Mentoring Undergraduate Research Students in Mathematical Modeling

Glenn Ledder

Department of Mathematics
University of Nebraska-Lincoln
gledder@unl.edu

October 20, 2025

Part 1: Undergraduate Research

1. Undergraduate research is a process, not a product (ie. a verb, not a noun).
   - Some characteristics of professional research must be modified for undergraduates.

2. Successful projects depend on positive student attributes.

3. Undergraduates need their mentors to structure their experience.

4. Dissemination is different for undergraduates, but still crucial.

A Process, not a Product

Professional Research

- The ultimate goal is publication, so publishability is critical.
- There are no deadlines that limit scope or thoroughness.

Undergraduate Research

- The ultimate goal is the research experience, so publishability is not required.
- Projects have predetermined end dates.

Research Characteristics: Direction and Originality

- Professional research is open-ended.
  - The research direction is determined by the questions and intermediate results.
- Undergraduate research is also open-ended.
  - Mentors cannot script a research project like they do for classroom projects.
- Professional research is original.
  - Only results not previously published count.
- Undergraduate research is also original, but ...
  - "Original" means only that the students discover the results for themselves.

Research Characteristic: Importance

- Professional research is important.
  - Reviewers recommend publication only if they think the work is a significant contribution to the literature.
- Undergraduate research only needs to be interesting.
  - Some undergraduate research is whimsical.
    - Modeling of zombie vs humans scenarios has a large literature, including undergraduate research as well as professional pedagogy.

Research Characteristic: Scope

- Professional research is thorough.
  - No loose ends that could be tied up with a little more work.
- Undergraduate research is (usually) not thorough.
  - Work stops at the deadline.
  - Partial results are acceptable.
  - Problems can be restricted to special cases.
    - Example: Flake, Hoang, and Perrigo completed work on just one of two subcases of their problem.[1]

[1] C. Flake, T. Hoang, E. Perrigo, A predator-prey model with disease dynamics, Rose-Hulman Undergraduate Math Journal, 4 (2003)

My Motivating Example

When I was a college freshman, I wanted to rank the color groups in Monopoly from best to worst. I knew I needed to know the expected value of each property with a hotel. The problem was finding the probability of landing on each square. I found this by constructing a Markov chain transition matrix and writing a computer program to iterate it to convergence. I had never heard of Markov chains or matrices; in effect, I developed a limited version of the mathematical concepts myself.

1. My problem was open-ended.
2. It was clearly not important, but at least a little interesting.
3. It was original in the sense of my discovering it for myself.

- Probably no mathematics professor in 1974 would have said I had done research. It never occurred to me to ask.

Attributes for Success: Commitment

- Many students take on a larger workload than they can manage: too many credits, work hours, and social activities.
- There are three good solutions to this problem:
  1. Do research in the summer if funding is available.
  2. Do research for independent study credit during an academic semester.
  3. Design a research course that can be delivered to small classes.[2]

[2] See Radu, Taking Math Outside of the Classroom: Math in the City, *PRIMUS*, **23**, 538–549 (2013)

Attributes for Success: Preparation

- When possible, pick projects that require relatively little background beyond calculus.
- Even if students have had the necessary background coursework, this does NOT mean they have the necessary background knowledge and skills.
- **Mentors should plan to start a project with a crash course on the necessary knowledge and skills.**
    - Example: For analysis of continuous-time dynamical systems, students need to understand equilibria and stability and be able to calculate a Jacobian and use the Routh-Hurwitz conditions.[3]

[3] See Ledder, *Mathematical Modeling for Epidemiology and Ecology*, Springer (2023), for example

Attributes for Success: Ownership

- Most students have only worked on problems assigned by an authority figure.
  - The authority figure is the owner of the problem.
  - Students get much more out of research if they are the owner(s).
- Foster ownership by giving students a role in choosing their project.
  - Choose a general topic and collaborate with the student(s) to choose the specific problem.
- If students are working in groups, it is alright for each to focus on a particular task, but make sure each knows enough about the whole project to give a brief presentation.

Attributes for Success: Maturity

- Most research students have only done textbook problems, which have a clear plan of attack and a clear record of successful solution.
- Not knowing how to attack a problem and not knowing a problem has been solved before can be discouraging.
- Partial solution: Require a sequence of progress reports.
  - Multiple submissions allows students to give themselves credit for small accomplishments.
  - A multiple submission plan helps mentors structure projects as a sequence of problems, starting simple and adding complexity.

Dissemination and Writing

- Dissemination is essential. It need not be a mentor/student professional paper; other good choices are a student-only paper, an oral conference presentation, a poster, or even just a Math Club presentation.
  - Whichever is the final deliverable, there should be both oral and written reports (even if brief) during the project.
- Assume that your students do not know how to write about mathematics, even if they are good writers in general.
  - Provide short examples of A and C+ student writing for students to critique.
    - Don't identify which is which, but assign the task of identifying which is better and why to the students.

Part 2: Mathematical Modeling

1. Reasons for choosing projects in mathematical modeling
2. Mentors for modeling projects do *not* need expertise in modeling.
    - They *do* need to understand "modeling theory," the key points of which are
        - Mathematical models are functions that map a set of parameters to one or more outcomes.
        - Results about models are obtained from simulations and parameter studies.

Why Choose **Modeling** Projects?

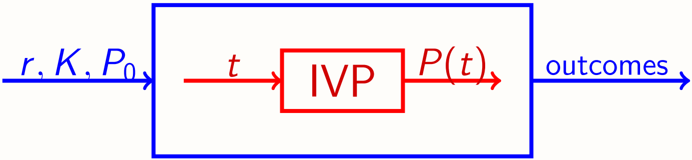
1. Many modeling topics require only ordinary differential equations and matrix theory as background.
    - A crash course of about 10 hours can teach the necessary material to students with a Calculus 1 background.
        - Equilibria and stability of systems of ODEs.
        - Jacobians and the Routh-Hurwitz conditions.
    - Some meaningful projects do not even require calculus.
2. Modeling has a broad appeal to intellectual curiosity if the topic is of interest to the general public.
    - Epidemiology and immunology are especially good.
    - Whimsical topics like zombies vs humans appeal to many students.

Models as Functions of Parameters

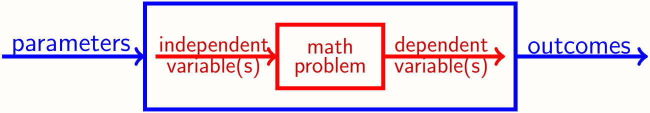
How do we view the model

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{K}\right), \quad P(0) = P_0 > 0, \qquad r, K > 0?$$

- **Narrow** view:
  Initial value problem for $P(t)$, with parameters r, K, and P_0.
- **Broad** view:
  Function that maps parameters r, K, and P_0 to outcomes.



Models as Functions of Parameters



- **Narrow** view: Math problem with **fixed parameters**.
  - The narrow view is used to determine the outcomes.
    - Narrow view questions are trivial: "Given $K = 10$, $R = 1$, and $P_0 = 1$, when does the population reach $P = 5$?"
- **Broad** view: Outcomes as **functions of parameters**.
  - The important questions are in the broad view.
    - Do solutions with any initial condition always approach K?
    - At what point is the population growth the fastest?

A Textbook-Style **Modeling** Problem

The susceptible and recovered population fractions in the SIR disease model (cite worldbank.org article) are related by the initial value problem

$$\frac{dS}{dR} = -\mathcal{R}_0 S, \qquad S(1-s_0) = s_0, \quad S(1-s_\infty) = s_\infty,$$

where s_0 and s_∞ are the initial and final susceptible population fractions, R is the recovered population fraction, and $\mathcal{R}_0$ is the *basic reproduction number* of the disease.

1. Explain why the initial and terminal conditions are evaluated at $1 - s_0$ and $1 - s_\infty$.
2. Plot a graph showing how the final susceptible population depends on $\mathcal{R}_0 \in (1, 8]$ for the case where the entire population is initially susceptible.

Footnote: Some Textbooks with a Modeling Focus

Many textbooks for ODE courses use "Modeling" in their titles. Don't be fooled. Most have "applications" but no real modeling.[4]

- Bryan, *Differential Equations: A Toolbox for Modeling the World*, 2nd ed., SIMIODE ($39 online)
  - The only ODE textbook with my personal seal of approval.
    - I'm not being paid, and I only met the author last week.
- Ledder, *Mathematical Modeling for Epidemiology and Ecology*, 2nd ed., Springer (2023)
  - Can be used for a mathematical biology course or as a resource for ODE course projects.

[4] "Applications" present equations without identifying them as models, ask only narrow view questions, and do not incorporate reality checks.

Project Design: Modeling Tasks

- **B**uild
  - Use mechanistic assumptions to construct a model of a scenario or experiment.
- **A**nalyze
  - Use *qualitative methods* to obtain outcomes such as equilibrium solutions and stability.
- **S**imulate
  - Use software such as MATLAB or R to obtain numerical solutions.
- **I**nterpret
  - Address general questions in the context of the model setting.
- **C**ritique
  - Discuss possible model limitations.

Project Design: Choosing Outcomes

Epidemiology Example

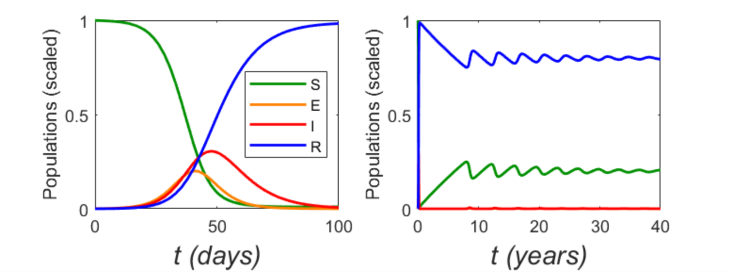
- Maximum number of new infections?
- Maximum number of hospitalizations per million? (compared to an average of 2800 hospital beds per million)
  - Serves as a measure of the stress on the health care system
- Percent deaths? (0.06% is 200,000 people)
  - Serves as a measure of the human cost
- Final fraction of susceptibles?
  - Serves as a measure of the risk of a new outbreak
- Times for any of these events?

Project Design: Asking and Addressing Questions

- Models must be designed to answer specific questions.
  - If we want to know the impact of COVID-19 on health care resources, we need to modify a standard disease model to track hospitalizations and/or ICU patients.
- Some common question types:
  - Is a specific claim supported by modeling or not?
  - What effect does parameter x have on outcome y?
- Strategies for addressing questions
  - Run simulations for several values of a parameter and compare simulation plots.
  - Calculate an outcome y for a large set of values of parameter x and plot y vs x.

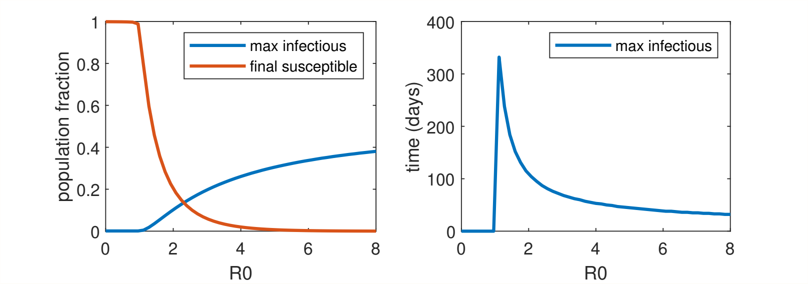
Simulations

- Simulations require known parameter values.
  - Plots show model behavior.
  - Results only guaranteed to apply to the specific parameter set.
  - Multiple simulations can *help* us understand the generality of the results and the effects of the parameters.



Parameter Studies

- Parameter studies systematically explore the impact of a parameter on one or more outcomes.
    - They show outcome vs parameter, not individual simulation results.



Part 3: Two Examples

1. A Mathematical Model of the Immune System
   - Requires topics from ODEs and Matrix Theory.
   - Drawn from a collection of exercises in my modeling book.
   - Used very successfully with a very capable sophomore who had completed an ODE course.
2. HPSR Epidemiological Model
   - Agent-Based Modeling; does not require any calculus.
   - Appears in a volume of contributed articles on projects without calculus (reference given below).
   - Used very successfully with teachers in a continuing education course.

Immune System: Origin and Background

Origin:

- I heard a research talk on the subject and found the associated research paper.
- I created a simplified version by decreasing the number of components.
- I included the simplified version in *Mathematical Modeling for Epidemiology and Ecology* as an extended exercise set.

Background:

- general: dynamical system construction, scaling
- DE background: equilibria, stability, nullcline plots
- matrix theory: Jacobian, eigenvalues, Routh-Hurwitz conditions

Immune System: Model Construction

- Three populations: a (p)athogen, a (g)eneralist macrophage, and a (s)pecialist macrophage
- p grows logistically, g are produced at a constant rate, s are produced at a rate that increases as p increases
- gp encounters kill g as well as p, but sp encounters only kill p.

After scaling, we get the model

$$p' = p(1 - p - rg - ks)$$

$$g' = b(1 - g - mgp)$$

$$s' = a\left(\frac{p}{p+h} - s\right)$$

Immune System: Model Analysis

$$p' = p(1 - p - rg - ks)$$

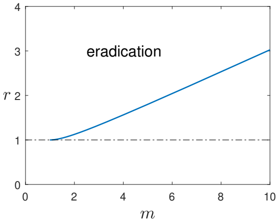
$$g' = b(1 - g - mgp)$$

$$s' = a\left(\frac{p}{p+h} - s\right)$$

- Do *pg* system.
  1. Use algebra to obtain a relationship between equilibrium p and the parameters.
  2. Use nullcline analysis and/or the Routh-Hurwitz conditions to determine stability for different cases based on inequalities in the parameter space.
  3. Determine the relationship between the 'good' parameters r and k and the 'bad' parameters m and h for eradication.
- Repeat for *ps* system and *pgs* systems.

pg Immune System: Results and Interpretation

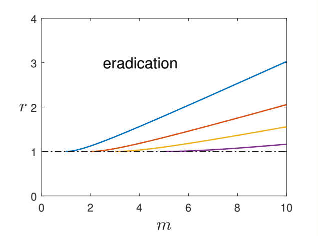
- Eradication requires $r > 1$ if $m < 1$; as m increases, the requirement for r increases, eventually becoming linear.



- The generalist can defeat some pathogens, but not all.
- *All results are for the model, not the real system.*
  - To assess the model, we need to ask experts if the model predictions match the real world.

ps and *pgs* Immune Systems: Results and Interpretation

- Eradication is impossible in the *ps* system. In the *pgs* system, eradication is easier than the *pg* system.



1. The specialists cannot defeat any pathogens on their own because they stop being produced as the pathogen level decreases.
2. With both present, the specialist significantly reduces the required generalist effectiveness.

Agent-Based Models

Definition

Agent-based models[5] (ABMs) consist of a database of individuals, each identified by one or more attributes that can change over time, and a set of stochastic rules that update the attributes of each individual at each time step.

- Attributes are qualitative or quantitative properties, with each individual having a single value at any time.
- ABMs are often implemented in NetLogo.
  - Visual Basic, R, or MATLAB can be used when there are no spatial attributes.

[5] Also called "individual-based-models (IBMs)"

Basic HPSR Agent-Based Model

Properties of the basic HPSR[6] model:

- One attribute: "status"
  1. '**H**ealthy' individuals have not yet been infected;
  2. '**P**re-symptomatic' individuals have been infected and can transmit the disease, but do not show symptoms;
  3. '**S**ick' individuals have been infected, can transmit the disease, and do show symptoms;
  4. '**R**ecovered' individuals are no longer sick, cannot transmit the disease, and cannot be reinfected.
- Progression is linear: **H** $\rightarrow$ **P** $\rightarrow$ **S** $\rightarrow$ **R** (or stay as **H**).

[6]This is the common SEIR model, with different class names because we don't want students doing internet research on 'SEIR'.

Status-Change Rules in the Basic HPSR ABM

1. At each time step, individuals in the model are randomly assigned to pairs.
2. H individuals paired with a P or S change to P with probablility 5/6.
3. P and S individuals change to S and R, respectively.

- Model output is a record of the class counts for each time step.
- These rules can be implemented as a physical simulation, either in person or virtually.

Exploring the HPSR

1. Do several simulations and describe the random differences that you see.
2. Do 10000 runs, each time recording H_∞, the number of H individuals at the end of the simulation. Plot a histogram and find the mean, standard deviation, and extreme values.
3. Discuss how random events can alter the outcome of a disease scenario. In particular, how does randomness have the greatest impact on the results.

Parameter Studies with the HPSR ABM

For each of these changes, find the mean final healthy population over 10000 runs. Some of these require a range of values for one parameter. In these cases, plot the mean H_∞ against the varying parameter.

1. Vary the transmission probability b from 0.5 to 1.
2. Add a second class of Sick individuals (SS) to represent people who are sick for two days instead of one. Instead of automatic progression from S to R, have a fraction p of S individuals transition to SS and the rest go to R directly. Have all SS individuals go to R.
3. Add isolation of the sick to the model by leaving a fraction q of the Sick individuals out of the daily pairing of contacts.

Project Resources (write to gledder@unl.edu for more)

- *Foundations for Undergraduate Research in Mathematics*, Springer book series
  - Includes *Mathematics Research for the Beginning Student, Volume 1*, ed. Goldwyn, Ganzell, Wootton (2022)
    - Includes Mathematical Epidemiology, Ledder and Homp
- *SIMIODE*, at qubeshub.org
  - A repository for projects.
- *CODEE* (at https://www.scholarship.claremont.edu)
  - A journal of DE case studies, including
    - Ledder, Qualitative Analysis of a Resource Management Model, (2021)
- *UMAP Journal* (at https://www.comap.com)
  - A journal of case studies, including
    - Ledder, Mathematical Model for a Mission to Mars, (2002)

Other Resources

Textbook (for a modeling or math bio course)

- Ledder, *Mathematical Modeling for Epidemiology and Ecology*, 2nd ed., Springer (2023)
  - Contains a large number of case studies, projects, and extended projects distibuted over multiple sections.

Undergraduate Research Mentoring

- Ledder, Mentoring Undergraduate Research in Mathematical Modeling, *Bull. Math. Biol.*, 84:77, (2022)
  - doi 10.1007/s11538-022-01040-4

Consultation

- gledder@unl.edu