COURSE SYLLABUS

BIOLOGY 450/550, Sec. M04 Biological Modelling Fall 2007

Instructor: Vince Gutschick, Prof. of Biology
Room 103, Biology Annex Phone 6-5661 (office) or 571-2269 (cell)

Class meets for the first session in Foster Hall, room 481, 4:15-5:05, Wednesday, 22 August. We will then find the best time and place for everyone for the succeeding meetings.

Welcome to the fourth and final offering of the BIOL 550 course on biological modelling. This is a graduate course intended to give you a familiarity with a broad range of biological models and to give you the experience of devising your own models. In short, I would like you to know how to make and use models. Similarly, I would like you to gain a critical view, so that you can determine if other researchers are using models appropriately and what these models offer as insight. Let me first outline the intellectual basis of the course. At the end I will indicate what work will be required and how grades will be assigned.

THE SCOPE OF THE COURSE

The exact scope of models to be surveyed will be set by all of us in consultation. With my help and that of your classmates, we will show how models can be posed conceptually, then posed mathematically, and finally solved either analytically or numerically. We will take some models to complete solution, including models that must be solved on the computer in an appropriate computer language.

The conceptual effort in modelling, as distinct from the mathematical effort

Your level of mathematical expertise should not be a barrier to successful development and solution of models. In biology, as in all sciences, we all use models all the time, at any level from verbal hypotheses to more quantitative statements and to full multi-process models. I hope to demonstrate that there are discrete steps in making and using models, in which much help is available and much reward can be had:

(a) Formulating the concepts. We may use verbal or mathematical statements, while in the end we will end up with mathematical statements. This is a hard step, but very enlightening. We are forced to say exactly what we know about the system, or hope to know. Inconsistencies are apparent once we have a mathematical statement, and they may then be rectified.

(b) Setting up a mathematical scheme for solving the model. This may mean setting up a differential equation, for example. We can do this on biological knowledge, primarily - that is, we can write down the descriptions of processes, such as how ions and water might move into plant or animal cells, where the boundaries lie, etc.

(c) Solving the mathematical scheme. Here, we may find the limits of our mathematical competence. Do we know how to solve differential equations of a given order? How about linear algebra, or other topics? Fortunately, there are several stages of help available:

(i) Our own math expertise.

(ii) That of our colleagues and classmates. Here I can help you, drawing on my background in mathematical physics. Your colleagues may be able to help, too. We may all learn together.

(iii) Literature examples of models of similar structure (not always easy to find, however).

(iv) Programming languages and packages, which often include packages to solve specific kinds of mathematical problems. For example, I rely on the Numerical Algorithms Group Fortran Library to provide algorithms to solve differential equations accurately and free of errors in setup. Many people rely on inclusive modelling languages, such as Stella or Mathematica. I hope we can share expertise beyond my own, the expertise that some of you have in specific computer languages.

Remember a basic principle of modelling in the computer age: computing power is almost never a problem. You may know only an inefficient or mathematically "inelegant" way to solve a problem, but if it works, use it. In fact, straightforward, slow math is often easier to comprehend than an elegant computational scheme (I can show you examples of both in my models). This means easier for both you and your colleagues to understand! In short, there are two kinds of complexity: mathematical (or conceptual), which we always strive to minimize, and computational. Never worry about the computational complexity, unless you attack such a large problem that your computer will take a month to give you your answer. Elegant math can be learned, if needed, or else borrowed.

What kinds of models will we look at?

We will choose models of interest to the group and some models specifically of interest to you individually. I hope you may become confident enough to work out your own models for research without first looking at the literature for guidance. The Nobel laureate physicist, Richard Feynman, always tried his hand before looking at what other people did; after all, it was the experts who led us into whatever impasse we find we must overcome now in our understanding.

We will focus on process models, which use known laws of system behavior (recombination rules for genetics; hydrodynamics and thermodynamics for water flow; etc.), as opposed to statistical models, which assume a rather nebulous (and often misleading) knowledge of how the system works. Classical examples of statistical models are linear (or nonlinear) regression models. The empirical coefficients mean nothing directly in terms of processes. Consequently, the models cannot be applied to new species, new geographic locations, new climates, or whatever, without a new fitting procedure. Clearly, this cannot be done for future climates, so that statistical models have a restricted utility. They are very useful for day-to-day analysis of data with a known underlying structure, but they cannot be used to extrapolate to new conditions. We will briefly discuss how we may introduce significant bias into our interpretation by our choice of model (linear additive vs. multiplicative, for example).

A few topics assume importance in process models. Commonly, systems have some conservation laws, as well as "constitutive" laws. For example, mass must be conserved in water flow, so that a powerful check on our models of water flow in cells or in plants or wherever is that we end up at every time with as much water as we started with.

GOALS

Let me state specific goals for the course:

(1) To illustrate how models can be used for any of several major purposes:

(a) Prediction of biological behavior (plant, animal, coupled system of biota with abiotic environmental processes, etc.). This is an ultimate objective that many of us set for our understanding, though it is rare that we know all the biological processes so well that we can make accurate predictions. Some predictive models have modest goals, such as models of light penetration to ground level through a canopy gap, based on hemispherical photographs and a knowledge of the sun’s daily and seasonal courses. Many predictive models are much more ambitious, hoping to predict the outcome of many coupled processes. At great effort, such complicated predictive models can be made to function accurately, of course. One example is GOSSYM, for cotton growth, yield, and quality under various environments and management practices. Another example is the set of models that predict macromolecular structure, based on thermodynamic rules derived nearly from first principles of chemistry. Yet another example is provided by pharmacokinetic models of drug delivery and action within a body, or related models of rational drug design that are based, for example, on substrate (drug) - protein interactions. Going to a large spatial and temporal scale, there are models for vegetation as a carbon sink and transpiration source in regional and global climate.

All such predictive models have limits of reliability, at which point they may inform us (by our careful analysis) of what processes we have described poorly or not at all. Then, the formerly predictive models become tools for seeking new hypotheses, as described in part (c) below.

Prediction may take several forms. Commonly, we think of prediction of the "state variables" of the problem - for example, what is the final population genetic structure in a given environment, or what is a crop yield. There are complementary ways to use a model, in addition. For example, one may ask under what range of environmental selection pressures we may expect two alleles to coexist indefinitely, rather than having one allele (or whole genotype) go extinct. The answer may also depend on the population genetic structure - how many alleles there are, how many loci exist, etc.

(b) Synthesis of our knowledge of separate processes, to seek and explain "emergent properties." Emergent properties are those properties not apparent in the individual parts. Phase transitions are the strongest example of emergent properties. On a less biological level, one may say that the ability of an airplane to fly, when none of its parts can fly, is an emergent property. On a more purely biological scale, we may ask if relatively simple combinations of rate processes in bulk solution and on membranes can generate a "biological clock," with properties of entrainment and ability to be reset in phase. Timing might be considered an inherent property of rate processes, but that self-perpetuating cycles and their resettability, etc. are emergent properties. What other surprises are in store? Can a visual processing system containing only about 20 neurons actually direct complex behavior in insect flight, including obstacle avoidance, prey recognition, and capture?

(c) Generation of experimentally testable hypotheses. This is by far the most common and productive use of biological models. Most of my published models fall into this category. One example is my assembly of a model of how whole-plant, whole-season yield and water-use efficiency (WUE) of alfalfa genotypes are controlled by mass per leaf area and leaf-internal setpoint for CO_2 concentration, or C_i. The model was used to ask if WUE could be substantially increased over that of present-day genotypes, and, if so, at what "cost" in yield. Surprising results on merged from the model, concerning constraints to improving WUE. The model was also field tested, with encouraging results. A second example (J. Exp. Bot. 44(1993):41-51, and 48: in press) is my model of how plant responses to low nutrient availability (increased root growth, increased kinetic capacity for uptake, changes in tissue nutrient concentration) enable plants to maintain relatively high growth rates under such stress. The model indicated that tissue nutrient content is likely to be a feature set passively by the balance between nutrient uptake and the nutrient's utility in photosynthesis. Furthermore, the contribution of each adaptive response to enabling high growth rate was quantifiable, with constraints also becoming apparent.

(d) Inverse use: taking the observed results, plus the driving variables, and finding out what structure the system must have to produce this. As an example, consider a stand of plants, with leaves placed in a variety of positions and at various angles, with a yet-unknown leaf reflectance. Sunlight is incident on the stand, including diffuse skylight with a specific angular distribution. Light is reflected at various angles, and it also penetrates probabilistically down to various depths in the canopy. A *direct* model takes the incident light and the canopy structure to predict the reflected or transmitted light. An *inverse* model takes the incident light and the reflected light distributions, and attempts to determine what canopy structure produced the result. Only a few cases are straightforward - for example, if light penetration probability P drops as an exponential with depth L, $P = e^{-KL}$, then we can determine L as $-(1/K) \ln P$ (if we also know the extinction coefficient K). Most inversion attempts are much more complicated. Because canopy-reflectance models are almost always numerical, there is no simple mathematical way to solve this problem. One ends up searching parameter space, in an enlightened fashion. Important questions remain, such as, Is the solution unique?

A special topic that a few of you might be interested in is neural networks, which might be described as finding patterns without process knowledge. Neural networks use the results of many repeated runs with varying initial conditions to become trained to recognize structural features. That is, one runs a canopy reflectance model 10,000 times with a variety of specified canopy structure and lighting conditions. The neural network tries to use the reflectance results to guess at the canopy structure; when it errs, it is corrected and randomly readjusts neural connections to try to improve its prediction success. We will also see neural networks in connection with optimization problems, in section (5) below. Neural networks are of restricted utility in biological modelling, but you may find one profitable to use later, so we will note them.

(2) To show the scope of biological phenomena to which models can be applied successfully and profitably. These phenomena range in scale from macromolecular behavior, through the levels of a cell (say, ion channel action), a tissue (such as skin viscoelastic properties), an organ (root water transport based on ion transport and irreversible thermodynamic driving forces), an individual (growth, cardiovascular development of an animal embryo), a population (recruitment dynamics of marine invertebrates), a community (species competition and the existence of invasion-proof strategies of inter-species competition), and ecosystem behavior (succession, trophic structure), and finally to the globe (climate as dependent on vegetation growth and gas exchange, or vice-versa). The models range from pure structure (assembly rules for membranes or ecosystems) to structure-function relations (consider the examples above). The structure and function need not be directly physico-chemical. For example, there are models of evolution in populations, based on processes at higher levels (genes, environmental selection, etc.). Animal behavior is modelled at a very high level, or, we may say, a very derived level, so as not to impute value to the scale.

In addition to models of how systems evolve with time or hit stable structures, we might consider game-theory models of competition as special cases. In these models, one considers two or more players (competing individuals, or an individual in a varying environment), with cost and benefits to various plastic responses. The set of responses deployed by an organism is entitled a strategy. Game theory computes the net benefit to each player when they interact. It can be used to decide if coexistence is possible, with fixed strategies. It can also be used to determine the optimal strategy for each player, and this strategy may be one that varies randomly with time, following a plan to deploy different strategies with different frequencies.

(3) Similarly, to show the scope of ways that models can be made. There are simple "back-of-the-envelope" models for how gene frequencies stabilize without (or even with) selection pressure, such as the Hardy-Weinberg law. There are models of relatively few coupled processes, such as the Farquhar- von Caemmerer- Berry model of photosynthesis in C_3 plants (which, for its relative simplicity, synthesized a great deal of knowledge and

showed common patterns in responses of many plant species to many environmental conditions). There are "monster models" of very many processes or inter-related structures, such as crop models or the more elaborate predator-prey models. These have very restrictive usefulness and are *not* the level for which we aim in general.

Part of the range of models is in their mathematical structure, which reflects their conceptual structure. For example, there are process models for time-dependent changes, vs. structural models for stable structures or for how structure emerges from simple rules. Process models may be discrete in structure (age classes in a population) or continuous (solute transport in a fluid). In continuous models, we typically get differential equations. These may be first-order or higher in the time derivative involved. Our knowledge may be of initial conditions or, alternatively, of boundary conditions (two points at the ends, such as our knowledge of sound wave propagation in the cochlea of the ear). The concepts set the model structure, ultimately. We may not want to solve a differential equation, so we might discretize a problem, but we must learn to recognize what structure is demanded by a given problem.

One important distinction is between deterministic models and stochastic models. In a deterministic model, there is a fixed structure to the system or a known course (such as blood-vessel branching pattern and elastic properties, even if they develop and age, if this occurs by known rules). There is, as well, a unique state of the driving variables (temperature of embryo development, nutritional supply, etc.). The end results are then unique. In stochastic models, either the system structure or the driving variables have inherently random variations. We are then modelling not a unique state, but a statistical distribution of states. Quite different mathematical techniques, and interpretations, are demanded. Some examples are relatively simple - computing the probability of light penetration into a canopy, or the mean time to extinction of an allele in a population undergoing genetic drift - and can be formulated explicitly, or "analytically." Others are more complex and require sophisticated numerical solutions.

There is a meeting ground between determinism and stochasticity, namely, chaos. This is a hot topic today. Chaos is behavior that becomes highly unpredictable at long times for small differences in starting conditions. Weather is a classical example. If we knew to very high precision the state of motion of all parcels of air and ocean, etc. at one time, we would predict a certain state two weeks later. If we had a very slightly different initial state, we might predict a very different state two weeks later. Chaos arises from nonlinear processes only. In linear systems, rates of change of any variable, say, x_1, depend only on linear combinations of variables x_1, x_2,... and not on x_1^2 or x_1x_2. These systems have an unvarying periodic behavior, so that future states are predictable from combinations of "eigenstates" with precisely known behavior in time. This brings up an important point about mathematical structure, namely, that one often "linearizes" models of processes for simplicity of solving them and because one knows that only a modest error is introduced. However, nonlinearity that is more significant can be the basis of chaotic behavior that is real - in weather, in opening of plant stomata, etc. We must be aware that we are not biasing our predictions to get "nice" answers without chaos. Note that, even with chaos, the long-term average patterns may be highly predictable, such as storm tracks, while the individual run may be wholly unpredictable.

One field of possible interest is self-organizing systems, wherein orderly structures arise from initially random initial conditions, such as membrane vesicles forming from dispersed lipids. Another example is oscillatory patterns in time and space in some chemical reactions (taken as crude examples of biological clocks, none of which has yet been identified in real organisms). Self-organization involves nonlinearity of interactions. I am no expert here but can guide you to the literature and provide some discussion if this is an interest of yours.

We are unlikely to get into the subject of controllability of a system, which is typically an engineering problem. However, the topic does apply in some biological systems, such as biomechanics of animal movement or neural control of hormone systems. It is strongly related to nonlinearity, as well as to time lags - both can induce uncontrollable behavior - either truly chaotic, or not within designed limits of departure from a desired setpoint. If you are interested, I can direct you to some readings.

(4) To give you some guiding rules, and even more intuition, about how one should balance empiricism with process details. Consider how one might describe stomatal conductance in plant leaves. We know that it responds to light, temperature, plant water status, etc. Should we make conductance equal to a function of light, times a function of temperature, times a function of humidity, etc.? One would then try to measure many parameters for each response; indeed, many people have tried this. However, is this the way that light, temperature, etc. really interact? What about additive models, or inbetween? Let me cut to the answer: to a very close approximation, conductance is proportional to photosynthetic rate (it is paced to photosynthetic rate, by sensors in leaf cells that respond to the drawdown of CO_2 concentration), multiplied by the relative humidity at the leaf surface. If one attempts to formulate conductance as a response to light times a response to temperature (as it varies over a day, for example), then one will confuse two effects of temperature: its activation of photosynthesis, and its effect on relative humidity. One will get a complicated model of limited accuracy. If, in contrast, one pays attention to what is known about processes (conductance is proportional to photosynthesis over a wide range of conditions, and it responds to transpiration rate in a negative feedback loop), then one can get a model with rather few parameters that will be accurate to a high degree. Shall we go further, to show how the apparent response to humidity derives from a response to transpiration rate and a balancing ability of the plant to conduct water to support transpiration? One will gain very little in accuracy of describing conductance, while adding many parameters needing to be measured. The only reason for the great effort is to prove that the two balancing processes do explain conductance, which is a separate inquiry from making and using models to predict conductance under a wide variety of environments, say, to predict vegetation effects on climate processes.

A broader topic about complexity in models is computability. In most process models, this is not a subtle issue - if the model is a differential equation, there is a known scaling of computational effort with the size of the problem (number of variables, number of time steps, etc.). With some models, especially optimization models or inverse models, the issue of computability can arise. That is, the time or effort needed to solve a problem rises faster than the number of items (variables, game players, etc.) raised to a simple power.

(5) To show you the power - and the limitations - of optimizing models. I have already mentioned that I have made a model of how water-use efficiency and yield (or growth rate) is controlled by two plant traits. We might postulate economic values (or Darwinian fitness values) to growth and yield, to get a single "objective function" that we seek to maximize (or minimize). We seek to get the maximum or minimum by choosing the best combination of system parameters, such as the two plant traits noted in the example earlier. All of the cost-benefit models of plant or animal responses, management practices, etc. are in this class of optimizing models.

Optimization models stand in contrast to simply descriptive models, which describe how the system functions *as presently constituted*. Optimization models can guide us to economic benefit, obviously. They can also tell us if organisms have evolved to an adaptive endpoint. Conversely, when organisms (or populations, etc.) fail to meet our optimization criteria, this tells us that we are overlooking an important selection pressure that acts on a separate measure of fitness (say, mineral nutrition success in a plant, while we were paying attention originally to water use alone). Failure to be at an optimum phenotype may also indicate that selection has not had enough time to act to change the genotype and phenotype distribution. One example is plants not responding adaptively (optimally) to elevated CO_2 concentrations in the air, rising over only a few generation times, esp. for trees! An example of competing optimizations is that of plants having more chlorophyll and thicker leaves than will give maximal photosynthesis and growth in monospecific stands - they trade off some growth against the ability to shade competitors, it appears. One of my mild obsessions is the opposite of optimization problems: if an optimum (say, in plant water relations, combined with mineral nutrition) is so obvious, why aren't all plants of one genotype? Why is there persistent genetic variability in most genetic loci in most organisms? The challenge is to find time-varying selection, which favors one genotype, then another, or frequency-dependent selection, or whatever. We often look at average conditions and neglect the rare occurrence of extreme conditions that are critical in molding a population. I will show a few examples, as in plants surviving drought stress.

Optimizing models require a special conceptual set-up. For example, we must know what variables (genetic traits, e. g.) can be adjusted, and whether or not any two variables can be adjusted independently of each other (vs. being constrained to co-vary). We learn a lot by thinking about which possibilities are allowed. Optimizing models also can call for special methods of mathematical solution. In simple cases, one can set up the condition that the derivative of the objective function with respect to each trait is zero (for continuous traits). The math is well defined. In other cases, the set of equations becomes so complicated that the optimum must be sought by "brute-force" searches in the "space" of parameter values. The computation time easily becomes insupportable, so that one must use some innovative schemes to get approximate optima. Among the methods of solution are the use of neural networks and the use of "genetic algorithms" (programs that compete with each other to specify the state of a system; they replicate with more or less success, according to which algorithms are closer to an optimal solution. They have other uses, too).

(6) To show some other aspects of modelling that are not readily categorized in the above 5 items. For example, there is fuzzy-systems theory, to use when we have a knowledge of the system that is more qualitative than quantitative, but which nonetheless can yield firm decisions (as in control actions).

(7) To instill in your work what we might call a "5-D" approach to making and using models:

> **D***efine* the state variables and the parameters clearly, carefully. What entity controls which other entity? This is usually straightforward, as in deciding to model how leaf photosynthesis is controlled by leaf temperature. It gets complicated when there are feedbacks. For example, photosynthesis then determines stomatal conductance, which determines transpirational cooling rate and thus the leaf temperature. The real driving variables are then air temperature, solar radiation, and humidity. Feedback loops can require some thought and some attention to stable mathematical solution, but there is always a way to handle them.
>
> Do a **D***imensional analysis*. When you write an equation, the units had better be the same among all terms, and on both sides of the equation. If one term is in units of per day (such as relative growth rate) but another term is in grams per day, then you have made a mistake - either typographically, or in neglecting a process. Dimensional analysis is a powerful way to check all models.
>
> **D***ebug* the formulas and the computer program. Make sure that output is computed properly numerically, by performing hand-calculations. This is something of an art, and no one ever finds all the errors in a big program, but one can become confident that the unfound errors occur only in rare conditions. Even the initial equations can be debugged. Is the predicted behavior reasonable? For example, one may look at a formula and ask if a plant‘s relative growth rate increases with its nutrient content. If it doesn’t, then you probably have an error.
>
> This brings up a strong opinion of mine, about model "validation." For large, numerical models, one cannot check all the intermediate states. For example, an ecosystem productivity model might involve soil water balance, plant phenology, air and soil temperatures, etc. at many spatial locations and times. In using the model, one commonly observes only some coarse weather conditions and final biomasses, not the detailed course of soil water, etc. One then hopes that the agreement of the final results alone, between the model and the real system, is a check of how accurate the model is. This is not true, and the situation is to be avoided if at all possible. Never make a model bigger than you can afford to check in detail against reality. However, climate modellers and others are in this bind. One can at least guard against a biased attitude. The very term "validation" implies that you are going to prove your model is correct. A more fruitful attitude is to check, without bias, if your model is correct, or even to try to show how and when your model is wrong!
>
> **D***ocument* the model and the computer program. I guarantee that one year after you write a computer program, you will have great difficultly even remembering

what it does overall, much less what each calculation inside it is for, if you do not have very many descriptive comments in the code (as many as several comment lines per line of code) and/or in a narrative text.

D*isplay* the results effectively. You may like graphs, or you may like columns of numbers, but you need to find the most effective selection of results or combinations of results to display. This does not mean you need to learn fancy three-dimensional plotting techniques. There is a good book by Tufte, called *The Visual Display of Quantitative Information*, which is very informative.

WORK WE WILL PERFORM IN THE COURSE, AND METHOD OF GRADING

This is the first time I have taught this course, which presents an unusual blend of lectures, group discussions, math, and computer usage. I propose the following scheme, but we may modify it if we all agree to a new scheme.

I will lecture on a variety of topics, to provide a background in biological concepts and in the math. For each topic, each of you will then look through some of the primary research literature, which is mostly in journals. Most often, you will try a literature search on a database, such as Science Citation Index or Agricola on-line. We will meet to find out what specific topics under the broader topic are most interesting (say, an article on blood-flow control in tuna, under the broader topic of animal physiology). One person (or perhaps two people) will volunteer to (a) lead a group discussion of the paper and (b) offer a conceptual model that condenses or extends the ideas of the paper, which the group will then work out to the stage of a mathematical model. The person(s) who did the original presentation will then develop a computer model and solution, in consultation with me and with anyone else who wishes to contribute; this work will be done out of class, typically. By the end of the semester, everyone will have volunteered to lead such an effort at least once. Your grade will be assigned based on:

(1) Your effort in leading a discussion and modelling session (50%). This will include:

(a) The effectiveness of your discussion: did you give adequate depth, with good organization, and with reasonable accuracy in presenting the ideas? (25%)

(b) The computer model: did you give good conceptual input to the model development? To the extent of your ability with math and computing, did you contribute well to completing the math and computing development? (25%)

(2) Your participation in class discussions and developing the concepts of models that others are presenting. (50%) The criteria are as in part (1).

The usual diligence in a graduate course will get you a high grade, as in other graduate courses. This course is novel, and it should be fun and stimulating, with a realistic level of effort on your part.

SOME OTHER MECHANICS OF THE COURSE

Office hours: With all of our research schedules being so varied, we should arrange office hours once we all convene in class. Hours during two afternoons a week are likely.

Attendance: Of course, attending every class is desirable, for your contribution to the class and my ability to grade your performance. If you have to miss a class, please tell me in advance or as soon as practical. We can work out an alternative way for you to make your contribution to the classwork.

Schedule of topics: We will develop this in class, to meet the needs of all the people in the class. We will balance topics of immediate interest and applicability with the breadth of topics that you should gain familiarity with.

Mutual help: Each of you will be good at particular topics in biology, math, and computing. Some of you might be willing to help your classmates on particular topics. For those of you who wish to be available, I will circulate your names and contact information, with a list of topics on which you are willing to be consulted. This is purely voluntary; if you don't want to do it, it won't adversely affect your grade. Also, if the burden gets to be more than you want, we can easily inform people that you will have to limit your availability.

COMPUTING SKILLS AND ACCESS

Each of you can, and should, get a free-of-charge computer account from the campus computing organization, ICT. I will find out how the system works this year. If you choose to use the UNIX systems, including my own machine, wombat, I can give you tutorials on basic commands (the shell) and on compilers. You can also do much of the work on your own PC or Mac. Many problems can be programmed in Excel or the OpenOffice equivalent, but eventually you will find this to be limiting. There are free or cheap compilers for C, C++, and (with more digging), Fortran.

It would be nice if you had experience in a programming language, such as Fortran, BASIC, Pascal, C, C++, Stella, CSMP, Stella, COBOL, etc. If you have none, we will work through it; I can teach you to do simple programs and help you to write larger ones, and this won't affect your grade; your willingness to learn is what counts.

TEXTS

No single text provides a good introduction to the scope of biological modelling. Some books that title themselves such actually have only a limited repertoire of techniques, such as stock-and-flow models, vs., say, differential equations, linear algebra, root-finding, etc. We will thus refer to a number of books and to the original research literature. Among the books you may find helpful are the following (where * indicates a more general book; the others can be hard to read without a good background in the field). *This is a long list;* I don't expect you to read even a modest fraction of these. It is also an old list, from 1997. Only one potent modelling book has come to my attention since, Gershenfeld's *The Nature of Mathematical Modelling* (Cambridge, 1999)...but it assumes a heady mathematical background. We might refer to it for the techniques, which we can then search for background info on (even Wikipedia works for many cases, as does the sosmath.com (once you know which math techniques you want to use!). The list is just here for your guidance if you want to pursue a topic. I haven't put these books on reserve, as there are too many to reserve for a relatively small class. I don't anticipate that you'll have any problems getting the book you want. If you have problems and the library can't help, please let me know; I may be able to get the book or I may have it myself.

In plant physiology and ecology:

* Gaylon S. **Campbell**. 1977. *An Introduction to Environmental Biophysics.* Springer, New York ff. vii+159 pp. ISBN 0-387-90228-7 (PAPER). Library call no. QH505 C34

* Thomas J. **Givnish** (ed.). 1986. *On the Economy of Plant Form and Function.* Cambridge U. Press. xvii+717 pp. ISBN 0-521-26296-8 (I have a copy, library doesn't)

* Vincent P. **Gutschick**. 1987. *A Functional Biology of Crop Plants.* Croom Helm, London and Sydney. x+230 pp. ISBN 0-7099-3819-5. Library call no. SB91 G87 1987

* Ian R. **Johnson** and John H. M. **Thornley**. 1990. *Plant and Crop Modelling: A Mathematical Approach to Plant and Crop Physiology*. Clarendon Press, Oxford. Library call no. QK711.2 T483 1990 (Naveen)

* Park S. **Nobel**. 1991. *Physicochemcial and Environmental Plant Physiology*. Academic, San Diego. xx + 635 pp. ISBN 0-12-520021-8. Library call no. QK711.2 N623 1991

In animal physiology and ecology:

* Peter **Calow** (ed.). 1987. *Evolutionary Physiological Ecology.* Cambridge U. Press. ix+239 pp. ISBN 0-521-32058-5. Library call no. QH541 E87 1987

In biomechanics (mostly animal):

Susan J. **Hall**. 1991. *Basic Biomechanics*. Mosby Year Book, St. Louis. Library call no. QP303 H35 1991

In cellular physiology:

Wilfred D. **Stein**. 1986. *Transport and Diffusion across Cell Membranes.* Academic Press, San Diego ff. xvii+685 pp. ISBN 0-12-664661-9 (PAPER). Library call no. QH509 S74 1986

* Wilfred D. **Stein**. 1990. *Channels, Carriers, and Pumps: An Introduction to Membrane Tranport*. Academic, San Diego. Library call no. QH509 S73 1990

In genetics:

* Bruce **Wallace**. 1981. *Basic Population Genetics*. Columbia University Press, New York. xii + 688 pp. Library call no. QH455 W34

In math, generally:

* Edward **Batschelet**. 1979. *Introduction to Mathematics for Life Scientists*. Springer, Berlin. xv + 643 pp. Library call no. QH323.5 B37 1979b

* Leah **Edelstein-Keshet**. 1988. *Mathematical Models in Biology.* Random House, New York. xvi + 583 pp. ISBN 0-394-35507-5 (Michael Cain has a copy)

Frank C. **Hoppensteadt** and Charles S. **Peskin**. 1992. *Mathematics in Medicine and the Life Sciences.* Springer, New york. xii + 252 pp. ISBN 0-387-97659-6 (Steve Warburton has a copy)

Jon **Mathews** and R. L. **Walker**. 1970. *Mathematical Methods of Physics*, 2nd Ed. W. A. Benjamin, New York. xiii + 501 pp. Library call no. QA401 M42 1970

Sol I. **Rubinow**. 1975. Introduction to Mathematical Biology. Wiley, New York. xiii + 386 pp. Library call no. QH323.5 R8

Douglass J. **Wilde**. 1964. *Optimum Seeking Methods.* Prentice-Hall, Englewood Cliffs, NJ. xii+202 pp. LC 63-20039 (I have a copy, library doesn't)

In game theory:

* Karl **Sigmund**. 1993. *Games of Life: Explorations in Ecology, Evolut9ion, and Behaviour.* Oxford University Press, Oxford. Library call no. QH313 S585 1993

In stochastic modelling:

Roe **Goodman**. 1988. *Introduction to Stochastic Models*. Benjamin/Cummings, Menlo Park, CA. Library call no. QA273 G655 1988

In chaos theory:

* Friedrich **Cramer**. 1993. Chaos and Order: The Complex Structure of Living Systems. Wienheim, New York. Library call no. Q172.5 C45 C713 1993

James B. **Gleick**. 1987. *Chaos: Making a New Science*. Viking, New York. Library call no. Q172.5 C45 G54 1987

In self-organizing systems in biology:

Stuart B. **Kauffman**. 1993. *The Origins of Order: Self Organization and Selection in Evolution*. Oxford University Press, New York. Library call no. QH325 K39 1993

See also Hoppensteadt and Peskin, above

In inverse modelling:

Dilip N. **Ghosh Roy**. 1991. *Methods of Inverse Problems in Physics*. CRC Press, Boca Raton, FL. Library call no. QC20.7 D5 G48 1991

In control theory:

Tom **Milhorn**. 1970?. *The Application of Control Theory to Physiological Systems.* Saunders, Philadelphia. (Steve Warburton has a copy)

Douglas S. **Riggs**. 1970. Control Theory and Physiological Feedback Mechanisms. Williams and Wilkins, Baltimore. Library call no. QH508 R5

In computing and in display of results:

* Merlin L. **James**. 1985. *Applied Numerical Methods for Digital Computation.* Harper & Row, New York. Library call no. QA297 J3 1985 (*For CSMP language*)

* Michael **Metcalf**. 1985. *Effective Fortran 77.* Clarendon Press, Oxford. ISBN 0-19-853709-3 (I have a copy, library doesn't)

* Robert J. **Traister**. Programming in C. Prentice-Hall, Englewood Cliffs, NJ. ISBN 0-13-729641-X (I have a copy, library doesn't)

* E. R. **Tufte**. 1983. *The Visual Display of Quantitative Information.* Graphics Press,

Cheshire, CT. 197pp. (no ISBN #) Library call no. QA90 T84 1983