

# Simulação do COVID-19

---

Professores:

Carlos Henrique Costa Ribeiro

Filipe Alves Neto Verri

Alunos:

Eric Toshio Endo Soares

Flávio Luiz dos Santos Souza

## 1. Introdução

Com a pandemia do novo coronavírus (COVID-19), o mundo tem enfrentado um dos momentos mais difíceis dos últimos séculos. Além das milhares de mortes causadas pela doença, a necessidade de medidas de isolamento e fechamento de escolas e comércio têm trazido prejuízos sociais e econômicos comparáveis aos tempos de guerra. Com o advento de uma vacina em um horizonte relativamente próximo, as autoridades públicas precisam preparar estratégias otimizadas de imunização da população e, nesse contexto, modelos matemáticos podem ser de grande auxílio. O ano de 2020 será lembrado como o ano no qual a pandemia do COVID-19 marcou a história da humanidade. Milhares de mortos foram registrados e implicações severas foram causadas na vida da maior parte da população mundial.

Esse contexto impulsionou várias pesquisas na área de modelos epidemiológicos de rede. Esses estudos buscam apresentar representações teóricas de difusão e imunização da sociedade que poderiam auxiliar e até prever o comportamento de doenças infecciosas com o desenvolvimento e implementação de modelos computacionais de propagação epidêmica.

Este projeto tem como objetivo apresentar uma simulação da difusão do COVID-19 dentro do território nacional. Para isso, criou-se um modelo de rede espacial e um modelo de difusão epidêmica que buscam aproximar-se da realidade. Posteriormente, os modelos são comparados com um dataset real que apresenta a quantidade de casos confirmados.

## 2. Metodologia

O projeto foi desenvolvido com a linguagem de programação Python 3.8 com o auxílio de várias bibliotecas. Dentre elas, destacam-se a *networkx* [1], que possibilitou a geração e análises de redes, e a *ndlib* [2], que auxilia na simulação de processos de difusão em redes complexas. Além disso, o código-fonte do projeto está disponível no Github [3].

### 2.1. Dataset

O dataset utilizado no projeto pode ser obtido no site Kaggle [4]. O foco deste dataset é apresentar o número de casos confirmados de COVID-19 por cidade e com uma granularidade diária. Suas grandes vantagens são que os dados são atualizados constantemente e são necessários poucos tratamentos sobre eles. Além disso, o dataset apresenta também tabelas auxiliares com a localização e o número de habitantes de cada cidade.

### 2.2. Modelo de rede espacial

O modelo de rede espacial adotado neste projeto tem como objetivo de se aproximar da realidade de cidades do interior que muitas vezes possuem pouco fluxo de pessoas. Para isso, propõem-se a criação de uma rede com comunidades onde cada uma delas representa uma cidade em uma região do interior. A cidade escolhida, para o referido experimento, foi a cidade de Matão/SP e as cidades vizinhas em um raio máximo de 50 Km de distância, o que totaliza, consequentemente, 13 cidades no total.

A quantidade de vértices em cada cidade é proporcional ao número de habitantes. Já as arestas dentro de uma mesma comunidade foram atribuídas aleatoriamente de forma a gerar um grau esperado constante  $K_{in}$ . Por outro lado, as arestas entre um vértice e todas as outras comunidades possui o grau esperado  $K_{out}$  multiplicado por um fator inversamente proporcional à distância entre as cidades. A Equação 1 apresenta o cálculo desse grau esperado, onde  $dist$  representa distância entre as cidades e  $max$  é a máxima distância que uma cidade pode estar.

$$Z_{out} \times \left(1 - \frac{dist}{2 \times max}\right) \quad (1)$$

A grande questão é qual o valor a ser atribuído a  $Z_{in}$  e  $Z_{out}$ . Para auxiliar nessa escolha, utilizou-se como base a pesquisa [5] no qual apresenta que a quantidade de amigos próximos é cerca de 20.

Nesse sentido, atribui-se o valor de  $Z_{in}$  igual a 15 e  $Z_{out}$  igual a 5 a fim de o gráfico de distribuição de graus ficasse plausível. O gráfico dessa distribuição dos graus é apresentado na Figura 1, que evidencia que a maioria dos vértices ficou bem distribuído para valores menores que 30.

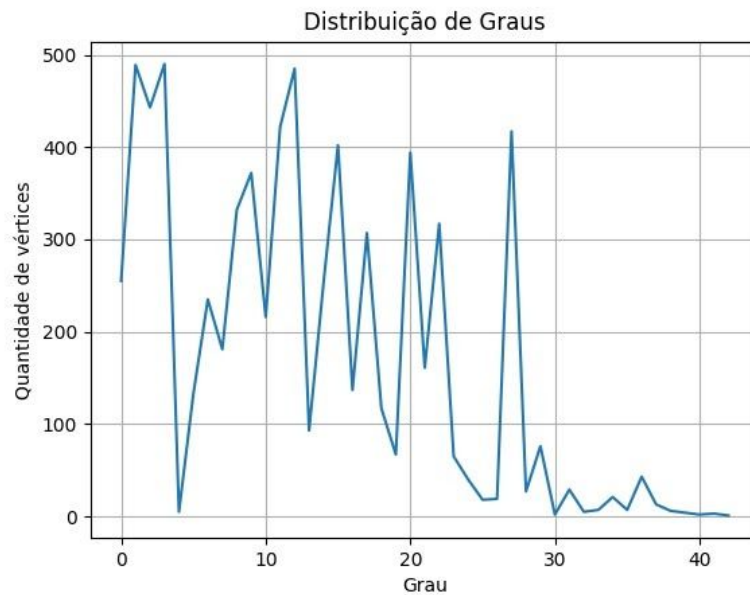


Figura 1 - Distribuição de Graus na rede proposta

Para exemplificar esse modelo espacial, gerou-se a Figura 2 que apresenta a rede complexa utilizando Matão como cidade principal e uma distância máxima de 30 Km, que engloba 7 cidades ao todo. Além disso, a fim de reduzir o número de nós foi utilizado uma proporção de 500 pessoas por vértice.

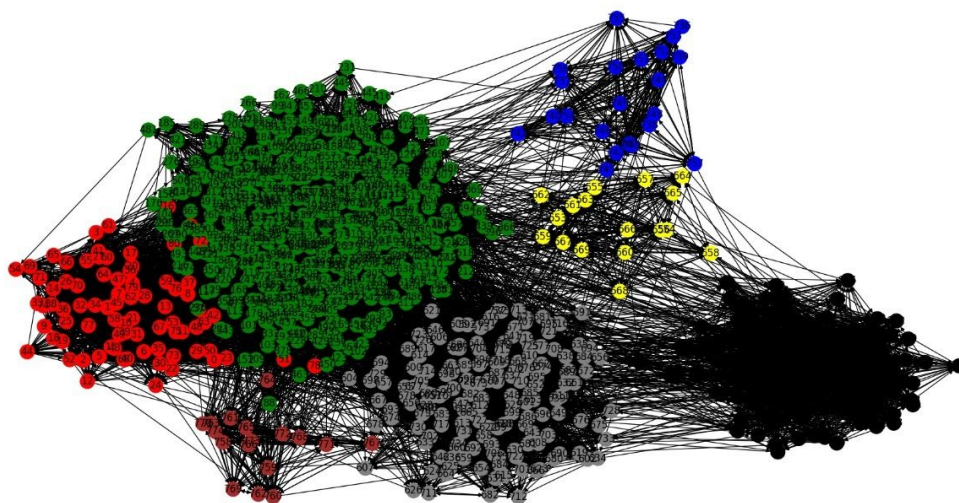


Figura 2 - Exemplificação de rede espacial utilizando Matão

## 2.3. Modelo de difusão epidêmica

### 2.3.1. Modelo SWIR

O modelo de rede de difusão usado como base do projeto foi o modelo SWIR (*Susceptible, Weakened, Infected and Removed*) apresentado no artigo de Lee D. [6]. Nesse modelo são propostos quatro tipos de grupo: suscetíveis, enfraquecidos, infectados e removidos.

Em uma rede modelada por SWIR, inicialmente apresenta vértices suscetíveis (S) e alguns poucos infectados (I). Um vértice infectado (I) pode apenas alterar o estado de seus vizinhos suscetíveis (S) para infectados (I) ou enfraquecidos (W). Além disso, os vértices infectados (I) também podem alterar o estado dos enfraquecidos (W) para infectados (I). A cada iteração os vértices originalmente infectados (I) são reclassificados para removidos (R). A Figura 3 representa um exemplo desse processo de difusão.

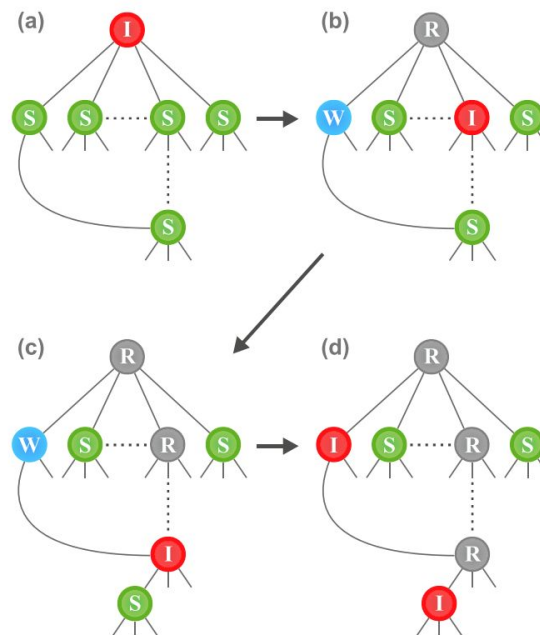


Figura 3 - Processo de propagação da epidemia no modelo SWIR[6]

Na Figura 3a, o único vértice infectado (I) transformou um vértice, originalmente suscetível (S), em enfraquecido (W) e outro em infectado (I). Na Figura 3b, o vértice inicialmente infectado (I) foi reclassificado como removido (R). A difusão de propagação continua seguindo o modelo SWIR.

### 2.3.2. Modelo SWIR modificado

A partir do modelo SWIR tradicional, foi proposto modificações para aproximá-lo das características conhecidas de disseminação do

COVID-19. A Figura 4 ilustra as modificações propostas, a qual a descrição das modificações são:

- O estado enfraquecido (W) passa a representar os casos assintomáticos do COVID-19;
- Por questões de adequação na nomenclatura o estado removido (R) foi renomeado para recuperando mantendo a simbologia (R);
- Um vértice suscetível (S) apenas poderá passar para o estado de assintomáticos (W) com uma probabilidade  $\mu$  ou para o estado de infectados (I) sob uma probabilidade  $\beta$ ;
- Os vértices categorizados com o estado assintomáticos (W) poderá contaminar os suscetíveis (S), igualmente aos vértices infectados (I);
- Os vértices com os estados de assintomático (W) e infectados (I) passam a ter uma probabilidade de  $\gamma$  de recuperação alterando o seu estado para recuperado (R) após algumas iterações;
- Os vértices assintomáticos (W) deixam de progredir para o estado de infectado (I).

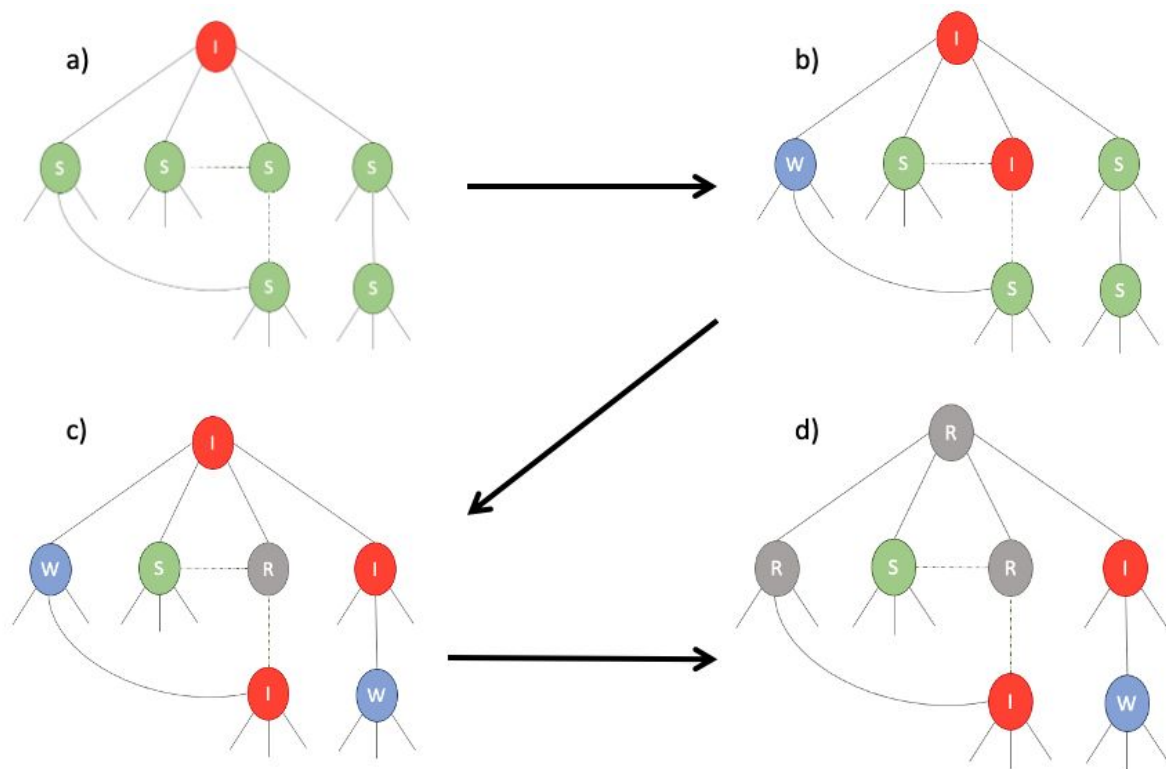


Figura 4 - Processo de propagação da epidemia no modelo SWIR Modificado

### 2.3.3. Parâmetros de simulação e estudo analítico

Para a definição dos valores para cada um dos parâmetros do modelo, foi elaborado um estudo analítico. O parâmetro  $TR$ , proposto para o modelo SWIR Modificado, que representa o Tempo Médio de recuperação do COVID-19 foi apresentado três cenários em que um indivíduo infectado pode transmitir o vírus para pessoas suscetíveis correspondente a 5, 14 e 21 dias. De acordo com [7] a probabilidade de recuperação  $\gamma$  ao COVID-19 é dada pelo inverso de  $TR$ . [7] também explica os parâmetros utilizados na modelagem matemática comumente utilizadas no COVID-19 com base na taxa de reprodução ( $R_0$ ). O segundo [9] identificou 12 estudos que estimaram o número  $R_0$ . A qual variaram entre 1,4 a 6,49, apresentando uma média de 3,28, e a mediana de 2,79 e um intervalo interquartil (IQR) de 1,16 para o  $R_0$  do COVID-19. A probabilidade de infecção  $\beta$ , segundo [7], pode ser definida com o produto de  $R_0$  e  $\gamma$ .

Tabela 1 - Parâmetro e valores de simulação do Modelo SWIR Modificado

Parâmetro	Descrição	Valor(es)	Fonte
$TR$	Tempo médio de recuperação do COVID-19	14.7 dias	[7]
$\gamma$	Probabilidade de recuperação do COVID-19	$1/TR$	[7]
$R_0$	Número básico de reprodução	2.79 á 3.77	[9]
$\beta$	Probabilidade de infecção	$R_0 * \gamma$	[7]
$P_w$	Porcentagem de assintomáticos (W)	15% á 20%	[8]
$\mu$	Probabilidade de infectar um assintomáticos (W)	$\beta * P_w$	-
$FI$	Fração de infectados inicialmente	0.01	-
$N_{int}$	Número de iterações	30	-

O trabalho [8] estima a duração do período de incubação do COVID-19 baseadas em pessoas que desenvolveram sintomas e descreve suas implicações. Os dados utilizados e apresentados em [8] sugerem que 15% a 20% das infecções são assintomáticas que corresponde ao parâmetro  $P_w$  do modelo proposto. O parâmetro  $P_w$  que é utilizado no

produto com  $\beta$  para a definição da probabilidade de infecção de um assintomáticos  $\mu$ .

Por fim os parâmetros  $FI$  e  $N_{int}$  são os mais simplistas do modelo proposto que corresponde, respectivamente, a Fração de infectados inicialmente no modelos e o número entre as iterações do modelo. Ambos os valores, de  $FI$  e  $N_{int}$ , tiveram a definições arbitrárias com base apenas nos experimentos com o modelo.

A Tabela 1 apresenta os parâmetros e os respectivos valores utilizados nas repetições de simulação.

### 3. Resultados

Os resultados obtidos da simulação correspondem a 4 cidades brasileiras:

- Agudos/RS
- Matão/SP
- Pio XII/MA
- Saubara/BA

As escolhas das cidades foram aleatórias com o propósito de validar o modelo proposto e demonstrar a aderência a uma diversidade de contexto e realidade intrínseca aos contexto urbanísticos e civis das cidades/regiões simuladas. Todas as simulações tiveram 10 repetições.

#### 3.1. Gráfico de estado dos vértices

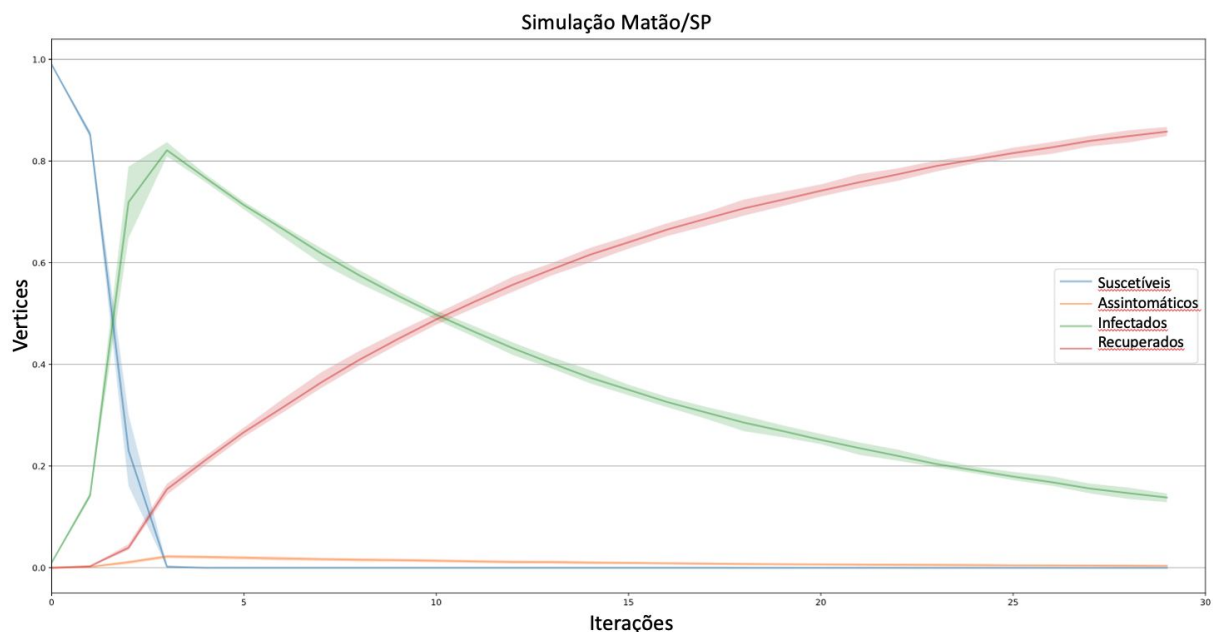


Figura 5 - Simulação epidêmica para Matão/SP

A Figura 5 apresenta os resultados de simulação epidêmica do modelo proposto para a cidade de Matão/SP. A curva em azul corresponde aos casos dos suscetíveis (S), em poucas iterações o percentual de suscetíveis reduz significativamente. A curva em laranja corresponde aos casos assintomáticos (W) da COVID-19, que na simulação apresentaram pouca representatividade. Com um papel similar ao assintomático, os infectados (I), representados pela curva em verde, apresentam na simulação uma onda tipicamente característica das infecções do COVID-19. A curva em vermelho é correspondente ao recuperados (R) que apresentam um resultado mais próximo do observado dos modelos teóricos clássicos.

### 3.2. Gráfico de casos confirmados para a região modelada

A partir do dataset gerou-se um gráfico que apresenta a porcentagem da população confirmada como infectada nas treze cidades consideradas. Abaixo é possível notar que na região considerada o número de infectados aproximou-se de cerca de 20% da população total na região.

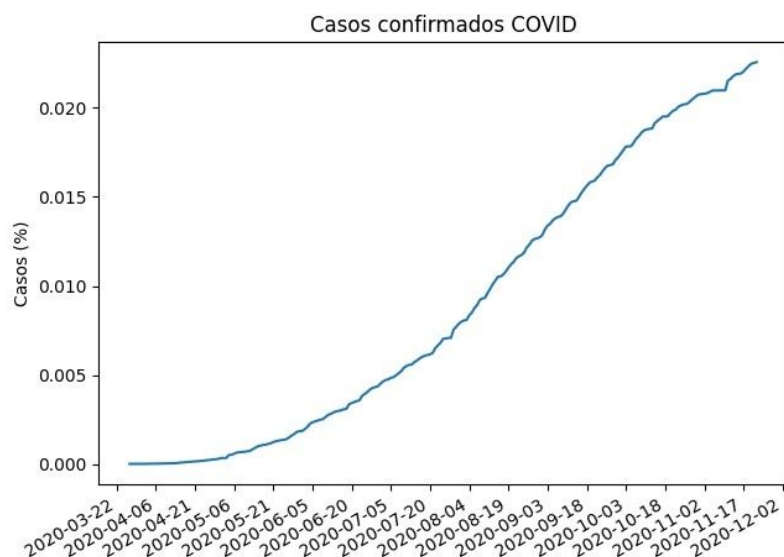


Figura 6 - Infectados reais em Matão/SP

Apesar de a porcentagem de casos confirmados não necessariamente representar a quantidade real infectados, é possível comparar a curva de casos com a curva de vértices em estado removido do modelo de difusão proposto. As duas curvas apresentaram formatos semelhantes, o que representa que as modelagens teóricas conseguiram em parte simular o comportamento real.



A Figura 7 apresenta as redes geradas para a simulação considerando outros contextos urbanísticos como as cidades de Saubara/BA, Pio XII/MA e Agudo/RS respectivamente. Seguido da respectiva distribuição de grau. Para cada contexto urbanístico o quantitativo de cidades vizinhas (comunidades) também altera para cada contexto. Para o contexto urbanístico simulado para a cidade de Saubara/BA foram consideradas 21 comunidades em um raio de 50km. Para a simulação da cidade de Pio XII/MA 7 comunidades para o mesmo raio. 3 comunidades são observadas, na Figura 7, para a simulação da cidade de Agudo/RS.

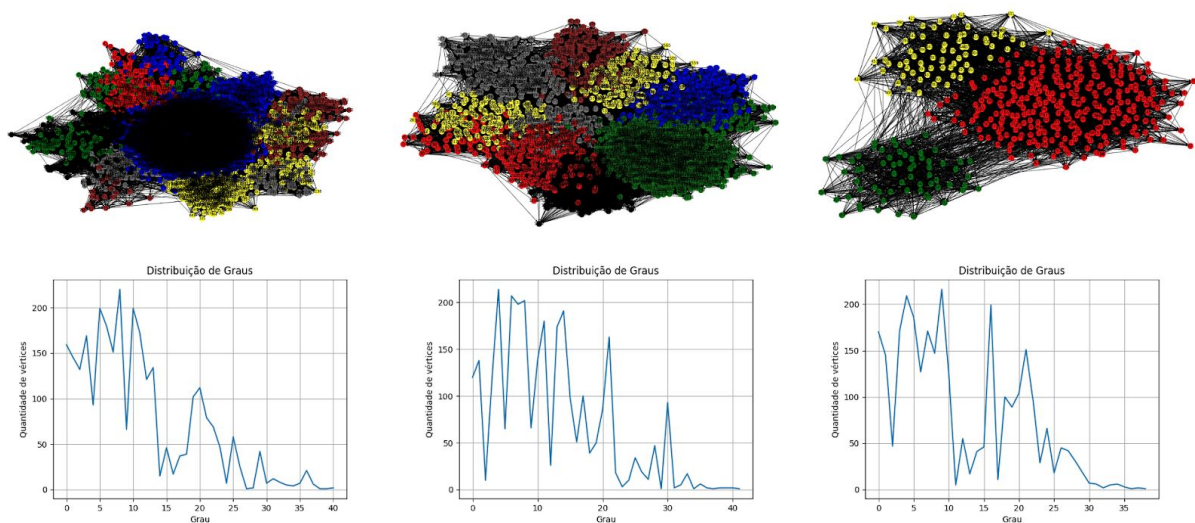


Figura 7 - Rede e distribuição de grau para Saubara/BA, Pio XII/MA e Agudo/RS

A Figura 8 apresenta os resultados gerados pelo modelo SWIR modificado para as cidades de Saubara/BA, Pio XII/MA e Agudo/RS respectivamente, seguido dos casos reais apresentados nas mesmas cidades.

Observa-se que para os resultados da simulação do modelo, na Figura 8, a simulação epidêmica para a cidade de Agudo/RS apresentou a maior variância entre os casos simulados. No entanto, é possível observar que para os 4 casos simulados a infecção afetou mais de 80% da população. Como observado no segundo conjunto de gráficos da Figura 8, onde são apresentados o acumulativo de infectados para as respectivas regiões, em todos os casos o percentual de infectados não passa de 2%.

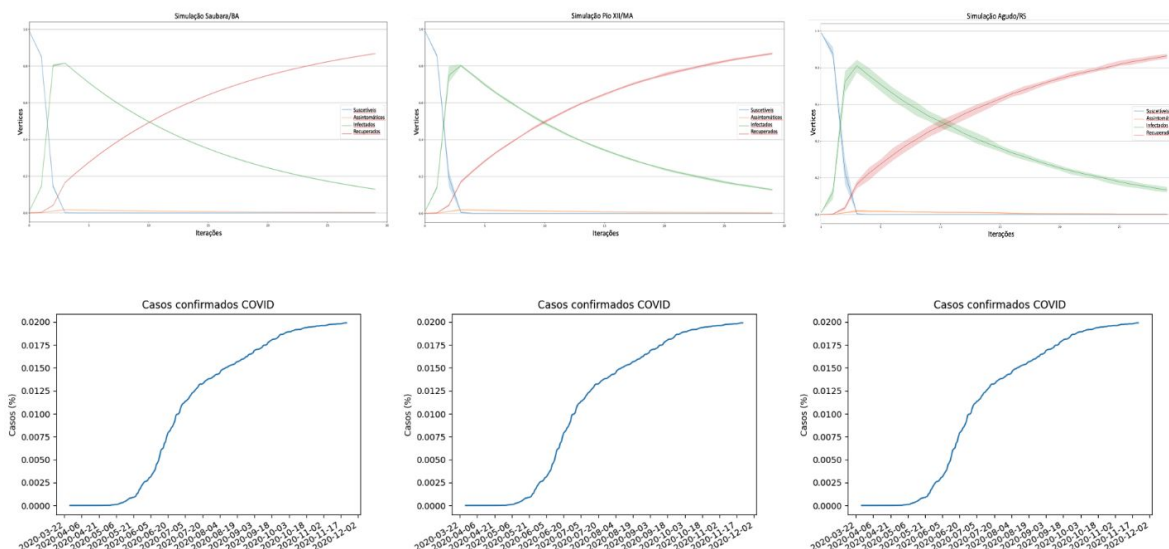


Figura 8 – Resultados do modelo proposto e dos casos reais confirmados para Saubara/BA, Pio XII/MA e Agudo/RS

## 4. Conclusão

O projeto propiciou um desafio prático interessante de tentar modelar uma doença infecciosa real. Esse trabalho pode ser dividido em duas grandes partes: a modelagem de uma rede espacial e a modelagem de um método de difusão.

Em ambos os casos, devido à grande complexidade do mundo real, foram necessárias adotar simplificações. Nesse sentido, a questão é como escolher as características relevantes a serem consideradas na modelagem. Mesmo com vários dataset e informações sobre as propriedades da doença, nem sempre a escolha dessas propriedades fica evidente.

No caso do COVID-19, existe o conhecimento de que há pessoas assintomáticas que transmitem a doença. Os dados estatísticos de taxa de transmissão dessas pessoas e suas proporções em relação a quantidade de infectados é uma grande incerteza. Isso tudo torna ainda mais desafiante a modelagem do COVID-19. Neste estudo, aplicando o modelo SWIR modificado aos diferentes casos de contextos urbanísticos de 4 cidades brasileiras, a difusão epidêmica do COVID-19 são próximas e similares para todos os casos.

No decorrer deste estudo, alguns insights epidemiológicos surgiram que podem trazer contribuições e melhoria para o estudo e nos resultados:

- Para o Modelo de redes espaciais, pode-se propor um aprimoramento na escolha do número de arestas com o propósito de representar um contexto de *Lockdown* ou isolamento social ou ciclos de quarentena;

- Ainda no Modelo de redes espaciais, reconsiderar a equivalência de proporcionalidade de um vértice por parte da população;
- No Modelo de difusão epidêmica proposto, SWIR Modificado, pode-se considerar a utilização de parâmetros que variam ao longo do tempo ( $R_0$ );
- Também seria de uma boa contribuição utilizar de técnicas de identificação paramétrica para a simulação.

## 5. Referência bibliográfica

- [1] <https://networkx.org/>
- [2] <https://ndlib.readthedocs.io/>
- [3] <https://github.com/flaviol-souza/simulacao-covid>
- [4] [Coronavirus - Brazil](#)
- [5] [How Many Friends Does the Average Person Have?](#)
- [6] Lee D, Choi W, Kertész J, Kahng B. Universal mechanism for hybrid percolation transitions. Sci Rep. 2017;7(1):5723. Published 2017 Jul 18. doi:10.1038/s41598-017-06182-3
- [7] A evolução epidêmica do COVID-19 - Modelo SIR, Grupo de Dispersão de Poluentes & Engenharia Nuclear da UFPel, <https://wp.ufpel.edu.br/fentransporte/2020/04/09/a-evolucao-epidemic-a-do-covid-19-modelo-sir/>
- [8] The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application, Annals of Internal Medicine, 2020, 10.7326/M20-0504
- [9] The reproductive number of COVID-19 is higher compared to SARS coronavirus, Ying Liu, Albert A Gayle, Annelies Wilder-Smith, Joacim Rocklöv, Journal of Travel Medicine, Volume 27, Issue 2, March 2020, taaa021, <https://doi.org/10.1093/jtm/taaa021>, Published: 13 February 2020