To appear in: MAA Notes Volume "Undergraduate Mathematics for the Life Sciences: Processes, Models, Assessment, and Directions" edited by Jenna Carpenter, Glenn Ledder and Tim Comar. 2009

Some lessons from fifteen years of educational initiatives at the interface between mathematics and biology: the entry-level course

Louis J. Gross
Departments of Ecology and Evolutionary Biology and Mathematics
University of Tennessee
Knoxville, TN 37999-1610

Abstract: In 1992 I led a workshop devoted to the issues of undergraduate education in quantitative approaches for life science students. One outcome of that workshop was a suggested set of topics to include in an entry-level math course for life science undergraduates. With NSF support, I proceeded to develop and pilot a two-semester course that followed this recommended topic coverage, and represented a considerable departure from the emphasis on calculus alone that was prevalent (and still is to a large degree) for undergraduates in the sciences. This course sequence consisted of two semesters, involved an hypothesis testing and data-rich view of mathematics, included a variety of computer-based projects, and has been taught by a diverse set of instructors, with widely varying backgrounds and levels of biological expertise, over the time since the pilot. It has become the major math sequence for biological science undergraduates at UTK, with approximately 350 students enrolled each semester, and was included as one of the main case studies in the National Research Council's Bio2010 report. I provide details on the components of the course, both with regard to topic coverage and student assessment methods, discuss the goals of the course (a focus on conceptual development rather than strictly on skills), and note how the course sequence has changed since it was initiated. I also discuss the challenges we have faced with regard to diversity of backgrounds of instructors.

I. Background

The history of courses in mathematics in the U.S. designed specifically for undergraduate students with interests in the biological sciences dates back at least to the early 1970's with the text by Batschelet (1971). In this period a number of texts and courses appeared, with the majority emphasizing elementary calculus and its prerequisites (Arya and Lardner, 1979; Cullen, 1983; DeSapio, 1977; Gentry 1978; Levin, 1975). By the early 1990's many of these courses had disappeared, with the remaining ones mainly ongoing at institutions with research emphases in mathematical biology. In 1991, I began a project, funded by the National Science Foundation (USE-9150354) with a goal of developing a quantitative curriculum for life science students. The initial focus was on entry-level math courses and started with a survey of quantitative course requirements for life science students at U.S. universities. The results indicated that about 80% of institutions had some calculus requirements while the remaining required only pre-calculus or had no explicit requirement.

In 1992 I convened a Workshop attended by many leading mathematical and quantitative biologists with interests in undergraduate education to discuss all aspects of undergraduate mathematics preparation for life science students. Among the recommendations from that workshop was a strong suggestion that students are not well-served by being exposed only to a calculus course, noting that experiences with a much broader array of quantitative concepts were necessary for students to be successful in careers in modern biology. We had been teaching at UTK an elementary calculus for life sciences course since the late 1970's, and as an outcome of this workshop, I revised the course to incorporate a much more diverse set of topics, and move the focus towards conceptual understanding rather than hand calculations of standard calculus rules.

II. Structure of the Entry-Level Course

The 1992 Workshop provided specific suggestions regarding topic coverage for an entry-level course, including the addition of descriptive statistics, matrix algebra, discrete probability and difference equations. Additional suggestions included ensuring that the calculus portion included an introduction to differential equations and incorporated the notions of equilibria and stability. Many additional suggestions were provided to incorporate specific biological concepts as part of this course, with the objective being to encourage problem-based learning rather than regurgitation. The full workshop report (Gross, 1992) also provided suggestions for upper division courses and methods to link the mathematics and biological portions of the curriculum.

Based upon the recommendations of the Workshop, I proceeded to implement a pilot version of the course in place of the two-semester, 3-credit hour course we had been teaching at UTK. There had been concern expressed at the Workshop that it would be extremely difficult to squeeze in all the suggested material into two 4-credit courses, let alone 3-credit ones. Based on our experiences over the 15 years that this course sequence has been offered however, it is most definitely feasible. The initial pilot course followed very closely the suggestions of the Workshop, which had focused mostly on particular mathematical topics. However over the course of the next several years, I developed a set of concepts rather than explicit mathematical methods, which served as a focus for much of the material included in this course sequence. These have been expanded upon elsewhere (NRC, 2003), so I will simply list the concepts here: Rate of Change, Scale, Equilibria, Stability, Structure, Interactions, Stochasticity, and Visualizing. The math that is needed to formalize these concepts provided the motivation for the revisions to the entry-level course that have occurred since the pilot version was offered. The first semester of this course covers the discrete mathematics topics through the concept of limits of sequences, and the second semester covers the calculus up to first order linear and simple non-linear differential equations.

The general goals of the current course are:

1. Develop the ability to quantitatively analyze problems arising in the biological areas of interest to the variety of students attending the course.
2. Illustrate the great utility of mathematical models to provide answers to key biological problems.
3. Develop appreciation of the diversity of mathematical approaches potentially useful in the life sciences.
4. Provide experience using computer software to analyze data, investigate mathematical models and provide some exposure to programming.

The course methods to meet these goals include:

1. Encouraging hypothesis formulation and testing for both the biological and mathematical topics covered.
2. Encouraging investigation of real-world biological problems through the use of data in class, for homework, and in examinations.
3. Reduceing rote memorization of mathematical formulae and rules through the use of software such as MATLAB and Maple.
4. Providing biological motivation for each main mathematical component of the course through a central example that is returned to regularly.

I will here give a brief example of each of these methods. Regarding hypothesis formulation, the first course begins with each class section going to the "field" to measure leaf lengths and widths in small groups of 3 students (we do not provide details as to how to do this, but leave the explicit design up to each group) and encourage each group to make an hypothesis concerning the relationships between the variables they are measuring. The class then immediately goes to a computer lab where we provide an introduction to basic Matlab commands, have each group enter their data, and produce histograms and scatter plots to evaluate their hypotheses. This requires a complete lab session (1 1/4 hours to complete) but serves an essential purpose to point out that mathematics connects to data, that hypothesis evaluation is part of science and that mathematics can assist in this scientific process.

Regarding real-world biology, the class is presented at the start of many sessions with a recently published paper (typically from Science, Nature or PNAS) that includes mathematics related to the topic of the day. Although few of the students are capable of reading these papers in detail, they point out the current importance of mathematical ideas in biological research, and some of the data from these papers are used directly to motivate the mathematics as well as appearing on tests and quizzes in the course. The course web page includes links to the papers as well as a brief summary I compose that discusses the results and how they relate to the course topics.

On the use of software, there are several objectives beyond simply reducing the need for hand calculation. One of these is the ability to include much larger data sets or deal with much more complex (and realistic) projects than could be included otherwise. For example, one project asks students to make an hypothesis regarding whether their height changes overnight, which the class proceeds to evaluate by collecting data on themselves for several nights. These data are then combined and analyzed using MATLAB to illustrate histograms and regressions, point out the importance of looking carefully at data for "bad" values (e.g. data which imply a student gained 10 cm in height overnight!), and to discuss the obligation to deal carefully with human data (through discussion of the Institutional Review Boards that oversee human data collection at every institution). Finally, we have chosen to use MATLAB as the software for the initial portion of the course (we move to Maple for much of the calculus portion) as it introduces students to basic coding and logic, as well as the idea of an algorithm.

Regarding a motivating central example, for the matrix algebra section, students are shown images of landscapes, asked how to characterize these and requested to develop hypotheses about how these characterizations might change as these images were retaken over several decades. This leads naturally to the use of vectors to characterize the distribution of states across the landscape and provides an introduction to the concept of ecological succession, which is typically familiar to students. It serves as a way to motivate students to develop the notion of a Markov chain, which they can readily do themselves, though not with any detail on how the mathematics might work. The concept of an eigenvector as the long-term fraction of the landscape in each "state" such as urban, forest, or agricultural then arises naturally and the notion of stability of such an equilibrium is readily motivated through consideration of different initial landscape state distributions.

III. Student Assessment

Over time we have developed three major and one minor assessment method to evaluate student performance in this course sequence. This includes are a formal set of written exams covering the basic concepts, as well as a comprehensive final exam. Together these account for 60% of the course grade, and the exams often include some examples taken from data or research papers discussed in class. Students are encouraged to utilize standard calculators as they wish in these exams, though we do not emphasize the use of such calculators in most aspects of the course assignments. The exams are structured and evaluated however in such a manner that the use of a calculator is not necessary. Students find it a helpful crutch however if their arithmetic skills are weak (and our focus in this sequence is not at all on arithmetic skills).

The course includes a set of weekly quizzes that cover the material assigned for homework (which we do not collect or grade). These quizzes are supplemented currently by a required "at the board" solution of a homework problem that counts as an additional quiz for each student. Together these count 20% of the course grade. Finally, we have a set of computer-based projects, some of which require students to collect their own data, but most of which involve writing a formal laboratory report on a set of "simulation experiments" carried out using MATLAB or Maple. In some cases we supply the code for this and the students have to modify it in a few locations, while in other cases the students must develop their own code. The more complex of these projects involve essentially research-level projects (such as the impact of random environmental variation on the dynamics and long-term structure of matrix population models) the motivation for which is readily understood by the students, although carrying out a full analysis would require extensive graduate-level math.

IV. Challenges

Since its development in 1993, this sequence has been taught by approximately 80 different instructors, ranging from regular mathematics faculty, to full and part-time instructors and graduate students. While the majority of these have been individuals with some exposure to mathematical biology (through our long-standing graduate education programs in this field) quite a few have not. This has created challenges and typically means that there is less emphasis on the biological aspects of the course when taught by an individual with little biological background. We have dealt with this in a couple of ways, one of which was to develop a course supplement and instructors guide which assists in providing motivating examples for the course. Additional lecture notes which develop the biological connections for the course, as well as providing the various links and explanations of research papers used to motivate sections of the course, are provided on the course web site. A set of modules are available to assist biologically naïve instructors, having been developed to motivate many of the concepts in a general biology course through a connection with mathematics (Harrell, Beals and Gross, 2002).

It is the first course in this sequence that is the most difficult for those with little biological background to teach, as the second course focus is on calculus. The first course has been taught recently in large-lecture format (about 200 students with smaller sections meeting for a once-a-week lab) by myself or another instructor with an extensive math biology background. This large-lecture, small-lab section format has allowed for only limited small-group discussion except in the weekly lab sessions, and with only two meetings a week in lecture leads to a much more rapid pace than might be preferable.

An additional challenge has been the lack an appropriate text for this sequence. We have relied upon one of the older texts (Cullen, 1983) supplemented by a variety of notes and on-line materials. Although several new texts have appeared that superficially may seem to be appropriate, these have a different topic coverage (often much more extensive than is feasible for this course) and have very little emphasis on data and its relation to the course topics. Although a text for this course has been planned, it is several years from fruition.

V. Lessons and Suggestions

First and foremost, in many ways this course is substantially more challenging, both for the students and the instructors, than either the standard science and engineering calculus sequence or the standard two-semester calculus for business and social science students. It covers a vast array of topics very quickly, and students who are not well-prepared find it quite difficult. The clientele includes a majority of students in pre-health sciences however, who are extremely motivated, realize the importance of working hard in order to be accepted by various professional schools, and often appreciate the connection between this course and their other biology courses. The class typically includes an array of students with varying levels of biological experience (as some of them delay taking this course until much later in the undergraduate program than we prefer). I suggest viewing this as beneficial since these students can then point out to their peers how the topics relate to those covered in more advanced biology courses. An example is the Hardy-Weinberg equilibrium which students see in their general genetics course, but which we derive mathematically.

Second, despite the fact that all students taking this course have had a pre-calculus background (mostly in high school) and have taken a math skills placement exam to be recommended for this course, their comprehension of certain pre-calculus concepts is weak. For this reason, we have explicitly focused the initial section of the course on descriptive statistics, including semi-log and log-log graphs, with application to biological growth and decay processes and allometry. This serves as a means to refresh the student's understanding of exponentials and logarithms, motivate this through biological observations and introduce new scaling concepts. Providing a new perspective, based upon observations, of a topic that never really “clicked” for many students, fosters the possibility that students will develop a conceptual understanding of non-linear scaling and why logarithms appear in so many contexts in biology.

Third, despite much experience with word-processing software and computer games, and some exposure to spreadsheets, few students enter this course with any understanding of programming or the basic logic that underlies all of computer coding. It requires considerable effort on their part to follow basic MATLAB codes provided to them, and they have great difficulty developing their own programs. Initially, we alleviated this by providing much of the MATLAB codes required for the course to the students, but we have since re-assessed the situation and require the students to develop their own codes for at least the initial, simpler computer projects. Otherwise we found that students had little understanding of the steps in the code. We assist in this process by including a variety of basic guides to MATLAB commands associated with each project.

VI. Conclusion

A major caveat concerning this course sequence is the lack of a planned protocol to assess its impact. Well over two thousand students have taken these courses over the past fifteen years, but we have had essentially no evaluative procedures developed to track students who take this sequence as compared with those who take the standard science and engineering calculus sequence. There are a range of evaluations by students completed each year, often using different assessment forms, and none of these have been compiled in a manner that would allow for comparisons either across time or between students in different courses. If there were a single suggestion I would have for those beginning to teach a course such as this, it would be to design a course assessment methodology before implementing the course, find funding to implement it over a number of years, and track student success and behavior as part of this. At this point, all I can point to is anecdotal information, including one student who told me that he was instructed by his MD to take the course as it had been beneficial in their medical career. As gratifying as this may be, we need to do better in assessing the impact of courses that fall at the interface between mathematics and biology.

VII. References:

Arya, J. C. and R. W. Lardner. 1979. **Mathematics for the Biological Sciences.** Prentice-Hall, Englewood Cliffs, NJ

Batschelet, E. 1971. **Introduction to Mathematics for Life Sciences.** Springer-Verlag, NY.

Cullen, M. R. 1983. **Mathematics for the Biosciences.** PWS, Boston.

DeSapio, R. 1977. **Calculus for the Life Sciences.** Freeman, San Francisco.

Gentry, R. D. 1978. **Intrduction to Calculus for the Biological and Health Sciences.** Addison-Wesley, Reading, Massachusetts.

Gross, L. J. 1992. Summary Report: Workshop on a Quantitative Sciences Curriculum for Life Science Students.
http://www.tiem.utk.edu/~gross/WorkshopReport92.txt

Harrell, S., M. Beals and L. J. Gross. 1999. General Biology Modules.
http://www.tiem.utk.edu/~gross/bioed/modulelist.html

Levin, S. A. 1975. **Mathematics for the Biological Sciences.** Cornell University.

National Research Council. 2003. **BIO2010: Transforming Undergraduate Education for Future Research Biologists.** National Academies Press, Washington.