How to predict an epidemic of Zika virus?

A challenge in nonlinear stochastic dynamics

Americo Cunha Jr

Universidade do Estado do Rio de Janeiro – UERJ
NUMERICO – **Nú**cleo de **M**odelagem e **E**xperi**m**entação **Co**mputacional
http://numerico.ime.uerj.br

In collaboration with: Eber Dantas (UERJ)
Michel Tosin (UERJ)

COLMEA 2019 - IM-UFRJ
March 28, 2019
Rio de Janeiro, Brazil

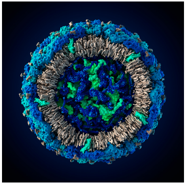
Outline

Section 1

Introduction

Zika virus (ZIKV)

- Member of *Flaviviridae* virus family
- First isolated in 1947 at Uganda, Africa
- Mainly spread by *Aedes* mosquitoes
- W.H.O declared it a public health emergency of international concern
- More than 140,000 confirmed cases in Brazil since 2015
- International consensus that ZIKV is a cause of:
  - Guillain–Barré syndrome
  - Microcephaly

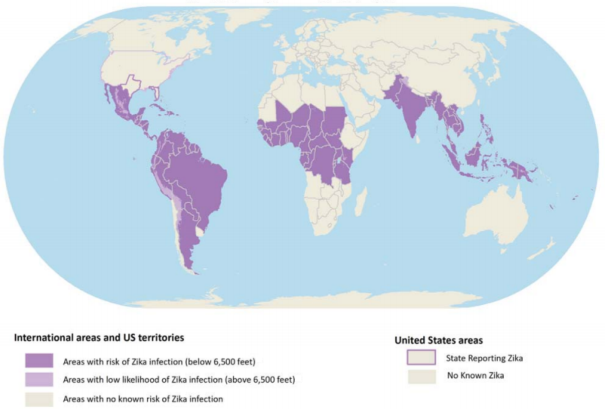


Zika virus

Aedes aegypti

Global outbreak of Zika virus

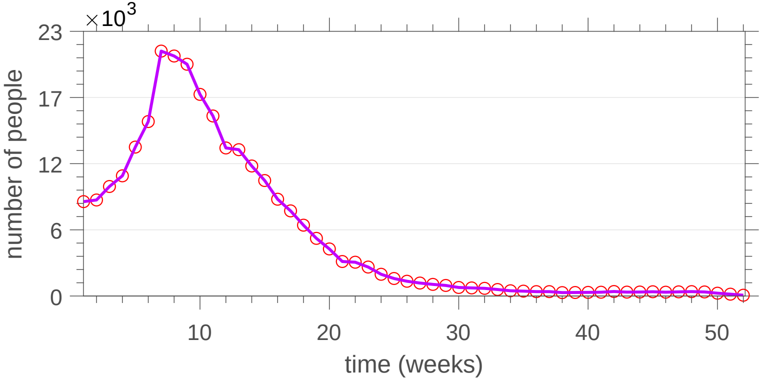
World Map of Areas with Risk of Zika




Centers for Disease Control and Prevention, *World Map of Areas with Risk of Zika, March 2018.*

Zika virus outbreak in Brazil

New cases in Brazil by epidemiological week of 2016



Ministério da Saúde. *Obtenção de número de casos confirmados de zika, por município e semana epidemiológica.* `https://bit.ly/2OVgGGt`

Dengue virus (DENV)

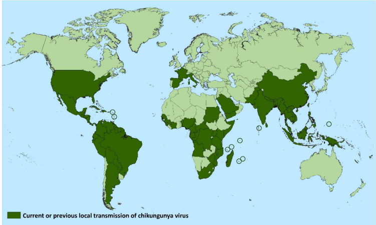
- Member of *Flaviviridae* virus family
- Mainly spread by *Aedes* mosquitoes, as in the case for Zika virus
- Probable cases in Brazil:
  - $>$ 170,000 in 2018
  - $>$ 250,000 in 2017
  - $>$ 3 million in 2016 and 2015

Dengue virus

Aedes aegypti

Other Arbovirus

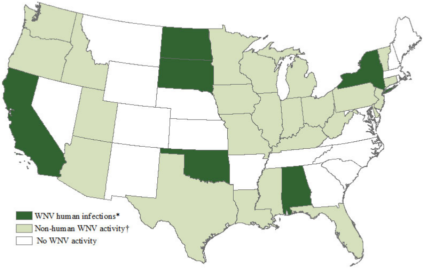
- **AR**thropod-**BO**rne virus
- Yellow Fever: South America and Africa
  (261 deaths in Brazil in 2017)
- Chikungunya: worldwide
  ($>$ 204,000 confirmed cases in Brazil since 2015)
- Rift Valley fever: Africa and Arabian Peninsula
  (ongoing outbreak in Kenya by June 2018)



`Chikungunya cases (May 2018)`

Other Arbovirus

- Japanese encephalitis: Southeast Asia, Western Pacific
- West Nile virus: widely established from Canada to Venezuela
- Both transmitted by the *Culex* mosquitoes



West Nile virus activity in USA (July 2018)

Typical questions to be answered

- How many people will the outbreak potentially infect?
- How far and how quickly will the disease spread?
- What areas and people are at highest risk, and when are they most at risk?
- How can we best make use of limited resources?
- How can we best slow or prevent the outbreak and protect vulnerable populations?

C. Manore and M. Hyman, *Mathematical Models for Fighting Zika Virus*, SIAM News, May 2016.

Typical questions to be answered

- How many people will the outbreak potentially infect?
- How far
- What ar ... hen are they most at
- How car
- How can we best slow or prevent the outbreak and protect vulnerable populations?

Mathematical models to simulate Zika virus spread can provide important guidance and insight to these questions.

C. Manore and M. Hyman, *Mathematical Models for Fighting Zika Virus*, SIAM News, May 2016.

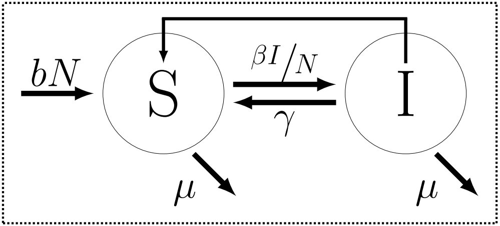
Research objectives

- Develop an epidemic model to describe the recent outbreak of Zika virus in Brazil
- Verify (qualitatively and quantitatively) the epidemic model capacity of prediction
- Calibrate this epidemic model with real data to obtain reliable predictions
- Construct a stochastic model to deal with data uncertainties and made more robust predictions

Section 2

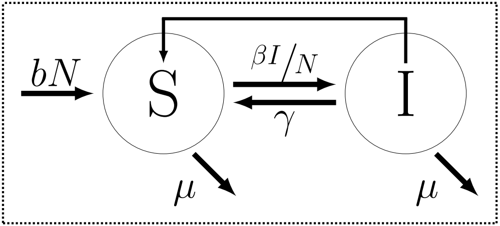
Dynamic Model

SIS model



S - Population of susceptible
I - Population of infected
N - Total population
β - Transmission rate
γ - Recovery rate
b - Birth rate
μ - Mortality rate

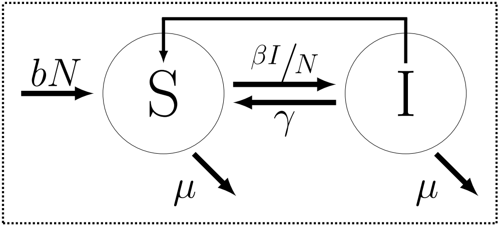
SIS model



S - Population of susceptible
I - Population of infected
N - Total population
β - Transmission rate
γ - Recovery rate
b - Birth rate
μ - Mortality rate

Rate of change of S = Input of S - Output of S

SIS model

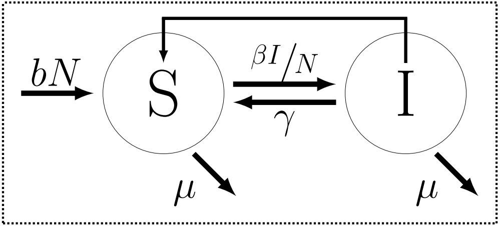


S - Population of susceptible
I - Population of infected
N - Total population
β - Transmission rate
γ - Recovery rate
b - Birth rate
μ - Mortality rate

Rate of change of S = Input of S - Output of S

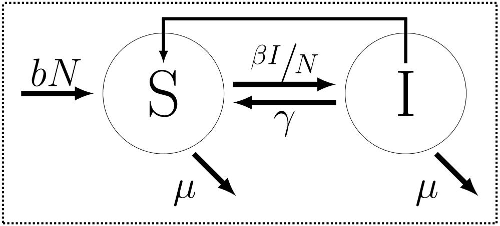
$$\frac{dS}{dt} = \Big(\underbrace{bN}_{\text{Births}} + \underbrace{\gamma I}_{\text{Recovery}} \Big) - \Big(\underbrace{\beta \frac{I}{N} S}_{\text{Infections}} + \underbrace{\mu S}_{\text{Mortality}} \Big)$$

SIS model



S - Population of susceptible
I - Population of infected
N - Total population
β - Transmission rate
γ - Recovery rate
b - Birth rate
μ - Mortality rate

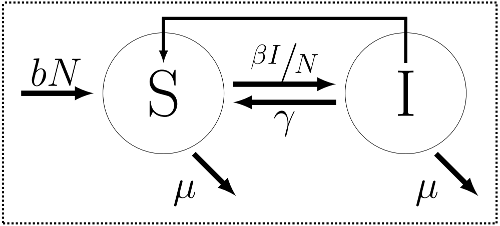
SIS model



S - Population of susceptible
I - Population of infected
N - Total population
β - Transmission rate
γ - Recovery rate
b - Birth rate
μ - Mortality rate

Rate of change of I = Input of I - Output of I

SIS model



S - Population of susceptible
I - Population of infected
N - Total population
β - Transmission rate
γ - Recovery rate
b - Birth rate
μ - Mortality rate

Rate of change of I = Input of I - Output of I

$$\frac{dI}{dt} = \underbrace{\beta \frac{S}{N} I}_{\text{Infections}} - \left(\underbrace{\gamma I}_{\text{Recovery}} + \underbrace{\mu I}_{\text{Mortality}} \right)$$

SIS model dynamical system

$$\frac{\mathrm{d}S}{\mathrm{d}t} = bN + \gamma\, I - \left(\beta\, \frac{I}{N} + \mu\right) S$$

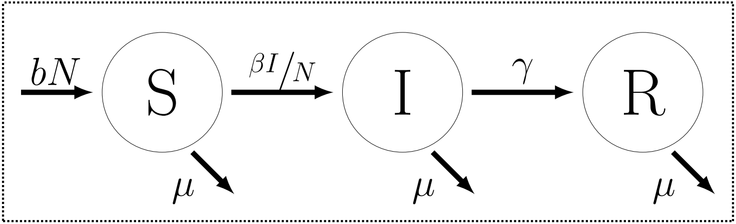
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta\, \frac{I}{N}\, S - (\gamma + \mu)\, I$$

\+ initial conditions

S - Population of susceptible
I - Population of infected
N - Total population

β - Transmission rate
γ - Recovery rate
b - Birth rate
μ - Mortality rate

SIR model



SIR model dynamical system

$$\frac{dS}{dt} = bN + \gamma I - \left(\beta \frac{I}{N} + \mu\right) S$$

$$\frac{dI}{dt} = \beta \frac{I}{N} S - (\gamma + \mu) I$$

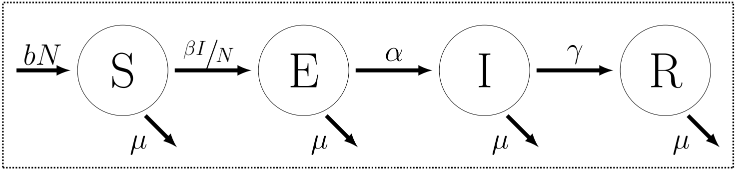
$$\frac{dR}{dt} = \gamma I - \mu R$$

+ initial conditions

S - Population of susceptible
I - Population of infected
R - Population of recovered
N - Total population

β - Transmission rate
γ - Recovery rate
b - Birth rate
μ - Mortality rate

SEIR model



SEIR model dynamical system

$$\frac{dS}{dt} = bN - \beta \frac{I}{N} S - \mu S$$

$$\frac{dE}{dt} = \beta \frac{I}{N} S - (\alpha + \mu) E$$

$$\frac{dI}{dt} = \alpha E - (\gamma + \mu) I$$

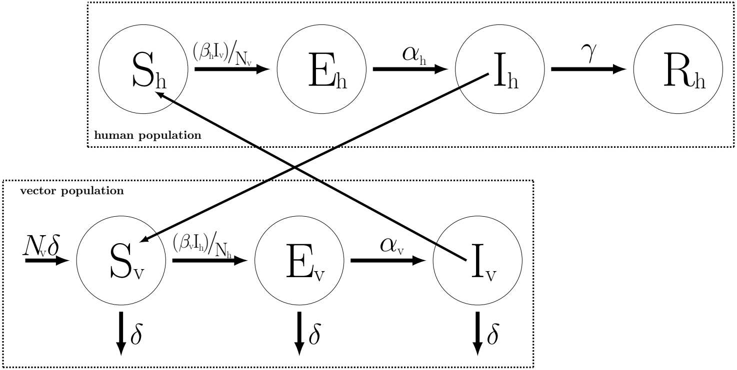
$$\frac{dR}{dt} = \gamma I - \mu R$$

$+$ initial conditions

S - Population of susceptible
E - Population of exposed
I - Population of infectious
R - Population of recovered
N - Total population

α - Incubation ratio
β - Transmission rate
γ - Recovery rate
b - Birth rate
μ - Mortality rate

SEIR-SEI model for Zika virus dynamics



A. J. Kucharski et al. *Transmission Dynamics of Zika Virus in Island Populations: A Modelling Analysis of the 2013–14 French Polynesia Outbreak.* **PLOS Neglected Tropical Diseases, 2016.**

Associated dynamical system

$$\frac{d S_h}{dt} = -\beta_h \frac{I_v}{N_v} S_h$$

$$\frac{d E_h}{dt} = \beta_h \frac{I_v}{N_v} S_h - \alpha_h E_h$$

$$\frac{d I_h}{dt} = \alpha_h E_h - \gamma I_h$$

$$\frac{d R_h}{dt} = \gamma I_h$$

$$\frac{d S_v}{dt} = \delta - \beta_v S_v \frac{I_h}{N_h} - \delta S_v$$

$$\frac{d E_v}{dt} = \beta_v S_v \frac{I_h}{N_h} - (\delta + \alpha_v) E_v$$

$$\frac{d I_v}{dt} = \alpha_v E_v - \delta I_v$$

$$\frac{d C}{dt} = \alpha_h E_h$$

$+$ initial conditions

S - Population of susceptible
V - Population of vaccinated
E - Population of exposed
I - Population of infectious
R - Population of recovered

C - Infected humans cumulative
N - Total population
α - Incubation ratio
β - Transmission rate
γ - Recovery rate

δ - Vector lifespan ratio
σ - Infection rate of vaccinated
ν - Fraction of vaccinated
h - Human-related
v - Vector-related

A. J. Kucharski et al. *Transmission Dynamics of Zika Virus in Island Populations: A Modelling Analysis of the 2013–14 French Polynesia Outbreak.* **PLOS Neglected Tropical Diseases, 2016.**

Model parameters and outbreak data

- open scientific literature

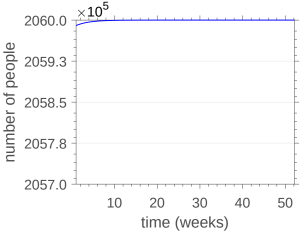
- Brazilian health system

MINISTÉRIO DA SAÚDE

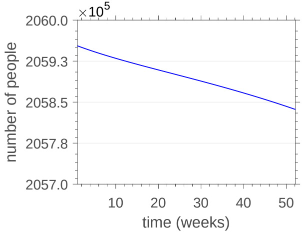
FIOCRUZ
Fundação Oswaldo Cruz

parameter	value	unit
α_h	$1/5.9$	days^{-1}
α_v	$1/9.1$	days^{-1}
γ	$1/7.9$	days^{-1}
δ	$1/11$	days^{-1}
β_h	$1/11.3$	days^{-1}
β_v	$1/8.6$	days^{-1}
N	206×10^6	people

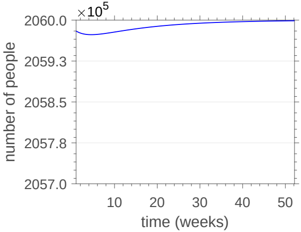
Time series of susceptible humans



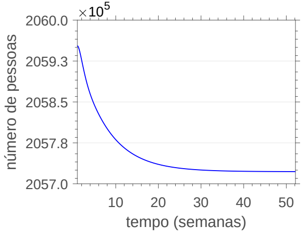
SIS model



SEIR model

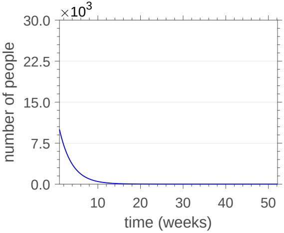


SIR model

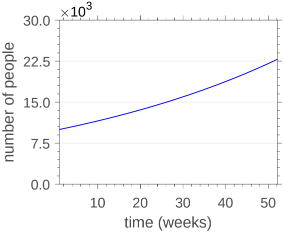


SEIR-SEI model

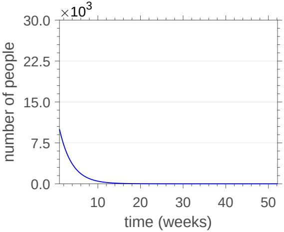
Time series of infectious humans



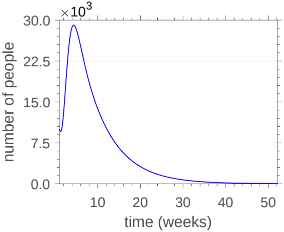
SIS model



SEIR model

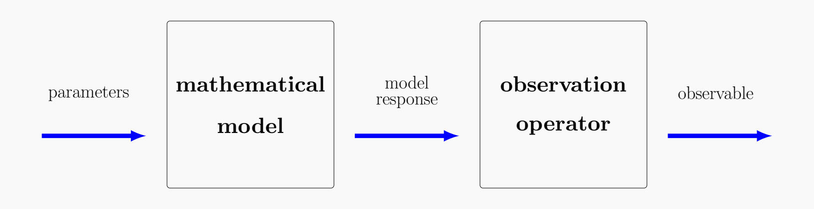


SIR model



SEIR-SEI model

Quantities of interest (QoI)



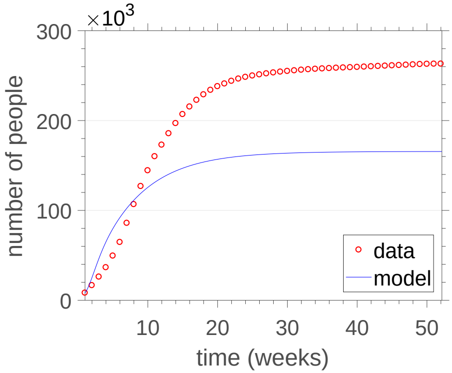
QoI 1: cumulative number of infectious

$$C_t = \int_{\tau=0}^{t} \alpha_h \, E_h(\tau) \, d\tau$$

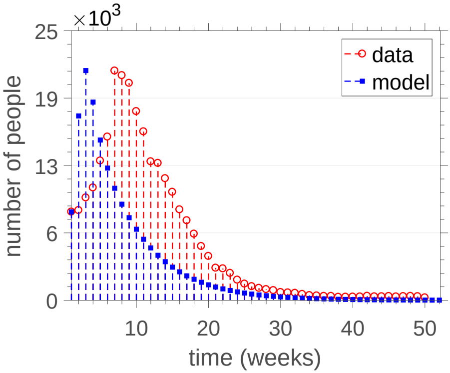
QoI 2: new infectious cases

$$\mathcal{N}_w = C_w - C_{w-1}, \quad (w = 2, 3, \cdots, 52)$$
$$\mathcal{N}_1 = C_1$$

Time series for QoI's (SEIR-SEI model)

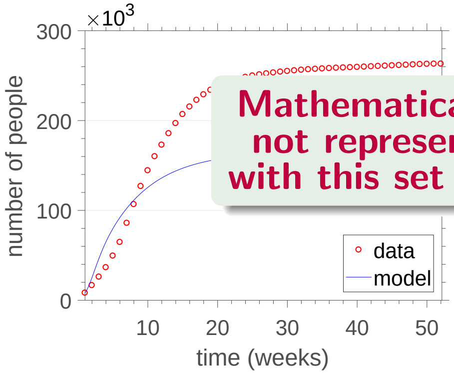


cumulative number of infectious

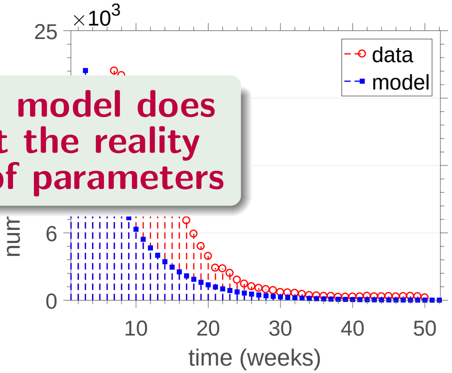


new infectious cases

Time series for QoI's (SEIR-SEI model)



cumulative number of infectious



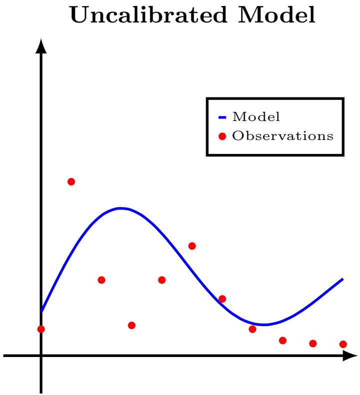
new infectious cases

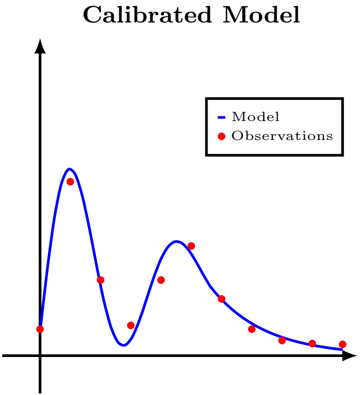
Mathematical model does not represent the reality with this set of parameters

Section 3

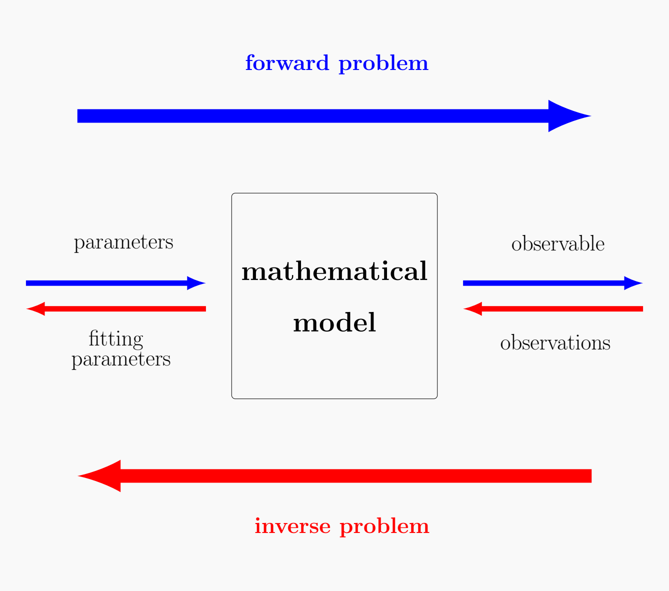
Inverse Problem

Calibration of the model





Forward and inverse problem



Inverse problem formulation

- data space: $F = \mathbb{R}^M$
- parameter space: $C = \left\{ \boldsymbol{\alpha} \in \mathbb{R}^{12} \mid \boldsymbol{\alpha}_{min} \leq \boldsymbol{\alpha} \leq \boldsymbol{\alpha}_{max} \right\}$
- observation vector: $\boldsymbol{y} = (y_1, y_2, \cdots, y_M) \in F$
- prediction vector: $\boldsymbol{\phi}(\boldsymbol{\alpha}) = (\phi_1, \phi_2, \cdots, \phi_M) \in F$
- misfit function:

$$J(\boldsymbol{\alpha}) = ||\boldsymbol{y} - \boldsymbol{\phi}(\boldsymbol{\alpha})||_F^2 = \sum_{m=1}^{M} \left| y_m - \phi_m(\boldsymbol{\alpha}) \right|^2$$

Find a vector of parameters such that

$$\boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha} \in C}{\arg\min}\ J(\boldsymbol{\alpha}).$$

Inverse problem formulation

- data space: $F = \mathbb{R}^M$
- parameter space: $C = \left\{ \boldsymbol{\alpha} \in \mathbb{R}^{12} \mid \boldsymbol{\alpha}_{min} \leq \boldsymbol{\alpha} \leq \boldsymbol{\alpha}_{max} \right\}$
- observation vector: $\boldsymbol{y} = (y_1, y_2, \cdots, y_M) \in F$
- prediction vector: $\boldsymbol{\phi}(\boldsymbol{\alpha}) = (\phi_1, \phi_2, \cdots, \phi_M) \in F$
- misfit function:

$$J(\boldsymbol{\alpha}) = ||\boldsymbol{y} - \boldsymbol{\phi}(\boldsymbol{\alpha})||_F^2 = \sum_{m=1}^{M} \left| y_m - \phi_m(\boldsymbol{\alpha}) \right|^2$$

Find a vector of parameters such that

$$\boldsymbol{\alpha}^* = \arg\min_{\boldsymbol{\alpha} \in C} J(\boldsymbol{\alpha}).$$

$\Longrightarrow$ Q-wellposed: existence, uniqueness, unimodality and local stability

Inverse problem formulation

- data space: $F = \mathbb{R}^M$
- parameter space: $C = \left\{ \boldsymbol{\alpha} \in \mathbb{R}^{12} \mid \boldsymbol{\alpha}_{min} \leq \boldsymbol{\alpha} \leq \boldsymbol{\alpha}_{max} \right\}$
- observation vector: $\boldsymbol{y} = (y_1, y_2, \; \cdots \;, \; y_M) \in F$
- prediction vector: $\boldsymbol{\phi}(\boldsymbol{\alpha}) = (\phi_1, \phi_2, \; \cdots \;, \; \phi_M) \in F$
- misfit function:

$$J(\boldsymbol{\alpha}) = ||\boldsymbol{y} - \boldsymbol{\phi}(\boldsymbol{\alpha})||_F^2 = \sum_{m=1}^{M} \left| y_m - \phi_m(\boldsymbol{\alpha}) \right|^2$$

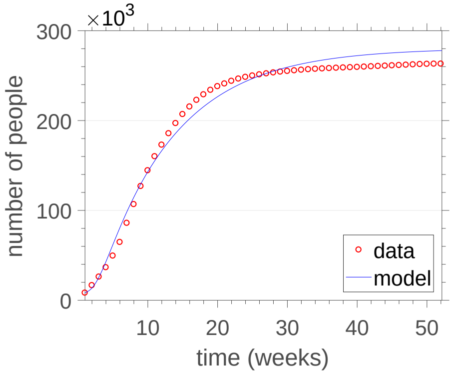
Find a vector of parameters such that

$$\boldsymbol{\alpha}^* = \arg\min_{\boldsymbol{\alpha} \in C} J(\boldsymbol{\alpha}).$$

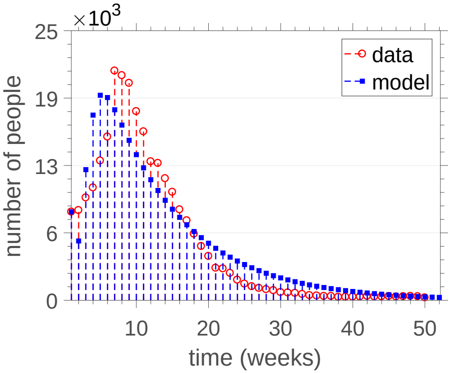
$\Longrightarrow$ Q-wellposed: existence, uniqueness, unimodality and local stability

$\Longrightarrow$ Solution algorithm: bounded trust-region-reflective

Calibrated model response

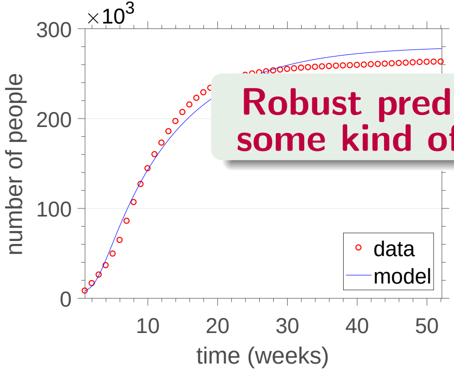


cumulative number of infectious

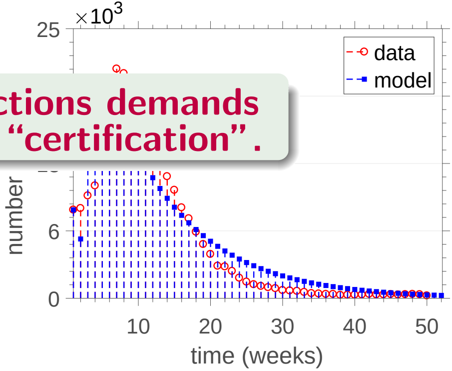


new infectious cases

Calibrated model response



cumulative number of infectious



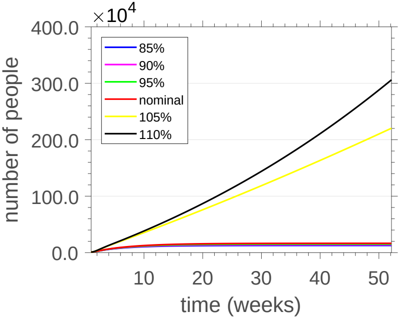
new infectious cases

Robust predictions demands some kind of "certification".

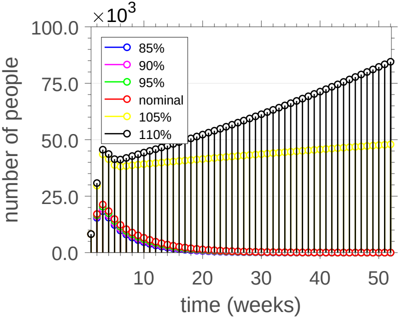
Section 4

Sensitivity Analysis

Parametric study for β_h

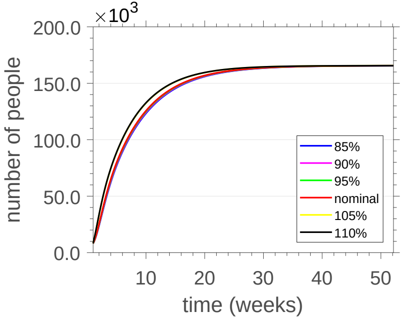


cumulative number of infectious

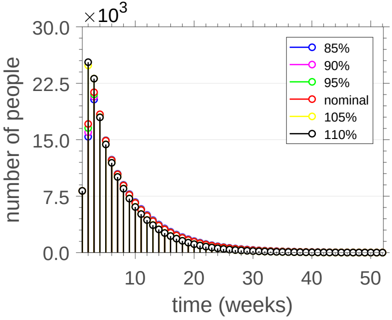


new infectious cases

Parametric study for α_h

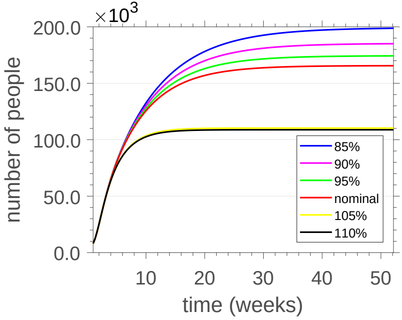


cumulative number of infectious

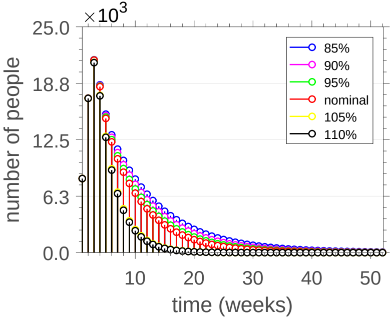


new infectious cases

Parametric study for γ

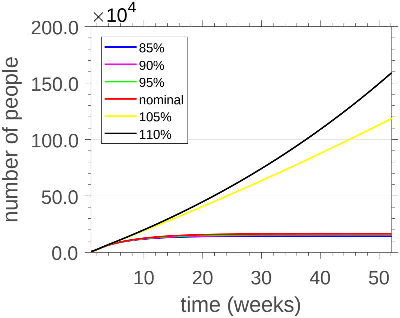


cumulative number of infectious

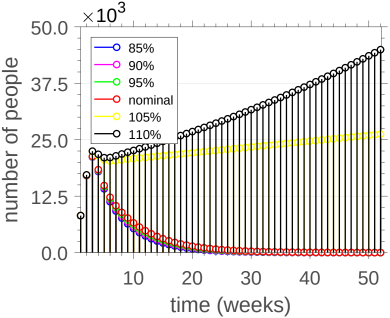


new infectious cases

Parametric study for β_v

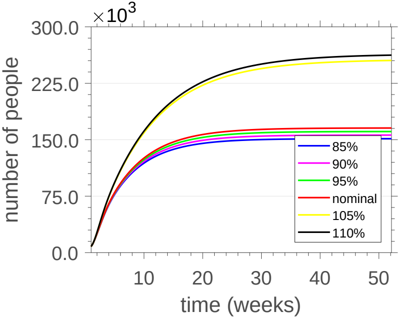


cumulative number of infectious

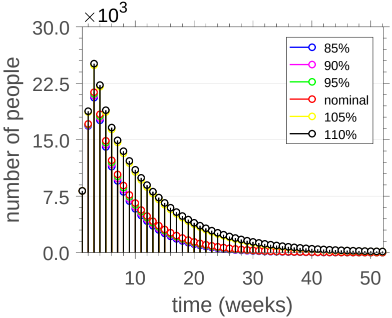


new infectious cases

Parametric study for α_v

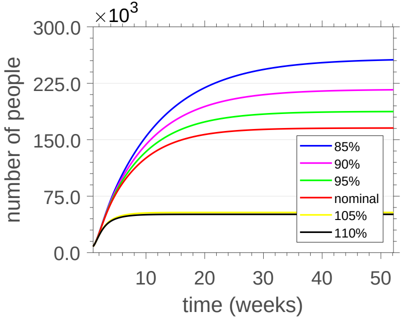


cumulative number of infectious

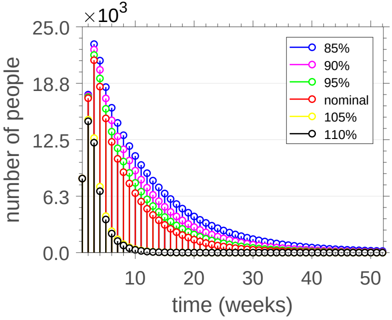


new infectious cases

Parametric study for δ



cumulative number of infectious



new infectious cases

Variance-based sensitivity analysis

Mathematical model:

$$Y = \mathcal{M}(\mathbf{X}), \quad X_i \sim \mathcal{U}(0,1), \quad \text{(i.i.d.)}$$

Hoeffding-Sobol' decomposition:

$$Y = \mathcal{M}_0 + \sum_{1 \leq i \leq n} \mathcal{M}_i(X_i) + \sum_{1 \leq i < j \leq n} \mathcal{M}_{ij}(X_i, X_j) + \cdots + \mathcal{M}_{1 \cdots n}(X_1 \cdots X_n)$$

An orthogonal decomposition in terms of conditional expectations:

- $\mathcal{M}_0 = \mathbb{E}\left\{Y\right\}$
- $\mathcal{M}_i(X_i) = \mathbb{E}\left\{Y|X_i\right\} - \mathcal{M}_0$
- $\mathcal{M}_{ij}(X_i, X_j) = \mathbb{E}\left\{Y|X_i, X_j\right\} - \mathcal{M}_0 - \mathcal{M}_i - \mathcal{M}_j$
- etc

Sobol' indices

Total variance:

$$D = Var\left[\mathcal{M}(\boldsymbol{X})\right] = \sum_{u \subset \{1,\cdots,k\}} Var\left[\mathcal{M}_u(\boldsymbol{X}_u)\right]$$

First order Sobol' indices:

$$S_i = Var\left[\mathcal{M}_i(X_i)\right]/D$$

(quantify the additive effect of each input separately)

Second order Sobol' indices:

$$S_{ij} = Var\left[\mathcal{M}_{ij}(X_i, X_j)\right]/D$$

(quantify interaction effect of inputs X_i and X_j)

Metamodelling via Polynomial Chaos

Assuming $Y = \mathcal{M}(\mathbf{X})$ has finite variance, then it admits a Polynomial Chaos expansion

$$Y = \sum_{\alpha \in \mathbb{N}^k} y_\alpha \, \Phi_\alpha(\mathbf{X})$$

where

- $\Phi_\alpha(\mathbf{X})$: multivariate orthonormal polynomials
- y_α: real-valued coeficients to be determined

D. Xiu, and G. Karniadakis, *The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations.* **SIAM Journal on Scientific Computing**, 24: 619-644, 2002.

PC-based Sobol' indices

For computational purposes, a truncated PCE is employed

$$Y \approx \sum_{\alpha \in \mathcal{A}} y_\alpha \, \Phi_\alpha(\mathbf{X})$$

Thus, Sobol' indices are given by

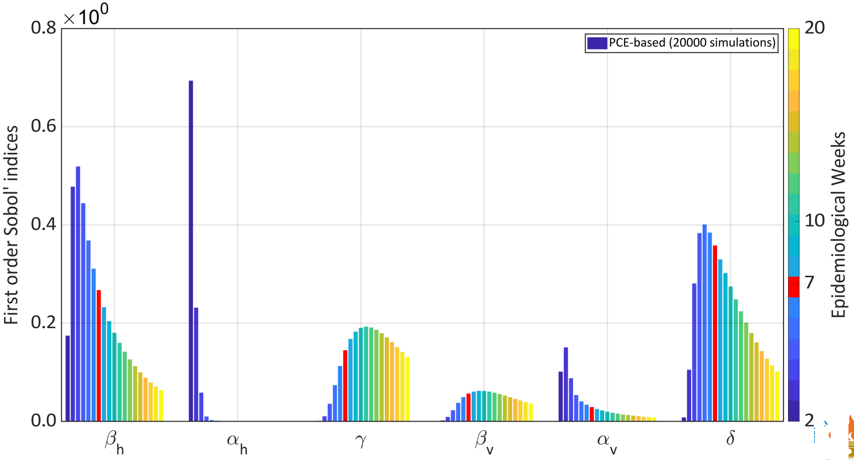
$$S_u = D_u / D = \sum_{\alpha \in \mathcal{A}_u} y_\alpha^2 \Bigg/ \sum_{\alpha \in \mathcal{A} \setminus 0} y_\alpha^2$$

$$\mathcal{A}_u = \{\alpha \in \mathcal{A} : i \in u \Longleftrightarrow \alpha_i \neq 0\}$$

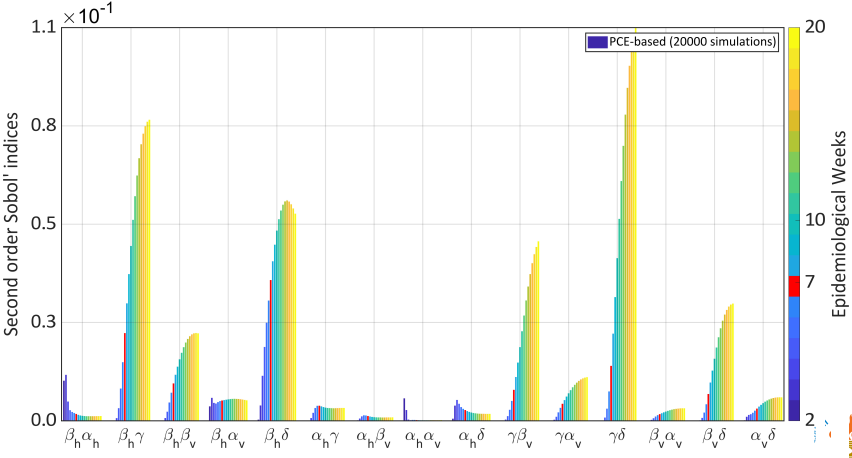
Sobol' indices of any order can be obtained, analytically, from the coefficients of the PC expansion!

B. Sudret, *Global sensitivity analysis using polynomial chaos expansions.* **Reliability Engineering & System Safety, 2016**, 93(7): 964–979, 2008.

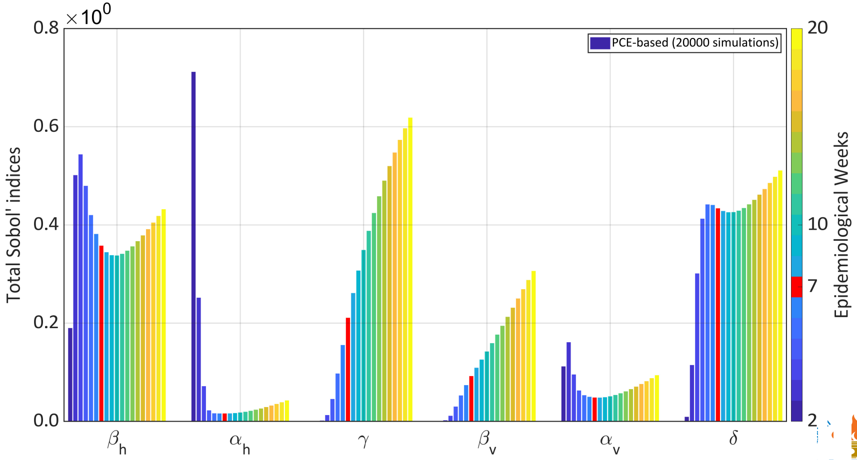
Global sensitivity analysis: first order



Global sensitivity analysis: second order



Global sensitivity analysis: total order



Global sensitivity analysis: general overview

$\rightarrow$ Two most relevant: δ and β_H (75% variance around 7th EW)

$\rightarrow$ Third most, γ, mainly by nonlinear interactions with δ and β_H

$\rightarrow$ Parameters limited to nonlinear interactions have, in general, delayed effects (significant for $EW > 15$)

$\rightarrow$ (*sparsity-of-effects principle*) Higher order interactions have minor effect: 1st and 2nd are 99.8–96.7% variance on 5–10th EW

Around 7th EW $\rightarrow$ uncertainty propagation of $\{\beta_h, \delta\}$

Section 5

Uncertainty Quantification

Uncertainty Quantification (UQ) framework

Mathematical model:

$$Y = \mathcal{M}(\mathbf{X})$$

General steps for UQ:

1. Stochastic modeling
   $\longrightarrow$ characterization of inputs uncertainties
   (MaxEnt Principle)
2. Uncertainty propagation
   $\longrightarrow$characterization of output uncertainties
   (Monte Carlo Method)
3. Response certification
   $\longrightarrow$ specification of reliability levels for predictions
   (Nonparametric Statistical Inference)

C. Soize, *Uncertainty Quantification: An Accelerated Course with Advanced Applications in Computational Engineering*, Springer, 2017.

Maximum Entropy Principle (MaxEnt)

Among all the probability distributions, consistent with the known information about a random parameter, choose the one which corresponds to the maximum of entropy (MaxEnt).

MaxEnt distribution = most unbiased distribution

Entropy of the random variable X is defined as

$$\mathcal{S}\left(p_X\right) = -\int_{\mathbb{R}} p_X(x) \ln p_X(x)\, dx,$$

"measure for the level of uncertainty"

MaxEnt optimization problem

Maximize

$$\mathcal{S}\left(p_X\right) = -\int_{\mathbb{R}} p_X(x) \, \ln p_X(x) \, dx,$$

respecting $N+1$ constraints (known information) given by

$$\int_{\mathbb{R}} g_k\left(X\right) p_X(x) \, dx = m_k, \qquad k = 0, \cdots, N,$$

where the g_k are known real functions, with $g_0(x) = 1$.

MaxEnt optimization problem

Maximize

$$\mathcal{S}\,(p_X) = -\int_{\mathbb{R}} p_X(x)\,\ln p_X(x)\,dx,$$

respecting $N+1$ constraints (known information) given by

$$\int_{\mathbb{R}} g_k\,(X)\,p_X(x)\,dx = m_k, \qquad k = 0, \cdots, N,$$

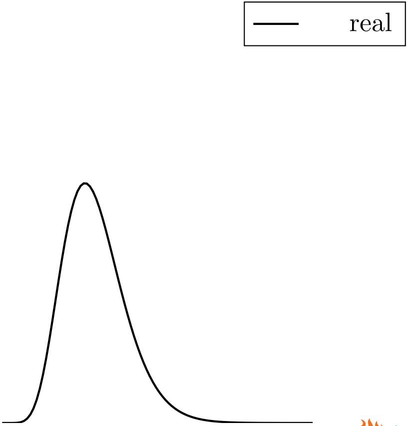
where the g_k are known real functions, with $g_0(x) = 1$.

MaxEnt general solution

$$p_X(x) = \mathbb{1}_{\mathcal{K}}(x)\,\exp\left(-\lambda_0\right)\,\exp\left(-\sum_{k=1}^{N} \lambda_k\,g_k(x)\right)$$

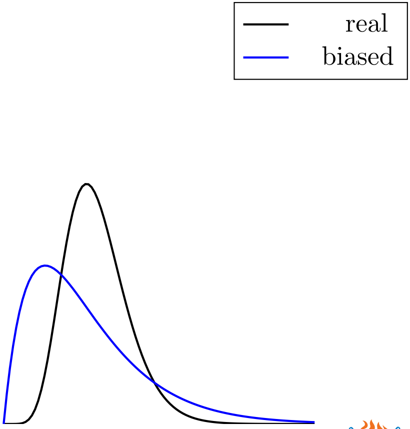
Philosophy of MaxEnt Principle

- The parameter of interest has a unknown distribution



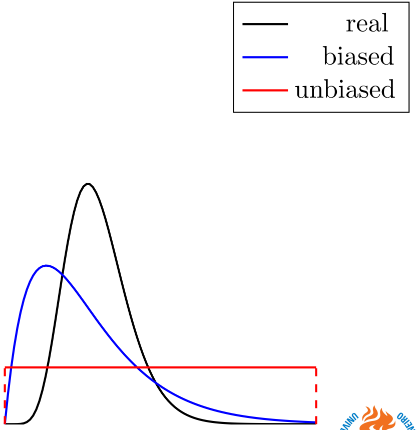
Philosophy of MaxEnt Principle

- The parameter of interest has a unknown distribution
- Distributions arbitrarily chosen can be coarse and biased



Philosophy of MaxEnt Principle

- The parameter of interest has a unknown distribution
- Distributions arbitrarily chosen can be coarse and biased
- A conservative strategy is to use the most unbiased (MaxEnt) distribution



Uncertainty propagation through the model

Monte Carlo Method

pre-processing	processing	post-processing
generation of scenarios	solution of model equations	computation of statistics
$\boldsymbol{X}_1$ ⋮ $\boldsymbol{X}_M$	$\boldsymbol{U} = h(\boldsymbol{X})$	$\boldsymbol{U}_1 = h(\boldsymbol{X}_1)$ ⋮ $\boldsymbol{U}_M = h(\boldsymbol{X}_M)$
→	**computational model**	→
known $F_{\boldsymbol{X}}$		estimated $F_{\boldsymbol{U}}$
generator of random vector $\boldsymbol{X}$	deterministic solver of $\boldsymbol{u} = h(\boldsymbol{x})$	statistical inference to estimate convergence and distribution of $\boldsymbol{U}$

Probabilistic model 1

Random variables: β_h and δ

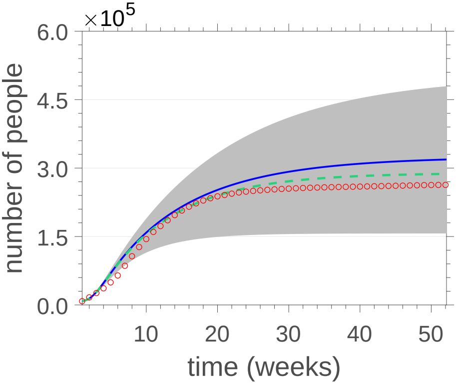
Available information: support and mean (nominal) value

MaxEnt distribution

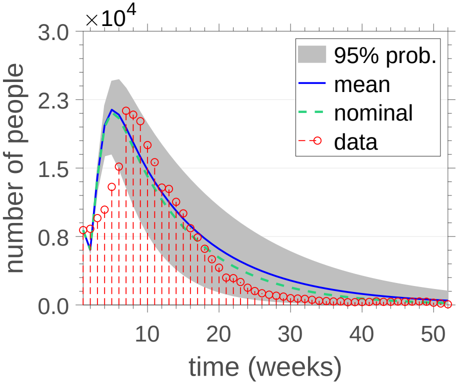
$$p_X(x) = \mathbb{1}_{[a,b]}(x) \, \exp\left(-\lambda_0 - \lambda_1 \, x\right)$$

“truncated exponential (2 parameters)”

Confidence band for the QoIs



cumulative number of infectious



new infectious cases

Probabilistic model 2

Random variables: β_h and δ

Available information: support, mean (nominal) value and dispersion

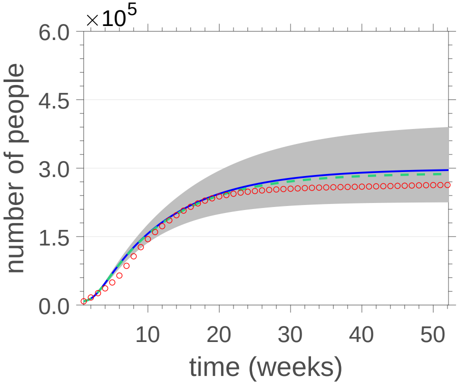
MaxEnt distribution

$$p_X(x) = \mathbb{1}_{[a,b]}(x)\, \exp\left(-\lambda_0 - \lambda_1\, x - \lambda_2\, x^2\right)$$

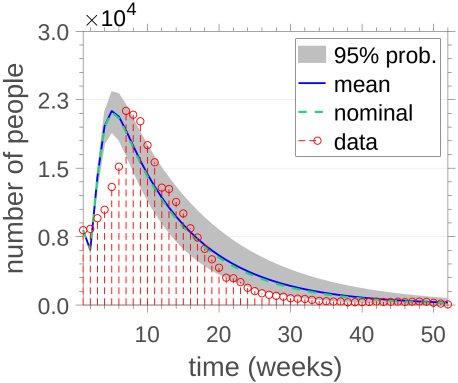
"truncated exponential (3 parameters)"

Confidence band for the QoIs

β_h dispersion = 5% , δ dispersion = 5%



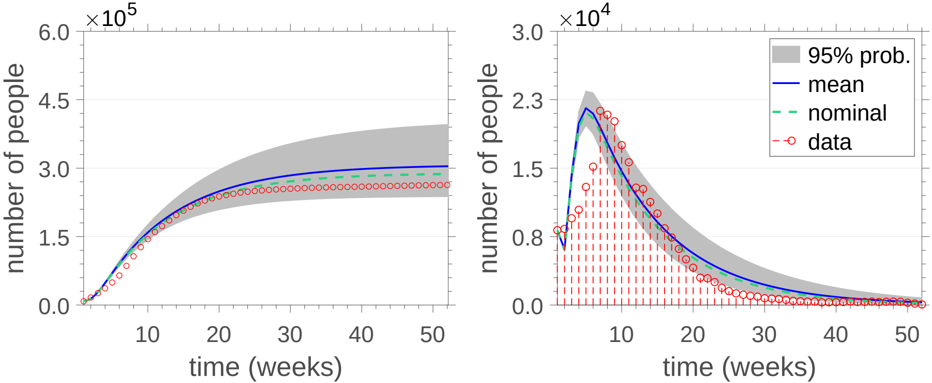
cumulative number of infectious



new infectious cases

Confidence band for the QoIs

β_h dispersion $= 10\%$, δ dispersion $= 5\%$

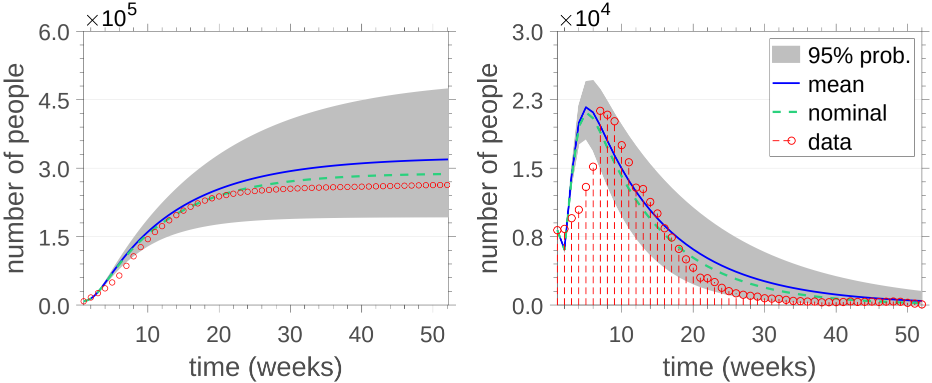


cumulative number of infectious

new infectious cases

Confidence band for the QoIs

β_h dispersion = 10% , δ dispersion = 10%



cumulative number of infectious

new infectious cases

Probabilistic model 3

Random variables: β_h, δ and σ

Available information for β_h and δ: support, mean (nominal) value

Distribution for β_h and β_v

$$p_X(x) = \mathbb{1}_{[a,b]}(x)\, \exp\left(-\lambda_0 - \lambda_1\, x - \lambda_2\, x^2\right)$$

Available information for σ: support

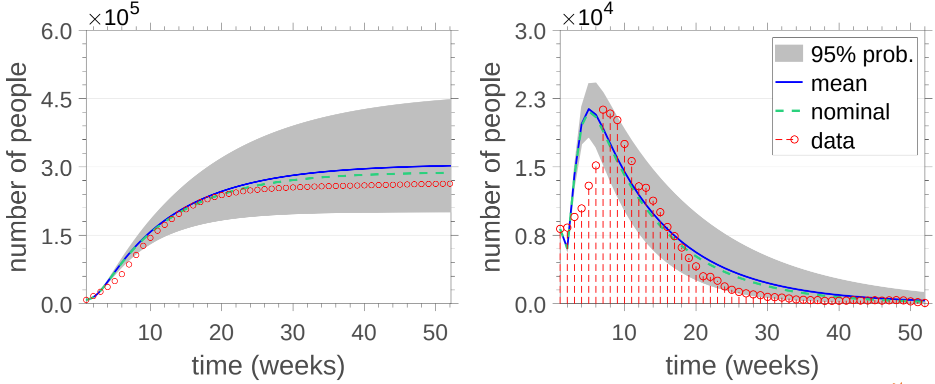
MaxEnt distribution for σ

$$p_X(x) = \mathbb{1}_{[a,b]}(x)\, \frac{1}{b-a}$$

"uniform"

Confidence band for the QoIs

random dispersion $\sim U(5\%, 10\%)$

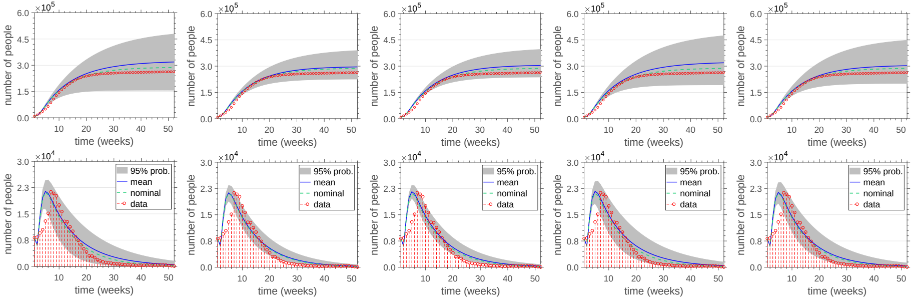


cumulative number of infectious

new infectious cases

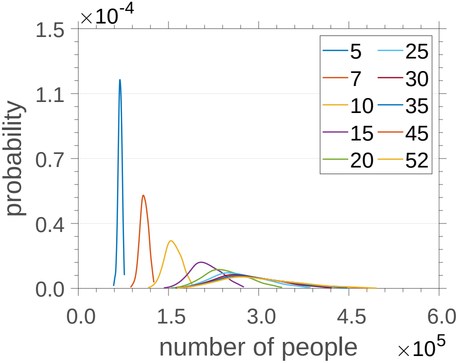
Confidence band for the QoIs

no dispersion | $\sigma = \{5\%,5\%\}$ | $\sigma = \{10\%,5\%\}$ | $\sigma = \{10\%,10\%\}$ | $\sigma \sim U(5\%, 10\%)$

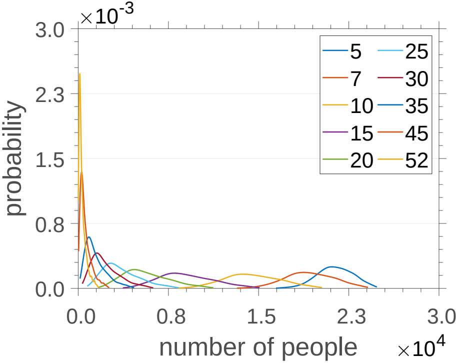


Evolution of QoIs PDFs

random dispersion $\sim U(5\%, 10\%)$

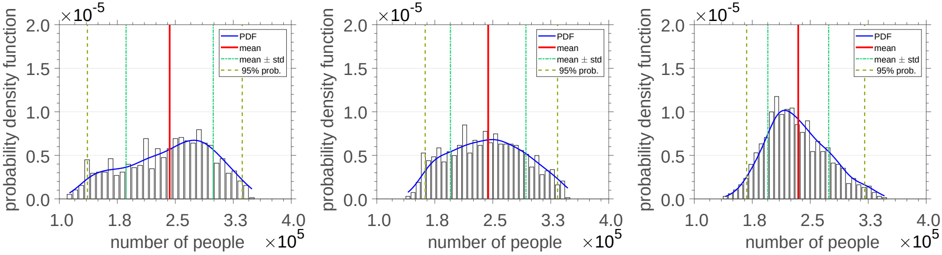


cumulative number of infectious



new infectious cases

Time-averaged cumulative infectious



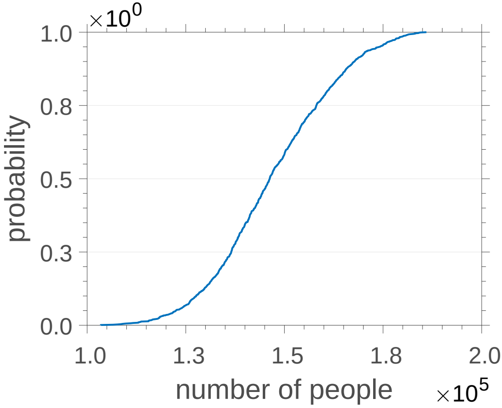
no dispersion

$\sigma = \{10\%, 10\%\}$

$\sigma \sim U(5\%, 10\%)$

(mean) Cumulative infectious CDF until EW 20

$\sigma \sim U(5\%, 10\%)$



Statistics of C

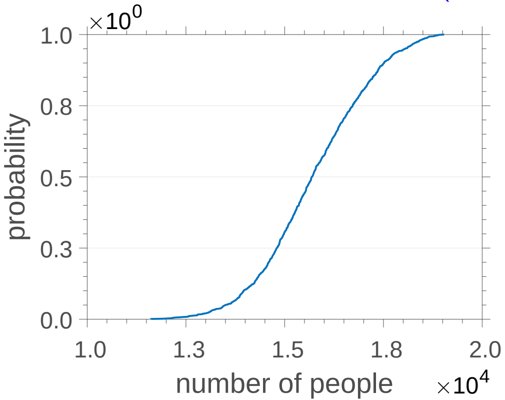
$$\begin{aligned} \text{mean} &= 1{,}47 \times 10^5 \\ \text{std. dev.} &= 1{,}53 \times 10^4 \\ \text{skewness} &= 0.084 \\ \text{kurtosis} &= 2.605 \\ P(C \geq c^*) &= 87.10\% \end{aligned}$$

$c^* = 130{,}000$

Half the maximum C (data)

(mean) New cases CDF until 10th EW

$\sigma \sim U(5\%, 10\%)$



Statistics of $\mathcal{N}_w$

$$\begin{aligned} \text{mean} &= 1{,}57 \times 10^4 \\ \text{std. dev.} &= 1{,}35 \times 10^3 \\ \text{skewness} &= -0.032 \\ \text{kurtosis} &= 2.656 \\ P(\mathcal{N}_w \geq NC^*) &= 83.40\% \end{aligned}$$

$NC^* = 14{,}440$

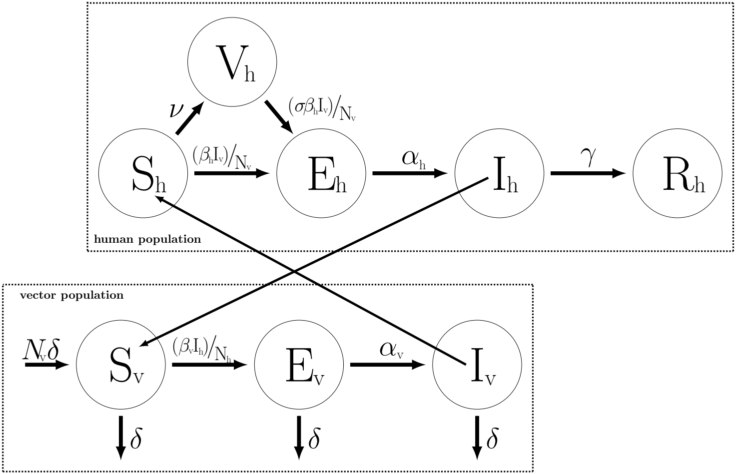
average NC (data) until EW 10

Section 6

Ongoing

Investigation of control strategies

SVEIR-SEI model for Zika virus dynamics



H. S. Rodrigues et al. *Vaccination models and optimal control strategies to dengue*. **Mathematical Biosciences, 247 (2014) 1–12.**

Associated dynamical system

$$\frac{d S_h}{dt} = -\left(\beta_h \frac{I_v}{N_v} + \nu\right) S_h$$

$$\frac{d V_h}{dt} = \nu S_h - \sigma \beta_h \frac{I_v}{N_v} V_h$$

$$\frac{d E_h}{dt} = \beta_h \left(S_h + \sigma V_h\right) \frac{I_v}{N_v} - \alpha_h E_h$$

$$\frac{d I_h}{dt} = \alpha_h E_h - \gamma I_h$$

$$\frac{d R_h}{dt} = \gamma I_h$$

$$\frac{d S_v}{dt} = \delta - \beta_v S_v \frac{I_h}{N_h} - \delta S_v$$

$$\frac{d E_v}{dt} = \beta_v S_v \frac{I_h}{N_h} - (\delta + \alpha_v) E_v$$

$$\frac{d I_v}{dt} = \alpha_v E_v - \delta I_v$$

$$\frac{d C}{dt} = \alpha_h E_h$$

+ initial conditions

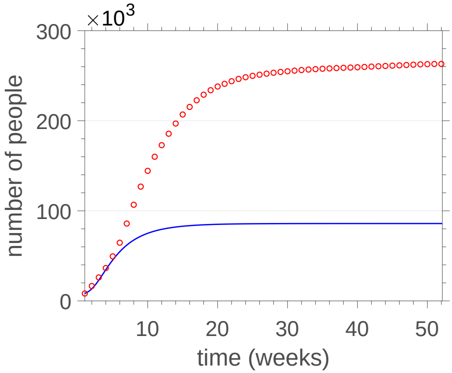
S - Population of susceptible
V - Population of vaccinated
E - Population of exposed
I - Population of infectious
R - Population of recovered

C - Infected humans cumulative
N - Total population
α - Incubation ratio
β - Transmission rate
γ - Recovery rate

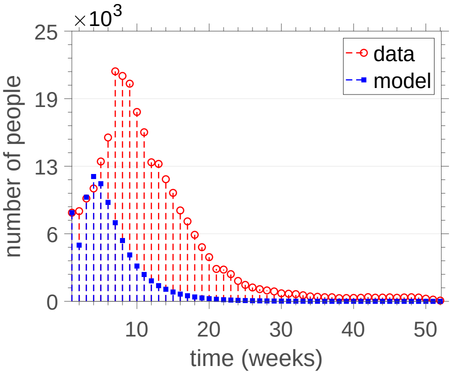
δ - Vector lifespan ratio
σ - Infection rate of vaccinated
ν - Fraction of vaccinated
h - Human-related
v - Vector-related

H. S. Rodrigues et al. *Vaccination models and optimal control strategies to dengue*. **Mathematical Biosciences, 247 (2014) 1–12.**

Time series for QoI's (SVEIR-SEI model)



cumulative number of infectious



new infectious cases

Quantification of model discrepancy

Calculation of model discrepancy

Conventional statistical calibration:

$$\underbrace{y}_{truth} = \underbrace{f(x, \mathbf{p})}_{model} + \underbrace{\varepsilon}_{error}$$

Novel approach:

$$\underbrace{y}_{truth} \approx \underbrace{f\,(x, \mathbf{p}_\varepsilon)}_{model}, \quad \mathbf{p}_\varepsilon = \sum_k \alpha_k \, \Phi_k(\xi)$$

Bayesian inversion to identify α

$$\underbrace{\pi(\text{model} \mid \text{data})}_{posterior} \propto \underbrace{\pi(\text{data} \mid \text{model})}_{likelihood} \times \underbrace{\pi(\text{model})}_{prior}$$

K. Sargsyan, H. N. Najm and R. Ghanem, *On the statistical calibration of physical models*.

International Journal of Chemical Kinetics, 47 (2015) 246-276.

Section 7

Final Remarks

Concluding remarks

Contributions:

- Development of an epidemic model to describe Brazilian outbreak of Zika virus
- Calibration of this model with real epidemic data
- Construction of a parametric probabilistic model of uncertainties

Ongoing research:

- Investigate the effectiveness of different control strategies
- Quantify model discrepancy in a nonparametric way

Future directions:

- Scenarios exploration with active subspace method
- Data-driven identification of epidemiological models

Acknowledgments

Invitation for the talk:

- Prof^{a}. Maria Eulalia Vares
- Prof. Leandro Pimentel

Financial support:

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

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

CAPES

UERJ
UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO

Thank you for your attention!

americo@ime.uerj.br

www.americocunha.org

E. Dantas, M. Tosin and A. Cunha Jr,
Calibration of a SEIR–SEI epidemic model to describe Zika virus outbreak in Brazil,
Applied Mathematics and Computation, 338: 249-259, 2018.
https://doi.org/10.1016/j.amc.2018.06.024

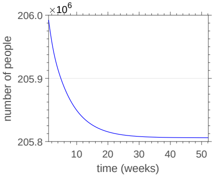
E. Dantas, M. Tosin and A. Cunha Jr,
Uncertainty quantification in the nonlinear dynamics of Zika virus, 2019 (in preparation).
https://hal.archives-ouvertes.fr/hal-02005320

nominal parameters

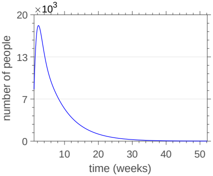
Nominal parameters and initial conditions

$\boldsymbol{\alpha}$	value	unit
α_h	$1/5.9$	days^{-1}
α_v	$1/9.1$	days^{-1}
γ	$1/7.9$	days^{-1}
δ	$1/11$	days^{-1}
β_h	$1/11.3$	days^{-1}
β_v	$1/8.6$	days^{-1}
N	206×10^6	people
S_h^i	205,953,959	people
E_h^i	8,201	people
I_h^i	8,201	people
R_h^i	29,639	people
S_v^i	0.99956	——
E_v^i	2.2×10^{-4}	——
I_v^i	2.2×10^{-4}	——

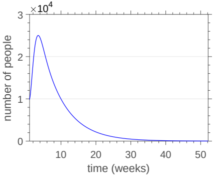
Model response with nominal parameters



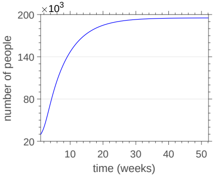
Susceptible humans



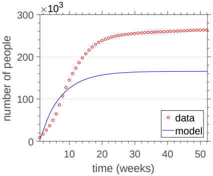
Exposed humans



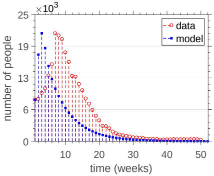
Infectious humans



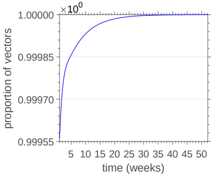
Recovered humans



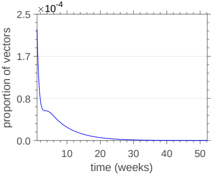
Cumulative infectious



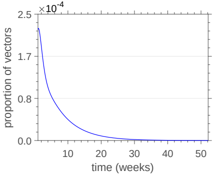
New cases



Susceptible vectors



Exposed vectors



Infectious vectors

Inverse Problem

Well-posedness

Let the forward map $\phi : E \to F$ associates to each parameter vector x, restricted to be on the set of admissible values C in the parameter space E, an observable vector in the data space F. The NLS problem is Quadratically (Q-) wellposed if, and only if, $\phi(C)$ possesses an open neighborhood ϑ such that

1. Existence and uniqueness: for every $z \in \vartheta$, the inverse problem has a unique solution $\hat{x}$
2. Unimodality: for every $z \in \vartheta$, the objective function $x \rightsquigarrow J(x)$ has no parasitic stationary point
3. Local stability: the mapping $z \rightsquigarrow \hat{x}$ is locally Lipschitz continuous from $(\vartheta, ||\cdot||_F)$ to $(C, |\cdot||_E)$.

G. Chavent. *Nonlinear Least Squares for Inverse Problems: Theoretical Foundations and Step-by-Step Guide for Applications.* **Springer, 2010.**

Well-posedness

Theorem

Let the follow finite dimension minimum set of hypothesis hold:

$$\begin{cases} E = \textit{finite dimensional vector space, with norm } ||\cdot||_E, \\ C = \textit{closed, convex subset of } E, \\ C_\eta = \textit{convex open neighborhood of } C \textit{ in } E, \\ F = \textit{Hilbert space, with norm } ||\cdot||_F, \\ z \in F, \\ \phi : C_\eta \rightsquigarrow F \textit{ is twice differentiable along segments of } C_\eta, \\ \textit{and: } V = \dfrac{\partial}{\partial t}\phi((1-t)x_0 + tx_1), A = \dfrac{\partial^2}{\partial t^2}\phi((1-t)x_0 + tx_1) \\ \quad \textit{are continuous functions of } x_0, x_1 \in C_\eta \textit{ and } t \in [0,1]. \end{cases}$$

Then, if moreover C is small enough for the deflection condition $\theta \leq \pi/2$ to hold, x is OLS-identifiable on C, or equivalently: the NLS problem is Q-wellposed on C.

G. Chavent. *Nonlinear Least Squares for Inverse Problems: Theoretical Foundations and Step-by-Step Guide for Applications.* **Springer, 2010.**

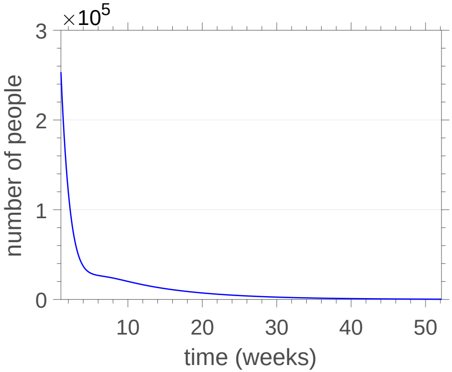
Calibrated parameters and initial conditions

$\boldsymbol{\alpha}$	TRR input	**lb**	**ub**	TRR output
α_h	1/5.9	1/12	1/3	1/12
α_v	1/9.1	1/10	1/5	1/10
γ	1/7.9	1/8.8	1/3	1/3
δ	1/11	1/21	1/11	1/21
β_h	1/11.3	1/16.3	1/8	1/10.40
β_v	1/8.6	1/11.6	1/6.2	1/7.77
S_h^i	205,953,959	$0.9 \times$ N	N	205,953,534
E_h^i	8,201	0	10,000	6,827
I_h^i	8,201	0	10,000	10,000
S_v^i	0.9996	0.99	0.999	0.999
E_v^i	2.2×10^{-4}	0	1	4.14×10^{-4}
I_v^i	2.2×10^{-4}	0	1	0

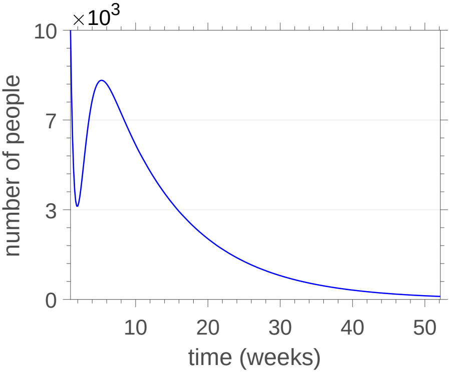
Remarks on the calibration

- Reasonable parameters
- cumulative number of infectious overshoots data by only 5.74%
- Initial infectious humans is approximately 10,000 individuals
- Peak value of new infectious cases differs from the data maximum by 10.57%
- Peak of new infectious cases occurs two weeks before the peak of the data

Comparison of infectious humans curves

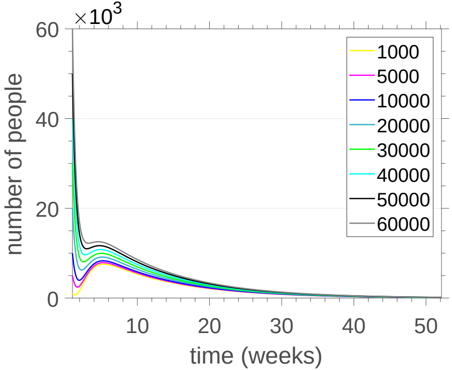


First calibration

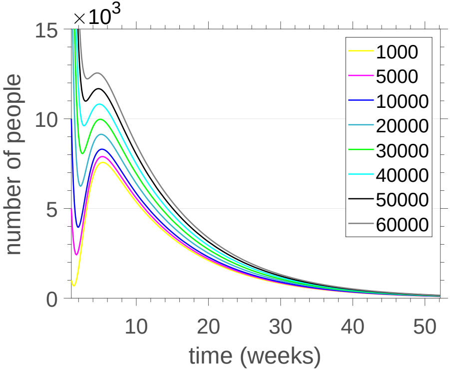


Second calibration

Comparison of infectious humans curves

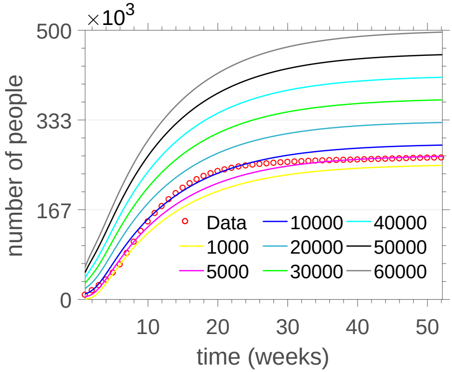


Curves for various initial infectious humans values

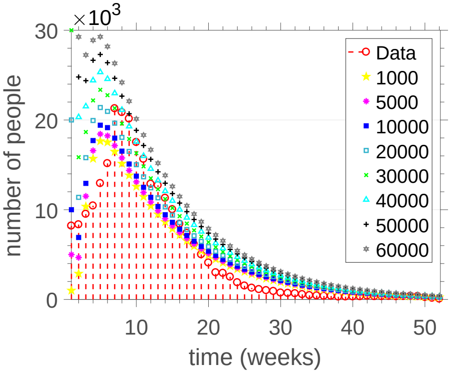


Zoom in the local peak region of the image to the left

Comparison of cumulative and new infectious curves

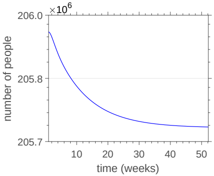


cumulative number of infectious

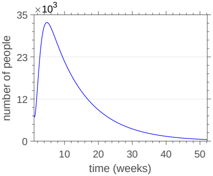


new infectious cases

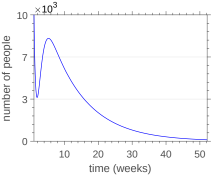
Calibrated model response



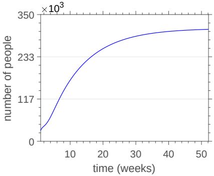
Susceptible humans



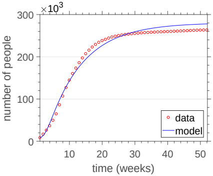
Exposed humans



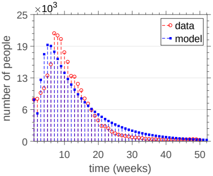
Infectious humans



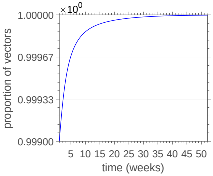
Recovered humans



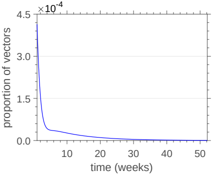
Cumulative infectious



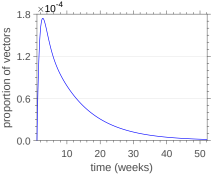
New cases



Susceptible vectors



Exposed vectors



Infectious vectors

Monte Carlo convergence

Study of convergence for MC simulation

Stochastic dynamic model:

$$\dot{\boldsymbol{U}}(t,\omega) = f(\boldsymbol{U}(\omega,t))$$

Convergence metric for of Monte Carlo simulation:

$$\mathrm{conv}(n_s) = \left(\frac{1}{n_s} \sum_{n=1}^{n_s} \int_{t_0}^{t_f} \| \boldsymbol{U}(t,\omega_n) \|^2 \; dt \right)^{1/2}$$

C. Soize, *A comprehensive overview of a non-parametric probabilistic approach of model uncertainties for predictive models in structural dynamics.* **Journal of Sound and Vibration**, 288: 623–652, 2005.

Study of convergence for MC simulation

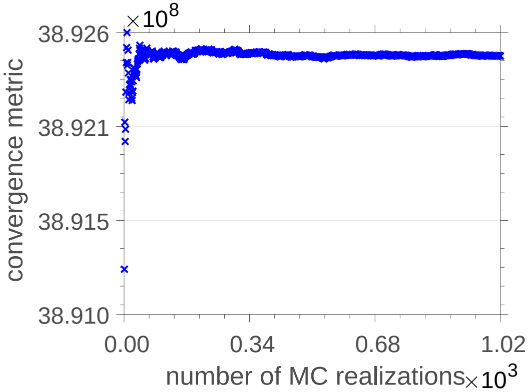


Figure: MC convergence metric as function of the number of realizations.