Modeling/Simulation Theory/App (CSE 561) MCS Portfolio Report

Kaden Brown
Arizona State University
Tempe, United States
ktbrown6@asu.edu

I. Introduction

Modeling and Simulation Theory and Application is concerned with the generation of models for systems that represent real-world problems which can then be experimented upon to draw conclusions about the behavior of the system. This broad definition is necessitated by the diversity of the subject which lends itself to applications across a vast number of unique fields. This portfolio project concerns epidemiology through the lens of modeling and simulation.

Though given a large amount of autonomy to design the project, foundational expectations were laid out to guide progress and determine grading criteria. The first of these was the necessity of establishing relevance and background information behind the idea of the project. Secondarily, the system itself had to be described. The system encompasses both structural and behavioral components while including constraints on abstraction in order to reduce the complexity of modeling. Each component must operate in terms of a time frame: discrete time, continuous time, or discrete event; while simultaneously having its inputs and output trajectories defined along with any behavior or data manipulations. This is what will be represented by a model — the description of what actual components and variables will be manipulated during simulation.

After definitively designing both the system and model, the model was required to be implemented using the Discrete Event System Specification (DEVS) library in Java. Different components needed to be built in independent classes and their interactions, or couplings, defined with one another. These were further required to be organized into frames, or collections of components that execute linearly (or consecutively) with a defined input and output set.

Following the implementation of the model, simulation experiments were required to demonstrate the observed behavior of the model under various conditions. Multiple scenarios were required to be executed and their performance evaluated across a variety of metrics such as summary and discussion of simulation outputs. This is the element of modeling and simulation that can be used to make insights about the larger system which was previously designed. Given that the system is representative of a real-world situation or structure, an aim of this project was to show an event or process might actually unfold — in order to do so data needed to be generated and analyzed for this purpose.

II. Solution Explanation

This section will explore how the aforementioned criteria set out for the project were met in the context of this particular project instance. This project problem statement was unique and not replicated by any other team.

A. Establishing Relevance

The selected system for modeling and simulation was a small yet urban town that had been exposed to a viral agent similar to that of COVID-19. This unique project aimed to explain how disease transmits through a population and its resulting effects. This is useful for demonstrating why institutional responses dictating social behavior might be helpful in stopping the spread of disease and preventing suffering.

B. System Description

A system of a densely populated town was defined with a population that was subject to variable trait characteristics. These are reflected as demographic traits. Other factors were included as well, such as social trends, behaviors, and disease strain. Each of these has the capability to affect infectivity and morbidity. In order to represent a town and its population, the number of people was normalized so that percentages of the population were used rather than integers. Each component reflects a discrete-time based data trajectory, allowing for the incremental advancements of one-week time units. In each iteration, individuals belonging to the population will become infected, perish from disease, recover from disease, or remain healthy if beginning the cycle healthy.

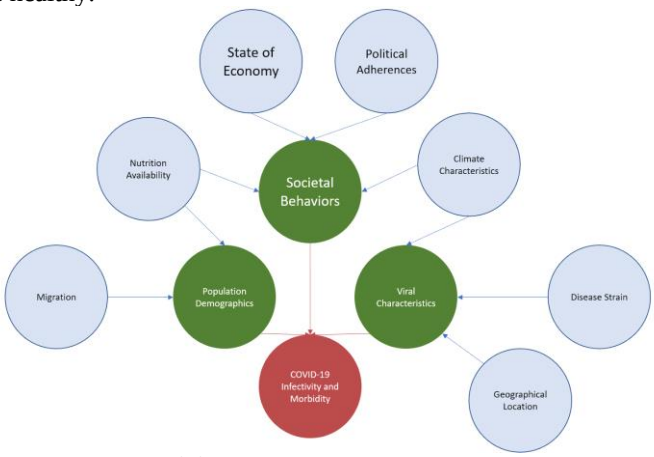

Fig 1. Abstraction of features. Not all possible variables may be considered due to constraints on development time — this demonstrates a hierarchy of variable organization.

Due to the epidemiological nature of this project, a high level of abstraction was necessary in order to reduce the dimensionality of factors impacting the population. Several key population demographic features were isolated in order for analysis including age distribution, innate immunity, asymptomatic rate, gender distribution, obesity rate, and income distribution. Selected social trends and behaviors for

analysis included vaccination rate, mask compliance level, population density, household size, and quarantine compliance. Finally, with respect to disease-specific characteristics, viral infectivity, vaccine resistance, and mutation rate were selected for testing through simulation.

Though this list might seem extensive, note that this abstracts other features such as access to adequate food supplies, cleanliness of water, use of antiviral drugs, etc. There are an infinite number of variables not considered, and those that were selected were chosen in order to reduce model complexity given the scale of the project.

Other abstractions include the assumption that only living vectors of disease (infected individuals) may spread illness, viral spread is limited to 10% of the population at a maximum within a one-week period, and that no travel in or out of the community may occur during the simulation. Additionally, no individuals are born or pass away from non-virus-related causes. It is also important to recognize that immunity is conferred if a recovery is made — this means that an individual cannot be infected twice. Lastly, it is assumed that at the end of one discrete-time iteration (one week) that a previously infected individual will make a full recovery if they do not perish.

C. Model Description

The model is specified in terms of each component involved in the representation of the system; each has some input and output trajectory along with a "react" set that identifies its behavior given an input that produces output. Before explaining these components, first the selected features must be re-visited to explain their functionality throughout the various components. In order to analyze the effects of manipulating system characteristics, two elements are isolated: transmission rate and morbidity rate. Transmission rate reflects how infectious a disease is in a population scenario while morbidity reflects how deadly it will be. Each feature is generally considered to either impact transmission or morbidity rather than both simultaneously (with few exceptions). By running scenarios through this model design, information regarding the speed of spread, extent of spread, and severity of disease within the population may be observed.

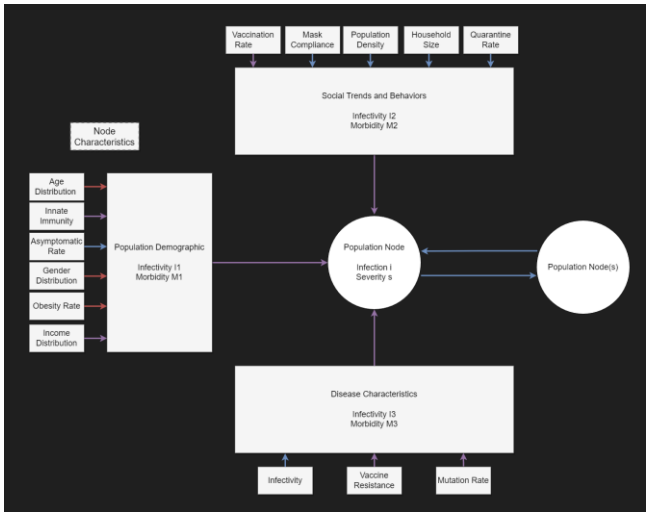

Fig 2. Organization of variables / features. Demographic features are on the right, behavioral on top, and viral on the bottom. Edge colors reflect effects on either infectivity, morbidity, or both.

The first element of the population features is age distribution which affects morbidity. As this distribution trends towards higher values, the morbidity rate increases [1]. The second is innate immunity which decreases the infectivity in the population as some individuals are not susceptible. Next comes asymptomatic rate — this increases the infectivity because individuals who do not know that they are ill are more likely to spread disease as they are less likely to quarantine / observe more cautious behavior. Gender distribution has been found to profoundly impact mortality with males experiencing significantly higher risk than females. Obesity rate is also correlated with increased morbidity with increasing rates of obesity linked to increasing morbidity. Finally, income distribution increases both infectivity and morbidity when lower, and has the inverse effects when higher[2].

In addition to population factors, features describing social behaviors are critical to understanding the system. The first of these is the population's vaccination rate which affects the morbidity rate within the model[3]. Similarly, as the feature mask rate increases in value representing adherence to masking policies, infectivity decreases[4]. Population density describes how clustered the population is and is positively correlated with infectivity[5]. Household size functions in a similar manner, reflecting how many people tend to live under one roof[6]. Finally, quarantine rate is also positively correlated with infectivity .

The third category of variable features are disease characteristics. The raw infectivity is set here — this feature has the initial and therefore most profound determining influence on transmission rate. Mutation rate is considered to affect both transmission rate and morbidity rate with a relationship of positive correlation. Lastly, vaccine resistance also is positively correlated with both transmission and morbidity.

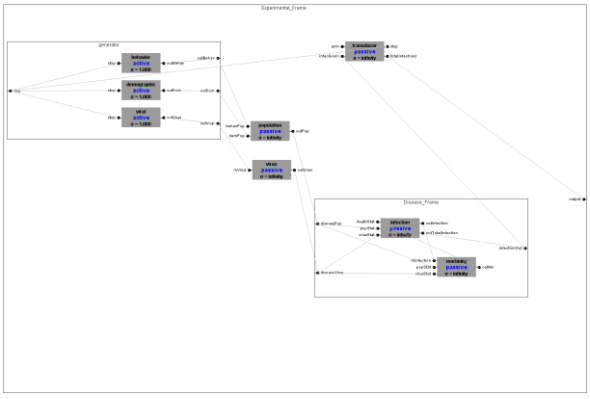

Fig 3. Model design in DEVS. The leftmost box represent generator within the generator frame. The two middle rectangles show the populationTraits and viralTraits components. The rightmost box represents the Disease_Frame and its Infection and Mortality components. The top-right rectangle is the Transducer.

All of these described features will differ in each model construction in order to explore the different possible scenarios. To accomplish this variable setting of values, the first component, an input generator, is defined. The overall generator is a frame composed of three individual sub-generators with each corresponding to the features from the population demographics, social behaviors, and disease

characteristics. These sub-generators are designated as demGenerator, behavGenerator, and viralGenerator. Each takes an input trigger along its singular input port to signal that it must create a value (random or specified through manual selection) between 1 and 10 to describe the danger reflected by the features it encompasses. A higher value reflects a more at-risk population, behavior set, or viral profile. This value is passed out of the overarching generator frame for each one individually and is used to inform the next set of components.

The next stage of model processing involves two components: populationTraits and viralTraits. These are responsible for setting the values of all the individual features previously described that will affect both morbidity and infectivity. Granted the output values from both demGenerator and behavGenerator, populationTraits will set the values for all demographic and behavioral features. Similarly, viralTraits will receive the value produced by viralGenerator. With each individual feature now provided a value, a calculation is performed to determine how the demographics, behaviors, and viral traits of the model will impact transmission and morbidity. Both populationTraits and viralTraits have one output channel, however populationTraits must take two separate inputs from different components and thus has two input channels while viralTraits has only one. The features set here can be manually determined, however by default are randomly assigned with weighted probabilities determined on their component inputs.

These components transition the model from generated input values used in simulations to the "Disease_Frame" which represents how the virus actually moves through and affects a population. This frame has two input ports, connected to both the populationTraits and viralTraits output ports. The frame has only one output port which contains data corresponding to the ultimate execution statistics of a completed simulation including total mortality and number of rounds to reach completion. The bulk of computation occurs within this frame through the cooperation of two components: Infection and Mortality.

The Infection component receives values along three input ports — one corresponds to the output from populationTraits, another to the output of viralTraits, and the last to the output of the Mortality component which will be explained next. Upon its first iteration, the population will be set to 1, representing that no deaths have occurred and all individuals are healthy. A subset will become infected, with traits incorporated from the outputs of populationTraits and viralTraits used to set weighted probabilities for how many infections will occur in each iteration. Each round/ iteration, the healthy population statistic will be relayed to the Mortality component.

The Mortality component takes both the outputs from populationTraits and viralTraits along two input channels as well as the current portion of healthy citizens from the Infection component along another input channel. Using the data from the first two, weighted probabilities are used to determine the morbidity risk in the first round. In each subsequent round when passed the portion of the population that is healthy, the amount that is infected can be calculated using the previous healthy count (because of the assumption that all infected who do not die recover in the next round). Using the initially set morbidity risk, some individuals perish from this healthy count, which is relayed along Morbidity's singular output port back to the Infection component.

In subsequent rounds, the Infection component will continue to infect more of the population until, due to random weight setting, the disease disappears from the population as no more infections occur in a cycle. Thus, the model will continually iterate between the Infection and Morbidity components until this is reached as more people are infected and die in each round. Once no more virus is propagated in a round, the second output port of Infection is activated, sending data to the output port of the Disease_Frame frame.

The final component of the model is known as the Transducer. This keeps track of the total number of infections that occurred over the duration of a simulation. It is also responsible for the ending of a simulation, triggering the stoppage generators from producing new input. This is done using an output port which connects as input to the generator frame. Finally, the second output port of the Transducer exits the experimental frame to deliver the simulation data. This concludes the design of the model and how it executes to produce data.

III. Results Description

The goal of this system and model construction was to execute simulations in order to demonstrate the effects of various population and viral features on disease infectivity and severity. By observing trends in the spread of a virus in a simulated population along with how it might inflict fatalities, the significance of certain factors such as mask-wearing-adherence might be clearly demonstrated and used to inform decision making processes on regulatory levels[7]. In order to assess the model, three scenario types were designed and each one iterated over numerous times to gather data.

The first scenario involves a town with a demographically strong population. This reflects a population that has a relatively low age distribution, high natural immunity rate, and low rates of asymptomatic infection. In order to test the effects of demographic traits (which are largely immutable) against behavioral traits, this scenario has a very low social behavior score, indicating low levels of mask-wearing and vaccination. Individuals are unlikely to quarantine here and live in much more densely packed areas than regular towns. Demographic factors are expected to reduce the spread and morbidity, while the behavioral are expected to increase both[8].

The second scenario aims to assess the opposite type of situation. Here the population is unhealthy — individuals are older, have pre-existing conditions, and have low rates of innate immunity. However, they have excellent adherence to quarantine and mask protocols and are nearly all vaccinated. By observing infection progress and severity measured through death count, the results of this scenario may be compared against the first to determine which between demographic and behavioral traits predict a less damaging viral response.

The third scenario is concerned with measuring the effects of viral variation. Using moderate settings for both demographic and viral traits, a run-of-the-mill population is

defined that should have no extreme variation in its vulnerability. Rather than altering these, the viral characteristics including initial infectivity, mutation rate, and vaccine resistance will be changed sporadically to assess their impact on transmittance and death count.

After running each of these scenarios alongside one another, the data included in figure 4 was observed. The first plot demonstrates that a broad sample space was covered along all the conducted simulations, showing a profile ranging from low-risk populations (low demographic risk and behavior risk) and to low virus risk to high-risk populations and high-risk virus. The second plot shows that, interestingly, the most severe scenarios observing high amounts of population deaths incur relatively short virus life periods. This means that the more deadly a virus is in a given scenario (whether due to population or viral factors), the shorter it will generally last in the population before being eliminated through immunity. This is likely because of the assumption that deceased individuals cannot spread illness.

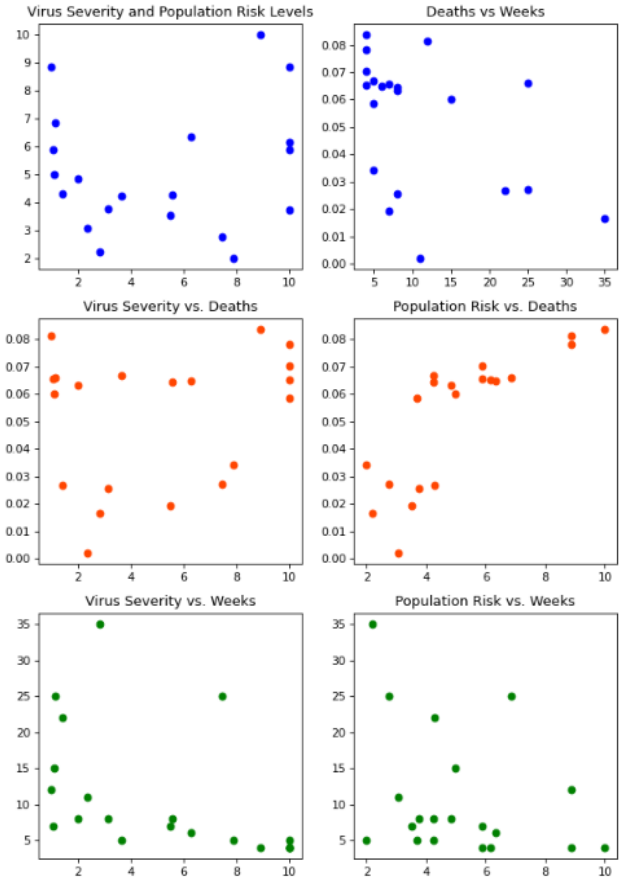

Fig 4. Six plots are shown here representing the findings of the simulation scenario executions. Each point is a simulation scenario, with the axes varying by plot (described in each header, X vs Y).

The third plot represents the correlation (which does not appear strong) between the severity of a virus and the amount of deaths. The severity is representative of the level of scores observed among the three viral variables. The fourth plot, however, demonstrates the clearest trend observed in this project. There is a clear relationship between the risk of a population (set by demographic trait scorings) and the amount of deaths that will be observed. This matched expectations — the poorer the health of a population the more it will suffer regardless of the virus' characteristics.

The fifth plot demonstrates the relationship between viral trait severity and the number of weeks that it can survive in a population. Generally, it appears that the more severe a virus is the fewer weeks it can remain in a population. This was an unexpected finding as it was thought that a more virulent virus would be able to continue to spread longer. It appears that the deadlier viruses, however, struggle to remain active. The final plot shows a weaker correlation between population risk (across both demographic and behavioral traits) and disease propagation time. Again, this was not expected — it was thought that the weaker a population was the longer the virus could exist, however the opposite was observed[9].

Across all observations, it was clear from the simulations across all scenarios that demographic features were the best predictors of the amount of infections as well as deaths. However, they were not indicative of the length of time a virus could spend in a population active. This was better determined by the viral characteristics themselves. While the behavioral characteristics certainly had some ability to alter the observed severity in terms of both infections and deaths, they were not as correlated or pronounced as demographic variables. Thus this project's findings suggest that the most stringent prevention methods and behaviors might need to be enforced for at-risk populations who suffer from demographic issues such as higher age distributions or low immunity rates. In contrast, such measures might be less necessary in environments where the individuals themselves are not at high risk.

IV. Individual Contributions

I offered key contributions to this project throughout the entire development process including the subject matter research phase, selection of features, overall model design, component design, implementation process, and scenario execution and data collection stage. During the research / feature selection phase, I was responsible for collecting information for the population demographic features and how these might impact both infection and morbidity. The accuracy of these features to real-world figures was critical in order to ensure the accuracy of the simulation results, thus after finding the data for these features I was also responsible for checking over the features included in the behavioral section. I also collected and assimilated the data required for accurate performance of the viral characteristics.

During the model design, I produced the figures reflecting potential abstractions as well as feature influences on mortality and infectivity. In this way I was able to generate adequate assumptions and abstractions to guide the project through its developmental stages.

In the component design phase, I was responsible for the generator frame as well as all three of its sub-generators. I also designed the populationTraits as well as its logic governing the weighted random distributions that selected for feature values. Within the Disease_Frame, I designed the Mortality component and the mechanism by which Mortality and Infectivity interacted. After designing these, it was my responsibility to build out their respective implementations in DEVS. The connection of components was largely a shared task, though I was heavily involved in the linking of ports between all components.

In the scenario execution stage, I was responsible for scenarios one and three, assessing the difference between population demographics and behavior. To do so I manually ran and recorded specific scenario settings. I also generated a majority of the figures seen in Figure 4 for the purpose of data visualization.

V. Skills Acquired

Over the course of this project, I became much more skilled in the derivation and construction of system models. I have not worked previously in a modeling or simulation environment and I expect the hands-on experience that I now have in DEVS will aid me in my career moving forwards. The design process for components was perhaps the most important dimension of both the course and this project, demonstrating the difficulties that come along with increasing complexity as well as the necessity of establishing assumptions and abstractions.

At the onset of this project, the only constraints given were that the final deliverable needed to be a functioning model of a system that could facilitate simulations to produce some conclusive data. The problem statement proposal, therefore, determined the subject matter and direction of the project — it was up to myself and my teammate to design the requirements and build on top of these to create what is presented in this paper. I have become more well versed in requirements management as well as system design throughout this solution development.

By designing multiple of the components and frames as well as implementing them in DEVS, I am now much more comfortable with the governing ideas of modeling and simulation engineering especially with regards to controlling the interactions between components via input and output trajectories. Finally, this project was helpful in improving my ability to work with libraries that I am unfamiliar with and test out functionalities that I have never seen before — this was my first time working with DEVs and though I am not an expert, I am proud of what I was able to accomplish in such a short time.

VI. Aknowledgements

This course project was completed in groups ranging from 1-3 students, most groups comprised of 2. This group was a team of 2; the other student was Animesh Gupta — he will also be using this project for his portfolio submission. Any work not explicitly stated to have been completed by me in section IV was his contribution to the project.

VII. References

[1] Centers for Disease Control and Prevention. (2022, September 16). *Risk for COVID-19 infection, hospitalization, and death by age group error processing SSI file*. Centers for Disease Control and Prevention. Retrieved October 22, 2022, from https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-age.html#:~:text=Sample%20interpretation%3A%20Compared%20with%20ages,to%2029%20years%20age%20group.)

[2] Meyerowitz, E. A., Richterman, A., Gandhi, R. T., & Sax, P. E. (2020, September 17). Transmission of SARS-COV-2: A review of viral, host, and environmental factors. Annals of internal medicine. Retrieved September 14, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7505025/

[3] Frazer, J. (2021, December 16). *The risk of vaccinated COVID transmission is not low*. Scientific American. Retrieved October 22, 2022, from https://www.scientificamerican.com/article/the-risk-of-vaccinated-covid-transmission-is-not-low/

[4] Andrejko, K. L., & Pry, J. M. (2022, February 10). Effectiveness of face mask or respirator use in indoor public settings for prevention of SARS-COV-2 infection - California, February–December 2021. Centers for Disease Control and Prevention. Retrieved October 22, 2022, from https://www.cdc.gov/mmwr/volumes/71/wr/mm7106e1.htm

[5] Ganasegeran, K., Jamil, M. F. A., Ch'ng, A. S. H., Looi, I., & Peariasamy, K. M. (2021, September 18). *Influence of population density for covid-19 spread in Malaysia: An ecological study*. International journal of environmental research and public health. Retrieved October 22, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8468130/

[6] Ge Y;Martinez L;Sun S;Chen Z;Zhang F;Li F;Sun W;Chen E;Pan J;Li C;Sun J;Handel A;Ling F;Shen Y; (2021, October 1). Covid-19 transmission dynamics among close contacts of index patients with covid-19: A population-based cohort study in Zhejiang Province, China. JAMA internal medicine. Retrieved September 14, 2022, from https://pubmed.ncbi.nlm.nih.gov/34424260/

[7] Zhang, Y., Tao, Y., Shyu, M.-L., Perry, L. K., Warde, P. R., Messinger, D. S., & Song, C. (2022, February 23). Simulating COVID19 transmission from observed movement. Nature News. Retrieved September 14, 2022, from https://www.nature.com/articles/s41598-022-07043-4

[8] Simoes, E. J., Schmaltz, C. L., & Jackson-Thompson, J. (2021, December). *Predicting coronavirus disease (COVID-19) outcomes in the United States early in the epidemic*. Preventive medicine reports. Retrieved October 22, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8545716/#:~:text=The%20reduced%20model%20had%20statistically,unit%20increase%20in%20population%20density

[9] World Health Organization. (2020, February 24). *WHO China joint mission on COVID-19 final report*. World Health Organization. Retrieved October 22, 2022, from https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf