

Andrews University

Digital Commons @ Andrews University

Honors Theses

Undergraduate Research

3-4-2021

Modeling of COVID-19 Utilizing Various Compartmental Models to Predict Infection Rates Throughout Michigan

Colleen M. Staniszewski

Andrews University, colleens@andrews.edu

Follow this and additional works at: <https://digitalcommons.andrews.edu/honors>



Part of the [Applied Mathematics Commons](#), and the [Disease Modeling Commons](#)

Recommended Citation

Staniszewski, Colleen M., "Modeling of COVID-19 Utilizing Various Compartmental Models to Predict Infection Rates Throughout Michigan" (2021). *Honors Theses*. 258.

<https://digitalcommons.andrews.edu/honors/258>

This Honors Thesis is brought to you for free and open access by the Undergraduate Research at Digital Commons @ Andrews University. It has been accepted for inclusion in Honors Theses by an authorized administrator of Digital Commons @ Andrews University. For more information, please contact repository@andrews.edu.

J.N. Andrews Honors Program
Andrews University

HONS 497
Honors Thesis

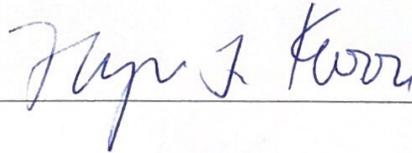
Modeling of COVID-19 Utilizing Various Compartmental Models to Predict Infection Rates
Throughout Michigan

Colleen M. Staniszewski

Date: March 4, 2021

Advisor: Dr. Hyun Kwon

Primary Advisor Signature: _____

A handwritten signature in black ink, appearing to read "Hyun J. Kwon", written over a horizontal line.

Department: Engineering

Modeling of COVID-19 Utilizing Various Compartmental Models to Predict Infection Rates Throughout Michigan (March 2021)

Colleen M. Staniszewski, Dr. Hyun Kwon

Department of Engineering, Andrews University

I. ABSTRACT

Compartmental modeling is a method of employing math to create a visual representation of a disease interacting with a select population, typically used in epidemiology analyses. This project applies compartmental modeling equations to data collected on the various aspects of COVID-19 in Michigan. Comparing current data to past predictive models, as well as the visual representations that were developed through the various compartmental modeling methods, allows an assessment of the effects of the preventative measures taken by the state, the various rates at which the infection is able to spread, as well as the potential path and spread of the virus in the future.

II. INTRODUCTION

The first reported incident of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) occurred at the end of December of 2019, in Wuhan, China. This would be the start of the COVID-19 pandemic that overtook the world in 2020 and would change day to day interactions as we know it. The severity and sudden impact of the virus on the population of the world has made being able to predict its spread rates crucial. Creating mathematical models for infection rates allows one to see how the preventative measure taken by the state positively or negatively affected the infection curve, allow comparison of previous models to present

models, and also help one to better understand the path the virus will take in the future.

II. a. COVID-19

The virus, more commonly called COVID-19 (corona virus disease 2019), is an infectious disease that typically manifests itself in the form of mild or moderate respiratory illnesses, however, will lead to more serious issues in people with other underlying medical conditions such as heart issues, diabetes, and respiratory diseases. [1] COVID-19 is spread from person to person through droplets, which means anything from as obvious as a cough or as simple as a breath can transfer the virus. The virus is spherical in shape, with branched surface proteins that latch onto healthy cells in the lungs once it has been transferred to its host. [2] Because of how quick and easy it is to transfer the virus from one person to another, various preventative measures have been taken by people everywhere in order to lessen the chance of contracting the virus in person-to-person interactions. Because of how rapidly the virus was able to spread from its origin point and how many people the virus affected, it was essential to begin characterizing the disease and the way it spread throughout a population.

II. b. COMPARTMENTAL MODELING AND EQUATIONS

A common method of predicting the spread of a disease throughout a specific population is compartmental modeling, used most frequently in the field of epidemiology. Compartmental modeling can be described broadly as the use of equations and data to create a visual representation of the different aspects, or “compartments”, of a disease’s interaction with a population. The math models are typically in the format of a standard cartesian graph on the positive axes, where x indicates time and y indicates the population. Using mathematical modeling in the context of predicting infection rates has existed since the mid-1700s, when Daniel Bernoulli attempted to model how vaccination decreased the rate of death due to smallpox. [3] Many more equations have since been derived to describe additional compartments such as quarantine and population recovery rate. These compartments are labeled with a single letter to denote their name. The three base compartments used when modeling infection rates are S, I, and R. Each of these letters indicate a compartment as well as a equation that models how that compartment changes over time in a population. “S” indicates the portion of the population that is susceptible to a disease, “I” represents the number of actively infectious individuals of a population, and R is the number of individuals that have recovered from the disease. The following (Equations 1, 2, and 3) are the equations that the compartments S, I, and R, represent, respectively:

$$\frac{dS}{dt} = \frac{-\beta IS}{N} \quad [4, \text{eq. (1)}]$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I \quad [4, \text{eq. (2)}]$$

$$\frac{dR}{dt} = \gamma I \quad [4, \text{eq. (3)}]$$

In these cases, β indicates the contact rate between humans, $1/\gamma$ is the average infectious period for the disease, and N is the total population of the group being observed. These variables are often assumed to be constants to make calculations simpler, however are often actually equations themselves.

Furthermore, introducing additional equations and compartments accounts for more aspects of data and may change the “base” SIR equations. For example, the SEIR model. The “E” compartment accounts for the number of people who are exposed to the virus, but not infectious. This changes the equation for compartment “I” to account for the number of exposed but non-infectious. The SIER model also includes the variable “ μ ”, to account for birth and death rates in a fluctuating population.

II. c. MODELING AND ANALYSIS PROGRAMS

While programs exist for the generation of compartmental models, these mathematical models can also be made in programs like COMSOL and MATLAB. With the proper programming, both of these programs are able to solve multivariable equations and plot the results on an x-y graph in the fashion of a standard compartmental model. These programs are commonly used by scientists and engineers, and the ability to mathematically model the distribution rate of something is valuable for more than just epidemiology studies. Math modeling with a particular interest in rates and interaction are valuable specifically in the chemical engineering field when one wants to study the rates at which, for example, chemical A and chemical B interact to form chemical C, or how quickly heat or cooling can be added to a chemical reactor to prevent a reaction on an industrial level from getting out of control.

II. d. OBJECTIVES AND VALUE

The primary goal of this research endeavor was to explore the various ways that COVID-19 has affected specifically the population of Michigan. This study was chosen for three main reasons. The first reason was its relevance. COVID-19 has affected the lives of almost everyone around the world, and an understanding of how the virus is interacting with a population is valuable in understanding the virus itself and how to prevent it from hurting more people. The second cause for this research was to observe and gather data specifically pertaining to the state of Michigan. While there is data available for predictions in larger populations such as China, Italy, or the U.S., more specific analysis is atypical of these studies. Observing only Michigan allows for more in-depth research considering the preventative measures taken specifically by the state, and fewer unaccountable variables that may affect the results. The final reason this project was chosen was to be able to display how mathematical modeling is a useful skill with many applications in many different fields.

III. METHODOLOGY

There were several steps in the research process. The first step can be considered the preliminary research step, which involved

finding credible sources for data collection, understanding the nature of the virus, and finding other compartmental models that had been developed analyzing the spread of COVID-19. The second step was the actual collection of data. The third step was programming MATLAB and COMSOL to complete parameter estimation tasks and mathematical modeling in order to view predictions based on data. The final step was comparing and contrasting the generated data to existing or previously predicted data.

III. a. PREVIOUSLY DEVELOPED MODELS FOR COVID-19 IN MICHIGAN

While there are a number of predictive models for COVID-19 travel rates in the United States, finding data specific to Michigan is difficult. Furthermore, models may use different predictive methods from the SIR to evaluate infection rates, making it difficult to compare models. Figure 1 displays an SIR model developed by Jathin Desan for the period of January 22nd to April 14th for 19 different states, including Michigan. [5] Desan gives a very brief and over-simplified analysis of the SIR model, stating that “the infected rate is found to be

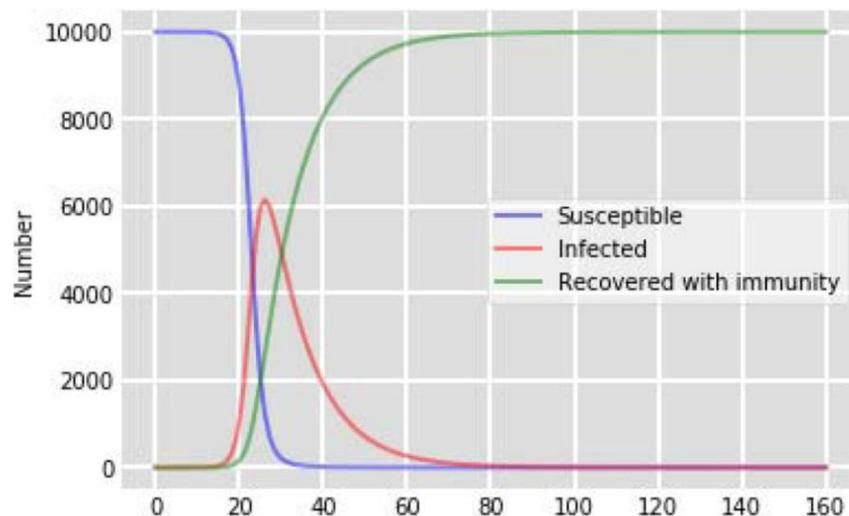


Figure 1: SIR Derived Model of Michigan, the x-axis denotes time in days [5]

moderate, and this rate also gradually decreased as days and time passed.” [5] This observation can be interpreted as, the virus is no more infectious than most viruses, and as fewer people are susceptible to the virus, fewer people will be infected with the virus. However, from the graph in Fig. 1, the numerical data that can be gathered is that in a population of 10,000,000 (assuming the graph is scaled by a factor of 1000), around 34 days into the virus, if no preventative measures are taken, we would hit a maximum of 6,000,000 actively infected cases. Allowing this to happen would destroy the economy, the hospitals would be overrun with patients, supplies would run low, and quality of care would decrease, causing tragedies in even the simplest of COVID-19 cases.

Figure 2 shows the work of Z. Wang et. al., and their application of an SIRD (Susceptible-Infected-Recovered-Deceased) model to COVID-19 in the state of Michigan. Their goal was to accurately simulate the

various characteristics of COVID-19 in Michigan, while considering “the evolving characteristics of testing, quarantine and treatment protocols” over the observed time (March 23rd to after June 1st). [6] Wang observed that the infection rates (top right graph in Figure 2) were unfortunately higher than the predicted rates should have been, considering quarantine. However, on a positive note, the recovery rate was also higher than predicted. While death rates began as higher than predicted, they eventually flattened to become as predicted. One thing to note is that in regard to infectious, Wang is assuming that all those who test positive for COVID-19 are infectious, when this is not the case.

Simply observing Fig 1 and Fig 2, it is easy to see how data can be misconstrued, misinterpreted, or unintentionally misrepresented regarding the virus. Furthermore, it is shockingly difficult to find consensus between different studies. The reasons for such difficulties will be described in further

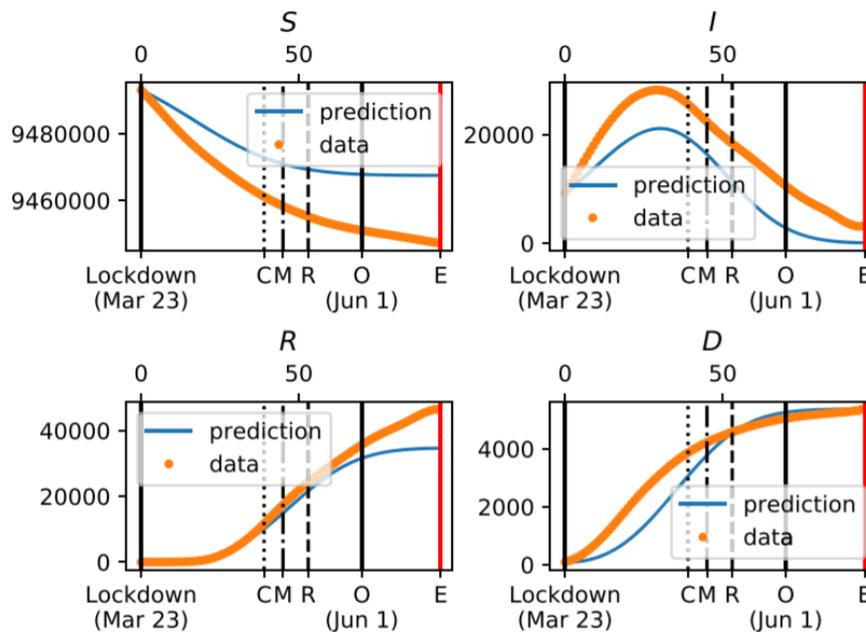


Figure 2 SIRD models compared with Michigan Data [6]

detail in the Section V., or the conclusions section of this document.

III. b. METHODS OF DATA COLLECTION

In late April to early May of 2020 data was collected on several different aspects, of COVID-19 in Michigan. The data included factors such as infectious individuals, tested individuals, deaths, and recovered individuals. Other aspects were also researched; however, the previously listed variables were the most prevalent and pertinent to the project. At the start of the research, validity was determined by comparing the numbers of multiple sites to each other. The source that supplied the numerical data regarding infectious, deaths, etc. was The COVID Tracking Project. [7]

III. c. COMSOL AND MATLAB GENERATED DATA IN COMPARISON TO ACTUAL DATA

The first step in being able to analyze data is the estimate the parameters β (contact rate) and γ ($1/\gamma$ being infectious period). COMSOL was programmed to simultaneously solve the SEIR equations as seen in equations 4-7:

$$\frac{dS}{dt} = \mu N - \mu S - \frac{\beta IS}{N} \quad [4, \text{eq. (4)}]$$

$$\frac{dE}{dt} = \frac{\beta IS}{N} - (\mu - \alpha)E \quad [4, \text{eq. (5)}]$$

$$\frac{dI}{dt} = \alpha E - (\gamma + \mu)I \quad [4, \text{eq. (6)}]$$

$$\frac{dR}{dt} = \gamma I - \mu R \quad [4, \text{eq. (7)}]$$

¹ These β and γ values do not necessarily indicate accurate infection rates or infectious period for medical analysis. The number is only accurate for a small portion of somewhat linear data, the actual

The variable μ helps to account for birth and death rates in a fluctuating population and $1/\alpha$ is equal to the approximate incubation period of the COVID-19 virus.

Overlaying the data collected of actual infected members of the Michigan population with the SEIR approximation and using trial and error to find the best fit line, parameter β was found to be 0.6×10^{-8} and parameter γ was found to be 0.0173¹.

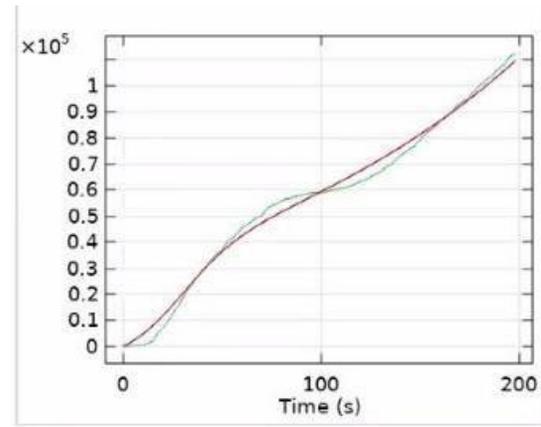


Figure 3: COMSOL graph of SEIR approximation over actual data from mlive.com [8]

This data was then programmed into MATLAB using an SIR (see equations 1-3) solver similar to the one developed by Giovanni Valentini (see Appendix) [9]. The graph of Figure 4 displays three significant lines. The blue line indicates the number of infectious individuals, that is, the number of individuals who are positive for the virus and are also able to spread the virus to infect another person. This line was gathered from data from the COVID Tracking Project. [7] The magenta line indicates the number of people who have tested positive, this data includes both infectious and non-infectious

infection rate and period are much more complicated and difficult to determine, as described in the conclusion.

cases and was gathered also from the COVID Tracking Project. [7] The final, red, line displayed is the line generated when solving for equation dI/dt (eq. 6), based on the experimental β and γ values gathered. This line is by far the most drastic looking of the three and indicates the number of infectious people over time. Note that this line does not account for exposed but not infectious persons and assumes that all exposed are infectious.

Figure 5 shows a more complicated model based on the SEIRQD (Susceptible-Exposed-Infectious-Recovered-Quarantined-Deceased) model. The equations for the SEIRQD model are as seen in equations 8-13.

$$\frac{dS}{dt} = \frac{\beta IS}{N} \quad [4, \text{eq. (8)}]$$

$$\frac{dE}{dt} = \sigma E - \frac{\beta IS}{N} \quad [4, \text{eq. (9)}]$$

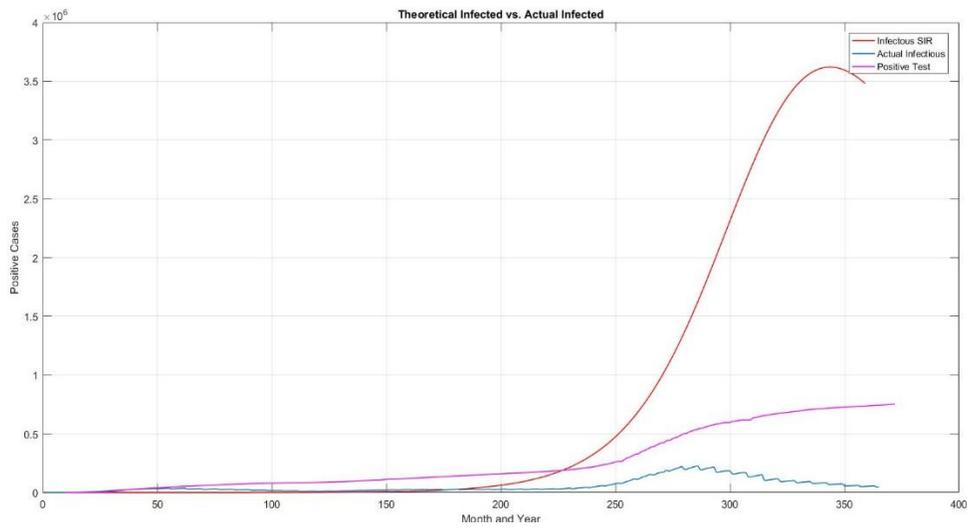


Figure 4: SIR Model for variable I in comparison to the actual number of positive cases, and actual infectious cases over the course of one year.

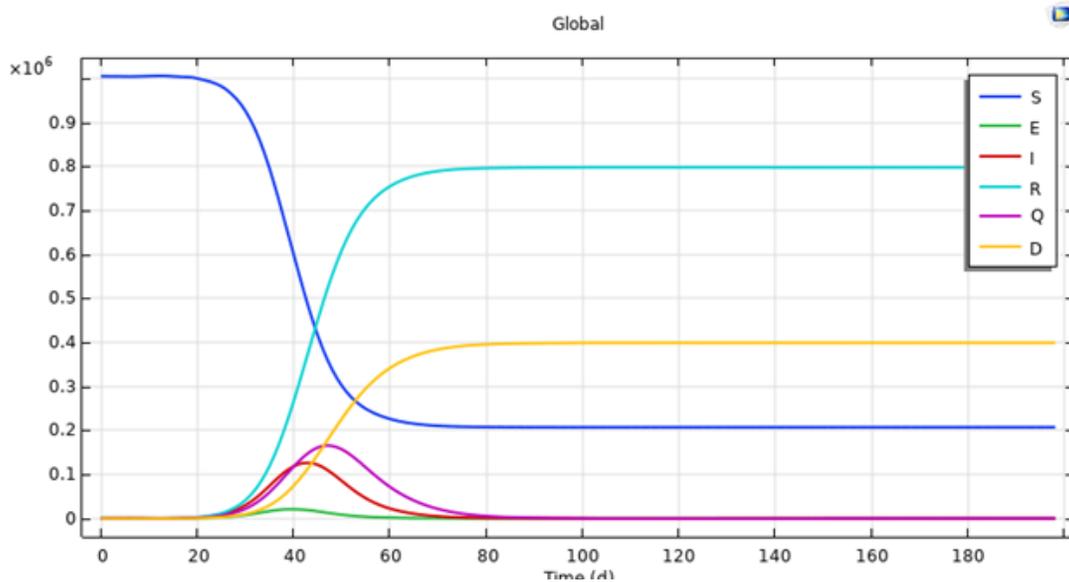


Figure 5: SEIRQD Model Generated from Equations in COMSOL

$$\frac{dI}{dt} = \gamma I - \sigma E \quad [4, \text{eq. (10)}]$$

$$\frac{dR}{dt} = \gamma I \quad [, \text{eq. (11)}]$$

$$\frac{dQ}{dt} = \lambda Q + \kappa Q \quad [10, 11, \text{eq. (12)}]$$

$$\frac{dD}{dt} = -\kappa Q \quad [10, 11, \text{eq. (13)}]$$

Using the parameters as previously found, as well as programming COMSOL to find κ and λ through the least-squares method, a graph was able to be generated for each of the variables.

This model is intended to show an ideal version of the infection rates if an “ideal” quarantine had been set in place at the early onset of the virus. However, for economic and practical reasons, a “perfect” quarantine would not actually be an ideal one by any means. Since there is always a need for frontline workers (as made doubly evident by the events of 2020) and that at a certain point the economic downfall of a population is a

death-sentence itself, there will never be such a thing as a perfect quarantine. [12]

Figure 6 shows the actual data as collected from the COVID Tracking Project over the course of one year (March 2020 - March 2021). [7] The following dates should be noted in the timeline of the graph. [13]

March 10th, 2020: The first COVID-19 cases are identified in Michigan.

March 11th, 2020: Suspension of all college classes, or movement to online classes, after spring break.

April 9th, 2020: “Stay Home, Stay Safe” Michigan Quarantine begins.

August 2020: Increased testing due to students returning to school causing rise in data.

IV. RESULTS

The most interesting graph to observe is Figure 4. The peak of the infectious line is estimated to be at 37,000,000, while the

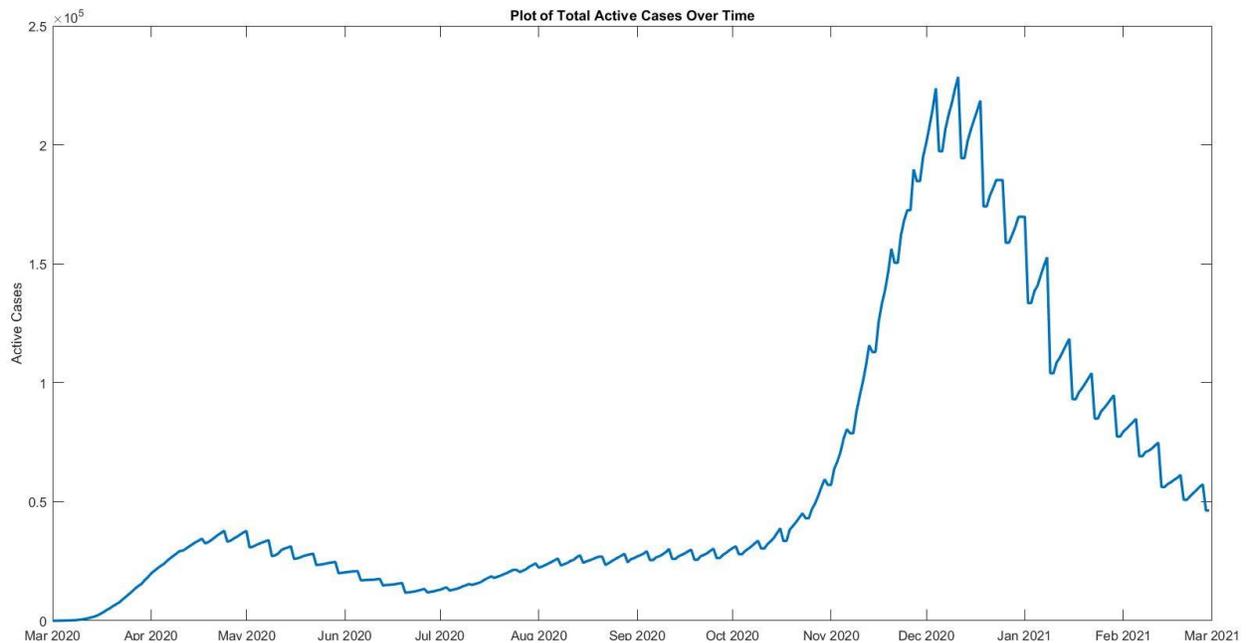


Figure 6: Actual data of all infectious cases in Michigan of COVID-19. [7]

actual number of infectious people is much closer to 240,000, only 0.6% of the estimated amount. Since the data does not account for exposed by not infectious, the line of all positive cases (infectious and non-infectious) was added for reference, approximately 750,000, only 2% of the estimated amount, given no change in our day-to-day interactions. This indicates the value of the Stay Home, Stay Safe order and the added effectiveness of the mask policy.

There are quite a few limitations to the developed data in this thesis. For example, looking at Figure 5, it states that a quarantine should have had the number of infectious people peak at one and a half months into the quarantine, which explains the assumed length of Michigan's initial Stay Home Stay Safe orders; however, this is clearly not the case, considering that Michigan is currently still accumulating new cases a year later. There are several issues that may contribute to this. The β and γ values were determined from the first 170 days of COVID-19 in Michigan. Observing Figure 6, it is evident than the actual account of how much covid had affected the community was not available until mandatory testing for students and faculty returning to schools were in place. The curve seen at the end of march and beginning of April is very different from the curve seen in November. Since the data is actually not linear, β and γ could actually be equations themselves, and not variables. Determining these parameters in the context of an equation is much more complex and goes beyond the scope of this thesis project.

V. CONCLUSION

The data gathered in this research should not be taken without cautious interpretation. Furthermore, the data of scientific articles

regarding COVID-19 and its effects on a population should be analyzed carefully before decidedly extracting data. One example of data that can easily be misinterpreted is when searching on the Google search engine "COVID-19 cases in Michigan" the number listed is 768,000, a fairly accurate number of *total* cases, however not an accurate representation of *infectious* cases which are less than half that amount. [14] Furthermore, maximum number of hospitalized cases at one time was 4,326 of a population of ~10,000,000. [7] By no means are these numbers to be taken lightly, however it is imperative to be sure one is reading and interpreting data correctly, especially in a period where the research surrounding a disease is politically charged.

There are two primary conclusions to the research, the first is that disease data and factoring in the action of people is difficult to model accurately. Researchers have come to many different conclusions and must be extremely careful when analyzing and distributing data. Even so, we can also conclude that the quarantine, stay home stay safe, and mask policies were all beneficial to the community.

VI. ACKNOWLEDGEMENTS

I would like to thank Dr. Hyun Kwon as my primary advisory during this research project, helping me to learn how to utilize COMSOL and MATLAB, as well as the additional guidance provided in next steps with research. Furthermore, I would like to thank the Engineering Department for providing access to COMSOL and MATLAB. Lastly, I thank Dr. Pittman and the Honors Department for supporting and helping to instigate this research project.

VII. REFERENCES, IN ORDER OF APPEARANCE

- [1] “Coronavirus,” *World Health Organization*. [Online]. Available: https://www.who.int/health-topics/coronavirus#tab=tab_1.
- [2] H. D. Bhargava, “Coronavirus: What Happens To People's Body If They Get Infected,” *WebMD*, 03-Feb-2021. [Online]. Available: <https://www.webmd.com/lung/coronavirus-covid-19-affects-body#1>.
- [3] K. Dietz, “The First Epidemic Model: A historical Note on P.D. En’Ko” *Institute of Medical Biometry, Tübingen University*, (1988) Available: <https://dokumen.tips/documents/the-first-epidemic-model-a-historical-note-on-pd-enko.html>.
- [4] “Compartmental models in epidemiology,” *Wikipedia*, 03-Apr-2021. [Online]. Available: https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology.
- [5] J. desan, “Covid-19 Prediction in USA using modified SIR derived model,” *medRxiv*, 01-Jan-2020. [Online]. Available: <https://www.medrxiv.org/content/10.1101/2020.12.20.20248600v1.full>.
- [6] Z. Wang, X. Zhang, G. H. Teichert, M. Carrasco-Teja, and K. Garikipati, “System inference for the spatio-temporal evolution of infectious diseases: Michigan in the time of COVID-19,” *Computational Mechanics*, vol. 66, no. 5, pp. 1153–1176, 2020.
- [7] The Atlantic, “Michigan,” *The COVID Tracking Project*. [Online]. Available: <https://covidtracking.com/data/state/michigan>.
- [8] J. P. Hicks, “Michigan closes in on 10,000 confirmed coronavirus cases with largest single-day jump,” *mLive*, 01-Apr-2020. [Online]. Available: <https://www.mlive.com/public-interest/2020/04/michigan-closes-in-on-10000-confirmed-coronavirus-cases-with-largest-single-day-jump.html>.
- [9] VG. Valentini, “SIR Epidemic Spread Model” MathWorks File Exchange. (Apr 2020) [Online]. Available: <https://www.mathworks.com/matlabcentral/fileexchange/75100-sir-epidemic-spread-model>
- [10] D. Baqaee, E. Farhi, M. J. Mina, and J. Stock, “Policies for a second wave,” *Brookings*, 06-Jul-2020. [Online]. Available: <https://www.brookings.edu/bpea-articles/policies-for-a-second-wave/>. [Accessed: 04-Apr-2021].
- [11] S. Saha, G. P. Samanta, and J. J. Nieto, “Epidemic model of COVID-19 outbreak by inducing behavioural response in population,” *Nonlinear Dynamics*, vol. 102, no. 1, pp. 455–487, 2020.
- [12] E. Bayraktar, A. Cohen, and A. Nellis, “A Macroeconomic SIR Model for COVID-19,” 2020.

- [13] K. Haddad, "Michigan coronavirus timeline: Key dates, COVID-19 case tracking, state orders," *WDIV*, 27-Apr-2020. [Online]. Available: <https://www.clickondetroit.com/health/2020/03/24/michigan-coronavirus-timeline-key-dates-covid-19-case-tracking-state-orders/>.
- [14] *Google Search*. [Online]. Available: https://www.google.com/search?q=COVID-19%2Bcases%2Bin%2Bmichigan&rlz=1C1CHBF_enUS829US829&oq=COVID-19%2Bcases%2Bin%2B&aqs=chrome..69i59j69i57j69i59j0i27113j69i60l2.5308j0j7&sourceid=chrome&ie=UTF-8.

VIII. APPENDIX

Matlab function for solving for SIR, as developed by Giovanni Velentini [9]

```
function [S,I,R] = sir_model(beta,gamma,delta,N,I0,T,dt)

    % if delta = 0 we assume a model without immunity loss

    S = zeros(1,T/dt);

    S(1) = N-I0;

    I = zeros(1,T/dt);

    I(1) = I0;

    R = zeros(1,T/dt);

    for tt = 1:(T/dt)-1

        % Equations of the model

        dS = (-beta*I(tt)*S(tt)) * dt;

        dI = (beta*I(tt)*S(tt) - gamma*I(tt)) * dt;

        dR = (gamma*I(tt) - delta*R(tt)) * dt;

        S(tt+1) = S(tt) + dS;

        I(tt+1) = I(tt) + dI;

        R(tt+1) = R(tt) + dR;

    end

end
```