1

Teaching Mathematical Biology in a Summer School for Undergraduates

Gerda de Vries and Thomas Hillen

Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, T6G 2G1, CANADA
devries@math.ualberta.ca and thillen@math.ualberta.ca

Summary. For the past four years, the University of Alberta has hosted a summer school on mathematical biology, aimed at undergraduate students who have completed 2–3 years of study in mathematics or a similar quantitative science. The aim of this summer school is to introduce the students to mathematical modelling and analysis applied to real biological systems. In the span of 10 days, students attend lectures and exercise sessions, learn how to set up mathematical models, and use analytical and computational tools to relate them to biological data. They experience the modelling process by working on a research project. In this paper, we explain our teaching philosophy, share some unique features of our summer school, and exemplify key course components.

Key words: education, undergraduate, modelling

1.1 Introduction

Mathematical modelling of biological systems requires a wide variety of methods and skills from multiple disciplines. Traditionally, these skills are taught separately in standard courses in mathematics, biology, computer science, etc., but rarely are they integrated into a single undergraduate-level course that focusses on the modelling process.

At the Centre for Mathematical Biology at the University of Alberta, inspired by a similar summer school at the University of Tübingen, we have developed a 10-day summer school for motivated undergraduate students that integrates the methods of applied mathematics in the context of mathematical biology.

The summer school addresses questions of utmost importance to the mathematical biologist, namely what type of model to use, how to develop a mathematical model, how to relate a model to experimental data, and how to validate and/or evaluate the model.

Our primary goal then is to teach the development of models based on biological observations and experimental data, and the interpretation of model results in order to make predictions, suggestions for further experiments, suggestions for control measures or treatments. We include qualitative model analysis, model simulation, and model validation.

In this paper, we describe how we accomplish all of the above in the short time of 10 days. In section 1.2, we provide an overview of the summer school, describing the schedule, typical participants, prerequisites for the course, etc. The bulk of the paper is devoted to describing the course content, in section 1.3. We conclude with our perspectives on the benefits and impact of the summer school in section 1.4, and a brief discussion in section 1.5.

1.2 Overview of the Summer School

We aim to introduce undergraduate students to mathematical modelling and analysis in the context of real biological systems, and provide them with a simulated research experience. To that end, our summer school consists of three integrated components:

1. Through **lectures and exercises**, focussing on four main topics (difference equations, ODEs, PDEs, and stochastic models and parameter estimation), students are introduced to various techniques of mathematical modelling.
2. Through a **self-guided tutorial**, students learn how to use Maple to simulate mathematical models and relate them to biological data.
3. Through **projects**, students experience the modelling process.

The summer school consists of 10 instructional days. The schedule we use is shown in Table 1.1. During the first 5 days, students attend lectures, work through exercises, and complete the self-guided computer tutorial. The last 5 days are devoted to the research projects.

The summer school is attended by 15–25 students each year. In addition to lecturing, our summer school involves a significant amount of one-on-one mentoring and interaction with students. For that reason, we cap enrolment at approximately 25.

Typical students have completed 2–3 years of undergraduate study in mathematics or a similar quantitative science. While the majority of attendees major in mathematics, some already are enrolled in a degree program combining mathematics or computer science and biology. Undergraduates in their third year are especially encouraged to attend. Beginning graduate students in the biological and medical sciences interested in mathematical modelling are welcome as well, and in fact serve as an excellent complement to the group composition.

The prerequisites for attending our summer school are a basic knowledge of calculus, linear algebra, and differential equations. Although not necessary, we have found that some knowledge of computer programming is extremely helpful.

Table 1.1. Schedule for the mathematical biology summer school.

	9:00–10:30	**11:00-12:30**	**1:30–3:00**	**3:30–5:00**	**Evening**
Day 1	Introduction	Discrete I	Maple Lab	Discrete II	Homework
Day 2	Exercises	ODE I	Maple Lab	ODE II	Homework
Day 3	Exercises	PDE I	Maple Lab	PDE II	Homework
Day 4	Exercises	Stochastic I	Maple Lab	Stochastic II	Homework
Day 5	Exercises	Maple Lab	Maple Lab	Maple Lab	Project Intro
Day 6	DAY OFF				
Days 7–10	Projects	Projects	Research Lecture	Projects	
Day 11	Presentations	Presentations	Presentations	Presentations	Graduation

We typically run the summer school with four core instructors, one high-profile guest instructor, secretarial support staff, and approximately 20 volunteers (graduate students and postdoctoral fellows). The core instructors give the lectures and guide the exercise sessions during the first half of the summer school, and serve as primary project consultants for 2-3 student teams during the second half. The guest instructor attends the summer school for the last few days only, delivers a keynote address, and serves as roaming project consultant, interacting with each student team. The graduate students and postdoctoral fellows help with all aspects of the summer school.

1.3 Course Content

In this section, we elaborate on the three course components, namely lectures and homework, the computer tutorial, and research projects.

Part I: Lectures and Homework

We integrate theory and modelling in the lectures and homework component of the summer school. We begin with a brief presentation on the history of mathematical biology. The modelling lectures begin with a review of the importance of distinguishing between dependent and independent variables, and probabilities and rates, as well as an overview of the most common model classes.

We elaborate on four of the model classes (difference equations, ODE's, PDE's, and stochastic models and parameter estimation) in four units of lectures and homework sessions. Each unit consists of 3 hours of lecturing and 1.5 hours of tutorial sessions during which homework problems (assigned during the lectures) are discussed. Students receive extensive course notes that fill in details not covered in the lectures. The course notes have been edited and published [**?**].

The unit on difference equations covers scalar and two-dimensional systems, both linear and nonlinear. Students are introduced to the concept of a fixed point, as well as the notions of stability and instability. For scalar equations, we teach both graphical stability analysis (cobwebbing) and linear stability analysis. The latter is extended to two-dimensional systems. We give a full treatment of the discrete logistic equation, including period-doubling and the Feigenbaum diagram, thereby introducing students to the concept of a bifurcation.

The unit on ordinary differential equations builds on the previous unit, and covers direction fields, nullclines, and phase plane analysis. Applications include the investigation of 2-species interaction models such as predator-prey or competition models, as well as standard SIR epidemiological models.

The unit on partial differential equations covers an age-structured population model and reaction-diffusion equations. In particular, we focus on the critical domain size problem and travelling waves.

The last unit covers stochastic modeling and parameter estimation. The section on stochastic models covers random walk models and birth-death processes. The section on parameter estimation includes the log-likelihood method, the Akaike Information Criterion (AIC), and the likelihood ratio test.

The following exercise, drawn from the lectures on difference equations, illustrates our approach to integrating theory and modelling:

> *Consider the following model for drug prescription:*
>
> $$a_{n+1} = a_n - ka_n + b,$$
>
> *where a_n is the amount (in mg) of a drug in the bloodstream after administration of n dosages at regular hourly intervals.*
>
> *(a) Discuss the meaning of the model parameters k and b. What can you say about their size and sign?*

(b) Perform cobwebbing analysis for this model. What happens to the amount of drug in the bloodstream in the long run? How does the result depend on the model parameters?

(c) Sketch a graph of a_n versus n. How should b be chosen to ensure that the drug is effective, but not toxic?

In the above, students are given the equation of the model, but they are asked to figure out the structure of the model themselves, by determining the meaning of each of the model parameters. In other exercises, students are asked to construct their own model, based on explicitly stated assumptions. The experience gained from working through exercises such as these prepare students for the project work later during the summer school.

Part II: Self-guided Maple Tutorial

The structure of the Maple tutorial follows the lectures. The tutorial supports the material discussed in the lecture, and provides a different perspective on biological problems, but also covers topics not discussed in the lectures. In particular, we use the tutorial to introduce students to data analysis, covering linear regression and dealing with data sets, as well as numerical solutions of differential equations.

The computational software of our choice is Maple. The point is not to learn Maple *per se*, but to extend the range of interesting problems within the grasp of students through computation. Other software, such as Mathematica, or Matlab, or even C++, could be used instead.

Each student has access to a computer, and works through the tutorial document at his/her own pace. We typically schedule about 12 hours of lab time for the Maple tutorial. An instructor and several teaching assistants are present during the Maple lab sessions to provide help when needed.

The tutorial is extensive, and contains many examples and exercises, ranging from trivial to challenging in difficulty. We do not expect every student to complete the entire tutorial, although each year there are a few who do (primarily advanced students with some computer programming background). Occasionally, a student already has some background with Maple. Those students are free to skip sections of the tutorial with material that is familiar to them. We have found that the tutorial is sufficiently wide-ranging and challenging to keep the interest of those students.

Following is a sample exercise from the Maple tutorial illustrating the use of data analysis in the modelling process:

Consider the Ricker model, written in the following form:

$$x_{n+1} = ae^{-bx_n}x_n.$$

(a) Fit the Ricker model to Barlow's data on the number of nests per hectare for a population of the common wasp Vespula vulgaris *[**?**], shown below.*

1988	1989	1990	1991	1992
8.6	31.1	7.0	11.7	10.2

That is, find the values of the parameters a and b that best fit the data.

(b) Check the fit for the data two ways: (1) in a plot of $\ln(x_{n+1}/x_n)$ *versus* x_n*, and (2) in a plot of* x_{n+1} *versus* x_n*.*

(c) What behaviour is predicted for the wasp population, based on the results of your earlier bifurcation analysis of the Ricker model?

The above exercise is one in which we expect students to integrate knowledge obtained in preceding exercises. In particular, students learned linear regression in an earlier set of exercises. Here, students are expected to recognize the linear relationship $\ln(x_{n+1}/x_n) = \ln(a) - bx_n$, and then use linear regression to obtain the best-fit values for $\ln(a)$ and b. In another set of preceding exercises, students were asked to systematically investigate the behaviour of the Ricker model written in the form $y_{n+1} = ry_ne^{-y_n}$ over a range of the parameter value r, and summarize the behaviour in a bifurcation diagram. Here, they are asked indirectly to determine the relationship between the two forms of the Ricker model, and use the results of the bifurcation analysis to predict the behaviour of the model obtained for the wasp population.

Part III: Research Projects

Students choose a modelling problem from a set of approximately 25 project descriptions, loosely grouped in four topic areas (epidemic models, population dynamics, models for spatial spread, and physiology).

Students work in teams of 2–3, under the guidance of one of the instructors. Students are expected to develop a model, analyze and/or simulate their model, and prepare a presentation.

Many of the problems have not been studied previously with a mathematical model, and are open-ended, with no “right solution” *per se*. Because of the open-ended nature of the research projects, instructors are flexible. Very often, students take the project in different, sometimes better, directions than the instructor might have.

In many cases, students will need to simplify their problem, and build a hierarchy of models, each model incorporating additional realism from the original problem at hand. We emphasize to students that success is not measured in terms of the end product, but in terms of the amount of learning that is taking place during the model development and analysis. It is not uncommon that the team apparently making the slowest progress actually is learning the most.

For some of the projects, we provide some supplementary reference materials. For others, students easily can obtain additional information from the internet. We do not require students to study the biological topic at length (we believe that initial efforts in mathematical modelling require only the identification of basic mechanisms). The point of the project work is not to produce publishable results, but rather for students to experience the modelling process.

Sample project topics include the spread of HIV in Cuba, cholera in South Africa, the extinction of a wolf population in Sweden, the pupil control system, and radiation treatment of cancer. The complete set of project descriptions can be found in [**?**].

Following is a sample project description, dealing with the outbreak of Yellow Fever in Senegal:

> *Yellow Fever (YF) is a viral disease transmitted to primates (including humans) by infected mosquitoes. The disease is endemic in populations of monkeys living in the jungle. The disease is spread into the human population in three stages:*
>
> 1. *Sylvatic transmission occurs when mosquitoes which have fed on infected monkeys next bite a human working in the jungle.*
> 2. *Intermediate transmission occurs when mosquitoes pass the virus among humans living in rural areas.*
> 3. *Urban transmission occurs when mosquitoes pass the virus among humans living in urban areas.*
>
> *Below is a data set of YF cases reported during an outbreak in the city of Touba in Senegal in 2002 [**?**]. As soon as the virus was identified (October 11), a vaccination program was started. YF vaccine is safe and effective, and provides immunity within one week in 95% of those vaccinated.*

Date	Jan 18	Oct 4	Oct 11	Oct 17	Oct 24	Oct 31	Nov 20	Nov 28
Cases (total)	18	12	15	18	41	45	57	60
Deaths (total)	0	0	2	2	4	4	10	11

> *Develop a model for the three stages of YF as outlined above, including vaccination in urban areas, and fit your model to the data. Would you expect that the disease dies out or that it becomes persistent? What would have happened without vaccination?*

The student team that tackled this problem in one of our recent summer schools decided to simplify the problem significantly, and focussed on the outbreak of YF within urban areas only. They constructed a model describing the transmission of the virus within the population of mosquitoes and from mosquitoes to humans. The mosquito population was divided into two classes (susceptible and infective), and the human population was divided into five classes (susceptible, exposed, infective, recovered, and vaccinated). Parameterizing a model of this size is a daunting task, even for experienced researchers. The students were able to obtain estimates for a number of parameters (such as the biting rate of mosquitoes) from the internet and journal articles, and adjusted the value of remaining parameters with a “fit-by-eye” procedure. They obtained a reasonable fit of the model to the data supplied, and then investigated the predictive power of the model in terms of suggesting strategies for controlling the outbreak of the disease.

Projects such as the above allow students to experience the entire modelling process, from model development to simulation/analysis to the interpretion of results.

1.4 Benefits and Impact of the Summer School

Students are very enthusiastic about the summer school. They are grateful for the challenging extracurricular experience, and appreciative of the opportunity to interact with other bright and motivated students from varied educational backgrounds. Following are examples of typical comments received from students upon the completion of the summer school:

> *I think the biggest thing I got out of the workshop was an appreciation of the wide variety of modeling applications (especially through the projects) – and also the immense power of a relatively limited set of techniques.*

> *This workshop not only helped me to gain experience in Mathematical Biology, but also to decide my direction in my academic career.*

> *It gave me a good overall look at math modeling . . . I now know what a "mathematical model" is. It's a phrase I hear a lot, but wasn't exactly sure what that meant. I also now have a clearer vision of a direction that I'd like to take in graduate studies.*

The comments quoted above illustrate that the summer school is highly useful to these students in guiding them towards future studies or careers in mathematical biology. Indeed, a significant number of past participants have continued their graduate studies in mathematical biology. To our knowledge, at least 12 out of 64 participants have found a place at different Canadian or US institutions.

The immeditate benefits of the summer school for both the core and guest instructors are infectious enthusiasm and exposure to potential graduate students committed to the field of mathematical biology. But there are additional benefits. In particular, the instructional materials developed for the summer school have grown over the years, and now are published as an undergraduate textbook [**?**]. Also, we are starting to use modules developed for the summer school in our regular classes, and have plans to offer an undergraduate course on mathematical biology in the near future using the same philosophy that we use in the summer school.

Last but not least, the summer school contributes significantly to the career development of the many graduate students and postdoctoral fellows who help run the summer school. Both graduate students and postdoctoral fellows participate as teaching assistants for the exercise sessions and Maple labs during the first half of the summer school. In addition, the postdoctoral fellows serve as project consultants during the second half.

1.5 Discussion

Our summer school is an academically stimulating program that teaches the following applied math skills in the context of mathematical biology: theory, modelling, analysis, computation, data fitting, and making predictions. Through lectures and exercises, the Maple tutorial, and project work, participants not only gain a wide knowledge of mathematical biology, but also are introduced to research.

Although our formal instructional time with the students is very limited, we cover enough aspects of modelling that students can find and learn required advanced techniques while working on their projects. In fact, students relish the challenge provided in the projects, and it is during this part of the summer school that we observe a phenomenal increase in knowledge and skill development. Participants are highly motivated and often achieve much more than expected.

For more information about the summer school, and to view student work, we invite readers to visit our website [**?**].

Acknowledgements

We thank our colleagues Mark Lewis (U Alberta), Frithjof Lutscher (U Ottawa), Johannes Müller (TU Munich), and Birgitt Schönfisch (U Tübingen), who all have contributed significantly to the development of the summer school. We are grateful to the guest instructors: Pauline van den Driessche (U Victoria), Leah Edelstein-Keshet (UBC) Michael Mackey (McGill), Rebecca Tyson (UBC Okanagan), Gary de Young (Kings U College), Michael Li (U Alberta), and Alex Potapov (U Alberta). We also thank our graduate students, postdoctoral fellows, and secretarial support staff for their helpful suggestions and feedback. Most of all we thank the enthusiastic participants of the summer school who have made it a highlight of our academic year. Last but not least, we are grateful to the institutions and organizations who have provided funding for the workshop, namely the Pacific Institute for the Mathematical Sciences (PIMS), Alberta Innovation and Science (ASRA), the Mathematics for Information Technology and Complex Systems (MITACS) National Centre for Excellence, and the Centre for Mathematical Biology and the Department of Mathematical and Statistical Sciences at the University of Alberta.