"First-principles machine learning modelling of COVID-19" - A Research Case Study

Siva Nanda Reddy Pokala
Danielle Gagné
Chaitanyabhargavi Gonuguntla

Universität Trier Department of Mathematics
Supervisor: Professor Dr. Volker Schulz

March 4, 2021

Table of Contents

COVID-19 Data

SIRD Model

The Neural Network

Overview of the Process

Results

Table of Contents

COVID-19 Data

SIRD Model

The Neural Network

Overview of the Process

Results

COVID-19 Data

- COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins Unversity
- URL: https://github.com/CSSEGISandData/COVID-19/
- Data includes number of:
  1)total Confirmed cases
  2)Recovered cases
  3)Death cases
  of each day for 188 countries
- Generating new column for active cases with name Infected
  Infected = Confirmed - Recovered - Deaths

COVID-19 Data

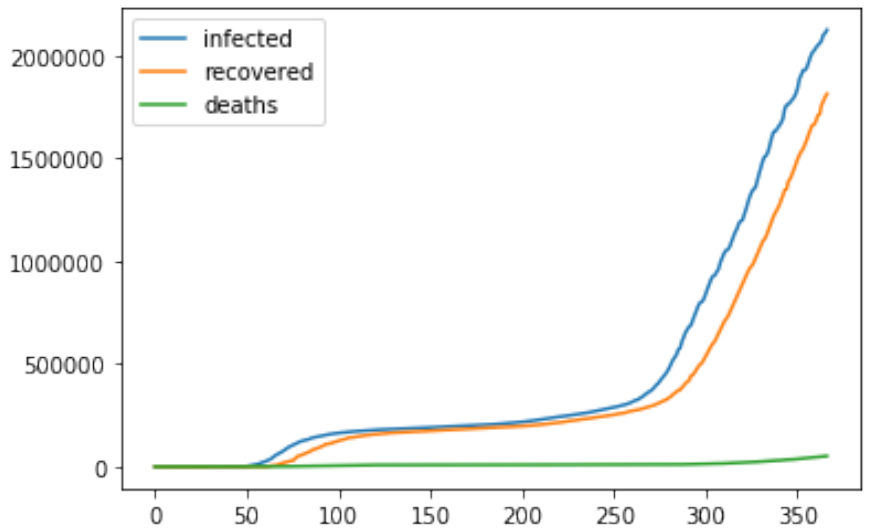


Figure: Cases in Germany

Table of Contents

COVID-19 Data

SIRD Model

The Neural Network

Overview of the Process

Results

The SIRD model

- To describe the behaviour of an epidemic, we consider a compartmentalized SIRD model to fit the data (groups are disjoint)
- Population N = S+I+R+D remains constant

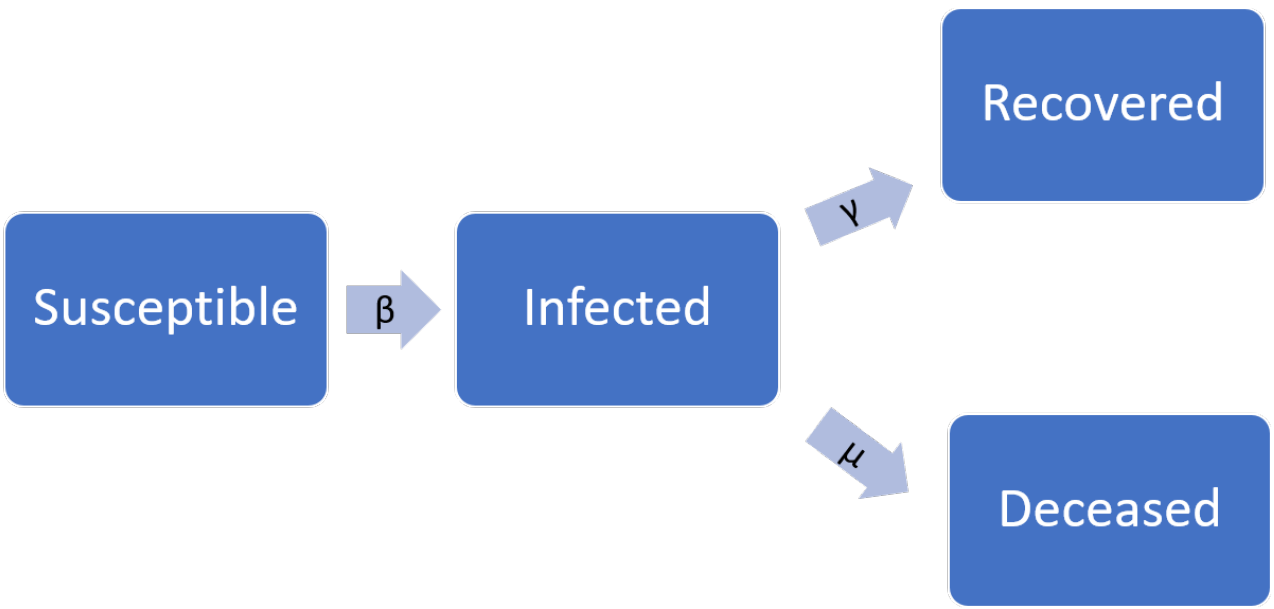


Figure: Flow through the SIRD Model

The SIRD model

- The change in the SIRD model over time is represented using a system of ODEs:

$$\begin{aligned}
\frac{\partial S}{\partial t} &= \frac{-\beta I S}{N} \\
\frac{\partial I}{\partial t} &= \frac{\beta I S}{N} - (\mu + \gamma) I \\
\frac{\partial R}{\partial t} &= \gamma I \\
\frac{\partial D}{\partial t} &= \mu I
\end{aligned}$$

- Where β, γ, and μ are rates of infection, recovery, and death, respectively.
- β, γ, and μ are time dependent.

Table of Contents

COVID-19 Data

SIRD Model

The Neural Network

Overview of the Process

Results

The Neural Network

- Objective: learn parameters β, γ, and μ, as well as the SIRD model state over time

- A fully connected neural network with 3 hidden layer and sigmoid activation function.
- Input size: 2(N_t), Output size: 3(N_t)
- First hidden layer with 32 neurons
- Second hidden layer with 16 neurons
- Third hidden layer with 8 neurons
- Output is reshaped into (N_t,3)

Initialization

- Run a trial using Optuna hyperparameter optimization framework to output a constant guess of β_0, γ_0, μ_0
- β_0, γ_0, μ_0 is found by minimizing the below equation

$$\sum_{t=0}^{N_t}(I_c(t) - \hat{I}(t))^2 + 100\sum_{t=0}^{N_t}(D_c(t) - \hat{D}(t))^2$$

- Initialization used mean-squared-error with Adam optimizer (stochastic gradient descent), 3000 iterations

The Discrete Time SIRD Model

- We solve the ODE system numerically using Euler's method with step size 1 in order to time advance our data. This results in a discrete time SIRD model:

$$
\begin{aligned}
S(t+1) &= S(t) - \frac{\beta(t)I(t)S(t)}{N} \\
I(t+1) &= I(t) + \frac{\beta(t)I(t)S(t)}{N} - (\mu(t)+\gamma(t))I(t) \\
R(t+1) &= R(t) + \gamma(t)I)(t) \\
D(t+1) &= D(t) + \mu(t)I)(t)
\end{aligned}
$$

- with I(0) and D(0) being the number of confirmed and deaths of the first confirmed day, respectively
- S(0) = N - I(0) - D(0) - R(0)

Loss Function and Training

- The loss function L of the neural network is given by

$$L = \sum_{t=0}^{N_t} \Big(\big(\log(I_c(t) - \log(\hat{I}(t)) \big)^2 + \big(log(D_c(t) - log(\hat{D}(t)) \big)^2 \Big) +$$

$$0.01 \frac{log(max(I_c))}{max(I_c)} \sum_{t=0}^{N_t} \Big(\big((I_c(t) - (\hat{I}(t)) \big)^2 + \big((D_c(t) - (\hat{D}(t)) \big)^2 \Big) +$$

$$100 \frac{log(max(I_c))}{max(\alpha_0)} \sum_{t=0}^{N_t - 1} (\hat{\beta}(t) - \hat{\beta}(t+1)^2 + (\hat{\gamma}(t) - \hat{\gamma}(t+1)^2$$

$$+ 100(\hat{\mu}(t) - \hat{\mu}(t+1)^2 +$$

$$100 \frac{log(max(I_c))}{max(\alpha_0)} \Big((\hat{\beta}(0) - \beta_0)^2 + (\hat{\gamma}(0) - \gamma_0)^2 + 100(\hat{\mu}(0) - \mu_0)^2 \Big)$$

- Magri and Doan (2020)

Loss function and training

- We trained on the countries Germany, Australia, India, Canada, Russia, and China
- We trained the network using Adam optimizer, with learning rate 5x10^{-5} with 3000 iterations and $1x10^{-5}$ with 3000 iterations for $N_t = 300$

Table of Contents

COVID-19 Data

SIRD Model

The Neural Network

Overview of the Process

Results

Architecture Overview

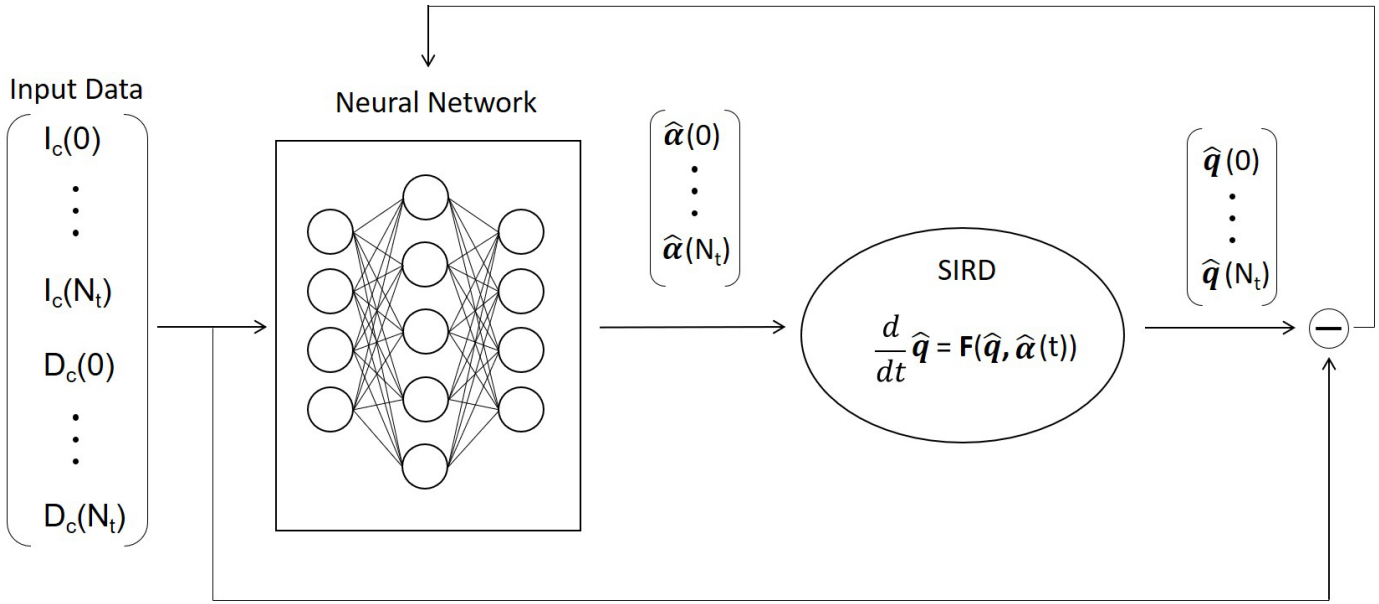


Figure: Pictorial representation of machine learning for COVID-19

- $I_c(t)$: Infected = Confirmed - Recovered - Deaths
- $D_c(t)$: Deaths
- Predicted parameters: $\hat{\boldsymbol{\alpha}}(t) = [\hat{\beta}(t), \hat{\gamma}(t), \hat{\mu}(t)]^T$
- Predicted state: $\hat{\boldsymbol{q}}(t) = [\hat{S}(t), \hat{I}(t), \hat{R}(t), \hat{D}(t)]^T$

Table of Contents

COVID-19 Data

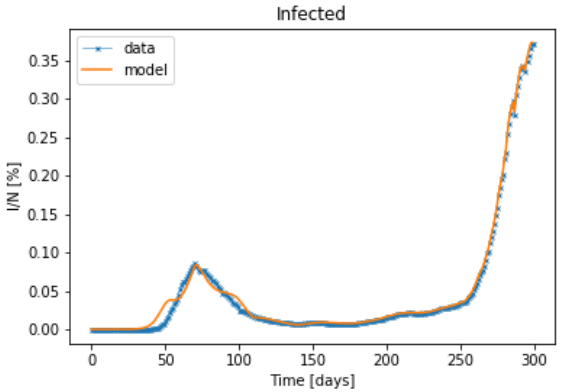
SIRD Model

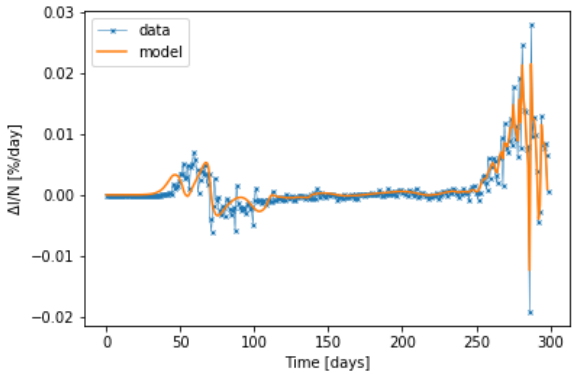
The Neural Network

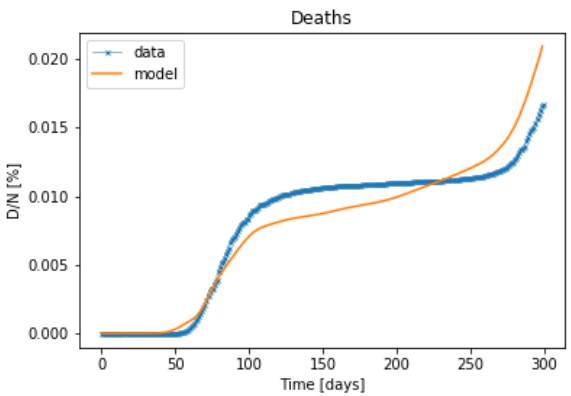
Overview of the Process

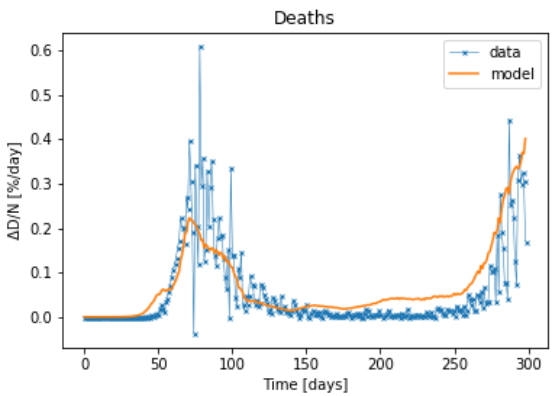
Results

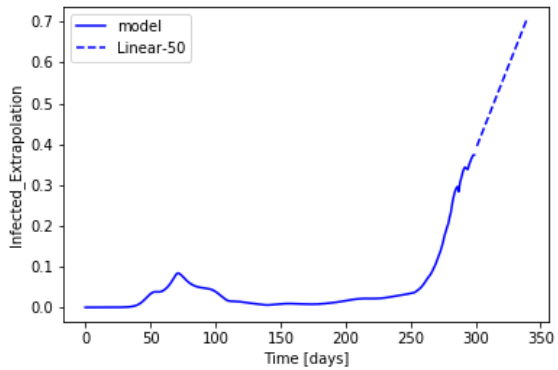
Results (Germany)

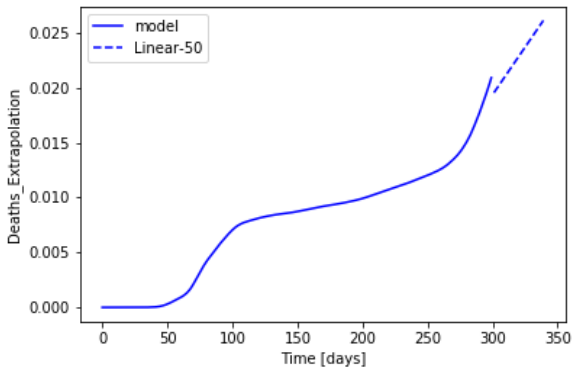


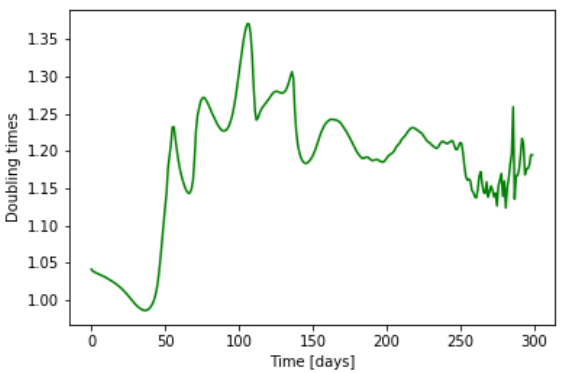


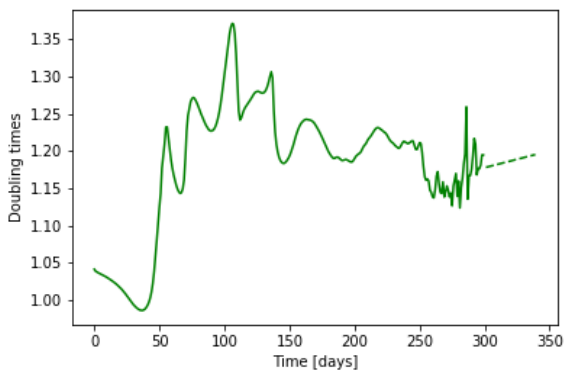




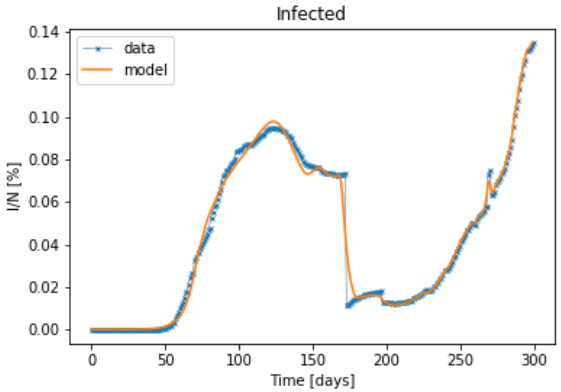

model
Linear-50
Infected_Extrapolation
Time [days]

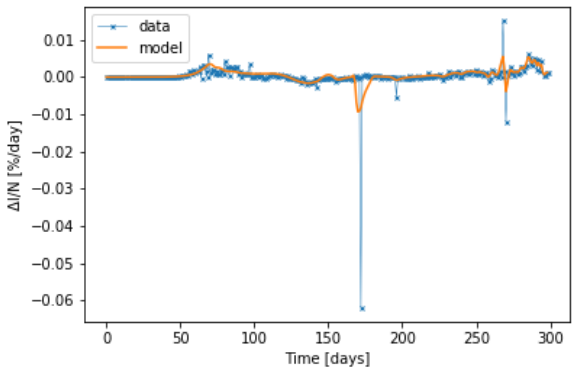

model
Linear-50
Deaths_Extrapolation
Time [days]

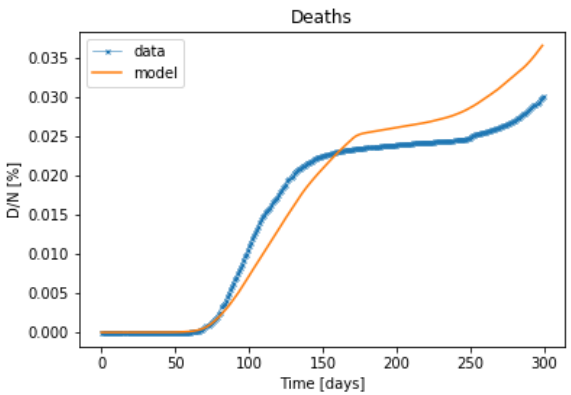

Doubling times
Time [days]

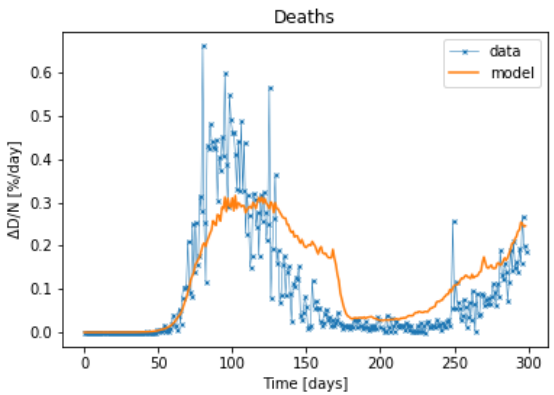

Doubling times
Time [days]

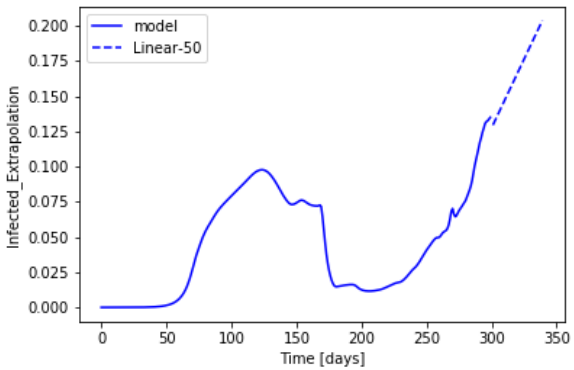
Results (Canada)

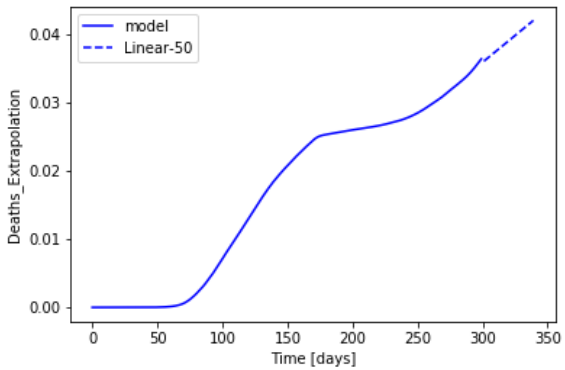


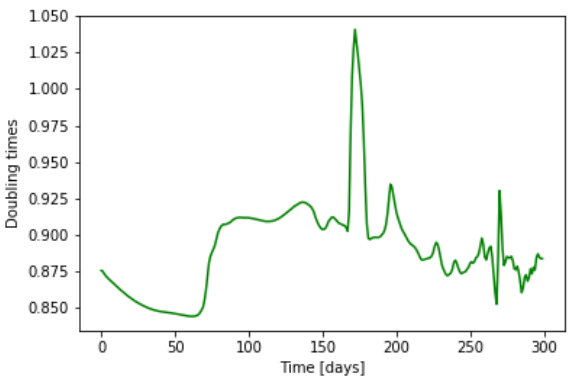


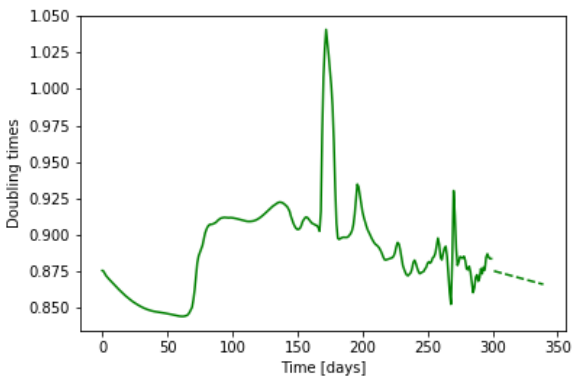




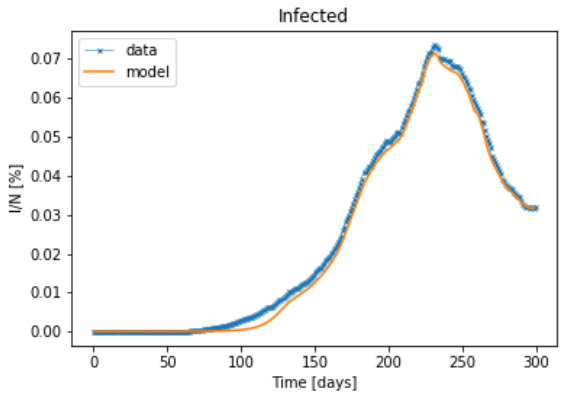

model
Linear-50
Infected_Extrapolation
Time [days]

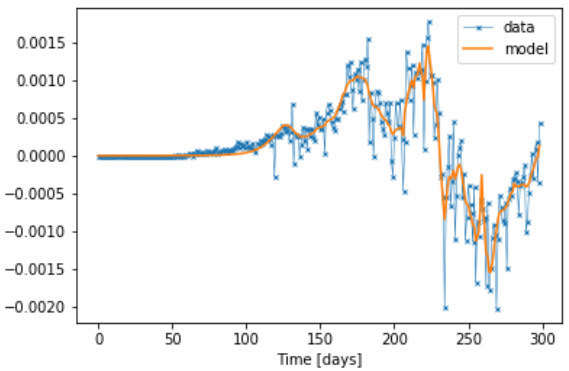

model
Linear-50
Deaths_Extrapolation
Time [days]

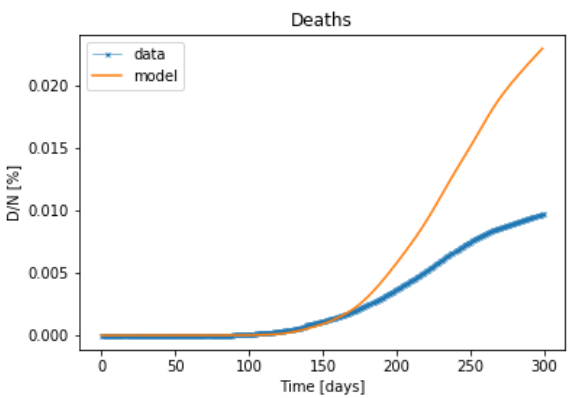

Doubling times
Time [days]

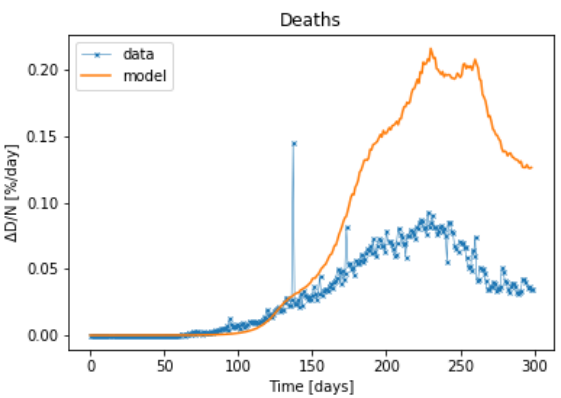

Doubling times
Time [days]

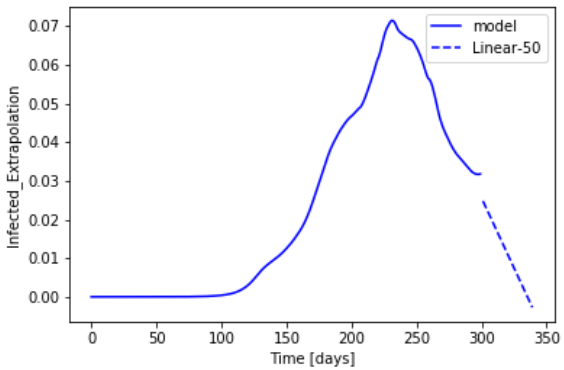
Results (India)

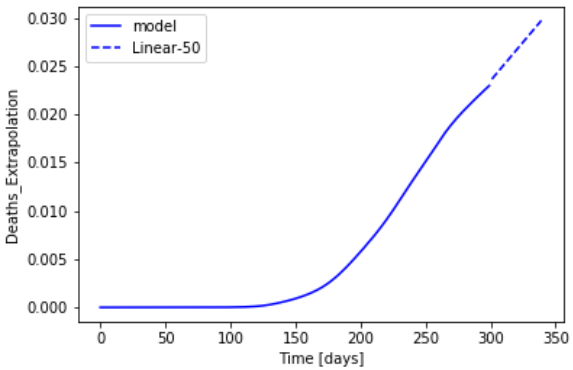


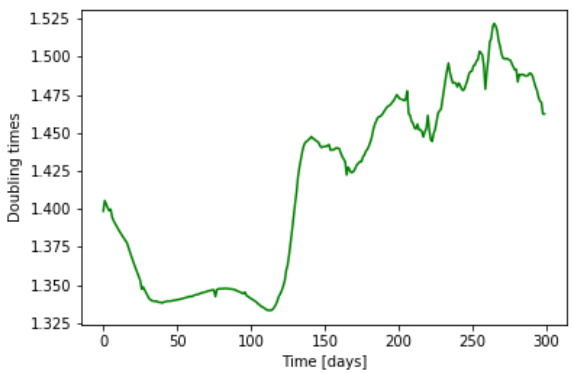


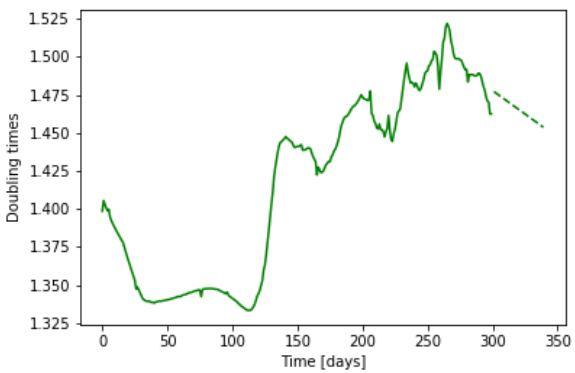





model
Linear-50
Infected_Extrapolation
Time [days]


model
Linear-50
Deaths_Extrapolation
Time [days]


Doubling times
Time [days]


Doubling times
Time [days]

Thank you for your attention