

Understanding of COVID-19 Spread with Mathematical Modeling

Haolin Li

May 15, 2020

Abstract

The SEIR (Susceptible-Exposed-Infectious-Recovered) model is a system of differential equations designated for the epidemic that not only can describe the epidemic, but also can predict the spread of the epidemic. In this paper, we want to provide a breakdown with details of the SEIR model and especially how it works within the pandemic disease, COVID-19. We covered the background of the SEIR model and how the model is a good fit to the coronavirus (COVID-19) epidemic. We experimentally evaluated the spread of COVID-19 by building new models based on the SEIR model, and adding new parameters to the models, to make prediction of the spread with the given information and data. Our experimental results showed the potential of the model and how they match with the current epidemic. The ultimate goal of this paper is to provide an insight of the SEIR model and the fit of the SEIR model to the COVID-19.

1 Introduction

On December 31, 2019, a Chinese city called Wuhan reported the outbreak of the novel coronavirus. The Chinese government wasted no time, they immediately started to lockdown all cities and strictly controlled people's movement. On March 11, 2020, the World Health Organization (WHO) has declared the novel coronavirus (COVID-19) outbreak a global pandemic. Many countries have started quarantine, shelter-in-place policy and social distancing in response to the rapid increase of infections and deaths caused by the coronavirus. The virus is spreading rapidly and drawing everyone's attention. In April 2020, more than 2 millions people were infected and almost 200 thousands people died from COVID-19 worldwide, the number of confirmed infections were reaching high on the chart and continually growing. Coronavirus has caught people's attention in a very short amount of time. In this paper, we will discuss the spread of the coronavirus with tools from mathematics and modeling to help people understand the spread of the coronavirus and how we can all stop the spread by following guidelines from the authorities.

2 The Problem

The spread of COVID-19 shows no sign of slowing down. The numbers of confirmed cases and deaths are increasing dramatically. Even though the spread of the disease seems impossible to slow down in a short period of time, the spread of COVID-19 can actually be predicted with a mathematical model. A well-designed mathematical model can help to predict the spread of COVID-19 and show how to slow down and stop the spread. Another benefit of utilizing a well-designed model is that it can reveal the most helpful strategies to control the spread of the disease. Therefore, this paper is about the mathematical model we used to predict the spread, it is called SEIR model. SEIR model takes in the account of four categories, the susceptible, exposed, infectious and recovered, basically the four most important groups of people in the population regarding the epidemic. Also, there are three parameters, the infectious rate, the incubation rate and the recovery rate. In a whole, they are to calculate and make prediction of the change of COVID-19. In order to improve the

prediction number and find out the most promising strategy to control the spread, we may consider other more factors that may affect the COVID-19 as well, such as the behavior of isolation (social distancing) and the behavior of testing COVID-19. In this document, we firstly explained how the SEIR model works and the fundamentals of the model, then we added social distancing policy, assuming COVID-19 can be transmitted by the exposed population and the behavior of testing for each group of people, to show the impact of these conditions on the model. To address the work for this project, our goal is to show the fit of the SEIR model with the coronavirus public data such as the death rate, recovery rate and more, and how this model is relevant for this epidemic.

3 Why is this Interesting

Even though the authorities shows the spread of the COVID-19 daily, people still lack understanding of the virus, how severe it is and how fast it grows. They do not realize how government policies are helpful to limit the spread of the disease. Many people ignore the policy of quarantine mainly because they underestimate the infectiousness of coronavirus. People watch the report everyday about the disease but have no idea of how professionals come out with the information, the numbers and the graphs. People deserve answers regarding everything about the disease because we are all in this together and having a better understanding of COVID-19 would definitely help to end it sooner. Therefore, it is interesting to show the mathematics behind, to show people how severe is this disease and how mathematics can predict the spread and improve the situation.

4 The Summary of the Proposed Solution

The ultimate goal of this paper is to provide an insightful look of the SEIR mode and the fit of the SEIR model to the COVID-19. We mainly separated this document into three parts where the first part talks about the basics of the SEIR model, then the second part we included a new parameter to represent the social distancing policy and the third part we included the behavior of people testing COVID-19 into the model. The basics of the SEIR model included the background of the SEIR model and the breakdown of the model, showing the evolution of the model helps to understanding it in details because each category within the model contributes to the result significantly. At the end of this paper, we hope people understand the spread of COVID-19 and how each factor can be explained in mathematics, as well as how this model can improve the future spread by providing strategies to slow and stop the disease.

5 Background of the SEIR Model

SEIR model falls into the compartmental models in mathematical modelling of infectious disease. The origin of such models is in the early 20th century[Wikipedia, 2017], according to Wikipedia. The model is composed of ordinary differential equations which are deterministic. SEIR model is to predict the spread of a disease or the number of infected individuals,

duration of the epidemic, or to estimate the parameters within the model. One of the powerful parts of this model is it can show how different public health policies may impact the spread and outcome of the epidemic, such as the recent policy of shelter in place and social distancing. There are four members in the model, susceptible, exposed, infectious and recovered. The members in the model represent the important members of an epidemic within a population. There are three parameters in the model, the infectious rate, incubation rate and the recovery rate. The three rates are important components of the model because they are the relationship between the members.

6 Composition of SEIR Model

In this section, we will discuss the SEIR model by breaking down the model by each member in the population, to show the relationship between the members and how the model is formed. Breaking down the model by each member can also reveal how the model is formed based on the reality of an epidemic.

6.1 SI Model

In this model, we assumed there are only two types of members in the population of an epidemic, susceptible and infectious. Members of the susceptible are the healthy individuals, and members of the infectious are the individuals carrying the virus, which is the infected people. In the differential equations, we will use S to represent the susceptible and I to represent the infectious. Since there are only two types of members in a population, the whole population denoted as N in the differential equations, is $N = S + I$. There is one parameter in the model, β , to represent the probability that a susceptible to become infectious. In this case we have one more component of the model, r , to represent the daily number of contact of an infectious. To summarize the differential equations in a whole, we can say that there are I number of infectious walking in the population N , they have close contact with r people daily, therefore we have a probability β that a susceptible to become infectious. Then the differential equations become:

$$\begin{aligned}\frac{dI}{dt} &= \frac{r\beta IS}{N} \\ \frac{dS}{dt} &= \frac{-r\beta IS}{N}\end{aligned}$$

The relationship between the susceptible and the exposed can be visualized in the diagram shown on the next page:

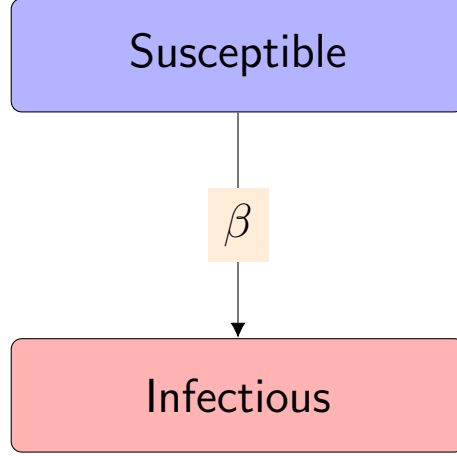


Figure 1: Visualization of the SI Model.

6.2 SIR Model

Now we can move onto the next model, enhance the model by adding a member to the model and introducing a new parameter. This model is slightly complicated than the SI model we introduced previously. In this model, we have three members in the population of the epidemic, the susceptible, infectious and the recovered. The recovered is denoted as R in the system of equations. Since we denoted the recovered as R in the differential equations, the whole population now becomes $N = S + I + R$. The case here is that once an infectious becomes recovered, they will not become infectious again, in other words, recovered people has immunization to the virus. We have a new parameter, denoted as γ , to represent the probability of recovery. There is the case where the recovered will be infected again, however, for the COVID-19, there is very few to none cases of reinfections. Therefore, in this model we do not consider the possibility of reinfection. Then, the differential equations become:

$$\begin{aligned}\frac{dI}{dt} &= \frac{r\beta IS}{N} - \gamma I \\ \frac{dS}{dt} &= \frac{-r\beta IS}{N} \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

The new relationship between the members of the population as the recovered joined the population can be now visualized in a diagram shown in the next page:

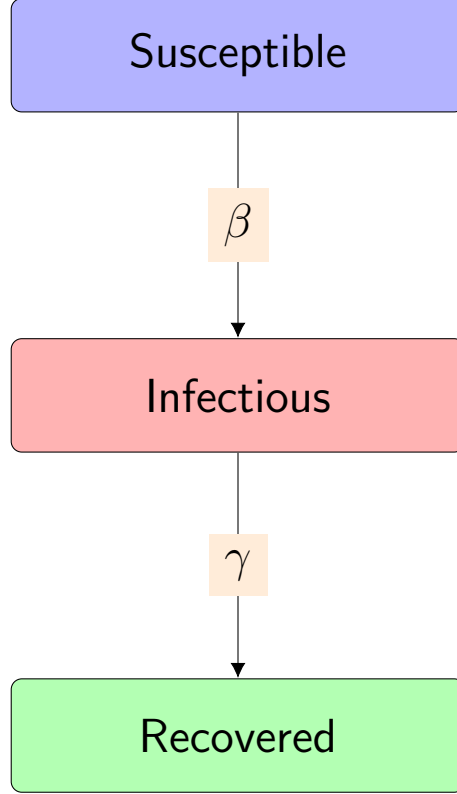


Figure 2: Visualization of the SIR Model.

6.3 SEIR Model

The previous two models were simple but not quite relevant to the actual epidemic. The reality is that it is more complicated to model the spread of the disease. If a member from the susceptible is infected by an infectious, they do not immediately become infected, they become a member of the exposed population because they experience the incubation period. Only after the incubation period, they then become either susceptible or infectious. Therefore, we have a new coefficient can be added to the model, which is called the incubation rate, the rate of latent individuals becoming infectious, we denoted such rate in the differential equations as σ . Since we have a new coefficient, we also have a new type of member in the population, the exposed, denoted as E in the equation, which is the group of individuals contacted with the infectious and possibly to become infectious after a period of time. The whole population now becomes $N = S + E + I + R$. Similar to the SI model was modified based on the SI model, we now obtained the SEIR model based on the SIR model. The differential equations now become:

$$\begin{aligned}\frac{dI}{dt} &= \sigma E - \gamma I \\ \frac{dS}{dt} &= \frac{-r\beta IS}{N} \\ \frac{dR}{dt} &= \gamma I \\ \frac{dE}{dt} &= \frac{r\beta IS}{N} - \sigma E\end{aligned}$$

The final model can be now visualized as the following diagram with four members of the population and three coefficients representing the rates:

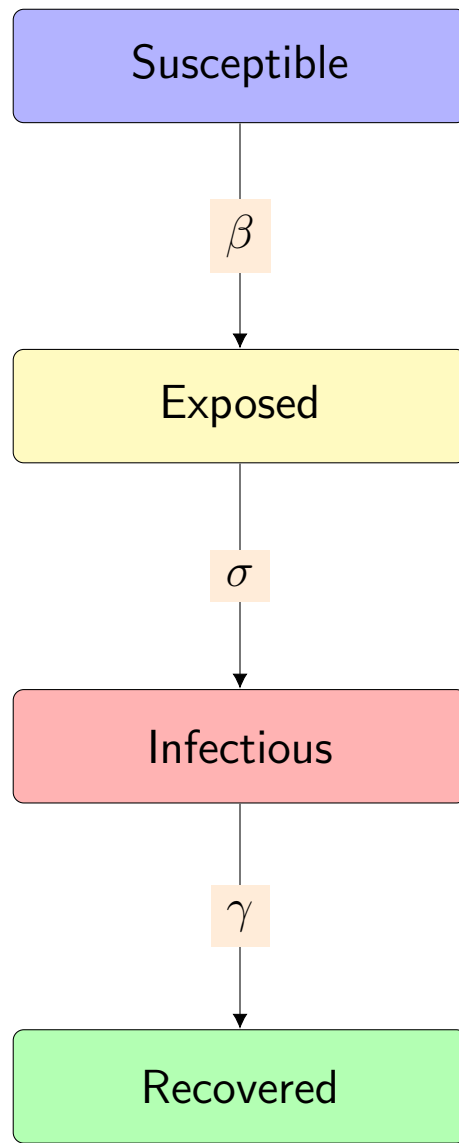


Figure 3: Visualization of the SEIR Model.

6.4 Model Visualization

The diagrams shown above are for the purpose of understanding how the SEIR model is generated to fit the pandemic outbreak as each member plays an important role in the model, and the significant relationship between the members that we need to be concerned about, as well as the order of an infection. The visualization of the SEIR model diagram is inspired by the explanation of the model from the documents by the Institute for Disease Modeling[IDM, 2019]. The SI and SIR models are simply the decomposition of the SEIR model, their purposes here are to show the process of generating a mathematical model (SEIR Model) for the epidemic in an appropriate order.

7 Basic SEIR Model on COVID-19

In this section, we will experiment the SEIR model with the data of COVID-19 to see the spread of this epidemic. We will experiment the model with only the basic information given in the model, in the hope of to further explain the SEIR model in reality.

7.1 Assumptions

In this model, there are three assumptions made. We assumed that the total population remains unchanged, which means that the total population is still $N = S + E + I + R$. The total population excluded births and deaths caused by other events because within a closed population, a more understandable model can be formed and better to show the fit of the model to the virus. Another assumption is that for the COVID-19, we only focused on the person to person transmission despite in reality, there are three modes of transmission for COVID-19. The reason we only focused on one transmission is also due to the simplicity of forming the model. The last assumption we made is the recovered population do not experience reinfection, this is due to the fact that COVID-19 has a very few to none cases of reinfections, as mentioned in the above SIR model description. With these three assumptions made, we now can move forward on building the model.

7.2 Member of the Model

For the members of the population in the model, we focused only on the US population for the sake of simplifying the problem. As of July 1, 2019, the population of the United States is 331,883,985 people according to U.S. Census Bureau[Census, 2019]. Therefore, before the COVID-19 outbreak in the United States, there are 328 millions of healthy individuals in the country, the number of susceptible of the model should be the population minus 1. The reason that the number of susceptible is the U.S. population minus 1 is that we need at least one individual carrying the virus in order to infect other people. Therefore, the number of infectious should be 1. (Since we want to estimate the spread of the disease, and the goal of this section is to show the basic idea of how the SEIR model applies in the COVID-19 epidemic, exact date of the outbreak is not crucial here yet.) We want to

estimate the spread of the disease within 100 days from the start. Therefore, the initial number of recovered should be 0 since only one infectious exists and the outbreak is only at its beginning. The number of exposed should also be 0 because similar to the reason for the recovered, we estimate the spread from the start and no exposed case should be appeared on day one. Overall, the number for each member of the model should be:

$$\text{Number for Members of the SEIR Model} \begin{cases} S = 331,883,985 \\ E = 0 \\ I = 1 \\ R = 0 \end{cases}$$

7.3 Parameters of the Model

Since the outbreak of COVID-19 is the first of its kind, many of the data are missing for this disease. However, human history experienced many other epidemics such as SARS, Ebola and H7N2, conducted many experiments and collected many useful data, we can extract data from these cases since COVID-19 share similarities with these previous diseases. In this case, we have three coefficients, β , γ and σ which represent the infectious rate, the recovery rate and the incubation rate. Thanks to the researchers from Stanford University School of Medicine, a group of researchers published a research paper online in JAMA regarding the infectious rate of the COVID-19 and came out with the number of 0.21[Beusekorn, 2020]. We calculated the recovery rate based on the daily report number of COVID-19 from the data collected by Atlantic and is 0.81[Atlantic, 2020]. The incubation rate is typically calculated as 1/the incubation time. According to a research article from American College of Physicians, the average of incubation period is about 5 days, 5.2 days to be exact, therefore the incubation rate should be $1/5.2$ [ACP, 2020]. Despite we have all three coefficients regarding the COVID-19, the most important piece of information we need to know is the number of people have close contact with daily. Unfortunately, United States has no yet started contact tracking for the infectious, therefore we obtained the number of contacts from the previous epidemic called H7N2, which in average, an infectious person contacted 18 people daily[Eames, 2010]. Overall, we obtained the information for our final inputs:

$$\text{Number for Coefficients of the SEIR Model} \begin{cases} \beta = 0.21 \\ \gamma = 0.81 \\ \sigma = 1/5.2 \\ r = 18 \end{cases}$$

7.4 Results

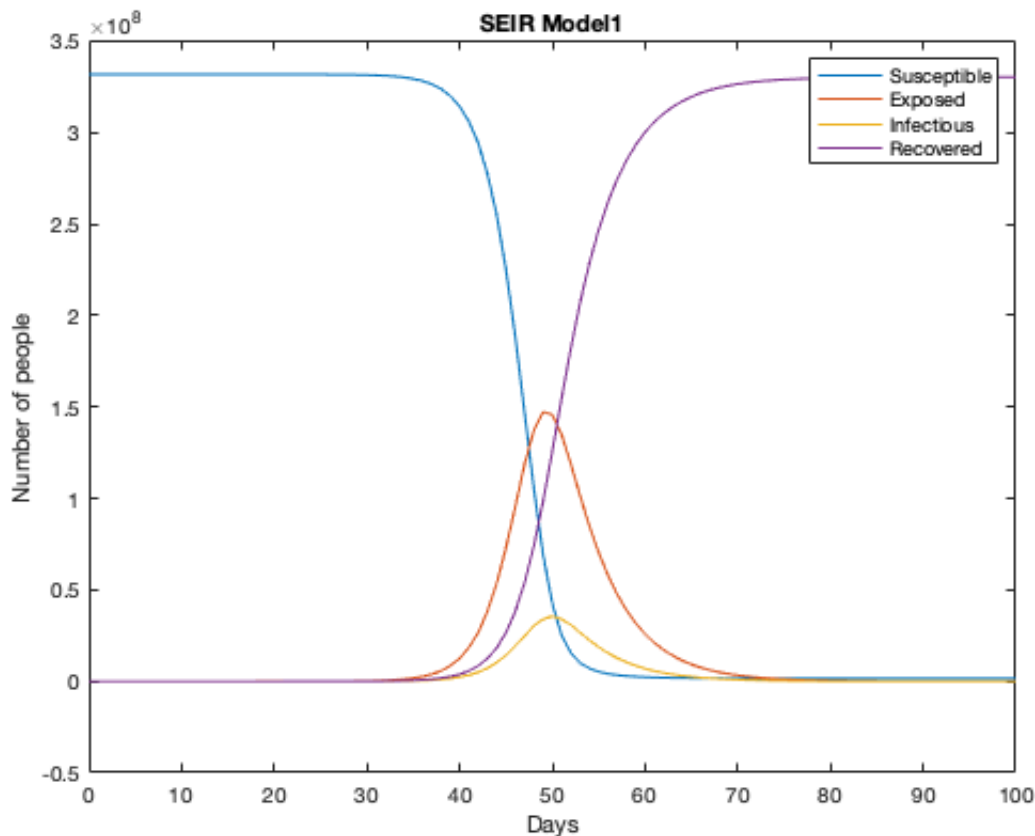


Figure 4: Visualization of the spread of COVID-19 with the SEIR model.

From the resulting plot, it shows the change of the members in the models. The blue line is representing the susceptible in the population. On day 0, the population of the susceptible is close to the whole U.S. population because the number of infectious is as low as 1 and no cases of exposed and recovered population yet. Then there is a change at about day 40, where the number of susceptible drops and the numbers of recovered, exposed and infectious increase. The line of susceptible drops significantly starting at about day 40, this is due to the skyrocket of recovered, infectious and exposed. From day 40 to 50, the increasing speed of the infected is very fast with the maximum of the infected reaches to about 35 millions at day 51, which is about 10.7% of the whole population.

By comparing the change of the members, there are some points worth to discuss. The change of recovered is higher than the change of infectious due to the recovery rate is higher than the infectious rate. The change of exposed is higher than the change of recovered and infectious, because each infectious person daily contact with 18 people, which means for every infectious there is an 18 units increase of exposed everyday, this is significantly large comparing to other rates. As the number of recovered continually increases, there is a decrease of exposed and infectious, and the whole population to be recovered on about day 80.

7.5 Conclusion

Overall, this whole section displays how the SEIR model fits into the COVID-19 epidemic by showing how the members and parameters of the model match with the information regarding the coronavirus and displaying visually the change of the COVID-19 in response to time. The plot shows that COVID-19 indeed grows rapidly and widely. Despite we are able to generate a plot of change of the SEIR model for COVID-19, in reality the spread and growth on the plot can be much different. In the next section, we again use the SEIR model on COVID-19, but we will add additional parameters into the model regarding the events of the COVID-19.

8 Social Distancing Policy in SEIR Model

In this section, we will furthermore experiment the SEIR model and add the social distancing policy of COVID-19 into the model, to predict the spread of this epidemic. We will also make an adjustment to the number of contacts for the members of the population, to match with the reality of the situation. We want to see if social distancing policy is an efficient strategy to slow and stop the disease.

8.1 Background of Social Distancing Policy

According to Wikipedia, the origin of social distancing can be traced back to the 18th century when the Lazzaretto of Ancona was constructed on an island to serve as a quarantine station for Ancona, Italy[Wikipedia, 2020]. The first case of social distancing in the United States was the New York City polio epidemic back in 1916 where public gatherings were banned, meetings were cancelled, and people were told to avoid crowded places such as parks and beaches. According to Centers for Disease Control and Prevention, to practice social distancing, people need to stay at least 6 feet from each other, do not gather in groups, stay out of crowded places and avoid mass gatherings[CDC, 2020]. The main goal of this act is to limit close contact with others outside the household in indoor and outdoor spaces.

8.2 Assumptions

In this mode, we made two assumptions related to the COVID-19. The first one is we added the social distancing policy into the model to see if social distancing policy is a promising strategy for COVID-19. The second assumption is that the exposed can infect the susceptible in the population. The reason we are assuming the exposed can infect the susceptible is the fact that COVID-19 can be transmitted from the exposed population during the incubation period, according to WebMd, people who are infected can spread it to others 2 to 3 days before symptoms start[WebMD, 2020].

8.3 Change of the Model

Since we want to see if applying social distancing policy and assuming the exposed population can infect others, will affect the original model we predicted from the previous section, many

of the features we remain unchanged. Our goal is to compare the two models and see if there are any differences. Recall from section 2, the problem, we want to improve the prediction number and find out the most promising strategy to control the spread. Therefore, at the end of this section, we hope to see social distancing policy is one of the improving strategies that slow and stop the spread of COVID-19.

8.3.1 Member of the Model

The members of the model remain the same, there are susceptible, infectious, exposed and recovered, represented by S, E, I, R respectively. The input for each member also remains unchanged. The reason that these number are unchanged is that social distancing policy and assuming exposed can infect others, do not affect the original number of people in the population. These inputs need to be unchanged to simulate the model to start from the beginning. Since we want to predict the spread from day 0, the input for the members are the same from the previous model:

$$\text{Number for Members of the Modified SEIR Model} \begin{cases} S = 331,883,985 \\ E = 0 \\ I = 1 \\ R = 0 \end{cases}$$

8.3.2 Parameters of the Model

The parameters of the model remain mostly unchanged, however, there are new parameters introduced to this modified model. Before the social distancing policy, the number of people each person contacts with daily is 18. During the social distancing policy, the number of contacts should be decreased since the policy limits people's contact to others, and we assumed this number to be 7. The new parameter is introduced as $r2$ in the equations.

We also have a new parameter called k in this model, it is the number of contacts for the exposed population. The reason we have a different number of contacts for the exposed is that we are assuming during the incubation period, the exposed population can transmit the coronavirus but not as strong as the infectious population does, therefore they should infect less people and contact less people, and this value is lower than r . During the social distancing policy, this parameter should be decreased, so we called it $k2$ in the model.

The exact number of contacts is not very significant here in the model, the main idea is that during the social distancing policy, people have less contact with each other and therefore the number of contacts should be dropped. So, we made $r2$, k and $k2$ to be 7, 8 and 2. Additionally, since the United States has not yet started contact tracing for COVID-19, we do not have efficient data for this disease and therefore we obtained the input based on the previous epidemic, H7N2. Overall, the information of the parameters for our modified model is:

$$\text{Number for Coefficients of the SEIR Model} \left\{ \begin{array}{l} \beta = 0.21 \\ \gamma = 0.81 \\ \sigma = 1/5.2 \\ r = 18 \\ r2 = 7 \\ k = 8 \\ k2 = 2 \end{array} \right.$$

8.4 The Model

Since we are assuming the exposed population has the ability to infect others, we need to modify the SEIR model to fulfill this assumption. From the model shown below, we displayed the modified model with two colors. The way we achieve the assumption of the exposed in the differential equations is to add the blue highlighted parts into the derivative of S and E. The red highlighted part is the original model 1. Derivative of R and I remain unchanged. The blue highlighted part is the probability of susceptible to become exposed, times the number of contacts of the exposed, which is k. Adding the blue part into the original equation increases the potential of having a more exposed population from the susceptible.

$$\frac{dE}{dt} = \frac{r\beta IS}{N} + \frac{k\beta ES}{N} - \sigma E$$

$$\frac{dS}{dt} = \frac{-r\beta IS}{N} - \frac{k\beta ES}{N}$$

$$\frac{dR}{dt} = \gamma I$$

$$\frac{dI}{dt} = \sigma E - \gamma I$$

The above set of differential equations are for before the social distancing policy. Recall that during the social distancing policy, the numbers of contacts for the susceptible and exposed decreased, and we called them $r2$ and $k2$. Below is the set of differential equations during the social distancing policy, with the cases of stating the change for number of contacts.

$$\frac{dE}{dt} = \frac{r2\beta IS}{N} + \frac{k2\beta ES}{N} - \sigma E$$

$$\frac{dS}{dt} = \frac{-r2\beta IS}{N} - \frac{k2\beta ES}{N}$$

$$\frac{dR}{dt} = \gamma I$$

$$\frac{dI}{dt} = \sigma E - \gamma I$$

change of r and k during quarantine $\begin{cases} r_2 < r \\ k_2 < k \end{cases}$

8.5 Results

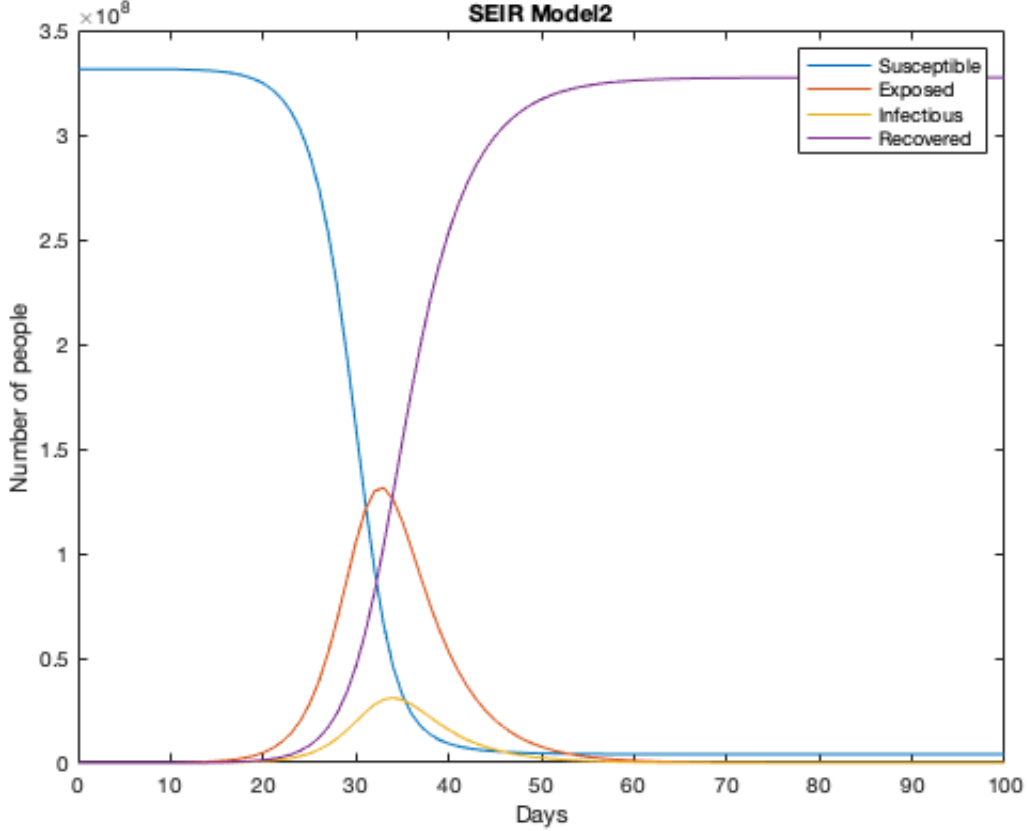


Figure 5: Visualization of the spread of COVID-19 with Quarantine Policy.

From the plot shown above, there are changes and similarities between this plot and the plot of the SEIR model 1. Same as the last plot, we have blue line representing the susceptible and the purple line for the recovered. The trend of this plot is the same the the last plot. The line of susceptible drops as the line of recovered, infectious and exposed increase. The changed of exposed is still higher than the change of recovered and infectious, this is due to the unchanged numbers of parameters. And at the end, susceptible reaches to the bottom and recovered reaches to the top.

The changes of this plot comparing to the last plot is obvious. We have the whole trend moved to an earlier number of days. From the last plot, the drop and rise of the susceptible and recovered happen at around day 40, but in this plot, the changes happen at about day 20. Since the dropping of the susceptible and the increasing of the recovered happen early, they also end early. The end day for plot 1 is about at day 80 while for plot 2 is about at day 60. These changes are mainly due to the changes for number of contacts during the social distancing policy, because people have less contact, the transmission slows down

and the spread slows down. There is one more change worthy to discuss is the change that the assumption of the exposed can infect others made, it pushes the whole trend to starts early because since the exposed population can infect others, the power of infecting increases and more susceptible become exposed, which matches the characteristic that it increases the potential of having a more exposed from the susceptible.

8.6 Conclusion

Overall, this section extents from the previous section about the basic of the SEIR model, however, the main difference between this section and the previous one is that we took social distancing policy into consider of the model and we assumed the exposed population can carry the virus and transmit. From the results shown above, social distancing policy indeed helps to slow down the spread and end the epidemic sooner than not having it, this means that social distancing policy is one of the most promising strategy to stop the spread of COVID-19. In the next section, we again use the modified SEIR model to predict the spread, but we will add additional parameters and inputs into the model to further improve the model.

9 Testing in SEIR Model

In this section, we will again further experiment the SEIR model. We will modify the model based on model 2, which means we will also include social distancing policy and the assumption about the exposed population. We will add the behavior of testing for the infectious, susceptible and exposed in this model. The goal here is we want to see if behavior of testing will change the spread of the coronavirus, and we want to add testing rate into the model to make it more relevant to the reality.

9.1 About Testing

According to an article from San Francisco Chronicle, written by Dominic Fracassa, on the week of May 19, three new COVID-19 testing locations will open to public in San Francisco, as well as including the city's first mobile testing site[Fracassa, 2020]. As of today, there are 10 locations where you can get tested for COVID-19 in San Francisco. The reason why that there are many testing locations in the city, as well as in other cities, is that without testing, there would be no data about the spread of COVID-19 and we would not know how severe the virus is and how close it is to us. If there is no testing, we cannot understand the epidemic since we need to know more about the virus, then we know what to do to prevent it. Therefore, having more testing locations encourage people to get tested for COVID-19, which then leads to having more data about the disease. Hence, testing is important because it gives information about the virus and the more information we have, the better we can control the virus.

9.2 Assumptions

In this model, we made four assumptions regarding the testing for COVID-19 in the population. The first assumption is that the distribution of the population for susceptible, exposed and infectious is 0.8, 0.1 and 0.1 respectively. The second assumption is that the test result for the susceptible is always negative, which means the susceptible will always test negative for COVID-19 since they are healthy. Then the third assumption is the test result for the infectious is always positive. The reason we assumed the test result for susceptible and infectious are always negative and positive is we want to simplify the calculation in the model. The last assumption is the isolation of the exposed and infectious population. We assumed that their number of contacts to be extremely low because they are either isolated by authorities or self-isolated. With all the assumptions made, we can now form the model.

9.3 Change of the Model

Since we wanted to see if behavior of testing will affect the SEIR model of COVID-19 we obtained from the previous section, we kept most of the features and rates unchanged, but we added significant rates into the model 3 to achieve the testing behavior in the COVID-19.

9.3.1 Members of the Model

In this model, the members of the model remain the same as we still have four group of people, susceptible, infectious, exposed and recovered, represented by S, E, I, R respectively. Since the test rate will not affect the starting value of each group, we kept their inputs unchanged. Therefore, we still have the members of the model as:

$$\text{Number for Members of the Third SEIR Model} \left\{ \begin{array}{l} S = 331,883,985 \\ E = 0 \\ I = 1 \\ R = 0 \end{array} \right.$$

9.3.2 Parameters of the Model

The essential parameters of the model remain unchanged, this is due to that in this model, we only introduced new parameters to represent the behavior of testing. The essential parameters are the infectious rate, the recovery rate, the incubation rate, the number of contacts for before and during the social distancing policy. They are same parameters mentioned in the model 2.

$$\text{Number for Coefficients of the SEIR Model} \left\{ \begin{array}{l} \beta = 0.21 \\ \gamma = 0.81 \\ \sigma = 1/5.2 \\ r = 18 \\ r2 = 7 \\ k = 8 \\ k2 = 2 \end{array} \right.$$

There is a huge change in the parameters of the model. Despite we still have the essentials parameters unchanged, there are new parameters we introduced to the third model. From the Coronavirus Pandemic Data, we did our own calculation to extract the accuracy of the COVID-19 testing, which for the test performed rate and test positive rate, we have 0.0496% (0.000496) and 0.0064% (0.000064) respectively.

$$\text{From the data} \left\{ \begin{array}{l} \text{test performed rate} = 0.0496\%(0.000496) \\ \text{test positive rate} = 0.0064\%(0.000064) \end{array} \right.$$

Then we used these two numbers to obtain the test behavior of each group. For the susceptible, we have its test performed rate as its distribution (0.8) times the COVID-19 test performed rate, then the non-tested rate is 1 minus the test performed rate.

$$\text{Susceptible} \left\{ \begin{array}{l} t = 0.8 * 0.000496 = 0.0003968 \\ nt = 1 - t = 0.9996032 \end{array} \right.$$

Because we have the behavior of testing for the susceptible as tested individuals and non-tested individuals, we want to introduce new parameters of number of contacts for the susceptible. We want to introduce rt and kt for model 3, rt is to the susceptible who are tested and for the non-tested susceptible, we use r , then we have kt for the tested susceptible and k for the non-tested susceptible. The difference between rt and kt is the same idea between r and k , kt is less than rt because kt is for the exposed individuals, they are less isolated than the infectious, they have a lower number of contacts.

$$\text{number of contacts for tested and non-tested individuals} \left\{ \begin{array}{l} rt = 16, \text{ where } rt < r \\ kt = 6, \text{ where } kt < k \end{array} \right.$$

For the infectious, since we assume their results are always positive, the test positive rate is its distribution (0.1) times the COVID-19 test performed rate, and the non-tested rate is 1 minus the test positive rate.

$$\text{Infectious} \left\{ \begin{array}{l} tp = 0.1 * 0.000496 = 0.0000496 \\ nt = 1 - tp = 0.9999504 \end{array} \right.$$

For the exposed population, we have the test rate, test positive rate, test negative rate and non-test rate. For the test rate, we have its distribution (0.1) times the COVID-19 test performed rate. For the test positive rate, we use the test positive rate obtained from the Coronavirus Pandemic Data, 0.000064, to get the positive test rate for the exposed by solving the equation. The idea of the equation is distribution of having test positive rate = probability of the infectious positive result * the proportion of the infectious tested + probability of the exposed positive result * the proportion of the exposed tested. Then we multiply the result with the test performed rate of COVID-19, then we have the test positive rate for the exposed as 0.000014384. Then the test negative rate and non-tested rate are obtained by the fundamental of equations.

$$\text{Exposed} \begin{cases} t = 0.1 * 0.000496 = 0.0000496 \\ tn = (1 - 0.29) * 0.0000496 = 0.000035216 \\ tp = 1 * 0.0000496 + E_{tp} * 0.0000496 = 0.000064 \rightarrow 0.29 * 0.0000496 = 0.000014384 \\ nt = 1 - tn - tp = 0.9999504 \end{cases}$$

The definitions of the abbreviations of the test rates are shown below:

$$\begin{cases} t = \text{tested rate} \\ nt = \text{non-tested rate} \\ tp = \text{test positive rate} \\ tn = \text{test negative rate} \end{cases}$$

Since we made an assumption for the isolation of the exposed and infectious population, we assumed their numbers of contacts, r , to be 1 and 2, which 1 is to the tested positive individuals because they are infected and completely isolated by the laws and authorities, and 2 is to the tested negative individuals because they are aware of being infected, they self-isolated themselves.

$$\text{isolation} \begin{cases} 1, \text{Tested positive individuals} \\ 2, \text{Tested negative individuals} \end{cases}$$

Because we are building this model based on the model 2, we are including the social distancing policy in model 3 as well. To achieve having social distancing policy, the same idea as in model 2 applies here: we had less number of contacts representing the change of contacts for the social distancing. In model 3, we have several parameters representing the number of contacts, they are r , rt , k and kt . Despite isolation falls into the category of number of contacts, we remained them unchanged because of the fact that once the individuals are infected or exposed, they are isolated whether there is social distancing or not. During the social distancing policy, for the change of number of contacts, we made them to be $r2$, $rt2$, $k2$ and $kt2$. The cases below showed the characteristic of the parameters.

$$\text{Number of Contacts during Social Distancing} \begin{cases} r2, \text{ where } r2 < r \\ rt2, \text{ where } rt2 < rt \\ k2, \text{ where } k2 < k \\ kt2, \text{ where } kt2 < kt \end{cases}$$

And their values during the social distancing policy are:

$$\text{inputs for number of Contacts during Social Distancing} \begin{cases} r2 = 7 \\ rt2 = 5 \\ k2 = 2 \\ kt2 = 1 \end{cases}$$

9.4 The Model

For the system of differential equations, the idea is to multiply the test behavior accordingly into the equations from model 2. Similar to model 2, we remained the derivative of R and I unchanged because in model 3, the only changes we made only happens in the change of the exposed and susceptible. Therefore the derivative of R and I are:

$$\frac{dR}{dt} = \gamma I$$

$$\frac{dI}{dt} = \sigma E - \gamma I$$

For the derivative of S and E, we displayed them with three color parts for better understanding. The red highlighted part is the isolation regarding the number of contacts for the infectious with positive tested result, exposed with positive tested result and exposed with negative tested result. The inputs for the isolation are displayed in the parameter section. The blue highlighted part is for the parameter of number of contact regarding the non-tested individuals. In this case we have rt is smaller than r and kt is smaller than k , the idea is rt and kt is to the tested susceptible population, they are self-isolated and more aware of contacting other people, so their numbers of contacts are less than the non-tested susceptible. The pink highlighted part is the behavior of testing we conducted in the parameter section, we added them accordingly into the equations matching the group of individuals. Overall, the system of differential equations now become:

$$\begin{aligned} \frac{dE}{dt} = & 1 * \beta * I_{tp} * \frac{S}{N} * I \\ & + (S_t * rt + S_{nt} * r) * \beta * I_{nt} * \frac{S}{N} * I \\ & + 1 * \beta * E_{tp} * \frac{S}{N} * E \\ & + 2 * \beta * E_{tn} * \frac{S}{N} * E \\ & + (S_t * kt + S_{nt} * k) * \beta * E_{nt} * \frac{S}{N} * E \\ & - \sigma * E \end{aligned} \tag{1}$$

$$\begin{aligned}
\frac{dS}{dt} = & -1 * \beta * I_{tp} * \frac{S}{N} * I \\
& -(S_t * rt + S_{nt} * r) * \beta * I_{nt} * \frac{S}{N} * I \\
& -1 * \beta * E_{tp} * \frac{S}{N} * E \\
& -2 * \beta * E_{tn} * \frac{S}{N} * E \\
& -(S_t * kt + S_{nt} * k) * \beta * E_{nt} * \frac{S}{N} * E
\end{aligned} \tag{2}$$

For the social distancing policy, we mentioned we changed the number of contacts of the equations are declared them as $r2$, $rt2$, $k2$ and $kt2$. Therefore, during the social distancing policy, the system of equations are:

$$\begin{aligned}
\frac{dE}{dt} = & 1 * \beta * I_{tp} * \frac{S}{N} * I \\
& +(S_t * rt2 + S_{nt} * r2) * \beta * I_{nt} * \frac{S}{N} * I \\
& +1 * \beta * E_{tp} * \frac{S}{N} * E \\
& +2 * \beta * E_{tn} * \frac{S}{N} * E \\
& +(S_t * kt2 + S_{nt} * k2) * \beta * E_{nt} * \frac{S}{N} * E \\
& -\sigma * E
\end{aligned} \tag{3}$$

$$\begin{aligned}
\frac{dS}{dt} = & -1 * \beta * I_{tp} * \frac{S}{N} * I \\
& -(S_t * rt2 + S_{nt} * r2) * \beta * I_{nt} * \frac{S}{N} * I \\
& -1 * \beta * E_{tp} * \frac{S}{N} * E \\
& -2 * \beta * E_{tn} * \frac{S}{N} * E \\
& -(S_t * kt2 + S_{nt} * k2) * \beta * E_{nt} * \frac{S}{N} * E
\end{aligned} \tag{4}$$

9.5 Results

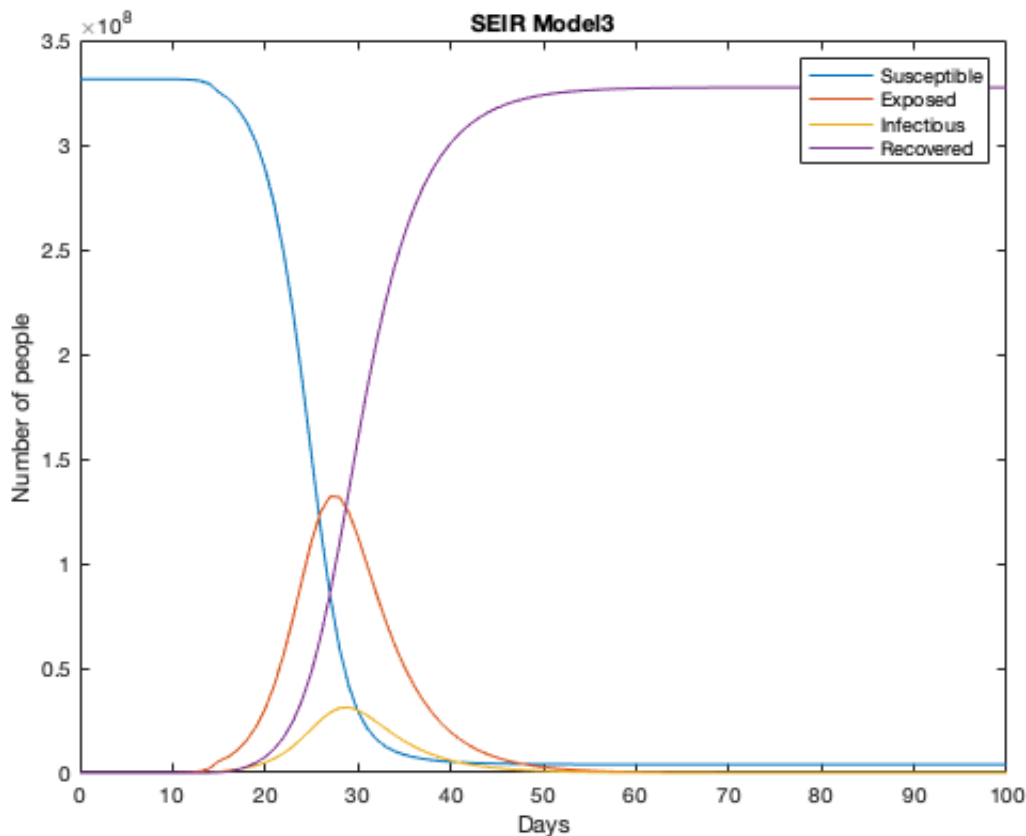


Figure 6: Visualization of the spread of COVID-19 with Behavior of Testing and Quarantine Policy.

From the plot shown above, despite there is not much difference between it and the last plot, there are still points that are worthy to discuss regarding the change of the model and the impact of adding behavior of testing made. Overall about the model, the trending of the model did not change, where the line of susceptible drops at around day 20, as well as the increase of the recovered. Then the change of exposed and the change of recovered and infectious are still sharing the same characteristic as the last plot where the change of exposed is still the highest one. At the end, the overall pandemic ends at around day 80.

Surprisingly, the overall trend of the model has not changed with adding the behavior of testing into the model. However, the numbers regarding the model reflected the changes. From the second model, the change of the susceptible starts at day 20, with the number as 327,191,048 people. However, for the third model, at day 20, the number of people for the susceptible is 327,181,154, which is 9,894 less people than the second model. The way we can interpret this change is that having behavior of testing in the model detects more cases of exposed and infectious, therefore the number of susceptible drops. With more tested cases, it speeds up ending the pandemic. This point can be observed from the recovered data, where at day 60, the day that both plot approximately showed the end day of the pandemic, the number of recovered from the second model is 326,129,824 and from the third model,

the number is 326,136,781. The difference between the two recovered number is 6,954, which means the third model does slightly better than the second model with more recovered cases and pushes an earlier end date.

Additionally, the effect of adding behavior of testing can also be observed from the numbers. The maximum numbers of infectious for the second and third model are 31,061,924 and 31,065,660, which the third model has 3,736 more cases of infectious than the second model does. Same phenomena in the exposed population, the third model has 2,612 more exposed cases than the second model does. These can be interpreted as the effect of the behavior of testing, because of there is testing, there are more detected cases of the coronavirus.

9.6 Conclusion

Overall, this section continues our model building from the second SEIR model, we added behavior of testing into the third model, made some assumptions regarding the testing and calculated test rates for each member of the population. From the plot of the third model, despite the changes of the third model were not obvious, the numbers of changes gave a detailed look of the model and showcased the differences between it and the second model. Even though the numbers of changes are not large, they still make a difference in the pandemic and reveal the efficiency of the behavior of testing. Therefore, having testing is also one of the strategies to help slowing and ending the spread of COVID-19 as it shows a sign of improvement.

10 Conclusion

In this report, we decomposed the SEIR model into SI and SIR model to learn about the characteristics and the fundamentals of the model. Then we built three models to fit with the current pandemic, COVID-19. The first model we used the basic SEIR model to obtained the changes of the pandemic, and displayed the general trend of COVID-19. The second model we went deeper as we added social distancing policy and we made the exposed population can carry and transmit the virus. For the third model, we added behavior of testing into the model, and then finalized our experiment of the SEIR model in COVID-19. The goal of this report is to provide an insightful look of the SEIR which we accomplished in the decomposition of SEIR model and model 1. Then the three models showcased their fits to the COVID-19 and revealed that social distancing policy and behavior of testing are the promising strategies to slow and end the pandemic. Finally, this report showed that the spread of COVID-19 can be predicted with a well-formed mathematical model and the SEIR model is one of them.

11 Recommendations

Despite in this paper, we were able to modified the SEIR model and generated the final model for COVID-19, there are many places where we could improve on such that we could be more precise with the rates for behavior of testing. We assumed that the test results for the susceptible and infectious are always negative and positive, where in reality the number

is close to 0 and 1, but not exactly 0 and 1. We could also be more precise with the number of contacts, we only took the number based on one study of the previous pandemic called H7N2. We could look at more articles and get a more promising number for the contacts. Overall, the SEIR model is designed to model infectious diseases, it can be modified better to provide a more accurate prediction of the spread of COVID-19. There are many behaviors we did not take account into the model but are significant to the spread of COVID-19, such that the limitation of traveling and wearing masks. We recommend adding more parameters and factors regarding the pandemic to get a better model in the future.

References

- [ACP, 2020] ACP (2020). Median estimated incubation period of covid-19 is five days. <https://www.acpjournals.org/aim/fullarticle/2762808/incubation-period-coronavirus-disease-2019-covid-19-from-publicly-reported>. Accessed on 2020-05-06.
- [Atlantic, 2020] Atlantic, T. (2020). The covid tracking project. <https://covidtracking.com/api>. Accessed on 2020-05-06.
- [Beusekorn, 2020] Beusekorn, M. V. (2020). Researchers report 21% covid-19 co-infection rate. <https://www.cidrap.umn.edu/news-perspective/2020/04/researchers-report-21-covid-19-co-infection-rate>. Accessed on 2020-05-06.
- [CDC, 2020] CDC (2020). Social distancing. <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html>. Accessed on 2020-05-06.
- [Census, 2019] Census, U. S. (2019). U.s. census bureau quick facts. www.census.gov/quickfacts/fact/table/US/PST045219. Accessed on 2020-05-06.
- [Eames, 2010] Eames, K.T., W. C. T. K. (2010). Assessing the role of contact tracing in a suspected h7n2 influenza a outbreak in humans in wales. <https://doi.org/10.1186/1471-2334-10-141>. Accessed on 2020-05-06.
- [Fracassa, 2020] Fracassa, D. (2020). San francisco to open 3 new covid-19 testing locations this week; oakland adds one. <https://www.sfchronicle.com/bayarea/article/San-Francisco-to-open-3-new-COVID-19-testing-15278019.php>. Accessed on 2020-06-02.
- [IDM, 2019] IDM (2019). Seir and seirs models. <https://www.idmod.org/docs/hiv/model-seir.html>. Accessed on 2020-05-06.
- [WebMD, 2020] WebMD (2020). Coronavirus incubation period. <https://www.webmd.com/lung/coronavirus-incubation-period>. Accessed on 2020-05-06.
- [Wikipedia, 2017] Wikipedia (2017). Compartmental models in epidemiology. https://en.wikipedia.org/wiki/compartmental_models_in_epidemiology. Accessed on 2020-05-06.
- [Wikipedia, 2020] Wikipedia (2020). Social distancing. https://en.wikipedia.org/wiki/Social_distancingHistory. Accessed on 2020-05-06.