Mathematical Biology

A proposed update for the 2015 CUPM Guide to Undergraduate Programs in Mathematics Sciences

March 2022

MAA
MATHEMATICAL ASSOCIATION OF AMERICA

Mathematical Biology

Author Information

Lisette DePillis, Harvey Mudd College
Jennifer Galovich, College of St. Benedict/St. John's University
Louis Gross, University of Tennessee, Knoxville
Raina Robeva, Randolph-Macon College
Jonathan Rubin, University of Pittsburgh

Overview

The confluence of mathematics and biology is central to scientific advancement in the 21st century. Challenges as diverse as global climate change, pharmaceutical design, emergent diseases, and genomics-age medicine all require scientists and mathematicians with expertise in both fields.

An integrated view of life science education was heralded in the NRC Bio2010 report [BIO2010: Transforming Undergraduate Education for Future Research Biologists (2003). The National Academies Press, Washington, DC] and explicitly argued for incorporation of a diversity of mathematical topics throughout biology courses and not simply isolated in the mathematics and statistics courses undergraduates in the life sciences were being required to take. Thus, the state of biology education with respect to mathematics is largely settled. The state of mathematics education with respect to biology however, is harder to pin down.

This committee is especially cognizant of the need to present recommendations that respect departmental autonomy, so as to accommodate a variety of institutional attitudes toward interdisciplinary curricula as well as departmental (and interdepartmental) staffing challenges. To that end, we outline several foundational courses, indicate some directions for more advanced undergraduate study, and present a list of fundamental mathematical competencies. We conclude with some recommendations regarding biological competencies.

Throughout, we have in mind as our audience departments of mathematics at liberal arts colleges, research and comprehensive universities, and community colleges. Such departments may wish to construct a new major or minor, a concentration within an existing major or an interdisciplinary concentration attached to an existing major. We encourage collaboration and cooperation with other departments that share these interests. Our hope is that by presenting competencies rather than prescribed mathematical content, we will provide the flexibility needed for multiple routes to success, based on local capabilities. For the same reason, we have steered away from offering specific examples of successful programs. This is especially important given the enormous breadth of biology as a discipline and the myriad of possibilities for institutions to design programs that would address their needs and objectives effectively.

Student Audience

Mathematical biology programs will benefit students with a range of career goals, educational interests, and mathematical backgrounds. The motivation for some students to pursue mathematical biology education will relate to an interest in medical and health-related directions. The 2009 report from the American Association of Medical Colleges and Howard Hughes Institute, ``Scientific Foundations for Future Physicians," (American Association of Medical

Colleges, Washington, DC; https://www.aamc.org/download/271072/data/scientificfoundationsforfuturephysicians.pdf) includes the recommendation that physicians be able to "Apply quantitative knowledge and reasoning - including integration of data, modeling, computation, and analysis - and informatics tools to diagnostic and therapeutic clinical decision making." An increased emphasis on quantitative competencies in medical education and admissions leads some students toward mathematical biology, with a longer-term goal of a medical career. A related ongoing trend that yields a similar effect is the major role of quantitative techniques in developing technological and medical advances, and mathematical biology students with such interests may aspire to biotech careers. Some of these advances involve the use of large data sets; for example, databases are used to predict regional risks of opioid overdoses and disease outbreaks. Awareness of these efforts as well as other interests in health-related biostatistics careers lead some students toward mathematical biology training. A compendium of some programs and considerations of their impacts is Ledder et al. (2013), which presents case studies from over twenty institutions for several types of undergraduate programs, including formal courses, research experiences, and modules.

Other students are drawn to mathematical biology through basic science interests. Some of these students are motivated to study specific biological fields that are either historically or increasingly quantitative, such as neuroscience, genetics, evolution, ecology, and molecular biology. Still others have more general scientific interests. Some of these students may aim for graduate study and research and/or teaching careers in biology or in an explicitly interdisciplinary area such as computational biology. Others may foresee employment in consulting, governmental, or other regulatory capacities.

Some students with a primary interest in mathematics also end up interested in mathematical biology. Compelling examples of the use of mathematics to study biological systems can arise to demonstrate topics throughout many areas of the mathematics curriculum, including calculus, linear algebra, and differential equations. Some students encountering such examples, perhaps in class or in the context of class projects, become hooked and develop a strong interest in mathematical biology. The career aspirations of many such students are similar to those of the students drawn to math biology for other reasons, while others may aim to go on to research and/or teaching careers in mathematics.

One of the challenges of designing a mathematical biology program is that the range of motivations and backgrounds of the students who study the subject is quite diverse, much broader than those found across a typical distribution of mathematics majors. Some students who have struggled with mathematics in general may nonetheless become strongly motivated to learn more mathematics when they come to understand its relevance to an application area of interest to them. Many of the students in mathematical biology have relatively broad interests, which means that they aim to juggle courses from many different areas, and as a result some have less mathematical breadth than typical mathematics majors. One of the benefits of students' inter-departmental experiences is that some will already have a familiarity with topics that are beneficial in mathematical biology as well, such as the use of certain software packages, aspects of the handling of data, and how to effectively communicate technical ideas, a skill that may be emphasized more in some areas of scientific training than in many mathematics programs.

Students who have completed a program that includes the foundational course material, and who possess the fundamental competencies should be well poised to embark on graduate work in this field. The names and types of programs vary enormously. Here are some examples:

Bioinformatics; Bioengineering; Mathematical Biology; Biostatistics; Computational Biology; Computational Neuroscience; GBCB (Genetics, Bioinformatics, and Computational Biology). Some programs are interdisciplinary in the sense that the students are officially in a mathematics graduate program, but their research is located in another department. Moreover, graduate programs in many biology subdisciplines may welcome students with the kind of preparation we have described.

Prerequisite Skills and Knowledge

A central question for any interdisciplinary program is how much foundational knowledge is required from each discipline. For a mathematical biology program, we assume that during the course of study students will complete some standard courses from the mathematics and biology majors. On the mathematics side, we assume that students will gain competence in logical reasoning, basic methods of proof, and the standard topics associated with the typical mathematics major: Calculus, Linear Algebra, and ODEs. The choice of courses on the biology side depends on the biology subdiscipline of interest. One or two introductory courses required for the biology major will be an appropriate start, followed by at least one upper division course in a biology subdiscipline. More details can be found under *Biology Training for Mathematical Biology Majors* below.

We also note and endorse the additional and complementary competencies recommended by *Scientific Foundations for Future Physicians,* a report prepared jointly by AAMC and HHMI: https://www.hhmi.org › default › files › Programs › aamc-hhmi-2009-report (see pp. 19 – 24) and the suggestions in the AAAS Vision and Change reports: https://visionandchange.org

This report focuses on the quantitative components of an undergraduate program, but a well-rounded undergraduate program that intends to prepare students for either life science related careers upon completion or graduate/professional education would best incorporate a breadth of exposure to biology. At the least, this should include some of the introductory biology courses (typically for first and second year students) that cover the basic concepts of molecular and cell biology, organismal biology, and ecology and evolution. Students could be encouraged, depending upon their desired career path, either through formal courses or through laboratory and/or field experience to enhance their background in some biological sub-discipline area. Such further experience in one biological area will enhance the capability for the student to move into an appropriate graduate program.

Specific Student Learning Outcomes

In an interdisciplinary endeavor it is not natural to break down activities into disjoint categories. There will necessarily be overlap among categories. In the interest of communicating our recommendations, however, we have attempted to group the 23 numbered learning outcomes below into four more general categories of SLOs labeled by the letters M, N, D and A. We envision that based on their programmatic goals and objectives departments may choose to use only a subset of these recommended competencies for assessment. Below, we give the general description of each SLO category and then list the SLOs that belong to that category.

(SLO M) Modeling: Students should be able to create, describe/explain, utilize and critique mathematical models of biological phenomena.

- (M1) Given a biological question, be able to create a mathematical model appropriate to address the question.
- (M2) Given a model and information about what its variables and parameters denote, be able to explain what the terms represent and what assumptions are encoded in the model.
- (M3) Be able to develop and analyze a set of models that address different levels of aggregation and different scales for a biological system.
- (M4) Given data or other information about a biological system, be able to infer the structural relationships present and represent this information in a mathematical or statistical model.
- (M5) Recognize, describe, and give examples of the distinction between discrete and continuous processes and mathematical formulations, and be able to justify the use of a continuous-space dynamical system versus a finite-state dynamical system for modeling.
- (M6) Recognize and identify potential sources of stochasticity in systems and gain familiarity with ways to model and simulate stochastic effects.
- (M7) Convey mathematical ideas (e.g., the results of mathematical analysis of a model system) clearly in writing as appropriate for different, possibly interdisciplinary audiences.
- (M8) Be able to apply mathematical solution techniques in the context of analyzing a mathematical model.
- (M9) Be able to evaluate how effectively a model meets the goals for which it was constructed.

(SLO N) Numerics/Coding: Students should be able to select and apply numerical methods for approximation and model analysis, as well as to write and debug computer programs so as to effectively implement mathematical algorithms and minimize numerical error.

- (N1) Recognize, describe, and give examples of the powerful synergies between mathematical/statistical models and simulation approaches.
- (N2) Be able to apply methods to efficiently and accurately compute approximations to functions and approximate solutions to classes of mathematical equations.
- (N3) Be able to apply methods to visualize function behaviors, data, information obtained from models, results of numerical computations, and results of statistical analysis.
- (N4) Be able to use numerical methods to analyze effects of parameter values on system behaviors.
- (N5) Recognize, describe, and give examples of sources of error in numerical computations.
- (N6) Develop skills in algorithmic thinking.
- (N7) Be able to design, code, debug, and document programs using techniques of good programming style and structure.

(SLO D) Data: Students should be able organize data in suitable formats, analyze biological data, and fit mathematical functions to data.

- (D1) Recognize, describe, and give examples of issues associated with biological data arising due to variability, heterogeneity and errors introduced by interactions and other processes.
- (D2) Develop skills in fitting mathematical functions to given data.
- (D3) Know how to organize data, as well as search, sort, and mine existing databases.

(SLO A) Analysis: Students should be able to work with linear and nonlinear systems, optimization techniques~~,~~ and spatiotemporal aspects of biological and physical systems.

- (A1) Recognize, describe, and give examples of how nonlinear systems differ from linear ones and the process of local linearization for nonlinear systems, and be able to justify the use of a nonlinear system versus a linear system for modeling.
- (A2) Recognize, describe, and give examples of the concepts of an optimum, constrained vs. unconstrained optimization, and the choice of an objective function for optimization purposes.
- (A3) Be able to find and analyze stability of equilibria of dynamical systems.
- (A4) Be able to analyze spatial and spatiotemporal aspects of biological and physical systems.

Foundational Coursework

The following courses are essential for training in mathematical biology. We recognize that different institutions may structure courses somewhat differently; in fact, the essential element is not that these specific courses should be given, but rather that the targeted SLOs should be achieved via entry level course work. In terms of pedagogy, we feel it is important that students be presented with ample opportunity to participate in collaborative work -- as part of a team to analyze problems arising from formal course material, a research experience program, and/or lab/field work. Ideally, such work will also be interdisciplinary, with faculty and students from mathematics, biology, statistics, and computer science employing their skills together. These courses could also be used as an appropriate venue for students to develop communication and presentation skills. Such approaches would be appropriate for many of the courses below, and there are many ways in which instructors could choose to use them. We expect that the Learning Outcomes accompanying each course would be mastered at a level appropriate for the level of courses and the students at the institution where they are being offered.

Regarding appropriate tools, we believe that students in these programs should gain experience with at least one computational tool and learn to use it to at least a level of comfort that would allow the student to use it to start a project, even if not they do know completely how to proceed. There are many generally applicable tools, including R, RStudio, Tidyverse, Matlab, Octave, Mathematica, Maple, C, C++, Python, Java, Excel, Wolfram Alpha. Students should be encouraged to realize that no one tool does everything, and there are benefits of at least knowing the limitations of each tool and what might be better for some tasks. For any particular tool it is important to be aware of the multiple ways that related information and help can be found to solve new problems. Students might also be prepared for frustration - becoming proficient in any of these tools is not easy and though some basics transfer from one tool or system to another, it may not be trivial to figure out the nomenclature/method to do so.

Introduction to Modeling - A course focused on (1) the process by which real systems are abstracted into mathematical formulations, and (2) methods to draw, communicate, and defend reasoned conclusions about the underlying systems from these representations. **Cognitive Recommendations Addressed:** *CR1, CR2, CR3, CR4.* **Learning Outcomes Addressed:** *SLOM1, SLOM2, SLOM3, SLOM5, SLOM6, SLOM7, SLOM8, SLOM9, SLON1, SLON3, SLON4, SLON5, SLOD2, SLOA1, SLOA3.*

Recommended topics:

- (SLOM2) interpretation of terms, including classification into inputs and outputs, in an existent mathematical model

- (SLOM1, SLOM2, SLOM3, SLOM5, SLOM6) modeling terminology, such as discrete and continuous, deterministic and stochastic, and various types of mathematical expressions (e.g., mass-action, exponential growth and decay, logistic growth, and Michaelis-Menten/Hill function kinetics) that often appear in models
- (SLOM2, SLOM9) how to identify and articulate the assumptions, and the limitations, inherent in a mathematical model
- (SLOM1, SLOM2, SLOM3, SLOM9) how to evaluate the suitability of a model for addressing a question being asked
- (SLOM2, SLON4) differences between, and roles of, variables and parameters in mathematical models
- (SLOM1, SLOM3, SLOM5, SLOM6, SLOM9, SLON4, SLOA1) the process of model creation including how to choose variables and parameters to describe a system
- (SLOM9, SLON5, SLOD2) how to fit a mathematical function to data and to evaluate the error in such a fit
- (SLOM8, SLON1) standard solution and/or simulation approaches associated with the classes of mathematical models being discussed
- (SLOM7, SLOM9, SLON1, SLON3) how to display and communicate results of model analysis/simulation in a way that is useful for a given modeling goal or question
- (SLOA1, SLOA3) calculation of equilibrium states of simple discrete and continuous dynamic models and evaluation of their stability
- (SLON4) analysis of sensitivity of model behavior, including model equilibria, to parameter values

Numerical Analysis - A course devoted to methods for generating accurate approximate solutions to mathematical problems are presented, analyzed, and implemented on a computer. **Cognitive Recommendations Addressed:** *CR1, CR2, and CR3.* **Learning Outcomes Addressed:** *SLON2, SLON3, SLON5, SLON6, SLON7, SLOD2.*

Recommended topics:

- (SLON6) algorithms, convergence
- (SLON5) sources of error in numerical computations and methods for describing bounds on error
- (SLON3, SLON5, SLON7, SLOD2) methods of data interpolation and data fitting
- (SLON2, SLON5, SLON6, SLON7) methods to estimate roots of nonlinear functions
- (SLON2, SLON5, SLON6, SLON7) numerical estimation of frequently computed quantities including derivatives and integrals
- (SLON2, SLON5, SLON6, SLON7) numerical methods (direct and iterative) to solve linear systems of equations
- (SLON2, SLON5, SLON6, SLON7) numerical methods to estimate eigenvalues and eigenvectors
- (SLON2, SLON5, SLON6, SLON7) methods to solve initial value problems for ordinary differential equations
- (SLON2, SLON5, SLON6, SLON7) fast Fourier transforms

Introduction to statistics A course covering the fundamentals of statistical investigation, with necessary concepts from probability. The frequentist perspective that is traditional for introductory courses would ideally be complemented with an introduction to Bayesian approaches. **Cognitive Recommendations Addressed:** *CR1, CR2, CR3, CR4.* **Learning Outcomes Addressed:** *SLOM6, SLOM7, SLOM9, SLON1, SLON3, SLON4, SLOD1, SLOD2.*

Recommended topics:

- (SLOM6, SLOD1): Data collection -- sampling, randomization, observational studies, experimental design, observational units, variables, variable types, paired data.
- (SLOM7, SLON3): Elementary data analysis -- descriptive statistics (mean, median, standard deviation, quartiles, dotplots, histograms, boxplots)
- (SLOM6, SLON1, SLON4, SLOD1, SLOD2): Statistical inference -- simulations, sampling distribution, parameter estimation, confidence intervals, significance level, null and alternative hypotheses, types of errors, power of tests, p-value, margin of error, comparing means and proportions, correlation and regression
- (SLOM2, SLOM9): Model utility analysis -- how to identify and articulate the assumptions and limitations inherent in a statistical model
- (SLOM9): Alternative perspectives: frequentist and Bayesian approaches -- similarities and differences, resampling methods

Introduction to Programming and Computer Science A first course in programming and the fundamentals of algorithmic thinking. **Cognitive Recommendations Addressed:** *CR2, CR3.* **Learning Outcomes Addressed:** *SLON6, SLON7.*

Recommended topics:

- (SLON7) how to break a problem into smaller parts and develop unambiguous sequences of steps that are guaranteed to produce the desired outcome
- (SLON6) the formulation of algorithms in a format appropriate for computer implementation
- (SLON7) program design and documentation following established software engineering principles.
- (SLON7) coding and debugging using techniques of good programming style and structure
- (SLON7) classes of variables and objects
- (SLON6, SLON7) standard coding constructs such as array descriptors, loops and conditionals, and how these are described in pseudo-code and in particular programming languages
- (SLON6, SLON7) use of version control methods

Advanced Coursework Options

In this section we present a list of more-advanced course options that may be beneficial to a subset of students pursuing a degree in mathematical biology. Not all institutions will have those courses as part of their academic offerings, and in many cases the courses may be cross-listed as upper-level undergraduate/low-level graduate courses. For each of the courses listed, the topics we include are those that we believe can be particularly useful for students in mathematical and computational biology. Depending on institutional needs and preferences, the courses we present here may not include all suggested topics, and some of the listed topics may be omitted. In addition, there are many courses typically not taught in math units that could be appropriate for undergraduate math bio students. Our general advice is that students consider taking as many courses from the list below as feasible and as many courses in biology as possible, based on interest and advisor recommendations.

Modeling applied to biological systems -- A course focusing on (1) the nature of information that can be obtained about various biological systems and how this information factors into the development of corresponding mathematical models, and (2) methods to draw, communicate, and defend reasoned conclusions about the underlying biological systems using such models. **Cognitive Recommendations Addressed:** *CR1, CR2, CR3, CR4.* **Learning Outcomes Addressed:** *SLOM1, SLOM2, SLOM3, SLOM4, SLOM5, SLOM6, SLOM7, SLOM8, SLOM9, SLON1, SLON3, SLON4, SLON5, SLOD1, SLOD2, SLOA1, SLOA4.*

Some possible topics for consideration at the undergraduate level:

- dimensional analysis, scaling, and proportionality
- non-dimensionalization of a mathematical model
- fitting data with least squares
- modeling biological systems with difference equations and mathematical analysis of difference equations
- linear biological models such as age-structured population dynamics models, Leslie matrices
- Markov chains and their use in biological modeling
- chemical reactions occurring within biological systems, especially within cells, and the roles of substrates, enzymes, and products within these reactions
- ordinary differential equations that represent a system of reactions using the law of mass action, and the roles of reaction rates in such models
- Gillespie's algorithm for systems of reactions
- the relationship between deterministic and stochastic representations of chemical reaction systems
- the quasi-steady state assumption
- Michaelis-Menten rate law and Hill functions and their roles in modeling biochemical reaction systems
- positive and negative feedback effects within dynamic models
- concepts of game theory such as two-player games, payoff matrices, evolutionary stable states, Nash equilibria and their relevance to biological systems
- Moran processes and master equations and their use in biological modeling
- the mathematical representation of diffusion
- alternative approaches to spatial modeling - discrete grid space, patches, GIS, PDE
- spatio-temporal growth or pattern formation models and linear stability analysis for transitions from uniform to patterned states in biological systems

Dynamical systems -- A course dedicated to the theory and applications of time-dependent systems and methods to analyze their behavior, in both the discrete and continuous cases.

Cognitive Recommendations Addressed: *CR1, CR2, CR3.* **Learning Outcomes Addressed:** *SLOM1, SLOM2, SLOM4, SLOM5, SLOM7, SLOM8, SLOM9, SLON3, SLON4, SLOA1, SLOA3.*

Some possible topics for consideration at the undergraduate level:

- difference equations vs differential equations (maps vs flows)
- modeling the evolution of complex systems using dynamical systems; use of continuous time vs. discrete time; use of continuous-space vs. finite-state models
- one-dimensional maps, interval maps, fixed points, cobweb graphs, stability of fixed points, periodic points, the logistic map, bifurcation
- fractals, self-similarity and scaling, construction of self-similar fractals, strange attractors
- two-dimensional maps, transition graphs, fixed points, types of fixed points, sinks, sources, and saddles, linear maps vs non-linear maps
- chaos in one and two-dimensional maps, Lyapunov exponent, Lyapunov dimension, Markov partitions, basins of attraction, chaotic attractors, conjugacy
- Boolean and finite dynamical systems, linking structure with dynamics, factors affecting long-term behavior, feedback loops, structural analysis
- synchronous vs asynchronous updates, model assumptions that guide the choice of update schemes - slow vs. fast variables, insufficient information regarding order of variable updates, noise factors
- autonomous systems of differential equations, phase plane analysis, linearization
- fixed points and limit sets for systems of differential equations, classification of fixed points, bifurcation, stability, types of stability, methods for stability analysis, Lyapunov functions, stability of linearization, topological equivalence and conjugacy
- oscillating systems of differential equations, periodic orbits, limit cycles, Dulac's criterion, the Poincare-Bendixson theorem
- chaos in continuous dynamical systems, strange attractors, the Lorenz attractor, Lyapunov exponents for flows

PDEs -- A course devoted to methods to derive and solve or analyze basic classes of partial differential equations, with a focus on linearity at the undergraduate level. **Cognitive Recommendations Addressed:** *CR1, CR2.* **Learning Outcomes Addressed:** *SLOM8, SLON3, SLOA1, SLOA4.*

Some possible topics for consideration at the undergraduate level:

- basic terminology associated with PDEs (order, linear vs. nonlinear, superposition, homogeneous vs. inhomogeneous, initial conditions, boundary conditions)
- well-posedness (existence, uniqueness, stability) for PDEs
- separation of variables
- method of characteristics
- derivations of the heat/diffusion and wave equations and Laplace's equation/Poisson equation
- fundamental solution and Green's function for the heat equation
- maximum principle and uniqueness of solutions to the heat equation and Laplace's equation

- Dirichlet and Neumann boundary conditions, linking these with separation of variables, and Duhamel's principle
- Fourier series and associated concepts (orthogonality, truncation, concepts associated with convergence, periodic extension of a function, sine and cosine series)
- uniqueness theorem for the wave equation
- d'Alembert's solution for the wave equation IVP and the method of images
- principle of causality, domain of dependence, and domain of influence for the wave equation
- Laplace's equation in various coordinate systems and solutions of Laplace's equation on corresponding domains

Stochastic Modeling -- A course devoted to the theory and applications of systems that take into account uncertainty (discrete and continuous) and methods to analyze their behavior. **Cognitive Recommendations Addressed:** *CR1, CR2, CR3, CR4.* **Learning Outcomes Addressed:** *SLOM1, SLOM2, SLOM3, SLOM4, SLOM6, SLOM7, SLOM8, SLOM9, SLON1, SLON3, SLON4, SLON6, SLON7, SLOD1.*

Some possible topics for consideration at the undergraduate level:

- Discrete and continuous random variables and associated probability distributions
- Important particular probability distributions - binomial, normal, poisson, exponential, gamma
- Means, variances, probability generating functions
- Law of large numbers and Central Limit theorem
- Definition of stochastic process
- Markov chains
- Limiting distributions and stationary distributions
- Birth and death processes (discrete and continuous time), master equations
- Branching processes
- General Markov processes
- Simulating stochastic processes
- Time series models
- Cellular automata and agent-based models
- Environmental stochasticity and demographic stochasticity in population biology models
- Applications to conservation biology, enzyme kinetics, epidemiology, genetic sequence analysis

Optimization -- A course providing mathematical foundations and methods of constrained and unconstrained optimization and applications. **Cognitive Recommendations Addressed:***CR1, CR2, CR3, CR4* **Learning Outcomes Addressed:** *SLOM7, SLON2, SLON3, SLON4, SLON5, SLON6, SLON7, SLOD2, SLOA2.*

Some possible topics for consideration at the undergraduate level:

- Identification of the key components of an optimization problem (what can be controlled, what will be affected by the controls)
- Evaluation of the interplay between the underlying model and the choice of objective
- Motivation for the concept and choice of "optimum" (e.g., what are the costs and pay-offs, what should be maximized and minimized in an optimal strategy? For example, in

treatment planning, do we maximize overall well-being? Maximize dosing to the threshold of tolerance?)
- Constrained vs unconstrained optimization, boundary conditions (free or fixed), etc.
- Linear programming/Simplex method
- Lagrange multipliers
- Parameter fitting and sensitivity analysis
- Dual linear programming/dual perspective
- Nonlinear programming
- Concave and convex functions
- Unconstrained optimization
- Network programming, combinatorial optimization (e.g. traveling salesperson problem and so on)
- Challenges of stochasticity in data and how to manage stochasticity.
- Dynamic programming

Introduction to Data Science -- A course designed to provide a general introduction to extracting useful information from large datasets by using relevant principles, theory, and tools from mathematics, statistics, informatics, and computer science. A course with particular focus on biological data would be ideal, but any data science course where biology is used as a disciplinary domain of application would be beneficial. **Cognitive Recommendations Addressed:** *CR1, CR2, CR3, CR4.* **Learning Outcomes Addressed:** *SLOM4, SLOM6, SLOM7, SLON3, SLON6, SLOD1, SLOD1, SLOD3.*

Some possible topics for consideration at the undergraduate level:

- big data; data collection and integration; heterogeneous data; structured vs. unstructured data
- exploratory data analysis; data graphs and plots; summary statistics
- data storage and databases; searching, sorting, and mining existing biological databases; information retrieval
- descriptive vs. predictive modeling
- supervised vs. unsupervised learning
- probability distributions; statistical modeling and analysis
- networks and network analysis
- recommendation systems; mathematical foundations (dimensionality reduction, singular value decomposition, principal component analysis)
- specific machine learning methods (linear regression, clustering methods, hidden Markov models, random forests, support vector machines, neural networks, Markov Chain Monte Carlo methods)
- techniques for effective visualization of results

Bioinformatics -- A course designed to provide an introduction to bioinformatics algorithms and the mathematical and statistical methods underlying them. Ideally, this material would be discussed in the context of evolutionary biology, in particular, molecular evolution and phylogenetics. **Cognitive Recommendations Addressed:** *CR1, CR2, CR3,CR4.* **Learning Outcomes Addressed:** *SLOM1, SLOM2, SLOM4, SLOM6, SLON3, SLON4, SLON6, SLOD1, SLOD3.*

Some possible topics for consideration:

- models of nucleic acid evolution, starting with the Jukes-Cantor model
- models of protein sequence evolution, e.g., PAM
- algorithms for sequence alignment (local and global)
- clustering methods, both hierarchical (e.g., distance based) and non-hierarchical (e.g., K means)
- phylogenetic methods, e.g., neighbor joining and maximum likelihood
- machine learning methods such as hidden Markov, neural networks, support vector machines
- experience with nucleic acid and protein databases
- genomic analysis

Capstone Experience and Extracurriculars

All students should have the chance to put their knowledge to work and to see the huge jump from well-posed classroom problems to the challenges of asking questions and groping for methods in open-ended research. Therefore, in our view, it is essential that students be provided with an undergraduate research opportunity or other capstone experience. The multiple avenues for such opportunities include senior thesis work, interdisciplinary work involving collaboration with a lab, addressing problems supplied by industrial partners, cooperative education programs, formal REU programs, and study abroad experiences. Mentors should also seize this opportunity to help students develop scientific communication skills, both oral and written.

Consistency with Content Goals

(The 2015 CUPM Curriculum Guide lists nine "Content Goals" (see pages 10-13) for all major programs in the mathematical sciences.)

A program in mathematical biology may include less exposure to traditional, theoretical mathematical proofs than some other programs in the mathematical sciences. Students in mathematical biology, however, should gain experience with mathematical arguments of increasing depth relating to the validity of models for representing certain biological scenarios or data, to the derivation of models, and to the derivation of solutions and other useful mathematical frameworks, as they progress through the major.

Biology Training for Mathematical Biology Majors

One of the curricular problems that we confront at the interface of biology and mathematics is trying to put together "packages" of biological content areas and mathematical skills in such a way that they make sense. However, the mathematical skills (and biological content) needed for bioinformatics may be quite different from those needed for biogeochemistry, for example. So for mathematically oriented undergraduates, we believe it would be more appropriate to take an introductory biology course, equivalent to that required for biology majors, then some upper level course in a particular biological subdiscipline. Which upper division courses and how many will depend on the interests of the student and availability of courses and faculty. Some topics that are particularly relevant include:

- Genetics and evolution: DNA, RNA, and the four "forces" of evolution (selection, mutation, migration, and drift).
- Cell biology: cell cycle, central dogma, gene regulation, how cell membranes regulate communication via electrochemical gradients.

- Biological systems and networks: the immune system, the brain, the circulatory and respiratory systems, gene regulatory networks, ecosystems, foodwebs. Systems biology takes a "holistic" approach to study interactions among the components of a biological system and the impact of those interactions on the overall function and behavior of the system.
- Population biology: How populations grow and the main forms of interaction (competition, predation/parasitism, mutualism).
- Biomedical science: Cancer biology as a link between cell biology and evolutionary biology; drug design and testing; epidemiology as an application of population biology

Ultimately, the goal is not just knowledge of content, but also an introduction to biology as a discipline. What constitutes an "interesting" question in biology? How do biologists think about those questions? What are the "standard" approaches, if any? To be successful as mathematical biologists, students must be able to have productive conversations with biologists and biological texts. We further recommend that students have some biology laboratory and/or field experience. Such an experience will also increase student understanding of the challenges involved in designing experiments and collecting data to validate a model.

Implementing a Mathematical Biology Program

A mathematical biology program need not be implemented as a fully formed plan. There are many different routes to arrive at a program in mathematical biology in steps, and a gradual implementation offers advantages, such as the opportunity to make adjustments based on learning from experience and any changes in focus or goals that arise. The extent of the program – major, minor, or just a cluster of courses – may depend strongly on the courses available within a department, as well as the specific goals and learning objectives that a department may set for the program. Developing a mathematical modeling course or a mathematical biology course, if a math modeling course already exists, may be a good first step. In many institutions, a mathematical biology program may arise organically from research collaborations between mathematics and biology faculty and their students. A program may also be created by a partnership between the biology and mathematics departments, starting by designing opportunities for connecting certain mathematics and biology courses – e.g., by offering joint lectures, labs, and projects -- to highlight the value and importance of collaborative efforts between the two disciplines.

References and Resources

Here we provide a short list of links to resources relating to undergraduate training in mathematical biology. The resources in this area are rapidly expanding and constantly in flux, so this list simply represents a small set of starting points.

A. Professional organizations and institutes

Most of these have links to information about mathematical biology curriculum development:

- SMB: http://smb.org/index.shtml
- BioSIGMAA: https://qubeshub.org/community/groups/biosigmaa
- MBI: http://mbi.osu.edu/
- NIMBioS: http://nimbios.org/
- BioQUEST: http://bioquest.org/

- SIAM Activity Group on the Life Sciences: https://www.siam.org/membership/Activity-Groups/detail/life-sciences
- Intercollegiate Biomathematics Alliance (IBA):
- https://about.illinoisstate.edu/iba/
- QUBES: Quantitative Undergraduate Biology Education: https://qubeshub.org

B. Online References

These include links to publications or organizations especially concerned with mathematical biology curriculum development:

- Bio2010: https://www.nap.edu/catalog/10497/bio2010-transforming-undergraduate-education-for-future-research-biologists
- Math/Bio 2010: https://www.maa.org/press/maa-reviews/math-bio-2010-linking-undergraduate-disciplines
- HHMI/AAMC Scientific Foundations for Preparing Future Physicians: http://www.hhmi.org/grants/pdf/08-209_AAMC-HHMI_report.pdf
- Vision and Change in Undergraduate Biology Education: http://visionandchange.org/
- Data Science for Undergraduates: Opportunities and Options: http://sites.nationalacademies.org/cstb/currentprojects/cstb_175246

C. Books

There is no standard set of books agreed upon for use in this field. The following books represent a small sample of resources that have proved useful to the authors of this report.

Calculus Oriented toward Biology:

- F. R. Adler, *Modeling the Dynamics of Life*, Cengage Learning 2012.
- Bodine, Erin, Suzanne Lenhart and Louis Gross. 2014. *Mathematics for the Life Sciences*. Princeton Univ. Press
- James L. Cornette and Ralph A. Ackerman, *Calculus for the Life Sciences: A Modeling Approach* (Available online only at http://cornette.public.iastate.edu/CLS.html)
- J. M. Mahaffy and A. Chavez-Ross, Calculus: *A Modeling Approach for the Life Sciences*, Vol. I and II, (2009 and 2005), Pearson Custom Publishing . Course materials also available on line at http://www-rohan.sdsu.edu/~jmahaffy/courses.html
- Neuhauser, Claudia and Roper. *Calculus for Biology and Medicine*.

Data Analysis:

- N. J. Gotelli and A. M. Ellison, *A Primer of Ecological Statistics*, Second Edition Sinauer 2012
- M.C. Whitlock and D. Schluter, *The Analysis of Biological Data*, W. H. Freeman; Second edition (2014
- J. H. Zar, *Biostatistical Analysis* (5th Edition) Pearson 2009

Modeling Relating to Biological Directions:

- E. Allman and J. Rhodes, *Mathematical Models in Biology*, Cambridge 2003.
- G. De Vries, T. Hillen, M. Lewis, J. Müller & B. Schönfisch, B. (2006). *A course in mathematical biology: quantitative modeling with mathematical and computational methods*. Society for Industrial and Applied Mathematics.

- L. Edelstein-Keshet (2005). *Mathematical models in biology*. Society for Industrial and Applied Mathematics.
- S. Ellner and J. Guckenheimer, *Dynamic Models in Biology*, Princeton 2006.
- Alan Garfinkel, *Modeling Life: The Mathematics of Biological Systems*, Springer, 2017
- Glenn Ledder, Jenna P. Carpenter and Timothy D. Comar (2013) *Undergraduate Mathematics for the Life Sciences Models Processes and Directions*, Mathematical Association of America.
- S. Otto and T. Day, *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*, Princeton University Press, 2007
- R. Robeva and T. Hodge, *Mathematical Concepts and Methods in Modern Biology*, Academic Press 2013