

## ME 2450 Lab 08: System Modeling

### Objective

The transmission of pathogens in agricultural and natural plant ecosystems can have a profound impact on the security of the food supply, the economics of agricultural production, and the general health and well-being of the population. Examples ranging from the Irish Potato Famine of 1845, the banana wilting disease outbreak of the 1890s, to the more recent outbreaks of Citrus Greening impacting orange production in Florida all illustrate the severe economic and social costs plant pathogens can have on society.

This lab will introduce concepts related to modelling the spread of an epidemic in a susceptible population. Our epidemic model will take the form of a system of linked ordinary-differential equations (ODEs) which must be integrated to represent the spread of disease. Here, we will perform this integration using Euler's method.

### References

- An introduction to function handles, *FunctionHandles.pdf*.\*
- Optional arguments in Matlab and Python, *OptionalArgs.pdf*.\*
- Lecture 13 in the course lecture notes.
- Chapter 1.1, *Chapra, Numerical Methods for Engineers (7th Edition)*.

### Model Description:

Starting with the pioneering work of van der Plank, modeling epidemic outbreaks caused by plant pathogens has been a continuous area of study. Here, we will learn the basics of how to model the spread of a pathogen starting in a homogeneous population with fixed environmental conditions. The basic model starts with the idea that our host population  $P$  (plants, animals, etc.) can be divided into four categories: *Susceptible*, *Latent*, *Infected*, or *Removed* forming the basis for a *SLIRP* model (a variant of the classical *SIR* model, note  $E$  for exposed is sometimes used instead of  $L$  and the  $P$  for population is sometimes omitted). The susceptible population are those members of a population that have the potential to be infected while the latent are those in the population that have been infected by the pathogen but are not yet infectious and the infected are those that have moved past the latent period and are now infectious and spreading the disease to others in the population. The removed/recovered category is used for those who can no longer be infected because of immunity either prior to the beginning of the epidemic or gained during the epidemic (e.g., because of age) or because they have been infected, and the infection has run through its effective lifecycle in the pathogen resulting in either death or immunity.

For plant pathogen systems, the total population  $P$  is governed by the plant's growth rate. The growth of most plants follows a logistic equation, and we can write that the change (increase) of plant material in general as

$$\frac{dP}{dt} = k(T, t)$$

where  $P$  is the population size with units of surface area and  $k$  is the rate of growth which is a function of at least the temperature  $T$  and for some models also on the number of days  $t$  since a particular event (e.g., budbreak for deciduous plants).

The exact form of  $k(T, t)$  depends on both the type of plant to be modeled and on the source of data used to prescribe constants and fit empirical functions to. In this lab we will use one of the simplest and most common forms, a degree day model (or growing degree day, or thermal time, etc.). A degree day is typically defined as the integrated amount of temperature during which a plant is actively growing so that the degree day temperature (or thermal unit) is defined as

$$T_{\text{unit}} = \begin{cases} (T - T_{\min}) & \text{if } T_{\min} < T < T_{\max} \\ 0 & \text{if } T < T_{\min} \text{ or } T > T_{\max} \end{cases}$$

and our total population equation is

$$\frac{dP}{dt} = kT_{\text{unit}} \quad (1)$$

where  $T_{\min}$ ,  $T_{\max}$ , and  $k$  are all dependent on the type of plant. For many plants  $T_{\min} = 10^\circ\text{C}$  and  $T_{\max} = 35^\circ\text{C}$ , with  $k$  taking a range of values depending on how fast a particular plant grows, soil moisture and nutrient conditions, pest and pathogen pressures, etc.

The other components of the system can be modeled similarly as rate equations (ordinary differential equations). The rate of increase of infected population is assumed to follow a logistical model (e.g., interest earned) so that the susceptible population can be represented as:

$$\frac{dS}{dt} = -(\beta SI + e) + \frac{dP}{dt} \quad (2)$$

where  $S$  is the fraction of the population with the potential to be infected,  $I$  is the fraction of the population that is currently infected and spreading,  $\beta$  is the transmission rate (days) of the disease,  $e$  is the rate of new infections from external sources, and  $dP/dt$  comes from equation (2). It is worth noting that in this simplified framework the distinction between the pathogen and the disease it causes is lost.

The population fraction that is newly infected moves to the latent phase where it has been exposed to the pathogen but has not yet become symptomatic and started to spread to other members of the susceptible population. The equation for the latent population  $L$  mirrors equation (3) for  $S$  but with an additional term to represent the transition from latent to infectious:

$$\frac{dL}{dt} = (\beta SI + e) - \mu_L^{-1}L \quad (3)$$

where  $\mu_L$  is the length of time before the latent population becomes infectious (e.g., if the latent period is 5 days,  $\mu_L^{-1} = 0.2/\text{day}$ ).

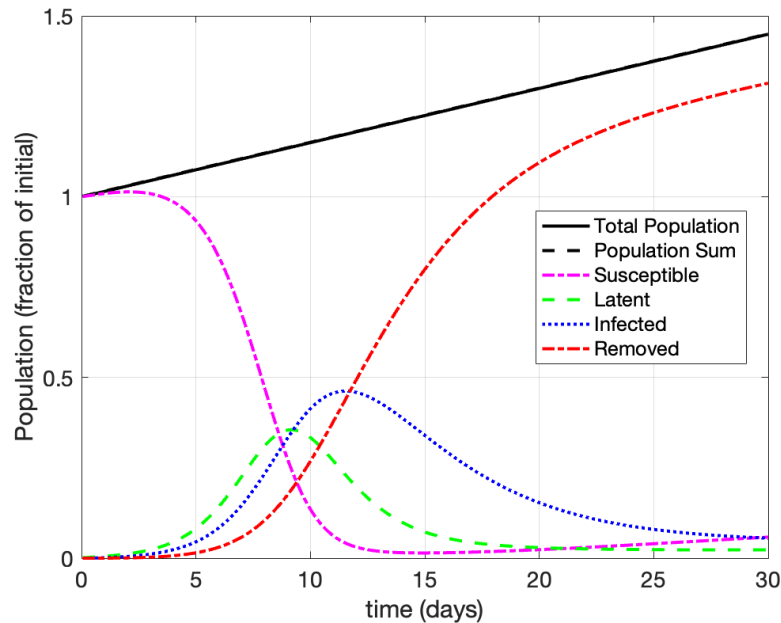
The infectious population  $I$  links directly to the latent population through the 2<sup>nd</sup> term on the left-hand side of equation (4):

$$\frac{dI}{dt} = \mu_L^{-1}L - \mu_I^{-1}I \quad (4)$$

where one new parameter is introduced related to the length of the infectious period  $\mu_I$ . Strictly,  $\mu_I$  is how long it takes before the infectious population moves to the removed/recovered stage which on average is equal to the length of an infection. Although not strictly needed to complete the model, the rate of increase for the removed portion of the population can be defined as

$$\frac{dR}{dt} = \mu_I^{-1}I \quad (5)$$

which corresponds to the last term on the right-hand side of equation (5). If we sum the contribution from the  $SLIR$  categories, we recover the total population so that  $P = S + L + I + R$ . An example of this system for a set of parameters (see figure caption) is given in **Figure 1**.



**Figure 1.** Example plot of the SILRP model system for  $\beta = 2$ ,  $\mu_L = 2$ ,  $\mu_I = 4$ ,  $k = 0.001$ , and  $e = 0.005$ . The value of  $\beta$  controls the rate of spread (steepness of the  $S$  and  $I$  curves) and the values of  $\mu_L$  and  $\mu_I$  the shift in when the falloff in  $S$  (and rise in  $R$ ) happens and the lag (separation) between the  $L$  and  $I$  curves. The growth rate of the population is governed by  $k$  which determines the slope of the total population line which is linear here because  $k$  and  $T$  are constants.

Many aspects of disease transmission are omitted from a model like this, but it is a useful first description of the dynamics of a plant disease epidemic (or any epidemic for that matter). In the project we will subsequently introduce common aspects including heterogeneity in the susceptible population in both distribution and resistance to disease, nonlocal and nonlinear transmission of the pathogen, and lastly realistic changes in the environmental conditions (e.g., wind, temperature, humidity) that the pathogen and host develop in.

## Numerical Methods

### Euler's Method

The model consists of a set of four linked ordinary differential equations (ODEs) plus one that defines the fraction of the population that is recovered/removed ( $R$ ). The four ODEs must be solved simultaneously with an appropriate numerical integration scheme. Because this system is zero dimensional (0-D, time dependence only) we can use a small timestep with a relatively straightforward method. Here, we will use the Euler-step method.

The basic Euler method can be written as:

$$y(t_{i+1}) = y(t_i) + \left. \frac{dy}{dt} \right|_{t=t_i} (t_{i+1} - t_i)$$

where  $t_i$  is the time at time point  $i$  and  $y$  is our variable of interest (what we are integrating, i.e.,  $S, L, I, R$  or  $P$ ).

Our equations are linked, and therefore we must satisfy them all at each time point simultaneously. Our discretized system is:

$$\begin{aligned} S(t_{i+1}) &= S(t_i) + \left. \frac{dS}{dt} \right|_{t=t_i} (t_{i+1} - t_i) \\ L(t_{i+1}) &= L(t_i) + \left. \frac{dL}{dt} \right|_{t=t_i} (t_{i+1} - t_i) \\ I(t_{i+1}) &= I(t_i) + \left. \frac{dI}{dt} \right|_{t=t_i} (t_{i+1} - t_i) \\ P(t_{i+1}) &= P(t_i) + \left. \frac{dP}{dt} \right|_{t=t_i} (t_{i+1} - t_i) \end{aligned} \tag{6}$$

where the time rate of change of  $S, L, I$ , and  $P$  are given by equations (1)-(4), respectively. Equation (5) can be solved as part of this system or afterwards using saved values of  $I(t_i)$ .

### Lab Assignment:

In this assignment you will develop a code to solve the SLIRP system of equations and examine the sensitivity to parameters of interest. You should turn in the following:

1. Write a code that solves equations (1)-(5). Your code should be modular and use functions. You may use functions/code developed as part of your previous labs and/or homework assignments. Because you will later integrate this with a larger system, the use of functions is highly desirable. You should develop:
  - a. (10 Pts.) A function to calculate  $\left.\frac{dS}{dt}\right|_{t=t_i}$ ,  $\left.\frac{dL}{dt}\right|_{t=t_i}$ ,  $\left.\frac{dI}{dt}\right|_{t=t_i}$ ,  $\left.\frac{dR}{dt}\right|_{t=t_i}$ , and  $\left.\frac{dP}{dt}\right|_{t=t_i}$  based on equations (1)-(5) with the appropriate functional dependencies. The function should take as inputs the variables  $S$ ,  $I$ ,  $L$ ,  $R$ , and  $P$  at time  $t_i$ , and the parameters  $\beta$ ,  $k$ ,  $T$ ,  $T_{\min}$ ,  $T_{\max}$ ,  $e$ ,  $\mu_L$ , and  $\mu_I$ . It should output the derivatives. This function will need to be compatible with your Euler function developed in part b.
  - b. (8 Pts.) A function that uses Euler's method to integrate the system given by equations (1)-(5) (see equation (6) for the discretized version). As inputs the function should accept a function handle to the system of slope equations (ie part a.), initial conditions  $S_{\text{init}}$ ,  $L_{\text{init}}$ ,  $I_{\text{init}}$ ,  $R_{\text{init}}$ , and  $P_{\text{init}}$ , and simulation parameters for the total length of the simulation and the time step (or equivalently time points for integration).
2. Create a driver script (or function) that calls your Euler integration function and use it to explore the impact of the five epidemiological parameters ( $\beta$ ,  $\mu_L$ ,  $\mu_I$ ,  $k$ , and  $e$ ) on the development of an epidemic. For this lab we will assume that  $T_{\min} = 10^\circ\text{C}$ ,  $T_{\max} = 35^\circ\text{C}$ , and that the temperature  $T$  is constant throughout the simulation at  $25^\circ\text{C}$ . To match Figure 1 use initial conditions:  $P_{\text{init}}=1$ ,  $S_{\text{init}}=P_{\text{init}}$ ,  $L_{\text{init}}= 0.001S_{\text{init}}$ ,  $I_{\text{init}}= 0$ , and  $R_{\text{init}}= 0$ .
  - a. (6 Pts.) creation of a driver script. Your script should call the Euler function for the SLIRP system for a given parameter set with an appropriate simulation length (number of days) and timestep. The length of a simulation should be enough to observe the full progression of the epidemic (e.g., peak of the infectious population and its decline). Your function should also make a plot similar to Figure 1 for later analysis.
  - b. (3 Pts.) using the parameter set from Figure 1 ( $\beta = 2$ ,  $\mu_L = 2$ ,  $\mu_I = 4$ ,  $k = 0.01$ , and  $e = 0.005$ ) as a basis, run your simulation for three different values of  $\beta$ . You should have three plots each with a different  $\beta$  value while other parameters are held constant at the base parameter values.
  - c. (3 Pts.) using the parameter set from Figure 1 ( $\beta = 2$ ,  $\mu_L = 2$ ,  $\mu_I = 4$ ,  $k = 0.01$ , and  $e = 0.005$ ) as a basis, run your simulation for three different values of  $\mu_L$ . You should have three plots each with a different value of  $\mu_L$  while other parameters are held constant at the base parameter values.

- d. (3 Pts.) using the parameter set from Figure 1 ( $\beta = 2$ ,  $\mu_L = 2$ ,  $\mu_I = 4$ ,  $k = 0.01$ , and  $e = 0.005$ ) as a basis, run your simulation for three different values of  $\mu_I$ . You should have three plots each with a different  $\mu_I$  value while other parameters are held constant at the base parameter values.
  - e. (3 Pts.) using the parameter set from Figure 1 ( $\beta = 2$ ,  $\mu_L = 2$ ,  $\mu_I = 4$ ,  $k = 0.01$ , and  $e = 0.005$ ) as a basis, run your simulation for three different values of  $e$ . You should have three plots each with a different  $e$  value while other parameters are held constant at the base parameter values.
3. (4 Pts.) Using the 12 plots you created in 2., discuss the impact of each of the four parameters ( $\beta, \mu_L, \mu_I, e$ ) on the development of an epidemic. Describe what each parameter does to change the evolution of the epidemic. Based on your observations of the impact of these parameters, which parameter would you target to reduce the spread of the disease?

**Submission:**

Your submission to Gradescope should include at least two files. One file should be a PDF that includes all required figures (12 total) with proper captions and the answer to question 3. Your code file (or files) should be included as a separate file(s) in its native format (e.g., .m, .py). To receive full credit, you must include both the code file (or multiple files if you save functions separately) and the PDF of the figures with captions and explanations. For an example caption see Figure 1 of this document.