Saturation in SIR model

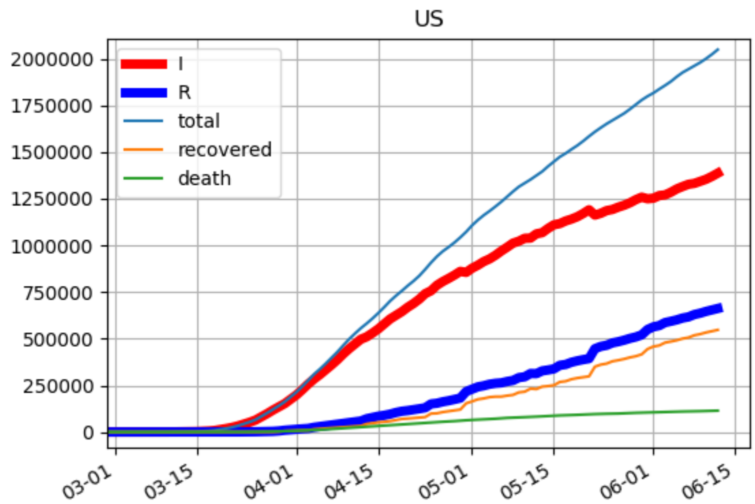
Theodore Kolokolnikov

Dalhousie University

- SIR model:

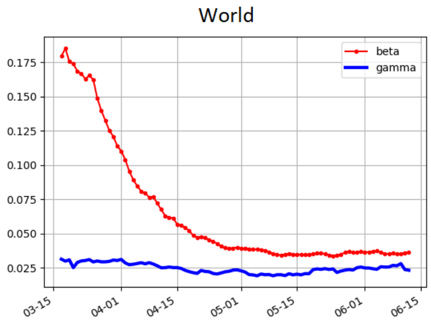
$$\frac{dS}{dt} = -\frac{\beta}{N}SI; \qquad \frac{dI}{dt} = \frac{\beta}{N}SI - \gamma I; \qquad \frac{dR}{dt} = \gamma I \tag{1}$$

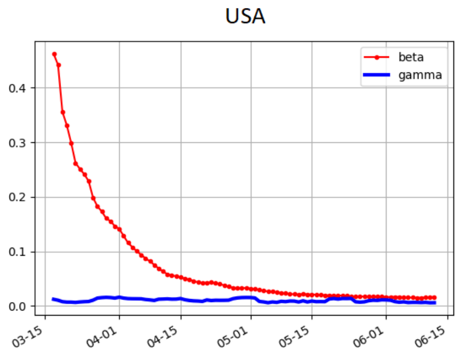
- Here is US data:

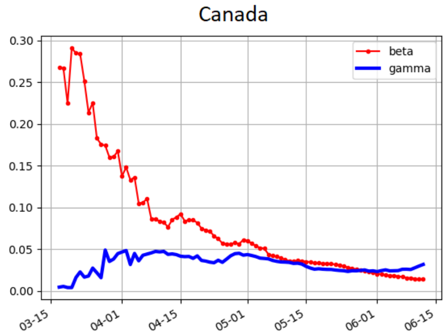


- Does SIR model fit coronavirus data?
  - Think of $\beta = \beta(t)$ and $\gamma = \gamma(t)$.
  - Estimate $\beta(t)\ \gamma(t)$ from equations and data: $\beta(t) \approx -\dfrac{NS'(t)}{SI}; \quad \gamma(t) \sim \dfrac{R'}{I}$

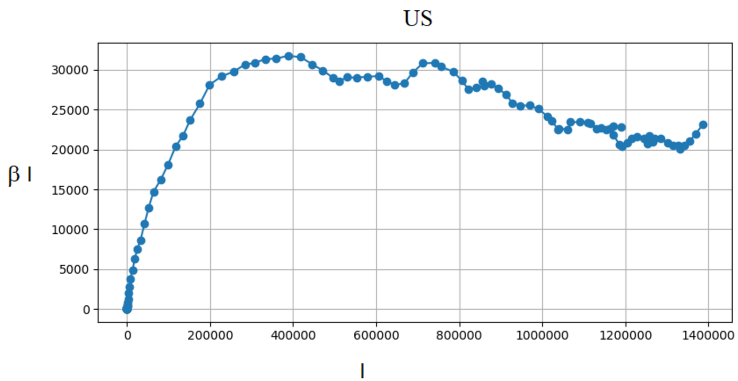
- Here is an example of plot of $\beta(t), \ \ \gamma(t)$ from data:







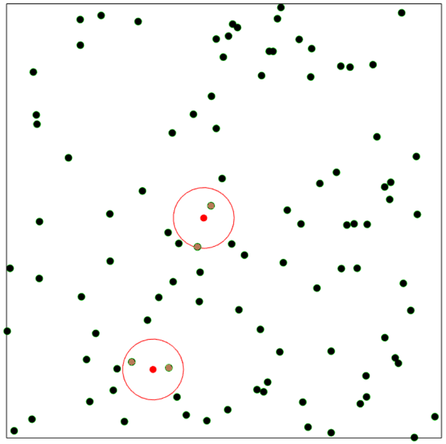
- Note that γ doesn't change much, but $\beta(t)$ decreases steadily throughout the infection.
- The decrease of $\beta(t)$ is consistent for multiple datasets.
- Alternatively, we can plot βI as a function of I :



- Note how βI saturates as I increases. This suggests that some saturation may occur.

Agent-based SIR model

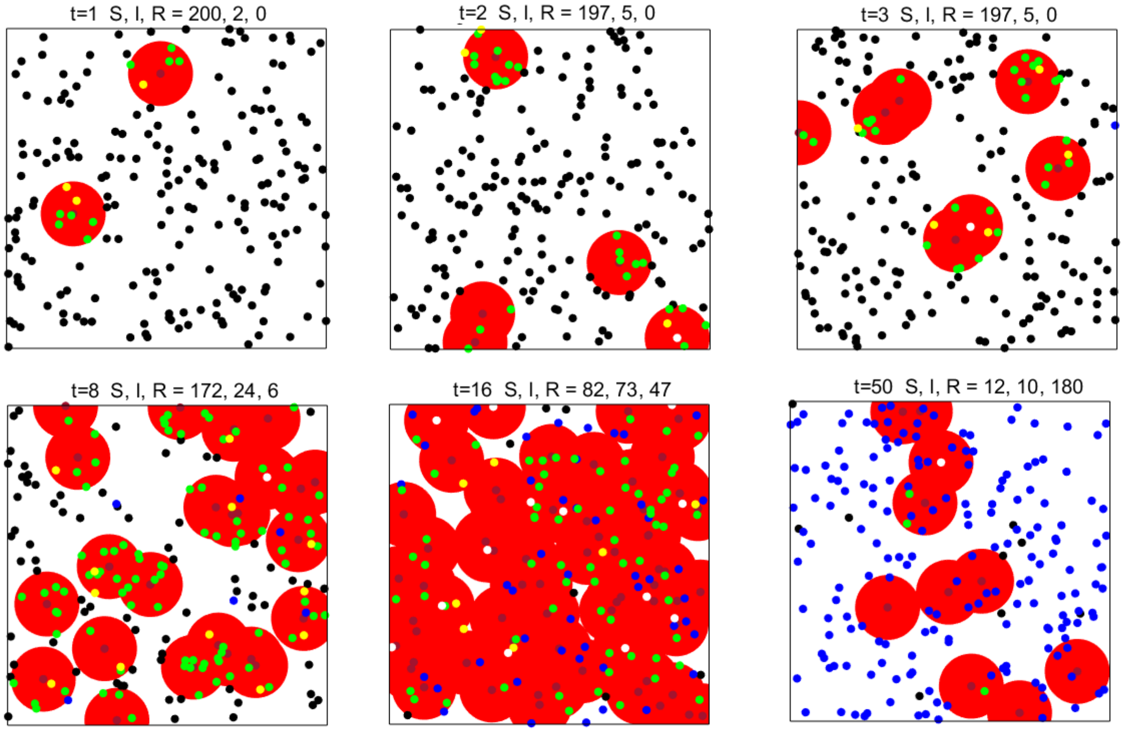
- Assume there is an "infection radius" r; infection can only happen when susceptible comes within the distance r of infected individual

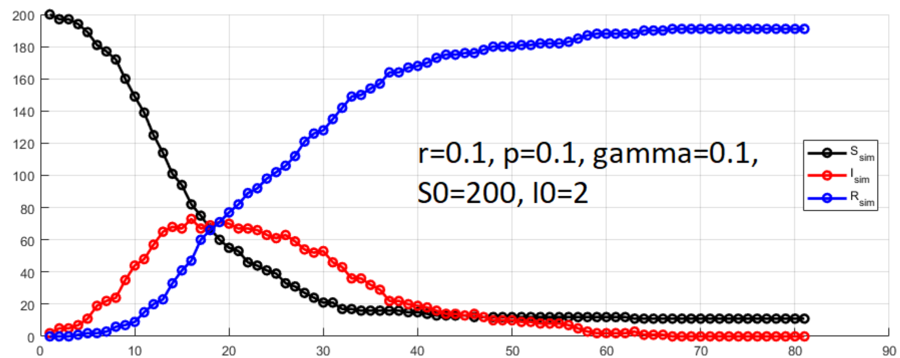


- An infection occurs with some probability p if a susceptible person comes within a radius r of the infected person

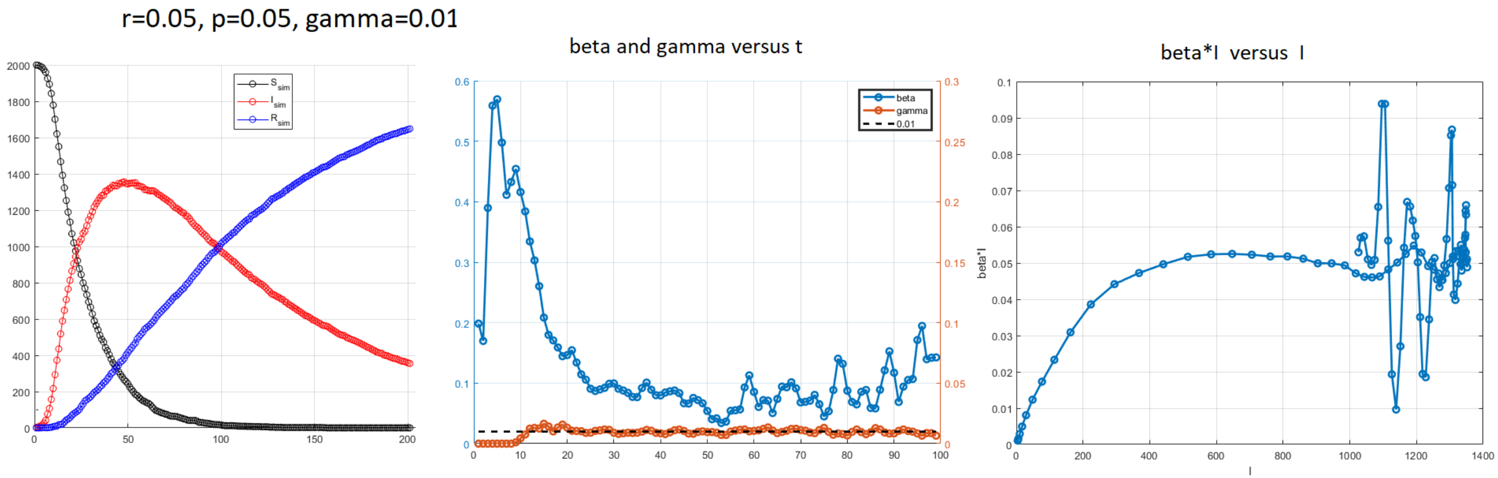
- Assume people are "well-mixed": all are located at random within a unit box.

- Infected individuals are removed with some rate γ.


t=1 S, I, R = 200, 2, 0
t=2 S, I, R = 197, 5, 0
t=3 S, I, R = 197, 5, 0
t=8 S, I, R = 172, 24, 6
t=16 S, I, R = 82, 73, 47
t=50 S, I, R = 12, 10, 180


r=0.1, p=0.1, gamma=0.1,
S0=200, I0=2
S_{sim}
I_{sim}
R_{sim}
0
20
40
60
80
100
120
140
160
180
200
0
10
20
30
40
50
60
70
80
90

- Let's estimate $\beta \approx -\dfrac{S'(t)}{SI}; \quad \gamma \sim \dfrac{R'}{I}$ we get this curve:



- Gamma is roughly constant ≈ 0.01; but β is clearly non-constant.
- There is a clear saturation of βI, when plotted versus I.
- This is similar to observed data for COVID in US, Canada, World and others.
- The reason is that when enough people are infected, they "cover" the whole area so introducing even more infecteds doesn't

Continuum limit of ABM

- Given I infected individuals, the probability of infection is: $p * E$, where E is the probability that you are inside one of I circles.

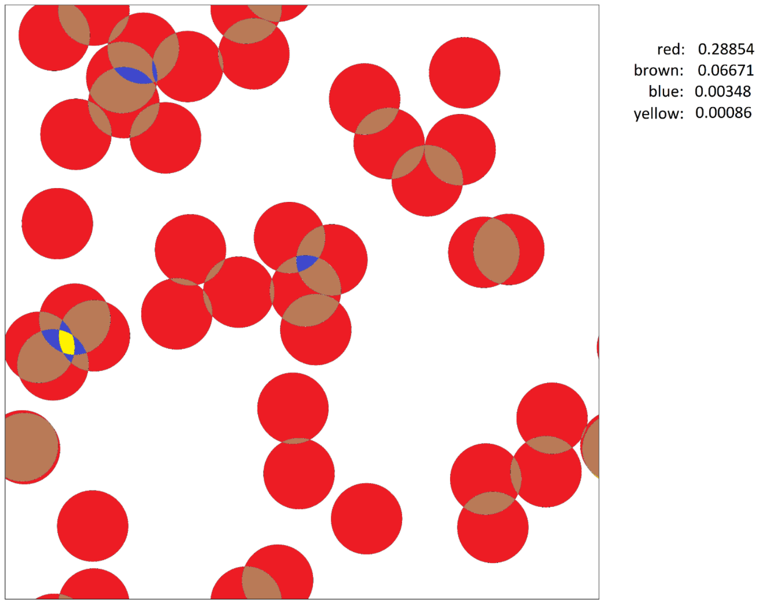
- E is the fraction of the area occupied by the union of I disks

- For small I and small r, we expect $E \approx aI$, where $a = \frac{\pi r^2}{A}$ is the area fraction for a single disk of radius r inside a domain of area A.

- But as I is increased, some cicles intersect, so eventually E saturates to 1.

- ***Key question: what is the*** **expected** ***area fraction of I disks?***

E : expected area fraction of union of I disks



- Add I areas, then substract $\binom{I}{2}$ pairwise intersections, then add back $\binom{I}{3}$ triple intersections....
- $E = IE_1 - \binom{I}{2} E_2 + \binom{I}{3} E_3 + \ldots - (-1)^I E_I$ where:

- $E_1 = a = \frac{\pi r^2}{A}$ is the area fraction of a single disk

- E_2 is the expected area fraction of intersection of two random disks of radius a

- etc

- CLAIM: $E_j = a^j$

  - Proof: either by direct integration, or by scaling argument: $E_j = C\left(a\right)^j$ for small a. Also $E_j = 1$ when $a = 1$, so $C = 1$.

- So we have obtain: $E = -\sum_{j=1}^{I} \begin{pmatrix} I \\ j \end{pmatrix} (-a)^j = 1 - (1-a)^I$.

- **Conclusion:** $E = 1 - (1-a)^I \sim (1 - \exp\left(-Ia\right))$, where $a = \frac{\pi r^2}{A}$ is the area fraction for a single disk of radius r inside a domain of area A.

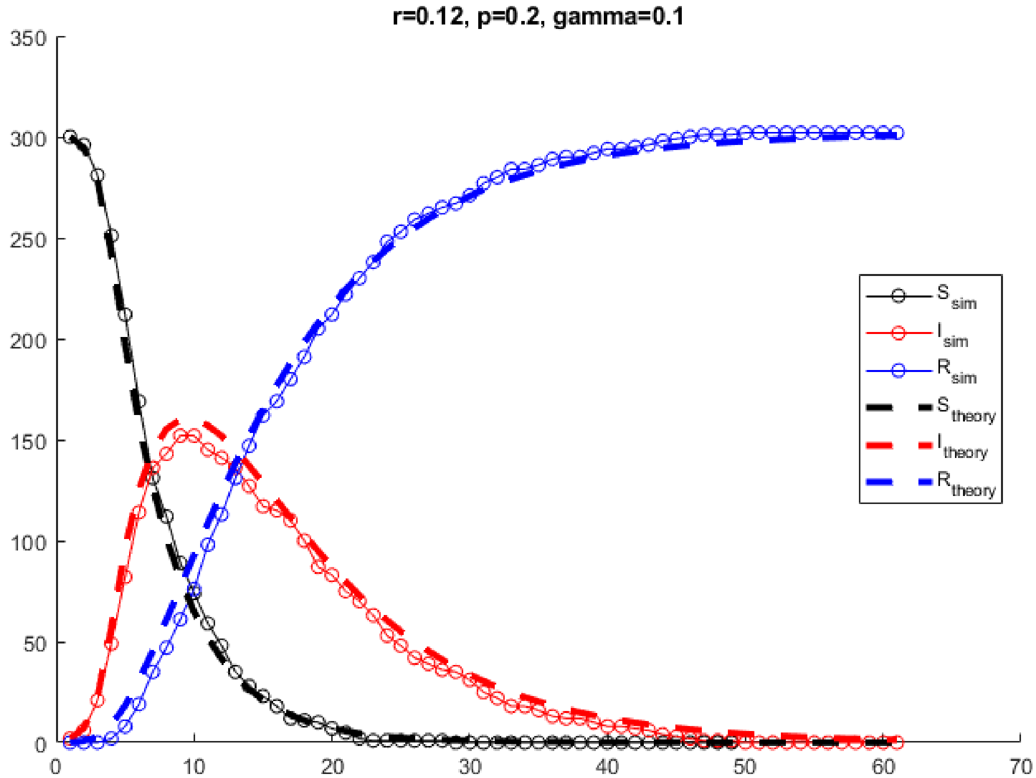
- SIR model becomes:

$$\boxed{\frac{dS}{dt} = -pS\left(1 - \exp\left(-Ia\right)\right); \qquad \frac{dI}{dt} = pS\left(1 - \exp\left(-Ia\right)\right) - \gamma I; \qquad \frac{dR}{dt} = \gamma I} \tag{2}$$

- Note that for small aI, we recover mass-action law: $pS\left(1 - \exp\left(Ia\right)\right) \sim paSI$

Direct comparison: ABM vs ODE:

$$\frac{dS}{dt} = -pS\left(1 - \exp\left(-Ia\right)\right); \qquad \frac{dI}{dt} = pS\left(1 - \exp\left(-Ia\right)\right) - \gamma I; \qquad \frac{dR}{dt} = \gamma I \qquad (3)$$



Estimating parameters from data

$$\frac{dS}{dt} = -\beta S \left(1 - \exp\left(-I\frac{\alpha}{N}\right)\right)\frac{1}{\alpha}; \quad \frac{dI}{dt} = \beta S \left(1 - \exp\left(-I\frac{\alpha}{N}\right)\right)\frac{1}{\alpha} - \gamma I; \quad \frac{dR}{dt} = \gamma I$$

- We use nonlinear least-squares fit to minimize the error in both I and R :

  $$\min_{\alpha,\beta,\gamma} \sum_j \left(I(t_j) - I_j\right)^2 + \left(R(t_j) - R_j\right)^2$$

  where S_j, R_j is the data from CSSEGIS or COVID19Tracking websites.

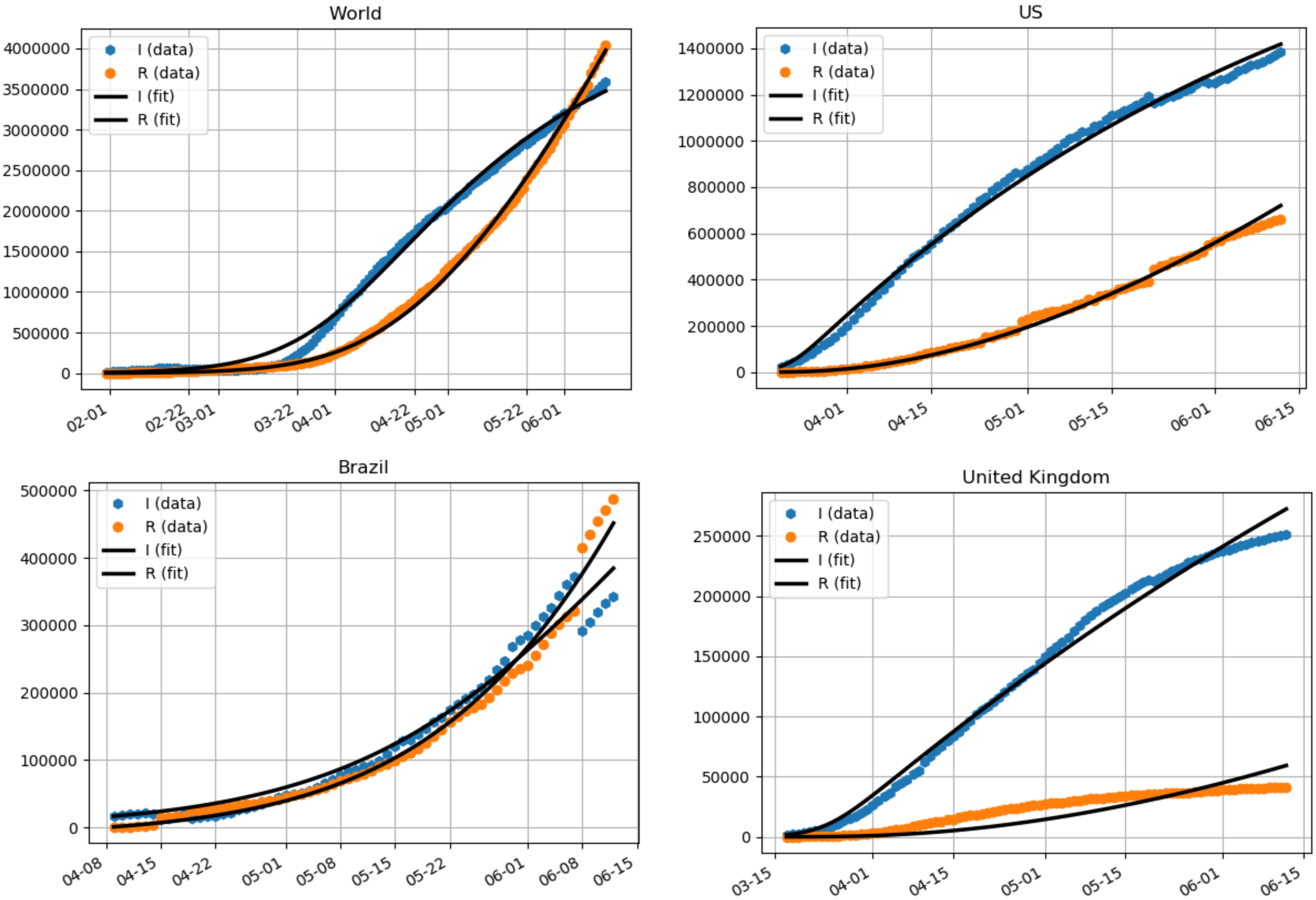
- Recovered = recovered+death; infected = total-recovered

- Used python's `optimize` routine using reasonable initial conditions (it "just works").

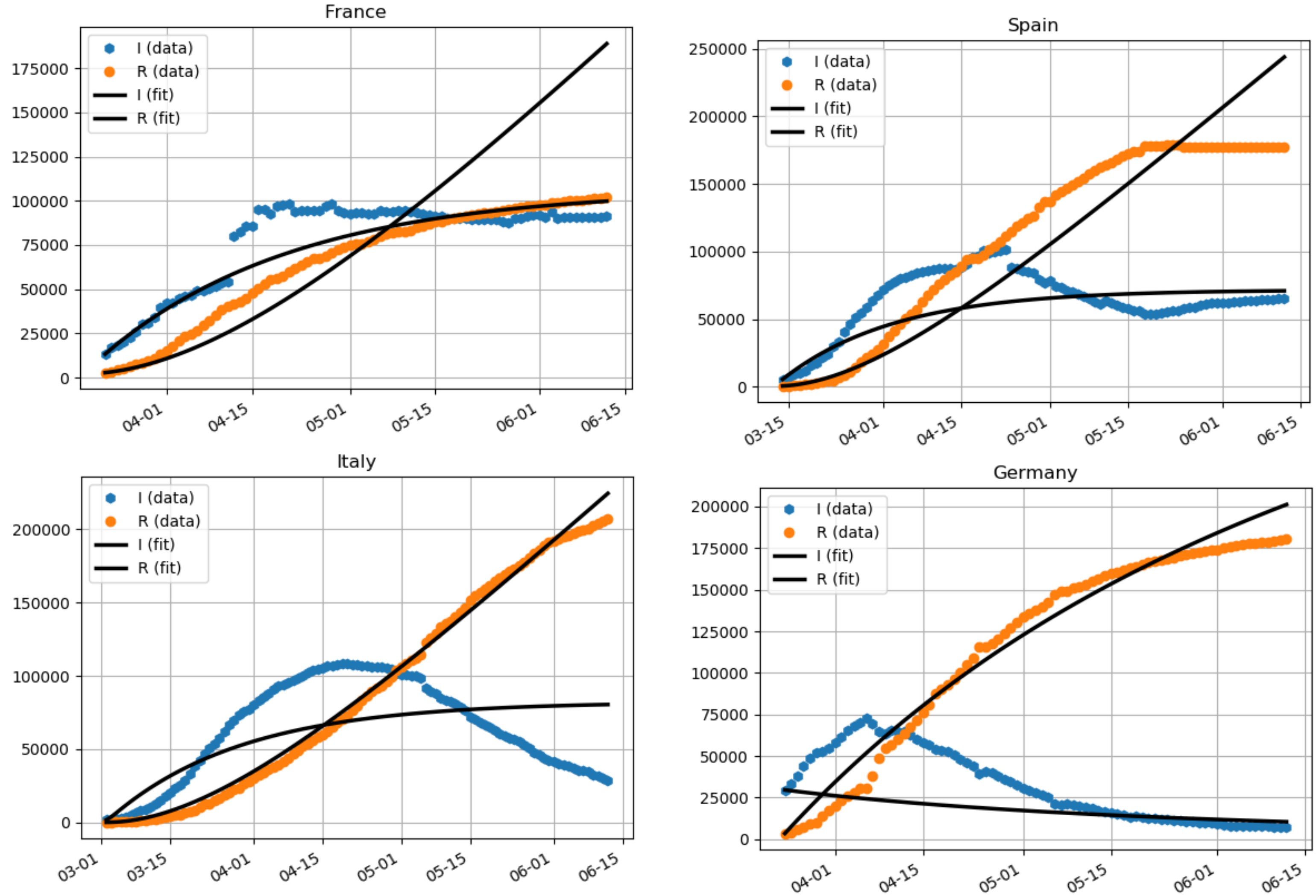
- We optimize over all parameters including gamma; but we fit to both I and R.

- Used data up to June 13, 2020, and from the time of 100 recovered individuals.

- Here are results:

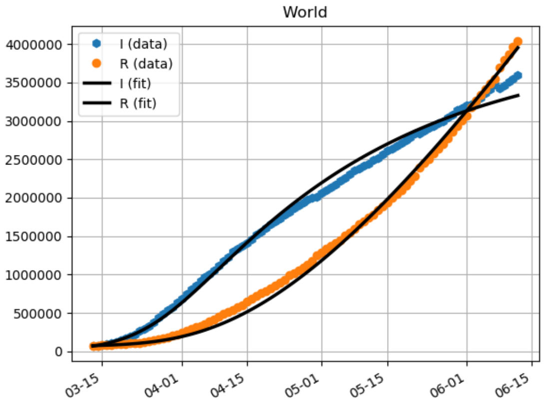

World
I (data)
R (data)
I (fit)
R (fit)
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
02-01
02-22
03-01
03-22
04-01
04-22
05-01
05-22
06-01
US
I (data)
R (data)
I (fit)
R (fit)
1400000
1200000
1000000
800000
600000
400000
200000
0
04-01
04-15
05-01
05-15
06-01
06-15
Brazil
I (data)
R (data)
I (fit)
R (fit)
500000
400000
300000
200000
100000
0
04-08
04-15
04-22
05-01
05-08
05-15
05-22
06-01
06-08
06-15
United Kingdom
I (data)
R (data)
I (fit)
R (fit)
250000
200000
150000
100000
50000
0
03-15
04-01
04-15
05-01
05-15
06-01
06-15

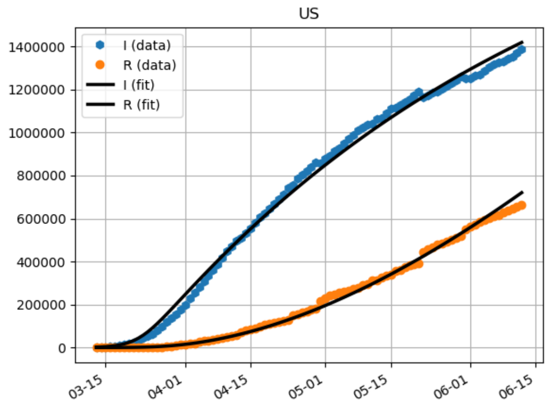
Sometimes it doesnt work:



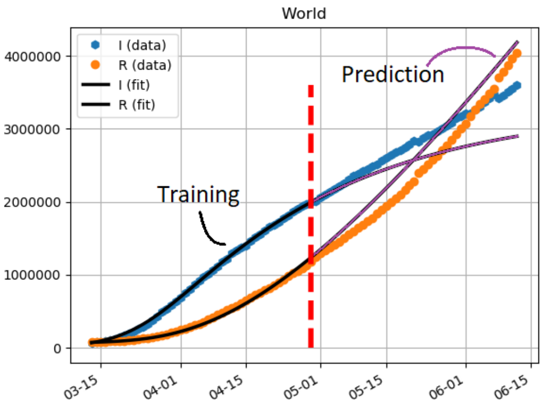
Predictive value

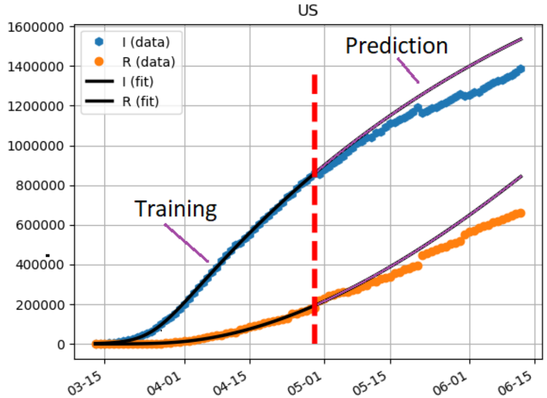
Training on all of data



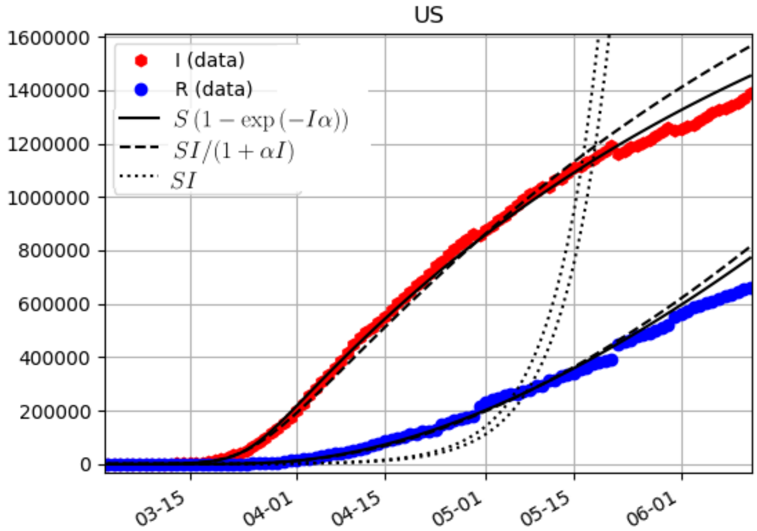


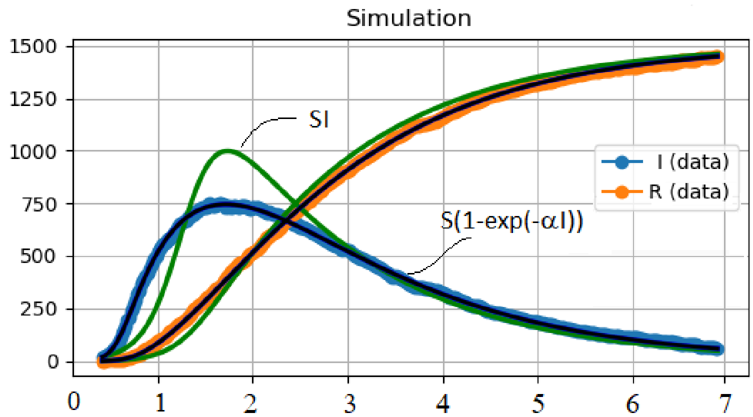
Training on half of data, predicting other half





Comparison: saturation vs. without





Effect of population density on R_0

- SIR with saturation

$$\frac{dS}{dt} = -pS\left(1 - \exp\left(-Ia\right)\right); \qquad \frac{dI}{dt} = pS\left(1 - \exp\left(-Ia\right)\right) - \gamma I; \qquad \frac{dR}{dt} = \gamma I \quad (4)$$

- For small aI, we linearize $1 - \exp\left(Ia\right) \sim aI$ so the model becomes

$$\frac{dS}{dt} = -paSI; \qquad \frac{dI}{dt} = paSI - \gamma I; \qquad \frac{dR}{dt} = \gamma I \quad (5)$$

- Here, $a = \pi r^2 / A$, where A is total area and r infection radius.

- In terms of population density, $A = N/\rho$ so we have:

$$a = \pi r^2 \frac{\rho}{N}$$

- Compared to the usual SIR model, $pa = \beta/N$ and

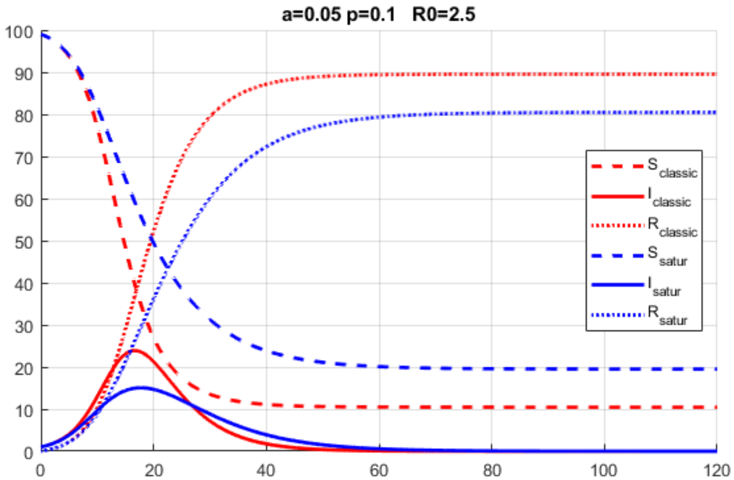
$$R_0 = \frac{\beta}{\gamma} = \frac{pNa}{\gamma} = \rho \frac{pr^2 \pi}{\gamma}$$

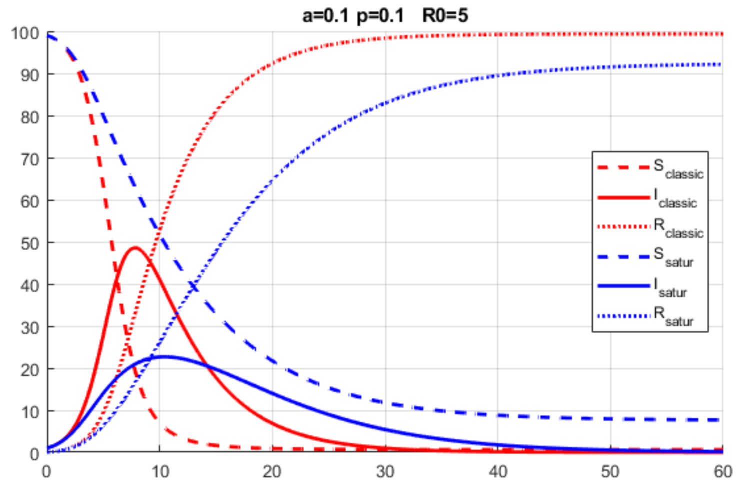
  where ρ is the population density.

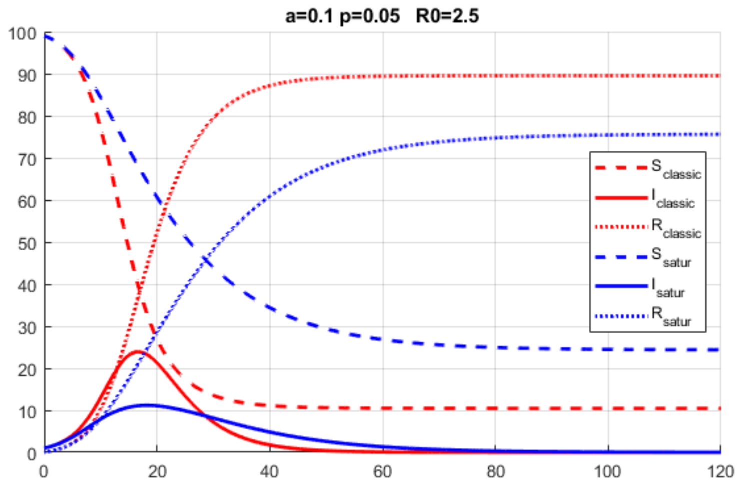
- **Conclusion:** R_0 ***is proportional to the population density at the onset of the outbreak.***

SIR classical vs. SIR saturated

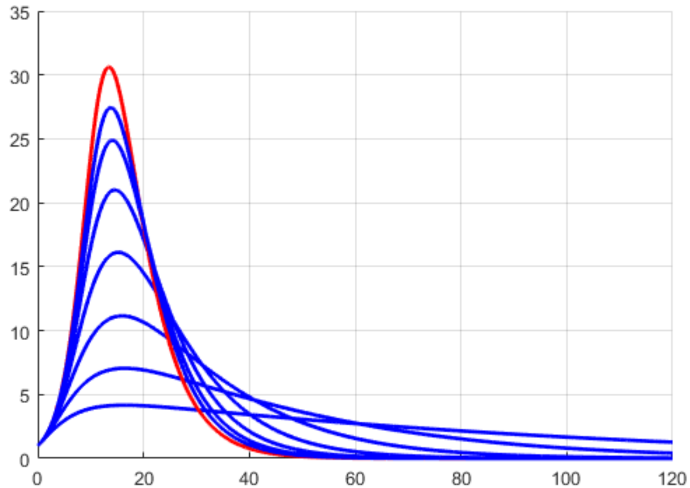
	SIR with saturation	Classical SIR
Equations	$\frac{dS}{dt} = -pS\left(1 - \exp\left(-Ia\right)\right);$ $\frac{dI}{dt} = pS\left(1 - \exp\left(-Ia\right)\right) - \gamma I$	$\frac{dS}{dt} = -\beta SI/N;$ $\frac{dI}{dt} = \beta SI/N - \gamma I$
Parameters	$a = \frac{\rho}{N} r^2 \pi$ includes $\left\{ \begin{array}{c} \text{pop density } \rho \\ \text{infection radius } r \end{array} \right\}$ p : probability of infection $\gamma \equiv$ recovery rate	$\beta \equiv$ infection rate $\gamma \equiv$ recovery rate
R_0	$\rho p \frac{r^2 \pi}{\gamma}$	β/γ
$I_{\max}/N$	$I_{\max,sat} < I_{\max,classical}$	$1 + \frac{1}{R_0}\left(\log \frac{1}{R_0} - 1\right)$
Attack rate r_∞	$r_{\infty,sat} < r_{\infty,classical}$	$\ln\left(1 - r_\infty\right)/r_\infty = R_0$

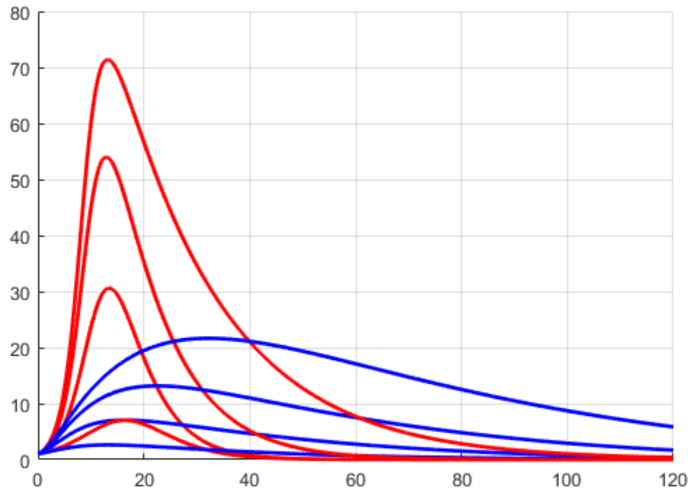

a=0.05 p=0.1 R0=2.5
S classic
I classic
R classic
S satur
I satur
R satur


a=0.1 p=0.1 R0=5
S classic
I classic
R classic
S satur
I satur
R satur


a=0.1 p=0.05 R0=2.5
S classic
I classic
R classic
S satur
I satur
R satur

Red = I (classic), blue =I (saturation)





p=0.01x2^k, k=0..6

a=0.006/p, gamma=0.2, N=100

a=0.3, p=0.02,

gamma=0.05x2^k, k=0..3

Social distancing

- Suppose that each person has a ***"hard-core" repulsion radius*** h in addition to ***"infection radius"*** r.

- Let E be the expected area fraction for I infected individuals. Then

$$E = Ia - \binom{I}{2}\frac{(a-b)^2}{1-2b} + \binom{I}{3}\frac{(a-2b)^3}{(1-3b)^2}\cdots$$

$$E = \sum_{j=1}^{I} -(-1)^j \binom{I}{j} I_j$$

where

$$I_j = \begin{cases} \dfrac{(a-(j-1)\,b)^j}{(1-jb)^{j-1}}, & \text{if } a > (j-1)\,b \\ 0, & \text{if } a < (j-1)\,b \end{cases}$$

$$a = \frac{\pi r^2}{A}, \quad b = \frac{\pi h^2}{A}$$

- The probability of coming into contact with one of I infected individuals is then given by:

$$F(I; a, b) := \frac{E(k, a, b) - Ib}{1 - Ib}$$

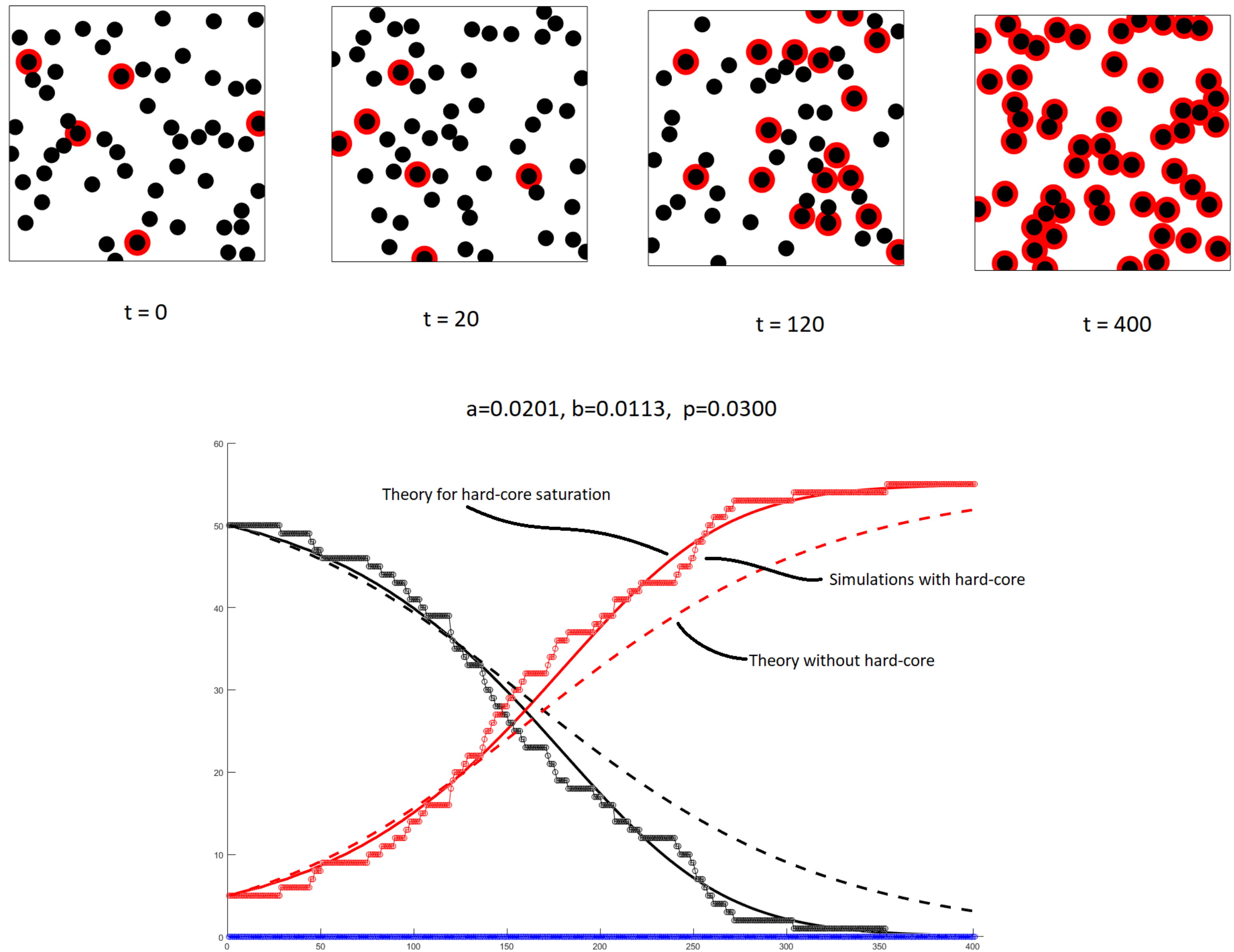
- Note that $F \sim (1 - \exp(-Ka))$ when $b = 0$ so this generalizes the $b = 0$ case.

- Note that $F \sim I\,(a - b)$ for small I, which reduces to the usual law of mass action (linear in I)

- **Conclusion:**

$$\text{Replace } \frac{\beta}{N} SI \;\text{ by } \beta F(I, \frac{\alpha}{N}, \frac{\mu}{N})$$

$$\frac{dS}{dt} = -\beta SF(I, \frac{\alpha}{N}, \frac{\mu}{N}); \qquad \frac{dI}{dt} = \beta SF(I, \frac{\alpha}{N}, \frac{\mu}{N}) - \gamma I; \qquad \frac{dR}{dt} = \gamma I \tag{6}$$


t = 0
t = 20
t = 120
t = 400
a=0.0201, b=0.0113, p=0.0300
Theory for hard-core saturation
Simulations with hard-core
Theory without hard-core
60
50
40
30
20
10
0
0
50
100
150
200
250
300
350
400

Future exploration

- Motion of agents (not fully-mixed), spatial propagation
- Piecewise parameter fit: Especially for European countries (e.g. Germany)
- Data fitting with parameters
- Decaying kernels
- Effect of social distancing

Thank you!