The price of uncertainty during COVID-19

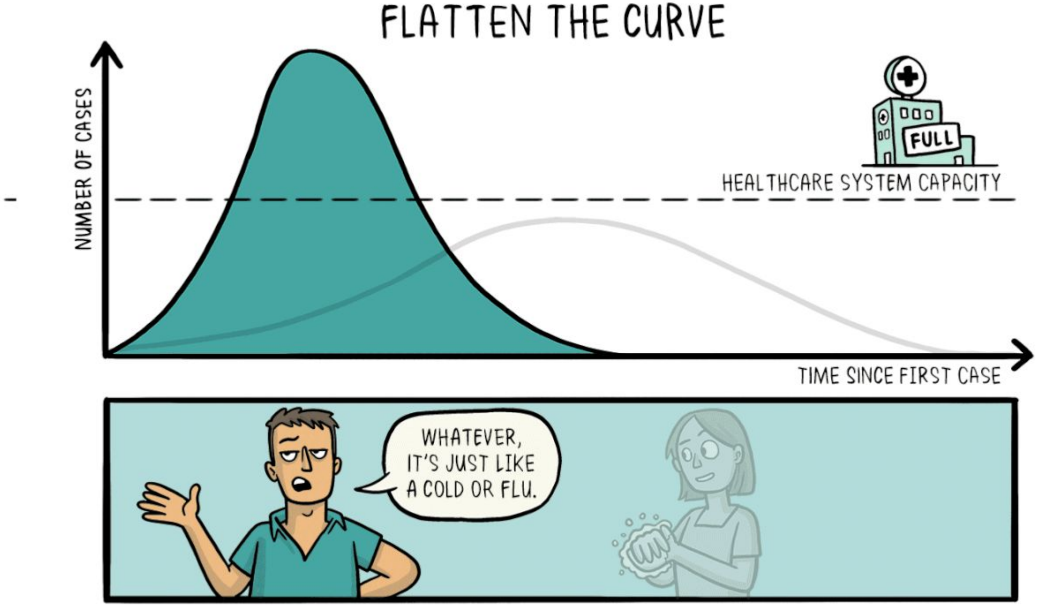
A sensitivity analysis on an extended infection model using UQpy

Teo Bucci
Flavia Petruso

POLITECNICO
MILANO 1863

Computational Statistics (8 CFU)
MSc. Mathematical Engineering
Politecnico di Milano

March 28th, 2023



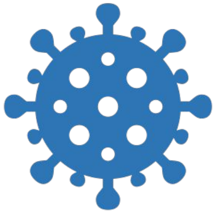
Summary

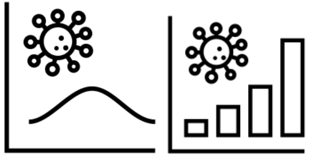
SIR models and beyond

The UQpy library

Results

Takeaways, conclusions, perspectives





Project goal

Reproduce the results of a sensitivity analysis from a previously published work, using UQpy.

Focus: the dynamics of the first 180 days of a pandemic of COVID-19.

2458 IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 26, NO. 6, JUNE 2022

Understanding Dynamics of Pandemic Models to Support Predictions of COVID-19 Transmission: Parameter Sensitivity Analysis of SIR-Type Models

Chunfeng Ma, *Member, IEEE*, Xin Li, *Senior Member, IEEE*, Zebin Zhao, Feng Liu, Kun Zhang, Adan Wu, and Xiaowei Nie

Crash course on compartmental models for epidemiology

Compartmental models in epidemiology

- Models in which people from a population of N individuals **move between compartments**
- The dynamics is essentially governed by the **initial conditions** and the **transition rates**

Compartmental models in epidemiology

- Models in which people from a population of N individuals **move between compartments**
- The dynamics is essentially governed by the **initial conditions** and the **transition rates**

Simplest case: **SIR** (Susceptible, Infectious, Recovered)

$$\text{Susceptible} \xrightarrow{\beta IS} \text{Infectious} \xrightarrow{\lambda I} \text{Recovered}$$

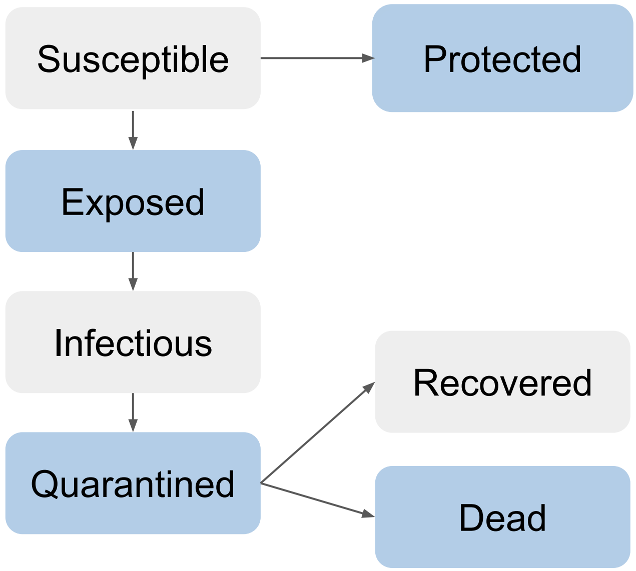
$$\frac{dS(t)}{dt} = -\beta \frac{S(t)\, I(t)}{N}$$
$$\frac{dI(t)}{dt} = \beta \frac{S(t)\, I(t)}{N} - \lambda I(t)$$
$$\frac{dR(t)}{dt} = \lambda I(t)$$

β: infection rate

λ: recovery rate

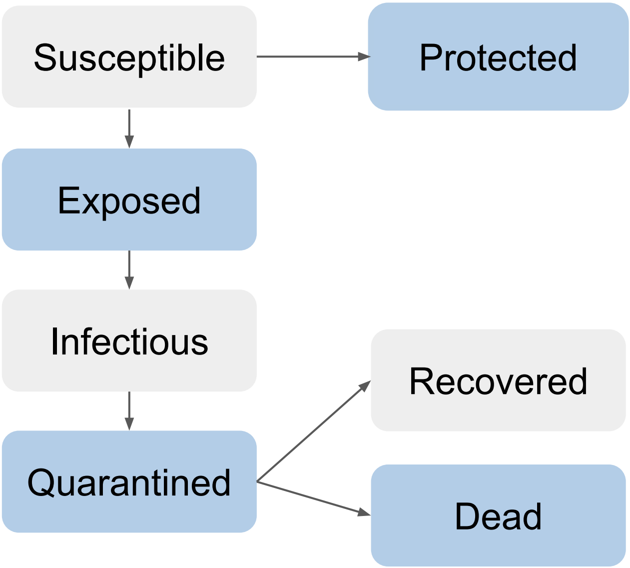
Beyond SIR: SEIQRDP models

- **New compartments** to make the dynamics more realistic



Beyond SIR: SEIQRDP models

- **New compartments** to make the dynamics more realistic
- **New governing equations**

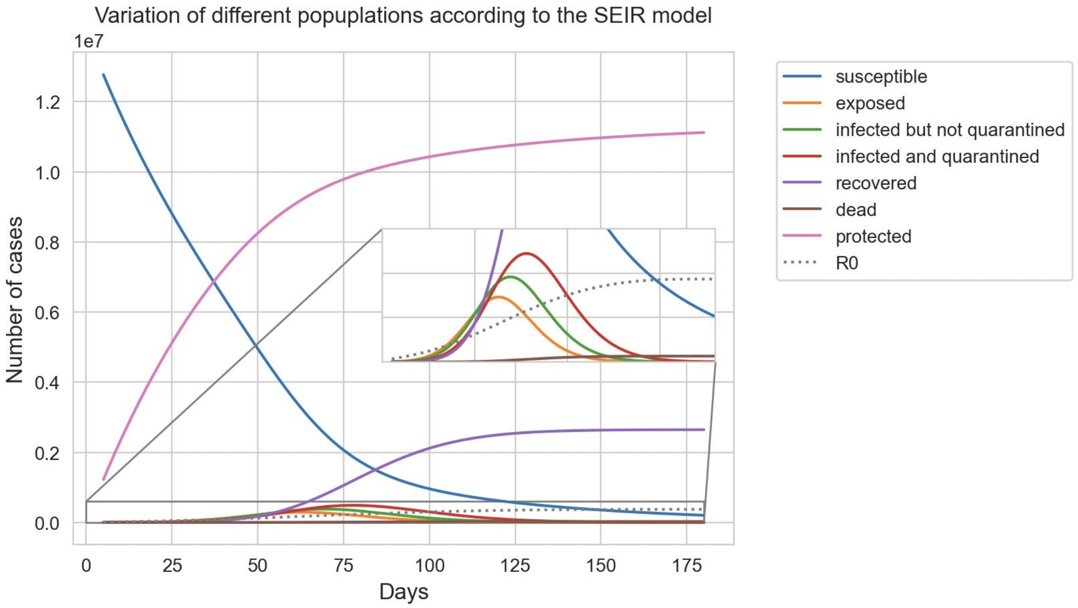


$$\frac{dS(t)}{dt} = -\beta\frac{S(t)\,I(t)}{N} - \alpha S(t)$$
$$\frac{dE(t)}{dt} = \beta\frac{S(t)\,I(t)}{N} - \gamma E(t)$$
$$\frac{dI(t)}{dt} = \gamma E(t) - \delta I(t)$$
$$\frac{dQ(t)}{dt} = \delta I(t) - \lambda Q(t) - \kappa Q(t)$$
$$\frac{dR(t)}{dt} = \lambda Q(t)$$
$$\frac{dD(t)}{dt} = \kappa Q(t)$$
$$\frac{dP(t)}{dt} = \alpha S(t)$$

N: total number of people
α: protection rate
β: infectious rate
γ^{-1}: average incubation time
δ^{-1}: average quarantine time
κ: mortality rate
λ: cure rate

Ranges and values were set according to the paper.

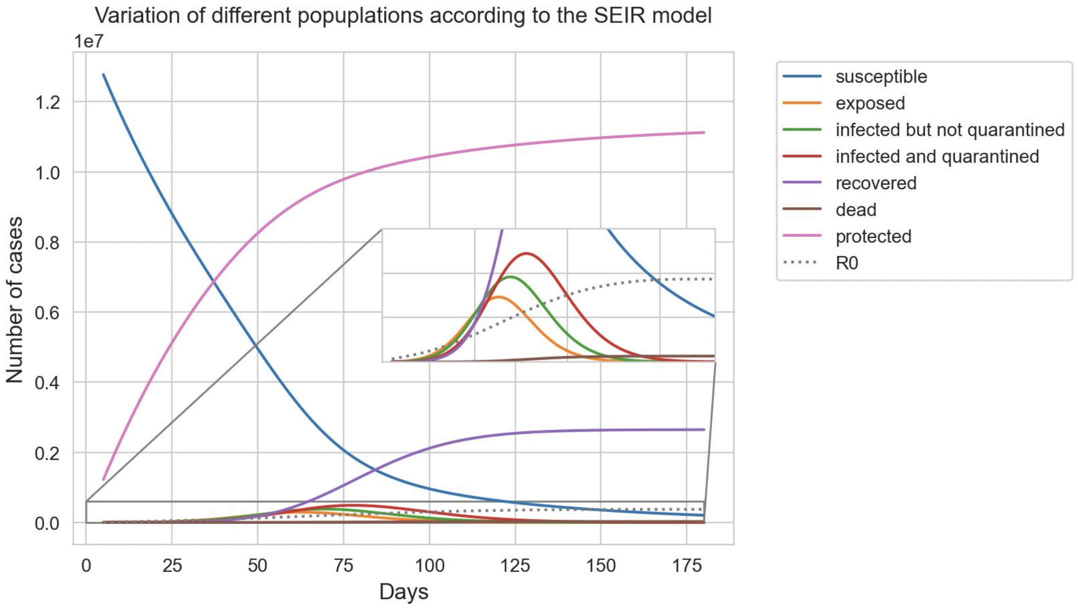
Reproducing the pandemic dynamics



The dynamics are close to the SIR model:

- the **initial** compartment S is slowly emptied
- the **intermediate** compartments E, I, Q present a bell shaped curve
- the **ending** compartments R, D, P are monotone non decreasing functions

Reproducing the pandemic dynamics

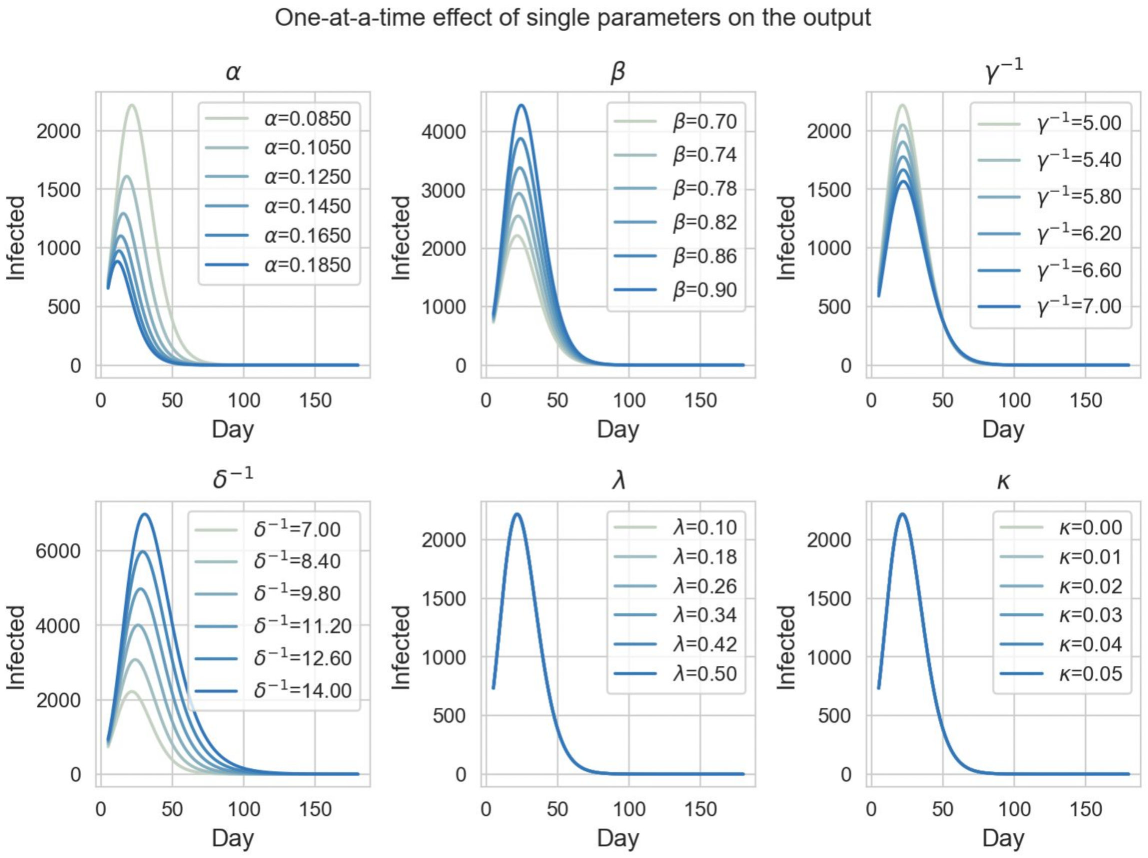


The dynamics are close to the SIR model:

- the **initial** compartment S is slowly emptied
- the **intermediate** compartments E, I, Q present a bell shaped curve
- the **ending** compartments R, D, P are monotone non decreasing functions

Following the approach of the paper, from now on our output of interest will be the number of the **infectious cases I**

Varying parameters one-at-a-time



- With the light-blue parameter combination, the **peak of infectious cases** is around 2220, and is reached around day 16
- The **mortality rate κ** and **cure rate λ** do not seem to be influential on the infected cases
- The other parameters exert a higher influence, especially the **protection rate α**

On with the Global Sensitivity Analysis!

Quick recap of Sobol' indices

Exploiting the decomposition of variance, they quantify how much variability in the output each input accounts for.

First order

$$S_i = \frac{V_i}{V} = \frac{\mathrm{Var}_{\mathbf{x}_{\sim i}}(\mathrm{E}(f(\mathbf{x})|\mathbf{x}_i))}{\mathrm{Var}(f(\mathbf{x}))}$$

Quantifies the **individual** effect of X_i on Y, **without interactions**

Total order

$$S_{T_i} = 1 - \frac{\mathrm{Var}_{\mathbf{x}_{\sim i}}(\mathrm{E}(f(\mathbf{x})|\mathbf{x}_i))}{\mathrm{Var}(f(\mathbf{x}))}$$

Quantifies the **overall** effect of X_i on Y, **including interactions**

Quick recap of Sobol' indices

Exploiting the decomposition of variance, they quantify how much variability in the output each input accounts for.

First order

$$S_i = \frac{V_i}{V} = \frac{\mathrm{Var}_{\mathbf{x}_{\sim i}}(\mathrm{E}(f(\mathbf{x})|\mathbf{x}_i))}{\mathrm{Var}(f(\mathbf{x}))}$$

Total order

$$S_{T_i} = 1 - \frac{\mathrm{Var}_{\mathbf{x}_{\sim i}}(\mathrm{E}(f(\mathbf{x})|\mathbf{x}_i))}{\mathrm{Var}(f(\mathbf{x}))}$$

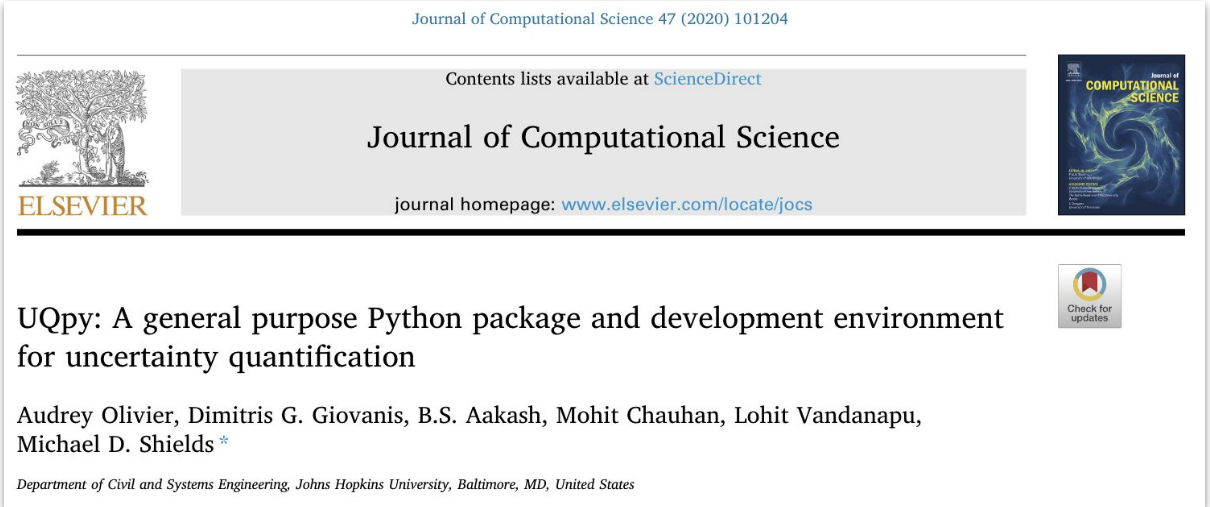
Main goals:

- factor prioritization
- factor fixing

The heavy artillery: UQpy

General purpose Python **toolbox** for modeling uncertainty in physical and mathematical systems, exploiting a black-box approach

Several tools for sensitivity analysis, including computation of **Sobol indexes**


Journal of Computational Science 47 (2020) 101204

Contents lists available at ScienceDirect

Journal of Computational Science

journal homepage: www.elsevier.com/locate/jocs

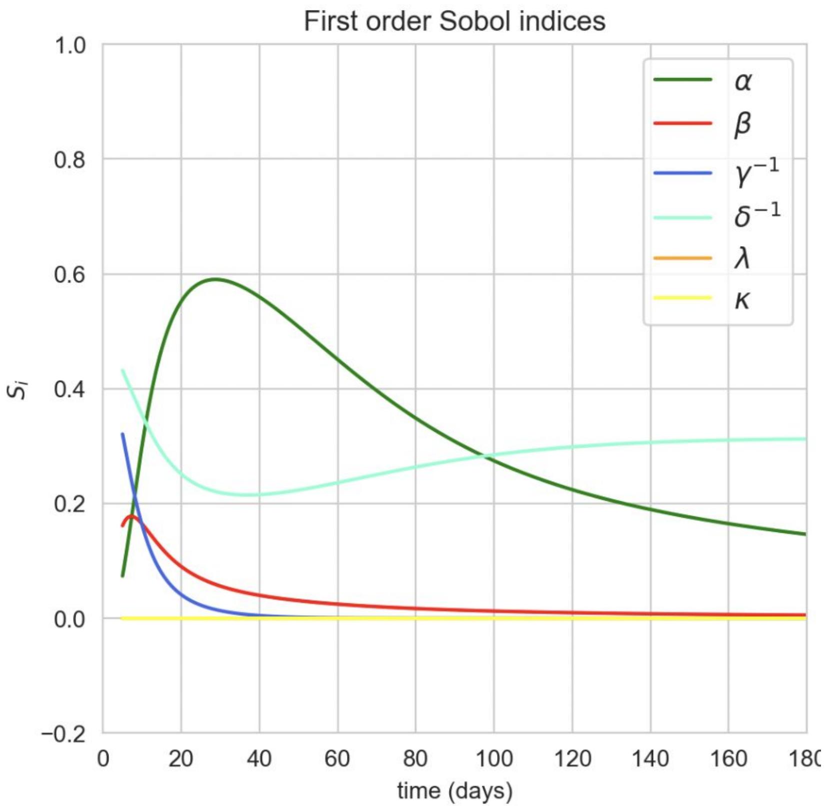
ELSEVIER

UQpy: A general purpose Python package and development environment for uncertainty quantification

Audrey Olivier, Dimitris G. Giovanis, B.S. Aakash, Mohit Chauhan, Lohit Vandanapu, Michael D. Shields *

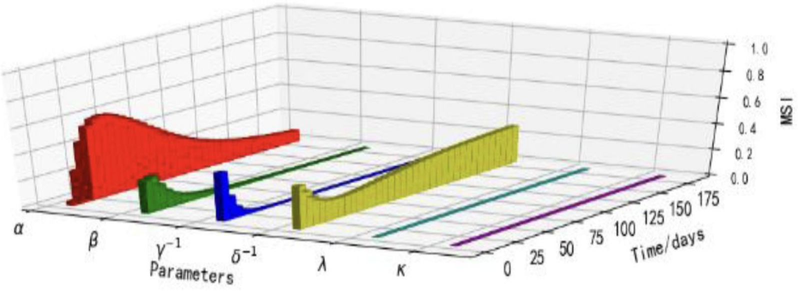
Department of Civil and Systems Engineering, Johns Hopkins University, Baltimore, MD, United States

First-order Sobol' indices on Infectious

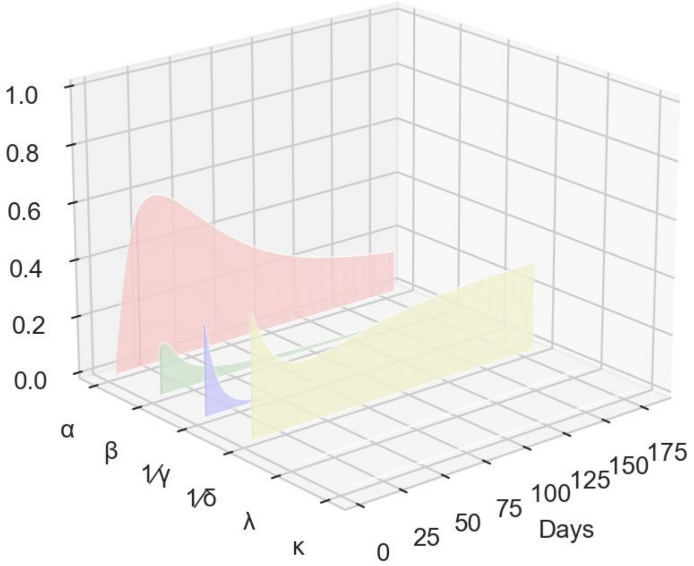


- As expected, the **protection rate α** and the **average quarantine time δ^{-1}** are highly influential on the infectious cases
- The influence of the **infectious rate β** and **average incubation time γ^{-1}** quickly decrease to 0
- The **mortality rate κ** and the **cure rate λ** always remain close to 0

First-order Sobol' indices on Infectious



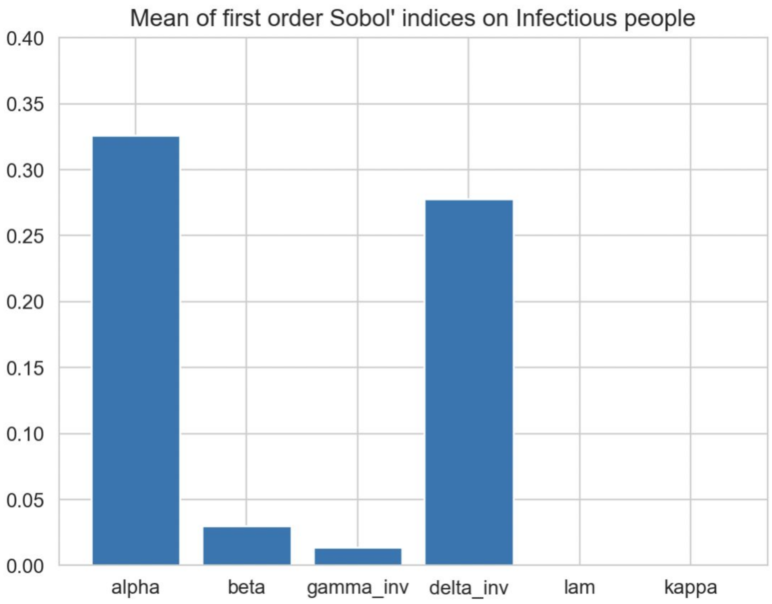
Original results



Reproduced results

Factor prioritization using Sobol' indices

We computed the **mean over time** of the first-order Sobol indexes.

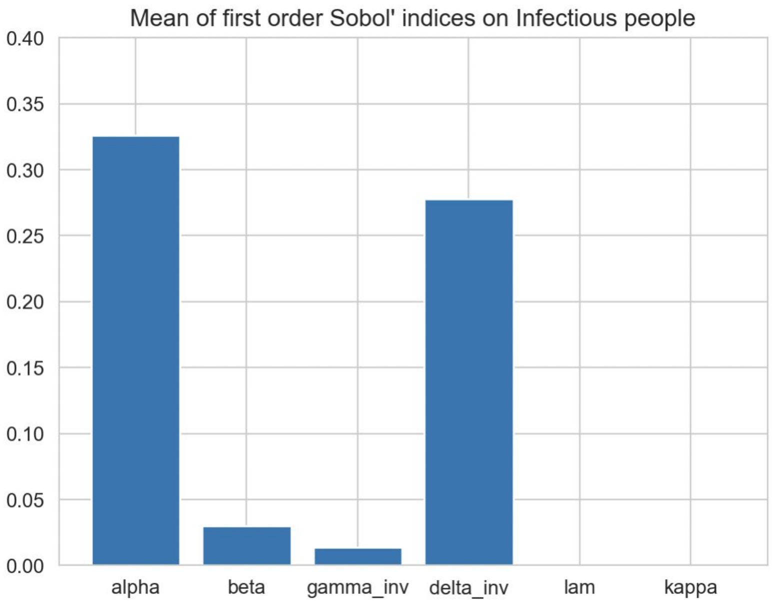


Factor ordered by influence:

1. **α**: protection rate
2. **δ^{-1}**: average quarantine time
3. **β**: infectious rate
4. **γ^{-1}**: average incubation time
5. **κ, λ**: mortality and cure rate

Factor prioritization using Sobol' indices

We computed the **mean over time** of the first-order Sobol indexes.



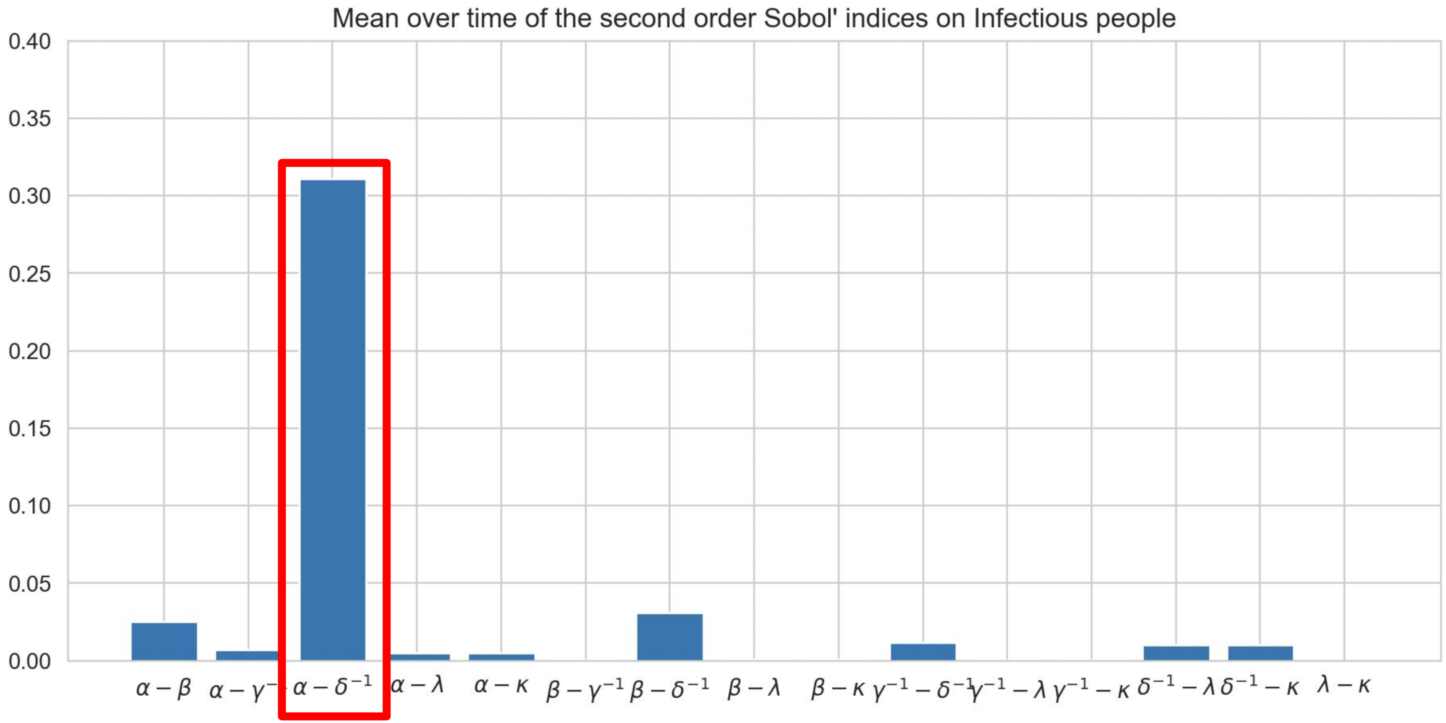
Factor ordered by influence:

1. **α**: protection rate
2. **δ^{-1}**: average quarantine time
3. **β**: infectious rate
4. **γ^{-1}**: average incubation time
5. **κ, λ**: mortality and cure rate

Same conclusions obtained in the paper

Extending the results of the paper: second order interactions

Mean of second-order Sobol indexes



The most influential interaction is between the **protection rate α** and the **average quarantine time δ^{-1}**, while the other interactions are less relevant.

We’re talking about pandemics,
so what about R0?

The basic reproduction number R0

The average number of secondary infected cases generated by a primary case, crucial for transmissibility:

- **R0 < 1** → infection-free steady state is reached asymptotically
- **R0 > 1** → **the pandemic will outbreak**

Here computed as:

$$R0 = \left(1 + \frac{\ln(I(t)/t)}{\gamma}\right)\left(1 + \frac{\ln(I(t)/t)}{\lambda}\right)$$

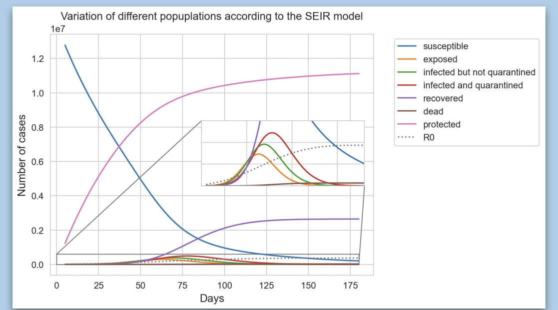
The basic reproduction number R0

The average number of secondary infected cases generated by a primary case, crucial for transmissibility:

- **R0 < 1** → infection-free steady state is reached asymptotically
- **R0 > 1** → **the pandemic will outbreak**

Here computed as:

$$R0 = \left(1 + \frac{\ln(I(t)/t)}{\gamma}\right)\left(1 + \frac{\ln(I(t)/t)}{\lambda}\right)$$



Remember this image?

The values for R0 do not really make sense (**R0 plateauing around 10^6**)

R0: an unsolved mystery

After many attempts and doubts on the formula we **contacted the author**.

He confirmed that everything was correct.

Dear Teo and Flavia,

Sorry for the delayed reply--I was doing fieldwork in a remote mountainous area with no communication signal.

(1) According to Equation 11 and reference 11, Y(t) is the total number of infected people given by I(t), without Q(t).

(2) inside Log.

(3) No, gamma and lambda are variables to be analyzed for their sensitivity.

Hope this reply is clear.

Please hesitate to contact us if you have further questions.

Best regards,

Chunfeng Ma

More investigating on this topic is needed.

Conclusions

- **Main takeaways**
  - In the SEIR model, the **protection rate α** is the most influential on the infectious cases, together with the **quarantine time δ^{-1}**, and their **interaction**
  - Other factors such as the **mortality rate κ** and the **cure rate λ** can be considered as **fixed** for modeling purposes

Conclusions

- **Main takeaways**
  - In the SEIR model, the **protection rate α** is the most influential on the infectious cases, together with the **quarantine time δ^{-1}**, and their **interaction**
  - Other factors such as the **mortality rate κ** and the **cure rate λ** can be considered as **fixed** for modeling purposes

- **Future improvements**
  - **Solving the mystery of R0**
  - Extend the analysis to other **quantities of interest** such as the number of quarantined cases (Q)
  - Apply different methods (*e.g.,* **moment-independent sensitivity measures**)

Thank you for the attention!

References

- **Main reference paper**
  Ma, C., Li, X., Zhao, Z., Liu, F., Zhang, K., Wu, A., & Nie, X. (2022). Understanding Dynamics of Pandemic Models to Support Predictions of COVID-19 Transmission: Parameter Sensitivity Analysis of SIR-Type Models. *IEEE Journal of Biomedical and Health Informatics*, *26*(6), 2458-2468. https://doi.org/10.1109/JBHI.2022.3168825

- **UQpy paper**
  Olivier, A., Giovanis, D. G., Aakash, B. S., Chauhan, M., Vandanapu, L., & Shields, M. D. (2020). UQpy: A general purpose Python package and development environment for uncertainty quantification. *Journal of Computational Science*, *47*, 101204. https://doi.org/10.1016/j.jocs.2020.101204

Any questions?

UQpy methods

Several methods to compute the Sobol indexes

- **First-order:**
  - Jensen (2014)
  - **Saltelli (2002)**
  - Sobol (1983)
- **Second-order:**
  - **Saltelli (2002)**
- **Total order:**
  - Homma (1996)
  - **Saltelli (2002)**

Sobol Method

Rationale: for each X_i, we generate 2 x N *i.i.d.* samples X_i,1 and X_i,2, plus N samples of all other input variables, $X_{\sim i}$. We estimate the normalized individual variance on the means and subtract the estimated normalized variance of the mean of the interactions.

$$\bar{f}_{i,1} = \frac{1}{n}\sum_{j=1}^{n} f(\boldsymbol{X}_{i,1,j}, \boldsymbol{X}_{\sim i,j})$$

$$\bar{f}_{i,2} = \frac{1}{n}\sum_{j=1}^{n} f(\boldsymbol{X}_{\sim i,j}, \boldsymbol{X}_{i,2,j})$$

$$\bar{f}_{i,\text{joint}} = \frac{1}{n}\sum_{j=1}^{n} f(\boldsymbol{X}_{i,1,j}, \boldsymbol{X}_{i,2,j}, \boldsymbol{X}_{\sim i,j}).$$

$$\bar{f} = \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{X}_i)$$

$$s_f^2 = \frac{1}{n-1}\sum_{i=1}^{n}(f(\boldsymbol{X}_i) - \bar{f})^2$$

$$\boxed{S_i = \frac{\mathrm{Var}_{\boldsymbol{X}_{i,1}}(\bar{f}_{i,1}) + \mathrm{Var}_{\boldsymbol{X}_{i,2}}(\bar{f}_{i,2})}{2s_f^2} - \frac{\mathrm{Var}_{\boldsymbol{X}_{i,1},\boldsymbol{X}_{i,2}}(\bar{f}_{i,\text{joint}})}{2s_f^2}}$$

Sobol method

Computational strategy

$$S_i = \frac{\mathbb{V}\left[E\left(Y \mid X_i\right)\right]}{\mathbb{V}(Y)} = \frac{(1/N) Y_A \cdot Y_{C_i} - f_0^2}{(1/N) Y_A \cdot Y_A - f_0^2}$$

$$y_A = f(A), \quad y_{C_i} = f(C_i), \quad f_0^2 = \left(\frac{1}{N}\sum_{j=1}^{N} y_A^{(j)}\right)^2$$