MA 331
Fall 2017
Due: 12/01/17

Semester Project

Summary: The goal of this project is to apply everything learned this semester in MA 331: Differential Equations for the Life Sciences towards a real-world application of your interest. Depending on the specific biological phenomenon that you choose to model, you will use one or more of the following techniques: finding the solution, vector fields, phase planes, determining equilibria and their stabilities, and parameter estimation.

Background: Throughout the semester, we have discussed the applications of differential equations to various biological systems such as tumor growth, predator-prey dynamics, population dynamics, and epidemiology. Within these models, we have discussed the assumptions made by the model. We have solved the differential equations, where possible, and introduced qualitative methods of system analysis (finding equilibria and determining stability, phase planes, vector fields, etc) when we are unable to directly compute the solution. During the latter portion of the course, we introduced the concept of the *inverse problem*, where we attempt to estimate the parameters of with our model by fitting the model to data.

Project Requirements: Although there is some flexibility in this project (you may choose a topic of your interest), the backbone of the project will be similar. The following will be required for your project:

1.) *Introduction.* Explain the biological system you are examining. Why do you want to understand the dynamics of the system? How do you think being able to mathematically model such a system will be beneficial to biological understanding?

2.) *Model Formulation.* What is the mathematical model that you will be using? What are the assumptions that are being made (i.e., exponential growth, logistic)? Why do you feel these assumptions are valid for the system you are studying? Explain the data you are using, how it was obtained, and why the assumptions you made are valid for the data. What are the limitations to the assumptions you have made? How realistic do you believe the assumptions are?

3.) *Model Analysis.* Are you able to solve the differential equation(s) outright? If so, document your solution process. Consider qualitative approaches: what are your equilibria and their stabilities? Draw phase line (or planes). Under what parameter conditions will you get differing behavior?

4.) *Parameter Estimation.* Use data to find the parameters that best fit your model (using simbiology). Describe how you performed the parameter estimation. *Include units.* What is your confidence in those parameters? What is the sensitivity of the parameters you optimized? Given your set of best fit parameters, what is the long-term behavior of your system (you may draw a phase plane/vector field with your best-fit parameters)?

5.) *Conclusions.* What did you learn about the system? How good was your fit to data? What improvements could be made for future models?

Structure: The final report should be approximately 5–10 pages (including figures), written in type. You may write your report in any word-processing software you choose, but you will submit your final report as a .pdf via email. Equations should be numbered (or labeled) where appropriate. Figures and/or tables should have captions, (e.g., "Figure 1. Phase plane analysis for model equations (equation 1) using..."). All mathematics should be clearly explained and all steps shown as in homework. All cited references (if applicable) should be listed at the end in MLA format.

Grading: The grading of this project is based on the clarity of your explanations, the correctness of your mathematics, and inclusion of all requested materials. You will *not* be graded on how well your model fits the data.

	POOR (1)	ACCEPTABLE (2)	GOOD (3)	OUTSTANDING (4)
Introduction	Student does not motivate or introduce their biological system. Little or no outline of the report.	Student perfunctorily introduces the system, does not discuss why they are interested or does not outline the report.	Student discusses their biological system but does not elaborate on why it interests them or does not outline the report.	Student properly motivates their biological system and discusses importance. Student outlines the remainder of the report.
Model Formulation	Model assumptions and/or equations are not given. Parameters and variables are not defined and/or units not provided.	Model assumptions are unclearly stated. Variables and parameters are discussed but units may be missing. Very little discussion of justification or limitations of assumptions.	Model assumptions are stated. Variables and parameters are discussed and units provided. Limited discussion of justification or limitations of assumptions.	Model assumptions are stated and justified. Variables and parameters are discussed and units provided. Discusses possible limitations of assumptions.
Model Analysis	Model analysis is incomplete and/or incorrect. Little or no discussion of biological implications of results.	Model analysis consists of one method. Student may or may not discuss biological implications of their results. Analysis may be performed with non-generic parameters.	Student uses multiple methods to analyze model (i.e., drawing phase plane and determining equilibria and stabilities) but does not relate their results back into biological understanding	Student uses multiple methods to analyze model (i.e., drawing phase plane and determining equilibria and stabilities) and relates their results back into biological understanding.
Parameter Estimation	Parameter estimates are provided but units are incorrect/missing and methodology is not clear	Parameter estimates are provided but units are incorrect/missing or methodology not clearly explained.	Parameter estimates are provided with units, but student does not sufficiently extend their results or does not discuss methodology	Student clearly explains how they obtained parameter estimates. Units are provided, student extends their results.
Conclusions	Student recap is disorganized, disconnected, or incomprehensible. Little or no future directions suggested.	Student briefly states results but fails to make connections from the mathematics to the biology. Little or no suggestions for future work.	Student recaps the biological system and summarizes results but does not discuss future avenues to explore.	Student recaps the biological system, summarizes results and provides future directions which could be investigated
Style	Paper is disorganized. Figure and table captions may be missing. Major spelling or grammatical errors. References not cited.	Paper is somewhat organized. Figures and tables are given non-descriptive captions. Minor spelling and grammatical errors. References may not be cited correctly.	Paper is clearly organized and flows. Figures and tables are given captions. Minor spelling and grammatical errors. References cited appropriately.	Paper is clearly organized and flows. Figures and tables are given captions. Few or no spelling and grammatical errors. References cited appropriately

Project Topics

Here are some examples of project ideas that you can use for your final project. If you are interested in a different project topic, please make sure to *see me* by **November 16**, so we can discuss finding/acquiring data for you to use. I highly encourage you to pursue an alternate project if there is some system you are very interested in.

Within each project topics, I have also listed possible project ideas, but feel free to pursue a topic of your interest and choice. These are just starting blocks – you may extend these ideas or choose entirely different ones.

Project 1: Predator-Prey
The Canadian lynx preys almost exclusively on the snowshoe hare. In the 19th century, Hudson's Bay Company began tracking the populations of pelts obtained by trappers for both the lynx and the hare. The dataset, from 1845-1903 can be found here.

- Extend the predator-prey system that was discussed in class to incorporate logistic growth. Analyze the model and fit to the data.
- Compare the ability of various predator-prey models to fit the data. Which generates the best fit? Why do you think it does?

Project 2: Epidemiology
In 2014 an ebola outbreak occurred in West Africa. Ebola virus disease (EVD) is high-mortality illness that is transmitted by human-to-human contact. On average, fatalities occur in approximately 50% (range 25% to 90%) of human cases. The World Health Organization compiled a list of total cases and total deaths from EVD in Guinea, Liberia, and Sierra Leone from March 1, 2014 to February 17, 2016 [1]. Here are the datasets for Guinea, Liberia, and Sierra Leone.

- Use one model to fit all three datasets independently. How do the parameters change for each country? What does that mean in a biological sense?
- Use multiple models to fit one of the datasets. Which model generates a good fit?

Project 3: Ecology
Daphnia magna is a species of water flea used in toxicity testing because they are sensitive to environmental change. The data included here investigate the effect of food restriction on daphnid growth (as measured by the major axis length of the animal) [2]. The experiment is performed over 14 days and measurements occurred daily. The mean over approximately 30 individual daphnids are reported. Here are the datasets for control, medium, and low food levels.

- Use one model to fit all three datasets independently. How do the parameters change for each food group? What does that mean in a biological sense?
- Use multiple models to fit one of the datasets. Which model generates a good fit? Why do you think this model is successful in obtaining a good fit?

Project 4: Cancer Growth

According to the American Cancer Society, in 2017, approximately 22,440 women will be diagnosed with ovarian cancer and about 14,080 women will die from ovarian cancer [3]. The data provided come from an experiment where tumor cells were injected into mice [4] and tumor volumes were measured throughout time. When tumors start growing large, they produce angiogenic factors which induce vascularization, providing more nutrients to the tumor. Data were also collected for mice who underwent anti-angiogenic treatment (blocking the vascularization from forming). Here are the datasets for control and treatment.

- Use one model to fit the untreated cancer dataset. Then extend your model to incorporate a term modeling death due to treatment. Using the optimized parameters from the untreated cancer dataset, try to fit the new model to the treatment cancer dataset.

- Use multiple models to fit the untreated cancer dataset. Which model generates a good fit? Why do you think this model is successful in obtaining a good fit?

References

[1] "Ebola (Ebola Virus Disease)." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 17 Feb. 2016, www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/cumulative-cases-graphs.html.

[2] Rutter, Erica M., et al. "Continuous Structured Population Models for Daphnia Magna." *Bulletin of Mathematical Biology*, 2017, doi:10.1007/s11538-017-0344-8.

[3] "What Are the Key Statistics About Ovarian Cancer?" *American Cancer Society*, The American Cancer Society, www.cancer.org/cancer/ovarian-cancer/about/key-statistics.html.

[4] Mesiano, Sam, Napoleone Ferrara, and Robert B. Jaffe. "Role of vascular endothelial growth factor in ovarian cancer: inhibition of ascites formation by immunoneutralization." *The American Journal of Pathology* 153.4 (1998): 1249-1256.