Chapter Title: What Are Dynamic Models?

Book Title: Dynamic Models in Biology
Book Author(s): Stephen P. Ellner and John Guckenheimer
Published by: Princeton University Press

Stable URL: https://www.jstor.org/stable/j.ctvcm4h1q.6

1   What Are Dynamic Models?

Dynamic models are simplified representations of some real-world entity, in equations or computer code. They are intended to mimic some essential features of the study system while leaving out inessentials. The models are called dynamic because they describe how system properties change over time: a gene's expression level, the abundance of an endangered species, the mercury level in different organs within an individual, and so on. Mathematics, computation, and computer simulation are used to analyze models, producing results that say something about the study system. We study the process of formulating models, analyzing them, and drawing conclusions. The construction of successful models is constrained by what we can measure, either to estimate parameters that are part of the model formulation or to validate model predictions. Dynamic models are invaluable because they allow us to examine relationships that could not be sorted out by purely experimental methods, and to make forecasts that cannot be made strictly by extrapolating from data.

Dynamic models in biology are diverse in several different ways, including

- the area of biology being investigated (cellular physiology, disease prevalence, extinction of endangered species, and so on),
- the mathematical setting of the model (continuous or discrete time and model variables, finite- or infinite-dimensional model variables, deterministic or stochastic models, and so on),
- methods for studying the model (mathematical analysis, computer simulation, parameter estimation and validation from data, and so on),
- the purpose of the model ("fundamental" science, medicine, environmental management, and so on).

Throughout this book we use a wide-ranging set of case studies to illustrate different aspects of models and modeling. In this introductory chapter we describe

and give examples of different types of models and their uses. We begin by describing the principles that we use to formulate dynamic models, and then give examples that illustrate the range of model types and applications. We do only a little bit of mathematics, some algebra showing how a model for enzyme kinetics can be replaced by a much simpler model which can be solved explicitly. We close with a general discussion of the final "dimension" listed above—the different purposes that dynamic models serve in biology. This chapter is intended to provide a basic conceptual framework for the study of dynamic models—what are they? how do they differ from other kinds of models? where do they come from?—and to flesh out the framework with some specific examples. Subsequent chapters are organized around important types of dynamic models that provide essential analytical tools, and examples of significant applications that use these methods. We explain just enough of the mathematics to understand and interpret how it is used to study the models, relying upon computers to do the heavy lifting of mathematical computations.

1.1 Descriptive versus Mechanistic Models

Salmon stocks in the Pacific Northwest have been in steady decline for several decades. In order to reverse this trend, we need to know what's causing it. Figure 1.1 shows data on salmon populations on the Thomson River in British Columbia (Bradford and Irvine 2000). Figure 1.2 presents the results from statistical analyses in which data from individual streams are used to examine how the rate of decline is affected by variables describing the impacts of humans on the surrounding habitat. Straight lines fitted to the data—called a linear regression model—provide a concise summary of the overall trends and quantify how strongly the rate of decline is affected by land use and road density. These are examples of *descriptive* models—a quantitative summary of the observed relationships among a set of measured variables.

Figure 1.2 provides very useful information, and descriptive models such as these are indispensable in biology, but it also has its limitations.

- It says nothing about *why* the variables are related the way they are. Based on the results we might decide that reducing road density would help the salmon, but maybe road density is just an indicator for something else that is the actual problem, such as fertilizer runoff from agriculture.
- We can only be confident that it applies to the river where the data come from. It might apply to other rivers in the same region, and even to other regions—but it might not. This is sometimes expressed as the "eleventh commandment for statisticians": Thou Shalt not Extrapolate Beyond the Range of Thy Data. The commandment is necessary because we often *want* to extrapolate beyond the range
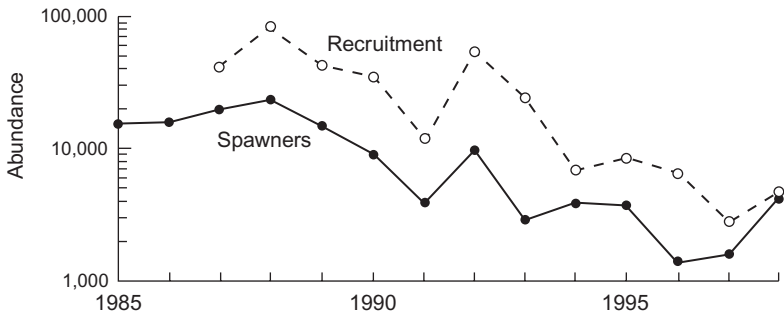
Figure 1.1 Decline in Coho salmon stocks on the Thomson River, BC (from Bradford and Irvine 2000).

(a)

$r = -0.50, P = 0.001$

Proportion of land used

(b)

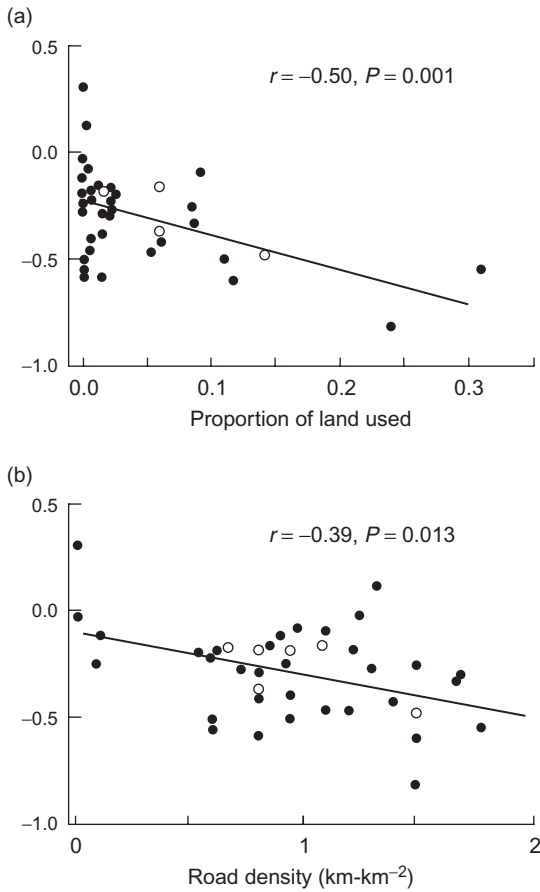$r = -0.39, P = 0.013$

Road density (km-km^{-2})

Figure 1.2 Rate of decline on individual streams related to habitat variables (from Bradford and Irvine 2000).

of our data in order to make useful predictions—for example, how salmon stocks might respond if road density were reduced.

The second limitation is related to the first. If we knew *why* the observed relationships hold in this particular location, we would have a basis for inferring where and when else they can be expected to hold.

In contrast, a *dynamic* model is mechanistic, meaning that it is built by explicitly considering the processes that produce our observations. Relationships between variables emerge from the model as the result of the underlying process. A dynamic model has two essential components:

- A short list of *state variables* that are taken to be sufficient for summarizing the properties of interest in the study system, and predicting how those properties will change over time. These are combined into a *state vector X* (a *vector* is an ordered list of numbers).
- The *dynamic equations*: a set of equations or rules specifying how the state variables change over time, as a function of the current and past values of the state variables.

A model's dynamic equations may also include a vector *E* of *exogenous variables* that describe the system's environment—attributes of the external world that change over time and affect the study system, but are not affected by it.

Because it is built up from the underlying causal processes, a dynamic model expands the Range of Thy Data to include any circumstances where the same processes can be presumed to operate. This is particularly important for projecting how a system will behave in the future. If there are long-term trends, the system may soon exceed the limits of current data. With a dynamic model, we still have a basis for predicting the long-term consequences of the processes currently operating in the study system.

1.2 Chinook Salmon

As a first example, here is a highly simplified dynamic model for the abundance of Chinook salmon stocks, based on some models involved in planning for conservation and management of stocks in the Columbia River basin (Oosterhout and Mundy 2001, Wilson 2003). Most fish return as adults to spawn in the stream where they were born; we will pretend for now that *all* of them do, so that we can model on a stream-by-stream basis. The state variable in the model for an individual stream is $S(t)$, the number of females spawning there in year t. The biology behind the simple models is as follows.

1. Fish return to spawn at age four or five, and then die. Of those that do return, roughly 28% return at age four, 72% at age five; this actually varies somewhat over time, but it turns out not to matter much for long-term population trends.
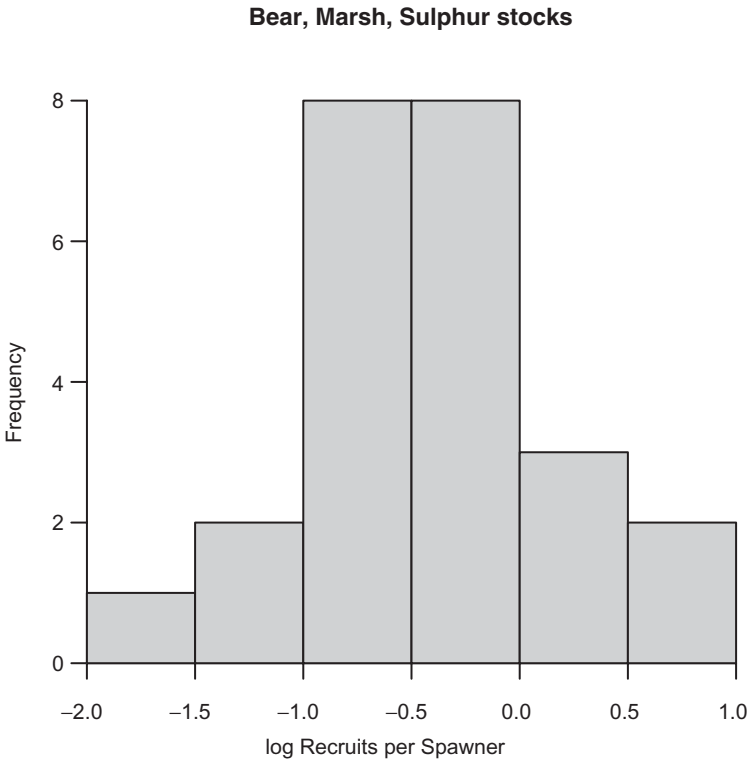
Bear, Marsh, Sulphur stocks

Figure 1.3 Frequency distribution of recruits per spawner for the Bear, Marsh, and Sulphur Creek stocks of spring/summer run Chinook salmon in "good" years.

2. Spawning success E is highly variable (Figure 1.3). In "bad years," E is essentially nil (about 0.05). About 18% of years (2/11 in the available data) are bad. In "good years," E can be large or small, and on average it is 0.72 (standard deviation 0.42). What makes years good versus bad may be related to El Niño (Levin et al. 2001).

For this simple illustration, we will ignore the fact that E varies over time and use instead a single "typical" value (in later chapters we discuss models that include variability over time). The median value of E is 0.56 and the average is 0.58, so for the moment let us take $E = 0.57$. The model is then

$$S(t) = Ep_4 S(t-4) + E(1-p_4)S(t-5)$$
$$= 0.16S(t-4) + 0.41S(t-5).$$

[1.1]

This is the dynamic equation that allows us to predict the future. Current spawners $S(t)$ come from two sources: those offspring of spawners four years ago who return at age four, and those offspring of spawners five years ago who return at age five. The model is nothing more than bookkeeping expressed in mathematical notation. So, to predict the number of spawners in 2002 from data in previous

years, we would use the dynamic equation with $t = 2002$:

$$S(2002) = 0.16S(1998) + 0.41S(1997),$$ [1.2]

and we can then continue on forward,

$$S(2003) = 0.16S(1999) + 0.41S(1998)$$

$$S(2004) = 0.16S(2000) + 0.41S(1999)$$

[1.3]

$$S(2005) = 0.16S(2001) + 0.41S(2000)$$

$$S(2006) = 0.16S(2002) + 0.41S(2001).$$

Notice that in the last line of [1.3], $S(2002)$ comes from the model. Having forecast a value for 2002, we can proceed to forecast beyond that time. Under the assumption that the processes represented by the dynamic equations will continue to hold in the future, we can legitimately project into the future based on conditions in the present.

For management purposes, this model is only a starting point—we also want to explore potential ways to improve the situation. For that we need a more detailed model that represents the separate stages of the salmon life cycle, because management actions have stage-specific effects—for example, protecting nests from predators, improving food availability when juveniles are growing, or modifying dams to decrease mortality during upstream and downstream migrations. By expanding the model to include relationships between management actions and stage-specific survival, it can then be used to predict which approaches are most likely to succeed (e.g., Wilson 2003; Ellner and Fieberg 2003). Such models and their use in conservation planning are a main topic of Chapter 2.

1.3 Bathtub Models

Our simple salmon model [1.1] was nothing more than bookkeeping—who were the parents of this year's spawners?—expressed in equations. The same is true of many models in biology, even complicated ones. A very important example is the *bathtub* with state variable $W(t)$ = amount of water in the tub (Figure 1.4). Bear with us, this is *really* true, and models with $W(t)$ = amount of protease inhibitor in the bloodstream are exactly the same in principle.

A dynamic equation for this model has to tell us how $W(t)$ changes over time. Time is a continuous variable in this case, but to derive the model we begin by considering a small interval of time from now (time t) until a little bit later (time $t + h$). We take h short enough that any changes in $I(t)$ and $O(t)$ over the time interval are small enough to be ignored. Then

$$W(t + h) = W(t) + \text{Inflow rate} \times \text{time elapsed} - \text{Outflow rate} \times \text{time elapsed}$$

[1.4]

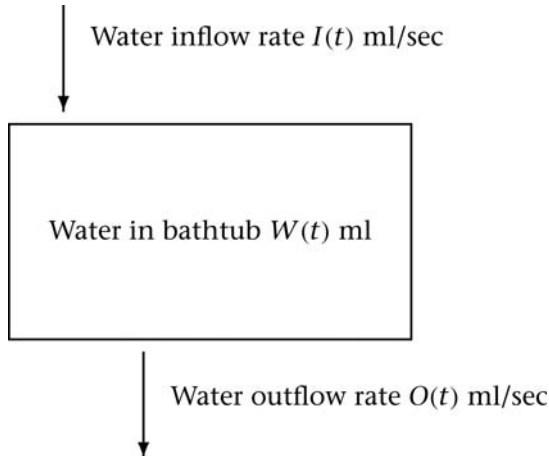$$= W(t) + I(t) \times h - O(t) \times h.$$

Figure 1.4 Compartment diagram of the bathtub model.
The rectangle denotes the state variable—the amount of
water in the tub—and the arrows denote flows.

We want to let $h \to 0$. To do that, we rearrange [1.4] into

$$\frac{W(t+h) - W(t)}{h} = I(t) - O(t). \qquad [1.5]$$

The left-hand side of [1.5] is a *difference quotient* for the derivative of W with respect to time. So we can now let $h \to 0$ and we get

$$\frac{dW}{dt} = I(t) - O(t). \qquad [1.6]$$

In words: Change in W = amount coming in − amount going out, with all amounts being *amounts per unit time*.

1.4 Many Bathtubs: Compartment Models

If we connect several bathtubs, so that the water (or whatever) flowing out of one tub can then flow into one or more other tubs, we get a *compartment model*. The state variables in a compartment model are the amount of some stuff in a number of distinct locations or categories ("compartments") within the system. Despite their simplicity, or perhaps because of it, compartment models are very widely used in biology.

The "stuff" in the model can be essentially anything. Sometimes it really is the amount of some material in a particular location—the level of lead in blood, liver, brain, and so on, or the amount of nitrogen in different layers of soil. It can also be amounts or numbers in a particular state or category: gene copies in the active versus inactive state; infected versus uninfected T-cells in the immune system; ion channels in the open versus closed state in a neuron; small, medium, and large sea turtles. The key assumption, as in our simple bathtub, is that items
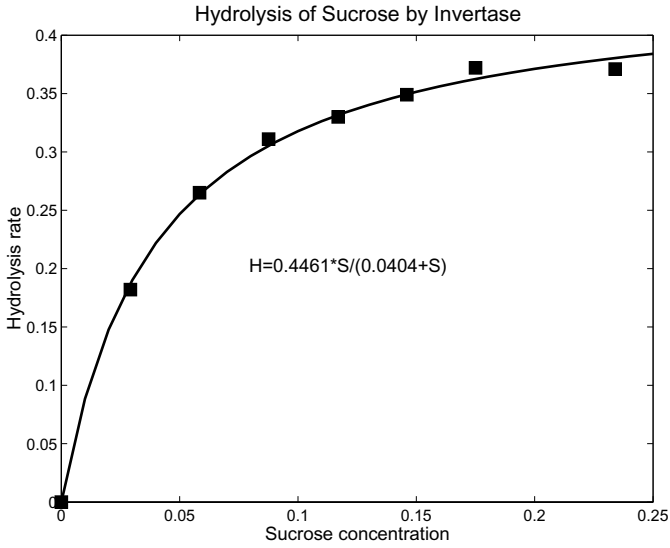
Figure 1.5 Data on the rate of hydrolysis of sucrose by invertase (from dickson 1998), and the fitted Michaelis-Menten equation with parameters estimated by least squares.

within a compartment are indistinguishable from each other—this is expressed by saying that compartments are "well mixed." Thus we only need to keep track of the quantity of material in each compartment—how much water is in the bathtub, but not the time when each molecule arrived or which other bathtub it came from.

1.4.1 Enzyme Kinetics

A simple example of a compartment model is an enzyme-mediated biochemical reaction. This is a classical example (Michaelis and Menten 1913). The reaction diagram, as chemists write it, is

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} SE, \quad SE \overset{k_2}{\rightarrow} P + E \tag{1.7}$$

where S = substrate, E = enzyme, and P = reaction product.

Figure 1.5 shows some data with sucrose as the substrate and invertase as the enzyme. Without a model we can partially understand what's going on. At low concentrations of sucrose the reaction is substrate limited: if we double the amount of sucrose, the hydrolysis rate doubles. At high concentrations of sucrose the reaction is enzyme limited: double the amount of sucrose and the hydrolysis rate stays the same. Assuming a smooth transition between these regimes, we can deduce the qualitative shape of the dependence—but nothing quantitative, or how the curve is determined by the underlying reaction rate constants.

To make a quantitative connection between process (the reactions) and pattern (the plotted data), we convert the reaction diagram into a compartment model. Apart from the reaction diagram, the only assumption is the *Law of Mass Action*: The rate of a chemical reaction is proportional to the product of the concentrations of the reactants. The k's in the diagram are the constants of proportionality.

To express these assumptions as a dynamic model, let s, e, c, p denote the concentrations of S, E, SE, P. The diagram then says that

- S and E combine to form SE at rate $k_1 se$
- SE separate to $S + E$ at rate $k_{-1} c$, and to $P + E$ at rate $k_2 c$

We need dynamic equations for each of the four "bathtubs" s, e, c, p. An S is lost whenever S and E combine to form an SE (which is like water flowing out of the bathtub), and gained whenever an SE separates into $S + E$. The principle is still the same:

rate of change = inflow rate − outflow rate.

We can construct a compartment diagram of this system by looking at the reaction diagram to find all the processes that create or destroy each chemical species, and then represent those as inputs to, and outputs from, each of the bathtubs (Figure 1.6). From the compartment diagram we can read off the rest of the dynamic equations:

$$ds/dt = k_{-1}c - k_1 es$$

$$de/dt = (k_{-1} + k_2)c - k_1 es$$

$$dc/dt = k_1 es - (k_{-1} + k_2)c$$

$$dp/dt = k_2 c.$$

[1.8]

The starting point for the process is to have only some substrate and enzyme present at concentrations s_0 and e_0, respectively:

$$s(0) = s_0, \quad e(0) = e_0, \quad c(0) = p(0) = 0.$$

Our interest is in the rate of product formation, so we do not really need the equation for p: the rate of product formation is $k_2 c(t)$. We can get rid of another dynamic equation by noting that

$$dc/dt + de/dt = 0.$$

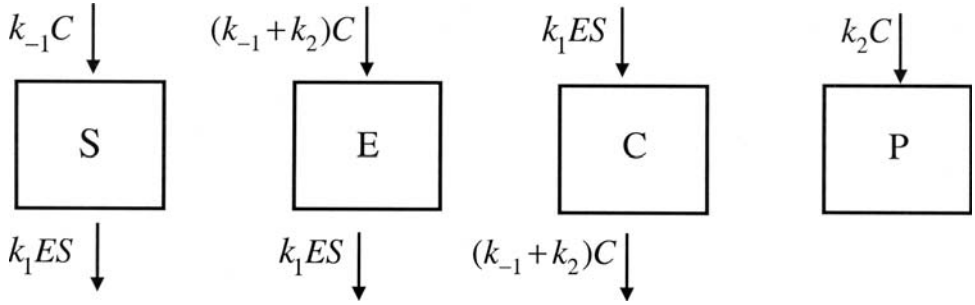This implies that $c(t) + e(t) = c(0) + e(0) = e_0$, so

$$e(t) = e_0 - c(t).$$

[1.9]

Figure 1.6 Compartment diagram for the enzyme-mediated biochemical reaction. The compartments are $S =$ unbound substrate, $E =$ unbound enzyme, $C =$ bound substrate-enzyme complex, and $P =$ reaction product. A single reaction can create or destroy molecules of several different chemical species, so an arrow in the reaction diagram can imply several different arrows in the compartment diagram.

This makes physical sense, because e and c are the unbound and bound forms of the enzyme, which is not created or destroyed in this reaction scheme (whereas substrate is irreversibly converted to product).

We can now substitute $e_0 - c$ for e into the equations for s and c, leaving us with a system of only two equations:

$$ds/dt = k_{-1}c - k_1(e_o - c)s$$
$$dc/dt = k_1(e_0 - c)s - (k_{-1} + k_2)c. \qquad [1.10]$$

We cannot solve these (except numerically on the computer) but we can get the information we want by a simple approximation. The key property of enzymes is that they are effective at very low concentrations, only a few percent or less of the substrate concentration. So we can assume that e_0 is *small*. Because [1.9] implies that $c(t) \le e_0$, we infer that $c(t)$ is also small.

Taking advantage of this assumption requires a change of variables. Let

$$v = c/e_0 \quad \Rightarrow c = e_0 v.$$

Then

$$dv/dt = (1/e_0)dc/dt = k_1(1 - c/e_0)s - (k_{-1} + k_2)c/e_0$$
$$= k_1(1 - v)s - (k_{-1} + k_2)v. \qquad [1.11]$$

More interesting is what happens to the other equation:

$$ds/dt = k_{-1}e_0 v - k_1(e_0 - e_0 v)s$$
$$= e_0(k_{-1}v - k_1(1 - v)s). \qquad [1.12]$$

Comparing [1.11] and [1.12], we see that s changes much more slowly because e_0 is small—so in [1.11] we can proceed *as if* s were constant. This brings us down

to the single equation

$$dv/dt = k_1 s - (k_{-1} + k_2 + k_1 s)v. \qquad [1.13]$$

It is easy to see graphically how [1.13] behaves: dv/dt is positive at $v = 0$ and decreases linearly with increasing v (Figure 1.7). So the behavior of solutions is that $v(t)$ approaches the value v^* at which $dv/dt = 0$:

$$k_1 s = (k_1 s + k_{-1} + k_2)v^* = (k_1 s + k_{-1} + k_2)c^*/e_0$$

$$k_1 e_0 s = (k_1 s + k_{-1} + k_2)c^* \qquad [1.14]$$

$$c^* = \frac{k_1 e_0 s}{k_1 s + k_{-1} + k_2} = \frac{e_0 s}{K + s}, \text{ where } K = (k_{-1} + k_2)/k_1.$$

c^* is called a *stable equilibrium point*: the system tends toward it, and then stays there.

Now we bring back the equation for p, which says that the rate of product formation is $k_2 c^*$. So defining $V_{max} = k_2 e_0$, the rate of product formation is

$$\frac{V_{max} s}{K + s}. \qquad [1.15]$$

Equation [1.15] is important enough to have a name, the Michaelis-Menten equation, named for the originators of the theory presented above. The importance comes from the very general prediction for the rate at which an enzyme-mediated reaction yields its final product.

In addition we have learned something important about the reaction system itself—the c equation is "fast." One value of models is that the things you know may turn out to imply other things that you did not realize before. The c equation describes how enzyme moves between bound and unbound states, as a function of how much substrate is around. It says that the fractions of bound versus unbound enzyme quickly reach an equilibrium value determined by the current amount of substrate (Figure 1.7). The amount of unbound enzyme then determines the rate at which substrate is converted to product via the intermediate step of binding to enzyme.

1.4.2 The Modeling Process

The development of our enzyme kinetics model illustrates the process of dynamic modeling.

1. A *conceptual model* representing our ideas about how the system works ...
2. ... is expressed visually in a *model diagram*, typically involving boxes (state variables) and arrows (material flows or causal effects).
3. Equations are developed for the rates of each process and are combined to form a *mathematical model* consisting of dynamic equations for each state variable.
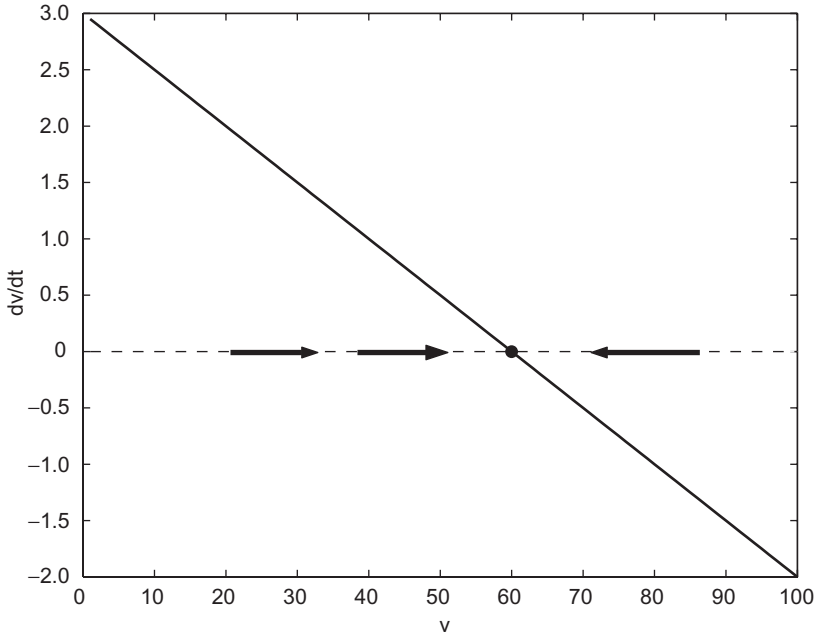
Figure 1.7 Graphical analysis of $dv/dt = k_1 s - (k_{-1} + k_2 + k_1 s)v$.

4. The dynamic equations can then be studied mathematically or translated into computer code to obtain *numerical solutions* for state variable trajectories.

An important part of the conceptual model is identifying which state variables and processes are important for your purposes and need to be included in the model. The conceptual model may be derived from known properties of the system. It may also be a set of hypotheses suggested by what is currently known about the system, and the goal of the model is to test the hypotheses by comparing model behavior against data. In this latter situation, a single model is less useful than a set of models representing different hypotheses. Instead of having to decide if your model fits the data, which is hard and subjective, you can see which model fits the data best, which is easier and more objective.

Rate equations can be based on various sorts of knowledge.

1. *First principles:* Fundamental laws of physics or chemistry that apply to the situation at hand. This is rarely enough; more often you need
2. *First principles + data:* The form of the equation is known from first principles or prior experience, but parameter values are not known. For example, if a Michaelis-Menten form is known to be appropriate, the two parameters in the model can be estimated from a small number of experimental measurements.
3. *Process rate data:* Curves are fitted to experimental measurements of how a specific process rate changes as a function of the model's state and exogenous variables,

using statistical methods that make few assumptions about the form of the equation.

4. *Previous models:* If it worked for somebody else, it might work for you, too. Then again ...

5. *System dynamics data:* Rate equations are deduced by trying to make the model produce state variable trajectories similar to experimental observations. This "inverse method" is technically difficult and used only as a last resort: a state variable is generally affected by many processes, and it is often not possible to disentangle one from another.

We will return to the modeling process in more detail in Chapter 9, and in particular we will describe methods for developing rate equations and fitting them to data.

1.4.3 Pharmacokinetic Models

More contemporary and practical examples of compartment models are models for the absorption, redistribution, and transformation of drugs or other ingested substances within the body, a set of processes collectively referred to as "pharma-cokinetics." By appropriately modifying parameter values in a model for these processes, data on one species (such as rats) can be used to make predictions for another (such as humans), for example about how much of the substance reaches any particular organ. Similarly, effects observed at low doses can be extrapolated to potential effects at higher doses.

Figure 1.8 shows a relatively simple example from Easterling et al. (2002), a model for arsenic transport and metabolism in rat hepatocytes. Arsenic is a natu-ral contaminant of drinking water, occurring in many areas at levels high enough to increase the risk of adverse effects including cancers of the skin, lung, liver, kidney, and bladder. This model is simple in part because it represents only one component in the development of a full-body model for arsenic.

The model was developed and tested with experimental data on individual rat hepatocytes incubated in media with various concentrations of arsenic. Each variable name in the model diagram (Figure 1.8) corresponds to a compartment in the model, arsenic (iAS) or one of its metabolites either in the cell or in the surrounding medium (MMA = methylarsonous acid, DMA = dimethylarsinic acid, p-iAS = protein-bound arsenic). Arrows represent transport and metabolic transformations that result in transfers between compartments, and the dotted line represents the inhibitory effect of arsenic on the rate of formation of DMA from MMA. The *p*'s and *k*'s are rate constants for these processes.

This simple model was already useful for understanding intracellular arsenic metabolism. In order to adequately model the data, it was necessary to assume that transformation from DMA to MMA was inhibited by arsenic. Three models
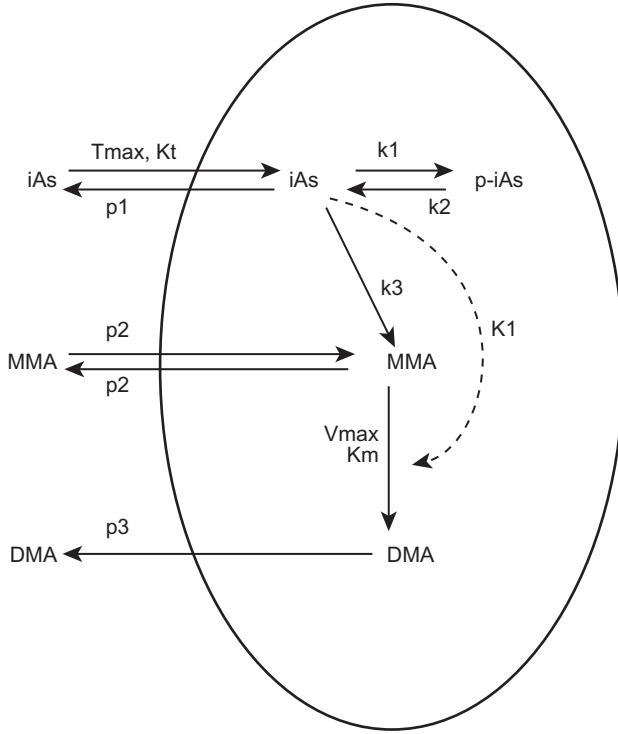
Figure 1.8 Compartment model for arsenic in rat hepatocytes
(from Easterling et al. 2002).

for inhibition were compared, and one of them was found to fit better than the others. Simulations of the three models were then used to identify experiments that could distinguish decisively between them, in particular measurements of MMA within the cell 46 hours after initial exposure to arsenic at $1.4\ \mu$M.

A much more complicated pharmacokinetic model, because it is complete and intended for practical use, is the Integrated Exposure Uptake Biokinetic Model for Lead in Children (IEUBK) developed and distributed by the U.S. Environmental Protection Agency (EPA) Technical Review Workgroup for Lead (EPA 2002; White et al. 1998). Figure 1.9 shows a diagram of the model.

Despite the elimination of leaded gasoline and lead-based paints in the United Stated, lead exposure remains a problem. A recent front-page headline in our local paper (Bishop 2002) concerned the cost overrun for cleaning up lead-contaminated soil surrounding a local recreation site used for fishing and swimming. "An environmental cleanup near Ithaca Falls that initially was projected to cost \$2 million and last about a year is now expected to cost twice as much and last twice as long." The contamination remained from previous manufacturing in the area. To bring lead exposure down to an acceptable level, a layer of soil was literally being vacuumed up off the ground and trucked off to a landfill.
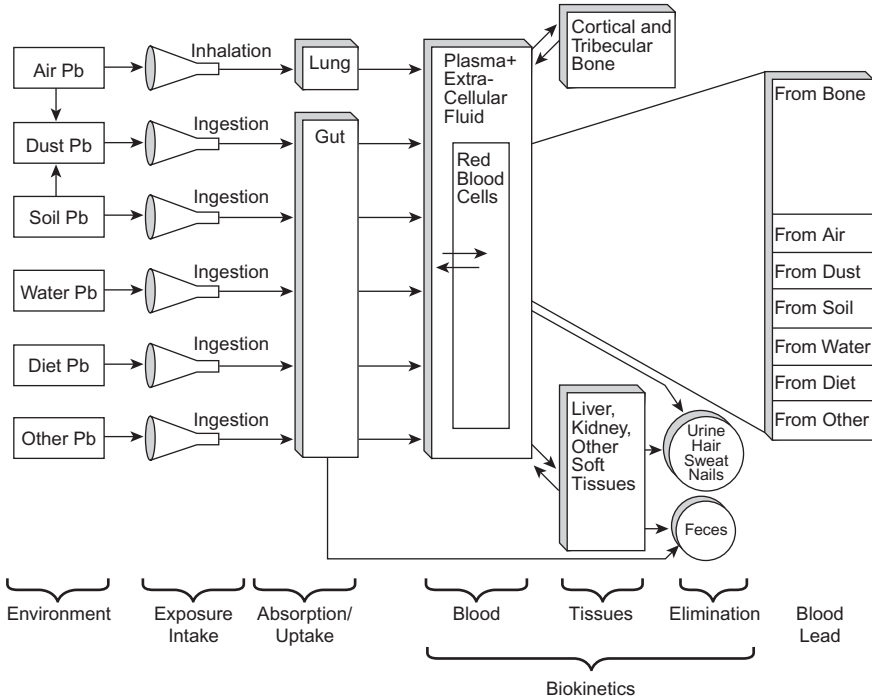
Figure 1.9 Compartment diagram for the IEUBK model for lead in children (from EPA 1994a). Plain rectangles represent environmental levels of lead, shaded rectangles are model compartments (forms of lead in the body that are distinguished in the model), and circles are losses of lead from the body.

And what should be an "acceptable level" of lead in the soil near a family swimming hole, based on the expected health impacts? The IEUBK model was developed to answer such questions. Producing version 0.99 required over a decade of effort by several different EPA programs (EPA 1994b). It was developed as an alternative to descriptive statistical models, because those models were difficult to generalize beyond the specific situations for which data had been collected. This is an example of the "curse of dimensionality" for descriptive models: the impact of lead exposure on a child depends on so many variables that it is infeasible to collect data spanning the full range of possibilities for all of them.

The full IEUBK model has three components: Exposure, Uptake, and Biokinetics. The exposure and uptake components are descriptive: a static set of equations that predict the amount of lead entering the child's body and transferring into the bloodstream, as a function of how much lead the child encounters in various daily activities but without reference to the biological processes involved. The biokinetic component is a dynamic compartment model. It starts by calculating the volumes and weights of model compartments (e.g., the volume of blood plasma) as a function of the child's age and the initial amounts of lead in

each compartment. The model equations specify the rates of lead transfer between compartments and the rate of lead loss through elimination—the rates of flow between the "bathtubs"—and the equations are solved numerically, from birth to age 84 months. Body fluid volumes and organ weights are also changed as the child ages, but this is modeled descriptively, based on published data on child growth (EPA 1994b). Putting all this together, lead levels in the soil around Fall Creek can be translated into the expected lead level in the blood (given an assumed frequency and duration of visits) and then into the expected health impacts.

The model is quite complex—a listing of its equations requires 22 printed pages (EPA 1994b), and there are roughly 100 parameters. Models intended for practical use are often complex, because accuracy is more important than aesthetics so the model grows to match the available data. However, it is important to remember that the model is only complex because it has many parts. All of its pieces are simple bathtubs obeying the basic balance equation: rate of change = (total input rate) − (total loss rate).

1.5 Physics Models: Running and Hopping

In compartment models the underlying physical law is conservation of mass. In this section we introduce a model based on the physical laws of mechanics. The variables in the model are positions and velocities, and the dynamic equations come from Newton's laws of motion.

Many animals use legs to move. The number of legs and the gaits they employ differ markedly. To study locomotion, *biomechanical* models that represent the animal as a rigid skeleton connected by joints are a starting point. Gravity and muscles exert forces on the skeleton, causing the animal to move. By measuring the physical properties of the animal—its shape, mass distribution, and muscle strength—we want to be able to analyze and predict the motion. With humans, our purpose may be to improve athletic performance or to restore function following an injury.

Different gaits have different mechanical characteristics. When humans walk, our feet never lose contact with the ground and we alternate between having both feet on the ground and a swing phase in which one foot is on the ground and the second leg swings like a pendulum. When we run, we alternate between a flight phase in which both feet are off the ground and a stance phase in which one foot is on the ground. Kangaroos hop with a flight phase alternating with a stance phase in which both feet are on the ground simultaneously.

A pogo stick with a single leg provides a simple model for running and hopping (Figure 1.10). The animal is regarded as a "point mass" with a springy leg attached to the mass. As a first approximation, you can think of the spring as the tendons
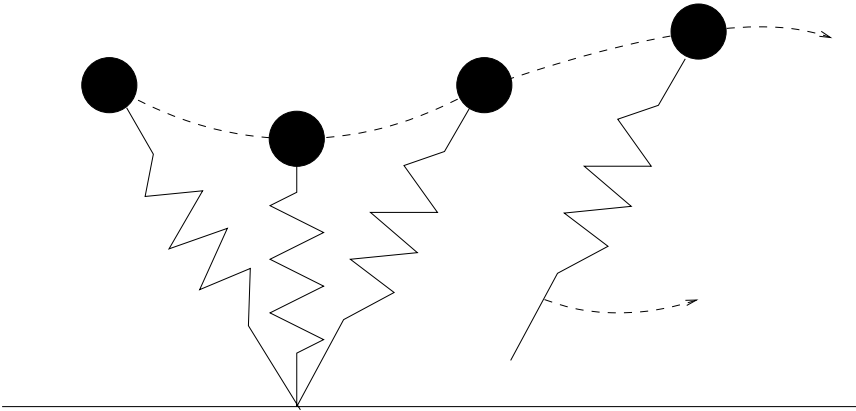
Figure 1.10 A cartoon of the pogo stick or monopode.

in the leg. By contracting muscles, the animal changes the force of the leg spring, enabling it to bounce off the ground. When in flight, we assume that the animal is able to swing the leg so that it will point in a new direction when the animal lands on the ground. At landing, the leg shortens, compressing the spring. The compressed spring exerts a vertical upward force that together with additional force exerted by the muscles propels the animal into its next flight phase.

We construct a system of equations that embodies this verbal description of a bouncing "monopode." For simplicity, we consider only the motion of the monopode in a vertical plane. There are two components to the virtual animal: a body that is assumed to be a point mass with mass m, and a massless leg whose (variable) length is denoted by r. Euclidean coordinates (x, y) will be used to label the position of the body: we write $(x(t), y(t))$ for the position of the body at time t. The leg is attached to the body by a "hip" joint that is free to rotate, so we label the angle of the leg with the downward vertical as θ. The position of the foot relative to the body is $(r \sin(\theta), -r \cos(\theta))$ so its position in "world coordinates" is $(x + r \sin(\theta), y - r \cos(\theta))$. The ground is assumed to be the x-axis $y = 0$.

When the foot of the monopode is above the ground, we say that the monopode is in "flight." During flight, we ignore friction and assume that gravity is the only force acting on the body, so Newton's laws of motion ($F = ma$) are just those for a particle moving under the influence of gravity:

$$m\ddot{x} = 0$$
$$m\ddot{y} = -mg.$$
[1.16]

Here $\ddot{x}$ and $\ddot{y}$ denote the acceleration (second derivative with respect to time) of the monopode's body in the x and y directions, and g is the gravitational force constant 9.8 m/s^2. The motion of the body in flight is along a parabola determined by the position and velocity of the body at the instant when the

flight begins. The position of the leg has no influence on the body motion until the foot hits the ground.

When the foot makes contact with the ground, we say that touchdown occurs and the stance phase of the motion begins. The response of the monopode to the impact is determined by additional physical properties that have to be specified to complete the model. We make the assumption that the foot sticks at the point of impact, rather than slipping along the ground or bouncing into the air. At touchdown, the downward motion of the body begins to compress the springy leg, so there is an additional force on the body exerted through the spring. Thus the equations of motion for the stance phase are more complicated than those for the flight phase. They are expressed most simply in a coordinate system in which the origin is placed at the foot (Figure 1.11). We denote the position of the body relative to the foot by (u, v). The force exerted on the body by the spring is proportional to the deviation between the body's location and its "resting location" (u_0, v_0)—where it would be if the spring were at its resting length r_0. The force vector is therefore

$$\vec{F} = f(r)(u_0 - u, v_0 - v) \qquad\qquad\qquad [1.17]$$

where $f(r)$ is the "spring constant" and $(u_0, v_0) = (r_0/r)(u, v)$. For a linear spring $f(r)$ is constant, but in general it can depend on how far the spring has been compressed or stretched. The equations of motion for the body are therefore

$$m\ddot{u} = f(r)(u_0 - u)$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad [1.18]$$
$$m\ddot{v} = f(r)(v_0 - v) - mg.$$

We assume that the motion of the body reaches a positive lowest height during stance: the spring force becomes larger than the force of gravity, slows the vertical motion of the body, and then propels the body upward. If the body hits the ground, then we regard the monopode as having fallen, and stop. We also assume that there is a maximal leg length l and that when the leg reaches this length, the foot comes off the ground and a new flight phase begins.

There is still one additional item that must be specified to obtain a complete model. Since the leg is assumed to be massless, there is no energetic cost in moving the leg relative to the body during flight. Changes in the angle θ do not affect the flight phase of the body, but θ affects foot position and therefore influences when touchdown occurs. We assume that our monopode can point the leg during flight as one of two things it can do to control its motion. It does not matter when or how this happens as long as the foot does not make contact with the ground during repointing, so we assume that the repointing occurs instantaneously when the body reaches its apex (highest point) during the flight phase. The second way in which we assume that the monopode can control its motion is to change the spring constant $f(r)$ during stance. This corresponds to using muscles to push along the leg or to make the leg stiffer.
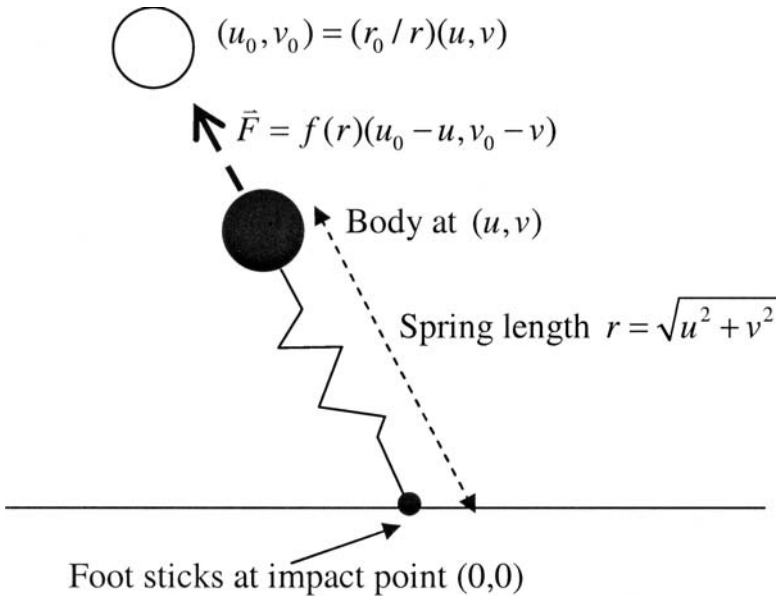
$$(u_0, v_0) = (r_0 / r)(u, v)$$

$$\vec{F} = f(r)(u_0 - u, v_0 - v)$$

Body at (u, v)

Spring length $r = \sqrt{u^2 + v^2}$

Foot sticks at impact point (0,0)

Figure 1.11 The monopode in stance phase. The dynamic equations are written in coordinates (u, v) giving the position of the body relative to the point where the foot meets the ground. The open circle shows the body's "resting position" corresponding to the leg at resting length r_0, and the bold dashed arrow indicates the resulting force vector generated by the spring.

There are many questions that we can ask about this model of running animals. First, we can try to fit observational data to such a model. Blickhan and Full (1993) show that the model works well for animals as "diverse as cockroaches, quail and kangaroos." Even though these animals are mechanically far more complicated than the monopode, it is possible that the behaviors that they actually use to control their motion do not fully exploit this complexity, and are similar to what would occur if the organisms were simple monopodes.

Second, we can ask what kind of motion results from different *control strategies*—the rules used for adjusting position and movements on the basis of the current situation. A goal that we may want to achieve is to produce a regular stable running motion in which each stride is the same as the previous one, except for being further along in the direction of motion. Koditschek and Buehler (1991) showed that the model does not always do this. A fixed control strategy can give rise to irregular, "chaotic" motions in which apex height continues to fluctuate from one hop to the next. Carver (2003) investigated "deadbeat" control strategies for which regular running motion at a desired speed can be obtained from a large region of initial positions after a single stride. Third, we can ask whether the model might help with the design of better legged robots. Raibert (1986) utilized models like this one in the design of monopode, biped, and quadruped robots. Similar principles are being utilized by Buehler and co-workers in advancing the

performance of hexapod robots (Saranli et al. 2001, or see http://www.rhex.org for movies of the robot in action).

The monopode model is just a starting point in understanding animal locomotion. More elaborate models with a large number of independently movable body parts have more fidelity to the actual biomechanics of animal locomotion, but producing simulations that are indistinguishable from "motion capture" of actual animals requires a better understanding of neuromuscular control than we currently have.

1.6 Optimization Models

Biologists also use *optimization* models, in which the modeler assumes that the study organisms are trying to achieve some particular goal. For example, we might posit that an animal will gather food in the habitat that gives it the best chance of getting enough food to survive, without itself getting eaten by its predators. Or we might ask how a set of neurons *ought to* be connected, and how they should affect each other, in order to best achieve some important task like identifying the direction a sound is coming from. An optimization model does not worry about how the goal is achieved in practice—how the animal evaluates different habitats, or how neurons accomplish their growth and connection patterns during development. We assume that natural selection will find a way, and see if the actual behavior matches the predictions based on the assumed goal.

Optimization models are especially informative when different assumptions lead to different predictions that can be compared with observations. For example, Segrè et al. (2002) predicted the changes in cell metabolism in *Escherichia coli* caused by a gene deletion ("knockout") that eliminated a particular metabolic pathway, based on two different assumed goals. The first was that certain parameters of the metabolic reaction network had values such that the net growth rate of knockout mutant cells was the highest possible given the deletion. The second was that network parameters minimized the disruption in the flux rates on the remaining pathways—that is, that the cell has evolved to be highly robust against mutations. Although the outcome was not uniform, overall the second model was far more successful at predicting the effects of several knockout mutations. Clutton-Brock et al. (1996) compared predictions about size-dependent reproductive decisions in Soay sheep—whether to have zero, one, or two offspring in a given year—based on two different assumptions about how well females can predict changes in population density and their effect on mortality rates. Again, neither model was exactly right, but the model assuming that females could make reliable short-term predictions did far worse than one assuming that females could not. In these studies the models were not *really* making predictions—the data were already in hand. But by asking what observable features of the study

system would result from different assumptions about unobservable features, it was possible to make inferences about properties that could not be observed directly.

1.7 Why Bother?

A dynamic model is built up from equations representing the processes thought to account for the patterns observed in the data, whereas a descriptive model only has to represent the patterns themselves. That difference typically means that constructing a dynamic model is a lot more work than constructing a descriptive statistical model. So why would anyone bother?

1. *Scientific understanding.* A model embodies a hypothesis about the study system, and lets you compare that hypothesis with data.

- Can processes A and B account for pattern C? Having observed some pattern in our study system, a model lets us ask whether it is *logically* possible for our hypothesized causes to have produced it, by creating (on the computer) an artificial world in which A and B are the only things happening, and seeing if and when C can be the outcome (e.g., Harris-Warrick et al. 1995). Hypotheses that fail this test can be rejected without taking the trouble to do experiments.

- "Biological detective" work: Which of several contending sets of assumptions is best able to account for the data? Scientists are often clever enough to invent several sets of causes (A_i, B_i) that are logically capable of producing something like observed pattern C. One way to decide between them, or at least narrow the field of possible explanations, is to implement each hypothesis as a quantitative model, and see which of them can match the available empirical data (Hilborn and Mangel 1997; Kendall et al. 1999).

- Given that processes A and B occur, what consequences do we expect to observe? Patterns may not be seen until they have been predicted to occur.

- Where are the holes in our understanding? Trying to build a model is a great way to see if you really have all the pieces.

Paradoxically, when a dynamic model is being used as a research tool to increase our scientific understanding, it is often most useful when it *fails to fit the data*, because that says that some of your ideas about the study system are wrong.

2. *Using our scientific understanding to manage the world*

- Forecasting disease or pest outbreaks

- Designing man-made systems, for example, biological pest control, bioengineering

- Managing existing systems such as agriculture or fisheries

- Optimizing medical treatments or improving athletic performance

In this book, and already in this chapter, there are many examples where models have been put to practical uses like these. Forecasting and management can be done, and often are done, using descriptive models. But a lot of effort also goes into building mechanistic models for forecasting, design, and management. Why is this?

3. *Experiments are small, the world is big.* Reviewing studies of plant competition, David Tilman (Tilman 1989) found

- Nearly half were on a spatial scale of $\leq 1 \text{ m}^2$

- About 75% on a scale $\leq 10 \times 10 \text{ m}^2$

- About 85% ran for ≤ 3 years

When large-scale (in space or time) experiments are infeasible, dynamic models can be built using the processes studied by small-scale experiments and then used to derive their larger-scale implications.

4. *There are experiments that you would rather not do*

- Endangered species management by trial and error

- Setting dosages for clinical trials of new drugs on humans

- Setting "safe" limits for exposure to toxic substances

A model can never be a complete substitute for experiments, but if the experiment is infeasible or undesirable, a model based on current scientific knowledge is far better than guesswork. Models can also be used to help plan risky experiments to increase the odds of success, such as drug trials or recovery plans for an endangered species.

5. *The curse of dimensionality.* Sometimes a purely experimental approach is not feasible because of the *curse of dimensionality*: the data requirements for estimating a statistical model with many variables grow rapidly in the number of variables.

As an example, one of our projects concerned the design of fish-farming systems (Krom et al. 1995; Ellner et al. 1996). The goal was to raise marine fish in a limited area of artifical ponds, allowing very little pollution due to nutrient export, which would have gone into the sea near a coral reef. The limiting factor in the system was the buildup of ammonia (NH_4), which reaches toxic levels if too many fish are excreting into a limited volume of water. One way to limit ammonia buildup is to have water flow through the system, but this releases a lot of nutrient-
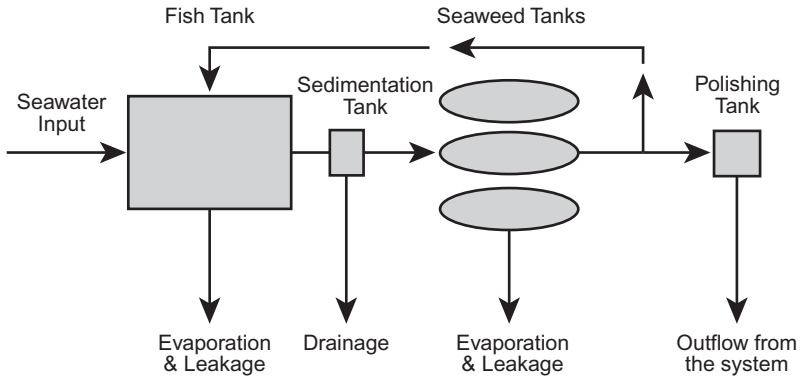
Figure 1.12 Diagram of the experimental fishpond system showing water flows between compartments.

rich pollution. So we designed a system that circulated water between ponds containing the fish and tanks containing seaweed that could be sold as a second crop, followed by a final "polishing" tank with seaweed before the effluent is released (Figure 1.12). Variables under the designer's control include

- Volumes of the fish ponds, seaweed tanks, and polishing tank

- Stocking densities (number of fish, weight of seaweed) in each tank

- Fish feeding schedule (how much, and when)

- Rate of water input

- Rate of water recirculation

- Degree of shading on fish and seaweed tanks to control phytoplankton (phytoplankton consume ammonia—which is good—but compete with seaweed, which is not)

How can we find the optimal combination of values for each of these twelve design parameters? If we want to determine experimentally where a function of one variable takes its maximum value, we need enough data points that one will be close to the maximum. So if we're *really* lucky, three data points will be enough. With two variables, to find the maximum with the same accuracy requires a 3×3 array of data points; hence nine experiments. But with twelve variables under our control and three values for each of these, that would be $3^{12} = 531,141$ experiments. This crude estimate turns out to be qualitatively correct, in the following sense. If we assume only that the function is differentiable, then the average estimation error ϵ using optimal methods scales asymptotically as $\epsilon \sim N^{-1/(2+n)}$ (Stone 1982), where n is the number of variables and N the number

of data points. This inverts roughly to $N \sim C^{2+n}$ data points to achieve a given estimation error, where $C = 1/\epsilon$.

So instead of a purely experimental approach, we built a dynamic model for the flow of inorganic nitrogen through the system. We quantified the processes causing N to be enter and leave each component of the system, resulting in a model that could predict ammonia levels and the nutrient export rate as a function of all the control variables. We could do this with fewer experiments because each process in the model depends on only a few variables—often only one. Also, many experiments could be done at the lab bench, such as measuring how fish size affects the ammonia excretion rate. That made it possible to get data spanning any values that might occur in the real system. Moreover, for many processes it was reasonable to assume on mechanistic grounds that the equation describing the process rate was a line with zero intercept, which can be estimated from a single data point. A dynamic model was feasible within our time and budget constraints, whereas a purely statistical approach would not have been. And in subsequent trials, the resulting system design used far less water than a system without recirculation.

For these kinds of practical applications the model needs to be consistent with all available data, so that you have some grounds for trusting its predictions about situations where data are not available. Therefore, an interaction between modeling and experiments is useful:

- Given a model, you can do lots of experiments quickly and cheaply on the computer.

- Those results can be used to identify potentially good designs or management options.

- Model predictions can be checked by doing a few well-chosen experiments.

1.8 Theoretical versus Practical Models

Although dynamic models are used for many purpose, we can put them under two broad headings: *theoretical understanding* of how the system operates, and *practical applications* where model predictions will play a role in deciding between different possible courses of action (Table 1.1). It is important to keep this distinction in mind because it affects how we build and evaluate models.

A theoretical model—such as the monopode model—has to be simple enough that we can understand why it is doing what it does. The relationship between hypotheses (model assumptions) and conclusions (properties of model solutions) is what provides understanding of the biological system. Replacing a complex system that we don't understand with a complex model that we also don't un-

Practical Models	Theoretical Models
Main goals are management, design, and prediction	Main goals are theoretical understanding and theory development
Numerical accuracy is desirable, even at the expense of simplicity	Numerical accuracy is not essential; the model should be as simple as possible
Processes and details can be ignored only if they are numerically unimportant	Processes and details can be ignored if they are conceptually irrelevant to the theoretical issues
Assumptions are quantitative representations of system processes	Assumptions may be qualitative representations of hypotheses about the system, adopted conditionally in order to work out their consequences
System and question specific	Applies to a range of similar systems

Table 1.1 Classification of models by objectives

derstand has not increased our understanding of the system. Theoretical models are often expressed as a few dynamic equations (e.g., matrix or differential equations) representing the most relevant processes for the particular phenomena of interest.

A practical model—such as the IEUBK model—sacrifices simplicity in order to make more accurate predictions for a particular system. There may be a lot of equations with terms based on quantitative data. As a result, practical models are usually too complicated for anything but computer simulation. However, theoretical models are nowadays typically studied by a combination of mathematical analysis, approximations, and computer simulation. Even simple-looking non-linear models may be beyond the reach of totally rigorous mathematical analysis. More often, a model is explored by simulation and we then try to use mathematics to understand what the computer has shown us.

The goal-centered nature of practical models makes it conceptually easy to evaluate them: a good model is one that lets you make better decisions. "Theoretical understanding" is not something we can quantify, so model evaluation is more difficult. A quail is not a monopode, but a monopode model can describe its gait well: good model or bad? (Answer: That's the wrong question).

The model types described above and in Table 1.1 are extremes of a continuum, and many models fall somewhere in between. Constraints on time and effort, data limitations, and the need to communicate with stakeholders place limits on the complexity of practical models. Theoretical models are often challenged by comparisons with experimental data. For that purpose models often remain

theoretical in the sense of including only a few "essential" processes, but the rate equations are allowed to become more complicated, so that a model with the right basic structure should survive a quantitative comparison with data.

1.9 What's Next?

This chapter has tried to introduce the distinctive features of dynamic models, some of the considerations involved in creating them, and some of the ways in which they are useful. What's next is the process of taking you, gradually, from being a spectator to acting as a participant in dynamic modeling. The remaining chapters in this book will introduce you to some important types of dynamic models that have stood the test of time and are likely to be around for a long time to come. New ideas are hard to come by, but fortunately a lot of new problems can be solved using old ideas. Cutting-edge science is still being done with centuries-old modeling frameworks, augmented by modern computers. Along the way, we will introduce mathematical and computational methods for working with dynamic models and understanding what they do. The final chapter then comes back to look in more detail at some of the issues we have touched on here about the process of formulating and testing dynamic models.

Exercise 1.1. Find a scientific research paper published within the last five years that uses a dynamic model in your area of biology (your major, concentration, or any area of particular interest to you), and read the paper to answer the following questions.

(a) Give the complete citation for the paper: authors, date, journal, pages.

(b) What was the purpose of the model—that is, what was accomplished by building and using the model?

(c) What are the state variables of the model?

(d) Identify one of the model's simplifying assumptions—some known aspect of the real world that the model omits or simplifies.

Exercise 1.2. This exercise refers to the dynamic model for salmon stocks. Table 1.2 gives some estimated spawner counts by age (the estimates are not integers because they are derived by scaling up from sample data). Use these data to estimate

(a) the value of $p_4(t)$ for as many years as possible

(b) the value of $E(t)$ for as many years as possible

Note that it may be possible to estimate $p_4(t)$ and/or $E(t)$ for years other than those appearing in the leftmost column of the table. Recall that for fish spawning in year t, $E(t)$ is the total number of offspring per spawner that will return to spawn, at either

Year	Four-year-old spawners	Five-year-old spawners
1985	84.2	210.4
1986	72.6	151.5
1987	116.3	340.0
1988	26.0	1083.3
1989	27.8	63.1
1990	72.4	112.7
1991	23.7	157.0
1992	47.2	125.8
1993	44.7	664.4

Table 1.2 Estimated spawner counts by age for the Bear Valley stock of Chinook Salmon on the Snake River

age four or age five; $p_4(t)$ is the fraction of those that return at age four. *Hint*: The first line in the table says that in 1985 there were $S = (84.2 + 210.4)$ spawners. Which entries in the table tell you how many offspring of those spawners returned eventually to spawn?

Exercise 1.3. Find the differential equations for the diagrams in Figure 1.13 representing continuous-time compartment models. This exercise uses some notational conventions for compartment models. $\rho_{i,j}$ is the flow rate (amount/time) *from* compartment j *to* compartment i. If an outflow rate is proportional to the amount in the compartment of origin, this is called *linear donor control* and the constant of proportionality is $a_{i,j}$. That is, $\rho_{i,j} = a_{i,j}q_j$ where q_j = amount in jth compartment.
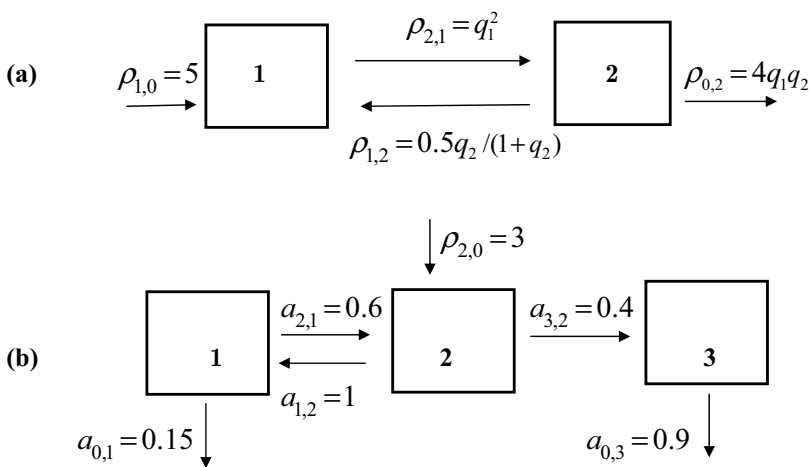

Figure 1.13 Compartment diagrams for Exercise 1.3.

Exercise 1.4. Another convention is that production within a compartment is represented as an input from outside, and any internal losses (consumption, degradation, etc.) are represented as a loss to the outside. For example, in a diagram of the Ellner et al. (1996) fishpond model, excretion of ammonia-N by fish into the fishpond water column would be drawn like $\rho_{2,0}$ in Figure 1.13(b), pointing into the box for the state variable "ammonia-N in the fishpond."

   (a) Using this convention, draw a compartment diagram for the Ellner et al. (1996) fishpond model, omitting the polishing tank so that there are six state variables (ammonia-N and ToxN in the fish tank, sedimentation tank, and seaweed tanks). Label each arrow with a ρ (subscripted appropriately), and below the diagram list what each flow is. So for example, if the fishpond is compartment 1, then your list could start with

   $\rho_{1,0}$ = ammonia-N excretion by fish in the fishtank

   (b) Which of the flows that you listed above are linear donor controlled in the model?

Exercise 1.5. Draw compartment diagrams (with flows labeled as in Figure 1.13) that correspond to the following dynamic equations.

   (a) $\dfrac{dq_1}{dt} = 1 - q_1, \quad \dfrac{dq_2}{dt} = q_1 - 2q_2,$

   (b) $\dfrac{dq_1}{dt} = -q_1, \quad \dfrac{dq_2}{dt} = q_1 - q_2 q_3 + q_2/(1 + 5q_2), \quad \dfrac{dq_3}{dt} = 3 - q_2/(1 + 5q_2).$

Exercise 1.6. We wrote *A quail is not a monopode, but a monopode model can describe its gait well: good model or bad? (Answer: That's the wrong question).*

   (a) Why is that the wrong question?

   (b) What was the right question, and why?

1.10 References

Bishop, L. 2002. Cleanup estimates to double. Ithaca Falls project will run over in time, money. *Ithaca Journal* (http://www.ithacajournal.com), October 4, 2002.

Blickhan, R., and R. Full. 1993. Similarity in multilegged locomotion: Bouncing like a monopode. Journal of Comparative Physiology A 173: 509–517.

Bradford, M. J., and J. R. Irvine. 2000. Land use, fishing, climate change, and the decline of Thompson River, British Columbia, coho salmon. Canadian Journal of Fisheries and Aquatic Science 57: 13–16.

Carver, S. 2003. Control of a Spring-Mass Hopper. Ph.D. Thesis, Cornell University.

Clutton-Brock, T. H., I. R. Stevenson, P. Marrow, A. D. MacColl, A. I. Houston, and J. M. McNamara. 1996. Population fluctuations, reproductive costs, and life-history tactics in female Soay sheep. Journal of Animal Ecology 65: 675–689.

Dickson, T. R. 1968. *The Computer and Chemistry; An Introduction to Programming and Numerical Methods*. W. H. Freeman, San Francisco.

Easterling, M. R., M. Styblo, M. V. Evans, and E. M. Kenyon. 2002. Pharmacokinetic modeling of arsenite uptake and metabolism in hepatocytes—mechanistic insights and implications for further experiments. Journal of Pharmacokinetics and Pharmacodynamics 29: 207–234.

Ellner, S., A. Neori, M. D. Krom, K. Tsai, and M. R. Easterling. 1996. Simulation model of recirculating mariculture with seaweed biofilter: Development and experimental tests of the model. Aquaculture 143: 167–184.

Ellner, S. P., and J. Fieberg. 2003. Using PVA for management in light of uncertainty: Effects of habitat, hatcheries, and harvest on salmon viability. Ecology 84: 1359–1369.

EPA (U.S. Environmental Protection Agency). 1994a. Guidance Manual for the Integrated Exposure Uptake Biokinetic Model for Lead in Children. Publication number 9285.7-15-1, EPA 540-R-93-081, PB93-963510. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency Washington, DC 20460. Online at http://www.epa.gov/superfund/programs/lead/products.htm

EPA (U.S. Environmental Protection Agency). 1994b. Technical Support Document: Parameters and Equations Used in the Integrated Exposure Uptake Biokinetic (IEUBK) Model for Lead in Children (v 0.99d). Publication number 9285.7-22, EPA 540/R-94/040, PB94-963505. Online at http://www.epa.gov/superfund/programs/lead/products.htm

EPA (U.S. Environmental Protection Agency). 2002. User's Guide for the Integrated Exposure Uptake Biokinetic Model for Lead in Children (IEUBK). Windows version, 32 bit version. Publication number EPA 540-K-01-005. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, DC 20460. Online at http://www.epa.gov/superfund/programs/lead/products.htm

Harris-Warrick, R., L. Coniglio, N. Barazangi, J. Guckenheimer, and S. Gueron. 1995. Dopamine modulation of transient potassium current evokes phase shifts in a central pattern generator network. J. Neuroscience 15: 342–358.

Hilborn, R. and M. Mangel. 1997. *The Ecological Detective: Confronting Models with Data*. Princeton University Press, Princeton, NJ.

Kendall, B. E., C. J. Briggs, W. W. Murdoch, P. Turchin, S. P. Ellner, E. McCauley, R. M. Nisbet, and S. N. Wood. 1999. Inferring the causes of population cycles: A synthesis of statistical and mechanistic modeling approaches. Ecology 80: 1789–1805.

Koditschek, D. and M. Buehler. 1991. Analysis of a simplified hopping robot. International Journal of Robotics Research 10: 587–605.

Krom, M. D., S. Ellner, J. van Rijn, and A. Neori. 1995. Nitrogen and phosphorus cycling in a prototype "non-polluting" integrated mariculture system, Eilat, Israel. Marine Ecology Progress Series 118: 25–36.

Levin P. S., R. W. Zabel, and J. G. Williams. 2001. The road to extinction is paved with good intentions: Negative association of fish hatcheries with threatened salmon. Proceedings of the Royal Society of London Series B 268: 1153–1158.

Michaelis, L., and M. L. Menten. 1913. Die Kinetik der Invertinwerkung. Biochemische Zeitschrift 49: 333–369.

Oosterhout, G. R., and P. R. Mundy. 2001. The Doomsday Clock 2001: An Update on the Status and Projected Time to Extinction for Snake River Wild Spring/Summer Chinook Stocks. Report prepared for Trout Unlimited. Online at http://www.decisionmatrix.net/models.htm and http://www.tu.org/pdf/newsstand/library/doomsday_clock_2001_report.pdf

Raibert, M. 1986. *Legged Robots that Balance*. MIT Press, Cambridge, MA.

Saranli, U., M. Buehler, and D. E. Koditschek. 2001. RHex: A simple and highly mobile hexapod robot. International Journal of Robotics 20: 616–631.

Segrè, D., D. Vitkup, and G. M. Church. 2002. Analysis of optimality in natural and perturbed metabolic networks. Proceedings of the National Academy of Sciences USA 99: 15112–15117.

Stone, C. J. 1982. Optimal rates of convergence for nonparametric regression. Annals of Statistics 10: 1050–1053.

Tilman, D. 1989. Ecological experimentation: Strengths and conceptual problems. pages 136–157 in G. E. Likens (ed.), *Long-Term Studies in Ecology: Approaches and Alternatives.* Springer-Verlag, New York.

White, P., P. Van Leeuwen, B. Davis, M. Maddaloni, K. Hogan, A. Marcus, and R. Elias. 1998. U.S. Environmental Protection Agency: The conceptual structure of the integrated exposure uptake biokinetic model for lead in children. Environmental Health Perspectives 106 (Suppl. 6): 1513.

Wilson, P. H. 2003. Using population projection matrices to evaluate recovery strategies for Snake River spring and summer Chinook salmon. Conservation Biology 17: 782–794.