

Dynamics of Zika Virus in Brazil

Model Calibration and Uncertainty Quantification

Americo Cunha Jr

Universidade do Estado do Rio de Janeiro (UERJ)
NUMERICÓ – Nucleus of Modeling and Experimentation with Computers
numerico.ime.uerj.br

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

CNMAI 2017
October 23-25, 2017
Rio de Janeiro - RJ, Brazil



Outline

- 1 Introduction
- 2 SEIR-SEI Model
- 3 V&V and Calibration
- 4 Uncertainty Quantification
- 5 Final Remarks



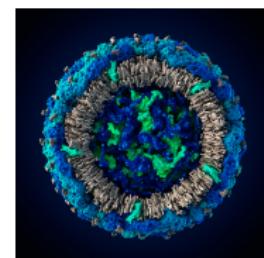
Section 1

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 130,000 confirmed cases in Brazil since 2015
- International consensus that ZIKV is a cause of:
 - Guillain–Barré syndrome
 - Microcephaly



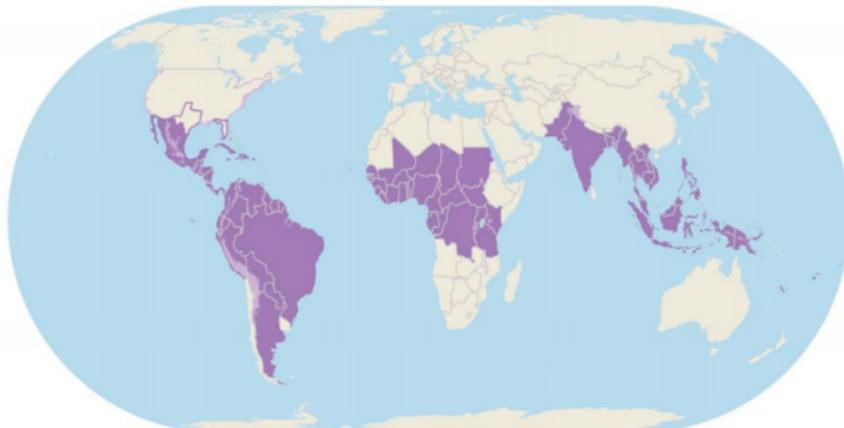
Zika virus



Aedes aegypti

Global outbreak of Zika virus

World Map of Areas with Risk of Zika



International areas and US territories

- [Dark Purple] Areas with risk of Zika infection (below 6,500 feet)
- [Medium Purple] Areas with low likelihood of Zika infection (above 6,500 feet)
- [Light Beige] Areas with no known risk of Zika infection

United States areas

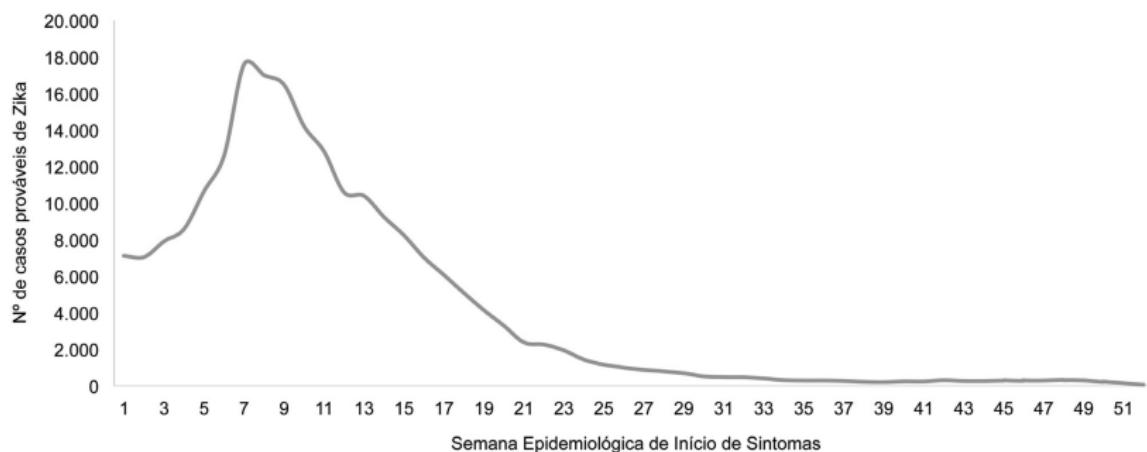
- [Light Purple] State Reporting Zika
- [Light Beige] No Known Zika



Centers for Disease Control and Prevention, *World Map of Areas with Risk of Zika, August 2017.*

Zika virus outbreak in Brazil

New cases in Brazil by epidemiological week of 2016



Secretaria de Vigilância em Saúde. *Zika virus - Boletim Epidemiológico v. 48, n. 14, 2017.*

ISSN: 2358-9450



Dengue virus (DENV)

- Mainly spread by *Aedes* mosquitoes as well as Zika virus
- More than 219,000 probable cases reported in Brazil until Epidemiological Week 35 of 2017
- 310 confirmed deaths until EW 35 of 2017 or in investigation



Dengue virus



Aedes aegypti

Research objectives

The objectives of this research are:

- Develop an epidemic model to describe the Zika virus outbreak of 2016 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

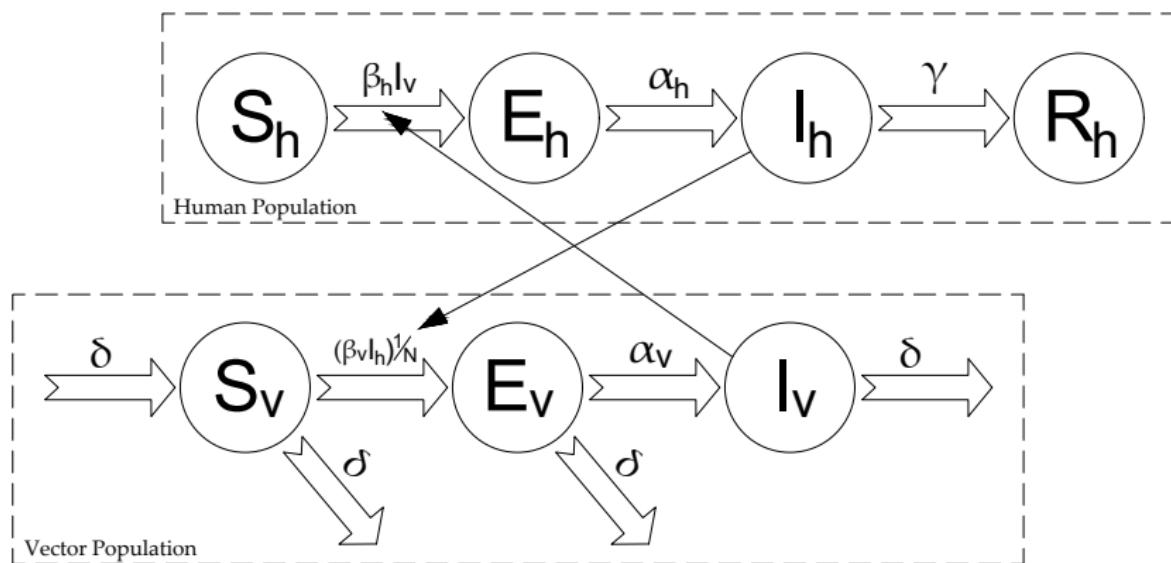


Section 2

SEIR-SEI Model



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.



Dynamic system associated to the SEIR-SEI model

$$\frac{dS_h}{dt} = -\beta_h S_h I_v$$

$$\frac{dS_v}{dt} = \delta - \beta_v S_v \frac{I_h}{N} - \delta S_v$$

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

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

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

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

$$\frac{dR_h}{dt} = \gamma I_h$$

$$\frac{dC}{dt} = \alpha_h E_h$$

+ initial conditions

S - Population of susceptible

E - Population of exposed

I - Population of infected

R - Population of recovered

N - Population of humans

C - Infected humans cumulative

α - Incubation ratio

δ - Vector lifespan ratio

β - Transmission rate

γ - Recovery rate

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.



SEIR-SEI model parameters

Model parameters and outbreak data are obtained from:

- open scientific literature



- Brazilian health system

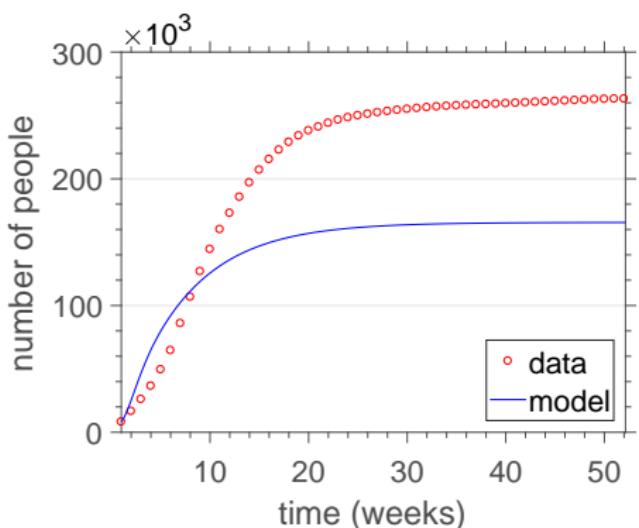


FIOCRUZ
Fundação Oswaldo Cruz

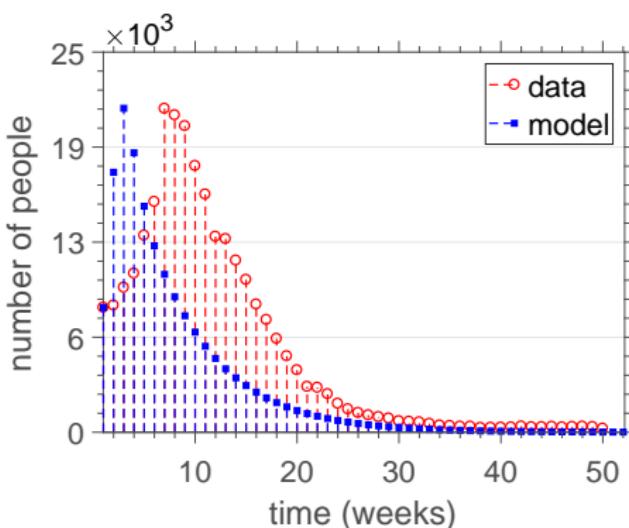
parameter	value	unit
α_h	1/5.9	days ⁻¹
α_v	1/9.1	days ⁻¹
γ	1/7.9	days ⁻¹
δ	1/11	days ⁻¹
β_h	1/11.3	days ⁻¹
β_v	1/8.6	days ⁻¹
N	206×10^6	people



Cumulative infectious and new infectious



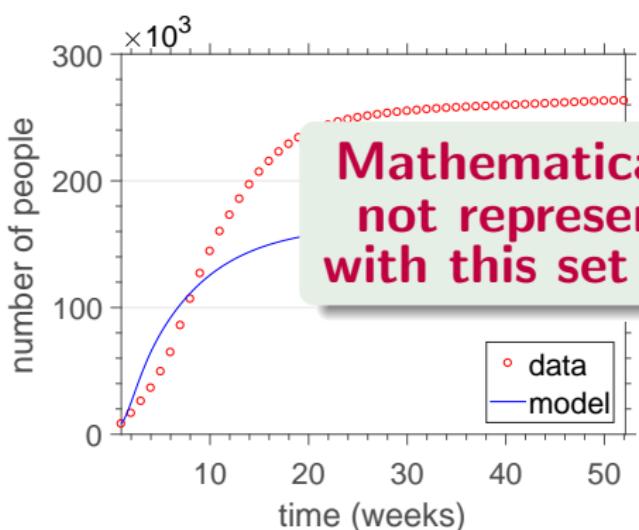
Cumulative number of infectious



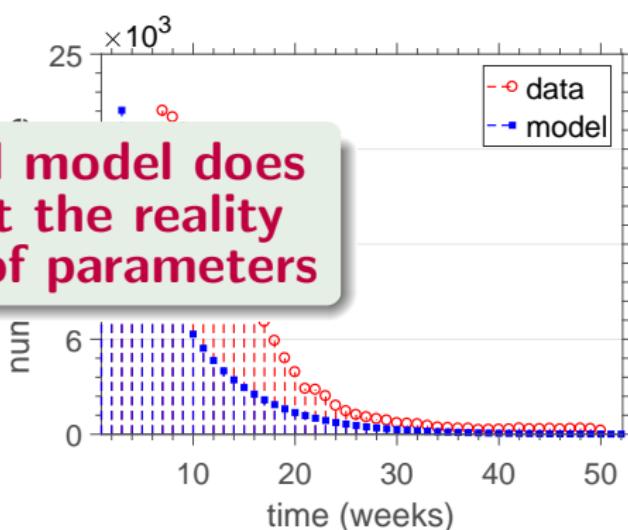
New infectious cases



Cumulative infectious and new infectious



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



Cumulative number of infectious

New infectious cases

Section 3

V&V and Calibration



Verification and Validation (V&V)

- **Verification**

Are we solving the equation *right*?

- **Validation**

Are we solving the *right* equation?



G. Iaccarino *Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods*,

VKI Lecture Series, Stanford University, 2008



Verification and Validation (V&V)

- **Verification**

Are we solving the equation *right*?

It is an exercise in *mathematics*.

- **Validation**

Are we solving the *right* equation?



G. Iaccarino *Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods*,

VKI Lecture Series, Stanford University, 2008



Verification and Validation (V&V)

- **Verification**

Are we solving the equation *right*?

It is an exercise in *mathematics*.

- **Validation**

Are we solving the *right* equation?

It is an exercise in *physics*.

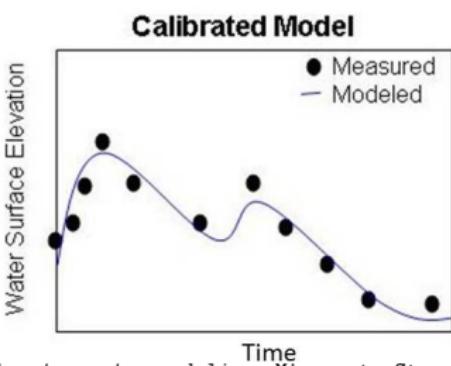
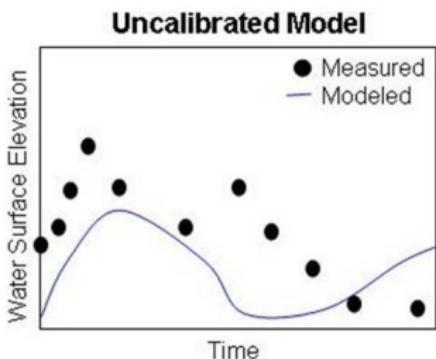


G. Iaccarino *Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods*,

VKI Lecture Series, Stanford University, 2008



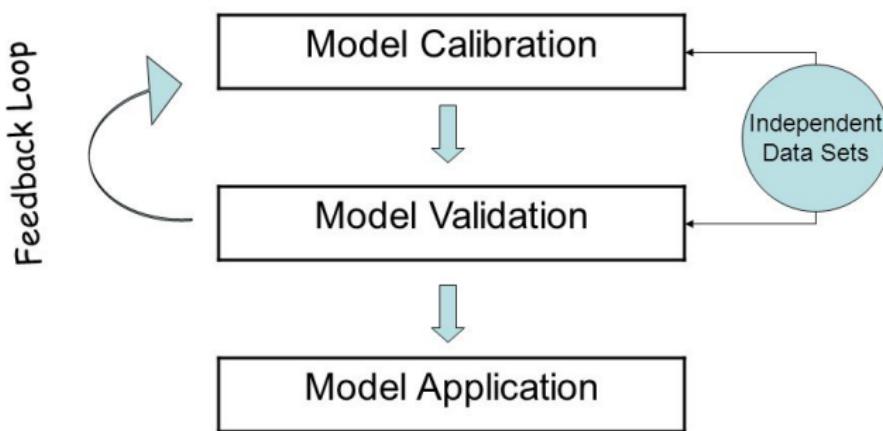
Calibration of the model



* Pictures from *Introduction to stormwater modeling*, Minnesota Stormwater Manual

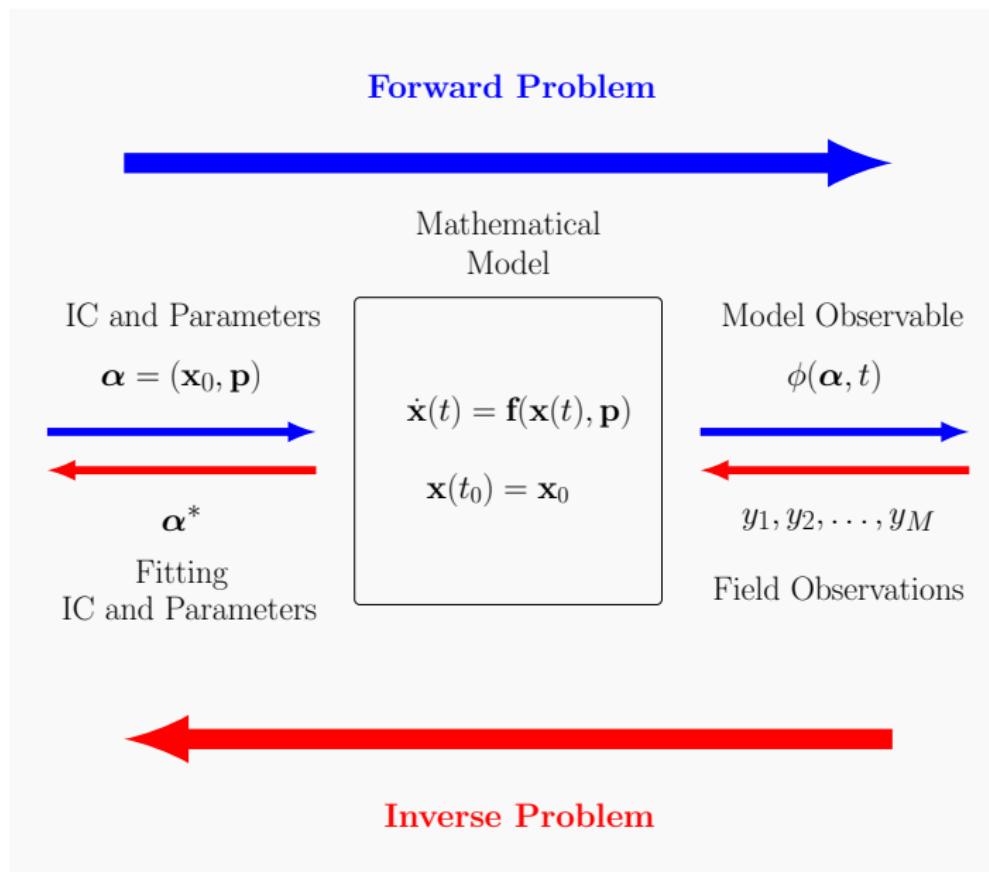
Calibration vs Validation

Model Calibration and Validation



* Picture from NHI course on Travel Demand Forecasting

Forward and inverse problem for SEIR-SEI dynamic system



An inverse problem for SEIR-SEI model calibration

Given a collection of field data

$$\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_M,$$

find a set of fitting parameters and initial conditions such that

$$\boldsymbol{\alpha}^* = \arg \min_{\boldsymbol{\alpha}} \left\{ \sum_{n=1}^M \left\| \mathbf{y}_n - \phi(\boldsymbol{\alpha}, t_n) \right\|_2^2 \right\}.$$

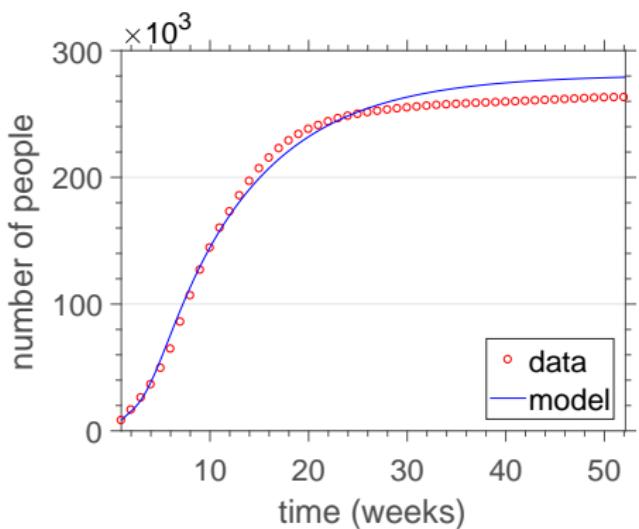
Solved using a bounded trust-region-reflective algorithm.



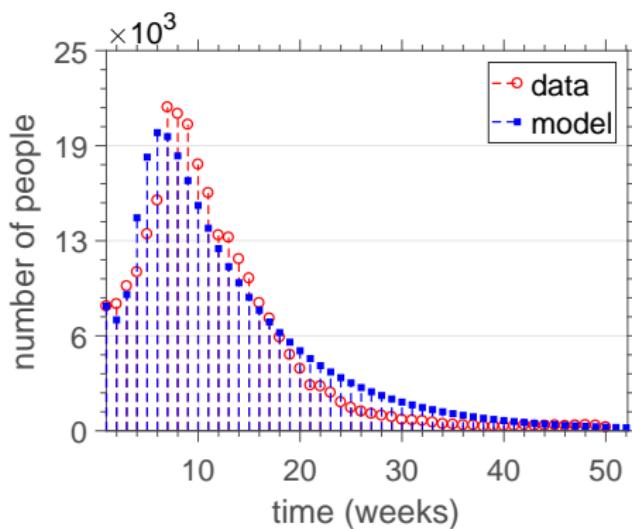
Coleman, T.F., and Li, Y., *An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds*. *SIAM Journal on Optimization*, 6:418–445, 1996.



Cumulative infectious and new infectious



Cumulative number of infectious



New infectious cases

First calibration

First calibration skills

- Reasonable parameters
- Cumulative number of infectious overshoots data by only 6%
- Peak value of New Cases differs from the data maximum by 7.87%
- Human IC sum differ 9.82% from the total population

$$S_h^i + E_h^i + I_h^i + R_h^i \neq N$$

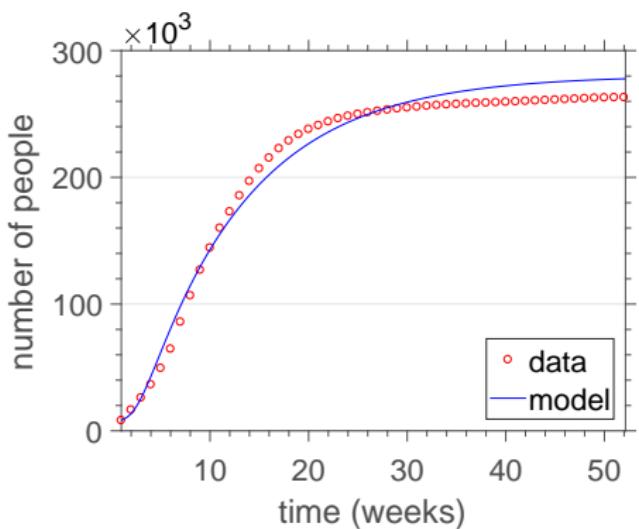
- Vector IC sum to 0.99

$$|S_h^i + E_h^i + I_h^i - 1| > \text{tol}$$

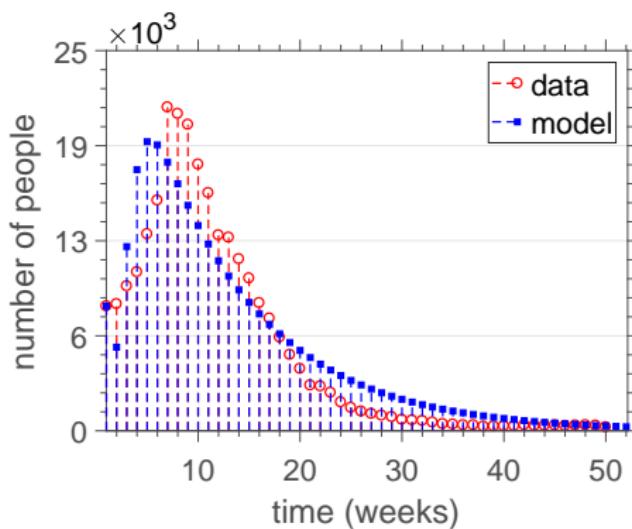
- Initial infectious humans is approximately 253,360 individuals



Cumulative infectious and new infectious



Cumulative number of infectious



New infectious cases

Second calibration



Second calibration skills

- Reasonable parameters
- Cumulative number of infectious overshoots data by only 5.74%
- Vector IC sum to 0.999

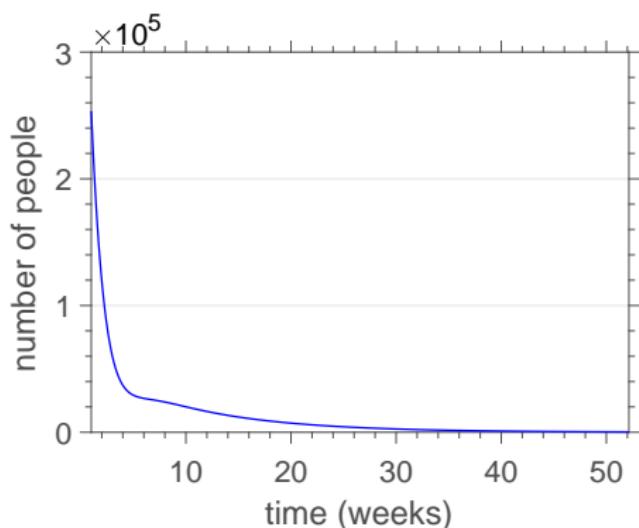
$$|S_h^i + E_h^i + I_h^i - 1| < \text{tol}$$

- Initial infectious humans is approximately 10,000 individuals
- Human IC sum is equal to total population

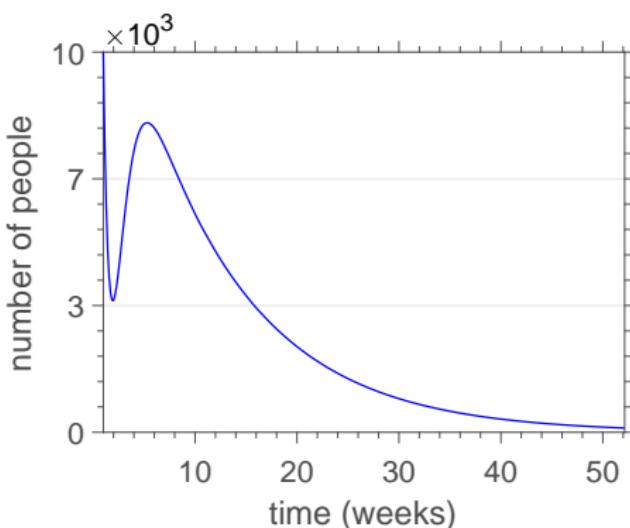
$$S_h^i + E_h^i + I_h^i + R_h^i = N$$

- 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

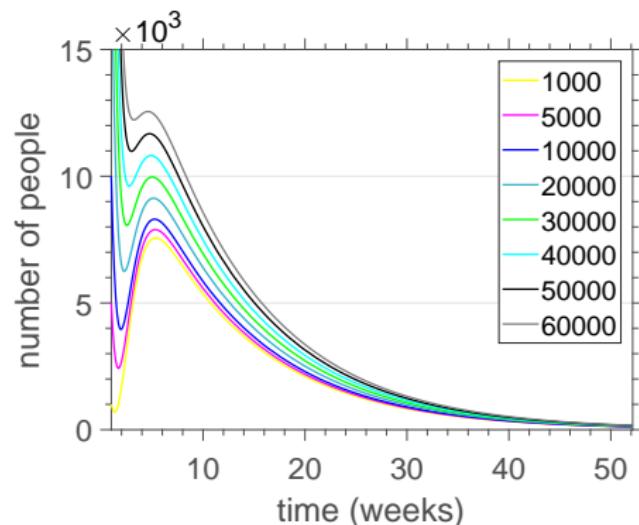
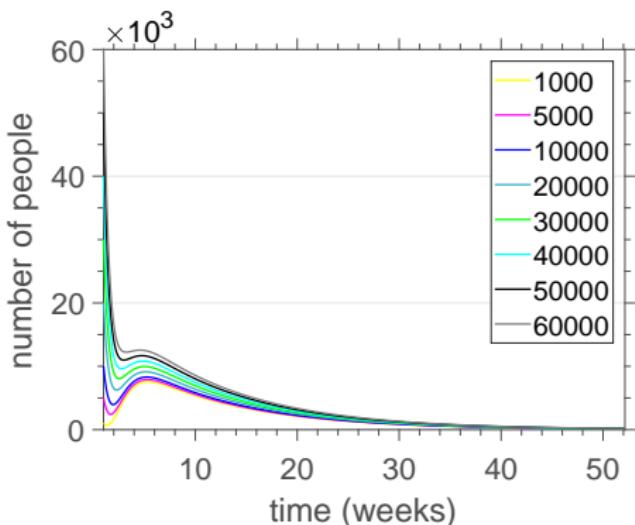


First calibration



Second calibration

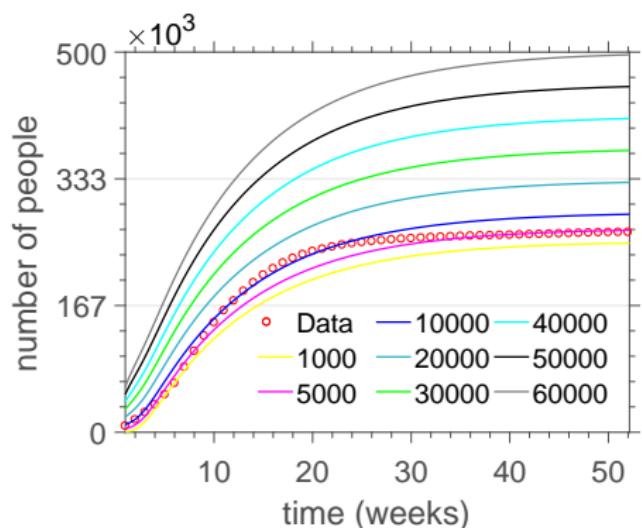
Comparison of infectious humans curves



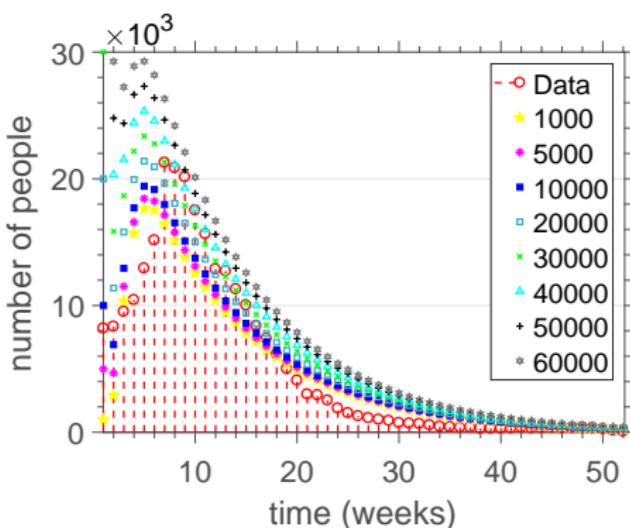
Curves for various initial infectious humans values

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

Comparison of cumulative and new infectious curves



Cumulative number of infectious



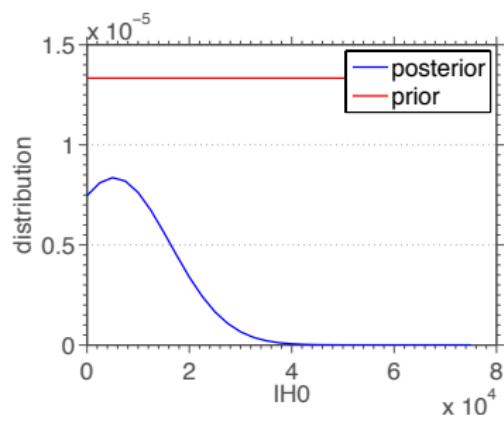
New infectious cases



Ongoing research: Bayesian updating

Bayesian formalism for model calibration:

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



uniform prior

Section 4

Uncertainty Quantification

Uncertainties in the mathematical model

The mathematical model is subjected to uncertainties:

- **model uncertainty**

- due to lack of knowledge of physics
- ignored in a first analysis

- **data uncertainty**

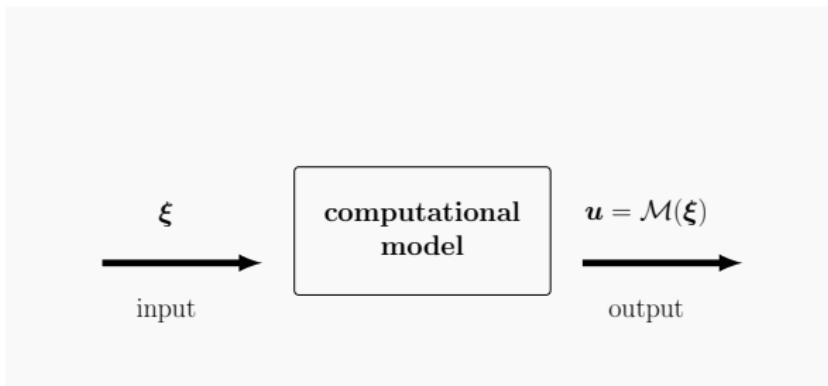
- due to model parameters variabilities
- initial infected humans
(most pronounced)



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



Computation under uncertainty



- ① Data Representation: characterize inputs uncertainties
- ② Uncertainty Propagation: quantify output uncertainties
- ③ Certification: establish acceptable levels of uncertainty

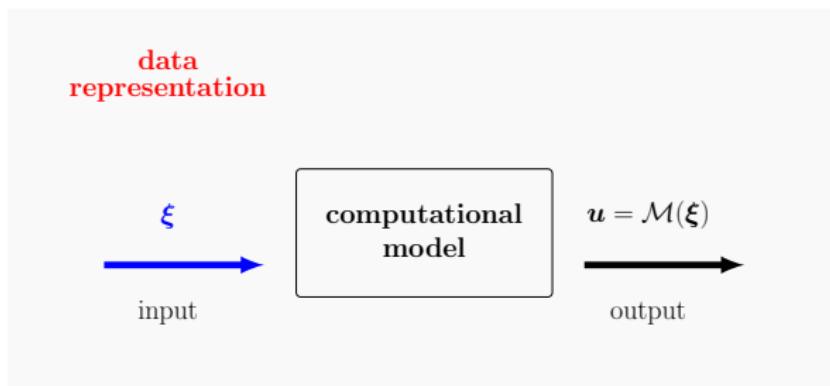


G. Iaccarino *Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods*,

VKI Lecture Series, Stanford University, 2008



Computation under uncertainty



- ① Data Representation: characterize inputs uncertainties
- ② Uncertainty Propagation: quantify output uncertainties
- ③ Certification: establish acceptable levels of uncertainty

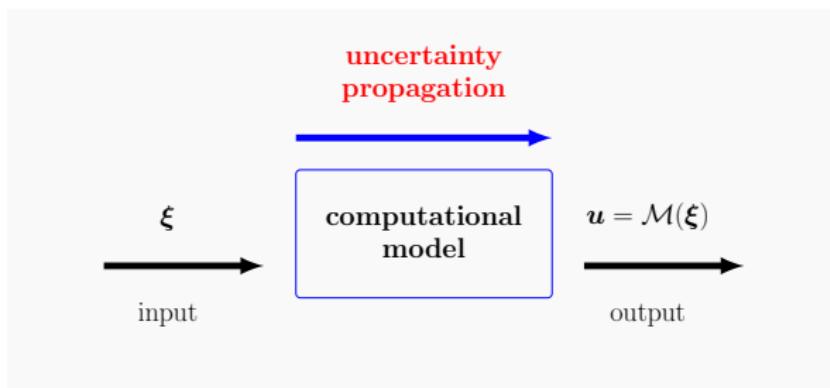


G. Iaccarino *Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods*,

VKI Lecture Series, Stanford University, 2008



Computation under uncertainty



- ① Data Representation: characterize inputs uncertainties
- ② Uncertainty Propagation: quantify output uncertainties
- ③ Certification: establish acceptable levels of uncertainty

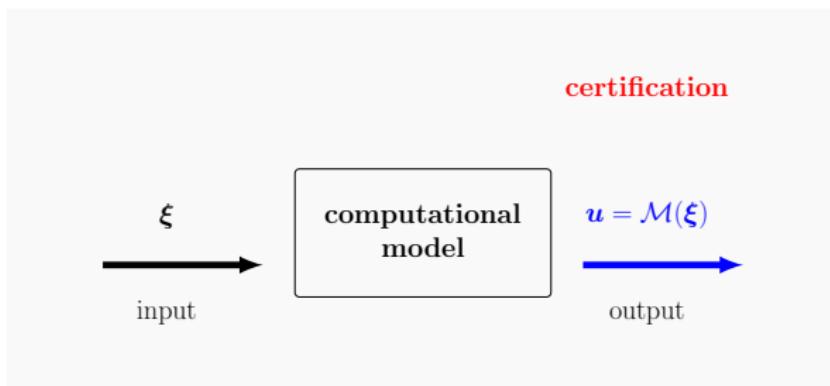


G. Iaccarino *Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods*,

VKI Lecture Series, Stanford University, 2008



Computation under uncertainty



- ① Data Representation: characterize inputs uncertainties
- ② Uncertainty Propagation: quantify output uncertainties
- ③ Certification: establish acceptable levels of uncertainty

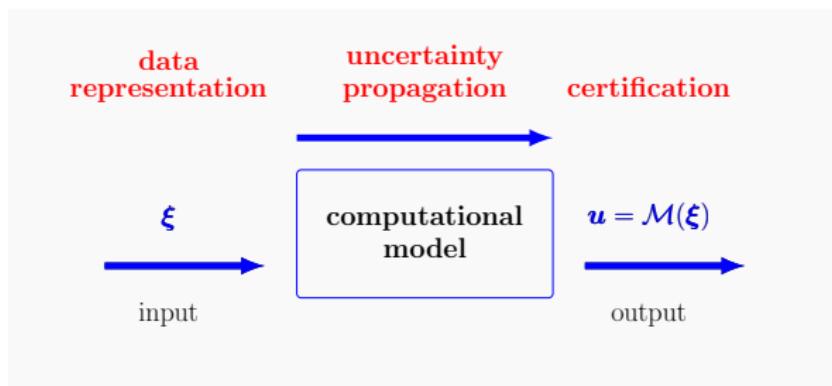


G. Iaccarino *Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods*,

VKI Lecture Series, Stanford University, 2008



Computation under uncertainty



- ① Data Representation: characterize inputs uncertainties
- ② Uncertainty Propagation: quantify output uncertainties
- ③ Certification: establish acceptable levels of uncertainty



G. Iaccarino *Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods*,

VKI Lecture Series, Stanford University, 2008



Construction of a stochastic model

Parametric probabilistic approach: $(\Omega, \Sigma, \mathcal{P})$

- initial infected humans I_{h0} is a random variable

$$\omega \in \Sigma \mapsto X(\omega) \in \mathbb{R}$$

- model response is a random process

$$(\omega, t) \in \Sigma \times \mathcal{T} \mapsto \mathbf{U}(\omega, t) \in \mathbb{R}^8$$

- model response is function of model parameter

$$\mathbf{U} = h(X)$$



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

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

(measure for the level of uncertainty)



MaxEnt optimization problem

Known information about X:

$$\text{Supp } p_X = [a, b] \subset [0, +\infty) \implies \int_a^b p_X(x) dx = 1$$

Constrained optimization problem:

Maximize

$$\mathcal{S}(p_X) = - \int_a^b p_X(x) \ln(p_X(x)) dx,$$

such that

$$\int_a^b p_X(x) dx = 1.$$



MaxEnt optimization problem

Lagrangian:

$$\mathcal{L}(p_X, \lambda_0) = - \int_{x=a}^b p_X(x) \ln(p_X(x)) dx - (\lambda_0 - 1) \left(\int_a^b p_X(x) dx - 1 \right)$$



MaxEnt optimization problem

Lagrangian:

$$\mathcal{L}(p_X, \lambda_0) = - \int_{x=a}^b p_X(x) \ln(p_X(x)) dx - (\lambda_0 - 1) \left(\int_a^b p_X(x) dx - 1 \right)$$

Necessary conditions for an extreme:

$$\frac{\partial \mathcal{L}}{\partial p_X}(p_X, \lambda_0) = 0$$

and

$$\frac{\partial \mathcal{L}}{\partial \lambda_0}(p_X, \lambda_0) = 0$$



MaxEnt optimization problem

$$\frac{\partial \mathcal{L}}{\partial p_X} (p_X, \lambda_0) = 0 \quad \Rightarrow \quad p_X(x) = \mathbb{1}_{[a,b]}(x) e^{-\lambda_0}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_0} (p_X, \lambda_0) = 0 \quad \Rightarrow \quad \int_{x=a}^b p_X(x) dx = 1$$

MaxEnt optimization problem

$$\frac{\partial \mathcal{L}}{\partial p_X} (p_X, \lambda_0) = 0 \quad \Rightarrow \quad p_X(x) = \mathbb{1}_{[a,b]}(x) e^{-\lambda_0}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_0} (p_X, \lambda_0) = 0 \quad \Rightarrow \quad \int_{x=a}^b p_X(x) dx = 1$$

Hence,

$$\int_{x=a}^b \mathbb{1}_{[a,b]}(x) e^{-\lambda_0} dx = 1 \quad \Rightarrow \quad e^{-\lambda_0} = \frac{1}{b-a},$$



MaxEnt optimization problem

$$\frac{\partial \mathcal{L}}{\partial p_X} (p_X, \lambda_0) = 0 \quad \Rightarrow \quad p_X(x) = \mathbb{1}_{[a,b]}(x) e^{-\lambda_0}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_0} (p_X, \lambda_0) = 0 \quad \Rightarrow \quad \int_{x=a}^b p_X(x) dx = 1$$

Hence,

$$\int_{x=a}^b \mathbb{1}_{[a,b]}(x) e^{-\lambda_0} dx = 1 \quad \Rightarrow \quad e^{-\lambda_0} = \frac{1}{b-a},$$

so that

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



MaxEnt optimization problem

$$\frac{\partial \mathcal{L}}{\partial p_X}(p_X, \lambda_0) = 0 \quad \Rightarrow \quad p_X(x) = \mathbb{1}_{[a,b]}(x) e^{-\lambda_0}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_0}(p_X, \lambda_0) = 0 \quad \Rightarrow \quad \int_{x=a}^b p_X(x) dx = 1$$

Hence,

$$\int_{x=a}^b \mathbb{1}_{[a,b]}(x) e^{-\lambda_0} dx = 1 \quad \Rightarrow \quad e^{-\lambda_0} = \frac{1}{b-a},$$

so that

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

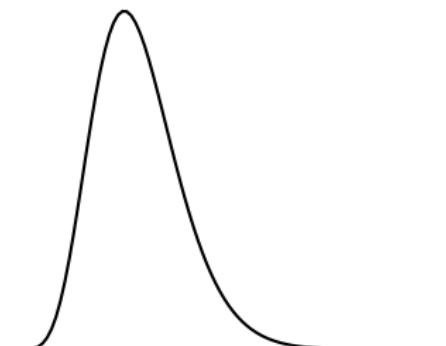
X has uniform distribution on [a, b].



Philosophy of MaxEnt Principle

- The parameter of interest has a unknown distribution

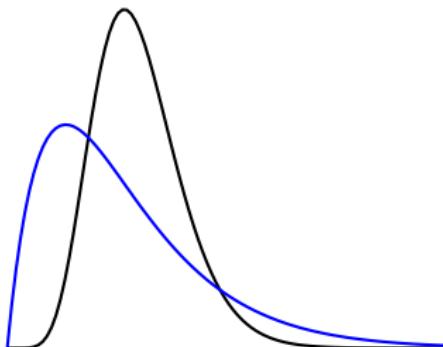
— real



Philosophy of MaxEnt Principle

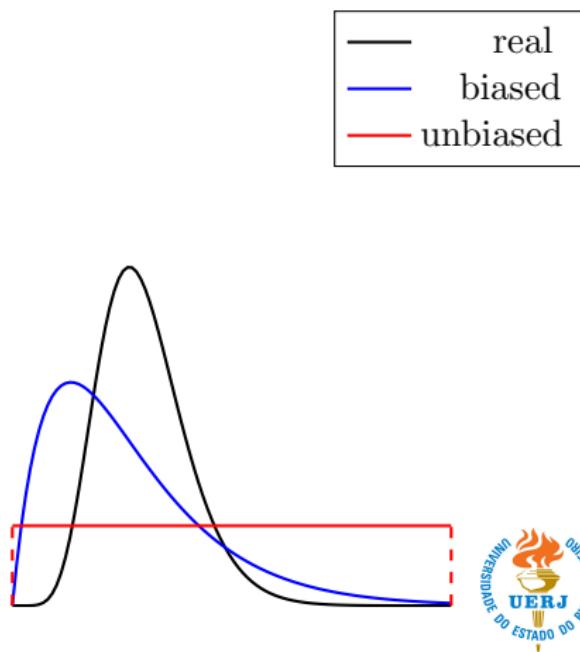
- The parameter of interest has a unknown distribution
- Distributions arbitrarily chosen can be coarse and biased

— real
— 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

$$\boldsymbol{X}_1$$

\vdots

$$\boldsymbol{X}_M$$



known $F_{\boldsymbol{X}}$

solution of
model equations

$$\boldsymbol{U} = h(\boldsymbol{X})$$

**computational
model**

computation
of statistics

$$\boldsymbol{U}_1 = h(\boldsymbol{X}_1)$$

\vdots

$$\boldsymbol{U}_M = h(\boldsymbol{X}_M)$$



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}

Study of convergence for MC simulation

Stochastic dynamic model:

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

Convergence metric for Monte Carlo simulation:

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



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

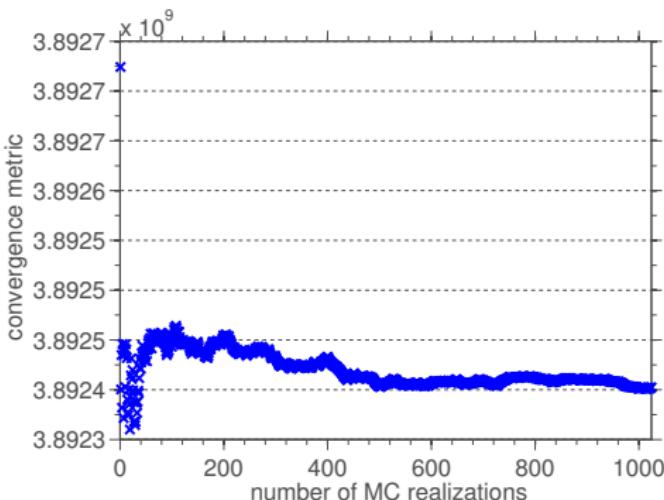
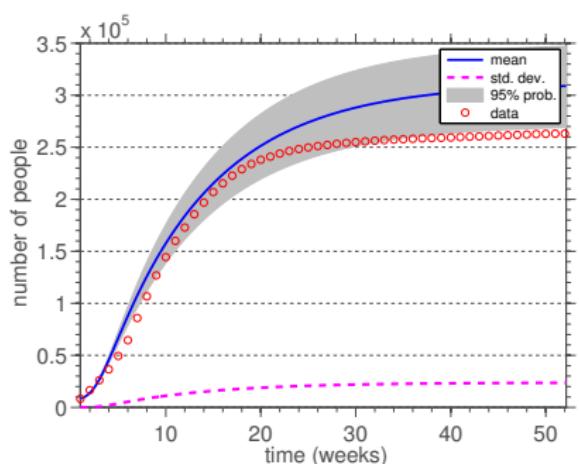


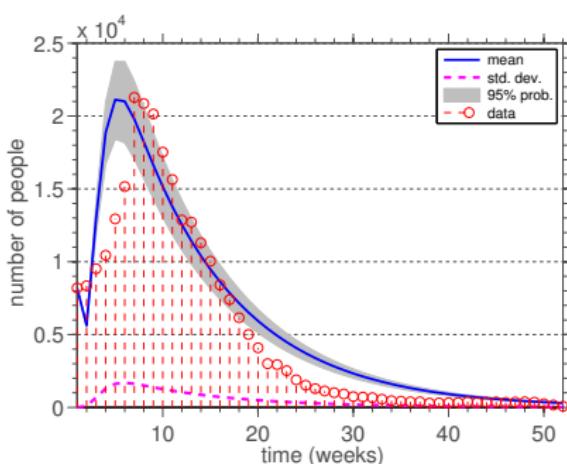
Figure: MC convergence metric as function of the number of realizations.

Cumulative and new infectious confidence band

$$X \sim Uniform(0, 75) \times 10^3$$



Cumulative number of infectious



New infectious cases

Section 5

Final Remarks

Concluding remarks

Contributions and conclusions:

- Development of a SEIR-SEI epidemic model to describe Zika virus outbreak in Brazil
- Calibration of this SEIR-SEI model with epidemic real data
- Construction of a probabilistic model to describe uncertainties

Ongoing research:

- Bayesian updating to improve the model calibration
- Describe model and other parameters uncertainties

Future directions:

- Design of experiments via Active Subspace
- Data-driven identification of epidemiological models

Acknowledgments

Financial support given to this research:



Uncertainties 2018



<http://icvramisuma2018.org>



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, 2017* (under review).



Number of new cases of infectious per epidemiological week



Number of new cases per week

A set of 52 points to represent the number of new infectious cases of Zika fever at each week is defined as follows:

$$\begin{aligned}\mathcal{N}_w &= C_w - C_{w-1}, \\ \mathcal{N}_1 &= C_1,\end{aligned}$$

where C_w is the cumulative number of infectious humans in the w th epidemiological week.



Nominal parameters and initial conditions



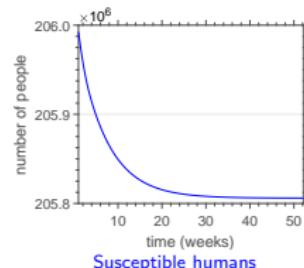
Nominal parameters and initial conditions

α	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}	—

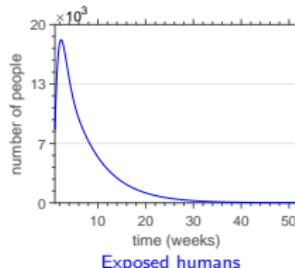


Curves for nominal inputs

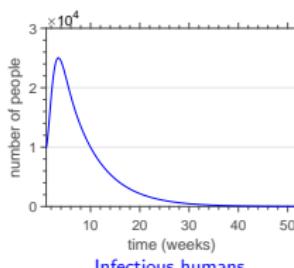
Curves for nominal inputs



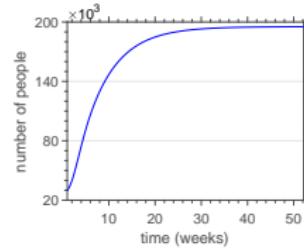
Susceptible humans



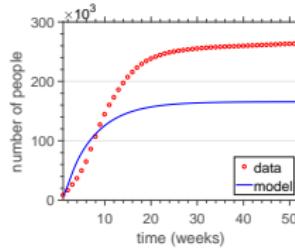
Exposed humans



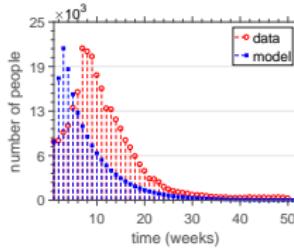
Infectious humans



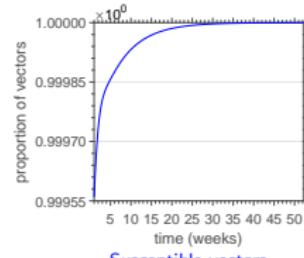
Recovered humans



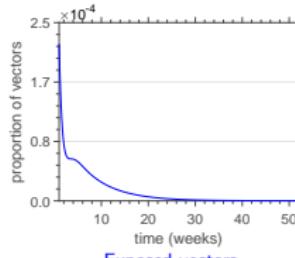
Cumulative infectious



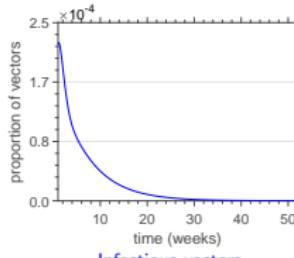
New cases



Susceptible vectors



Exposed vectors



Infectious vectors



First calibration parameters and initial conditions

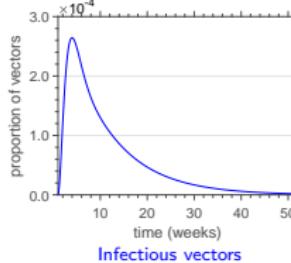
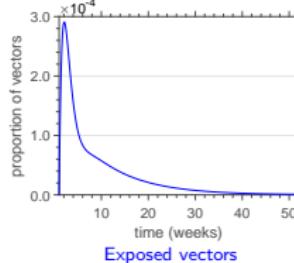
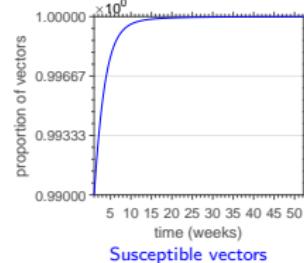
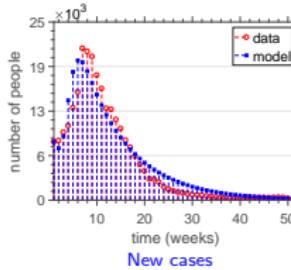
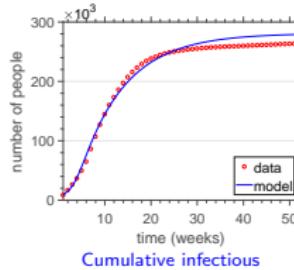
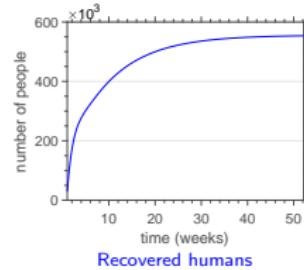
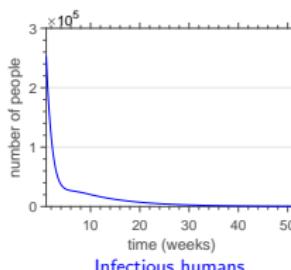
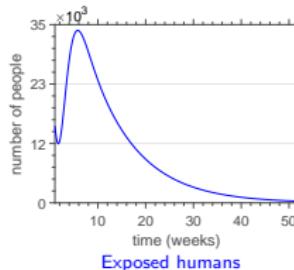
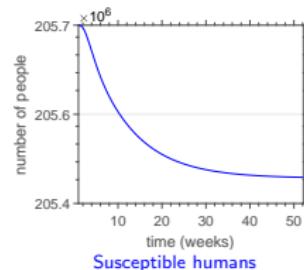


First calibration parameters and IC

α	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	1/8.8
δ	1/11	1/21	1/11	1/16.86
β_h	1/11.3	1/16.3	1/8	1/16.3
β_v	1/8.6	1/11.6	1/6.2	1/11.6
S_h^i	205,953,959	$0.9 \times N$	N	205,700,000
E_h^i	8,201	0	N	15,089
I_h^i	8,201	0	N	253,360
S_v^i	0.99956	0.99	0.999	0.99
E_v^i	2.2×10^{-4}	0	1	0
I_v^i	2.2×10^{-4}	0	1	0

Curves for first calibration inputs

Curves for first calibration inputs



Second calibration parameters and initial conditions

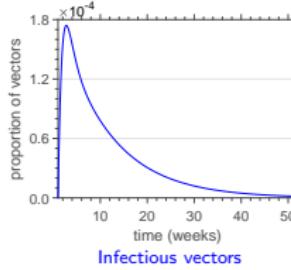
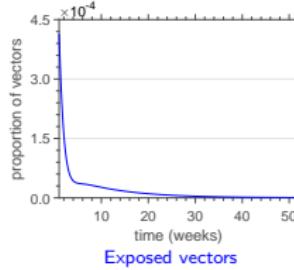
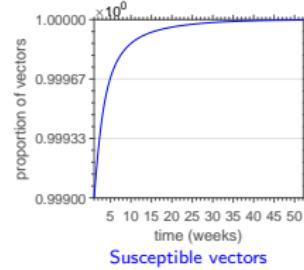
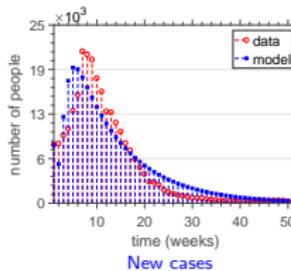
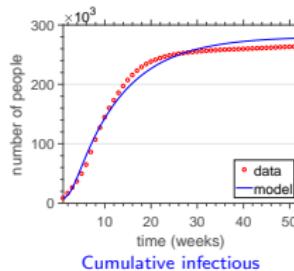
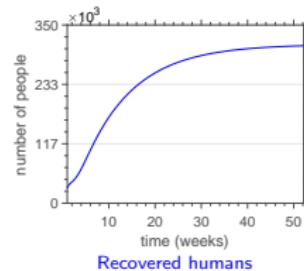
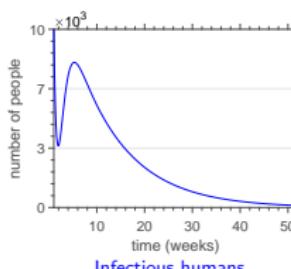
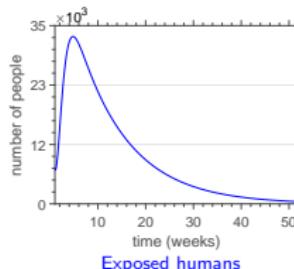
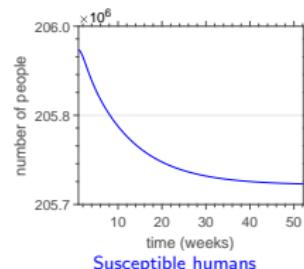


Second calibration parameters and IC

α	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

Curves for second calibration inputs

Curves for first second inputs



References from images and data



References from images and data

-  Zika Virus 3D Model by *visual-science. com* at [goo.gl/CwHe6v](https:// goo.gl/CwHe6v)
-  Hi-resolution female Aedes aegypti mosquito by *CDC* at [goo.gl/WxWrjz](https:// goo.gl/WxWrjz)
-  World Map of Areas with Risk of Zika, April 2017, by *CDC* at [goo.gl/5U6pdL](https:// goo.gl/5U6pdL)
-  Zika Virus 3D Model by *hhmi. org* at [goo.gl/6MvNMP](https:// goo.gl/6MvNMP)
-  Aedes aegypti mosquito by *denguevirusnet. com* at [goo.gl/gHXoSA](https:// goo.gl/gHXoSA)
-  New Cases data in Brazil/2016 by *Secretaria de Vigilância em Saúde* at [goo.gl/JbLzWL](https:// goo.gl/JbLzWL)