

Predictive Research on COVID-19 using the Compartmental Model

Haotian Guo

*Department of Statistics and Operations Research, University of North Carolina at Chapel Hill,
27514, Chapel Hill, NC, U.S.A.*

Keywords: Compartmental Model, COVID-19, SIR Model, Epidemiology.

Abstract: Compartmental models appeared in our field of vision in the early 20th century, with the contributions mainly of Ronald Ross, Hilda Hudson, David Kendall and other scientists. These models turned out to be one of the most important techniques to model real world issues and to provide most of the answers that people seek. In epidemiology, in particular, these are the essential modeling techniques to the modeling of infectious diseases. Compartmental models have helped epidemiologists solve problems like the initial spread of the disease, the number of people infected, and the level of risk the disease would bring to people. In the early 2019, a highly infectious kind of pneumonia, later named COVID-19 by the World Health Organization, was firstly discovered in China, and later the doctors and epidemiologists would find out that this disease was actually due to a new kind of virus that had never been studied or seen before. In study this “new” virus, epidemiologists revealed to us how important and effective compartmental models can be even when dealing with a virus that scientists did not have full knowledge on its properties. For this project, several models were also developed with the toll of compartmental model based on the data extracted from the website of the World Health Organization of the spread of COVID-19 in countries like Vietnam, Spain, and the United States of America. The goal was trying to do was to simulate the spread of the data and predict the trend of the spread in such countries. Based on the analysis, it's concluded that COVID may be still prevalent in countries like these for a very long period of time, and because of that, there is a possibility that COVID may never end worldwide.

1 INTRODUCTION

Compartmental models have greatly affected the world of epidemiology and scientists of other aspects, and even people like us since they were first invented and used in the 1910s. Compartmental models, as indicated by their name, break the issue of one's interests into different compartments and then evaluate and analyze people or other elements in different compartments and how do they change between compartments.

Epidemiologists would use ODEs (ordinary differential equations) and stochastic models and framework as a platform when trying to analyze diseases with compartmental models. The advantages of ODEs would be that the process of solving equations would be faster, and it would give a deterministic solution to the equations scientists found. Then we can use these differential equations to plot graphs that would help visualize the models. On the other hand, the platform of stochastic

framework is more suited for analysis of real-world problems. Stochastic, which means random, would assume that everything happens randomly following current incident. For example, if there are three states, A, B, and C, that a person can be in an event and he's currently in state A, then it's completely random whether he would be in state B, state C, or remain in state A in the next time slot, and the choices would not be affected by the states he had been in in the previous state, which is also known as the Markov property (Alshomrani et al., 2021). ODEs would rather assume that things would happen according to the will of scientists.

With the outbreak of COVID-19 and the pressing demand for analysis of such disease because of the damage it brought to the world, it is realized that we can also analyze COVID-19 using SIR model, although the pathogenic mechanism remained unclear to scientists. That's one of the many advantages of such a model, which is that we can analyze epidemic with just the presence of the data of the spread of the disease. Now that COVID-19 has

been affecting us for more than a year, the question that most scientists and general public would want answers for is how long such epidemic would persist and more radically. Would it ever end? So several SIR models were created in order to find an answer to this question using ordinary differential equations to simulate the progress of COVID-19 in several countries with data from the website of the World Health Organization, and a prediction of the trend of development of the disease was finally successfully made.

Thus, the knowledge of compartmental models and SIR model in particular provided a useful tool to analyze the disease and predict number of people would be infected and how long the disease would last. However, it is unsettling that the results tended to support the idea that COVID-19 may not end worldwide when predicting using the SIR model (Merchant, 2020). Such prediction is extremely important in that epidemiologists and health scientists could focus on how to deal with this disease in a long-term fashion, and not as an emergent outbreak anymore. Also, general public would need to realize that COVID would eventually just be like flu and common cold, and they need to take action like vaccinate and wear masks as well in order to cope with this disease.

2 RESULTS AND ANALYSIS

Fig. 1 below showed what a typical SIR model would look like. The blue, red, and green lines represent susceptible (S), infected (I), and recovered (R) individuals as time evolves. In the short term, we would assume that $S+I+R$ would equal a fixed number, N , the total population size. In other words, it means that the whole population in a given area would be divided into the three groups as identified. However, when it comes to longer terms like several months or years, it would not be so accurate to assume that these are the only three groups that existed. For example, in the case of COVID-19, millions of people died after infection. Although it is rare for people to be infected again and thus susceptible after recovery, several people did show signs of infection after recovery, and in this case it might be better to use SIS model. But for purposes of simplicity, most epidemiologists assume for a SIR model at the initial stages of research.

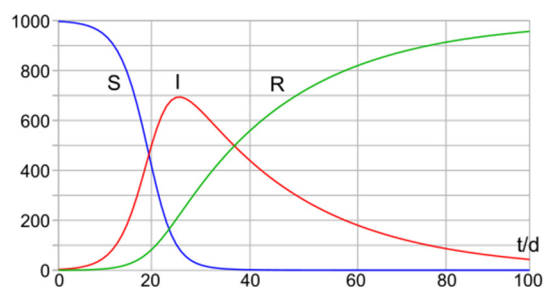


Figure 1: SIR Model.

As we can see in the diagram, the red line, or the infected individuals, has a right skewed distribution, which means that the disease can be highly infectious because people infect quickly at first, but it takes time to recover, so the infected amount would decrease with a smaller rate than the increasing rate. And as assumed by the model, once infected individuals are recovered, they would not be susceptible to the disease again, so when no one is susceptible to the disease, the epidemic or pandemic is considered to be finished (Bailey, 1975). A typical example would be the spread of COVID-19 in China, as shown below (Fig. 2). It is easy to see that by the time 100, which roughly corresponds to 220 days after the first case, the red curve shows no sign of increasing anymore and roughly no more infected individuals present. It is then safe to say that COVID-19, in a short period of time, would not surge in China again (Yang et al., 2020).

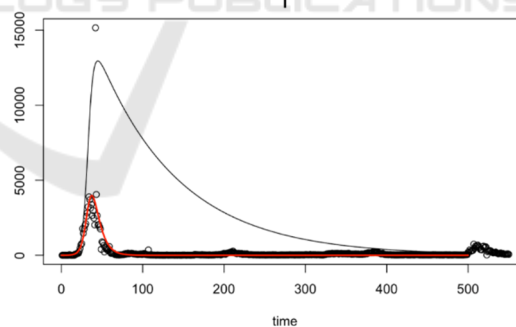


Figure 2: COVID-19 Spread in China.

However, China might be one of the few cases that would show a sign of ending COVID-19. For example, Fig. 3, 4, and 5 below show the change of infected population over time in the United States of America, Spain, and Vietnam respectively. In the case of the United States, we can see that there are multiple peaks in the graph, which would correspond to multiple SIR models. However, if we split the full graph into several models, we would discover that each SIR model does not indicate a full recovery. In

other words, the disease did not quench before the next breakout happened. In such countries, they need to first achieve the final SIR and then finally end the disease, and that would cost even years to come. For Spain, it resembles the case of the US, so one particular peak was cut out to be analyzed. In this case, it shows a left skewed distribution, and also, if we focus on the right tail, we can see that it has not achieved a flat tail, meaning that it also lacks the stability required by the SIR model to consider the pandemic to be finished. In the case of Vietnam, at first it seemed like it showed a complete peak and showed signs of dying down, but it suddenly had more cases surging out. So, we need to make sure that we have a really flat right tail, which means the newly infected individuals become mostly zero for a long time before saying that this epidemic ended in this area. These three countries were chosen in a way to represent the three of the five continents on earth, America, Europe, and Asia. And more countries displayed similar traits of these countries, which would finally lead to the conclusion that it might be years before COVID-19 end worldwide, and it is even possible that it may never end.

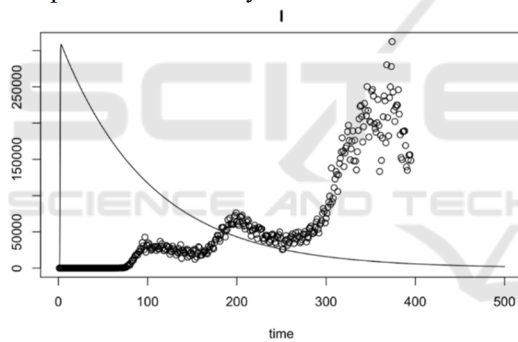


Figure 3: COVID-19 Spread in the United States of America.

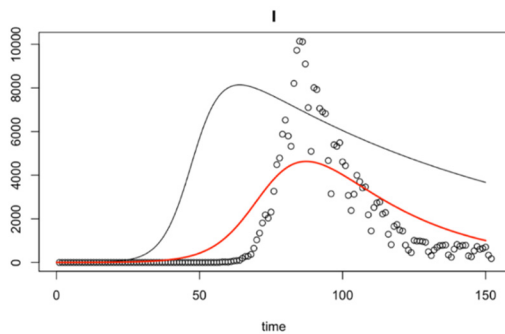


Figure 4: COVID-19 Spread in Spain.

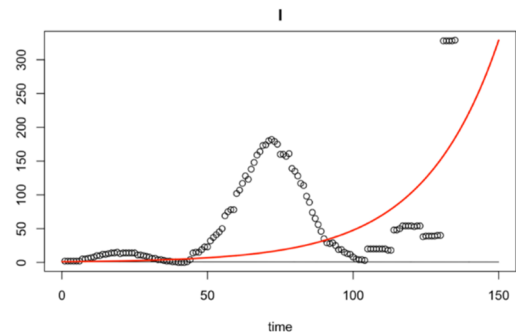


Figure 5: COVID-19 Spread in Vietnam.

3 CONCLUSION

Based on the analysis of these models, it has also been found that the R_0 is less than one for all cases analyzed, which means that the virus should not spread as quickly as before and such value does not signify an outbreak of the epidemic, but it certainly was not the case shown by these models. So alternatively it can be concluded that a short period of a small R_0 achieved is not full proof that the disease would end permanently. We also need to see a clear sign of dying out for a long period of time before having such a conclusion. Thus, as the statistics have shown, for countries like Vietnam, Spain, and the United States of America, we don't have proof that COVID would subside in a short time. Also, it is clear that even if there was only one country that still had COVID going on, it would be a disaster to the whole world because of the high infectivity of COVID. Actually that was how the breakout first happened. After its first emergence in Wuhan, China, although China had done great effort to quarantine people and implement lockdown, the virus failed to be locked down and quickly spread to other countries. There is no way to isolate virus in modern world. So, in our case where most countries or even continents showed signs that COVID would still affect people there, it would be months and even years to reach the state where only one or few countries are still affected by COVID, and even if we reached that state it would still cost great effort to eliminate it. This is where we made such a conclusion that epidemiologists, scientists, and common people need to be aware of the possibility that we may need to coexist with this virus in the end.

ACKNOWLEDGMENT

Through the writing of this paper I received a great deal of help and assistance. I would first like to thank Dr. Otto X. Cordero for being my supervisor and mentor. Your expertise had given me great insight on how to approach problems of my interest and I could never finish my project without your methodology. I would also like to acknowledge my tutor Shiqi Wen and Mandy for their excellent guidance throughout my studies. With the tools they provided I was able to conduct my research in a more timely and effective manner.

REFERENCES

- Alshomrani, A.S., Ullah, M.Z. & Baleanu, D. (2021). Caputo SIR model for COVID-19 under optimized fractional order. *Adv Differ Equ* 2021, 185. <https://doi.org/10.1186/s13662-021-03345-5>.
- Bailey, N. (1975). *The mathematical theory of infectious diseases and its applications* (2nd ed.). London: Griffin.
- Kröger, M. & Schlickeiser R. (2020). Analytical solution of the SIR-model for the temporal evolution of epidemics. Part A: time-independent reproduction factor. *J. Phys. A: Math. Theor.* 53 505601.
- Merchant, H. (2020). CoViD-19 may not end as predicted by the SIR model. *The BMJ*, 369, [m1567/rr]. <https://www.bmj.com/content/369/bmj.m1567/rr>.
- Ronald, R. (1916). An application of the theory of probabilities to the study of a priori pathometry.—Part I. *Proc. R. Soc. Lond. A* 92: 204-230. <http://doi.org/10.1098/rspa.1916.0007>.
- Yang, W., Zhang, D., Peng, L., Zhuge, C. & Hong, L. (2020). Rational evaluation of various epidemic models base on the COVID-19 data of China. [arXiv:2003.05666 \[q-bio.PE\]](https://arxiv.org/abs/2003.05666).