

Identifying COVID-19 spread with a SEIRD Model

Stefania Patsou, *Postgraduate Student, DS2190014*

Abstract—COVID-19, also called SARS-CoV-2, is a novel coronavirus that has been spread to all over the world, causing a pandemic. Its epicentre was in Hubei Province of People's Republic of China where also the morbidity and mortality rates are high in contrast with the remaining Asia, Europe and North America [1]. Pneumonia is an initial clinical sign that someone is suffering from COVID-19. Reports also claim that gastrointestinal symptoms and asymptomatic infections describe this disease, mostly occurred within children [1]. In this research I will mostly emphasis on the epidemiology of the COVID-19 virus, by defining a SEIRD model [2], [3]. Then, I will use the COPASI tool [4] in order to instantiate my model. After describing the model in detail and further calibrating it, I was going to use the data included in the COVID-19 data repository of Johns Hopkins University [5] in order to test whether the simulation behaves similarly with the time-series data collected for Greece and Italy, for the same time period. Then, I would try to predict the course of the pandemic for Greece and Italy and compare my results with the actual data. Finally, I would depict the results of my model and state some future recommendations where possible, regarding the parameter estimation and model redefinition.

Index Terms—COVID19, Epidemiology, SEIRD model, System Dynamics, Parameter estimation, COPASI implementation, Greece, Italy.

I. INTRODUCTION

CORONAVIRUSES are enveloped, positive single-stranded large RNA viruses that infect humans, but also a wide range of animals. SARS-CoV-2 apparently succeeded in making its transition from animals to humans on the Huanan seafood market in Wuhan, China. Reports so far estimate a mean incubation period of seven days and a median incubation period of three days. The population of the asymptomatic people has not yet been assessed [1].

For this research, I used the model that performs Ordinary Differential Equations (ODEs) with deterministic compartmental models. The deterministic compartmental models, based on a system of initial value ODEs, have been studied by W.O. Kermack and A. G. MacKendrick [6], proposing the basic Susceptible-Infected-Removed (SIR) model. Since then, many extensions have been developed to study the epidemics of different infectious diseases. Some models take other compartments into account when defining the population, like the Exposed, the Vaccinated etc. Since I believe that relevant information concerns not only infected and exposed but also Recovered and Dead populations, I choose to split removed population into Recovered and Dead, obtaining the SEIRD model [2], [3].

According to the above, COPASI tool has been used to define the SEIRD model. The system introduced is then going to be used as a stepping stone prediction mechanism for future

outbreaks of the COVID-19 pandemic. This can be utilized by other researchers to determine a more robust model, that its outcome takes into account many compartments and properties that they have yet not been defined in this study. Communities could take into their consideration this model, if it was completed, further down the line, so that to estimate their COVID-19 epidemiology state and act accordingly, smoothening the infected cases curve.

A. Epidemiology

1) *Initial Image*: The epidemiology curve can roughly be divided into three phases.

- i. The first phase is characterized by the first cases found in the Huanan seafood market in December 2019 to the emergence of new cases outside Wuhan by January 13, 2020, a total of 41 cases were confirmed. Person-to-person transmission had occurred by close contact.
- ii. The second phase initiated on January 13, marked by rapid expansion and spread of the virus within hospitals and by family transmission. In this phase the epidemic spread from Wuhan to other areas. Already by January 23, 29 provinces, plus six foreign countries, had reported a total of 846 confirmed cases, an approximately 20-fold increase from the first phase. Meanwhile, Wuhan city implemented a 'lock-down' (i.e., shutting down all movement within and out of the city). Unfortunately, this period coincided with the traditional mass movement of people, a form of 'home-coming', before Chinese New Year and thus more than 5 million people had already left Wuhan.
- iii. The third phase started on January 26, which is marked by the rapid increase of cluster cases that accounted for 50–80% of all confirmed cases in Beijing, Shanghai, Jiangsu, and Shandong. On January 30, the number increased 240-fold, reaching 9826 confirmed cases, and the WHO declared this epidemic a PHEIC. February 3 seems to be a tipping point of the epidemic, as the confirmed cases outside Hubei began to decline.

[7]

2) *Generally in Europe*: In Europe, the first case of COVID-19 was recorded in the beginning of January 2020. During January 2020 to July 2020, the highest number of cases were recorded in April (850,681). A further increase of cases occurred in August (752,209), with highest number of cases were recorded in November (7,360,373). As of week 5 of 2021, Eastern Europe has been the most severely hit region with 9,856,455 cases and 209,230 deaths, followed by 8,369,494 cases and 193,963 deaths in Western Europe. The second highest number of deaths were reported in Southern Europe (197,392).

There were six countries in Europe, which accounted for 61.14% of total COVID-19 cases in Europe: Russian Federation (3,983,197, 12.69%), United Kingdom (3,959,784, 12.61%), France (3,337,048, 10.63%), Spain (2,989,085, 9.52%), Italy (2,636,738, 8.4%), and Germany (2,288,545, 7.29%). The six countries also accounted for 65.61% of COVID-19 related deaths in Europe: Russian Federation (77,068, 10.45%), United Kingdom (112,798, 15.29%), France (78,965, 10.7%), Spain (62,295, 8.44%), Italy (91,273, 12.37%), and Germany (61,675, 8.36%) [8].

II. METHODOLOGY

The COVID-19 strand depends on many factors, specifically the susceptible population, the exposed, the infected, the recovered, the dead, the vaccinated, the immunes etc. Some of these factors can be defined using a simple System Dynamic (SD) SEIRD approach. This model is rather simplistic but we can use this as a stepping stone in order to further improve it down the line.

A. The Model

The SIR model proposed is consisted of three main compartments:

1. **Susceptible $S(t)$:** the population that is not infected but it can become infected. A susceptible person can become infected or remain susceptible. As the virus spreads, so is the number of susceptible individuals (surge).
2. **Infected $I(t)$:** the population that is infected. These individuals can transmit the disease to other susceptible individuals. An infected person may remain infected, recover or die. An infected individual can be symptomatic or asymptomatic but in my model, this compartment describes both.
3. **Removed $R(t)$:** the population that is removed from the disease, whether by recovery and further immunity, or death.

[6]. COVID-19 infection can be more realistically defined, if the incubation period is introduced into the system. The incubation period refers to the population that it is infected but not yet infectious [9]. Such an assumption can be incorporated in the SIR model as well by adding the exposed compartment, denoted by E. In addition, the removed population can be further divided to recovered and dead individuals.

The SIR model can now be extended and include:

1. **Exposed $E(t)$:** the population that is infected but cannot infect the susceptible individuals just yet. The exposed individual become infected after a period of time.
2. **Recovered $R(t)$ and Dead $D(t)$:** the population that is removed, however now the removal is either by recovery or death. The recovered individuals can become again susceptible to the disease or be indefinitely immune. Dead individuals are unquestionably removed from the population.

proposing a SEIRD model [2], [3].

The differential equations describing the SEIRD model are:

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

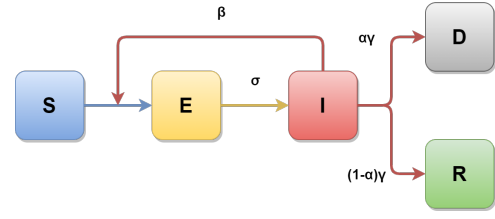


Fig. 1. SEIRD Model Diagram. The model defined in this research. It is consisted of five compartments S, E, I, R and D.

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

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

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

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

where each parameter can be described as:

1. β : the rate at which infectious people interact with others,
2. γ : the estimated duration of illness,
3. α : the infection fatality rate (IFR),
4. σ : the estimated incubation period of the disease.

Taking into account all the mentioned characteristics of the model, it can be depicted as in Fig.1. In addition, a variable N should be introduced, stating the general population.

B. Assumptions and Calibration

As the model is very simple-defined, the assumptions include the parameters of the system according to literature. For each parameter:

1. N : it is constant in time and it is defined by the equation: $N = S + E + I + R + D$. It is the population size of a given country. Population of Greece is 10,423,054 [10] and Italy's 60,461,826 [11].
2. R_0 : it is the average number of secondary infections produced by an infectious individual in a population where everyone is considered susceptible [12]. The mean and median of this variable is constant and it is approximately 2.38 for Greece [13] and 3.29 for Italy [14].
3. β : $\beta = R_0 \gamma$. Approximately 0.38 for Greece and 0.329 for Italy.
4. γ : $\frac{1}{6}$ for Greece [13] and $\frac{1}{10}$ for Italy [15].
5. α : 0.03 for Greece [16] and 0.057 for Italy [17].
6. σ : $\frac{1}{5}$ for Greece [13] and $\frac{1}{4}$ for Italy [15].
7. $R(0)$: 0 in the beginning.
8. $D(0)$: 0 in the beginning.
9. $S(0)$: $N - I(0) - E(0) - R(0) - D(0)$.
10. $E(0)$: 1 for simplicity reasons.
11. $I(0)$: 0 in the beginning, for simplicity reasons.
12. Errors: no errors are introduced in the system.
13. Birth and death rates: are constant and therefore removed from the system.

Keeping in mind that the data should start from when the pandemic started on each country, I would have used time-series data that start at 24 February 2020 and end at 26 April 2020 for Greece and Italy. I would have used these periods so that to minimize any distortions before the pandemic started on each country. In addition, if the data includes lockdown phases, then the model does not behave as intended, meaning that the plots would contain mostly outliers in these periods.

III. RESULTS

Results could not be obtained within the time limit of the assignment. The only thing, regarding the actual COPASI implementation that was completed, was the definition of the model, the ODEs and the fixed values. Also, a plot was produced regarding the time source that it needs to be further checked, making it unacceptable in the context of the assignment.

IV. DISCUSSION

The SEIRD approach does not offer an accurate model that describes the COVID-19 pandemic in Greece and Italy. To achieve a better model, it will be necessary to extend the model, taking into account the asymptomatic infected population rather than summing all the infected, symptomatic and asymptomatic to the same group/compartment. In addition, vaccinated population should also be defined in the system, as well as temporary immunity. Another aspect that needs to be further searched is the effect of birth rates and death rates, because my model makes the assumption that these are static, non changeable. A further improvement would be to introduce into the system the hospitalized population in order to weight the recovery time more.

The system produced allows us to define credible models that can be further expanded and validated. The initial parametrization is based on the literature, as mentioned. The containment compartment should also be introduced in the SEIRD model so that to depict better the asymptomatic cases and to smoothen the curve of the infected cases.

The purpose of this study is to delve into the parametrization of the COVID-19 pandemic. For this reason, it is clear that a more exhaustive analysis is needed, taking into account more countries, a more robust exploration of the parameter space and a selected time-specific data space. This could allow us to better infer the values of the model.

The data used are strictly derived with the time-period in mind, specifically from 24th of February to 26th of April, if fitting and prediction would be completed within the time limits of the assignment. This can introduce error in the system if the pandemic changes course, as the data are static. Keeping that in mind, using real-time data to define the system's properties is another thing that it should be considered for future enhancement. With this improvement, the model would provide better previsions.

V. CONCLUSION

In this study, I have proposed a SEIRD model and used the COPASI tool for the analysis of the COVID-19 outbreak

evolution in Greece and Italy, for the time period between 24th of February 2020 and 26th of April 2020. In this model, I have taken into account the literature in order to define the properties of the system. I would use the data from the Johns Hopkins University in order to validate my model, if the COPASI implementation was completed. I did not take into account the temporary immunity and the vaccinations, only the rate of mortality.

I would utilize the model in order to predict the COVID-19 pandemic for the next week, aka 7 days, because my model is not dynamic, meaning that the system would produce many errors for the predictions. To validate my model, I would collect the actual data and compared them with my predictions. The model I proposed seems to miss the adaptation of the epidemic stages, due to lack of important properties used to describe the model.

Further study, analysis and re-calibration should be done regarding the model's properties, compartments and differential equations. Concerning the epidemic evolution of Greece and Italy, which is very different regarding the population size, it is of crucial meaning to develop a model which will carry out very accurate predictions so as to prevent possible future outbreaks.

VI. SUPPLEMENTARY MATERIAL

More information regarding the COPASI implementation can be viewed here <https://github.com/PiStefania/COVID-19-COPASI>. It includes a well-defined README section, containing all the important instructions for reproducing the results of the research.

REFERENCES

- [1] T. P. Velavan and C. G. Meyer, "The covid-19 epidemic," *Tropical Medicine & International Health*, vol. 25, no. 3, p. 278–280, 2020.
- [2] M. J. Keeling and P. Rohani, *Modeling infectious diseases in humans and animals*. Princeton University Press, 2008.
- [3] F. Brauer, "Mathematical epidemiology: Past, present, and future," *Infectious Disease Modelling*, vol. 2, no. 2, p. 113–127, 2017.
- [4] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, U. Kummer, and et al., "Copasi—a complex pathway simulator," *Bioinformatics*, vol. 22, no. 24, p. 3067–3074, 2006.
- [5] "Covid-19 data repository by the center for systems science and engineering (csse) at johns hopkins university."
- [6] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 115, no. 772, p. 700–721, 1927.
- [7] J. Sun, W.-T. He, L. Wang, A. Lai, X. Ji, X. Zhai, G. Li, M. A. Suchard, J. Tian, J. Zhou, and et al., "Covid-19: Epidemiology, evolution, and cross-disciplinary perspectives," *Trends in Molecular Medicine*, vol. 26, no. 5, p. 483–495, 2020.
- [8] M. Nanda, A. Singhal, and R. Sharma, "A paradox of best healthcare systems and large covid-19 burden: Review of 1 year of covid-19 epidemiology and public health response in europe," *SSRN Electronic Journal*, 2021.
- [9] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler, "The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: Estimation and application," *Annals of Internal Medicine*, vol. 172, no. 9, p. 577–582, 2020.
- [10] "Greece demographics."
- [11] "Italy demographics."
- [12] P. L. Delamater, E. J. Street, T. F. Leslie, Y. T. Yang, and K. H. Jacobsen, "Complexity of the basic reproduction number (r_0)," *Emerging Infectious Diseases*, vol. 25, no. 1, p. 1–4, 2019.

- [13] V. Sypsa, S. Roussos, D. Paraskevis, T. Lytras, S. Sotirios Tsiodras, and A. Hatzakis, "Modelling the sars-cov-2 first epidemic wave in greece: social contact patterns for impact assessment and an exit strategy from social distancing measures," *Emerging Infectious Diseases*, 2020.
- [14] M. Allietta, A. Allietta, and D. Rossi Sebastiano, "Covid-19 outbreak in italy: estimation of reproduction numbers over 2 months prior to phase 2," *Journal of Public Health*, 2021.
- [15] J. Chen, H. Lu, G. Melino, S. Boccia, M. Piacentini, W. Ricciardi, Y. Wang, Y. Shi, and T. Zhu, "Covid-19 infection: the china and italy perspectives," *Cell Death & Disease*, vol. 11, no. 6, 2020.
- [16] "Covid-19 pandemic in greece," Jun 2021.
- [17] G. De Natale, V. Ricciardi, G. De Luca, D. De Natale, G. Di Meglio, A. Ferragamo, V. Marchitelli, A. Piccolo, A. Scala, R. Somma, and et al., "The covid-19 infection in italy: A statistical study of an abnormally severe disease," *Journal of Clinical Medicine*, vol. 9, no. 5, p. 1564, 2020.