

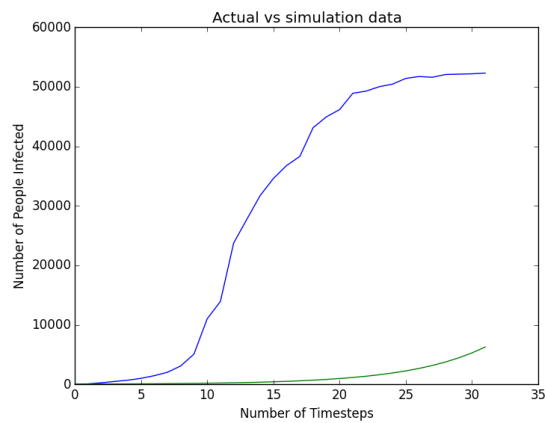
In recent years there has been an interest in using computer science as a tool in biology – including the development of breast cancer detection, interpreting the human genome for diseases, and modeling the growth of metastasis in humans. Based on this, computer science could also be used to model the spread of a disease in the United States. This project aimed to model the growth of the flu in ten different regions in the United States over the annual flu cycle. Generally, any disease can be modeled through equation (1):

$$growth = (p)Current\ Population \quad (1)$$

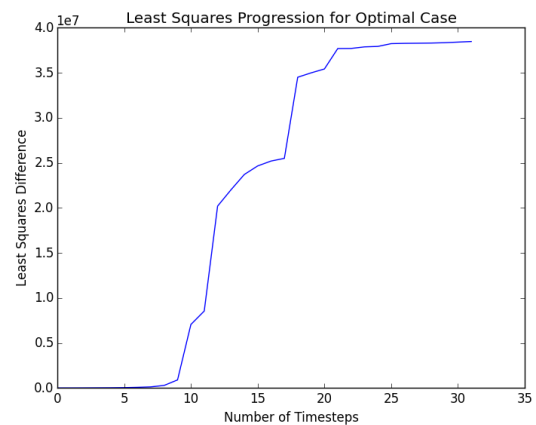
Where  $p$  is the probability that an individual will transmit the disease to another person.

For the simulation, the actual datasets of the number of people who had the flu were obtained from the Center for Disease Control (CDC), and the first week flu count was used as the starting population in the simulation. The model attempted to achieve the optimal growth of a disease by iterating through probabilities in the range of 0-1 for each region, where with each growth cycle in the region, the probability remained constant. The optimization was achieved through a minimization of the least squares values, in which the growth values of the calculation were compared against the actual growth values of the flu in the United States (equation 2):

$$Least\ Squares\ Value = (Actual\ growth - Simulated\ growth)^2 \quad (2)$$



(a)



(b)

**Fig 1 a)** Graphs the cumulative number of people who have had the flu over time, where the blue trace represents the real population in the United States and the green trace is the calculated optimal population **b)** Depicts the progression of the least squares value over time.

The optimal probability case is depicted in figure 1 a) and the progression of the least squares value over time is depicted in figure 1 b). Both indicate that there is a significant difference between the calculated cumulative flu growth and the actual growth of the flu over time.

Overall, the study attempted to model the growth of the flu over time in the United States. However, due to the simplicity of the model, an exponential function, the model failed to well describe the growth of the flu – as the true model of the growth of the flu is a logistic function. A better simulation to conduct in the future would be to depict each person with the flu as a probability density function of spreading the flu, rather than a discrete number. Through the stochastic simulation, a better result and a better least squares value could be achieved. Although the project did fail to model the disease well, it is still possible to model the disease spread with a better methodology.