Λ otimizados e cálculo aproximado do valor de $\mathcal{R}_0$ em cada período p_i , referente ao modelo 24

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
β	2.4241	2.9600	1.1691	0.1578	2.6543	2.9832	0.6275	2.9955
μ_1	0.1932	0.2600	0.1401	0.1994	0.0748	0.2587	0.0021	0.1050
Λ	4.1337	4.7218	22.0322	0.0040	8.7956	2.0235	26.5018	25.8725
μ_2	0.2453	0.2647	0.1338	0.0845	0.1931	0.0646	0.0756	0.0650
$N(0)$	41.7439	3.0691	9.8011	74.3608	1.1673	53.0267	55.2042	3.0049
γ	0.0619	0.0575	0.0016	0.0573	0.0136	0.0294	0.0021	0.0045
ξ	0.1173	0.1794	0.2497	0.2302	0.2255	0.0059	0.0025	0.2984
$\mathcal{R}_0$	6.6851	6.6658	3.0482	0.5012	6.3409	42.3484	8.0288	8.2434
Min(FO)	0.0006	0.0006	0.0022	0.1366	0.8741	0.6214	2.7106	2.7961

Figura 12 – (a) Frequência de buscas relativas realizadas no Google no período analisado e curva de ajuste considerando os parâmetros do modelo 24 via AGRP; (b) Gráficos dados GT $\times$ dados GT (bissetriz), dados GT $\times$ dados simulados e correlação de Pearson.

4.3 Para além do modelo SIR - Modelo tipo SIR - Aberto sem interação Infectado - Recuperado

Ainda com o objetivo de, por meio do modelo SIR clássico na Seção 2.3, inserir/retirar novos termos baseando-nos que, no processo infodêmico juntamente com o contágio social e a limitação natural e dimensional aqui posta, considera-se que processos como estes carregam intrinsecamente, outros desdobramentos e fenômenos, optamos por considerar no modelo 24, uma outra abordagem mais próxima de certos modelos baseados no SIR com dinâmica vital, que foi de considerar o valor de $\gamma = 0$ ou seja, impor a taxa de ajuda por influência de recuperados nula. Desta forma teremos as seguintes equações

$$\begin{cases} \frac{dS}{dt} = -\beta \frac{SI}{N} - \mu_1 S + \Lambda \\ \frac{dI}{dt} = \beta \frac{SI}{N} - \xi I - \mu_2 I \\ \frac{dR}{dt} = \xi I \end{cases} \quad (28)$$

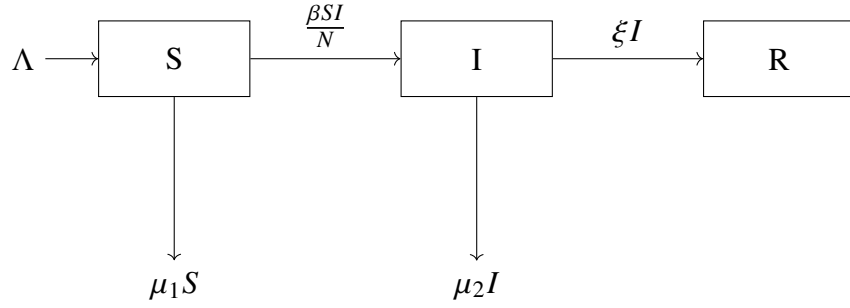


Figura 13 – Diagrama do modelo 28

Vale ressaltar que as descrições e significados dos parâmetros referentes ao modelo 28 são os mesmos, visto que este é um refinamento do modelo 24, desta forma optamos em omiti-las aqui nesta seção.

Analogamente ao caso do modelo 24 uma vez que temos sua descrição normalizada em 25, fazendo $\gamma = 0$ chegamos, apenas para título de abordagem extra, ao sistema normalizado:

$$\begin{cases} \frac{ds}{dt} = (1-s) \frac{\Lambda}{N} + (\mu_2 - \beta) si + \mu_1 s^2 - \mu_1 s \\ \frac{di}{dt} = (\beta + \mu_1) si - \left(\frac{\Lambda}{N} + \xi + \mu_2 \right) i + \mu_2 i^2 \\ \frac{dr}{dt} = \xi i + \mu_1 rs + \mu_2 ri - \frac{\Lambda}{N} r \end{cases} \quad (29)$$

4.3.1 Estudo da estabilidade - Modelo 28

Para este modelo, analogamente ao do outro caso Modelo 24, temos um único ponto de equilíbrio que refere-se ao caso livre de doença², à saber, $P_1 = \left(\frac{\Lambda}{\mu_1}, 0, 0 \right)$.

E para este modelo, sua matriz Jacobiana calculada no ponto P_1 é dada por:

$$J(P_1) = \begin{bmatrix} -\mu_1 & -\frac{\beta\Lambda}{N_0\mu_1} & 0 \\ 0 & \frac{\beta\Lambda}{N_0\mu_1} - (\xi + \mu_2) & 0 \\ 0 & \xi & 0 \end{bmatrix} = \begin{bmatrix} -\mu_1 & -(\mu_2 + \xi) R_e(0) & 0 \\ 0 & (\mu_2 + \xi) (R_e(0) - 1) & 0 \\ 0 & \xi & 0 \end{bmatrix} \quad (30)$$

² Ressalta-se que para essa afirmação este fato está sendo considerado, já que, $\forall R^* \geq 0$ os pontos $P^* = \left(\frac{\Lambda}{\mu_1}, 0, R^* \right)$ anulariam o sistema.

De forma análoga ao outro caso, o polinômio característico deste sistema em P_1 será dado por

$$p(\lambda) = -\lambda^3 + ((\xi + \mu_2)(R_e(0) - 1) - \mu_1)\lambda^2 + \mu_1(\mu_2 + \xi)(R_e(0) - 1)\lambda \quad (31)$$

Aplicando o critério de Routh-Hurwitz sobre o polinômio $\tilde{P}(\lambda) = -p(\lambda)$:

λ^3	1	$\mu_1(\xi + \mu_2)(1 - R_e(0))$
λ^2	$\mu_1 - (\xi + \mu_2)(R_e(0) - 1)$	0
λ^1	$\mu_1(\xi + \mu_2)(1 - R_e(0))$	0
λ^0	$\mu_1(\xi + \mu_2)(1 - R_e(0))$	

Tabela 6 – Tabela de Routh-Hurwitz referente ao ponto de equilíbrio P_1 e polinômio $\tilde{P}(\lambda) = -p(\lambda)$ - Equação 31

Observação: Foi utilizada na Tabela 6 na última linha o polinômio auxiliar visto que a última linha da tabela de Routh resultou-se numa linha nula (caso especial).

Se $1 - R_e(0) > 0$ temos que $\mu_1 - (\xi + \mu_2)(R_e(0) - 1) > 0$ logo o sistema no ponto P_1 é estável e se $1 - R_e(0) < 0$ temos que $\mu_1(\xi + \mu_2)(1 - R_e(0)) < 0$ o que implica que o sistema em P_1 neste caso é instável, em outras palavras temos a seguinte conclusão sobre o ponto de equilíbrio P_1 :

- Se $R_e(0) < 1$ o sistema é estável;
- Se $R_e(0) > 1$ o sistema é instável;
- Se $R_e(0) = 1$ temos uma bifurcação transcítica, ou seja, onde há troca de estabilidade.

Vale a pena ressaltar que no instante $t = 0$ temos $R_e(0) = \mathcal{R}_0$. O que implica que as condições anteriores podem ser reescritas da seguinte forma:

- Se $\mathcal{R}_0 < 1$ o sistema é estável;
- Se $\mathcal{R}_0 > 1$ o sistema é instável;
- Se $\mathcal{R}_0 = 1$ temos uma bifurcação transcítica, ou seja, onde há troca de estabilidade.

É possível perceber também que a análise da estabilidade para este modelo 28 é idêntica ao caso do modelo 24, visto que existe apenas o ponto de equilíbrio P_1 também para este, decorrente da não existência de um segundo ponto de equilíbrio endêmico, devido à 3ª equação $\left(\frac{dR}{dt} = \xi I\right)$ se anular se, e somente se, $I = 0$ considerando o fato de $\xi \neq 0$. E ainda pelo fato da matriz jacobiana $J(P_1)$ (26) não depender de γ o que remete obter o mesmo polinômio característico para este sistema. Mas como a retirada deste termo não refletirá uma mudança no sistema? Haverá uma mudança sim na dinâmica do sistema visto que o fato da presença de elementos recuperados

influenciarem positivamente na população de infectados por si só já remete algo, isso ao longo do tempo refletirá numa convergência populacional diferente com o não esgotamento das três populações, mas especificamente no modelo com interação IR a população I tende a se anular como esperado de uma forma mais rápida acelerando à recuperação. É possível identificar estes efeitos por meio da simulação numérica, subfigura 11(b) (Figura 11) e subfigura 14(b) (Figura 14), na evolução destes modelos fazendo $t \rightarrow \infty$ o que mostra que nem todos os indivíduos susceptíveis serão infectados o que faz bastante sentido pois em um processo epidêmico nem todos susceptíveis são acometidos da infecção e em certos cenários de retroalimentação como os casos aqui colocados sempre terão população infectada.

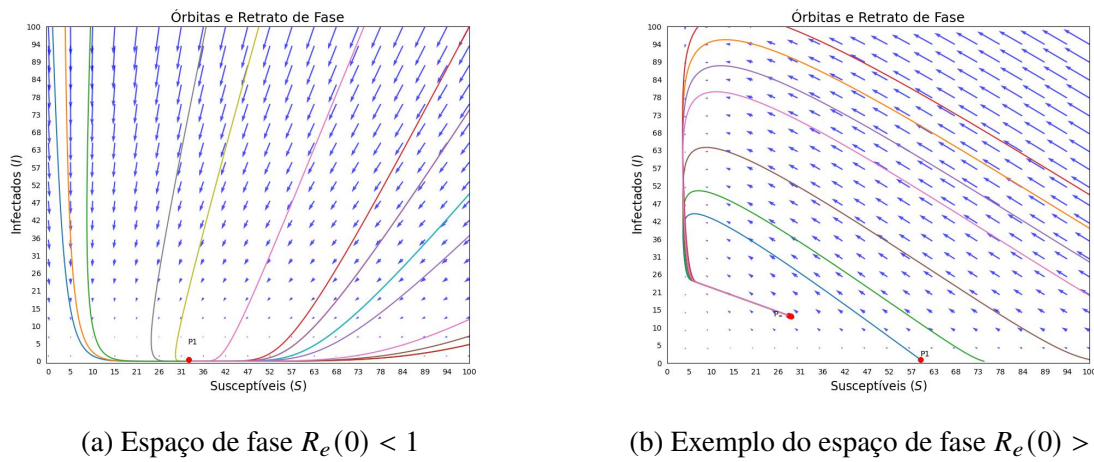


Figura 14 – Exemplo do Espaço de Fase Modelo 28

4.3.2 Resultados Algoritmo Genético

Os valores obtidos via AGRP para o modelo 23 estão descritos como dito anteriormente na Tabela 3 bem como os valores de $\mathcal{R}_0$ em cada período de engajamento e de desengajamento. E para o modelo desta seção, modelo 28, os resultados estão apresentados na Tabela 7 e o resultado gráfico se encontra na figura 15.

Tabela 7 – Valores dos parâmetros β , μ_1 , μ_2 , $N(0)$, ξ e Λ otimizados e cálculo aproximado do valor de $\mathcal{R}_0$ em cada período p_i referente ao modelo 28.

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
β	2.3382	0.1008	2.8830	0.8871	2.9992	1.9366	0.6132	0.1003
μ_1	0.2044	0.3000	0.1718	0.2995	0.0615	0.2584	0.0011	0.2985
Λ	0.9423	0.0053	9.8639	0.0072	3.7189	0.0384	29.9852	0.0191
μ_2	0.2259	0.3000	0.1036	0.2998	0.0363	0.0858	0.0247	0.0379
$N(0)$	50.0915	99.9998	30.1508	46.4348	1.0001	1.3528	99.9876	13.9503
ϵ	0.0974	0.3000	0.2158	0.2998	0.2385	0.2764	0.0011	0.0010
$\mathcal{R}_0$	7.2320	0.1680	9.0271	1.4797	10.9114	5.3473	23.7644	2.5795
Min(FO)	0.0003	0.7351	0.3472	0.1527	0.8535	0.4905	2.6319	2.4767

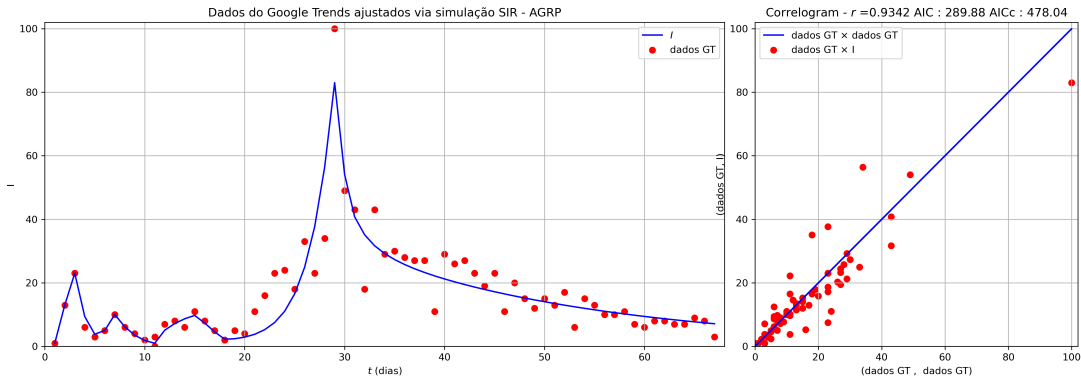


Figura 15 – (a) Frequência de buscas relativas realizadas no Google no período analisado e curva fit considerando os parâmetros do modelo 28 via AGRP; (b) Gráfico de dispersão entre os dados obtidos pelo modelo 28 via AGRP e os dados do Google Trends.

Seguem tabelas demonstrativas (Tabelas 8 e 9), explicitando os números básicos de reprodução $\mathcal{R}_0$ em cada período p_i e os valores AIC e AICc, para cada um dos modelos 23, 24 e 28, conforme experimentações descritas nesta seção.

Tabela 8 – Valores dos números básicos de reprodução $\mathcal{R}_0$ em cada período referente ao Modelo 23, Modelo 24 e Modelo 28, ajustado aos dados do Google Trends por meio do AGRP

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
Modelo 23	19.9464	0.0173	4.7878	0.5169	0.5734	0.7475	5.9822	0.0428
Modelo 24	6.6851	6.6658	3.0482	0.5012	6.3409	42.3484	8.0288	8.2434
Modelo 28	7.2320	0.1680	9.0271	1.4797	10.9114	5.3473	23.7644	2.5795

Tabela 9 – Valores dos Coeficientes de correlação de Pearson (r), AIC (19) e AICc (20) ao Modelo 23, Modelo 24 e Modelo 28, ajustado aos dados do Google Trends por meio do AGRP

	Modelo 23	Modelo 24	Modelo 28
r	0.8637	0.9298	0.9342
AIC	379.78	308.94	289.88
AICc	403.78	684.47	478.04

Capítulo 5

Modelagem da Polarização Social e Política

As interações sociais são fenômenos de natureza complexa e, portanto, são de difícil tratamento por meios reducionistas e determinísticos (FIGUEIREDO et al., 2020). Entretanto, a Sociologia é uma ciência que busca observar as regularidades nestes fenômenos. Algumas destas regularidades envolvendo comportamentos e interações sociais admitem abordagem com métodos estatísticos que envolvem Matemática e Física. Estas abordagens se dão via a codificação matemática de conceitos identificados no fenômeno complexo para produzir modelos que tentam explicar aspectos do seu comportamento. Como resultado, é possível estabelecer parâmetros para mensurar no tempo o próprio fenômeno social de interesse. Esta abordagem pode favorecer medidas educacionais e de informação em políticas públicas para mitigar eventuais impactos sociais.

A possibilidade de se produzir modelos matemáticos sobre um sistema complexo aluda a codificar os conceitos presentes no fenômeno. Também pode ser útil para estabelecer parâmetros de mensuração do fenômeno no tempo em que ocorrem. Esta abordagem pode favorecer medidas educacionais e de informação para mitigar impactos sociais.

Neste capítulo, trataremos o fenômeno da polarização, engajamento (ou desengajamento), visibilidade, legitimidade e ressonância do ponto de vista epidêmico.

5.1 Contextualizações sociais e exemplos.

Diante da mensagem muito controversa como esta: “Todos tem o direito de matar alguém”, é certo que a maioria da sociedade teria uma reação de repulsa. A maioria dos cidadãos compartilham da ideia de que subtrair a vida de alguém é uma atitude socialmente inaceitável.

Gera-se portanto, um consenso social negativo. Isso é justificável pois “Quando nos pomos a questão de subtrair a vida de outro, a primeira reação que geralmente nós (homens e mulheres considerados comuns) temos, é de ojeriza”(CHAGAS, 2016). Neste sentido, uma mensagem como essa é altamente ilegítima e certamente não possui grande poder de repercussão por parte da mídia. Mais precisamente, se um determinado grupo, começar a requerer por meio desta mensagem esse “direito”, tal iniciativa não terá a força necessária para ser objeto de veiculação pela mídia, justamente pela falta de legitimidade e mesmo que gerasse alguma repercussão, seria repudiada rapidamente. Em síntese, isso corresponde a um cenário onde temos uma mensagem ilegítima com baixa ressonância.

Em contrapartida, a maioria das pessoas, concordam que a mensagem voltada para a

educação: "A relação professor-aluno deve ser saudável, respeitosa, com empatia, de forma a favorecer ativamente o engajamento do aluno, por meio de ações que favoreçam o suporte tanto acadêmico quanto emocional, promovendo assim um ambiente educacional enriquecedor e produtivo." é uma mensagem que nos remete a um consenso positivo, haja vista que a maioria das pessoas entende e compreende o seu significado. Neste sentido, temos um exemplo de consenso social positivo e, por isso, teremos uma mensagem altamente legítima e de baixa ressonância.

Agora uma mensagem requerendo o direito ao procedimento de Eutanásia levaria a uma discussão muito mais abrangente e seria um assunto com alta visibilidade pela mídia e certamente possuiria uma alta ressonância e que possivelmente nos levaria a um sistema polarizado. Isso poderia ser justificado por exemplo, pelo fato de que "No Direito Brasileiro, a eutanásia caracteriza homicídio, pois é conduta típica, ilícita e culpável." (DODGE, 2009), mas em alguns outros países tal procedimento é aceitável e legal. Ou seja, o tema por si só é controverso e depende de crenças intrapessoais e até mesmo culturais. Este assunto ou a exigência da aplicação deste direito é ressoante. Citemos, a título de exemplo, o ocorrido na Espanha no ano de 2021 e que ganhou visibilidade no referido ano, que foi a notícia referente a Aprovação pela regularização do procedimento da Eutanásia. Esta ganhou os holofotes em março de 2021, no site da CNN España, onde foi respaldada pelo atual presidente da Espanha (CNN, 2021). O resultado da votação por meio dos representantes foi de 141 votos contrários e 202 votos a favor. Ou seja, o assunto foi divisivo gerou debate no parlamento e, consequentemente, permaneceu por semanas em discussão. Este é um exemplo de tema cuja discussão se dá longe de consensos e, portanto, é um exemplo de grande ressonância.

Mensagens como estas relacionadas a assuntos como políticas de asilo, legalização do aborto (MIGUEL, 2012), união entre pessoas de mesmo sexo, dentre outros tendem a gerar muita repercussão e, portanto, podem ser utilizadas como ferramentas políticas ao serem exploradas politicamente para gerar visibilidade pela sua capacidade de causar ressonância.

É importante ressaltar que, no contexto mais amplo, a mensagem que é utilizada com o intuito de aumentar a visibilidade não necessariamente favorece majoritariamente a imagem do ator ou do grupo político que a utiliza. O que ocorre é a projeção por meio de uma repercussão nos meios de comunicação de massa e, consequentemente, um aumento de visibilidade destes atores.

5.1.1 Associações e Relações

Pelas características de crescimento e decrescimento da ressonância em função da legitimidade a relação entre estas podem ser ilustradas por meio de um gráfico parabólico, onde se destacam três possíveis cenários (GIANNINI, 2021).

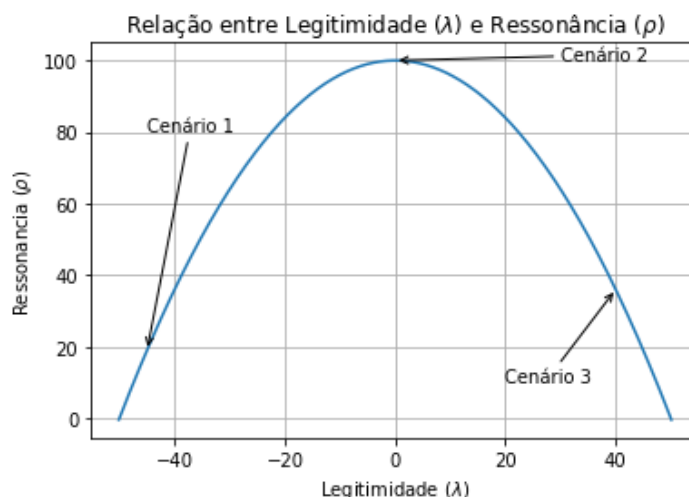


Figura 16 – Relação entre Legitimidade λ e Ressonância ρ

No cenário 1, temos uma legitimidade negativa portanto uma mensagem altamente ilegítima com uma ressonância baixa, ou seja, neste sentido esta poderia ser uma boa pauta para a discussão mas no aspecto de propagação da mensagem, provavelmente não haveria uma grande repercussão. Podemos associar esse cenário ao exemplo descrito na seção anterior e que abordara o direito de matar.

Já no cenário 2, temos uma legitimidade nula indicando baixo consenso global. Essa situação tem potencial para gerar muito debate e polarização. O maior engajamento neste debate caracteriza um cenário de máxima ressonância. Temos aí uma legitimidade maior em comparação com o cenário 1. Isso implicaria numa situação onde a mensagem, é amplamente discutida, por grupos que a aceitam e por grupos que não a aceitam e dada a sua natureza controversa chegará ao conhecimento de mais pessoas, ou seja, criando um ambiente de maior ressonância e consequentemente maior visibilidade. Por associação à seção anterior, temos o caso do direito ao procedimento da eutanásia, por exemplo.

Já, no cenário 3, o que temos é uma legitimidade positiva portanto maior e uma ressonância mais baixa em comparação ao cenário anterior, neste sentido temos uma mensagem mais positivamente consensual, com baixo poder de repercussão, em outras palavras com baixa ressonância, o que nos remete a associação com a descrição do direito pedagógico referente a relação professor-aluno, explanada na seção anterior.

Cenários como o 1 e o 3 aplicados em ciências políticas, gerariam baixa visibilidade, já que possuem alto consenso e consequentemente uma baixa ressonância, em contrapartida, cenários como o 2, gerariam uma alta visibilidade, já que possuem baixo consenso. Políticos vem procurando abrir mão de certos graus de legitimidade com objetivo de aumentar a ressonância para, então, aumentar a sua visibilidade (HAM; ELSAS, 2024).

Neste trabalho, desejamos fundamentar esta teoria através de um modelo matemático para futuras aplicações e previsões. Utilizamos, para isto, o modelo contínuo SIR, modelo 23 adaptado com duas classes de infectados, I_+ e I_- , que neste caso poderiam ser vistas como populações que

reagem de maneiras distintas a um tema ou mensagem, uma positivamente e a outra que reage negativamente. A função de ressonância, descrita por ρ , descreve a interação entre os grupos e é modelada pela soma cumulativa de toda efervescência entre as populações (Modelo 32).

5.2 Modelo Sociofísico Baseado em SIR para Polarização

Este modelo foi construído baseando-se no modelo clássico SIR modelo 23, considerando suas equações populacionais e uma equação diferencial que descreve uma nova variável de interesse com característica não populacional, a variável ρ .

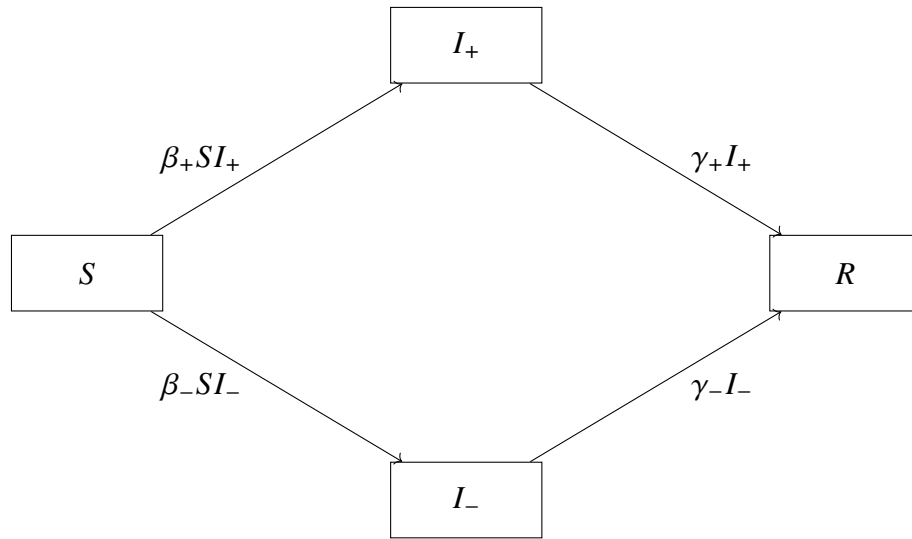


Figura 17 – Diagrama referente as populações S , I_+ , I_- e R do modelo 32

$$\left\{ \begin{array}{l} \frac{dS}{dt} = -\beta_+ S I_+ - \beta_- S I_- \\ \frac{dI_+}{dt} = \beta_+ S I_+ - \gamma_+ I_+ \\ \frac{dI_-}{dt} = \beta_- S I_- - \gamma_- I_- \\ \frac{dR}{dt} = \gamma_+ I_+ + \gamma_- I_- \\ \frac{d\rho}{dt} = \kappa I_- I_+ \end{array} \right. \quad (32)$$

Neste, S representa a população de susceptíveis que estão por decidir entre dois tipos de infectados I_+ ou I_- em resposta a uma mensagem. Podem ser entendidos como correntes de opinião polarizadas, uma em cada sentido e com taxas de adesão β_+ e β_- respectivamente. A população R representa os indivíduos recuperados que aqui são aqueles indivíduos que já esqueceram do conteúdo do tema em pauta ou que se informaram melhor e não vêm mais sentido na polêmica. Neste sentido, as populações I_+ e I_- não mudam de ideia, ou seja, uma vez decidido por aceitar ou não o teor do tema, não mudam de opinião, até se tornarem indivíduos recuperados.

A variável ressonância ρ , é identificada neste modelo como sendo a intensidade da interação entre os dois grupos I_+ e I_- e modelada pela soma cumulativa de toda efervescência. Desta forma para cada escolha de λ , temos ρ obtida pela integral :

$$\rho(t) = \int_0^t \kappa I_+(\tau) I_-(\tau) d\tau \quad \tau \in [0, t] \quad \kappa = cte \quad (33)$$

As taxas de adesão são escolhidas segundo a definição:

$$\beta_+ + \beta_- = \beta \quad \beta_+, \beta_- > 0 \quad (34)$$

Os parâmetros do modelo SIR são escolhidos por simplicidade de modo que as taxas de recuperação, ou qualificação da mensagem, γ_+ e γ_- , no presente caso, satisfaçam as condições:

$$\gamma_+ + \gamma_- = \gamma \quad \gamma_+, \gamma_- > 0 \quad (35)$$

O número de reprodução definido no tempo t é dado por $\mathcal{R}_0(t) = \frac{\beta(t)}{\gamma(t)}$ (SIMOS et al., 2021). Considerando o fato de que temos independentemente, dois compartimentos de infectados I_+ e I_- com suas respectivas taxas de contágio e de recuperação, considere as notações $\mathcal{R}_0^+(t) = \frac{\beta_+(t)}{\gamma_+(t)}$ e $\mathcal{R}_0^-(t) = \frac{\beta_-(t)}{\gamma_-(t)}$.

Para descrever a Legitimidade $\lambda(t)$ a consideramos como sendo um número real entre -1 e 1 e dada, considerando as descrições de $\mathcal{R}_0^+(t)$ e $\mathcal{R}_0^-(t)$, da seguinte forma:

$$\lambda(t) = \frac{\mathcal{R}_0^+(t) - \mathcal{R}_0^-(t)}{\mathcal{R}_0^+(t) + \mathcal{R}_0^-(t)} \quad (36)$$

Tal definição de λ reflete o comportamento de que quanto mais positiva esta for, maior engajamento por parte da população que reage positivamente e quanto mais negativa for a legitimidade, maior engajamento por parte da população que reage negativamente.

Em decorrência destas definições, a ressonância ρ depende da legitimidade λ , já que cada escolha dos parâmetros β_+ e γ_+ , fixados β e γ , corresponderá a uma realização do modelo proposto, com uma dinâmica distinta e uma conjugação de respostas.

5.2.1 Pontos de Equilíbrio do Modelo de Polarização e Análise de estabilidade

Este modelo apresenta duas cepas (duas classes infecciosas I_+ e I_-) e portanto fragmentaremos esta abordagem e estudo considerando cada uma separadamente e depois um estudo conjunto.

Primeiramente considere $I_- = 0$, o $\mathcal{R}_0^+$ referente a cepa I_+ é obtido da seguinte forma:

$$\frac{dI_+}{dt} = \beta_+ S I_+ - \gamma_+ I_+ \approx \beta_+ N I_+ - \gamma_+ I_+ = (\beta_+ N - \gamma_+) I_+$$

onde foi utilizada a condição de homogeneidade, ou seja, $S(0) \approx N$ no início da epidemia. Assim, $\mathcal{R}_0^+ = \frac{\beta_+ S(0)}{\gamma_+} = \frac{\beta_+ N}{\gamma_+}$. De forma análoga, considerando $I_+ = 0$ obtemos analogamente que $\mathcal{R}_0^- = \frac{\beta_- S(0)}{\gamma_-} = \frac{\beta_- N}{\gamma_-}$. Em comparação com a descrição dada anteriormente temos duas observações à saber, estamos em nosso modelo usando as quantidades absolutas das populações (embora que nas simulações considerou-se as porcentagens populacionais) e considerando os fatores, taxas de infecção e de recuperação constantes, o que justifica carregar na descrição do número básico de reprodução dependendo do $S(0) \approx N$ e consequentemente constante.

É importante frisar também que a abordagem pode ser calculada inicialmente analisando cada cepa separadamente ignorando a presença da outra, visto que o cenário analisado é o do início da evolução, mas que isso não é aconselhável justamente pelo fato de que pode-se (e o que normalmente ocorre) ter a presença de ambas classes infecciosas durante todo o processo.

Considerando ambas as cepas e escolhendo o método da matriz da próxima geração (NGM) dado por (DRIESSCHE, 2017), o $\mathcal{R}_0$ do sistema, será o raio espectral (maior autovalor) da matriz $K = FV^{-1}$, ver seção 3.4. O sistema possui dois compartimentos infectados, I_+ e I_- , desta forma $F(I_+, I_-) = \begin{bmatrix} \beta_+ N & 0 \\ 0 & \beta_- N \end{bmatrix}$ e $V(I_+, I_-) = \begin{bmatrix} \gamma_+ & 0 \\ 0 & \gamma_- \end{bmatrix}$, sendo a matriz F referente a novas infecções e a matriz V a matriz referente a transferência de indivíduos para o outro compartimento. A matriz da próxima geração será dada por

$$K = F \cdot V^{-1} = \begin{bmatrix} \beta_+ N & 0 \\ 0 & \beta_- N \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\gamma_+} & 0 \\ 0 & \frac{1}{\gamma_-} \end{bmatrix} = \begin{bmatrix} \frac{\beta_+ N}{\gamma_+} & 0 \\ 0 & \frac{\beta_- N}{\gamma_-} \end{bmatrix} = \begin{bmatrix} \mathcal{R}_0^+ & 0 \\ 0 & \mathcal{R}_0^- \end{bmatrix}$$

Dado ao fato da matriz K ser diagonal, seus autovalores serão os termos da sua diagonal ou seja, os números básicos de reprodução independentes e consequentemente, o $\mathcal{R}_0$ do sistema será

$$\mathcal{R}_0 = \max \{ \mathcal{R}_0^+, \mathcal{R}_0^- \}$$

Procuremos primeiramente analisar o ponto de equilíbrio livre de doença ou seja considerando $I_+ = I_- = 0$, com isso temos que $N = S + I_+ + I_- + R = S + R = S$, visto que sem população infecciosa $R = 0$. Assim o ponto de equilíbrio livre de doença é $P_1 = (N, 0, 0, 0)$, observe que neste caso temos ainda, mesmo que não seja uma condição necessária do ponto de vista populacional, a equação de energia de iteração de infectados $\frac{d\rho}{dt}$ também nula neste ponto.

Por outro lado, busquemos agora os pontos de equilíbrio endêmicos, ou seja, quando a população de infectados é não nula. É possível perceber utilizando a equação populacional de R que para termos $\frac{dR}{dt} = 0$ com as condições naturais de que γ_+, γ_- positivas e I_+ e I_- não negativas que a única possibilidade de anulamento desta equação é que $I_+ = I_- = 0$, o que mostra que este sistema não possui pontos de equilíbrio não triviais.

5.2.2 Estudo da Estabilidade

Para estudarmos a estabilidade deste modelo utilizaremos a linearização em torno do ponto de equilíbrio. Considere as seguintes descrições:

$$F_1(S, I_+, I_-) = -I_+S\beta_+ - I_-S\beta_-, F_2(S, I_+, I_-) = I_+S\beta_+ - I_+\gamma_+ \text{ e } F_3(S, I_+, I_-) = I_-S\beta_- - I_-\gamma_-.$$

Temos que a matriz jacobiana do sistema será dada por:

$$J(S, I_+, I_-) = \begin{bmatrix} -I_+\beta_+ - I_-\beta_- & -S\beta_+ & -S\beta_- \\ I_+\beta_+ & \gamma_+S\beta_+ - \gamma_+ & 0 \\ I_-\beta_- & 0 & S\beta_- - \gamma_- \end{bmatrix}. \text{ Assim temos:}$$

$$\bullet P_1: J(P_1) = \begin{bmatrix} 0 & -N\beta_+ & -N\beta_- \\ 0 & \gamma_+ \left(\frac{N\beta_+}{\gamma_+} - 1 \right) & 0 \\ 0 & 0 & \gamma_- \left(\frac{N\beta_-}{\gamma_-} - 1 \right) \end{bmatrix} \text{ também podemos reescrevê-la em função}$$

dos números básicos de reprodução da seguinte forma:

$$J(P_1) = \begin{bmatrix} 0 & -R_0^+\gamma_+ & -R_0^-\gamma_- \\ 0 & \gamma_+ (R_0^+ - 1) & 0 \\ 0 & 0 & \gamma_- (R_0^- - 1) \end{bmatrix} \text{ o que nos leva ao polinômio característico}$$

$$P_1(\lambda) = \gamma_+\gamma_-\lambda (R_0^+ \cdot (1 - R_0^-) + R_0^- - 1) - \lambda^3 + \lambda^2 (\gamma_+ (R_0^+ - 1) + \gamma_- (R_0^- - 1)).$$

Aplicando o critério de Routh-Hurwitz sobre o polinômio $\tilde{P}_1(\lambda) = -P_1(\lambda)$ temos:

λ^3	1	$\gamma_+ (R_0^+ R_0^- \gamma_- - R_0^+ \gamma_- - R_0^- \gamma_- + \gamma_-)$
λ^2	$\gamma_+ (1 - R_0^+) + \gamma_- (1 - R_0^-)$	0
λ^1	$\gamma_+\gamma_- (R_0^+ - 1) (R_0^- - 1)$	0
λ^0	$\gamma_+\gamma_- (R_0^+ - 1) (R_0^- - 1)$	-

Tabela 10 – Tabela de Routh-Hurwitz - Modelo 32

Observação: Para a construção da última linha foi utilizado a derivada do polinômio auxiliar (considerando a penúltima linha), visto que a última linha da tabela de Routh resultou-se numa linha nula (caso especial).

É possível perceber que se:

- * Se $(R_0^+ - 1) > 0$ e $(R_0^- - 1) > 0$ o sistema em P_1 é instável;
- * Se $(R_0^+ - 1) > 0$ e $(R_0^- - 1) < 0$ o sistema em P_1 é instável;
- * Se $(R_0^+ - 1) < 0$ e $(R_0^- - 1) > 0$ o sistema em P_1 é instável;
- * Se $(R_0^+ - 1) < 0$ e $(R_0^- - 1) < 0$ o sistema em P_1 é estável;

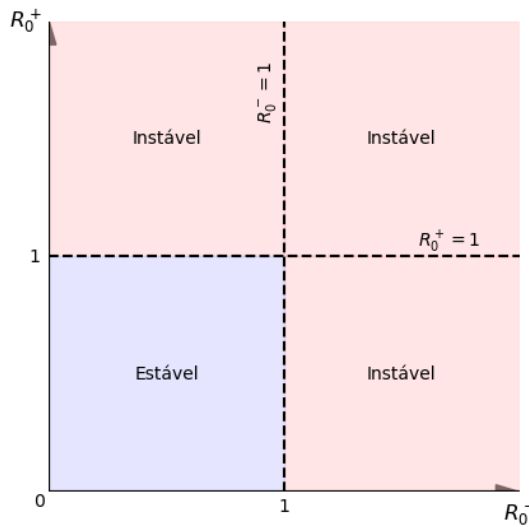
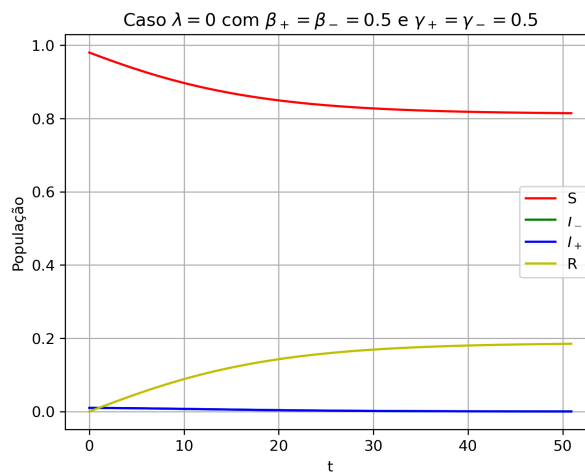


Figura 18 – Descrição de regiões de Estabilidade/Instabilidade do modelo 32

5.2.3 Simulações referentes as populações

Nas simulações consideramos $\kappa = 1.0$, as mesmas condições iniciais, de 1% da população total fixa, isto é, $I_-(0) = I_+(0) = 0.01$.

Para o caso geral de simulação da curva ρ consideramos 51 valores de λ entre -1 e 1 . Por construção, o modelo possui simetria em relação a $\lambda = 0$. Esta característica é desejável pois não quisemos a priori que houvesse distinção entre a qualidade da mensagem entre as populações I_+ e I_- . Seguem as simulações para este caso em particular, ou seja, $\lambda = 0$, considerando ainda $\beta_+ = \beta_- = 0.5$ Figuras 19, 20 e 21.

Figura 19 – Exemplo - Variações populacionais - $\lambda = 0$, com populações I_+ e I_- superpostas.

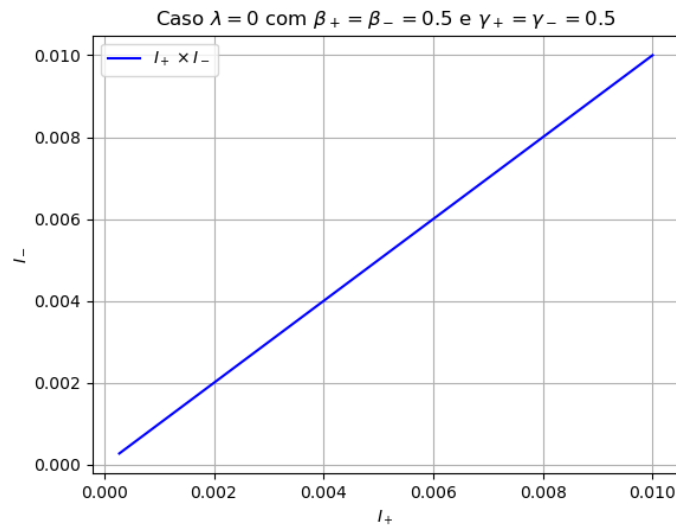


Figura 20 – Resultado da evolução do modelo SIR com duas populações reagindo a uma mensagem política caracterizada pelo valor de $\lambda = 0$. A trajetória das populações I_+ e I_- estão representadas na projeção plana do espaço de estados.

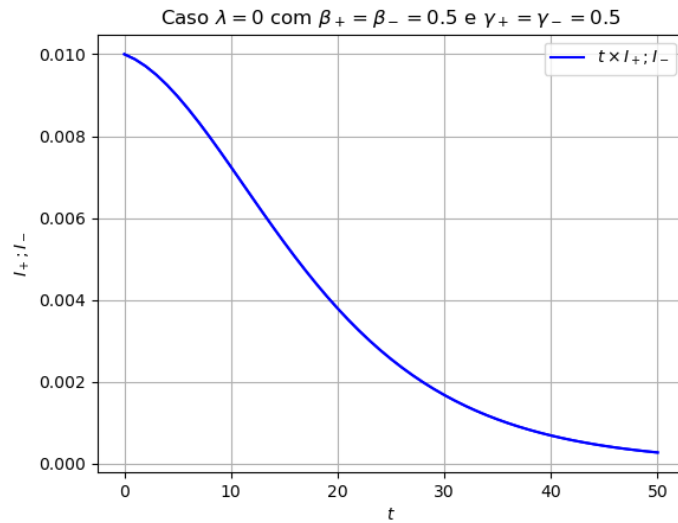


Figura 21 – Resultado da evolução do modelo SIR com duas populações reagindo a uma mensagem política caracterizada pelo valor de $\lambda = 0$. A trajetória das populações I_+ e I_- estão representadas na projeção plana do espaço de estados.

Mantendo as mesmas condições iniciais, de 1% da população total fixa, isto é, $I_-(0) = I_+(0) = 0.01$, considerando os valores $\gamma_+, \gamma_- \in [0.01, 0.99]$, isto é, variando os valores de γ_+ e de γ_- , considerando $\beta_+ = 0.3$ e $\beta_- = 0.7$ temos a Figura 22.

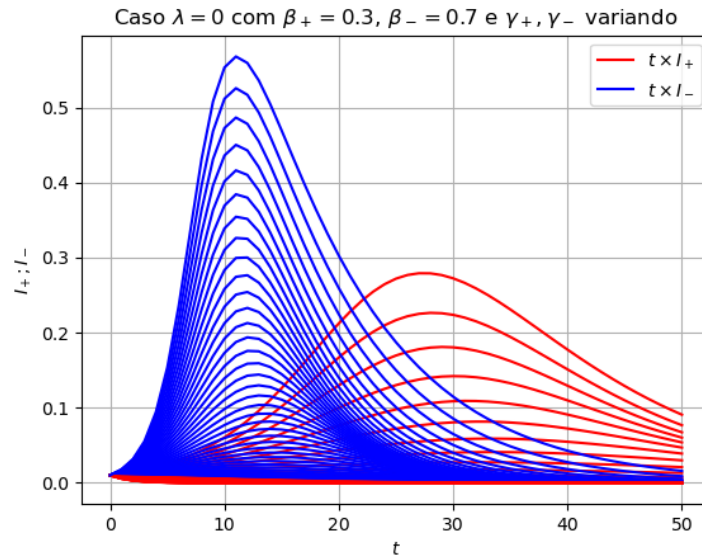


Figura 22 – Evoluções das populações I_+ e I_- , considerando variação de γ_+ e de γ_- de 0.01 a 0.99 e $\beta_+ = 0.3$ e $\beta_- = 0.7$. Evolução mais acentuada da população I_- em comparação com I_+ induzida do fato de que taxa de adesão de β_- é superior a taxa de adesão de β_+ .

Agora, considerando ainda as mesmas condições iniciais, de 1% da população total fixa, isto é, $I_-(0) = I_+(0) = 0.01$, os valores $\beta_+, \beta_- \in [0.0, 1.0]$, isto é variando os valores de β_+ e de β_- , considerando $\gamma_+ = 0.3$ e $\gamma_- = 0.7$ temos a Figura 23.

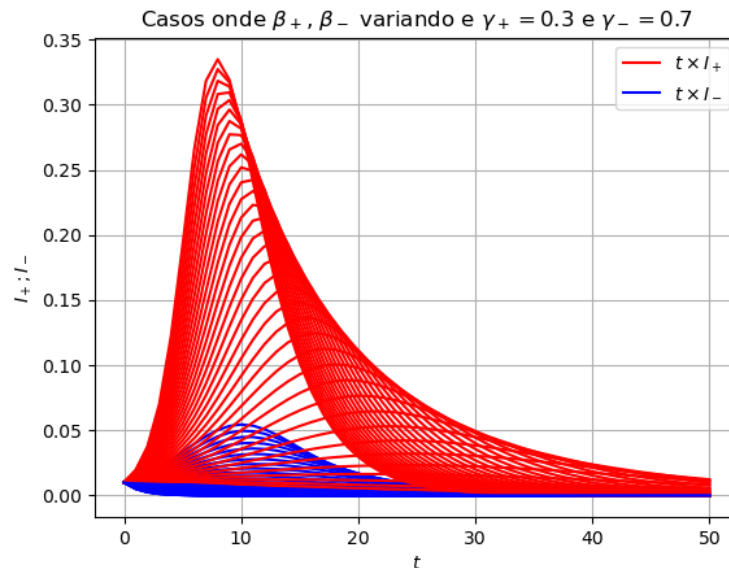


Figura 23 – Evoluções das populações I_+ e I_- , considerando variação de β_+ e de β_- de 0.0 a 1.0 e $\gamma_+ = 0.3$ e $\gamma_- = 0.7$. Evolução mais acentuada da população I_+ visto que a taxa de recuperação γ_- é superior a γ_+ , ou seja, a população I_- se recupera mais rapidamente.

Outra simulação que fizemos foi de considerar ambos os parâmetros variando dentro dos domínios escolhidos Figura 24.

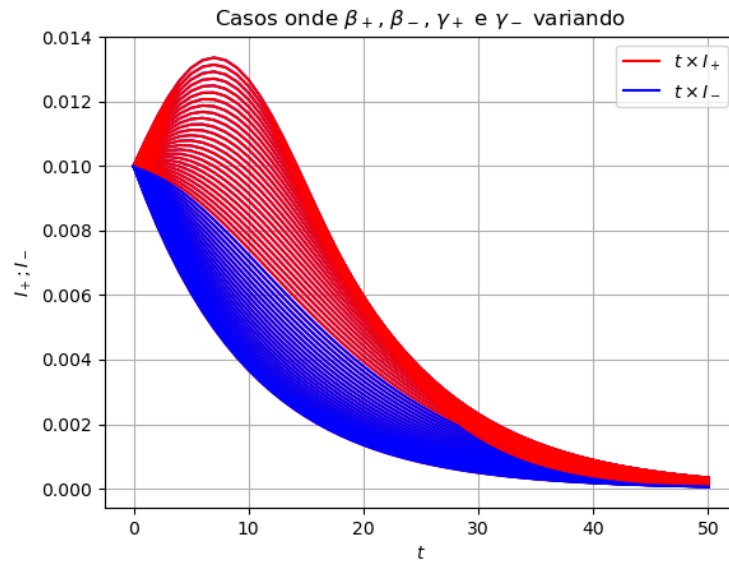


Figura 24 – Evoluções das populações I_+ e I_- , considerando γ_+ , γ_- , variando de 0.01 a 0.99 e β_+ e β_- variando de 0.0 a 1.0.

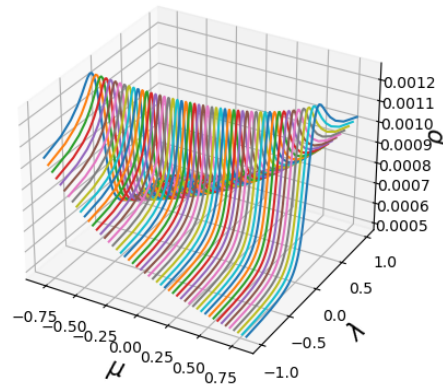
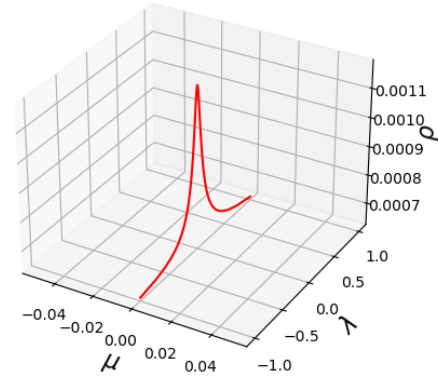
5.2.4 Desengajamento, Legitimidade e Ressonância

Para descrever a variável desengajamento (e correspondentemente o engajamento) agregamos aos parâmetros e termos já definidos no modelo, uma variável μ que descreve a taxa de desengajamento da população I_+ , que reage positivamente em relação a população I_- que reage negativamente, esta é dada pela diferença entre as taxas de recuperação γ_+ e γ_- , ou seja,

$$\mu = \gamma_+ - \gamma_- \quad \gamma_+, \gamma_- > 0 \quad (37)$$

Essa descrição pode ser interpretada de modo que $\mu > 0$ significa que a população I_+ está menos empenhada do que a população I_- , ou seja, os indivíduos da população I_+ tendem a esquecer mais rapidamente da mensagem do que da população I_- , já que a taxa de recuperação de I_+ é maior. Já o caso $\mu < 0$ significa que a população I_- está menos empenhada do que a população I_+ , ou seja, os indivíduos desta população tendem a esquecer mais rapidamente da mensagem do que da população I_+ , já que a taxa de recuperação de I_- é maior. E para o caso $\mu = 0$ interpretamos que ambas populações estão empenhadas o que remete que a tendência é de que os indivíduos de cada uma destas esqueçam do tema de forma semelhante.

Com o intuito de analisar o que classificamos aqui como Grau de Polarização, consideramos 51 variações de β_+ e consequentemente de β_- , considerando $\beta = 1.0$ e de forma análoga consideramos também as variações de γ_+ e consequentemente de γ_- , considerando $\gamma = 1.0$, dadas respectivamente pelas equações 34 e 35. Assim, geramos as curvas de evolução $(\mu, \lambda, \rho(t))$ que descrevem o comportamento das populações em função destas três variáveis.

Desengajamento $I_+ \rightarrow I_- (\mu)$ X Legitimidade (λ) X Ressonância (ρ)(a) Desengajamento X Legitimidade X Ressonância ($\mu \times \lambda \times \rho$) $\mu \times \lambda \times \rho$ - Curva central ($\mu = 0$)(b) Mesma intensidade de Desengajamento X Legitimidade X Ressonância ($\mu \times \lambda \times \rho$)Figura 25 – Gráficos dos parâmetros: Desengajamento, Legitimidade e Ressonância com evidenciamento da curva central caso $\mu = 0$.

5.2.5 Conclusões sobre a experimentação referente ao modelo 32

A função ρ conforme a definição neste trabalho, trata-se de uma medida do nível de ressonância social ou, ausência de consenso que causa uma mensagem. Reflete o fato de que quanto maior a energia de interação entre as populações, maior a ressonância. A polarização política é o fenômeno de abrir mão da legitimidade e do consenso para causar ressonância e assim obter visibilidade. Neste sentido, evidenciamos que a polarização ocorrerá, quando o engajamento das populações tiverem intensidades aproximadas, máxima em torno de ou em $\lambda = 0$ o que provocará a ressonância ρ máxima. Há de fato uma semelhança entre o resultado do modelo (figuras 25(a) e 25(b)) e a curva (Figura 16) proposta de forma intuitiva e que descreve os conceitos de legitimidade e de ressonância tratados aqui como variáveis, conforme (GIANNINI, 2021) e (KOOPMANS, 2004). Importante perceber também a concavidade da curva obtida por nosso modelo (figura 25(b)) que descreve um comportamento social interessante, visto que ao aproximar de $\lambda = 0$ a ressonância tende a aumentar mais rapidamente, o que caracteriza o 'boom' informacional (alta visibilidade) ocorrendo de forma repentina. Portanto, podemos dizer que o modelo constitui-se de uma representação matemática para a teoria de Koopmans sobre polarização política e pode ser usado para valoração deste fenômeno ou como parte de um sistema mais completo de estudos de resposta social.

Futuramente consideraremos colocar outras classes de populações (indivíduos mediadores), isto é, os que poderiam por meio da interação, mudar de opinião, amenizando assim a polarização e consequentemente o seu efeito, utilizando para isso equações estocásticas.

5.2.6 Uma nova definição para a função Legitimidade

Um dos estudos que estamos fazendo é de descrever a função Legitimidade por meio do fator de crescimento/decrescimento de uma população I através do número de reprodução basal $\mathcal{R}_0$ mas não de forma direta como descrito anteriormente 36. Desta forma a ideia foi, considerando uma equação do tipo $\frac{dI}{dt} = \beta \frac{S}{N} I - \gamma I$ e supondo $S \approx N$ utilizar o fator de crescimento/decrescimento dado por $\gamma_+ \delta \mathcal{R}_0^+ = \gamma_+ (\mathcal{R}_0^+ - 1)$ referente a população I_+ e $\gamma_- \delta \mathcal{R}_0^- = \gamma_- (\mathcal{R}_0^- - 1)$ referente a população I_- . Por meio desta descrição a função Legitimidade será dada como:

$$\lambda = \frac{\gamma_-}{\gamma_- + \gamma_+} \delta \mathcal{R}_0^- - \frac{\gamma_+}{\gamma_- + \gamma_+} \delta \mathcal{R}_0^+$$

Assim, considerando esta descrição temos a seguinte curva que descreve a legitimidade para o caso $\mu = 0$, onde é possível observar que a ressonância máxima será atingida quando $\lambda = 0$ e que possui uma maior semelhança com a curva descrita intuitivamente (Figura 16).

$\mu \times \lambda \times \rho$ - Curva central ($\mu = 0$)

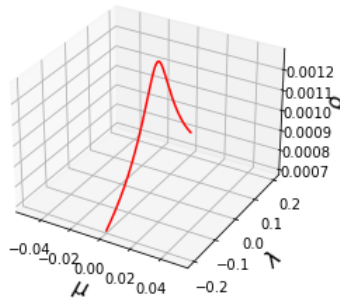


Figura 26 – Curva Legitimidade descrita via nova função λ

5.3 Modelo de Polarização com competição de notícias baseado no SIR e no Lotka-Volterra (LVSIR)

Considere a figura 27 construída a partir de dados obtidos no Google Trends que descrevem as frequências de buscas pelo tema “Vacinação Infantil” nos períodos de 03/04/2022 a 19/02/2023, considerando o período em uma escala semanal. É possível perceber que sua estrutura é complexa mas é possível destacar certos picos. Podemos assim associá-los aos fenômenos de engajamento e de ressonância na sociedade, já que estes picos demonstram que o tema estava mais em evidência nestes momentos. Desta forma é possível modelar as notícias em duas categorias: as notícias qualificadas (verídicas) e as notícias falsas (fake-news).

- Considera-se para a identificação das populações de presas e de predadores, as frequências das postagens das notícias falsas e das notícias verdadeiras;

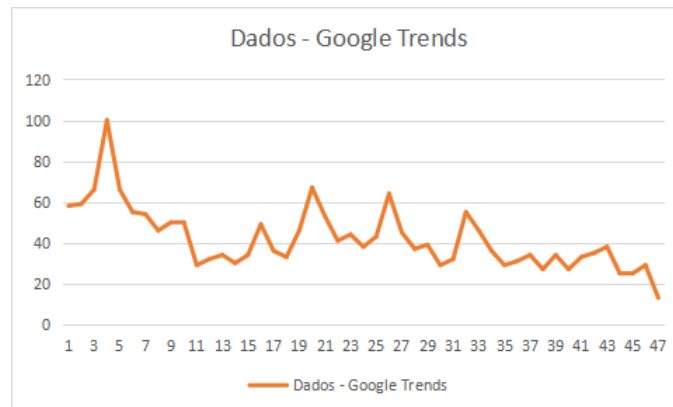


Figura 27 – Dados Google Trends - Vacinação Infantil - 03/04/2022 a 19/02/2023 (semanas)

- Estamos supondo que as notícias falsas surgem exponencialmente na ausência de notícias verdadeiras;
- Estamos supondo que na ausência de notícias falsas o surgimento de notícias verdadeiras tendem a desaparecer.

Embora as duas suposições, no âmbito geral, parecem não descrever sucintamente o que realmente ocorre nos meios de comunicação, considerando uma situação particular onde para cada surgimento de notícias falsas é necessária uma intensificação de notícias verídicas, podemos entender aqui que F (população de presas) e N (população de predadores) interagem com N predando F assim essa relação e dinâmica, se mostra de forma simplista, porém plausível.

Juntamente com estas equações, acoplamos equações diferenciais baseadas no modelo SIR, que descrevem uma dinâmica envolvendo por meio da Lei da ação das massas, encontros entre as populações de susceptíveis (S), infectados (I) e recuperados (R) com a ‘população’ de notícias falsas F e notícias verdadeiras N refletindo encontros exclusivos apenas com estas duas classes, ou seja, não há como de costume nos modelos epidemiológicos, encontros entre indivíduos infectados e susceptíveis mais especificamente o processo de infecção e recuperação se dá respectivamente por meio de encontros entre notícias falsas (F) e indivíduos susceptíveis (S) e recuperação entre notícias verdadeiras (N) e indivíduos infectados (I). Tal proposta, mesmo estranha do ponto de vista epidemiológico, por meio da abstração e construção matemática, juntamente com a Lei da ação das massas é plausível. Neste sentido estamos considerando que o modelo presa-predador está isolado nesta versão e que apenas influencia da forma descrita anteriormente na dinâmica populacional envolvendo as populações S , I e R . Por uma questão óbvia, não denotaremos estas populações com os adjetivos herdados do modelo SIR por não fazer sentido mas consideramos as seguintes descrições:

- S := classe de indivíduos suscetíveis à influência de notícias falsas;
- I := classe de indivíduos influenciados pelas notícias falsas;

- R := classe de indivíduos recuperados das notícias falsas por influência das notícias verdadeiras e que não mais podem ser influenciados por notícias falsas;

Desta forma formulamos:

$$\begin{cases} \frac{dS}{dt} = -\beta SF \\ \frac{dF}{dt} = aF - bFN \\ \frac{dN}{dt} = -cN + dFN \\ \frac{dI}{dt} = \beta SF - \gamma NI \\ \frac{dR}{dt} = \gamma NI \end{cases} \quad (38)$$

5.3.1 Pontos de equilíbrio do Modelo baseado no SIR e no Lotka-Volterra (LVSIR)

Considerando o modelo 38 aqui proposto, como dito anteriormente reflete encontros entre três tipos de populações S, I, R dados respectivamente por S susceptíveis a receber a mensagem/notícia falsa, I indivíduos que por influência de notícias falsas as receberam por um período e R os indivíduos que por influência das notícias verdadeiras se recuperaram. Desta forma, duas classes populacionais descritas para este modelo, a classe de pessoas (gente) dada por S, I e R , com a seguinte condição $S + I + R = M$ onde M é a população total (pessoas) e a outra classe destinada a dinâmica de notícias verdadeiras N e falsas F . Diferentemente da outra classe esta, tirando o fato de serem não negativas, não possui uma condição fixa como $S + I + R = M$.

Sendo assim busquemos os pontos de equilíbrio do tipo (S, F, N, I, R) para este modelo. É possível perceber das equações da classe notícias, do ponto exclusivamente algébrico da segunda e terceira equações que se $F = 0$ teremos $N = 0$ e vice-versa o que nos fornece dois tipos de pontos de equilíbrio que são da forma $P_1 = (S^*, 0, 0, I^*, M - S^* - I^*)$ e sendo não nulos temos necessariamente $S = I = 0$ e assim temos um outro tipo de ponto de equilíbrio $P_2 = (0, \frac{c}{d}, \frac{a}{b}, 0, M)$. Desta forma temos o ponto de equilíbrio livre de doença como caso particular de P_1 , à saber, $P_0 = (S^*, 0, 0, 0, M - S^*)$, os pontos gerais do tipo de P_1 com $I^* > 0$ e outro tipo P_2 que reflete a dinâmica cíclica do modelo Lotka-Volterra.

A matriz jacobiana deste modelo será da forma:

$$J = \begin{bmatrix} -F\beta & -S\beta & 0 & 0 & 0 \\ 0 & -Nb + a & -Fb & 0 & 0 \\ 0 & Nd & Fd - c & 0 & 0 \\ F\beta & S\beta & -I\gamma & -N\gamma & 0 \\ 0 & 0 & I\gamma & N\gamma & 0 \end{bmatrix}$$

Desta forma temos

$$J(P_0) = \begin{bmatrix} 0 & -S^*\beta & 0 & 0 & 0 \\ 0 & a & 0 & 0 & 0 \\ 0 & 0 & -c & 0 & 0 \\ 0 & S^*\beta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

forneendo como autovalores $\lambda_0 = a > 0$, $\lambda_1 = 0$ com multiplicidade igual a 3 e $\lambda_2 = -c < 0$ o que nos mostra que neste ponto de equilíbrio o sistema é instável. De forma análoga, considerando $I^* > 0$ temos que

$$J(P_1) = \begin{bmatrix} 0 & -S^*\beta & 0 & 0 & 0 \\ 0 & a & 0 & 0 & 0 \\ 0 & 0 & -c & 0 & 0 \\ 0 & S^*\beta & -I^*\gamma & 0 & 0 \\ 0 & 0 & I^*\gamma & 0 & 0 \end{bmatrix}$$

e esta possui os mesmos autovalores $\lambda_0 = a > 0$, $\lambda_1 = 0$ com multiplicidade igual a 3 e $\lambda_2 = -c < 0$ o que nos mostra que para qualquer ponto da forma de P_1 será um ponto de equilíbrio instável. Por outro lado, considerando os pontos de equilíbrio da forma de P_2 temos:

$$J(P_2) = \begin{bmatrix} -\frac{\beta c}{d} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{bc}{d} & 0 & 0 \\ 0 & \frac{ad}{b} & 0 & 0 & 0 \\ \frac{\beta c}{d} & 0 & 0 & -\frac{a\gamma}{b} & 0 \\ 0 & 0 & 0 & \frac{a\gamma}{b} & 0 \end{bmatrix}$$

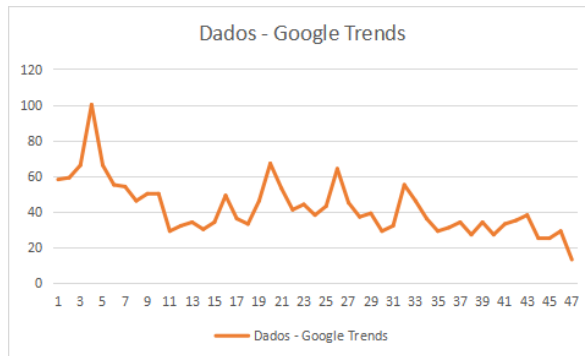
cujos autovalores são $\lambda_0 = -\frac{\beta c}{d} < 0$, $\lambda_1 = -\sqrt{-ac}$, $\lambda_2 = \sqrt{-ac}$, $\lambda_3 = -\frac{\gamma a}{b} < 0$ e $\lambda_4 = 0$, ou seja, duas reais negativas mostrando uma dinâmica de estabilidade, outros dois complexos com parte real nula, o que caracteriza como já era de se esperar a dinâmica oscilatória derivada da parte acoplada vinda do Lotka-Volterra e um nulo o mostra que o sistema não é assintoticamente estável mas que as soluções ficam próximas do ponto de equilíbrio. Desta forma podemos concluir que o sistema para neste ponto de equilíbrio é estável no sentido de Lyapunov, ou seja, começando suficientemente próximo do ponto de equilíbrio, o sistema permanece sempre próximo dele, sem necessariamente convergir para ele. É estável mas não assintoticamente.

5.3.1.1 Número de Reprodução Básico $\mathcal{R}_0$ para o modelo LVSIR

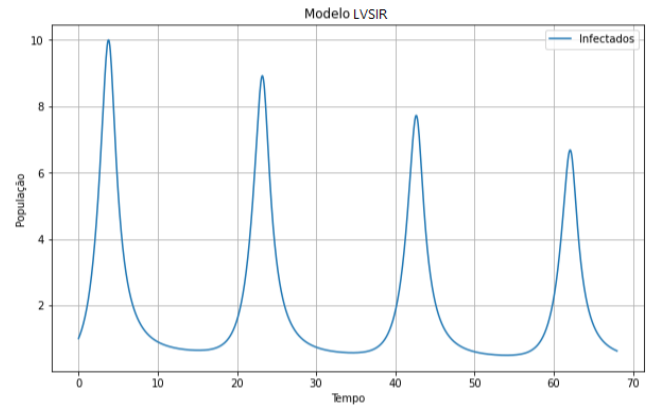
Um dos parâmetros mais importantes em modelos epidemiológicos e que permeia essa tese como referenciado na seção 3.4 é o número básico de reprodução. Para este modelo, devido ao fato da 'população' F agir de forma independente de I , não é possível definir quantas infecções secundárias em F , I gerará, ou seja, não faz sentido buscar ou definir o número de reprodução para este modelo. Mais especificamente, o termo não linear βSF por não depender de I impossibilita esta definição e abordagem.

5.3.2 Exemplo de simulação para o modelo LVSIR

Obtemos para este modelo 38, a seguinte simulação, a título de exemplo da evolução temporal do mesmo:



(a) Dados GT - Vacinação Infantil - 03/04/2022
a 19/02/2023 (semanas)



(b) Infectados - Modelo LVSIR

Figura 28 – Dados do Google Trends (Gráfico 28(a)) com potencial resposta do modelo (Ilustração 28(b)), descrevendo picos descendentes com ondas de engajamento/desengajamento.

Cujo comportamento é relativamente semelhante ao que ocorre neste caso particular obtido do Google Trends.

Capítulo 6

Discussão

Para os três modelos, Modelo tipo SIR - clássico (23), Modelo tipo SIR - Aberto com interação IR (24) e Modelo tipo SIR - Aberto sem interação IR (28), os parâmetros foram obtidos através da otimização em cada subintervalo, distinguindo os períodos de engajamento e desengajamento, conforme a suposição inicial de se tratar de períodos, de desinformação, rumores e notícias falsas, seguidos de períodos de informação mais qualificada.

Uma observação importante que devemos ter em mente é de que pela descrição de R_0 seção 2.3 valores estratosféricos algumas vezes podem ocorrer, não no sentido biológico mas no sentido da experimentação computacional e isso remete que estes não devem ser considerados puramente quantitativos e sim qualitativos visto que as variações e intervalos de buscas dos parâmetros foram escolhidos de forma aleatória buscando manter uma mínima coerência, mesmo assim, pode-se fugir dos limites biológicos e comportamentais, como são os casos de alguns parâmetros colocados nas Tabelas 2 e 8. Ainda, temos por premissa que estes modelos não são biológicos visto que a dinâmica de infecção se dá no âmbito computacional por contágio socio-computacional e que isto pode justificar estes altos valores de R_0 . Ou seja, busca-se por meio desta análise, ao ocorrer este tipo de situação, que para dentro da proposta o modelo é cabível de adaptações e que o fenômeno poderá ser descrito utilizando novas dinâmicas juntamente ao modelo SIR básico inicial. Estes valores irreais podem surgir devido a vários fatores tais como o processo de subdeterminação natural derivada da escolha apenas da curva $I(t)$ o que obviamente não permitiria recuperar as curvas populacionais $S(t)$ e $R(t)$ relacionadas ao fenômeno. O que no máximo conseguiríamos, via simulações como as aqui estudadas, seriam bons ajustes cujas diferentes combinações de parâmetros levariam a resultados parecidos. Desta forma justifica-se a análise menos quantitativa e mais qualitativa destes parâmetros e ajustes aqui obtidos.

Ao decidir pela construção via AGRP, consideramos como parte importante destas simulações e consequentemente resultados, o coeficiente de correlação de Pearson (r) como indicador inicial da qualidade de ajuste, o Critério de Informação de Akaike - AIC e dada a melhor aplicabilidade também a sua versão corrigida para amostras pequenas AICc, descritas na seção 3.3, nestes casos aqui apresentados, visto que permitem o balanço de qualidade e complexidade dos modelos, com o objetivo de parcimônia como princípio orientador, assegurando que os modelos selecionados sejam suficientemente explicativos e sem complexidade excessiva. Neste sentido, observou-se que o modelo 28 apresentou um $r = 0.9342$ superior ao do modelo 23, $r = 0.8637$ e ligeiramente superior ao modelo 24, $r = 0.9298$ o que remete uma melhor capacidade de representação da variabilidade dos dados observados, embora que o seu AICc (478.04) tenha

sido superior ao do modelo 23, (403.78), ou seja pior, o AICc deste foi substancialmente melhor (ou seja inferior) ao do modelo 24, (684.47). Vale ressaltar ainda que o valor AIC do modelo 28 também foi dos três modelos o menor, indicando que de fato, este modelo oferece uma melhoria no ajuste sem incorrer nos riscos elevados de sobreajuste (overfitting) evidenciados no modelo 24 (Tabela 9). A escolha do modelo 28 se justifica pela busca de um equilíbrio entre qualidade de ajuste e complexidade estrutural. Ou seja, demonstra-se suficientemente parcimonioso, evitando o excesso de parâmetros, e ao mesmo tempo foi capaz de capturar de forma eficaz, as nuances presentes nos dados, quando comparado com o modelo 23.

Pôde-se perceber (Tabela 8), conforme nossa hipótese inicial, que a condição de $R_0 > 1$ (embora em alguns períodos com ordem muito superior a casos reais de epidemias $R_0 \gg 1$) ocorre pelo menos uma vez em um dos períodos de análise e em todos os três modelos e método, o que endossa ao que é indicado em (CINELLI et al., 2020), que reflete que houve nos períodos estudados um processo infodêmico em torno do assunto de vacinação infantil.

Já para o modelo 32, referente a polarização social e política, considerando as três variáveis envolvidas no processo de polarização, desengajamento (μ), legitimidade (λ) e ressonância (ρ), obtivemos as figuras 25(a) e 25(b). Por meio da análise numérica temos que:

- (i) $\mu > 0 \rightarrow \gamma_+ > \gamma_- \rightarrow$, população I_- mais engajada do que I_+ o que significa que a ressonância máxima ocorrerá à direita de $\lambda = 0$ ($\lambda > 0$), ou seja as discussões serão difundidas dentro da população I_+ ;
- (ii) $\mu < 0 \rightarrow \gamma_+ < \gamma_- \rightarrow$, população I_+ mais engajada do que I_- o que significa que a ressonância máxima ocorrerá à esquerda de $\lambda = 0$ ($\lambda < 0$), ou seja as discussões serão mais difundidas dentro desta população I_- ;
- (iii) $\mu = 0 \rightarrow \gamma_+ = \gamma_- = 0.5$, e neste caso, teremos um cenário onde ambas as populações estarão engajadas com intensidades iguais e a ressonância máxima ocorrerá quando $\lambda = 0$, ou seja, quando tivermos um sistema polarizado;
- (iv) $\lambda = 0$ ou próxima de 0 temos os casos de ressonância máxima, o que descreve o comportamento das populações como esperávamos.
- (v) A variação destas máximas ressonâncias estão ocorrendo em torno de $\lambda = 0$, o que é um indicador de que o sistema está mais polarizado quando a legitimidade é aproximadamente nula. Figuras 25(a) e 25(b).
- (vi) No quadrante onde $\mu > 0$ e $\lambda < 0$, isto é, quando o desengajamento de I_+ for maior do que I_- e a legitimidade negativa, teremos uma relativa ressonância baixa para I_+ e alta para I_- o que nos mostra que haverá pouca discussão da mensagem por parte desta população I_+ e uma intensa discussão por parte da população I_- , invertendo estes cenários a medida que fazemos μ negativo mantendo $\lambda < 0$;

- (vii) No quadrante onde $\mu < 0$ e $\lambda > 0$, isto é, quando o desengajamento de I_+ for menor do que I_- e a legitimidade positiva, teremos uma relativa ressonância baixa para I_- e alta para I_+ o que nos mostra que haverá pouca discussão da mensagem por parte desta população I_- e uma intensa discussão por parte da população I_+ , invertendo estes cenários a medida que fazemos μ positivo mantendo $\lambda > 0$;

Por meio da experimentação computacional procuramos validar o fenômeno de que o sistema, ou relações entre essas duas classes populacionais I_+ e I_- , haverá mais ressonância, menos consenso, estabelecendo-se a polarização, quando $\lambda = 0$, ou próxima de 0. Percebemos que isso ocorrerá para $\lambda = 0$ caso o engajamento de ambas tiverem a mesma intensidade ou próximas.

Uma possibilidade diferente das que já foram ditas ao longo da tese que pode ser visto como possibilidade de trabalhos futuros é a utilização de técnicas de inteligência computacional como redes neurais artificiais para a identificação automática de subintervalos. Isto possibilitaria trabalhar com um conjunto de dados mais amplo, permitindo por meio da identificação automática detectar padrões nos dados e identificar os momentos em que ocorram mudanças significativas no processo em análise.

Outra possibilidade seria referente ao que foi exposto na [Seção 5.3](#) é a utilização de equações para descrever uma notícia falsa de uma notícia verdadeira utilizando para isso equações baseadas no modelo presa-predador e com isso procurar descrever as oscilações que ocorrem num processo como este, inserindo por exemplo no compartimento das notícias falsas F outros fatores como sugerido na [Seção 4.3](#) e outras dinâmicas oscilatórias por exemplo no compartimento de suscetível como uma fonte por exemplo.

No que se refere ao modelo destinado à análise da polarização social e política, uma de nossas intenções, e que apresentamos na [Subseção 5.2.6](#) foi de avaliar uma outra forma de descrever a função legitimidade, com o intuito de que a curva originada do modelo fosse mais próxima de uma curva mais suave contrastando com a primeira curva descritiva da função Legitimidade obtida que possui nitidamente um ponto de mudança de inclinação abrupta. E como objetivo maior, empregá-lo utilizando conjuntos de dados reais em futuras pesquisas, para avaliar sua eficácia e sua capacidade de identificar tendências e padrões nesses contextos complexos de polarização envolvendo populações antagônicas, tão real nos contextos contemporâneos.

Capítulo 7

Conclusão

Por meio do modelo SIR em sua forma mais simples, pôde-se descrever de forma satisfatória o processo infodêmico considerando este estudo de caso em particular (Figura 4) e isso pode servir para analisar o caráter epidêmico em torno da divulgação de informações. A adoção de hipóteses envolvendo a partição do tempo de observação deste processo complexo em subintervalos e períodos de engajamento e desengajamento mostrou-se necessário e eficaz. Também é uma interpretação sensata do que realmente acontece: uma mensagem polêmica torna-se viral e é seguida de um processo de informação qualificada. Ao considerar as generalizações do modelo SIR, modelos 24 e 28, percebemos a possibilidade juntamente com o método AGRP de descrever o fenômeno de engajamento e desengajamento por períodos de forma satisfatória, pela análise visual da curva de aproximação juntamente com o coeficiente de correlação (r), o Critério de Akaike (AIC), o Critério de Akaike Corrigido (AICc), os gráficos de convergência (apêndice C) e os parâmetros estatísticos (apêndice D), onde de forma equilibrada e parcimoniosa, surge como novidade o modelo 28 sendo um elo entre o modelo epidemiológico clássico SIR e outro modelo epidemiológico com muitos parâmetros cuja população não é fechada. Um fato relevante é a obtenção em alguns períodos do número de reprodução básico $R_0 \gg 1$ o que nos leva a concluir que é necessário um refinamento destes modelos ou até mesmo a alteração destes e/ou do método escolhido para estudar tais fenômenos se lançarmos mão de processos epidemiológicos mais realistas.

Já o modelo 32 que descreve a polarização entre duas populações antagônicas sugere o comportamento que esperávamos Figuras 22, 23 e 24, onde a evolução de uma população sendo superior a da outra, em determinado momento se tornam significativas descrevendo uma polarização entre as populações.

Por outro lado, o modelo 38 possui particularidades muito interessantes visto que é a articulação entre dois modelos clássicos que são: o modelo base de toda essa pesquisa que é o modelo epidemiológico SIR (23) e o modelo Lotka-Volterra (11), o que permite uma análise mais abrangente de ciclos de infecção e de recuperação, o que normalmente ocorre em processos infodêmicos.

Desta forma conclui-se que, embora estes estudos tenham contribuído para a compreensão da dinâmica informacional de temas ressoantes e que apresentaram engajamentos e desengajamentos caracterizados nas curvas do Google Trends, os fatos das amostras serem limitadas e os modelos construídos apresentarem resultados que quantitativamente fogem um pouco do que normalmente é encontrado na literatura, os modelos aqui apresentados poderão contribuir com

importantes *insights* para abordagens futuras, bem como outros desdobramentos e intervenções eficazes. Assim, mesmo que existam outras maneiras pelas quais a metodologia deste trabalho pode ser refinada, o método científico utilizado e apresentado, é um instrumento eficiente de análise social e consequentemente verificação emergente de Infodemias.

Referências

- BAPTISTA, J. P.; GRADIM, A. A working definition of fake news. **Encyclopedia**, MDPI, v. 2, n. 1, 2022. Citado na página 13.
- BASSANEZI, R. C. Ensino-aprendizagem com modelagem matemática. 2ª reimpressão são paulo. **Contexto**, 2010. Citado na página 25.
- BRAUER, F. Mathematical epidemiology: Past, present, and future. **Infectious Disease Modeling**, Elsevier, v. 2, n. 2, p. 113–127, 2017. Citado na página 21.
- BURNHAM, K. P.; ANDERSON, D. R. Multimodel inference: understanding aic and bic in model selection. **Sociological methods & research**, Sage Publications SAGE CA: Thousand Oaks, CA, v. 33, n. 2, p. 261–304, 2004. Citado na página 36.
- CASTRO, L. N. de et al. Computação natural: Uma breve visão geral. In: **Workshop em nanotecnologia e Computação Inspirada na Biologia**. [S.l.: s.n.], 2004. Citado na página 43.
- CHAGAS, J. S. das. O direito de matar a partir do pensamento de peter singer. **Revista Vertentes do Direito**, v. 3, n. 2, p. 75–93, 2016. Citado na página 53.
- CINELLI, M. et al. The covid-19 social media infodemic. **Scientific reports**, Nature Publishing Group, v. 10, n. 1, p. 1–10, 2020. Citado 3 vezes nas páginas 15, 18 e 71.
- CNN. **Congreso de Espana Aprueba la Regularizacion de la Eutanasia**. 2021. Disponível em: <<https://cnnespanol.cnn.com/2021/03/18/congreso-de-espana-aprueba-la-regularizacion-de-la-eutanasia/>><https://cnnespanol.cnn.com/2021/03/18/congreso-de-espana-aprueba-la-regularizacion-de-la-eutanasia/>. Citado na página 54.
- COLOMBO, C.; DIAMANTI, M. The smallpox vaccine: The dispute between bernoulli and d’alembert and the calculus of probabilities. **Lettera Matematica**, Springer, v. 2, n. 4, p. 185–192, 2015. Citado na página 21.
- CONTRERAS, S. et al. Real-time estimation of r t for supporting public-health policies against covid-19. **Frontiers in public health**, Frontiers Media SA, v. 8, p. 556689, 2020. Citado na página 23.
- DODGE, R. E. F. Eutanásia-aspectos jurídicos. **Revista Bioética**, v. 7, n. 1, 2009. Citado na página 54.
- DRIESSCHE, P. Van den. Reproduction numbers of infectious disease models. **Infectious disease modelling**, Elsevier, v. 2, n. 3, p. 288–303, 2017. Citado 2 vezes nas páginas 36 e 58.
- FERREIRA, A. G. C.; CAREGNATO, S. E. Visibilidade de revistas científicas: um estudo no portal de periódicos científicos da universidade federal do rio grande do sul. **TransInformação**, SciELO Brasil, v. 26, p. 177–190, 2014. Citado na página 14.
- FIGUEIREDO, T. da C. et al. Capítulo sistemas complexos. **Introdução à Complexidade do movimento Humano**, p. 17, 2020. Citado na página 53.
- FORTE, R. S. M. C. **O discurso do dumping social no sistema comercial internacional**. [S.l.]: Editora Dialética, 2021. Citado na página 19.

- FORTIN, F.-A. et al. Deap: Evolutionary algorithms made easy. **The Journal of Machine Learning Research**, JMLR, org, v. 13, n. 1, p. 2171–2175, 2012. Citado 2 vezes nas páginas 34 e 43.
- FRENCH, J. R.; RAVEN, B. et al. The bases of social power. **Studies in social power**, Ann Arbor, MI, v. 150, p. 167, 1959. Citado na página 14.
- GIANNINI, L. Mensagens compartilhadas: Uma análise teórica do processo de construção, compreensão e propagação de mensagens em redes sociais. *Revista temática*, 2021. Citado 2 vezes nas páginas 54 e 64.
- GOV.BR. **Consulta Pública - Inclusão de crianças de 5 a 11 anos na campanha de vacinação contra a Covid-19**. <https://www.gov.br/saude/pt-br/aceso-a-informacao/participacao-social/consultas-publicas/2021/consulta-publica-para-esclarecimentos-quanto-a-vacinacao-de-criancas-de-5-a-11-anos>. Acesso em: 13 mar 2024: [s.n.], 2021. Citado na página 28.
- GRANCHI, B. B. G. **Vacinação de crianças no Brasil: o que sabemos até agora**. <https://www.bbc.com/portuguese/brasil-59999518>. Acesso em: 13 mar 2024: [s.n.], 2022. Citado na página 28.
- HAM, C. V.; ELSAS, E. van. When legitimacy becomes the object of politics: the politicization of political support in european democracies. **Frontiers in Political Science**, Frontiers Media SA, v. 6, p. 1363083, 2024. Citado na página 55.
- JOHNSON, C.; DOWD, T. J.; RIDGEWAY, C. L. Legitimacy as a social process. **Annu. Rev. Sociol.**, Annual Reviews, v. 32, p. 53–78, 2006. Citado 2 vezes nas páginas 14 e 19.
- KLEIN, E. **Why we're polarized**. [S.l.]: Simon and Schuster, 2020. Citado na página 19.
- KOOPMANS, R. Movements and media: Selection processes and evolutionary dynamics in the public sphere. **Theory and society**, Springer, v. 33, n. 3-4, p. 367–391, 2004. Citado 5 vezes nas páginas 13, 14, 19, 20 e 64.
- LIPPITT, R.; POLANSKY, N.; ROSEN, S. The dynamics of power: A field study of social influence in groups of children. **Human Relations**, Sage Publications Sage CA: Thousand Oaks, CA, v. 5, n. 1, p. 37–64, 1952. Citado na página 14.
- MIGUEL, L. F. Aborto e democracia. **Revista Estudos Feministas**, SciELO Brasil, v. 20, p. 657–672, 2012. Citado na página 54.
- MYUNG, I. J. Tutorial on maximum likelihood estimation. **Journal of mathematical Psychology**, Elsevier, v. 47, n. 1, p. 90–100, 2003. Citado na página 34.
- NUNES, F.; TRAUMANN, T. **Biografia do abismo**. [S.l.]: HARLEQUIN, 2023. Citado na página 19.
- NUTI, S. V. et al. The use of google trends in health care research: a systematic review. **PloS one**, Public Library of Science San Francisco, USA, v. 9, n. 10, p. e109583, 2014. Citado na página 15.
- PACKER, A. L.; MENEGHINI, R. Visibilidade da produção científica. **Comunicação & produção científica: contexto, indicadores e avaliação**. São Paulo: Angellara, p. 237–259, 2006. Citado na página 14.

- POZO, A. et al. Computação evolutiva. **Universidade Federal do Paraná, 61p.(Grupo de Pesquisas em Computação Evolutiva, Departamento de Informática-Universidade Federal do Paraná)**, 2005. Citado na página 43.
- RAVEN, B. H. Legitimate power, coercive power, and observability in social influence. **Sociometry**, JSTOR, v. 21, n. 2, p. 83–97, 1958. Citado na página 14.
- REKKER, R. The nature and origins of political polarization over science. **Public Understanding of Science**, SAGE Publications Sage UK: London, England, v. 30, n. 4, p. 352–368, 2021. Citado na página 13.
- ROCHA, C.; SILVA, M. Experiência social e ressonância cibernética: juventude e a onipresença na rede. **Rocha, C.; Santaella, L. A onipresença nas redes**, p. 13–29, 2015. Citado na página 14.
- SILVA, L. d. S. B. et al. Otimização mono-objetivo no controle do mosquito aedes aegypti por meio de um modelo de duas populações com influência da precipitação. **TEMA (São Carlos)**, SciELO Brasil, v. 20, p. 197–216, 2019. Citado 2 vezes nas páginas 32 e 43.
- SIMOS, T. E. et al. Real-time estimation of r_0 for covid-19 spread. **Mathematics**, MDPI, v. 9, n. 6, p. 664, 2021. Citado na página 57.
- SIQUEIRA, J. C. A noção do termo "informação": revisão de literatura. **Brazilian Journal of Information Science**, Universidade Estadual Paulista, <https://dialnet.unirioja.es/servlet/articulo?codigo=3840268> - Acessado em 24 de abr. de 2024, v. 5, n. 1, p. 69–92, 2011. Citado na página 12.
- SKAZA, J.; BLAIS, B. Modeling the infectiousness of twitter hashtags. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 465, p. 289–296, 2017. Citado na página 15.
- SORJ, B. et al. Sobrevivendo nas redes: guia do cidadão. **São Paulo: Plataforma Democrática**, 2018. Citado na página 14.
- SOUSA, E. Da importância da legitimidade política (i). **Análise Psicológica**, Instituto Superior de Psicologia Aplicada, v. 13, p. 385–394, 1995. Citado na página 14.
- TAKAHASHI, R. H. Otimização escalar e vetorial. **Notas de aula**, p. 309, 2007. Citado 3 vezes nas páginas 33, 34 e 43.
- THEY, N. H. Uma breve linha do tempo. **UFRGS Litoral**, 2020. Citado na página 12.
- VOSOUGHI, S.; ROY, D.; ARAL, S. The spread of true and false news online. **science**, American Association for the Advancement of Science, v. 359, n. 6380, p. 1146–1151, 2018. Citado 2 vezes nas páginas 13 e 18.
- WEBER, M. **Economy and society: An outline of interpretive sociology**. [S.l.]: University of California press, 1978. v. 2. Citado 2 vezes nas páginas 14 e 19.
- (WHO), W. H. O. et al. Understanding the infodemic and misinformation in the fight against covid-19. **Retrieved March**, v. 10, p. 2021, 2020. Citado na página 18.
- ZAROCOSTAS, J. How to fight an infodemic. **The lancet**, Elsevier, v. 395, n. 10225, p. 676, 2020. Citado na página 18.

ZHAO, Z. et al. Fake news propagates differently from real news even at early stages of spreading. **EPJ data science**, Springer Berlin Heidelberg, v. 9, n. 1, p. 7, 2020. Citado 2 vezes nas páginas 13 e 18.

Apêndices

APÊNDICE A

Artigo publicado

Infodemia: abordagem da epidemia de desinformação com modelos epidemiológicos

Infodemic: approaching the disinformation epidemic with epidemiological models

Infodemia: aproximación a la epidemia de desinformación con modelos epidemiológicos

DOI: 10.55905/oelv22n4-156

Originals received: 03/15/2024

Acceptance for publication: 04/01/2024

Júlio César de Jesus Onofre

Doutorando do Programa de Pós-Graduação em Modelagem Matemática e Computacional (PPGMMC)

Instituição: Centro Federal de Educação Tecnológica de Minas Gerais (CEFETMG)

Endereço: Belo Horizonte, Minas Gerais, Brasil

E-mail: julio.onofre@gmail.com

José Luiz Acebal Fernandes

Doutor em Física área de concentração em Teoria de Partículas e Campos pela Universidade Federal de Minas Gerais (UFMG)

Instituição: Programa de Pós-Graduação em Modelagem Matemática e Computacional do Centro Federal de Educação Tecnológica de Minas Gerais (PPGMMC – CEFETMG)

Endereço: Belo Horizonte, Minas Gerais, Brasil

E-mail: acebal@cefetmg.br

RESUMO

De acordo com a Organização Mundial da Saúde (OMS), o fenômeno social contemporâneo caracterizado pela disseminação de informações em excesso, sem critérios, por boatos notícias falsas que dificultam o acesso às fontes confiáveis é chamado de *infodemia*. Esta denominação refere-se ao fato de ser possível identificar diversos elementos da fenomenologia das epidemias usuais no fenômeno moderno de disseminação de informações, principalmente em temas de saúde pública. Neste trabalho, exploramos esta relação para descrever através do modelo epidemiológico SIR a intensa polêmica ocorrida durante a pandemia de Covid-19, em torno do tema “vacinação infantil”. Os dados para o estudo correspondem à série temporal das frequências de buscas sobre o tema no *Google Search* obtidos através da plataforma *Google Trends*. A metodologia de estudo possibilita o uso dos parâmetros típicos das epidemias para caracterizar processos de comunicação social, possibilitando um estudo quantitativo sobre o engajamento ao debate em casos polêmicos e a conscientização da sociedade por fontes de informação segura.

Palavras-chave: Infodemia, Desinformação, Epidemia, Modelo SIR, Google Trends, Vacinação Infantil.

ABSTRACT

According to the World Health Organization (WHO), the contemporary social phenomenon characterized by the excessive dissemination of information without criteria, rumors, and fake news that hinder access to reliable sources is called "infodemics." This denomination refers to the fact that it is possible to identify several elements of the phenomenology of regular epidemics in the modern phenomenon of information dissemination, especially in the context of public health. In this work, we explore this relationship to describe, through the SIR epidemiological model, the intense controversy that occurred during the Covid-19 pandemic regarding the topic of 'childhood vaccination'. The data for the study correspond to the time series of search frequencies on the topic in *Google Search* obtained through the *Google Trends* platform. The study methodology enables the use of typical epidemic parameters to characterize social communication processes, allowing for a quantitative study of engagement in controversial cases and society's awareness of reliable sources of information.

Keywords: Infodemic, Misinformation, Epidemic, SIR Model, Google Trends, Childhood Vaccination.

RESUMEN

Según la Organización Mundial de la Salud (OMS), el fenómeno social contemporáneo caracterizado por la difusión de información excesiva, sin criterios, por rumores y noticias falsas que dificultan el acceso a fuentes confiables se llama infodemia. Este nombre se refiere al hecho de que es posible identificar diversos elementos de la fenomenología de las epidemias que son comunes en el fenómeno moderno de la difusión de información, principalmente en temas de salud pública. En este trabajo exploramos esta relación para describir a través del modelo epidemiológico SIR la intensa polémica que se produjo durante la pandemia de COVID-19, en torno al tema "vacunación infantil". Los datos del estudio corresponden a las series temporales de frecuencias de búsqueda de Google Search obtenidas a través de Google Trends. La metodología de estudio permite utilizar los parámetros típicos de las epidemias para caracterizar los procesos mediáticos, posibilitando un estudio cuantitativo sobre la participación del debate en casos controvertidos y la conciencia de la sociedad por parte de fuentes de información segura.

Palabras clave: Infodemia, Desinformación, Epidemia, Modelo SIR, Tendencias Google, Vacunación infantil.

1 INTRODUÇÃO

O desenvolvimento recente da tecnologia digital que garantiu o acesso amplo da sociedade à internet nas últimas décadas trouxe inegáveis benefícios, porém, trouxe também preocupações a respeito da qualidade e da rápida e intensa exposição à informação disseminada. No cenário de intensa disseminação, as informações sem critérios rigorosos, com interesses políticos, econômicos, ou ainda, com interesse em apenas de atrair audiência têm causado prejuízos às sociedades. Este prejuízo torna-se dramático quando o tema é a saúde pública. A baixa qualidade da informação manifesta-se de diversos modos, tais como: no excesso de informação, na confiabilidade, no caráter tendencioso, na forma de boatos e notícias falsas etc.

De acordo com a Organização Mundial da Saúde (OMS), o fenômeno social caracterizado por um excesso de informações, que dificultam o acesso às fontes confiáveis e às diretrizes de políticas públicas é chamado de infodemia. A expressão faz referência à capacidade de disseminação de informação inadequada em temas de saúde pública e sugere uma analogia destes processos informacionais com as epidemias (World Health Organization et al, 2020; Zaracostas, 2020; Gupta et al, 2022). Há, contudo, a possibilidade de se generalizar o conceito, estendendo-o para temas que ultrapassam assuntos de saúde pública. De fato, a sociedade vem sofrendo verdadeiro contágio por informações de baixa qualidade em diversos assuntos. Este fenômeno tem produzido extremismo, polarização, intolerância e violência. Assim, pode-se entender que a própria sociedade esteja adoecida pelo processo informacional. Portanto, de forma indistinta da natureza da informação, podemos identificar o processo como uma questão de saúde pública.

O cenário para o surgimento do conceito atual de infodemia foi a pandemia da Covid-19 que resultou em importantes impactos para a população global, acarretando enormes perdas socioeconômicas. Diante daquele quadro, pesquisadores em todo o mundo têm empenhado esforços para mitigar o impacto e prevenir a disseminação do vírus. Nos diversos países, o período crítico compreendido entre o surgimento da

epidemia até os resultados mais sensíveis das campanhas de vacinação foi de enorme engajamento das pessoas às mídias sociais.

Em meio às discussões sobre possíveis tratamentos para a Covid-19, as primeiras vacinas surgiram em menos de um ano. Rapidamente, a eficácia dos tratamentos e das próprias vacinas tornaram-se objetos de intensos debates marcados por desinformação e politização. No processo, houve intensa disseminação de notícias falsas por meio de redes sociais, até mesmo veículos de mídia tradicional contribuíram com a divulgação de notícias sem critérios e pautadas pelo sensacionalismo. O processo de informação e desinformação deu-se de modo análogo à disseminação de doenças em epidemias, formando uma epidemia informacional dentro da pandemia, o que contribuiu para o agravamento do quadro da pandemia de Covid-19.

O primeiro modelo em epidemiologia matemática reconhecido se deveu a Daniel Bernoulli, publicado em 1760. Seu modelo estudava a vantagem relativa ao risco de se fazer inoculação contra a varíola (Brauer, 2017). A técnica de inoculação tratava-se de uma 'vacinação rudimentar' através do contágio intencional com uma cepa branda da varíola. A "variolação", portanto, buscava produzir imunidade com risco de complicações e morte minorados. Na evolução subsequente dos modelos epidemiológicos, pode-se destacar a contribuição de P.D. En'ko (En'ko, 1989) (artigo original em 1889), seguida das contribuições de Sir R.A. Ross, W. H. Hamer, A.G. McKendrick e W.O. Kermack e J. Brownlee, entre 1900 e 1935, que desenvolveram a chamada abordagem por 'modelos compartimentais'. Esta abordagem consiste da partição da população em subconjuntos dinâmicos como de indivíduos 'suscetíveis' ao contágio, 'infectados' e 'recuperados' a interagir entre si. W. H. Hamer, em 1906, propõe o princípio de ação de massas que supõe que o número de novas infecções é proporcional aos produto do número $S(t)$ de indivíduos suscetíveis pelo número de indivíduos infectados $I(t)$. Surge assim o modelo SIR, um modelo compartimental de três subpopulações: suscetíveis, infectados e recuperados que é atribuído a Kermack e McKendrick e a Hamer e Ross. De fato, a forma moderna do modelo SIR se deve em conjunto às contribuições deste grupo de cientistas, no início do Século XX.

Pode parecer intrigante a possibilidade de se compreender o fenômeno infodêmico a partir de modelos desenvolvidos para epidemias, cujos agentes etiológicos mais comuns são de natureza biológica. Contudo, é possível identificar todos os elementos da fenomenologia das epidemias nos processos informacionais modernos. Na infodemia, o agente etiológico pode ser identificado como a informação de baixa qualidade, falsa, *fakenews*, ou até excessiva que trataremos genericamente como “informação desqualificada”¹. O “contágio” se dá através de contatos em mídias sociais, trocas de mensagens e veículos de informação tradicionais ou não. Podemos caracterizar o contágio pelo despertar do interesse em relação a um assunto polêmico através da troca de informação em uma “experiência de comunicação”.

Uma série de fenômenos observados mais recentemente concorrem para estabelecer esta interpretação epidêmica dos fenômenos comunicacionais.

A eficiência na disseminação representada na reprodutividade basal R_0 guarda relação com resultados de estudos que apontam que notícias falsas e informações imprecisas têm disseminação mais rápida e ampla do que as notícias baseadas em fatos (Vosoughi et. al, 2018; Zhao et. al, 2020). Além disso, a eficiência da disseminação depende do algoritmo de seleção e entrega de conteúdos usado pelas plataformas, que tendem a entregar informações sobre temas sobre os quais os usuários mostraram interesse previamente (Cinelli et. al, 2020). Este mecanismo reforça o chamado “viés de confirmação”, fenômeno em que as pessoas tendem a conferir credibilidade maior às informações que chegam recorrentemente e que é potencializado quando oriundas de diferentes fontes (Skaza e Blais, 2017).

A sintomatologia da infodemia está associada a uma série de efeitos adversos observados na sociedade. O processo de confirmação sistemática de ideias às quais se tem afinidade e a rejeição de ideias antagônicas culmina no sintoma da organização da sociedade em grupos exclusivos com crenças e cosmovisões semelhantes, formando as chamadas “bolhas”. A dinâmica de troca de informação no interior das bolhas se dá na forma de ‘câmaras de eco’, reforçando o viés de confirmação e causando a polarização

¹ Por simplicidade usaremos o termo ‘informação desqualificada’ para tratar qualquer tipo de mensagem, declaração, notícia, de caráter excessivo, duvidoso, impreciso, manipulado ou literalmente falsa.

da sociedade, principalmente em relação a temas mais polêmicos como política, economia e costumes (Flaxman et. al, 2016; Min et. al, 2021). O comportamento coletivo e a polêmica favorecem outro sintoma que é o processo de superengajamento aos temas polêmicos nas redes sociais e demais meios de comunicação, retroalimentando todo o sistema. O resultado global destes fenômenos pode ser entendido como característico de uma sociedade adoecida, conduzida a extremismos devido ao processo de comunicação. A abordagem do processo infodêmico com o formalismo teórico-matemático das epidemias fornece uma série de indicadores e métodos usados nas epidemias que possibilitam o estudo e a mensuração do fenômeno de comunicação social.

A manifestação mais evidente do conjunto de fenômenos sociais envolvidos nos processos comunicacionais modernos é a chamada “polarização política” que pode ser entendido como a organização da sociedade em dois grupos manifestadamente antagonistas em relação a um tema. O processo de posicionamento parece ser bastante influenciado por afinidades às posições majoritárias nas respectivas bolhas. Estudos indicam que o processo de polarização é evolucionário ou dinâmico, na medida que a ressonância social em relação ao debate sobre o tema em disputa ocorre com aumento do engajamento, seguido de desengajamento, com evolução para posições de maior ou menor legitimidade e consenso na sociedade (Koopmans, 2004). A dinâmica pode repetir quando há reposicionamento dos grupos envolvidos ou novos fatos são trazidos ao debate.

Um caso bastante característico deste cenário dinâmico evolutivo foi o aparecimento do tema “vacinação infantil” em dezembro de 2021 durante a pandemia de Covid-19. Ao longo de três meses houve uma sucessão de períodos de intenso debate sobre a segurança da vacinação. A dinâmica detalhada mostrou períodos de aumento explosivo do engajamento, seguido de períodos de desengajamento ao assunto, à medida que a discussão atingia veículos de mídia tradicionais. Esta dinâmica de engajamento e desengajamento se repetiu algumas vezes no período quando novos fatos reacqueceram o debate.

A busca por informação sobre um assunto na página de buscas *Google Search* é um recurso universalmente usado para acessar informações de naturezas quaisquer na internet. Através desse recurso, pode-se buscar referências a um assunto de interesse em

sites corporativos, serviços de mídias sociais, bem como, diversos canais de informação que vão de veículos de mídia corporativos e independentes a plataformas de redes sociais etc. Portanto, é possível supor que durante o processo dinâmico de intenso engajamento e desengajamento em torno de algum assunto polêmico, haverá aumento proporcional das buscas na plataforma *Google* pelo tema em debate. Estas buscas, em princípio, ocorrerão de forma independente do tipo de acesso que vai se dar em seguida pelo usuário, seja a sites corporativos, mídias sociais, seja a páginas de mídias corporativas e outras. É possível supor, portanto, que o número de buscas relativas a um determinado assunto, em determinado período de tempo, se comporte como um indicador amostral do interesse grau de ressonância social sobre este tema.

A companhia *Google LCC* disponibiliza a plataforma *Google Trends* que é um serviço gratuito para registrar as séries temporais das pesquisas realizadas ao longo de um período de tempo e em um estado, região ou país determinados. A ferramenta mostra a série temporal da frequência relativa de buscas em relação a um tema como uma proporção da frequência máxima alcançada que é fixada em 100 pontos (Fig.2). Estes dados de buscas podem fornecer informações sobre o comportamento da população e fenômenos comunicacionais sociais em relação aos diversos assuntos (NUTI et. al, 2014). Através da série temporal relativa a um assunto, é possível evidenciar a dinâmica de engajamento, o pico e o desengajamento desencadeados por fatos nas mídias sociais e, posteriormente, enfraquecidos por campanhas de informação.

Processos infodêmicos já foram abordados pelo uso do modelo epidemiológico SIR sobre dados de redes sociais específicas como Twitter, Instagram, YouTube, Reddit e Gab (Cinelli et. al, 2020; Skaza e Blais, 2017). No presente trabalho, de forma mais geral, usamos a plataforma *Google Trends* sobre o caso polêmico da "vacinação infantil" que apresentou forte ressonância entre dezembro de 2021 e março de 2022 para estudar através do modelo epidemiológico SIR os processos de engajamento e desengajamento em processos infodêmicos. Na seção 2, apresentamos informações básicas sobre o modelo SIR. Os dados e métodos são discutidos na seção 3. Os resultados e discussão são fornecidos na seção 4. Na seção 5, apresentamos as conclusões.

2 O MODELO SIR

Pode-se identificar na fenomenologia dos processos infodêmicos diversos elementos presentes na fenomenologia das epidemias. O agente etiológico da infodemia é informação desqualificada. O contágio se dá em um processo de comunicação, ou seja, em um contato entre um indivíduo suscetível e um indivíduo infectado em que há efetiva troca de mensagens.

Como nas epidemias, as taxas de contágio dependem do comportamento social, do ambiente e da natureza do agente etiológico. As notícias falsas e informações imprecisas têm disseminação mais rápida e mais ampla do que as notícias baseadas em fatos (Vosoughi *et. al.*, 2018; Zhao *et. al.*, 2020). Provavelmente, o fenômeno da disseminação mais eficiente de notícias falsas se deve à predileção das pessoas por temas sensacionalistas. A eficiência na disseminação também depende de fatores ambientais no ecossistema das plataformas em que os métodos de seleção de conteúdos tendem a entregar informações sobre temas sobre os quais os usuários já mostraram interesse anteriormente, reforçando o viés de confirmação (Cinelli *et. al.*, 2020; Skaza e Blais, 2017).

Como resultado do contágio, há a formação de grupos sociais, frequentemente polarizados, organizados em bolhas em que a troca de informação se dá na forma de câmaras de eco. A troca intensa neste ambiente retro-alimenta o sistema (Flaxman *et. al.*, 2016; Min *et. al.*, 2021). Uma vez que toda a fenomenologia das epidemias está presente, é natural supor que a descrição da dinâmica do acometimento das populações possa ser feita via modelos epidemiológicos como nas epidemias. Esta abordagem tem potencial para fornecer uma série de indicadores e métodos usados nas epidemias para quantificação e estudo do fenômeno social da infodemia.

O modelo SIR é uma representação matemática da evolução de epidemias, pertencente à classe de modelos compartimentais ou populacionais. Ao longo de décadas, esse modelo e suas variações tem mostrado úteis devido à sua simplicidade e potencial para a descrição da disseminação de doenças contagiosas em populações, bem como na inferência sobre epidemias e métodos de controle. No modelo SIR, são consideradas três

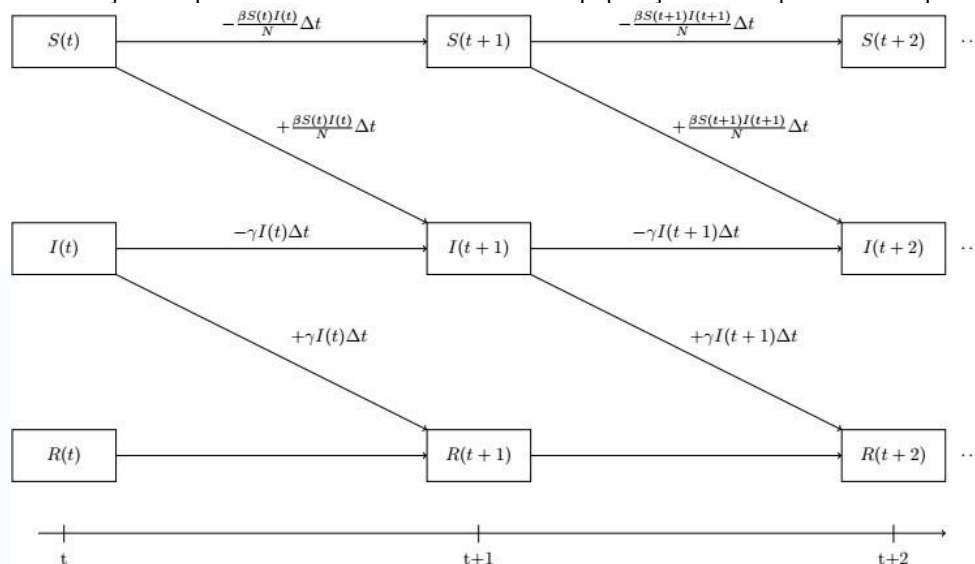
populações: $S(t)$, $I(t)$ e $R(t)$. Essas populações representam, respectivamente, os indivíduos suscetíveis, infectados e recuperados da doença. Quando o sistema é considerado fechado, a soma dessas três populações é constante, ou seja, $S(t) + I(t) + R(t) = N$, onde N representa o número total da população. A Figura 1 apresenta o modelo SIR na sua versão em tempo discreto $t \in \mathbb{Z}$. As transições dos indivíduos entre as subpopulações ou compartimentos S , I e R correspondem a frações dessas populações em um intervalo de tempo Δt . Assim, em um determinado momento t , a interação não linear da infecção ocorre entre os indivíduos das populações $S(t)$ e $I(t)$. Os valores dessas populações compõem a taxa $\frac{\beta S(t)I(t)}{N}$ de indivíduos suscetíveis que se tornam infectados no intervalo de tempo e passam para a população de infectados. O parâmetro β é proporcional à taxa de contato social multiplicada pela fração de contatos que de fato resultam em contágio. Da mesma forma, a população $I(t)$ perde uma fração de seus indivíduos com taxa $\gamma I(t)$ para a população de recuperados. O parâmetro γ é a taxa de recuperação, ou seja, é o inverso do tempo médio de recuperação da doença.

A dinâmica do modelo SIR em tempo discreto pode ser representada por meio de um sistema de equações de diferenças:

$$\begin{aligned} S(t+1) &= S(t) - \left[\frac{\beta S(t)I(t)}{N} \right] \Delta t \\ I(t+1) &= I(t) + \left[\frac{\beta S(t)I(t)}{N} - \gamma I(t) \right] \Delta t \\ R(t+1) &= R(t) + [\gamma I(t)] \Delta t \end{aligned} \quad (1)$$

O passo de tempo Δt nas equações é, em princípio, arbitrário e pode ser menor ou igual ao intervalo de tempo entre os dados a serem ajustados.

Figura 1: Diagrama esquemático do modelo SIR mostrando o fluxo de indivíduos entre compartimentos em função dos parâmetros do modelo e de suas subpopulações com o passar do tempo.



Fonte: os autores.

Além dos parâmetros do modelo, para determinar a intensidade do processo infeccioso, é comumente usado o Número de Reprodução Basal, $R_0 = \frac{I(1)}{I(0)}$, que representa o número de novos infectados em um período subsequente a partir de um único indivíduo infectado no período anterior inicial do processo epidêmico (Contreras et. al, 2020). Assim, as populações no início são aproximadamente $S(0) \approx N$, $I(0) \approx 1$ e $R(0) = 0$. Nesse caso, segue-se da equação para $I(t)$ no modelo 1 e da definição de R_0 que

$$R_0 = \frac{I(1)}{I(0)} = 1 + \left[\frac{\beta S(0)}{N} - \gamma I(0) \right] \Delta t \approx 1 + \left[\frac{\beta}{\gamma} - 1 \right] \gamma \Delta t, \quad (2)$$

em que aplicamos a condição do início da pandemia. Quando a escala de tempo é escolhida de forma que $\gamma \Delta t = 1$, temos que

$$R_0 = \frac{\beta}{\gamma} \quad (3)$$

Dessa forma, o valor desse número adimensional é expresso em relação aos parâmetros do modelo. Se os parâmetros associados são tais que $R_0 < 1$, a epidemia

descrita pelo modelo 1 tende a extinguir-se. Se $R_0 > 1$, têm-se um crescimento exponencial no número de infectados $I(t)$ na fase inicial, o que caracteriza o surgimento de uma epidemia, ou, no caso, de um processo infodêmico (Cinelli et. al, 2020). Podemos usar o valor de R_0 para estudar o impacto de notícias, informações em determinados assuntos e seu grau de ressonância social.

3 DADOS E MÉTODOS

Um estudo de caso foi conduzido para avaliar o processo infodêmico associado ao tema “vacinação infantil”, utilizando a coleta de dados do Google Trends com um filtro geográfico definido para o Brasil, no período de estudo de 01/12/2021 a 27/02/2022. Foram obtidos desde a plataforma o gráfico da série temporal da frequência das pesquisas diárias sobre esse tema e a planilha com os dados correspondentes (Figura 2). A análise inicial dos dados revelou uma certa complexidade, sem identificar inicialmente um processo epidêmico típico, mas possivelmente a sobreposição de várias ondas epidêmicas pulsantes, formando uma sequência de eventos (Figura 3).

A série temporal das frequências relativas das buscas pelo assunto 'vacinação infantil' foi interpretada como número de infectados pelo debate, correspondendo à variável dinâmica $I(t)$ do modelo SIR (Seção 2). Para estudo do processo social infodêmico através da determinação do parâmetro R_0 o modelo foi aplicado aos dados.

Não há garantias de que o número total de indivíduos N potencialmente expostos ou envolvidos com o tema “vacinação infantil” permaneça constante ao longo do período de observação. De fato, é esperado que o sistema em questão seja aberto com fluxo de indivíduos a entrar e sair. Dessa forma, trataremos o número N de pessoas envolvidas no processo como contínuo por partes, variando de um período para outro, caracterizado de acordo com os eventos de engajamento crescente ou de desengajamento. De forma análoga, consideramos que o fenômeno epidêmico do sistema aberto está exposto a contínuas alterações ambientais como novas notícias, ambiente político etc. Assim, as taxas β e γ do modelo SIR também podem variar em cada período. Esse efeito de variação dos parâmetros ao longo do tempo é esperado em epidemias reais, uma vez que ações de

controle, como distanciamento social, medidas sanitárias e vacinas e outras alterações ambientais estão sempre presentes. Com estes pressupostos, efetivamente, cada período de observação será caracterizado por uma epidemia particular.

A aplicação desta metodologia se deu com a divisão da série temporal em intervalos cujas sequências lógicas estão relacionadas a processos de engajamento (aumento das buscas) e desengajamento (diminuição das buscas) em cada período (ver Figura 3). Foram observados quatro picos ou ondas epidêmicas caracterizadas pelo aumento, seguido da diminuição do engajamento. Cada uma destas ondas epidêmicas foi dividida em dois subintervalos consecutivos: um subintervalo de engajamento com a sequência de dados de buscas crescente, seguido de um subintervalo de desengajamento com a sequência decrescente. A partição se deu conforme a Figura 3 e a Tabela 1. A observação indica que o período de engajamento é mais frequentemente marcado por exposição a informações desqualificadas com desinformação, rumores e notícias falsas e, o período de desengajamento, pelo processo de conscientização do público por meios de comunicação, evoluindo para consensos e informações com maior legitimidade com ações educacionais de informação e notícias através da mídia e agências.

Para cada subintervalo de aumento ou diminuição de engajamento, o compartimento de infectados $I(t)$ do modelo SIR foi ajustado via otimização aos dados de buscas. No processo uma sequência de parâmetros ótimos N , β e γ foi determinada para cada subintervalo. Associamos, em cada subintervalo β à taxa em que os indivíduos se engajam no assunto polêmico, frequentemente provocada pela aparição de informação desqualificada e γ , à taxa de recuperação, frequentemente associada à reação de instituições, mídias e sociedade comprometidos com informação de maior qualidade.

A implementação do modelo e do processo de otimização foi feito em linguagem Python. Para a resolução do modelo descrito como um sistema de equações diferenciais ordinárias (EDOs) foi utilizado a função *odeint* do pacote *scipy.integrate*. Foi utilizado o método Trust Region Reflective disponível no pacote *scipy.optimize* para resolver o

problema de mínimos quadrados ou seja, o problema de minimizar a soma dos quadrados das diferenças entre os dados observados e os valores previstos na curva de ajuste².

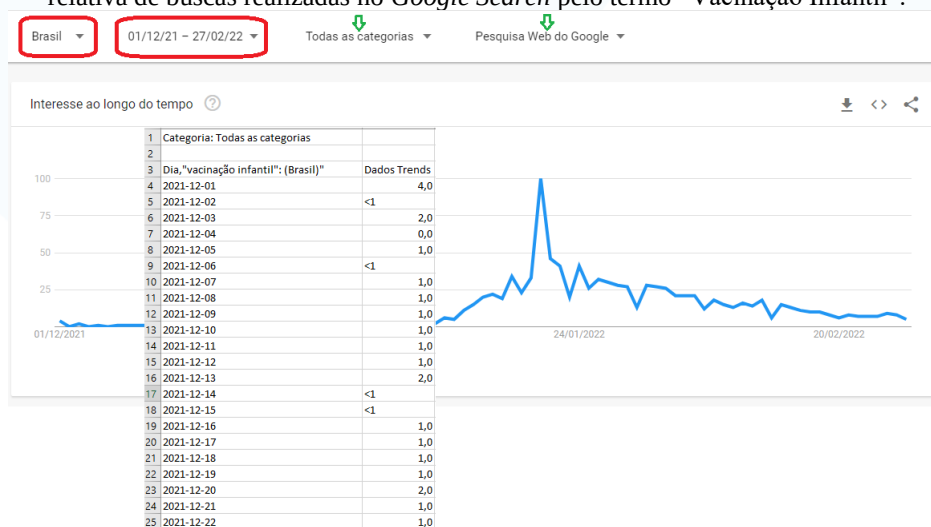
Tabela 1: Descrição da partição do período de estudo entre 01/12/2021 e 27/02/2022 em intervalos correspondentes a picos ou ondas epidêmicas, por sua vez, divididos em subintervalos de engajamento e desengajamento.

Período	Dias	Pico	Período de
P1	22° ao 24°	24/12/21 (24° dia)	Engajamento
P2	24° ao 26°	24/12/21 (24° dia)	Desengajamento
P3	26° ao 28°	28/12/21 (28° dia)	Engajamento
P4	28° ao 31°	28/12/21 (28° dia)	Desengajamento
P5	31° ao 37°	06/01/22 (37° dia)	Engajamento
P6	37° ao 40°	06/01/22 (37° dia)	Desengajamento
P7	40° ao 51°	20/01/22 (51° dia)	Engajamento
P8	51° ao 89°	20/01/22 (51° dia)	Desengajamento

Fonte: os autores

Após a determinação dos parâmetros, o número básico de reprodução R_0 de cada caso é calculado para caracterizar a intensidade com a qual o processo de infodemia se desenvolveu em cada período analisado.

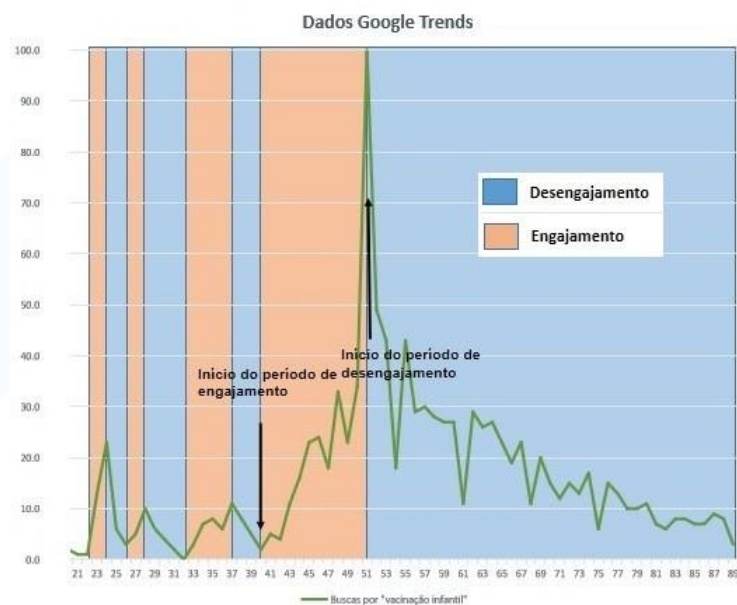
Figura 2: Imagem e planilha obtidas através da Plataforma *Google Trends* que fornece a frequência relativa de buscas realizadas no *Google Search* pelo termo "Vacinação Infantil".



Fonte: Google Trends (adaptado).

² Ambos pacotes e suas descrições estão disponíveis em (<https://scipy.org/>).

Figura 3: Série temporal obtida através dos dados do *Google Trends* referente a frequência relativa de buscas (máximo: 100) pelo termo "vacinação infantil" realizadas no *Google Search* entre 01/12/2021 e 27/02/2022.



Fonte: os autores sobre dados do Google Trends.

4 RESULTADOS

Os parâmetros do modelo SIR, N , β e γ , foram obtidos através da otimização pelo método de otimização Trust Region Reflective em cada período caracterizado por ondas epidêmicas e seus picos, distinguindo os subintervalos de engajamento e desengajamento. Em cada um desses períodos de engajamento e desengajamento, os parâmetros do modelo SIR foram otimizados sobre os dados de buscas utilizando o método de otimização Trust Region Reflective e tiveram valores de R_0 e o Erro Médio Quadrático Normalizado (NRMSE) em relação à média dos dados no período. Os valores dos parâmetros obtidos estão listados na Tabela 2.

As taxas de contágio β permaneceram na mesma ordem de grandeza nos subintervalos p3 a p7, indicando que o processo de alimentação da polêmica permaneceu ativo ao longo destes períodos. No período p1, temos em β o resultado de uma ordem de grandeza acima, indicando que talvez tenha sido o efeito explosivo inicial do surgimento do tema. Já, nos subintervalos p2 e p8 os valores de β foram várias ordens de grandeza

menores. Isto pode ser interpretado como o desligamento do processo de produção de polêmica.

Os valores da taxa de recuperação γ , entendida aqui como a evolução para consensos, teve a mesma ordem de grandeza nos períodos p1, p2 e p4 a p6, indicando um processo de conscientização social equilibrado entre estes subintervalos. No subintervalo p8 a taxa foi de uma ordem de grandeza menor. Contudo, em comparação com a taxa de contágio β do mesmo período, seu valor é de 3 ordens de grandeza maior, o que indica uma recuperação social do debate efetiva, porém de prazo longo.

Os números de indivíduos N obtido pelo método nos subintervalos devem ser interpretados de forma proporcional entre seus valores nos intervalos e não de modo absoluto. Estes valores devem ser entendidos como proporcionais à fração da sociedade que tomou conhecimento do assunto naquele subintervalo. Seus valores variaram de duas a cinco ordens de grandeza. O valor obtido no subintervalo p2 de quatro ordens de grandeza, possivelmente se deu a haver poucos pontos no intervalo e a curva de desengajamento se assemelhar a uma reta. Nos subintervalos p7 e p8 os valores de N mostram-se altos o que é compatível com o momento de máximo da amostra $I(51)=100$.

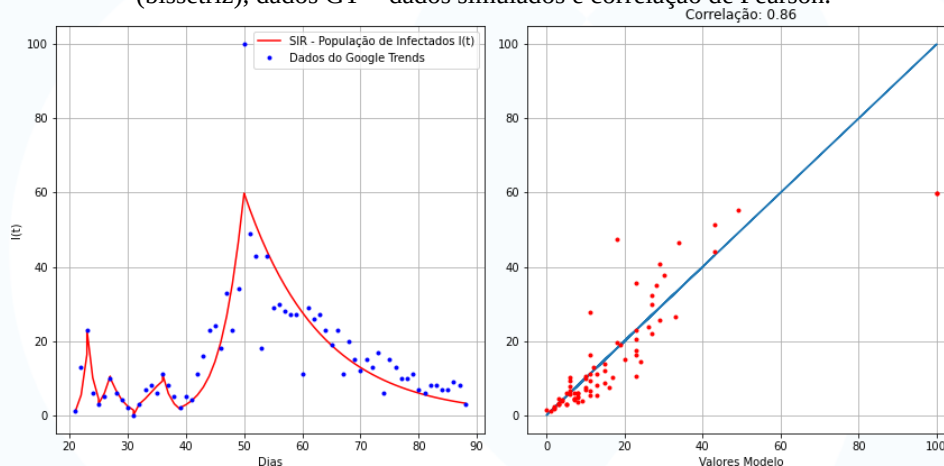
Os valores de R_0 possibilitam a compreensão do crescimento ou decrescimento do processo infodêmico considerando de forma combinada parâmetros β e γ . Os valores de R_0 variaram ao longo de várias ordens de grandeza. Por este motivo, na Tabela 2, optamos por representar estes valores em termos do $\log(R_0)$. Esta variabilidade é atípica se comparada aos casos epidemias comuns. Ela pode ser explicada por que o ambiente da disseminação de notícias é cibernético; diferentemente do ambiente físico-material, o número de contatos, o comportamento dos indivíduos e virulência dos processos digitais podem se alterar muito e contribuir para esta variabilidade em ordens de grandeza. Nos subintervalos p3 e p7 os altos valores de R_0 são explicados pela baixíssima taxa de recuperação γ , evidenciando que, nesses períodos, a disseminação de informação desqualificada não havia encontrado ainda uma reação adequada dos meios seguros de informação.

Tabela 2: Valores dos parâmetros N , β e γ otimizados em cada subintervalo de engajamento e desengajamento de ondas epidêmicas ou picos com os respectivos valores de $\log(R_0)$ e Erro Médio Quadrático Normalizado (NRMSE).

	p1	p2	p3	p4	p5	p6	p7	p8
N	43	4463	29	19	22	11	202	36705
β	2.310	9.787×10^{-4}	7.987×10^{-1}	0.882	0.643	0.654	3.399×10^{-1}	1.148×10^{-5}
γ	0.452	8.295×10^{-1}	3.757×10^{-26}	0.736	0.145	0.659	4.548×10^{-24}	7.699×10^{-2}
NRMSE	0.472	0.231	0.100	0.160	0.261	0.228	0.570	0.475
$\log R_0$	0.709	-3.921	25.33	0.077	0.648	-0.004	22.87	-4.000

Fonte: os autores.

Figura 4: (a) Curva do modelo SIR com parâmetros N , β e γ otimizados sobre os dados de frequência relativa de buscas realizadas no Google no período de estudo. (b) Gráficos dados GT $\times$ dados GT (bissetriz), dados GT $\times$ dados simulados e correlação de Pearson.



Fonte: os autores.

5 CONCLUSÕES

O modelo SIR (Suscetíveis-Infectados-Recuperados), em sua forma mais simples, foi utilizado neste estudo de caso para analisar o fenômeno social da infodemia. O modelo conseguiu descrever de maneira satisfatória o processo observado, o que fortalece a compreensão da disseminação de informações como um fenômeno epidêmico. Contudo, um ponto central neste trabalho é que não se quer demonstrar através da boa qualidade do ajuste do modelo aos dados que a interpretação epidêmica usando o modelo SIR é aplicável. Pelo contrário, a aplicabilidade dos modelos epidemiológicos ao fenômeno é um pressuposto do trabalho, devidamente argumentada no texto com base na literatura do assunto “infodemia”.

O objetivo principal foi o de usar os parâmetros obtidos pelo modelo como metodologia para mensurar os processos de comunicação social. Neste sentido, o trabalho foi exitoso. Uma etapa mais refinada desta metodologia poderá incorporar um estudo sobre as notícias contemporâneas ao período de estudo. Esses resultados reforçam a utilidade dessa abordagem no entendimento e análise da dinâmica da infodemia, bem como na importância de estratégias de prevenção e controle eficazes para lidar com a disseminação de informações enganosas ou falsas.

A abordagem feita ao dividir o tempo de observação do processo complexo em subintervalos e períodos de engajamento e desengajamento foi feita de forma baseada na experiência e observação pessoal do fenômeno comunicacional. Os resultados demonstraram que este pressuposto de divisão em subintervalos foi eficaz. Os resultados sugerem boa consistência comportamental do modelo em relação aos dados, sugerindo que este pressuposto seja uma razoável representação do fenômeno.

Existem possíveis aprimoramentos na metodologia deste trabalho que podem transformar o método em um instrumento de análise social. Uma opção é a identificação automática dos subintervalos por meio de inteligência computacional. Além disso, podem ser utilizados outros métodos de otimização, como Algoritmos Genéticos, para lidar com dados mais complexos. Essas melhorias podem aumentar a acurácia e a eficácia do método, possibilitando uma análise mais aprofundada do fenômeno em estudo e contribuindo para um entendimento mais abrangente dos processos sociais envolvidos na disseminação de informações.

REFERÊNCIAS

- BRAUER, Fred. Mathematical epidemiology: Past, present, and future. **Infectious Disease Modelling**, v. 2, n. 2, p. 113-127, 2017.
- CINELLI, Matteo et al. The COVID-19 social media infodemic. **Scientific reports**, v. 10, n. 1, p. 1-10, 2020.
- CONTRERAS, Sebastián et al. Real-time estimation of R_t for supporting public-health policies against COVID-19. **Frontiers in public health**, v. 8, p. 556689, 2020.
- EN'KO, P. D. On the course of epidemics of some infectious diseases. **International journal of epidemiology**, v. 18, n. 4, p. 749-755, 1989.
- FLAXMAN, Seth; GOEL, Sharad; RAO, Justin M. Filter bubbles, echo chambers, and online news consumption. **Public opinion quarterly**, v. 80, n. S1, p. 298-320, 2016.
- GUPTA, Ashish et al. Understanding patterns of COVID infodemic: A systematic and pragmatic approach to curb fake news. **Journal of business research**, v. 140, p. 670-683, 2022.
- KOOPMANS, Ruud. Movements and media: Selection processes and evolutionary dynamics in the public sphere. **Theory and society**, v. 33, n. 3-4, p. 367-391, 2004.
- MIN, Yong et al. The role of degree correlation in shaping filter bubbles in social networks. **Physica A: Statistical Mechanics and its Applications**, v. 584, p. 126366, 2021.
- NUTI, Sudhakar V. et al. The use of google trends in health care research: a systematic review. **PloS one**, v. 9, n. 10, p. e109583, 2014.
- SKAZA, Jonathan; BLAIS, Brian. Modeling the infectiousness of Twitter hashtags. **Physica A: Statistical Mechanics and its Applications**, v. 465, p. 289-296, 2017.
- VOSOUGHI, Soroush; ROY, Deb; ARAL, Sinan. The spread of true and false news online. **science**, v. 359, n. 6380, p. 1146-1151, 2018.
- WORLD HEALTH ORGANIZATION et al. Understanding the infodemic and misinformation in the fight against COVID-19. **Published Online at https://iris.paho.org/bitstream/handle/10665.2/52052/Factsheetinfodemic_eng.pdf**, 2020.
- ZAROCOSTAS, John. How to fight an infodemic. **The lancet**, v. 395, n. 10225, p. 676, 2020.
- ZHAO, Zilong et al. Fake news propagates differently from real news even at early stages of spreading. **EPJ data science**, v. 9, n. 1, p. 7, 2020.

APÊNDICE B

Algoritmos

Algoritmo 1: Algoritmo otimização Modelo SIR Clássico via Trust Region Reflective

Entrada: D, E, P, Θ

Resultado: Solução otimizada via Trust Region Reflective (`scipy.optimize.least_squares`)

D : Conjunto de períodos de tempo; E : Número de execuções por período; P : Conjunto de dados para simulação; Θ : Espaço de busca para o vetor de parâmetros θ .

Definir parâmetros do sistema de EDO controlados por θ ;

Definir função objetivo

$$J(\theta) = \frac{1}{m} \sqrt{\sum_{i=1}^m (y_i - u(x_i, t_i; \theta))^2}$$

onde $u(\cdot, \cdot; \theta)$ é a solução do sistema de ED;

para cada período $D_t \in D$ **faça**

para execução $e \leftarrow 1$ até E **faça**

 Inicializar chute inicial $\theta^{(0)} \in \Theta$;

$\theta^{(e)} \leftarrow \text{least_squares}(\text{função } \theta \rightarrow u(x, t; \theta), \theta^{(0)}, \text{method='trf', bounds} = \Theta)$;

 Ajustar a condição inicial para o próximo período D_t aplicando uma suavização no último vetor de estado em D_{t-1} ;

fim

fim

retorna $\theta^* = \arg \min_{\theta^{(e)}} J(\theta)$;

Algoritmo 2: Algoritmo otimização Modelo clássico SIR**Entrada:** D, G, N, E, P, Θ **Resultado:** Solução otimizada via AGRP D : Conjunto de períodos de tempo; G : Número de gerações por execução evolutiva; N :Tamanho da população; E : Número de execuções evolutivas por período; P : Conjunto de dados para simulação; Θ : Espaço de busca para o vetor de parâmetros θ .Definir parâmetros do sistema de EDO (SIR) controlados por θ ;Definir função objetivo $J(\theta) = \frac{1}{m} \sqrt{\sum_{i=1}^m (y_i - u(x_i, t_i; \theta))^2}$ onde $u(\cdot, \cdot; \theta)$ é a solução do sistema de EDO (SIR)Definir operadores genéticos: seleção $S(\cdot)$, cruzamento $C(\cdot)$, mutação $M(\cdot)$, reflexão $R(\cdot)$;**para** cada período $D_t \in D$ **faça****para** execução $e \leftarrow 1$ até E **faça**Inicializar população $Pop^{(0)} \subset \Theta$ com N indivíduos;**para** geração $g \leftarrow 1$ até G **faça****para** cada indivíduo $\theta \in Pop^{(g-1)}$ **faça**Resolver numericamente o sistema de EDOs com parâmetros θ para obter $u(x, t; \theta)$;Calcular $J(\theta)$ usando os valores de u nos pontos $\{(x_i, t_i)\}$;**fim**Selecionar pais $P_s \leftarrow S(Pop^{(g-1)})$;Gerar filhos $Child \leftarrow C(P_s)$;Aplicar mutação $Child \leftarrow M(Child)$;Refletir filhos $Child \leftarrow R(Child, \Theta)$;Formar nova população $Pop^{(g)} \leftarrow Child \cup \{\arg \min_{\theta \in Pop^{(g-1)}} J(\theta)\}$;**fim**Armazenar $\theta^{(e)} = \arg \min_{\theta \in Pop^{(G)}} J(\theta)$;Ajustar a condição inicial para o próximo período D_t pegando o último vetor de estado em D_{t-1} ;**fim****fim****retorna** $\theta^* = \arg \min_{\theta^{(e)}} J(\theta)$;

Algoritmo 3: Algoritmo otimização Modelo tipo SIR Aberto com interação IR e Modelo tipo SIR Aberto sem interação IR

Entrada: D, G, N, E, P, Θ
Resultado: Solução otimizada via AGRP

D : Conjunto de períodos de tempo; G : Número de gerações por execução evolutiva; N : Tamanho da população; E : Número de execuções evolutivas por período;

P : Conjunto de dados para simulação; Θ : Espaço de busca para o vetor de parâmetros θ .

Definir parâmetros do sistema de EDO (SIR) controlados por θ ;

Definir função objetivo $J(\theta) = \frac{1}{m+w} \sqrt{\sum_{i=1}^m (y_i - u(x_i, t_i; \theta))^2 + w^2 (y_0 - vr)^2}$

onde $u(\cdot, \cdot; \theta)$ é a solução do sistema de EDO (SIR), $w = \sqrt{m}$, $vr = \mu u_{D_{t-1}}^{\text{final}} + (1 - \mu) y_0$ com $\mu \in [0,1]$.

Definir operadores genéticos: seleção $S(\cdot)$, cruzamento $C(\cdot)$, mutação $M(\cdot)$, reflexão $R(\cdot)$;

para cada período $D_t \in D$ **faça**

para execução $e \leftarrow 1$ até E **faça**

 Inicializar população $Pop^{(0)} \subset \Theta$ com N indivíduos;

para geração $g \leftarrow 1$ até G **faça**

para cada indivíduo $\theta \in Pop^{(g-1)}$ **faça**

 Resolver numericamente o sistema de EDOs com parâmetros θ para obter $u(x, t; \theta)$;

 Calcular $J(\theta)$ usando os valores de u nos pontos $\{(x_i, t_i)\}$;

fim

 Selecionar pais $P_s \leftarrow S(Pop^{(g-1)})$;

 Gerar filhos $Child \leftarrow C(P_s)$;

 Aplicar mutação $Child \leftarrow M(Child)$;

 Refletir filhos $Child \leftarrow R(Child, \Theta)$;

 Formar nova população $Pop^{(g)} \leftarrow Child \cup \{\arg \min_{\theta \in Pop^{(g-1)}} J(\theta)\}$;

fim

 Armazenar $\theta^{(e)} = \arg \min_{\theta \in Pop^{(G)}} J(\theta)$;

 Ajustar a condição inicial para o próximo período D_t aplicando uma suavização no último vetor de estado em D_{t-1} ;

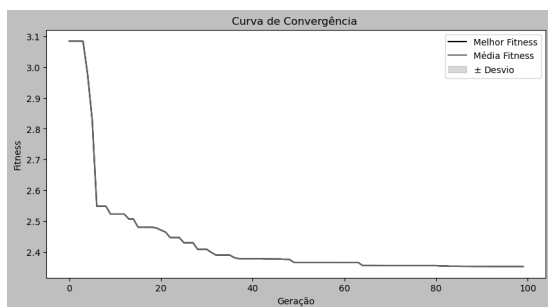
fim

fim

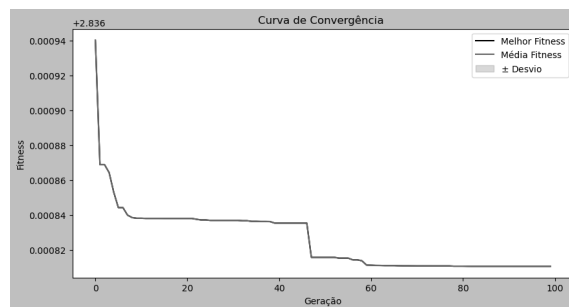
retorna $\theta^* = \arg \min_{\theta^{(e)}} J(\theta)$;

APÊNDICE C

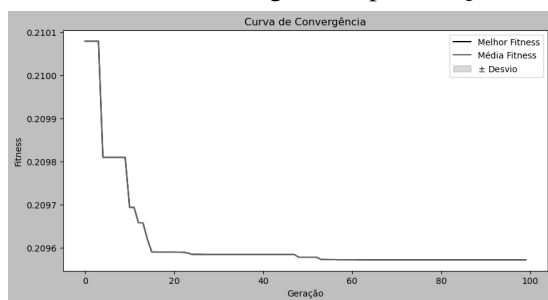
Curvas de Convergência e Boxplots



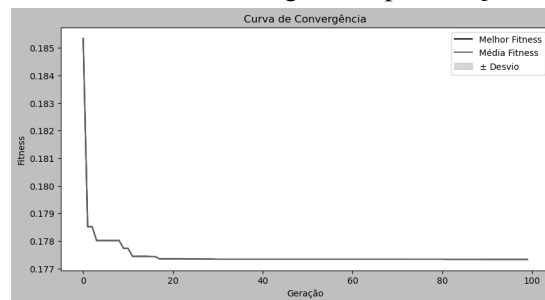
(a) Curva de Convergência - período p_1 .



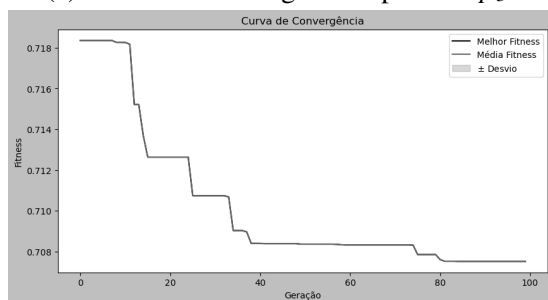
(b) Curva de Convergência - período p_2 .



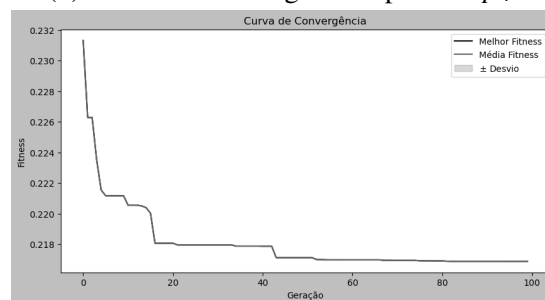
(c) Curva de Convergência - período p_3 .



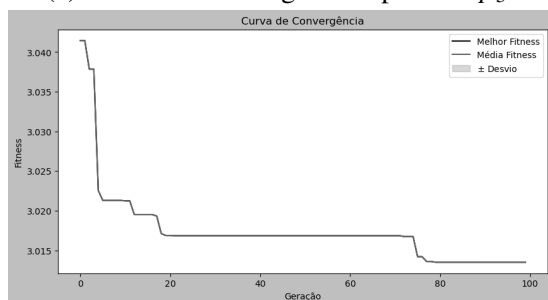
(d) Curva de Convergência - período p_4 .



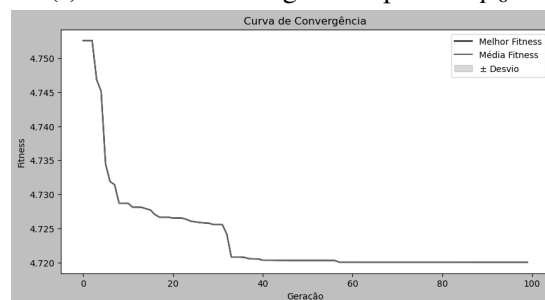
(e) Curva de Convergência - período p_5 .



(f) Curva de Convergência - período p_6 .

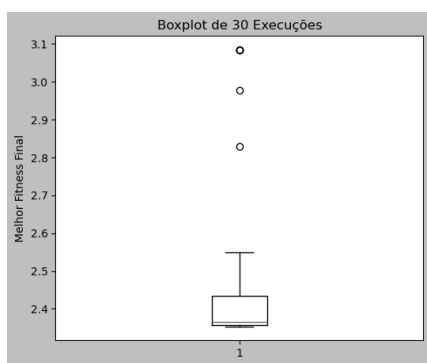


(g) Curva de Convergência - período p_7 .

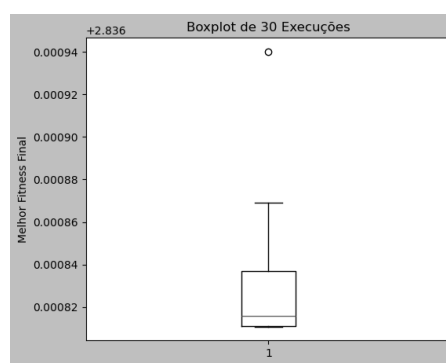


(h) Curva de Convergência - período p_8 .

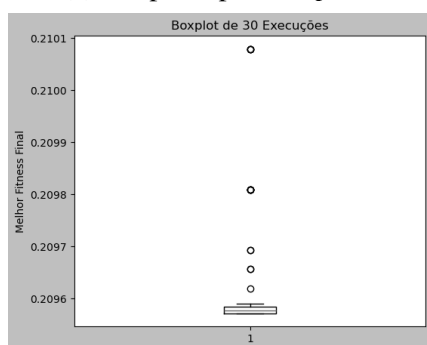
Figura 29 – Curvas de Convergência referente ao modelo 23



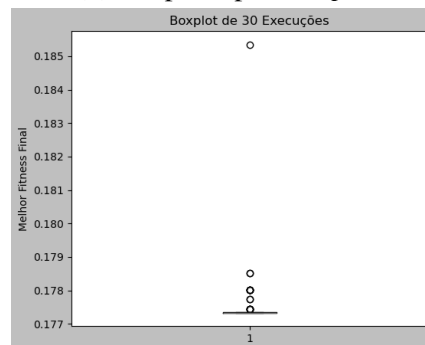
(a) Boxplot - período p_1 .



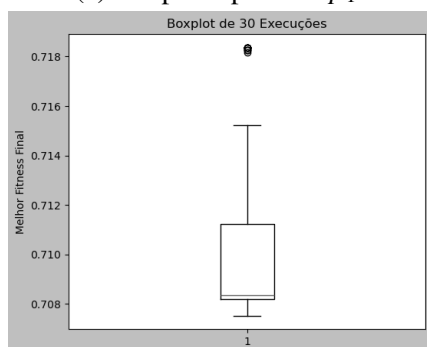
(b) Boxplot - período p_2 .



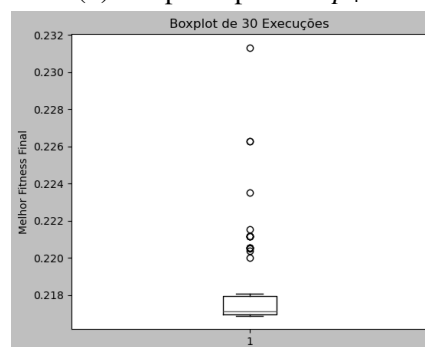
(c) Boxplot - período p_1 .



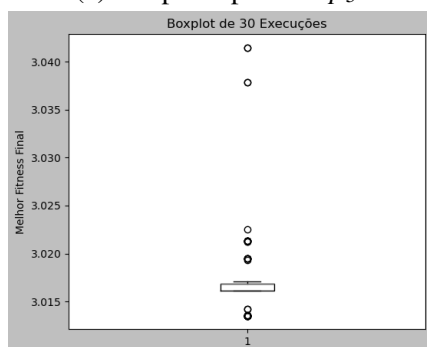
(d) Boxplot - período p_4 .



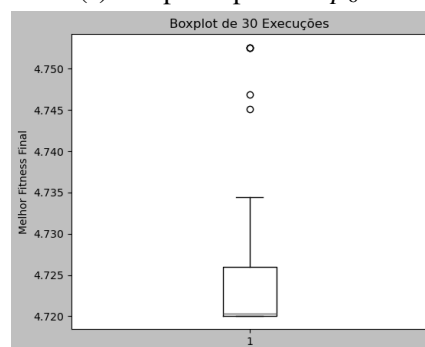
(e) Boxplot - período p_5 .



(f) Boxplot - período p_6 .



(g) Boxplot - período p_7 .



(h) Boxplot - período p_8 .

Figura 30 – Boxplots referentes ao modelo 23

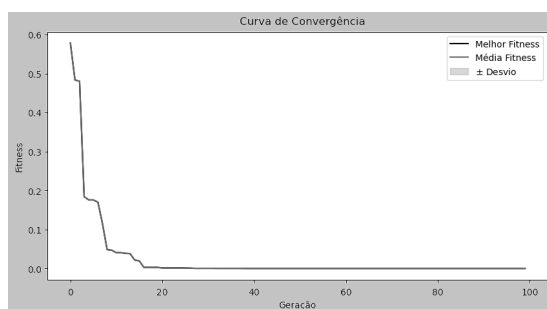
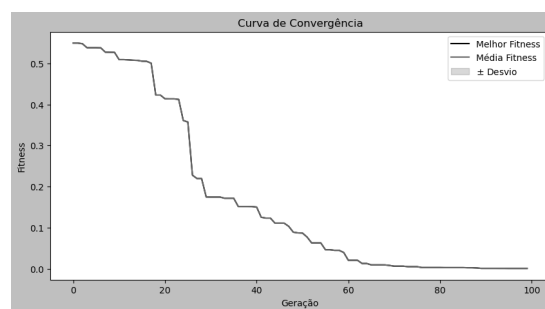
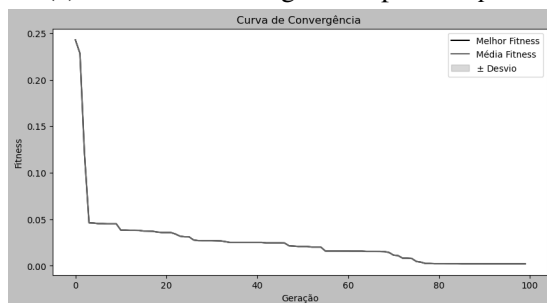
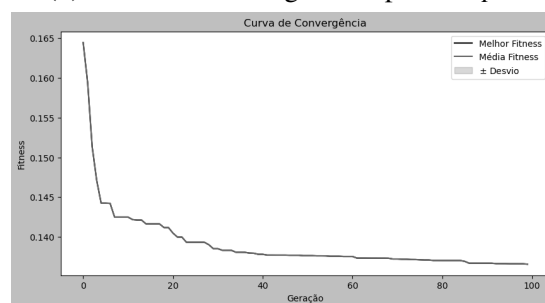
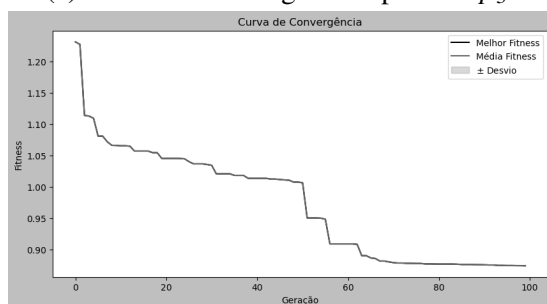
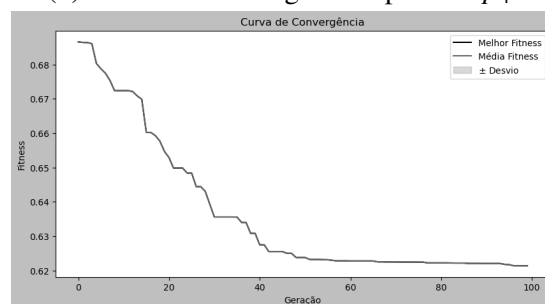
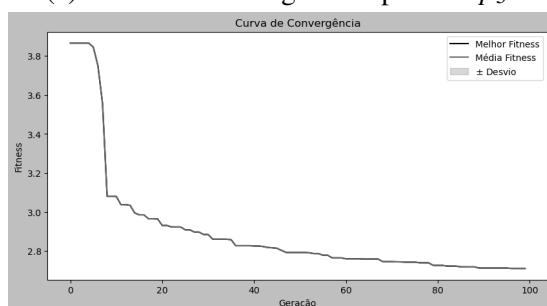
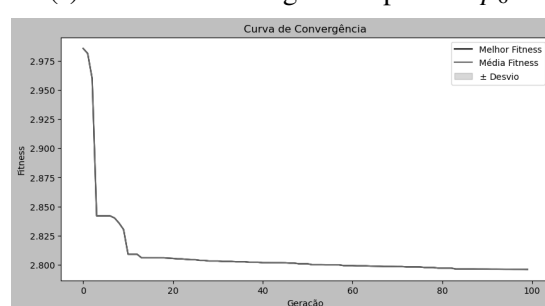
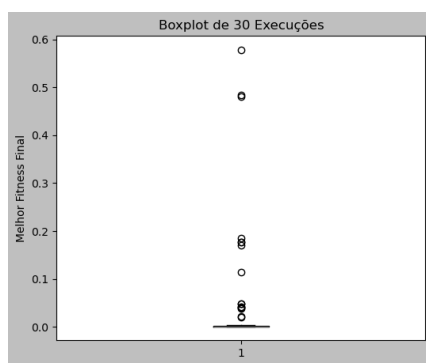
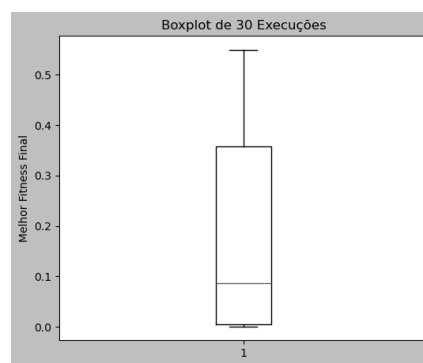
(a) Curva de Convergência - período p_1 .(b) Curva de Convergência - período p_2 .(c) Curva de Convergência - período p_3 .(d) Curva de Convergência - período p_4 .(e) Curva de Convergência - período p_5 .(f) Curva de Convergência - período p_6 .(g) Curva de Convergência - período p_7 .(h) Curva de Convergência - período p_8 .

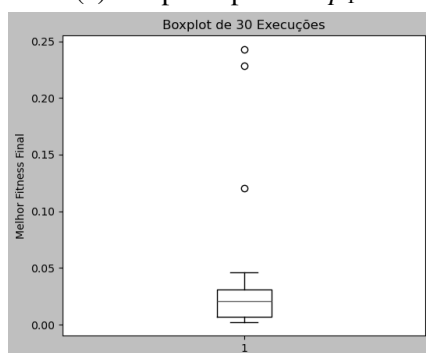
Figura 31 – Curvas de Convergência referente ao modelo 24



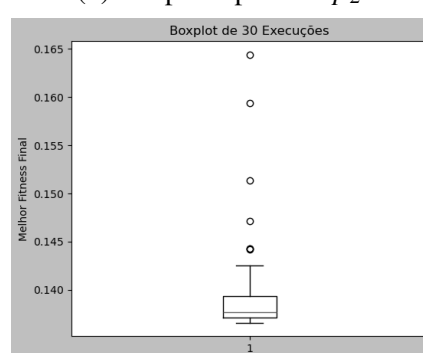
(a) Boxplot - período p_1 .



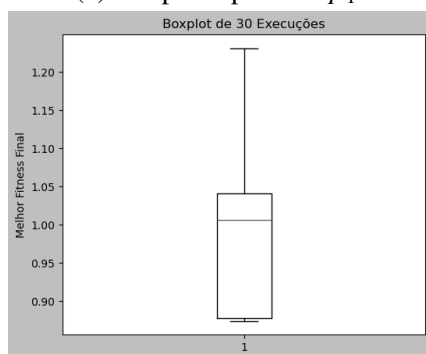
(b) Boxplot - período p_2 .



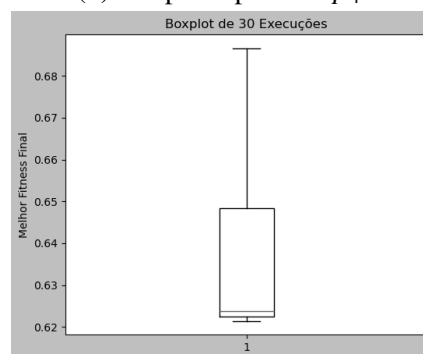
(c) Boxplot - período p_1 .



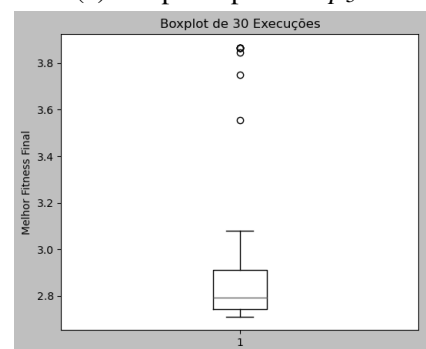
(d) Boxplot - período p_4 .



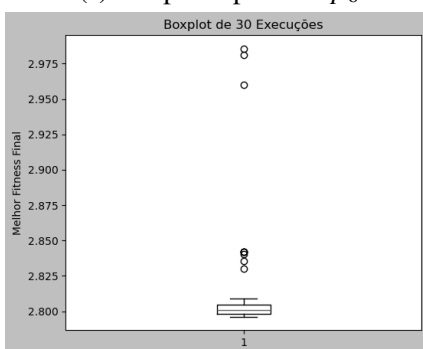
(e) Boxplot - período p_5 .



(f) Boxplot - período p_6 .



(g) Boxplot - período p_7 .



(h) Boxplot - período p_8 .

Figura 32 – Boxplots referentes ao modelo 24

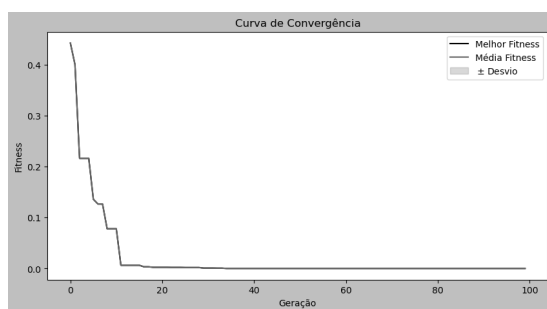
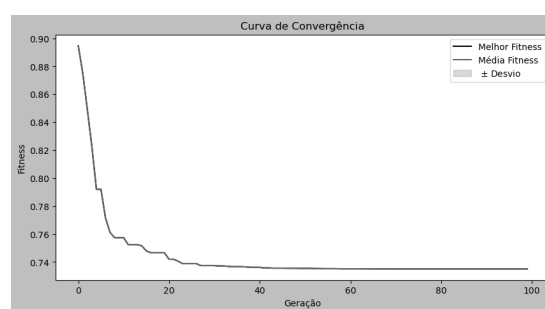
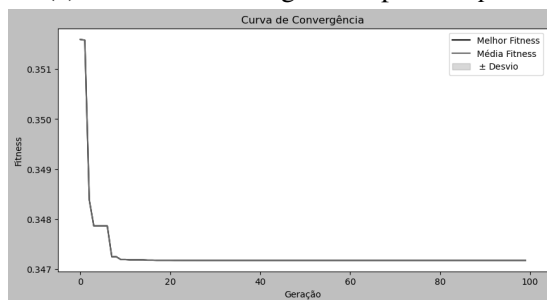
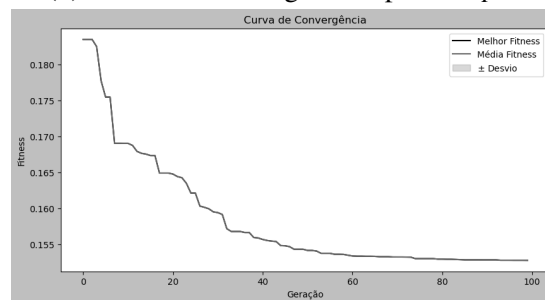
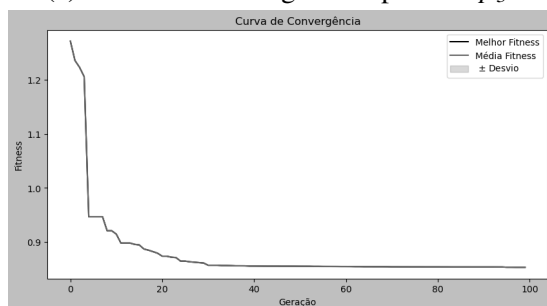
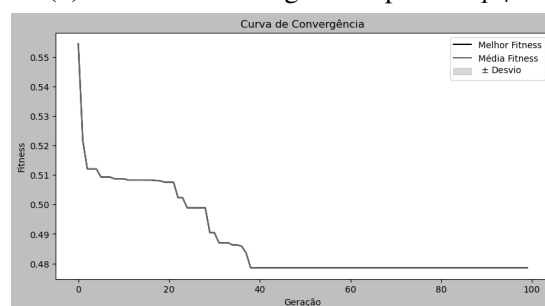
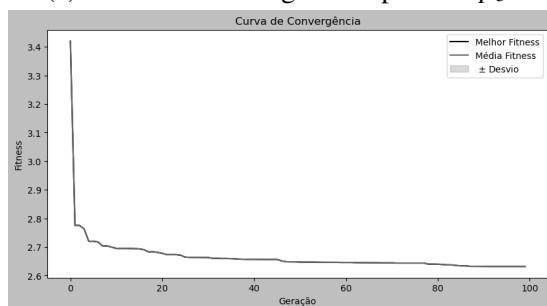
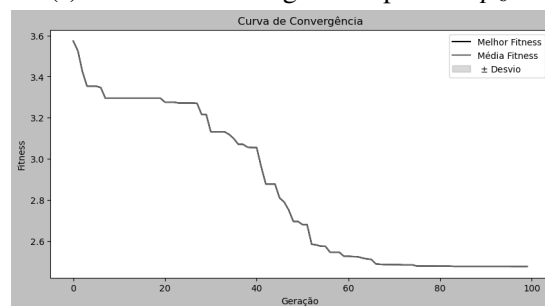
(a) Curva de Convergência - período p_1 .(b) Curva de Convergência - período p_2 .(c) Curva de Convergência - período p_3 .(d) Curva de Convergência - período p_4 .(e) Curva de Convergência - período p_5 .(f) Curva de Convergência - período p_6 .(g) Curva de Convergência - período p_7 .(h) Curva de Convergência - período p_8 .

Figura 33 – Curvas de Convergência referente ao modelo 28

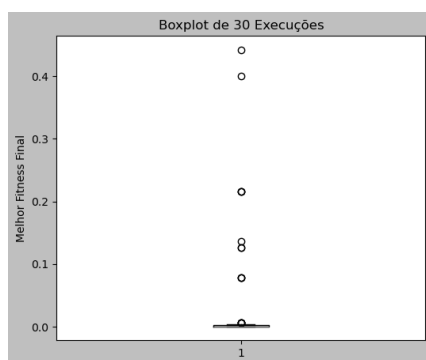
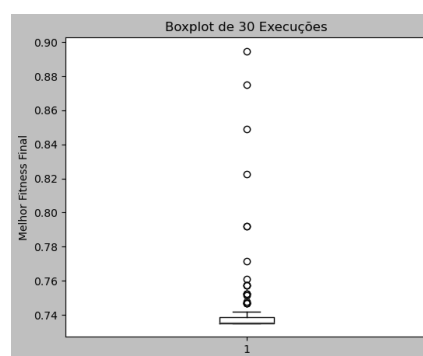
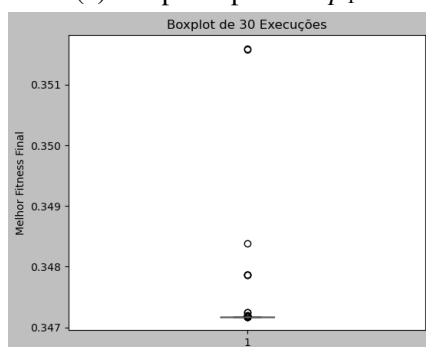
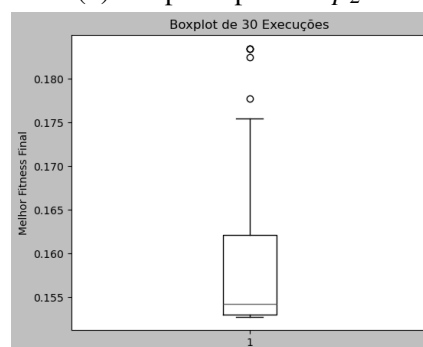
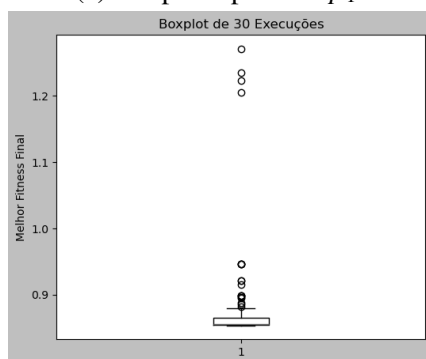
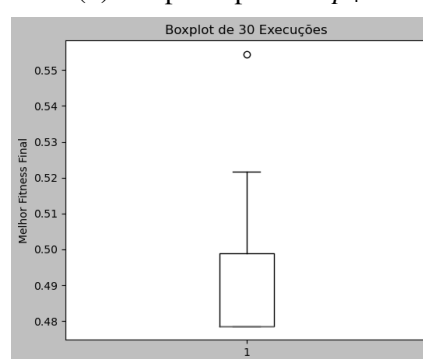
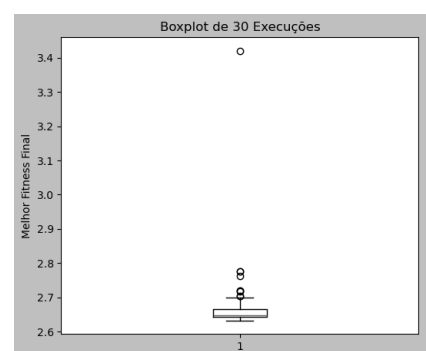
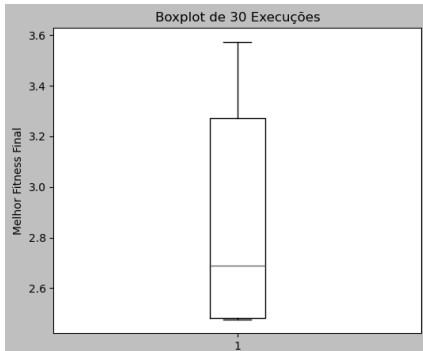
(a) Boxplot - período p_1 .(b) Boxplot - período p_2 .(c) Boxplot - período p_1 .(d) Boxplot - período p_4 .(e) Boxplot - período p_5 .(f) Boxplot - período p_6 .(g) Boxplot - período p_7 .(h) Boxplot - período p_8 .

Figura 34 – Boxplots referentes ao modelo 28

APÊNDICE D

Parâmetros Estatísticos dos AGRP

Tabela 11 – Média, Desvio Padrão e Coeficiente de variação do valor ótimo da função objetivo encontrado, por meio do AGRP, em cada período p_i referente ao Modelo 23

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
Média	2.3790	2.8370	0.2399	0.1840	0.7590	0.2186	3.0386	4.7214
Desvio Padrão	0.0992	0.0018	0.1044	0.0318	0.2427	0.0116	0.2092	0.0105
CV	0.0417	0.0006	0.4353	0.1728	0.3197	0.0529	0.0689	0.0022

Tabela 12 – Média, Desvio Padrão e Coeficiente de variação do valor ótimo da função objetivo encontrado, por meio do AGRP, em cada período p_i referente ao Modelo 24

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
Média	0.1109	0.0208	0.0347	0.1390	0.8969	0.6252	2.7666	2.8128
Desvio Padrão	0.3777	0.1508	0.1366	0.0098	0.0817	0.0220	0.1706	0.1597
CV	3.4056	7.2558	3.9336	0.0706	0.0911	0.0351	0.0616	0.0568

Tabela 13 – Média, Desvio Padrão e Coeficiente de variação do valor ótimo da função objetivo encontrado, por meio do AGRP, em cada período p_i referente ao Modelo 28

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
Média	0.3353	0.7396	0.3991	0.1537	0.8774	0.7711	2.6381	2.5114
Desvio Padrão	0.7235	0.0149	0.1236	0.0075	0.0927	0.7342	0.0292	0.1619
CV	2.1578	0.0202	0.3096	0.0491	0.1056	0.9521	0.0111	0.0645