

RESEARCH

Open Access



# Towards Dynamic lockdown strategies controlling pandemic spread under healthcare resource budget

Satyaki Roy<sup>1\*</sup> , Ronojoy Dutta<sup>2</sup> and Preetam Ghosh<sup>3</sup>

\*Correspondence:

satyakir@unc.edu

<sup>1</sup> University of North Carolina, Chapel Hill, USA

Full list of author information is available at the end of the article

## Abstract

COVID-19 is one of the deadliest pandemics in modern human history that has killed nearly a million people and rapidly inundated the healthcare resources around the world. Current lockdown measures to curb infection spread are threatening to bring the world economy to a halt, necessitating dynamic lockdown policies that incorporate the healthcare resource budget of people in a zone. We conceive a dynamic pandemic lockdown strategy that employs reinforcement learning to modulate the zone mobility, while restricting the COVID-19 hospitalizations within its healthcare resource budget. We employ queueing theory to model the inflow and outflow of patients and validate the approach through extensive simulation on real demographic and epidemiological data from the boroughs of New York City. Our experiments demonstrate that this approach can not only adapt to the varying trends in contagion in a region by regulating its own lockdown level, but also manages the overheads associated with time-varying dynamic lockdown policies.

## Introduction

COVID-19 is the latest addition to the long list of pandemics that scarred human history during the last millennium (Coronavirus 2020). COVID-19 has followed a similar course like the plague, flu and Ebola and claimed nearly 1.34 million lives globally as of November 2020, while its severity continues to burgeon in the US, UK, Brazil and parts of Asia (Coronavirus 2020; Mortality 2020). In the absence of any credible vaccination treatment (Adhikari et al. 2020), social distancing and lockdown measures emerged as the modus operandi to negate the surge in infection numbers. However, the projected slump in the world economy has compelled the policymakers in the developing nations to consider easing the mobility restrictions.

The primary concern of the health officials regarding this lockdown relaxation is the increased social interaction leading to a possible spike in infected counts which the healthcare facilities of even the most developed nations may be ill-equipped to combat (Coronavirus 2020). Such fears are a throwback to April 2020 when the national media reported a shortfall in intensive care unit beds and rising fatalities in the wake of the projected wave of COVID-19 cases (Coronavirus 2020). Recently, as many as 14 US

states have been compelled to reconsider their decision to ease lockdown restrictions, as the infection numbers soared and threatened to overwhelm the available healthcare resource (Weeks 2020). The present state of affairs necessitates the design of dynamic lockdown policies that incorporate the economic and epidemiological ramifications of future pandemics and, more importantly, the healthcare resource budget of a region.

### **Related works**

#### ***Susceptible-exposed-infected-death (SEIRD) epidemic model***

SEIRD (see “Data” section) has been used to model the effects of immunity, demography as well as social distancing on the spread of COVID-19. Gharakhanlou applied SEIRD to create an agent-based simulation to show the effects of social contact and propose potential mitigation measures to contain the spread of COVID-19 in Urmia city, Iran (Gharakhanlou and Hooshangi 2020). Ghanam et al. present a bayesian method to estimate the parameters for the SEIRD model and quantify the impact of government intervention measures on infection spread (Ghanam et al. 2020). Lattanzio et al. studied the interrelationship of lockdown and mobility in Lombardy and London as well as the ill-effects of flouting social distancing regulations (Lattanzio and Palumbo 2020). Keeping in mind, the debate over whether the recovered individual can be reinfected, Malkov et al. utilized SEIRD to study the effects of mitigation measures on reinfection and no-reinfection scenarios (Malkov 2019). Piccolomini et al. adapted the SEIRD with time-varying transmission rates to model restrictions imposed by the government to combat COVID-19 (Loli Piccolomini and Zama 2020; Piccolomini and Zama 2020).

#### ***Machine learning approaches***

The lack of prior knowledge on COVID-19 leaves the policymakers ill-equipped to design mitigation strategies. Epidemiologists, health experts and computer scientists have joined forces to identify the socioeconomic factors and their implications on contagion as well as economic downturn (Adhikari et al. 2020)—this includes using machine learning (ML) to build prediction models on epidemiological and clinical data. Given existing clinical data, prediction models (Wynants 2020) and therapeutic approaches can help identify vulnerable groups (Alimadadi et al. 2020; Randhawa et al. 2020). Epidemiologists are trying to identify spread dynamics of COVID-19. Holmdahl and Buckee (2020) analyzed the pros and cons of forecasting models that make predictions through curve fitting or mechanistic models, while supervised and unsupervised ML is helping trace the trends in infection dynamics (Wang et al. 2020). Khan et al. used regression tree analysis, cluster analysis and principal component analysis on Worldometer infection count data to gauge the variability and effect of testing in the prediction of confirmed cases (Khan et al. 2020). Also, Roy et al. performed regression analysis to identify pre-lockdown factors that affect the post-lockdown pandemic numbers (Roy and Ghosh 2020).

#### **Issues in vaccine production and supply**

There is a mistrust brewing over the efficacy of the vaccines. The public at large is sceptical about the “rush” to put out the vaccine before adequate bouts of clinical testing (Mistrust 2020). Many believe that the undue optimism in releasing the vaccine can have

adverse health ramifications. Moreover, governments continue to plan to expedite this process by parallelizing the steps of research and trials as well as industrial-scale manufacture of vaccines (Testimony 2020). Third, vaccines must be affordable and accessible by all irrespective of the social or economic strata they belong to. This poses a policy-making problem to guarantee the equity of resource allocation. There are several aspects to vaccine allocation from the standpoint of policymaking per se. There is an economic angle associated with the distribution of vaccines. Given that the vaccines will be stored at warehouses, it becomes crucial to minimize the economic overhead of transporting vaccines to the affected zones. Finally, there are political and market forces that may obtrude the vaccines from reaching the worst-hit states, making fairness a key factor for regulations.

Other epidemiological factors, such as population density, number of susceptible individuals and the infected ratio, play a role in the dynamics of infection spread (Farman et al. 2018). Population density governs the “contact with susceptible individuals” resulting in contagion (Tarwater and Martin 2001; Rocklöv and Sjödin 2020). Also, the spread and sustenance of an epidemic is contingent on whether there is an adequate number of susceptible hosts in the total population (Principles 2020). Finally, since the spread of infection depends on contact between a susceptible host and infected individual (Korolev 2020), high percentage of infected people in the total population (i.e., infected ratio) leads to contacts contributing to the spread of the outbreak. Thus, the policymaker also needs to factor in some or all of these aspects in determining the vaccine distribution policies across regions. Most importantly, in absence of adequate clinical trials, the allocation strategies must take into consideration the innate uncertainty in the extent of immunity such vaccination can achieve (Lurie et al. 2020).

*Contributions.* In this work, we conceive a dynamic pandemic lockdown strategy that factors in public health infrastructure of a geographical region. The proposed approach built upon reinforcement learning (RL) allows agents to take decisions to maximize reward, while adapting to a complex and uncertain environment (Tuyls and Weiss 2012; Pecka and Svoboda 2014). We create an agent-based simulation environment running the ordinary differential equation-based SEIRD epidemic model (Hethcote 2000) (discussed in “Scenario” section). A geographic region, modeled as an agent, is classified into zones, and each zone has a healthcare budget commensurate with its gross domestic product (GDP). Each agent (or zone) periodically invokes the RL model to select a discrete lockdown level based on two different models (1) *average velocity* of the individuals in that zone and (2) *contact index*: a measure of the average contact of individuals within a borough independent of the demographic factors like population density, both of which affect the rate of human contact. Both these models showcase the generalizability of our proposed framework which can readily be extended to other factors that may affect the contact rates of infected individuals with the susceptible population. We employ the queueing model to ensure that the number of hospitalizations is constrained by its available healthcare resource budget.

We design a simulation environment using the Python Simpy library (Matloff 2008) that operates on the real demographic and epidemiological data from the 5 boroughs of New York City, namely Manhattan, Bronx, Brooklyn, Queens and Staten Island. We introduce a realistic model that employs real mobility traces and epidemic status of

individuals in each zone to determine inter-zone mobility. Moreover, we formulate a KL divergence minimization problem to learn the epidemiological parameters that yield the daily infected curve. The healthcare budget of each borough comprises a quanta of hospital facilities, each modeled, as per single server queueing theory, as a server with an inflow and outflow of patients. Our experiments show that despite the heterogeneity in infection dynamics, each borough effectively modulates its mobility to curb infection spread and consequent hospitalization. We also demonstrate how several simulation parameters can help regulate the overall cost associated with the time-varying dynamic lockdown strategy.

## Materials and methods

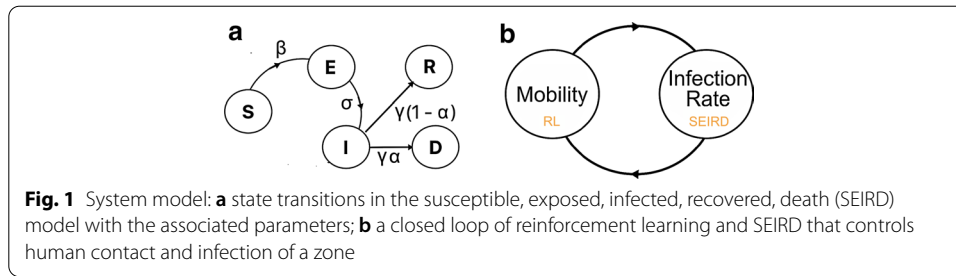
Let us discuss the data used and the components of the proposed approach.

### Data

- We consider 5 towns, called *boroughs*, in New York City (NYC), viz. Bronx, Brooklyn, Manhattan, Queens and Staten Island. We obtain the borough data, such as Gross Domestic Product (GDP), population density, etc., from Wikipedia (Neighborhoods 2020). COVID infection and deaths are taken from The City (Coronavirus 2020) based on records of Department of Health and Mental Hygiene. We use NYC Health records (Nyc health 2020) that show daily infected from March-August 2020 from New York Department of Health.
- We source the mobility data of NYC traffic from NYCOpenData (Nycopendata 2020)—a repository for fields ranging from city government, education, environment, health to public safety, recreation, social services and transportation. The stated data (spanning a period from 2014 to 2019), collected by the Department of Transportation of New York Metropolitan Transportation Council (NYMTC), has the following fields: *ID*, *road name*, *source and destination intersecting street name*, *compass direction*, *date* and *time*. We calculate the transition matrix (see “Inter-zone mobility model” section) that captures the probability of travelling within and across boroughs.

### SEIRD epidemic model

We adapt the susceptible-exposed-infected-recovered-death (SEIRD) model (see Fig. 1a) (Hethcote 2000). The *susceptible* (*S*) class comprises individuals who are not exposed to the infection. Once exposed to infected individuals, they may transfer to the *exposed* (*E*) category, and this transition is controlled by a value  $\beta$  (that is not necessarily a probability). The *E* class are asymptomatic or untested individuals, who transition to the (tested) *infected* (*I*) with probability  $\sigma$ . The individuals in *I* transition to another state with a probability  $\gamma$ , either *recovered* (*R*) or *dying* (*D*) with probabilities  $1 - \alpha$  and  $\alpha$ .



**Ordinary differential equations**

Ordinary Differential Equations (ODE) model estimates the state transitions  $S \rightarrow E \rightarrow I \rightarrow R \rightarrow D$  by solving the following set of ordinary differential equations (Korolev 2020). We implement the ODE model on Python Odeint library (Ahnert and Mulansky 2011).

$$\frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N} \tag{1}$$

$$\frac{dE(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \sigma E(t) \tag{2}$$

$$\frac{dI(t)}{dt} = \sigma E(t) - \gamma I(t) \tag{3}$$

$$\frac{dR(t)}{dt} = (1 - \alpha)\gamma I(t) \tag{4}$$

$$\frac{dD(t)}{dt} = \alpha\gamma I(t) \tag{5}$$

Here  $\beta$  is the contact rate,  $N$  is the total population,  $\sigma$  is the incubation period,  $\gamma$  is the duration of infection and  $\alpha$  is the fatality rate. In Eq. 1,  $\beta = p \times C$ , where  $p$  is the infection probability and  $C$  is the individual contact rate. Since, contact rate  $C$  can vary for different zones, we represent it as a product of a constant terms (density  $\rho$ ) and a variable term (*contact index k*), i.e.  $C = \rho \times k$ . We assume that the social contact and interaction among individuals at a zone follows the collision among ideal gas molecules in a homogeneous mixture. We can also calculate  $C = \sqrt{2}\pi d^2 \rho v$ , where  $d$  is the collision diameter,  $\pi d^2$  is the cross-sectional area,  $\rho$  is the population density and  $v$  is the mean velocity (Hu et al. 2013). The COVID-19 specific epidemic parameters for ODE used in our experiments are discussed at the beginning of “Results” section.

**Scenario**

We create an agent-based simulation environment using the Python Simpy library (Matloff 2008), where each zone (termed *borough*) is an agent with a predefined initial population of susceptible, exposed, infected, recovered and dead individuals. Each zone also

has an initial lockdown level  $l_i$  (where  $1 \leq i \leq \chi$ ); the higher the lockdown level, the lower is the mobility (measured in terms of the average velocity  $v$  of individuals) in a zone. Specifically, the assumption is that the average velocity of individuals at a particular zone is proportional to their distance covered, which in turn correlates with higher overall social contact, mixing and contagion. The new infected count  $I$  is determined by periodically (i.e., after interval  $\eta$ ) invoking the ODE SEIRD model (discussed in “[Ordinary differential equations](#)” section). Since higher mobility leads to greater contact and contagion, each zone invokes the reinforcement Q-learning module (see “[Reward function](#)” section) to learn and determine an updated  $v$ . The objective is to maximize mobility, while ensuring that the number of hospitalizations is within the healthcare resource budget of the zone. We assume that the healthcare resource budget of borough  $b$  (is commensurate with the number of hospital beds and) is measured in terms of its overall GDP using the equation below:

$$|H_{b_i}| = \frac{GDP_{b_i} - \mu(GDP)}{\sum_{b_j \in B} GDP_{b_j}} \times bG \tag{6}$$

Here  $GDP_{b_i}$  is the GDP of borough  $b_i$  and  $bG$  is the baseline hospital bed count. This formulation ensures that the number of beds allotted to each borough is proportional with its GDP. Finally, we measure the number of hospitalizations as a fraction, say  $k$ , of the newly infected population. Finally, Fig. 1b shows that the proposed system is a closed loop of reinforcement learning and SEIRD models controlling the mobility and infection of a zone.

**Inter-zone mobility model**

Given a region with a set of geographical sub-regions (or zones)  $B$ , the frequency matrix  $F \in \mathbf{M}_{|B| \times |B|}(\mathbb{R})$  is created from the human mobility traces, where  $f_{i,j} \in F$  denotes the number of trips made from zone  $b_j \in B$  to  $b_i \in B$ . We generate a transition matrix  $A \in \mathbf{M}_{|B| \times |B|}$  performing column normalization of  $F$ . Each element of the matrix  $a_{i,j} \in A$  is the probability of making a trip from  $b_j$  to  $b_i$ . The frequency (and transition matrix) captures the overall mobility trends within and across zones of any given region. A pre-assigned number of people migrate from one region to another based on the following inter-zone mobility procedure employing the transition matrix  $A$ .

---

**Algorithm 1** Inter-zone mobility model

---

**Input.** Transition matrix  $A$ , Migration rate  $\zeta$ ,  $b \in B$ ,  $N_b = \sum_{state} n_b(state)$   
 $h = \lceil N_b \times \zeta \rceil$   
 $i = 0$   
**while**  $i < h$  **do**  
     $p(dest) \sim Multinomial(A_{*,b})$   
     $p(st) \sim Multinomial(\frac{n_b(st)}{N_b - n_b(dead)}) \quad \forall st \neq dead$   
     $i++$   
     $n_{dest}(st) = n_{dest}(st) + 1$   
     $n_b(st) = n_b(st) - 1$   
**end**

---

**Description.** Procedure 1 is invoked periodically by each borough  $b$  of population  $N_b$ , where  $h = \lceil \zeta * N_b \rceil$  people move, where migration rate  $\zeta$  ranges between 0 and 1. For a moving person, the destination borough  $dest \in B$  is chosen based on a multinomial distribution on the  $b$ -th column of transition matrix  $A$ , i.e.,  $A_{*,b}$ . Similarly, state of the alive person  $st$  is proportional to the fraction of people within that state, i.e.  $\frac{n_b(st)}{N_b - n_b(dead)}$ . Following this, the algorithm decrements the number of people in  $b$  with state  $st$ ,  $n_b(st)$ , and increments  $n_{dest}(st)$  to reflect migration from borough  $b$  to  $dest$ .

**Minimization of Kullback–Leibler divergence**

The Kullback–Leibler (KL) divergence measures the difference between one probability distribution from another reference probability distribution (Kullback and Leibler 1951). Given two probability distributions  $P$  and  $Q$ , it is measured as:

$$D_{KL}(P||Q) = \sum_{i=1}^n P(x_i) \times (\log P(x_i) - \log Q(x_i)) \tag{7}$$

Given  $N, \sigma, \alpha, p, \rho$  and  $\gamma$ , we learn the two parameters  $\nu$  and  $E(t = 0)$  to generate a reference curve that fits actual infected curve  $I_f$  having the least KL divergence from the actual (or reference) daily infected curve for a given region  $I_a$ . The fitting optimization problem is formulated as:

$$\min_{\nu, E(t=0)} D_{KL}(I_a||I_f(N, \sigma, \alpha, p, \rho, \gamma)) \tag{8}$$

**Modeling hospitalization queue**

A borough  $b \in B$  has a healthcare capacity  $H_b$ . We model a hospitalization facility  $h \in H_b$  as a single server with patient arrival rate  $a$  and treatment (or service) rate  $r$ . As per single-server queueing model, a waiting line or a *queue* is formed when a server has more than 1 person in the system. Let  $p^h(i)$  be the probability that there are  $i$  people in the system of server  $h$ ; then probability of a queue forming in front of facility  $h$  is given by  $p_{queue}^h = 1 - p_0^h - p_1^h$ . We assume that the *arrival-to-service* ratio  $\delta = \min(1, \frac{a}{r})$ , making  $p_0^h = 1 - \delta$  and  $p_1^h = p_0^h \times \delta$  as per single-server queueing model. At any given time  $t$ , we estimate the mean probability of queue as  $p_{queue} = \frac{\sum_h p_{queue}^h}{|H_b|}$ . Once hospitalized, a patient may transitions to dead (D) or recovered (R) categories with *hospital fatality rates*  $\alpha_h$  and  $1 - \alpha_h$ , respectively.

**Reinforcement Q-learning model (RL)**

Q-learning (Watkins and Dayan 1992), invoked every  $W$  hours, allows agents (i.e., zones) to take decisions to maximize reward while adapting to an uncertain environment. Given a set of possible actions  $A = \{a_1, a_2, \dots\}$ , each agent maintains a Q-table that records the past rewards the agent has received for an action. Thus, the Q-table,  $Q$ , is a matrix  $\mathbb{R}^{|A| \times |A|}$ . We modulate the *exploration vs exploitation* factor allowing the RL model to pick a random action with a probability  $ep$ . Note that  $ep$  undergoes a decay by a factor  $dc$  ( $e, d \in [0, 1]$ ) after each run of the model. Next, we discuss the action space and reward.

**Action space**

The rows and columns of  $Q$  represent current action and next action, respectively. In addition to the  $\chi$  lockdown levels  $l_1, l_2, \dots, l_\chi$  (explained in “Scenario” section), the probability of queue  $p_{queue}$  is discretized into  $\omega$  levels  $w_1, w_2, \dots, w_\omega$  ranging from low to high. This makes the search space  $A$  a set of combinations of lockdown levels and  $p_{queue}$  levels, i.e.,  $A = \{(l_1, w_1), (l_1, w_2), \dots, (l_\chi, w_\omega)\}$ . It is noteworthy that the RL model of each zone can only control the lockdown level  $l$  (i.e., velocity). Consequently, a zone can transition from current action  $a = (l_y, w_j)$  to another state  $a = (l_z, w_j)$ , where  $1 \leq y, z \leq \chi$ , and the new  $p_{queue}$  will be determined by the number of new infections spawned by the change in lockdown level.

**Reward function**

Recall from our discussion in “Ordinary differential equations” section, the rate of human contact is controlled by the contact rate  $C = \sqrt{2\pi} d^2 \rho v = \rho k$ . Since  $C$  is a function of both average velocity  $v$  and contact index  $k$ , we devise the two reward function that incentivizes conflicting goals of (1) high human contact based on either  $C$  or  $k$  and (2) low hospital occupancy (i.e., low  $p_{queue}$ ). It is calculated as:

$$R = \frac{v}{\max(\mathbf{V})} \times e^{-p_{queue}} \text{ or } \frac{k}{\max(\mathbf{K})} \times e^{-p_{queue}} \tag{9}$$

Here, the first term is the permitted velocity  $v \in V$  (or contact index  $k \in K$ ) of a borough normalized by the maximum velocity  $\max(\mathbf{V})$  (or contact index  $\max(\mathbf{K})$ ) and the second term penalizes high hospital occupancy of a zone. Later in Fig. 3b, we show that the  $e^{-p_{queue}}$  drops with the increase in  $p_{queue}$ .

**Pearson correlation coefficient**

It captures the linear relationship between two variables. The values of 1 and  $-1$  represent high positive and negative correlations, while 0 represents uncorrelated variables. Given two distributions  $X$  and  $Y$ , it is  $\frac{cov(X, Y)}{\sigma_X \times \sigma_Y}$ , where  $\sigma_X$  and  $cov(X, Y)$  are standard deviation and covariance (Benesty et al. 2009).

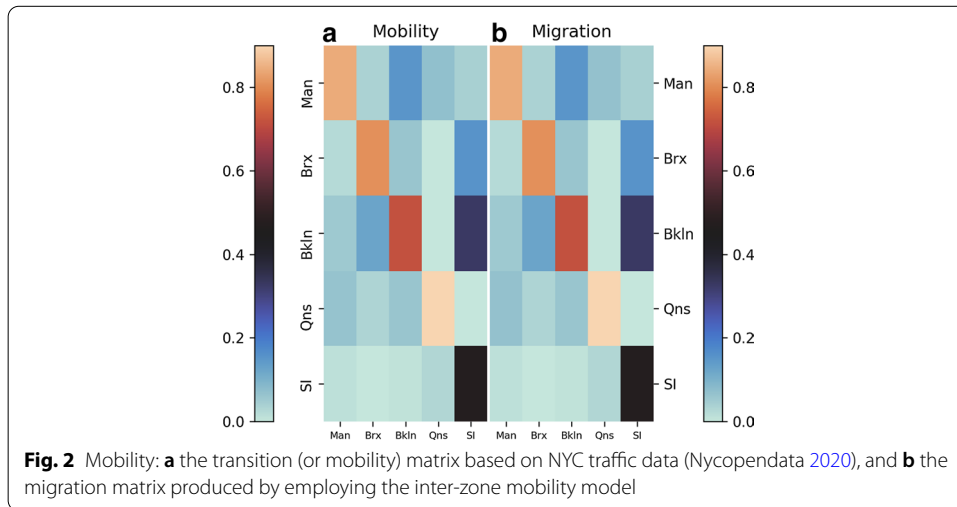
**Overhead of lockdown**

It is imperative to recognize that imposing time-varying lockdown is cost-intensive, as the news of the updated lockdown level must be disseminated among the public through



**Table 1** Default parameter values

Parameter	Notation	Value
Number of iterations	–	100
Simulation duration	$T$	180 days
Number of boroughs	$ \mathcal{B} $	5
Number of lockdown levels	$\chi$	6
Interval for invoking ODE	$\eta$	12 hours
SEIRD parameters(Korolev 2020)	$\sigma, \gamma, \alpha$	0.25, 0.1, 0.05
Interval and initial number of new infection	$\zeta, k$	30 days, 200
Migration rate	$\zeta$	0.01
Treatment rate	$r$	0.0029
Hospital fatality rate	$\alpha_h$	0.2
SEIRD infection probability	$p$	0.01
Collision diameter	$d$	1 m
RL probability of random action	$ep$	0.75
RL decay factor	$dc$	0.99
RL transition window	$W$	5.5 days
Number of probability of queue levels	$\omega$	3
Levels in probability of queue	$w_1, w_2, w_3$	0 – 0.33, 0.33 – 0.66, 0.66 – 1.0



**Fig. 2** Mobility: **a** the transition (or mobility) matrix based on NYC traffic data (Nycopendata 2020), and **b** the migration matrix produced by employing the inter-zone mobility model

electronic and print media. Moreover, since lockdown affects every aspect of human life, temporal lockdowns can have wide-ranging social and economic implications. We assume that the cost of lockdown is directly proportional to the number of transitions in lockdown levels  $l_i$ . This cost can be controlled by regulating two parameters: (1) *lockdown window*,  $\tau$  – duration (in hours) before the RL model is invoked and  $l_i$  is re-evaluated and (2) *lockdown threshold*,  $\alpha$  – real value between 0 and 1, such that RL model is invoked by a borough only if  $|p_{queue}(t) - p_{queue}(t - 1)| > \alpha$ .

## Results

The simulation environment is implemented in Python. The demographic, epidemiological and human mobility data sources for NYC boroughs are discussed in “Data” section. We define the following 4 lockdown levels in terms of contact index:  $l_1, l_2, l_3, l_4$  with contact index  $k = 4.4 \times 10^{-6}, 1.7 \times 10^{-5}, 3.1 \times 10^{-5}$  and  $4.4 \times 10^{-5}$ , respectively. The results section has been organized into three broad headings: (1) inter-zone mobility model, (2) relationship among contact index (or velocity), healthcare capacity and reward and (3) dynamics among contact index, infection and hospital capacity in boroughs. All the parameters (and their default values) are summarized in Table 1.

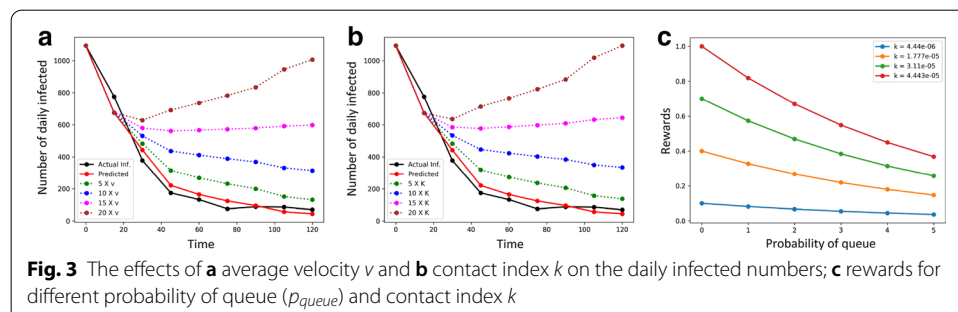
### Inter-zone mobility model

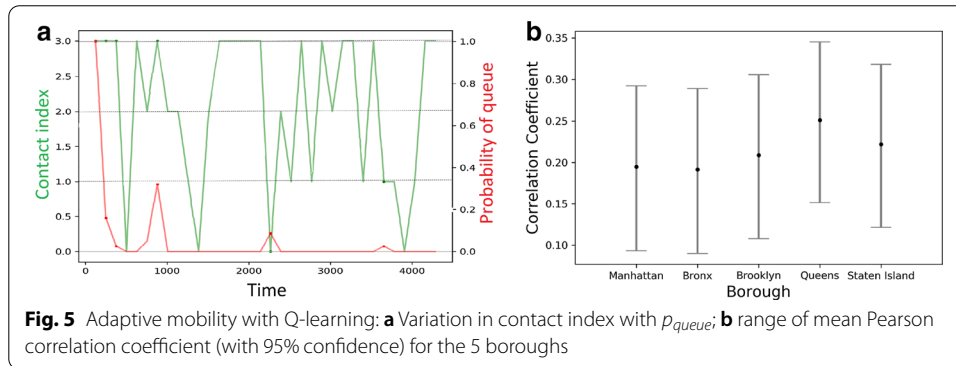
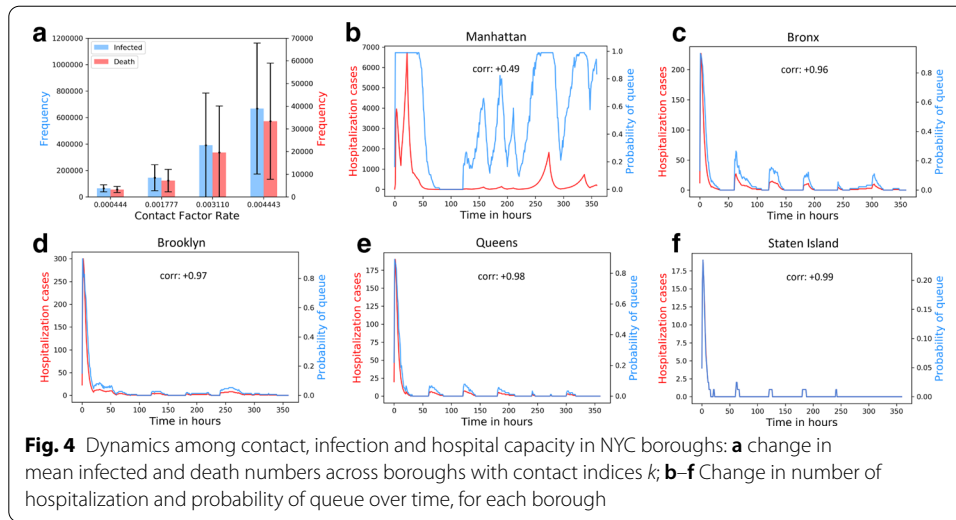
The migration of people across boroughs are dictated by the inter-zone mobility model (refer “Inter-zone mobility model” section) to mimic the transition matrix. We generate this migration matrix where each element  $(i, j)$  is the number of trips made from source borough  $j$  to  $i$  normalized by the column sum. Figure 2 shows the migration matrix (right) where each element has the same color (i.e., mobility probabilities) as the corresponding element from the transition matrix.

### Relationship among contact index (or velocity), healthcare capacity and reward

We first apply KL divergence minimization (explained in “Minimization of Kullback–Leibler divergence” section) to learn three SEIRD parameters (average velocity  $v$ , fraction of initial exposed population  $E(t = 0)$  and infection duration  $\gamma$ ) for each borough based on COVID-19 daily confirmed cases (discussed in “Data” section). (The values of all other parameters used in the ODE model (Eq. 1 - 5) have been taken from Korolev (2020).)

Figure 3a shows the fit line (shown in solid red line) obtained by solving the optimization on the *post-lockdown* daily infected numbers (solid black line); the corresponding parameters are  $v = 0.1$  km/h and  $E(t = 0) = 1.82 \times 10^{-4} \times$  total population ( $N$ ). In Fig. 3b, we apply the same fitting to obtain an equivalent contact index  $k = 4.63 \times 10^{-6}$ . For either case, we show the surge in the projected daily infected numbers for lower lockdown levels (i.e., higher velocities and corresponding contact index levels) shown in different colors. This shows that the proposed approach is fairly generalizable, as the roles of  $k$  and  $v$  are interchangeable in the RL reward function.



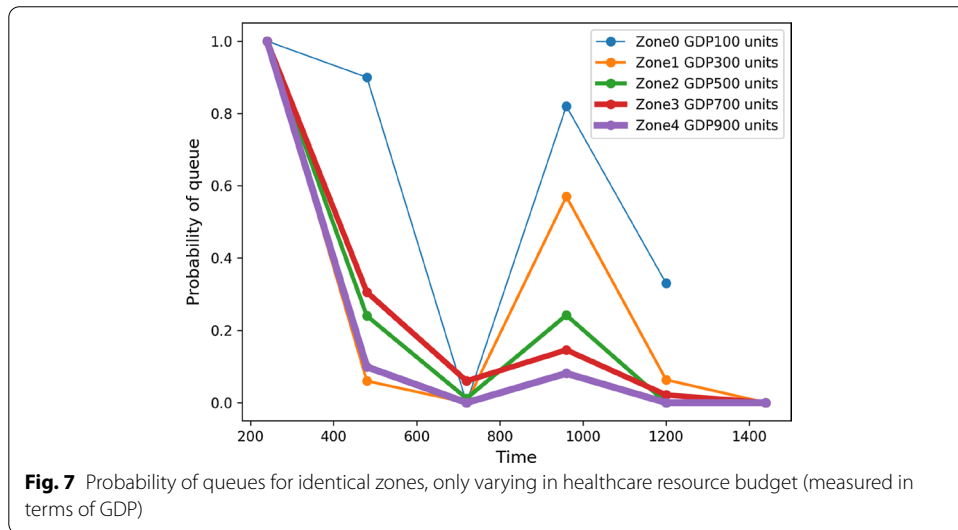
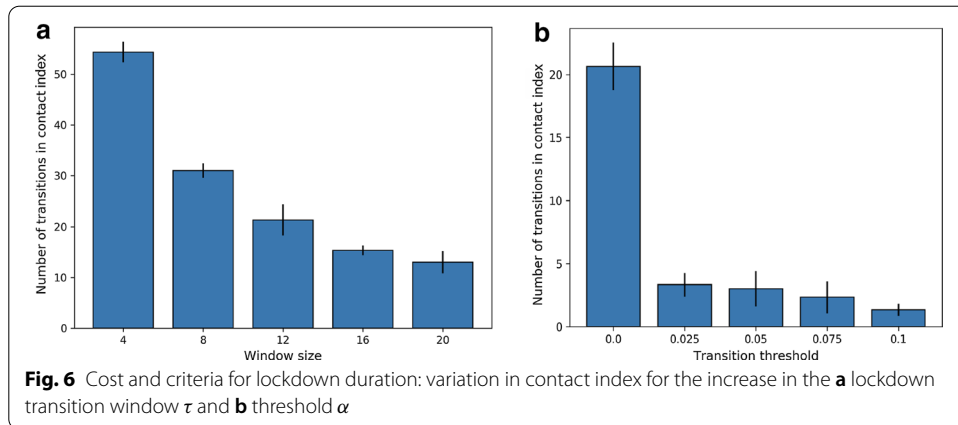


**Reward function**

We analyze the variation in the reinforcement Q-learning reward (Eq. 9) for different input parameter values of hospital capacity (measured in terms of probability of a hospital queue  $p_{queue}$ ) and the lockdown levels (estimated by velocity  $v$  and contact index  $k$ ) in each borough. Figure 3c shows that the reward function balances the trade-off between  $k$  and  $p_{queue}$ . Evidently, the reward is low if any of the two conditions hold true:  $k$  is low or  $p_{queue}$  is high.

**Dynamics among velocity, infection and hospital capacity in boroughs**

We study how high contact (or  $k$ ) affects infected number, which in turn affects hospital capacity. Figure 4a shows that the mean total infected and death numbers across boroughs increase with velocity. Both the mean infected and death numbers exhibit a fairly high deviation from the mean, suggesting that there is a high variation in dynamics of infection spread across NYC boroughs. Figure 4b–f shows the time-varying number of hospitalizations for each borough is understandably correlated with the probability of queue  $p_{queue}$ .



### Adaptive mobility with reinforcement Q-learning

We study how boroughs invoke the RL model (discussed in “Reward function” section) to adapt its contact index (i.e.,  $k$ ) with changing hospital queue  $p_{queue}$ . Given 4 equispaced contact index levels  $k$  ranging between  $4.4 \times 10^{-6}$  to  $4.4 \times 10^{-5}$ . Figure 5a shows the overall change in  $k$  with the dynamics of  $p_{queue}$ . The phase changes in  $p_{queue}$  and the corresponding phase changes in  $k$  are denoted by red and green curves, respectively. Evidently, the RL model is able to adapt mobility to keep the infection counts (and the associated hospitalizations) under check (with mean absolute difference between the levels of  $k$  ( $l_k$ ) and  $p_{queue}$  on a scale of 0 to 1  $\sum_t |0.33 \times l_k(t) - p_{queue}(t)| = 0.63$ ). Next, we record the mean correlation between 4 levels of  $k$  and  $-p_{queue}$  across 100 iterations. Figure 5b shows that the range of mean correlation coefficient (with 95% confidence) varies from 0.1 to 0.35, suggesting that mobility restrictions are indeed higher when the number of hospitalization rises.

### **Cost and criteria for lockdown duration**

We discuss in “[Overhead of lockdown](#)” section that the duration of lockdown can be controlled by regulating the lockdown transition window  $\tau$  and threshold  $\alpha$ . In Fig. 6a we show how the cost of lockdown (measured in terms of the number of transitions) in contact indices ( $k$ ) decrease with  $\tau = \kappa \times W + \frac{W}{2}$  hours, where  $\kappa = 2, 4, 6, 8, 10$ . Figure 6b shows that the decrease in  $\alpha$  has a similar effect on transitions in  $k$ , increasing the overall cost of enforcing lockdown.

### **Conclusions**

We present an approach for dynamic time-varying lockdown strategy based on the healthcare budget and epidemic spread of a geographical region. This approach models each zone as an agent that applies reinforcement learning (RL) to periodically select a lockdown level that maximizes mobility, while constraining the number of hospitalizations to its healthcare resource budget. Through extensive simulation experiments on the real demographic and epidemiological data from the 5 boroughs of New York City, we demonstrate the efficacy of the approach. Each borough not only adapts to changing infection numbers by regulating its lockdown level, but also efficiently manages the overall cost associated with the time-varying dynamic lockdown strategy.

We are currently exploring how the RL model can employ epidemic model parameters, apart from mobility, that can realistically model infection spread via social contact. Second, here we assumed GDP as a measure of the healthcare resource available to a zone. To understand, how GDP affects the probability of hospital queue, we carry out a preliminary analysis where we create 5 zones that are identical in all respect, except have GDP 100, 200,  $\dots$ , 500 units, respectively. Figure 7 shows that regions with high GDP have a lower overall probability of queue. Taking a cue from this result, we shall devise collaborative strategies where neighboring zones with disparate healthcare budgets can pool their resources to avoid patient waiting times. This will require us to include additional considerations such as the distance between two zones as well as the quality of healthcare facility available to each zone based on some standard zone-level health index measures. Finally, we considered healthcare budget of a zone exclusively with respect to COVID-19 patients. This assumption may not always hold over long periods of time, making it imperative to include the effect of patients with other conditions into the RL model who may compete for the shared hospital resources.

### **Abbreviations**

NYC: New York City; RL: Reinforcement learning; GDP: Gross domestic product; SEIRD: Susceptible exposed infected recovered dead; ODE: Ordinary differential equations; KL: Kullback–Leibler.

### **Acknowledgements**

The authors would like to acknowledge the editor/reviewers for critically assessing the materials and providing suggestions that significantly improved the presentation of the paper. Furthermore, they acknowledge the Department of Computer Science, Virginia Commonwealth University for its computational resources.

### **Authors' contributions**

SR and PG conceived of the idea presented in this paper. SR developed the theory and RD performed the necessary experiments. SR and PG verified the methods and results. All authors discussed the results and contributed to the final manuscript.

### **Funding**

This work is partially supported by National Science Foundation (CBET-1802588).

**Availability of data and materials**

All relevant data (epidemiological and demographic data related to the boroughs of New York City) as well as the Python scripts are made available at [https://github.com/satunr/COVID-19/tree/master/Dynamic\\_Lockdown](https://github.com/satunr/COVID-19/tree/master/Dynamic_Lockdown)

**Competing interests**

The authors have declared that no competing interests exist.

**Author details**

<sup>1</sup> University of North Carolina, Chapel Hill, USA. <sup>2</sup> Deep Run High School, Glen Allen, VA, USA. <sup>3</sup> Virginia Commonwealth University, Richmond, USA.

Received: 24 August 2020 Accepted: 23 December 2020

Published online: 07 January 2021

**References**

- Adhikari S, Meng S, Wu Y, Mao Y, Ye R, Wang Q, Sun C, Sylvia S, Rozelle S, Raat H et al (2020) Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (covid-19) during the early outbreak period: a scoping review. *Infect Diseases Poverty* 9(1):1–12
- Ahnert K, Mulansky M (2011) Odeint–solving ordinary differential equations in c++-. In: AIP conference proceedings, American Institute of Physics, vol 1389, pp 1586–1589
- Alimadadi A, Aryal S, Manandhar I, Munroe P, Joe B, Cheng X (2020) Artificial intelligence and machine learning to fight covid-19. *Physiol Genom* 52(4):200–202
- Benesty J, Chen J, Huang Y, Cohen I (2009) Pearson correlation coefficient. In: Noise reduction in speech processing, Springer, Berlin, pp 1–4
- Coronavirus (2020) Coronavirus in new york city – tracking the spread of the pandemic. [https://projects.thecity.nyc/2020\\_03\\_covid-19-tracker/](https://projects.thecity.nyc/2020_03_covid-19-tracker/)
- Coronavirus (2020) How dangerous is lifting lockdown? <https://www.bbc.com/news/health-52878816>
- Coronavirus (2020) What have been the worst pandemics and epidemics in history? [https://en.as.com/en/2020/04/18/other\\_sports/1587167182\\_422066.html](https://en.as.com/en/2020/04/18/other_sports/1587167182_422066.html)
- Coronavirus Cases (2020) These states face biggest potential shortfalls in hospital icu beds. <https://www.cnbc.com/2020/04/06/coronavirus-cases-states-with-biggest-hospital-bed-shortfalls.html>
- Coronavirus World Map (2020) which countries have the most cases and deaths? <https://www.theguardian.com/world/2020/may/06/coronavirus-world-map-which-countries-have-the-most-cases-and-deaths>
- Farman M, Saleem M, Ahmad A, Ahmad M (2018) Analysis and numerical solution of seir epidemic model of measles with non-integer time fractional derivatives by using laplace adomian decomposition method. *Ain Shams Eng J* 9(4):3391–3397
- Ghanam R, Boone E, Abdel-Salam A (2020) Seird model for qatar covid-19 outbreak: a case study. *arXiv preprint arXiv:2005.12777*
- Gharakhanlou N, Hooshangi N (2020) Spatio-temporal simulation of the novel coronavirus (covid-19) outbreak using the agent-based modeling approach (case study: Urmia, iran). *Inf. Med. Unlocked* 20:100403
- Hethcote H (2000) The mathematics of infectious diseases. *SIAM Rev* 42(4):599–653
- Holmdahl I, Buckee C (2020) Wrong but useful-what covid-19 epidemiologic models can and cannot tell us. *N Engl J Med*
- Hu H, Nigmatulina K, Eckhoff P (2013) The scaling of contact rates with population density for the infectious disease models. *Math Biosci* 244(2):125–134
- Khan N, Naushad M, Fahad S, Faisal S, Muhammad A (2020) Covid-2019 and world economy. *COVID-2019 and World Economy*
- Korolev I (2020) Identification and estimation of the seird epidemic model for covid-19. Binghamton University. <https://doi.org/10.2139/ssrn.3569367>
- Kullback S, Leibler R (1951) On information and sufficiency. *Ann Math Stat* 22(1):79–86
- Lattanzio S, Palumbo D (2020) Lifting restrictions with changing mobility and the importance of soft containment measures: a seird model of covid-19 dynamics
- Lurie N, Saville M, Hatchett R, Halton J (2020) Developing covid-19 vaccines at pandemic speed. *N Engl J Med* 382(21):1969–1973
- Malkov E (2020) Simulation of coronavirus disease 2019 (covid-19) scenarios with possibility of reinfection. *Chaos, Solitons Fractals* 139:110296
- Matloff N (2008) Introduction to discrete-event simulation and the simpy language. Dept Comput Sci Univ Calif Davis, Davis, CA 2:1–33
- Mistrust (2020) Mistrust of a coronavirus vaccine could imperil widespread immunity. <https://www.nytimes.com/2020/07/18/health/coronavirus-anti-vaccine.html>
- Mortality Analysis (2020) Coronavirus resource center, johns hopkins university and medicine. <https://coronavirus.jhu.edu/data/mortality>
- Neighborhoods (2020) Neighborhoods in New York city. [https://en.wikipedia.org/wiki/Neighborhoods\\_in\\_New\\_York\\_City](https://en.wikipedia.org/wiki/Neighborhoods_in_New_York_City)
- Nyc Health (2020) <https://www1.nyc.gov/site/doh/covid/covid-19-data.page>
- Nycopendata (2020) <https://data.cityofnewyork.us/Transportation/Traffic-Volume-Counts-2012-2013-/p424-amsu>

- Pecka M, Svoboda T (2014) Safe exploration techniques for reinforcement learning—an overview. In: International workshop on modelling and simulation for autonomous systems, Springer, Berlin, pp 357–375
- Piccolomini EL, Zama F (2020) Monitoring italian covid-19 spread by an adaptive seird model. medRxiv
- Piccolomini EL, Zama F (2020) Preliminary analysis of covid-19 spread in italy with an adaptive seird model. arXiv:2003.09909
- Principles (2020) Principles of epidemiology in public health practice, third edition an introduction to applied epidemiology and biostatistics. <https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section11.html>
- Randhawa G, Soltysiak M, El Roz H, de Souza C, Hill K, Kari L (2020) Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: Covid-19 case study. PLoS ONE 15(4):e0232391
- Rocklöv J, Sjödin H (2020) High population densities catalyse the spread of covid-19. J Travel Med 27(3):038
- Roy S, Ghosh P (2020) Factors affecting covid-19 infected and death rates inform lockdown-related policymaking. PLoS ONE 15(10):e0241165
- Tarwater P, Martin C (2001) Effects of population density on the spread of disease. Complexity 6(6):29–36
- Testimony on Operation Warp Speed (2020) Researching, manufacturing, and distributing a safe and effective coronavirus vaccine. <https://www.nih.gov/about-nih/who-we-are/nih-director/testimony-operation-warp-speed-researching-manufacturing-distributing-safe-effective-coronavirus-vaccine>
- Tuyls K, Weiss G (2012) Multiagent learning: basics, challenges, and prospects. Ai Mag 33(3):41–41
- Wang P, Zheng X, Li J, Zhu B (2020) Prediction of epidemic trends in covid-19 with logistic model and machine learning technics. Chaos, Solitons Fractals 139:110058
- Watkins C, Dayan P (1992) Q-learning. Mach Learn 8(3–4):279–292
- Weeks after lockdowns were lifted, surging numbers of covid-19 cases prompt some u.s. states to renew restrictions. <https://ktla.com/news/coronavirus/weeks-after-lockdowns-were-lifted-surging-numbers-of-covid-19-cases-prompt-some-u-s-states-to-renew-restrictions/> (2020)
- Wynants L et al (2020) Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ 369:1

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---