

Universidade de São Paulo
Instituto de Física

Dinâmica livre de escala da COVID-19

Josué Mendes Pinheiro Policarpo


Dissertação de mestrado apresentada ao Instituto de Física da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestre em Ciências. Orientador: Airton Deppman

Banca Examinadora:

Prof. Dr. Airton Deppman (Orientador) - IFUSP

Profa. Dra. Ester Cerdeira Sabino - FMUSP

Prof. Dr. Constantino Tsallis - CBPF



São Paulo

2023

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Policarpo, Josué Mendes Pinheiro

Dinâmica livre de escala da COVID-19. São Paulo, 2023.

Dissertação (Mestrado) - Universidade de São Paulo. Instituto de Física. Depto. de Física Experimental

Orientador: Prof. Dr. Airton Deppman

Área de Concentração: Física Nuclear

Unitermos: 1. Sistemas complexos; 2. Mecânica estatística; 3. Fractal;
4. Dinâmica epidêmica

USP/IF/SBI-015/2023

University of São Paulo
Physics Institute

Scale-free dynamics of COVID-19

Josué Mendes Pinheiro Policarpo

Dissertation submitted to the Physics Institute
of the University of São Paulo in partial
fulfillment of the requirements for the degree of
Master of Science. Supervisor: Airton Deppman

Examining Committee:

Prof. Dr. Airton Deppman (Supervisor) - IFUSP

Profa. Dra. Ester Cerdeira Sabino - FMUSP

Prof. Dr. Constantino Tsallis - CBPF

São Paulo

2023

Dedicado a minha família, professores, amigos e a todos
que de alguma forma me ajudaram.

Agradecimentos

Agradeço a Deus pela minha vida e por me manter até aqui, sustentando e proporcionando tudo aquilo que precisei durante esse período.

A meus pais que me incentivaram e contribuíram, através de muito esforço, com minha educação, acreditando nos meus sonhos e ensinando os valores que tenho.

Sou grato a minha esposa pelo apoio e por deixar essa caminhada mais alegre.

Pelos inúmeros conselhos, horas de aula, exemplo de cientista e professor, deixo a minha gratidão ao meu orientador Airton Deppman.

Agradeço a Universidade de São Paulo por tudo que me proporcionou, em especial à Faculdade de Medicina pelo apoio financeiro.

I think the next century will be the century of complexity.

Stephen Hawking

Resumo

Neste trabalho, o processo de infecção epidemiológica é investigado, procurando-se uma descrição mais realista. O modelo obtido resulta em um conjunto de três equações diferenciais acopladas que generalizam o conhecido modelo epidemiológico SIR (Suscetíveis - Infectados - Recuperados). O número de contatos próximos, N , é um parâmetro do modelo, e é considerado sempre finito. As soluções das equações tem a forma de q -exponenciais, onde o índice entrópico, q , depende de N na forma $(q - 1)^{-1} = N - 1$. Os testes realizados para a verificação da abordagem contaram com a utilização de dados sobre a COVID-19 da cidade de São Caetano, mostrando que o uso de funções do tipo q -exponencial melhoram o ajuste entre a curva teórica e os dados reais de novos infectados. O modelo também foi capaz de estimar o número de contatos próximos N e indicar mudanças no comportamento social. Durante a epidemia, quando a população flexibilizou o isolamento social, o número de contatos próximos passou de $N = 3$ para $N = 6$.

Além disso, foram feitas análises da relação entre o número de infectados e a população de alguns países e cidades, como USA, França, Itália, Espanha, China e São Paulo. Essa relação se dá por meio de uma lei de potência com expoente próximo de 1.15, valor semelhante ao encontrado na literatura para diversas grandezas sociais, associado a invariância de escala na propagação de vírus, e trouxe indícios de aspectos fractais.

Palavras-chave: Sistemas Complexos. Mecânica Estatística. Fractal. Dinâmica Epidêmica

Abstract

In this paper, the epidemiological infection process is investigated, seeking a more realistic description. The model obtained results in a set of three coupled differential equations that generalize the well-known epidemiological model SIR (Susceptible - Infected - Recovered). The number of close contacts, N , is a parameter of the model, and is always considered finite. The solutions of the equations have the form of q -exponentials, where the entropic index, q , depends on N in the form $(q - 1)^{-1} = N - 1$. The tests performed to verify the approach relied on data on COVID-19 from the city of São Caetano, showing that the use of functions of the q -exponential type improve the fit between the theoretical curve and the actual data of newly infected. The model was also able to estimate the number of close contacts N and indicate changes in social behavior. During the epidemic, when the population relaxed social isolation, the number of close contacts increased from $N = 3$ to $N = 6$.

In addition, analyses of the relationship between the number of infected and the population of some countries and cities, such as USA, France, Italy, Spain, China and São Paulo, have been made. This relationship occurs through a power law with an exponent close to 1.15, a value similar to that found in the literature for several social quantities, associated with scale invariance in virus propagation, and brought indications of fractal aspects.

Keywords: Complex System. Statistical Mechanics. Fractal. Epidemic Dynamics

Lista de ilustrações

Figura 1 – Curva de Koch e sua interação em 3 passos (FALCONER, 2004).	21
Figura 2 – Esquema ilustrando o método box-counting retirado de (WU et al., 2020).	22
Figura 3 – Fluxo de pessoas entre os compartimentos no modelo SIR.	25
Figura 4 – Simulação utilizando as equações 2.7. Os parâmetros escolhidos foram: $N = 750$, $\beta = 0.17$ e $\gamma = 0.05$	28
Figura 5 – Simulação usando as equações 2.7 para diferentes R_0	29
Figura 6 – Modos de contágio em um pequeno grupo de quatro indivíduos ((L)esquerda, (C)central, (D)direita e (B)fundo). As figuras mostram esquematicamente os diferentes modos, diretos (figura esquerda) ou indiretos (figuras central e direita), que um indivíduo suscetível no grupo pode ser infectado a partir de um indivíduo inicialmente infectado (vermelho).	32
Figura 7 – Ilustração do aspecto fractal do modelo em diferentes escalas. O tamanho do grupo e dos agentes variam conforme a escala muda. No círculo 1, cada agente é formado por uma única pessoa, $N = 2$, enquanto no círculo 4, cada agente é formado por 27 pessoas, $N = 27$	34
Figura 8 – Cálculos numéricos da evolução temporal da população infectada e suscetível para grupos de: (Esquerda) 250000 indivíduos; (Direita) 2500 indivíduos.	36
Figura 9 – Sequência de dez curvas mostrando a variação das curvas de incidência em função da população. Os valores da população inicial de suscetíveis, s_0 , variou entre 500 a 10000 pessoas, κ foi fixado por uma lei de potência: $\kappa = 0.017s_0^{-0.833}$; $\kappa^* = 0.05$ e $q = 0.75$	39
Figura 10 – (Esquerda) Comportamento da largura de cada pico em função da população inicial de suscetíveis s_0 . Os parâmetros são os mesmos que o da figura 9. (Direita) Comportamento da largura de cada pico em função do parâmetro q . Nessa simulação a população inicial foi mantida fixa e com valor 800; $\kappa = 0.017s_0^{-0.833}$; $\kappa^* = 0.05$ e $0.5 < q \leq 1$	39
Figura 11 – Comportamento do número de casos de contaminação em função da população nas regiões a) China, b) EUA c) França, Alemanha, Espanha e Itália juntos e d) Estado de São Paulo, por população da cidade. Os intervalos de confiança são desenhados com nível de confiança de 95%.	46

Figura 12 – A figura acima representa o ajuste utilizando a equação 3.13 para um conjunto de dados separados em diferentes bins. A figura superior esquerda representa bins de 30 dias; superior direita bins de 15 dias; inferior esquerda bins de 7 dias; inferior direita bins de 4 dias. As linhas tracejadas indicam a contribuição de cada grupo enquanto a linha contínua é a soma de todas as contribuições. A linha horizontal tracejada foi deixada para um melhor ajuste e pode mostrar uma contaminação de fora dos grupos.	47
Figura 13 – Ajuste linear dos parâmetros Δt e κ em função do número de indivíduos do grupo (contatos) na forma de lei de potência.	48
Figura 14 – A figura acima representa o ajuste utilizando a equação 3.13 para um conjunto de dados separados em diferentes bins e fixando a lei de potência $\kappa = \kappa_0 s_0^{-0.833}$. A figura superior esquerda representa bins de 30 dias; superior direita bins de 15 dias; inferior esquerda bins de 7 dias; inferior direita bins de 4 dias. As linhas tracejadas indicam a contribuição de cada grupo enquanto a linha contínua é a soma de todas as contribuições. A linha horizontal tracejada foi deixada para um melhor ajuste e pode mostrar uma contaminação de fora dos grupos.	49
Figura 15 – A figura acima representa o ajuste utilizando a equação 3.13 para um conjunto de dados da segunda onda separados em diferentes bins e fixando a lei de potência $\kappa = \kappa_0 s_0^{-0.833}$. A figura esquerda representa bins de 4 dias; a direita bins de 7 dias.	50
Figura 16 – A figura acima representa o ajuste utilizando a equação 3.13 para um conjunto de dados da primeira e segunda onda separados em bins de 4 dias e com lei de potência $\kappa = \kappa_0 s_0^{-0.833}$	50
Figura 17 – Na esquerda os valores ajustados para o parâmetro q para a primeira onda (azul) e segunda onda (vermelho). A barra de erro não é visível, pois são pequenas em relação à escala do desenho, ver tabela 4. A linhas contínuas representam os valores médios ajustados.. A imagem à direita mostra a média estimada do isolamento social no período analisado.	51
Figura 18 – Na esquerda, a distribuição da menor distância entre um novo caso e todos os anteriores. A análise foi feita entre 22-03-2020 e 30-09-2020. Na direita, a distribuição de novos casos, usando o método de separação por clusters, através de mapa de calor. Cada cluster foi associado a uma determinada concentração de pessoas. Escola Técnica - próximo do Cluster 1, Parque - próximo do Cluster 4, Estádio Municipal - próximo dos Clusters 2 e 3 , Hospital - próximo do Cluster 5, Estação de Trem Universidade Municipal de São Caetano - próximo do Cluster 6, Fábrica de carros - próximo dos Clusters 5 e 6.	53

Figura 19 – Ajuste com dados em bins de 4 dias: (esquerda) $k^* = 0$ e os demais parâmetros livres; (direita) todos os parâmetros livres. Valores dos ajustes estão nas duas primeiras linhas da tabela 5.	53
Figura 20 – Ajuste com a população fixa em $5.0E4$ e todos os demais parâmetros livres (esquerda); índice $q = 1$ fixo e todos os demais parâmetros livres (direita).	54
Figura 21 – A linha contínua vermelha representa o ajuste com o modelo FSIR usando o valor $\kappa = 0.2$ e três picos (linhas coloridas tracejadas). A curva ponto-linha preta é o ajuste com o modelo FSIR usando o valor $\kappa = 0.2$ e $q = 1$	55

Lista de tabelas

Tabela 1	– Alguns exemplos de sistemas complexos, seus agentes, interações e comportamentos emergentes. (BAR-YAM, 1997)	15
Tabela 2	– Valores do expoente β para diversas grandezas sociais. (BETTENCOURT et al., 2007)	18
Tabela 3	– Parâmetros da equação 3.13 ajustados com o conjunto de dados separados por bins de 4, 7, 15 e 30 dias, como descrito em 4.3.2.1. O período estudado foi entre 18-03-2020 e 19-09-2020 mostrado na figura 12. O background c é uma constante que representa contaminação externa. .	48
Tabela 4	– Parâmetros otimizados usando a equação 3.13 e dados da primeira e segunda onda. O período analisado está entre 19-09-2020 e 19-11-2020. O parâmetro χ^2 está sendo comparado com o número de graus de liberdade. A coluna Δt_j representa a largura de cada fico, ou seja, o tempo de que cada grupo fica contaminado. O τ é o produto de κ e Δt .	52
Tabela 5	– Parâmetros ajustados para: $\kappa^* = \mathbf{0}$ fixo (primeira linha); todos os parâmetros ajustados (segunda linha); população fixa 5.0E4 , 5.0E5 e 7.5E5 (linhas 3 a 5); $\mathbf{q} = \mathbf{1}$ fixo (última linha).	54
Tabela 6	– Primeiras três linhas: Parâmetros ajustados com as equações 3.20 com $\mathbf{k}^* = \mathbf{0.2}$ e 3 agentes. Última linha: único agente, $\mathbf{k}^* = \mathbf{0.2}$ e $\mathbf{q} = \mathbf{1}$. .	54

Lista de abreviaturas e siglas

SI: Modelo Epidemiológico Suscetível-Infetado

SIR: Modelo Epidemiológico Suscetível-Infetado-Recuperado.

FSI:Variação do modelo SI usando a hipótese fractal.

FSIR:Variação do modelo SIR usando a hipótese fractal.

Lista de símbolos

$s_j(t)$	População suscetível do grupo j no tempo t .
$i_j(t)$	População infectada do grupo j no tempo t .
$r_j(t)$	População recuperada do grupo j no tempo t .
κ_j	Taxa de contágio do agente j .
κ_j^*	Taxa de removidos do agente j .
q	Índice entrópico.
λ	Parâmetro de escala que determina o tamanho do agente.
t_{0j}	Tempo inicial do primeiro infectado no grupo j .
Δt_j	Largura do pico que representa o tempo em que há infectados no grupo j .
k_i	Grau do nó i ou número de ligações que chegam no nó i .
σ	Número de novos casos (Incidência)

Sumário

1	INTRODUÇÃO	14
1.1	Considerações iniciais	14
1.2	Ciência das cidades	17
1.3	Fractais	20
2	MODELOS EPIDEMIOLÓGICOS	24
2.1	Modelos compartimentados	24
2.1.1	Suscetíveis-Infetados-Recuperados	24
2.2	Cálculo do R_0	26
2.2.1	Comportamento das populações no SIR	27
2.3	Variações do modelo SIR	29
2.3.1	Modelo SIS	29
2.3.2	Modelo SI	30
3	MODELO FRACTAL	31
3.1	Motivação	31
3.1.1	Dependência temporal	33
3.2	Modelo matemático e aproximações	33
3.2.1	Uma aproximação para um modelo do tipo Suscetível-Infetado	34
3.2.2	Uma aproximação para um modelo do tipo Suscetível - Infetado - Recuperados	36
4	METODOLOGIA	40
4.1	Dados	40
4.1.1	USA, São Paulo, China e Europa	40
4.1.2	São Caetano	40
4.1.3	Isolamento social	41
4.2	Programa	41
4.3	Etapas do ajuste	42
4.3.1	Lei de potência nos dados do COVID	42
4.3.2	Teste inicial com o modelo SI fractal (FSI)	42
4.3.2.1	Ajuste inicial	42
4.3.2.2	Ajuste usando $\kappa_j = \kappa_0 s_{0j}^\beta$	43
4.3.3	Localização geográfica dos novos casos	43
5	RESULTADOS	45
5.1	Ajuste com lei de potência	45

5.2	Teste inicial com o modelo FSI	45
5.3	Teste inicial com o modelo FSI usando lei de potência	48
5.3.1	Análise geográfica	51
5.3.2	Comparação do modelo FSIR e SIR	51
6	CONCLUSÃO	56
	REFERÊNCIAS	58

1 Introdução

1.1 Considerações iniciais

Na natureza existem diversos fenômenos que podem ser descritos através de modelos matemáticos que, por meio de aproximações, permitem realizar previsões coerentes com dados experimentais e entender muito do comportamento do sistema estudado. Um exemplo conhecido é o gás ideal que considera as moléculas com dimensões nulas, sem interação entre seus pares e podendo apenas sofrer colisões perfeitamente elásticas (CALLEN, 1998; GREINER; NEISE; STÖCKER, 2012). Tais aproximações são irrealistas e não correspondem a nenhum sistema composto por gases, mesmo os mais simples, pois sempre haverá interações entre as moléculas ou átomos e suas dimensões tem efeitos importantes. Contudo, esse modelo tem como resultado uma equação de estado que, para gases com densidades e pressões baixas, oferece uma excelente descrição das variáveis termodinâmicas pressão, volume e temperatura. Podemos ainda usar modelos do tipo Ising (ONSAGER, 1944), campo médio (KADANOFF, 2009) ou método Monte Carlo (BENYOUSSEF et al., 2020) para estudar transições de fase em modelos um pouco mais elaborados. Todas essas abordagens recorrem a aproximações que, muitas vezes, não conseguem representar todos os sistemas reais, como o comportamento de uma pessoa vivendo em uma grande cidade ou o aparecimento de fenômenos climáticos devido à dinâmica da nossa atmosfera. Há muitas situações que tornam tais aproximações difíceis, normalmente quando temos muitos agentes interagindo fortemente, fazendo surgir características e fenômenos não deduzidos a partir do estudo de seus constituintes isolados. Esses sistemas são chamados de sistemas complexos e tem como característica o número elevado de constituintes, com propriedades iguais ou não, interações mútuas de diversas naturezas e apresentam comportamento emergente que não podem ser estudados a partir dos seus agentes isolados (WEST, 2018; BERTIN, 2016; ANDERSON, 2011).

Atualmente, muitos estudos mostram diversos sistemas que podem ser considerados complexos, em diversas áreas, como economia (ARTHUR; BEINHOCKER; STANGER, 2020; ANDERSON, 2018; ARTHUR, 2018), ecologia (PASCUAL et al., 2006; ODUM, 1964), Biologia e Evolução (UYEHARA, 2008; RAMOS; ABRAHAM, 2004), cidades (WEST, 2018; CHOWELL et al., 2003; SCHLAPFER et al., 2014) entre outros.

De modo geral, podemos definir os sistemas complexos através de suas principais propriedades (MITCHELL, 2009; HOLLAND, 2014; AMARAL; OTTINO, 2004):

- Número elevado de constituintes: Partículas ou agentes que podem ser de mesma natureza ou não, que interagem com seus vizinhos ou com todos do sistema. Podemos

citar como exemplos as colônias de insetos, nosso sistema nervoso, gases da atmosfera e sua interação com os seres vivos etc.

- **Comportamento emergente:** O sistema apresenta um comportamento macroscópico que não é trivial quando comparado com o de suas partes. A dinâmica do sistema é mais complexa do que a soma dos efeitos de suas partes (ANDERSON, 1972). Um dos mais importantes comportamentos emergentes é a nossa consciência. Mesmo conhecendo a estrutura dos neurônios (agente), suas sinapses (interações), a manifestação da consciência surge dando origem aos mais diversos pensamentos.
- **Auto-organização:** Não existe um elemento central que comanda ou é responsável pela dinâmica do sistema, mesmo assim, há uma organização a partir de estímulos internos ou externos que organiza o sistema para responder aos mais diversos propósitos, como a adaptação ao meio.
- **Adaptação:** A partir de estímulos externos o sistema evolui para se adaptar às novas condições e parâmetros. A evolução biológica é um dos maiores exemplos dessa propriedade.
- **Hierarquia:** Sistemas complexos que se juntam formando novos sistemas complexos. Nossas células são sistemas que se juntam para formar órgãos que formam nossos corpos. As pessoas vivem em sociedade, constituindo um novo sistema complexo que interage com o meio ambiente que possui características complexas. Assim, mesmo não existindo um comando central ou uma parte mais importante que a outra, os sistemas podem interagir com outros, formando novos sistemas com comportamento emergente ainda mais imprevisível.

Tabela 1 – Alguns exemplos de sistemas complexos, seus agentes, interações e comportamentos emergentes. (BAR-YAM, 1997)

Sistema	Elementos	Tipo de Interação	Emergência
Proteínas	Aminoácido	Ligações Químicas	Atividade Enzimática
Sistema Nervoso	Neurônios	Sinapse	Consciência
Fisiologia	Células	Informação Química	Desenvolvimento Biológico
Vida	Organismos	Reprodução Competição	Evolução
Sociedade	Pessoas Empresas	Cooperação Relacionamentos	Evolução Social

A tabela 1 mostra alguns exemplos de sistemas complexos. Vemos que são formados pelos mais diferentes tipos de agentes, mostrando a necessidade de ferramentas multidisci-

plinares aplicadas nos estudos desses temas. O comportamento emergente é mostrado na última coluna da tabela. Mesmo com todo o conhecimento relacionado aos agentes e a forma como eles interagem, não é possível deduzir, a partir da descrição microscópica, o comportamento em grande escala. Esse exemplo também mostra uma hierarquia entre eles. Veja que a sociedade, última escala da tabela 1, é resultado da interação de diversos outros agentes que estão descritos na linha acima, formando um conjunto hierárquico entre eles. Com tantos fenômenos de diferentes naturezas, dinâmicas complicadas e propriedades muito distintas entre si, faz-se necessário encontrar uma maneira mais geral de trabalhar com esses sistemas. Para descrever muitas dessas propriedades em diferentes escalas, cientistas recorrem às redes.

Rede é um sistema composto por um conjunto de vértices ou nós, com conexões entre eles, chamadas de arestas, que irão representar os agentes e suas interações.

Esses sistemas podem ser representados, matematicamente, por meio de grafos. Alguns dos exemplos comuns classificados como redes são: Internet, interações sociais de diversos tipos, sistemas neurais e sistemas biológicos, em geral (ALBERT; BARABÁSI, 2002).

O estudo de redes tem grande importância em diversas áreas, pois consegue capturar comportamentos e padrões em sistemas de alta complexidade. Muitos desses padrões podem revelar comportamentos universais, como em Ciência das Cidades e suas leis de potência. Além disso, estudos como esse tem uma capacidade de simplificação bastante conveniente em muitos casos.

Para fazer uso dessa análise é necessário estabelecermos algumas definições e desenvolvermos ferramentas matemáticas para descrever, por exemplo, as relações entre dois nós ou como que distribuições de probabilidade podem ser aplicadas.

Um grafo é definido como um par ordenado $G = (V(G), E(G))$ onde $V(G)$ é um conjunto não nulo de elementos, chamados de vértices ou nós, e $E(G)$ um conjunto de pares de vértices (não necessariamente diferentes) chamados de arestas ou links (BONDY, 2008). A teoria dos grafos tem grandes aplicações em sistemas complexos, tornando importante o desenvolvimento de ferramentas para quantificar e descrever as principais propriedades.

- Grau: Representa uma propriedade existente em cada nó e indica o número de ligações que cada agente possui. Representamos por k_i o grau do vértice i dentro da rede.
- Distribuição de Probabilidade: p_k é a probabilidade de escolhermos aleatoriamente um nó e ele possuir o grau k . Como toda a distribuição, temos:

$$\sum_{i=1}^N p_i = 1. \quad (1.1)$$

Além disso, a média é definida como:

$$\langle k \rangle = \sum_{i=1}^N k p_i. \quad (1.2)$$

- Matriz de Adjacência: Podemos representar um grafo através de uma matriz que informa se o nó i está conectado com o nó j .

$$A_{ij} = \begin{cases} 1, & \text{se há ligação entre o nó } i \text{ e } j \\ 0, & \text{se não há ligação entre os nós } i \text{ e } j \end{cases} \quad (1.3)$$

Para um grafo não direcionado, ou seja, que não há diferença no sentido da ligação, $i \rightarrow j = j \rightarrow i$:

$$k_i = \sum_{j=1}^N A_{ij}. \quad (1.4)$$

Para o caso de grafos com ligações direcionadas, com K_i^{entrada} representando o número de ligações que chegam em i e $K_i^{\text{saída}}$ o número de ligações que saem de i , temos:

$$K_i^{\text{entrada}} = \sum_{j=1}^N A_{ij}, \quad (1.5)$$

$$K_i^{\text{saída}} = \sum_{i=1}^N A_{ij}. \quad (1.6)$$

A distribuição de probabilidade dos graus pode assumir várias formas dependendo do tipo de rede. Pode assumir a forma de uma distribuição binomial ou Poisson, no caso de uma distribuição completamente randômica, ou uma distribuição em forma de lei de potência. Esta última são as mais comuns para sistemas reais. Ela indica que alguns agentes terão mais ligações que outros, promovendo o aparecimento de clusters, que é encontrado, por exemplo, na internet (ALBERT; JEONG; BARABÁSI, 1999).

1.2 Ciência das cidades

O comportamento das pessoas mudou muito em relação as suas escolhas de moradia e trabalho. Por volta da década de 50, menos da metade da população vivia em cidades, enquanto atualmente, cerca de 80% da população mundial é considerada urbana e com previsão de acréscimo de mais de 2 bilhões até 2050 (BETTENCOURT; WEST, 2010). Essa alteração na densidade populacional, em muitos lugares, geram problemas, como o aumento da poluição, maior criminalidade, falta de infraestrutura adequada e outros exemplos característicos de um crescimento urbano mal planejado. Por outro lado, grandes centros urbanos tem um maior número de inovações, mostrado pelo número de patentes e maior média salarial (WEST, 2018; CHANG et al., 2015). A ciência das cidades busca

entender a relação do tamanho da população com as grandezas sociais medidas nos grandes centros. Uma maneira de observar as características das cidades é coletando dados sobre o consumo e produção de seus habitantes. O número de novas estradas, quantidade de ligações, produção intelectual, consumo de alimentos e muitos outros tipos de informações que caracterizam as cidades e as diferenciam umas das outras.

As cidades são compostas por regiões, bairros, conjuntos habitacionais, famílias que por sua vez são compostas por indivíduos. Cada membro da cidade interage com outros da mesma classe, seja por relações familiares, trabalho, amizade ou simples interação em seu espaço comum. O comportamento dos indivíduos é extremamente complexo e na maioria da vezes imprevisível. Logo, poderíamos supor que as características e comportamento da cidade, em uma escala maior, não tenham padrão ou ordem, mas como mostrado em (WEST, 2018; BETTENCOURT et al., 2007), muitas grandezas sociais apresentam uma lei de escala por meio de uma lei de potência com coeficientes bem definidos e compatíveis com os dados.

$$Y \propto N^\beta . \quad (1.7)$$

Na equação 1.7, Y é uma grandeza social, dependente do tamanho N da população, que pode ser a renda por pessoa, número de estradas, número de patentes ou mesmo alguma medida de contato entre os indivíduos. O expoente β é obtido através da análise dos dados. O que normalmente é percebido é que muitas das grandezas sociais escalam com um $\beta \approx 1.15$.

Tabela 2 – Valores do expoente β para diversas grandezas sociais. (BETTENCOURT et al., 2007)

Y	β	País e ano
Novas Patentes	1.27	EUA - 2001
Inventores	1.25	EUA - 2001
Empregos de Pesq. e Desenv.	1.26	EUA - 2002
Empregos Criativos	1.15	EUA - 2003
Salários	1.12	EUA - 2002
Crimes Graves	1.16	EUA - 2003
Total de Habitações	1.00	EUA - 1990
Consumo doméstico de água	1.01	China-2002
Consumo elétrico doméstico	1.00	Alemanha-2002
Consumo elétrico doméstico	1.05	China-2002
Postos de gasolina	0.77	EUA - 2001
Comprimento da Malha Elétrica	0.87	Alemanha-2002
Tamanho da Superfície das Estradas	0.83	Alemanha-2002

A tabela 2 mostra o comportamento de β para diversas grandezas sociais. Este expoente indica a dinâmica e natureza da grandeza. Podemos separar essas informações em

três grupos. O primeiro representa quantidades que independem da quantidade de pessoas, $\beta = 1$, como a quantidade de consumo de água, comida ou energia elétrica consumida por pessoa. Nesse caso, temos um comportamento linear. Para $\beta < 1$, vemos que a dinâmica é sublinear, indicando uma maior eficiência. Ou seja, quanto maior a população, menos de determinada grandeza sublinear, por exemplo, o número de postos de gasolina, será necessária para atender a demanda da cidade. Quando determinado processo depende do número de contatos, o que aumenta com o tamanho da população, temos $\beta > 1$ e um comportamento superlinear, como mostrado na tabela 2 para salário, número de patentes e crimes graves (DEPPMAN, 2019).

A lei de potência modelando as grandezas de uma cidade tem relevância, pois fornece um modelo para identificar e descrever, ao longo do tempo, como elas variam. Por exemplo, se adicionarmos na equação 1.7 uma dependência temporal e uma constante de normalização Y_0 , temos:

$$Y(t) = Y_0 N(t)^\beta. \quad (1.8)$$

Se considerarmos um recurso $Y(t)$ que tem, em média, um valor α gasto por pessoa em um certo intervalo de tempo e um fator de consumo ϕ devido ao acréscimo na população $\dot{N}(t)$, temos a seguinte expressão para a quantidade do recurso no tempo t :

$$Y(t) = \alpha N(t) + \phi \frac{dN(t)}{dt}. \quad (1.9)$$

Isolando o termo $\dot{N}(t)$, podemos verificar a taxa de crescimento da população (BETTENCOURT et al., 2007).

$$\frac{dN(t)}{dt} = \frac{Y(t)}{\phi} - \frac{\alpha}{\phi} N(t). \quad (1.10)$$

Usando a lei de potência 1.8:

$$\frac{dN(t)}{dt} = \frac{Y_0}{\phi} N(t)^\beta - \frac{\alpha}{\phi} N(t). \quad (1.11)$$

No artigo (BETTENCOURT et al., 2007) é mostrado o comportamento da população em função dos valores de β . Para uma dinâmica linear, $\beta = 1$, a população cresce de forma exponencial. Com $\beta < 1$ a população tem um valor máximo e depois se estabiliza. Por fim, para $\beta > 1$, obtemos um tempo crítico onde o crescimento é insustentável. Entender como o tamanho população evolui no tempo em função dos parâmetros da equação 1.9 é uma importante aplicação, desde que se conheça o parâmetro β , o que é estimado de diversas formas utilizando dados reais.

De maneira geral, os sistemas complexos associados às cidades representam um conjunto muito grande de fenômenos e necessitam de uma abordagem multidisciplinar. Um exemplo

atual disso são os esforços de físicos, médicos, matemáticos e vários outros cientistas de diversas áreas para tentar entender como quantificar e descrever a propagação do vírus da COVID-19. Isso porque o vírus se propaga de acordo com o comportamento das pessoas. Se elas tiverem uma movimentação e interação grande umas com as outras, o vírus se espalha mais rápido do que uma população em distanciamento social. Esse problema exige técnicas das mais diversas e tem como objeto de estudo as interações sociais e como as cidades se organizam formando um sistema complexo e de alta conectividade.

Sabendo das relações entre o tamanho da população e das grandezas mensuráveis nas cidades foi feita uma análise com dados de diversos países e verificamos que o número de novos casos varia com a população seguindo uma lei de potência (ABBASI et al., 2020) e isso motivou a análise com dados precisos do número de novos casos de COVID-19 na cidade de São Caetano, o que será analisado nos próximos capítulos.

Um dos modelos mais conhecidos e utilizados em epidemiologia é o SIR (Suscetíveis-Infetados-Recuperados) que é baseado em uma ideia de modelagem compartimentada, onde a população é separada em grupos (MARATHE, 2013). Uma das aproximações por trás desse modelo é de uma contaminação homogênea que indica uma probabilidade de contaminação igual em toda a população. Porém, é fácil perceber que em uma situação real, indivíduos com maior contato com um contaminado tem probabilidade de contaminação muito maior do que indivíduos com menor contato. Isso é resultado de uma propagação não-homogênea.

A análise segue com a utilização de um modelo de contaminação baseado em grupos menores de contato que variam de tamanho conforme a escala da população analisada. Grupos de uma escala pequena pode ser representado por famílias e grupos maiores por bairros. A ideia central é que a forma de propagação entre os grupos se dá de forma auto-semelhante quando comparado com grupos de outros tamanhos.

Com alguns desenvolvimentos a serem detalhados na dissertação, seguimos com o que denominamos de modelo FSIR ou modelo SIR fractal e a partir disso fizemos uma série de ajustes para avaliar o modelo.

1.3 Fractais

A geometria sempre ocupou um lugar importante na matemática e em suas aplicações. Sua origem e desenvolvimento estão ligados diretamente com a evolução do pensamento abstrato, desde a resolução de problemas com relação prática, como cálculo ou descrição de áreas, volumes e formas, até a formulação de modelos cosmológicos que se desenvolveram ao longo de toda a história da ciência.

A classificação das figuras geométricas e seu uso através do cálculo oferece uma grande variedade de ferramentas para a descrição da natureza. Porém, uma grande parte da

geometria do nosso mundo natural não encontra correspondência exata ou aproximada com a geometria clássica. Podemos citar a forma da costa dos continentes, padrões em plantas e árvores, nossos vasos sanguíneos, sistema nervoso ou até mesmo o desenho de um relâmpago no céu. Esses são apenas alguns exemplos de geometrias difíceis de classificar. Em 1975 Benoît B. Mandelbrot publica um trabalho importante sintetizando ideias e definindo propriedades dos objetos, agora chamados, de fractais. Os fractais são os objetos matemáticos que descrevem a complexidade dessas formas. A palavra fractal vem do latim *fractu* que significa quebrado, fração, irregular.

Ao longo dos anos muitas ferramentas foram desenvolvidas para analisar os fractais e suas propriedades. Isso trouxe avanços em diversas áreas, como a matemática, biologia, ciências sociais, física e química.

Assim como fizemos com os sistemas complexos, podemos definir os fractais a partir de algumas das suas propriedades (FALCONER; FALCONER, 1997):

- Possui uma estrutura fina, podendo ser regular ou não, exibindo detalhes mesmo em escalas pequenas.
- As suas formas não podem ser descritas pela geometria tradicional.
- A sua geometria possui auto-semelhança, exata ou aproximada, em diferentes escalas.
- Tem dimensão D fora do conjunto dos inteiros.
- Pode ser gerada por métodos de recursão.

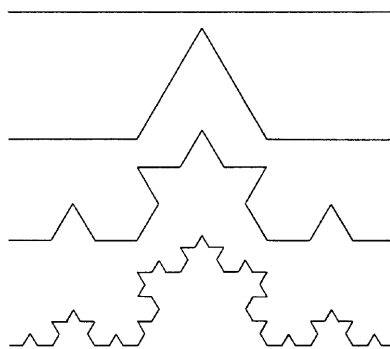


Figura 1 – Curva de Koch e sua interação em 3 passos (FALCONER, 2004).

A curva de Koch, figura 1, é um exemplo de fractal. A sua construção pode ser obtida a partir de uma sequência de interações que podem seguir indefinidamente. A primeira etapa consiste em dividir a reta com comprimento L em 3 partes de tamanho $r = L/3$ cada uma e substituir o segmento central por duas partes, como mostrado na figura 1. O mesmo procedimento é repetido em cada segmento. Observe que em cada

escala (passos da interação) a figura mantém a mesma forma (auto-semelhança).

Existem mais de uma forma de determinar a dimensão fractal dos objetos. A mais comum é a dimensão de Hausdorff D_f (MANDELBROT; MANDELBROT, 1982):

$$D_f = \frac{\ln N}{\ln \frac{1}{r}}. \quad (1.12)$$

É fácil perceber que em figuras regulares como linhas, quadrados ou cubos, a definição de D_f é equivalente a dimensão topológica D . Dividindo uma linha em N partes com um fator de redução de $r = 1/N$, temos $D_f = \frac{\ln N}{\ln \frac{1}{N}} = 1$. Para um quadrado, reduzimos cada lado pelo mesmo fator r e teremos $N = (\frac{1}{r})^2$ quadrados. Se substituirmos na equação de D_f obtemos o valor de 2. Na curva de Koch temos, em cada interação, um fator de redução de $r = 1/3$ e $N = 4$. Fazendo o cálculo temos $D_f = \frac{\ln 4}{\ln 3} \simeq 1.26$.

A dimensão fractal nos diz o quanto a figura ocupa o espaço no qual ele faz parte D_i (ASSIS et al., 2008). A costa de um país é um objeto fractal (MANDELBROT; MANDELBROT, 1982) e não ocupa uma superfície mas também não pode ser contida em uma linha. Sua dimensão é algo entre 1 e 2. Em geral, a dimensão fractal segue a seguinte relação: $D < D_f < D_i$.

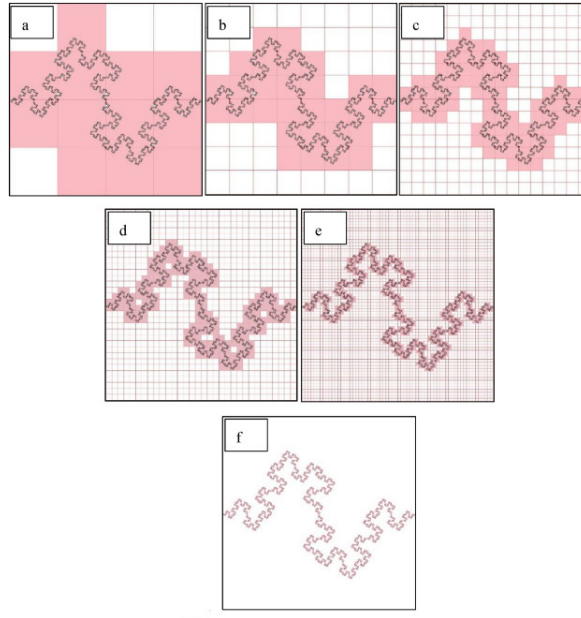


Figura 2 – Esquema ilustrando o método box-counting retirado de (WU et al., 2020).

Um método prático e muito usado para se determinar a dimensão fractal é o "box-counting" (WU et al., 2020), que tem a mesma ideia usada na dimensão de Hausdorff, porém, usando uma técnica de contagem de caixas. O método consiste em escolher caixas com lados de tamanho ϵ que irão cobrir toda a figura, como em a) da figura 2. Em seguida,

a figura é novamente coberta com novas caixas possuindo lados menores. Em cada etapa o tamanho delas diminuem, $\epsilon \rightarrow 0$, e o procedimento é repetido, como mostrado na figura 2. A partir disso, uma nova definição de dimensão é utilizada, muito parecida com a equação 1.12, mas utilizando o número de caixas necessárias para cobrir toda a imagem.

$$D_{box} = \lim_{\epsilon \rightarrow 0} \frac{\ln N(\epsilon)}{\ln \frac{1}{\epsilon}}. \quad (1.13)$$

Na equação 1.13, $N(\epsilon)$ representa o número de caixas, com lados de tamanho ϵ , necessárias para cobrir o fractal. Fazendo o plot log-log de $N(\epsilon)$ em função de ϵ , D_{box} será numericamente igual ao coeficiente angular da reta encontrada (KEMPKES et al., 2019; FOROUTAN-POUR; DUTILLEUL; SMITH, 1999).

2 Modelos epidemiológicos

2.1 Modelos compartimentados

A propagação de doenças causadas por vírus, fungos ou bactérias trazem grandes impactos negativos quando não controlados. Recentemente, a vida das pessoas ao redor do mundo foi alterada com o início da maior pandemia da história recente causada pelo novo coronavírus (SARS-CoV-2). As ações tomadas para conter o avanço da COVID-19 envolvem altos custos, desenvolvimento tecnológico e restrições sociais difíceis de serem implantadas, como quarentenas e distanciamento social.

Um conjunto de ações bem sucedidas exigem do poder público uma fundamentação técnica e científica de como o vírus se propaga para, a partir disso, estabelecer medidas que diminuam o número de infectados com um menor impacto social.

Os modelos matemáticos são de extrema importância para quantificar a dinâmica da propagação e prever possíveis cenários. Neste capítulo irei apresentar dois modelos bastante utilizados na epidemiologia: modelo SIR (Suscetíveis-Infectados-Recuperados) e modelo SI (Suscetíveis-Infectados).

2.1.1 Suscetíveis-Infectados-Recuperados

Os trabalhos relacionados a modelagens em epidemiologia não são um assunto recente. Há trabalhos feitos por Bernoulli, por volta de 1760, para o estudo da varíola (BERNOULLI, 1760). Uma aproximação importante nesses estudos foi a separação da população em compartimentos (KERMACK; MCKENDRICK, 1927). Essa abordagem é usada até hoje para descrever epidemias. De forma geral, os modelos podem ser divididos em duas categorias: Estocásticas e Compartimentais.

Os modelos estocásticos usam processos randômicos, redes e variáveis discretas para descrever seus processos (JAFARI; MALINOWSKI; EBADI, 2021; NAKAMURA et al., 2017). Já os modelos compartimentados usam variáveis contínuas e trabalham com a aproximação de que a população é composta por uma mistura homogênea de agentes. Isso significa que qualquer indivíduo pode interagir com qualquer outro agente da população com igual probabilidade, independente da distância, posição social, sexo ou idade. Esse modelo não é realista, pois as nossas interações sociais acontecem em um raio de ação, normalmente, muito menor que uma cidade (WEISS, 2013; WORMSER; POURBOHLOUL, 2008; WEISS, 2013).

Para começar a dedução do modelo, consideremos a população dividida em 3 compartimentos; suscetíveis (S), infectados (I) e recuperados (R). Aqui não estamos contando com o número de nascimentos ou morte de indivíduos. Além disso, há um fluxo bem definido

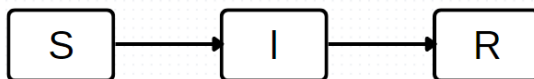


Figura 3 – Fluxo de pessoas entre os compartimentos no modelo SIR.

nesses compartimentos.

Considere que um determinado indivíduo tem uma média de k contatos por unidade de tempo e que há interação entre ele e os demais membros do grupo em um curto intervalo de tempo dt . Então podemos calcular o número de contatos como:

$$k \frac{I}{N} dt. \quad (2.1)$$

O número N é constante e representa a população total e I é o número de infectados. A razão entre eles é a fração de infectados.

Definindo a probabilidade de contaminação em um dos contatos como a e $(1 - a)$ para a não contaminação, podemos calcular a probabilidade de contaminação. Cada contato é independente do outro, ou seja, podemos usar o produto das probabilidades para encontrar a probabilidade de contaminação após determinado número de contatos. Por conveniência, a probabilidade de não contaminação será $1 - \delta q$. Para $k \frac{I}{N} dt$ contatos, temos:

$$1 - \delta q = (1 - a)^{k \frac{I}{N} dt}. \quad (2.2)$$

Definindo $\beta = k \ln(1 - a)$ e substituindo na equação acima:

$$\delta q = 1 - e^{-\beta \frac{I}{N} dt}. \quad (2.3)$$

Para transformar a probabilidade δq em uma taxa de transmissão, usamos uma expansão de Taylor em primeira ordem e fazemos $dt \rightarrow 0$.

$$\frac{dq}{dt} = -\beta \frac{I}{N}. \quad (2.4)$$

O termo $\frac{dq}{dt}$ é a taxa de transmissão. Considerando o fluxo de pessoas nos compartimentos, sabemos que os novos infectados representam a variação do número de suscetíveis. Usando essa taxa na população de suscetíveis, temos:

$$\frac{dS}{dt} = -\beta \frac{I}{N} S. \quad (2.5)$$

Dividindo ambos os lados por N e mudando o nome das variáveis: $I/N \rightarrow I$ e $S/N \rightarrow S$:

$$\frac{dS}{dt} = -\beta IS. \quad (2.6)$$

A equação 2.6 representa a taxa de pessoas que saem do grupo de suscetíveis e vão para o compartimento dos infectados. Como estamos considerando N constante e apenas 3 compartimentos, essa variação de pessoas, de acordo com a figura 3, necessariamente irá para os infectados.

O número de novos infectados dependem das populações S e I , como na 2.6. No entanto, os novos casos de recuperados depende, proporcionalmente, apenas do número de pessoas infectadas. Definindo γ como sendo a taxa de recuperação e que o número de recuperados dependa de I de forma linear, escrevemos a variação de cada compartimento através das equações diferenciais acopladas 2.7:

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I(t) S(t), \\ \frac{dI(t)}{dt} = \beta I(t) S(t) - \gamma I(t), \\ \frac{dR(t)}{dt} = \gamma I(t). \end{cases} \quad (2.7)$$

O parâmetro $\beta > 0$ é a taxa de transmissão do vírus por pessoa com unidade $(\text{pessoa} \cdot \text{tempo})^{-1}$. O parâmetro $\gamma > 0$ carrega uma informação valiosa e é a taxa com que os indivíduos se recuperam. Sendo T o período da infecção, $\gamma = \frac{1}{T}$ com unidade de $(\text{tempo})^{-1}$.

2.2 Cálculo do R_0

Os valores de $S(t)$, $I(t)$ e $R(t)$ possuem valores limites desde que N permaneça constante.

$$0 \leq S(t) \leq S(0) \leq N, \quad (2.8)$$

$$0 \leq R(0) \leq R(t) \leq N, \quad (2.9)$$

$$0 \leq I(0) \leq I(t) = N - S(t) - R(t). \quad (2.10)$$

Para qualquer valor de t .

Além disso;

$$S(t) + I(t) + R(t) = N \Rightarrow \frac{dS(t)}{dt} + \frac{dI(t)}{dt} + \frac{dR(t)}{dt} = 0. \quad (2.11)$$

Uma informação importante é saber o quanto uma epidemia pode ser acentuada em sua curva de infectados, ou mesmo se um vírus pode dar início a uma pandemia. Para chegarmos a algum parâmetro quantitativo, vamos analisar o início da contaminação, em outras palavras, quando o tempo $t = 0$.

No início temos: $I(0) = I_0 > 0$, $S(0) = S_0 > 0$ e $R(0) = 0$.

$$\frac{dI(t=0)}{dt} = \beta I_0 S_0 - \gamma I_0, \quad (2.12)$$

$$\frac{dI(t=0)}{dt} = I_0(\beta S_0 - \gamma). \quad (2.13)$$

Em uma epidemia o termo $\frac{dI}{dt} > 0$.

$$I_0(\beta S_0 - \gamma) > 0 \Rightarrow I_0 \beta S_0 > \gamma I_0, \quad (2.14)$$

$$S_0 > \frac{\gamma}{\beta}. \quad (2.15)$$

Olhando para a primeira linha da equação 2.7 vemos que $\frac{dS(t)}{dt} \leq 0$, ou seja, $S_0 \geq S(t)$. Isso significa que:

$$\frac{dI(t=0)}{dt} = I_0(\beta S_0 - \gamma) \geq 0, \quad (2.16)$$

para qualquer valor de t . Caso contrário, $S_0 < \frac{\gamma}{\beta}$, implicará em $\frac{dI(t=0)}{dt} < 0$ e não haverá epidemia.

Um parâmetro bastante comum na literatura é o R_0 , taxa de reprodução, que por definição é:

$$R_0 = \frac{S_0 \beta}{\gamma}. \quad (2.17)$$

Se $R_0 < 1$, não haverá epidemia. Esse parâmetro também serve para mensurar a força de uma epidemia. Quanto maior for esse número, mais inclinada será a curva dos infectados, como na figura 4

2.2.1 Comportamento das populações no SIR

Na figura 4 é mostrado o comportamento das populações $S(t)$, $I(t)$ e $R(t)$ no modelo SIR através de integração numérica. A curva dos infectados mostra uma distribuição aproximadamente simétrica, indicando um valor máximo no seu centro. As curvas dos

suscetíveis e recuperados apresentam derivada aproximadamente constante no intervalo entre 40 e 70 dias. Em 5 a curva azul, com $R_0 < 1$ não aparece, conforme mostrado acima. Para a curva com $R_0 = 1.52$ o crescimento é bem suave e demora a chegar em seu máximo, condição favorável para medidas de prevenção. Já a curva com $R_0 = 4$ apresenta um pico alto e crescimento rápido.

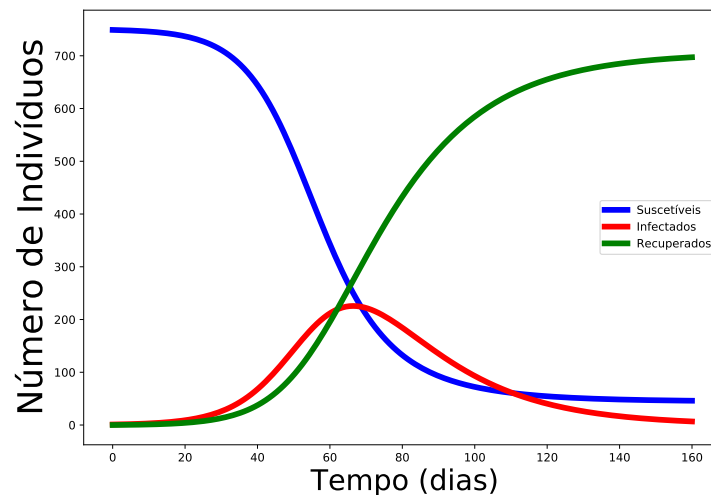


Figura 4 – Simulação utilizando as equações 2.7. Os parâmetros escolhidos foram: $N = 750$, $\beta = 0.17$ e $\gamma = 0.05$.

Um dos primeiros casos modelados foi a peste Bubônica, em 1896, pelos trabalhos de W. O. Kermack e A. G. McKendrick (KERMACK; MCKENDRICK, 1927). Eles mostraram que as curvas do modelo descrevem bem os dados do número de mortos. Mesmo não tendo a variável para a mortalidade, o número de mortos foi tão elevado que foi possível, através da aproximação do número de mortos $M(t)$ com o número de recuperados $R(t)$, fazer um ajuste adequado com os dados.

$$\frac{dR(t)}{dt} \approx \frac{dM(t)}{dt}. \quad (2.18)$$

Usando essa aproximação chegaram a uma solução analítica aproximada para o número $\frac{dM(t)}{dt}$.

$$\frac{dM(t)}{dt} = 890 \operatorname{sech}^2(0.2t - 3.4). \quad (2.19)$$

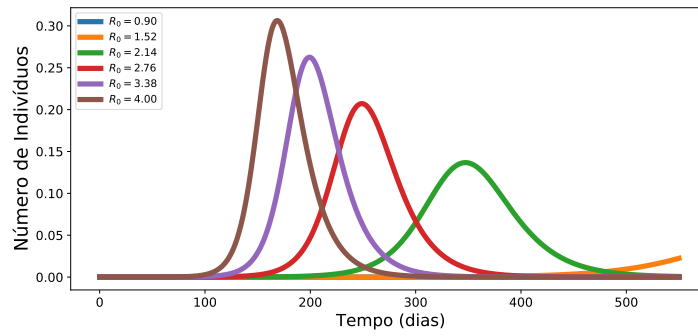


Figura 5 – Simulação usando as equações 2.7 para diferentes R_0 .

2.3 Variações do modelo SIR

2.3.1 Modelo SIS

O modelo SIS (Suscetíveis - Infectados - Suscetíveis) se baseia em uma população homogênea e compartimentada. Diferente do SIR não há recuperados. O fluxo segue dos suscetíveis para os infectados e novamente para os suscetíveis. Esse modelo é útil para epidemias onde não há imunidade para aqueles que se recuperam.

Ele é uma variação do modelo SIR apresentada também por Kermack e McKendrick (KERMACK; MCKENDRICK, 1932; KERMACK; MCKENDRICK, 1933).

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I(t) S(t) + \gamma I(t) \\ \frac{dI(t)}{dt} = \beta I(t) S(t) - \gamma I(t) \end{cases} . \quad (2.20)$$

A maioria das considerações continua válidas neste modelo, como: o número da população N é constante; $S(t) + I(t) = N$; não há nascimentos ou mortes durante o período; a quantidade $\beta I(t) S(t)$ tem o mesmo significado do modelo SIR; e $\gamma I(t)$ é a quantidade de indivíduos que retornam para o grupo dos infectados.

Podemos juntar as equações 2.23 em uma única equação.

$$S(t) + I(t) = N \Rightarrow S(t) = N - I(t) . \quad (2.21)$$

Substituindo esse resultado na equação de $\frac{dI(t)}{dt}$:

$$\frac{dI(t)}{dt} = \beta I(t) (N - I) - \gamma I(t) = (\beta N - \gamma) I(t) \left(1 - \frac{I(t)}{N - \frac{\gamma}{\beta}} \right) . \quad (2.22)$$

Assim, ficamos com apenas uma equação para descrever o número de infectados.

2.3.2 Modelo SI

Existem casos onde os indivíduos contaminados não se curam e nem se tornam suscetíveis. É o caso de epidemias como do vírus HIV. Nesse caso é possível modelar usando apenas duas populações: Suscetíveis e Infectados.

Continuamos com uma população homogênea, β igual à taxa de contaminação e $S(t) + I(t) = N$ fixo. As equações que descrevem a dinâmica das populações são:

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I(t) S(t) \\ \frac{dI(t)}{dt} = \beta I(t) S(t) \end{cases} . \quad (2.23)$$

Diferente do caso SIS, não temos o fator de saída $\gamma I(t)$. Com duas equações e duas variáveis, podemos simplificar juntando as equações em apenas uma.

$$S(t) + I(t) = N \Rightarrow S(t) = N - I(t) . \quad (2.24)$$

Substituindo na equação de $S(t)$ em $\frac{dI(t)}{dt}$:

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

Esse modelo considera que todos os habitantes deverão ser contaminados. Isso porque não consideramos nenhuma barreira para o vírus, como vacina, cura ou métodos de não infecção.

No equilíbrio, temos $\frac{dI(t)}{dt} = 0$, o que resulta que ou $I_0 = 0$, nesse caso não há nenhum agente infectado. Ou $I(t) = N$, que significa que toda a população se infectou.

3 Modelo fractal

3.1 Motivação

Os modelos epidemiológicos tradicionais consideram uma população homogênea e com probabilidade de contaminação igual para todos os agentes, desprezando a existência de grupos próximos (WEISS, 2013; ANDERSON, 1991). O modelo fractal considera uma população heterogênea e busca capturar detalhes importantes sobre a dinâmica do vírus ao ser transmitidos entre grupos de diferentes escalas de tamanho. Normalmente, quando olhamos para a curva de infectados ou de novos casos diários, percebemos muitos detalhes, fazendo com que o comportamento da curva não seja suave. Se é feita uma separação ou média semanal, parte desses detalhes desaparecem. O oposto também acontece; ao diminuirmos o tempo de coleta da informação, analisando o número de novos casos diários e não apenas semanais, há a possibilidade de verificar muito mais informações. Esse comportamento é similar ao dos fractais e é essa hipótese que deu início ao trabalho.

Para começar a construção do modelo é necessário formular as maneiras como se dá o contato entre os agentes. Ou seja, identificar os possíveis caminhos do vírus: se há contaminação direta ou indireta e qual a probabilidade de o vírus percorrer cada caminho. Na figura 6 é mostrado um esquema de como é contado cada maneira em que uma pessoa pode ser contaminada em um grupo, chamada de modo de contágio. Considere que o agente contaminado (em vermelho) pode ser contaminado de três diferentes formas: A primeira é quando há o contágio de forma direta, L-C (o indivíduo L infecta o indivíduo C). A segunda é indireta e com um agente intermediando, L-B-C. E por último uma maneira indireta com dois agentes intermediando o processo, L-B-R-C. Cada processo desse é chamado de modo de contágio.

Nesse exemplo, o grupo é composto por $N = 4$ agentes e contamos $(N - 1)$ modos de contágio, mas esse não é o total de configurações possíveis, pois o primeiro contaminado pode ser o B ou o R e isso daria outros caminhos. Repare que quanto mais o grupo é interagente, mais modos de contágio ele terá. É possível generalizar esse resultado para um N qualquer e incluir todas as possibilidades de contágio. Para iniciar essa generalização, começamos por entender os modos de contágio por um grupo com $N = 4$. Podemos escrever o número de modos de contatos diretos (L-C, B-C, R-C) em função de $\alpha = (N - 1)$, levando em consideração que o primeiro contaminado no grupo pode ser L, B ou R, já que o C não pode passar o vírus para ele mesmo. Para o caso de contaminação direta temos α modos. Para o caso de contaminação indireta com dois estágios, temos $\frac{\alpha!}{(\alpha-2)!2!}$ modos de contato e para três estágios $\frac{\alpha!}{(\alpha-3)!3!}$ modos.

Definindo τ como a fração de contaminados por contato e assumindo $\tau \ll 1$, podemos

escrever o número de contaminados em função de τ e do número de modos de contato através de uma expansão em τ . Para um grupo de 4 agentes, temos:

$$\nu = 1 + 3\tau + 3\tau^2 + \tau^3. \quad (3.1)$$

O primeiro termo representa o primeiro infectado do grupo. Generalizando para um grupo com tamanho qualquer, podemos escrever o número de infectados da seguinte forma:

$$\nu = \sum_{k=0}^{\alpha} \frac{\alpha!}{(\alpha - k)!k!} \tau^k, \quad (3.2)$$

ou de maneira simplificada (ABBASI et al., 2020; DEPPMAN; SEGUNDO, 2021)

$$\nu = (1 + \tau)^{N-1}. \quad (3.3)$$

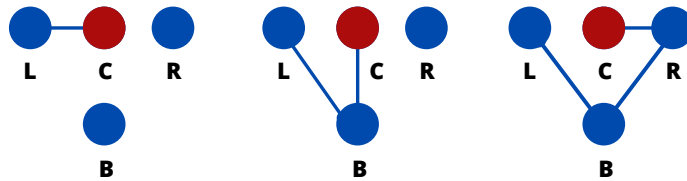


Figura 6 – Modos de contágio em um pequeno grupo de quatro indivíduos ((L)esquerda, (C)central, (D)direita e (B)fundo). As figuras mostram esquematicamente os diferentes modos, diretos (figura esquerda) ou indiretos (figuras central e direita), que um indivíduo suscetível no grupo pode ser infectado a partir de um indivíduo inicialmente infectado (vermelho).

O expoente $\alpha = N - 1$ pode ter duas interpretações: a primeira continua sendo a descrita acima, o número de agentes do grupo menos o primeiro contaminado; a segunda é particularmente mais interessante, pois é possível atribuir ao valor α uma média de contatos próximos que uma pessoa do grupo pode ter. Na segunda interpretação, o valor de α pode variar de acordo com a intensidade da interação no grupo, afetada com ações sociais como quarentenas ou períodos de comemoração.

A situação descrita pela equação 3.3 acontece em um período onde os agentes do grupo conseguem entrar em contato uma ou mais vezes. Durante esse período é estimado o número de infectados e qual é a fração deles que foram infectados pelo primeiro contaminado,

pelo segundo e assim por diante.

3.1.1 Dependência temporal

O passo seguinte foi introduzir um parâmetro temporal e um fator de escala para tentar descrever como será a dinâmica de propagação.

A primeira hipótese é sobre o parâmetro τ . É natural supor que ao decorrer do tempo a chances de um infectado passar o vírus diminui, pois além da interação diminuir, devido aos sintomas, há um tempo bem definido de transmissibilidade. A hipótese é que τ não é fixo, mas varia no tempo e de forma linear.

$$\tau = \kappa \Delta t, \quad (3.4)$$

sendo κ uma constante com dimensão de tempo⁻¹ e $\Delta t = (t_f - t_0)$ a diferença entre o tempo t_f , onde termina a infecção, e tempo inicial t_0 do primeiro contágio do vírus.

A partir disso, a dependência temporal do número de infectados será explícita $\nu \rightarrow i(t)$ e considerado, para melhor compreensão com os modelos epidemiológicos compartmentais. Além disso, identificamos uma população de suscetíveis, $s(t)$, que também dependerá do tempo.

Para introduzir o aspecto fractal do modelo, fazemos uma nova hipótese. O tamanho dos grupos variam conforme mudamos a escala com que observamos o fenômeno. Na figura 7 é mostrado um esquema da alteração da escala e como os grupos são formados. É possível obter grupos de tamanho familiar, onde cada agente é um membro dessa família. Também podemos ampliar a escala e olharmos grupos maiores, por exemplo, bairros onde os agentes agora são as famílias e estender isso para outros tamanhos: empresas, conjuntos habitacionais, cidades, etc..

Neste trabalho será usado o conceito de grupo e agente. O que cada um representa dependerá da escala analisada. Um grupo pode ser formado por regiões onde os agentes são os bairros. Também podemos considerar uma família onde os agentes são os membros dessa família e assim por diante.

3.2 Modelo matemático e aproximações

A partir disso podemos inserir um fator de escala λ para dar suporte a ideia fractal e relacionar o número de pessoas em cada grupo com sua escala. A hipótese baseia-se em considerar o escalonamento do número de suscetíveis da seguinte forma:

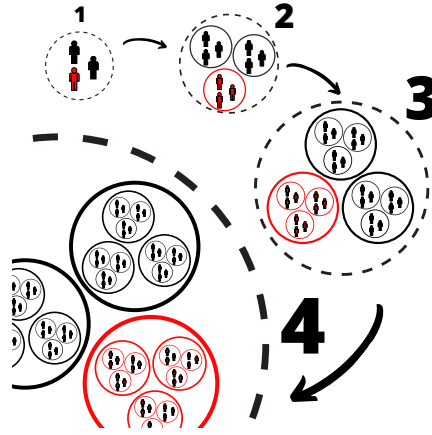


Figura 7 – Ilustração do aspecto fractal do modelo em diferentes escalas. O tamanho do grupo e dos agentes variam conforme a escala muda. No círculo 1, cada agente é formado por uma única pessoa, $N = 2$, enquanto no círculo 4, cada agente é formado por 27 pessoas, $N = 27$.

$$s(t) = \lambda(N - 1). \quad (3.5)$$

Para grupos de famílias o $\lambda = \lambda_0 = 1$, para grupos de bairros possuindo casas com 5 habitantes, em média, $\lambda = 5$, e assim por diante.

Por fim, foi introduzida a transformação $1 - q = (N - 1)^{-1}$, sendo q um parâmetro a ser ajustado. Reescrevendo a equação 3.3 com os novos parâmetros:

$$\nu \rightarrow i(t),$$

$$\tau \rightarrow \kappa(t - t_0),$$

$$N - 1 \rightarrow \frac{1}{1 - q},$$

$$i(t) = \left[1 + (1 - q) \frac{\kappa s(t)(t - t_0)}{\lambda} \right]^{\frac{1}{1 - q}}. \quad (3.6)$$

Com $\frac{s(t)}{\lambda(N - 1)} = 1$.

3.2.1 Uma aproximação para um modelo do tipo Suscetível-Infetado

A equação 3.6 fornece o número de infectados em determinado intervalo de tempo e com ela é possível obter a variação de infectados com relação ao tempo.

$$\frac{di(t)}{dt} = \frac{1}{1-q} \left[1 + (1-q) \frac{\kappa s(t)(t-t_0)}{\lambda} \right]^{\frac{q}{1-q}} \frac{(1-q)}{\lambda} [\kappa s(t) + \kappa(t-t_0)\dot{s}(t)]. \quad (3.7)$$

Simplificando a equação e fazendo $\lambda = \lambda_0 = 1$, obtemos:

$$\frac{di(t)}{dt} = \kappa i^q(t) s(t) + \kappa i^q(t) \dot{s}(t)(t-t_0) = \kappa i^q [s(t) + \dot{s}(t)(t-t_0)]. \quad (3.8)$$

Fazendo as aproximações $s(t) \gg \dot{s}(t)(t-t_0)$ e $\dot{s}(t) = \text{constante}$, podemos incluir a influência desse termo na constante κ e simplificarmos a expressão de $\frac{di(t)}{dt}$.

$$\frac{di(t)}{dt} = \kappa i^q(t) s(t). \quad (3.9)$$

O modelo SI tradicional possui dois compartimentos. Nele toda a população é contaminada e depois recuperada. Usando a equação 3.9 e a condição de uma população com um número fixo N de indivíduos, temos:

$$i(t) + s(t) = N, \quad (3.10)$$

$$\Rightarrow \frac{di(t)}{dt} + \frac{ds(t)}{dt} = 0 \Rightarrow \frac{di(t)}{dt} = -\frac{ds(t)}{dt}. \quad (3.11)$$

Por fim, aplicando o resultado 3.6:

$$\begin{cases} i(t) = \left[1 + (1-q)\kappa s(t)(t-t_0) \right]^{\frac{1}{1-q}} \\ s(t) = s_0 - i(t) \end{cases}. \quad (3.12)$$

As equações 3.13 descrevem o comportamento do vírus em um determinado grupo. Os parâmetros κ e q podem ser obtidos através da comparação com dados reais ou simulação numérica. Cada grupo possui um número inicial de indivíduos s_0 que podem se infectar. Para uma grande população, formada por mais de um grupo, a equação 3.13 deverá ser aplicada em cada um deles.

$$\begin{cases} i_j(t) = \left[1 + (1-q_j)\kappa_j s_j(t)(t-t_{0j}) \right]^{\frac{1}{1-q_j}} \\ s_j(t) = s_{0j} - i_j(t) \end{cases}, \quad (3.13)$$

com o subíndice j especificando o grupo analisado.

Para determinar o comportamento geral da contribuição de cada grupo deverá ser somada.

$$i(t) = \sum_j i_j(t) + c, \quad (3.14)$$

sendo a constante c possíveis casos de infectados de origem exterior ao grupo, como casos de turistas ou pessoas que interagem de alguma forma no grupo, mas estão fora da contagem inicial.

Na figura 8 é mostrado um exemplo do comportamento das equações 3.13. É possível perceber que a curva de infectados não é simétrica como no modelo SI tradicional. Aqui o número de infectados sobe rapidamente no início e vai caindo suavemente.

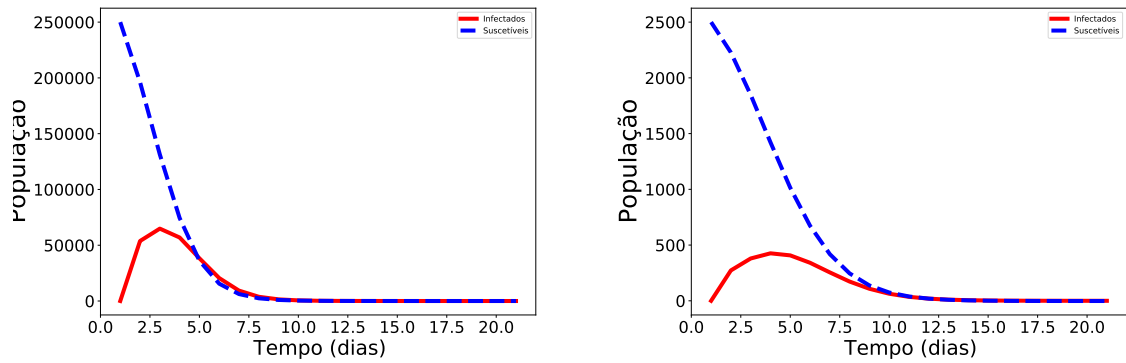


Figura 8 – Cálculos numéricos da evolução temporal da população infectada e suscetível para grupos de: (Esquerda) 250000 indivíduos; (Direita) 2500 indivíduos.

3.2.2 Uma aproximação para um modelo do tipo Suscetível - Infectado - Recuperados

A equação 3.8 diz como o número de infectados varia conforme o tempo passa. No modelo fractal e com vários grupos, denotamos $di_j(t)$ como sendo a taxa de variação do grupo j .

Comparando os resultados obtidos com o modelo SIR tradicional, o primeiro termo da equação 3.8 corresponde aos novos infectados, enquanto o negativo do segundo termo corresponde ao número de novos recuperados. Com essa comparação, a equação 3.8 pode ser reescrita como um sistema de equações, seguindo o seguinte significado:

$$\left\{ \begin{array}{l} \frac{ds_j(t)}{dt} = -\kappa_j i_j^q(t) s_j(t) \\ \frac{di_j(t)}{dt} = \kappa_j i_j^q(t) s_j(t) + \kappa_j i_j^q(t) (t - t_{oj}) \dot{s}_j(t) , \\ \frac{dr_j(t)}{dt} = -\kappa_j i_j^q(t) (t - t_{oj}) \dot{s}_j(t) \end{array} \right. , \quad (3.15)$$

com j representando o grupo do sistema.

Por definição, consideramos $s_j(t_0) = s_0$ como a população de suscetíveis no grupo j no instante inicial t_0 . Assim como no modelo SIR, no modelo fractal, há três estágios durante a epidemia: suscetíveis, infectados e recuperados. No instante inicial, o número de suscetíveis $s_j(t_0)$ é igual ao número de agentes do grupo; $i_j(t_0) = 1$ e $r_j(t_0) = 0$.

O número $s_{0,j}$ é um parâmetro importante a ser encontrado através do ajuste com os dados. Os parâmetros κ_j , q e $t_{0,j}$ também podem ser encontrados através de ajustes e se relacionam, como será mostrado mais abaixo.

As duas últimas equações do sistema 3.15 estão em função da variação do número de suscetíveis. Isso permite abrir caminho para escrevê-las de forma mais simples, através de apenas duas equações. Para isso, $\dot{r}(t)$ é escrita como:

$$\frac{dr_j(t)}{dt} = -\kappa_j i_j^q(t)(t - t_{0j})\dot{s}(t) = \kappa_j i_j^q(t)(t - t_{0j})\langle |\dot{s}(t)| \rangle. \quad (3.16)$$

Para simplificar as equações foi usado uma aproximação para a variação do número de suscetíveis. Agora fazemos uma diferenciação entre os termos Δt e $t - t_0$. O primeiro é a largura do pico, em unidade de tempo, de cada grupo, representando o período que o grupo teve infectados. O segundo é o tempo em que o grupo está contaminado, a partir de um tempo inicial t_0 . A relação entre eles é $(t - t_0) \leq \Delta t$. Aproximando $(t - t_0)$ por uma constante:

$$\langle \kappa(t - t_0) \rangle = \kappa \langle (t - t_0) \rangle \simeq \kappa \frac{\Delta t}{2}. \quad (3.17)$$

Considerando uma aproximação linear para o termo $\dot{s}(t) = \text{constante}$ e sabendo que $\dot{s}(t) < 0$, temos:

$$\frac{dr_j(t)}{dt} = i_j^q(t) \kappa_j (t - t_{0j}) \langle |\dot{s}(t)| \rangle = \kappa^* i_j^q(t). \quad (3.18)$$

Usando a definição $\tau = \kappa \Delta t$

$$\kappa^* = \frac{\tau}{2} \langle |\dot{s}(t)| \rangle. \quad (3.19)$$

No modelo FSI toda a população de suscetíveis, s_0 , é contaminada. Para que o modelo possa ser comparado com o SIR tradicional, é necessário incluir um fator de saída no termo $\frac{di(t)}{dt}$. Com as aproximações feitas acima e as mudanças nas variáveis, podemos escrever as equações do modelo FSIR de forma análoga, como na equação 3.20.

$$\left\{ \begin{array}{l} \frac{ds_j(t)}{dt} = -\kappa_j i_j^q s_j \\ \frac{di_j(t)}{dt} = \kappa_j i_j^q s_j - \kappa_{qj}^* i_j^q \\ \frac{dr_j(t)}{dt} = \kappa_{qj}^* i_j^q \end{array} \right. \quad (3.20)$$

Embora esse resultado venha de princípios diferentes, fica clara a semelhança dos modelos, diferindo apenas pela forma como $i(t)$ evolui no tempo. As equações 3.20 foram propostas de maneira independente e com uma abordagem distinta em (TSALLIS; TIRNAKLI, 2020; TIRNAKLI; TSALLIS, 2020). Espera-se que o expoente q quantifique e modele de forma mais realista a complexidade da dinâmica do vírus.

Outro aspecto importante é que o modelo FSIR generaliza o SIR tradicional, pois quando $q \rightarrow 1$ o conjunto de equações 3.20 retoma a forma do modelo SIR.

A figura 10 representa uma simulação com todos os parâmetros fixos, excetuando o valor de q . O resultado mostra que o tempo de contaminação, Δt , em cada grupo, não é linear com q . Conforme aumentamos a interação, a infecção acontece em um tempo menor. Isso sugere a possibilidade de o modelo ser sensível a mudanças no comportamento social, como a não adesão ao distanciamento social. Outro comportamento interessante é a relação em lei de potência entre a largura dos picos e o tamanho da população inicial, mostrado na figura 10. Se ajustarmos a reta na forma de $\Delta t = \Delta t_0 s_0^\alpha$, obteremos: $\alpha \approx 0.21$. O Δt é o período entre o primeiro infectado e o final da infecção.

Assim como no caso do modelo SIR tradicional, podemos calcular o análogo de R_0 , utilizando o mesmo procedimento. Para que a propagação do vírus ocorra é necessária a seguinte condição:

$$\frac{di_j^q(t)}{dt} > 0, \quad (3.21)$$

$$\Rightarrow i_{0j}^q (\kappa_j s_{0j} - \kappa_j^*) > 0. \quad (3.22)$$

Nas condições iniciais, i_{0j} e s_{0j} são o número de infectados e suscetíveis, respectivamente, no momento t_0 . Definindo R_{0j} como sendo a taxa de reprodução no grupo j .

$$R_{0j} = \frac{\kappa_j s_{0j}}{\kappa_j^*} > 1. \quad (3.23)$$

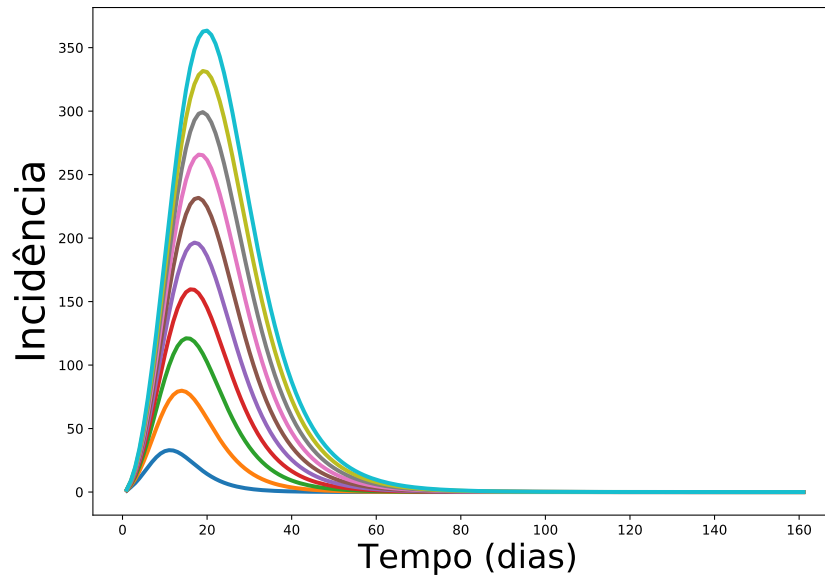


Figura 9 – Sequência de dez curvas mostrando a variação das curvas de incidência em função da população. Os valores da população inicial de suscetíveis, s_0 , variou entre 500 a 10000 pessoas, κ foi fixado por uma lei de potência: $\kappa = 0.017s_0^{-0.833}$; $\kappa^* = 0.05$ e $q = 0.75$.

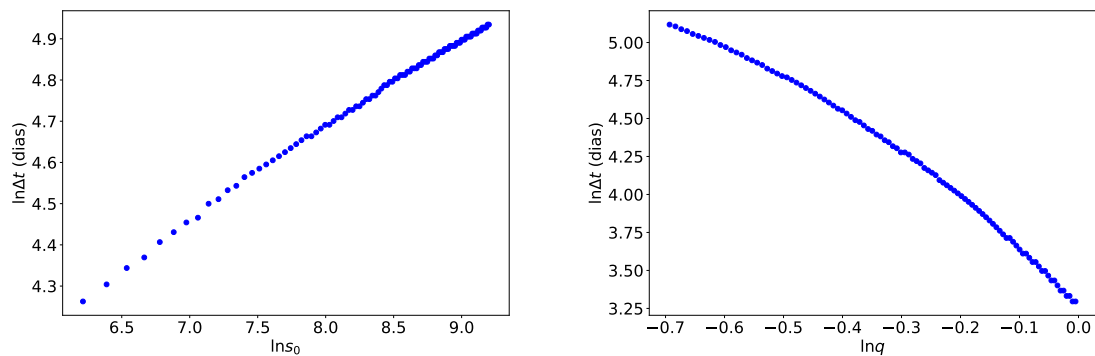


Figura 10 – (Esquerda) Comportamento da largura de cada pico em função da população inicial de suscetíveis s_0 . Os parâmetros são os mesmos que o da figura 9. (Direita) Comportamento da largura de cada pico em função do parâmetro q . Nessa simulação a população inicial foi mantida fixa e com valor 800; $\kappa = 0.017s_0^{-0.833}$; $\kappa^* = 0.05$ e $0.5 < q \leq 1$.

Se $R_{0j} > 1$ o vírus irá se espalhar e contaminar a população, caso contrário não haverá epidemia.

Tanto o modelo 3.13 e 3.20 podem ser resolvidas de forma recursiva e ajustadas em dados reais para determinar os parâmetros κ e κ^* .

4 Metodologia

Neste capítulo descrevo as etapas e ferramentas usadas para validar o modelo fractal. A metodologia usada foi um conjunto de análises que contou com pesquisa e tratamento de dados, ajustes das equações que descrevem a dinâmica do vírus com os dados de novos infectados (incidência) e discussão do significado dos parâmetros do modelo.

No primeiro momento o objetivo foi verificar o comportamento do número de infectados com o tamanho da população. Esse procedimento buscou testar a hipótese de que essa relação é livre de escala. Para isso, foram escolhidos dados públicos, disponibilizados por organizações de saúde de diversos países. Em seguida, os dados foram ajustados com uma distribuição do tipo lei de potência.

O trabalho seguiu utilizando dados sobre São Caetano, o que permitiu, além do ajuste, uma ideia do comportamento dos infectados no espaço geográfico da cidade.

Os ajustes foram realizados com um programa desenvolvido durante este trabalho e o critério utilizado foi a minimização do χ^2 . Os parâmetros otimizados foram comparados com resultados de outros métodos, como a quantificação da adesão ao isolamento social durante o período estudado.

4.1 Dados

4.1.1 USA, São Paulo, China e Europa

Para a primeira parte da análise foi utilizado dados públicos, disponibilizados pelos governos dos países ou por instituições de pesquisas. Os dados sobre a China foram coletados no site The Humanitarian Data Exchange (EXCHANGE, 2020); Center for Disease Control (CDC) e Worldometers usados para os USA (CONTROL, 2020); em São Paulo foi usado o site do governo do estado (ESTADO, 2020). Como os países da Europa são pequenos em relação aos demais países, os dados sobre a França (FRANCA, 2020), Alemanha (ALEMANHA, 2020), Espanha (ESPANHA, 2020) e Itália (ITALIA, 2020) foram utilizados juntos.

4.1.2 São Caetano

Já na segunda parte do trabalho, ajuste dos parâmetros, as informações utilizadas para testar o modelo foram obtidas a partir de um programa da cidade de São Caetano, Brasil, que tinha por objetivo registrar as informações sobre a propagação do vírus na cidade, cedida pela Faculdade de Medicina da USP (FMUSP). A primeira etapa do

programa consistia na entrevista por meio do preenchimento de formulário online ou entrevista telefônica e a partir disso, feita uma triagem para identificar possíveis casos de COVID. Após a triagem, uma equipe acompanhava o entrevistado e realizava um acompanhamento por meio de testes, visitas na residência e ligações telefônicas. Esse banco de dados coletado durante o programa é detalhado, contendo o sexo, idade, resultado do teste para detectar a presença da infecção e as coordenadas geográficas das residências dos entrevistados (LEAL et al., 2021). As informações consideradas foram de pessoas com idade superior a 12 anos e com teste PCR positivo e o período estudado foi entre 18-03-2020 e 19-11-2020.

4.1.3 Isolamento social

Para quantificar a adesão da população ao isolamento social, o estado de São Paulo adotou o índice de Isolamento Social (DADOS...), Os dados são fornecidos por 4 operadoras de celular: VIVO, TIM, CLARO e Oi. Elas fornecem a localização dos usuários através do sinal de celular.

A casa de cada indivíduo é determinada pelo sinal entre 22h e 2h, supondo que nesse horário as pessoas estão em suas residências. A partir dessa localização, é estabelecido um raio de 200m que corresponde a área do isolamento. Quando o sinal vem de fora da área estabelecida entende-se que o isolamento foi quebrado.

Outro banco de dados é o fornecido pela empresa de tecnologia InLoco. O procedimento usado para determinar a quebra do isolamento é semelhante ao anterior, porém com um raio de 460m.

4.2 Programa

O programa usado para analisar os dados foi escrito em linguagem python e contou com a utilização do método de minimização através da biblioteca iminuite (JAMES; ROOS, 1975). Ela permite comparar uma função de muitas variáveis com os dados, minimizar os parâmetros e estimar as incertezas dos ajustes por meio do comando `migrad`, usando um algoritmo proposto por Fletcher (FLETCHER, 1970).

O objetivo do programa foi fazer a leitura e separação dos dados, plotar e ajustar com a função escolhida.

Como mostrado nas equações 3.13 e 3.20, o modelo usa um sistema de equações diferenciais acopladas e isso requer a escolha de umas das funções para o ajuste. Como a ênfase dos dados foi no número de novos casos, conhecido na literatura como incidência σ , a função escolhida foi $i(t)$ para o caso de 3.13 e o negativo da variação dos suscetíveis para 3.20. O sinal negativo indica que a incidência é obtida a partir do módulo do número da variação do compartimento dos suscetíveis. Isso porque na equação dos infectados há dois termos: o primeiro se refere aos agentes que chegam do compartimento de suscetíveis; e o segundo

é o termo que vai para os recuperados. No programa foi usado o método de Euler para fazer as interações (BURDEN; FAIRES; BURDEN, 2015).

$$-\frac{ds(t)}{dt} = \kappa i_j^q s(t) = \sigma. \quad (4.1)$$

Como explicado na seção 3.1, o número total de suscetíveis, infectados ou recuperados é a soma das contribuições de cada grupo. Considerando N grupos e um conjunto de dados sobre a incidência σ , a função a ser ajustada seria do tipo:

$$\sigma = \sum_{j=1}^N \kappa i_j^q s_j(t). \quad (4.2)$$

Repare que para obter σ é necessário calcular as interações de todo o sistema 3.20. Com isso é possível obter, através do método dos mínimos quadrados (MMQ), o conjunto de parâmetros s_{0j} , κ_j , k_j^* , q_j e t_{0j} .

4.3 Etapas do ajuste

4.3.1 Lei de potência nos dados do COVID

Para prosseguir com a ideia de um modelo fractal, um banco de dados foi montado com a atualização quase diária de dados sobre a evolução de novos infectados pelo mundo. Conforme indicado na descrição dos dados, usamos informações oficiais para verificar a relação entre o número de novos casos e o tamanho da população dos respectivos países. Era esperado encontrar uma relação do tipo $Y \propto N^\beta$ que caracteriza uma grande quantidade de grandezas sociais. A partir da confirmação desse comportamento a característica fractal pode ser considerada de forma mais evidente.

4.3.2 Teste inicial com o modelo SI fractal (FSI)

4.3.2.1 Ajuste inicial

Inicialmente os dados foram separados de modo a considerar apenas o número de indivíduos com resultado do teste PCR positivo e com idade acima de doze anos. Esse grupo foi identificado com infecção inicial a partir do dia dos primeiros sintomas, informação relatada pelo próprio entrevistado que consta no banco de dados.

Para entender melhor o comportamento dos dados e simplificar o ajuste, os dados foram separados em bins de 30, 15, 7 e 4 dias. Isso permitiu um ajuste gradual com um número diferente de grupos, como mostrado no capítulo dos resultados.

Nessa primeira etapa o parâmetro q foi fixado, enquanto κ , t_0 e s_0 otimizados através do MMQ usando a equação 3.13. Foram feitos ajustes com 12 grupos de tamanhos variados

que resultou em um conjunto de parâmetros. Por fim, isso possibilitou encontrar um comportamento do tipo $\kappa_j = \kappa_0 s_{0j}^\beta$ e $\Delta t_j = \Delta t_0 s_{0j}^\alpha$, com $\Delta t_j =$ tempo em que há infectados no grupo, α e β constantes determinadas pelo ajuste em forma de lei de potência.

4.3.2.2 Ajuste usando $\kappa_j = \kappa_0 s_{0j}^\beta$

Com os ajustes preliminares feitos, o valor de κ foi fixado de acordo com uma lei de potência, $\kappa_j = \kappa_0 s_{0j}^\beta$ e os parâmetros t_0 , s_0 e q deixados livres. Os dados foram separados por bins de 30, 15, 7, 4 dias e usados em ajustes com as equações 3.13. O primeiro ajuste foi realizado usando bins de 30 dias. Isso possibilitou o uso de apenas um grupo (pico). Em seguida, o bin foi reduzido em 15 dias e ajustado com 2 grupos. O procedimento se repetiu com bins de 7 e 4 dias e 3 e 6 grupos, respectivamente. Em cada etapa da divisão, foi realizada a tentativa de acrescentar grupos, porém o programa ajustava alguns picos como aproximadamente uma constante, ou seja, muito achatados.

Devido ao comportamento dos dados sugerirem dois períodos onde o número de novos casos variou significativamente, foi separado dois grandes períodos, classificados como 1° onda e 2° onda. O objetivo dessa divisão foi comparar o comportamento do parâmetro q nos dois períodos e verificar se há alguma relação com a interação entre os agentes, através de dados sobre o isolamento social.

4.3.3 Localização geográfica dos novos casos

Uma das possibilidades que os dados de São Caetano trouxe para a análise foi a de mapear e acompanhar a propagação do vírus por meio da localização geográfica. Com acesso às longitudes, latitudes e o tempo do início dos sintomas foi possível verificar como os clusters de contaminação surgiram e se há alguma ligação entre eles.

Também foi utilizado um algoritmo para identificar se um novo caso pertence ao cluster e estabelecer uma relação de distância entre eles.

A visualização foi feita através de gráficos da superfície da cidade. A distância entre cada coordenada é dada por:

$$d_{mn} = \sqrt{(|\vec{r}_m| - |\vec{r}_n|)^2}, \quad (4.3)$$

sendo d_{mn} a distância entre o ponto n com vetor posição \vec{r}_n e o ponto m com vetor \vec{r}_m . Por definição:

$$\vec{r} = R_T L_o \hat{i} + R_T L_a \cos(L_o) \hat{j}. \quad (4.4)$$

Com a longitude L_o na direção \hat{i} (norte-sul) e latitude L_a na direção \hat{j} (leste-oeste) medidos na superfície terrestre de raio R_T .

A medida de interesse é a proximidade dos novos casos com os já infectados. Seja I_j a coordenada do centro de um cluster com $j = 1, 2, \dots$, a distância do cluster mais próximo será:

$$d_m = \min_{\substack{j \in I_j \\ m \neq j}} \{d_{mj}\}. \quad (4.5)$$

Para cada novo caso é atribuído uma probabilidade p de ele pertencer ou não a um determinado cluster. Seja L um valor entre 0 e 1, se $p > L$ para um determinado caso, ele será considerado nesse cluster.

A probabilidade do ponto m pertencer ao cluster j será:

$$p_{jm}(\vec{r}) = A_j d_{mj}^{-\delta} (t - t_j)^{-\omega}, \quad (4.6)$$

A_j , δ e ω são constantes.

Além disso;

$$\int_0^\infty p_{jm}(\vec{r}) d\vec{r} = 1. \quad (4.7)$$

5 Resultados

5.1 Ajuste com lei de potência

Considerando Y como o número total de infectados e N a população da região; C uma contante de normalização e β o expoente da lei de potência; temos:

$$Y(N) = CN^\beta. \quad (5.1)$$

Para o ajuste usamos a forma liner log-log:

$$\ln Y(N) = \beta \ln N + \ln C. \quad (5.2)$$

Os resultados da figura 11 confirmam um comportamento em forma de lei de potência para as regiões analisadas. Esses resultados demonstram que a propagação do vírus segue um comportamento super linear em USA, China e São Paulo, pois $\beta > 1$. Os valores do expoente ficam próximos de 1.15, resultado encontrado em muitas outras análises (BATTY; LONGLEY, 1994; SCHLÄPFER et al., 2014; CHOWELL et al., 2003). A característica encontrada é de extrema importância devido ao fato de podermos ter um resultado quantitativo sobre um comportamento de origem social, econômico e de infraestrutura. Esses resultados sugerem uma estrutura fractal e o próximo estudo tem como objetivo testar o modelo com dados mais precisos.

5.2 Teste inicial com o modelo FSI

Os ajustes foram realizados usando um programa escrito em linguagem python. A figura 12 mostra o resultado dos ajustes.

A incerteza dos dados atribuída a contagem n de novos casos é igual a \sqrt{n} . O programa consegue, a partir das incertezas dos dados, estimar as incertezas nos parâmetros e possibilitar o melhor ajuste usando o critério do menor χ^2 .

Na tabela 3 os valores dos parâmetros são apresentados com suas respectivas incertezas. O valor total da população dos grupos variou entre 2615 pessoas para bins de 4 dias e 2054 para bins de 30 dias.

Conforme se diminui o tamanho dos bins, mais detalhes sobre os dados aparecem. Para um espaçamento maior, um pico (grupo) é o suficiente para descrever os dados. Se a largura for dividida pela metade, um pico apenas não é o suficiente, mas dois picos fazem um melhor ajuste, usando o critério do menor χ^2 . Repetindo o procedimento, dividindo a

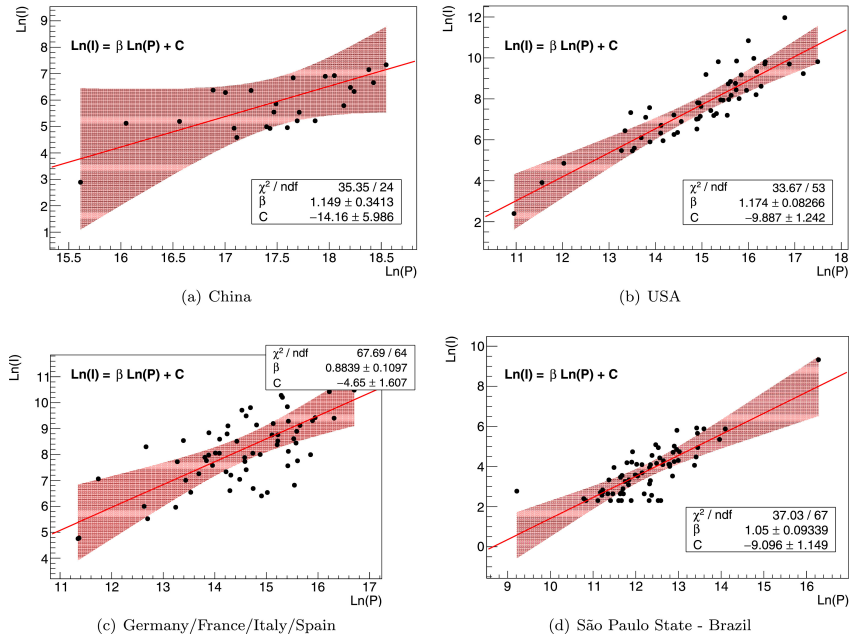


Figura 11 – Comportamento do número de casos de contaminação em função da população nas regiões a) China, b) EUA c) França, Alemanha, Espanha e Itália juntos e d) Estado de São Paulo, por população da cidade. Os intervalos de confiança são desenhados com nível de confiança de 95%.

largura dos bins e duplicando o número de picos, 4 ao todo, aparece no ajuste um grupo que tem um comportamento aproximadamente constante. Isso se repetiu nos demais ajustes e sugeriu uma tentativa usando um background c , constante, e o resultado do χ^2 diminuiu. Essa constante está sendo interpretada como possíveis casos de origem externa à cidade ou mesmo de pessoas fora dos grupos. A partir disso, a constante c foi colocada em todos os ajustes.

O valor do expoente da lei de potência para diversas grandezas envolvendo comportamentos e dinâmicas sociais é estudado e tem um valor aproximado de 1.15, o que é consistente com os ajustes da 11. Para fixar o valor do parâmetro q usamos a equação 3.6, igualando o expoente a 1.15.

$$\frac{1}{1-q} = 1.15, \quad (5.3)$$

$$\Rightarrow q = 0.13. \quad (5.4)$$

Fixado o valor de q é possível verificar, através da figura 13 o comportamento de κ e $\Delta t = (t_j - t_{0j})$ em comparação ao tamanho de cada grupo. A ideia é verificar se o comportamento escala com uma lei de potência. Para isso os valores foram plotados em

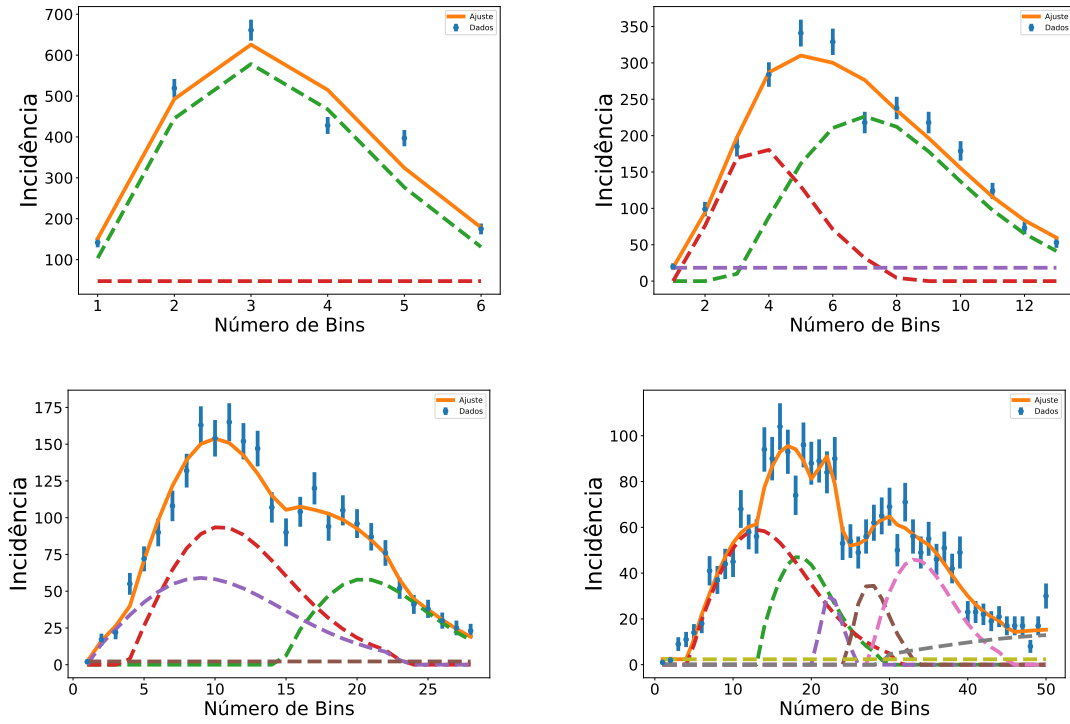


Figura 12 – A figura acima representa o ajuste utilizando a equação 3.13 para um conjunto de dados separados em diferentes bins. A figura superior esquerda representa bins de 30 dias; superior direita bins de 15 dias; inferior esquerda bins de 7 dias; inferior direita bins de 4 dias. As linhas tracejadas indicam a contribuição de cada grupo enquanto a linha contínua é a soma de todas as contribuições. A linha horizontal tracejada foi deixada para um melhor ajuste e pode mostrar uma contaminação de fora dos grupos.

escala logaritma e feito um ajuste linear.

Conforme o ajuste o ajuste:

$$\kappa_j = \kappa_0 s_j^\beta \Rightarrow \beta = -1.207 \pm 6.5E - 2 \quad (5.5)$$

Com $\kappa_0 = 1.50 \pm 6.6E - 1$. Também obtemos os valores para Δt :

$$\Delta t_j = \Delta t_0 s_j^\alpha \Rightarrow \alpha = 0.833 \pm 1.3E - 2 \quad (5.6)$$

Com $\Delta t_0 = 0.4252 \pm 8.9E - 3$.

No modelo FSI há apenas uma população de suscetíveis $s(t)$ e outra de infectados. O parâmetro inicial, $s(t_0)$, obtido no ajuste e mostrado na tabela 3, possui um significado diferente do modelo SI tradicional. Esse valor não é a população total da cidade ou país, mas representa o número de contatos próximos no qual o vírus irá se espalhar. Isso explica os valores de $s(t_0)$ serem menores do que o valor da população de São Caetano.

Tabela 3 – Parâmetros da equação 3.13 ajustados com o conjunto de dados separados por bins de 4, 7, 15 e 30 dias, como descrito em 4.3.2.1. O período estudado foi entre 18-03-2020 e 19-09-2020 mostrado na figura 12. O background c é uma constante que representa contaminação externa.

Bin	s_{oj}	κ_j	Δt_j
	852(191)	6(1)E-04	103
	418(80)	1.6(4)E-03	60
4-dias	100(16)	1.2(2)E-02	18
$c = 2 \pm 1$	183(40)	4.7(8)E-03	32
$s_{total} = 2615(233)$	459(36)	1.2(1)E-03	68
	603(92)	6(1)E-05	125
7-dias	772(180)	2.5(6)E-04	149
$c = 2 \pm 1$	985(152)	3.7(8)E-04	130
$s_{total} = 2310(241)$	553(50)	4.5(5)E-04	109
15-dias	671(44)	7.0(4)E-04	92
$c = 18 \pm 1$	1453(110)	2.1(2)E-04	174
$s_{total} = 2124(118)$			
30-dias	2054(79)	1.8(1)E-04	196
$c = 48 \pm 11$			

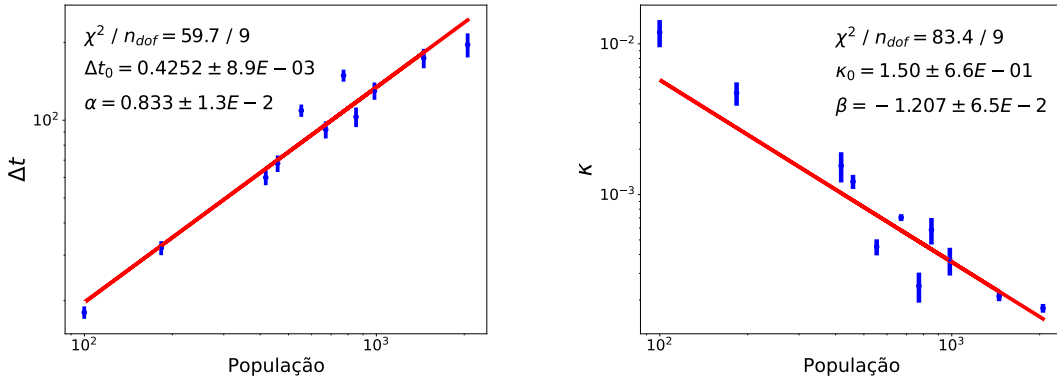


Figura 13 – Ajuste linear dos parâmetros Δt e κ em função do número de indivíduos do grupo (contatos) na forma de lei de potência.

5.3 Teste inicial com o modelo FSI usando lei de potência

Nessa etapa foi usado o resultado obtido do ajuste mostrado na figura 13 para restringir o parâmetro κ na forma de lei de potência: $\kappa = \kappa_0 s^\beta$. Isso porque o objetivo foi verificar a influência de q nos ajustes.

Na estatística de Tsallis o índice q tem um significado bem definido para os sistemas físicos. Ele denota o grau de interação entre os agentes. No modelo fractal isso também acontece quando ele se aproxima de 1. Veja que o expoente da equação 3.13, $\frac{1}{1-q}$, tende ao infinito. Como explicado no final da seção 5.2, o expoente $N - 1$ representa o número de contatos próximos. Espera-se que quanto maior for a interação, maior será esse expoente que com

a mudança de variável, fica $\frac{1}{1-q}$. Esse crescimento faz com que o número de infectados aumente, além de diminuir a largura dos picos, o que, na prática, representa uma situação epidêmica onde há muitos infectados em um curto período, podendo causar forte impacto na saúde pública.

Na figura 14 é mostrado o resultado dos ajustes com $\kappa_j = \kappa_0 s_j^{-0.833}$. Os parâmetros κ_0 , s_{0j} , t_{0j} e q ficaram livres para serem ajustados na curva dos dados.

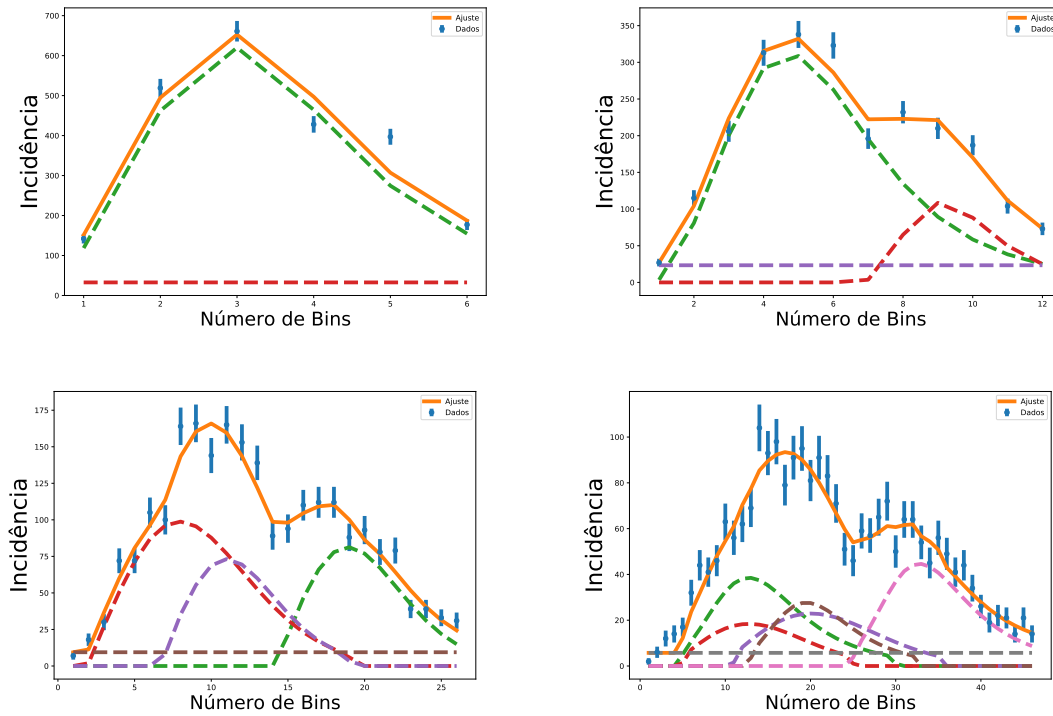


Figura 14 – A figura acima representa o ajuste utilizando a equação 3.13 para um conjunto de dados separados em diferentes bins e fixando a lei de potência $\kappa = \kappa_0 s_0^{-0.833}$. A figura superior esquerda representa bins de 30 dias; superior direita bins de 15 dias; inferior esquerda bins de 7 dias; inferior direita bins de 4 dias. As linhas tracejadas indicam a contribuição de cada grupo enquanto a linha contínua é a soma de todas as contribuições. A linha horizontal tracejada foi deixada para um melhor ajuste e pode mostrar uma contaminação de fora dos grupos.

Como no caso anterior, a complexidade aumenta conforme o tamanho dos bins diminuem. O mesmo número de picos foi utilizado e o ajuste ficou com um χ^2 menor, quando comparado com o do item anterior.

Acrescentando dados da segunda onda, novos ajustes foram feitos com bins de 4 e 7 dias. A ideia é comparar os valores de q e testar se o modelo consegue capturar uma maior interação. Como a segunda onda tem picos com largura menor, é de se esperar um aumento nas interações, medido por q . Na tabela 4 é mostrado o resultado dos ajustes e na figura 17 os valores de q e sua média.

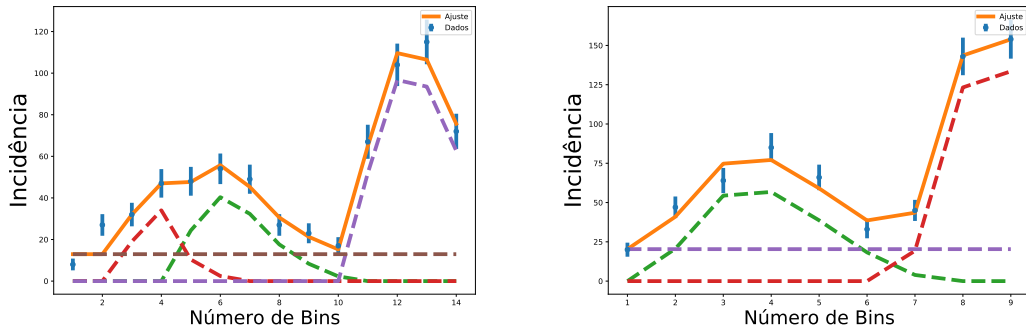


Figura 15 – A figura acima representa o ajuste utilizando a equação 3.13 para um conjunto de dados da segunda onda separados em diferentes bins e fixando a lei de potência $\kappa = \kappa_0 s_0^{-0.833}$. A figura esquerda representa bins de 4 dias; a direita bins de 7 dias.

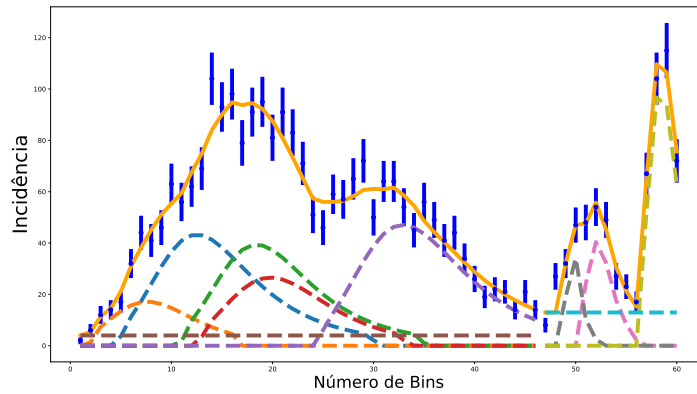


Figura 16 – A figura acima representa o ajuste utilizando a equação 3.13 para um conjunto de dados da primeira e segunda onda separados em bins de 4 dias e com lei de potência $\kappa = \kappa_0 s_0^{-0.833}$.

Os resultados dos ajustes mostram um aumento do parâmetro q durante a segunda onda. Isso confirma que é possível, com esse modelo, identificar alterações na intensidade das interações dos grupos. Na figura 17 as médias desses valores são mostradas com linhas contínuas. A imagem à esquerda representa os valores ajustados enquanto que na direita são valores estimados sobre o isolamento social.

Os valores encontrados na primeira onda são $0 < q_j < 0.6$ enquanto na segunda $q_j > 0.6$. Para o valor médio $q \approx 0.5$ obtemos $N = 3$ no começo da pandemia. Já na segunda onda temos $q \approx 0.8$ e $N = 6$. Esse resultado é coerente com a diminuição do índice de isolamento.

Na primeira parte deste trabalho usamos a hipótese de $\tau = \kappa \Delta t = \text{constante}$. Ao analisar a tabela 4 verificamos que esse valor sofre pequenas variações com média $\langle \tau \rangle = 0.03625(7)$, sendo uma boa aproximação. No entanto, quando olhamos para a segunda onda, essa

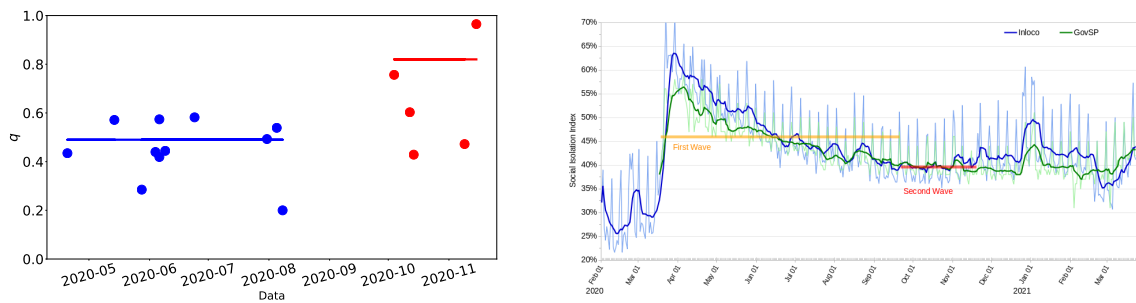


Figura 17 – Na esquerda os valores ajustados para o parâmetro q para a primeira onda (azul) e segunda onda (vermelho). A barra de erro não é visível, pois são pequenas em relação à escala do desenho, ver tabela 4. A linhas contínuas representam os valores médios ajustados.. A imagem à direita mostra a média estimada do isolamento social no período analisado.

aproximação não é o suficiente. Uma possível explicação pode ser a mudança não apenas do comportamento das pessoas, analisada pelo índice q , mas também no próprio vírus. Um exemplo disso, é o surgimento da variante P2 que se originou da linhagem B.1.1.28 em outubro de 2020 (MICHELON, 2021).

5.3.1 Análise geográfica

No início da pandemia houve 5 casos com grande distanciamento e a maioria dos demais se aglutinou em torno dessas coordenadas. Na primeira tentativa, acrescentamos os novos casos nesses clusters através do critério descrito na seção 4.3.3 com $L = 0.4$. Posteriormente, o cluster central foi dividido e ao todo ficaram 6 clusters e $L = 0.55$. Os ajustes de δ e ω resultaram nos valores 3.5 e 2, respectivamente. Um ponto importante nesta análise foi a relação da incidência com a distância entre os clusters. Na figura 18 a maior parte dos casos estão próximos, com distância média menor que 100m. Isso mostra a importância de modelos epidemiológicos que considerem uma população não homogênea.

5.3.2 Comparação do modelo FSIR e SIR

O modelo FSIR é uma generalização para os dois modelos tradicionais, abordados no texto. Quando $q = 1$ e $\kappa^* = 0$ temos o modelo SI. Se $q = 1$ e $\kappa^* \neq 0$ retornamos para o SIR. Para $q \neq 1$ é possível termos $\kappa^* = 0$ ou $\kappa^* \neq 0$, transformando as equações no FSI e FSIR, respectivamente.

Com essa generalização, temos disponíveis 4 modelos, onde podemos comparar a adequação deles com os dados de São Caetano. Para isso, usamos a incidência da primeira onda separada por bins de 4 dias e ajustamos as curvas para otimizar os parâmetros.

Na figura 19 é mostrado ajustes considerando $\kappa^* = 0$ e κ^* livre. O resultado é que mesmo

Tabela 4 – Parâmetros otimizados usando a equação 3.13 e dados da primeira e segunda onda. O período analisado está entre 19-09-2020 e 19-11-2020. O parâmetro χ^2 está sendo comparado com o número de graus de liberdade. A coluna Δt_j representa a largura de cada fico, ou seja, o tempo de que cada grupo fica contaminado. O τ é o produto de κ e Δt .

Parâmetros Globais	s_{oj}	q_j	Δt_j	τ_j
1° onda				
Bin 4 dias	554(1)	0.5712(4)	82(1)	0.0297(2)
$\kappa_0 = 0.06998(49)$	163(1)	0.435(2)	50(1)	0.0501(3)
$c = 4 \pm 1$	485(1)	0.57(1)	78(1)	0.0314(2)
$\chi^2/n_{\text{dof}} = 36/29$	320(1)	0.4444(2)	69(1)	0.0395(3)
$s_{\text{total}} = 2084(2)$	652(1)	0.5388(3)	89(1)	0.0282(2)
Bin 7 dias	569(35)	0.19(1)	76(3)	0.0398(9)
$\kappa_0 = 0.08723(43)$	1170(1)	0.2850(4)	90(1)	0.0281(3)
$c = 9 \pm 2$	402(27)	0.44(7)	72(1)	0.0401(1)
$\chi^2/n_{\text{dof}} = 35/15$				
$s_{\text{total}} = 2141(44)$				
Bin 15 dias	1701(355)	0.42(4)	134(1)	0.020(4)
$\kappa_0 = 0.07(1)$	342(150)	0.5(2)	67(1)	0.04(1)
$c = 23 \pm 1$				
$\chi^2/n_{\text{dof}} = 13/4$				
$s_{\text{total}} = 2043(385)$				
Bin 30 dias	2255(173)	0.6(2)	190(1)	0.012(6)
$\kappa_0 = 0.04(2)$				
$c = 33 \pm 29$				
$\chi^2/n_{\text{dof}} = 34/1$				
2° onda				
Bin 4 dias	114(11)	0.429(4)	15(1)	0.14(1)
$\kappa_0 = 0.48071(15)$	67(1)	0.756(8)	9(1)	0.1370(7)
$c = 13 \pm 1$	376(1)	0.472(1)	24(1)	0.0819(2)
$\chi^2/n_{\text{dof}} = 14/3$				
$s_{\text{total}} = 557(11)$				
Bin 7 dias bin	217(2)	0.602(1)	33(1)	0.0645(9)
$\kappa_0 = 0.172520(13)$	408(1)	0.9647(7)	34(1)	0.03922(4)
$c = 23 \pm 1$				
$\chi^2/n_{\text{dof}} = 6/1$				
$s_{\text{total}} = 625(2)$				

deixando todos os parâmetros livres, o menor χ^2 é obtido quando $\kappa^* \simeq 0$ e $q \simeq 0.56$. Isso mostra que o modelo flexibiliza as possibilidades do ajuste e pode descrever melhor os dados. Com todos os parâmetros livres, poderíamos também obter resultados com o FSIR ou SIR, mas o χ^2 aumentaria.

A população desses ajustes ficou bem abaixo do real, aproximadamente 162 mil habitantes para São Caetano, o que é interpretado no modelo fractal, como o número de contatos próximos.

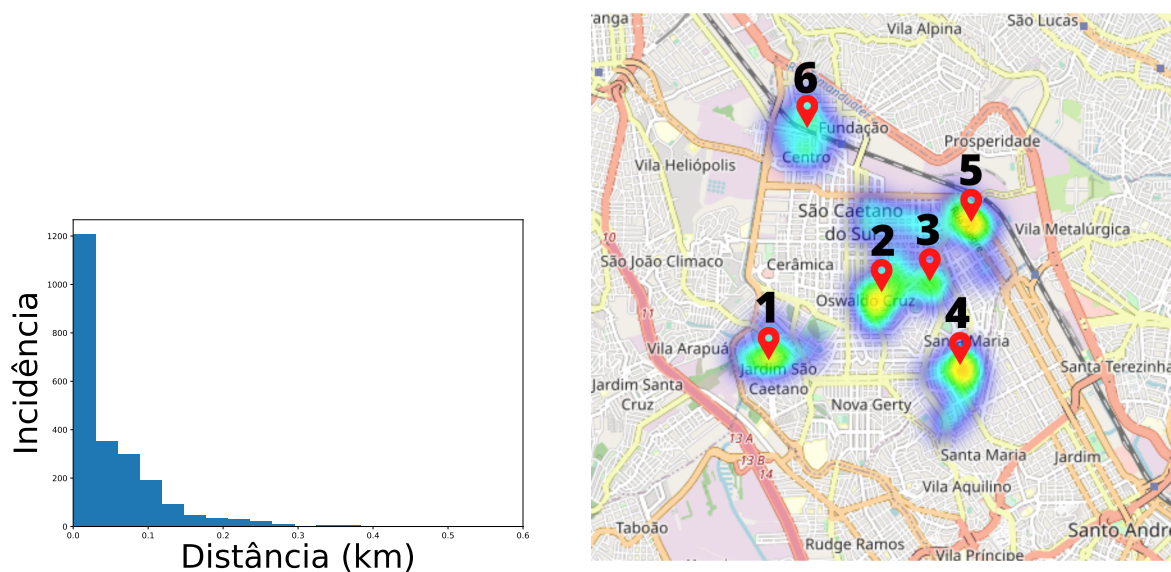


Figura 18 – Na esquerda, a distribuição da menor distância entre um novo caso e todos os anteriores. A análise foi feita entre 22-03-2020 e 30-09-2020. Na direita, a distribuição de novos casos, usando o método de separação por clusters, através de mapa de calor. Cada cluster foi associado a uma determinada concentração de pessoas. Escola Técnica - próximo do Cluster 1, Parque - próximo do Cluster 4, Estádio Municipal - próximo dos Clusters 2 e 3, Hospital - próximo do Cluster 5, Estação de Trem Universidade Municipal de São Caetano - próximo do Cluster 6, Fábrica de carros - próximo dos Clusters 5 e 6.

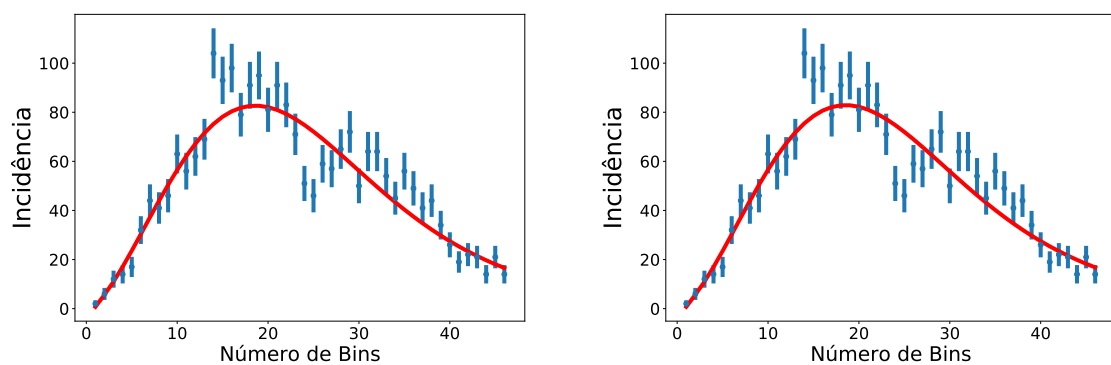


Figura 19 – Ajuste com dados em bins de 4 dias: (esquerda) $k^* = 0$ e os demais parâmetros livres; (direita) todos os parâmetros livres. Valores dos ajustes estão nas duas primeiras linhas da tabela 5.

Para verificar a influência do tamanho da população, mais ajustes foram feitos, mas agora com a população fixa. Os resultados são mostrados na figura 20 e nas linhas 3 a 5 da tabela 5. Conforme aumentamos o valor da população para usar algo mais próximo da população real da cidade, o valor de κ^* aumenta, se afastando de zero e o índice $q \rightarrow 1$, aproximando para um comportamento do tipo SIR. Contudo, esse método não melhora o

ajuste, como mostrado na última coluna da tabela 5.

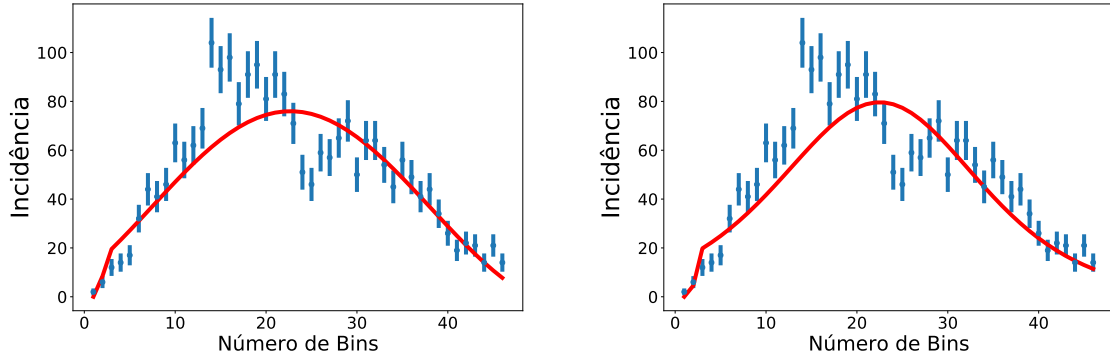


Figura 20 – Ajuste com a população fixa em $5.0E4$ e todos os demais parâmetros livres (esquerda); índice $q = 1$ fixo e todos os demais parâmetros livres (direita).

Tabela 5 – Parâmetros ajustados para: $\kappa^* = 0$ fixo (primeira linha); todos os parâmetros ajustados (segunda linha); população fixa $5.0E4$, $5.0E5$ e $7.5E5$ (linhas 3 a 5); $q = 1$ fixo (última linha).

s	κ	κ^*	q	χ^2/n_{dof}
2430(55)	3.0(4)E-4	0	0.56(2)	81/42
2446(77)	3.0(5)E-4	0.05(14)	0.57(3)	81/41
5.0E4	8.22(6)E-5	4.02(3)	0.605(7)	138/42
5.0E5	1.731(1)E-5	8.635(7)	0.999(2)	194/42
7.5E5	1.2026(7)E-5	9.01(6)	0.999(2)	318/42
1.56837(1)E5	2.936(6)E-5	4.574(9)	1	162/42

Ainda é possível comparar a adequação do modelo FSIR com vários grupos e o SIR tradicional. Utilizando o modelo fractal temos um conjunto maior de parâmetros. Se ajustarmos a curva com 3 grupos, teremos 12 parâmetros. Com o objetivo de melhorar o ajuste, o κ^* será fixado em 0.2.

Tabela 6 – Primeiras três linhas: Parâmetros ajustados com as equações 3.20 com $k^* = 0.2$ e 3 agentes. Última linha: único agente, $k^* = 0.2$ e $q = 1$.

s_j	κ_j	q_j	R_{0j}	χ^2/n_{dof}
461(90)	0.00136(40)	0.66(9)	3(1)	
1366(91)	0.000377(32)	0.74(3)	2.5(3)	36/33
658(62)	0.00324(79)	0.38(6)	10(3)	
3555(93)	0.000072(2)	1	1.29(5)	771/43

Os resultados da tabela 6 mostram que, embora o modelo FSIR possua mais picos e conseqüentemente mais parâmetros, o χ^2 reduzido é menor, conseguindo capturar mais detalhes sobre a curva de dados. Assim, podemos usar os dois modelos, FSI e FSIR, com a ideia de grupos.

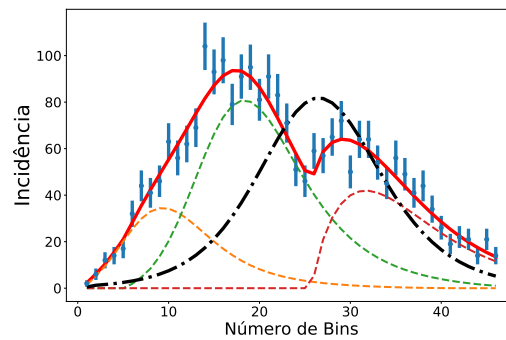


Figura 21 – A linha contínua vermelha representa o ajuste com o modelo FSIR usando o valor $\kappa = 0.2$ e três picos (linhas coloridas tracejadas). A curva ponto-linha preta é o ajuste com o modelo FSIR usando o valor $\kappa = 0.2$ e $q = 1$.

Outro aspecto importante é que a curva de infectados deixa de ser simétrica, apresentando um maior crescimento no início da propagação. Embora o modelo FSIR tenha um parâmetro a mais, a análise mostra que essa abordagem captura aspectos importantes, como o grau de interação entre os agentes do sistema. O valor de q se aproxima de 1 conforme o número de contatos aumenta. Essa flexibilidade que o modelo permite é compatível com um comportamento subexponencial, devido à formação de cluster de contatos e variações no comportamento da população (SZENDROI; CSÁNYI, 2004; VIBOUD; SIMONSEN; CHOWELL, 2016; CHOWELL et al., 2016).

6 Conclusão

Essa pesquisa apresentou conceitos interdisciplinares e atuais, como ciência das cidades, epidemiologia e sistemas complexos para construir um modelo capaz de descrever a dinâmica de propagação de doenças causadas por vírus. Também foi mostrado os resultados dos estudos feitos com dados de alguns países. Essa análise mostrou que o número de infectados pelo vírus da COVID-19 pode ser descrito por uma lei de potência e que os coeficientes ficam próximos de 1.15. Esse valor é coerente com outros resultados encontrados na literatura e indica uma certa universalidade nesses expoentes. O ajuste em lei de potência indica a possibilidade de fractalidade e motiva os estudos nessa área. A lei de potência indica que os dados não tem um tamanho ou escala característica, propriedade típica de fractais.

Usando uma abordagem original, desenvolvemos um modelo para descrever a difusão do vírus baseado em primeiros princípios e considerando uma população heterogênea. Esse modelo mudou a forma da solução de exponencial para q -exponencial, o que possibilitou um ajuste melhor com os dados.

Foram feitos diversos ajustes com dados reais sobre a pandemia e isso resultou em um conjunto de parâmetros otimizados. O modelo também permitiu acessar informações sobre as interações sociais através do índice q , comparando os resultados do ajuste com dados sobre isolamento social, representado pelo índice de isolamento social, obtido através de informações de empresas de telefonia. Além disso, conseguimos estimar o número de contatos próximos: $N = 3$ e $N = 6$ para a primeira e segunda onda, respectivamente. Nesses mesmos ajustes identificamos a dependência de alguns parâmetros com a população: $\kappa = \kappa_0 s^\beta$ e $\Delta t = \Delta t_0 s^\alpha$.

Fazendo modificações no modelo, incluímos uma variável dos recuperados, $r(t)$, que tornou a comparação com o SIR mais direta. O resultado foi um modelo que generaliza o SIR e o SI. Como discutido no capítulo Modelo Fractal, o comportamento de q e κ^* determinam o tipo de modelo usado. Para $q = 1$ e $\kappa^* = 0$ obtemos o SI tradicional. Se $q = 1$ e $\kappa^* \neq 0$, temos o SIR tradicional. A generalização é útil para descrever comportamentos subexponenciais.

A hipótese fractal sobre a heterogeneidade da população ficou bastante evidente através do método de clusterização que foi desenvolvido. Foi possível identificar os clusters e associá-los a locais de fluxo alto de pessoas.

Todos esses resultados permitiram duas publicações, uma delas na revista *Chaos, Solitons and Fractals* (ABBASI et al., 2020) e outra na *Clinical Immunology and Immunotherapy* (POLICARPO et al., 2020). Outro artigo escrito com colaboração de pesquisadores do Brasil e do exterior está sendo analisado pela revista *Applied Mathematical Modelling*.

A ideia da propagação do vírus pode ser estendida para outros tipos de informações que se espalham na sociedade por meio de contato entre seus agentes, como notícias e boatos (SONG; WANG; LI, 2021; CHANG et al., 2021), propagação e difusão de informação em uma network (DEPPMAN; SEGUNDO, 2021) e há indícios de que o número de homicídios tem aspectos fractais (DEPPMAN, 2019).

Os próximos passos deste estudo incluem o teste em diferentes dados de epidemias a fim de comparar o modelo fractal com os modelos tradicionais. Como descrito no texto, o FSIR generaliza os modelos SIR e SI, além de fornecer uma dinâmica estruturada em grupos, agentes de diferentes escalas e uma ideia fractal.

Referências

- ABBASI, M.; BOLLINI, A.; CASTILLO, J.; DEPPMAN, A.; GUIDIO, J.; MATUOKA, P.; MEIRELLES, A.; POLICARPO, J.; RAMOS, A.; SIMIONATTO, S.; VARONA, A.; ANDRADE-II, E.; PANJEH, H.; TREVISAN, L. Fractal signatures of the covid-19 spread. *Chaos, Solitons & Fractals*, v. 140, p. 110119, 2020. ISSN 0960-0779. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0960077920305166>>.
- ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. *Reviews of modern physics*, APS, v. 74, n. 1, p. 47, 2002.
- ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Diameter of the world-wide web. *nature*, Nature Publishing Group UK London, v. 401, n. 6749, p. 130–131, 1999.
- ALEMANHA. *ALEMANHA*. 2020.
Urlhttps://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html.
- AMARAL, L. A.; OTTINO, J. M. Complex networks. *The European physical journal B*, Springer, v. 38, n. 2, p. 147–162, 2004.
- ANDERSON, P. W. More is different: broken symmetry and the nature of the hierarchical structure of science. *Science*, American Association for the Advancement of Science, v. 177, n. 4047, p. 393–396, 1972.
- ANDERSON, P. W. *More is different*. [S.l.: s.n.], 2011.
- ANDERSON, P. W. *The economy as an evolving complex system*. [S.l.]: CRC Press, 2018.
- ANDERSON, R. M. Discussion: the kermack-mckendrick epidemic threshold theorem. *Bulletin of mathematical biology*, Springer, v. 53, p. 1–32, 1991.
- ARTHUR, W.; BEINHOCKER, E.; STANGER, A. *Complexity economics: Proceedings of the Santa Fe Institute's 2019 fall symposium (dialogues of the applied complexity network)*. [S.l.]: SFI Press, 2020.
- ARTHUR, W. B. *The economy as an evolving complex system II*. [S.l.]: CRC Press, 2018.
- ASSIS, T. A. d.; MIRANDA, J. G. V.; MOTA, F. d. B.; ANDRADE, R. F. S.; CASTILHO, C. M. C. d. Geometria fractal: propriedades e características de fractais ideais. *Revista Brasileira de ensino de física*, SciELO Brasil, v. 30, p. 2304–1, 2008.
- BAR-YAM, Y. *Dynamics of Complex Systems*. [S.l.]: Addison-Wesley, 1997.
- BATTY, M.; LONGLEY, M. Fractal cities - a geometry of form and function. *Academic Press, London*, 01 1994.
- BENYOUSSEF, S.; AMRAOUI, Y. E.; EZ-ZAHRAOUI, H.; MEZZANE, D.; KUTNJAK, Z.; LUK'YANCHUK, I.; MARSSI, M. E. Mean field theory and monte carlo simulation of phase transitions and magnetic properties of a tridimensional fe7s8 compound. *Physica Scripta*, IOP Publishing, v. 95, n. 4, p. 045803, 2020.

- BERNOULLI, D. Essay on a new analysis of mortality caused by smallpox and the advantages of inoculation in its prevention. *Histoire de l'Acad., Roy. Sci. (Paris) avec Mem., Des Math. And Phis., Mem.*, p. 1–45, 1760.
- BERTIN, E. *Statistical Physics of Complex Systems*. [S.l.]: Springer, 2016.
- BETTENCOURT, L.; WEST, G. A unified theory of urban living. *Nature*, Nature Publishing Group, v. 467, n. 7318, p. 912–913, 2010.
- BETTENCOURT, L. M.; LOBO, J.; HELBING, D.; KÜHNERT, C.; WEST, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences*, National Acad Sciences, v. 104, n. 17, p. 7301–7306, 2007.
- BONDY, U. M. J. *Graph Theory*. [S.l.]: Springer, 2008.
- BURDEN, R.; FAIRES, J.; BURDEN, A. *Numerical Analysis*. Cengage Learning, 2015. ISBN 9781305465350. Disponível em: <<https://books.google.com.br/books?id=9DV-BAAAQBAJ>>.
- CALLEN, H. B. *Thermodynamics and an Introduction to Thermostatistics*. [S.l.]: American Association of Physics Teachers, 1998.
- CHANG, Y. S.; LEE, Y.-T.; KIM, H. K.; NAM, S. W. Scaling relationship of patent counts vs. population size-regional and population subgroup analyses of 50 states in the us. *International Journal of Technology, Policy, and Management (2016) Vol*, v. 16, p. 13–11, 2015.
- CHANG, Z.; JIANG, H.; YU, S.; CHEN, S. Dynamic analysis and optimal control of iscr rumor propagation model with nonlinear incidence and time delay on complex networks. *Discrete Dynamics in Nature and Society*, Hindawi Limited, v. 2021, p. 1–20, 2021.
- CHOWELL, G.; HYMAN, J. M.; EUBANK, S.; CASTILLO-CHAVEZ, C. Scaling laws for the movement of people between locations in a large city. *Physical Review E*, APS, v. 68, n. 6, p. 066102, 2003.
- CHOWELL, G.; VIBOUD, C.; SIMONSEN, L.; MOGHADAS, S. M. Characterizing the reproduction number of epidemics with early subexponential growth dynamics. *Journal of The Royal Society Interface*, The Royal Society, v. 13, n. 123, p. 20160659, 2016.
- CONTROL, W. W. Center for D. *Center for Disease Control, Worldometers*. 2020. Url<https://www.cdc.gov/coronavirus>, <https://www.worldometers.info/coronavirus/>.
- DADOS Isolamento Social. <<https://www.saopaulo.sp.gov.br/coronavirus/isolamento/>>. Accessed: 2020-09-30.
- DEPPMAN, A. Scaling properties of firearm homicides in brazilian cities. *arXiv preprint arXiv:1905.11483*, 2019.
- DEPPMAN, A.; SEGUNDO, E. O. A. *Flux of information in scale-free networks*. 2021.
- ESPANHA. *ESPANHA*. 2020. Ur-
l<https://cneocovid.isciii.es/covid19/distribuci%C3%B3ngeogr%C3%A1fica>.
- ESTADO governo do. *governo do estado*. 2020. Url<https://www.cdc.gov/coronavirus>.

- EXCHANGE, T. H. D. *The Humanitarian Data Exchange*. 2020. Url<https://data.humdata.org>.
- FALCONER, K. *Fractal geometry: mathematical foundations and applications*. [S.l.]: John Wiley & Sons, 2004.
- FALCONER, K. J.; FALCONER, K. *Techniques in fractal geometry*. [S.l.]: Wiley Chichester, 1997. v. 3.
- FLETCHER, R. A new approach to variable metric algorithms. *The Computer Journal*, v. 13, n. 3, p. 317–322, 01 1970. ISSN 0010-4620. Disponível em: <<https://doi.org/10.1093/comjnl/13.3.317>>.
- FOROUTAN-POUR, K.; DUTILLEUL, P.; SMITH, D. L. Advances in the implementation of the box-counting method of fractal dimension estimation. *Applied mathematics and computation*, Elsevier, v. 105, n. 2-3, p. 195–210, 1999.
- FRANCA. *Franca*. 2020. Url<https://www.bing.com/covid/local/france?vert=graph>.
- GREINER, W.; NEISE, L.; STÖCKER, H. *Thermodynamics and statistical mechanics*. [S.l.]: Springer Science & Business Media, 2012.
- HOLLAND, J. H. *Complexity: A very short introduction*. [S.l.]: OUP Oxford, 2014.
- ITALIA. *Italia*. 2020. Url<https://github.com/pcm-dpc/COVID-19/tree/master/datiregioni>.
- JAFARI, H.; MALINOWSKI, M. T.; EBADI, M. Fuzzy stochastic differential equations driven by fractional brownian motion. *Advances in Difference Equations*, SpringerOpen, v. 2021, n. 1, p. 1–17, 2021.
- JAMES, F.; ROOS, M. Minuit - a system for function minimization and analysis of the parameter errors and correlations. *Computer Physics Communications*, v. 10, n. 6, p. 343–367, 1975. ISSN 0010-4655. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0010465575900399>>.
- KADANOFF, L. P. More is the same; phase transitions and mean field theories. *Journal of Statistical Physics*, Springer, v. 137, n. 5, p. 777–797, 2009.
- KEMPKES, S. N.; SLOT, M. R.; FREENEY, S. E.; ZEVENHUIZEN, S. J.; VANMAEKELBERGH, D.; SWART, I.; SMITH, C. M. Design and characterization of electrons in a fractal geometry. *Nature physics*, Nature Publishing Group UK London, v. 15, n. 2, p. 127–131, 2019.
- KERMACK, W. O.; MCKENDRICK, A. G. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, The Royal Society London, v. 115, n. 772, p. 700–721, 1927.
- KERMACK, W. O.; MCKENDRICK, A. G. Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity. *Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character*, The Royal Society London, v. 138, n. 834, p. 55–83, 1932.

- KERMACK, W. O.; MCKENDRICK, A. G. Contributions to the mathematical theory of epidemics. iii.—further studies of the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, The Royal Society London, v. 141, n. 843, p. 94–122, 1933.
- LEAL, F. E.; MENDES-CORREA, M. C.; BUSS, L. F.; COSTA, S. F.; BIZARIO, J. C. S.; SOUZA, S. R. P. de; THOMAZ, O.; TOZETTO-MENDOZA, T. R.; VILLAS-BOAS, L. S.; SILVA, L. C. de Oliveira-da; GRESPAN, R. M. Z.; CAPUANI, L.; BUCCHERI, R.; DOMINGUES, H.; ALEXANDER, N.; MAYAUD, P.; SABINO, E. C. Clinical features and natural history of the first 2073 suspected covid-19 cases in the corona são caetano primary care programme: a prospective cohort study. *BMJ Open*, British Medical Journal Publishing Group, v. 11, n. 1, 2021. ISSN 2044-6055. Disponível em: <<https://bmjopen.bmj.com/content/11/1/e042745>>.
- MANDELBROT, B. B.; MANDELBROT, B. B. *The fractal geometry of nature*. [S.l.]: WH freeman New York, 1982. v. 1.
- MARATHE, M. V. Mathematical tools for understanding infectious disease dynamics. princeton series in theoretical and computational biology. by odo diekmann, hans heesterbeek and tom britton. xiv+502 pp. princeton, nj: Princeton university press. 2014. \$90.00 (cloth and e-book). *American Journal of Human Biology*, v. 25, n. 5, p. 710–711, 2013. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/ajhb.22436>>.
- MICHELON, C. M. Principais variantes do sars-cov-2 notificadas no brasil main sars-cov-2 variants notified in brazil. 2021.
- MITCHELL, M. *Complexity: A guided tour*. [S.l.]: Oxford university press, 2009.
- NAKAMURA, G. M.; MONTEIRO, A. C. P.; CARDOSO, G. C.; MARTINEZ, A. S. Efficient method for comprehensive computation of agent-level epidemic dissemination in networks. *Scientific reports*, Nature Publishing Group, v. 7, n. 1, p. 1–12, 2017.
- ODUM, E. P. The new ecology. *BioScience*, JSTOR, v. 14, n. 7, p. 14–16, 1964.
- ONSAGER, L. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Physical Review*, APS, v. 65, n. 3-4, p. 117, 1944.
- PASCUAL, M.; DUNNE, J. A.; DUNNE, J. A. et al. *Ecological networks: linking structure to dynamics in food webs*. [S.l.]: Oxford University Press, 2006.
- POLICARPO, J.; MEIRELLES, A.; RAMOS, A.; DEPPMAN, A. Fractal aspects of epidemic spread. *J Clin Immunol Immunother*, v. 6, n. 052, p. 2, 2020.
- RAMOS, V.; ABRAHAM, A. Evolving a stigmergic self-organized data-mining. *arXiv preprint cs/0403001*, 2004.
- SCHLAPFER, M.; BETTENCOURT, L.; GRAUWIN, S.; RASCHKE, M.; CLAXTON, R.; SMOREDA, Z.; WEST, G.; RATTI, C. The scaling of human interactions with city size. jr soc. *Interface*, v. 11, n. 98, p. 20130789–20130789, 2014.
- SCHLÄPFER, M.; BETTENCOURT, L. M. A.; GRAUWIN, S.; RASCHKE, M.; CLAXTON, R.; SMOREDA, Z.; WEST, G. B.; RATTI, C. The scaling of human interactions with city size. *Journal of The Royal Society Interface*, v. 11, n. 98, p.

- 20130789, 2014. Disponível em: <<https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2013.0789>>.
- SONG, G.; WANG, Y.; LI, Y. Dynamic mathematical model of information spreading on news platform. *Wireless Communications and Mobile Computing*, Hindawi Limited, v. 2021, p. 1–11, 2021.
- SZENDROI, B.; CSÁNYI, G. Polynomial epidemics and clustering in contact networks. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, The Royal Society, v. 271, n. suppl_5, p. S364–S366, 2004.
- TIRNAKLI, U.; TSALLIS, C. Epidemiological model with anomalous kinetics: Early stages of the covid-19 pandemic. *Frontiers in Physics*, Frontiers Media SA, v. 8, p. 613168, 2020.
- TSALLIS, C.; TIRNAKLI, U. Predicting covid-19 peaks around the world. *Frontiers in Physics*, Frontiers Media SA, v. 8, p. 217, 2020.
- UYEHARA, K. Self-organization in social insects. Citeseer, 2008.
- VIBOUD, C.; SIMONSEN, L.; CHOWELL, G. A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks. *Epidemics*, Elsevier, v. 15, p. 27–37, 2016.
- WEISS, H. H. The sir model and the foundations of public health. *Materials mathematics*, p. 0001–17, 2013.
- WEST, G. *Scale: The universal laws of life, growth, and death in organisms, cities, and companies*. [S.l.]: Penguin, 2018.
- WORMSER, G. P.; POURBOHLOUL, B. *Modeling Infectious Diseases in Humans and Animals By Matthew James Keeling and Pejman Rohani Princeton, NJ: Princeton University Press, 2008. 408 pp., Illustrated*. [S.l.]: The University of Chicago Press, 2008.
- WU, J.; JIN, X.; MI, S.; TANG, J. An effective method to compute the box-counting dimension based on the mathematical definition and intervals. *Results in Engineering*, v. 6, p. 100106, 2020. ISSN 2590-1230. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2590123020300128>>.