Uncertainty Quantification in Mechanistic Epidemic Models

Americo Cunha Jr

Rio de Janeiro State University — UERJ

XLII Congresso Nacional de Matemática Aplicada e Computacional

Bonito - MS, 18 a 22 de setembro de 2023

@AmericoCunhaJr @AmericoCunhaJr @AmericoCunhaJr

Acknowledgments

- Collaborators

David A. W. Barton

Thiago G. Ritto

- Institutional and financial support

UERJ – Universidade do Estado do Rio de Janeiro

University of Bristol

UFRJ

CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico

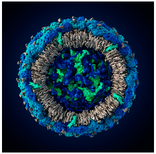
CAPES

FAPERJ – Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro

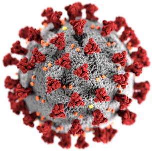
EPSRC – Engineering and Physical Sciences Research Council

Epidemics are recurrent in history ...

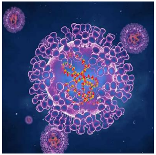
Dengue



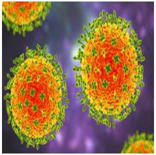
Zika virus



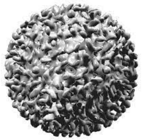
COVID-19



Monkeypox



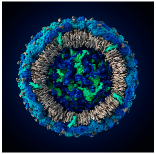
Langya



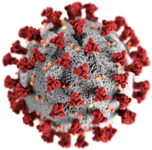
Next ?

Epidemics are recurrent in history ...

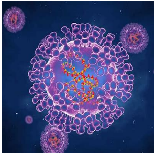
Dengue



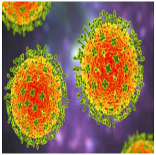
Zika virus



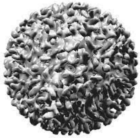
COVID-19



Monkeypox

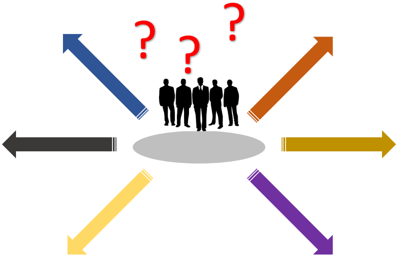


Langya



Next ?

They may cause a lot of trouble!



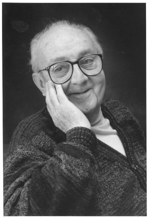
Computational Model + Real World Data = Tool to aid with Decision Making

Mathematical models are approximations!

"All models are wrong but some are useful"

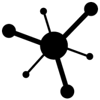
George E. P. Box

Mathematical models are approximations!



"All models are wrong but some are useful"

George E. P. Box

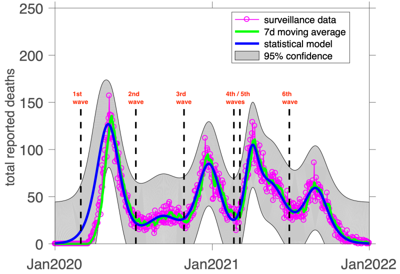
 × 

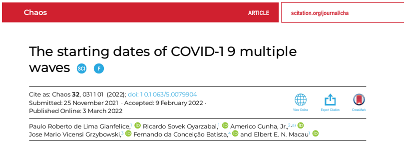
This idea has a more pronounced

meaning in computational epidemiology than in physics!

"First principles of epidemiology are unknown!"

$$I(t) = \sum_{i=1}^{N} \frac{r_i\, K_i\, e^{-r_i\,(t-\tau_i)}}{\left(1 + e^{-r_i\,(t-\tau_i)}\right)^2}$$




Chaos ARTICLE scitation.org/journal/cha

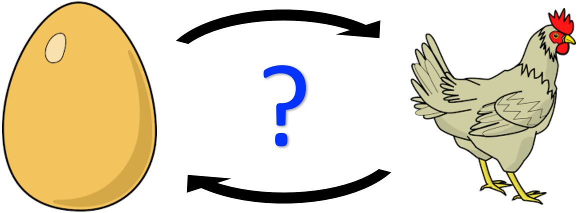
The starting dates of COVID-19 multiple waves

Cite as: Chaos **32**, 031101 (2022); doi: 10.1063/5.0079904
Submitted: 25 November 2021 · Accepted: 9 February 2022 ·
Published Online: 3 March 2022

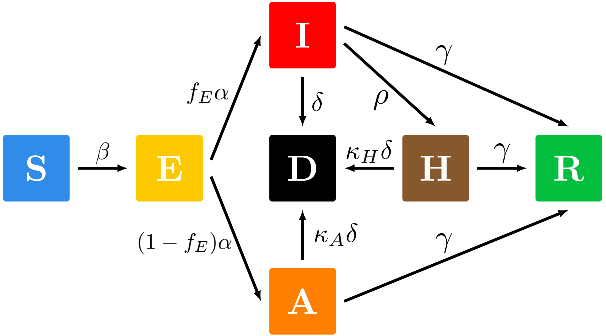
Paulo Roberto de Lima Gianfelice,[1] Ricardo Sovek Oyarzabal,[1] Americo Cunha, Jr.,[2,a)] Jose Mario Vicensi Grzybowski,[3] Fernando da Conceição Batista,[4] and Elbert E. N. Macau[1]

P. R. L. Gianfelice et al., *The starting dates of COVID-19 multiple waves*, **Chaos**, 32:031101, 2022

Mechanistic models are necessary!
(cause-effect relationship)



SEIR(+AHD) mechanistic epidemic model



EPIDEMIC
EPIDEMIOLOGY EDUCATIONAL CODE

- Dynamic model:

$$
\begin{aligned}
\dot{S} &= -\beta(t)\, S\, (I + A + \epsilon_H\, H)/N \\
\dot{E} &= \beta(t)\, S\, (I + A + \epsilon_H\, H)/N - \alpha\, E \\
\dot{I} &= f_E\, \alpha\, E - (\gamma + \rho + \delta)\, I \\
\dot{R} &= \gamma\, (I + A + H) \\
\dot{A} &= (1 - f_E)\, \alpha\, E - (\kappa_A\, \delta + \gamma) \\
\dot{H} &= \rho\, I - (\gamma + \kappa_H\, \delta)\, H \\
\dot{D} &= \delta\, (I + \kappa_A\, A + \kappa_H\, H) \\
\dot{N} &= -\dot{D}
\end{aligned}
$$

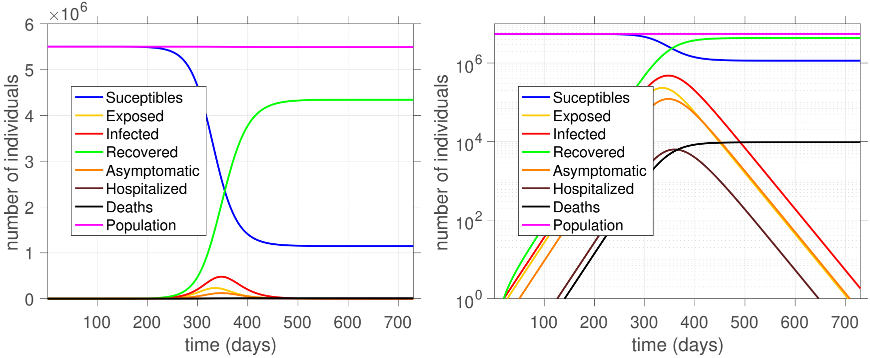
+ initial conditions

- Rate of transmission:

$$\beta(t) = \beta_0 + \frac{(\beta_\infty - \beta_0)}{2} \left(1 + \tanh\left(\eta\, \frac{(t - t_\beta)}{2} \right) \right)$$

SEIR(+AHD) model response (virgin population)

In this example the transmission rate is fixed, i.e., $\beta(t) = \beta$



- Dynamic model:

$$\frac{d}{dt}\boldsymbol{u}(t) = F(t,\ \boldsymbol{u}(t),\ \boldsymbol{x})$$

- Vector field:

$$(t,\ \boldsymbol{u}(t),\ \boldsymbol{x}) \in \mathbb{R} \times \mathbb{R}^8 \times \mathbb{R}^{12} \mapsto F(t,\ \boldsymbol{u}(t),\ \boldsymbol{x}) \in \mathbb{R}^8$$

- State vector:

$$\boldsymbol{u}(t) = (S, E, I, A, H, R, D, N)$$

- Parameters vector:

$$\boldsymbol{x} = (\beta_0, \alpha, f_E, \gamma, \rho, \delta, \kappa_A, \kappa_H, \epsilon_H, \beta_\infty, \eta, t_\beta)$$

- QoI vector:

$$\boldsymbol{y} = [H(t_1),\ \cdots\ ,\ H(t_n),\ D(t_1),\ \cdots\ ,\ H(t_n)]$$

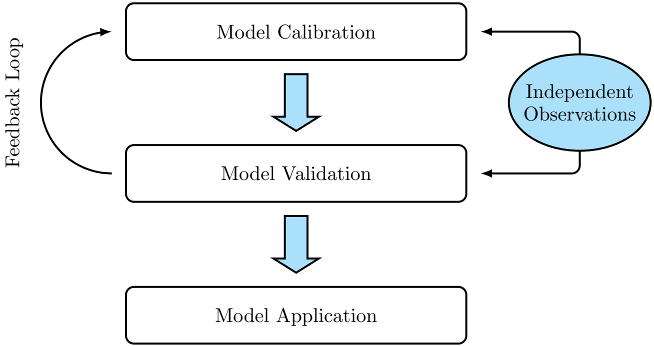
- Computational model:

$$\boldsymbol{y} = \mathcal{M}(\boldsymbol{x})$$

Model Parametrization: expert knowledge, literature, ansatz, etc

Model Calibration: data + model inversion

Model Validation: other data + error metrics



How to obtain consistent initial conditions ?

Initial conditions (almost surely) are unknown for epidemic systems!

Initial conditions (almost surely) are unknown for epidemic systems!

Ansatz $\longrightarrow$ biased and inconsistent

How to obtain consistent initial conditions ?

Initial conditions (almost surely) are unknown for epidemic systems!

Ansatz	$\longrightarrow$	biased and inconsistent
Inversion	$\longrightarrow$	hard and expensive

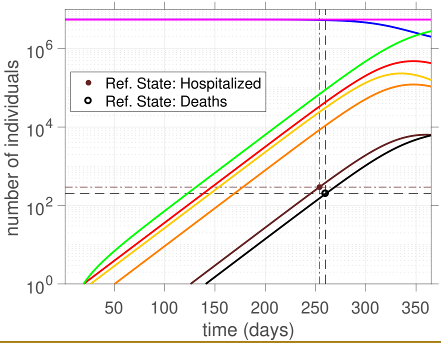
Initial conditions (almost surely) are unknown for epidemic systems!

Ansatz $\longrightarrow$ biased and inconsistent

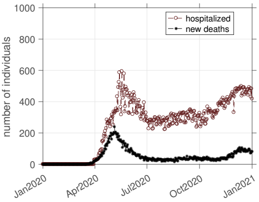
Inversion $\longrightarrow$ hard and expensive

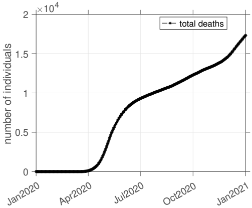
Dynamic inspired $\longrightarrow$ $\boldsymbol{u}(0) = \omega\, \boldsymbol{u}(0)^{H}_{ref} + (1-\omega)\, \boldsymbol{u}(0)^{D}_{ref}$

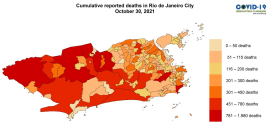
"convex combination of admissible reference states"



- Surveillance data (Rio de Janeiro city):





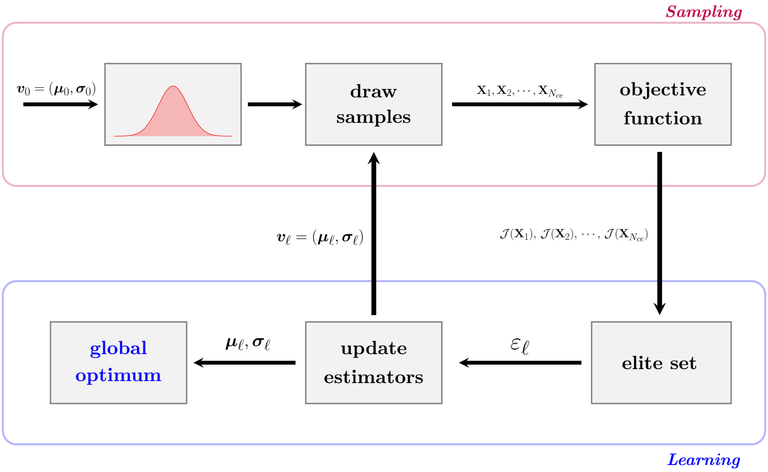


- Data partition:

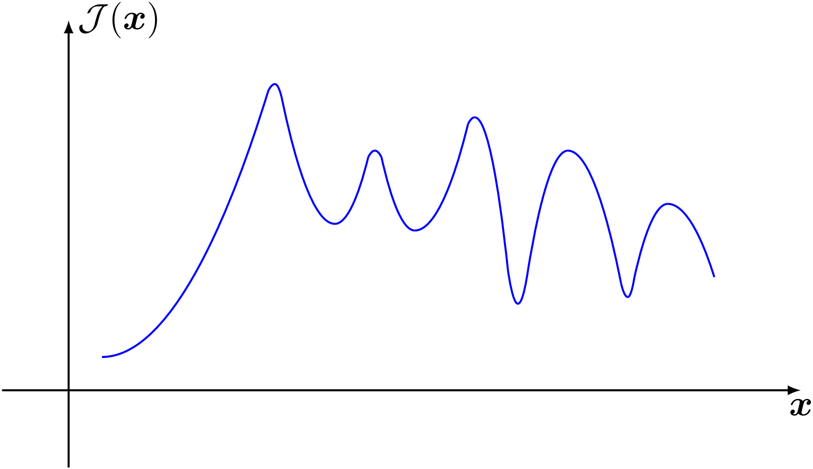
$$\mathbf{y}_{data} = [\mathbf{y}^H_{data} \quad \mathbf{y}^D_{data}]$$

- Misfit function:

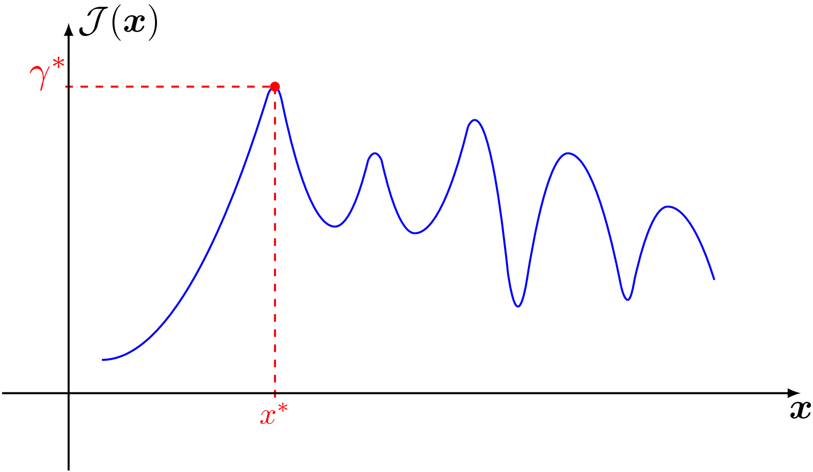
$$\mathcal{J}(\mathbf{x}) = \omega \, \frac{||\mathbf{y}^H_{data} - \mathbf{y}^H(\mathbf{x})||^2}{||\mathbf{y}^H_{data}||^2} + (1-\omega) \, \frac{||\mathbf{y}^D_{data} - \mathbf{y}^D(\mathbf{x})||^2}{||\mathbf{y}^D_{data}||^2}$$

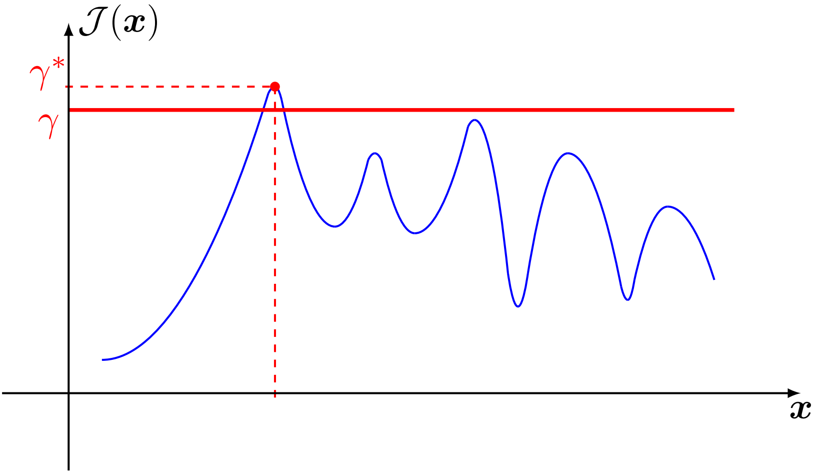

Sampling
$v_0 = (\mu_0, \sigma_0)$
draw samples
$X_1, X_2, \cdots, X_{N_{ce}}$
objective function
$v_\ell = (\mu_\ell, \sigma_\ell)$
$\mathcal{J}(X_1), \mathcal{J}(X_2), \cdots, \mathcal{J}(X_{N_{ce}})$
global optimum
μ_ℓ, σ_ℓ
update estimators
$\mathcal{E}_\ell$
elite set
Learning

Cross-entropy method in action!

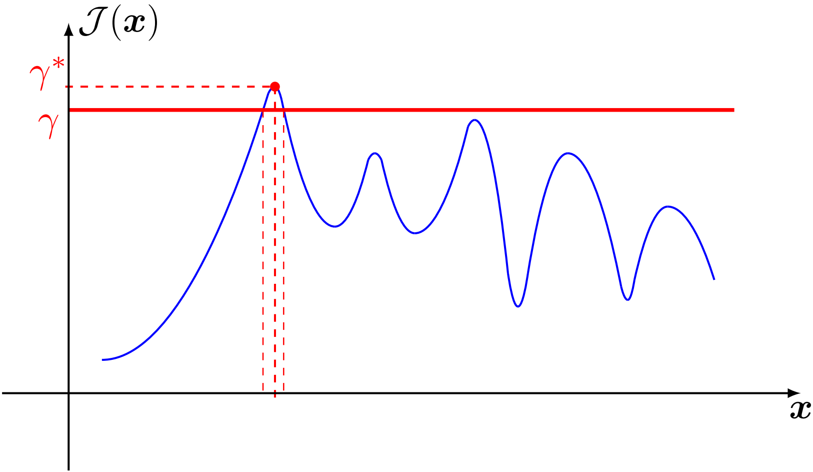


Cross-entropy method in action!

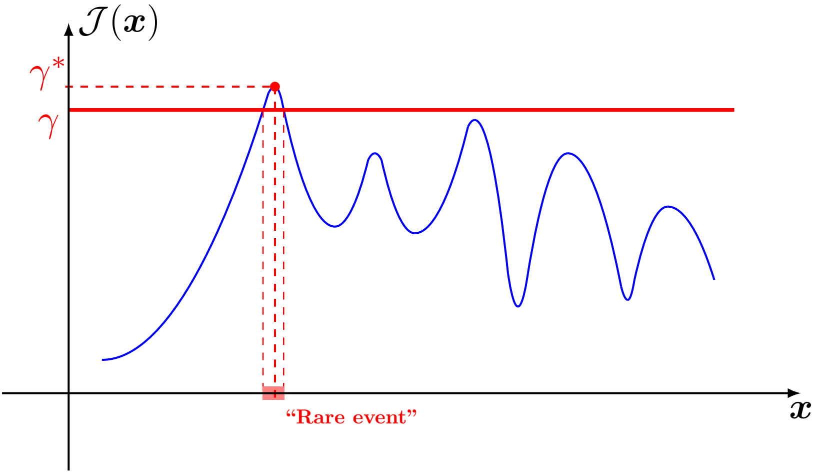


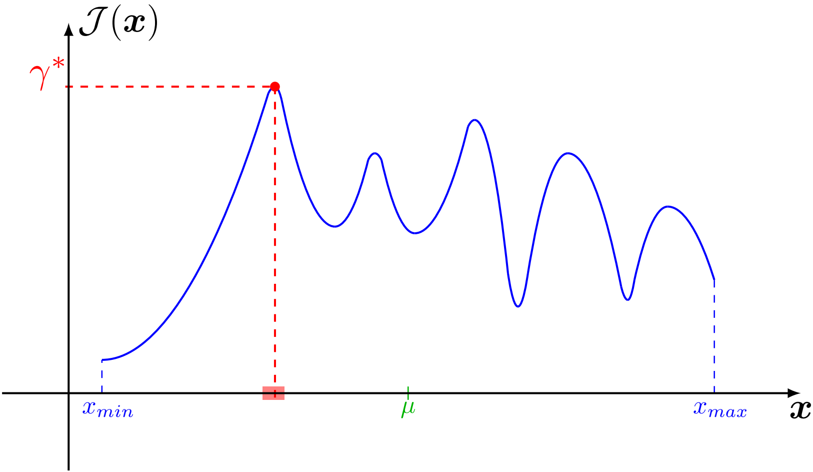


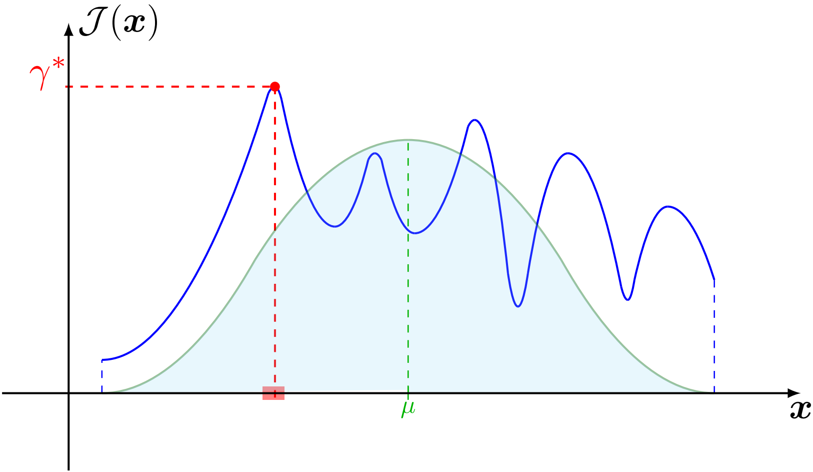
$\mathcal{J}(\boldsymbol{x})$
γ^*
γ
$\boldsymbol{x}$


$\mathcal{J}(\boldsymbol{x})$
γ^*
γ
$\boldsymbol{x}$

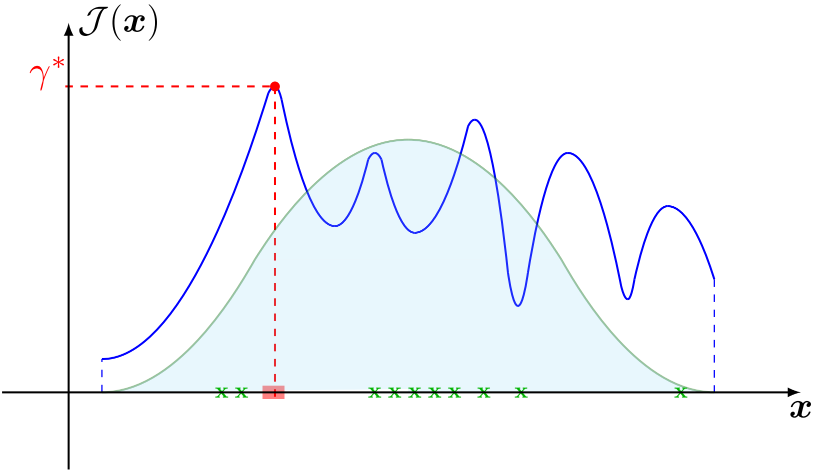
Cross-entropy method in action!



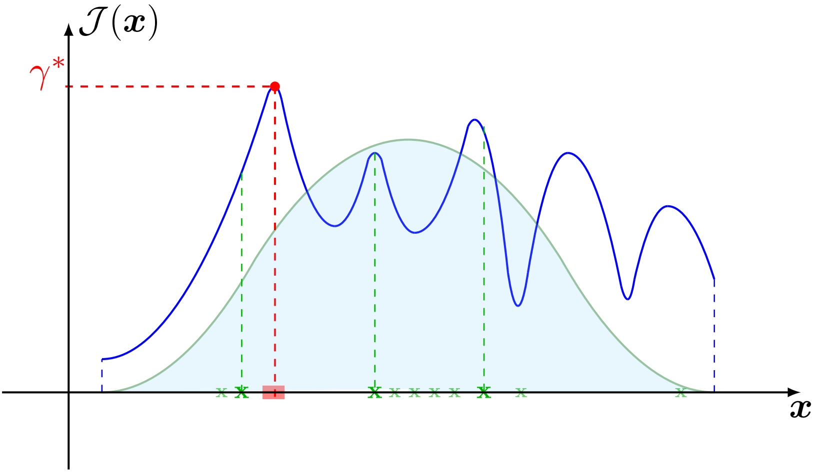

$\mathcal{J}(\boldsymbol{x})$
γ^*
x_{min}
μ
x_{max}
$\boldsymbol{x}$


$\mathcal{J}(\boldsymbol{x})$
γ^*
μ
$\boldsymbol{x}$

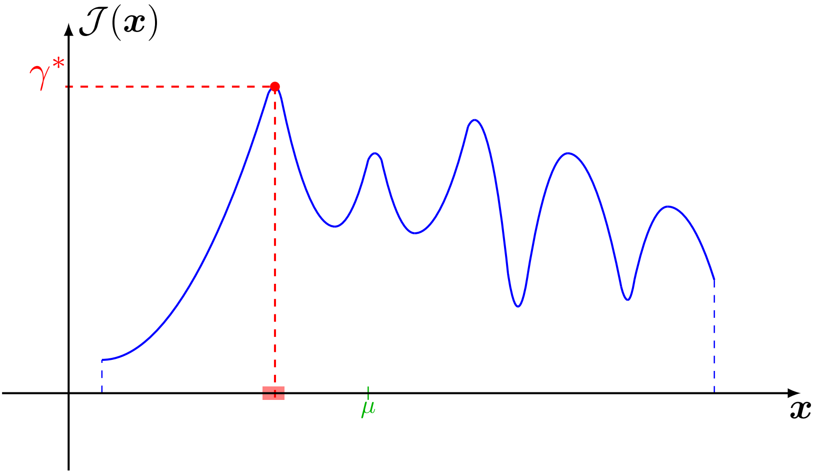
Cross-entropy method in action!

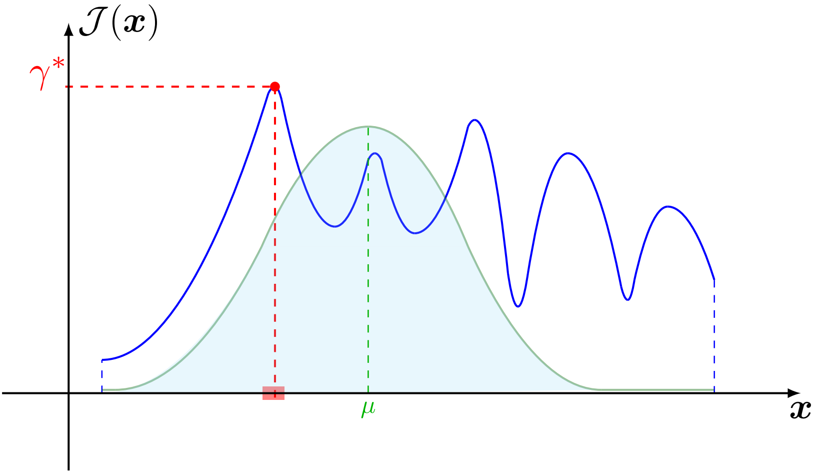


Cross-entropy method in action!

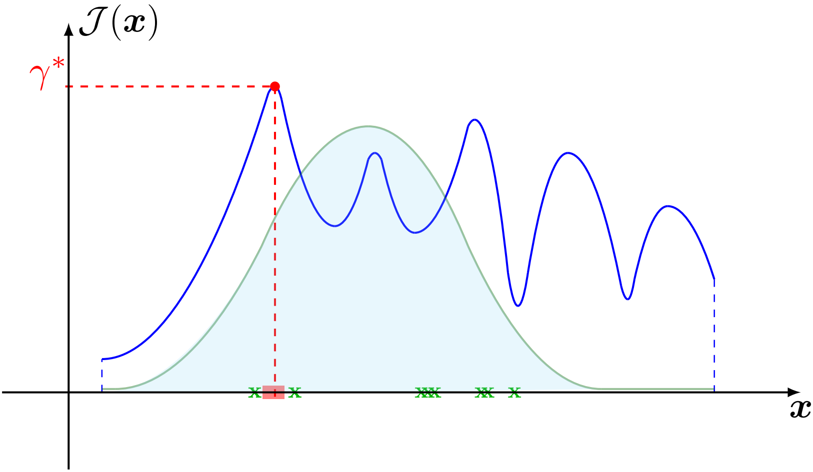


Cross-entropy method in action!

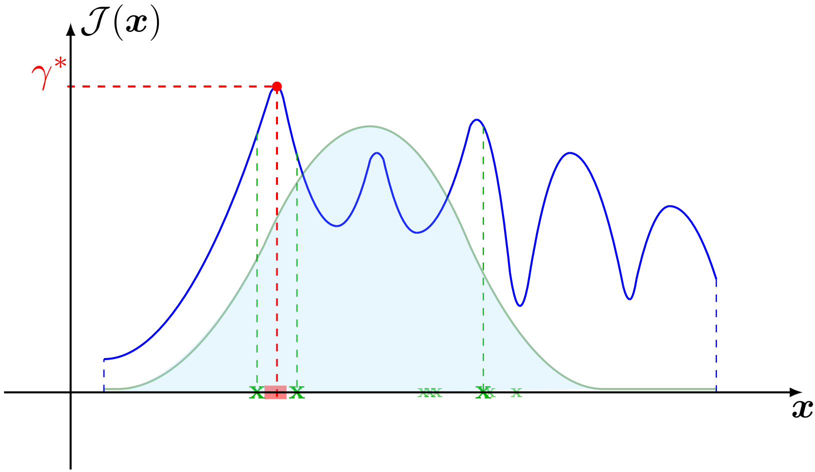


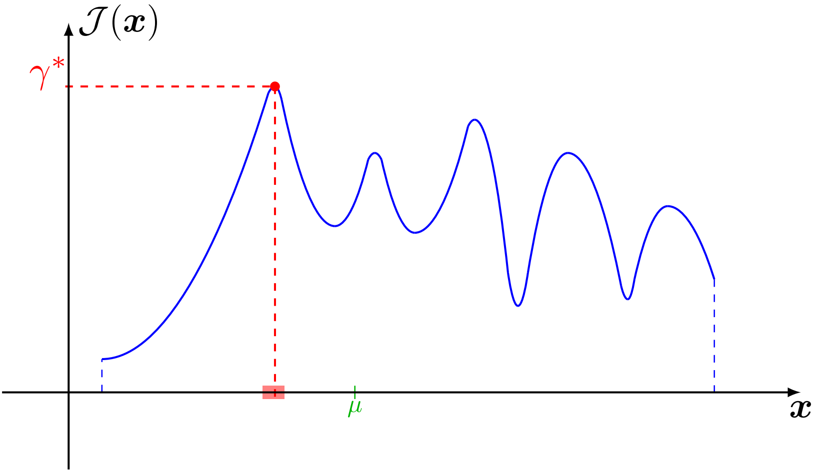

$\mathcal{J}(\boldsymbol{x})$
γ^*
μ
$\boldsymbol{x}$

Cross-entropy method in action!

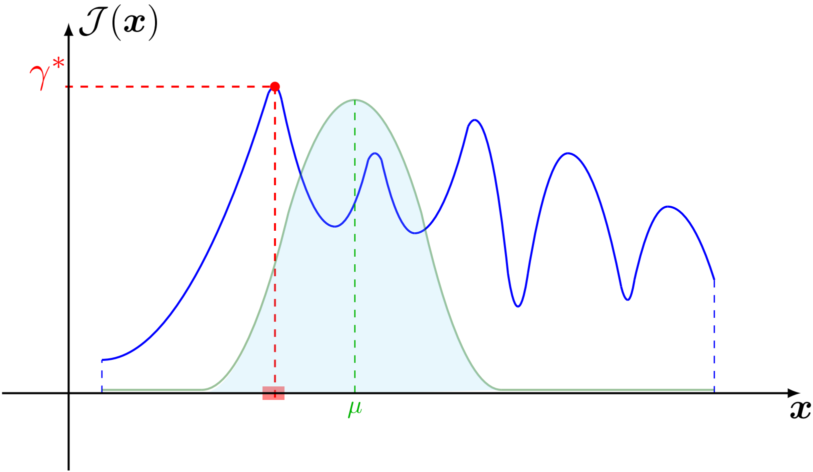


Cross-entropy method in action!

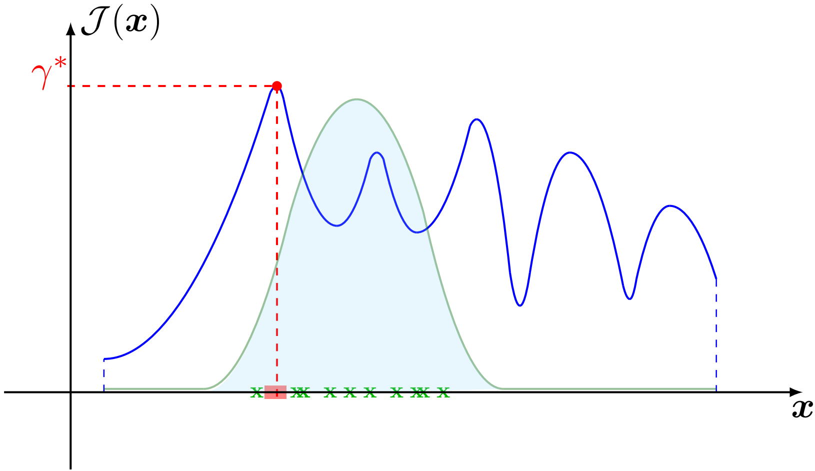


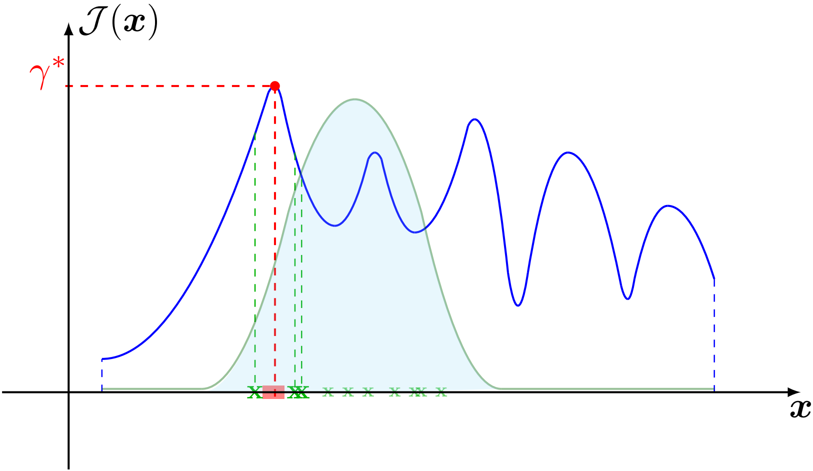


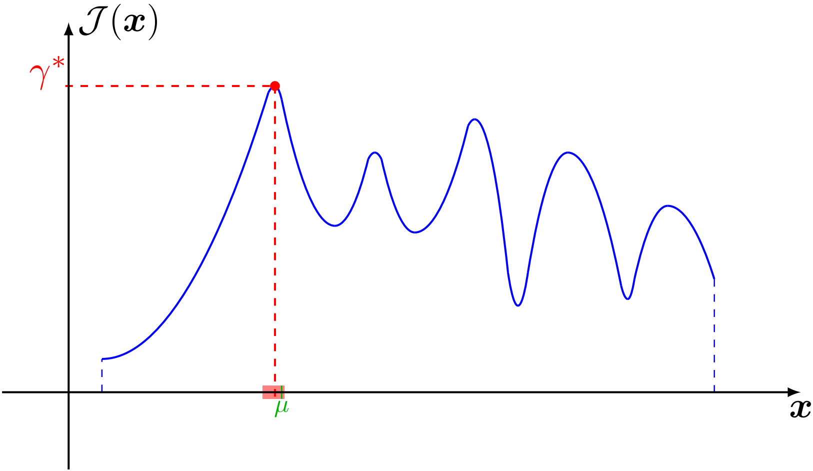
$\mathcal{J}(\boldsymbol{x})$
γ^*
μ
$\boldsymbol{x}$


$\mathcal{J}(\boldsymbol{x})$
γ^*
μ
$\boldsymbol{x}$

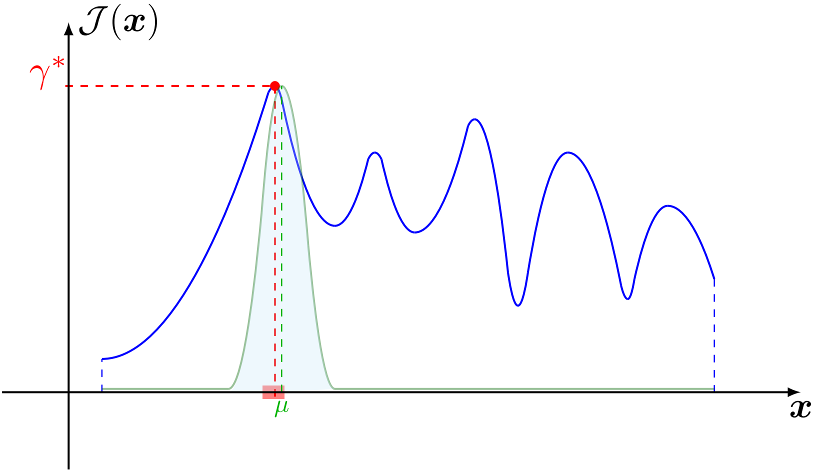
Cross-entropy method in action!

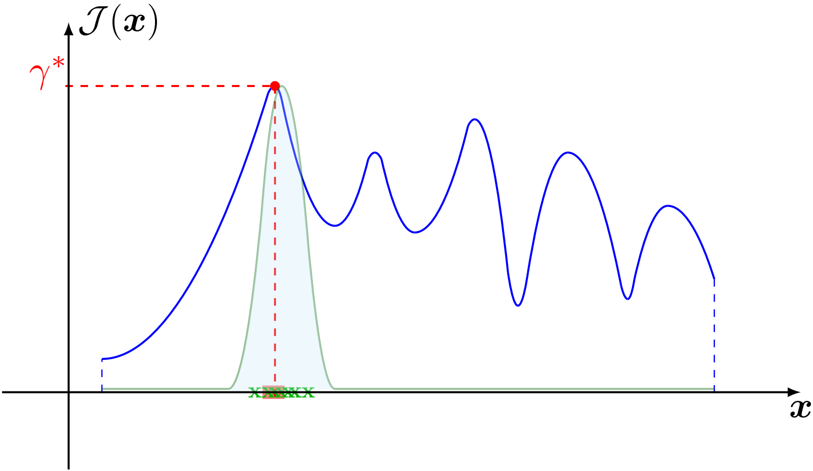


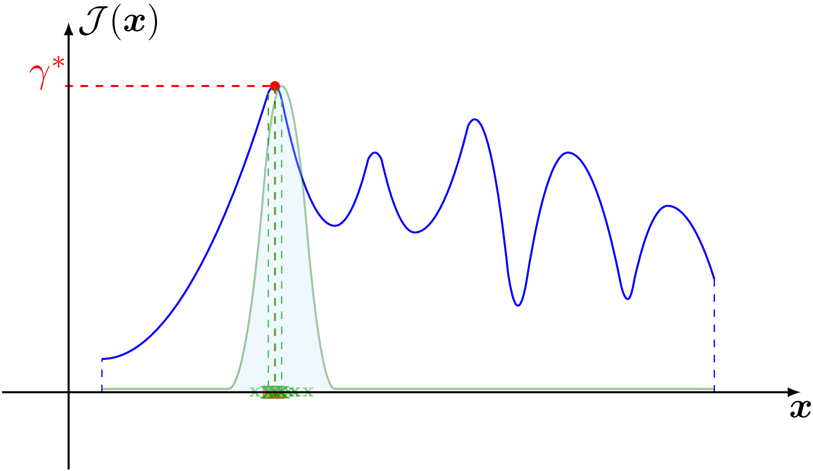

$\mathcal{J}(\boldsymbol{x})$
γ^*
$\boldsymbol{x}$



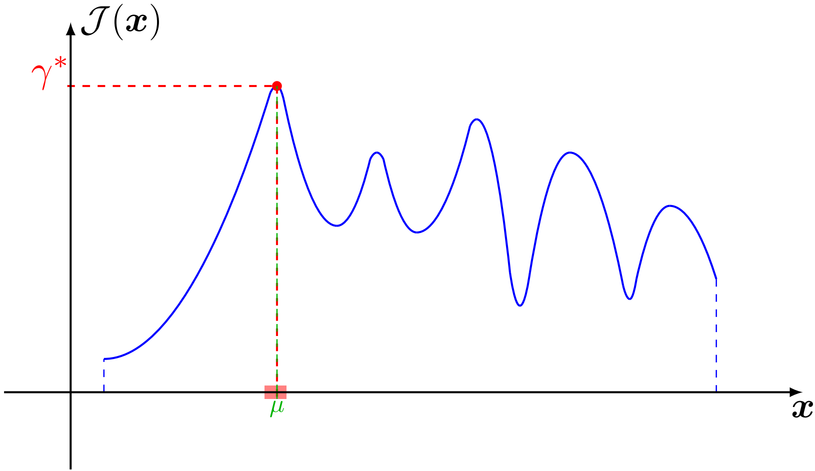
$\mathcal{J}(\boldsymbol{x})$
γ^*
μ
$\boldsymbol{x}$


$\mathcal{J}(\boldsymbol{x})$
γ^*
μ
$\boldsymbol{x}$

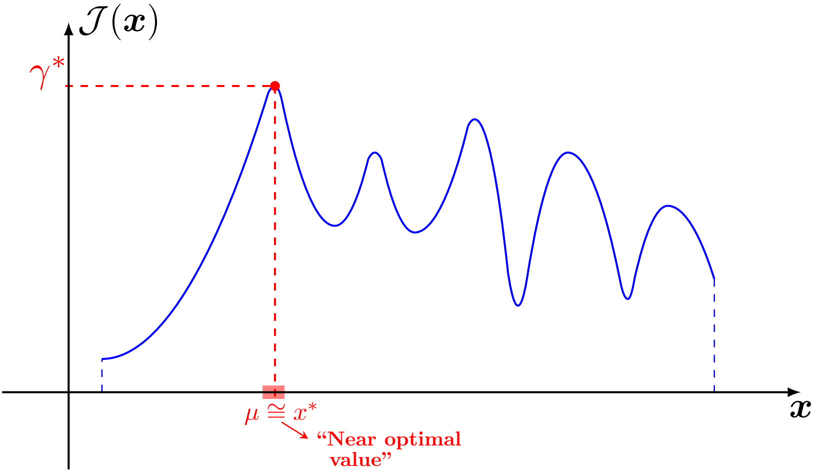

$\mathcal{J}(\boldsymbol{x})$
γ^*
$\boldsymbol{x}$


$\mathcal{J}(\boldsymbol{x})$
γ^*
$\boldsymbol{x}$

Cross-entropy method in action!



Cross-entropy method in action!

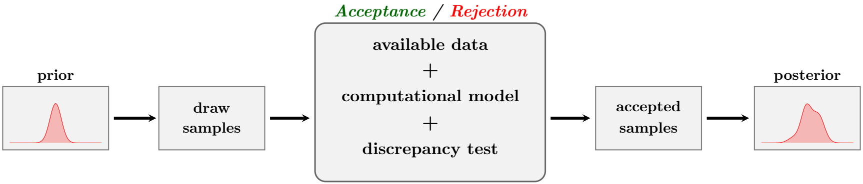


Bayesian Inference

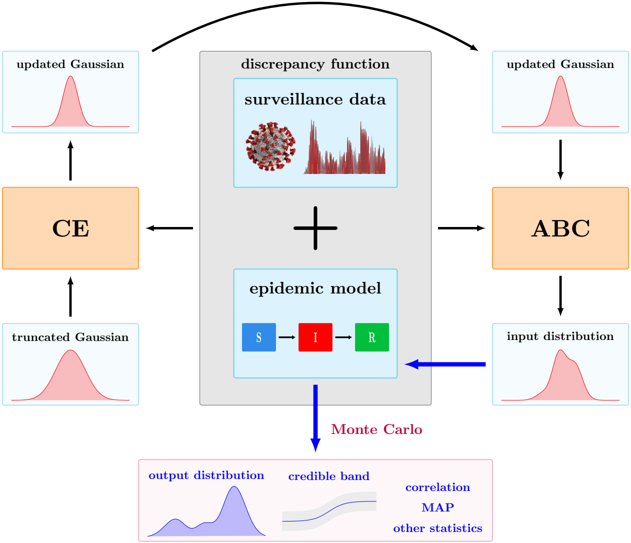
- Traditional Bayesian Computation

$$\underbrace{\pi\big(\text{model} \,\big|\, \text{data}\big)}_{\textit{posterior}} \propto \underbrace{\pi\big(\text{data} \,\big|\, \text{model}\big)}_{\textit{likelihood}} \times \underbrace{\pi\big(\text{model}\big)}_{\textit{prior}}$$

- Approximate Bayesian Computation

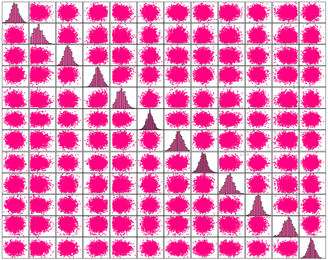


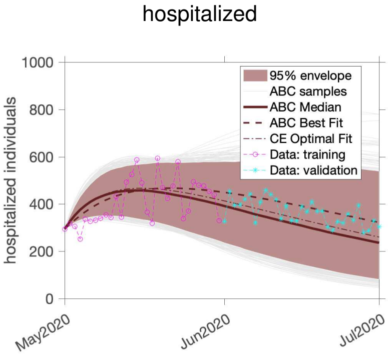
Cross-Entropy Approximate Bayesian Computation

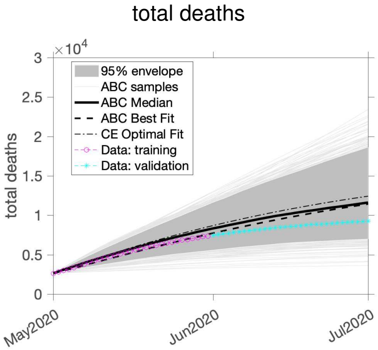


Calibration of the SEIR(+AHD)$_\beta$ model parameters

	Unit	CE Optimal	CE std dev	ABC Best	ABC std dev
β or β_0	1/day	0.12	0.02	0.13	0.02
α	1/day	0.20	0.07	0.27	0.06
f_E	—	0.81	0.03	0.84	0.03
γ	1/day	0.13	0.01	0.12	0.01
ρ	1/day	0.0006	0.0001	0.0005	0.0001
δ	1/day	0.0021	0.0004	0.0015	0.0004
κ_A	—	0.0026	0.0008	0.0027	0.0008
κ_H	—	0.0563	0.0130	0.0575	0.0128
ϵ_H	—	0.25	0.07	0.33	0.07
β_∞	1/day	0.31	0.06	0.43	0.06
η	1/day	5.8	1.9	6.2	1.8
τ_β	day	146	7	153	7



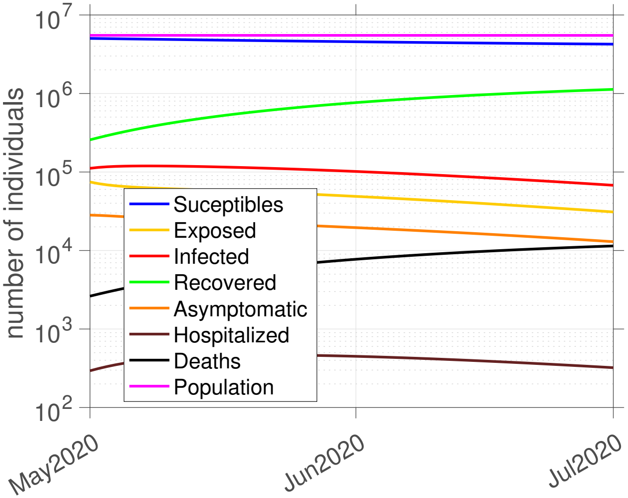


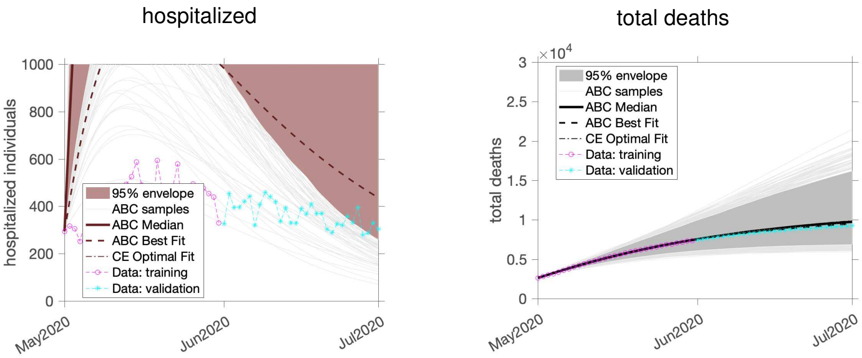


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

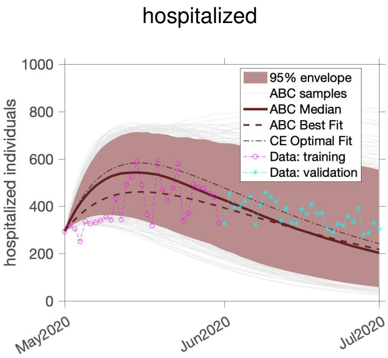
Acceptance rate: 87%

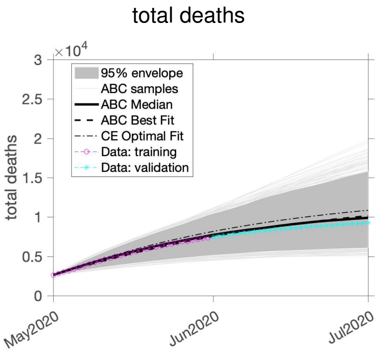
Model exploration of latent variables





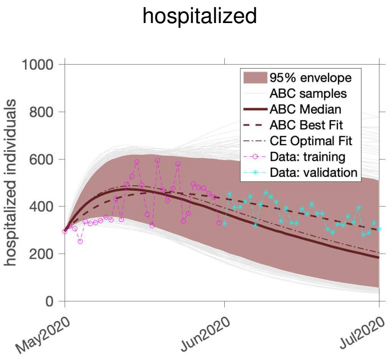
$\omega = 0.00 \quad N_{ce} = 100 \quad N_{abc} = 2000$

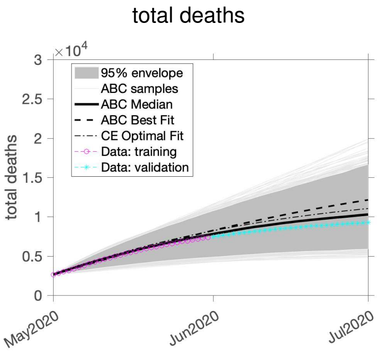




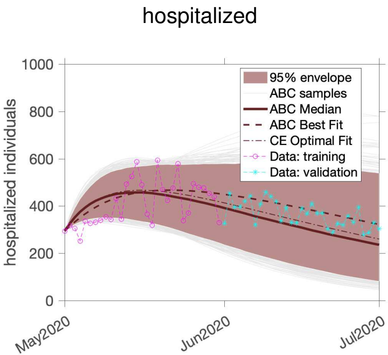
$\omega = 0.25 \quad N_{ce} = 100 \quad N_{abc} = 2000$

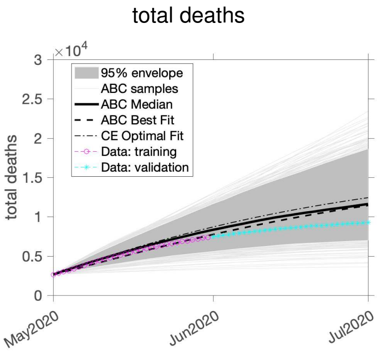
Influence of the weight parameter ω



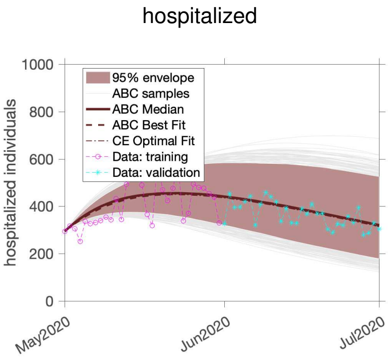


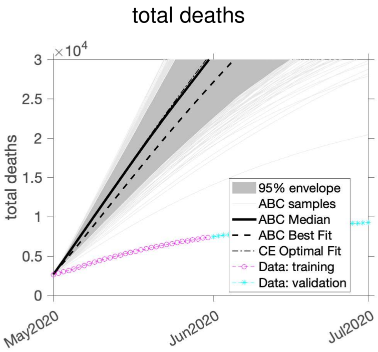
$\omega = 0.50 \quad N_{ce} = 100 \quad N_{abc} = 2000$



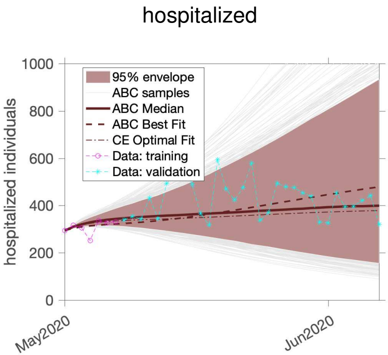


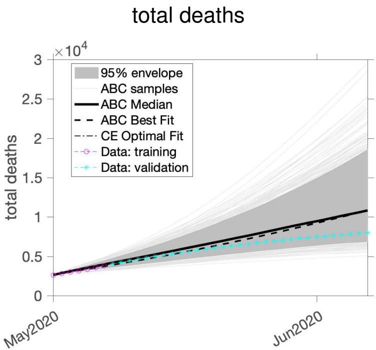
$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$





$\omega = 1.00 \quad N_{ce} = 100 \quad N_{abc} = 2000$

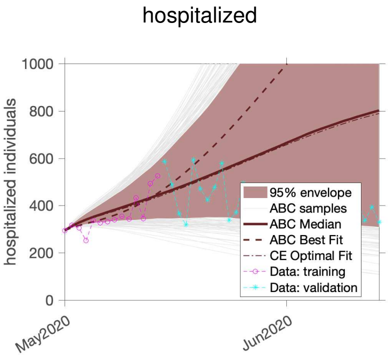


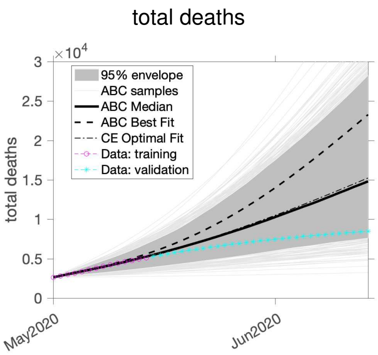


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - May 7, 2020

Influence of the dataset size

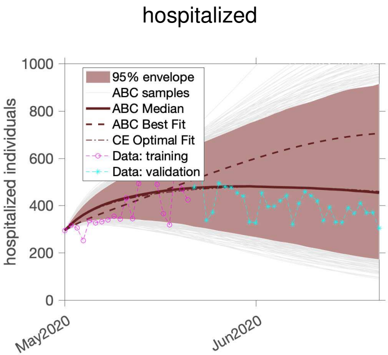


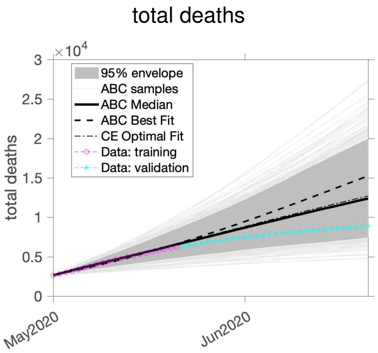


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - May 14, 2020

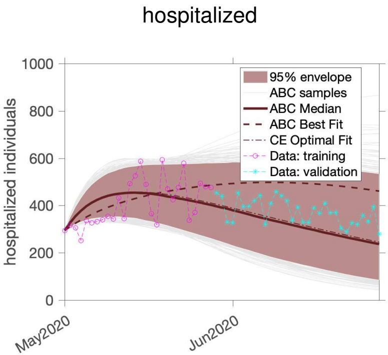
Influence of the dataset size

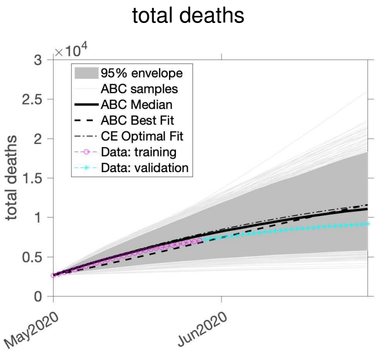




$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - May 21, 2020

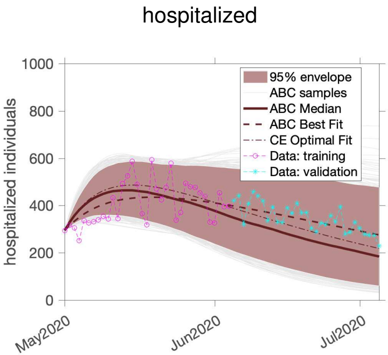


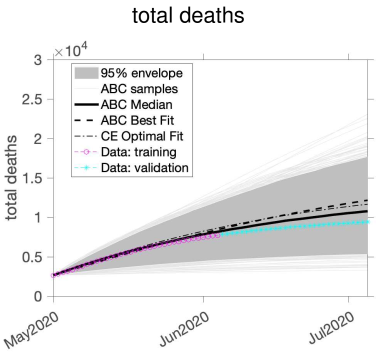


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - May 28, 2020

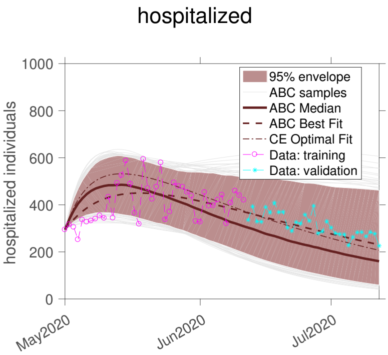
Influence of the dataset size

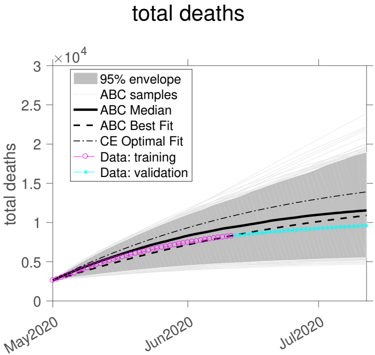




$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

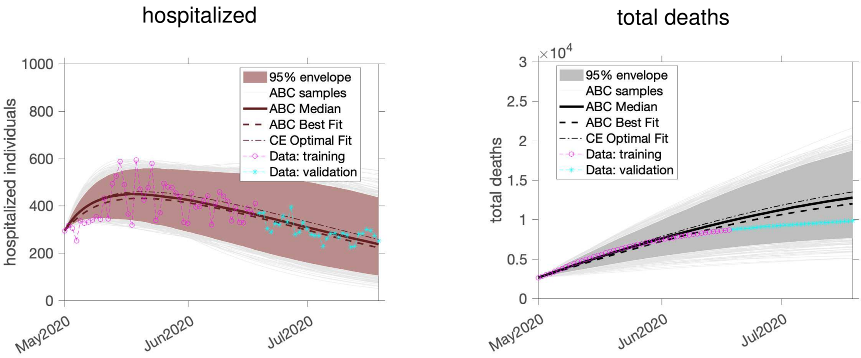
May 1 - June 4, 2020





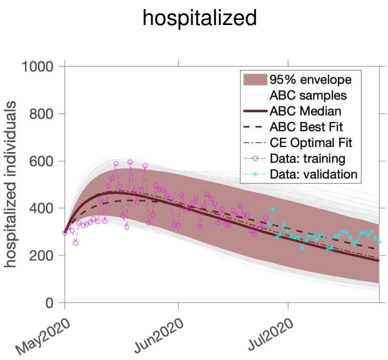
$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

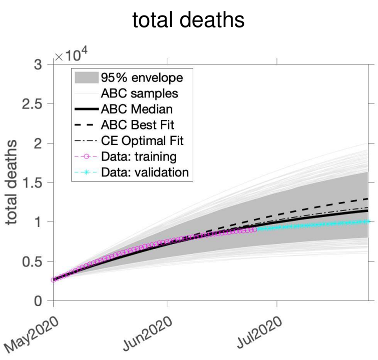
May 1 - June 11, 2020



$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - June 18, 2020

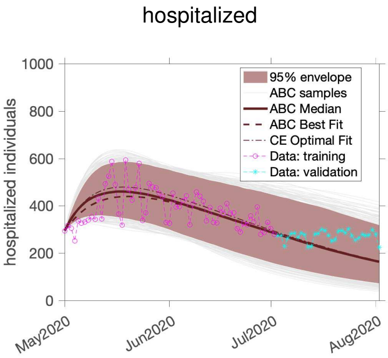


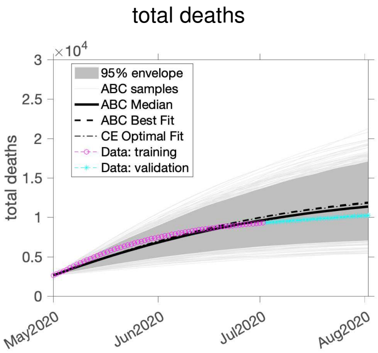


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - June 25, 2020

Influence of the dataset size

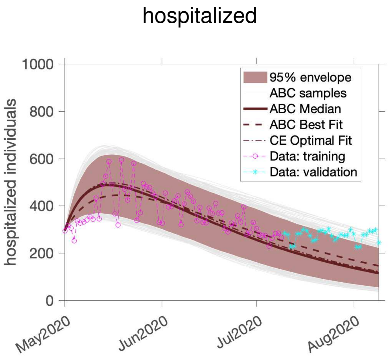


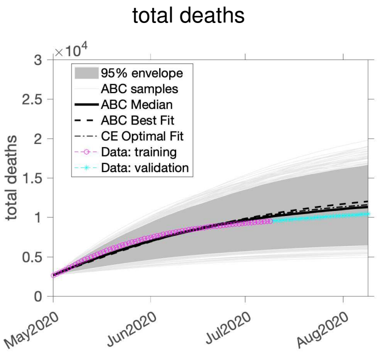


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - July 2, 2020

Influence of the dataset size



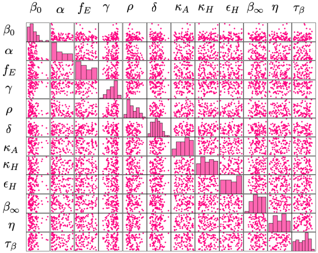


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

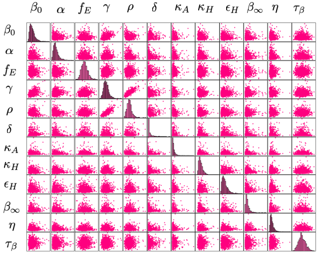
May 1 - July 9, 2020

Comparing CE-ABC and classical ABC

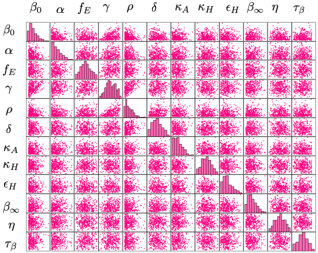
ABC uniform (acceptance < 1% / 100k samples)



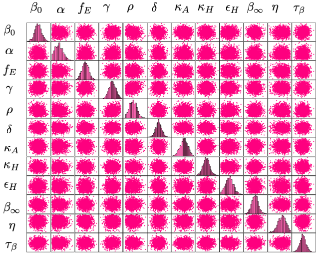
ABC lognormal (acceptance ~ 1% / 100k samples)



ABC truncated Gaussian (acceptance < 1% / 100k samples)

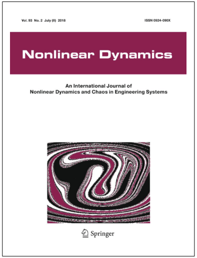


CE-ABC truncated Gaussian (acceptance ~ 87% / 2k samples)



Final remarks

- Data-driven computational models may be a useful tool to aid with decision making during the course of an epidemic outbreak
- A Cross-Entropy Approximate Bayesian Computation (CE-ABC) framework for model calibration and uncertainty quantification is proposed
- CE-ABC + suitable dynamic model + reliable epidemic data may provide a very helpful insights (quantitative and qualitative) about an ongoing outbreak
- Our data-driven COVID-19 dynamic model for Rio de Janeiro:
  - “high-fidelity” predictability horizon: until one-week
  - “reasonable” predictability horizon: until one-month
- Model oriented decisions demand an interdisciplinary panel of experts:
  - Scientific
  - Ethical
  - Humanistic



Nonlinear Dyn (2023) 111:9649–9679
https://doi.org/10.1007/s11071-023-08327-8

Check for updates

ORIGINAL PAPER

Uncertainty quantification in mechanistic epidemic models via cross-entropy approximate Bayesian computation

Americo Cunha Jr · David A. W. Barton · Thiago G. Ritto

CE-ABC

A. Cunha Jr, D. A. W. Barton and T. G Rritto, *Uncertainty quantification in mechanistic epidemic models via cross-entropy approximate Bayesian computation*, **Nonlinear Dynamics**, 111:9649-9679, 2023

Thank you for your attention!

www.americocunha.org

@AmericoCunhaJr

@AmericoCunhaJr

@AmericoCunhaJr