Um modelo SEIRD generalizado com mecanismo implícito de quarentena para a disseminação da COVID-19: uma abordagem Bayesiana

Regina C. Almeida[1]

Diego Tavares Volpatto[1,2], Anna Claudia Mello de Resende[1],Lucas dos Anjos[1], João Vitor Oliveira Silva[1], Claudia Mazza Dias[3], Sandra Mara Cardoso Malta[1]

COLMEA, UERJ

[1] Laboratório Nacional de Computação Científica
[2] ESSS, Engineering Simulation and Scientific Software
[3] Universidade Federal Rural do Rio de Janeiro

25/06/2020

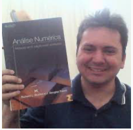
UFRRJ UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO

Laboratório Nacional de Computação Científica

Equipe

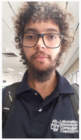
Sandra Malta, D.Sc., LNCC

Claudia M. Dias, D.Sc., UFRRJ



Diego T. Volpatto, M.Sc., LNCC

Anna C.M. Resende, M.Sc., LNCC



Lucas dos Anjos, D.Sc., LNCC

João V.O. Silva, B.Sc., LNCC

Planejamento

- Introdução
- Modelagem Matemática & Métodos
  - Modelo
  - Dados
  - Calibração Bayesiana
  - Análise de Sensibilidade
- Resultados
- Observações Finais

Introdução - A COVID-19 no mundo

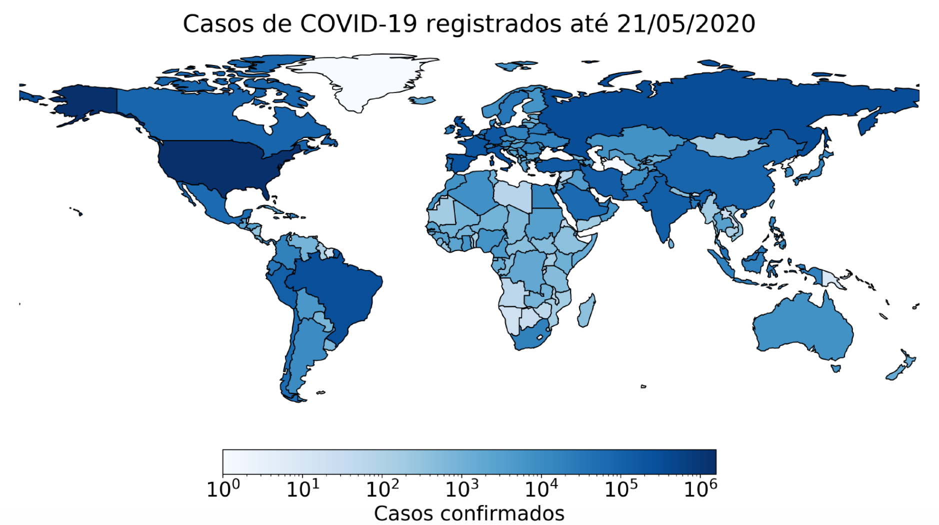


Figura 1: Distribuição mundial dos casos confirmados de COVID-19 até 21/05/2020. Dados retirados do repositório da Johns Hopkins University.

Introdução - A COVID-19 no Brasil

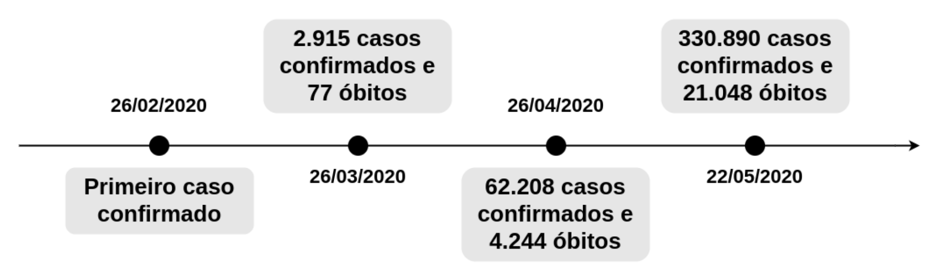
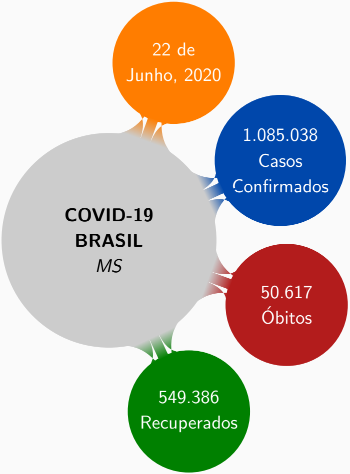


Figura 2: Linha do tempo da evolução de casos confirmados e total de óbitos para o Brasil até 21/05/2020. Dados do Ministério da Saúde - MS.

Introdução



Números: Brasil

- "falta" de dados confiáveis
- capacidade de testagem limitada
- política de testagem restrita
- subnotificações
- heterogeneidade sócio-demográfica

Modelo Matemático (BR/RJ)

- generalização do modelo SEIRD
- quarentena (distanciamento social)
- incertezas

Introdução - Por que modelo do tipo SEIRD?

SEIRD: Susceptible-Exposed-Infected-Recovered-Dead

Características

- equações diferenciais ordinárias
- homogeneidade espacial
- $S(t) + E(t) + I(t) + R(t) + D(t) = N \rightarrow$ população total
- "simples"

O que (não) se conhece sobre a COVID-19?

- tempo de latência: 3 a 7 dias
- indivíduos infectados podem ou não apresentar sintomas: infectados sintomáticos e assintomáticos
- tempo de recuperação: cerca de 14 dias
- imunidade

Introdução - A dinâmica da COVID-19

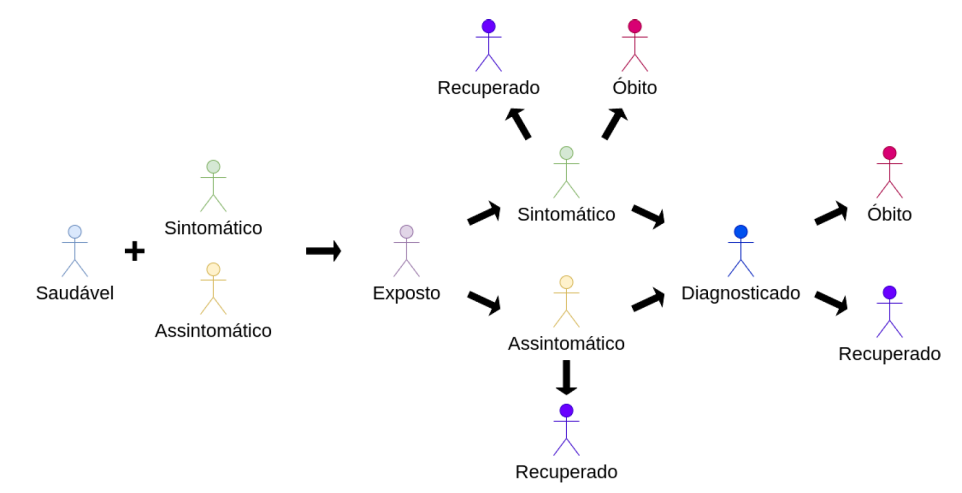
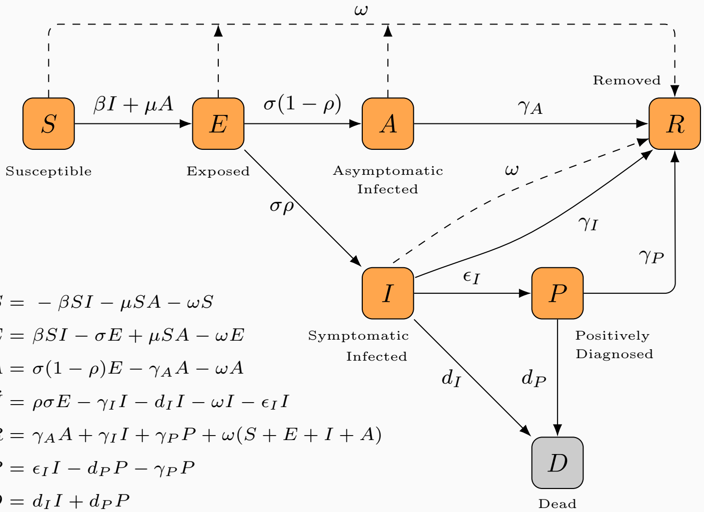


Figura 3: Evolução do estado de um indivíduo inicialmente saudável/suscetível.

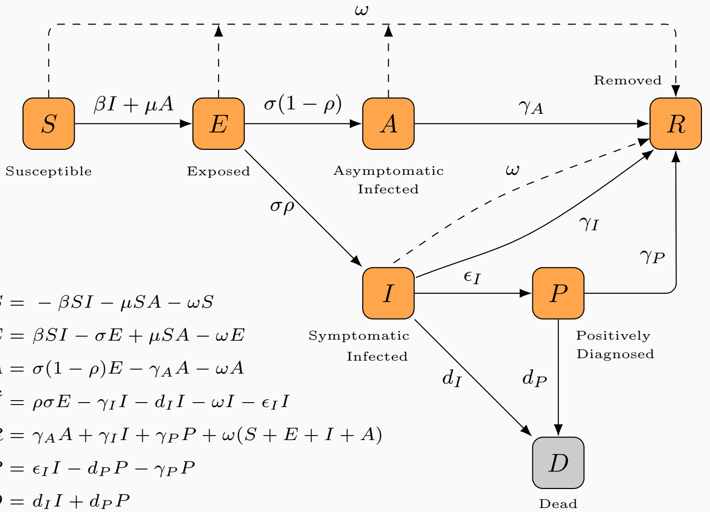
- “Essentially, all models are wrong, but some are useful.” George Box



$$
\begin{aligned}
\dot{S} &= -\beta SI - \mu SA - \omega S \\
\dot{E} &= \beta SI - \sigma E + \mu SA - \omega E \\
\dot{A} &= \sigma(1-\rho)E - \gamma_A A - \omega A \\
\dot{I} &= \rho\sigma E - \gamma_I I - d_I I - \omega I - \epsilon_I I \\
\dot{R} &= \gamma_A A + \gamma_I I + \gamma_P P + \omega(S + E + I + A) \\
\dot{P} &= \epsilon_I I - d_P P - \gamma_P P \\
\dot{D} &= d_I I + d_P P
\end{aligned}
$$

Figura 4: Descrição esquemática do modelo SEAIRPD-Q.

- β, μ: Taxas de transmissão por contato com indivíduos infectados I e A;

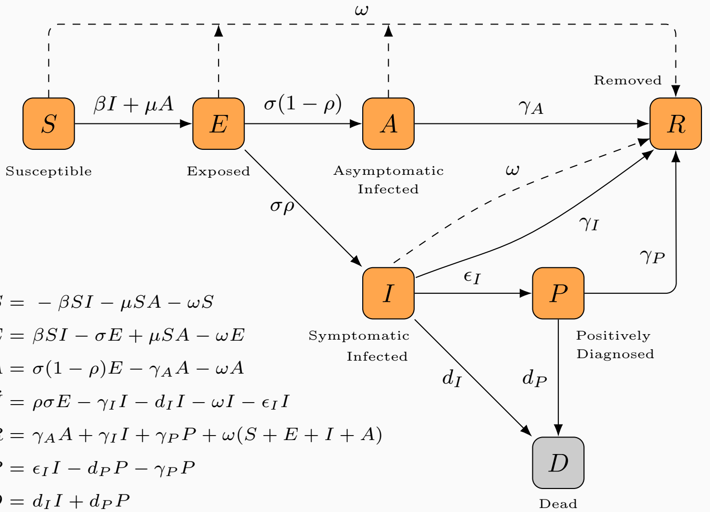


$$
\begin{aligned}
\dot{S} &= -\beta SI - \mu SA - \omega S \\
\dot{E} &= \beta SI - \sigma E + \mu SA - \omega E \\
\dot{A} &= \sigma(1-\rho)E - \gamma_A A - \omega A \\
\dot{I} &= \rho\sigma E - \gamma_I I - d_I I - \omega I - \epsilon_I I \\
\dot{R} &= \gamma_A A + \gamma_I I + \gamma_P P + \omega(S + E + I + A) \\
\dot{P} &= \epsilon_I I - d_P P - \gamma_P P \\
\dot{D} &= d_I I + d_P P
\end{aligned}
$$

Figura 4: Descrição esquemática do modelo SEAIRPD-Q.

Modelo Matemático: SEAIRPD-Q

- σ: Taxa de transição de E para I e A;

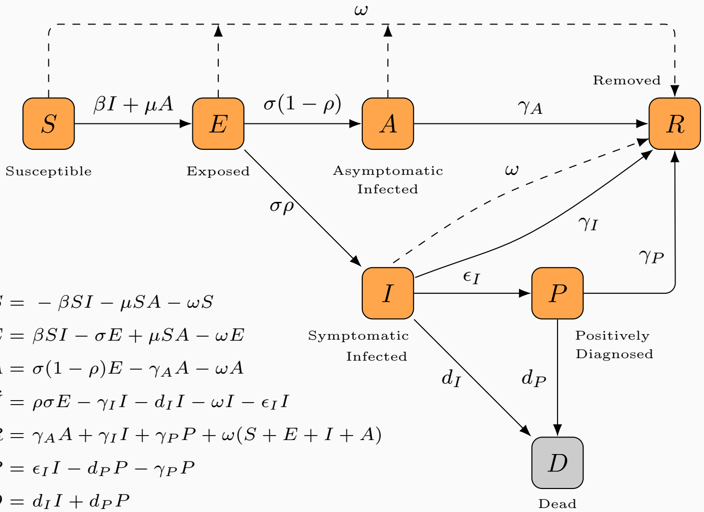


$$\dot{S} = -\beta S I - \mu S A - \omega S$$
$$\dot{E} = \beta S I - \sigma E + \mu S A - \omega E$$
$$\dot{A} = \sigma(1-\rho)E - \gamma_A A - \omega A$$
$$\dot{I} = \rho\sigma E - \gamma_I I - d_I I - \omega I - \epsilon_I I$$
$$\dot{R} = \gamma_A A + \gamma_I I + \gamma_P P + \omega(S + E + I + A)$$
$$\dot{P} = \epsilon_I I - d_P P - \gamma_P P$$
$$\dot{D} = d_I I + d_P P$$

Figura 4: Descrição esquemática do modelo SEAIRPD-Q.

Modelo Matemático: SEAIRPD-Q

- ρ: Proporção de indivíduos infectados que apresentam sintomas;

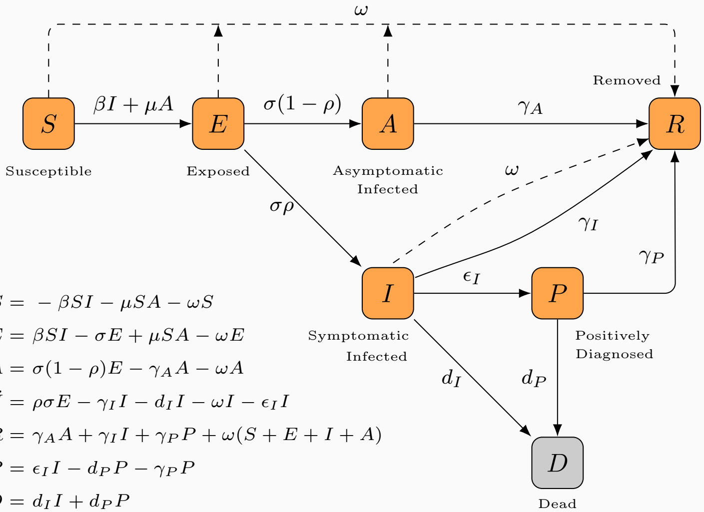


$$
\begin{aligned}
\dot{S} &= -\beta SI - \mu SA - \omega S \\
\dot{E} &= \beta SI - \sigma E + \mu SA - \omega E \\
\dot{A} &= \sigma(1-\rho)E - \gamma_A A - \omega A \\
\dot{I} &= \rho\sigma E - \gamma_I I - d_I I - \omega I - \epsilon_I I \\
\dot{R} &= \gamma_A A + \gamma_I I + \gamma_P P + \omega(S + E + I + A) \\
\dot{P} &= \epsilon_I I - d_P P - \gamma_P P \\
\dot{D} &= d_I I + d_P P
\end{aligned}
$$

Figura 4: Descrição esquemática do modelo SEAIRPD-Q.

Modelo Matemático: SEAIRPD-Q

- ε_I: Taxa de diagnóstico (de indivíduos infectados qua apresentam sintomas);

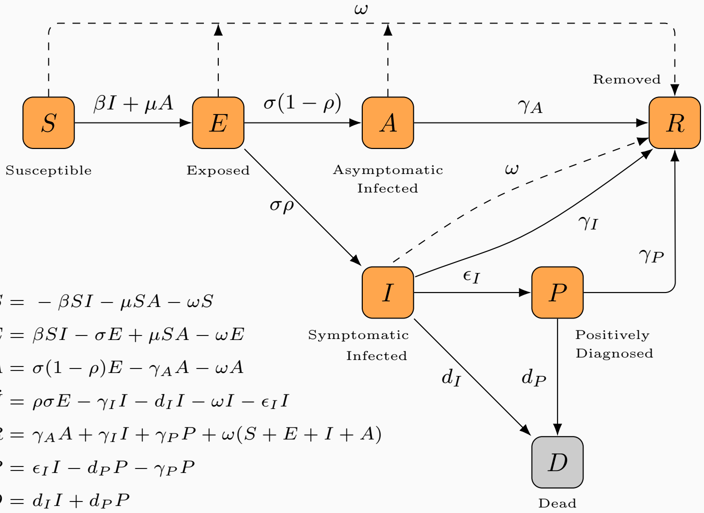


$$
\begin{aligned}
\dot{S} &= -\beta S I - \mu S A - \omega S \\
\dot{E} &= \beta S I - \sigma E + \mu S A - \omega E \\
\dot{A} &= \sigma(1-\rho)E - \gamma_A A - \omega A \\
\dot{I} &= \rho\sigma E - \gamma_I I - d_I I - \omega I - \epsilon_I I \\
\dot{R} &= \gamma_A A + \gamma_I I + \gamma_P P + \omega(S + E + I + A) \\
\dot{P} &= \epsilon_I I - d_P P - \gamma_P P \\
\dot{D} &= d_I I + d_P P
\end{aligned}
$$

Figura 4: Descrição esquemática do modelo SEAIRPD-Q.

Modelo Matemático: SEAIRPD-Q

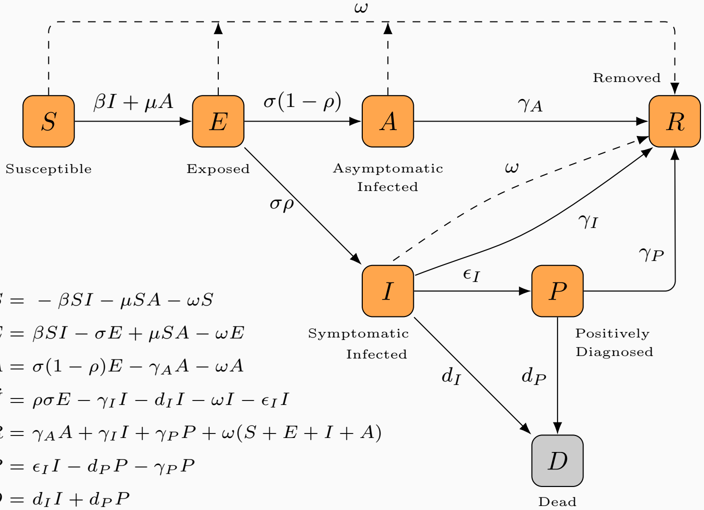
- γ_A, γ_I, γ_P: Taxas de recuperação de A, I e P, respectivamente;



$$
\begin{aligned}
\dot{S} &= -\beta SI - \mu SA - \omega S \\
\dot{E} &= \beta SI - \sigma E + \mu SA - \omega E \\
\dot{A} &= \sigma(1-\rho)E - \gamma_A A - \omega A \\
\dot{I} &= \rho\sigma E - \gamma_I I - d_I I - \omega I - \epsilon_I I \\
\dot{R} &= \gamma_A A + \gamma_I I + \gamma_P P + \omega(S + E + I + A) \\
\dot{P} &= \epsilon_I I - d_P P - \gamma_P P \\
\dot{D} &= d_I I + d_P P
\end{aligned}
$$

Figura 4: Descrição esquemática do modelo SEAIRPD-Q.

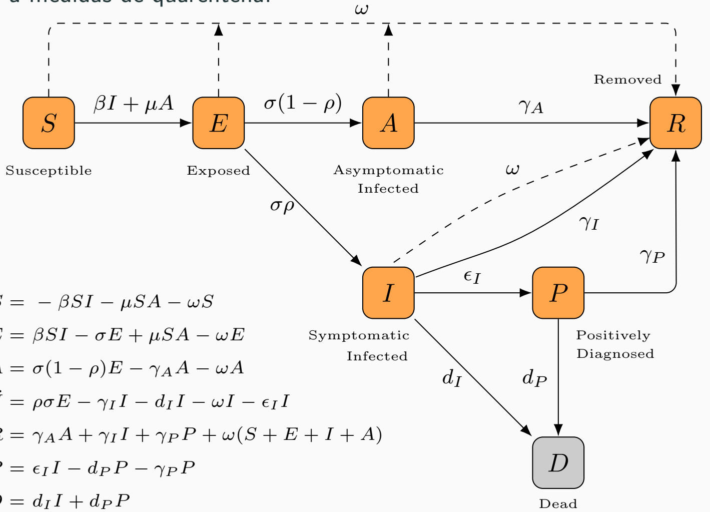
- d_I, d_P: Taxas de mortalidade de I e P;



$$
\begin{aligned}
\dot{S} &= -\beta SI - \mu SA - \omega S \\
\dot{E} &= \beta SI - \sigma E + \mu SA - \omega E \\
\dot{A} &= \sigma(1-\rho)E - \gamma_A A - \omega A \\
\dot{I} &= \rho\sigma E - \gamma_I I - d_I I - \omega I - \epsilon_I I \\
\dot{R} &= \gamma_A A + \gamma_I I + \gamma_P P + \omega(S + E + I + A) \\
\dot{P} &= \epsilon_I I - d_P P - \gamma_P P \\
\dot{D} &= d_I I + d_P P
\end{aligned}
$$

Figura 4: Descrição esquemática do modelo SEAIRPD-Q.

- ω: Taxa em que S, E, I, e A são removidos (colocados em isolamento) devido à medidas de quarentena.



$$
\begin{aligned}
\dot{S} &= -\beta SI - \mu SA - \omega S \\
\dot{E} &= \beta SI - \sigma E + \mu SA - \omega E \\
\dot{A} &= \sigma(1-\rho)E - \gamma_A A - \omega A \\
\dot{I} &= \rho\sigma E - \gamma_I I - d_I I - \omega I - \epsilon_I I \\
\dot{R} &= \gamma_A A + \gamma_I I + \gamma_P P + \omega(S + E + I + A) \\
\dot{P} &= \epsilon_I I - d_P P - \gamma_P P \\
\dot{D} &= d_I I + d_P P
\end{aligned}
$$

Figura 4: Descrição esquemática do modelo SEAIRPD-Q.

Número de Reprodução

Calculado usando o método *Next Generation Matrix* [1]:

- vetor: $\mathbf{x}^{\mathsf{T}} = [E, I, A]$
- $\dfrac{d\mathbf{x}}{dt} = (T + \Sigma)\mathbf{x}$, $T = \begin{bmatrix} 0 & \beta S & \mu S \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ e $-\Sigma = \begin{bmatrix} \sigma + \omega & 0 & 0 \\ -\sigma\rho & \gamma_I + d_I + \omega + \epsilon_I & 0 \\ -\sigma(1-\rho) & 0 & \gamma_A + \omega \end{bmatrix}$
- $\mathcal{R}_0$ é o autovalor dominante da matriz $K = -T\Sigma^{-1}$

Número básico de reprodução

$$\mathcal{R}_0 = \left(\frac{\sigma(1-\rho)\mu}{(\sigma+\omega)(\gamma_A+\omega)} + \frac{\beta\sigma\rho}{(\sigma+\omega)(\gamma_I + d_I + \omega + \epsilon_I)} \right) S(0)$$

Número efetivo de reprodução

$$\mathcal{R}(t) = \left(\frac{\sigma(1-\rho)\mu}{(\sigma+\omega)(\gamma_A+\omega)} + \frac{\beta\sigma\rho}{(\sigma+\omega)(\gamma_I + d_I + \omega + \epsilon_I)} \right) S(t)$$

[1] Diekmann et al., Journal of The Royal Society Interface, 7:873–885, 2010.

Dados

Casos acumulados de confirmados e óbitos

- BR: *https://covid.saude.gov.br/*
  dados epidemiológicos de 63 dias: de 05/03 a 06/05
- RJ: *https://covid19br.wcota.me*
  dados epidemiológicos de 53 dias: de 10/03 a 01/05

Dia 0 epidemiológico (t_0): pelo menos 5 casos confirmados

SBMAC: A Matemática Aplicada na pandemia de COVID-19

- Notebook (Python): *https://www.kaggle.com/resendeacm/covid-19-rio-de-janeiro-city*
- Dataset: *https://www.kaggle.com/resendeacm/covid-19-rio-de-janeiro-city*

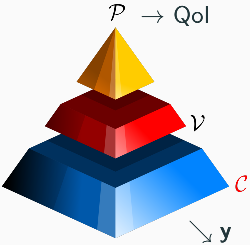
Calibração Bayesiana

Model Calibration

$\boldsymbol{\theta}$:	Vetor de "parâmetros"
$\mathbf{y}$:	Medidas experimentais
π :	Função densidade de probabilidade

$$\pi_{\text{post}}(\boldsymbol{\theta}|\mathbf{y}) = \frac{\pi_{\text{like}}(\mathbf{y}|\boldsymbol{\theta}) \cdot \pi_{\text{prior}}(\boldsymbol{\theta})}{\pi_{\text{evid}}(\mathbf{y})}$$

Hierarquia de cenários



- π_{prior}: conhecimento inicial sobre $\boldsymbol{\theta}$
- π_{like}: função de verossimilhança de $\mathbf{y}$ dado $\boldsymbol{\theta}$
- π_{evid}: informação contida em $\mathbf{y}$
- π_{post}: conhecimento atualizado sobre $\boldsymbol{\theta}$ dado $\mathbf{y}$.

Calibração Bayesiana

- Parâmetros calibrados: $\{\beta = \mu, d_I, d_P, \omega\}$ e $\{\sigma_C, \sigma_D\}$
- Ruídos nos dados: $\mathcal{N}(0, \sigma_C^2)$ e $\mathcal{N}(0, \sigma_D^2)$
- Quantidades observadas C e D:

$$C(t_i) := y^C_{\text{model}}(t_i) = \int_{t_0}^{t_i} \varepsilon_I I(t) dt \quad \text{e} \quad D(t_i) := y^D_{\text{model}}(t_i) = \int_{t_0}^{t_i} \left[d_I I(t) + d_P P(t)\right] dt$$

Função de Verossimilhança

$$\pi_{\text{like}}(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{j \in \{C,D\}} \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \left(\frac{y^{(j)}(t_i) - y^{(j)}_{\text{model}}(t_i)}{\sigma_j} \right)^2 \right)$$

- Calibração realizada usando `PyMC3`
- Resolução do problema inverso utiliza o método CATMIP (*Cascading Adaptive Transitional Metropolis in Parallel*)

Análise de Sensibilidade

Por que?

- Como as incertezas nos parâmetros afetam a QoI selecionada?
- Contribui para um melhor entendimento sobre o modelo.

Número Efetivo de Reprodução

$$\mathrm{QoI}_1(t) = \mathcal{R}(t)$$

Soma Quadrada Normalizada

$$\mathrm{QoI}_2(t) = \sqrt{C(t)^2 + D(t)^2}$$

- **Elementary Effects Method** (Morris, 1991[2])
- Hipótese: para cada parâmetro $\bar{\theta}_i$ analisado: $\theta_i \sim \mathcal{U}(0.5\bar{\theta}_i, 1.5\bar{\theta}_i)$

[2] Campolongo et al. An effective screening design for sensitivity analysis of large models. Environmental modelling & software, 22(10):1509–1518, 2007

Resultados - Calibração para o estado do Rio de Janeiro

Tabela 1: Condições Iniciais.

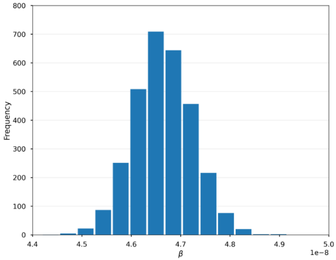
Classe	População Inicial (Ind.)	Fonte
N	17.264.943	IBGE 2019
$P(0)$	8	Dados do dia 10-03-2020
$E(0)$	$10 \cdot P(0)$	
$I(0)$	$5 \cdot P(0)$	
$A(0)$	$P(0)$	
$D(0)$	0	Dados do dia 10-03-2020
$R(0)$	0	Dados do dia 10-03-2020
$S(0)$	17.264.837	$S(0) = N - \{E(0) + A(0) + I(0) + P(0) + R(0) + D(0)\}$

Calibração para o estado do Rio de Janeiro

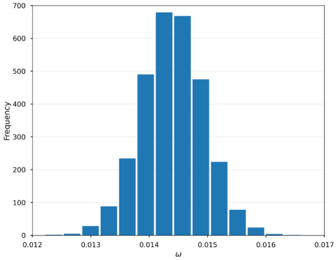
Tabela 2: Valores fixados e mais prováveis (MAP) (u.a.).

Parâmetro	Valor/MAP (RJ)	π_{prior}	Source
$\beta = \mu$	4.658646×10^{-8}	$\mathcal{U}(0, 1 \times 10^{-5})$	CATMIP
σ	$1/5$	–	*
ρ	0.85	–	*
ε_I	$1/3$	–	*
$\gamma_A, \gamma_I, \gamma_P$	$1/14$	–	*
d_I	5.549912×10^{-4}	$\mathcal{U}(0, 0.1)$	CATMIP
d_P	1.315661×10^{-2}	$\mathcal{U}(0, 0.1)$	CATMIP
ω	1.970400×10^{-2}	$\mathcal{U}(0, 1)$	CATMIP
$\mathcal{R}_0$	2.830033	–	NGM
σ_C	148	$\mathcal{U}(1, 10000)$	CATMIP
σ_D	22	$\mathcal{U}(1, 10000)$	CATMIP

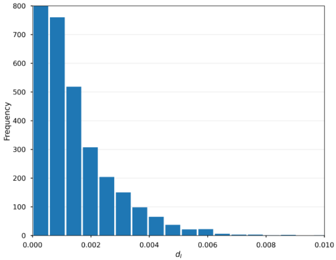
Distribuições a posteriori



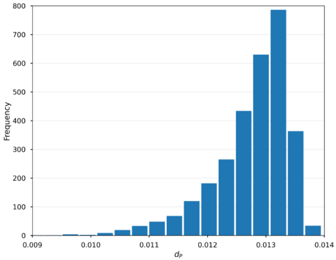
(A) Calibrated parameter: β.



(B) Calibrated parameter: ω.



(C) Calibrated parameter: d_I.



(D) Calibrated parameter: d_P.

Projeções

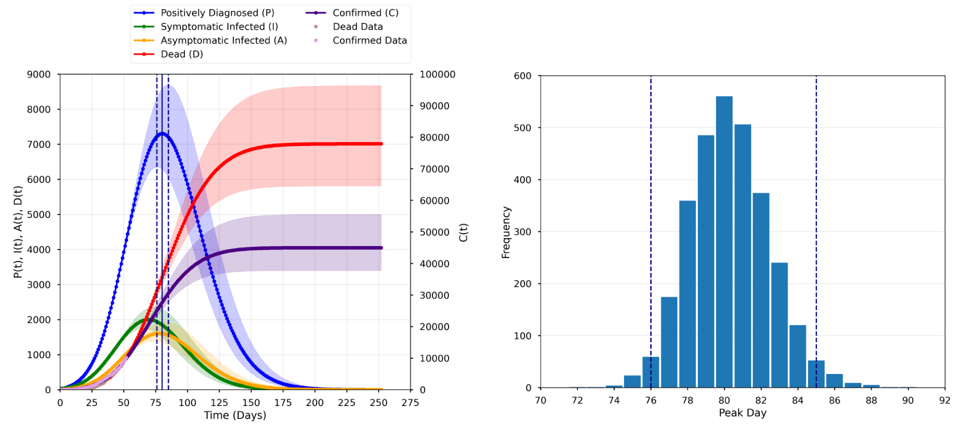


Figura 5: Predições de P, I, A, D e C para o Rio de Janeiro. As linhas verticais pontilhadas indicam o intervalo de credibilidade de 95% para o pico dos infectados, cuja distribuição é apresentada à direita.

Projeções

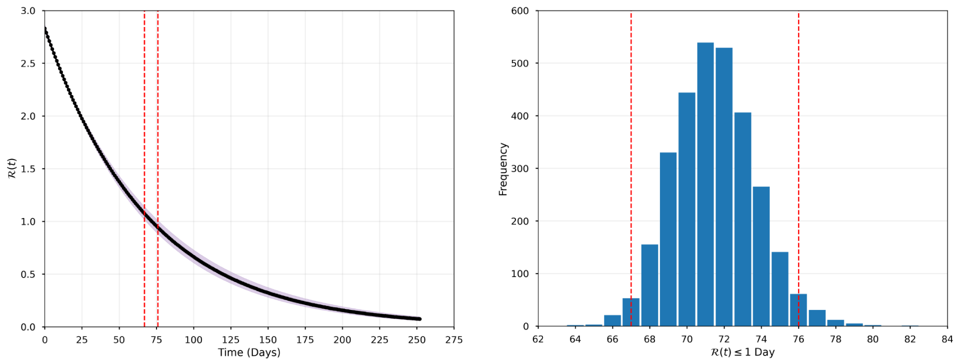
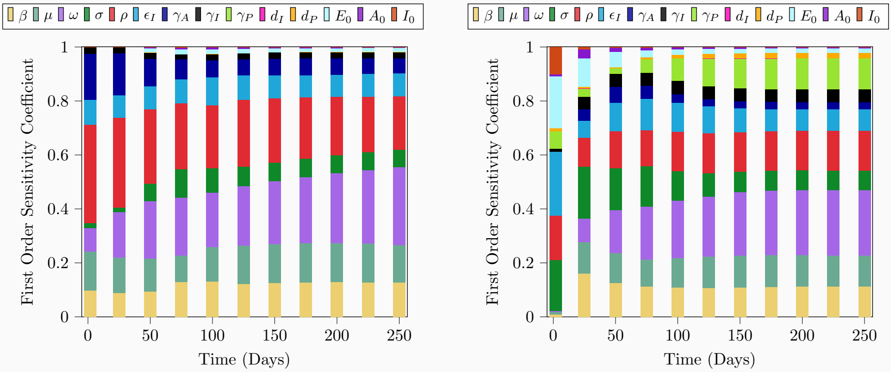


Figura 6: Evolução do $\mathcal{R}(t)$ e histograma para o dia em que $\mathcal{R}(t) = 1$.



(a) $\mathrm{QoI}_1(t) = \mathcal{R}(t)$ **(b)** $\mathrm{QoI}_2(t) = \sqrt{C(t)^2 + D(t)^2}$

Figura 7: Alterações temporais do índice de sensibilidade de primeira ordem.

Cenários hipotéticos: alterando as medidas de distanciamento social

Taxa de remoção

$$\omega_r = \omega e^{-\lambda(t-t_d)}$$

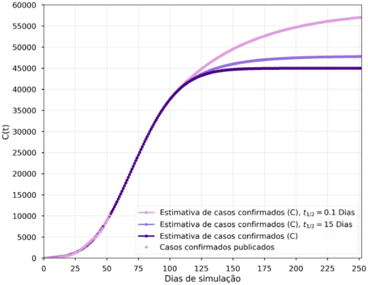
Constante de decaimento: $\lambda = \ln 2/t_{1/2}$

Dia da implementação: $t_d = 100$ dias

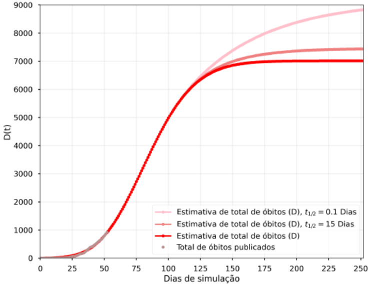
COVID-19

Cenários considerados

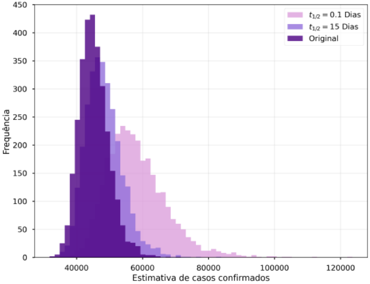
- Relaxamento brusco, com meia-vida de $t_{1/2} = 0.1$ dias
- Relaxamento gradual, com meia-vida de $t_{1/2} = 15$ dias
- Situação original, com taxa de remoção ω



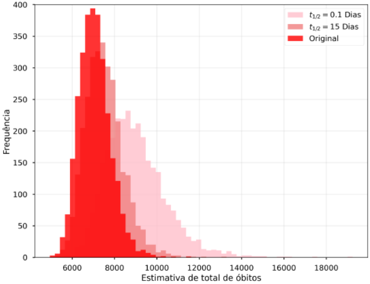
(a) Estimativas dos valores médios de casos confirmados.



(b) Estimativas dos valores médios de total de óbitos.



(c) Distribuição da estimativa de casos confirmados no último dia de simulação.



(d) Distribuição da estimativa de total de óbitos no último dia de simulação.

Cenários hipotéticos: alterando as medidas de distanciamento social em $t_d = 100$ dias

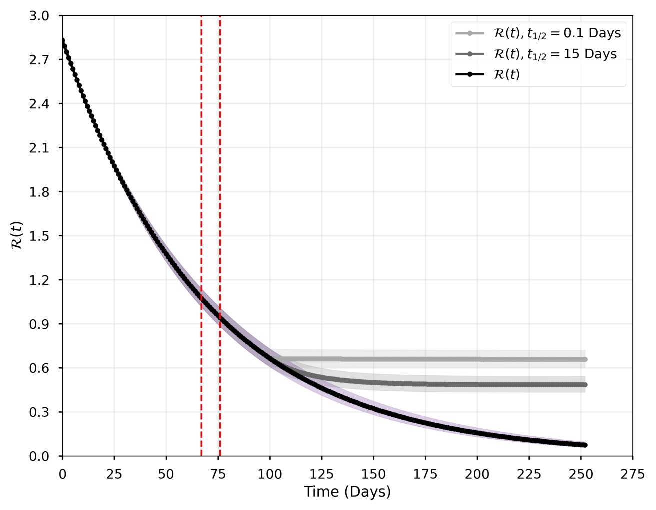


Figura 8: Evolução do $\mathcal{R}(t)$ in BR. Linhas pontilhadas indicam intervalo de credibilidade de 95% em torno de $\mathcal{R}(t) = 1$.

Um outro cenário hipotético (mais crítico!)

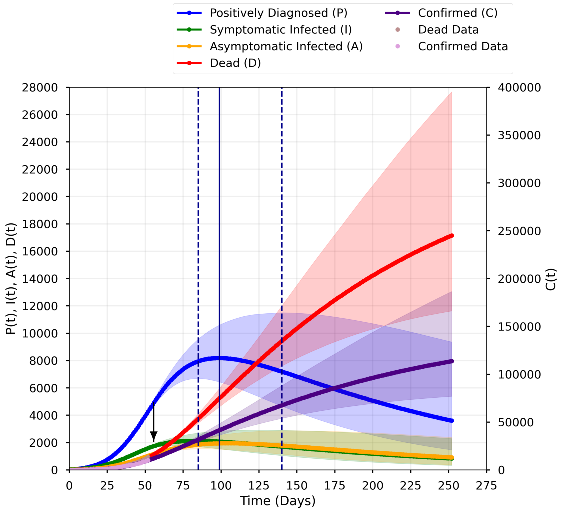


Figura 9: Predições considerando liberação da quarentena com $t_{1/2} = 20$ dias realizada 20 dias antes da ocorrência do pico dos casos infectados ativos. Situação em 22/06/2020: 97572 confirmados e 8933 óbitos acumulados (semana 26).

Observações Finais

SEAIRPD-Q

- Desenvolvemos um modelo SEIRD generalizado que leva em conta quarentena;
- Caso de estudo: COVID-19 no RJ;
- Dados reais e incertezas foram consideradas.

Sumário dos resultados

- AS identificou os parâmetros ω e ρ como os mais influentes;
- Resultados apontam para a necessidade de manter as políticas de distanciamento social; tanto o dia quanto a forma de relaxamento da quarentena devem ser cuidadosamente planejados;
- AS aponta a importância de uma política de testagem mais ampla, capaz de identificar os infectados sintomáticos e assintomáticos;
- As incertezas nas predições aumentam de forma significativa ao longo do tempo evidenciando a importância de "alimentar" continuamente o modelo.

Perspectivas

- Incorporação das novas informações sobre subnotificação e percentual de assintomáticos;
- Desenvolvimento de novos modelos (extensões/simplificações - família de modelos) e de procedimentos sistemáticos de seleção de modelos;
- Extensão para heterogeneidade espacial: modelagem de metapopulações e em redes complexas.

Referências

Diego T. Volpatto, Anna C. M. Resende, Lucas dos Anjos, João V. O. Silva, Claudia M. Dias, Regina C. Almeida, and Sandra M. C. Malta.
Spreading of COVID-19 in Brazil: Impacts and uncertainties in social distancing strategies.
medRxiv:10.1101/2020.05.30.20117283, p. 1–34, 2020.

Diego T. Volpatto, Anna C. M. Resende, Lucas dos Anjos, João V. O. Silva, Claudia M. Dias, Regina C. Almeida e Sandra M. C. Malta.
Avaliação de estratégias de relaxamento do distanciamento social para o Brasil e Estado do Rio de Janeiro.
SciELOPreprints:10.1590/SciELOPreprints.595, p. 1–13, 2020.

Diego T. Volpatto, Anna C. M. Resende, Lucas dos Anjos, João V. O. Silva, Claudia M. Dias, Regina C. Almeida, and Sandra M. C. Malta.
pydemic: Scripts for the BR/RJ social distancing study.
Zenodo:10.5281/zenodo.3865004