The “getSCurve” portion of our code proved to be the most troublesome but also had the most potential for creativity. After several tries, this is what we came up with.

We first construct the necessary variables, including the final SDLM double array used to calculate the future cases and the “factor” array used to create a starting point for the SDLM calculations. The “factor” array holds four values for SDLM, respectively, that are arbitrary and serve as increments for the for-loops to perform on. Each variable begins at the lower bound value assigned to it via the “paramLowerBound” double array. Using a for-loop, these values are increased by the respective “factor” and runs as such until it has reached the equivalent of its assigned upper bounds. This creates a range of SDLM values that can then be tested after each iteration. Once these have been created, the SDLM values are plugged into the following equation:

$$s(d) = S + \frac{M}{1 + e^{-L(d-D)}}$$

to calculate the accumulated amount of infected at day ‘i’. We then subtract this from the actual value from that date assigned as pastData[i]. This repeated the every value of past data we have, in this case around 100 days. Here, we are beginning to calculate mean square error (MSE) between our projected S-curve against the past data we already have using the equation:

$$MSE = \frac{\sum_{d\ in\ \Lambda}\left(s(d) - r(d)\right)^2}{\Lambda}$$

where “Λ” is the range of days, day is the day, s(d) is our projected amount of infected on day d and r(d) is the actual value on day d. Once these for loops have run, we compare all of the MSE values from all of the SDLM values which were calculated from the range between the upper and lower bounds. Once we have found the set of SDLM values to with the smallest MSE, they become a part of the final SDLM array mentioned earlier. These values are then used to calculate forward into time from day hundred plus however many is input when the method is called.

Interpretations of the SDLM values are useful for completing this task. S is essentially the initial value of infected on the first day of the data set. D is the day that the center of the S-curve occurs, or essentially where the rate of increase is maximized. L refers to the steepness of the curve. M is the value the determines the upper bound as M+S will never be reached by this equation.