Mathematical modelling of COVID-19 and analysis of Rio de Janeiro city's outbreak in 2020

Lucas M. Moschen[1], Maria S. Aronna[1] and Roberto Guglielmi[2]
[1]Fundação Getulio Vargas, Brazil; [2]University of Waterloo; Canada

The authors acknowledge the support of FAPERJ, CNPq, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Development of Mathematics and Science (FGV CDMC, Brazil).

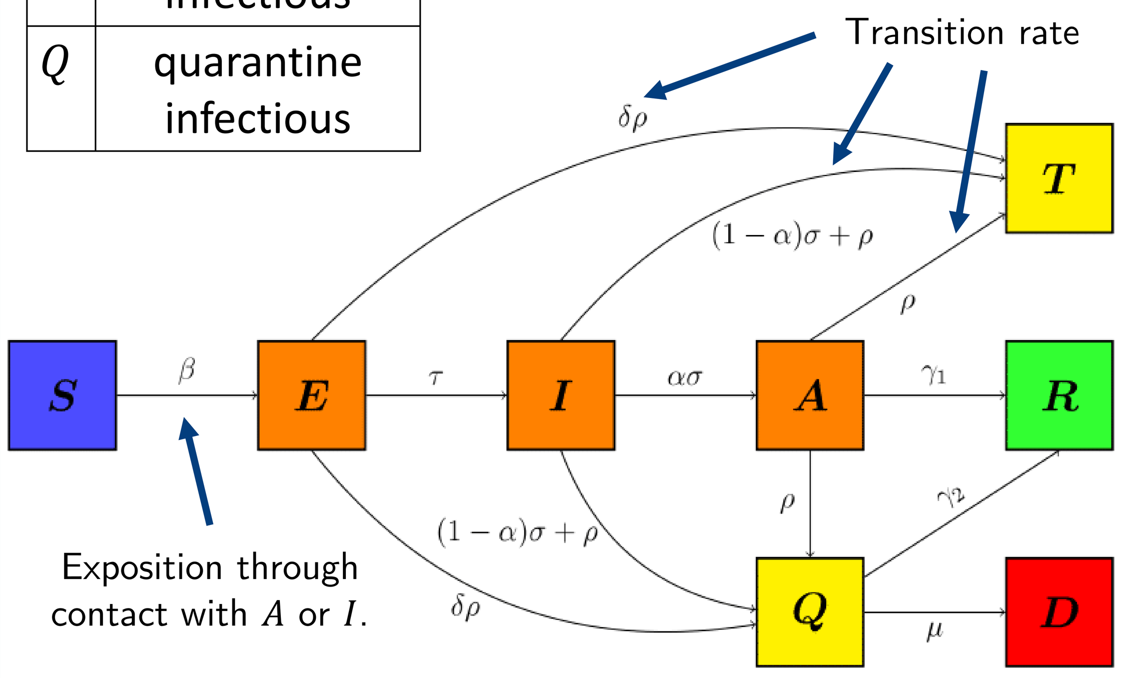
Highlights

- Model the spread of the coronavirus SARS-CoV-2 (COVID-19) in 2020 with **non-pharmaceutical interventions**.
- Analyze the data from **the first outbreak in the Rio de Janeiro** city.
- Estimate relevant parameters with emphasis to the **unreported rate** and the effective reproduction curve, which are relevant to measure the **real extent** and **impact** of the disease.
- Every reported case by the health system **corresponds 9–10 not detected cases**.

SEIAQR model

- **Compartmental model simplified** from [1] without the lockdown for non-essential workers. The diagram flow is

S	susceptible	R	recovered
E	exposed	D	death
I	infectious	T	total positive tests
A	asymptomatic infectious		
Q	quarantine infectious		



- The parameter $\beta(t)$ measures the **effective contact rate**, which depends on the average contact rate and the probability of infection given a contact.
- α is the **proportion of infectious and asymptomatic** individuals (undetected without testing). We consider this parameter as a proxy to the **unreported rate** of infections.
- ρ is the **testing rate** among infected people with mild or no symptoms. We estimated it using IBGE data.
- The **mortality rate** among detected cases is $\mu(t)$ and varies in periods of distress for the health system and lack of testing kits.
- The other rate parameters are related to the pathogen and taken from the literature.
- The **basic reproduction number** is

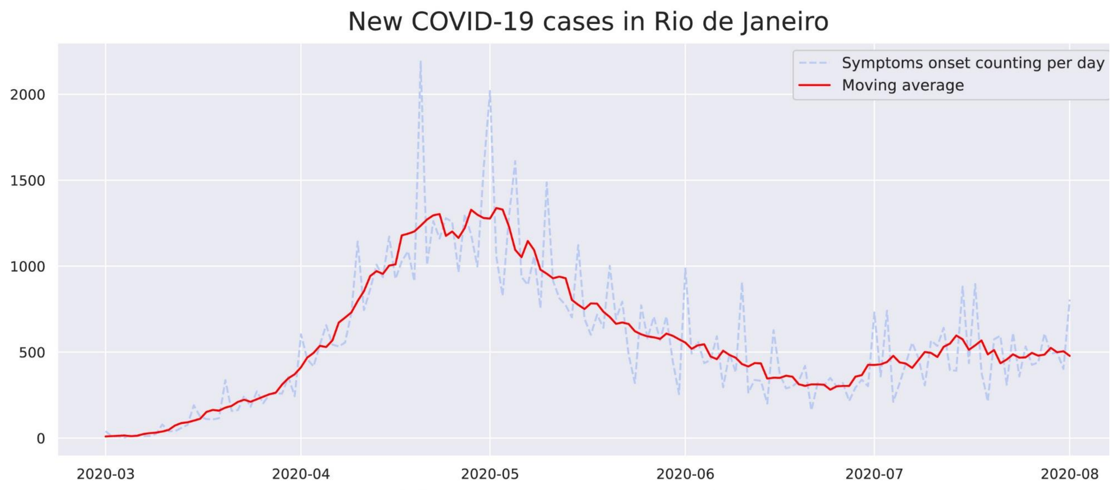
$$R_0 = \frac{1}{2}\left(\varphi + \sqrt{\varphi^2 + \frac{4\sigma\alpha}{\rho+\gamma_1}\varphi}\right), \varphi = \frac{\beta\tau}{(\rho\delta+\tau)(\sigma+\rho)}.$$

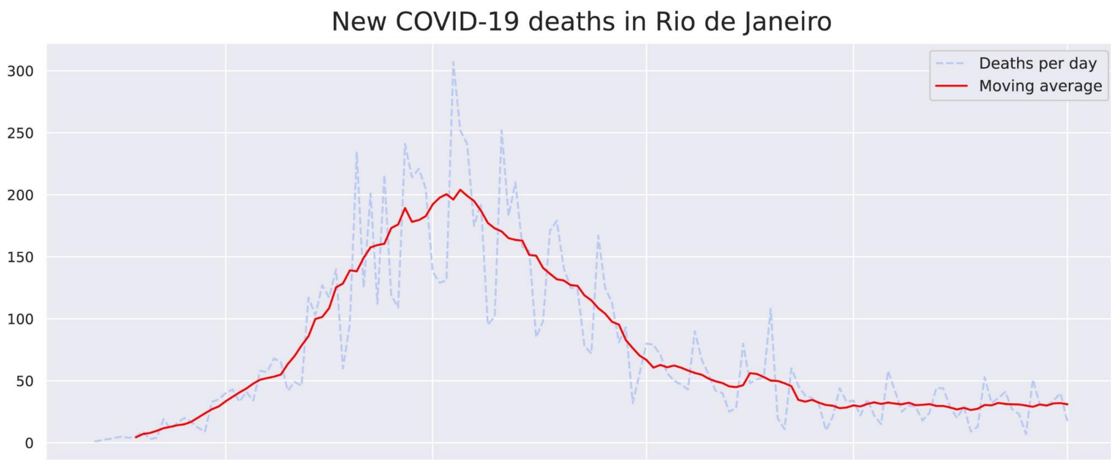
- As the epidemic evolves, the **effective reproduction curve** R_t has the same expression of R_0 with the change that

$$\varphi(t) = \frac{\beta(t)\tau S(t)}{(\rho\delta+\tau)(\sigma+\rho)}.$$

COVID-19 in Rio de Janeiro

- In March 2020, the state of Rio de Janeiro declared **public health emergency** and **social distancing measures,** but the outbreak could not be avoided:





- Scarce testing of infected people with **mild or no symptoms**.
- Data smoothed with a **7-day moving average**.

Methods

- Fit the **SEIAQR model** with $\beta(t), \mu(t)$ and α being unknown.
- The functions β and μ are written in a basis of polynomials of degree $k-1$, the **B-splines**:

$$\beta(t) \approx \sum_{j=1}^{s} \beta_j B_{j,k}(t), \qquad \mu(t) \approx \sum_{j=1}^{r} \mu_j B_{j,k}(t),$$

equally spaced knots (where the polynomials meet). The order and the number of coefficients are chosen using the Akaike Information Criterion: four coefficients for both with $k=2$ for β and $k=1$ for μ.

- The parameter estimation assumes that the observations have an **uncorrelated normally distributed error**:

$$\begin{cases} NewCases(i) = T(i) - T(i-1) + er_T(i) \\ NewDeaths(i) = D(i) - D(i-1) + er_D(i)' \end{cases}$$

and minimizes a weighted quadratic error, balancing new detected cases and deaths.

- A **parametric Bootstrap approach** is used to calculate confidence intervals. The parameters are estimated in new datasets generated by sampling the errors from the normal distribution and summing them to the observed new cases/deaths.

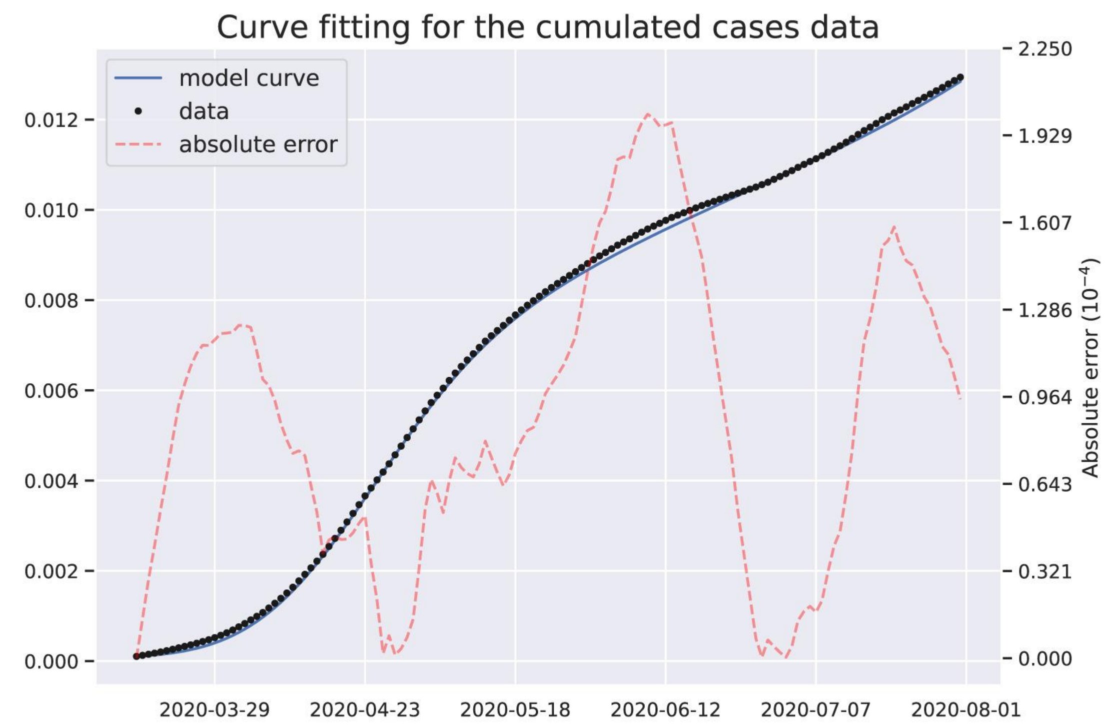
Identifiability

Can the parameters **be estimated in a unique way** from the available measurements?

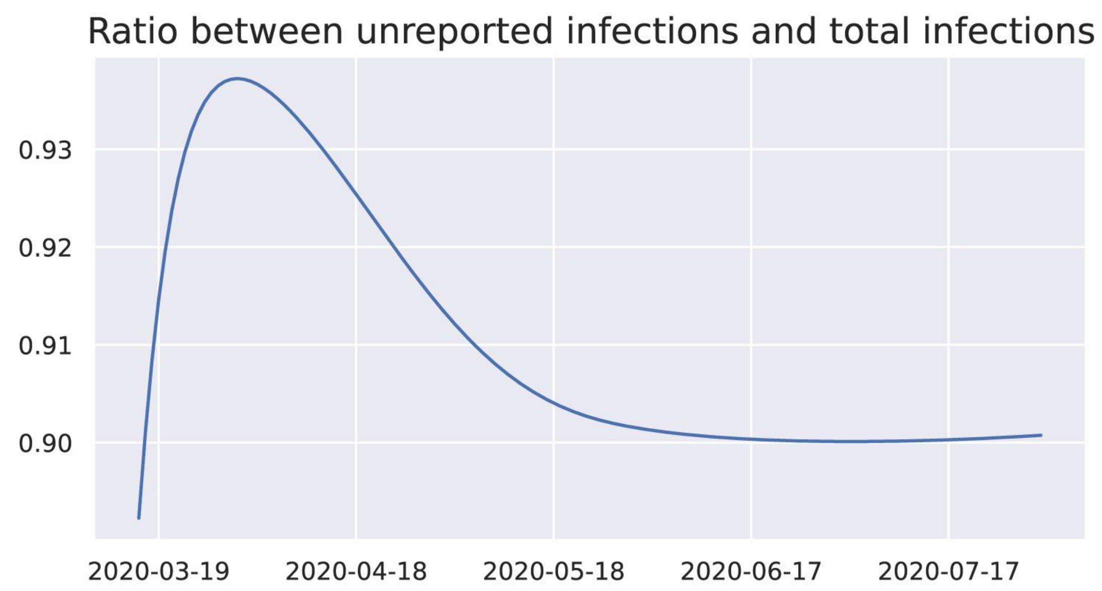
1. **Structural:** theoretical analysis of the structure of the model. We could not guarantee structural identifiability using the available methods in the literature.
2. **Practical:** considers the noisy data. We analyzed the correlation between the estimated parameters using the **Fisher Information Matrix (FIM)** and expected low values for the entries of its inverse.

Results

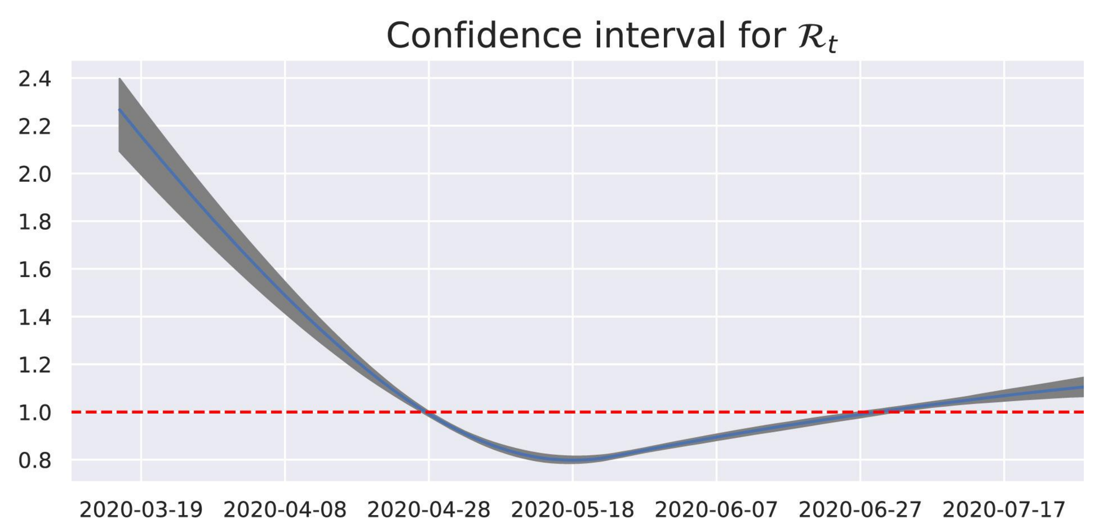
- The fitting presents a **good matching**:



- We had evidence that the residuals are **normally distributed but correlated**, which requires more attention in future works.
- We estimated $\hat{\alpha} = 0.9$ **with confidence interval** $(0.85, 0.93)$. We verified that the values of the non-estimated parameters do not impact the estimation of α.
- The evolution of **the non-tested positive cases over the total number of positive cases** is



- The estimated R_t curve with confidence interval is



- The FIM indicated correlation between α and β_2, which **limits the identifiability of the system**.

Discussion and conclusions

- The estimated range of values for α and the curve R_t over time **are in line with other works**.
- The identifiability of the model is a matter to study in the future.
- **Data analysis** techniques have been deployed to deal with the low quality of data.
- The rate of unreported positive cases is about 90% (85%-93%), that is, every case reported by the health system corresponds to about **9–10 cases that were not detected.**
- This estimation provides a useful tool to determine periods of growth or decrease of the epidemic force.

References

1) M.S. Aronna, R. Guglielmi, and L.M. Moschen, A model for COVID-19 with isolation, quarantine and testing as control measures. Epidemics, 2021.
2) M.S. Aronna, R. Guglielmi, and L.M. Moschen, Estimate of the rate of unreported COVID-19 cases during the first outbreak in Rio de Janeiro. Infectious Disease Modelling, 2022.
3) Municipal Health Departament, City Hall of Rio de Janeiro, Individual data of COVID-19 confirmed cases in the city of Rio de Janeiro, 2021.