CODEE Journal

Volume 14 *Engaging Learners: Differential Equations in Today's World*

Article 5

3-15-2021

Facing the Pandemic Together: Forming a Collaborative Research Group

Michael C. Barg
Niagara University

Facing the Pandemic Together: Forming a Collaborative Research Group

Michael C. Barg
Niagara University

Keywords: SIR, epidemiological modeling, COVID-19, pandemic, research group
Manuscript received on September 2, 2020; published on March 15, 2021.

Abstract: This is an account of how a reading and writing project in an introductory differential equations course was transitioned to a professor-student research group collaborative project, in response to the global COVID-19 pandemic. Adapting on the fly to the ever-evolving pandemic, we collected data, estimated parameters in our models, and computed numerical solutions to SIR-based systems of differential equations. This is a description of what we did and how we found comfort in the project in this time of great uncertainty. The collaboration yielded successes and more questions than we had answers for, but the situation provided an opportunity of a lifetime for my students to engage in a real-world developing situation.

1 Introduction

Nestled quietly in the northwest corner of New York State, my Spring 2020 introductory differential equations course began on January 22, 2020 like it had in so many previous spring semesters. We watched a *YouTube* clip of the movie *Contagion*, and I mentioned how my favorite part of *Contagion* and similar films is the cut to the scientific lab where we see a computer model predicting the spread of some deadly virus across the globe. Little were we to know that such images would saturate our lives not three months later. We discussed rates of change and the well-known SIR compartmental model. My course includes a group reading and writing project, and I encouraged students to choose their groups and begin thinking about topics of interest to them. Apart from a few first day jitters amongst the students, all was well.

In late January, the novel coronavirus was lurking in the headlines, and in the second week of the semester, I brought in a pre-print of a research article by Chen et al. [1], in which a detailed compartmental model is developed for the transmission of the novel coronavirus SARS-CoV-2, to share with the students as a way to further explain what I wanted them to do with their group projects. The deadline for group proposals came and went, and nobody chose a topic involving epidemiological models.

Fast forward six weeks to March 11. Many colleges and universities in the northeastern United States had announced plans to close or move education online, and my students

were told that their upcoming spring break had been extended by one week. We all know that things moved quickly after this point, and much of the next two weeks was spent with preparations for completing the semester in an online environment. I immediately thought about my students' group projects, and I took steps to offer an alternative option in order to complete the project. Five of my seven students took me up on the offer of forming a research group to explore variations on the standard SIR model. This is our story.

In putting together the pitch for this collaborative project, I found it necessary to temper the structure of my traditional project with the ever-evolving and constantly changing COVID-19 pandemic. There were some tasks that all group members completed, and there were clear ways in which certain parts could be delegated to individuals who would then report back their findings to the larger group. A regular Wednesday evening meeting served as a requirement that all research group members had to complete. My students embraced this opportunity to work together. We collected data, estimated parameters in our models, and computed numerical solutions to SIR-based systems of differential equations. In this time of great uncertainty, we found comfort in the project and sought ways to understand our models and their relationships to a real-life developing situation. While I hope that such a devastating scenario does not present itself again, the situation provided the opportunity of a lifetime for my students.

2 Our Story

The original reading and writing project required that the students create a five-page paper and a group presentation. While other educators adapted such group presentations to an online format, I felt it was necessary to abandon the existing structure so that the course could better align with a rapidly evolving current event. All too often, colleagues in other departments talk about how their students are reading about and learning about hot topics from the current news cycle. I fear that mathematics is sometimes left behind in this context and viewed by students and the general population as a rather stale discipline.

This moment, a global pandemic, was our time to shine. I felt that I would be remiss in not addressing epidemiological models in my course. I clearly remember thinking that, if I did not somehow change my current plans, I would regret the decision in the future. I asked myself the question "How could I adapt the course project in a way to harness our limited preliminary knowledge of ODEs to explore dire questions of disease spread?" I dove into the rapidly expanding literature. It seemed that multitudes were drafting research articles aimed at understanding the transmission and control of COVID-19. Many were seeking to predict its course. I thought to myself "Is there any reason why my students and I cannot be involved with this process?" Fearlessly and humbly, I decided that the answer was "No." I set to work formulating a plan for how we could gather data, process the data, and share in the pandemic experience as a research group. While I had supervised groups of students in course projects and as research assistants, I had never led a research group with students whose mathematical preparation consisted, in some cases, of only a two-semester sequence in calculus. Based on the vast amounts of freely available COVID-19 data being reported daily, I knew that the revised project should include a data

collection component. After finding De Castro [2], I also knew that I wanted to have the students perform parameter estimation, a topic that I had never included in my first-level ODE course. Now, I am a staunch believer in adhering to the syllabus as an agreement between my students and myself, and I only very rarely make adjustments to those items in the syllabus that were essentially agreed upon by all at the beginning of the term. As such, I decided a fair alternative would be to give my students an option for their project completion. Classes were slated to resume remotely after the extended spring break on Monday, March 30. Nearly a week beforehand, I emailed my class. I invited all course students to a video conference in order to discuss course changes, and I gave students a choice:

- Project Option #1: Continue the group project from before without the group presentation.
- Project Option #2: Engage in collaborative work concerning an epidemiological model for COVID-19.

We discussed my ideas for Project Option #2. Student groups were to notify me their preference the next week. All but one group selected Project Option #2. We were all set to go!

In my search of the growing literature, I selected a handful of articles and pre-prints including

- Chen et al. [1], "A Mathematical Model for Simulating the Phase-based Transmissibility of a Novel Coronavirus," Infection Diseases of Poverty, (2020) 9:24.
- De Castro [2], "SIR Model for COVID-19 Calibrated with Existing Data and Projected for Colombia," pre-print accessed from arXiv:2003.11230 [q-bio.PE] 29 March 2020.
- Sameni [3], "Mathematical Modeling of Epidemic Diseases; A Case Study of the COVID-19 Coronavirus," pre-print accessed from arXiv:2003.1137v1 [q-bio.PR] 25 March 2020.

The paper by Chen et al. is a slightly revised version of the pre-print that I had shared with the entire class in January during my explanation of the course project. Deciding that all students in the class, not just the five of seven in our research group, could benefit from a more in-depth analysis of SIR and infectious disease models, we all discussed parts of the work by Sameni during the first week of remote instruction after the extended spring break.

Sameni's article presents an interesting and applicable modification of the SIR model to include an exposed, but not symptomatic, group. Dubbed SEIR in the paper, the model allows for the exposed compartment to increase in size from interactions between susceptible individuals and either symptomatic or asymptomatic individuals. Moreover, a passed-away group is included, and perhaps we might refer to the model as SEIRP. Figure 1 is a schematic for the SEIRP model showing the flow between compartments. As we will see in the details that follow, one reason why I decided to share SEIRP with the entire class is because it gave a description for how a model parameter could be

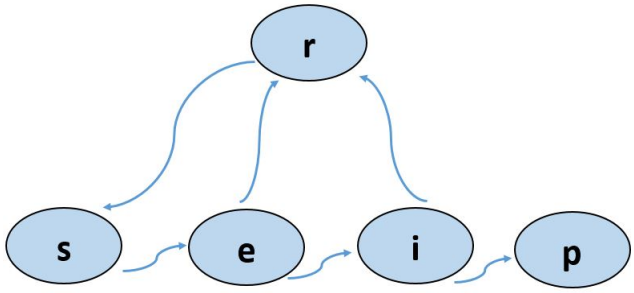

Figure 1: Schematic for SEIRP model adapted from Sameni [3, Figure 5].

estimated using the fourteen-day quarantine that we had all been hearing about in our news feeds. In particular, if one only considers the flow from the exposed compartment to the infected compartment, the differential equation for the exposed compartment is simply the exponential equation

$$\frac{de}{dt} = -\kappa e$$

with solution $e(t) = e(0)\mathrm{e}^{-\kappa t}$. One can verify that after approximately $6\kappa^{-1}$ time units, the size of the exposed population decreases by 99.75%. With a 14-day quarantine period, we might then argue that $14 = 6\kappa^{-1}$ which gives $\kappa \approx 0.43$ (inverse days).

Our newly formed research group would eventually work with SEIRP and consider its many other parameters, but we needed to back up a bit first. Students had read the details of Option #2, and those who had decided to join the group had some preliminary decisions to make. In particular, Figure 2 is a portion of the document that I had distributed to the entire class. There were tasks for every member to complete, and there was also a division of labor for other parts of the work.

A shared Excel file *COVID_19_model* would become the main working spreadsheet for our group. At this point, it mainly consisted of data that I had started to collect for Niagara County, where Niagara University is located, and the spreadsheet listed the number of positive cases as reported via the New York State's Department of Health COVID-19 Tracker (see Appendix A.1). I also included a file that served as an example for how to implement Euler's method to obtain numerical solutions to an SIR model. This file had been shared with the entire class earlier in the semester when we had previously considered Euler's method.

In preparation for our first research group meeting, students were to each select one location and time period. This part of the project was tailored to the fact that the group consisted of six total individuals, the five students and myself. In an email to my research group, I wrote "The distinction between early and late dates will be something for us to decide as a group based on when we think the early and late stages of the particular outbreak took place. Also, we should decide which US location we wish to focus on. Should we focus on all of NY? Should we focus on Niagara County? Should we focus on all of NY except NYC? Should we focus on NYC? There are lots of options. Please let me know your preference."

Tasks for Everyone:

1. Read the article Chen et al. [1], "A Mathematical Model for Simulating the Phase-based Transmissibility of a Novel Coronavirus" available on Canvas.
2. Read the article De Castro [2], "SIR Model for COVID-19 Calibrated with Existing Data and Projected for Colombia" available on Canvas.
3. Download Excel files *COVID_19_model* and *numericalSolsEx1* for personal exploration.
4. Participate in Wednesday group meetings.

Individual Tasks - Choose 1:

1. Compile data for *Confirmed, Deaths*, and *Recovered* from selected location in China during appropriate time periods.
   (a) Early dates
   (b) Dates near the peak of the epidemic - present
2. Compile data for *Confirmed, Deaths*, and *Recovered* from selected location in Italy during appropriate time periods.
   (a) Early dates
   (b) Dates near the peak of the pandemic - present
3. Compile data for *Confirmed, Deaths*, and *Recovered* from selected location in the United States during the appropriate time periods.
   (a) Early dates
   (b) Dates near the peak of the pandemic - present

Figure 2: Portion of document describing Project Option #2.

3 The First Meeting

Our first meeting took place on Wednesday, April 8 via Zoom. It was comforting to see each other, and many students expressed their eagerness to engage in the work with the group. One of the first questions that arose was how to specifically define the early dates and the dates near the peak of the pandemic for the various geographical regions. We also discussed some of the limitations of the basic SIR model, including the fact that it assumes a fixed population size, recovery from the illness, and subsequent immunity from re-infection. To better address the fixed population size, I asked participants to refine their geographical location provided that the relevant data was available. Students chose Hubei in China, New York City (NYC) in the US, and the entirety of Italy. Our main source for data was the "COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University" (see Appendix A.2). For brevity, we occasionally refer to this data as the "JHU CSSE COVID-19 Data."

The discussion about time periods was interesting, and students had much to offer. One student suggested looking at a site (see Appendix A.3) that had charts like Figure 3. Some of the suggested ideas for time period cut-offs included finding dates at which a certain percentage of the peak number of infected individuals had occurred and just an *ad hoc* selection of a cut-off date based on when a certain number of infected individuals had been reached. Since the peak number of infected individuals might not have been

t

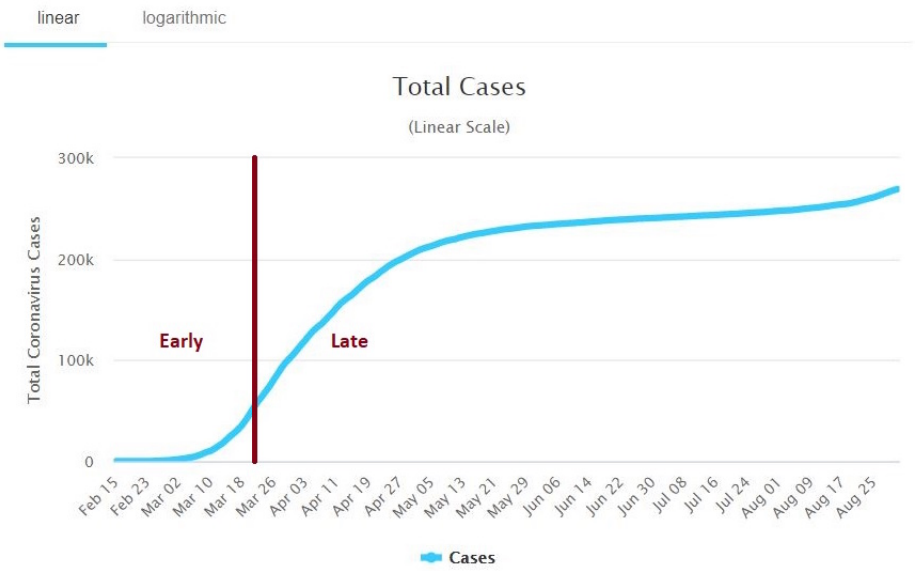

Figure 3: Total coronavirus cases in Italy as reported by Worldometers.info [4].

reached at the time of this first meeting, we abandoned the idea of selecting the cut-off date based on the timing of the peak. Collectively, we agreed to use the rule that the early time period ends when the total number of cases in a given location surpasses 50,000 individuals. This value seemed reasonable to use as a cut-off based on our review of the reported numbers of cases in our regions. As an illustration, the early Italy time period is February 15 through March 20 since the total number of cases on March 20 was 47,044, and the total number of cases on March 21 was 53,598. The other time periods are shown in Table 1, together with the starting population size values, N_0, as described in Section 4.

Table 1: Time Periods and Population Sizes

Region	Time Period	N_0
Early China	(01/22/20 - 02/11/20)	57,237,740
Late China	(02/12/20 - 04/13/20)	57,236,672
Early Italy	(02/15/20 - 03/20/20)	60,461,826
Late Italy	(03/21/20 - 05/03/20)	60,457,794
Early US	(03/02/20 - 03/22/20*)	8,550,971
Late US	(03/23/20 - 04/13/20)	8,550,748

Note that the total number of cases in the US had not yet exceeded 50,000 by the time of our discussion about cut-off dates. As such, we used 10,000 as the cut-off value to obtain the distinction between early and late US time periods. The end dates for the late time periods vary depending on when each student last collected data for their region. While

there are arguably better ways to define the time periods of interest, my students and I could get started with our data collection using this procedure. With so much uncertainty surrounding pandemic modeling, we decided to see what conclusions we would be led to with this decision. We could review our time period definition later if necessary to see if the data seemed to warrant a change. The main task for the group for the next week was individual data collection. Specifically, student learning objectives for the data collection phase of the project were to locate relevant data for their location and time period from the JHU CSSE COVID-19 Data and to compile the cleaned data in a spreadsheet with daily values of confirmed cases, deaths, and recovered cases.

4 The Second Meeting

Our second group Zoom conference took place in the evening on Wednesday, April 15. Students had collected data and sent me their spreadsheets beforehand. As an example, let us look at some data that was collected for the late Italy time period. Figure 4 is an image of part of the Excel file created by one of the students.

	A	B	C	D	E
1	Date	Confirmed	Deaths	Recovered	Active
2	3/21	53578	4825	6072	
3	3/22	59138	5476	7024	46638
4	3/23	63927	6077	7432	50418
5	3/24	69176	6820	8326	54030
6	3/25	74386	7503	9362	57521
7	3/26	80589	8215	10361	62013
8	3/27	86498	9134	10950	66414
9	3/28	92472	10023	12384	70065
10	3/29	97689	10779	13030	73880
11	3/30	101739	11591	14620	75528
12	3/31	105792	12428	15729	77635

Figure 4: Portion of initial spreadsheet with cumulative totals for late Italy time period.

The Confirmed, Deaths, and Recovered columns were filled in by hand from the JHU CSSE COVID-19 Data. During our second meeting, we looked at the data. Thinking it wise to start with a previously presented model, we discussed the De Castro [2] paper in which a traditional SIR model is employed. With only the Infected class I (from the Active data) and the Recovered class R (which included both the Deaths and Recovered columns) data available at first, we all quickly noticed the need to create a corresponding Susceptible class, S, data for each date. Students were tasked with this process for the next meeting. For the students working with the early time periods, a starting population size N was required. We decided that students would find this value via an internet search while noting their source. Using the fixed population size assumption $S + I + R = N$, students could then compute an S value for each day with $S = N - I - R$. Students charged with the late time periods decided to find their starting N value as the final S value from the corresponding early group. A more realistic approach is to select the starting N value for the late time periods by calculating the difference between the early time period N value and the total number of deaths during the early time period. The starting population size, N_0, values reported for the late time periods in Table 1 are computed this way.

Since it is the unlikely case that N is fixed, we were led to discuss other limitations of DeCastro. De Castro [2] In particular, students pointed out that the traditional SIR model assumes full recovery and immunity for those who transition from the I to R compartments. Certainly the deceased data is different than recovered data, and immunity for COVID-19 is still an unknown question. Despite these drawbacks to a traditional SIR approach, the work by De Castro [2] gave our group a way to approximate the two parameters in the model, namely the transmission coefficient ψ and the recovery rate δ. With this notation, the SIR system in De Castro [2] is written as

$$\begin{aligned}\frac{dS}{dt} &= -\psi SI\\ \frac{dI}{dt} &= \psi SI - \delta I\\ \frac{dR}{dt} &= \delta I.\end{aligned}$$

Taking the time step as $\Delta t = 1$ day (consistent with our available data) and approximating the derivatives $\frac{dS}{dt}$ and $\frac{dI}{dt}$ with finite backward differences, coefficients ψ_n and δ_n could be computed for each day via

$$\psi_n = -\frac{S_n - S_{n-1}}{S_n I_n} \text{ and } \delta_n = \frac{\psi_n S_n I_n - (I_n - I_{n-1})}{I_n},$$

where the subscript n refers to the nth day in one's data set. Letting N_d be the total number of days, ψ and δ can then be approximated as numerical means

$$\psi = \frac{\sum_{n=1}^{N_d} \psi_n}{N_d} \text{ and } \delta = \frac{\sum_{n=1}^{N_d} \delta_n}{N_d}.$$

After demonstrating these computations to the group during our second meeting, students were asked to perform the computations with their own data for the next week. Students seemed to agree that these computations were straightforward to complete. We forged on. Equipped with ψ and δ, we then used Euler's method to compute numerical solutions for the SIR model. Not surprisingly, as I assured the group, the numerical solution predictions were vastly different than the data. Since we were just getting started with the project, we did not stress about this too much. I followed-up with a numerical solution from Matlab's ODE45 solver that uses a Runge-Kutta method. While more accurate, the Matlab approximation for the number of infected individuals on the nth day was still quite different than I_n.

5 The Third Meeting

The students and I were largely unconcerned with the inaccuracies of our SIR model since we had plans to consider the SEIRP model which was anticipated to be more accurate. We met next on April 22. After receiving and reviewing the requested computations of S, ψ, and δ from each student, I decided that in our third meeting we would look

instead at the SEIRP model in Sameni [3]. As is standard in the analysis of an SIR model, we divide each of S, E, I, R, and P by the total population size N in order to form the proportions in each class. Lowercase letters are used to label the proportions, i.e., $s = S/N, e = E/N, i = I/N, r = R/N$, and $p = P/N$. The SEIRP system from Sameni [3, Eq. (12)] is

$$\begin{aligned} s' &= -\alpha_e se - \alpha_i si + \gamma r \\ e' &= \alpha_e se + \alpha_i si - \kappa e - \rho e \\ i' &= \kappa e - \beta i - \mu i \\ r' &= \beta i + \rho e - \gamma r \\ p' &= \mu i. \end{aligned}$$

Students agreed that this model seemed more realistic as it contained the additional compartments E for exposed, asymptomatic individuals and P for those who had died. Based on what we had all been hearing about COVID-19, SEIRP seemed like a better choice as a model. News sources frequently discussed exposed, asymptomatic individuals, and everyone regularly heard updated death counts. While including more compartments seemed to be appropriate, the SEIRP model included more parameters which needed to be estimated. The group decided on a few different methods for the selection of the parameters. I summarized the plan in a document to the students, and it is copied in Figure 5 (on the next page) for convenience. Item 4 and Item 6 in Figure 5 serve as explicit student learning objectives for the modeling portion of the project. Parameter values for our regions are given in Table 2.

Table 2: Calculated Parameter Values

Region	Recovery Rate β	Mortality Rate μ
Early China	0.01007	0.01054
Late China	0.08561	0.00198
Early Italy	0.02623	0.02103
Late Italy	0.01893	0.00693
Early US	0	0.00157
Late US	0	0.00604

For our Euler's method numerical solution, group members noted that initial conditions were needed for all of the compartments. The glaringly difficult one to estimate was E_0, since E_0 is impossible to know exactly for a given location and starting time. We discussed how we might instead use E_0 as a variable in our computations. In other words, we could compute an estimate to I_n and check our accuracy for different values of E_0. In this way, the group decided that we might be able to discover some acceptable E_0 values.

Parameter and Initial Condition Estimation/Exploration
Here are some things to do with our data in order to compute numerical solutions to the SEIRP model.

1. Perhaps use $\kappa = 0.43$ as in Sameni's article.
2. Explore the use of different values for α_i and α_e. Note that we expect $\alpha_e > \alpha_i$.
3. Explore different γ values. Note that $\gamma = 0$ corresponds to the case where the disease confers immunity.
4. Estimate the recovery rate β by performing computations with your data like we did for the SIR model in DeCastro's article. You should be able to compute a sequence of recovery rates β_n with the formula

$$\beta_n = \frac{r_{n+1} - r_n}{i_n},$$

where the subscript n refers to the nth day. Then define β as the average

$$\beta = \frac{\sum_{n=1}^{N_d} \beta_n}{N_d},$$

where N_d is the total number of days of data.
5. As in Sameni's work, make the parameter ρ of the same order as β. Explore the use of different ρ values.
6. Estimate the mortality rate μ by performing computations with your data like we did for the SIR model in DeCastro's article. You should be able to compute a sequence of mortality rates μ_n with the formula

$$\mu_n = \frac{p_{n+1} - p_n}{i_n},$$

where the subscript n refers to the nth day. Then define μ as the average

$$\mu = \frac{\sum_{n=1}^{N_d} \mu_n}{N_d},$$

where N_d is the total number of days of data.

Figure 5: Portion of document describing parameter selection and computation.

6 The Fourth Meeting

We convened as a research group for the final time on Monday, May 4. I had previously compiled recent modifications to all of the student's individual spreadsheets. In this updated spreadsheet, I also provided formulas for an Euler's method solution to the SEIRP model using parameter values as in Sameni [3, Examples 3-5]. One such solution can be seen in Figure 6.

Students were also given a file containing additional tasks to complete in order to finish the project. In particular, Figure 7 is a portion of that document.

While I hoped that many students would complete the tasks, I realized that it might be necessary to relax my expectations for the project completion. Our fourth meeting would provide me with an opportunity to assess how far along we had come. We started our discussion by re-visiting Sameni [3]. Students had attempted to compute the mortality rate and recovery rates, but not everyone had succeeded. Additionally, attempts by the students to compute numerical solutions to the SEIRP model were largely unsuccessful. I decided that it was necessary to spend more time discussing the examples that I had provided to the group. During this final meeting, we spent time looking at the Euler's method formulas that I had implemented in our shared Excel file. This discussion seemed to clear up some of the misunderstandings that students had, and I encouraged them to pursue our shared goals from the list of additional tasks on their own time.

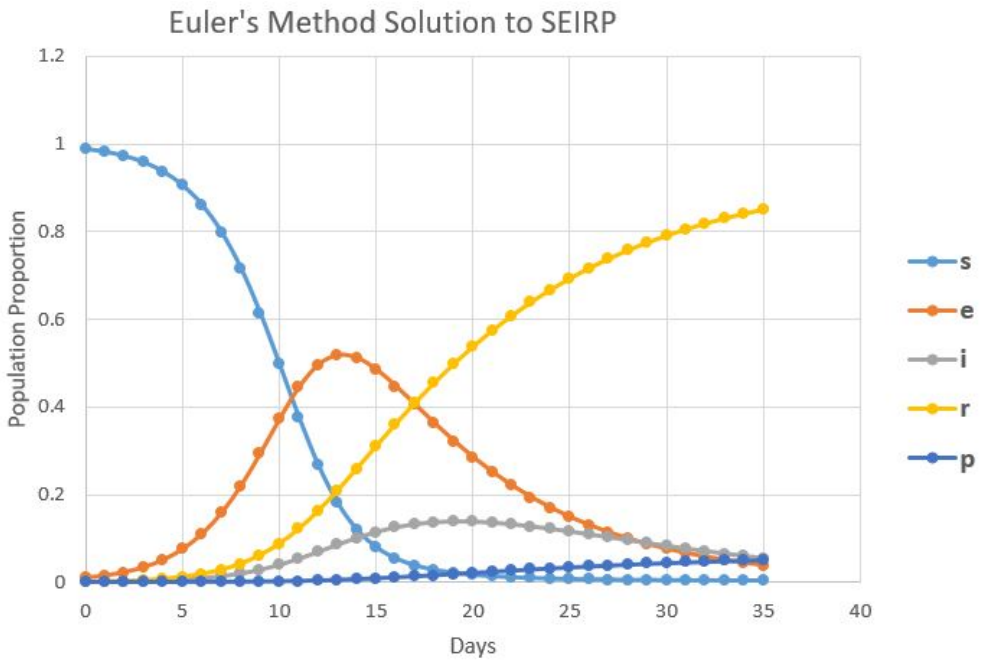

Figure 6: Numerical Solutions to SEIRP with parameter values as in Sameni [3, Ex. 3].

In order to enforce the constraint $s+e+i+r+p=1$, solve for $s=1-e-i-r-p$ and substitute this expression into the e' equation in place of s. The resulting system has one fewer equation. We no longer need the equation for s' since we can determine s by knowing the other four variable values.

$$\begin{aligned} e' &= (1-e-i-r-p)(\alpha_e e+\alpha_i i)-\kappa e-\rho e \\ i' &= \kappa e-\beta i-\mu i \\ r' &= \beta i+\rho e-\gamma r \\ p' &= \mu i \end{aligned}$$

Note that this reduced system is Eq. (27) in Sameni's article. The SEIRP tab in our *COVID_19_model* Excel spreadsheet is set up to use Euler's method to solve this reduced system for Examples 3, 4, and 5 of Sameni's article. Values for the parameters are given in those examples in such a way as to illustrate a number of different features of the (reduced) SEIRP model.

Our task is to use the data that we found in order to

- Estimate and explore relevant parameter values for COVID-19 in our different locations.
- With our parameter values, assess how well the SEIRP model fits the observed data.
- Use acceptable parameter values in order to predict when the peak of the infection will occur and how many individuals will be infected at that time.

Figure 7: Portion of document describing final project analysis.

The student working with the late Italy time period made good progress with implementing the Euler's method formulas to solve the SEIRP system. Table 3 lists the values that were used for the late Italy time period parameters. To set the initial conditions, the students computed $e(0) = 10,000/N, i(0) = i_0/N, r(0) = r_0/N, p(0) = p_0/N$, where the subscript refers to the first actual data value (March 21) for the respective compartment.

The peak of the infection was found to occur on April 18, with 13,528,012 people infected. Comparing this with the data, we see good agreement with the actual date of the peak (April 19), but poor agreement with the actual number of infected individuals (108,257). A more accurate numerical solver will give better results, but it is likely that the parameter values in the model need to be adjusted for better agreement with the observations.

Table 3: Late Italy Parameter Values

β	μ	α_e	α_i	κ	ρ	γ
0.01893	0.00693	0.6	0.4	0.43	1	0

7 Project Assessment

Assignment of project grades remained as a final task for me to complete. The course project accounted for 15% of the total course grade. Project proposals submitted before the pandemic interruption were still included in the project grade. For the remaining portion of the project grade, I decided to use a modified version of a presentation rubric that I had used in previous semesters. In the original rubric, three items were assessed on a five point scale (0-4): Time requirements are met at a comfortable pace, Voice is loud and clear, and Style and content complement each other. The modified rubric used the five point scale for three different items: Requested components were submitted in a timely fashion, Student actively participated in Wednesday meetings, and Reasonable mathematics was used within a student's spreadsheet submissions, respectively. Overall, my small research group performed well.

8 Conclusions

That is our story. I am happy to share it with you, and I hope that it reminded you of your own journey through these tumultuous times or that it has given you inspiration to forge your own path into real-time modeling with an undergraduate class. We learned a lot through the process. We had fun, we scratched our heads a lot, and we shared some of our feelings of anxiety and loss during an extremely challenging semester. Most importantly, my students obtained a real-life experience that led to a better appreciation for how mathematics is alive and relevant in the world today.

Appendix A Links to Data and Files

A.1 This is the New York State's Department of Health COVID-19 Tracker:

https://covid19tracker.health.ny.gov/views/NYS-COVID19-Tracker/NYSDOHCOVID-19Tracker-Map?%3Aembed=yes&%3Atoolbar=no&%3Atabs=n

A.2 Our main source for data was the "COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University" (JHU CSSE)

https://github.com/CSSEGISandData/COVID-19.

A.3 Some charts like Figure 3 can be found at

https://www.worldometers.info/coronavirus/#countries.

References

[1] T. Chen, J. Rui, Q. Wang, Z. Zhao, J. Cui, and L. Yin. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infectious Diseases of Poverty*, page 9:24, 2020.

[2] C. De Castro. SIR model for COVID-19 calibrated with existing data and projected for colombia. 2020. Pre-print accessed from arXiv:2003.11230[q-bio.PE].

[3] R. Sameni. Mathematical modeling of epidemic diseases; a case study of the COVID-19 coronavirus. 2020. Pre-print accessed from arXiv:2003.1137v1[q-bio.PR].

[4] Worldometers.info. http://worldometers.info, 2020. Accessed: 2020-09-01.