Mathematical Biology Education:
Modeling Makes Meaning

J. R. Jungck

Department of Biology, Beloit College, Beloit, WI 53511, USA

Abstract. This special issue of *Mathematical Modelling of Natural Phenomena* on biomathematics education shares the work of fifteen groups at as many different institutions that have developed beautiful biological applications of mathematics that are different in three ways from much of what is currently available. First, many of these selections utilize current research in biomathematics rather than the well-known textbook examples that are at least a half-century old. Second, the selections focus on modules that are intended for instant classroom adoption, adaptation, and implementation. Instead of focusing on how to overcome the challenges of implementing mathematics into biology or biology into mathematics or on educational research on the effectiveness of some small implementation, the authors develop individual biological models sufficiently well such that they can be easily adopted and adapted for use in both mathematics and biology classrooms. A third difference in this collection is the substantive inclusion of discrete mathematics and innovative pedagogies. Because calculus-based models have received the majority of the biomathematics modeling attention until very recently, the focus on discrete models may seem surprising. The examples range from DNA nanostructures through viral capsids to neuronal processes and ecosystem problems. Furthermore, a taxonomy of quantitative reasoning and the role of modeling *per se* as a different practice are contextualized in contemporary biomathematics education.

Key words: biomathematics education, discrete mathematics, modeling
AMS codes: 92-01

1. Narrative

Mathematics plays an enormous role in the life sciences (Cohen, 2004)[1]. However, the amount and kind of mathematics used in most general biology education is woefully inadequate at most

1

universities. Recently, I co-edited a special issue of *CBE Life Science Education* (Jungck and Marsteller, 2010)[2, 3] to celebrate the progress made by numerous institutions since the 2003 publication of the U.S. National Research Council of the National Academies' epoch-making book: *Bio 2010: Transforming Undergraduate Education for Future Research Biologists's* clarion call for more mathematics in the preparation of biologists [4]. While we were able to share many curricular successes, this issue did not afford us many opportunities to detail specific rich applications to biologists and mathematicians in a way that allowed easy adoption and adaptation into readers' classrooms. Thus, I welcomed the invitation from Editor Vitaly Volpert of *Mathematical Modelling of Natural Phenomena* to edit this special issue devoted to mathematical biology education. In particular, the congruence of interests of readers of this journal and the biomathematics education community are enormous. Besides such interesting biologically-oriented issues of *MMNP* on cancer, morphogenesis and development, plant structure, epidemiology, etc., *MMNP*'s explicit focus on modeling *per se* lets us celebrate the tremendous importance of learning to model in the context of biomathematics education. Several articles herein, particularly by Weisstein, [5], Gaff *et al.* [6], Neuhauser and Stanley [7], and Koch [8], provide accessible approaches to introducing modeling to biology students with little or no prior experience in modeling.

The considerable progress that has been made in addressing the challenges laid down in Bio 2010 [4] includes the mutual effort of biologists, mathematicians, and mathematics-and-science education researchers in the development of new curricular materials, majors, collaborative research initiatives, national institutes, professional development workshops, journals, national science digital libraries, awards, etc. These initiatives have been funded by granting agencies such as NSF, HHMI, NIH, and Sloan as well as by professional societies including: AAAS, AIBS, MAA, SMB, ASM, BSA, and ASCB. In addition to celebrating the progress which has resulted from the investment of considerable resources, we need to plan how should we address recent reports: *AAAS Vision and Change in Undergraduate Biology Education* (2010)[9], *A New Biology for the 21st Century* (2010)[10], *Towards 2020 Science* (2006)[11], *Math and Bio 2010: Linking Undergraduate Disciplines* (2005)[12], *The Fourth Paradigm: Data-Intensive Scientific Discovery* (2009)[12], and *Scientific Foundations for Future Physicians: Report of the AAMC-HHMI Committee* (2009)[13]. These recent reports all allude to several specific challenges that are the basis of my introduction. Twenty-first-century scientists will work with terabytes of data, and we need to prepare both our biology and mathematics students to process, visualize, and interpret huge quantities of multi-dimensional data. We need to move beyond calculus to include more contemporary models, including frameworks that are applicable to analyzing this tsunami of biological data. Future scientists will create innovative models that may include nonlinear, multi-causal, multi-scale, or stochastic components. Additionally, significant changes in pedagogy from teacher-centered to more learner-centered classrooms demand much more attention to minds-on and hands-on activities.

2.   Data Challenge: the Power of Modeling in Biomathematics Education in an Era of Tera

The challenge of data is an old but particularly important one to the fruitful interaction of biologists and mathematicians. At a recent workshop of scientists and mathematicians, the two principal requests from scientists of their mathematician colleagues were (1) data and (2) units. Woodger's (1929) [15] defense of "theoretical biology" stated the need thusly: "The continual heaping up of data is worse than useless if interpretation does not keep pace with it. This is all the more deplorable because it leads us to slur over what is characteristically biological in order to reach hypothetical 'cause'." Compare this to the contemporary claim by Chris Anderson, Editor in Chief of *Wired* magazine: "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete" [16]. Philosopher Massimo Pigliucci rejoins: "If Anderson had been talking about advertizing and how companies should select their targets or fine-tune their merchandizing, he would have been right. But he makes it clear that, '[t]he big target here isn't advertising, though. It's science. [...] Scientists are trained to recognize that correlation is not causation [...] There is now a better way. Petabytes allow us to say: 'Correlation is enough.' We can stop looking for models' But, if we stop looking for models and hypotheses, are we still really doing science? Science, unlike advertizing, is not about finding patternsalthough that is certainly part of the processit is about finding explanations for those patterns" [17]. Gary An (2010) [18] explicitly addresses these misconceptions by showing where the breakdown is occurring in the usual scientific cycle at the experimental evaluation of causal hypotheses. In order to address this problem in the petabyte data world, modeling has been described by multiple authors as the only solution (e.g., see An and Christley, 2011 [19]).

Ever since Richard Levins (1966) [20] classic paper on mathematical biology's 'satisficing' between three disparate goals in modeling (Figure 1), the role of theory is not so easily dismissed as mere prediction. Biologists may be interested in whether a model is robust across a broad phylogenetic spectrum, across spatial and temporal scales, or with respect to multiple forms of causation. Furthermore, the plausibility of causes may be dependent on culturally-complex factors. The nuances of what is a persuasive model to biologists may not be the elegance of the mathematics nor the power of a prediction for a specific case, but may instead be its power for developing new experimental designs, identifying different factors to measure or phenomena to observe, or the ability to conceptualize interacting variables with easily interpretable visualizations.

Clarke and Wiebe (2000) [21] have suggested that the argument between data-driven and hypothesis-driven science can be refocused by examining the difference between "Data-Driven Versus Concept-Driven Visualizations." They write: "Concept-driven visualizations are very common during the pre-activity (lab) phase when the science concept that is the focus of the activity is introduced. This can be a textbook image or from another source. In most cases, ... it represented a simplified or 'idealized' representation of a phenomena or idea. Students will then conduct an experiment/activity where they may be collecting data and creating a visualization. This empirically-driven visualization, as part of their post-lab activities, needs to be part of a reflective process where they link the 'messy' real-world results of their study back to the idealized concept-
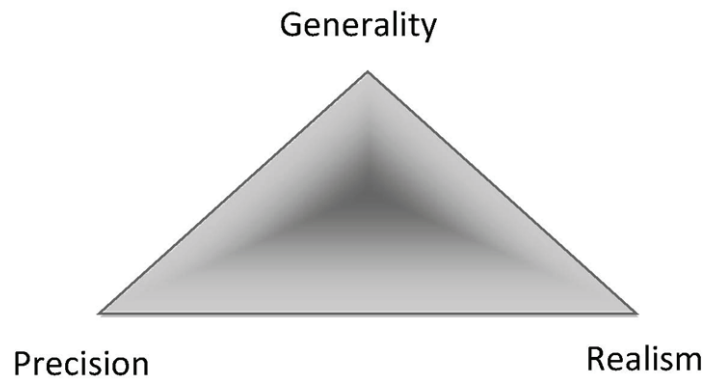
Generality

Precision   Realism

Figure 1: Levins' (1966) [20] model claimed that multiple, simultaneous factors must be considered when judging a mathematical model of a biological phenomenon.

driven visualization. This can be a huge challenge for both the teacher and the student. [We're] very interested in how students make these connections between concept-driven and data-driven visualizations, both of which are becoming an increasingly common part of science lab activities at all levels." Thus, a mathematical analysis that includes abstraction of the empirical data is a crucial part of the synthesis.

The iterative relationship between experimental design, observed data, modeling, and hypothesis testing is explicitly discussed in this special issue on biomathematics education by Neuhauser and Stanley [7]: "One of the fundamental tools of making inferences from data is hypothesis testing. ... Without any prior knowledge of hypothesis testing, however, this [canned] approach does little to convey to the student the principles behind reaching the conclusion. Instead, hypothesis testing remains a rote experience. ... the process is both passive and not informative. ... The problems with this approach are manifold: (1) the appropriate statistical test and test statistic are not motivated; (2) the model assumptions for calculating the distribution of the test statistic under the null hypothesis are not made explicit; and (3) the meaning of the rejection region or the p-value remain obscure." They then elaborate the use of resampling statistics in the context of an actual biological context (seed development in plants) to illustrate how all three of these challenges can be addressed. "In particular, [their] approach emphasizes the mathematical model that underlies the null hypothesis." Throughout this issue, you will be struck by the congruence of the authors with respect to their concern that students unpack the process of modeling, hypothesis formation, and the rigorous testing of these conceptual foundations.

The use of statistics is widespread in biology. Powerful statistical packages, spreadsheets, and many web tools are easily available to compute statistics. Also, very inexpensive calculators have built-in two-variable statistical functions for mean, variance, standard deviation, regression, and Pearson product-moment correlation. As Goldstein (2006) [22] says, "developments in IT have revolutionized the interactive ability to utilize statistics, graphically and computationally, so that very large, complex and hence, intrinsically interesting data sets, can be handled within the cur-

riculum." Thus, our classroom focus is no longer on the statistical computations, but instead on the development of statistical judgment. Before data collection begins, we want students to ask, "What test should I plan to use?" After data collection, we want students to observe patterns in the data and ask: "How do I test if this pattern is accurate?" We want students to develop causal models, and ask, "Is my sample size from this experiment big enough to test my new hypothesis?" "How can I be sure that random bad luck didn't cause me to reject a valid hypothesis or support a false hypothesis?" Clifford Konold and colleagues developed a variety of statistical tools based on constructivist learning principles that were heavily influenced by our belief that the best way to motivate students to learn and use statistical techniques is through exploring issues of concern to them. [They were] designed to be computationally powerful yet simple to use. ... The software was intended for use in introductory courses stressing exploratory data analysis of fairly large data sets. It encouraged students to make initial judgments of relationship by visually comparing plots" (Konold and Higgins, 2003) [23]. Furthermore, we want our students to intelligently consider the options of including or excluding outliers in their experimental data. Statistics builds on all of the counting, measuring, graphing, data mining, analysis, problem solving, etc. activities described heretofore as well as developing keen judgment to deal with ambiguous, uncertainty, imprecision, inaccuracy, unwarranted assumptions, etc. Hence, as with problem-solving, we believe in practice, practice, practice–the "Carnegie Hall" effect in *Science* (Haak *et al.*, 2011) [24] and the value of many different types of problems that employ statistical analysis of data.

When biologists test their models, they have different criteria in addition to Levins' [20] three criteria for weighing alternatives in constructing models. I have developed the triangle below (Figure 2) in an attempt to be parallel to Levins' triangle and in order to eliminate confusion due to mixing categories. First, most biologists simply look at the *a posteriori* probability of a T-test, F-test, or Chi-squared goodness-of-fit tests. In some cases, such as regression, they simply report an R^2 without even looking up a probability. Ziliak and McCloskey (2008) [25] have excoriated the over-reliance on such "statistical significance" to evaluate scientific claims. Second, in addition to evaluating the improbability that a certain event has occurred by chance alone, we also need to evaluate the *possibility* that the event can indeed occur. For example, if a rare phenomenon has occurred already, its possibility is one even though its probability is low. Fuzzy logic (Zadeh, 1978 [26]), Bayesian statistics, hidden Markov models, and artificial intelligence (neural networks, perceptrons, genetic algorithms, Darwinian programming, Dempster-Schafer theory) approaches often emphasize this distinction. Third, although plausibility is infrequently set on a quantitative scale, it is so deeply embedded in the biologists' intuitive knowledge of the causal bases of their biological system that plausibility will almost certainly be included in their criteria for evaluating mathematical models. Friedman and Halpern (2001) [27] have introduced some plausibility measures" that are more formal than those embraced here, but I would simply posit that many articles herein employ biological "plausibility" as a criterion for evaluating the worth of their models rather than computing a statistical significance or predicting a numerical value for a particular biological phenomenon.

Possibility

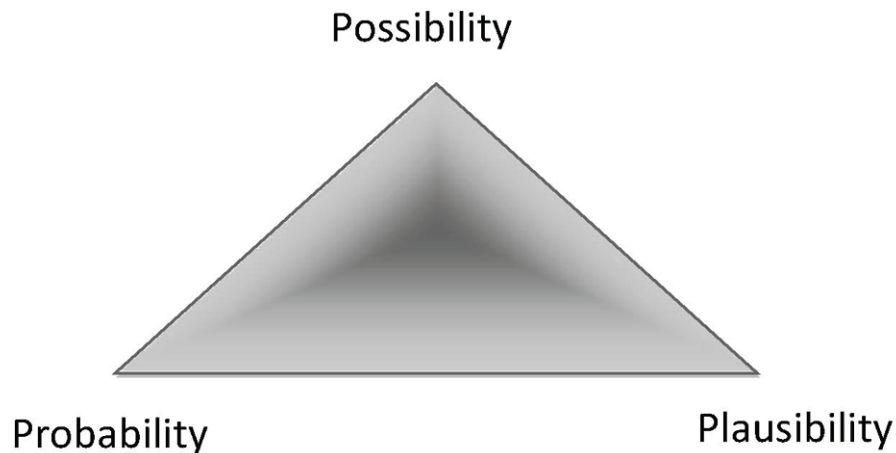Probability   Plausibility

Figure 2: Three criteria for evaluating a model: Probability, Possibility, and Plausibility. As with Levins' original triangle for distinguishing criteria that biologists use in modeling as opposed to traditional models proposed by physicists that emphasized precision, accuracy, repeatability, etc., biologists effectively weigh multiple constraints that they satisfice in their integrated evaluation. See Qi (2008) [28] for a review of formal measures of iterated belief revision.

Thus, while data describe an event that has happened, mathematical models allow students to explore a larger world of what might happen. Schwarz *et al.* (2009) [29] emphasize: "Models make it possible to go beyond observables and imagine a world not yet seen. Models enable predictions of the form "if ... then ... therefore" to be made in order to test hypothetical explanations." Or, as the paleontologist Dolf Seilacher says, "I wouldn't have seen it unless I believed it." Or, on the other hand, Cantor said "I see it, but I don't believe it" (in suggesting the need for a better mathematical proof). If mathematics is to play a more significant role in biology education, the power of modeling as a means to make sense of data, to suggest data acquisition that has the potential to test causal hypotheses, and to construct broad theoretical syntheses of disparate kinds of data must be emphasized and instantiated in educational activities such as found herein:

> A successful model assumes attributes which last far beyond the problem-definition stage or its use in specific research. The model may become the dominating paradigm in the field, not only in the sense of directing scientific inquiry but in the way the student and the public think about an issue. ... Because of the nature of models the teacher of science or mathematics must be alert, since students are easily seduced. Students need constant reinforcement of the fact that it is reality, the phenomenon, which if the focus of study. The model is a tool and modeling a process. ... Modeling is one of the most basic tools in the emerging information society. Modern decision-making, at all levels, tend to be based on modeling. The better the model, the better the decision. The future citizen, whether a scientist or an office worker, should understand the strengths and mechanisms of models but the limitations and pitfalls as well. The modeling process may be the best way to develop this understanding. (Ost, 1987 [30])

Of course, models which help us to draw conclusions from data are only a subset of the wide array of modeling approaches that are valuable in extending our understanding of biology. So, simply, why model? Odenbaugh (2006) [31] identifies five pragmatic uses for models in biology:

1. Simple models help biologists explore complex systems;

2. Models can be used to explore possibilities;

3. Models can lead to the development of conceptual frameworks;

4. Models can make accurate predictions; and,

5. Models can generate explanations.

While globally modeling can contribute enormously to theory construction, students can iteratively modify their assumptions in the process of building models in order to make their models better fit their expectation of biological behavior. Indeed, as students, or for that matter scientists, develop a mechanistic model that is based on clear assumptions of behavior and that produces seemingly robust results, a particularly powerful phenomenon occurs when the model outputs a counterintuitive result that causes the model creator to deeply rethink both the assumptions that are considered as input as well as the assumption of what are plausible results. As George Box (1979) [32] said: "All models are wrong; some models are useful." In his article: "False models as means to truer theories." the philosopher of biology, William C. Wimsatt (1987) [33], asks that while we know every model that we ever build is false even before we test it as a result of the variety of necessary simplifying assumptions, are there deeper reasons than simple utilitarian criteria for considering modeling valuable? Wimsatt explores how modeling can be so important to contemporary scientific practice, and he emphasizes that good models capture most of the variance with a minimal number of variables. On the other hand, he details the value of "bad" models: "One often designs a false model deliberately for use as a template to compare with data when one is interested not in how well it fits the data, but in the form of the residualswhere it does not fit, how and why. This can be a powerful tool in the construction of new theory." Thus, there must be some other criteria that scientists use in judging the heuristic value of the modeling process:

> Any model must make some assumptions and simplifications, many of which are problematic, so the best working hypothesis would be that there are no bias free models in science. This observation has a parallel in the question, 'What variables must be controlled for in an experimental design?' There are no general specifications for what variables should be controlled, since (1) what variables which should be controlled or factored out through appropriate attempts to isolate the system, (2) what variables should be measured, (3) what errors are acceptable, and (4) how the experiment should be designed are all functions of the purpose of the experiment. Similarly, (1) what models are acceptable, (2) what data are relevant to them, and (3) what counts as a 'sufficiently close fit' between model and data is a function of the purposes for which the models and data are employed. (As one referee for this paper pointed out,

many people who would applaud the use of appropriate controls and isolations in an experimental design inconsistently turn on mathematical or causal models of the system and criticize them for doing the same thing! This activity is desirable and necessary in either case.) (Wimsatt, 1987 [33])

More recently, Harry Bhadeshia (2008) [34] has identified four aspects of modeling:

  (i) those which lead to an unexpected outcome that can be verified;

 (ii) those which are created or used in hindsight to explain diverse observations;

(iii) existing models which are adapted or grouped to design materials or processes; and,

(iv) models used to express data, reveal patterns, or for implementation in control algorithms.

   ...Like ordinary science, the proper use of models also leads to insight, but by tackling complexity at the level that is posed, it can reveal issues which are lost during the simplification characteristic of ordinary science.

   I believe that to some degree or another, all four of these aspects of modeling are illustrated in the articles herein and thus represent the power of modeling in understanding problems across the diversity of the life sciences.

   If we return to students engaged in modeling, we can benefit from a long-lasting science education research project entitled MUSE: Modeling for Understanding in Science Education. After Jim Stewart and two of his former doctoral students (2010) [35] observed how middle and high school students solve genetics and evolution problems, they developed a Taxonomy of Problem Solving that posits four levels of Problem Solving:

   Model-Less Problem Solving
   Model-Using Problem Solving
   Model-Revising Problem Solving
   Model-Building Problem Solving

   Just as with Perry's developmental scheme of cognitive growth in student appreciation of what education is trying to achieve, Stewart *et al.*'s taxonomy helps us recognize that when students are able to build models, they are also able to have an enhanced ability to solve problems at an expert rather than at a novice level. Similarly, numeracy has been partitioned by several authors to involve equation literacy (see especially Weisstein herein [5]), spatial and temporal literacy, and abstraction. These trajectories are intended to help students develop beyond arithmetic and algorithmic skills to heuristics for making inferences, persuading peers, dealing with complexity, ambiguity, and uncertainty. Julie Svoboda (personal communication), a science education researcher from the University of California Davis, found that students developed a much better sense of epistemology

(how do we know that we know) and that their cognitive payoffs included the ability to do meta-analysis of multiple models and thereby they were able to transfer their modeling insights to other subject areas.

How do we scale up modeling education to adequately prepare students for dealing with the shift to terabytes of data per day? Numerous educational programs in bioinformatics (see our project: BEDROCK: Bioinformatics Education Dissemination: Reaching Out, Connecting, and Knitting-together <http://bioquest.org/BEDROCK>) (Jungck, 2009) [36], geographic information systems (see My World GIS <www.myworldgis.org/>), medical informatics, ecoinformatics (<ecoinformatics.org>), biocomplexity (<http://knb.ecoinformatics.org/index.jsp>), biodiversity informatics (<http://www.gbif.org/>), and others have already been developed. With terabytes of data arising daily, many more initiatives are almost sure to be developed. Thus, while there will surely be sufficient data available to support educational programs, our challenge is in the development of learning environments that enable students to develop original hypotheses and to use powerful software and hardware to investigate these rich datasets. Furthermore, since such databases almost inevitably are associated with social and ethical controversies such as stem cell usage, global warming, coral reef bleaching, overfishing, deforestation, extinction, invasive species, and pollution, a contemporary education will need to better prepare students for operating in such a contentious and environment. Ademir Donizeti Caldeira (2009) [37] reports that it is possible to achieve both engagement in mathematical modeling and community issues:

> The study also showed that it is possible, in the process of learning about mathematics and environmental issues, for teachers to develop a critical competence, without it being imposed, but rather constructed through the experiences of the actors themselves; and that it is also possible to discuss a curriculum that questions the functions of school contents, the interests and assumptions reflected in these contents, and finally, the possible social interventions they can catalyze (or slow down), and under what circumstances. The approach adopted for the study, based on the theoretical-methodological assumptions of mathematical modeling, showed us that it is possible for teachers to become involved in situations in their social context, and that the problems encountered there, when perceived as relevant, and because they are objectively existent social problems, lead to the critical engagement of the teachers in their communities.

This special issue devoted to mathematical biology education of *Mathematical Modelling of Natural Phenomena* focuses extensively on modeling. The activities described herein engage students in the reflective practices of building, using, and evaluating models. Our hope is that these activities will spur similar excitement, engagement, and evaluation amongst other students about the power and beauty of insight gained through mathematical modeling.

3. Mathematical Challenges in Biology Education

While many individual approaches for linking mathematics to specific biological problems have been developed, there has been a lack of systematic approaches to developing mathematical reasoning in biology curricula. Too often, mathematical examples are incorporated only to emphasize biological topics with no attention to simultaneously developing students' mathematical reasoning. Lynn Steen (2003) [38] states: "Were schools to design mathematics programs expressly for work and careers, the selection of topics, the order in which they are taken up, and the kinds of examples employed would be substantially different." One of the authors, Grant Wiggins (2003) [39], suggests criteria for "Quantitative Reasoning" or "Quantitative Literacy," namely: "A. Confidence with Mathematics; B. Cultural Appreciation; C. Interpreting Data; D. Logical Thinking; E. Making Decisions; F. Mathematics in Context; G. Number Sense; H. Practical Skills; I. Prerequisite Knowledge; and J. Symbol Sense." While these concepts may be valuable to identify as goals, I believe they do very little for educators who want students to achieve these goals. Also, in the same volume, his colleagues Randall M. Richardson and William G. Mccallum (2003) [40] single out biology as expecting that quantitative reasoning means: "Life Sciences: Throughout these recommendations, the definition of mastery of a mathematical concept recognizes the importance of both conceptual understanding at the level of definition and understanding in terms of use/implementtation/ computation." Koch (2011) [8] states an alternative challenge as: "As we teach our students various topics in mathematics, how often do we give them the opportunity to really *participate and do mathematics*?" In an attempt to address the challenge of infusing a more relevant mathematical education based upon participation and "doing mathematics" into contemporary biology education, in the following I suggest an implementable sequence of steps (analogous to Bloom's taxonomy for learning [41]) that is intended to develop students' metacognitive growth in mathematical thinking through explicit activities . These activities engage students in practices ranging from (1) counting, measuring, and calculating to (2) graphing, mapping, and ordering, to (3) problem solving, to (4) analysis, to (5) hypothesis testing, and, to (6) modeling (Figure 3). Each of these six sets of procedures most be reached through opportunities for students to engage in open-ended investigation. While many mathematical tools within the investigations may be elementary, it is imperative that the investigations are not perceived as a remedial mathematics activity. It is equally essential that the students cannot be passive recipients of information. I contend that it is only through serious commitment to inclusion of investigations addressing all six procedures that we will be able to develop the hands-on and minds-on involvement that will result in science graduates who have a deep sense of where data come from, how to actively draw inferences from data, and how to build and use models appropriately.

Through a variety of projects, NUMB3R5 COUNT (Numerical Undergraduate Mathematical Biology Education <http://bioquest.org/numberscount/>), Microbes Count! (<http://bioquest.org/microbescount/>), BEDROCK (Bioinformatics Education Dissemination: Reaching-Out, Connecting, and Knitting-together <http://bioquest.org/bedrock/>), and the Biological ESTEEM Project

(Excel Simulations and Tools for Exploratory, Experiential Mathematics <http://bioquest.org/esteem/index.php>), we (the BioQUEST Curriculum Consortium), will share interactive learning materials to engage students with each of these six processes.
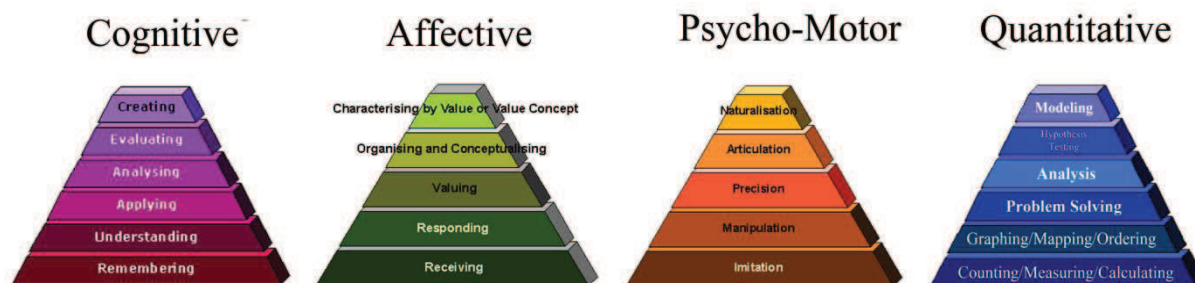

Figure 3: Bloom's taxonomies. Bloom [see Krathwohl's (2002). A Revision of Bloom's Taxonomy: An Overview; [41]] originally identified three hierarchies: (a) Cognitive (revised version); (b) Affective; and (c) Psycho-motor. (d) I have developed a fourth taxonomy for "Quantitative Reasoning in Biology" that is intended to help biologists identify procedures and skills that they are trying to help their students learn and to have a better appreciation for a variety of ways in which mathematics deeply informs biological science.

I have been influenced very much by the Dutch mathematician and theorist of mathematics education, Hans Freudenthal. His "Weeding and Sowing" book published in 1980 [42] particularly had a profound influence on my teaching and research. I believe that the focus on activities in the classroom that build from the everyday experiences of biology students to the more abstract process of modeling is very consistent with Freudenthal's admonition to: "allow learners to come to regard the knowledge they acquire as their own, personal knowledge, knowledge for which they themselves are responsible." Ironically, Freudenthal was an extremely harsh critic of Bloom who "he accused ... of conceiving of learning as a process in which knowledge is poured into the heads of students" (Gravemeijer and Terwel, 2000) [43]. Instead Freudenthal recognized that "educational processes [are] discontinuous: from rich, complex structures of the world of everyday-life to abstract structures of the world of symbols and not the other way around. Starting points should be found in situations that 'beg to be organized' where .... categories are not predefined but are developed by the learners themselves, and need to be accommodated to their needs" (Gravemeijer and Terwel, 2000) [43].

Thus, my pyramid of quantitative reasoning in biology emphasizes activities that students are engaged in and the production of scientific visualizations, tests, and models that they "own" rather than predetermined objects mandated by instructors or a sort of prescriptive learning hierarchy from simple to complex. Considerable iteration, recursion, and reflective interaction of activities identified in the pyramid would reasonably take place in a feedback and feedforward cybernetic loop rather than as a simple linear progression. The hierarchy is in terms of mathematical metacognition rather than privileging one type of scientific activity over another. In other words,

the pyramid is meant to reflect a process where the students' construction of meaning in the world grows from the concrete, procedural to the abstract, formal because the growth of reflection and interpretation as the students participate in more "mathematical thinking" that encourages them to embrace self-critical aspects of creativity, originality, curiosity, and introspection. Again in Freudenthal's aspirations: "allow learners to come to regard the knowledge they acquire as their own, personal knowledge, knowledge for which they themselves are responsible" (Gravemeijer and Terwel, 2000) [43]. Randa Khattar and Carol Anne Wien (2010) [44] state that in "most traditional teaching, we can easily see why little new can emerge in such restricted contexts. Paired with the fact that ideas must be able to interact is the notion that unilateral teacher control of structure and collective results must be relinquished. One must give up control if complexity is going to happen. Decentralized control is a necessary condition of emergence." It is just this dual process of independence in exploration and responsibility for their own work that provides a tension essential to their development as responsible citizens and as potential scientists.

Mette Andresen (2009) [45] from Denmark argues that in Freudenthal's approach: "A new mathematical reality is created at each level. Reflections substantiate the progressive mathematizing" (Figure 4) and "Students' modeling processes take place when the student or, most commonly, a group of students, start with a more or less authentic word problem and build a mathematical model on their own, solve the problem mathematically and transfer the solution back to the real world situation" [Andresen (2007), [46]]. In a similar vein, George Gadanidis and Vince Geiger (2010) [47] argue: "that the confluence of these developments means that it is necessary to re-examine what it means to learn and do mathematics and proposes that it is now possible to view learning mathematics as an activity that is performed rather than passively acquired." Doorman and Gravemeijer (2009) [48] assert that activities and modeling are central in developing students' reasoning: "The conjectured process of teaching and learning is supposed to ensure that the mathematical and physical concepts will be rooted in students' understanding of everyday-life situations. Students' inventions are supported by carefully planned activities and tools that fit their reasoning. The central design heuristic of the instructional sequence is emergent modeling."

If we begin to use the pyramidal hierarchy above (Figure 3) as a lens to understanding the motives of the authors in this issue of *Mathematical Modelling of Natural Phenomena*, Neuhauser and Stanley (2011) [7] demonstrate that counting can lead to many unsuspected directions of inquiry by testing contradictory selection pressures and deriving distributions from sample data. They take students from their practical every day experience through several levels of hypothesis testing and modeling. Given the focus of *Mathematical Modelling of Natural Phenomena* on modeling *per se*, it is not surprising that most authors have chosen to simply focus on educating students about modeling. Nonetheless the richness of their approaches is vast. Not only do they encompass a variety of mathematics, it is important to the biological community that they range from a focus on the following topics:

Molecular Biology: Konrashov (2011) [52] models the dynamics of the protein calmodulin as a coupled oscillator with differential equations in order to understand such phenomena as the
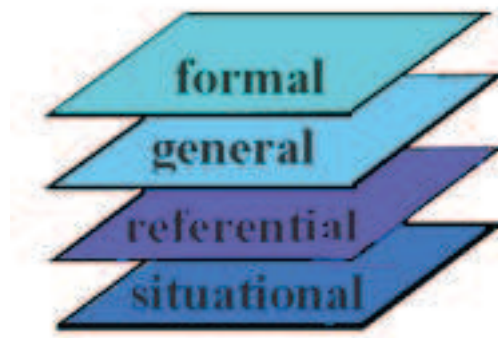
Figure 4: Four levels of Freudenthal's approach to realistic mathematics education from Gravemeijer, and Stephan (2002) [50]. Emergent models as an instructional design heuristic. In Gravemeijer *et al.* (2002), pp145–169 [51], it is the learner's growth of reflection and abstraction that is being stressed by these multiple mathematics educators from different countries.

opening and closing of an active site of an enzyme and how that might be used to predict intermediate configurations of catalysis; Ellis-Monaghan and Pangborn (2011) [53] use graph theory to offer physical builders of DNA structures how to chose sequences that will self-assemble into particular three-dimensional lattices such as cubes, tetrahedra, and octahedral; and, Robic and Jungck (2011) [54] demonstrate the power of mathematical manipulatives in helping students understand topological processes accompanying replication, recombination and repair.

Structural Biology: In another system involving self-assembly, namely that of full viral capsids, Kerner (2011) [55] investigates the combinatorics and symmetry of protein dimers in order to form all possible icosahedral viral capsids as well as develop a good classification system (affinity tables) for the variety of icosahedral viral capsids. He helps students begin to address questions about the susceptibility of configurational change to mutation and which viral capsids might be closer evolutionarily to one another.

Cellular Biology: Robeva, Kirkwood, and Davies (2011) [56] investigate genetic regulation of the lack operon with a clever Boolean logic approach that helps students develop, validate, and refine a dynamic understanding of positive and negative feedback as well as associated allosteric conformational changes in repressors and relating them to mathematical steady states, fixed points, limit cycles, and systems of polynomial equations. Gill, Shaw, Rountree, Kehl, and Chiel (2011) [57] have developed a simulation of lattice models in order to demonstrate "how autopoietic behavior, i.e., behavior of biochemical networks that is self-maintaining, may have emerged spontaneously from the fundamental rules of physics and chemistry." While they have particular expertise on the electrochemical properties of neurons and their networks, they illustrate how their alternative methodologies and a focus on spatial heterogeneity might be extended from the biochemical to cellular to ecological levels.

Organismal Biology: Koch (2011) [8] has intrigued her students with the counterintuitive results produced in pharmacokinetics. By using half-lives of a spectrum of drugs and one- and two-compartment pharmacokinetic models, she helps students explore such questions as what happens if someone misses just one birth control pill (Aygestin®), how long will it be before they are "safe" again? She shares a database of half-lives, recommended dosages, and time intervals of a variety of over-the-counter and prescription drugs. Pedagogically, she shares that: "There may be too much freedom in allowing students to select a drug of choice, as students may choose to model recreational drugs. It is more important to me to have the lab be meaningful to my students; if they get to model a drug of choice, they *own* that model." Similarly, Milton, Radunskaya, Ou, and Ohira (2011) [58] help interdisciplinary student research teams explore questions about themselves related to maintaining balance by using data collected in the lab on human motion and modeling inverted pendulums with time-delay models, introducing noise, and using dimension reduction. They motivate their students with a societal challenge: "the impending epidemic of falling in aging Western societies. How is the risk of falling for a subject to be estimated without causing the person to fall?" In addition to their focus on mathematical biology education, they are particularly concerned that students have the opportunity to learn skills essential for success in today's workplace including productive team work, critical thinking, problem solving, project management, and effective communication ... The success rate of these programs becomes particularly high when projects are relevant to student interests and the skills are transferable to other problem settings."

Ecology: Cozzens (2011) [59] uses graph theory principles such as interval graphs, competition graphs, boxicity, and weighted edges to help students explore food webs. In addition to helping students understand such community ecology principles such as resource partitioning into niches and trophic levels, she helps them see how competition between predators for common food sources and for places to raise and protect their offspring can structure communities. Current issues around maintaining biodiversity can be addressed by students as they explore the stability of ecosystems to biotic and physical change. Hartvigsen (2011) [60] also explores the power of graph theory as a lens to analyze complex networks such as more complicated food webs, plant-herbivore communities, interactomes, and vaccinations' impact on disease transmission. He lays out explicit instructions for using the free statistical package R along with igraph to represent and analyze such complex networks. He considers Watts-Strogatz small-work networks, Barabasi scale-free networks, and Erdós-Rényi random networks. The power law distribution of degrees per vertex is used powerfully to discriminate between the three models. Knisley (2011) [61] applies the Poisson, exponential, and Erlang distributions and compartmental models to investigate migratory dynamics Meanwhile he demonstrates the power of Bernoulli approximations, multivariable Taylor series, and the Maclaurin series representation of the exponential for biological problems and introducing concepts of calculus. He suggests how a model of bird migration might be altered to investigate problems ranging from protein-protein interactions to fossils to massive migration of land mammals. In this sense, Knisley argues that this model may serve as a canonical example for biology in much the same way that the harmonic oscillator does for physics education. An important inclusion of plant modeling, photosynthesis, respiration, and resource partitioning

of carbohydrates to different plant parts by Grossman, Berdanier, Custic, Feeley, Peake, Saenz, and Sutton (2011) [59] invites students to grow their own plants under different conditions and to compare actual plant growth with their model. Besides performing an analysis of covariance (ANCOVA) and multiple regressions of observed and expected values over time for different plant parts under different light conditions, they also perform a sensitivity analysis to infer that respiration rates estimates influenced behavior of their model much less than photosynthesis rates. Students could better understand the tradeoffs made by breeders who developed new plant strains during the "Green Revolution" and valued the expertise that they gained in the modeling experience. The authors suggest that users of their model could explore contemporary topics such as the growth of biofuels and using forests as carbon sinks to reduce global warming.

Epidemiology: Both Weisstein (2011) [5] and Gaff, Lyons, and Watson (2011) [6] focus on building epidemiological models. However, their approaches and techniques are very different. Weisstein (2011) focuses on "equation literacy" and moves students through a series of: (1) Graphical Models to Encapsulate Biological Knowledge, (2) Writing "Word Equations," (3) Translating Word Equations into Formal Equations, (4) Implementing the Model, and finally, (5) Extending the Model. His activity emphasizes the processes of identifying key biological assumptions, translating those assumptions into simple discrete-time recursion equations, and explaining the specific mechanisms by which changes in each parameter affect the simulated outcome. ... [in order] to equip students with a powerful and broadly applicable tool for approaching scientific questions, and to empower students to unpack mathematical relationships into explicit hypotheses that are subject to critical evaluation. More fundamentally, however, [he intends] this modeling activity to help students build a more mature understanding of scientific processes – an understanding that embraces stochasticity and competing processes, spans multiple levels of biological organization, and yields quantitatively testable predictions."

On the other hand, Gaff, Lyons, and Watson (2011) [6] take a very kinesthetic approach to learning modeling by the use of manipulatives to build an iterative and intuitive conception of the modeling and revision process. Their article gives concise and clear instructions for how to engage students in a hands-on exercise in basic epidemiology. The variations on the exercise should greatly strengthen students' understanding of the meanings and ramifications of the various SIR model configurations. The reader of the article could easily implement the exercise in classes with greatly differing goals and student preparation levels.

Therefore, the fifteen articles in this special issue of *Mathematical Modelling of Natural Phenomena* on biomathematics education easily demonstrate a breadth of biology from molecular to the population levels and a similar breadth of mathematics from graph theory to topology to differential equations. It is our collective hope that these serve to stimulate others to share their classroom expertise and innovations with a broader community in order to address some of the challenges that we face in the current and forthcoming "era of tera."

4.   Conclusion: Mathematical Biology Education

While each article in this special issue of *Mathematical Modelling of Natural Phenomena* on biomathematics education develops an activity or series of activities, we urge readers to take a more synthetic approach than simply incorporating one or two activities into their individual courses. Herein we have tried to raise larger issues about how students learn to connect biology and mathematics in deep ways such that they will see natural interdependencies of both fields and move beyond a strictly utilitarian approach so that they appreciate deep questions, aesthetic perspectives, and formal modes of reasoning in different disciplines. For faculty, we hope that they will develop new courses, curricula, and collaborative research projects. While models for each of these are presented in this introductory essay, they are presented to stimulate others to develop their own approach to student development and the intersection of biology and mathematics. Only when mathematics is seamlessly interwoven throughout the tapestry of biology education and biology is seamlessly interwoven throughout the tapestry of mathematics education will we be able to reflect upon which models serve our students best in terms of being prepared for living with the deluge of data and complex issues that we currently face.

Acknowledgements

advice for improving each and every one of the manuscripts. Finally, for over four decades, numerous undergraduates have helped me explore a rich variety of mathematical biology problems. Our co-learning has been a joy.

References

[1] J. Cohen. *Mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better.* PLOS Biology, 2, (2004), 439.

[2] J. Jungck, P. Marsteller, editors. Bio 2010: Mutualism of biology and mathematics. A special issue of CBE Life Science Education, 9, (2010) No. 3. Available from: (http://www.lifescied.org/content/vol9/issue3/index.dtl).

[3] J. Jungck. *Ten equations that changed biology.* Bioscene, 23 (1997), No. 1, 11–36.

[4] Board on Life Sciences. National Research Council. BIO2010: Transforming undergraduate education for future research Bbologists. National Academies Press: Washington, D.C., 2003.

[5] A. Weisstein. *Building mathematical models and biological insight in an introductory biology course.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 198–214.

[6] H. Gaff, M. Lyons, G. Watson. *Classroom manipulative to engage students in mathematical modeling of disease spread: 1+1 = Achoo!.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 215–226.

[7] C. Neuhauser, E. Stanley. *The p and the peas: An intuitive modeling approach to hypothesis testing.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 76–95.

[8] G. Koch. *Drugs in the classroom: Using pharmacokinetics to introduce biomathematical modeling.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 227–244.

[9] AAAS Vision and Change in Undergraduate biology education: A call To action. American Association for the Advancement of Science, Washington, D.C., 2011.

[10] National Research Council. A New Biology for the 21st Century: Ensuring that the United States Leads the Coming Biology Revolution. National Academies Press, Washington, D.C., 2009.

[11] S. Emmott, S. Rison, Editors. Towards 2020 science. Microsoft Corporation, Cambridge, 2006, http://research.microsoft.com/en-us/um/cambridge/projects/towards2020science/downloads/t2020s_report.pdf

[12] L. Steen, Editor. Math and Bio 2010: Linking Undergraduate Disciplines. Mathematics Association of America, Washington, D.C., 2005.

[13] T. Hey, St. Tansley, K. Tolle, Editors. The fourth paradigm: Data-intensive scientific discovery. Microsoft: Redmond, Washington, 2009. (http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf).

[14] Scientific Foundations for Future Physicians: Report of the AAMC-HHMI Committee. Association of American Medical Colleges, Washington, D.C., 2009. (http://www.hhmi.org/grants/pdf/08–209_AAMC-HHMI_report.pdf).

[15] J. Woodger. Biological principles : a critical study. Harcourt, Brace, London, 1929.

[16] C. Anderson. *The end of theory: The data deluge makes the scientific method obsolete.* Wired, 16 (2008) 7.

[17] M. Pigliucci. *The end of theory in science?* EMBO Reports, 10 (2009), 534.

[18] G. An. *Closing the scientific loop: bridging correlation and causality in the petaflop age.* Sci Transl Med., 2 (2010), No. 41, 34.

[19] G. An, S. Christley. *Agent-based modeling and biomedical ontologies: a roadmap.* Computational Statistics, 3 (2011) No. 4, 343-356.

[20] R. Levins. *The strategy of model building in population biology.* American Scientist, 54 (1966) 421–431.

[21] A. Clark, E. Wiebe. *Scientific visualization for secondary and post-secondary schools.* Journal of Technology Studies, 26 (2000), No. 1.

[22] H. Goldstein. *The future of statistics within the curriculum.* Teaching statistics, 29 (2006), No. 1, 8–9.

[23] C. Konold, T. Higgins. *Reasoning about data.* In J. Kilpatrick, W. Martin, D. Schifter (Eds.), A research companion to principles and standards for school mathematics, Reston, VA, National Council of Teachers of Mathematics, (2003), 193–215.

[24] D. Haak, J. Hille, R. Lambers, E. Pitre, S.Freeman. *Increased structure and active learning reduce the achievement gap in introductory biology.* Science, 332 (2011), 1213–1216.

[25] S. Ziliak, D. McCloskey. The cult of statistical significance. The University of Michigan Press, Ann Arbor, 2008.

[26] L. Zadeh. *Fuzzy sets as a basis for a theory of possibility.* Fuzzy Sets and Systems, 1 (1978), 3–28.

[27] N. Friedman, J. Halpern. *Plausibility measures and default reasoning.* Journal of ACM, 48 (2001), No. 4, 648-685.

[28] G. Qi. *A semantic approach for iterated revision in possibilistic logic.* AAAI, (2008), 523–528.

[29] C. Schwarz, B. Reiser, E. Davis, L. Kenyon, A. Acher, D. Fortus, Y. Shwartz, B. Hug, J. Krajcik. *Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners.* Journal of Research in Science Teaching. 46 (209), No. 6, 632–654.

[30] D. Ost. *Models, modeling and the teaching of science and mathematics.* School Science and Mathematics, 87 (1987), No. 5, 363–370.

[31] J. Odenbaugh. *The strategy of model building in population biology.* Biology and Philosophy, 21 (2006), 607–621.

[32] G. Box. *Robustness in the strategy of scientific model building.* (May 1979) in R. Launer, G. Wilkinson, Editors, Robustness in Statistics: Proceedings of a Workshop, 1979.

[33] W. Wimsatt. *False models as means to truer theories.* In M. Nitecki, editor, Neutral models in biology; Oxford University Press, Oxford, (1987), 23–55.

[34] H. Bhadeshia. *Mathematical models in materials science.* Materials Science Technology, 24 (2008), 128–136.

[35] J. Stewart, C.Passmore, J. Cartier. *Project MUSE: Involving high school students in evolutionary biology through realistic problems and causal models.* Biology International, 47 (2010), 78–90.

[36] J. Jungck. *Genetic codes as codes: Towards a theoretical basis for Bioinformatics.* In R. Mondaini (Universidade Federal do Rio de Janeiro, Brazil), Editor. BIOMAT 2008. World Scientific, Singapore, (2009), 300–331.

[37] A. Caldeira. *Mathematical modeling and environmental education.* Proceedings of the 11th International Congress on Mathematics Education, Monterrey, Mexico, July 6 - 13, 2008, (20009), (http://tsg.icme11.org/document/get/493).

[38] L. Steen. *Data, shapes, symbols: Achieving balance in school mathematics.* In B. Madison, L. Steen, Editors, Quantitative literacy: Why numeracy matters for schools and colleges. Mathematics Association of America, Washington, DC., (2003), 53–74.

[39] G. Wiggins. *Get real! assessing for quantitative literacy.* In B. Madison, L. Steen, Editors, Quantitative literacy: Why numeracy matters for schools and colleges. Princeton, NJ, National Council on Education and the Disciplines, (2003), 121–143.

[40] R. Richardson, W. Mccallum. *The third R in literacy.* In B. Madison, L. Steen, Editors, Quantitative literacy: Why numeracy matters for schools and colleges. Mathematics Association of America, Washington, DC., (2003), 99–106.

[41] D. Krathwohl. *A revision of Bloom's Taxonomy: An overview.* Theory Into Practice, 41 (2002), No. 4, 212–218.

[42] H. Freudenthal. Weeding and sowing: Preface to a science of mathematics education. Dordrecht, Netherlands, 1980.

[43] K. Gravemeijer, J. Terwel. *Hans Freudenthal: a mathematician on didactics and curriculum theory.* J. Curriculum Studies, 32 (2000), No. 6, 777–796.

[44] R. Khattar, C.Wien. *Review of complexity and education: Inquiries into learning, teaching, and research* by B. Davis, D. Sumara, 2006. New York and London: Lawrence Erlbaum Associates. Complicity, 7 (2010), No. 2, 122–125.

[45] M. Andresen. *Teaching to reinforce the bonds between modelling and reflecting.* In M. Blomhoj, S. Carreira, Editors, Mathematical applications and modelling in the teaching and learning of mathematics. Proceedings from Topic Study Group 21 at the 11th International Congress on Mathematical Education in Monterrey, Mexico, July 6-13, 2008, (2009), 73–83. (Available at http://diggy.ruc.dk:8080/retrieve/14388#page=77).

[46] M. Andresen. *Modeling with the software 'Derive' to support a constructivist approach to teaching.* International Electronic Journal of Mathematics Education, 2 (2007), No. 1, 1–15.

[47] G. Gadanidis, V. Geiger. *A social perspective on technology-enhanced mathematical learning: from collaboration to performance.* ZDM, 42 (2010), No. 1, 91–104.

[48] L. Doorman, K. Gravemeijer. *Emergent modeling: discrete graphs to support the understanding of change and velocity.* ZDM, 41 (2009), No. 1/2.

[49] K. Gravemeijer, M. Doorman. *Context problems in realistic mathematics education: A calculus course as an example.* Educational Studies in Mathematics, 39 (1999), 111–129.

[50] K. Gravemeijer, M. Stephan. *Emergent models as an instructional design heuristic.* In Gravemeijer *et al.*, (2002), 145–169.

[51] K. Gravemeijer, R. Lehrer, L. Verschaffel, B. Van Oers (Eds.). Symbolizing, modeling, and tool use in mathematics education. Dordrecht, Netherlands, Kluwer, 2002.

[52] D. Kondrashov. *Using normal modes analysis in teaching mathematical modeling to biology students.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 278–294.

[53] J. Ellis-Monaghan, G. Pangborn. *Using DNA self-assembly design strategies to motivate graph theory concepts.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 96–107.

[54] S. Robic, J. Jungck. *Unraveling the tangled complexity of DNA: Combining mathematical modeling and experimental biology to understand replication, recombination and repair.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 108–135.

[55] R. Kerner. *Self-assembly of icosahedral viral capsids: the combinatorial analysis approach.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 136–158.

[56] R. Robeva, B. Kirkwood, R. Davies. *Boolean biology: Introducing boolean networks and finite dynamical systems models to biology and mathematics courses.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 39–60.

[57] J. Gill, K. Shaw, B. Rountree, Ca. Kehl, H. Chiel. *Simulating kinetic processes in time and space on a lattice.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 159–197.

[58] J. Milton, A. Radunskaya, W. Ou, T. Ohira. *A team approach to undergraduate research in biomathematics: Balance control.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 260–277.

[59] M. Cozzens. *Food webs, competition graphs, and habitat formation.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 22–38.

[60] G. Hartvigsen. *Using R to build and assess network models in biology.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 61–75.

[61] J. Knisley. *Compartmental models of migratory dynamics.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 245–259.

[62] Y. Grossman, A. Berdanier, M. Custic, L. Feeley, S. Peake, A. Saenz, K. Sitton. *Integrating photosynthesis, respiration, biomass partitioning, and plant growth: Developing a Microsoft Excel-based simulation model of Wisconsin Fast Plants (Brassica rapa, Brassicaceae) growth with students.* Math. Model. Nat. Phenom., 6 (2011), No. 6, 295–313.