

Modelagem Epidemiológica do COVID-19

Modelos de compartimento estratificado e espacial

Força tarefa do ITA para simulação avançada do COVID-19

Instituto Tecnológico de Aeronáutica

Departamento de Física

Relatório para comunicação das investigações teóricas voltadas à modelagem epidemiológica do COVID-19. Versão 4.5.

23 de março de 2020

Sumário

1	Objetivo	1
2	Desenvolvimento Teórico - Modelo Global	1
2.1	Modelo SIR	1
2.2	Modelo SEIR	2
2.3	Modelo SEAIR	3
2.4	Relação entre γ e μ_{cov}	3
3	Desenvolvimento Teórico - Propagação epidêmica sobre um território e por faixa etária	4
3.1	SIR - Espacial	6
3.2	SEIR - Espacial	6
3.3	Modelagem dos fluxos populacionais	7
3.4	Transformação para cidades/microrregiões	8
3.5	Inclusão de faixa etária	9
3.5.1	Modelo global com faixas etárias	12
3.5.2	Matriz de contato	13
3.6	Cenários	14
3.6.1	Nenhuma intervenção. Cenário 1	14
3.6.2	Fechamento de escola apenas. Cenário 2	14
3.6.3	Fechamento de escola e distanciamento social. Cenário 3	14
3.6.4	Fechamento de escola, distanciamento no trabalho e distanciamento social. Cenário 4	15
3.6.5	Isolamento dos idosos. Cenário 5.	15
3.6.6	Fechamento de escola, distanciamento no trabalho, distanciamento social e isolamento dos idosos. Cenário 6	15
3.6.7	Knock Down com isolamento de idosos. Cenário 7	16
3.6.8	Normalização	16
3.7	Modelagem do fluxo entre as microrregiões	16

3.8	Cálculo de grandezas úteis: número de casualidades (C), hospitalizações (H), leitos ocupados (L) e número final total de infectados (M)	17
3.8.1	Casualidades	17
3.8.2	Hospitalizações	18
3.8.3	Leitos ocupados	18
3.8.4	Número total casos M	18
3.8.5	Outras saídas importantes	19
4	Desenvolvimento Teórico - Modelo SEAHIR-Q_{AI}	20
4.1	Generalização do modelo modelo SEAHIR- Q_{AI}	22
5	Parâmetros dos modelos compartimentados	23
5.1	Metodologia para estimativa dos parâmetros	23
6	Benchmark utilizando o modelo de Prem <i>et al</i>	29

1 Objetivo

A presente versão deste relatório tem como objetivo a rápida comunicação da modelagem matemática para a descrição quantitativa da propagação da contaminação pelo COVID-19. Apresento dois tipos de modelo. O primeiro, mais simples, é um modelo global, pois considera a população como um todo e despreza a dinâmica local e a difusão territorial da epidemia. O segundo modelo resulta de um esforço para modelar a dinâmica espacial do vírus sobre o território de um país. A razão para considerar os dois casos está no fato de que o primeiro é de rápida implementação e permite captar as características gerais do problema. O segundo é mais realista, mas de implementação mais dispendiosa.

A abordagem “0-D” segue o padrão conhecido como SIR (susceptible-infected-recovered) e SEIR (susceptible-exposed-infected-recovered) (seção 2). Nós mostramos como essas abordagens podem ser extendidas para incluir a propagação regional das doenças infecciosas (seção 3).

2 Desenvolvimento Teórico - Modelo Global

2.1 Modelo SIR

No modelo SIR considera-se a evolução temporal da população susceptível (S), infectada (I) e recuperada (R). As equações diferenciais que regem a variação dessas populações é dada por

$$\frac{dS}{dt} = \Lambda N - \mu_{eq}S - \beta \frac{S}{N}I \quad (1)$$

$$\frac{dI}{dt} = \beta \frac{S}{N}I - \mu_{cov}I - \mu_{eq}I - \gamma I \quad (2)$$

$$\frac{dR}{dt} = \gamma I - \mu_{eq}R \quad (3)$$

$$\frac{dN}{dt} = \Lambda N - \mu_{eq}N - \mu_{cov}I \quad (4)$$

$$N = S + I + R \quad (5)$$

onde μ_{eq} é o coeficiente de taxa de mortalidade em condições normais e μ_{cov} é o coeficiente de taxa de mortalidade do COVID-19. Os outros parâmetros são Λ , o

coeficiente de taxa natalidade típica da população, β , coeficiente de taxa de contaminação e γ , o coeficiente de taxa típica de recuperação do COVID-19.

Note que a equação (4) resulta da soma das equações (1) a (3), não sendo portanto linearmente independente delas. Trata-se da equação para a população total em sua forma mais geral, admitindo aumento da população por conta da taxa de natalidade e diminuição por conta da taxa de mortalidade. As simplificações mais usuais habitualmente consideram uma situação de equilíbrio; no entanto, pelo conhecimento prévio sobre a doença, acredito ser necessário avaliar o impacto sobre a taxa de mortalidade.

2.2 Modelo SEIR

No modelo SEIR considera-se a evolução temporal da população susceptível (S), exposta(E), infectada (I) e recuperada (R). As equações diferenciais que regem a variação dessas populações é dada por

$$\frac{dS}{dt} = \Lambda N - \mu_{eq}S - \beta \frac{S}{N}I \quad (6)$$

$$\frac{dE}{dt} = \beta \frac{S}{N}I - \mu_{eq}E - aE \quad (7)$$

$$\frac{dI}{dt} = aE - \mu_{cov}I - \mu_{eq}I - \gamma I \quad (8)$$

$$\frac{dR}{dt} = \gamma I - \mu_{eq}R \quad (9)$$

$$\frac{dN}{dt} = \Lambda N - \mu_{eq}N - \mu_{cov}I \quad (10)$$

$$N = S + E + I + R \quad (11)$$

onde μ_{eq} é o coeficiente de taxa de mortalidade em condições normais e μ_{cov} é o coeficiente de taxa de mortalidade do COVID-19. Os outros parâmetros são Λ , o coeficiente de taxa natalidade típica da população, β , coeficiente de taxa de contaminação, γ , o coeficiente de taxa típica de recuperação do COVID-19. Note que a equação (10) resulta da soma das equações (6) a (9), não sendo portanto linearmente independente delas. As mesmas considerações feitas anteriormente se aplicam a este caso.

2.3 Modelo SEAIR

No modelo SEAIR tem sido utilizado frequentemente na literatura recente sobre o COVID-19 [1], e consiste basicamente na introdução de uma nova variável dinâmica, a dos assintomáticos ou sub-clínicos. Os assintomáticos também podem infectar e também provém dos expostos, mas tem uma taxa de infecção menor, nenhuma mortalidade e em tese podem ter uma taxa de recuperação menor. Nesse caso, as equações dinâmicas podem ser escritas como:

$$\frac{dS}{dt} = \Lambda N - \mu_{eq}S - \beta \frac{S}{N}I - \alpha\beta \frac{S}{N}A \quad (12)$$

$$\frac{dE}{dt} = \beta \frac{S}{N}I + \alpha\beta \frac{S}{N}A - \mu_{eq}E - aE \quad (13)$$

$$\frac{dI}{dt} = a\rho E - \mu_{eq}(t)I - \mu_{cov}I - \gamma I \quad (14)$$

$$\frac{dA}{dt} = a(1 - \rho)E - \mu_{eq}A - \gamma^{(A)}A \quad (15)$$

$$\frac{dR}{dt} = \gamma I + \gamma^{(A)}A - \mu_{eq}R \quad (16)$$

$$\frac{dN}{dt} = \Lambda N - \mu_{eq}N - \mu_{cov}I \quad (17)$$

$$N = S + E + A + I + R \quad (18)$$

nesse modelo, é introduzido o fator α , que representa o quanto a infectividade do assintomático é menor do que a do simptomático (α deve ser um valor positivo menor ou igual a 1), a probabilidade ρ , que é a probabilidade de um exposto se tornar sintomático, a , o coeficiente de taxa de conversão de expostos para infectados ou assintomático, e $\gamma^{(A)}$, o coeficiente de taxa de recuperação de assintomáticos.

2.4 Relação entre γ e μ_{cov}

Os modelos acima mencionados permitem alguma arbitrariedade com relação aos parâmetros, mas algumas relações entre eles precisa ser obedecida para que o modelo seja consistente com o comportamento observado da doença e da epidemia. É necessário garantir, por exemplo, que a porcentagem total de mortos e de recuperados obedeça a proporção observada. Definindo T_R como a razão entre recuperados e o total de infectados e T_C a relação entre casualidades e o total de infectados, que

são observáveis já conhecidas para a COVID-19, temos que os parâmetros γ e μ_{cov} precisam obedecer as seguintes relações

$$T_R = \frac{\gamma}{\gamma + \mu_{cov}} \quad (19)$$

e

$$T_C = \frac{\mu_{cov}}{\gamma + \mu_{cov}} \quad (20)$$

Dessa forma é possível escrever um dos parâmetros, digamos, μ_{cov} , em função dos outros

$$\mu_{cov} = \gamma \frac{T_C}{T_R} \quad (21)$$

3 Desenvolvimento Teórico - Propagação epidêmica sobre um território e por faixa etária

A generalização para considerar a propagação espacial (ou territorial) se baseia na equação da continuidade. Nesse caso, vamos trabalhar com a área total A_T de um território (digamos, o território brasileiro), e discretizá-lo em partes de igual área, A_i , que chamaremos de elemento de área. A forma das áreas é arbitrária, contanto que tenham sempre a mesma área, mas por simplificação podemos imaginá-la como sendo um quadrado. Cada elemento área A_i pode ser caracterizada pela coordenada geográfica central (latitude e longitude).

Em poucas palavras, o problema se resumirá a calcular a evolução das populações dentro dos elementos de área, levando-se em conta o fluxo da população para dentro e para fora dela. Antes disso, vamos introduzir alguma notação e definições importantes:

$A_T :=$ ‘Área total do território.’

$A_i :=$ ‘ i -ésimo elemento de área. Seu valor é simplesmente A .’

$$A_T = \sum_i A_i$$

$s_i = \frac{S}{A_i} :=$ ‘densidade de pessoas susceptíveis no i -ésimo elemento de área.’

$e_i = \frac{E}{A_i} :=$ ‘densidade de pessoas expostas no i -ésimo elemento de área.’

$y_i = \frac{I}{A_i} :=$ ‘densidade de pessoas infectadas no i -ésimo elemento de área.’

$r_i = \frac{r}{A_i} :=$ ‘densidade de pessoas recuperadas no i -ésimo elemento de área.’

$n_i = \frac{N}{A_i} :=$ ‘densidade total de pessoas no i -ésimo elemento de área.’

$\Gamma_n(A_i) :=$ ‘fluxo total de indivíduos através do i -ésimo elemento de área.’

$\Gamma_s(A_i) :=$ ‘fluxo de indivíduos susceptíveis através do i -ésimo elemento de área.’

$\Gamma_e(A_i) :=$ ‘fluxo de indivíduos expostos através do i -ésimo elemento de área.’

$\Gamma_y(A_i) :=$ ‘fluxo de indivíduos infectados através do i -ésimo elemento de área.’

$\Gamma_r(A_i) :=$ ‘fluxo de indivíduos recuperados através do i -ésimo elemento de área.’

Os fluxos tem tanto um contribuição positiva, isto é, de indivíduos que se deslocam de fora para dentro, $\Gamma^{\text{in}}(A_i)$, quanto uma contribuição negativa, isto é, de indivíduos que se deslocam de dentro para fora, $-\Gamma^{\text{out}}(A_i)$. Qualquer um dos fluxos acima definidos podem ser escritos como:

$$\Gamma_k(A_i) = \Gamma_k^{\text{in}}(A_i) - \Gamma_k^{\text{out}}(A_i)$$

onde k é índice que simplesmente representa uma das variáveis populacionais, isto é, $k = n, s, e, y, r$. Chamamos a atenção também para o fato de que o fluxo da população total é igual a soma dos fluxos de suas parcelas, isto é,

$$\Gamma_n(A_i) = \Gamma_s(A_i) + \Gamma_e(A_i) + \Gamma_y(A_i) + \Gamma_r(A_i) . \quad (22)$$

No modelo SIR basta omitir o segundo termo da soma. Vamos mostrar como esses fluxos podem ser modelados nas seções seguintes.

3.1 SIR - Espacial

As equações de taxa agora devem ser resolvidas para cada elemento de área, considerando-se os fluxos de saída e entrada.

$$\frac{ds_i}{dt} = \frac{\Gamma_s(A_i)}{A} + \Lambda n_i - \mu_{eq} s_i - \beta \frac{s_i}{n_i} y_i \quad (23)$$

$$\frac{dy_i}{dt} = \frac{\Gamma_y(A_i)}{A} + \beta \frac{s_i}{n_i} y_i - \mu_{eq} y_i - \mu_{cov} y_i - \gamma y_i \quad (24)$$

$$\frac{dr_i}{dt} = \frac{\Gamma_r(A_i)}{A} + \gamma y_i - \mu_{eq} r_i \quad (25)$$

$$\frac{dn_i}{dt} = \frac{\Gamma_n(A_i)}{A} + \Lambda n_i - \mu_{eq} n_i - \mu_{cov} y_i \quad (26)$$

$$n_i = s_i + y_i + r_i \quad (27)$$

Para obter a população de uma determinada parcela em todo território brasileiro, digamos, do número total de infectados, basta somar sobre todos os elementos de área e multiplicar o resultado pelo valor do elemento de área, *i.e.*

$$s = A \sum_i s_i \quad (28)$$

O acoplamento entre os diversas elementos de área é regido pelos fluxos de entrada, que como veremos mais a frente, é função dos fluxos de saída dos outros elementos de área.

3.2 SEIR - Espacial

As equações para o modelo SEIR podem ser escritas como

$$\frac{ds_i}{dt} = \frac{\Gamma_s(A_i)}{A} + \Lambda n_i - \mu_{eq}s_i - \beta \frac{s_i}{n_i} y_i \quad (29)$$

$$\frac{de_i}{dt} = \frac{\Gamma_e(A_i)}{A} + \beta \frac{s_i}{n_i} y_i - \mu_{eq}e_i(t) - e_i(t) \quad (30)$$

$$\frac{dy_i}{dt} = \frac{\Gamma_y(A_i)}{A} + e_i(t) - \mu_{eq}y_i - \mu_{cov}y_i - \gamma y_i \quad (31)$$

$$\frac{dr_i}{dt} = \frac{\Gamma_r(A_i)}{A} + \gamma y_i - \mu_{eq}r_i \quad (32)$$

$$\frac{dn_i}{dt} = \frac{\Gamma_n(A_i)}{A} + \Lambda n_i - \mu_{eq}n_i - \mu_{cov}y_i \quad (33)$$

$$n_i = s_i + e_i + y_i + r_i \quad (34)$$

É importante considerar que os parâmetros do modelo podem agora ser considerados localmente, se necessário, levando-se em conta as peculiaridades de cada região do território nacional.

3.3 Modelagem dos fluxos populacionais

O fluxo total de saída em um elemento de área i pode ser prescrito, com base em conhecimento prévio sobre a região, ou por meio de algum tipo de modelo. É razoável considerar, por exemplo, que o fluxo de saída é proporcional à densidade populacional:

$$\Gamma_n^{\text{out}}(A_i) = \xi n_i . \quad (35)$$

Sabendo o fluxo total, podemos considerar que os fluxos das parcelas é proporcional às parcelas, isto é,

$$\begin{aligned} \Gamma_s^{\text{out}}(A_i) &= \frac{s_i}{n_i} \Gamma_n^{\text{out}}(A_i) \\ \Gamma_e^{\text{out}}(A_i) &= \frac{e_i}{n_i} \Gamma_n^{\text{out}}(A_i) \\ \Gamma_y^{\text{out}}(A_i) &= \frac{y_i}{n_i} \Gamma_n^{\text{out}}(A_i) \\ \Gamma_r^{\text{out}}(A_i) &= \frac{r_i}{n_i} \Gamma_n^{\text{out}}(A_i) \end{aligned}$$

Para modelar os fluxos de entrada em um determinado elemento de área A_i , nós temos que considerar que todas os outros elementos de área contribuem para o fluxo.

Podemos escrever, de modo geral, o fluxo de entrada total em um elemento de área A_i como

$$\Gamma_n^{\text{in}}(A_i) = \sum_{j \neq i} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \quad (36)$$

onde a soma é feita sobre todo os elementos de área, menos o elemento i , d_{ij} é a distância entre o elemento de área j e i e $f(d_{ij})$ é uma função decrescente de d_{ij} . Essa função deve ter a seguinte propriedade

$$\sum_{i \neq j} f(d_{ij}) = 1 \quad (37)$$

A função $f(d_{ij})$ deve ser interpretada como a fração do fluxo de população em A_j que se desloca para A_i . Os fluxos parciais também podem ser considerados proporcionais às parcelas populacionais, ou seja,

$$\begin{aligned} \Gamma_s^{\text{in}}(A_i) &= \sum_{j \neq i} \frac{s_j}{n_j} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \\ \Gamma_e^{\text{in}}(A_i) &= \sum_{j \neq i} \frac{e_j}{n_j} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \\ \Gamma_y^{\text{in}}(A_i) &= \sum_{j \neq i} \frac{y_j}{n_j} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \\ \Gamma_r^{\text{in}}(A_i) &= \sum_{j \neq i} \frac{r_j}{n_j} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \end{aligned}$$

3.4 Transformação para cidades/microrregiões

No modelo apresentado anteriormente, definimos um elemento de área arbitrário do território nacional. Note que as equações diferenciais desse modelo podem ser escritas independentemente do valor da área, bastando multiplicá-las membro a membro pela densidade A . Em geral, podemos trabalhar com o número total de municípios ou microrregiões em vez de trabalhar com elementos de área arbitrários. Tal abordagem facilita a implementação, a integração das informações do IBGE e a interpretação dos resultados. As equações do modelo SEIR, por exemplo, tomam a seguinte forma

$$\frac{dS_i}{dt} = \Gamma_i^s + \Lambda N_i - \mu_{eq} S_i - \beta \frac{S_i}{N_i} I_i \quad (38)$$

$$\frac{dE_i}{dt} = \Gamma_i^e + \beta \frac{S_i}{N_i} I_i - \mu_{eq} E_i(t) - E_i(t) \quad (39)$$

$$\frac{dI_i}{dt} = \Gamma_i^y + E_i(t) - \mu_{eq} I_i - \mu_{cov} I_i(t) - \gamma I_i(t) \quad (40)$$

$$\frac{dR_i}{dt} = \Gamma_i^r + \gamma I_i(t) - \mu_{eq} R_i \quad (41)$$

$$\frac{dN_i}{dt} = \Gamma_i^n + \Lambda N_i - \mu_{eq} N_i - \mu_{cov} I_i(t) \quad (42)$$

$$N_i = S_i + E_i + I_i + R_i \quad (43)$$

Agora cada população ou parcela populacional (S_i, E_i, I_i, R_i) é indexada pela i -ésima microrregião ou i -ésimo município. A interpretação para os fluxos e para a função $f(d_{i,j})$ é análoga ao caso anterior. Note também que mudamos a notação do fluxo; o sub-índice i indica a microrregião/município e o índice superescrito indica a parcela populacional.

3.5 Inclusão de faixa etária

A inclusão da faixa etária pode ser realizada a partir da quebra da população em estratos de faixa etária. Dessa forma, uma dada variável de população, digamos, a população total $N_{i,k}$, é agora indexada pela sua i -ésima microrregião e sua k -ésima faixa etária. No Brasil, por exemplo, temos $M = 558$ microrregiões e podemos, por exemplo, trabalhar com faixas etárias espaçadas em cinco anos até 80 ou mais, totalizando $Q = 17$ faixas etárias. Considerando que temos quatro parcelas populacionais (S, E, I, R) , teremos no total $4 \times 558 \times 17 = 37944$ variáveis dependentes do tempo e $\times 558 \times 17 = 9486$ sistemas de equações diferenciais ordinárias a resolver. Apesar de parecer muito, é um cálculo factível para os padrões computacionais atuais. Antes de escrever as equações, é necessário introduzir uma nova notação.

$S_{ik} :=$ ‘número de indivíduos susceptíveis na i -ésima região e k -ésima faixa etária.’
 $E_{ik} :=$ ‘número de indivíduos expostos na i -ésima região e k -ésima faixa etária.’
 $I_{ik} :=$ ‘número de indivíduos infectados na i -ésima região e k -ésima faixa etária.’
 $R_{ik} :=$ ‘número de indivíduos recuperados na i -ésima região e k -ésima faixa etária.’
 $N_{ik} :=$ ‘número total de indivíduos na i -ésima região e k -ésima faixa etária.’

$\Gamma_{ik}^n :=$ ‘fluxo total de indivíduos através da i -ésima região, na k -ésima faixa etária.’
 $\Gamma_{ik}^s :=$ ‘fluxo total de susceptíveis através da i -ésima região, na k -ésima faixa etária.’
 $\Gamma_{ik}^e :=$ ‘fluxo total de expostos através da i -ésima região, na k -ésima faixa etária.’
 $\Gamma_{ik}^y :=$ ‘fluxo total de infectados através da i -ésima região, na k -ésima faixa etária.’
 $\Gamma_{ik}^r :=$ ‘fluxo total de recuperados através da i -ésima região, na k -ésima faixa etária.’

$$\begin{aligned}
S_i &= \sum_k S_{ik} & ; & & \Gamma_i^s &= \sum_k \Gamma_{ik}^s \\
E_i &= \sum_k E_{ik} & ; & & \Gamma_i^e &= \sum_k \Gamma_{ik}^e \\
I_i &= \sum_k I_{ik} & ; & & \Gamma_i^y &= \sum_k \Gamma_{ik}^y \\
R_i &= \sum_k R_{ik} & ; & & \Gamma_i^r &= \sum_k \Gamma_{ik}^r \\
N_i &= \sum_k N_{ik} & ; & & \Gamma_i^n &= \sum_k \Gamma_{ik}^n
\end{aligned}$$

As equações do modelo SEIR podem então ser escritas como:

$$\frac{d}{dt}S_{ik} = \Gamma_{ik}^s + \delta_{1k}\Lambda_{ik}N_{ik} - \mu_{eq,ik}S_{ik} - \sum_j \beta_{i,kj} \frac{S_{ik}}{N_i} I_{ij}(t) \quad (44)$$

$$\frac{d}{dt}E_{ik} = \Gamma_{ik}^e + \sum_j \beta_{i,kj} \frac{S_{ik}}{N_i} I_{ij}(t) - \mu_{eq,ik}E_{ik} + E_{ik}(t) \quad (45)$$

$$\begin{aligned} \frac{d}{dt}I_{ik} &= \Gamma_{ik}^y + E_{ik}(t) - \mu_{ik,eq}I_{ik}(t) - \mu_{ik,cov}I_{ik}(t) \\ &\quad - \gamma_{ik}I_{ik}(t) \end{aligned} \quad (46)$$

$$\frac{d}{dt}R_{ik} = \Gamma_{ik}^r + \gamma_{ik}I_{ik}(t) - \mu_{eq,ik}R_{ik} \quad (47)$$

$$\frac{d}{dt}N_{ik} = \Gamma_{ik}^n + \delta_{1k}\Lambda_{ik}N_{ik} - \mu_{eq,ik}N_{ik} - \mu_{ik,cov}I_{ik}(t) \quad (48)$$

$$N_{ik} = S_{ik} + E_{ik} + I_{ik} + R_{ik} \quad (49)$$

onde δ_{1k} é a função delta de Kronecker (1 se $k = j$ e 0 noutro caso) e $\beta_{i,kj}$ é o coeficiente de taxa de exposição que um indivíduo susceptível da faixa etária k tem a um indivíduo infectado da faixa etária j , na região i . Note na notação aqui adotada a dependência dos parâmetros com a região e faixa etária estão explicitadas por meio dos índices subscritos, por exemplo, $\mu_{eq,ik}$ é o coeficiente de taxa de mortalidade típica na faixa etária k na região i .

Em uma escala temporal maior do que a escala típica das epidemias, o envelhecimento da população produz um fluxo entre as faixas etárias, apenas positivo. Nas equação acima, não levamos em consideração esse fluxo contínuo, desprezível contanto que as faixas etárias não sejam mais estreitas do que um ano.

Por completeza, incluimos também as equações do modelo SEAIR:

$$\frac{d}{dt}S_{ik} = \Gamma_{ik}^s + \delta_{1k}\Lambda_{ik}N_{ik} - \mu_{eq,ik}S_{ik} - \sum_j \beta_{i,kj} \frac{S_{ik}}{N_i} I_{ij} - \sum_j \alpha_{i,j} \beta_{i,kj} \frac{S_{ik}}{N_i} A_{ij} \quad (50)$$

$$\frac{d}{dt}E_{ik} = \Gamma_{ik}^e + \sum_j \beta_{i,kj} \frac{S_{ik}}{N_i} I_{ij} - \sum_j \alpha_{i,j} \beta_{i,kj} \frac{S_{ik}}{N_i} A_{ij} - \mu_{eq,ik}E_{ik} - a_{ik}E_{ik} \quad (51)$$

$$\frac{d}{dt}A_{ik} = \Gamma_{ik}^a + (1 - \rho_{ik})a_{ik}E_{ik} - \mu_{ik,eq}A_{ik} - \gamma_{ik}^{(A)}A_{ik} \quad (52)$$

$$\frac{d}{dt}I_{ik} = \Gamma_{ik}^y + \rho_{ik}a_{ik}E_{ik} - \mu_{ik,eq}I_{ik} - \mu_{ik,cov}I_{ik}(t) - \gamma_{ik}I_{ik} \quad (53)$$

$$\frac{d}{dt}R_{ik} = \Gamma_{ik}^r + \gamma_{ik}I_{ik}(t) + \gamma_{ik}^{(A)}A_{ik} - \mu_{eq,ik}R_{ik} \quad (54)$$

$$\frac{d}{dt}N_{ik} = \Gamma_{ik}^n + \delta_{1k}\Lambda_{ik}N_{ik} - \mu_{eq,ik}N_{ik} - \mu_{ik,cov}I_{ik} \quad (55)$$

$$N_{ik} = S_{ik} + E_{ik} + A_{ik} + I_{ik} + R_{ik} \quad (56)$$

3.5.1 Modelo global com faixas etárias

É possível escrever as equações SEIR considerando as faixas etárias, desconsiderando no entanto sua distribuição espacial. Isso pode ser útil em modelos globais em que a informação sobre os coeficientes de taxa de infecção por faixa etária é importante. Nesse caso, os índices subscritos indicam apenas a faixa etária, e não mais a região.

$$\frac{d}{dt}S_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,k}S_k - \sum_j \beta_{kj} \frac{S_k}{N} I_j \quad (57)$$

$$\frac{d}{dt}E_k = \sum_j \beta_{kj} \frac{S_k}{N} I_j - \mu_{eq,k}E_k(t) - a_k E_k \quad (58)$$

$$\frac{d}{dt}I_k = a_k E_k - \mu_{k,eq}I_k - \mu_{cov,k}I_k - \gamma_k I_k \quad (59)$$

$$\frac{d}{dt}R_k = \gamma_k I_k - \mu_{eq,k}R_k \quad (60)$$

$$\frac{d}{dt}N_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,k}N_k - \mu_{cov,k}I_k \quad (61)$$

$$N_k = S_k + E_k + I_k + R_k \quad (62)$$

O modelo SIR pode ser escrito como:

$$\frac{d}{dt}S_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,k}S_k - \sum_j \beta_{kj} \frac{S_k}{N} I_j \quad (63)$$

$$\frac{d}{dt}I_k = \sum_j \beta_{kj} \frac{S_k}{N} I_j - \mu_{k,eq}I_k - \mu_{k,cov}I_k - \gamma_k I_k \quad (64)$$

$$\frac{d}{dt}R_k = \gamma_k I_k - \mu_{eq,k}R_k \quad (65)$$

$$\frac{d}{dt}N_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,ik}N_k - \mu_{k,cov}I_k \quad (66)$$

$$N_k = S_k + I_k + R_k \quad (67)$$

Por último, o modelo SEAIR:

$$\frac{d}{dt}S_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,k}S_k - \sum_j \alpha_j \beta_{kj} \frac{S_k}{N} A_j - \sum_j \beta_{kj} \frac{S_k}{N} I_j \quad (68)$$

$$\frac{d}{dt}E_k = \sum_j \beta_{kj} \frac{S_k}{N} I_j + \sum_j \alpha_j \beta_{kj} \frac{S_k}{N} A_j - \mu_{eq,k}E_k - a_k E_k \quad (69)$$

$$\frac{d}{dt}I_k = a_k \rho_k E_k - \mu_{k,eq}I_k - \mu_{k,cov}I_k - \gamma_k I_k \quad (70)$$

$$\frac{d}{dt}A_k = a_k(1 - \rho_k)E_k - \mu_{k,eq}A_k - \gamma^{(A)} A_k \quad (71)$$

$$\frac{d}{dt}R_k = \gamma_k I_k + \gamma^{(A)} A_k - \mu_{eq,k}R_k \quad (72)$$

$$\frac{d}{dt}N_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,ik}N_k - \mu_{cov,k}I_k \quad (73)$$

$$N_k = S_k + E_k + A_k + I_k + R_k \quad (74)$$

3.5.2 Matriz de contato

Um dos grandes problemas está em calcular a matriz de β_{kj} . Essa matriz pode ser escrita como a probabilidade p de um contato resultar na infecção vezes a chamada matriz de contato C_{kj} ($\beta_{kj} = pC_{kj}$). Essa matriz contabiliza o número de contatos que uma população da faixa k faz com uma população da faixa j . Felizmente, encontramos na literatura trabalhos recentes que calculam a matriz de contato para vários países, inclusive o Brasil (Prem *et al* [2]). Nesses trabalhos, a matriz de contato total resulta da soma de matriz de contato de casa C_{kj}^H , matriz de contato do trabalho

C_{kj}^W , matriz de contato de escola C_{kj}^S e matriz de contato de outras origens C_{kj}^O

$$C_{kj} = C_{kj}^H + C_{kj}^W + C_{kj}^S + C_{kj}^O \quad (75)$$

Como não sabemos o valor de p , vamos trabalhar com o valor de β , estimado ou pela regressão linear no início da curva ou estimado pelo valor de R_0 . Nesse caso, nós precisamos escalar a matriz, primeiro normalizando-a

$$\mathbf{C}_{\text{esc}} = \frac{\mathbf{C}}{(\text{Max}(\text{Re}(\text{Eig}(\mathbf{C}))))} \quad (76)$$

para então multiplicá-la por $\beta \sim \gamma \cdot R_0$ ($\beta_{kj} = \beta C_{esc,kj}$). Se o β foi obtido a partir do modelo não estratificado, ele pode ser posteriormente refinado para reajustar a curva.

3.6 Cenários

3.6.1 Nenhuma intervenção. Cenário 1

Deve-se utilizar a matriz de contato total dada pela equação 75.

3.6.2 Fechamento de escola apenas. Cenário 2

Nesse caso, considera-se que não há mais contato entre indivíduos no ambiente escolar. A matriz escolar é portanto zerada.

$$C_{kj} = C_{kj}^H + C_{kj}^W + 0 \cdot C_{kj}^S + C_{kj}^O \quad (77)$$

3.6.3 Fechamento de escola e distanciamento social. Cenário 3

Nesse caso, considera-se um aumento de 50% no contato dos jovens (0 a 20 anos) e um aumento de 10% das outras faixas etárias na matriz casa. A matriz escolar é zerada. Considera-se uma diminuição de 50% nos contatos dos jovens na comunidade.

$$C_{kj} = \mathbf{A} \cdot C_{kj}^H + C_{kj}^W + 0 \cdot C_{kj}^S + \mathbf{B} \cdot C_{kj}^O \quad (78)$$

onde \mathbf{A} é uma matriz diagonal, $A_{ii} = 1.5$ para $i = 1$ a 4 e $A_{ii} = 1.1$ para os outros valores de i , e \mathbf{B} é uma matriz diagonal, $B_{ii} = 0.5$ para $i = 1$ a 4 e $B_{ii} = 1.0$.

3.6.4 Fechamento de escola, distanciamento no trabalho e distanciamento social. Cenário 4

Nesse caso, considera-se um aumento de 50% no contato dos jovens (0 a 20 anos) e um aumento de 10% das outras faixas etárias na matriz casa. A matriz escolar é zerada. Considera-se uma diminuição de 50% nos contatos dos jovens na comunidade. Considera-se também uma diminuição de 50% nos contatos de trabalho.

$$C_{kj} = \mathbf{A} \cdot C_{kj}^H + \zeta_t \cdot C_{kj}^W + 0 \cdot C_{kj}^S + \mathbf{B} \cdot C_{kj}^O \quad (79)$$

onde ζ_t é um escalar, com $\zeta = 0.5$ qualquer que seja o $i = 1$. Ela representa uma diminuição de 50% dos contatos no trabalho. Para fechamento do trabalho, basta multiplicar a matriz de trabalho por 0.

3.6.5 Isolamento dos idosos. Cenário 5.

Nesse cenário, os contatos dos idosos ficam restritos à casa. Nesse caso deve-se realizar a seguinte transformação de simetria sobre as matrizes:

$$C_{kj} = C_{kj}^H + (\tilde{\mathbf{I}} \cdot C_{kj}^W \cdot \tilde{\mathbf{I}}) + (\tilde{\mathbf{I}} \cdot C_{kj}^S \cdot \tilde{\mathbf{I}}) + (\tilde{\mathbf{I}} \cdot C_{kj}^O \cdot \tilde{\mathbf{I}}) \quad (80)$$

onde $\tilde{\mathbf{I}}$ é uma matriz diagonal tal que $\tilde{I}_{kk} = 1$ se $k < i_{th}$ e $\tilde{I}_{kk} = 0$ se $k \geq i_{th}$. O índice fixo i_{th} é limiar de idade. Caso o isolamento comece na faixa de 55 anos para cima, $i_{th} = 12$.

3.6.6 Fechamento de escola, distanciamento no trabalho, distanciamento social e isolamento dos idosos. Cenário 6

Nesse caso, os contatos dos idosos ficam restritos à casa. Considera-se um aumento de 50% no contato dos jovens (0 a 20 anos) e um aumento de 10% das outras faixas etárias na matriz casa. A matriz escolar é zerada. Considera-se uma diminuição de 50% nos contatos dos jovens na comunidade. Considera-se também uma diminuição de 50% nos contatos de trabalho.

$$C_{kj} = \mathbf{A} \cdot C_{kj}^H + \zeta_t \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^W \cdot \tilde{\mathbf{I}}) + 0 \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^S \cdot \tilde{\mathbf{I}}) + \mathbf{B} \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^O \cdot \tilde{\mathbf{I}}) \quad (81)$$

3.6.7 Knock Down com isolamento de idosos. Cenário 7

Nesse cenário, os contatos dos idosos ficam restritos à casa. Considera-se um aumento de 50% no contato dos jovens (0 a 20 anos) e um aumento de 10% das outras faixas etárias na matriz casa. A matriz escolar é zerada. Considera-se uma diminuição de 25% nos contatos dos jovens e 50% nos contatos dos adultos na comunidade. Considera-se também uma diminuição de 25% nos contatos de trabalho.

$$C_{kj} = \mathbf{A} \cdot C_{kj}^H + \zeta_t \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^W \cdot \tilde{\mathbf{I}}) + 0 \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^S \cdot \tilde{\mathbf{I}}) + \zeta_c \mathbf{B} \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^O \cdot \tilde{\mathbf{I}}) \quad (82)$$

onde $\zeta_t = 0.25$ e $\zeta_c = 0.5$.

3.6.8 Normalização

A normalização ou procedimento de escalar a matriz deve ser feito após somar das matrizes não escalonadas no cenário sem intervenção. O determinante obtido é o fator de normalização para todos os outros cenários.

3.7 Modelagem do fluxo entre as microrregiões

Deixamos o problema do fluxo entre regiões em aberto, e agora o retomamos após termos feito a generalização do modelo para os estratos. O modelo que propomos supõe que o fluxo de saída de uma microrregião é dado pelo movimento pendular, que atinge, proporcionalmente, apenas as microrregiões adjacentes, e o fluxo de viagens aéreas, que podem atingir qualquer microrregião, contanto que existam voos diretos entre ponto de partida e destino. Uma abordagem parecida foi adotada no estudo de Coelho *et al* [3]. Assim sendo, vamos trabalhar com as seguintes definições:

- Fluxo aéreo total para fora de uma região i ($\Gamma_{ik}^{n,(aer-out)}$):

$$\Gamma_{ik}^{n,(aer-out)} = \frac{N_{ik}}{N_i} \sum_{j, aer(i)} \Gamma_{i \rightarrow j}^{n,(aer)} \quad (83)$$

onde $\Gamma_{i \rightarrow j}^{n,(aer)}$ é o fluxo aéreo (indivíduos por dia) indo de i para j , e a soma $\sum_{j, aer(i)}$ é definida apenas sobre as microrregiões j com conexão aérea com i no sentido $i \rightarrow j$.

- Fluxo aéreo total para dentro de uma região i , $\Gamma_{ik}^{n,(aer-in)}$:

$$\Gamma_{ik}^{n,(aer-in)} = \sum_{j, aer(i)} \frac{N_{jk}}{N_j} \Gamma_{j \rightarrow i}^{n,(aer)} \quad (84)$$

onde $\Gamma_{j \rightarrow i}^{n,(aer)}$ é o fluxo aéreo (indivíduos por dia) vindo de j para i , e a soma $\sum_{j,aer(i)}$ é definida apenas sobre as microrregiões j com conexão aérea com i no sentido $j \rightarrow i$.

- Fluxo de saída total ($\Gamma_{ik}^{n,(out)}$):

$$\Gamma_{ik}^{n,(out)} = \frac{N_{ik}}{N_i} \left(\Gamma_i^{n,(pend)} + \Gamma_i^{n,(ae-out)} \right) \quad (85)$$

o fluxo de saída pendular é um dado fornecido pelas estatísticas do IBGE, e existem valores para cada microrregião. Note que consideramos que o fluxo por faixa etária é proporcional à fração de indivíduos naquela faixa etária em uma dada região.

- Fluxo de entrada total ($\Gamma_{ik}^{n,(in)}$):

$$\Gamma_{ik}^{n,(in)} = \left(\sum_{j,adj(i)} \frac{N_{jk}}{\sum_{l,adj(i)} N_l} \Gamma_j^{n,(pend)} \right) + \Gamma_{ik}^{n,(aer-in)} \quad (86)$$

Aqui, a soma $\sum_{j,adj(i)}$ é definida apenas para as microrregiões j adjacentes a i .

Para obter o fluxo médio diário de passageiros entre as microrregiões pode-se consultar o guia OAG (Official Airline Guide) (<http://www.oag.com>) [4]. Para obter o movimento pendular, pode-se consultar os senso demográfico de 2010 [5].

3.8 Cálculo de grandezas úteis: número de casualidades (C), hospitalizações (H), leitos ocupados (L) e número final total de infectados (M)

Há certas grandezas importantes que não entram no modelo como grandezas dinâmicas em um primeiro momento, mas que podem ser calculadas com base nos parâmetros e nas variáveis dinâmicas do modelo. A seguir vamos mostrar como calcular essas variáveis nos modelos (SIR, SEIR, SEAIR) estratificados. A generalização para o caso territorial é imediata.

3.8.1 Casualidades

A taxa de mortalidade é simplesmente dada por:

$$\frac{dC_k}{dt} = \mu_{k,cov} I_k(t) \quad (87)$$

onde $\mu_{k,cov}$ é a taxa de mortalidade típica do COVID-19 por faixa etária. Para obter o número $C_k(t)$ em função do tempo, a equação acima deve ser integrada numericamente a partir da função I determinada numericamente. Note que a equação acima vale para os modelos SIR, SEIR e SEAIR. O número total é obtido fazendo a soma $C(t) = \sum_k C_k(t)$.

3.8.2 Hospitalizações

A taxa de hospitalização é dada pela equação

$$\frac{dH_k}{dt} = \phi_k \sum_j \beta_{kj} \frac{S_k}{N} I_j - \gamma_{H,k} H_k \quad (88)$$

onde ϕ_k é a porcentagem média de infectados que precisarão de hospitalização, $\gamma_{H,k}$ é o coeficiente de taxa de recuperação no hospital. Note que a equação acima vale para os modelos SIR, SEIR e SEAIR, sem levar em conta o tempo de incubação. O número total é obtido fazendo a soma $H(t) = \sum_k H_k(t)$.

3.8.3 Leitos ocupados

A taxa de ocupação de leitos é dada pela equação

$$\frac{dL_k}{dt} = \eta_k \sum_j \beta_{kj} \frac{S_k}{N} I_j - \gamma_{L,k} (1 + T_{L,C}/(100 - T_{L,C})) L_k \quad (89)$$

onde η_k é a porcentagem média de infectados que precisarão de leitos em UTI, $\gamma_{L,k}$ é o coeficiente de taxa de recuperação na UTI e $T_{L,C}$ é a porcentagem de mortalidade na UTI. Note que a equação acima vale para os modelos SIR, SEIR e SEAIR, sem levar em conta o tempo de incubação. O número total é obtido fazendo a soma $L(t) = \sum_k L_k(t)$.

3.8.4 Número total casos M

Definimos como número final total de infectados $M(t)$, ou acumulado de infectados, como o número de indivíduos que pegaram a doença e que ou se recuperaram ou morreram até o instante t . Nos modelos SIR e SEIR esse número é dado simplesmente

como

$$M_k(t) = I_k(t) + R_k(t) + C_k(t) \quad (90)$$

No modelo SEAIR, esse número é dado por

$$M_k(t) = A_k(t) + I_k(t) + R_k(t) + C_k(t) \quad (91)$$

O número total é obtido fazendo a soma $M(t) = \sum_k M_k(t)$. Nos modelos SIR e SEIR pode-se usar essa grandeza para comparação com os dados reportados, e $M - A$ no caso SEAIR, mas é necessário tomar alguns cuidados especiais, pois é sempre esperado que os dados sejam subnotificados, e não representem a totalidade dos infectados.

3.8.5 Outras saídas importantes

Outras saídas importantes do código são:

- $I_{\max}(t)$: número máximo de infecções;
- $t_{\max,I}$: instante em que $I_{\max}(t)$ ocorre;
- $L_{\max}(t)$: número máximo de leitos ocupados
- $t_{\max,L}$: instante em que $L_{\max}(t)$ ocorre;
- $H_{\max}(t)$: número máximo de hospitalizações;
- $t_{\max,H}$: instante em que $H_{\max}(t)$ ocorre;
- M_{\max} : número final de casos, t ;

4 Desenvolvimento Teórico - Modelo SEAHIR-Q_{AI}

Os modelos de compartimento permitem incluir aspectos mais detalhados do sistema por meio da inclusão de novas variáveis dinâmicas. Neste modelo, SEAHIR-Q_{AI}, consideramos sete diferentes compartimentos: Suscetíveis (S), Expostos (E), Infectados sintomáticos (I), Assintomáticos (A), assintomáticos quarentenados (Q_A), sintomáticos quarentenados (Q_I) e hospitalizados (H), que obedecem uma equação tipo SEIR, dado por:

$$\frac{dS}{dt} = -\frac{\beta IS}{N} - \frac{\alpha \beta AS}{N}, \quad (92)$$

$$\frac{dE}{dt} = \frac{\beta IS}{N} + \frac{\alpha \beta AS}{N} - aE, \quad (93)$$

$$\frac{dI}{dt} = \rho E - (\gamma_H + \gamma_{R,I} + \gamma_{Q,I})I, \quad (94)$$

$$\frac{dA}{dt} = (1 - \rho)E - (\gamma_{R,A} + \gamma_{Q,A})A, \quad (95)$$

$$\frac{dQ_I}{dt} = \gamma_{Q,I}I - \gamma_{H,Q_I}Q_I - \gamma_{R,Q_I}Q_I, \quad (96)$$

$$\frac{dQ_A}{dt} = \gamma_{Q,A}A - \gamma_{R,Q_A}Q_A, \quad (97)$$

$$\frac{dH}{dt} = \gamma_H I + \gamma_{H,Q_I}Q_I - \gamma_{H,R}H - \mu_{cov}H, \quad (98)$$

neste equação, são definidos novos coeficientes de taxa do tipo γ que controlam os fluxos entre os compartimentos. O coeficiente γ_H é o coeficiente de taxa de hospitalização de um infectado, e está relacionado com o tempo d_H , o coeficiente $\gamma_{R,I}$ é o coeficiente de taxa de recuperação do infectado não quarentenado. O coeficiente γ_{R,Q_I} é o coeficiente de taxa de recuperação do quarentenado sintomático, γ_{H,Q_I} é o coeficiente de taxa de hospitalização do sintomático quarentenado e $\gamma_{Q,I}$ é o coeficiente de taxa de movimento do infectado para a quarentena, em razão dos testes de confirmação. Da mesma forma, o assintomático também pode ser levado à quarentena, por força de teste positivo, e o coeficiente de taxa que o representa é $\gamma_{Q,A}$. O coeficiente de taxa de recuperação do assintomático não quarentenado é dado por $\gamma_{R,A}$ e a do assintomático quarentenado é γ_{R,Q_A} . Para simplificar a equação, foram omitidos os termos de taxa de natalidade e mortalidade de equilíbrio, mas no caso mais geral devem ser incluídos.

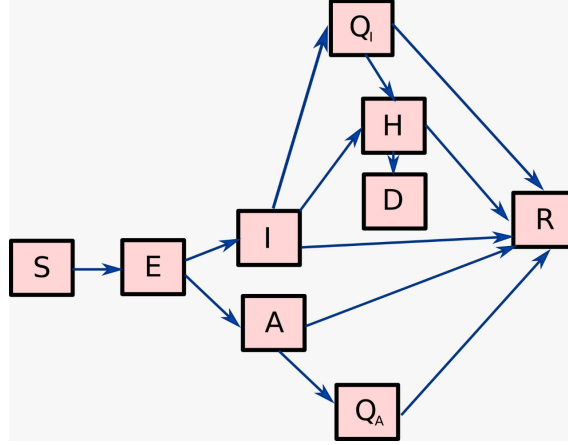


Figura 1: Relações entre compartimentos no modelo SEAHIR- Q_{AI} . O compartimento D representa o número acumulado de mortos até um instante t .

Os coeficientes $\gamma_{Q,I}$, γ_H , $\gamma_{R,I}$ estão relacionados entre si com a probabilidade de testar um infectado sintomático x_I :

$$x_I = \frac{\gamma_{Q,I}}{\gamma_H + \gamma_{R,I} + \gamma_{Q,I}}, \quad (99)$$

e os coeficientes $\gamma_{Q,A}$, $\gamma_{R,A}$ com a probabilidade de testar um infectado assintomático:

$$x_A = \frac{\gamma_{Q,A}}{\gamma_{R,A} + \gamma_{Q,A}}. \quad (100)$$

Os coeficientes γ_H , $\gamma_{R,I}$ e $\gamma_{R,A}$ são prescritos, e estão relacionados com os tempos típicos de recuperação e o tempo típico de hospitalização desde a manifestação dos sintomas. As probabilidades, por sua vez, são prescritas. Em uma política de testagem geral, x_I e x_A são diferentes de zero. Em uma política de nenhuma testagem de indivíduos não hospitalizados, ambos valores são zero. As probabilidades podem ser “ligadas” e “desligadas”, variando no tempo.

A figura 1 mostra a relação entre os compartimentos. Eliminando o teste ou a quarentena dos testados, o modelo de compartimento se reduz ao SEAHIR, que tem o compartimento H a mais com relação ao modelo SEAIR discutido na seção anterior.

No modelo SEAHIR- Q_{AI} o número acumulado de óbitos é calculado a partir da integração da equação:

$$\frac{dC}{dt} = -\mu_{cov} H. \quad (101)$$

Agora μ_{cov} é condicionado pelas porcentagem de hospitalizados que se recuperam $T_{L,R}$ e pela porcentagem de hospitalizados que morrem $T_{L,C}$ e é dado pela equação

$$\mu_{cov} = \gamma_{H,Q} \frac{T_{L,C}}{T_{L,R}} = \gamma_{H,Q} \frac{T_{L,C}}{100 - T_{L,C}} . \quad (102)$$

O número de leitos ocupados pode ser estimado por meio da expressão

$$L(t) = \frac{\eta}{\phi} H(t) . \quad (103)$$

Por fim, o número acumulado de infectados em um dado instante de tempo será dado por

$$M(t) = I(t) + A(t) + Q_I(t) + Q_A(t) + H(t) + C(t) + R(t) . \quad (104)$$

4.1 Generalização do modelo modelo SEAHIR-Q_{AI}

Por completeza, incluímos também as equações estratificadas e por faixa etária do modelo SEAHIR-Q_{AI}:

$$\begin{aligned} \frac{d}{dt} S_{ik} = & \Gamma_{ik}^s + \delta_{1k} \Lambda_{ik} N_{ik} - \mu_{eq,ik} S_{ik} - \sum_j \beta_{i,kj} \frac{S_{ik}}{N_i} I_{ij} \\ & - \sum_j \alpha_{i,j} \beta_{i,kj} \frac{S_{ik}}{N_i} A_{ij} \end{aligned} \quad (105)$$

$$\frac{d}{dt} E_{ik} = \Gamma_{ik}^e + \sum_j \beta_{i,kj} \frac{S_{ik}}{N_i} I_{ij} - \sum_j \alpha_{i,j} \beta_{i,kj} \frac{S_{ik}}{N_i} A_{ij} - \mu_{eq,ik} E_{ik} - a_{ik} E_{ik} \quad (106)$$

$$\frac{d}{dt} A_{ik} = \Gamma_{ik}^a + (1 - \rho_{ik}) a_{ik} E_{ik} - \mu_{ik,eq} A_{ik} - (\gamma_{ik}^{(R,A)} + \gamma_{ik}^{(Q,A)}) A_{ik} \quad (107)$$

$$\frac{d}{dt} I_{ik} = \Gamma_{ik}^y + \rho_{ik} a_{ik} E_{ik} - \mu_{ik,eq} I_{ik} - (\gamma_{ik}^H + \gamma_{ik}^{R,I} + \gamma_{ik}^{Q,I}) I_{ik} \quad (108)$$

$$\frac{d}{dt} Q_{ik}^I = \gamma_{ik}^{Q,I} I_{ik} - \gamma_{ik}^{H,Q_I} Q_{ik}^I - \mu_{eq,ik} Q_{ik}^I - \gamma_{ik}^{R,Q_I} Q_{ik}^I \quad (109)$$

$$\frac{d}{dt} Q_{ik}^A = \gamma_{ik}^{Q,A} A_{ik} - \gamma_{ik}^{R,Q_A} Q_{ik}^A - \mu_{eq,ik} Q_{ik}^A \quad (110)$$

$$\frac{d}{dt} H_{ik} = \gamma_{ik}^H I_{ik} + \gamma_{ik}^{H,Q_I} Q_{ik}^I - \gamma_{eq,ik}^{H,R} H_{ik} - (\mu_{eq,ik} + \mu_{cov,ik}) H_{ik} \quad (111)$$

$$\begin{aligned} \frac{d}{dt} R_{ik} = & \Gamma_{ik}^r + \gamma_{ik}^{R,I} I_{ik} + \gamma_{ik}^{R,A} A_{ik} + \gamma_{ik}^{R,Q_A} Q_{ik}^A + \gamma_{ik}^{H,Q_I} Q_{ik}^I + \gamma_{eq,ik}^{H,R} H_{ik} \\ & - \mu_{eq,ik} R_{ik} \end{aligned} \quad (112)$$

5 Parâmetros dos modelos compartimentados

5.1 Metodologia para estimativa dos parâmetros

Um problema chave na implementação dos códigos é o conhecimento sobre os parâmetros do modelo. Como a quantidade de parâmetros é muito grande, é impossível obter todos eles por ajuste pelo método dos mínimos quadrados ou outra metodologia utilizando apenas informação de infectados notificados no início da epidemia. Por isso, é importante que alguns parâmetros, supostos intrínsecos à natureza da doença, possam ser fornecidos ou estimados a partir de outros estudos. Nós vamos adotar essa abordagem, e fazer uso dos dados das curvas mais confiáveis para ajustar parâmetros quando necessário.

Na tabela 1 apresentamos os valores de referência dos parâmetros e algumas observações sobre sua aplicação. São parâmetros globais, baseados no conhecimento mais atualizado sobre a COVID-19. Os valores regionalizados e estratificados estão apresentados em uma planilha anexada como material suplementar.

Tabela 1: Tabelas de parâmetros para modelos SIR, SEIR e SAEIR

Parâmetro	Definição	Valor	Observação	Referência
β	coeficiente de taxa de infecção	$\beta \sim R0 \times \gamma$	A ser inferido a partir de valor de $R0$ dado, γ e outros parâmetros relevantes.	
$R0$	número de reprodução	3.28 (1.4-6.49)		[6–8]
C_{ij}	matriz de contato			
Λ	taxa de natalidade	$3.9 \cdot 10^{-5} \text{ dia}^{-1}$	Estimativa para 2015. A multiplicar pela população total.	IBGE, 2013
μ_{eq}	taxa de mortalidade típica	$1.6 \cdot 10^{-5}$	Estimativa para 2015. A multiplicar pela população total.	IBGE, 2013
γ	coeficiente de taxa de recuperação do infectado	$\sim 1 - e^{-(1/d_I)}$	Uso no modelo SIR, SEIR e SEAIR	
$\gamma_{R,I}$	coeficiente de taxa de recuperação do infectado sintomático	$\sim 1 - e^{-(1/d_I)}$		
$\gamma_{R,A}$	coeficiente de taxa de recuperação do infectado assintomático	$\sim 1 - e^{-(1/d_I)}$		
γ_{R,Q_A}	coeficiente de taxa de recuperação do infectado assintomático quarentenado	$\sim 1 - e^{-(1/d_I)}$	A ser usado no modelo SEAHIR-Q _{AI}	
Continuação...				

Tabela 1: Continuação.

Parâmetro	Definição	Valor	Observação	Referência
γ_{R,Q_I}	coeficiente de taxa de recuperação do infectado sintomático quarentenado	$\sim 1 - e^{-(1/d_I)}$	A ser usado no modelo SEAHIR-Q _{AI}	
$\gamma_{H,R}$	coeficiente de taxa de recuperação do hospitalizado.	$\sim 1 - e^{-(1/d_A)}$	A ser usado no modelo SEAHIR-Q _{AI}	
γ_{H,Q_I}	coeficiente de taxa de hospitalização do infectado quarentenado	$\sim 1 - e^{-(1/d_H)}$	A ser usado no modelo SEAHIR-Q _{AI}	
$\gamma_{Q,A}$	coeficiente de taxa de mudança de assintomático para quarentena	calcula-se de x_A	A ser usado no modelo SEAHIR-Q _{AI}	
$\gamma_{Q,I}$	coeficiente de taxa de mudança de sintomático para quarentena	calcula-se de x_I	A ser usado no modelo SEAHIR-Q _{AI}	
x_I	probabilidade de testar um infectado sintomático	valor prescrito, pode variar no tempo	A ser usado no modelo SEAHIR-Q _{AI}	
x_A	probabilidade de testar um infectado sintomático	valor prescrito, pode variar no tempo	A ser usado no modelo SEAHIR-Q _{AI}	
μ_{cov}	coeficiente de taxa de casualidades	$\mu_{cov} = \gamma T_C / T_R$	No modelo SEAHIR pode ser estimado como $\sim \gamma T_{L,C} / (100 - T_{L,C})$	
a	coeficiente de taxa de conversão de período de incubação	$a \sim 1 - e^{-(1/d_L)}$		
Continuação...				

Tabela 1: Continuação.

Parâmetro	Definição	Valor	Observação	Referência
T_R	porcentagem de recuperação	96,3(96,2-96,4)%		[9]
T_C	porcentagem de casualidades	3,7(3.6-3,80)%		[9]
d_I	tempo médio de infecção	3 a 7 dias		[10]
d_L	tempo médio de incubação	4.6 a 9 dias		[11]
d_C	tempo médio desde infecção até óbito	15,0 (12.5 - 17.8) dias		[12, 13]
d_H	tempo desde início da infecção até hospitalização	6(3-7)		[14, 15]
d_U	tempo desde início da infecção até internação	8(12-15)		[14, 15]
d_A	tempo típico em hospitalização até a alta	8(4-15)	Assumimos igual ao tempo de UTI	[14]
d_{AU}	tempo típico em UTI até a alta	8(4-15)		[14]
d_O	tempo típico até paciente vir a óbito uma vez hospitalizado	14		
ϕ_k	porcentagem de infectados que necessitam de hospitalização	0.15		[15, 16]
η_k	porcentagem de infectados que necessitam de internação em UTI	0.05		[15, 16]
$T_{L,C}$	porcentagem típica de internações que levam a óbito	???		
Continuação...				

Tabela 1: Continuação.

Parâmetro	Definição	Valor	Observação	Referência
α	fator de correção para coeficiente de infecção de subclínico	0.25		[17]
ρ_i	probabilidade de caso ser clínico	0.4 (menores de 55 anos)/ 0.6 (maiores de 55 anos)		[18]
I_0	Número inicial de infectados		A ser ajustado, entre mínimo $I_{0,min}$ = notificados oficialmente em t_0 , a $10 \cdot I_{0,min}$	

Tabela 2: Porcentagem de óbitos entre infectados. Estimativa para T_C , de acordo com estatística dos casos chineses corrigidos para subnotificação e perfil demográfico [9].

Faixa etária	dI(%)
0 a 9	0.0026
10 a 19	0.0148
20 a 29	0.060
30 a 39	0.146
40 a 49	0.295
50 a 59	1.25
60 a 69	3.99
70 a 79	8.61
80 ou mais	13.4

Tabela 3: Porcentagem de hospitalizações entre infectados. Estimativa para ϕ_k , de acordo com estatística dos casos chineses corrigidos para subnotificação e perfil demográfico [9].

Faixa etária	$\phi_k(\%)$
0 a 9	0.
10 a 19	0.04
20 a 29	1.04
30 a 39	3.43
40 a 49	4.25
50 a 59	8.16
60 a 69	11.8
70 a 79	16.6
80 ou mais	18.4

6 Benchmark utilizando o modelo de Prem *et al*

Neste seção, apresentamos o procedimento para fazer o benchmark do modelo aqui apresentado com o modelo de Wuhan. O procedimento é o seguinte:

- O modelo a ser comparado é o SEAHIR
- A matriz de contato inicial é a chinesa, fornecida em material complementar. Ela precisa ser simetrizada e escalada para o padrão populacional de Wuhan. O procedimento é o seguinte:

O primeiro passo é garantir a reciprocidade da matriz de contato chinesa para posterior transformação, aplicando a seguinte fórmula à matriz de contatos chinesa \mathbf{C} :

$$\mathbf{C}_{\text{sym}} = 0.5(\mathbf{C} + \mathbf{C}^T \mathbf{P}) \quad (113)$$

onde \mathbf{P} é uma matriz tal $P_{ii} = 1$, i sendo a i -ésima faixa etária, e $P_{ij} = p_i/p_j$, onde p_i é a população ou fração da população na faixa etária i em Wuhan. O passo seguinte é calcular a matriz \mathbf{C}_{new} de Wuhan fazendo:

$$\mathbf{C}_{\text{new}} = \mathbf{C}_{\text{sym}} \cdot \mathbf{P} \quad (114)$$

O procedimento realizado até agora precisa ser realizado para todas as matrizes (total, casa, escolar, trabalho, e comunidade). Depois, é necessário rescalas a matriz \mathbf{C}_{new} dividindo-a pelo seu autovalor dominante:

$$\mathbf{C}_{\text{esc}} = \frac{\mathbf{C}_{\text{new}}}{(\text{Max}(\text{Re}(\text{Eig}(\mathbf{C}_{\text{new}}))))} \quad (115)$$

onde $\text{Eig}()$ é o autovalor da matriz. Para se obter a matriz β_{ij} , a matriz escalada deve ser multiplicada por $R0$, prescrito dado no artigo e γ , também prescrito. Note que as matrizes parciais serão divididas pelo autovalor dominante da matriz total, e não pelos próprios autovalores dominantes.

O procedimento acima descrito pode ser utilizado também para obter a matriz de contato nas cidades ou regiões brasileiras.

A matriz \mathbf{P} é construída a partir da estrutura etária de Wuhan, dado em material suplementar. Para as condições iniciais de infectados, essa matriz é multiplicada pelo fator $f_0 = 0.0002$ ($I_0 = 200$) e $f_0 = 0.002$ ($I_0 = 2000$) - tal informação encontra-se no código e não é mencionada no artigo. A população inicial de susceptíveis é dada pela distribuição sem nenhum fator de correção. Os outros parâmetros podem ser encontrados na tabela do artigo de Prem *et al* [2] e as outras variáveis de compartimento são nulas.

O benchmark será realizado comparando os resultados obtidos com os resultados de Wuhan sem intervenção, em quatro casos:

- $d_I = 3$, $I_0 = 200$, $n_{\text{sim}} = 200$ fazendo $R0$ variar de 1.6 a 3.0.
- $d_I = 3$, $I_0 = 2000$, $n_{\text{sim}} = 2000$ fazendo $R0$ variar de 1.6 a 3.0.
- $d_I = 7$, $I_0 = 200$, $n_{\text{sim}} = 200$ fazendo $R0$ variar de 1.6 a 3.0.
- $d_I = 7$, $I_0 = 2000$, $n_{\text{sim}} = 200$ fazendo $R0$ variar de 1.6 a 3.0.

O resultado final das duzentas simulações pós processado calculando-se a mediana da amostra de 200 pontos a cada instante. Os gráficos para comparação estão no material suplementar. O artigo faz o gráfico da incidência acumulada, que corresponde ao $M(t)$ aqui definido. Ele também faz o gráfico de novos casos por dia, dado simplesmente por $I(t+1) + A(t+1) - I(t) - A(t)$. É necessário testar se o máximo de incidência e máximo de novas infecções obtidos é consistente em número e instante aos resultados dos gráficos do material suplementar.

Referências

- [1] K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, N. Davies, M. Jit, P. Klepac *et al.*, “The effect of control strategies that reduce social mixing on outcomes of the covid-19 epidemic in wuhan, china,” *Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working and Jit, Mark and Klepac, Petra, The Effect of Control Strategies that Reduce Social Mixing on Outcomes of the COVID-19 Epidemic in Wuhan, China (3/9/2020)*, 2020.
- [2] K. Prem, A. R. Cook, and M. Jit, “Projecting social contact matrices in 152 countries using contact surveys and demographic data,” *PLoS computational biology*, vol. 13, no. 9, p. e1005697, 2017.
- [3] F. C. Coelho, R. M. Lana, O. G. Cruz, C. T. Codeco, D. Villela, L. S. Bastos, A. P. y Piontti, J. T. Davis, A. Vespignani, and M. F. Gomes, “Assessing the potential impacts of covid-19 in brasil: Mobility, morbidity and impact to the health system,” *medRxiv*, 2020.
- [4] OAG, “Official Airline Guide,” <http://www.oag.com>, 2020, [Online; acessado em 31 de março de 2020].
- [5] IBGE – Instituto de Geografia e Estatística, “Caso demográfico 2010: Resultados gerais da amostra,” <https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html>, 2010, [Online; acessado em 31 de março de 2020].
- [6] Y. Liu, A. A. Gayle, A. Wilder-Smith, and J. Rocklöv, “The reproductive number of covid-19 is higher compared to sars coronavirus,” *Journal of travel medicine*, 2020.
- [7] Z. Guopeng and C. Chunhua, “A model simulation study on effects of intervention measures in wuhan covid-19 epidemic,” *medRxiv*, 2020.
- [8] A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R. M. Eggo, F. Sun, M. Jit, J. D. Munday *et al.*, “Early dynamics of trans-

mission and control of covid-19: a mathematical modelling study,” *The Lancet Infectious Diseases*, 2020.

- [9] R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. Walker, H. Fu *et al.*, “Estimates of the severity of covid-19 disease,” *medRxiv*, 2020.
- [10] R. Woelfel, V. M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M. A. Mueller, D. Niemeyer, P. Vollmar, C. Rothe, M. Hoelscher *et al.*, “Clinical presentation and virological assessment of hospitalized cases of coronavirus disease 2019 in a travel-associated transmission cluster,” *medRxiv*, 2020.
- [11] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler, “The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application,” *Annals of internal medicine*, 2020.
- [12] P. Sun, X. Lu, C. Xu, W. Sun, and B. Pan, “Understanding of covid-19 based on current evidence,” *Journal of medical virology*, 2020.
- [13] N. M. Linton, T. Kobayashi, Y. Yang, K. Hayashi, A. R. Akhmetzhanov, S.-m. Jung, B. Yuan, R. Kinoshita, and H. Nishiura, “Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data,” *Journal of clinical medicine*, vol. 9, no. 2, p. 538, 2020.
- [14] F. Zhou, T. Yu, R. Du, G. Fan, Y. Liu, Z. Liu, J. Xiang, Y. Wang, B. Song, X. Gu *et al.*, “Clinical course and risk factors for mortality of adult inpatients with covid-19 in wuhan, china: a retrospective cohort study,” *The Lancet*, 2020.
- [15] W.-j. Guan, Z.-y. Ni, Y. Hu, W.-h. Liang, C.-q. Ou, J.-x. He, L. Liu, H. Shan, C.-l. Lei, D. S. Hui *et al.*, “Clinical characteristics of coronavirus disease 2019 in china,” *New England Journal of Medicine*, 2020.
- [16] D. Baud, X. Qi, K. Nielsen-Saines, D. Musso, L. Pomar, and G. Favre, “Real estimates of mortality following covid-19 infection,” *The Lancet Infectious Diseases*, 2020.

- [17] Y. Liu, S. Funk, and S. Flasche, “The contribution of pre-symptomatic transmission to the covid-19 outbreak,” 2020.
- [18] Q. Bi, Y. Wu, S. Mei, C. Ye, X. Zou, Z. Zhang, X. Liu, L. Wei, S. A. Truelove, T. Zhang *et al.*, “Epidemiology and transmission of covid-19 in shenzhen china: Analysis of 391 cases and 1,286 of their close contacts,” *medRxiv*, 2020.