A Mathematical Model to Predict the Spread of COVID-19 in a Population and Evaluate the Effectiveness of Containment Strategies

Vahid S. Bokharaie

Abstract

The available data on the spread of COVID-19 in China has been used to estimate the parameters of epidemiological models in a population stratified based on age distribution. The models can then be adapted and used for other countries or populations with different age distributions. The resulting models can be used to calculate the trajectory of the spread of the disease in various age groups in different countries. More importantly, it can be used to predict the effects of different containment strategies on the spread of COVID-19.

I. Introduction

There is an abundance of mathematical models of various degrees of sophistication to model the spread of a disease in a population. But the common problem in using these models in real-world applications is that it is very difficult to come up with meaningful values for parameters of such models. To understand the reasons for that, consider the *contact rate*, which is an important parameter in many epidemiological models. Contact rate represents the probability that a disease is transmitted between two individuals or two groups in a population. When we model our network on the level of individuals, the probability that the disease is being transmitted from one individual to another depends on their distance, their health conditions, characteristics of the disease career agent, and quite possibly on their age and genetic background. Finding that probability is not possible, and even if it was, it is not ethical. One reasonable approach to get closer to an answer is to average such interactions over time and over individuals. One solution, which is used in this manuscript too, is to divide the population into groups, for example age groups. Let's assume we have divide the population into age groups, each group representing a 10-year range. How can we estimate the contact rate between the group of 30-40 years old with other age groups? Even in a simplified case that the disease can spread uniformly among individuals and groups, and contact rates depend only on the physical distance between individuals, how can we find their values? To do so, we need a detailed knowledge of the mobility of individuals or groups that even the Big Brother could not dream of knowing it.

V. S. Bokharaie is with Max Planck Institute for Biological Cybernetics, Tübingen, Germany (email: vahid.bokharaie@tuebingen.mpg.de)

To address this issue, I have used the available statistics on the prevalence of the confirmed cases of COVID-19 in China in early stages of its spread (early February), and have come up with values for contact rates that can be adapted and used for other countries too. I also show that the resulting model is robust with respect to the assumptions which are made in estimating the parameters.

To understand the mathematical model presented in the next section, knowing some basic concepts of mathematical epidemiology is necessary. To model the spread of a disease in a populations, a very logical choice is to use compartmental models. In such models, population is divided into compartments which capture the progress of that disease. The three most commonly used compartments are *Susceptible*, S, which includes those who are healthy and can catch a disease. *Infectives*, I, which includes those who are infected and can also transmit the disease. *Recovered*, R, which includes those who are recovered from the disease and are immune to it. There are also other compartments that sometimes are used, for example *Exposed*, E, which includes those who are infected but do not transfer the disease. Two most commonly used models are SIS and SIR. These terms represent the progress of the disease in a population. SIS means individuals are first in susceptible group, then and when they catch the disease they join the Infective compartment, and when they recover, they join the Susceptible compartment again, meaning that there is no immunity to the disease. SIR, on the other hand, is used for cases that the disease infers immunity, and people recovered from the disease join the compartment R, not the compartment S.

Which model is suitable for a certain disease depends no the nature of that disease, and whether or not the population can have immunity to that disease or if there is vaccination available for it. For COVID-19 virus, it is still too early to say which is the case. Although there have been at least one report case in which a patient recovered from COVID-19 is infected again [7]. But more information is needed to settle the issue. And even if there is no naturally occurring immunity to the disease, a vaccine can provide it. In view of this uncertainty, I present the mathematical formulation for both SIS and SIR cases. In both cases, each compartment is further divided into groups based on the ages of the individuals. There are two justification for such a choice: one being the increased vulnerability of older individuals to the COVID-19 virus. And the other being the availability of data per age groups for confirmed cases of COVID-19 virus. But there is nothing in the mathematical formulation presented in the following that prevents any other choice for such stratification.

The mathematical models and some relevant results are detailed in Section II. The optimisation scheme used to calculate the parameters of the model based on available data on the spread of COVID-19 is presented in Section III. Conclusions and future works are discussed in Section IV.

II. Mathematical background

A. Notations and Some Basic Definitions

$\mathbb{R}$ is the field of real numbers and $\mathbb{R}_+$ is The set of nonnegative real numbers. $\mathbb{R}^n$ is The space of column vectors of size n of real numbers and $\mathbb{R}^{n\times n}$ is The space of $n \times n$ matrices of real numbers. I use x_i to represent The ith entry of the vector x in $\mathbb{R}^n$, for $i \in \{1, \cdots, n\}$. Please note that x_0 is a vector in $\mathbb{R}^n$ that usually represents initial condition. Notation a_{ij} is used for (i, j) entry of the matrix A. $D = \operatorname{diag}(x)$ is an $n \times n$ diagonal matrix in

which $d_{ii} = x_i$ for all i. A^{-1} is The inverse of the matrix A. I is the identity matrix of proper dimensions and 0 is the zero matrix of proper dimensions. $\sigma(A)$ is the set of all eigenvalues (spectrum) of the matrix A. $\rho(A)$ is the spectral radius of the matrix A, i.e. the maximum of the absolute values of all eigenvalues. $\mu(A)$ is The spectral abscissa of the matrix A.

$A \gg B$ means $a_{ij} > b_{ij}$, for all $i, j \in \{1, \cdots, n\}$. It should not be mistaken with Positive Definite (PD) matrices. $A > B$ means $a_{ij} \geq b_{ij}$, for all $i, j \in \{1, \cdots, n\}$ and $A \neq B$ and $A \geq B$ means $a_{ij} \geq b_{ij}$, for all $i, j \in \{1, \cdots, n\}$. $\mathbb{R}^n_+$ is The positive orthant of $\mathbb{R}^n$, given by $\{x \in \mathbb{R}^n : x \geq 0\}$.

Knowing the following basic definitions can help in understanding the text better.

A matrix A is called *Hurwitz*, if $\mu(A) < 0$.

A real $n \times n$ matrix $A = (a_{ij})$ is *Metzler* if its off-diagonal entries are nonnegative.

The matrix A is *irreducible* if and only if for every non-empty proper subset K of $N := \{1, \cdots, n\}$, there exists an $i \in K$, $j \in N \setminus K$ such that $a_{ij} \neq 0$. When A is not irreducible, it is *reducible*.

For any subset $\mathcal{U}$ of $\mathbb{R}^n$, a point x_0 is called an *interior point* of $\mathcal{U}$ if there is an open ball around x_0 which is wholly contained in $\mathcal{U}$. The set of all interior points of $\mathcal{U}$ is called the interior of $\mathcal{U}$ and is denoted by $\operatorname{int}(\mathcal{U})$.

Consider a continuous-time nonlinear systems of the form:

$$\dot{x}(t) = f(x), \quad x(0) = x_0 \tag{1}$$

where $f : \mathcal{D} \mapsto \mathbb{R}^n$ is a nonlinear vector field on a subset $\mathcal{D}$ of $\mathbb{R}^n$ and $x_0 \in \mathcal{D}$ is called the initial condition.

Definition 2.1: System (1) is called positive, if

$$x(t, x_0) \geq 0 \text{ for all } t \geq 0, x_0 \geq 0 \tag{2}$$

In other words, if $\mathbb{R}^n_+$ is an invariant set for the system (1), then the system is positive.

Definition 2.2 (Monotonicity): Suppose $\mathcal{D} \subset \mathbb{R}^n$ is a forward invariant set for system (1). The system (1) is *monotone* in $\mathcal{D}$ if and only if $\forall x_0, y_0 \in \mathcal{D}$ with $x_0 \leq y_0$, it holds that $x(t, x_0) \leq x(t, y_0)$.

A vector field is said to be C^1, if it is continuous and so is its first derivative.

Definition 2.3 (Cooperativity): We say that the C^1 vector field $f : \mathcal{D} \to \mathbb{R}^n$ is *cooperative* on a closed subset $\mathcal{U}$ of $\mathcal{D}$ if the Jacobian matrix $\frac{\partial f}{\partial x}(a)$ is Metzler for all $a \in \mathcal{U}$. Also, system (1) is said to be cooperative, if f is cooperative.

It can be proved that every cooperative system defined on a suitable set is monotone.

B. SIS Model

The formulation presented in this section is adopted from [6]. In this model, the population of interest is first divided into two compartments S, susceptibles, and I, infectives. Please note the usage of term *infective* as opposed to *infected*, meaning the disease does not have a latent period. For COVID-19, if a latent period exists, it seems to negligible for practical purposes [8]. As already mentioned in Section I, each compartment is sub-divided into n

groups. These groups can represent different age groups, different health condition, professions, etc. In the following, I have divided the population into age groups, for two reasons: one being the reported dependence of the spread of COVID-19 on age, and the other being the availability of data for COVID-19 based on age groups, which is the basis of the analysis reported in Section III.

Let $I_i(t)$ and $S_i(t)$ be the number of infectives and susceptibles at time t in group i for $i = 1, \cdots, n$, respectively. Also, let $N_i(t) = S_i(t) + I_i(t)$ be the total population of group i. The total population of each group is assumed to be constant; formally, $N_i(t) = N_i$. This does not usually oversimplify the model, specially when the total population is significantly greater than the number of infectives, which is still the case for COVID-19 at the time of writing this manuscript. But even if that assumption is not realistic for a population, the formulation can be altered accordingly.

Here, β_{ij}, the contact rate between groups i and j, denotes the rate at which susceptibles in group i are infected by infectives in group j for $i, j = 1, \cdots, n$. Further, γ_i, the transfer rate, is the rate at which an infective individual in group i is cured. We also consider birth and death in the population, although to keep the total population constant, we assume for each group the birth and death rates are equal to the value μ_i. Using the mass-action law, the basic SIS model is then described as follows [6]:

$$\begin{cases} \dot{S}_i(t) = \mu_i N_i - \mu_i S_i(t) - \sum_{j=1}^{n} \beta_{i,j} \frac{S_i(t).I_j(t)}{N_i} + \gamma_i I_i(t) \\ \dot{I}_i(t) = \sum_{j=1}^{n} \beta_{i,j} \frac{S_i(t).I_j(t)}{N_i} - (\gamma_i + \mu_i) I_i(t) \end{cases} \tag{3}$$

Since the population of each group is constant, it is sufficient to know $I_i(t)$. If we set $x_i(t) = I_i(t)/N_i$ and $\tilde{\beta}_{i,j} = \beta_{i,j} N_j / N_i$ and $\alpha_i = \gamma_i + \mu_i$, we obtain the following differential equation:

$$\dot{x}_i(t) = (1 - x_i(t)) \sum_{j=1}^{n} \tilde{\beta}_{i,j} x_j(t) - \alpha_i x_i(t), \quad \forall i = 1, \cdots, n \tag{4}$$

Based on the definition, $x \in B_n$ where $B_n := \{x \in \mathbb{R}^n_+ : x \leq 1\}$. We can write the differential equation (4) in compact form as:

$$\dot{x} = [D + B - \text{diag } (x)B]x \tag{5}$$

where $D = -\text{diag } (\alpha_i)$ and $B = (\tilde{\beta}_{ij}) > 0$. Please note that in the simulations I have assumed the birth and death rate is negligible compared to transfer rate, in other words, $\alpha_i = \gamma_i$.

The following properties of (5) are easy to check.

(i) $f(x) = [D + B - \text{diag } (x)B]x$ with D and B defined as above is C^1 in $\mathbb{R}^n$, therefore, the solution for every initial condition in $\mathbb{R}^n$ exists and is unique for all $t \geq 0$.

(ii) Since D is diagonal and $B > 0$, it can be easily seen that system (5) is cooperative.

(iii) The origin is an equilibrium point of (5). This equilibrium is referred to as the disease-free equilibrium (DFE) of the system (5).

(iv) Since the system (5) is cooperative and has an equilibrium at the origin, hence the system (5) is positive.

(v) System (5) may have an equilibrium in $\text{int}\,(\mathbb{R}^n_+)$ (also referred to as an endemic equilibrium). Conditions for existence of endemic equilibrium for the system (5) depends on parameter R_0, explained below.

(v) Linearising the system (5) around the origin, we obtain the following linear system:

$$\dot{x}(t) = (D + B)x(t) \tag{6}$$

The system (4) (or equivalently system (5)) has another important property:

Lemma 2.1: Let $B_n := \{x \in \mathbb{R}^n_+ : x \leq 1\}$. Then B_n is an invariant set for the system described in (4).

One important parameter in an epidemiological model is the *basic reproduction number*, R_0. There are different definitions for the basic reproduction number. Probably the most common definition is as follows.

Definition 2.4 (Basic reproduction number): The basic reproduction number is the expected number of secondary cases produced, in a completely susceptible population, by a typical infective individual during its entire period of infectiousness [5].

For the SIS model (5), following the reference [6], it can be proved that $R_0 = \rho(-D^{-1}B)$. The reproduction number can be used to characterise the existence and stability of the equilibria of (5). The following result is Theorem 2.3 in [6].

Theorem 2.2: Consider the system (5). Assume that the matrix B is irreducible. The disease-free equilibrium of the system (5) is globally asymptotically stable if and only if $R_0 \leq 1$.

The next result considers the existence and stability of endemic equilibria and is a restatement of Theorem 2.4 and the discussion in Section 2.2 of [6].

Theorem 2.3: Consider the system (5) and assume that B is irreducible. There exists a unique endemic equilibrium $\bar{x}$ in $\text{int}\,(\mathbb{R}^n_+)$ if and only if $R_0 > 1$. In such a case, $\bar{x}$ is asymptotically stable with the region of attraction $B_n \setminus \{0\}$.

C. SIR Model

The SIR model is quite similar to SIS, with a minor difference, namely, those who are cured, join the recovered, R, population, not S. Hence, the formulation for an SIR model is as follows:

$$\begin{cases} \dot{S}_i(t) = \mu_i N_i - \mu_i S_i(t) - \sum_{j=1}^{n} \beta_{i,j} \dfrac{S_i(t).I_j(t)}{N_i} \\ \dot{I}_i(t) = \sum_{j=1}^{n} \beta_{i,j} \dfrac{S_i(t).I_j(t)}{N_i} - (\gamma_i + \mu_i) I_i(t) \\ \dot{R}_i(t) = \gamma_i I_i(t) \end{cases} \tag{7}$$

Again, assuming $N_i(t) = S_i(t) + I_i(t) + R_i(t)$ is constant, similar to what was done in the previous section, if we set $x_i(t) = I_i(t)/N_i$ and $y_i(t) = R_i(t)/N_i$ and $\tilde{\beta}_{i,j} = \beta_{i,j}N_j/N_i$ and $\alpha_i = \gamma_i + \mu_i$, we obtain the following differential equation:

$$\begin{cases} \dot{x}_i(t) = (1 - x_i(t)) \sum_{j=1}^{n} \tilde{\beta}_{i,j} x_j(t) - \alpha_i x_i(t) \\ \dot{y}_i(t) = \gamma_i x_i(t) \end{cases} \tag{8}$$

$\forall i = 1, \cdots, n$. In compact from, (7) can be written as follows:

$$\begin{cases} \dot{x} = [D + B - \text{diag}\ (x)B]x \\ \dot{y} = \Gamma x \end{cases} \tag{9}$$

where $D = -\text{diag}\ (\alpha_i)$ and $B = (\tilde{\beta}_{ij}) > 0$ and $\Gamma = \text{diag}\ (\gamma_i)$ for $i = 1, \cdots, n$.

III. Using available data on COVID-19 confirmed cases to estimate parameters of the Model

In order to solve ordinary differential equations (ODEs) (9) or (5), we need to have a reliable estimate of $\tilde{\beta}_{i,j}$ and γ_i for all i, j. γ_i is easy to estimate. If the average duration that individuals in a group i are infective is 20 days, and if we choose our time unit to be one day, then $\gamma_i = 1/20 = 0.05$. Estimating $\tilde{\beta}_{i,j}$ is what we discuss in this section.

Column C2 in Table I shows the distribution of confirmed cases of COVID-19 in different age groups in China as of Feb. 11th [2]. To normalise these numbers to the relative distribution of each age group in the general population, we can divide values in Column C2 to those of Column C1. We can further divide the resulting numbers to the smallest of them, which happens to be the first row. By doing so, we obtain the last column of Table I, that shows the normalised relative distribution of infective people in the Chinese population as of Feb. 11th.

We can then use these figures as the desired values for the ratio of different states in (9) or (5) such that at the time t_g, the ratio of the values of each of the states are approximately those reported in the last column of Table I. At the same time, we should keep the basic reproduction number R_0 to be equal to the estimated R_0 for the spread of COVID-19 in an unconstrained population. There are various estimates for R_0 for COVID-19, some are listed in [1]. Most estimates fall in $[1.5, 3.5]$ range. I have chosen $R_0 = 2.28$, as reported in [11].

Hence, the optimisation problem we need to solve is as follows: finding the matrix B (as defined in previous section) that satisfy the following two conditions:

(i) For a given dialogical matrix D and scalar R_0, $R_0 = \rho(-D^{-1}B)$

(ii) The solutions of (5) or (9) at a given time t_g and initial condition x_0 satisfy a certain condition.

But how to choose t_g and x_0? For that, I have relied on reports that the spread of the virus has probably started in a wet market in Wuhan city, in late November [10]. Hence I have set $t_g = 75$ days (meaning the spread of the virus has initiated 75 days before Feb 11th), and x_0 to be 0 for all groups except 0.0001 in the fourth age group, which corresponds to people aged 40-50. Later on we will see that the results are robust with respect to initial conditions and exact value of t_g. I have also set constrained for the minimum and maximum of the elements of matrix B to be 0.0001 and 0.1 to avoid solutions with unrealistic values.

Now that all the required parameters are set, we can solve the optimisation problem to find a B_{opt}. In order to solve the optimisation problem, I have used `globalsearch` function in Matlab© with `sqp` algorithm. Optimisation is done in two steps, in the first step, initial values for matrix B are chosen randomly from a uniform distribution. When the optimisation algorithm converges to a solution, the optimisation procedure is repeated, this time with the optimum value obtained in the first step as the initial values. The objective function in the optimisation scheme is the weighted sum of two terms. One being the 2-norm of the difference between ratio of trajectories of an SIS or SIR model at time t_g with initial condition x_0 with the desired ratio extracted from Table I. The second term is the difference between $\rho(-D^{-1}B)$ and the desired basic reproduction number of $R_0 = 2.28$.

Age Group	Population Ratio in % (C1)	Confirmed COVID-19 Ratio in % (C2)	C2/C1	C2/C1 normalised to first row
0-9	11.9	0.9	0.075	1.0
10-19	11.6	1.2	0.107	1.4
20-29	13.5	8.1	0.600	8.0
30-39	15.6	17.0	1.089	14.5
40-49	15.6	19.2	1.230	16.4
50-59	15.0	22.4	1.493	19.9
60-69	10.4	19.2	1.846	24.6
70-79	4.7	8.8	1.872	25.0
80+	1.7	3.2	1.882	25.1

TABLE I

POPULATION DISTRIBUTION IN CHINA AND THE DISTRIBUTION OF CONFIRMED COVID-19 CASES IN CHINA AS OF FEB. 11TH. TO NORMALISE THE DISTRIBUTION OF CONFIRMED CASES, WE CAN DIVIDE THE VALUES OVER THE POPULATION RATIO OF THAT AGE GROUP. THE RESULTING VALUES ARE THEN USED TO ESTIMATES PARAMETERS OF BOTH SIS AND SIR EPIDEMIOLOGICAL MODELS.

The optimisation algorithm runs in 5-15 minutes on 40 hyper-threaded cores of type Intel(R) Xeon(R) CPU E5-2687W v4 @ 3.00GHz.

Note 3.1: It should be noted that the values obtained from the optimisation schemes are $\tilde{\beta}_{i,j}$ which usable only for population distribution in China. But using the relationship $\beta_{i,j} = \tilde{\beta}_{i,j} N_i/N_j$, we can obtain universal values that can hold for all population densities, where N_i/N_j is the ratio of population ratio in age group i over that of age group j. For each other target population we can use the equation $\tilde{\beta}_{i,j} = \beta_{i,j} N_j/N_i$ to calculate the $\tilde{\beta}_{i,j}$ and then solve ODEs (5) or (9) to find the spread of the disease in each age group over time.

A. Sensitivity Analysis

The values set for x_0 and t_g are based on speculations. So it is necessary to check to what extent the results of the optimisation scheme depend on the choices of these variable.

In order to do so, in SIS case, I varied t_g in $[55, 95]$ days range, up to 20 days more or less than the 75 days guess. Then solving the ODE described in (4) in each case, we can calculate $x(t_g)$, normalised it to the value of the first element, and then compared the relative error between the resulting vector and the desired values as reported in the last column of Table I. The maximum relative error over all elements of $x(t_g)$ and all values of t_g was 1.96%. In other words, if the initial guess of $t_g = 75$ days was wrong and the spread of the disease started any time from 95 days to 55 days before the Feb. 11th date, the values we obtained $\tilde{\beta}_{i,j}$ would lead to a ratio between states that match the desired ratios with an maximum error of 1.96%.

As a reminder, to solve the optimisation problem, I assumed x_0 is 0 everywhere except its 4th element (corresponding to the age group 40-50) which is set to `1e-4`. Note that this value represent the ratio of the population of infective in that age group to the total population in that age group. So, if the age group [40-50] include millions of

individuals, which is a reasonable assumption for the regions dealing with COVID-19 in China as of Feb 11th, the absolute value would be of the order of hundreds of people. To test the sensitivity to x_0, in the first step, I assumed one age group among [20,30], [30,40], [40,50], [50,60] would be `1e-4` while other are 0. The total relative error over all such cases was 5.12%. Interestingly enough, when the first group, age group 0-10, was assumed to be `1e-4` and others 0, the maximum error jumped to 19.4%, which intuitively is what one expects from the completely different reported effect of COVID-19 on children as compared to adults. I then assume all four age groups from 20 to 60 have a value of `1e-4` while other are 0, and the maximum error was 1.2%, which is even less than the previous case, and I find peculiar.

I have done the same procedure for initial values of `1e-3` and `1e-5`. Once assuming only one age group in [20,30], [30,40], [40,50], [50,60] has such a value, and then all of them together. For the value of `1e-3`, maximum error was 6.69% over each of the four distinct cases, and 7.39% when all 4 age groups started from `1e-3`. When initial condition was set to `1e-5`, the maximum error in these two cases were 5.02% and 0.85% respectively.

For SIR case, maximum relative error when t_g changes in [55, 95] days range is 4.70%.

When any of the 3rd, 4th, 5th or 6th elements of x_0 is `1e-5`, `1e-4` or `1e-3`, maximum relative error in each case for the solution of the SIR system (9) over all four cases is 3.29%, 2.71% and 8.50% respectively. When all four age groups start from those initial values, the maximum relative errors are 0.67%, 2.09% and 20.77% respectively (more detailed analysis is needed to figure out why the last case has a relatively higher error).

To summarise, the sensitivity analysis shows that even if out guess for when the COVID-19 virus has started to spread in human population is off by 20 days, or if the initial number of infectives is of one order of magnitude lower or higher than what we have guessed, the model obtained from the optimisation scheme generally performs well in reaching the desired values presented in Table I.

IV. Conclusions

In this manuscript, I presented a method to estimate contact rates for an SIS or SIR model, with known recovery rate and basic reproduction number, based on available data for COVID-19. The main advantage of the method is that the resulting contact rates can be adapted and used in other countries or societies. But the implicit assumption is that the spread of COVID-19 in different countries depends only on the age distribution on that population. It might or might not be true. Also, if basic reproduction number in a country happens to be different than the values considered in the optimisation scheme, $R_0 = 2.28$, we can simply multiply the matrix B by a suitable scalar to change R_0. This follows simply from the definition of eigenvalues for a matrix.

The results can be used not only to calculate the changes in the number of infective people in each age group over time, but to examine the effect of different containment strategies. As an example, assuming only schools are shut downed in a population, that corresponds with changes in contact rates for 1st and 2nd states,i.e. $\tilde{\beta}_{1,j}$ and $\tilde{\beta}_{2,j}$ for all $j = 1, \cdots, 9$. Estimating the ratio in the number of individuals a pupil encounters with when going to school compared to when he/she stays home is not difficult.

This work can be extend in a number of ways. Most importantly, the groups in (9) or (5) need not to be age groups. These groups can be defined based on any characteristics of a population, such as profession, health condition, relative mobility and so on. But without reliable statistics on the spread of COVID-19 in such groups, we cannot use the methodology presented in this manuscript.

REFERENCES

[1] "2019 novel corona virus parameter estimates," https://github.com/midas-network/COVID-19/tree/master/parameter_estimates/2019_novel_coronavirus, accessed: 2020-03-14.

[2] Breakdown of 44,672 sample patients infected with novel coronavirus covid-19 in china as of february 11, 2020, by age group. Accessed: 2020-03-14. [Online]. Available: https://www.statista.com/statistics/1095024/china-age-distribution-of-wuhan-coronavirus-covid-19-patients/

[3] Population pyramid of china. Accessed: 2020-03-12. [Online]. Available: https://www.populationpyramid.net/china/2020/

[4] Y. Bai, L. Yao, T. Wei, F. Tian, D.-Y. Jin, L. Chen, and M. Wang, "Presumed Asymptomatic Carrier Transmission of COVID-19," *JAMA*, 02 2020. [Online]. Available: https://doi.org/10.1001/jama.2020.2565

[5] O. Diekmann, J. Heesterbeek, and J. Metz, "On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations," *Journal of mathematical biology*, vol. 28, no. 4, pp. 365–382, 1990.

[6] A. Fall, A. Iggidr, G. Sallet, and J. Tewa, "Epidemiological models and Lyapunov functions," *Math. Model. Nat. Phenom.*, vol. 2, no. 1, pp. 62–68, 2007.

[7] L. Lan, D. Xu, G. Ye, C. Xia, S. Wang, Y. Li, and H. Xu, "Positive RT-PCR Test Results in Patients Recovered From COVID-19," *JAMA*, 02 2020. [Online]. Available: https://doi.org/10.1001/jama.2020.2783

[8] C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, M. Seilmaier, C. Drosten, P. Vollmar, K. Zwirglmaier, S. Zange, R. Wlfel, and M. Hoelscher, "Transmission of 2019-ncov infection from an asymptomatic contact in germany," *New England Journal of Medicine*, vol. 382, no. 10, pp. 970–971, 2020. [Online]. Available: https://doi.org/10.1056/NEJMc2001468

[9] P. van den Driessche and J. Watmough, "Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission," *Math. Biosci.*, pp. 29–48, 2002.

[10] Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: summary of a report of 72 314 cases from the chinese center for disease control and prevention," *Jama*, 2020.

[11] S. Zhang, M. Diao, W. Yu, L. Pei, Z. Lin, and D. Chen, "Estimation of the reproductive number of novel coronavirus (covid-19) and the probable outbreak size on the diamond princess cruise ship: A data-driven analysis," *International Journal of Infectious Diseases*, 2020.