On a Coupled Time-Dependent SIR Models Fitting with New York and New-Jersey States COVID-19 Data

Benjamin Ambrosio * and M. A. Aziz-Alaoui

UNIHAVRE, LMAH, FR-CNRS-3335, ISCN, Normandie University, 76600 Le Havre, France
* Correspondence: benjamin.ambrosio@univ-lehavre.fr

check for updates

Abstract: This article describes a simple Susceptible Infected Recovered (SIR) model fitting with COVID-19 data for the month of March 2020 in New York (NY) state. The model is a classical SIR, but is non-autonomous; the rate of susceptible people becoming infected is adjusted over time in order to fit the available data. The death rate is also secondarily adjusted. Our fitting is made under the assumption that due to limiting number of tests, a large part of the infected population has not been tested positive. In the last part, we extend the model to take into account the daily fluxes between New Jersey (NJ) and NY states and fit the data for both states. Our simple model fits the available data, and illustrates typical dynamics of the disease: exponential increase, apex and decrease. The model highlights a decrease in the transmission rate over the period which gives a quantitative illustration about how lockdown policies reduce the spread of the pandemic. The coupled model with NY and NJ states shows a wave in NJ following the NY wave, illustrating the mechanism of spread from one attractive hot spot to its neighbor.

Keywords: COVID-19; New York; New Jersey; SIR models; Network
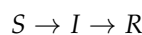
1. Introduction and Model

From a scientific perspective, the COVID-19 pandemic has highlighted the crucial role of mathematical and statistical models in providing guidance for health policies. Expressions such as "flatten the curve", the "apex", the "plateau" have been widely heard in media and employed by decision makers to explain their choices regarding rules and policies during this critical period. In this short article, we first introduce a simple Susceptible Infected Recovered (SIR) model, in which we adjust a key parameter *k* standing for a control on the Susceptible-Infected rate, and secondarily the death rate, in order to fit the data of the pandemic in New York (NY) state in March 2020, and provide predictions for a near future. Then, we add a node in the model to take into account the daily fluxes between NY and New Jersey (NJ) states. Note that these two close states, are, up to the day of publication of this article, the most severely hit by the pandemic in United States. Of course, the coupling may be extended to other states. However, in this article, we restrict ourselves to NY and NJ. Accordingly, the main key points of this article are that, (1) it highlights the dynamics and epidemiological characteristics which have been discussed in press and health policies; it highlights qualitatively how lockdown policies have decreased the spread of the virus and provides prediction and explanation of an upcoming apex, (2) it fits real data provided for the New York state and (3) it fits the data of NJ state by considering coupled equations taking into account the daily fluxes between NY and NJ. This provides a quantitative visualization of how the virus may spread from an attractive hot spot (New York City in NY state) towards close states trough the daily fluxes of commuters.

We especially focus on fitting the total number of cases tested positive for COVID-19 as well as the number of deaths in both NY and NJ states. We also give insights in prediction of the number of people needing hospitalization in NY state.

SIR models are very classical in literature. For some reader's convenience, we mention here a few contextual elements and references. The simplest classical SIR model is the Kermack–McKendrick (KMcK) model which goes back to 1927, see [1,2]. It is written

$$(KMcK) \begin{cases} S_t &= -kSI \\ I_t &= kIS - rI \\ R_t &= rI \end{cases} \tag{1}$$

In this original KMcK model, the population splits into three classes. The class S stands for susceptible, who can catch the disease, I stands for infective, who have the disease and can transmit it and R stands for the removed, namely, those who have or have had the disease but do not transmit it anymore. Note that our terminology is slightly different as explained below. In (1), the dynamics follow the scheme

$$S \to I \to R$$

with respective transfer rates between classes of k and r.

SIR type models, and more generally mathematical models of epidemics have in fact a significant history. We refer to [2] for a textbook on these models and a brief history of epidemics. Models have become more sophisticated and may include more compartments such as exposed, infective asymptomatic, infective with symptoms, and also reservoirs such as bats, and include stochastic dynamics. Recently, SIR type models have been widely used in the context of the COVID-19 pandemic, to model the spread all around the world, see for example [3–9] and reference therein cited. Here are also some examples of references for SIR models in other epidemic diseases: Dengue [10], Chikungunya [11,12] and Ebola [13]. See also references therein cited.

In the present article, we first consider the following model

$$\begin{cases} S_t &= -k(t)\frac{S}{I+S+R}I \\ I_t &= k(t)\frac{S}{I+S+R}I - r(t)I - d(t)I \\ R_t &= r(t)I \end{cases} \tag{2}$$

This simple model has classically three classes: S for susceptible, I for infected and R for recovered. Specifically, the class I is intended to represent all the people who bear effectively the virus at a given time, and can transmit it if in contact with other people. It includes all infected people with or without symptoms, reported or not. There are some differences with Equation (1). First, it includes a death rate $d(t)$. Even though the number of deaths does not appear explicitly as a variable, it is simply given by the integral $\int_0^t d(u)I(u)du$. Note that in expression

$$S_t = -k(t)\frac{S}{S+I+R}I,$$

the rate of contamination from S to I is proportional to the proportion of susceptible (S) in the whole population ($S + I + R$). This is a classical expression standing for the fact that the probability for each individual in the I class to spread the virus among the class S is proportional to the portion of S in the whole population, see for example [14,15] and references therein cited. This rate is corrected by a crucial coefficient $k(t)$ which is intended to fit the real transfer rate and which contains the effects inherent to the properties of the virus (for example change of propagation rate due to genetic mutation of the virus) or to specific policies (like quarantine, social distancing, lockdown). This time dependence allows us to adjust the dynamics to fit the data. This is a specificity of our model and turns it into a non-autonomous equation. This time dependence of k is obviously relevant in our model since the

rate of transfer from S to R is the main target of health policies and is subsequently subject to vary over time. Secondarily, we also allow the death rate to vary. Many internal or external factors may affect the death rate among which are concomitant lethal disease, temperature, hospital conditions. More significantly, one has to note that the rate transfers considered here are instantaneous transfer between compartments, and the function $d(t)$ is different from the Case Fatality Rate (CFR). Recall that the CFR is the death rate per confirmed case over a given period of time, and is a typical indicator for death rate. In South Korea, the country which led the highest number of tests, it has been reported to be of 1 percent, see [16]. In China as of 20 February, this rate varied between 3.8 in the region of Wuhan and 0.7 in others regions, see [17] and also [18]. Using Tables 1 and 2 gives a CFR in NY from 1 March to 1 April of $\frac{1941}{83804} \simeq 2.3$ per cent. Since our model fits the data, by definition, it fits the CFR. At the end of the epidemic, if the whole number of people that contracted the virus would be tested positive the CFR would provide the probability to die for an individual who catches the virus. However, during a growing phase such as the month of March considered here, the CFR has large variations. Moreover, there is a delay between the time a person is tested positive and the time he dies. The time between symptom onset and death has been reported to range from about 2 to 8 weeks, see [18], the typical average being 23 days according to [16]. One could for example introduce a time-integral death expression $\int_0^t \tilde{d}(u,t)I(u)du$ to take into account these informations. However, the time-window considered here is short and corresponds to the beginning of reported cases in NY. Furthermore, in this short article, we wanted to focus on a simple model able to fit data, highlight relevant dynamics and provide estimations. Since a person in the I compartment will either recover or die, above remarks on the function d hold for the coefficient r. In the present work, for sake of simplicity, we have set the r coefficient to the constant value 0.64. This is a simplification which is classical in SIR models. Note that setting the coefficient r to a constant value is equivalent to assume that people recovering between times $t-1$ and t, which is given by $R(t) - R(t-1)$ would also write $r \int_{t-1}^t I(u)du$ according to Equation (2). A more meaningful expression for $R(t) - R(t-1)$ would be $\int_0^t \tilde{r}(u,t)I(u)du$ standing to the fact that people recovering between time $t-1$ and t have been infected from a period ranging from 0 to t, with a transfer rate given by $\tilde{r}(u,t)$. This would lead to the equation $r = \frac{\int_0^t \tilde{r}(u,t)I(u)du}{\int_{t-1}^t I(u)du}$. Once r is set to a constant value, to fit data over a reasonable period, numerical tries provide a unique choice for constants k and d. It was still possible to use different values of r to fit the data. However, different values of r result in different dynamics over time, and notably different time for apex. Our choice of $r = 0.64$ was made to provide dynamics that seem relevant to us regrading the timely dynamics beyond the data. In particular, too smaller of values for r would provide dynamics with a late apex and less relevant regarding the timely effects of the disease. Other studies have considered models with a constant r, with different values. See for example [19] and references therein cited. Remarkably, varying r, and making it depend on time, provides some freedom to later fit data over a longer period of time, taking in account the end of the first wave in NY.

Upon the above discussions, our strategy is rather simple: set r to a constant value. Then choose a constant k which fits the data of positive cases during a period of time. Next, choose a constant d to fit deaths data for the same period of time. Note that since the parameter k has much more effect than the small parameter d, this procedure is possible and efficient. After, repeat the procedure over a subsequent period. The overall procedure results in a constant function r and two piecewise constant functions $k(t)$ and $d(t)$. For our model and the given data, the procedure was efficient allowing us to provide these choices by successive tries. Note that this could be done automatically by cooking an algorithm to set the parameters following these guidelines. Note finally that our assumptions do not a include birth rate for the susceptible population but rather focus on the short time effects of the disease.

Table 1. Total number of cases reported in NY state from 1 March to 1 April. See [20].

Day	3/1	3/2	3/3	3/4	3/5	3/6	3/7	3/8	3/9	3/10	3/11
Number of Cases	1	1	2	11	22	44	89	106	142	173	217
Day	3/12	3/13	3/14	3/15	3/16	3/17	3/18	3/19	3/20	3/21	3/22
Number of Cases	326	421	610	732	950	1374	2382	4152	7102	10356	15168
Day	3/23	3/24	3/25	3/26	3/27	3/28	3/29	3/30	3/31	4/1	
Number of Cases	20875	25665	33066	38987	44635	53363	59568	67174	75832	83804	

Table 2. Total number of deaths reported in NY state from 1 March to 1 April. See [20].

Day	3/1	3/2	3/3	3/4	3/5	3/6	3/7	3/8	3/9	3/10	3/11
Number of Deaths	0	0	0	0	0	0	0	0	0	0	0
Day	3/12	3/13	3/14	3/15	3/16	3/17	3/18	3/19	3/20	3/21	3/22
Number of Deaths	0	0	2	6	10	17	27	30	57	80	122
Day	3/23	3/24	3/25	3/26	3/27	3/28	3/29	3/30	3/31	4/1	
Number of Deaths	159	218	325	432	535	782	965	1224	1550	1941	

2. Numerical Simulations, Data and Dynamics

2.1. Fitting the Total Number of Infected People and the Number of Deaths

Data from total infected people and total number of deaths in New York state are available at the New York times web site, see [20]. We downloaded the data from there from 1 March to 1 April, which makes 32 days. We reported the number of total cases in Table 1 and the number of deaths in Table 2.

For numerical simulations of Equation (2), the parameter r was set to 0.64. Then, in order to fit the data of the total number of infected people, we chose:

$$k(t) = \begin{cases} k_1 = 1.057 & \text{if} \quad 0 \leq t < 21 \\ k_2 = 0.9 & \text{if} \quad 21 \leq t < 24 \\ k_3 = 0.67 & \text{if} \quad 24 \leq t < 27 \\ k_4 = 0.71 & \text{if} \quad 27 \leq t \leq 32 \end{cases} \tag{3}$$

and to fit the total number of deaths, we chose:

$$d(t) = \begin{cases} d_1 = 0.0016 & \text{if} \quad 0 \leq t < 21 \\ d_2 = 0.00232 & \text{if} \quad 21 \leq t < 24 \\ d_3 = 0.00232 & \text{if} \quad 24 \leq t < 27 \\ d_4 = 0.0068 & \text{if} \quad 27 \leq t \leq 32 \end{cases} \tag{4}$$

Initial conditions were set to

$$(19453556, 5, 0),$$

19453556 being the number of people living in NY state in 2019 according to USA Census bureau. Regarding reported data in Tables 1 and 2, as a matter of fact, not all infected people in March 2020 were tested. To take into account this fact in our model, we adjusted the parameters to fit the quantity $0.2 \times (I + R)$ from the model, with the total number of infected people (data in Table 1) minus the total number of deaths (data in Table 2), i.e., we assumed here that only 20 percent of living people that have contracted the virus has been tested positive. Note that shortly after the first submission of this work, NY state reported the result of a random antibody test over a sample of 3000 individuals. Statewide, 13.9 percent of people tested positive, ranging form 21.2 percent in New York City (NYC), to 3.6 percent in upstate counties. See [21,22]. Comparing with data from [20], this gave a ratio of

approximatively 10; only about 10 percent of people having caught the virus in NY would have tested positive. Statistical estimates of related ratio, depending on the number of estimates provide a range of [5, 3000], see [23]. Simulation of system (2), and comparison with data resulting from Tables 1 and 2 are plotted in Figure 1a. In Figure 1b, we plotted the total number of deaths from the model and compared it to the data. In Figure 1c, we have plotted four curves corresponding to various $I(t)$ for different simulations of (2).

- The curve $I_1(t)$ in red corresponds to the simulation of (2) with $k(t) = k_1 = 1.057$ and $d(t) = d_1 = 0.0016$ for all time.
- The curve $I_2(t)$ in green corresponds to the simulation of (2) with $k(t) = k_1 = 1.057$ and $d(t) = d_1 = 0.016$ for $0 \leq t \leq 21$, and $k(t) = k_2 = 0.9$ and $d(t) = d_2 = 0.00232$ for $21 \leq t < 24$.
- The curve $I_3(t)$ in pink corresponds to the simulation of (2) with $k(t)$ as given in (3), i.e., $k(t) = k_i$ and $d(t) = d_i$, $t \in [t_{i-1}, t_i)$, $i \in \{1, ..., 4\}$ with $t_0 = 0, t_1 = 21, t_2 = 24, t_3 = 27, t_4 = 32$.

For reader's convenience, these values are summarized in Table 3. This panel illustrates how the health policies flatten the curve. In Figure 1d, we have again plotted the solution $I(t)$ for $k(t)$ as in (3), for a longer period.
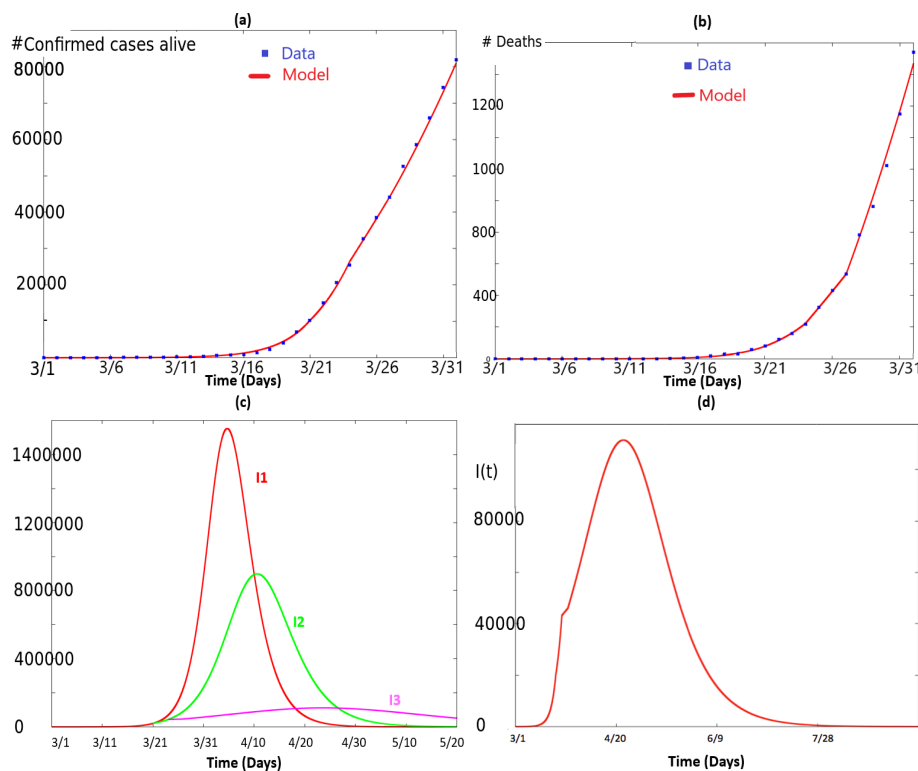

Figure 1. This figure illustrates the simulation of system (2) and how it fits the data. In (**a**), we have plotted the quantity $0.2 \times (I + R)$ as a function of time in red. The blue dots correspond to the data retrieved from [20]. Analogously, in (**b**), we have plotted the quantity $\int_0^t d(s)I(s)ds$, which represents the total number of deaths according with the model, as a function of time in red. The blue dots correspond to the data retrieved from [20]. In (**c**), we have illustrated the quantity $I(t)$ corresponding to different values of $k(t), d(t)$: the curve $I_1(t)$ in red corresponds to the simulation of (2) with $k(t) = k_1$ and $d(t) = d_1$ for all time. The curve $I_2(t)$ in green corresponds to the simulation of (2) with $k(t) = k_1$ and $d(t) = d_1$ for $0 \leq t \leq 21$, and $k(t) = k_2$ and $d(t) = d_2$ for $t \geq 21$. The curve $I_3(t)$ in pink corresponds to the simulation of (2) with $k(t)$ as given in (3), i.e., $k(t) = k_i$ and $d(t) = d_i$, $t \in [t_{i-1}, t_i)$, $i \in \{1, ..., 4\}$ with $t_0 = 0, t_1 = 21, t_2 = 24, t_3 = 27, t_4 = 32$. It illustrates how the health policies flatten the curve. In (**d**), we have again plotted the solution $I(t)$ for $k(t)$ as in (3), for a longer period.

Table 3. Summary of the values used in Figure 1c to obtain the curves $I_1(t)$, $I_2(t)$ and $I_3(t)$. Recall that $k_1 = 1.057, k_2 = 0.9, k_3 = 0.67, k_4 = 0.71, d_1 = 0.0016, d_2 = 0.00232, d_3 = 0.00232, d_4 = 0.0068, t_0 = 0,$ $t_1 = 21, t_2 = 24, t_3 = 27$ and $t_4 = 32$.

	$k(t)$	$d(t)$
$I_1(t)$	$k(t) = 1.057$	$d(t) = 0.0016$
$I_2(t)$	1.057 if $t < 21$; 0.9 otherwise	0.0016 if $t < 21$; 0.00232 otherwise
$I_3(t)$	k_i if $t_{i-1} \leq t < t_i, i \in \{1, ..., 4\}$	d_i if $t_{i-1} \leq t < t_i, i \in \{1, ..., 4\}$

Table 4. Total number of daily current hospitalizations reported in NY state from 16 March to 1 April. See [23,24].

Day	3/16	3/17	3/18	3/19	3/20	3/21	3/22	3/23	3/24	3/25
Total Number of Hospitalized	326	496	617	1042	1496	2043	2629	3343	4079	5327
Day	3/26	3/27	3/28	3/29	3/30	3/31	4/1			
Total Number of Hospitalized	6481	7328	8503	9517	10929	12226	13383			

2.2. Fitting the Total Number of People at Hospital

Next, we have retrieved data corresponding to the number of people at the hospital between 16 March and 1 April, see Table 4. During this period, and after, NY state officially reported daily useful charts and statistics, on the local spread. We then computed an estimation of people being effectively at hospital at a given date. We denote $H(t_i), t_i \in \{16, 17, ..., 32\}$ the total number of hospitalizations at a given date. In order to fit the solution of (2) with this data, we performed a linear regression between $(I(t_i)), t_i \in \{16, 17, ..., 32\}$ and $(H(t_i)), t_i \in \{16, 17, ..., 32\}$. The coefficients a and b of the linear regression were determined by the least-square method, leading to the formulas:

$$a = \frac{< x, y > - n\bar{x}\bar{y}}{||x||^2 - \bar{x}^2}$$

$$b = \bar{y} - a\bar{x}$$

with

$$x = (H(t_i), t_i \in \{16, 17, ..., 32\}$$

$$y = (I(t_i)), t_i \in \{16, 17, ..., 32\}$$

$$< x, y > = \sum_{i=1}^{n} x_i y_i, \ ||x||^2 = <x, x>, \ \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \ \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \ n = 14.$$

The result is plotted in Figure 2a. It clearly shows a good approximation by two distinct lines, corresponding to

$$a_1 = 10.3301, \ b_1 = -1227.61$$

and

$$a_2 = 1.91478, \ b_2 = 33560.3.$$

Then, a prevision of number of people needing hospitalization can be provided by the formula:

$$H(t) = \frac{I(t) - b_2}{a_2} \text{ if } I(t) \geq 41475.76, H(t) = \frac{I(t) - b_1}{a_1} \text{ otherwise.}$$
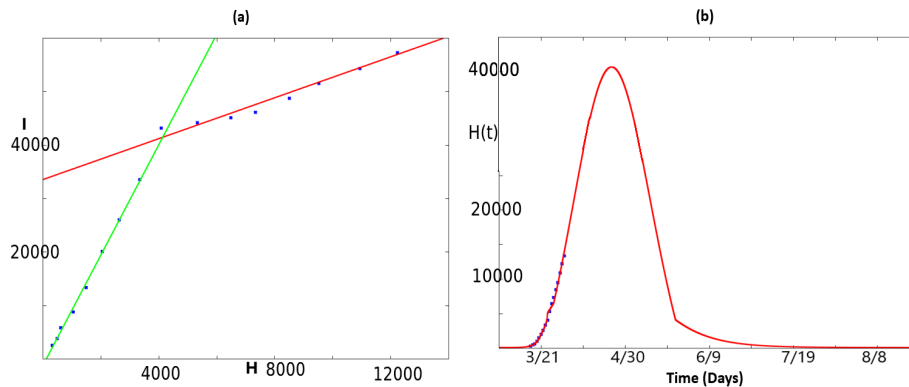
The result is plotted in Figure 2b.

Figure 2. This figure illustrates how to provide an estimation for people needing hospitalization thanks to Equation (2) and statistical methods. Panel (**a**) illustrates an approximation of $I(t_i), t_i \in \{16, 17, ..., 32\}$ by a vector $(a_1 H(t_i) + b_1), t_i \in \{16, 17, ..., 23\}$ and another vector $(a_2 H(t_i) + b_2), t_i \in \{24, 25, ..., 32\}$ where a_j and $b_j, j \in \{1, 2\}$ are the coefficients obtained thanks to the least-square method. Panel (**b**) then provides a prediction of people in need of hospitalization by plotting the quantity $H(t) = \frac{I(t) - b_2}{a_2}$ if $I(t) \geq 41,475.76, H(t) = \frac{I(t) - b_1}{a_1}$ otherwise.

2.3. Dynamics

Summing up the equations in (2) and looking for stationary solutions yield the following theorem.

Theorem 1. *We assume that initial condition of system (2) satisfies $S(0) > 0, I(0) > 0$ and $R(0) = 0$. Then for $t > 0$, all variables remain bounded and positive: there exists a positive constant $M < R(0) + I(0)$ such that for all $t > 0$:*

$$0 < S(t) < M, 0 < I(t) < M, 0 < R(t) < M$$

and

$$S(t) + I(t) + R(t) \leq S(0) + I(0) - \int_0^t d(s)I(s)ds.$$

Non-negative stationary solutions of the system are given by $(\bar{S}, 0, \bar{R})$, with $\bar{S} \geq 0$ and $\bar{R} \geq 0$. Furthermore, $S(t)$ is decreasing, $R(t)$ is increasing and the variation of $I(t)$ is given by the sign of

$$k(t)\frac{S(t)}{S(t) + I(t) + R(t)} - r - d(t). \tag{5}$$

Remark 1. *It is worth noting that this theorem, whose proof is relatively straightforward, provides two simple but relevant interpretations from the applicative point of view. The first thing is that the stationary solutions are given by $(\bar{S}, 0, \bar{R})$, with $\bar{S} \geq 0$ and $\bar{R} \geq 0$. This means that the stationary solutions, are all with 0 infected but may take arbitrary non-negative values (within a bounded interval) for susceptible and recovered. This reflects the reality of the disappearance of the virus. The second thing we want to mention is that the variation of I is given by the sign of expression (5). In particular, basically during a classical wave, this sign will be positive before the apex and negative after.*

3. Two Coupled SIR Systems Fitting COVID-19 for NY and NJ States

It is estimated that around 400,000 people used to commute from NJ to NY before the lockdown policies due to the pandemic. It is natural to integrate these effects in the model, since those commuters played the role of vector for virus before the lockdown. We therefore built a small network by coupling two nodes, representing NY (node 1) and NJ (node 2). To this end, we coupled two copies

of the local model (2) to take into account the daily fluxes between those two states. This led to the following model:

$$
\begin{cases}
S_{1t} &= -k_1(t)I_1\frac{S_1}{I_1+S_1+R_1} + l(t)\big(c_{12}^S(t)S_2 - c_{21}^S(t)S_1\big) \\
I_{1t} &= k_1(t)I_1\frac{S_1}{I_1+S_1+R_1} - rI_1 - d_1(t)I_1 + l(t)\big(c_{12}^I(t)I_2 - c_{21}^I(t)I_1\big) \\
R_{1t} &= rI_1 + l(t)\big(c_{12}^R(t)R_2 - c_{21}^R(t)R_1\big) \\
S_{2t} &= -k_2(t)I_2\frac{S_2}{I_2+S_2+R_2} + l(t)\big(-c_{12}^S(t)S_2 + c_{21}^S(t)S_1\big) \\
I_{2t} &= k_2(t)I_2\frac{S_2}{I_2+S_2+R_2} - rI_2 - d_2(t)I_2 + l(t)\big(-c_{12}^I(t)I_2 + c_{21}^I(t)I_1\big) \\
R_{2t} &= rI_2 + l(t)\big(-c_{12}^R(t)R_2 + c_{21}^R(t)R_1\big)
\end{cases}
\tag{6}
$$

The assumptions are analogous to those given Section 2, but we aimed here to take into account the daily fluxes between NJ and NY. For sake of simplicity, we considered that there was a flux of people coming from NJ to NY in the morning and returning home at night. The coupling functions, which are non-autonomous here are given by:

$$
\begin{pmatrix}
l(t)\big(c_{12}^S(t)S_2 - c_{21}^S(t)S_1\big) \\
l(t)\big(c_{12}^I(t)I_2 - c_{21}^I(t)I_1\big) \\
l(t)\big(c_{12}^R(t)R_2 - c_{21}^R(t)R_1\big)
\end{pmatrix}
$$

for node 1 and

$$
\begin{pmatrix}
l(t)\big(-c_{12}^S(t)S_2 + c_{21}^S(t)S_1\big) \\
l(t)\big(-c_{12}^I(t)I_2 + c_{21}^I(t)I_1\big) \\
l(t)\big(-c_{12}^R(t)R_2 + c_{21}^R(t)R_1\big)
\end{pmatrix}
$$

for node 2, where functions $+c_{ij}$ stand for the densities of population coming from mode j and going into node i. In the remaining of the manuscript we assume that these functions c_{ij} do not depend on S, I and R and drop the superscripts. Furthermore, we assume c_{12} and c_{21} to be periodic of period 1 (one day), with a Gaussian profile and respective apex at 8:30 am and 6:30 pm. Note that the amplitude c_{21} is multiplied by a coefficient greater than 1, in comparison with c_{12} to take into account the amount of population int NY and NJ. We want to emphasize here the attractivity of NYC, and therefore assume that the fluxes are mainly from NJ to NY in the morning and form NY to NJ in the evening. The function $l(t)$ integrates policies of lockdown: after March 23, in the model, the daily fluxes are divided by 1000. Therefore, the function $l(t)$ is a piecewise constant function given by:

$$
l(t) := \begin{cases} 1 & \text{if} \quad 0 \le t < 23 \\ 10^{-3} & \text{if} \quad t \ge 23. \end{cases}
\tag{7}
$$

Initial conditions were set to

$$
(19453556, 5, 0),
$$

for node 1 and

$$
(8882190, 0, 0),
$$

for node 2. Note there are no infected cases at initial time in NJ. This means that in our model, initial spread in NJ follows from infection in NY. The same methods as in Section 2 were used to fit the data for both NY and NJ for the network model (6). Illustrations are provided in Figure 3. It shows how the model fits the data. In Figure 3a, we have plotted the quantity $0.2(I_2 + R_2)$ as a function of time in red. Recall that, in the model $I_2 + R_2$, represents the number of people in the population which has been infected by the virus and are still alive. The blue dots correspond to the data retrieved from [20] and plots the total number of infected minus the number of total deaths. Analogously, in Figure 3b we have plotted the quantity $\int_0^t d_2(u)I_2(u)du$ as a function of time, in red. The blue dots

correspond to the data retrieved from [20]. Finally, Figure 3c illustrates $I_1 t)$ and $I_2(t)$, which represent respectively the infected population in NY and NJ. It shows how the curve of NJ follows the curve in NY, with a small attenuation.
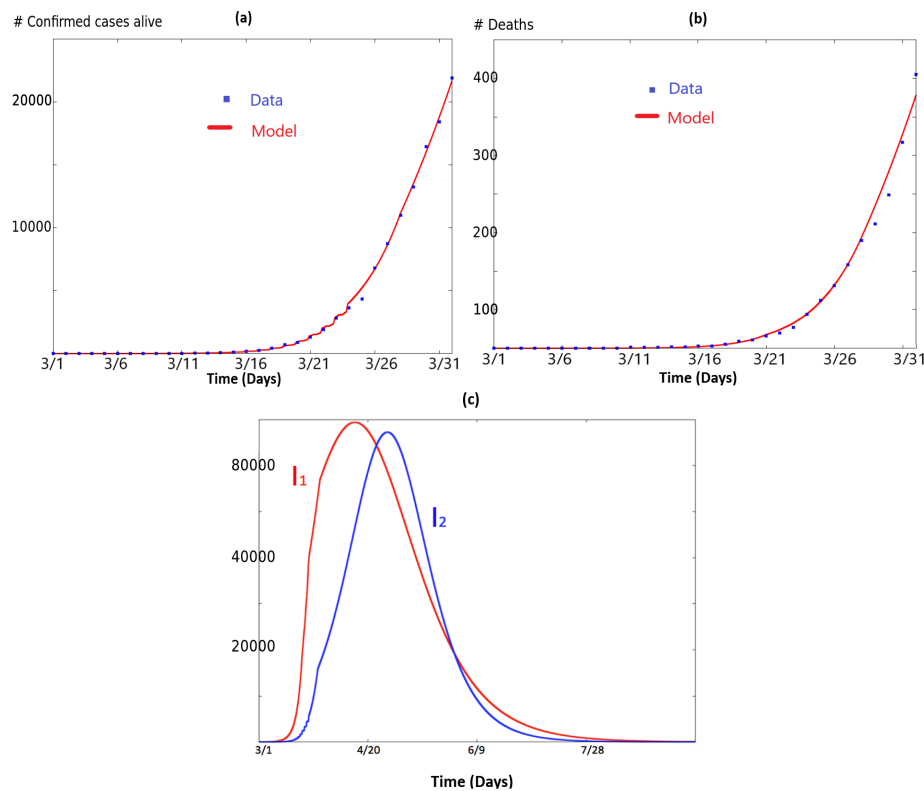

Figure 3. This figure illustrates the simulation of system (6) and how it fits the data. In (**a**), we have plotted the quantity $0.2(I_2 + R_2)$ as a function of time in red. Recall that, in the model $I_2 + R_2$, represents the number of people in the population which has been infected by the virus and are still alive. The blue dots correspond to the data retrieved from [20] and plots the total number of infected minus the number of total deaths. Analogously, in (**b**) we have plotted the quantity $\int_0^t d2(u)I_2(u)du$ as a function of time in red. The blue dots correspond to the total number of deaths in NJ according with data retrieved from [20]. Panel (**c**) illustrates $I_1(t)$ and $I_2(t)$, which represent respectively the infected in NY and NJ.

An analog theoretical result as in section 2 holds for solutions of (6).

Theorem 2. *We assume that initial condition of system* (6) *satisfies* $S_1(0) > 0, I_1(0) > 0, R_2(0) = 0,$ $S_2(0) > 0, I_2(0) = 0$ *and* $R_1(0) = 0.$ *Then for* $t > 0$, *all variables remain bounded and positive: there exists a positive constant* $M < R(0) + I(0)$ *such that for all* $t > 0$:

$$0 < S_i(t) < M, 0 < I_i(t) < M, 0 < R_i(t) < M, i \in \{1, 2\}$$

and

$$S_1(t) + I_1(t) + R_1(t) + S_2(t) + I_2(t) + R_2(t) \le S_1(0) + I_1(0) + S_2(0) - \int_0^t (d_1(s)I_1(s) + d_2(s)I_2(s))ds.$$

For any non-negative $S_1(0), R_1(0), S_2(0), R_2(0)$, the following functions satisfying for $t > 0$

$$I_1(t) = I_2(t) = 0$$
$$S_1(t) + S_2(t) = S_1(0) + S_2(0)$$
$$R_1(t) + R_2(t) = R_1(0) + R_2(0)$$
$$S_{1t} = -S_{2t} = l(t)\big(c_{12}(t)S_2 - c_{21}(t)S_1\big)$$
$$R_{1t} = -R_{2t} = l(t)\big(c_{12}(t)R_2 - c_{21}(t)R_1\big)$$

are solutions of (6). *Furthermore, $S_1(t) + S_2(t)$ is decreasing, $R_1(t) + R_2(t)$ is increasing and $I_1(t) + I_2(t)$ is non-increasing if and only if*

$$I_1(t)k_1(t)\frac{S_1(t)}{S_1(t)+I_1(t)+R_1(t)} + I_2(t)k_2(t)\frac{S_2(t)}{S_2(t)+I_2(t)+R_2(t)} \le I_1(t)(r + d_1(t)) + I_2(t)(r + d_2(t)). \quad (8)$$

Remark 2. *As in Remark 1, we want to point out some interpretation of this theorem in the context of the pandemic. Here, we have described some solutions with free epidemic component ($I_(t) = I_2(t) = 0$). Note that in this case however, S_i and R_i are not constant since they vary according to the fluxes between the two nodes. Again the last inequality, quantifies the idea that when the spread of the virus becomes lower than the death and recovery the infected population starts to decrease. The difference here is that we take into account the two nodes together.*

4. Conclusions

In this article, we have considered a simple non-autonomous SIR model to fit the data of COVID-19 in New York state. The model illustrates and quantifies how acting on the control $k(t)$ allows us to flatten the curve of infected people over the time. From the model, using classical statistical methods, it is then possible to provide predictions of the number of people in needs of hospitalization. Lastly, we have fitted data from NJ state thanks to a coupled SIR model taking into account the daily fluxes between NJ and NY. It allows us to predict similar dynamics in NY and NJ, with a delay and small attenuation. Note that, despite its simplicity, our model fits, as good as more sophisticated models, the available data during the growing phase of March 2020. In a forthcoming work, we aim to fit data with the model over a longer period . This would provide an accurate estimation of the different parameters.

Author Contributions: B.A. and M.A.A.-A. set up the mathematical model, performed the research, provided the mathematical analysis and wrote the paper; B.Ambrosio did the numerical simulations.

Conflicts of Interest: Not applicable.

References

1. Kermack, W.O.; McKendrick, A.G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Character* **1927**, *115*, 700–721. [CrossRef]
2. Murray, J. *Mathematical Biology*; Springer-Verlag: New York, NY, USA, 2010.
3. Hui, D.S.; Zumla, A. Severe Acute Respiratory Syndrome. *Infect. Dis. Clin. N. Am.* **2019**, *33*, 869–889. [CrossRef] [PubMed]
4. Lin, Q.; Zhao, S.; Gao, D.; Lou, Y.; Yang, S.; Musa, S.S.; Wang, M.H.; Cai, Y.; Wang, W.; Yang, L.; et al. A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. *Int. J. Infect. Dis.* **2020**, *93*, 211–216. [CrossRef] [PubMed]
5. Chen, T.M.; Rui, J.; Wang, Q.P.; Zhao, Z.Y.; Cui, J.A.; Yin, L. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infect. Dis. Poverty* **2020**, *9*, 24. [CrossRef] [PubMed]
6. Kuniya, T. Prediction of the Epidemic Peak of Coronavirus Disease in Japan, 2020. *J. Clin. Med.* **2020**, *9*, 789. [CrossRef] [PubMed]

7. Liu, Z.; Magal, P.; Seydi, O.; Webb, G. Understanding Unreported Cases in the COVID-19 Epidemic Outbreak in Wuhan, China, and the Importance of Major Public Health Interventions. *Biology* **2020**, *9*, 50. [CrossRef] [PubMed]

8. Liu, Z.; Magal, P.; Seydi, O.; Webb, G. Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data. *arXiv* **2020**, arXiv:2002.12298.

9. Tang, B.; Wang, X.; Li, Q.; Bragazzi, N.L.; Tang, S.; Xiao, Y.; Wu, J. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions. *J. Clin. Med.* **2020**, *9*, 462. [CrossRef] [PubMed]

10. Aziz-Alaoui, M.A.; Gakkhar, S.; Ambrosio, B.; Mishra, A. A network model for control of dengue epidemic using sterile insect technique. *Math. Biosci. Eng.* **2017**, *15*, 441–460. [CrossRef] [PubMed]

11. Moulay, D.; Aziz-Alaoui, M.; Cadivel, M. The chikungunya disease: Modeling, vector and transmission global dynamics. *Math. Biosci.* **2011**, *229*, 50–63. [CrossRef] [PubMed]

12. Moulay, D.; Aziz-Alaoui, M.; Kwon, H. Optimal control of chikungunya disease: Larvae reduction, treatment and prevention. *Math. Biosci. Eng.* **2012**, *9*, 369–392. [CrossRef] [PubMed]

13. Thompson, R.N.; Morgan, O.W.; Jalava, K. Rigorous surveillance is necessary for high confidence in end-of-outbreak declarations for Ebola and other infectious diseases. *Philos. Trans. R. Soc. B Biol. Sci.* **2019**, *374*, 20180431. [CrossRef] [PubMed]

14. Gao, L.; Hethcote, H. Disease transmission models with density-dependent demographics. *J. Math. Biol.* **1992**, *30*, 717–731. [CrossRef] [PubMed]

15. Yoshida, N.; Hara, T. Global stability of a delayed SIR epidemic model with density dependent birth and death rates. *J. Comput. Appl. Math.* **2007**, *201*, 339–347. [CrossRef]

16. Lachmann, A.; Jagodnik, K.M.; Giorgi, F.M.; Ray, F. Correcting under-reported COVID-19 case numbers: estimating the true scale of the pandemic. *medRvix* **2020**. [CrossRef]

17. WHO. *Report of the World Health Organization-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*; Report; WHO: Geneva, Switzerland, 16–24 February 2020.

18. Baud, D.; Qi, X.; Nielsen-Saines, K.; Musso, D.; Pomar, L.; Favre, G. Real estimates of mortality following COVID-19 infection. *Lancet Infect. Dis.* **2020**. [CrossRef]

19. Roques, L.; Klein, E.K.; Papaïx, J.; Sar, A.; Soubeyrand, S. Using Early Data to Estimate the Actual Infection Fatality Ratio from COVID-19 in France. *Biology* **2020**, *9*, 97. [CrossRef] [PubMed]

20. We're Sharing Coronavirus Case Data for Every U.S. County. *The New York Times*, 28 March 2020. Available online: https://www.nytimes.com/article/coronavirus-county-data-us.html?action=click&module=Spotlight&pgtype=Homepage (accessed on 15 April 2020).

21. Official oral report of New York Governor on COVID19. *PBS News*, 23 April 2020.

22. 21 Percent Of NYC Residents Tested In State Study Have Antibodies From COVID-19. *Gothamist*, 23 April 2020. Available online: https://gothamist.com/news/new-york-antibody-test-results-coronavirus (accessed on 23 April 2020).

23. The Institute for Health Metrics and Evaluation, University of Washington. COVID-19 Projections. 2020. Available online: https://covid19.healthdata.org/united-states-of-america/new-york (accessed on 1 May 2020).

24. Official oral report of New York Governor on COVID19. *PBS News*, 2 April 2020.