

# Modelagem matemática para epidemia de COVID-19 no Brasil

Gabriela Moutinho de Souza Dias<sup>\*,a</sup> e Luiz Henrique da Costa Araújo<sup>b</sup>

<sup>a</sup> Instituto Militar de Engenharia (IME) – Praça Gen. Tibúrcio, 80 – Urca, Rio de Janeiro – RJ, 22290-270

\* gabriela@ime.eb.br

**RESUMO:** Este artigo apresenta uma modelagem matemática para a epidemia de COVID-19 na população brasileira e um método para a obtenção de estimativas dos parâmetros do modelo. Para tal, foi utilizado como base o modelo matemático epidemiológico SIRD que permite acompanhar o desenvolvimento da doença e fazer projeções da evolução da epidemia. Resultados iniciais validados, com projeções de curto prazo, são apresentados. O presente estudo encontra-se em andamento, em constante atualização e aprimoramento, uma vez que foi desenvolvido durante o curso da epidemia.

**PALAVRAS-CHAVE:** covid-19, epidemia, pandemia, modelo matemático, SIRD.

**ABSTRACT:** This article presents a mathematical model for the covid-19 epidemic in the Brazilian population and a method to obtain estimates of the model parameters. To this end, the epidemiological mathematical model SIRD was used as a basis, which allows to monitor the evolution of the disease and to produce its projection. Initial results, with short-term projections, are presented. The current study is in progress, being constantly updated and improved once it was developed during the course of the epidemic.

**KEYWORDS:** covid-19, epidemic, pandemic, mathematical model, SIRD.

## 1. Introdução

No final do ano de 2019, o mundo acendeu um sinal de alerta com o início de uma nova doença infecciosa viral chamada COVID-19 (*Coronavirus Disease 2019*), causada por um vírus chamado coronavírus SARS-CoV-2 (*Severe Acute Respiratory Syndrome Coronavirus 2 of the Genus Betacoronavirus*) [1].

Conforme a epidemia foi se desenvolvendo no mundo, as autoridades médicas mundiais perceberam a rápida disseminação da doença e o seu potencial perigo para o colapso dos Sistemas de Saúde dos países afetados. A comunidade científica, então, se mobilizou em prol de buscar soluções para conter o avanço da doença. Os pesquisadores da área de modelagem e simulação, ao redor do mundo, têm envidado

esforços para buscar modelar a evolução da epidemia [2-5].

A modelagem matemática epidemiológica é bastante eficiente no auxílio do processo de tomada de decisões. A aplicação de modelos analíticos na análise do desenvolvimento de epidemias pode representar importante ferramenta para decisores, permitindo a análise equilibrada e embasada do impacto da epidemia na saúde, na economia, além de entre outras áreas da sociedade. Por essa razão, faz-se mister o desenvolvimento desses modelos objetivando fornecer embasamento para decisões acertadas.

Desde o surgimento do surto de COVID-19, diversos trabalhos científicos foram publicados com o objetivo de analisar o comportamento da epidemia na população [2-5].

Existem basicamente duas formas de anali-

sar esse tipo de fenômeno, por meio de modelos analíticos ou por meio de modelos de simulação computacional. O presente trabalho está inserido na classe de modelos analíticos.

Apesar de haver diversos estudos já divulgados versando sobre a evolução da epidemia e a comparação entre possíveis cenários, cada estudo é direcionado para um país específico. Tendo em vista que cada país tem suas peculiaridades e particularidades locais, é importante que haja um estudo especificamente focado na população brasileira.

Alguns trabalhos direcionados para o cenário específico da população do Brasil já foram divulgados [2,3]. O presente trabalho apresenta contribuições em relação aos anteriores, tendo em vista que os dados utilizados para obtenção dos resultados desta modelagem foram atualizados até a data de 10 de abril de 2020. Os primeiros resultados do presente estudo foram apresentados em [6]. Além disso, o método de estimação dos parâmetros do modelo adequa-se à característica dos dados reais disponibilizados até a escrita deste texto considerando o total de casos acumulado, como explicado na Seção 3 adiante. A constante atualização dos parâmetros e revisão da metodologia para estimá-los são de extrema importância para se obter resultados com maior acurácia e mais próximos da real situação da epidemia.

O restante deste artigo está organizado da seguinte forma. Na Seção 2 é apresentado o modelo SIDR (*Susceptible-Infected-Deceased-Removed*) aplicado nesta modelagem e uma análise teórica de dois cenários fictícios. A Seção 3 apresenta a metodologia utilizada para estimação dos parâmetros e a Seção 4, os resultados obtidos com este estudo. As principais conclusões do trabalho até a data de sua atualização do mesmo são apresentadas na Seção 5.

## 2. Modelos epidemiológicos

A modelagem matemática da propagação de doenças infecciosas começou por volta de 1760, com o trabalho de Daniel Bernoulli sobre a varíola. Em 1906, William Hamer foi o primeiro a prever o princípio da Lei de Ação das Massas para um modelo determinístico em tempo discreto [7]. Esse princípio, que incorpora o princípio da mistura homogênea, tem sido a base dos trabalhos subsequentes na área de epidemiologia matemática.

Desde então, diversos modelos matemáticos foram propostos para os mais diversos tipos de doenças infecciosas que afetam a população humana, seja em nível comunitário (epidemia local) ou mundial (pandemia), como a que o mundo enfrenta atualmente com a COVID-19.

O objetivo da modelagem matemática epidemiológica é analisar o comportamento e a evolução de uma doença dentre indivíduos de uma população ao longo do tempo. Essa análise tem o propósito de auxiliar no controle da propagação da doença, de modo a evitar o avanço de epidemias.

A literatura dessa área é bastante rica, principalmente por ser um assunto antigo e amplamente estudado. Existe uma grande quantidade de referências de trabalhos sobre modelos matemáticos para doenças infecciosas [8-16].

Dentre os diversos modelos existentes na literatura, para a modelagem descrita neste documento foi escolhido o modelo SIDR. Os principais motivos para tal escolha são a baixa complexidade e a eficiência do modelo. Em termos matemáticos, o modelo é relativamente simples, se comparado a outros modelos da área, pois trata-se de um sistema de equações diferenciais. Em termos

de eficiência, esse modelo já foi utilizado para modelar diversas doenças virais infecciosas no passado [17], mostrando-se altamente eficiente na estimação do andamento da epidemia na população, quando suas condições de aplicabilidade são satisfeitas.

O grande desafio em se trabalhar com modelos epidemiológicos é a estimação dos parâmetros. Essa etapa é de crucial importância, pois pequenas variações nos parâmetros são capazes de causar um grande impacto nos resultados finais.

É importante ressaltar que, independente do método de estimação utilizado, um bom resultado só pode ser obtido se houver uma quantidade adequada de dados acerca da população e do comportamento da doença nos indivíduos em particular e na população como um todo.

## 2.1 Modelo SIDR

Esta seção apresenta o modelo SIDR utilizado no presente trabalho, com destaque para as principais características do mesmo.

Em 1927, Kermack e McKendrick desenvolveram o clássico modelo SIR (*Susceptible-Infected-Removed*) para a propagação de doenças infecciosas [18,19]. Trata-se de um modelo determinístico de tempo contínuo, amplamente utilizado tanto na área epidemiológica quanto em outras áreas da ciência [20,21]. A partir do modelo SIR, diversas extensões foram propostas ao longo do tempo. Em termos gerais, esse modelo divide a população em três compartimentos: Suscetível, Infectado e Removido. O número de óbitos devido à doença está inserido no compartimento Removido. Porém, quando se torna relevante observar separadamente a evolução dos casos de óbitos, torna-se necessário criar um quarto compartimento, específico para tais casos.

A literatura, em geral, se refere a esse compartimento como *Deceased*, e neste trabalho o modelo que inclui tal compartimento é chamado de SIDR (*Susceptible-Infected-Deceased-Removed*).<sup>1</sup>

Seja  $N$  o número total de indivíduos na população. Considera-se que a população é distribuída de forma homogênea na região. Além disso, o modelo também considera uma população homogênea, isto é, todos os indivíduos possuem mesma probabilidade de serem infectados e as taxas de contato de cada indivíduo são iguais. A consideração de homogeneidade é válida para grandes populações, no âmbito da análise da modelagem.

Considere que o período de tempo em que uma epidemia ocorre seja muito menor que o tempo de existência da população, ou seja, o tamanho dessa pode ser considerado constante no período, e os nascimentos e mortes não relacionadas à doença podem ser ignorados. A qualquer instante de tempo, cada indivíduo encontra-se em um determinado estado com relação à doença, dividindo a população em quatro compartimentos:

- S - Suscetíveis: Indivíduos que estão suscetíveis a serem infectados;
- I - Infectados: Indivíduos que estão infectados com a doença e são capazes de transmiti-la a indivíduos suscetíveis;
- D - Óbitos: Indivíduos que morreram devido à doença<sup>2</sup>; e
- R - Removidos: Indivíduos que se recuperaram da doença ou foram isolados, não sendo capazes de transmitir a doença a outros indivíduos.

A transmissão da doença ocorre por contato

<sup>1</sup> Alguns autores se referem a este modelo como SIRD ou SIR com mortalidade induzida.

<sup>2</sup> Utiliza-se a letra D para este compartimento para acompanhar a notação da literatura, do inglês *Deceased*.

entre os indivíduos<sup>3</sup>. Um indivíduo suscetível se torna infectado após contato com um indivíduo infectado. O termo contato deve ser interpretado de forma ampla, pois refere-se ao evento responsável pela transmissão do vírus, podendo ser um contato físico propriamente dito (como aperto de mãos ou a proximidade física no momento de espirro ou tosse do infectado) ou outro tipo de situação de transmissão, como quando um suscetível tem contato com um local contaminado pelo vírus, por exemplo. O compartimento R engloba os indivíduos que já foram infectados mas não podem mais transmitir a doença. Considera-se que um indivíduo que se recupera da doença torna-se imune durante o período de duração da epidemia, não retornando ao compartimento S, mas sim sendo contabilizado no compartimento R. Indivíduos infectados que são isolados do contato com indivíduos suscetíveis (não sendo assim capazes de transmitir a doença) também são inseridos no compartimento R.

Sejam  $S(t)$ ,  $I(t)$ ,  $D(t)$  e  $R(t)$  o número de Suscetíveis, Infectados, Óbitos e Removidos, respectivamente, a cada instante de tempo  $t \geq 0$ . Assim, para todo  $t \geq 0$ , temos que:

$$S(t) + I(t) + D(t) + R(t) = N \quad (1)$$

A evolução do processo infeccioso de acordo com o modelo SIRD é definida pelo seguinte sistema de equações diferenciais [22]:

$$\frac{dS(t)}{dt} = -\beta S(t)I(t) \quad (2)$$

$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \frac{\gamma}{1-\rho} I(t) \quad (3)$$

$$\frac{dD(t)}{dt} = \frac{\rho}{1-\rho} \gamma I(t) \quad (4)$$

<sup>3</sup>Na literatura, há outros modelos para tipos de doença não transmissíveis por contato direto entre indivíduos, como a dengue, por exemplo.

$$\frac{dR(t)}{dt} = \gamma I(t) \quad (5)$$

Condição inicial:  $(S(0), I(0), D(0), R(0)) = (S_0, I_0, 0, 0)$ ,  $S_0, I_0 > 0$ .  $\beta, \gamma, > 0$ .

O parâmetro  $\beta$  é o **parâmetro de infecção** do processo. Sendo a doença transmitida por contato,  $\beta$  corresponde à taxa com que ocorrem contatos entre dois indivíduos quaisquer, isto é, o número de contatos entre dois indivíduos quaisquer por unidade de tempo. Muitos autores também se referem a  $\beta$  como “taxa de contato”, porém esse termo pode se confundir com a taxa de contato de cada indivíduo, isto é, número de contatos que um indivíduo faz com outros indivíduos na população por unidade de tempo. A **taxa de contato** de cada indivíduo é comumente representada na literatura como  $\alpha$ . A relação entre  $\alpha$  e  $\beta$  é dada por  $\alpha = (N - 1)\beta$ , podendo ser aproximada por  $N\beta$  quando  $N$  é grande.

O parâmetro  $\gamma$  é chamado **taxa de recuperação**, e representa a taxa com que um indivíduo infectado é recuperado (taxa de recuperação por indivíduo infectado).

O parâmetro  $\rho$  é chamado **probabilidade de mortalidade**, isto é, a probabilidade de um indivíduo infectado morrer devido à doença antes de se recuperar.

Na análise da evolução de uma epidemia, há duas importantes métricas de interesse: o **número de reprodutibilidade basal**  $R_0$  (*basic reproduction number*, em inglês<sup>4</sup>) e a **taxa de infecção** (*infection rate*, em inglês).

O número  $R_0$  representa o número médio (valor esperado) de casos secundários provenientes por um indivíduo infectado introduzido em uma

<sup>4</sup> Outras denominações: *basic reproduction ratio*, *basic reproductive number*, *basic reproductive ratio*.

população de indivíduos suscetíveis [23]. Esse conceito é válido para qualquer modelo epidemiológico, porém sua definição matemática depende do modelo específico e de seus parâmetros. No caso do modelo SIDR,  $R_0$  é dado pela razão  $\alpha(1 - \gamma)/\gamma$  [2].

A infecção dos indivíduos suscetíveis, levando à transição para o compartimento Infectados, é chamada na literatura de incidência horizontal (incidência vertical é a variação devido a nascimentos e/ou mortes não relacionados à doença). A incidência horizontal é a taxa de infecção de indivíduos suscetíveis por meio do contato desses com indivíduos infectados [14, 15]. A taxa de infecção varia ao longo do tempo e é dada por  $\beta S(t)I(t)$ .

## 2.2 Análise de longo prazo em dois cenários fictícios

A seguir, são apresentados resultados da modelagem para as variáveis  $S(t)$ ,  $I(t)$ ,  $D(t)$  e  $R(t)$ , com valores fictícios para os parâmetros  $\beta$ ,  $\gamma$ ,  $\rho$ . O objetivo desta apresentação é explicar como o modelo prevê a evolução de uma epidemia em uma população, e como é o comportamento das curvas das variáveis envolvidas.

Considere uma população com  $N = 100$  mil indivíduos, e que os parâmetros do modelo são dados por (em  $\text{dia}^{-1}$ ):  $\beta = 0,8$ ;  $\gamma = 0,183$ ;  $\rho = 0,2$ . O número  $R_0$  para esses parâmetros é de aproximadamente 3,5. Considere que em  $t = 0$ , a condição inicial é dada por  $(S(0), I(0), D(0), R(0)) = (N - 1, 1, 0, 0)$ . Vamos chamar esse cenário de C1.

A **figura 1** apresenta a evolução das variáveis  $S(t)$ ,  $I(t)$ ,  $D(t)$  e  $R(t)$  para o cenário C1. Podemos observar que o número de infectados aumenta até atingir um pico e depois decai. O processo converge para um período de estabilidade após um certo tempo.

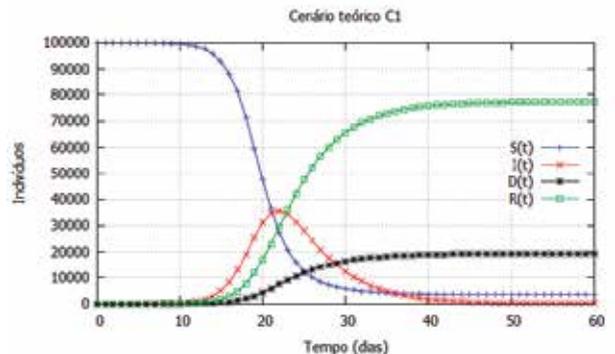


Fig. 1 –  $S(t)$ ,  $I(t)$ ,  $D(t)$  e  $R(t)$  para o cenário C1.

A **figura 2** apresenta a evolução das variáveis  $S(t)$ ,  $I(t)$  e o número de infectados acumulado ao longo do tempo. Essa variável é obtida pela diferença  $N - S(t)$  ou pela soma  $I(t) + D(t) + R(t)$ , ou seja, representa todos os indivíduos que foram acometidos pela doença, distribuídos entre os compartimentos I, D e R. Cabe observar a diferença em se analisar  $I(t)$  e o número de Infectados Acumulado. Nesse cenário fictício quase toda a população foi acometida pela doença.

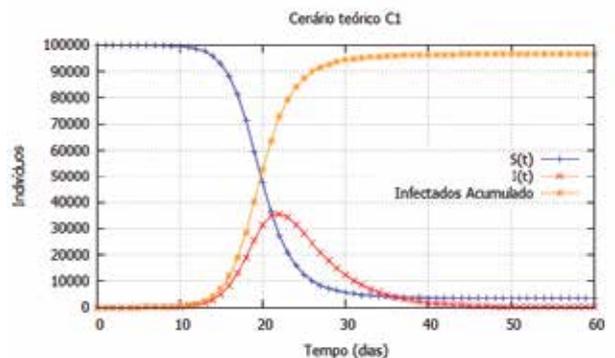
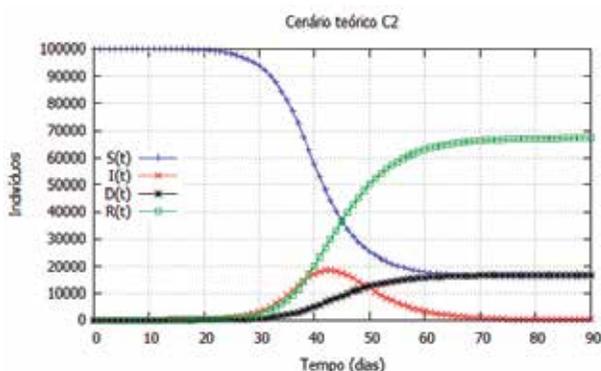


Fig. 2 –  $S(t)$ ,  $I(t)$  e Infectados Acumulado para o cenário C1.

Vamos observar a seguir o que acontece quando o parâmetro de infecção é menor do que no cenário anterior, mantendo-se as demais condições e parâmetros. Sendo assim, vamos supor

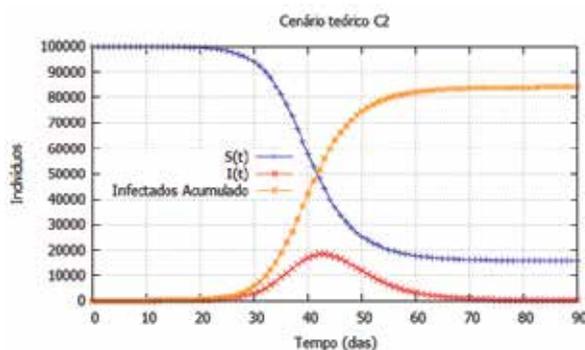
agora que os parâmetros do modelo são dados por (em  $\text{dia}^{-1}$ ):  $\beta = 0,5$ ;  $\gamma = 0,183$ ;  $\delta = 0,2$ . O número  $R_0$  para esses parâmetros é de aproximadamente 2,18. Considere que em  $t = 0$ , a condição inicial é dada por  $(S(0), I(0), D(0), R(0)) = (N - 1, 1, 0, 0)$ , igual a C1. Vamos chamar esse cenário de C2.

A **figura 3** apresenta a evolução das variáveis  $S(t)$ ,  $I(t)$ ,  $D(t)$  e  $R(t)$  para o cenário C2. Podemos observar que, apesar de ambos os cenários apresentarem o mesmo comportamento em termos de evolução das curvas, o pico no número de infectados é menor do que em C1. Além disso, tal pico leva mais tempo para ocorrer (em C2 ocorre após 40 dias, enquanto que em C1 ocorre próximo a 20 dias a contar do início do processo).



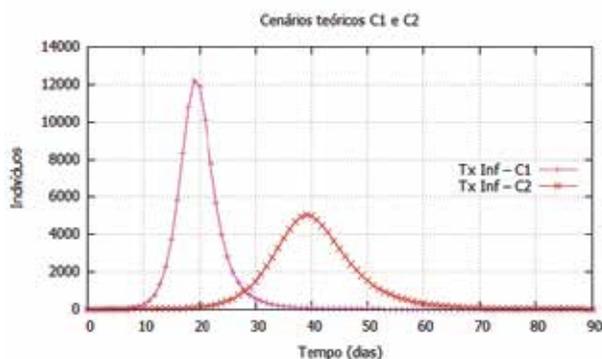
**Fig. 3** –  $S(t)$ ,  $I(t)$ ,  $D(t)$  e  $R(t)$  para o cenário C2.

A **figura 4** apresenta a evolução das variáveis  $S(t)$ ,  $I(t)$ , e o número de Infectados Acumulados ao longo do tempo para o cenário C2. Em comparação ao cenário anterior, pode-se observar que uma parcela menor da população é acometida pela doença.



**Fig. 4** –  $S(t)$ ,  $I(t)$  e Infectados Acumulados para o cenário C2.

Para finalizar essa apresentação teórica, a **figura 5** apresenta a taxa de infecção ao longo do tempo para os cenários C1 e C2.



**Fig. 5** – Taxa de infecção para os cenários C1 e C2.

Pode-se observar na comparação entre os cenários C1 e C2 a influência de medidas de controle de propagação da doença na população. Tais medidas são capazes de reduzir o parâmetro de infecção  $\beta$ . Quanto maior o parâmetro de infecção, maior a taxa com que indivíduos suscetíveis tornam-se infectados, como no cenário C1. O cenário C2 representa uma situação em que ocorre a adoção de medidas de controle. A medida de isolamento social, por exemplo, é uma medida de controle pois reduz o número de contatos entre os indivíduos de forma geral, e, con-

sequentemente, reduz o número de contatos entre suscetíveis e infectados. Por isso, medidas de controle de propagação são capazes de atrasar a evolução da epidemia, além de reduzir o número de infectados total.

É importante ressaltar que a comparação apresentada é teórica e objetiva, com base simplesmente no resultado matemático que o modelo apresenta. Este estudo não tem o objetivo de analisar impactos econômicos ou sociais de qualquer medida adotada pelas autoridades governamentais, e tão pouco reflete a opinião pessoal dos autores acerca da adoção ou não de qualquer medida. Cabe destacar que este tipo de comparação entre possíveis cenários, considerando a aplicação ou não de medidas de controle, é realizado na grande maioria dos trabalhos que tratam da análise de propagação de doenças infecciosas publicados na literatura há mais de cem anos, tanto em artigos científicos como em livros [9, 12].

### 3. Metodologia para estimação dos parâmetros

Como mencionado anteriormente, o desafio da modelagem epidemiológica é a estimação dos parâmetros. Nesta seção é apresentada a metodologia de estimação utilizada neste trabalho.

Os parâmetros do modelo que precisam ser estimados são  $\beta$ ,  $\gamma$ ,  $\rho$ . Um dos métodos de estimação mais utilizados na literatura é a minimização do erro médio quadrático para as variáveis  $I(t)$  e  $D(t)$ , como proposto em [9, 12], por exemplo.

Entretanto, os dados reais oficialmente divulgados pelo Ministério da Saúde do Brasil<sup>5</sup> referem-se a: número de casos (infectados) diário (casos novos por dia), número de casos (infectados)

<sup>5</sup>Disponível em <https://covid.saude.gov.br/>. Acesso em 10/04/2020.

acumulado ao longo do tempo, número de óbitos por dia e número de óbitos acumulado ao longo do tempo. Sendo assim, não há dados oficiais para se inferir com acurácia o número de indivíduos infectados no dia  $t$ . Quando se analisa um período curto da epidemia, como na obtenção dos resultados preliminares apresentados em [6], pode-se considerar que o número de infectados é igual ao número acumulado de infectados ao longo do tempo, retirando apenas os óbitos e considerando que nenhum infectado foi removido. Porém, quando analisamos um período maior do que o tempo médio previsto para um infectado sair desse estado (por meio de recuperação ou morte ou isolamento, por exemplo), esses dois valores não mais correspondem.

Sendo assim, devido a essa restrição na obtenção de dados reais para comparação e ajuste, neste trabalho o método utilizado para estimar os parâmetros  $\beta$ ,  $\gamma$ ,  $\rho$  do modelo, é a minimização do erro médio quadrático para as variáveis  $CI(t)$  e  $D(t)$ , onde  $CI(t)$  representa o número acumulado de infectados no período de tempo de zero a  $t$ . O objetivo é minimizar a seguinte expressão para o erro  $e_n$ :

$$e_n = \frac{1}{2} \left( \sum_{t=t_0}^{t_0+n} (CI_t - \hat{C}I_t)^2 + \sum_t (D_t - \hat{D}_t)^2 \right) \quad (6)$$

onde  $\hat{C}I_t$ ,  $\hat{D}_t$  são os valores estimados para os valores reais  $CI_t$ ,  $D_t$ , respectivamente, no tempo  $t$ . O erro  $e_n$  é obtido considerando um conjunto de dados de tamanho  $n + 1$ , ou seja, correspondendo a  $n + 1$  valores das variáveis, a partir do tempo  $t_0$ .

É importante ressaltar que o valor do erro  $e_n$  e o resultado do processo de minimização dependem fortemente da quantidade e da qualidade dos dados reais disponíveis. A estimação dos

parâmetros deve ser atualizada constantemente, e os valores obtidos dependem dos dados reais que serão usados na comparação com o resultado da modelagem. Quanto mais fidedignos e mais próximos à realidade da situação da epidemia forem os dados de entrada, melhor e mais ajustado será o resultado da modelagem em termos de projeções futuras. Cabe ressaltar que um bom ajuste do modelo aos dados reais não garante uma boa projeção da evolução futura. Um bom ajuste está relacionado ao processo de validação da modelagem. Porém, para que a modelagem, mesmo que validada, seja capaz de fornecer uma projeção futura que seja próxima à realidade, os dados de entrada (contabilização de casos reais) devem ser os mais fiéis possíveis. No curso deste estudo, o chamado “isolamento social” foi uma medida de controle exógena adotada com o objetivo de reduzir a propagação da doença na população. Um dos objetivos deste trabalho é fornecer projeções de curto prazo para a evolução da epidemia. Sendo assim, para a estimação dos parâmetros, os dados reais considerados para ajuste referem-se ao período a partir de 19/03/2020, quando pode-se afirmar com mais certidão que tal medida estava vigente.

Medidas de controle de propagação de uma doença podem afetar os parâmetros do modelo, em especial o parâmetro de infecção, como comentado na seção anterior. Com o isolamento social, o número de contatos realizados por cada indivíduo diminui, e em consequência também diminui o contato entre suscetíveis e infectados. Logo, a velocidade de propagação da doença na população diminui e isso se reflete na redução do valor do parâmetro  $\beta$  em comparação a uma situação onde tal medida não é adotada. Cabe destacar que, neste estudo, os parâmetros foram

estimados considerando a adoção de tal medida vigente e os valores dos mesmos englobam tal situação.

Com relação à condição inicial, o processo de propagação inicia com 1 indivíduo infectado e  $N - 1$  suscetíveis, ou seja,  $(S(0), I(0), D(0), R(0)) = (N - 1, 1, 0, 0)$ . O tempo zero corresponde ao dia 26/02/2020, quando ocorre a primeira notificação oficial de um infectado no Brasil, segundo informação do Ministério de Saúde do Brasil<sup>5</sup>.

A unidade de tempo considerada é de um dia, tendo em vista que os dados reais divulgados são diários.

Os resultados apresentados neste trabalho foram obtidos com base nos dados reais divulgados até a data de 10/04/2020.

O valor para o tamanho da população do Brasil foi obtido na página do Instituto Brasileiro de Geografia e Estatística (IBGE) na Internet<sup>6</sup>.

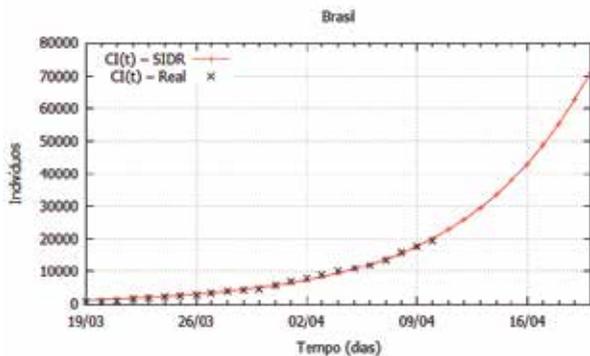
## 4. Resultados

A análise de modelagem de que trata este documento é um trabalho em andamento, em constante atualização e refinamento. Essa análise dinâmica é necessária pois, além da importância em atualizar os parâmetros conforme mais dados são divulgados, a situação da epidemia também muda ao longo do tempo. Novas medidas podem ser adotadas pela população, ou novos fatores podem vir a ser considerados. Nesta seção são apresentados resultados da modelagem obtidos até a data de 10/04/2020 para a população do Brasil, conforme a metodologia explicada na seção anterior.

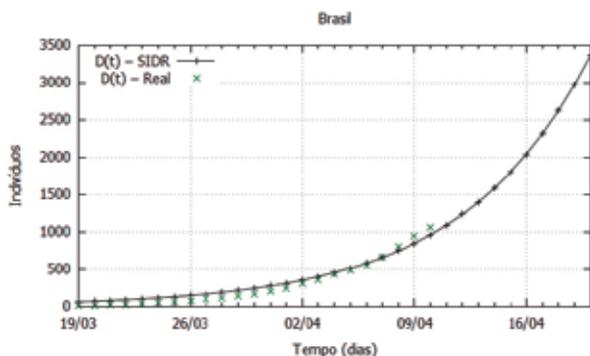
As **figuras 6 e 7** apresentam o número acumulado de indivíduos infectados e o número de óbitos, ao longo do tempo, no intervalo entre os

<sup>6</sup> Disponível em <https://www.ibge.gov.br/apps/populacao/projecao/>. Acesso em 02/04/2020.

dias 19/03/2020 e 20/04/2020, respectivamente. Pode-se observar nas figuras o bom ajuste entre os resultados da modelagem e os dados reais do histórico da epidemia, validando, assim, o modelo implementado. Nesses gráficos são indicadas também projeções de curto prazo, mostrando a tendência de crescimento dos casos. O valor do número de reprodutibilidade basal  $R_0$  no período de análise é de 1,21, aproximadamente.



**Fig. 6** – Ajuste do modelo aos dados reais. Número acumulado de indivíduos infectados no período de 19/03/2020 até 20/04/2020, com estimativa do modelo no período de 11/04/2020 a 20/04/2020 e dados reais até a data de 10/04/2020.

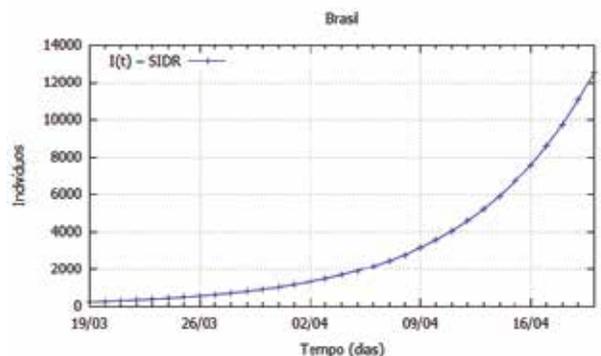


**Fig. 7** – Ajuste do modelo aos dados reais. Número de óbitos no período de 19/03/2020 até 20/04/2020, com estimativa do modelo no período de 11/04/2020 a 20/04/2020 e dados reais até a data de 10/04/2020.

Neste ponto da análise cabe destacar um alerta. O resultado da modelagem para a estimativa da evolução da doença quanto ao número de óbitos deve ser analisado com cautela. A ocorrência do óbito pode variar em um curto intervalo de tempo, por diversos motivos cuja discussão está fora do escopo deste estudo. Para exemplificar, alguns fatores que podem influenciar na mortalidade da doença no indivíduo infectado são as condições do tratamento disponibilizado, condições de saúde particulares, capacidade de atendimento hospitalar, tempo de diagnóstico, dentre outros.

A **figura 8** apresenta a estimativa do modelo para o número de indivíduos infectados no período de 19/03/2020 a 20/04/2020.

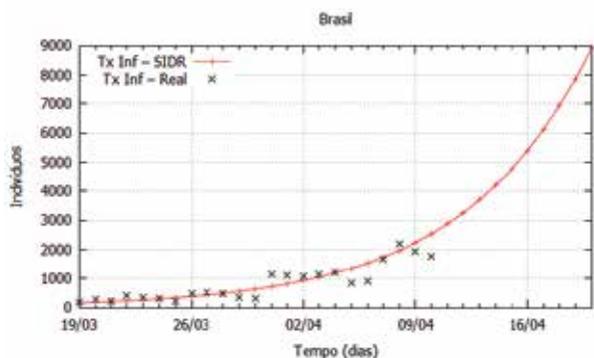
É importante lembrar que não há dados oficiais para comparação com essa métrica. No entanto, o gráfico é útil para se analisar como é o crescimento do número de indivíduos com a infecção ativa, ou seja, o número de indivíduos capazes de transmitir a doença



**Fig. 8** – Estimativa do modelo para o número de indivíduos infectados, no período de 19/03/2020 a 20/03/2020.

A **figura 9** apresenta a taxa de infecção no período de 19/03/2020 até 20/04/2020, com es-

timativa do modelo no período de 11/04/2020 a 20/04/2020 e dados reais até a data de 10/04/2020. Esse gráfico indica a curva de crescimento do número de novos casos da doença no curto prazo.



**Fig. 9** – Total de indivíduos infectados de 19/03/2020 até 20/04/2020, com estimativa do modelo no período de 11/04/2020 a 20/04/2020 e dados reais até a data de 10/04/2020

## 5. Conclusão

O presente artigo apresentou uma modelagem matemática para a epidemia de COVID-19 na população brasileira e um método para a obtenção de estimativas dos parâmetros do modelo. Foi utilizado como base o modelo matemático epidemiológico SIRD.

A partir de dados oficiais coletados do sistema de atendimento ao COVID19, é possível analisar o desenvolvimento da doença. Os resultados se referem a um cenário específico, população Brasileira, com as medidas de mitigação de propagação da doença não farmacológicas adotadas até o momento de escrita deste documento. Os dados reais utilizados para estimação dos parâmetros da modelagem são atualizados até a data de 10/04/2020.

A qualidade da previsão do desenvolvimento da epidemia está relacionada com a qualidade dos índices de notificação da doença pelo Sistema de Saúde, isto é, quanto mais eficiente for a testagem da

doença em casos suspeitos e a notificação para o Sistema de Saúde, melhor serão os resultados da modelagem, uma vez que a curva de desenvolvimento da epidemia ajusta-se aos números informados no processo de modelagem. Os parâmetros do modelo devem ser estimados com certa periodicidade, acompanhando os dados reais divulgados, de forma a atualizar os resultados da modelagem.

Na Seção 4, foram apresentados os resultados preliminares do presente estudo, com projeções de curto prazo para a evolução da epidemia no Brasil. Outros resultados de médio e longo prazos, além da comparação de cenários considerando diferentes medidas de controle de propagação, também foram obtidos e encontram-se em fase de consolidação e análise.

No processo de análise dos resultados, cabe ressaltar um alerta, já mencionado na Seção 4. A estimativa para o número de óbitos deve ser analisada com cautela, pois pode variar com as condições do tratamento disponibilizado ao paciente e a capacidade hospitalar do país, dentre outros fatores.

É importante ressaltar que os resultados apresentados neste documento são teóricos, obtidos de forma técnica e objetiva, com base simplesmente no resultado matemático que a modelagem fornece. Este estudo não tem o objetivo de analisar impactos econômicos ou sociais de qualquer medida adotada pelas autoridades governamentais, e também não reflete a opinião pessoal dos autores acerca da adoção ou não adoção de qualquer medida.

## Agradecimentos

Agradecemos a Anderson Fernandes Pereira dos Santos, Chefe da Seção de Engenharia de Computação do Instituto Militar de Engenharia, pelo apoio e incentivo na realização do presente trabalho.

## Referências bibliográficas

- [1] F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E. C. Holmes, Y.-Z. Zhang, A new coronavirus associated with human respiratory disease in china, *Nature*, vol. 579, 2020, pgs 265–269.
- [2] S. B. Bastos, D. O. Cajueiro, Modeling and forecasting the Covid-19 pandemic in Brazil, Universidade de Brasília, 2020.
- [3] R. M. Cotta, C. P. Naveira-Cotta, p. magal, Modelling the covid-19 epidemics in Brazil: Parametric identification and public health measures influence, doi:10. 1101/2020.03.31.20049130. URL <https://www.medrxiv.org/content/early/2020/04/03/2020.03.31.20049130>
- [4] K. Roosaa, Y. Leea, R. Luoa, A. Kirpicha R. Rothenberga, J. M. Hymanb P. Yanc, G. Chowell, Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020, *Infectious Disease Modelling*, vol. 5, 2020, pgs 256-263.
- [5] D. Fanellia, F. Piazza, Analysis and forecast of COVID-19 spreading in China, Italy and France, *Chaos, Solitons & Fractals*, vol. 134, 2020.
- [6] G. M. S. Dias, L. H. C. Araújo, Modelagem Matemática para Epidemia de COVID-19 e Estimativa de Casos no Brasil, Instituto Militar de Engenharia. Grupo de Assessoramento Científico, 2020. URL <http://www.comp.ime.eb.br/covid19/artigo1.pdf>
- [7] W. Hamer, The Milroy lectures on epidemic disease in England: the evidence of variability and of persistency of type, Bedford Press, 1906.
- [8] L. J. S. Allen, A. M. Burgin, Comparison of deterministic and stochastic sis and sir models in discrete time, *Mathematical Biosciences*, vol. 163, nr. 1, 2000, pgs 1–33. URL <http://www.ncbi.nlm.nih.gov/pubmed/10652843>
- [9] L. J. S. Allen, An Introduction to Stochastic Processes with Applications to Biology, Pearson Prentice Hall, 2003. URL <http://www.hit.ac.il/staff/leonidM/stochastic-processes/Chapter-2.pdf>
- [10] N. T. J. Bailey, A simple stochastic epidemic, *Biometrika*, vol. 37, nr. 3-4, 1950, pgs 193–202.
- [11] N. T. J. Bailey, The total size of a general stochastic epidemic, *Biometrika*, vol. 40, nr. 1-2, 1953, pgs 177–185. URL <http://biomet.oxfordjournals.org/cgi/doi/10.1093/biomet/40.1-2.177>
- [12] D. J. Daley, J. Gani, Epidemic Modelling: An Introduction, Cambridge Studies in Mathematical Biology, 2001.
- [13] N. C. Grassly, C. Fraser, Mathematical models of infectious disease transmission, *Nature Reviews Microbiology*, vol. 6, nr. 6, 2008, pgs. 477–487. URL <http://www.ncbi.nlm.nih.gov/pubmed/18533288>
- [14] H. Hethcote, The mathematics of infectious diseases, *SIAM review*, 2000, pgs 599–653.
- [15] A. Kaddar, A. Abta, H. T. Alaoui, A comparison of delayed sir and seir epidemic models, *Nonlinear Analysis: Modelling and Control*, vol. 16, nr. 2, 2011, pgs 181–190.
- [16] M. G. Roberts, J. A. P. Heesterbeek, Mathematical models in epidemiology, *Mathematical models. UNESCO-EOLSS*, 2004.
- [17] Vargas-De-León, Cruz. Modeling control strategies for influenza A H1N1 epidemics: SIR models. *Revista Mexicana de Física*, vol. 58, 2012, pgs 37–43.
- [18] W. O. Kermack, A. G. McKendrick, 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*, vol. 115, nr. 772, 1927, pgs 700–721. URL:<http://rspa.royalsocietypublishing.org/cgi/doi/10.1098/ rspa.1927.0118>
- [19] W. O. Kermack, A. G. McKendrick, Contributions to the mathematical theory of epidemics. ii. the

problem of endemicity, *Proceedings of the Royal Society of London Series A*, vol. 138, nr. 834, 1932, pgs 55–83. URL: <http://rspa.royalsocietypublishing.org/cgi/doi/10.1098/rspa.1932.0171>

- [20] G. M. S. Dias, *Modelo epidemiológico SIR aplicado a redes tolerantes a atrasos e desconexões*, Tese de Mestrado, Instituto Militar de Engenharia, 2013.
- [21] G. M. S. Dias, R. M. Salles, Epidemic SIR model applied to delay-tolerant networks, 30th Brazilian Telecommunications Symposium, SBrT'12, SBrT, 2012.
- [22] M. J. Keeling, P. Rohani, *Modeling Infectious Diseases in Humans and Animals*, Princeton University Press, 2011.
- [23] P. van den Driessche, Reproduction numbers of infectious disease models, *Infectious Disease Modelling*, vol. 2, nr. 3, 2017, pgs 288-303.