Control of COVID-19 outbreak using an extended SEIR model

Sean T. McQuade

Department of Mathematical Sciences and Center for Computational and Integrative Biology,
Rutgers University-Camden.
sn.mcquade@gmail.com

Ryan Weightman

Department of Mathematical Sciences and Center for Computational and Integrative Biology,
Rutgers University-Camden.
ryanj.weightman@gmail.com

Nathaniel J. Merrill

Pacific Northwest National Laboratory - PNNL.
njm.merrill@gmail.com

Aayush Yadav

Department of Psychology and Walter Rand Institute for Public Affairs, Rutgers
University-Camden.
ay331@scarletmail.rutgers.edu

Emmanuel Trélat

Sorbonne Université, Laboratoire Jacques-Louis Lions.
emmanuel.trelat@sorbonne-universite.fr

Sarah R. Allred

Department of Psychology and Walter Rand Institute for Public Affairs, Rutgers
University-Camden.
srallred@camden.rutgers.edu

Benedetto Piccoli

Department of Mathematical Sciences and Center for Computational and Integrative Biology,
Rutgers University-Camden.
b.piccoli@gmail.com

The outbreak of COVID-19 resulted in high death tolls all over the world. The aim of
this paper is to show how a simple SEIR model was used to make quick predictions
for New Jersey in early March 2020 and call for action based on data from China and
Italy. A more refined model, which accounts for social distancing, testing, contact tracing

and quarantining, is then proposed to identify containment measures to minimize the economic cost of the pandemic. The latter is obtained taking into account all the involved costs including reduced economic activities due to lockdown and quarantining as well as the cost for hospitalization and deaths. The proposed model allows one to find optimal strategies as combinations of implementing various non-pharmaceutical interventions and study different scenarios and likely initial conditions.

Keywords: COVID-19; SARS-CoV-2; SEIR compartmental models; non-pharmaceutical intervention; optimal control; data-fitting; New Jersey.

AMS Subject Classification: 22E46, 53C35, 57S20

1. Introduction

The COVID-19 outbreak gave rise to an unprecedented production of models and studies aimed at understanding the pandemic, predicting its evolution and designing measures to reduce its spread. A complete account would be fairly impossible and would likely become obsolete in short time. However, a review of some approaches is in order to better cast the contribution of this paper.

The birth of epidemiological models dates back to the pioneering work of Kormack and McKendrick [35], which proposed to divide the population in categories as susceptible, exposed, infected and recovered. Such models, using the first letter of populations, are called SEIR and are based on systems of differential equations. The infection and recovery rates dramatically affect the trajectories of model populations, making them key parameters in our model. Various generalizations of the SEIR model were used to model the COVID-19 spread, such as: 1) considering time-dependent infection rates, travel and zoonotic infections [16,47]; 2) including more sub-populations capturing different disease progressions and/or hospitalization [29]; 3) adding age-structure and spatial models [12,19,20,33,60,59]. Let us also briefly mention several other approaches not based on SEIR models. For instance, considering the space dynamics of infection gives rise to spatial models such as reaction-diffusion equations [5] and integro-differential ones [34]. On the other side, considering in-host as well as population dynamics can be achieved via multiscale models such as [8].

Every modeling approach presents advantages and disadvantages. We want to point out that the SEIR approach is useful for fast predictions, as shown in Section 2, as well as for large-scale optimization, as shown in Section 5. However, it has severe limitations in the lack of description of in-host dynamics including immune competition, limitation to only aggregated data, and difficulties in representing spatial component of the dynamics [36]. There has also been discussion on the role played by models for predictions [39,55]. The problem is particularly difficult as any model including human behavior in the system, see [40].

In this paper, we focus on modeling a combination of interventions to control the spread of the infection. Interventions include social distancing and other restrictive measures, testing for infected persons and contact tracing for persons who had contact with infected ones or traveled to areas with high rate of infections. The

action following testing and contact tracing is to mandate quarantine for a period of time. There has been effort in modeling and analyzing such control policies and, as said above, we report a subset of the feasible approaches to better understand effects of these interventions. Previous work in control of pandemic includes: quantifying the effect of containment measures [27], determining the controllability using daily data [14], considering individual reaction to non-pharmaceutical interventions [38], determining best timing of interventions [28,44], including testing and quarantining [6]. Moreover, some of the considered interventions were already modeled for other viruses such as HPV [13,48]. Finally, some papers focused on the economic cost considering uncertainty in data [30], cost of lockdown [1,4], hospital and ICU occupancy [18,49].

Our work developed over the past year starting with a simple SEIR model to predict the need of hospital beds in different social distancing scenarios. The model included discrimination between asymptomatic and symptomatic infected to better capture the spread dynamic and also provide an estimate of hospital bed occupancy and shortfall directly due to COVID-19. The model was instrumental in identifying the needs for the state of New Jersey as the spread was arriving from the neighboring New York, see [2] and was cited in a letter from New Jersey Governor Phil Murphy [41].

We then augmented the model to include three main control mechanisms: (1) general social distancing measures; (2) testing and (3) contact tracing. The focus of this work is on understanding the effects of the different controls and optimizing a cost function capturing the economic cost of the controls, healthcare, and population. More precisely our advanced SEIR-type model includes quarantined, hospitalized and deceased subpopulations. The dynamics introduces additional terms to model the effects of social distancing and lockdown measures, as well as the quarantined based on testing and contract tracing activities. The model is then tuned to data from New Jersey, separated into three groups of counties: Northern(Red), Central(Orange), and Southern(Yellow). This reflects the different characteristic of the pandemic evolution, which was affected by regional differences in population density and population structure. For instance the basic reproduction number R_0 was estimated to be higher in the northern and central regions compared to the southern one, while hospitalization rate was higher in the northern region than the central one.

We consider an optimal control problem using the NJ dataset and our model, wherein the cost is given by the sum of six terms: the economic cost of lockdown measures, the cost of testing, the cost of contact tracing, the economic cost of quarantining, the economic cost of hospitalization and, finally, the economic cost of deaths. The economic cost of lockdown and quarantining is estimated as a daily cost per person, while the economic cost of hospitalization is differentiated for bed cost when the occupancy is below or above the normal capacity. We simulated the scenario of detection of the infection spread after 15 days of appearance. The main results are the following: a) the lockdown should be as strict as possible for the first

60 days after detection; b) testing should be sustained at its maximum for the whole time horizon; c) the contact tracing should be activated only for two weeks after the detection. Interestingly enough, despite the difference in the characteristic of the pandemic spread in the three regions, the optimal policy appears to be reasonably uniform. We believe that the tools presented here could be useful in the management of the pandemic in different phases of its progression and for different states.

2. A SEIR Model for Quick Predictions of Hospital Bed Needs

In March 2020, a study [3] was developed by the Senator Walter Rand Institute and the Center for Computational and Integrative Biology, in collaboration with New Jersey Health Initiatives. This work compared the available hospital beds in eight counties of New Jersey, with the likely demand of hospital beds due to the COVID-19 outbreak, considering multiple virus spread scenarios. The focus was on using a simple model for a quick prediction of the imminent spread and an estimate of hospital beds needed to manage the spread.

2.1. *The SEIR model*

The work is based on an augmented form of a SEIR model [57] for COVID-19 in Wuhan and its international spread. Additional features were added to the model and parameters were fit with data from the Italian outbreak of February 2020. For each county, the following were estimated:

- Hopital bed availability;
- Expected day of peak in hospital bed demand considering three social distancing policies: minimal, moderate, and strong;
- Expected demand exceeding hospital bed capacity;
- Impact of social distancing on R_0 and lessening the spread of COVID-19.

The augmented SEIR model considered in (2.1) is given by:

$$\frac{dS}{dt} = -\frac{S}{N}\left(\frac{R_0}{D_I}I_A\right)$$

$$\frac{dE}{dt} = \frac{S}{N}\left(\frac{R_0}{D_I}I_A\right) - \frac{E}{D_E}$$

$$\frac{dI_A}{dt} = \alpha\frac{E}{D_E} - \frac{I_A}{D_I}$$

$$\frac{dI_S}{dt} = (1-\alpha)\frac{E}{D_E} - \frac{I_S}{D_I} \tag{2.1}$$

$$\frac{dH}{dt} = \sigma\frac{I_S}{D_I} - \frac{H}{D_H}$$

$$\frac{dR}{dt} = (1-\sigma)\frac{I_S}{D_I},$$

Name	Description	Estimate	Units
R_0	rate of infection	[2-6]	-
D_I	infectious period	5-7	days
D_E	latent period	4-14	days
D_H	hospitalization period	7-14	days
α	asymptomatic rate	0.81	-
σ	hospitalization rate	0.1-0.55	-

Table 1: Parameters for the initial SEIR model in (2.1).

where S are susceptible, E exposed, I_A infected asymptomatic, I_S infected symptomatic, R recovered, and $N = S + E + I_A + I_S + R$ is the total population. The main augmentation is the distinction between infected symptomatic I_S and infected asymptomatic I_A, so the total number of infected is given by $I = I_A + I_S$. The parameter α is the asymptomatic incidence estimated to be 0.81 by the study [58]. Also added is the hospitalization rate $\sigma = 0.55$ of symptomatic patients from the same study. The whole set of parameters are reported in table 1.

To fit the model, we used data from China and Italy. More precisely, data was gathered from a publication from the Chinese Center for Disease Control and Prevention on the largest case series to date of coronavirus disease [58], and from recent data at the time of writing about the impact on Italy [46] in order to find suitable model parameters. China's spread was already under control in March 2020 after strong social distancing and lockdown measures. Italy was the country with the highest infection rates in February and early March 2020 and data were publicly available. The number of hospitalizations in northern Italy quickly bumped from 200 on February 24 2020 to around 6,000 on March 11 2020 on a population of around 25 million affected in North Italy [46]). The resulting basic reproduction number was estimated as $R_0 = 3.9$. To perform simulations, we considered different choices of the basic reproduction number: $R_0 = 3.9$ as estimated by the Italian data, $R_0 = 2.68$ as suggested in the study [57] based on Wuhan data and $R_0 = 1.3$. The choices were reflecting the expected spread evolution in cases of different levels of intervention with social distancing measures. At the time, social distancing measures were still under design at state level as cases were quickly shifting from the neighboring New York State.

2.2. *Results of initial SEIR study*

The research brief [2] provided predictions based on possible "moderate intervention" and "strong intervention" with social distancing measures. A confidence interval of dates was given for each county, when the hospital bed demand was expected to

6   *S. McQuade, R. Weightman, N. J. Merrill,A. Yadav, E. Trélat,S. R. Allred, B. Piccoli*
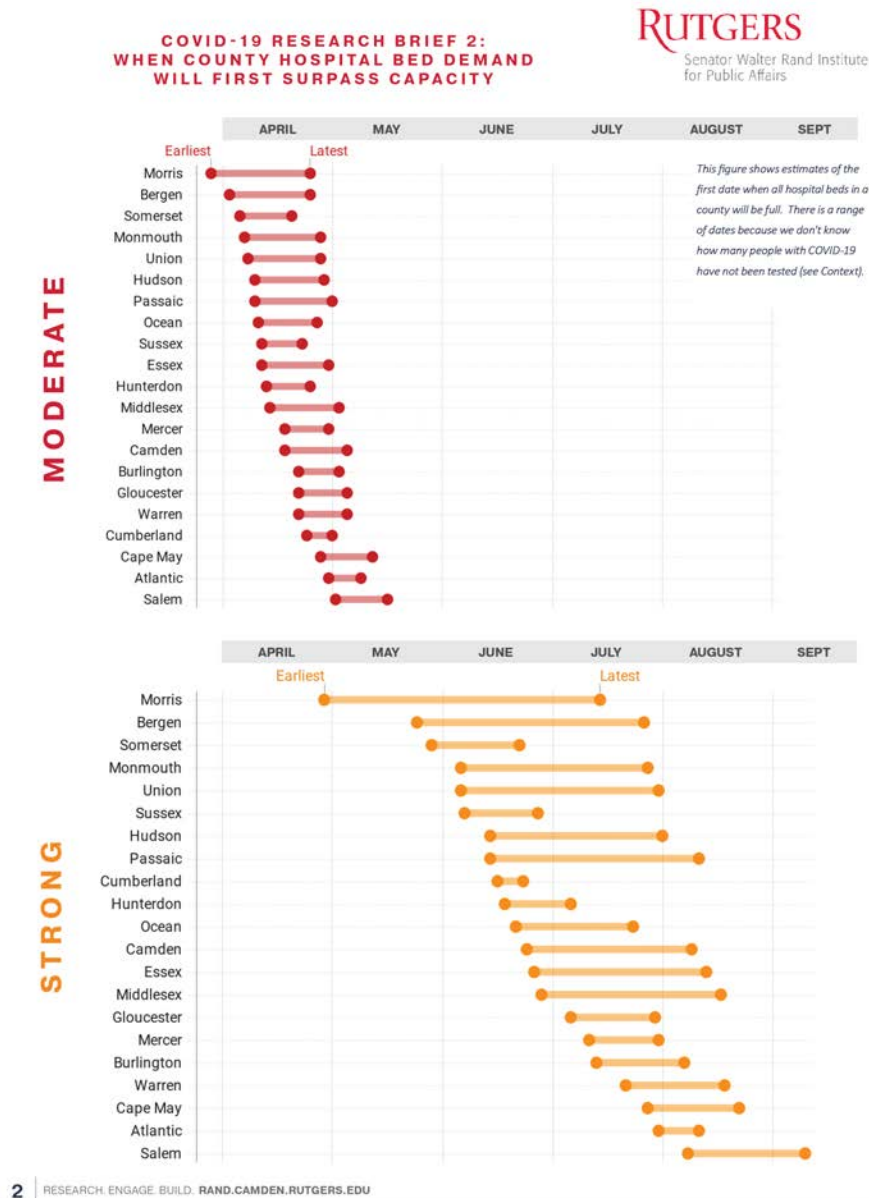

Fig. 1: Predictions to reach hospital bed capacity given implementing a moderate or strong social distancing policy from a March 16, 2020 research brief.

surpass capacity. Moreover, a range of parameterizations of the SEIR model (2.1) led to the prediction interval. At the time, and still now, it was not clear how the social distancing measures and lockdown would impact the basic reproduction number, thus we resorted to the two levels observed from Italian and Chinese data and the level $R_0 = 1.3$ corresponding to more robust measures. The motivation for this study was to forecast as early as possible the hospital bed shortfall in New Jersey from the COVID-19 pandemic. By quickly assembling a collaboration between the Center for Computational and Integrative Biology, the Senator Walter Rand Institute, and New Jersey Health Initiatives, we were able to release this briefing on March 16, 2020 and Gov. Murphy cited the study in a letter to President Trump [41]. Results are summarized in a graphic (Fig. 1).

3. Continued Model Fitting with New Jersey Data-set

In the following months the SEIR model (2.1) was continuously updated using data from New Jersey. To fit our model parameters we used three distinct sets of NJ counties where the virus spread differently. The zones were identified with a color (red, orange and yellow) and the map is shown in Figure 2 (below). The northern collection of counties are referred to as the "red zone" (population = 4,111,309). The central collection of counties are referred to as the "orange zone" (population = 2,835,111). The southern collection of counties are referred to as the "yellow zone" (population = 1,845,474). Figure 2 (above) reports the list of counties with onset dates, time for doubling of cases and deaths and other data.

3.1. *Fitting the reproduction number*

Estimates of the reproduction number were obtained using techniques developed by Thompson et al. [51]. This method deals with the *instantaneous reproduction number*, given by the average number of secondary cases that would arise from a primary case infected at time t given all causal factors remain fixed after time t [26]. Intuitively, the instantaneous reproduction number characterizes the "instantaneous" transmissibility at time t and does not require assumptions about the future [51], thus making it easier to estimate. Thompson et al. [51] also provide access to an online tool for the purpose of this task (the tool is available through the following URL [50]).

Serial intervals, defined as the time between successive cases in a single series of transmission, are typically interval-censored data studied in the form of approximate lower and upper bounds on the interval. These data are often reported by the way of household-interviews and analysis of hospital records such as in the case of Ebola [22]. Without an early stage contact-tracing protocol, accurately collecting data on transmission chains is a challenging task.

By initializing a gamma distribution prior for the serial intervals, we perform a Bayesian parameter estimation to obtain posterior samples of the serial interval distribution. Assuming that the number of daily (local) incidences is drawn from

8   *S. McQuade, R. Weightman, N. J. Merrill, A. Yadav, E. Trélat, S. R. Allred, B. Piccoli*

| countyName	StateAndCountyFIPS	JustCountyFIPS	nInThousandWith	nInThousandDead	doublingTimeForCases	doublingTimeForDeaths	Onset date March2020	Onset day
Union County	34039	39	13.86	0.5	11.55	8.07	10	0
Passaic County	34031	31	13.47	0.36	11.2	6.12	9	0
Hudson County	34017	17	13.42	0.49	14.85	5.65	9	0
Bergen County	34003	3	11.99	0.67	19.49	7.3	5	0
Essex County	34013	13	10.94	0.75	13.52	6.39	12	0
Middlesex County	34023	23	8.12	0.31	14.42	8.14	11	5
Morris County	34027	27	7.16	0.43	15.51	7.31	12	5
Ocean County	34029	29	6.97	0.29	15.85	11.03	14	5
Monmouth County	34025	25	6.54	0.26	23.42	6.69	10	5
Somerset County	34035	35	6.19	0.31	14.17	9.26	12	5
Mercer County	34021	21	5.06	0.21	11.58	5.1	15	5
Warren County	34041	41	4.31	0.28	11.6	6.2	19	0
Sussex County	34037	37	3.69	0.27	15.56	10.07	19	0
Camden County	34007	7	3.09	0.1	9.72	4.94	11	6
Burlington County	34005	5	2.81	0.1	14.58	3.3	11	6
Hunterdon County	34019	19	2.77	0.11	16.25	9.02	16	6
Gloucester County	34015	15	2.09	0.05	15.86	NaN	18	6
Cape May County	34009	9	1.83	0.1	10.32	NaN	19	6
Cumberland County	34011	11	1.34	0.02	9.29	NaN	20	6
Salem County	34033	33	1.27	0.06	6.56	NaN	22	6
Atlantic County	34001	1	1.17	0.05	8.33	NaN	17	6
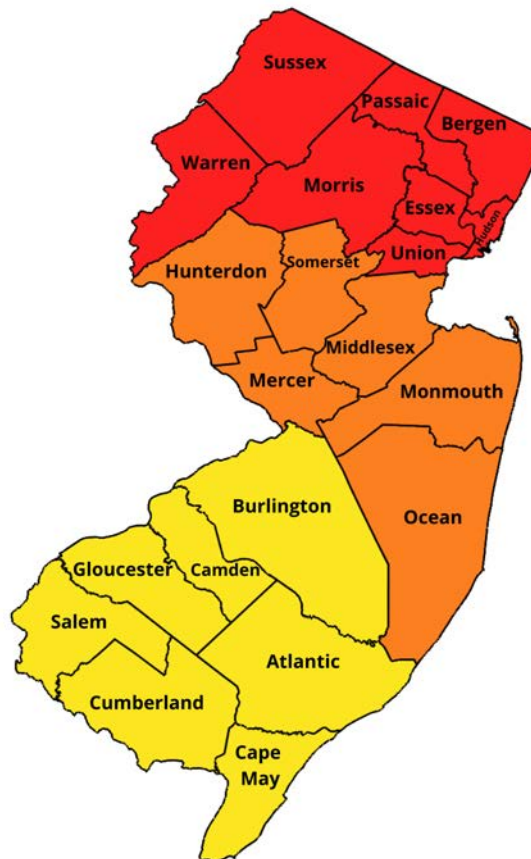
Fig. 2: Map of NJ counties according to colored zones (below) and table of data per each county (above).
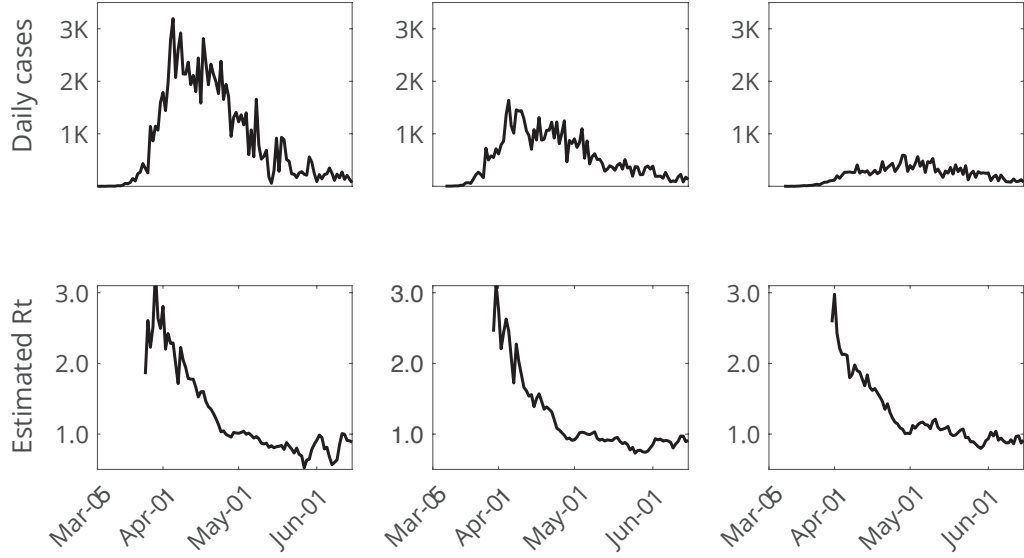
Fig. 3: Estimate of R_0 (Left:Red, Center:Orange, Right:Yellow) over time with confidence interval (day 1 is January 22nd).

a Poisson distribution, a gamma distributed prior (conjugate to the Poisson likelihood) can be used to obtain a fully analytical form of the posterior distribution of R_t, the time-varying reproduction number, given the incidence data up to time t and the serial interval distribution. Lacking any data on the transmission chains in the state of New Jersey at the time of our study, we parameterize the gamma distributed prior with a mean interval of 3.96 days with a standard deviation 4.75 days based on a study of 468 confirmed cases of coronavirus disease reported in China as of February 2020 [23]. The estimated R_t for the three regions is shown in Figure 3.

3.2. *Fitting the model with confirmed cases, hospitalizations and deaths*

After fitting the reproduction number R_t, we used further data from the Johns Hopkins public repository [21]: confirmed COVID-19 cases, COVID-19 related hospitalizations, and COVID-19 related deaths. As with R_0 we segregated R_t into the three sets of counties called "Zones": Red, Orange, and Yellow.

The confirmed cases, hospitalization, and deaths time series data was used to generate a seven day moving average of the daily cumulative totals; the moving average of the time series is used to account for irregularities in the data due to the day of the week. The deaths data was additionally processed; on June 26, 2020 many deaths were deemed resulting from COVID-19 that had previously been considered otherwise. This caused an abrupt increase in cumulative deaths from day

83 to 84.

To reflect the true cumulative deaths from COVID-19 over time, we took the difference from day 84 and 83 and spread this amount over day 1 to day 83, weighted by the proportion of deaths from day 1 to day 83. Let the increase in number of deaths from day 83 to day 84 for the red, orange, and yellow zones be, $J = 1126, J = 538, J = 132$ respectively. We find the proportions of J that we must add to every day before day 84. For cumulative deaths on day t, we have the raw deaths (the data with a large jump from day 83 to 84) and the normalized deaths (the data after we process it to smooth this jump) shown in Figure 4.

$$w(t) = D_R(t) - D_R(1)$$

$$D_N(t) = \begin{cases} J\frac{D_R(t)}{w(t)} & \text{if } t \leq 83 \\ D_R(t) & \text{otherwise} \end{cases}$$
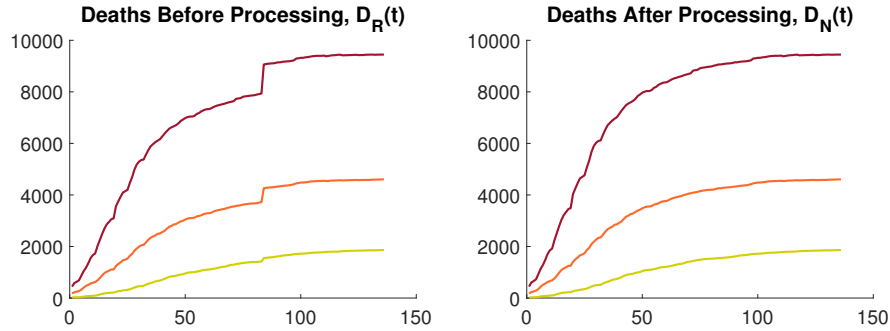

Fig. 4: Left: Artifact in raw Deaths dataset due to retroactively labeling a large group of deaths as resulting from COVID, and Right: the normalized dataset after we distribute the large jump over all previous days.

We considered three ways to interpret how confirmed cases data informs our model. A person who tests positive may not necessarily be sick already, in which case the positive test would move one person to the I_S compartment. If they were not yet sick, this means the positive test moves one person to I_A because they are asymptomatic. Because testing was scarce in the beginning of the pandemic, there is reason to believe that a vast majority of those testing positive were symptomatic and should therefore count towards the I_S population. In later months, increased testing suggests that there would be more test positive for asymptomatic people.

(1) All confirmed cases contribute to I_S.
(2) 100% of cases contribute to I_S until April 20, then cases contribute 50% I_S,

50% I_A by May 31. The contribution percentage is interpolated linearly between these two dates.

(3) All confirmed cases contribute to I_S until April 20, then no longer fit data for confirmed cases afterward.
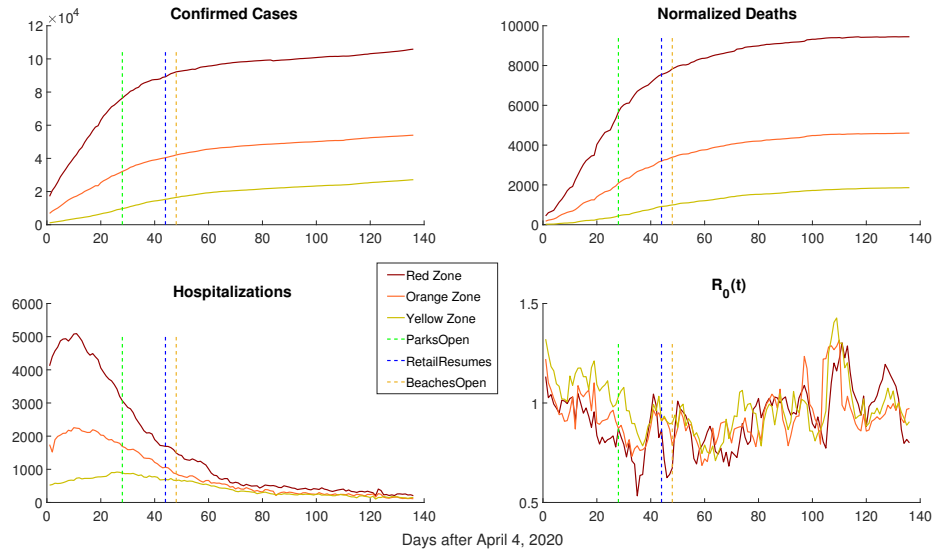

Fig. 5: Data starting from April 2020. Top Left: 7-day rolling average of confirmed cases; Top Right: 7-day rolling average of confirmed deaths. Bottom Left: Confirmed cases resulting in hospitalization; Bottom Right: Estimated reproduction rate.

With these four datasets: calculated R_t, confirmed cases, hospitalizations, and deaths, we now fit model parameters using a least squares optimization scheme. The optimization script was written with AMPL code ("A Mathematical Programming Language"). The optimization algorithm requires that we use a continuous function to represent each of the four time series datasets we've described above. We used matlab to interpolate all four datasets, transforming the time series into a degree 8 polynomial. Then, to find the best fitting parameters, we chose reasonable bounds in the parameter space over which the fitting was done. The bounds for D_I, D_E, and D_H are shown in the estimate column of Table 1 whereas the other parameters had the following intervals: $R_0 \in [0, 15]$, $\alpha \in [0, 1]$, $\sigma \in [0, 1]$, $r \in [0, 1]$, $E_0 \in [0, 8 \times 10^6]$, $I_{A_0} \in [0, 8 \times 10^6]$. Figure 6 reports the fitting results for the red zone.
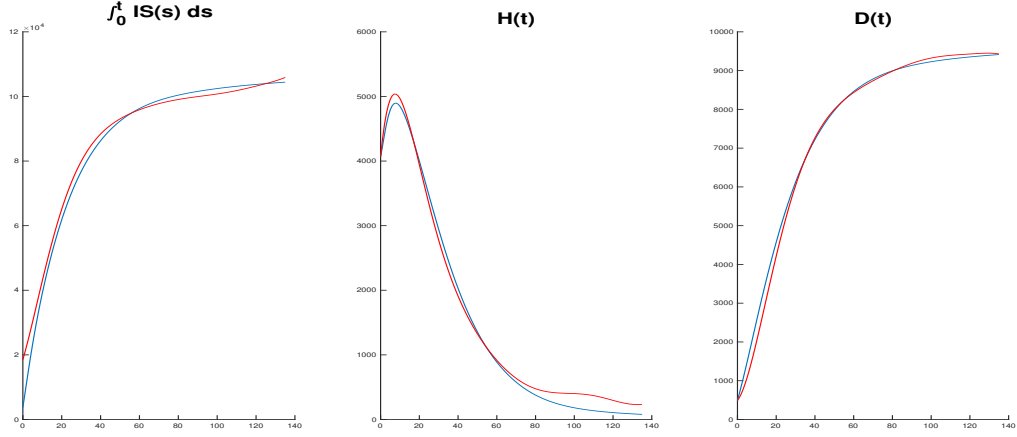
Fig. 6: Result of fitting infected, hospitalized and death data for the red zone. Blue line: data from the Johns Hopkins public repository. Red line: model trajectory using optimized parameters.

zone	max R_t	D_I	D_E	D_H	α	σ	r	E_0	I_{A_0}
Red	2.79	5.00	4.00	10.60	0.39	0.14	0.47	38, 682	1.85
Orange	3.26	5.00	4.01	10.87	0.35	0.17	0.43	15, 159.3	1.33136
Yellow	1.66	5.00	4.00	7.52	0.62	0.32	0.21	5141.2	407.32

Table 2: Values of parameters per zone.

The optimization to model parameters was done with chosen "reasonable" bounds for each parameter and we penalized max R_t if it was lower than 2.5 and higher than 3.5; max R_t was strictly bounded between 0 and 15; D_I,D_E,D_H were strictly between 5 and 7, 4 and 14, and 7 and 14 respectively; α was bound between 0 and 1.

The main takeaways after fitting parameters are as follows. The basic reproduction max R_t number was similar for the three zones with a higher value for the orange zone. The time duration of exposure D_E, infection D_I, and hospitalization D_H were reasonable values determined by the optimizer, and they were similar among the zones. The initial asymptomatic infected population I_{A_0} is smallest for the red zone and largest for the yellow zone. The initial Exposed E_0 was highest in the red zone , and smallest in the yellow zone which reflects the total population in these zones. σ was bound between 0.1 and 0.55; r was bound between 0 and 1; E_0 was bound between 0 and 8 million. We penalized I_{A_0} if it was greater than 500, with a strict upper bound of 8 million.

4. A Model for Social Distancing, Testing and Contact Tracing

Measures to keep the pandemic under control included: testing for the presence of the virus, contract tracing for people that had recent contact with infected people, and social distancing measures to reduce the reproduction number. Moreover, quarantining was put in effect for detected infected persons, but also via contact tracing.

To model these measure, we add three populations: S_Q, resp. E_Q, the population of susceptible, resp. exposed, that are quarantined by effect of contact tracing and I_Q of infected that are quarantined by either test or contact tracing. We also add the number of deceased D, which will serve later to estimate the economic cost of the pandemic and the controls.

The social distancing and lockdown measure are captured by a control variable $u \in [0, 1]$, which multiplies the basic reproduction number R_0. In other words, this represents the social behavior that limits number of asymptomatic infected exposing other populations to the virus. Moreover, two other controls are added: a parameter δ corresponding to tests and μ related to contact tracing. The final model reads:

$$
\begin{aligned}
\frac{dS}{dt} &= -\frac{S}{N}\left(\frac{u\,R_0}{D_I}(I_A)\right) - \mu\frac{S}{S+E+I_A}\frac{\delta I_A}{S+E+I_A} + \frac{S_Q}{D_Q} \\
\frac{dE}{dt} &= \frac{S}{N}\left(\frac{u\,R_0}{D_I}(I_A)\right) - \mu\frac{E}{S+E+I_A}\frac{\delta I_A}{S+E+I_A} - \frac{E}{D_E} \\
\frac{dI_A}{dt} &= \alpha\frac{E}{D_E} - \delta\frac{I_A}{S+E+I_A} - \mu\frac{I_A}{S+E+I_A}\frac{\delta I_A}{S+E+I_A} - \frac{I_A}{D_I} \\
\frac{dI_S}{dt} &= (1-\alpha)\frac{E+E_Q}{D_E} - \frac{I_S}{D_I} \\
\frac{dS_Q}{dt} &= \mu\frac{S}{S+E+I_A}\frac{\delta I_A}{S+E+I_A} - \frac{S_Q}{D_Q} \\
\frac{dE_Q}{dt} &= \mu\frac{E}{S+E+I_A}\frac{\delta I_A}{S+E+I_A} - \alpha\frac{E_Q}{D_Q} - (1-\alpha)\frac{E_Q}{D_E} \\
\frac{dI_Q}{dt} &= (1-\sigma)\frac{I_S}{D_I} + \delta\frac{I_A}{S+E+I_A} + \mu\frac{I_A}{S+E+I_A}\frac{\delta I_A}{S+E+I_A} - \frac{I_Q}{D_Q} \\
\frac{dH}{dt} &= \sigma\frac{I_S}{D_I} - \frac{H}{D_H} \\
\frac{dR}{dt} &= \frac{I_A}{D_I} + \alpha\frac{E_Q}{D_Q} + \frac{I_Q}{D_Q} + (1-r)\frac{H}{D_H} \\
\frac{dD}{dt} &= r\frac{H}{D_H}
\end{aligned}
\tag{4.1}
$$

The equation for susceptible S has two additional terms: the first is negative and reflects the quarantining as result of contact tracing and the second is positive and reflects the end of quarantining. Notice that the first term is obtained by multiplying the number of tested positive (i.e. $\frac{\delta I_A}{S+E+I_A}$) and the fraction of susceptible among individual quarantined because of contact tracing (i.e. $\mu\frac{S}{S+E+I_A}$). The second term

is simply obtained from S_Q using the quarantining duration in days D_Q.

The equation for exposed E has similar terms, while the equation for infected asymptomatic I_A has the additional negative term due to positive tests $\delta \frac{I_A}{S+E+I_A}$. Also, the correction for the infected symptomatic I_S is only due to the new population E_Q.

The equations for the new quarantined populations S_Q, E_Q and I_Q are as follows. S_Q has two terms corresponding those of S since in quarantine susceptible will not be infected. The equation for E_Q has the positive term corresponding to the negative for E and two negative terms as for transition either to recovered at the end of quarantine, i.e. D_Q days, in the asymptomatic case and to infected symptomatic for the symptomatic case, with the same ratio α as for E. Finally, I_Q represent the infected quarantined which are necessarily asymptomatic or mildly symptomatic. Therefore the equation has three positive terms for symptomatic which are mild and not hospitalized, and asymptomatic found by testing and contact tracing, and one negative term for quarantine ending.

The equation for hospitalized H is as before, the recovered include terms for the end of quarantine and a correction from hospitalized that transition to the new category D of deceased.

The bounds for the control variables are as follows:

- Control $u = 1$ means no intervention, while $u = 0$ would represent total lockdown. For our simulations, we subject $u(t)$ to constraints to ensure that $R_0 u(t) \in [0.8, R_0]$ for R_0 measured before any social distancing policy.
- δ is the amount of testing of infected not symptomatic population to detect Sars-Cov-19 presence. Given a maximum test-per-day availability M_1 we have the following constraint $\delta \leq M_1$.
- μ is the amount of contact tracing to quarantine persons which had contacts with I_A detected by test. Given a maximum tracing-per-day availability M_2 we have the following constraint $\mu \leq M_2$.

This model has new variables and parameters in addition to those of the initial SEIR model (2.1). The quarantine time was chosen as $D_Q = 14$ days, while the other model parameters were fit to data as explained in Section 3.

5. Optimal Control

In this section we propose an optimal control problem for the model (4.1) with controls representing social distancing, testing, contact tracing and quarantining. We use the model fit to the three New Jersey areas (Red, Orange and Yellow) and optimize the strategy over the economic cost of pandemic management under the scenario that the infection spread is detected after 15 days of first case.

We define the cost function $C = C(u, \delta, \mu, S, E, I_A, H)$ by

$$C = C_1 + \cdots + C_6,$$

where

$$C_1 = c_1 S_0(1-u), \quad C_2 = c_2\delta, \quad C_3 = c_3\mu,$$

$$C_4 = c_4(S_Q + E_Q + I_Q), \quad C_5 = c_5\psi(H), \quad C_6 = c_6 D$$

and the constants $c_1, \ldots, c_6$ and the function ψ are to be defined. The meaning of the cost function is as follows. The term C_1 reflects the cost of social distancing: $u = 1$ correspond to no measure so zero cost, while $u = 0$ corresponds to total lockdown so a loss of c_1 dollar per day per susceptible person. The term C_2 represent the cost of testing and C_3 of contact tracing. The cost C_4 is the cost of quarantining with a loss of c_4 dollars per day per quarantined person. The cost C_5 represent the cost of hospitalization. More precisely we define:

$$\psi = \begin{cases} H & \text{for } H \leq H_{\text{capacity}} \\ H_{\text{capacity}} + 10(H - H_{\text{capacity}}) & \text{for } H > H_{\text{capacity}} \end{cases} \quad (5.1)$$

where H_{capacity} is the number of available hospital bed before the pandemic onset. Therefore the cost of hospitalization is c_5 dollars per day per hospitalized person up to capacity. Then the cost is multiplied by a factor of 10 representing the need of creating new capacity by building hospitals. Finally, the term C_6 capture the social cost c_6 dollars per deceased person.

To obtain an explicit expression of the cost, we use the parameters in the following table:

$C_1 = c_1 S_0(1-u)$	$c_1 =$	70,	cost of social distancing,
$C_2 = c_2\delta$	$c_2 =$	100,	cost of testing,
$C_3 = c_3\mu$	$c_3 =$	100	cost of contact tracing,
$C_4 = c_4(S_Q + E_Q + I_Q)$	$c_4 =$	70	cost of quarantining,
$C_5 = c_5\psi(H)$	$c_5 =$	2700	cost of hospitalization,
$C_6 = c_6 D$	$c_6 =$	1,500,000	cost of mortality.

The values of c1-c6 shape the cost function; thus, choosing reasonable values for these constants is important. However, choosing appropriate constants is challenging because the actual costs associated with these constants (social distancing, testing, contact tracing, quarantining, hospitalization, and death) varied by jurisdiction and also changed over the course of the pandemic. To select values, we searched for estimates of each cost in New Jersey (or the country) during the pandemic. Here we explain the context for our choices.

As discussed, we cannot know true values of c1-c6 with complete certainty. Thus, confidence in the model rests on the assumption that the difference between selected c1-c6 and true c1-c6 values are in a range that would not change the model output. To test the security of that assumption, we conducted a sensitivity analysis (described below). This analysis showed model outputs are similar provided that true c1-c6 values are within roughly an order of magnitude of selected c1-c6. Because

the variation in estimates of each constant was well within an order of magnitude, we are confident in using the selected values of c1-c6 to inform the model.

To estimate the cost of social distancing (C1), we used two estimates of the economic cost during lockdown. The president of the St. Louis Federal Reserve bank estimated on April 16th, 2020 that the lockdown cost the US economy 25 billion per day [43]. With a US population of 331 million on that date [15], this puts the cost of social distancing at $76 per person per day. A second estimate from the Wall Street Journal suggests that the economic shutdown would decrease GDP in the US by 30% [32]. The World Bank estimates GDP per capita in the US in 2019 as 65,000 [7]. This puts the cost of shutdown per person per day as $53 (30%*65,000/365).

Although the cost of a COVID-19 test varied by provider and time during the pandemic, the New York Times estimates that the cost of a test (C2) averaged about $100 [52].

In the fall of 2020, New Jersey issued a 37 million contract for contact tracing for the six month period from October 2020-March of 2021 [31]. During that time period, New Jersey had 694,000 COVID-19 cases [21]. The percentage of contacts traced varied in that time. On 15 days sampled between October 1st and March 31st from the New Jersey COVID-19 dashboard, on average 62% had successful contact tracing [42]. At 37 million for 420,280 cases with contact tracing, this is $86 per person for contact tracing (c3).

The cost per person for quarantining (c4) can vary widely. Some individuals quarantine at home, while others are in hotels. In some cases quarantining also includes lost wages or unemployment, while in other cases individuals continue to work from home. In February of 2021, the estimated cost of quarantine per person was $430 for a 7 day quarantine, or $61 per person [10].

Hospitalization costs (c5) also vary widely. In New Jersey, the average estimated cost of hospitalization per day is 2,786 per person per day [24]. Once beds are full, the cost of constructing a new hospital is large. One estimate is $70 million for 3000 bed hospital, which puts the cost at about 23,000 per bed.

Finally, estimates of early mortality vary widely. For insurance and and policy guidelines, different entities place economic valuations on human life. When deciding on whether to recommending health interventions, for example, it is common to weigh the cost of an intervention against the years of quality life it helps achieve. While there is a range of estimates for COVID, one analysis puts the cost of an average COVID-19 death at $1.5 million [17].

We consider the optimal control problem in Bolza form over the time horizon $[0, T]$

$$\min_{u(\cdot),\delta(\cdot),\mu(\cdot)} \int_0^T C(u(\cdot),\delta(\cdot),\mu(\cdot),S(t),E(t),I_A(t),H(t)) \, dt \qquad (5.2)$$

for the dynamics (4.1).

5.1. *Optimization algorithm*

The optimal controls are included in two steps:

(1) Define bounds for the controls; bounds for $u(t)$ are calculated from the estimates of R_0.

$$u(t) \in [a, 1], \qquad \mu(t) \in [0, 1e4] \qquad \text{and} \qquad \delta(t) \in [0, 1e4],$$

where $a = \frac{0.8}{\max R_0(t)}$. The choice of a ensures that the lowest measured $R_0(t) = 0.8$ which is consistent with data.

(2) Optimize to find the levels of social distancing $(u(t))$, testing $(\delta(t))$, and contact tracing $(\mu(t))$ that will minimize the economic costs. These controls are introduced formally in Section 5.

To compute numerically the optimal solutions of our optimal control problem, we choose here a *direct method*. Very briefly, when dealing with an optimal control problem, one usually distinguishes between direct and indirect methods. Direct (transcription) methods consist in discretizing the whole problem (discretize the system, e.g., by some Runge-Kutta method; discretize the cost functional, e.g., by some trapezoidal rule) so that we end up with a high-dimensional, but classical, nonlinear optimization problem under equality and inequality constraints, which can then be handled thanks to an optimization solver implementing, e.g., a gradient-like method based on the KKT rule, or a dual method like Uzawa. In other words, in the direct method, we *first discretize, then optimize (or dualize)*. In the indirect approach, in contrast, we first apply a first-order necessary condition for optimality to the optimal control problem, i.e., we apply the *Pontryagin maximum principle* (see [11,37,45,54]), which leads to a *shooting problem* that can be solved, numerically, thanks to a Newton-like method (see [9] for well-posedness issues). In other words, in the indirect approach, we *first optimize (or dualize) and then discretize*. We refer to [53,54] for a survey on these methods and on the pros and cons of direct vs indirect approaches.

Here, we choose the direct transcription approach because our optimal control problem involves state constraints that would be difficult to handle in the Pontryagin approach. Moreover, direct methods are much softer insofar they allow to change the model very easily. We discretize the control system with the implicit RK2 scheme and the cost functional with the trapezoidal rule, on a regular subdivision of the time interval (we take: one step = one day).

The numerical implementation of the optimal control problem is done by combining the sophisticated modeling language AMPL [25] (which encapsulates automatic differentiation in a very efficient way) with the open-source expert optimization routine IpOpt [56].

The initialization of the algorithm is done in a very simple way by initializing the discretized controls to (any) constant value and the discretized states to a constant value that is corresponding to the initial condition of each state. This

rough initialization suffices to ensure convergence. Execution on a standard desktop machine is almost instantaneous.

5.2. *Results*

Using the optimization algorithm of Section 5.1, we identified the optimal policies for lockdown measures ($u(t)$), testing ($\delta(t)$), and contract tracing (δ) assuming an initial period of inaction of 15 days, due to lack of detection of the virus, and time horizon of 200 days. The algorithm was used for the three different zones, thus using different parameters specified in Table 2. Our main results are shown in Figures 7, 8 and 9.

In figure 7, the optimal policy for the red zone is shown, together with the evolution of all the variables of the model (4.1). Let us first analyze the time-evolution of the optimal controls. The lockdown captured by $u(t)$ evolves as follow. The first 15 days we have $u = 0$ due to lack of virus detection so no measure taken. Then for around 85 days a maximum lockdown is imposed, with the reproduction number pushed down to the chosen minimum of $R_t = 0.8$. After those 85 days, the lockdown is completely removed, thus $u = 1$ for the rest of the simulation. The testing $\delta(t)$ is kept at its maximum after the first 15 days of inaction and completely stopped just after the lockdown lift. On the other side the contact tracking $\mu(t)$ is activated at its maximum after 15 days, but is quickly reduced to zero in around three weeks.

The resulting effect of these control policies can be clearly seen in the evolution of the populations. The susceptible population S decreases due the infection in the first 15 days, then has a large dip due to quarantining, with a bounce back completed around day 50, due to the stop in contact tracing. A reversed evolution can be seen in S_Q and the other quarantined populations E_Q and I_Q. The exposed population E quickly decreases because of transition to infected and the control measures. The infected populations I_A and I_S exponentially increase the fist 15-20 days and then are kept completely under control. The evolution of the hospitalized population has consequently the same characteristic, with a peak above 3000 around day 20, with a decrease to close to zero around day 60. The recovered quickly increase for the first month but then saturate at around two months. Similarly deaths grow quickly in the first month, then saturate reaching a total of around 5000 at the end of the time horizon.

Lastly, it is interesting noticing the cost evolution. Costs C_1, C_2 and C_3 as expected grow linearly when the corresponding controls are active. Cost C_4 grows quickly due to heavy quarantining in first 50 days then saturate. The time-evolution of costs C_5 and C_6 is similar since hospitalizations and deaths are on the high end only for a couple of months.

It is interesting to notice that the maximum cost is given by the lockdown measures with final cost of the order of 10^{13}, i.e. around 200 billions. Despite the high level of this cost, the optimal strategy is to pay the high cost of lockdown to prevent

the infection from spreading and becoming uncontrollable. If that happens, then, despite the fact that the other costs are lower per day, they would accumulate over time giving rise to a higher final cost.

In figure 8, the optimal policy for the orange zone is reported. We do not notice big differences with respect to the red zone. An initial strong lockdown lifted around day 100 is combined with testing lifted around the same time as the lockdown, and contact tracing active only for around a month.

In figure 9, we report the optimal policy for the yellow zone. The most significant difference is the longer period of total lockdown prolonged up to around day 117, so for three months, while testing is stopped around day 75.

To explore the sensitivity of the optimal policy w.r.t. the chosen cost parameters, we performed various simulations for the yellow zone varying the cost coefficients as well as the initial populations. The optimal policy is very robust with initial strict lockdown (u at minimum value) lifted completely at a fixed date. Contact tracing and testing have similar behavior. In figure 10, we report the dependence of the date of lockdown lift as a function of some parameters: the cost c_1 of lockdown per person per day, the initial population of exposed and the economic cost c_6 of death. We noticed a stronger dependence on cost of lockdown which, when varied in our range from \$70 to \$700, diminished the lockdown duration by almost 40 days. In contrast, we see a milder dependence on the number of initial exposed and economic cost of deaths with a change of the order of few days when varying the initial exposed up to 10,000 and the economic cost of death up to \$ 10 millions.
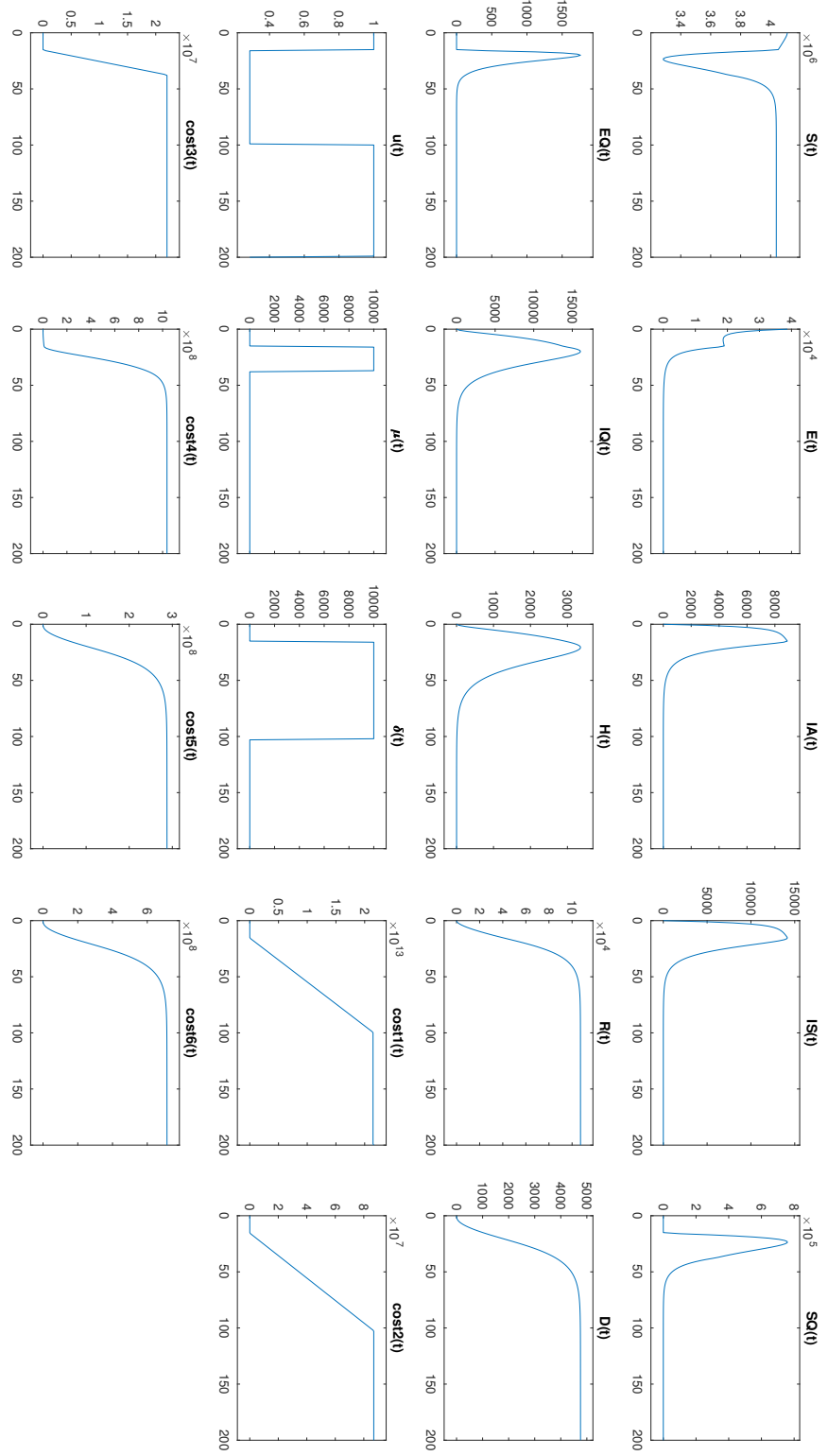
Fig. 7: Red zone results.
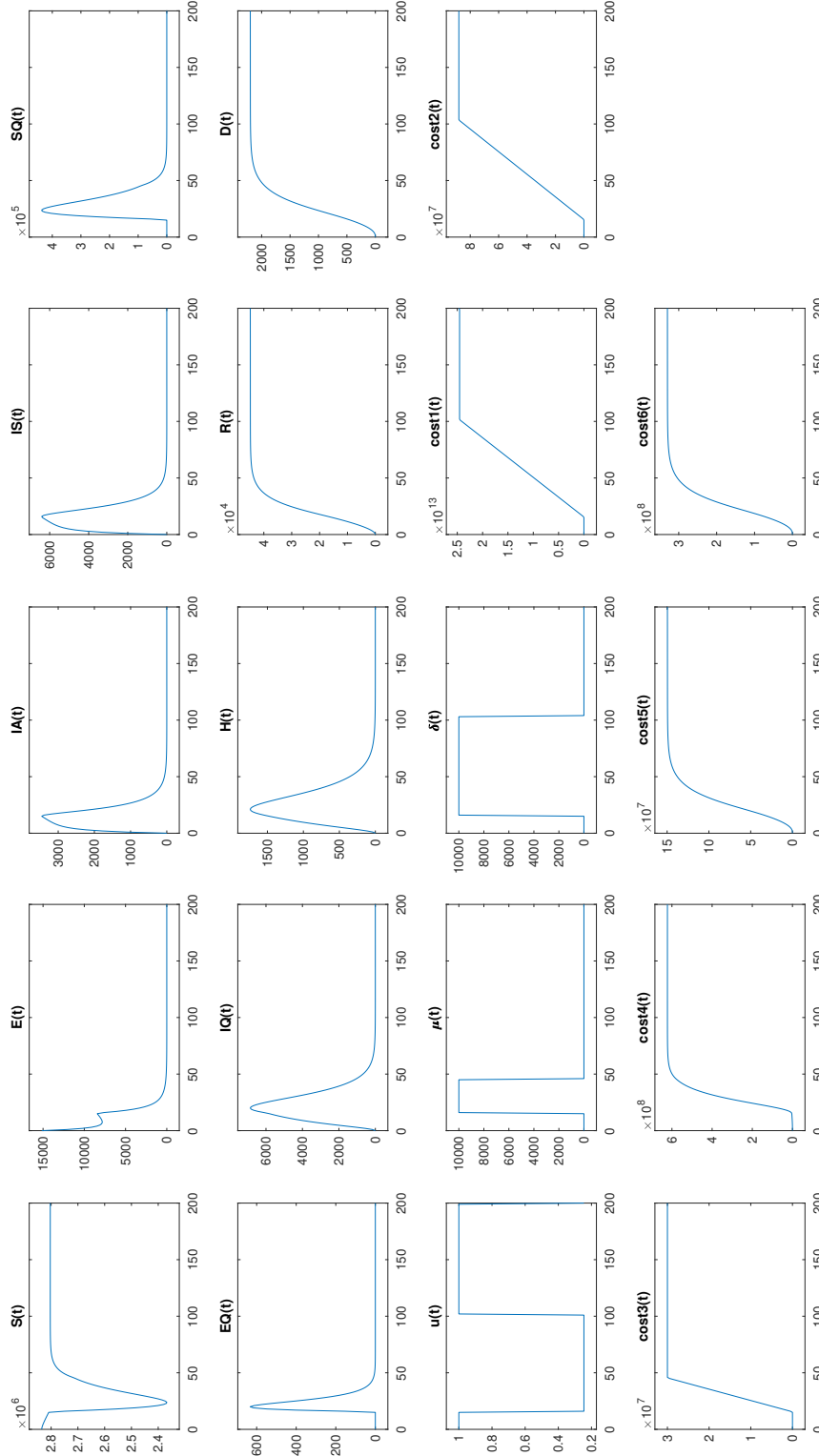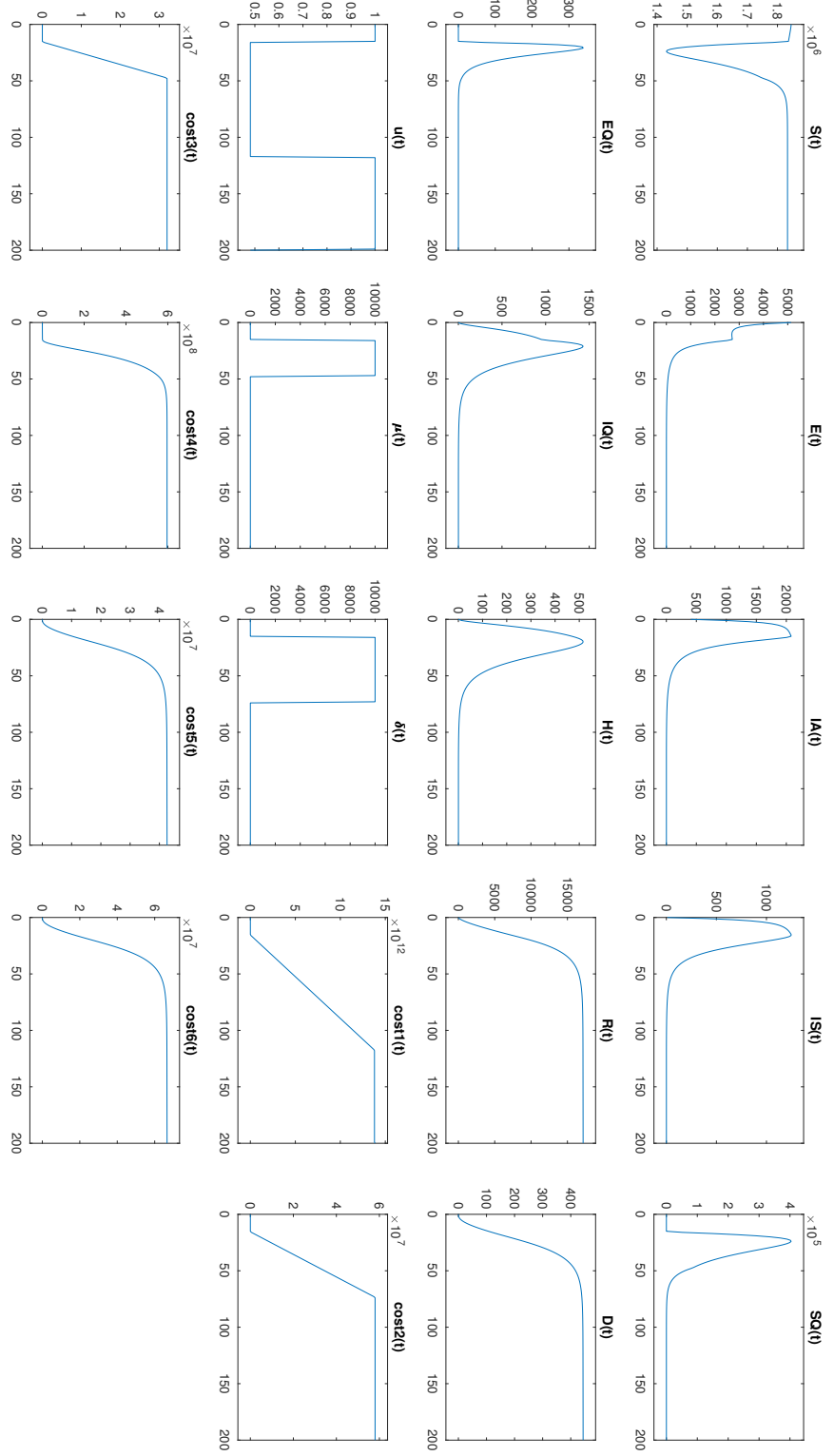
Fig. 8: Orange zone results.
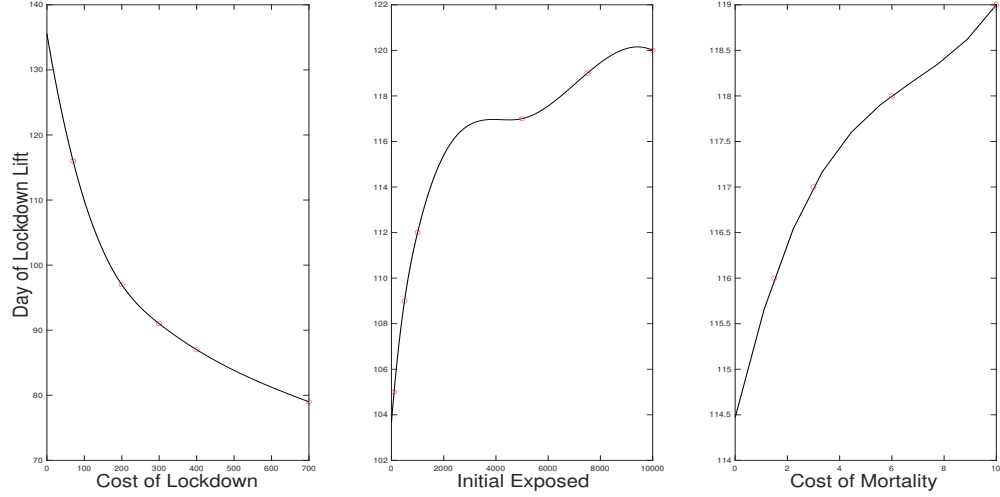
Fig. 9: Yellow zone results.

Fig. 10: Sensitivity analysis of lockdown lift for the yellow zone. Left: cost of lockdown per day. Center: initial population of exposed. Right: economic cost of death.

6. Conclusions

In this paper we showed two main results on the use mathematical modeling to fight the COVID-19 pandemic.

First we showed a simple SEIR model, including hospitalization, which was particularly useful at the onset of the spread in New Jersey. Such simple models are ready to use but are limited to very first estimates.

Second, we proposed a more complex model, still stemming from SEIR approach, including social distancing and lockdown measures, testing, contact tracing and quarantining. Our main idea was to estimate the economic cost of the pandemic taking into account the impact of social distancing and quarantining, but also of hospitalization, taking into account the limited capacity, and deaths. The main result is that the best strategy consists of an immediate and strict lockdown for two to three months followed by a reopening. The testing and consequent quarantining should be kept at capacity for around the same time as lockdown (except for one zone in New Jersey with earlier termination), while contact tracing is useful only at the onset for few weeks and then can be dismissed. Beside these results with specific tuning to New Jersey data, the model can be used to explore different scenarios and fit to data from different states and countries.

Acknowledgements

Authors' affiliations
Sean T. McQuade, Ryan Weightman, and Benedetto Piccoli: Department of Mathematical Sciences and Center for Computational and Integrative Biology, Rutgers University-Camden.
Nathaniel J. Merrill: Pacific Northwest National Laboratory - PNNL.
Aayush Yadav and Sarah R. Allred, Department of Psychology and Walter Rand Institute for Public Affairs, Rutgers University-Camden.
Emmanuel Trélat: Sorbonne Université, Laboratoire Jacques-Louis Lions.

References

1. Daron Acemoglu, Victor Chernozhukov, Iván Werning, and Michael D Whinston. *Optimal targeted lockdowns in a multi-group SIR model*, volume 27102. National Bureau of Economic Research, 2020.
2. Sarah R Allred, Sean T McQuade, Nathaniel J Merrill, Benedetto Piccoli, Darren Spielman, Carla Villacis, Ross Whiting, Aayush Yadav, Doug Zacher, and Devon Ziminski. Regional health system shortfalls with a novel covid-19 model. https://rand.camden.rutgers.edu/files/COVID-research-brief-with-curves03152020.pdf, 2020.
3. Sarah R Allred, Sean T McQuade, Nathaniel J Merrill, Benedetto Piccoli, Darren Spielman, Carla Villacis, Ross Whiting, Aayush Yadav, Doug Zacher, and Devon Ziminski. Regional health system shortfalls with a novel covid-19 model. 2020.
4. Fernando E Alvarez, David Argente, and Francesco Lippi. A simple planning problem for covid-19 lockdown. Technical report, National Bureau of Economic Research, 2020.
5. Sebastian Aniţa and Vincenzo Capasso. Reaction-diffusion systems in epidemiology. *An. Ştiinţ. Univ. Al. I. Cuza Iaşi. Mat. (N.S.)*, 66(2):171–196, 2020.
6. Maria Soledad Aronna, Roberto Guglielmi, and Lucas Machado Moschen. A model for covid-19 with isolation, quarantine and testing as control measures. *Epidemics*, 34:100437, 2021.
7. The World Bank. Gdp per capita. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD, 2020.
8. Nicola Bellomo, Richard Bingham, Mark A. J. Chaplain, Giovanni Dosi, Guido Forni, Damian A. Knopoff, John Lowengrub, Reidun Twarock, and Maria Enrica Virgillito. A multiscale model of virus pandemic: heterogeneous interactive entities in a globally connected world. *Math. Models Methods Appl. Sci.*, 30(8):1591–1651, 2020.
9. Bernard Bonnard, Jean-Baptiste Caillau, and Emmanuel Trélat. Second order optimality conditions in the smooth case and applications in optimal control. *ESAIM Control Optim. Calc. Var.*, 13(2):207–236, 2007.
10. Margaret Bourdeaux, Jessica Kaushal, Linda Bilmes, Annmarie Sasdi, Megan Mishra, and Anne Hoyt. Estimating the costs and benefits of supported quarantine and isolation in massachusetts: The missing link in covid-19 response. HKS Working Paper No. RWP21-003, 2021.
11. Alberto Bressan and Benedetto Piccoli. *Introduction to the mathematical theory of control*, volume 2 of *AIMS Series on Applied Mathematics*. American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.
12. Tom Britton, Frank Ball, and Pieter Trapman. A mathematical model reveals

the influence of population heterogeneity on herd immunity to sars-cov-2. *Science*, 369(6505):846–849, 2020.

13. V. L. Brown and K. A. Jane White. The role of optimal control in assessing the most cost-effective implementation of a vaccination programme: HPV as a case study. *Math. Biosci.*, 231(2):126–134, 2011.

14. Francesco Casella. Can the covid-19 epidemic be controlled on the basis of daily test reports? *IEEE Control Systems Letters*, 5(3):1079–1084, 2021.

15. US census. Us and world population clock. https://www.census.gov/popclock/, 2020.

16. Yi-Cheng Chen, Ping-En Lu, Cheng-Shang Chang, and Tzu-Hsuan Liu. A time-dependent SIR model for COVID-19 with undetectable infected persons. *IEEE Trans. Network Sci. Eng.*, 7(4):3279–3294, 2020.

17. The Apothecary Chris Conover. How economists calculate the costs and benefits of covid-19 lockdowns. https://www.forbes.com/sites/theapothecary/2020/03/27/how-economists-calculate-the-costs-and-benefits-of-covid-19-lockdowns/?sh=44a20e846f63, 2020.

18. M. Chyba, Y. Mileyko, O. Markovichenko, R. Carney, and A. E. Koniges. Epidemiological model of the spread of covid-19 in hawaii's challenging fight against the disease. In *The Ninth International Conference on Global Health Challenges GLOBAL HEALTH 2020*. IARIA, 2020.

19. Rinaldo M Colombo and Mauro Garavello. Well posedness and control in a nonlocal sir model. *Appl. Math. Optim.*, pages 1–35, 2020.

20. Rinaldo M. Colombo, Mauro Garavello, Francesca Marcellini, and Elena Rossi. An age and space structured SIR model describing the Covid-19 pandemic. *J. Math. Ind.*, 10:Paper No. 22, 20, 2020.

21. Ensheng Dong, Hongru Du, and Lauren Gardner. An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, 20(5):533–534, 2020.

22. Scott F. Dowell, Rose Mukunu, Thomas G. Ksiazek, Ali S. Khan, Pierre E. Rollin, and for the Commission de Lutte contre les Epidémies à Kikwit Peters, C. J. Transmission of Ebola Hemorrhagic Fever: A Study of Risk Factors in Family Members, Kikwit, Democratic Republic of the Congo, 1995. *The Journal of Infectious Diseases*, 179(Supplement 1):S87–S91, 02 1999.

23. Z. Du, X. Xu, Y. Wu, L. Wang, B. J. Cowling, and L. Meyers. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. *Emerging Infectious Diseases.*, 26(6):1341–1343, 06 2020.

24. The Henry J. Kaiser Family Foundation. Hospital adjusted expenses per inpatient day. https://www.kff.org/health-costs/state-indicator/expenses-per-inpatient-day/?currentTimeframe=0&sortModel=%7B%22colId   %22:%22Location %22,%22sort%22:%22asc%22%7D, 2019.

25. Robert Fourer, David M Gay, and Brian W Kernighan. *AMPL: A mathematical programming language*. AT & T Bell Laboratories Murray Hill, NJ 07974, 1987.

26. Christophe Fraser. Estimating individual and household reproduction numbers in an emerging epidemic. *PLOS ONE*, 2(8):1–12, 08 2007.

27. Marino Gatto, Enrico Bertuzzo, Lorenzo Mari, Stefano Miccoli, Luca Carraro, Renato Casagrandi, and Andrea Rinaldo. Spread and dynamics of the covid-19 epidemic in italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences*, 117(19):10484–10491, 2020.

28. Jana L. Gevertz, James M. Greene, Cynthia H. Sanchez-Tapia, and Eduardo D. Sontag. A novel COVID-19 epidemiological model with explicit susceptible and asymptomatic isolation compartments reveals unexpected consequences of timing social distancing. *J. Theoret. Biol.*, 510:110539, 25, 2021.

29. Giulia Giordano, Franco Blanchini, Raffaele Bruno, Patrizio Colaneri, Alessandro Di Filippo, Angela Di Matteo, and Marta Colaneri. Modelling the covid-19 epidemic and implementation of population-wide interventions in italy. *Nature medicine*, 26(6):855–860, 2020.

30. Christian Gollier. Pandemic economics: optimal dynamic confinement under uncertainty and learning. *The Geneva Risk and Insurance Review*, 45(2):80–93, 2020.

31. The New Jersey Herald. Nj contracted \$37m for contact tracing. was there a better way to use taxpayer dollars? https://www.njherald.com/story/news/politics/state/2020/09/22/nj-contracted-37m-for-contact-tracing-was-there-better-way-to-use-taxpayer-dollars/42678549/, 2020.

32. The Wall Street Journal. State shutdowns have taken a quarter of u.s. economy offline. https://www.wsj.com/articles/state-coronavirus-shutdowns-have-taken-29-of-u-s-economy-offline-11586079001, 2020.

33. V. Kala, K. Guo, E. Swantek, A. Tong, M.. Chyba, Y. Mileyko, C. Gray, T. Lee, and A. E. Koniges. Pandemics in hawaii: 1918 influenza and covid-19. In *The Ninth International Conference on Global Health Challenges GLOBAL HEALTH 2020*. IARIA, 2020.

34. Alexander Keimer and Lukas Pflug. Modeling infectious diseases using integro-differential equations: Optimal control strategies for policy decisions and applications in covid-19. Tech. rep., Friedrich-Alexander-Universität Erlangen-Nrnberg, 2020.

35. William Ogilvy Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity. *Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character*, 138(834):55–83, 1932.

36. Daewa Kim and Annalisa Quaini. Coupling kinetic theory approaches for pedestrian dynamics and disease contagion in a confined environment. *Mathematical Models and Methods in Applied Sciences*, 30(10):1893–1915, 2020.

37. E. B. Lee and L. Markus. *Foundations of optimal control theory*. John Wiley & Sons, Inc., New York-London-Sydney, 1967.

38. Qianying Lin, Shi Zhao, Daozhou Gao, Yijun Lou, Shu Yang, Salihu S Musa, Maggie H Wang, Yongli Cai, Weiming Wang, Lin Yang, et al. A conceptual model for the coronavirus disease 2019 (covid-19) outbreak in wuhan, china with individual reaction and governmental action. *International journal of infectious diseases*, 93:211–216, 2020.

39. C Jessica E Metcalf, Dylan H Morris, and Sang Woo Park. Mathematical models to guide pandemic response. *Science*, 369(6502):368–369, 2020.

40. Kelly R. Moran, Geoffrey Fairchild, Nicholas Generous, Kyle Hickmann, Dave Osthus, Reid Priedhorsky, James Hyman, and Sara Y. Del Valle. Epidemic Forecasting is Messier Than Weather Forecasting: The Role of Human Behavior and Internet Data Streams in Epidemic Forecast. *The Journal of Infectious Diseases*, 214(suppl. 4):S404–S408, 11 2016.

41. Philip D. Murphy. Letter to the President Donald J. Trump. http://d31hzlhk6di2h5.cloudfront.net/20200317/3c/e6/ea/5b/71a343 b469cf7732d3a12e0e/President_Trump_Ltr_re_COVID19_3.17.20.pdf, March 17th 2020.

42. The New Jersey Department of Health. Nj covid-19 dashboard. https://www.nj.gov/health/cd/topics/covid2019_dashboard.shtml, 2020.

43. Federal Reserve Bank of St. Louis. Bullard discusses impact of economic shutdown during a u.s. cham-

ber of commerce event. https://www.stlouisfed.org/from-the-president/speeches-and-presentations/2020/impact-economic-shutdown-us-chamber-of-commerce, 2020.

44. T. Alex Perkins and Guido España. Optimal control of the COVID-19 pandemic with non-pharmaceutical interventions. *Bull. Math. Biol.*, 82(9):Paper No. 118, 24, 2020.

45. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes.* Interscience Publishers John Wiley & Sons, Inc. New York-London, 1962. Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt.

46. Umberto   Rosini.   March   up-date: Data provided by the ministry of health and acquired by the regions in the operating board. *github:https://github.com/pcm-dpc/COVID-19/blob/master/schede-riepilogative/regioni/dpc-covid19-ita-scheda-regioni-20200302.pdf*, 2020.

47. Nick Warren Ruktanonchai, JR Floyd, Shengjie Lai, Corrine Warren Ruktanonchai, Adam Sadilek, Pedro Rente-Lourenco, Xue Ben, Alessandra Carioli, Joshua Gwinn, JE Steele, et al. Assessing the impact of coordinated covid-19 exit strategies across europe. *Science*, 369(6510):1465–1470, 2020.

48. Fernando Saldaña, Andrei Korobeinikov, and Ignacio Barradas. Optimal control against the human papillomavirus: protection versus eradication of the infection. *Abstr. Appl. Anal.*, pages Art. ID 4567825, 13, 2019.

49. Cristiana J. Silva, Carla Cruz, Delfim F. M. Torres, Alberto P. Muñuzuri, Alejandro Carballosa, Iván Area, Juan J. Nieto, Rui Fonseca-Pinto, Rui Passadouro, Estevão Soares dos Santos, Wilson Abreu, and Jorge Mira. Optimal control of the covid-19 pandemic: controlled sanitary deconfinement in portugal. *Nature Scientific Reports*, 11:3451, 2021.

50. RN Thompson, JE Stockwin, Rolina D van Gaalen, JA Polonsky, ZN Kamvar, PA Demarsh, Elisabeth Dahlqwist, Siyang Li, Eve Miguel, Thibaut Jombart, et al. Epiestim app. https://shiny.dide.imperial.ac.uk/epiestim, 2019.

51. RN Thompson, JE Stockwin, Rolina D van Gaalen, JA Polonsky, ZN Kamvar, PA Demarsh, Elisabeth Dahlqwist, Siyang Li, Eve Miguel, Thibaut Jombart, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics*, 29:100356, 2019.

52. The   New   York   Times.   coronavirus-test-cost-varies-widely.   https://www.nytimes.com/2020/06/16/upshot/coronavirus-test-cost-varies-widely.html, 2020.

53. E. Trélat. Optimal control and applications to aerospace: some results and challenges. *J. Optim. Theory Appl.*, 154(3):713–758, 2012.

54. Emmanuel Trélat. *Contrôle optimal.* Mathématiques Concrètes. [Concrete Mathematics]. Vuibert, Paris, 2005. Théorie & applications. [Theory and applications].

55. Alessandro Vespignani, Huaiyu Tian, Christopher Dye, James O Lloyd-Smith, Rosalind M Eggo, Munik Shrestha, Samuel V Scarpino, Bernardo Gutierrez, Moritz UG Kraemer, Joseph Wu, et al. Modelling covid-19. *Nature Reviews Physics*, 2(6):279–281, 2020.

56. Andreas Wächter and Lorenz T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.*, 106(1, Ser. A):25–57, 2006.

57. Joseph T Wu, Kathy Leung, and Gabriel M Leung. Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study. *The Lancet*, 395(10225):689–697, 2020.

58. Zunyou Wu and Jennifer M McGoogan. Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: summary of a re-

28   *S. McQuade, R. Weightman, N. J. Merrill,A. Yadav, E. Trélat,S. R. Allred, B. Piccoli*

port of 72 314 cases from the chinese center for disease control and prevention. *Jama*, 323(13):1239–1242, 2020.

59. Juanjuan Zhang, Maria Litvinova, Yuxia Liang, Yan Wang, Wei Wang, Shanlu Zhao, Qianhui Wu, Stefano Merler, Cécile Viboud, Alessandro Vespignani, et al. Changes in contact patterns shape the dynamics of the covid-19 outbreak in china. *Science*, 368(6498):1481–1486, 2020.

60. Juanjuan Zhang, Maria Litvinova, Wei Wang, Yan Wang, Xiaowei Deng, Xinghui Chen, Mei Li, Wen Zheng, Lan Yi, Xinhua Chen, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside hubei province, china: a descriptive and modelling study. *The Lancet Infectious Diseases*, 20(7):793–802, 2020.