Mathematical Modeling in Microbial Ecology

Chapman & Hall Microbiology Series

Physiology/Ecology/Molecular Biology/Biotechnology

Other Publications in the Chapman & Hall Microbiology Series

Methanogenesis; James G. Ferry, ed.
Acetogenesis; Harold L. Drake, ed.
Gastrointestinal Microbiology, Volume 1; Roderick I. Mackie and Bryan A. White, eds.
Gastrointestinal Microbiology, Volume 2; Roderick I. Mackie, Bryan A. White, and
        Richard E. Isaacson, eds.
Bacteria in Oligotrophic Environments; Richard Y. Morita

Forthcoming Titles in the Chapman & Hall Microbiology Series

Oxygen Regulation of Gene Regulation in Bacteria; Rob Gunsalus, ed.
Metal Ions in Gene Regulation; Simon Silver and William Walden, eds.

Mathematical Modeling in Microbial Ecology

Edited by

Arthur L. Koch

Indiana University
Department of Biology
Bloomington, IN

Joseph A. Robinson

Upjohn Laboratories
Biostatistic and Environmental Research
Kalamazoo, MI

George A. Milliken

Kansas State University
Department of Statistics
Manhattan, KS

Contents

Preface

Mathematical modeling has become a staple in the inferential tool kit used by microbial ecologists to discern patterns in data. The wide availability of powerful microcomputers and statistical software have contributed to this development. Microbial ecologists can now easily explore the opportunities offered by mathematical modeling such as using models as (1) heuristic devices, (2) means of generating hypotheses about biological systems, and (3) aids in making decisions (i.e., the testing of hypotheses statistically). The chapters in this text provide examples of each of these different, but complementary aims of mathematical modeling.

For microbial ecologists, the range of systems being subjected to modeling analyses runs the gamut from subcellar systems to ecosystems. Throughout, the objective is to express the system of interest in mathematical terms and to test whether the model provides an adequate representation of the system, depending on the researcher's objectives. The path between this objective and the development of a parsimonious mathematical model winds through various techniques, including parameter estimation, sensitivity analysis, model discrimination, and experimental design. It is hoped that the readers of this text will be served well by the information contained herein such that the biology remains the focus, when that focus is appropriate, and is not lost among the mathematical and statistical tools.

Contributors

R. L. Baldwin, Department of Animal Science
University of California–Davis
Davis, CA 95616

Francis H. Chapelle, Water Resources Division
U.S. Geological Survey
Stephenson Center, Suite 129
Gracern Road
Columbia, SC 29210

Robert R. Christian, Biology Department
East Carolina University
Greenville, NC 27834

Peter T. Cummings, Department of Chemical Engineering
University of Tennessee
Knoxville, TN 37996
and
Chemical Technology Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831

Michael A. Doherty, USDA-ARS-NRI-ECL
Building 001
BARC-West
10300 Baltimore Avenue
Beltsville, MD 20705

K. C. Donovan, Department of Animal Science
University of California–Davis
Davis, CA 95616

Roseanne M. Ford, Department of Chemical Engineering
University of Virginia
Charlottesville, VA 22903

Stephen P. Garabedian, U.S. Geological Survey
28 Lord Road, Suite 208
Marlborough, MA 01752

Arthur L. Koch, Department of Biology
Indiana University
Jordan Hall
Bloomington, IN 47405

Derek, R. Lovley, Department of
  Microbiology
Morrill Science Center IV
University of Massachusetts
Amherst, MA 01003

April J. Milliken-MacKinnon,
  Milliken Associates, Inc.
Manhattan, KS 66502

George A. Milliken, Department of
  Statistics
Kansas State University
Dickens Hall
Manhattan, KS 66506

David D. Myrold, Department of
  Crop and Soil Science
Oregon State University
Agriculture and Life Sciences
Building, Room 3017
Corvallis, OR 97331

Mariachiara Naldi, Dipartimento di
  Scienze Ambientali
Universita di Parma
Via Delle Scienze
43100 Parma, Italy
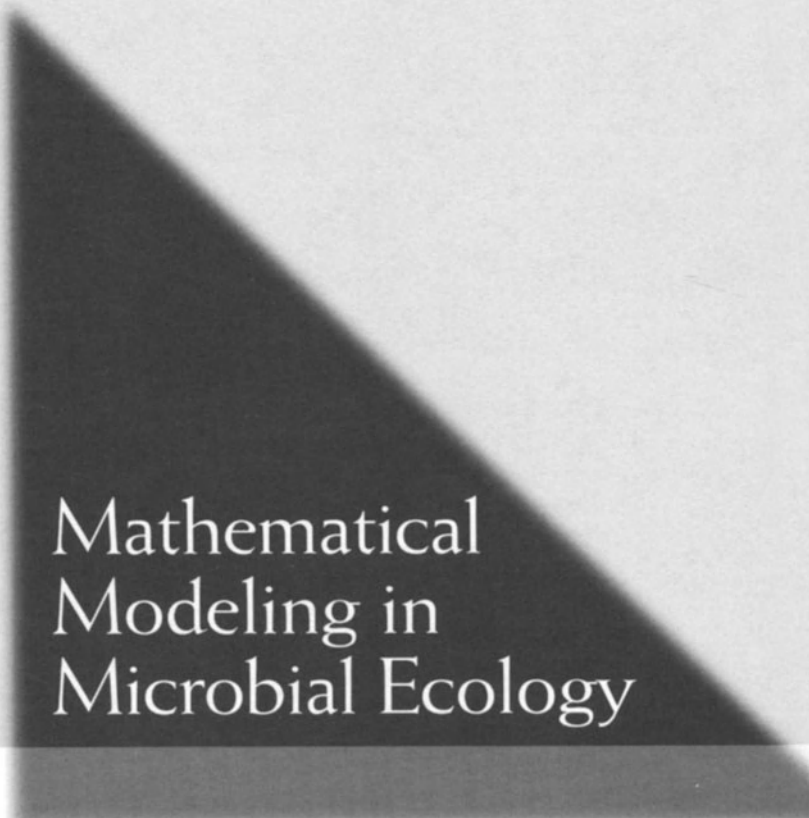
Timothy B. Parkin, National Soil
  Tilth Laboratory
USDA-ARS
2150 Pammel Drive
Ames, IA 50011

Joseph A. Robinson, Animal Health
  Product Development
7922-190-289
Pharmacia & Upjohn Co.
Kalamazoo, MI 49001

Daniel R. Shelton, USDA-ARS-
  NRI-ECL
Building 007
BARC-West
10300 Baltimore Avenue
Beltsville, MD 20705

Richard L. Smith, U.S. Geological
  Survey
3215 Marine Street
Boulder, CO 80303

Pierluigi Viaroli, Dipartimento di
  Scienze Ambientali
Universita di Parma
Via Delle Scienze
43100 Parma, Italy

Mathematical
Modeling in
Microbial Ecology

1

What Is Happening to
Microbial Ecology?

Arthur L. Koch

1. Introduction

This is a time in which very important advances are occurring on many fronts to change the field of microbial ecology. These include new experimental ways to measure and identify organisms in natural samples, methods and approaches for modeling rate processes in complex systems, ways of combining the various kinetic processes into comprehensible approximate equations and graphical representations, computer applications for the simulation of overall processes, and statistical methods for the reduction and analysis of data. We now have increased ability to analyze the microbe's *real* environment—the microenvironment. More important is the convergence of methodologies from the many and varied different field toward the multidisciplinary analysis of microbial ecology. Most importantly, although microbial ecology has been the invisible part of almost all ecosystems, it is crucial that ecologists come to recognize that more than half of any ecosystem is its microbial component.

At this time there are radical changes in microbiological ecology. At a number of levels new methods have been developed that can and are now being applied to both old and new topics in ecology. Past and current pollution of the environment is driving an interest, active work, and a need to understand the role of microbes in our biosphere and how the metabolic versatility of procaryotes can be modified to serve a key function in bioremediation. An appreciation of what has been tried and what has failed to solve environmental problems is now growing. This chapter lists some kinds of approaches that might work; some of these approaches are topics in this book.

This chapter, which presents an overview of the general problem, is divided

into four parts: Section 2 lists new analytical approaches and techniques, Section 3 discusses ways to model ecosystems, Section 4 reports developments in statistics; and Section 5 considers the use of Dutch uncle techniques to define and teach approaches that employ the scientific method to treat problems in microbial ecology. The rest of this book consists of chapters devoted to the current state of the art in different areas of microbial ecology with emphasis on the mathematical methods; the book as a whole; it is hoped, may help to lead to the future maturation of this very crucial intradisciplinary area.

2. Analytical Methods

Sometimes it is possible to focus on only a portion of a system and draw general conclusions about some particular process taking place in it. At other times it may be necessary to identify many of the types of organisms in an ecosystem and to ascertain the individual biomasses, the biological components, and the viability of the different species. Further, it may be necessary to estimate the collective living biomass and nonliving organic substances in the system. The partitioning of the biomass into living and nonliving is a very difficult feat. In many cases, the relevant chemical components of the ecosystem and the physical and geologic conditions must also be tracked.

There have been many advances in methodologies to establish, enumerate, and identify the organisms that grow and persist in various ecosystems. Better methods are becoming available to measure the pollutants, their metabolic products, and their metabolism in natural environments, which means that there are many new opportunities to analyze life processes in totality in a given habitat. This section just lists our new abilities and mentions a few of the virtues and limitations of these techniques.

2.1. Ribosomal RNA probes

Because the 16S ribosomal genes have now been sequenced for many organisms, it is possible to design probes to screen for broad and narrow groups of microorganisms in natural samples. This can be done even if the organisms cannot be cultivated. Probes can be labeled with radioactive tracers or they can be attached to fluorescent molecules. The latter strategy enables deeper and more detailed localization in both terrestrial and aqueous habitats. One goal of these procedures is to distinguish indigenous (autochthonous) classes of organisms from organisms that are not indigenous (allochthonous); for example, organisms introduced by man. The major impetus is to access microbial biodiversity. The merit of the these hybridization techniques is that organisms need not be cultured to carry out their identification; however, the question of the degree to which

the response to ribosomal probes actually represents living biomass, biomass that had been living a short time previously, or that which was living a long time beforehand is very important.

2.2. Polymerase Chain Reaction (PCR)

The power of the PCR is immense and its applicability is great and is growing. The availability of this technique means that extremely rare organisms in an ecosystem can be detected and identified in an analysis that can be carried out quickly. Once the gene has been amplified, part or all of the genome of the organism of interest can be quickly isolated for laboratory study. There are, however, many technical problems. Further progress is needed to improve ways to isolate essentially all the DNA of an ecological sample and purify it from the humic acid components that are a dominant part of many ecosystems and interfere with the PCR. There are many other problems with the PCR; most of these are understood, can be anticipated, and can be at least partially guarded against.

2.3. Fluorescent In Situ Hybridization (FISH)

With fluorescent-labeled probes, we can now target cells and identify the morphological shape of the organisms that contain the DNA that complements and hybridizes to the probe. This method also provides an opportunity to measure the distribution of a class of organisms within the ecosystem. Often patches or colonies of organisms are spottily distributed. Knowledge of such heterogeneity in space may be important to the proper analysis of a system.

2.4. Progress in Analytical Chemistry

Extensive progress has recently occurred in this field. Some of the advances depend on newer chemical methods, but others depend on the development of very sensitive physical methods. Particularly, the newer physical detection methods enable the measurement of classes and individual components rapidly and/or frequently in time and/or closely spaced in the physical system. High-performance liquid chromatography (HPLC) with different kinds of detectors and mass spectrographic methods with the many variations make it possible to learn very much detail very rapidly. A worker in this field has to deal with acronyms such as phospholipid ester-linked fatty acids (PLFA), gas chromatography (GC), Iontrap mass spectrometer (ITMS), mass spectrometry (MS), and electrospray (ES).

Among the classes of organic species, there are three groups of biocompounds that must be particularly highlighted. The first group consists of the phospholipids; these compounds are essential to the formation of the membranes of all cells. Polar lipids may be a measure of viable biomass because they are particularly

rapidly degraded. Because specific types of lipoidal molecules identify particular types of organisms, the measurement of these molecules and their degradation products can be very useful. The second class are the teichoic acids, which are constituents of the walls of Gram-positive bacteria. These are unique enough from their chemical constituents that it is possible to measure them by High Pressure Liquid Chromatography (HPLC) and deduce the level of Gram-positive organisms in the ecosystem; particularly this method has been applied to marine sediments. The third biochemical is adenosine triphosphate (ATP). By chemilumi-nescence methods, with luciferase and luciferin, it is possible to make ultrasensi-tive assays of ATP. ATP has a turnover time in living organisms of seconds; consequently the ATP level can be assumed to be a measure of something related to the totality of the living biomass in the sample. All three of these highlighted types of measurements, like those of many other chemical species, have advan-tages and limitations as measures of biological activity in the environment, and they all depend on the validity of the assumption made about their turnover in the environment.

2.5. Flow Cytometric Methods

Instruments of two different forms are commercially available for these meth-ods. The early flow cytometers were originally developed by biophysicists and engineers working in conjunction with the Hertzenbergs to help them explore some questions in immunology, but now the commercial instruments have a much wider sphere of applicability. Originally, with good signal-to-noise ratios, this technique could detect particles of the size of animal cells, but would fail with smaller cells, such as bacteria of the size of *Escherichia coli*. Instruments of this type are now better and can be aligned and operated so that they can be used to measure even very small bacteria. The second type of instrument is constructed with a commercial inverted microscope as an essential constituent part. It was designed specifically to be capable of counting and measuring bacteria. These two types of instruments, of course, have different limitations, but both have some of the same strong advantages. One important advantage of flow cytometry over other approaches is that it is capable of counting a large number of particles within a short time. This is done by detecting light-scattering pulses such that the intensity is related to the biomass of the particle passing in a flowing stream through the laser beam or illuminated by a high-pressure mercury arc. Simultaneously, fluorescent light from suitable probes can indicate some other parameter (or parameters) of the same particle.

In practice, bacteria are often stained with fluorescent molecules. By using a probe that binds to double-stranded nucleic acid (largely DNA), the lack of fluorescent signal can be used to identify and eliminate nonbiological particles. Such "dirt" particles are then excluded as not being biological particles. Additional

variation of these techniques can be used to estimate the internal pH and to estimate the viability of the bacteria.

2.6. Microelectrodes

Microelectrodes can be used to measure very local pH, redox, and other properties of natural environments. The development of these procedures has been extensive and they can now be considered available, reputable, and mature techniques. In many ecosystems, the pH and oxidation/reduction potential change micrometer by micrometer. Measuring these spatially changing relationships is key to understanding processes in many ecosystems. Although such spatial microheterogeneity is frequently ignored while analyzing ecosystems, generally it is where two adjacent or overlapping physical or chemical gradients meet that biomass formation is greatest and of most interest. In fact, it is the core fact of ecology that ecosystems flourish at the boundaries where one environment joins another. For microbial ecosystems, these boundaries are just as important as the interface where rivers run into oceans bearing nutrients and pollutants, but the relevant distance scales frequently are only micrometers instead of kilometers.

2.7. Fumigation Incubation (FI) and Fumigation Extraction (FE)

This technique is now becoming popular and, to some degree, has been quantitatively justified. The execution of it is simple: The environmental sample is treated with excess chloroform. This substance destroys living cells, but leaves enzymes functional so that they can degrade the biomass of the previously living cell. After 10 days the biomass has been reduced, possibly eliminated, and a corresponding amount of CO_2 is liberated. It seems surprising that such methods work at all, but they apparently do work surprisingly well, although there are many detractors among microbial ecologists. Some believe that fumigation does not kill all the bacteria and that the living cells mineralize the dead cells. With more understanding of the capabilities and the pitfalls, these methods may become the workhorses of microbial ecology.

2.8. Fluorescent Antibodies and Other Probes

Monoclonal antibodies, labeled with fluorescent molecules, can be sufficiently specific to identify and localize particular bacteria in a natural sample. Probe molecules can contain biotin or many other tags that can be used with ingenious techniques to sense the processes going on in these microcosms.

2.9. Confocal Microscopy

Together with old-fashioned staining, now phase, differential interference contrast (DIC), and fluorescence microscopy offer important tools. The most recently added and a very powerful tool is confocal microscopy. It enables one to focus on a narrow layer within a thicker sample and to do many other elegant tricks.

2.10. Modern Versions of Old-fashioned Microbiological Methods

Because the shapes of bacterial cells and their appendages are the only morphological characteristics readily accessible, more characteristics were needed as microbiology developed; it was the metabolic capabilities that became critical for taxonomic purposes. Today a variety of multitests are available. The most extensive system is BIOLOG®, which uses many test wells, has an elaborate computer backup, and can be used to quickly identify strains from nature.

3. Kinetic Aspects

It is not enough to measure the biology and chemistry of an ecosystem at an instant of time, it is also necessary to measure how the system progresses with time. The time in question includes the seasons and the time since humanity made changes in that environment. There are only a few conditions in nature where the biosystem approaches a steady state. When this does occur it is valid to assume that all the fluxes—chemical, geologic, and biological—in the ecosystem are constant and production and degradation balance. Then it is reasonable to measure the time-independent concentration of components and infer the rates of reaction based on the hypotheses that the net fluxes within the ecosystem creating and destroying every component and organism are equal. This, however, is only of a special and limited occurrence. More generally, the amounts of various components change with time, and from the time-dependencies (or in some cases concentration dependencies) information it becomes possible to measure the properties of the ecosystem. As in the previous section, I just list (inclusively) the general categories of phenomena that must be taken into account.

3.1. Growth Kinetics

Growth includes the utilization of environmental resources and the production of cell components, such as proteins, nucleic acids, etc. It also includes the production of waste products and the production of resources utilizable by other biological components of the ecosystem. The analysis of ecosystems involves many types of kinetic processes. The primary kinetic concepts are the first-,

second-, and, occasionally, third-order chemical reactions. For these, the reaction rate depends on the products of the concentration of one, two, or three constituent species or groups of species in the system with the appropriate numerical rate constants. Using only such terms then, the integration of the full rate expression with appropriate boundary conditions leads to a mathematical description of the development of the system.

Quite often cellular growth in isolated laboratory conditions depends on auto-catalytic kinetics (exponential growth). This occurs when the rate of formation of a component is proportional to the concentration of that component. But the simple exponential expression is sufficient only in the most elementary approximate treatments of an ecosystem. Frequently it is necessary to complicate and expand the formulation that led to the autocatalytic model to include lag phases, the approach to stationary phases, death phases, and many others. One hundred and twenty years ago, the approach to stationary phase would have been modeled according to the Verhulst model (i.e., the logistic model). This symmetric curve clearly is unsatisfactory for microbial cases, as first shown by Jacques Monod, but the Verhulst formulation is still the theoretical backbone of the ecology of higher plants and animals. Kinetic treatments of microbial populations, certainly of monocultures, could be approximated by the integrated Monod model. In the differential form, the Monod law is a hyperbolic equation identical in form to, but different in significance from, the Michaelis-Menten equation of enzyme kinetics. This Monod law relationship is the most used equation in microbial ecology. Possibly more relevant is the Blackman model, and even more realistic than either the Monod model or the Blackman model is the Best model. The Blackman and Best formulations enable a more accurate estimation of the growth of organisms in natural environments. I suggest switching to their use.

3.2. Diffusion

Much of biology depends on the kinetic process of diffusion. This is particularly relevant to microbiology because microbes may not be motile and certainly do not have kidneys, hearts, and lungs and therefore have few alternatives to move themselves or their internal parts and components around. Analysis of diffusion in the environment, through the cell wall (including facilitated and active transport) has occupied a good deal of interest in microbial physiology and gradually is becoming more appropriately taken into account in microbial ecology. An area of current interest has to do with the influence of glycocalyxes and capsules in isolating and separating organisms from other organisms (e.g., potential predators) and from the effects of convection and stirring in the environment. As important is the effect of grouping or aggregation of microorganisms interfering with uptake of each other and of the positive and negative factors associated with a microorganism inhabiting a nonbiological surface. A cell so attached may experi-

ence a different physical environment (such as a different pH), have a different supply of nutrients, cofactors, inhibitors, etc.

3.3. Intermediary Metabolism

For our purposes, the metabolism of microbes can be divided into three domains. First is metabolism for energy generation through a variety of largely understood pathways of intermediary metabolism. Most often, consideration of the availability in any environment of substances that can be used to generate biologically useful energy is the dominant ecological problem. Categorizing bacteria on the method of energy trapping is high on the list of topics in any microbiology text. Choice of resources to be converted into biomass is often a close second topic of interest and occasionally it is the prime problem. Natural ecosystems start with plant materials, which may be passed through animals, insects, and protozoa, before they are finally metabolized through a series of fungi and procaryotes. Thus successively different members of the microbial part of the ecosystem work and rework the biomass until it is all mineralized (assuming that current coal and petroleum generation is completely negligible in today's world). The farther along the food chain the biomass moves, the more recalcitrant are the residual molecules and the more drastic the methods required to metabolize it further. Especially significant in today's environments are the recalcitrant pollutants made by humans; some of these are very difficult to metabolize for evident chemical and biochemical reasons. Sometimes in a metabolically active ecosystem these recalcitrant molecules may be consumed by what is termed "secondary substrate metabolism," which is the term used if the bioprocesses of normal metabolism can also handle the pollutant.

Another process, categorized as "cometabolism," occurs when an enzyme or enzyme system is generated by an organism for one purpose and incidentally can act on the pollutant. This difference is subtle, but important. Cometabolism is especially significant for the process of transformation, or possibly, destruction of aromatic and chlorinated compounds. In many cases, the amount of resource that the actual pollutant represents is negligible from the point of view of the generation of microorganismal biomass. Thus only if there is some more massive resource that can facilitate the destruction of the pollutant by either of these two processes can degradation of a pollutant be accomplished.

3.4. Geologic and Hydrologic Considerations

In soil ecosystems, the key problem that needs to be appreciated and continuously considered is the movement of water through them. Availability, percolation, state of aeration, and flow are important. The fluxes of water, either downward or laterally, bear with them the chemical resources for the development of

biomass or bring kinds of biomass than then are remodeled. For terrestrial bio-masses, it is necessary to consider the geologic makeup of the soil, its components, the surface area of the particles, their absorptivity, and their exchange properties. How water percolates through the soil, how air moves through it, how substances differentially bind to the soil components, and the location of microbes in the soil particles are all important. Soil particles may be very complex structures. An important point, and one frequently overlooked, is that in the neighborhood of charged surfaces, where inorganic molecules are adsorbed, the effect of the surface charges on organic molecules, or living cells, can be very large, and can, for example, alter the effective pH by as much as several units.

4. Principles of Kinetic Modeling

This section presents a qualitatively different kind of list than is presented in Sections 2 and 3. It is comprised of stages or steps that need to be considered in analyzing a microbial ecosystem. I formulate this in the second person. You have a problem, maybe you have some data, you would like to be able to describe the relevant phenomenon and hopefully estimate some important parameters. You hope that the values of the parameters and the qualitative conclusions will be significant to more than just the few examples for which you have data and that you have studied and fitted. What shall you do? A clear and critical formulation of the detailed situation together with the appreciation of the general problem is the first, middle, and last step. All your background should be brought to bear, but that is not enough. Use friends, colleagues, biologists, and other experimental scientists mercilessly. Utilize statisticians; their knowledge of the analytical meth-ods of data reduction and their ability to estimate the validity of conclusions can be most helpful and usually is essential. If you are an experimentalist, being associated with them is critical to the reliable and speedy solution of the type of problems you have likely encountered. On the other hand, often the statistician has difficulty in appreciating the biological problem and he or she is of no help when the problem is not understood. Sometimes this is actually the statistician's fault, but in any case, you (the experimentalist) are at fault and must not have clearly explained and formulated the situation and the biological principles.

Formulate a mathematical model that is more general than your favorite intu-ition permits. Then you may be able to exclude models instead of just showing consistency with your gut feeling. This is key to the analysis. Science is no good if it does not exclude certain hypotheses that are reasonable *a priori*. Make a decision as to whether to require that the fitting procedure estimate a given parameter or, as an alternative, choose its value from the literature or from other experience; thus the other parameters can be more critically and accurately determined. As far as possible, you should adopt constants from outside your

data to be inserted into the model; choose those that are reasonable and are known to apply in other cases (later you will check to see whether this was a reasonable move). Use mathematical insight (yours or somebody else's) as to what parameters are only marginally relevant. Look at the mathematics and choose which factors to lump, which to neglect, and which to treat as being critical. Fit the model to the data. Compared with an earlier time, this fitting is now trivial because of the speed of personal computers and the availability of application programs that can be applied directly to a vast variety of cases.

Having carried out the preceding process, do not think that the project is finished. Indeed, you are not done, you have only begun. Besides restating your model with the perspective of having fitted the data, now it is incumbent on you to isolate the key factors, identify what circumstances control their variability, and reconsider whether the points that you glossed over really should have been given short shrift. Now is the time to see if the other models have been falsified by the data and now can be excluded and discarded.

5. Progress in Statistical Methods

5.1. Variability Varies from Situation to Situation

Physicists, chemists, and molecular biologists can often count on precise methods so that the experiments that they carry out do not have to be replicated either often or at all. Their measurement methods may be very precise, the experimental system accurately defined, and the experiment may be always reproducible. But often, even in these fields, this presumption of precision and accuracy is not wise or valid because many systems are much more complicated than intuition or hope may have presumed. On the other hand, no one thought that a vast range of biological situations would be trivially constant. Because of their inherent variability, it is not surprising that statistics emerged from the study of agricultural problems and in attempts to understand the abilities and behaviors of samples of people for eugenics purposes. Experimental microbial ecology is usually closer to the latter camp than to a rigidly reproducible field of "hard" science. This means that experiments must be repeated more often; but they must be repeated in a reasonable way to assess the variability appropriate for the purposes of the study. The Latin square arrangement is appropriate when the source of variability is not known. Pairing the treatments to give equal weight to known or assumed sources of variability, such as using equal numbers of males and females, has much merit. An alternative approach is to use only one sex, one strain, one growth environment, etc. However, then it is necessary to point out that although the conclusion is clearer and more valid, it has a restricting range of applicability. There are a large number of statistical tests that can compare two or more groups.

These methods must be used correctly and appropriately. But most important is the assessment and the validation of when variability from other causes is large or small.

5.2. Mathematical Form of Models for Processes

Most often one assumes (rightly or wrongly) that an adequate number of groups, treatments, strains, or conditions have been used and that the average values will serve to factor out the underlying biology. The usual question is "What is the temporal course?" or "What is the concentration dependence?" or "What is the role of the type of product being tested in a plus and minus fashion?" Here is where my opinion commonly diverges quite a lot from that of many professional statisticians. Rather than choose a model because it has good statistical properties, I insist on a model that is biologically relevant from a global (biological) point of view. Fitting the data to a hyperbola, as in the Monod equation, may be slightly in error, but it has the global biological features that (1) if there is no substrate, there will be no product and (2) if there is a very large substrate concentration, only a certain maximum level of product formation would be formed. If one substitutes infinity into a formula that is supposed to model and describe a biological situation and the result is plus or minus infinity or a negative number, then it is "no go," and one should go back to the drawing board. Power laws and various formulas or transformations of the data often correspond to an irrational and inappropriate biological model at extreme values of independent variables. One must insist that substitutions into the fitted expressions of zero and infinity and other conditions outside the range of the actual experiment make reasonable sense.

5.3. The Necessity of Experimenter and Statistician Planning Together before the Experiments are Conducted

Forethought is the most critical shortage in science. I tell my students that a day at the beach and lolling in the sun can speed the research immeasurably; that is, if as they lolled, (1) they were to think through the experiment to be performed or the data to be collected; (2) they were to think of the possible outcomes and what additional data will be needed in such cases to have a valid, interpretable, unambiguous result; (3) they considered that they should, and must, anticipate the necessary controls to verify conclusions; and (4) they realized that they should have tried to anticipate what part of the proposed experiment should be tightened down and then tighten it down before conducting the experiment or sampling a natural ecosystem. Students should then consider revising the experimental design so that certain possible sources of variability are eliminated before they can become a questionable factor that causes the final conclusion to

be doubted. In addition to the other merits of lolling, giving one's time to anticipating contingencies and then seeking all the intellectual help you can to guard against possibilities is important. This is so even though you know that this will not be enough. But statisticians are generally past masters at this type of game. Moreover, when the experimenter has to explain and rationalize what is to be done and tested to another willing, sentient human being who will think with him or her, then the holes in the protocol become clear. With adequate forethought, science will move much faster and more surely.

One can wonder how many highly productive individuals and research groups were able to create and analyze a phenomenon so effectively. I'll consider here only one creative individual: Max Delbrück. Truthfully he never did much, but he is the father of molecular biology, largely because he made grunts and criticized his colleagues and students because he could, or would, see what these other had not, or could not, see. He would notice that inclusive thoughts had not been formulated and that generalization was needed before the research experiment was to be conducted or the data collected. By the way, he taught me what a standard deviation was from first principles when I had only the foggiest idea of statistical concepts.

My next learning experience in statistics came from Snedecor's book. This book was awesome. Although I knew nothing about corn, pigs, or chickens, he propounded the ideas that enabled me (and many others) to understand how to deal with real problems. I hope the present book will serve as a primer for the next generation of researchers in microbial ecology no matter how their predilection bends them between the execution of experiments or the entirely different question of what the results mean after they have been obtained.

6. Conclusions

Microbial ecology is now undergoing a convergence of the developments of its own character and by the infusion from developments in other field. The same process has taken place for human biology. It is worthwhile noting the rays of that convergence. These fields include progress in biochemistry, analytical chemistry, molecular biology, medicine, genetics, statistics, mathematics, and the sequence of the human genome. I want to focus on the cooperative role of these disciplines. Medicine may produce the naturally occurring mutants; molecular biology supplies the markers for finding genetic regions in all parts of the human genome; genetics proscribes how pedigrees can be used; and finally, a new kind of mathematics, only a decade or so old, enables us to find the genetic target of interest. It is possible to find a gene that is part of a many gene system

or variable in its expression (multifactorial) for a disorder that is poorly defined and poorly recognized. In spite of these difficulties, when the rest of the procedure involves up-to-the-minute technology, it can be done. In this human case, all this has happened because the need to know is there. The same thing is happening, or is about to happen, in microbial ecology.

2

Modeling Microbial Processes: An Overview of Statistical Considerations

Joseph A. Robinson

1. Introduction

The development of a mathematical model is often a goal of experiments in which a dependent variable is measured as a function of other explanatory or independent variables (e.g., time, substrate concentration, dilution rate, pH, temperature). Models of interest to microbial ecologists run the gamut from linear ones, such as those used in linear regression and the general mixed linear model analysis of variance, to models that are nonlinear in their parameters or simply put, nonlinear models. Examples of nonlinear models include one-term exponential decay, Michaelis-Menten, Monod, and variants of these simpler equations. The methodology for fitting data to linear models is described in a multitude of statistics textbooks, for beginners to advanced practitioners alike, and is not dealt with herein. Rather, the focus of this chapter is the fitting of data to models that are nonlinear in their parameters.

Data are fit to nonlinear models using regression or nonlinear parameter estimation (NPE) methods. The purpose of this chapter is to highlight the theory of NPE and to illustrate the possibilities offered through use of the statistical methodology. This chapter is not a substitute for more thorough treatments of NPE such as those provided by Bates and Watts (1988), Ratkowsky (1983), and Ratkowsky (1990). The reader is advised to consult these texts for details on implementing the methods and assumptions of NPE.

NPE has its roots in regression theory, although many of the statistical results are approximate (Bates and Watts 1988; Draper and Smith 1981; Ratkowky 1983). The degree to which linear regression results are approximated depends on the combined effects of the nonlinear model being used and the dataset to be

fit to the model (Ratkowsky 1983). The more nonlinear the nonlinear model dataset combination, the more approximate the results. The uncertainties are compounded by the fact that there exists literally an infinite number of nonlinear models of potential interest. The practical result of this situation is that the methods of NPE should not be relied on to lead to a single mathematical function that best describes the ecological system of interest. For this reason, Ratkowsky (1983) argues that since many models can be found to describe a given dataset equivalently, it is important that models that exhibit close-to-linear behavior should be preferred over those with far-from-linear behavior, unless practical reasons dictate otherwise. The use of models with close-to-linear properties helps to ensure that the results of NPE more closely approximate linear regression results.

More often than not, the nonlinear model builder is left with a subset of models that fit the data equivalently. Practical criteria (i.e., what makes scientific sense) must then be used for model discrimination and/or selection. This situation is particularly true for nonlinear models that are not simple subsets of one another (i.e., non-nested), because the number of methods available for discriminating between competing models is limited or nonexistent. As Bates and Watts (1988) point out, "if there are scientific reasons for preferring one model over the others, strong weight should be given to the researcher's reasons because the primary aim of data analysis is to explain or account for the behavior of the data, not simply to get the best fit." Alexander (1994) extends this concern by asserting that it "is not always possible to distinguish among models, even with nonlinear regression techniques, so that the final choices among models are arbitrary." Thus, although this chapter closes with recommendations for statistically distinguishing between competing models, the recommendations should not be adopted blindly without consideration of their limitations and intended scope.

2. Model Identification versus Discrimination

Model identification is the process of fitting data to the model(s) of interest. The objective is to estimate the parameters of the model given a dataset. Model identification is sometimes referred to as the inverse problem (Robinson 1985). The use of NPE to estimate the parameters of the differential form of the Michaelis-Menten model, given initial velocity data, is an example of the model identification or inverse problem. In contrast, model discrimination is the process of eliminating candidate models to derive the subset of models (ideally one) that best describes the fitted dataset. The problem of model discrimination can be solved by applying various statistical criteria, although in some cases (e.g., non-nested models) appropriate methods may be lacking or even nonexistent (Bates and Watts 1988).

The problems of model identification and discrimination may be investigated

before or after collection of a dataset. Issues relevant to model identification are presented first in this chapter, since models that cannot be identified should not be considered in the model discrimination phase of model building. Model discrimination is discussed second, and two sets of criteria are proposed: one for comparisons between models that are nested and one for models that are not nested.

2.1. Model-Building Goals

The approach taken to mathematically represent microbial processes in the form of regression models depends on the researcher's goals (Fig. 2.1). In many situations, the experimenter has already decided to use a model which is identifiable and is supported by experimental evidence. One example is the pseudo-first-order rate model used by investigators to estimate a second-order rate coefficient for biodegradation (Paris and Rogers 1986). Such a model may be applied without considering whether the dataset collected supports an alternative equation. This assumes, of course, that the dataset generally supports the form of the model. One would not choose a model whose mathematical behavior was grossly at

Approaches to Model Building

1. **Collect initial data set**
2. **Fit a priori identifiable models**
3. **Select a subset of "good" models using a posteriori methods**
4. **Design discrimination experiments using a priori methods as aids**
5. **Collect discrimination data set**
6. **Select "best" model using a posteriori methods**
7. **Design optimal experiments for parameter estimation**
8. **Collect data for precise estimation of parameters of best model**
9. **Analyze data and summarize parameter estimates**

Complete characterization; precision of θ important

Partial Characterization; precision of θ less important

Exploratory work; routine estimation of θ

Figure 2.1 The model-building process. The number of steps completed depends on the researcher's goals. The algorithm shown is meant to highlight concepts and not to serve as a strict recipe to be followed. Theta (θ) is the 2×1 vector that includes the initial substrate concentration (S_0 and the first-order rate coefficient (k_1) for the one-term exponential decay model.

odds with the data. As another example, Michaelis-Menten-type models are often used to assess the affinity of microbially catalyzed processes (Robinson 1985). In both of these situations, it is the routine estimation of a parameter such as the second-order rate constant or a K_m that is the goal; deciding whether or not the data support other models is not the goal.

Model discrimination experiments (Fig. 2.1) can be designed and conducted when a more thorough characterization of the microbial process is of interest. Once the best model is selected (assuming this can be done), experiments can be designed to estimate the parameters of the best model with the highest possible precision (i.e., lowest variance) (Fig. 2.1). Experiments to estimate the parameters of a model with the highest possible precision should follow the discrimination phase (Fig. 2.1), and designing these optimal experiments is discussed later in this chapter. The approach shown in Figure 2.1 is not intended to be a recipe to be adhered to strictly. Rather, it suggests the elements of an approach that can be used in different phases of the model-building process. The value of the algorithm depicted in Figure 2.1 lies in the concepts identified not in its use as *the* only means of building parsimonious mathematical representations of microbial processes or systems.

2.2. Limits to Model Discrimination

It is often unrealistic to expect a single model to emerge from a collection of equations as the best one for a given dataset. This dilemma stems from the fact that it is easy to detect model failures, but a model is not correct merely because it provides a good fit to a dataset. Practically, one can use statistical criteria for model discrimination up to a point; beyond that, nonstatistical information regarding the ecological system or process of interest must provide the answer. The point is again that, in many cases, it is the researcher or subject-matter specialist who must decide between competing regression models using information outside of that obtained from the NPE analyses (Bates and Watts 1988).

3. The Least-Squares Criterion

The ordinary (unweighted) least-squares criterion generally is used when data are fit to either linear or nonlinear regression models (Bates and Watts 1985; Draper and Smith 1981; Robinson 1985). This criterion is satisfied when a set of unique parameter estimates are found that minimize the sum of the squared errors about the best fit curve. Other criteria such as weighted least-squares and maximum likelihood exist, but they are used less frequently, presumably because of lack of specific knowledge regarding the error structure in the data (Robinson 1985).

For models that are nonlinear in their parameters, initial estimates of the parameters are needed to minimize the sum of the squared errors. This situation does not exist for linear regression models. The need for initial estimates for NPE and means of obtaining these for some models of microbiological interest have been highlighted elsewhere (Robinson 1985). Most computer programs in current use for fitting data to nonlinear models allow the user to either specify initial estimates or employ a routine for searching a grid of parameter estimates to obtain more reliable starting values for the NPE routine. For some nonlinear models, success or failure in minimizing the sum of the squared errors is largely dependent on how close the initial estimates of the parameters are to the best estimates. This situation is another reason why nonlinear models with close-to-linear behavior are preferred over models with far-from-linear behavior (Ratkowsky 1983).

3.1. Weighting

Weighted least-squares analysis can be used when the variance of the measurement errors depends on the magnitude of the measured variable (Li 1983; Robinson 1985; Bates and Watts 1988; Silvert 1979). The exact form of the weighting function, i.e., whether the weighting function should be $1/Y$ or $1/Y^2$ may have little effect for some models. One means of checking this assumption is to use various weighting functions in the range from $1/Y$ to $1/Y^2$ to assess the extent to which the values of the estimated parameters are influenced by the value of the exponent in the denominator of the weighting function. Practical experience suggests that, when weighted least-squares analysis is appropriate, either $1/Y$ or $1/Y^2$ is adequate. However, the researcher is urged to verify the appropriateness of this recommendation for the model dataset combination being studied. One should not ignore means of incorporating information on the error structure of the data, when it is known, into the routine used to minimize the sum-of-squares function.

4. Model Identification

4.1. A Priori *Identifiability*

A model is identifiable if the parameters in the model (such as V_{max} and K_m in the differential form of the Michaelis-Menten equation) can be uniquely estimated from a dataset. If unique parameter estimates cannot be obtained for a nonlinear model dataset, then little (or no) statistical significance should be given to their estimated values (Robinson 1985). The following nonlinear model is not identifiable (based on a single substrate depletion curve):

$$S = S_0 \times \exp\,[-(V_{max}/K_m) \times t]. \tag{1}$$

This equation is the integrated form of the Michaelis-Menten model, but only for the first-order region of substrate concentration. It describes S as a function of time, t, given the initial substrate concentration, S_0. The ratio V_{max}/K_m is equivalent to a first-order decay coefficient in a one-term exponential decay equation. Model (1) is not identifiable because an infinite combination of V_{max} and K_m values can describe the same substrate concentration-versus-time curve. Thus, it is impossible to estimate these two parameters simultaneously from a single substrate depletion curve that exhibits first-order decay behavior.

In contrast to Equation (1), the integrated form of the Michaelis-Menten equation, solved for the complete range of substrate concentrations, can be identified. This implicit nonlinear equation has the following form:

$$V_{max} \times t = (S_0 - S) + K_m \times \ln\,(S_0/S). \tag{2}$$

If V_{max} is decreased by 50%, for example, the same substrate concentration-versus-time curve cannot be produced by decreasing K_m (within biologically meaningful limits). Thus, unique estimates of V_{max} and K_m may be obtained theoretically by fitting substrate depletion data to Equation (2).

4.2. Sensitivity Equations

Sensitivity equations provide an analytical assessment of whether or not a nonlinear model is identifiable. These equations describe how sensitive the measured variable (such as substrate utilization in the case of the Michaelis-Menten equation) is to changes in each of the parameters of the model (Robinson 1985). If the solutions to the sensitivity equations are proportional to one another, then the model cannot be identified regardless of how precisely the dependent variable (i.e., initial velocity) is measured (Robinson 1985). In other words, if one of the sensitivity equations for a two-parameter model can be converted to the second sensitivity equation simply by multiplying by a constant, then the model is not *a priori* identifiable.

Sensitivity equations also provide a definition for what constitutes a nonlinear model (Robinson 1985). When the sensitivity equations depend on the parameters of a model, then the model is said to be nonlinear in its parameters. In contrast, the sensitivity equations of linear models do not depend on the model parameters, only on the independent variable(s). A model is nonlinear even if only a single parameter in the model is a function of one or more of the other parameters in the model.

Consider the sensitivity equations for the first-order version of the integrated Michaelis-Menten model (Equation [1]),

$$d\ln(S)/d(V_{max}) = -(1/K_m) \times t, \qquad (3)$$
$$d\ln(S)/d(K_m) = -V_{max} \times t. \qquad (4)$$

These equations describe the sensitivity or rate of change (first derivative) of the dependent variable (substrate concentration, S) to changes in the parameters of the nonlinear model, V_{max} and K_m in this case. Equation (4) can be obtained from Equation (3) if Equation (3) is multiplied by $V_{max}K_m$. Thus, Equations (3) and (4) illustrate that unique (statistically reliable) estimates of V_{max} and K_m can not be obtained by fitting substrate depletion data to Equation (1).

In contrast to Equation (1), the sensitivity equations for the integrated Michaelis-Menten Equation (2) support an interpretation that the parameters of this nonlinear model can be uniquely estimated. These equations have the following forms:

$$dS/d(V_{max}) = t/[1 + (K_m/S_0)], \qquad (5)$$
$$dS/d(K_m) = -\ln(S_0/S)/[1 + (K_m/S_0)]. \qquad (6)$$

Neither Equation (5) nor Equation (6) can be obtained from the other by multiplying by a constant. Thus, the sensitivity equations for the integrated Michaelis-Menten model support an interpretation that unique estimates of V_{max} and K_m can be obtained by fitting substrate depletion data to Model (2). Whether unique parameter estimates can be obtained in practice is another matter. The latter depends on the measurement errors, and is the subject of the next section of this chapter.

4.3. A Posteriori *Identifiability*

Often a model that is *a priori* identifiable cannot be identified after data are collected. This is because dependent variables are not free of error. Noise in data is the bane of modelers everywhere. This issue can be illustrated by considering the following expression:

$$\text{Response variable} = \text{Mechanistic model} + \text{Variance model.} \qquad (7)$$

In Equation (7), the response variable is the measured quantity (substrate or product concentration, growth rate, etc.). The mechanistic model could be Equation (2), for example. The variance model is generally unknown and practically impossible to identify, and if it dominates the mechanistic model, then the parameters of the mechanistic model can not be uniquely estimated regardless of whether a model is *a priori* identifiable. The only solution in this case is to improve the precision with which the dependent variable is measured.

4.4. Limits to Identifiability

   There is a practical limit to the number of parameters that can be estimated from measurements on a single response variable. For example, consider the following equation:

$$[S] = S_0[\exp(-k_1 \times t) + \exp(-k_2 \times t) + \exp(-k_3 \times t)]. \qquad (8)$$

Equation (8) describes the disappearance of a substrate, S, resulting from three separate exponential decay processes as a function of time, t. Equation (8) has a total of four parameters, including the initial condition, i.e., the initial substrate concentration. If k_1, k_2, and k_3 equaled 1000, 100, and 10, respectively, then this model could likely be identified from a single substrate depletion curve. However, if these parameters equaled 5, 2, and 1, then it is unlikely that unique estimates of the parameters could be obtained from fitting Equation (8) to substrate concentration-versus-time data. In the latter case, a single first-order decay equation would accurately summarize the data. This result can be extended to models that are functionally different from Equation (8) but have a comparable number of parameters. In practice, it may not be possible to identify a nonlinear model that has four or more parameters from measurements made on a single response variable. This practical limit depends on whether the model exhibits close-to-linear versus far-from-linear behavior. Of course, the practical limit of how many parameters can be uniquely estimated depends on the magnitude of the measurement errors as well.
   The integrated version of the Monod equation is an example of a model that approaches the limits of identifiability (Robinson 1985; Robinson and Tiedje 1983). This model, equivalent to the integrated Michaelis-Menten model, is an implicit nonlinear function and it has the following form:

$$\mu_{max} \times t = C_1 \times \ln\{[Y(S_0 - S) + X_0]/X_0\} - C_2 \times \ln(S/S_0), \qquad (9)$$

where $C_1 = (YK_s + YS_0 + X_0)/(YS_0 + X_0)$ and $C_2 = (YK)_s/(YS_0 + X_0)$.
   Estimates of μ_{max} (maximum specific growth rate), K_s (half-saturation coefficient for growth), Y (yield coefficient), S_0, and X_0 (initial biomass concentration) obtained using Equation (9) may be suspect because the sensitivity equations are nearly proportional to one another (Robinson and Tiedje 1983). Some nonlinear models proposed to describe the kinetics of degradation of xenobiotics have five or more parameters (Alexander 1994; Brunner and Focht 1984; Schmidt et al. 1985; Simkins and Alexander 1984). These models can be difficult to identify statistically unless the number of data points is large (e.g., four to five times the number of parameters) and the response variable is measured with a high degree of precision. Alexander (1994) has succinctly described the difficulties associated

with the use of these multiparameter biodegradation models and pointed out the arbitrariness associated with attempts to distinguish one from another using substrate depletion or product formation data alone.

One means of improving the identifiability of a nonlinear model is to make it less nonlinear, i.e., reparameterize the equation into a form such that it exhibits more close-to-linear behavior. Ratkowsky (1983) presents a thorough treatment on how to assess the nonlinearity of a nonlinear model using Monte-Carlo methods and transformations (reparameterizations) that can improve the identifiability of a nonlinear model. In a second treatise, Ratkowsky (1990) provides an excellent compilation of various nonlinear models with recommendations on which ones to use, based on consideration of their intrinsic and parameter-effects-associated nonlinearities.

As an illustration of Ratkowsky's approach (1983), Equation (9) can be reparameterized to a model that exhibits improved estimations properties. This can be shown by considering first the differential form of Equation (9) and a reparameterized version. The differential form of equation (9) is

$$dS/dt = -[(\mu_{max} \times S)/(K_s + S)] \times \{[Y(S_0 - S) + X_0]/Y\}. \tag{10}$$

Equation (10) can be reparameterized (algebraic rearrangement) to yield the following differential equation:

$$dS/dt = -[k_1S^2 - (k_1S_0 + k_2) \times S]/(k_3 + S), \tag{11}$$

where $k_1 = \mu_{max}$, $k_2 = (\mu_{max}X_0)/Y$, and $k_3 = K_s$. The integrated form of Equation (11) is

$$k_1 \times t = [k_3/(S_0 + k_2)] \times \ln\{[(S_0 + k_2 - S) \times S_0]/(k_2S)\} \tag{12}$$
$$+ \ln[(S_0 + k_2 - S)/k_2].$$

Equation (12) exhibits less nonlinearity than Equation (9) because it has one fewer parameters. With fewer parameters, Equation (12) will exhibit more close-to-linear behavior, depending on the dataset fit to it. The only disadvantage here is that independent estimates of X_0 and Y can not be obtained by fitting substrate concentration-versus-time data to Equation (12), however, an improvement in the estimation properties of this model is probably a greater advantage than being able to obtain unique estimates of the original parameters (μ_{max}, K_s, Y, S_0, and X_0).

Microbial ecologists need to strive for simplicity in attempts to mathematically define processes and systems of interest because of the aforementioned limits to model identifiability. Adding more and more terms to a nonlinear model can only lead to an equation that can not be identified. Such a model is a poor

hypothesis about the way an ecosystem is presumed to function, particularly if it is based on measurements of a single variable. The need for parsimonious models is reinforced when the model discrimination problem is considered next. It is in this setting that the value of Occam's razor (plurality or complexity should not be assumed without necessity) is most evident.

5. Model Discrimination

A thorough modeling investigation should include a model discrimination phase (Fig. 2-1). To assume that a given dataset can be ascribed to a single nonlinear model is wishful thinking. As noted in the Introduction, there are essentially an infinite number of nonlinear models that may be of potential interest. Statistically defensible models can be developed through the use of model discrimination criteria (Bates and Watts 1988; Robinson 1985). But statistical criteria alone should not be relied on to lead to a single nonlinear model, particularly in the case of nonlinear models that are not nested within one another. The statistical criteria only inform. The researcher should decide which model fills the need for a mathematical description of the process of interest (Bates and Watts 1988).

5.1. A Priori *Discrimination*

The model discrimination problem entails answering the following question: How do we decide which model of several candidate models best describes a given dataset? An answer to this question can be approached prior to collecting data by plotting discrimination functions (Mannervik 1981). An initial dataset is needed since the model discrimination functions will depend on the estimates of the parameters of the competing models. However, guesses of the parameter values may be sufficient at this stage. The simplest type of discrimination function equals the absolute value of the differences in the predicted values of the response variable for two models (Mannervik 1981). The absolute difference between competing models can be compared against what is known regarding errors associated with measurements of the dependent variable. This process is particularly useful for deciding whether differences in models can be discerned in future studies.

Plotting discrimination functions underscores the difficulty of statistically discriminating between multiparameter models hypothesized to describe microbial processes or systems. As was true for model identification, discriminating between competing nonlinear models becomes more difficult as the number of parameters increases. Practically speaking, measuring the response variable precisely enough to distinguish between models with more than four or five parameters may be

difficult, unless the models describe dissimilar types of curves (e.g., S-shaped versus exponential).

Discrimination functions, although useful, suffer from the same limitation suffered by *a priori* methods of assessing model identifiability. The discrimination function ignores the variance model. If the variance model dominates the mechanistic model then discriminating among competing models may not be possible. Another limitations is that the number of discrimination functions that must be considered, in a pairwise fashion, grows rapidly as the number of nonlinear models of interest increases. But despite these limitations, the examination of discrimination functions can offer insights into predictions made by dissimilar models.

The following example illustrates an extreme situation in which dissimilar processes lead to models that are functionally equivalent. Consider two variations on the integrated Michaelis-Menten model, Equation (2). The following differential form of Equation (2) incorporates a term describing a threshold effect at which the rate of substrate utilization becomes zero:

$$dS/dt = -[(V_{max}S)/(K_m + S)] \times [(S - S_T)/S], \tag{13}$$

where S_T is the threshold concentration at which $dS/dt = 0$. The integrated form of Equation (13) describes a substrate concentration-versus-time curve that approaches an asymptote equal to the threshold concentration, S_T. The integrated form of Equation (13) is

$$V_{max} \times t = (S_0 - S) + (K_m + S_T) \times \ln[(S_T - S_0)/(S_T - S)]. \tag{14}$$

Now consider an alternative model that instead assumes no threshold effect, however, substrate is produced endogenously and constantly during Michaelian utilization of exogenously added substrate (Robinson and Characklis 1984). The differential form of this second modification of the integrated Michaelis-Menten equation is

$$dS/dt = -[(V_{max}S)/(K_m + S)] + R, \tag{15}$$

where R is the assumed constant rate of endogenously produced S. The integrated form of Equation (15) is

$$-\Phi \times t = S_0 - S + [(K_m\Phi - RK_m)/\Phi] \\ \times \ln[(RK_m + \Phi S_0)/(RK_m + \Phi S)], \tag{16}$$

where $\Phi = R - V_{max}$. Equation (16) is the correct integrated form of model (15) and not the integrated form shown in Robinson and Characklis (1984). Like Equation (14), Equation (16) describes a substrate depletion curve that approaches a nonzero asymptote. Thus, the integrated versions of both models describe similar types of curves. Indeed it can be shown that S_T is equivalent to $(RK_m)/(V_{max} - R)$. Both Equation (14) and Equation (16) simplify to the integrated Michaelis-Menten equation, Equation (2), when S_T and R equal zero, respectively. Discriminating statistically between models (14) and (16) would require additional experimental evidence, for example, controls to eliminate the possibility of substrate being produced endogenously during the concomitant utilization of exogenously added substrate. The preceding example illustrates the value of comparing predictions made by competing models, using discrimination functions or simple algebraic comparisons, prior to wasting time with regression analyses.

5.2. A Posteriori *Discrimination*

Once a dataset is in hand, various methods can be used to decide which of the candidate models best describes the data statistically. The first set of criteria presented herein are those for comparisons between models with different numbers of parameters that are nested. A model is nested within another more complex model if the simpler model can be obtained by setting a parameter(s) in the complete (more complex) model equal to zero. More theoretical work exists on this problem than on the problem of deciding which of two models is best when the models have the same number of parameters, or for models with dissimilar numbers of parameters that are not nested. The practical considerations associated with the *a posteriori* discrimination problem are covered in greater detail by Bates and Watts (1988).

The F test for model discrimination (Bates and Watts 1988; Robinson 1985) can and has been used to discriminate between models with different numbers of parameters (Bates and Watts 1988; Robinson and Characklis 1984; Robinson 1985; Simkins et al. 1986). The calculation of the F statistic is straightforward and has been reported elsewhere (Robinson 1985). This statistical test, however, has some weaknesses. First, it cannot be used to discriminate between models with the same number of parameters or models that are not nested (Bates and Watts 1988). A second limitation of the F test is that the significance of the test is approximate for nonlinear models (Bates and Watts 1988; Draper and Smith 1981; Ratkowsky 1983). The actual significance level of the test may not equal the α value presumed, such as 0.05 or 0.01. The more nonlinear a nonlinear model is, the more approximate the significance level (Ratkowsky 1983). Third, indiscriminate use of the F test for model discrimination leads to over-parameterized models, i.e., models with more terms and parameters than are necessary.

Two other easily calculated summary statistics can be used to discriminate

between models with different numbers of parameters. The first is the coefficient of determination (r^2). This statistic has its usefulness but it is limited since it is constrained to vary between zero and one. With precise data, and where the number of data points is high (greater than 50, say), r^2 values may range from only 0.90 to 1.00 depending on the model. The second easily calculated summary statistic is the Akaike information criterion (AIC) (Akaike 1974). The AIC is calculated using the following equation:

$$AIC = n \times \ln(RSS_i) + 2 \times \theta_i, \tag{17}$$

where n is the number of data pairs, RSS_i is the sum of squared errors for the ith model of interest, and θ_i is the number of parameters of the ith nonlinear model. The model that gives the lowest AIC value for a given dataset is considered the best. The limitation of the AIC is that it is not a probabilistic model selection method. It is not clear what constitutes a significant difference in AIC values, however, the AIC does offer the researcher the ability to compare the quality of fits empirically between models that have dissimilar numbers of parameters that are not nested. There is no advantage to calculating values of AIC for comparisons between models with the same number of parameters, because this simplifies to merely ranking the mean square errors (sum of the squared errors divided by n minus the number of parameters in the model).

A model can be judged inferior statistically if the standard errors of the estimated parameters are high (Mannervik 1981; Bates and Watts 1988). Most nonlinear regression packages approximate the standard errors of the parameter estimates in addition to calculating values for the parameters themselves. If the respective standard errors are greater than 50% of the parameter estimates, then the fitted model can be discriminated against (Mannervik 1981).

One useful tool for model discrimination is residuals analysis (Bates and Watts 1988; Cook and Weisberg 1982; Draper and Smith 1981; Robinson 1985). Residuals analysis allows one to check whether the model does an accurate job of describing the data over the range of the measurements of the dependent variable. Further, residuals analysis provides clues to the types of measurement errors present in an experiment. A residual equals the observed value of the measured variable minus the value predicted by the fitted nonlinear model. There is one residual for each value of the measured variable and typically these are presented graphically. When a model does a poor job of describing the data, systematic trends exist in the data. In contrast, those models that more accurately reflect the behavior of the data yield residuals plots with no systematic behavior. Ideally, the residuals plot should show a random scattering of points above and below the zero line (observed minus predicted = zero). The other type of useful residuals plot is the one in which the residuals are plotted versus the predicted values of the dependent variable. This plot indicates whether the measurement

errors vary with the magnitude of the measured variable. If some correlation exists, then weighting the data should be considered (Bates and Watts 1988).

Cross-validation (Picard and Cook 1984; Snee 1977; Stone 1974) is a computer-intensive method that when applied may reduce the likelihood of building over-parameterized models. The F test for model discrimination can lead to overparameterized models since the addition of one more parameter to a model can make the F test significant statistically, if enough data are collected. Cross-validation emphasizes the predictive nature of the model instead of focusing on reducing the sum of squared errors about the fitted equation.

In the simplest implementation of cross-validation, the data are split into a "fit" set and a "test" set. The fit set is then fitted to the models of interest. The values of the test set are then predicted from each of the fitted models, and a prediction error is calculated (observed minus predicted value). Each prediction error (one for each value of the measured variable) is dependent on the rest of observations in the fit set. In its most intensive form, a single value of the measured variable becomes the test set. The fitted models are used to predict the point left out, and a prediction error is calculated. This datum is returned to the fit set, the next point is placed in the test set, the models are refit to the new fit set, and the value of the second measured point is predicted from the candidate models. This process continues until the last point in the curve takes its turn in the test set. In the end, there are a number of prediction errors for each model equal to the number of data pairs (e.g., substrate concentration versus time) collected. Histograms of the prediction errors for the models can be compared and the best model selected based on these empirical prediction-error distributions. The technique is clearly computer intensive, however, it can be implemented on microcomputers.

For models with an equal number of parameters, the discrimination statistics and approaches mentioned are all applicable except for the F test for model discrimination. This test is appropriate only for comparisons of models with different numbers of parameters that are nested. As an example, a one-term first-order decay model is not nested within the integrated Michaelis-Menten equation, Equation (2). This result is not at first obvious for this example since the first-order model is a simpler version of the Michaelis-Menten equation when $S \ll K_m$, however, the first-order model can not be obtained by setting K_m equal zero.

In place of the F test, ranking the residual mean squares (RMS) obtained for each model can be used (Mannervik 1981) for comparisons of models with the same number of parameters or for models with different numbers of parameters that are not nested. The weakness of this approach is the same as that of the AIC: what constitutes a significant difference in RMS values for fits of several models to a given dataset? For these situations, common sense and the practical needs of the researcher come to the forefront (Bates and Watts 1988).

Based on the respective strengths and weakness of the a posteriori discrimina-

tion methods discussed, minimal sets of criteria can be suggested. The following list of criteria should provide a solution to the discrimination problem involving models with different numbers of parameters:

1. F test for model discrimination (nested models only),
2. lack of systematic trends in residuals,
3. parameter standard errors that are 50% or less of the parameter estimates themselves.

For models with the same number of parameters, a suggested list of discrimination criteria would omit the F test and replace it with ranking of RMS values. Cross-validation is not listed because it is computer intensive. Comparisons of much more than two to three models presumably would require an inordinate amount of effort. But its use could prevent the selection of overparameterized models.

No single summary statistic can, or should, be used to decide which of several nonlinear models best describes a given dataset. Each technique tells us different things about the data. If all of the discrimination criteria point to the same model, for several data sets, then the experimentalist should consider himself or herself fortunate.

6. Optimal Experiments for Parameter Estimation

If the precision of the parameters of the best model is important, then the microbial ecologist can proceed to that phase of model building (Fig. 2.1). In this phase, values for the independent variable (e.g., time, dilution rate, pH, substrate concentration) are chosen that yield parameter estimates with the lowest possible variances. Some statistical work done in this area has focused on the Michaelis-Menten equation and the resulting recommendations are dependent on the variance model (Currie 1982; Duggleby 1979; Endrenyi and Chan 1981). To obtain minimum variance estimates of V_{max} and K_m, only two substrate concentrations are replicated, one equal to the K_m value and the other as high as is practically possible.

This two-point experimental design for the Michaelis-Menten model produces minimum-variance estimates of V_{max} and K_m only if the measurement errors are independent of the measured variable (substrate utilization rate, in this case) (Endrenyi and Chan 1981). A different variance model, such as when the variance of the measurement errors is proportional to the magnitude of the initial velocity, changes the second design point to the lowest substrate concentration that can be measured experimentally (Endrenyi and Chan 1981). This change is not obvious from the sensitivity equations of the model. This difference emphasizes

the importance of knowing the variance model for optimal experimental design. Obtaining this knowledge may be difficult because of the large number of data points presumably required to elucidate the variance mode (e.g., 50 or more).

Aside from accurate knowledge of the variance model, optimal experimental design for parameter estimation has a drawback when models are considered that have time as the independent variable. As an example, a model that does not have time as an independent variable is the model that predicts the steady-state substrate concentration (S_{ss}) as a function of the dilution rate, D, in a chemostat:

$$S_{ss} = (K_s D)/(\mu_{max} - D). \tag{18}$$

There are two optimal design points for this model (as there are for all two-parameter models) and the experimentalist could readily obtain replicate estimates of S_{ss} at these points merely by setting the dilution rate equal to the optimal design points in turn and replicating the study. However, the researcher cannot control time or "dial it in" to obtain replicate estimates of the substrate concentration at the optimal design points for the integrated Michaelis-Menten model, Equation (2). For this reason, the application of optimal design criteria may be limited to situations where the microbial ecologist is working with models that contain controllable variables, such as dilution rate.

If the variance of the measurement errors is constant, then the sensitivity equations can be used to predict the optimal design points for a nonlinear model. As noted, the number of optimal design points for parameter estimation equal the number of parameters in a model. These points are found by finding the values of the independent variable where the sensitivity equations are maximal. These can either be determined from plots of the sensitivity equations (Robinson 1985; Robinson and Characklis 1984; Robinson and Tiedje 1983), or analytically by differentiating each sensitivity equation with respect to the independent variable, setting the derivatives equal to zero, and solving for the independent variable. Either way the assumption of constant variance for the measurement errors may be critical, depending on the model, and this assumption is rarely true.

7. Concluding Remarks

This chapter has highlighted the use of NPE methods for identifying and discriminating among nonlinear models of interest to microbial ecologists. Much of the emphasis has been on the model discrimination aspects of the problem. However, no apologies are offered for this emphasis, because model discrimination forces the model builder to consider differences in predictions made by competing models and whether the estimation properties of individual models are acceptable. Ratkowsky (1990) reinforced this sentiment in his statement that

"a far greater threat to valid nonlinear regression modeling is the belief that a complicated model is superior to a simple model." The model discrimination phase of NPE compels us to use Occam's razor and to not build a Rube Goldberg shaver.

References

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans. Auto. Control.* AC-19:716–723.

Alexander, M. 1994. *Biodegradation and Bioremediation.* pp 71–101. Academic Press, New York.

Bates, D. M., and D. D. Watts. 1988. *Nonlinear Regression Analysis and its Applications.* pp 67–133 John Wiley & Sons, New York.

Brunner, W., and D. D. Focht. 1984. Deterministic three half-order kinetic model for microbial degradation of added carbon substrates in soil. *Appl. Environ. Microbiol.* 47:1211–1216.

Cook, R. D., and S. Weisberg. 1982. *Residuals and Influence in Regression.* Chapman and Hall, Ltd., New York

Currie, D. J. 1982. Estimating Michaelis-Menten parameters: Bias, variance and experimental design. *Biometrics* 38:907–919.

Draper, N. R., and H. Smith. 1981. *Applied Regression Analysis.* John Wiley & Sons, New York.

Duggleby, R. G. 1979. Experimental designs for estimating the kinetic parameters for enzyme-catalyzed reactions. *J. Theor. Biol.* 81:671–684.

Endrenyi, L., and F.-Y. Chan. 1981. Optimal design of experiments for precise hyperbolic kinetic and binding parameters. *J. Theor. Biol.* 90:214–263.

Li, W. K. W. 1983. Consideration of errors in estimating kinetic parameters based on Michaelis-Menten formalism in microbial ecology. *Limnol. Oceanogr.* 28:185–190.

Mannervik, B. 1981. Design and analysis of kinetic experiments for discrimination between rival models. In L. Endrenyi (ed.), *Kinetic Data Analysis: Design and Analysis of Enzyme and Pharmacokinetic Experiments*, pp. 235–270. Plenum Press, New York.

Paris, D. F., and J. E. Rogers. 1986. Kinetic concepts for measuring microbial rate constants: effects of nutrients on rate constants. *Appl. Environ. Microbiol.* 51:221–225.

Picard, R. R., and R. D. Cook. 1984. Cross-validation of regression models. *J. Am. Stat. Assoc.* 79:575–583.

Ratkowsky, D. A. 1983. *Nonlinear Regression Modeling: a Unified Practical Approach.* Marcel Dekker, New York.

Ratkowsky, D. A. 1990. *Handbook of Nonlinear Regression Models.* Marcel Dekker, New York.

Robinson, J. A. 1985. Determining microbial kinetic parameters using nonlinear regression analysis: advantages and limitations in microbial ecology. *Adv. Microb. Ecol.* 8:61–114.

Robinson, J. A., and J. M. Tiedje. 1983. Nonlinear estimation of Monod growth parameters from a single substrate depletion curve. *Appl. Environ. Microbiol.* 45:1453–1458.

Robinson, J. A., and W. G. Characklis. 1984. Simultaneous estimation of V_{max}, K_m and the rate of endogenous substrate production (R) from progress curve data. *Microb. Ecol.* 10:165–178.

Schmidt, S. K., S. Simkins, and M. Alexander. 1985. Models for the kinetics of biodegradation of organic compounds not supporting growth. *Appl. Environ. Microbiol.* 50:323–331.

Silvert, W. 1979. Practical curve fitting. *Limnol. Oceanogr.* 170:767–773.

Simkins, S., and M. Alexander. 1984. Models for mineralization kinetics with the variables of substrate concentration and population density. *Appl. Environ. Microbiol.* 47:1299–1306.

Simkins, S., R. Mukheriee, and M. Alexander. 1986. Two approaches to modelling kinetics of biodegradation by growing cells and application of a two-compartmental model for mineralization kinetics in sewage. *Appl. Environ. Microbiol.* 51:1153–1160.

Snee, R. D. 1977. Validation of regression models: Methods and examples. *Technometrics* 19:415–428.

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. *J. Roy. Stat. Soc. B.* 36:111–147.

3

Analysis of Repeated Measures Data Using Nonlinear Models

George A. Milliken and April J. Milliken-MacKinnon

1. Introduction

Many microbiological studies involve repeatedly measuring the same system, process, sample, or experimental unit several times or in several places, thus generating a repeated measures dataset. For example, an experiment, here called an experimental unit, is set up. A number of such experiments are run simultaneously. Consider just one unit, it may be observed at a number of times or a number of different aspects of it may be measured, but it is a single unit because the experiment is only observed and not interfered with in any way. These observations and measurements are called repeated measures. In contrast, repeated measures are to be distinguished from replication or replicate measures in which the same set of conditions are observed to the best of the experimenters ability on additional occasions. Repeated measures are observed on each of these replicates, but the setup of the experiment itself provides the replication of the treatment or set of conditions. The repeated measures are not replications. More specifically, you can measure the size of a bacterial colony as a function of time on a given substrate by measuring the size at several selected time points. That is, these measurements at different time points on the same bacterial colony are repeated measures on that colony. Such experiments generally consist of comparing several treatments or conditions or treatment combinations. In the example, the treatments can be several different growth substrates. The typical experiment is to obtain several replicate runs or observations or experimental units from the same condition or set of conditions. The observations within a run from a repeated measures dataset are correlated and the methods derived based on an independent errors

assumption for modeling and comparing the models across conditions are not appropriate. For the bacterial growth example, several petri dishes (more than one) would be assigned to each substrate and the size of the colonies in each petri dish would be measured at several predetermined times. The usual methods of analysis are derived based on the assumptions that the observations within a run are independently distributed, i.e., not correlated. The measurements made on the same petri dish are correlated, thus the usual methods of modeling the growth and comparing the models across substrates are not applicable.

The situation discussed in this chapter consists of an experiment or study with t treatments, where each treatment is applied to m experimental units with n observations or repeated measures measured on each experimental units. The observations between runs or experimental units should not be correlated if an experimental design using proper randomization is utilized. The measured response variables are to be modeled as functions of independent or regressor variables measured within each run, process, or experimental unit. Time or depth are two possible regressor variables. For example, the integral form of the Michaelis-Menten enzyme kinetic equation may be appropriate to describe the decrease in concentration of a herbicide as a function of time. The analysis is to select a model that describes the repeated measures data collected from each experimental unit and then fit that model to the data from each experimental unit. The model is most likely selected to be representative of the biological process (rather than a polynomial regression model) where the parameters are of interest to the researcher. Parameters such as growth rates, decay rates, extent of growth or decay, etc. may be of interest to the biologist, rather than the data collected at each time point or depth. The parameters of the Michaelis-Menten enzyme kinetic equation are more likely to be of interest to the biologist than the amount of residual herbicide that has accumulated at a particular time. The process starts by fitting a model to the data from each experimental unit to provide estimates of the models parameters. The estimated parameters from each experimental unit's model, or some other characteristic of the model, are to be used as data for further statistical analyses. Several procedures are described for comparing the models across treatments including, comparing specific parameters of the models, comparing models evaluated at specific values of the independent or regressor variables, and comparing predicted values of the regressor variables required to provide a specified level of response. The treatments can also be compared by constructing confidence intervals for individual parameters and functions of parameters, constructing confidence bands for individual models, and constructing confidence bands for the difference of models from two different treatments. The next section discusses the assumptions, the data collected on each experimental unit, and the type of model needed to adequately describe that data.

2. The Model

The number of treatments to be compared are denoted by t, which can be different levels of the same type of factor (levels of nitrogen) or could be levels of different items such as different management systems or different processes to treat a soil sample. The experiment consists of randomly assigning r_i experimental units to the ith level of the treatment and making n repeated measurements on the jth experimental unit assigned to the ith level of the treatment. Let y_{ijk} denote the observed response or dependent variable, and $x_{1ijk}, x_{2ijk}, \ldots x_{pijk}$ denote the independent or regressor variables from the kth repeated measurement on the jth experimental unit assigned to the ith level of the treatment. A model to describe the relationship between the response variable and the regressor variables for each level of the treatment is

$$y_{ijk} = f(\boldsymbol{x}_{ijk}, \boldsymbol{\beta}_i) + e_{ij} + \epsilon_{ijk}, \quad i = 1, 2, \ldots t, \quad j = 1, 2, \ldots r_i, \tag{1}$$
$$k = 1, 2, \ldots n,$$

where $\boldsymbol{x}_{ijk}$ is the $p \times 1$ vector of the independent variables; $\boldsymbol{\beta}_i$ is the $q \times 1$ vector of unknown parameters corresponding to the model $f(\boldsymbol{x}_{ijk}, \boldsymbol{\beta}_i)$, a known function of $\boldsymbol{x}_{ijk}$ and $\boldsymbol{\beta}_i$, e_{ij} is the error term representing the variability of the experimental units assumed to be identically independently distributed (iid) $N(0, \sigma_e^2)$ and ϵ_{ijk} is the error term representing the variability in the repeated measures obtained from a given experimental unit. The $n \times 1$ vector of errors $\boldsymbol{\epsilon}_{ij} = (\epsilon_{ij1}, \epsilon_{ij2}, \ldots , \epsilon_{ijn})$, are assumed to be identically independently distributed as $N(0, \Sigma_\epsilon)$. The covariance matrix of the repeated measures, Σ_ϵ, can be of any form as long as it is a positive definite matrix.

The major assumption is the application of the levels of the treatment can have an effect on the mean of the model, but the application of the treatment cannot have an effect on the variance of the experimental units or on the covariance matrix of the repeated measures. (If the levels of the treatment have an effect on the variance as well as on the mean, an unequal variance procedure would have to be developed to solve the problem.) Here we consider only a univariate case of that very important problem. The $n \times 1$ vector of repeated measurements on ijth experimental unit, denoted by $\boldsymbol{y}_{ij} = (y_{ij1}, y_{ij2}, \ldots y_{ijn})$, can be expressed as the model

$$\boldsymbol{y}_{ij} = f(\mathbf{X}_{ij}, \boldsymbol{\beta}_i) + e_{ij}\boldsymbol{j}_n + \boldsymbol{\epsilon}_{ij}, \quad i = 1, 2, \ldots t, \quad j = 1, 2, \ldots r_i \tag{2}$$

where $\boldsymbol{j}_n$ is a $n \times 1$ vector of 1's and $\mathbf{X}_{ij}$ is the $n \times p$ matrix of the regressor variables with the vectors $\boldsymbol{x}_{ijk}$ as rows. The variance of $\boldsymbol{y}_{ij}$ is $\text{var}(\boldsymbol{y}_{ij}) = \sigma_e^2 \mathbf{J}_n + \Sigma_\epsilon$, where $\mathbf{J}_n$ is an $n \times n$ matrix of 1's. Thus observations from the same experimental

unit are correlated because they are all measured on the same experimental unit (through e_{ij} or $\sigma_e^2 \, \mathbf{J}_n$) and because they are repeated measures (through ϵ_{ijk} or $\mathbf{\Sigma}_\epsilon$). The usual assumption is $\mathbf{\Sigma}_\epsilon = \sigma_\epsilon^2 \, \mathbf{I}_n$ and, if Model (2) contains an intercept, the usual least-squares estimation procedures yield appropriate estimators for the model's parameters and provide suitable procedures for comparing treatments through regression models. If the form of $\mathbf{\Sigma}_\epsilon$, is unknown or is more complicated than $\sigma_\epsilon^2 \, \mathbf{I}_n$, the usual least-squares procedures will not provide suitable results for comparing characteristics of the treatments' models. Methods that do not rely on knowing the form of $\mathbf{\Sigma}_\epsilon$ are desirable. One such strategy is described in the next section.

3. Parameter Estimation

The strategy for carrying out the analysis of the data for Model (1) is to use ordinary least squares to obtain estimates of the parameters of the model fit to the data from each experimental unit and then use those estimates or functions of those estimates in an analysis of variance setting to compare the treatments. If the model is linear in the parameters, a linear least-squares computer code can be used to estimate the model's parameters for each experimental unit. If the model is a nonlinear function of the parameters, a nonlinear least-squares computer code can be used to estimate the model's parameters for each experimental unit. For example, a Michaelis-Menten model can be fit to the data from each experimental plot and an analysis of variance can be conducted on the integrated Michaelis-Menten constants to compare the treatments.

Ordinary least-squares estimators of the parameters of the model for the ijth experimental unit having r_i repeated measures are obtained by selecting the values of $\boldsymbol{\beta}_{ij}$ that minimize the quantity

$$Q(\boldsymbol{\beta}_{ij}) = \sum_{k=1}^{n}\left[y_{ij} - f(\mathbf{X}_{ij}, \boldsymbol{\beta}_i) \right]^2, \quad i = 1, 2, \ldots t, \quad j = 1, 2, \ldots r_i \qquad (3)$$

producing estimates $\hat{\boldsymbol{\beta}}_{ij}$ $i = 1, 2, \ldots t, j = 1, 2, \ldots r_i$. The asymptotic sampling distributions of the

$$\hat{\boldsymbol{\beta}}_{ij}, \ i = 1, 2, \ldots t, j = 1, 2, \ldots r_i, \text{ are } \boldsymbol{\beta}_{ij} \sim N(\boldsymbol{\beta}_i, \mathbf{\Sigma}_{ij}),$$
$$i = 1, 2, \ldots t, j = 1, 2, \ldots r_i$$

where

$$\Sigma_{ij} = \left(Z_{ij}'Z_{ij}\right)^{-1}Z_{ij}'\left(\sigma_e^2 J_n + \Sigma_e\right)Z_{ij}\left(Z_{ij}'Z_{ij}\right)^{-1}, \quad i = 1, 2, \ldots t, \quad j = 1, 2, \ldots r_i$$

The $n \times q$ matrix Z_{ij} consists of the derivatives of the ijth model with respect to the $q \times 1$ vector of parameters, β_i, computed as

$$\begin{bmatrix} \dfrac{\partial f(x_{ij1},\beta_i)}{\partial \beta_{i1}} & \dfrac{\partial f(x_{ij1},\beta_i)}{\partial \beta_{i2}} & \cdots & \dfrac{\partial f(x_{ij1},\beta_i)}{\partial \beta_{iq}} \\[2ex] \dfrac{\partial f(x_{ij2},\beta_i)}{\partial \beta_{i1}} & \dfrac{\partial f(x_{ij2},\beta_i)}{\partial \beta_{i2}} & \cdots & \dfrac{\partial f(x_{ij2},\beta_i)}{\partial \beta_{iq}} \\[2ex] \vdots & \vdots & \ddots & \vdots \\[2ex] \dfrac{\partial f(x_{ijn},\beta_i)}{\partial \beta_{i1}} & \dfrac{\partial f(x_{ijn},\beta_i)}{\partial \beta_{i2}} & \cdots & \dfrac{\partial f(x_{ij1n}\beta_i)}{\partial \beta_{iq}} \end{bmatrix}, \quad i = 1, 2, \ldots t, \quad j = 1, 2, \ldots r_i.$$

The preceding information can be used to construct a model for the estimates of the parameters from each experimental units model as the $\beta_{ij} = \hat{\beta}_i + a_{ij}$, $i = 1, 2, \ldots t, j = 1, 2, \ldots r_i$, where the $a_{ij} \sim N(0, \Sigma_{ij})$. The covariance matrices of the $\hat{\beta}_{ij}$ are not necessarily equal. This is the case since the covariance matrices Σ_{ij} $i = 1, 2, \ldots t, j = 1, 2, \ldots r_i$, depend on the values of the matrix of derivatives, Z_{ij}, which are functions of the independent variables evaluated at β_i. Since the values of the parameters are likely to be different and the values of the independent variables for each experimental unit are likely to be different, the covariance matrices are likely to be different. If the values of the independent variables are similar from experimental unit to experimental unit within a treatment, the covariance matrices of experimental units from the same treatment should be approximately equal. (You should plot the values of the response variable against the value of each of the independent variables to check for outliers. Also you could use multivariate outlier techniques if the number of independent variables is quite large.) The covariance matrices between treatments are most likely not equal since the parameters between treatments are not necessarily equal. The assumption made for the remainder of this chapter is that the independent variables for experimental units within a treatment group are similar. This assumption enables us to assume the covariance matrices of the repeated measures within a treatment group are equal.

4. Comparing the Treatments

There are many ways to compare the treatments. The method of comparison depends on the objective of the experiment. A general comparison is described, followed by examples to demonstrate specific types of comparisons. Let $g(x,\beta_i)$

be a known function of the unknown parameters $\boldsymbol{\beta}_i$ and of the specific values for the vector of independent variables $\boldsymbol{x}$, where it is of interest to test the hypothesis $H_0: g(\boldsymbol{x}, \boldsymbol{\beta}_1) = g(\boldsymbol{x}, \boldsymbol{\beta}_2) = \cdots = g(\boldsymbol{x}, \boldsymbol{\beta}_t)$ versus H_a: not H_0. For each experimental unit estimate $g(\boldsymbol{x}, \boldsymbol{\beta}_i)$ by evaluating the function at $\boldsymbol{\beta}_i = \hat{\boldsymbol{\beta}}_{ij}$, i.e., by computing the values of $g_{ij} = g(\boldsymbol{x}, \hat{\boldsymbol{\beta}}_{ij})$, $i = 1, 2, \ldots t, j = 1, 2, \ldots r_i$. The asymptotic sampling distributions of the g_{ij}, $i = 1, 2, \ldots t, j = 1, 2, \ldots r_i$ are independent $N(\mu_i, \sigma_{gi}^2)$, where $\mu_i = g(\boldsymbol{x}, \boldsymbol{\beta}_i)$, $\sigma_{gi}^2 = \boldsymbol{h}_i' \boldsymbol{\Sigma}_{ij} \boldsymbol{h}_i$, and $\boldsymbol{h}_i' = [\partial g(\boldsymbol{x}, \boldsymbol{\beta}_i)/\partial \beta_{i1}, \partial g(\boldsymbol{x}, \boldsymbol{\beta}_i)/\partial \beta_{i2}, \ldots, \partial g(\boldsymbol{x}, \boldsymbol{\beta}_i)/\partial \beta_{iq}]$, $i = 1, 2, \ldots t$. The values of the parameter vectors $\boldsymbol{\beta}_i$ are unknown and the values of the variances σ_a^2 and $\boldsymbol{\Sigma}_\epsilon$ are unknown, so the values of the σ_{gi}^2 are unknown.

To circumvent not knowing the values of the parameters, construct a new model for the g_{ij} using the preceding information. The model for the g_{ij} is

$$g_{ij} = \mu_i + a_{ij}, \quad i = 1, 2, \ldots t, \quad j = 1, 2, \ldots r_i \qquad (4)$$

where the a_{ij} are independently distributed $N(0, \sigma_{gi}^2)$. Model (4) is that of a one-way treatment structure in a completely randomized design structure with possibly unequal variances for each level of the treatment. The g_{ij} are independently distributed since each g_{ij} is computed from a different experimental unit and we have assumed the experimental units are all independent. The hypothesis $H_0: g(\boldsymbol{x}, \boldsymbol{\beta}_1) = g(\boldsymbol{x}, \boldsymbol{\beta}_2) = \ldots = g(\boldsymbol{x}, \boldsymbol{\beta}_t)$ versus H_a: (not H_0) is equivalent to the hypothesis $H_0: \mu_1 = \mu_2 = \ldots \mu_t$ versus H_a: (not H_0), which can be tested using the F test from an unequal variance the analysis of variance (see Chap. 2 of Milliken and Johnson (1992). Before the unequal variance analysis of variance procedure is used, the equal variance hypothesis should be tested. If the F test for equal treatment effects indicates there is sufficient evidence to believe the null hypothesis is false, i.e., reject H_0, the treatments can be compared through the $g(\boldsymbol{x}, \boldsymbol{\beta}_i)$ by constructing contrasts of the treatments or by constructing a multiple comparisons procedure to make pairwise comparisons.

The beauty of this procedure is that the form of the covariance matrix $\boldsymbol{\Sigma}_\epsilon$ does not need to be known, nor does the value of σ_e^2, to compare the treatments through the $g(\boldsymbol{x}, \boldsymbol{\beta}_i)$, i.e., using the assumed form of the model between the response variable and the independent variables.

The preceding discussion is general in that the form of $g(\boldsymbol{x}, \boldsymbol{\beta}_i)$ is general. Some specific forms of $g(\mathbf{x}, \boldsymbol{\beta}_i)$ that may be of interest to the researcher are as follows:

1. If you want to compare the treatments through one specific parameter, such as the growth rate, select $g(\boldsymbol{x}, \boldsymbol{\beta}_i) = \beta_{is}$, $i = 1, 2, \ldots, t, j = 1, 2, \ldots r_i$, i.e., select the sth parameter from each $\boldsymbol{\beta}_i$ to provide a comparison of individual parameters.
2. If you want to compare the models evaluated at a specific value of the vector of independent variables $\boldsymbol{x}$, such as the amount of estimated residual herbicide after 10 days, select $g(\boldsymbol{x}, \boldsymbol{\beta}_i) = f(\boldsymbol{x}, \boldsymbol{\beta}_i)$, $i = 1, 2, \ldots t, j = 1, 2, \ldots r_i$.

3. If your theory suggests some important nonlinear or linear function of the parameters is an interpretable function, then select that function as a basis for comparing the treatments. For example, that interesting function might be

$$g(x,\beta_i) = \frac{\beta_{i1} + \beta_{i2}\, e^{(x_1 - \beta_{i3})}}{\beta_{i4} + \beta_{i5}x_2}, \quad i = 1, 2, \ldots t, \quad j = 1, 2, \ldots r_1.$$

4. If you want to use the slopes of the regression models in the x_p direction evaluated at a specific value of x as a basis for comparing the treatments then select

$$g(x,\beta_i) = \frac{\partial f(x,\beta_i)}{\partial x_p}, \quad i = 1, 2, \ldots t, \quad j = 1, 2, \ldots r_i.$$

5. If you want to use the values of x that will yield a specific value of the model, say the 50th percentile, as the basis for comparing the treatments then select

$$x_0 = g(x_0, \beta_i, f_0) = f^{-1}(x_0, \beta_i, f_0), \quad i = 1, 2, \ldots t, \quad j = 1, 2, \ldots r_1.$$

This choice includes using the lethal dose 50% (LD50s or some other LD percentiles) to compare the models. LD (lethal dose) or ED (effective dose) is the dose or x value at which a given percent of the population will respond. Thus LD50 is the dose at which 50% of the population will respond.

Most any function of x and β_i can be selected as the basis for comparing the treatments' models. The important aspect of this process is to choose functions that relate to the objectives of the study or experiment.

5. Constructing Confidence Bands for the Models

The estimates of the parameters from the model fit to the data from each individual experimental unit can be used to construct confidence intervals about the functions of the model's parameters, including the models. Let $\hat{\beta}_{ij}$ denote the $q \times 1$ vector of estimates of the parameters from the jth experimental unit from the ith treatment group. The estimates from the ith treatment group $\hat{\beta}_{ij}, j = 1, 2, \ldots r_i$, are independently identically distributed as a multivariate normal with mean β_i and variance covariance matrix Σ_i, denoted by $\hat{\beta}_{ij}, j = 1, 2, \ldots r_i \sim \text{iid}N(\beta_i, \Sigma_i)$. The sample mean of the individual parameter estimates and their sample covariance matrix are the estimates β_i and Σ_i. Let b_i denote the vector

of sample means and $\mathbf{S}_i$ denote the matrix of sample variances and covariances computed as

$$\boldsymbol{b}_i = (b_{i1}, b_{i2}, \ldots b_{iq})' \quad \text{and} \quad \mathbf{S}_i = \begin{bmatrix} S_{i11} & S_{i12} & \ldots & S_{i1q} \\ S_{i21} & S_{i22} & \ldots & S_{i2g} \\ \vdots & \vdots & \ddots & \vdots \\ S_{iq1} & S_{iq2} & \ldots & S_{iqq} \end{bmatrix}$$

where

$$b_{ik} = \frac{\sum_{j=1}^{r_i} \hat{\beta}_{ijk}}{r_i} \quad \text{and} \quad s_{ikm} = \frac{\sum_{j=1}^{r_i} (\hat{\beta}_{ijk} - b_{ik})(\hat{\beta}_{ijm} - b_{im})}{r_i - 1}$$

Let $g(x,\boldsymbol{\beta}_i)$ be a known function of the unknown parameters $\boldsymbol{\beta}_i$ and of the specific values for the vector of independent variables x. We want to provide an estimate of $g(x,\boldsymbol{\beta}_i)$ and an estimate of an approximation to its standard error. The estimate of $g(x,\boldsymbol{\beta}_i)$ is obtained by evaluating the function at $\boldsymbol{\beta}_i = \mathbf{b}_i$ as $\hat{g}(x,\boldsymbol{\beta}_i) = g(x,\mathbf{b}_i)$. The asymptotic sampling distribution of the $g(x,\mathbf{b}_i)$ is iid $N [g(x,\boldsymbol{\beta}_i), \sigma^2_{gi}]$, where $\sigma^2_{gi} = \mathbf{h}'_i \, \boldsymbol{\Sigma}_i \mathbf{h}_i$ and $\mathbf{h}'_1 = [\partial g(x,\boldsymbol{\beta}_i)/\partial \beta_{i1}, \partial g(x,\boldsymbol{\beta}_i)/\partial \beta_{i2}, \ldots, \partial g(x,\boldsymbol{\beta}_i)/\partial \beta_{i2}]$ and iid denotes "independently identically distributed." The estimate of $\boldsymbol{\Sigma}_i$ is $\mathbf{S}_i$ and a large sample approximation to the estimate of σ^2_{gi} is $\hat{\sigma}^2_{gi} = \dfrac{\hat{\mathbf{h}}'_i \, \mathbf{S}_i \mathbf{h}_i}{r_i} = s^2_{gi}$ where $\hat{\mathbf{h}}_i$ is the vector of derivatives evaluated at $\boldsymbol{\beta}_i = \boldsymbol{b}_i$. An approximate $(1 - \alpha)\ 100\%$ confidence interval about $g(x,\boldsymbol{\beta}_i)$ is $g(x,\boldsymbol{b}_i) \pm t_{\alpha/2,(ri-1)} S_{gi}$, where $t_{\alpha/2,(ri-1)}$ is the $\alpha/2$ upper percentage point from a Student's t distribution with $r_i - 1$ degrees of freedom. An assumption that is tacitly made is the variances, σ^2_{gi}, $i = 1, 2, \ldots t$, are not necessarily equal and they have not been pooled. If the variances are not significantly different (using a test such as Levene's; see Chap. 2 of Milliken and Johnson (1992), the variances are poolable and they should be pooled into a single estimate and the pooled degrees of freedom should be pooled when selecting the percentage point. If the choice of $g(x,\boldsymbol{\beta}_i)$ is a function of x and you wish to construct a confidence band about $g(x,\boldsymbol{\beta}_i)$ for a range of values of x construct the preceding interval for a grid of x values, compute the confidence interval about $g(x,\boldsymbol{\beta}_i)$ for each x and connect the dots to provide the confidence band.

Confidence intervals and bands can be constructed about the difference of two treatments' models as $g(x,\boldsymbol{\beta}_i) - g(x,\boldsymbol{\beta}_h)$ for $i \neq h$ by

$$g(x,\boldsymbol{b}_i) - g(x,\boldsymbol{b}_h) \pm \sqrt{t_{\alpha/2,(r_{i-1})} s^2_{gi} + t_{\alpha/2,(r_h-1)} \, s^2_{gh}},$$

where $t_{\alpha/2(r_i-1)}$ and $t_{\alpha/2(r_h-1)}$ are the $\alpha/2$ upper percentage points from a Student's t distribution with $r_i - 1$ and $r_h - 1$ degrees of freedom, respectively, and $s_{gi}^2 = \text{var}[g(\mathbf{x},\mathbf{b}_i)]$ and $s_{gh}^2 = \text{var}[g(\mathbf{x},\mathbf{b}_h)]$.

The confidence interval about the difference can be expressed as

$$g(\mathbf{x},\mathbf{b}_i) - g(\mathbf{x},\mathbf{b}_h) \pm t_{\alpha/2,(v)} \sqrt{s_{gi}^2 + s_{gh}^2}$$

where $t_{\alpha/2,(v)}$ is the $\alpha/2$ upper percentage points from a Student's t distribution with v degrees of freedom. v is the approximated degrees associated with $s_{gi}^2 + s_{gh}^2$ computed using the Satterthwaite approximation, as described in Milliken and Johnson (1992). The value of v is computed as

$$v = \frac{(s_{gi}^2 + s_{gh}^2)^2}{\dfrac{s_{gi}^4}{r_i - 1} + \dfrac{s_{gh}^4}{r_h - 1}}$$

The confidence intervals can be extended to provide interval estimates for a linear combination of the $g(\mathbf{x},\boldsymbol{\beta}_i)$ as $\sum_{i=1}^{t} c_i\, g(\mathbf{x},\boldsymbol{\beta}_i)$, where the c_i are known constants.

An approximate $(1 - \alpha)$ 100% confidence interval about $\sum_{i=1}^{t} c_i\, g(\mathbf{x},\boldsymbol{\beta}_i)$ is

$$\sum_{i=1}^{t} c_i g(\mathbf{x},\mathbf{b}_i) \pm t_{\alpha/2,(v)} \sqrt{\sum_{i=1}^{t} c_i^2 s_{gi}^2}\,,$$

where v is the approximate degrees of freedom computed as

$$v = \frac{\left(\sum_{i=1}^{t} c_i^2\, s_{gi}^2\right)^2}{\sum_{i=1}^{t} \dfrac{c_i^4 s_{gi}^4}{r_i - 1}}$$

If the covariance matrices are not significantly different, the sample covariance matrices can be pooled by pooling the elements as

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1q} \\ s_{21} & s_{22} & \cdots & s_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ s_{q1} & s_{q2} & \cdots & s_{qq} \end{bmatrix}$$

where

$$s_{km} = \frac{\displaystyle\sum_{i=1}^{t} \sum_{j=1}^{r_1} (\hat{\beta}_{ijk} - b_{ik})(\hat{\beta}_{ijm} - b_{im})}{\displaystyle\sum_{i=1}^{t} (r_i - 1)}.$$

Using the pooled estimate of Σ, S, the large sample approximation to the estimate of σ_{gi}^2 is

$$\hat{\sigma}_{gi}^2 = \frac{\hat{h}_i \, S \hat{h}_i}{r_i} = s_{gi}^2$$

where $\hat{h}_i$ is the vector of derivatives evaluated at $\beta_i = b_i$. An approximate $(1 - \alpha)$ 100% confidence interval about $g(x, \beta_i)$ is $g(x, b_i) \pm t_{\alpha/2(v)} s_{gi}$, where $t_{\alpha/2(v)}$ is the $\alpha/2$ upper percentage point from a Student's t distribution with $v = \sum_{i=1}^{t} (r_i - 1)$ degrees of freedom.

Confidence intervals and bands can be constructed about the difference of two treatments' models as $g(x, \beta_i) - g(x, \beta_h)$ for $i \neq h$ by $g(x, b_i) - g(x, b_h) \pm t_{\alpha/2(v)} \sqrt{s_{gi}^2 + s_{gh}^2}$. An approximate $(1 - \alpha)$ 100% confidence interval about $\sum_{i=1}^{t} c_i \, g(x, \beta_i)$, where the c_i are known constants is

$$\sum_{i=1}^{t} c_i \, g(x, b_i) \pm t_{\alpha/2, (v)} \sqrt{\sum_{i=1}^{t} c_i^2 s_{gi}^2}.$$

The preceding procedures are demonstrated using two examples. The first example is from bakery science where characteristics of the cookie development curve are evaluated and compared for several cookie formulations. The second example compares three soil-handling and pesticide application methods by modeling the radioactive CO_2 production over time. For both of these examples the data collected from each experimental unit are repeated measures, but the experimental units are independent.

6. Example 1: Growing Cookies

The data in Table 3.1 are from an experiment involving the development of cookies where the cookies are from different cookie formulations or recipes (Doescher (1986)). The process is to (1) mix ingredients for a given formulation

Table 3.1 Data for the cookie development example where the columns under each formulation correspond to the three cookies.

Time	Soft Winter Wheat and Granular Sucrose			Time	Soft Winter Wheat and Fructose Syrup		
0	62.10	61.72	61.96	0	59.41	60.31	60.62
1	63.08	62.21	63.19	1	63.76	62.31	62.31
2	66.99	67.31	67.59	2	67.66	66.65	67.39
3	72.86	72.90	74.18	3	72.51	72.21	72.69
4	76.28	76.79	78.08	4	77.36	76.79	78.00
5	80.68	79.95	81.74	5	79.78	79.93	81.14
6	85.57	84.08	85.89	6	80.75	82.11	82.83
7	91.44	87.24	90.52	7	80.99	82.11	83.07
8	94.13	88.69	92.48	8	80.51	82.11	83.07
9	94.62	89.18	92.72	9	80.02	81.14	82.59
10	94.62	89.18	92.72	10	78.33	80.18	81.14

Time	Soft Winter Wheat and Glucose Syrup			Time	Hard Red Winter Wheat and Granular Sucrose		
0	59.65	61.34	61.23	0	60.61	60.94	61.74
1	62.08	63.27	62.92	1	62.05	62.62	63.21
2	65.47	66.17	68.00	2	65.42	65.49	67.13
3	70.81	70.76	74.29	3	70.71	71.19	72.52
4	76.14	76.55	78.89	4	74.79	74.57	76.19
5	80.02	80.90	82.04	5	76.96	77.44	79.13
6	81.96	82.35	83.97	6	78.40	79.35	81.83
7	82.45	82.83	84.94	7	80.09	80.06	82.56
8	82.69	82.83	84.94	8	79.36	80.06	82.56
9	82.45	82.83	84.94	9	79.36	79.83	82.32
10	82.21	82.35	84.94	10	79.36	79.83	82.32

Time	Soft Winter Wheat and Granular Sucrose		
0	·	60.98	60.11
1	63.19	61.47	61.79
2	68.07	64.85	65.14
3	72.96	71.15	71.13
4	78.81	77.20	76.88
5	83.94	82.28	81.91
6	87.59	87.36	85.50
7	90.52	90.02	87.42
8	91.25	91.96	89.09
9	91.25	91.23	89.81
10	91.25	91.96	89.81

into a cookie dough; (2) measure out a specified amount of cookie dough; (3) put the cookie dough in the oven with preset conditions; and (4) watch it bake, i.e., track the diameter of the cookie as it develops or grows. A camera was mounted on the oven door focused through the door's window with a calibration bar behind the cookie dough. A picture was taken every minute for 10 min, starting with zero time thus there are 11 measurements made on each cookie. The pictures were used to measure the diameter of the cookie after each minute of exposure to the baking temperature. The graph in Figure 3.1 displays the general pattern of diameter growth of a cookie, called the cookie development curve. The development curve shows that the cookie grows for a certain length of time, at which point it stops expanding or sets. The time at which the cookie stops expanding is called the *set time*. The slope of the first segment of the model is called the expansion rate or *growth rate* of the cookie, and the second segment or plateau part of the model is at a height of the *maximum diameter*. Once the cookie sets, the diameter changes very little (those cookies that set early, they started to burn, thus reducing the diameter a little bit). These 11 measurements on the same cookie are repeated measures and are highly correlated with each
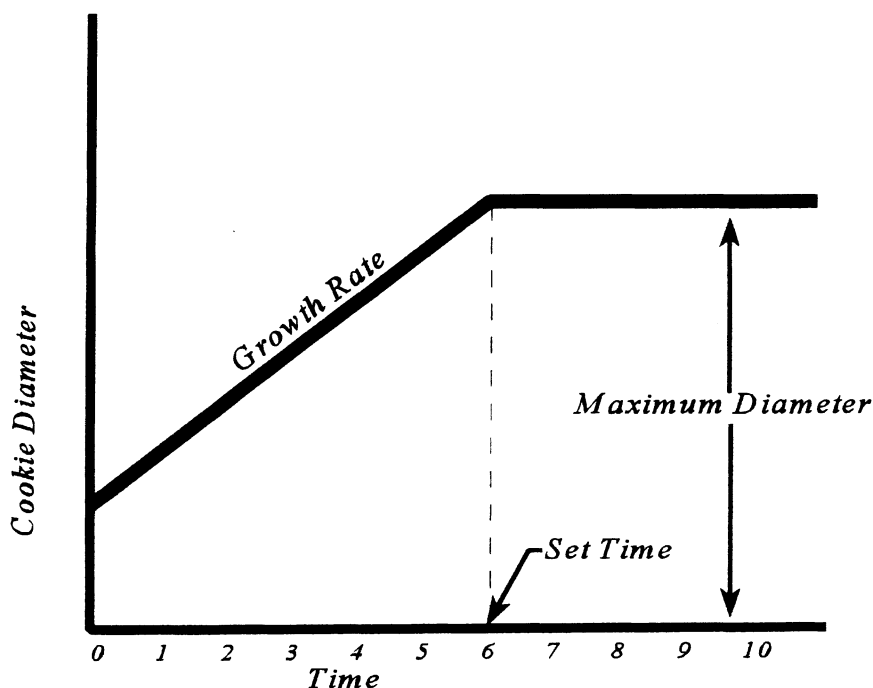

Figure 3.1 Graph of the cookie development function with important characteristics noted.

other, but are independent of the measurements made on other cookies. The repeated measurements are also not equally correlated, because those measurements of the growth stage of the process are not as highly correlated as those on the plateau stage of the process. The variability of the measurements is larger for the growth stage than for the horizontal stage. A repeated measures analysis of variance could be used to compare the cookie formulations at each time of exposure to the oven temperature. Such comparisons are not of interest to the bakery scientists. The bakery scientist is interested in the growth rates, the set times, and the maximum diameters of the cookies and comparisons of those characteristics among the formulations. Thus a model was constructed to describe the development pattern of each cookie. The two-stage linear-plateau model was selected to describe the observation from the jth cookie of ith formulation measured at the kth time point denoted by t_k, which can be expressed as

$$
y_{ijk} = \begin{cases} \delta_i + \beta_i(\gamma_i - t_k) + \epsilon_{ijk} & \text{if } t_k \leq \gamma_i \\ \delta_i + \epsilon_{ijk} \text{ if } t_k > \gamma_i \end{cases},
$$
$$
i = 1, 2, 3, 4, 5, \quad j = 1, 2, 3, \quad k = 1, 2, \ldots 10.
$$

The model parameter δ_i denotes the maximum diameter for the ith cookie formulation, the model parameter β_i denotes the growth rate for the ith cookie formulation, and the model parameter γ_i denotes the set time for the ith cookie formulation. The estimation procedure described by Hudson (1966) was used to fit the model to the data from each cookie of each formulation. Hudson's method is a search procedure consisting of partitioning the dataset into all possible sets of two adjacent partitions and then fitting a linear regression model to the first partition and a mean or model with zero slope to the second partition. If the two lines intersect in the interval separating the two partitions, the fits provide a possible solution. If the two lines do not intersect in the interval separating the two partitions, the two models are fit to the data where they are forced to intersect at one of the boundaries of the interval separating the two partitions. The partition of the data and the corresponding model with the smallest residual sum of squares is selected as the model to provide the least squares estimates of the parameters of the model. The cookie development experiment consisted of comparing five cookie formulations. Three cookies from each cookie formulation were baked for 10 min, thus the cookie is the experimental unit for cookie formulation. For each formulation, three batches of cookie dough were mixed, each independently of the others, and one cookie was extracted and baked from each batch. Thus each cookie was mixed and baked independently of the other cookies. There were 15 cookies, 3 from each formulation, and they were baked in a completely random order. Therefore the design structure of the experiment is a completely randomized design. The data are in Table 3.1.

Table 3.2 contains the estimates of the model parameters for each cookie, where Hudson's method was applied to obtain ordinary least-squares estimators (we did not attempt to determine the form of the covariance matrix). Since the cookies are mixed and baked independently of each other the estimated characteristics of the models are independent between cookies. Consequently, the slopes are all independently distributed, the set times are all independently distributed and the maximum diameters are all independently distributed. The analysis to compare the cookie formulations continues by comparing each of the model's characteristics using the analysis of variance procedure. The first step is to test the equality of the within cookie formulation variances for each characteristic using Levene's test. Levene's test (see Chap. 2 of Milliken and Johnson 1992) consists of (1) fitting the one-way analysis of variance model to the dataset, (2) computing the residuals for each observation in the dataset, (3) computing the absolute value of the residuals, and (4) performing a one-way analysis of variance on the absolute values of the residuals. If the resulting F test for treatments is significant at a preselected significance level, then reject the equal variance hypothesis and use a weighted analysis to compare the treatment means based on the original observations.

The results of the analysis of the three estimated parameters, the estimated slopes, the estimated set times, and the estimated maximum diameters are contained in Tables 3.3, 3.4, and 3.5, respectively. The top section of each table is the analysis of variance for that variable, the middle section of each table titled

Table 3.2 Estimates of the linear plateau model parameters for each cookie made from each formulation

Formulation	Slope or Growth Rate	Set Time	Maximum Diameter
Soft winter wheat with	4.30	8.09	94.6
granular sucrose	3.93	7.32	89.0
	4.28	7.53	92.6
Soft winter wheat with	4.21	4.88	80.1
fructose syrup	4.16	5.33	81.5
	4.43	5.25	82.5
Soft winter wheat with	4.27	5.62	82.4
glucose syrup	4.06	5.65	82.6
	4.52	5.49	84.7
Hard red winter wheat with	3.58	5.54	79.3
granular sucrose	3.54	5.64	79.8
	3.75	5.79	82.3
Soft winter wheat with	5.01	6.55	91.1
sucrose syrup	4.66	7.21	91.7
	4.58	6.77	89.0

Table 3.3 Analysis of the slopes for the cookie formulation example

| | | Slopes | | | |
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Formula	4	1.93677333	0.48419333	13.32	0.0005
Error	10	0.36340000	0.03634000		
Corrected total	14	2.30017333			

| | | Absolute Value of Slope Residuals | | | |
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Formula	4	0.01749926	0.00437481	0.67	0.6245
Error	10	0.06482222	0.00648222		
Corrected total	14	0.08232148			

T tests (LSD) for variable: Slope
Least significant difference = 0.3468
Means with the same letter are not significantly different

T Grouping	Mean	N Formula
A	4.7500	3 Sucrose syrup
B	4.2833	3 Glucose syrup
B	4.2667	3 Fructose syrup
B	4.1700	3 Granular sucrose
C	3.6233	3 Hard red winter

"Absolute Value of the Residuals of (variable name)" is the Levene's test for that variable, and the last section of each table consists of the means and comparisons of the means for the cookie formulations.

The results of Levene's tests indicate there is not sufficient evidence to conclude the variances are not equal ($p = 0.6254$ for slopes, $p = 0.1296$ for set times, and $p = 0.3973$ for maximum diameters). Thus, the one-way treatment structure in a completely randomized design structure model was used to compare the formulations through the characteristics. The top section of Tables 3.3, 3.4, and 3.5 displays the analysis of variance tables for each characteristic. There is strong evidence to conclude the formulations are different for all three characteristics ($p = 0.0005$ for slopes, $p < 0.0001$ for set times, and $p < 0.0001$ for maximum diameters). The formulations are compared using the 5% LSD (least significant difference) (see Chap. 3 of Milliken and Johnson 1992) values, which are in the bottom sections of Tables 3.3, 3.4, and 3.5. From Table 3.3, we see the sucrose syrup formulation grows significantly faster than glucose syrup, fructose syrup, granular sucrose, all of which grow faster than hard red winter wheat. From Table 3.4 we see the granular sucrose grows longer than sucrose syrup, which

Table 3.4 Analysis of the set times for the cookie formulation example

| Set Time | | | | |
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	12.81316000	3.20329000	45.49	0.0001
Error	10	0.70413333	0.07041333		
Corrected total	14	13.51729333			

| Absolute Value of Residuals of Set Time | | | | |
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	0.11713778	0.02928444	2.31	0.1296
Error	10	0.12703704	0.01270370		
Corrected total	14	0.24417481			

T tests (LSD) for variable: Set Time
Least significant difference = 0.4828
Means with the same letter are not significantly different.

T Grouping	Mean	N Formula
A	7.6467	3 Granular sucrose
B	6.8433	3 Sucrose syrup
C	5.6567	3 Hard red winter
CD	5.5867	3 Glucose syrup
D	5.1533	3 Fructose syrup

grows longer than the other three formulations. Hard red winter wheat also grows longer than fructose syrup, which does not grow significantly slower than glucose syrup. From Table 3.5, we see the maximum diameters of granular sucrose and sucrose syrup are not significantly different, but are significantly larger than the other three formulations. The other three formulations are not significantly different in the maximum diameter.

For the cookie development process, we learned a lot about the relationships between the five cookie formulations by using ordinary least squares to fit a simple, reasonable, and interpretable model to the data from each cookie. We are in essence letting the model fitting procedure provide the data that are used to compare the cookie formulations. This is the major point of the presentation. Even though the individual diameters measured over time from each cookie are highly and unequally correlated we can use the model to extract information about the growth process or about interesting characteristics of the growth curve, and then use the estimated characteristics to compare the cookie formulations.

Another possible approach is to use a multivariate analysis of variance to compare the formulations. For the multivariate analysis of variance to be applica-

Table 3.5 Analysis of the maximum diameters for the cookie formulation example

Maximum Diameter

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	350.03066667	87.50766667	27.83	0.0001
Error	10	31.44666667	3.14466667		
Corrected Total	14	381.47733333			

Absolute Value of Residuals of Maximum Diameter

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2.70696296	0.67674074	1.13	0.3973
Error	10	6.00518519	0.60051852		
Corrected total	14	8.71214815			

T tests (LSD) for variable: MAX DIA
Least significant difference = 3.2261
Means with the same letter are not significantly different.

T Grouping	Mean	N Formula
A	92.067	3 Granular sucrose
A	90.600	3 Sucrose syrup
B	83.233	3 Glucose syrup
B	81.367	3 Fructose syrup
B	80.467	3 Hard red winter

ble, there must be more multivariate observations per treatment group than characteristics measured. For the cookie example there are three characteristics and three cookies per formulation, thus the multivariate analysis of variance is not applicable. It should be noted that the maximum diameter is a function of the growth rate and set time. That is evident from the analysis as granular sucrose cookies grew longer than the sucrose syrup cookies, but the sucrose syrup cookies grew faster. But, the maximum diameters of the two sucrose cookie formulations were not significantly different. The dimension of this process in most likely two, i.e., knowing two of the characteristics, you will know about the third.

7. Example 2: Cumulative Radioactive CO_2 Production

The purpose of this laboratory study was to develop a method to investigate soil and environmental factors influencing spatial and temporal variability of pesticide degradation in field soils. Two specific factors were considered in the method development: (1) soil-handling process and (2) pesticide application technique. Three soil-handling/pesticide treatment combinations were evaluated

where treatment *A* was sieved soil/spray pesticide application, treatment *B* was injected pesticide application/sieved soil, and treatment *C* was injected pesticide application/intact core incubation. For treatment *A,* intact core segments sieved and carbofuran solution was applied using an atomizer. For treatment *B,* intact core segments injected with carbofuran at one location then sieved resulting in a more heterogeneous dispersion of pesticide. For treatment *C,* intact core segments were injected with carbofuran then incubated to more closely mimic field conditions by preserving soil structure and creating localized concentrations of carbofuran similar to the field result of banding of granular material at planting time. Ten replications of each treatment were prepared and the daily cumulative amount of radioactive CO_2 was measured. A plot of the cumulative CO_2 production data for each replication, revealed a sigmoidal type of relationship between the cumulative CO_2 production and the number of elapsed days. The treatments could be compared using a repeated measures analysis of variance where the daily measurements on each replication are the repeated measurements. But a more meaningful approach is to use a mathematical model to reduce dimensionality of the data from the number of repeated measures to the number of parameters in the model. This reduction enables the statistical evaluation of the treatments to be preformed in a more meaningful way by comparing the parameters of the models, where the parameters of the selected model have a direct biological or physical interpretation.

The logistic growth nonlinear model was used describe cumulative CO_2 production data for each replication as a function of number of elapsed DAYS. The logistic growth model used is

$$P_{ijk} = \frac{MAX_i}{1 + e^{(\alpha_i + \beta_i * DAYS)}} + r_{ij} + \epsilon_{ijk}, \quad i = A, B, C, \quad j = 1, 2, \ldots 10, \tag{5}$$

$$DAYS = 1, 2, \ldots, 11, 14, 16, 21, 28,$$

where P_{ijk} is cumulative CO_2 formation at the kth day of the jth replication of the ith treatment, MAX_i is the asymptote for the ith treatment, α_i is the location parameter for the ith treatment, and β_i is the rate coefficient for the ith treatment.

Radioactively labeled CO_2 production from all of the replications of the treatments exhibited sigmoidal kinetics. The process reaction was virtually complete after day 14 of a 30-day study and recovery of CO_2 ranged from 78% to 100%. Assessment of treatment differences by looking at the graphs of the data and fitted models were inconclusive, thus additional analyses of the data were required to enable decisions to be made about the models.

Model (5) was fit to the data for each replication of each treatment using PROC NLIN of the SAS® system to obtain the ordinary least-squares estimates

of the parameters of the logistic model. Table 3.6 contains the estimates for each replication of each treatment. The data from each replication is independently distributed from the data of all other replicates, thus the estimates of the parameters of the logistic model from each replication are independently distributed from the estimates of the parameters from all other replicates. Thus, a one-way treatment structure in a completely randomized design structure analysis of variance model in Equation (4) was used to compare the treatments through the characteristics of the model. The results of the analysis of variance for the parameter estimates are

Table 3.6 Estimates of the model's parameters for each replication of each treatment

| Treatment | Replication | Parameters | | |
		α	β	MAX
A	1	5.13927	−1.17853	64.0932
	2	5.27761	−1.00071	87.0985
	3	5.01829	−1.12973	87.0698
	4	4.94114	−1.08457	92.6543
	5	3.74832	−0.65412	78.9885
	6	5.03412	−1.06527	88.3552
	7	3.78861	−0.59266	84.0725
	8	4.47711	−0.75757	97.8895
	9	4.74424	−0.82437	98.6647
	10	5.46320	−1.06944	86.1166
B	1	3.69605	−1.05086	82.1051
	2	4.60804	−0.87880	92.2226
	3	4.59524	−0.64630	90.1819
	4	4.26751	−0.84761	90.9165
	5	3.90064	−0.76050	89.0598
	6	5.30245	−0.97201	94.2029
	7	5.76957	−1.10222	88.5235
	8	5.31191	−1.00131	89.4742
	9	4.03774	−0.58230	84.3643
	10	4.43914	−0.71483	92.0224
C	1	4.66019	−0.84229	89.6435
	2	4.09142	−0.64356	86.7260
	3	4.97761	−0.80121	91.8231
	5	3.69313	−0.59847	76.5151
	6	5.05195	−0.64119	78.3523
	7	3.63259	−0.48560	76.4144
	8	4.55604	−0.63734	86.6247
	9	4.53587	−0.74446	86.6653
	10	4.09640	−0.70131	91.0353

in Table 3.7, where there are no significance differences between the treatments for parameters α and MAX ($p = 0.3688$ and $p = 0.3956$, respectively). There are significant differences between the treatments for parameter β where treatments A and B are not different, but they are both significantly different from treatment C.

Predicted values for each replication of each treatment were computed at a selected set of DAYS using

$$P_{ijk} = \frac{\text{M}\hat{\text{A}}\text{X}_i}{1 + e^{(\hat{\alpha}_i + \hat{\beta}_i * DAYS)}}, \quad i = A, B, C, \quad j = 1, 2, \ldots 10. \tag{6}$$

Again, since the data from each replication are independently distributed from the data from all of the other replications, the predicted values for each replication evaluated after a specific number of elapsed DAYS are independent. Thus a one-way analysis of variance using Model (4) can be used to compare the treatments at each of those selected number of elapsed DAYS. The treatments are compared at each of the numbers of elapsed DAYS displayed in Table 3.8. Models for treatments A and B are significantly larger than the model for treatment C for DAYS 4 through 9 and the treatments are not significantly different for the other DAYS.

The number of DAYS of exposure required to yield 10%, 25%, 50%, and 75% of the estimated MAX for each replication was predicted using

$$\text{D}\hat{\text{A}}\text{YS}_{ijk} = \frac{-\hat{\alpha}_{ij} + log_e \left(\dfrac{\text{M}\hat{\text{A}}\text{X}_{ij}}{\%CO_2} - 1 \right)}{\hat{\beta}_{ij}}, \quad i = A, B, C, \quad j = 1, 2, \ldots 10 \tag{7}$$

The estimated numbers of days are in Table 3.9 and the analysis of variance results for comparing the treatments are in Table 3.10. The number of days to achieve 25%, 50%, and 75% cumulative CO_2 production are not significantly different for treatments A and B, which are significantly less than required by

Table 3.7 Comparisons of the treatments for the model parameters

Parameter	Significance Level	Treatment A	Treatment B	Treatment C
α	0.3688	4.7632	4.5928	4.3661
β	0.0102	-0.93570^a	-0.8557^a	-0.6773^b
MAX	0.3956	86.5002	89.3073	84.8666

[a]Values within a row with different letters are significantly different ($p = 0.05$).

Table 3.8 Comparisons of the treatments for the models evaluated at specific values of number of DAYS

Models Evaluated at DAYS	Significance Level	Treatment A	Treatment B	Treatment C
1	0.6603	1.960	2.390	2.211
2	0.5251	4.640	5.360	4.200
3	0.2135	10.800	11.500	7.844
4	0.0442	23.120^a	22.250^a	14.166^b
5	0.0077	40.950^a	37.590^a	24.000^b
6	0.0038	58.320^a	54.340^a	36.811^b
7	0.0059	70.690^a	67.960^a	50.355^b
8	0.0143	78.070^a	77.000^a	62.044^b
9	0.0378	82.120^a	82.450^a	70.711^b
10	0.0898	84.240^{ab}	85.580^a	76.467^b
11	0.1708	85.350	87.300	80.055
12	0.2456	85.900	88.260	82.122
15	0.3651	86.430	89.150	84.367
20	0.3925	86.510	89.300	84.822
25	0.3941	86.520	89.310	84.856
30	0.3941	86.520	89.310	84.856

[a]Values within a row with different letters are significantly different ($p = 0.05$).

treatment C. The treatments were not significantly different at $p = 0.05$ for the number of days required for 10% cumulative CO_2 production.

The information in Table 3.11 provides additional analyses for parameter β. This additional set of analyses provides tests for equality of variances among the treatment groups and then an analysis that fits an unequal variance model to the estimates of the β parameters. The SAS® system code in part (1) of Table 3.11 fits the one-way analysis of variance model to the estimated values of β, computes the residuals, indicated by the variable "rb," and stores the residuals in a data set called "res." The code in part (2) computes the absolute value of the residuals for each observation and the code in part (3) carries out the one-way analysis of variance on the absolute value of the residuals. The F value in the second section of Table 3.11 (dependent variable: ARB) corresponds to Levene's test of equality of variances, indicating the variances are most likely different ($p = 0.0403$). The Levene's test was used to test the equality of variances for all of the computed characteristics of the model and only the test for the estimates of β indicated a significant difference between the treatment variances ($p = 0.05$). The code in part (4) of Table 3.11 fits the unequal variance model to the data using the MIXED procedure of the SAS® system. The statement "REPEATED/GROUP=TRT" tells PROC MIXED to fit a different variance for

Table 3.9 Estimates of the numbers of days required to achieve a specified percentage of CO_2 accumulation for each replication

	Replication	Time to 10%	Time to 25%	Time to 50%	Time to 75%
Treatment A	1	2.93	3.98	5.44	·
	2	3.23	4.36	5.57	7.10
	3	2.63	3.64	4.71	6.06
	4	2.61	3.64	4.70	5.89
	5	2.78	4.55	6.56	10.20
	6	2.79	3.85	4.97	6.35
	7	3.01	4.94	7.04	9.96
	8	3.04	4.50	5.97	7.48
	9	3.11	4.44	5.79	7.15
	10	3.21	4.27	5.41	6.89
Treatment B	1	1.64	2.73	3.94	5.76
	2	2.85	4.12	5.44	6.92
	3	3.89	5.63	7.45	9.58
	4	2.57	3.89	5.27	6.86
	5	2.41	3.89	5.45	7.33
	6	3.26	4.41	5.58	6.86
	7	3.36	4.39	5.47	6.79
	8	3.23	4.36	5.54	6.95
	9	3.49	5.45	7.58	10.51
	10	3.27	4.83	6.45	8.28
Treatment C	1	3.07	4.40	5.81	7.47
	2	3.19	4.95	6.84	9.24
	3	3.59	4.99	6.44	8.08
	5	3.00	4.96	7.23	12.69
	6	4.88	6.70	8.76	12.73
	7	3.58	6.00	8.79	15.66
	8	3.95	5.73	7.64	10.07
	9	3.36	4.88	6.51	8.59
	10	2.86	4.46	6.12	8.04

each level of TRT or each level of the GROUP=variable. The option DDFM= SATTERTH requests that PROC MIXED use the Satterthwaite approximation to the degrees of freedom for all statistics concerning the comparison of the means of the treatments.

The results from PROC MIXED are in the bottom three sections of Table 3.11. The center section provides estimates of the variances for each of the treatments as 0.0443, 0.0313, and 0.0118 for treatments *A*, *B*, and *C*, respectively. The "Tests of Fixed Effects" provides a test of the equality of the treatment means of the estimated β values where the approximate degrees of freedom for

Table 3.10 Comparisons of the treatments for numbers of days required for a given percent of CO_2 to accumulate

Amount of Cumulative CO_2 Production	Significance Level	Treatment A	Treatment B	Treatment C
10%	0.0587	2.934	2.997	3.498
25%	0.0082	4.217[a]	4.370[a]	5.230[b]
50%	0.0052	5.616[a]	5.817[a]	7.127[b]
75%	0.0088	7.456[a]	7.584[a]	10.286[b]

Values with a row with different letters are significantly different ($p = 0.05$).

the denominator are 19.5, the F statistic is 7.53, and the significance level is 0.0038. This test for the equality of the treatment means indicates that the estimated β treatment means are most likely different. The fourth section contains the adjusted means for the treatments and the corresponding estimated standard errors. The degrees of freedom associated with each of the estimated standard errors corresponds to the sample size minus one, i.e., the degrees of freedom associated with the estimated of the variance from each of the treatments. The bottom section of Table 3.11 displays the pairwise comparisons of the treatments, indicating treatments A and B are not different, whereas both are significantly less than treatment C. The degrees of freedom were are computed using the Satterthwaite approximation discussed in Section 5.

As a contrast to the unequal variance model fit above, the estimates of β were analyzed using an equal variance model and the results in Table 3.12. The significance level of the equal variance F test for equality of treatment means is 0.0102, which is less than the significance level of the corresponding unequal variance F test, 0.0403. But, the major differences between the two analyses are in the estimated standard errors of the treatment effects. The estimated standard errors in Table 3.12 are equal (except treatment C's mean is based on 9 observations, whereas treatments A and B means are based on 10 observations). The estimated standard errors for the treatment effects in Table 3.11 reflect the magnitude of the respective estimates of the variances. When the variances are unequal, the unequal variance analysis of variance is the appropriate method to compare the treatments.

Tables 3.13, 3.14, and 3.15 contain the sample means, standard deviations, and covariance matrices of the estimates of α, β, and MAX from each replication for treatments A, B, and C, respectively. These statistics were used to construct confidence bands for each treatment's model and to construct confidence bands for the differences for each pair of models, using the methods in Section 5. The models are graphed in Figure 3.2 where MODA, MODB, and MODC are the

Table 3.11 Analysis of the estimates of β using the unequal treatment variance model

(1) proc glm data=est; class trt rep; model b = trt/solution; output out=res r=rb;
(2) data res; set res; arb=abs(rb);
(3) proc glm; clas trt; model arb=trt;
(4) proc mixed data=est; class trt; model b=trt/ddfm=satterth; repeated/group=trt;
     lsmeans trt/diff;

Dependent Variable: ARB

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	0.04634296	0.02317148	3.64	0.0403
Error	26	0.16536060	0.00636002		

Covariance Parameter Estimates (REML)

Cov Parm	Estimate	Std Error	Z	Pr > |Z|
DIAG TRT *A*	0.04430749	0.02088675	2.12	0.0339
DIAG TRT *B*	0.03129916	0.01475457	2.12	0.0339
DIAG TRT *C*	0.01183279	0.00591639	2.00	0.0455

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
TRT	2	19.5	7.53	0.0038

Least-Squares Means

Level	LSMEAN	Std. Error	DDF	T	Pr > |T|
TRT *A*	−0.93569631	0.06656387	9	−14.06	0.0001
TRT *B*	−0.85567437	0.05594565	9	−15.29	0.0001
TRT *C*	−0.67726959	0.03625954	8	−18.68	0.0001

Differences of Least Squares Means

Level 1	Level 2	Difference	Std. Error	DDF	T	Pr > |T|
TRT *A*	TRT *B*	−0.08002194	0.08695208	17.5	−0.92	0.3699
TRT *A*	TRT *C*	−0.25842672	0.07579910	13.8	−3.41	0.0043
TRT *B*	TRT *C*	−0.17840478	0.06666836	15.1	−2.68	0.0172

estimated models for treatments *A, B,* and *C,* respectively. Confidence bands for the models for treatments *A, B,* and *C* shown in Figures 3.3, 3.4, and 3.5, respectively. The confidence band for the model becomes a confidence interval about the MAX parameter as the number of DAYS becomes large. The confidence band for the differences of models for treatments *A* and *B* is in Figure 3.6, the band for the differences of models for treatments *A* and *C* is in Figure 3.7, and the band for the differences of models of treatments B and C is in Figure 3.8. The treatments are declared to produce different levels of the response at a

Table 3.12 Analysis of the estimates of β using the equal treatment group variance model

Dependent Variable: β

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
TRT	2	0.32812028	0.16406014	5.50	0.0102
Error	26	0.77512213	0.02981239		
Corr Total	28	1.10324242			

TRT	β, LSMEAN	Std. Err., LSMEAN	LSMEAN Number
A	−0.93569631	0.05460072	1
B	−0.85567437	0.05460072	2
C	−0.67726959	0.05755422	3

	Pr > |T| HO: LSMEAN (i) = LSMEAN (j)		
I/j	1	2	3
1	.	0.3096	0.0031
2	0.3096	.	0.0332
3	0.0031	0.0332	.

Table 3.13 Means, standard deviations, and covariances of the set of estimated parameters for treatment A

	$\hat{\alpha}$	$\hat{\beta}$	$\widehat{MAX}$
MEAN	4.7632	−0.9357	86.5003
STD	0.5896	0.2105	9.9120
Cov $\hat{\alpha}$	0.3476	−0.1114	−0.0160
Cov $\hat{\beta}$	−0.1114	0.0443	0.5135
Cov $\widehat{MAX}$	−0.0160	0.5135	98.2474

Table 3.14 Means, standard deviations, and covariances of the set of estimated parameters for treatment B

	$\hat{\alpha}$	$\hat{\beta}$	$\widehat{MAX}$
MEAN	4.5928	−0.8557	89.3073
STD	0.6771	0.1769	3.6545
Cov $\hat{\alpha}$	0.4585	−0.0605	1.2390
Cov $\hat{\beta}$	−0.0605	0.0313	0.0157
Cov $\widehat{MAX}$	1.2390	0.0157	13.3556

Table 3.15 Means, standard deviations, and covariances of the set of estimated parameters for treatment C

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\text{MAX}}$
MEAN	4.3661	−0.6773	84.8666
STD	0.5166	0.1088	6.1509
Cov $\hat{\alpha}$	0.2669	−0.0375	1.4345
Cov $\hat{\beta}$	−0.0375	0.0118	−0.5375
Cov $\hat{\text{MAX}}$	1.4345	−0.5375	37.8339

specific number of DAYS when the confidence band does not include zero. From Figure 3.6, there is possibly a difference between treatments A and B around DAYS = 5, but the models are not different elsewhere. The graph in Figure 3.7 shows that treatment A produces a larger response for DAYS 2 to 10 and from Figure 3.8, treatment B produces a larger response than treatment C for DAYS 2 through 13.

The preceding discussion of this example demonstrate several possible paths to take to compare the three treatments. The important aspect of this discussion is you should compute those characteristics of the model that are meaningful to your being able to understand the process being studied. Compute those characteristics and carry out an appropriate analysis on those characteristics.
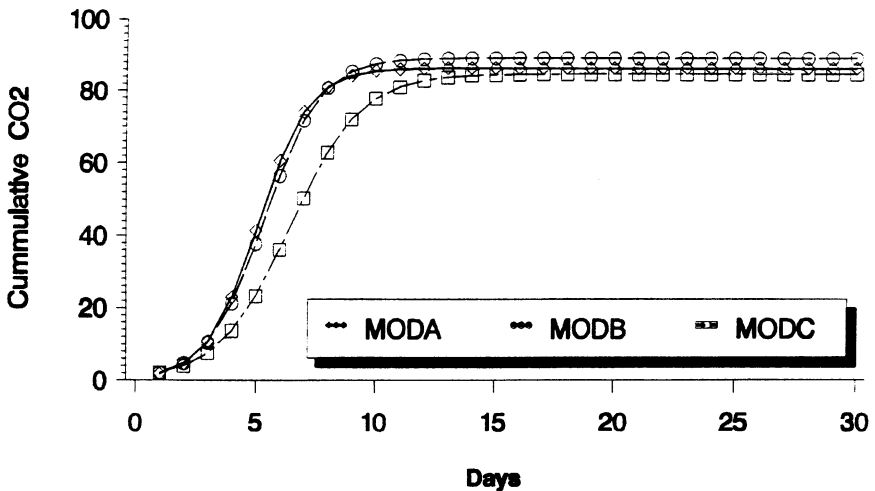
Figure 3.2 Overlay of all three treatments.
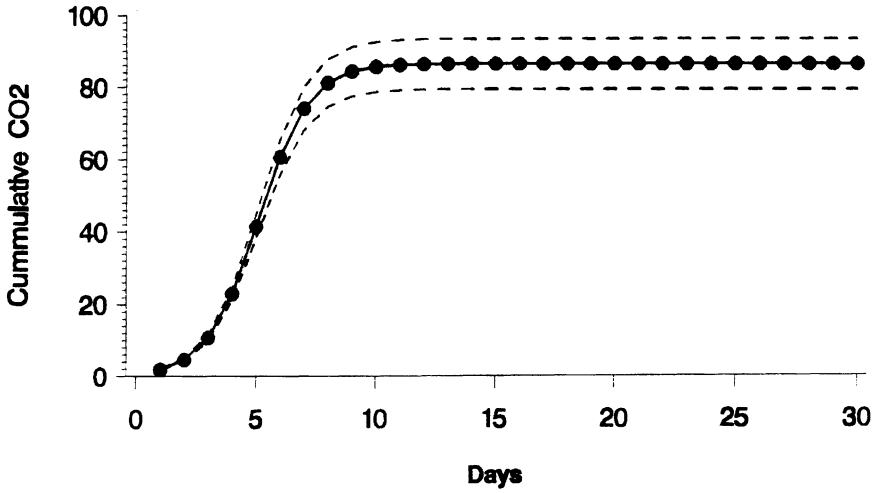
Figure 3.3 The 95% confidence band for treatment *A*.
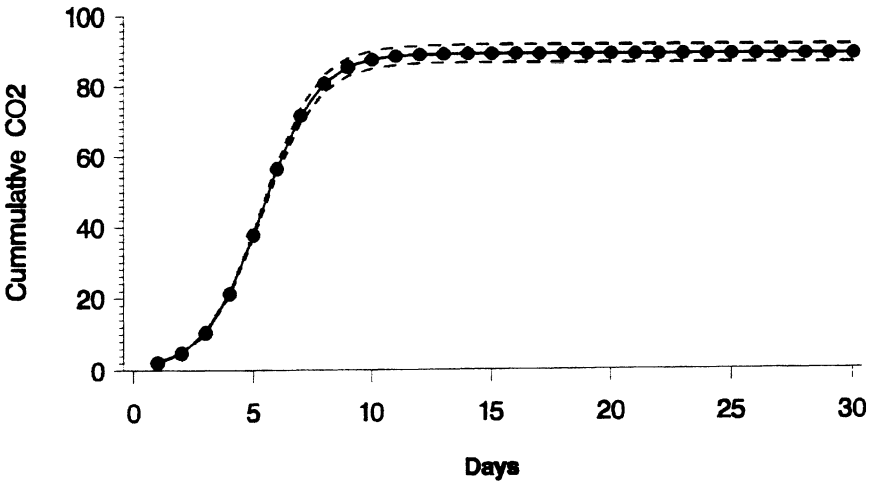
Figure 3.4 The 95% confidence band for treatment *B*.
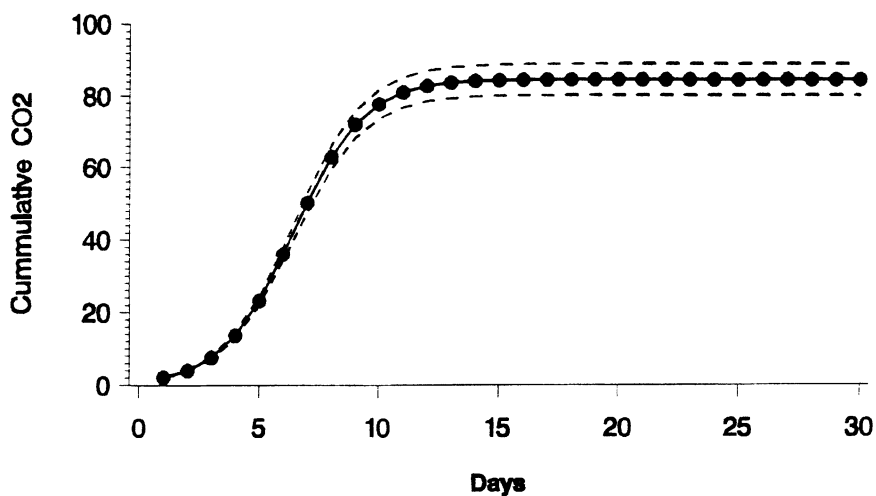
Figure 3.5 The 95% confidence band for treatment *C*.
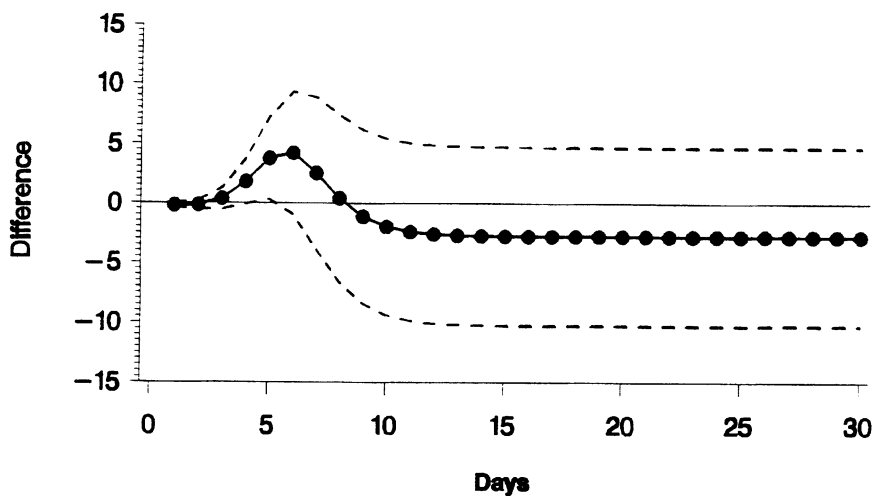
Figure 3.6 The 95% confidence band for the difference between treatments *A* and *B*.
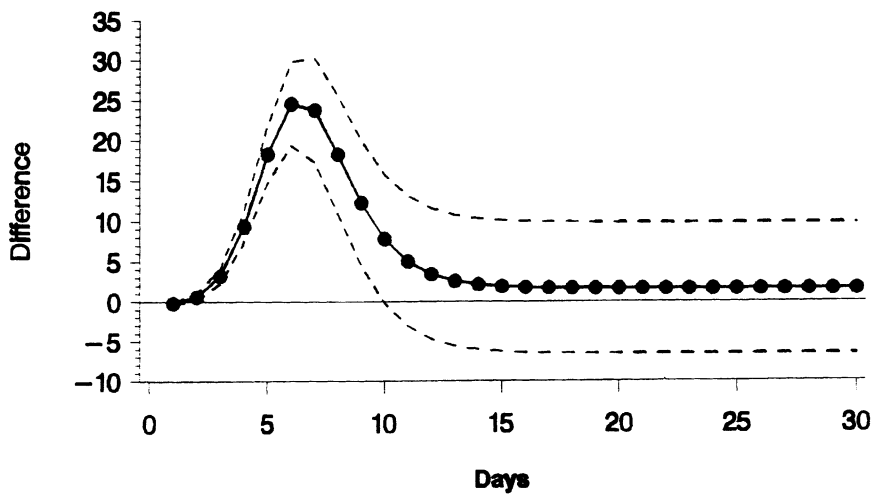
Figure 3.7 The 95% confidence band for the difference between treatments *A* and *C*.
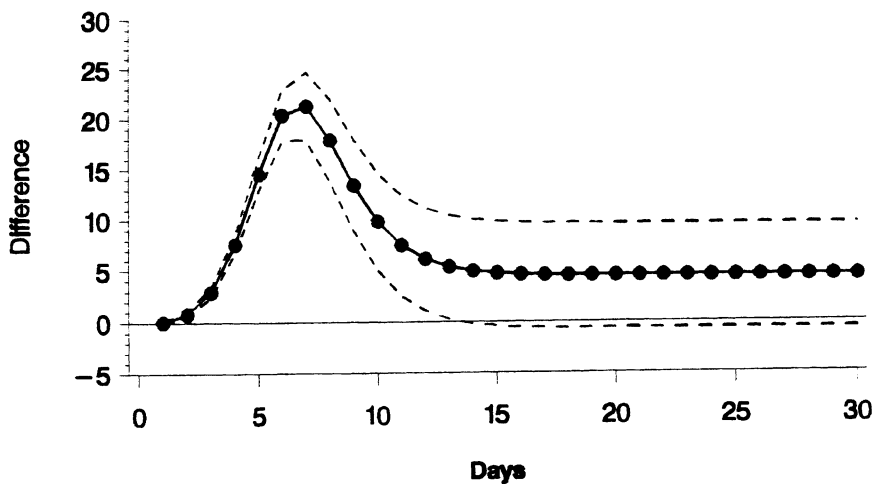
Figure 3.8 The 95% confidence band for the difference between treatments *B* and *C*.

8. Summary

Methods are described for analyzing data consisting of repeated measures from each of several experimental units subjected to one of a set of treatments under study. The process consists of fitting a model to the repeated measures obtained from each experimental unit using ordinary least squares. The important characteristics of the models are then computed for each experimental unit and the treatments are compared using an appropriate analysis of variance. Procedures are described for both equal treatment variances and unequal treatment variances. The unequal treatment variance case is expected to be the case for most situations, thus a test of the equal variance hypothesis is required. A confidence band for a model and a confidence band for the difference of two models are constructed for both the equal and unequal variance cases. The procedures are demonstrated by analyzing the data from two examples.

References

Doescher, L. C. 1986. Factors Affecting Surface Cracking in Sugar-Snap Cookies and a Mechanism for Cookie Dough Setting. Ph.D. Dissertation, Department of Grain Science and Industry, Kansas State University, Manhattan, KS.

Hudson, D. J. 1966. Fitting segmented curves whose join points have to be estimated. *J. Am. Stat. Assoc.* 61:1079–1129.

Milliken, G. A., and D. E. Johnson. 1992. *Analysis of Messy Data: Design Experiments, Vol. I.* Chapman & Hall, London.

Parkin, T. B., D. R. Shelton, and J. A. Robinson. 1991. Evaluation of methods for characterizing carbofuran hydrolysis in soil. *J. Environ. Quality* 20:763–769.

SAS Institute Inc. 1996. *SAS/STAT® Software: Changes and Enhancements through Release 6.11.* SAS Institute, Inc., Cary, NC.

4

The Monod Model and Its Alternatives

Arthur L. Koch

1. Jacques Monod: His Life and Work

1.1. The Logistic Equation

Jacques Monod started life as an ecologist. Logically, he follows after the Belgian ecologist, Verhulst. Verhulst (1838) developed the logistic equation for the growth of populations. It can be derived in a variety of contexts. Possibly the simplest way to understand his model of the mathematics of the population growth process is to think of a trypsinogen solution being converting into trypsin. If completely pure, the solution would be stable, but if one molecule of trypsin were present as a contaminant, then it would convert trypsinogen into trypsin and that would produce trypsin faster and faster—a simple autocatalytic process. Finally, production would be slowed by the consumption of the trypsinogen. These stages could be collectively modeled:

$$dT/dt = rTS \quad \text{or} \quad dT/dt = rT(K - T), \tag{1}$$

where T is the concentration of trypsin; S is the substrate concentration, which in this special case is the concentration of trypsinogen; r is the second-order rate constant for the conversion of trypsinogen to trypsin; and K is the initial concentration of the substrate, trypsinogen. On integration and evaluation of the boundary conditions, one obtains

$$T = T_0 K e^{rt}(K + T_0 e^{rt}), \tag{2}$$

where T_0 is the initial concentration of trypsin. After a long time, T becomes equal to the initial trypsinogen concentration K. I have used K because ecologists use this symbol and call it the carrying capacity. This equation is the famous autocatalytic, "logistic" equation that Raymond Pearl (Pearl and Reed 1920) used to successfully predict the population growth of the United States. A plot of it gives a symmetrical S-shaped curve. It is the basis, moreover, of much modern ecological thought, with r now equal to the maximal rate of net population growth and K, as mentioned, equal to the carrying capacity of the ecosystem; i.e., the population level when birth and death are equal (see, for example, Ricklefs 1990).

Different fields have a fixity on which symbols to use; here we deal with several fields so first we rewrite the logistic in symbols more common in bioengineering, where the biomass is taken to be X. The logistic becomes in differential and integral form:

$$dX/dt = rX(K - X)/K, \tag{3}$$
$$X = X_0 K e^{rt}/(K + X_0 e^{rt}), \tag{4}$$

and has limiting expressions of

$$X = X_0 K e^{rt} \quad \text{at early times,} \tag{5}$$
$$X = K \quad \text{at late times.} \tag{6}$$

This is shown in Figure 4.1.

1.2. Monod's Doctoral Thesis (Part 1): The Electronic Photometer; Bacteria Instead of Plants and Animals

In the late 1930s, Monod attempted to experimentally test this equation for his dissertation in zoology under Teissier at the Sorbonne. For plants or animals, the time scale was too long, but for bacteria a few hours would suffice. The number of plants or animals could be counted at different time intervals, but bacteria, being small, could not. Monod seized on the photoelectric colorimeter, which fortunately had recently become available. Although the instrument was made to measure the absorption of light by colored solutions, the light scattered by the turbid solution of bacteria was a signal that could be related to the mass of bacteria. This gave an additional advantage over counting individual organisms in a microscope counting chamber or enumerating colony forming units on petri plates. In turbidity measurements, since millions and millions of organisms were present in the sample, no random error from this source would accrue.

Using the primitive photometer, he studied the growth of assorted bacteria by inoculating various media with an infinitesimal number of bacteria and following
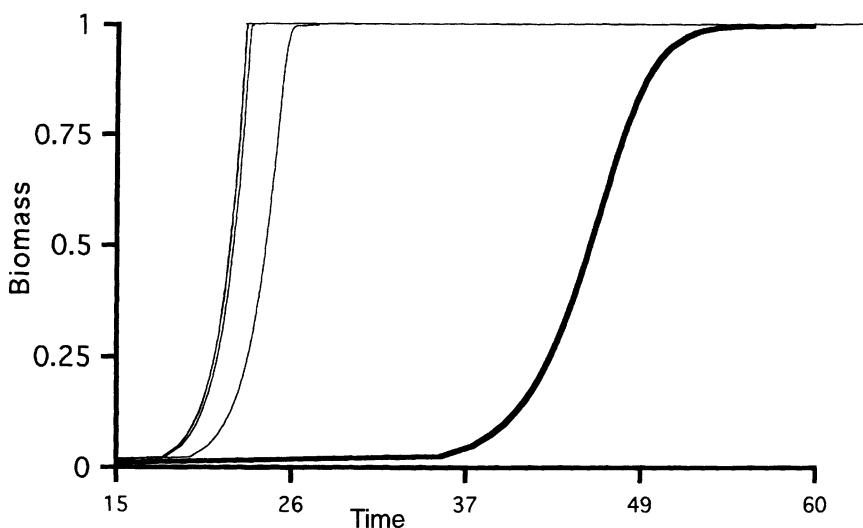
Figure 4.1 The logistic and Monod integral equations. Theoretical growth curves are shown for the logistic equation by a thicker line and the Monod model with thinner lines for various choices of the value of K_m. The Monod model approaches the logistic when K_m is large so that $\mu = \mu_{max}S/(K_m + S)$ approaches the first-order relationship, $\mu = (\mu_{max}/K_m)S$, where growth is directly proportional to substrate concentration. The K_m values for the first three curves are 10^{-6}, 10^{-5}, and 10^{-1}. The thick line is for $K_m = 10^4$ and also for K_m equal to infinity.

the growth until a stable population was reached. This plateau was due to the consumption in a minimal medium of a limiting single substance present in the medium, such as, for example, glucose. To his surprise growth did not follow the logistic equation. Instead the curve was not symmetrical, and growth continued at the same exponential rate well past the consumption of half of the glucose; then, however, it did abruptly slow and stop. World events temporarily slowed the experiments, but not the thinking about their significance.

1.3. Bacterial Population Growth

Perhaps the analogy to the trypsin/trypsinogen case was not carried far enough. In 1838, Verhulst had only a simplistic idea of the possible interactions, and, of course, had never heard of trypsinogen. We imagined earlier that the interaction of the substrate and the enzyme was a second-order process. But Michaelis and Menten (1913) had shown that invertase action on sucrose was only first order in the substrate at low concentrations and that the rate saturated (zero order in substrate) at high concentrations. Monod, a passionate Frenchman of Huguenot

ancestry by way of Milwaukee, Wisconsin, and Geneva, Switzerland, preferred to attribute the idea of enzyme kinetics to the nineteenth-century Frenchman, Henry. (My reading, however, gives Michaelis and Menten clear priority). With the enzyme as the paradigm, now we would attribute a maximum velocity, V_{max}, and a Michaelis/Menten constant, K_m, to the enzyme trypsin. Note that K_m is a constant for fixed conditions for a given simple irreversible enzyme and numerically is that concentration that gives half the maximal reaction velocity. Then r in Equations (1) through (5) is replaced with $V_{max}/(K_m + S)$, where S is the substrate concentration. Since if S is small this reduces to a constant V_{max}/K_m, the Verhulst equation would apply when very low amounts of trypsinogen were initially present. Conversely, it would give the property to the system that when a high initial concentration of trypsinogen was initially present there would be two stages of growth: exponential growth followed by quick slowing and stopping of further growth. This mathematical treatment of the improved model required more elaborate algebra (presented below), but did lead to results that fitted the data. Monod's idea and treatment became very popular and it is the Michaelis-Menten enzyme kinetic equation with new symbols and new meaning to the application to bacterial growth that is *the* Monod growth equation (Monod 1942, 1949); perhaps it is most commonly symbolized in microbiology, microbial ecology, and bioengineering by

$$\mu = \mu_{max}S/(K + S), \tag{7}$$

where μ is variously called the specific growth rate or growth rate constant. Here μ_{max} has the units of the amount of product created per unit time per unit amount of enzyme, since, in the case of bacterial growth, the product is the equivalent of more enzyme, the amounts cancel, and the dimensions are usually expressed as per hour, but other reciprocal units of time as well. Also, μ_{max} is the maximum growth rate at high, saturating substrate concentration.

By substituting this equation into the Verhulst equation (in doing so we equate μ for r, use X for the biomass of the system, and assume a yield coefficient of Y, equal to the ratio of dry weight of cells produced per gram of growth substrate), using the partial fraction technique and again substituting limits Monod (1942; or see Koch 1985) obtained the equivalent of

$$t\mu_{max} = [(K_m + S_0 + X_0/Y)/(S_0 + X_0/Y)] \ln(X/X_0) \tag{8}$$
$$- [K_m/(S_0 + X_0/Y] \, ln[(S_0 + X_0/Y - X/Y)/S_0].$$

Note that it is not possible to rewrite this equation to calculate the time course of X; rather we must try different values of X and calculate the times that correspond to them. Methods to fit this equation to parameters and their errors

have been developed by Robinson and Tiedje (1983) and by Simkins and Alexander (1985). The limiting forms of this equation are

$$dX/dt = \mu_{max} \; S_0/(K_m - S_0) \tag{9}$$

and

$$X = X_0 e^{\mu_{max} t/(K_m + S_0 + X_0/Y)} \tag{10}$$

at early times, and

$$X = K, \tag{11}$$

at late times. In any case, the problem is solved and solutions for various values of K_m are shown in Fig. 4.1, with the simplistic assumption that a bacterium behaves like a hydrolytic enzyme.

1.4. Monod's Thesis (Part II): Diauxie

Maybe the analogy between trypsin/trypsinogen and bacteria/glucose was too simplistic and a more diverse environment of substrates was appropriate. So Monod studied a variety of two-carbon-source media, and again he found unexpected results—not those anticipated from the results with single carbon sources nor from the logistic equation.

With two sources, such as lactose and glucose, *Escherichia coli* first consumed the glucose and then after an interval started to consume the lactose. This kind of result he observed with a number of carbon source pairs. With *B. subtilis* the result was even more dramatic. This is because starvation can lead this organism's autolysins to cause self-destruction of the bacteria (it is starting to sporulate). Thus, after the consumption of the one carbon source instead of having a hiatus in which the turbidity remained constant until new growth was commenced, there was a dip in the turbidity curve. These led Monod to coin a new term, "diauxie," meaning two foods. This was the phenomenon where the organism fully consumed one "preferred" source before starting on the "less palatable" resource. Although this is a very human response, unraveling the mechanism for it required the rest of Monod's lifetime with the help of enumerable colleagues and workers in other laboratories. However, this elucidation led to the foundations of modern molecular biology.

1.5. World War II and the Resistance Movement

Monod earned his doctorate in late 1940 during the German occupation of Paris. He became a serious worker in the resistance movement in the spring of 1943 largely because it took that time after the occupation for an effective movement to develop. He joined the Communist party to become active in the Franc-Tireurs et Partisans until the liberation. The work was dangerous, but important (see Judson 1996), but he retained bench space in Lwoff's laboratory at the Pasteur Institute, well away from the Sorbonne where the gestapo were looking for him.

1.6. The Origins of Molecular Biology

The phenomena of induced enzyme formation and the idea of messenger RNA came in good part from his studies with the lactose operon. The work demonstrated the existence of operator regions in the DNA, repressor genes, DNA promoter regions (Jacob and Monod 1961). All were discovered and the basis understood (wholly or in part) with the *lac* system. I will briefly recount only those developments relative to uptake and growth because the audience of this chapter is presumed not to be committed bacterial physiologists but committed microbial ecologists and biostatisticians. But the physiological basis in the various cases greatly affects the shape of the growth curve at both the populational and at the single-cell level, and defines and constrains the type of mathematical fitting to experimental data.

First, since the cells could "learn" to grow on lactose, it should be possible to find mutants that fail to do so, and vice versa. Actually it was known that such variants do occur naturally. Such strains were called *E. coli mutable* because frequently *lac*-positive cells were thrown off as the cells grew.

Initially Monod selected noninducible strains from populations of inducible cells. In fact, Monod isolated *E. coli* strains from the feces of his graduate advisor, Andre Lwoff (Monod and Audureau 1946). One strain, the third isolate was naturally *lac* negative, did not exhibit diauxie, and was revertible to *lac* positivity. It was designated ML3 (merde Lwoff #3). He isolated from ML3 a lactose-positive mutant, ML30, which was inducible for lactose utilization and exhibited diauxie. When grown on lactose, like most *E. coli* strains, it created β-galactosidase, which splits lactose into the readily utilizable sugars, galactose and glucose. β-galactosidase was readily shown to be a soluble enzyme that could be isolated, purified, and had the properties generally associated with proteins and enzymes.

Now he looked for mutants of ML30 that lacked the enzyme; i.e., revertants that were not *lac* positive on indicator (dye containing) petri plates also containing lactose and enough other substances so that all cells formed colonies. Most of these were indeed deficient in β-galactosidase. This was confirmation for bacteria

of the "one-gene:one-enzyme" hypothesis of Beadle and Tatum in 1940 formulated based on types of mutations in *Neurospora crassa.*

1.7. Discovery of Permease

During the war, with the help of his coworkers, Monod, went on to notice that there were a subset of the *lac*-negative strains that were different from the majority type. Although the colonies had the appropriate color on special plates, which showed that they did not ferment lactose, they did contain β-galactosidase. When grown on lactose, β-galactosidase could be demonstrated by lysing the cells in a mixture containing toluene to destroy the integrity of the cell envelope to a sufficient degree that it lets the substrates for β-galactosidase into the cell. What could this class of mutant be? The answer came in the key papers of Rickenberg et al. (1956) and Cohen and Monod (1957): There was a second protein, an integral membrane protein, necessary for the transport of lactose through the cell membrane. This protein did not chemically alter the lactose. Monod had a knack for words (his addition to the words now integral to the biological sciences include diauxie, induced enzyme, gratuity, permease, crypticity, operon, promoter, and allosteric enzyme) and dubbed this protein "permease" and the strains with this character as being "cryptic."

1.8. Gratuitous Inducers and Nonmetabolizable Substrates: Thiogalactosides And Uptake Kinetics

The problem of induced enzyme synthesis (then called "adaptive enzymes") was obscure until inducers were developed that could not be metabolized by the bacteria. Monod reasoned that if the critical oxygen glycosyl bridge between the two saccharides was replaced with a sulfur atom, the structural chemistry would be almost the same, but the biochemistry would be different and β-galactosidase would not be able to act to produce consumable sugars. Even though not substrates, some of these synthetic compounds made by Monod's colleague, Mel Cohn, served to induce the formation of β-galactosidase (Monod, Cohen-Bazire, and Cohn, 1951). Thus, such compounds would serve as "gratuitous" inducers; i.e., they would enable the study of induction without the complication of their serving also as an energy source. We do not follow the consequence of these studies in the development of molecular biology, but note that radioactive molecules of this class enabled the study of permease action. It was found that radioactive thio-methyl-galactoside, made by Georges Cohen (Rickenberg et al. 1956; Cohen and Monod 1957), was concentrated hundreds of times above the concentration in the medium by cells containing the permease, but the cellular concentration was not higher than that in the medium with the cryptic *lac*-negative mutants. Moreover, the galactoside was in a dynamic equilibrium in

that nonradioactive galactosides would cause the efflux of the radioactive galactoside, even if the non galactose part was chemically different.

With two genes, β-galactosidase and permease, serving in the same metabolic pathway, Monod's group went on to find the nature and organization of these two and regulatory genes that grouped together and allowed coordinate function; again we do not pursue the spin-off from these findings and the operon concept.

2. The Monod Model and Its Derivations

2.1. Uptake Kinetics

At the supply level, the organism must have uptake systems on its periphery to bring necessary substances inside in sufficient amounts for growth. The elementary formulation for the permease model can be made as follows: the substrate outside the cell, at concentration S_0, binds to the cytoplasmic-membrane-bound permease, P, to form a complex that "resolves" to become the product, which is the same substrate molecule but is now present inside the cell. The concentration of galactoside inside the cell is designated by S_i. This reaction could be studied because the workers now had in hand radioactive thiomethyl-galactoside which, like lactose, is a substrate for the permease, but is not cleaved by the β-galactosidase (because, as explained, of its sulfur-containing analogue of oxygen at the critical glycoside bond). The Institut Pasteur workers also had ways (by membrane filtration) to separate the cells from the medium faster than substrates could permeate the cell membrane.

The gene for permease, *lac*Y, has since been cloned, sequenced, mutated, and the version in other organisms studied. It has also been the subject of many other kinds of studies. Today we understand that the protein is a membrane protein that coils through the bilayer 10 times with its sequence of 200 amino acids. This structure has the properties of binding a galactoside with any of a range of aglycone residues in either the α or β configuration and binding stoichiometrically a proton on the same face of the cytoplasmic membrane as the galactoside binds. This form converts to the same structure; i.e., one with both the proton and the galactoside still in the *cis* arrangement but now on the opposite side of the inner membrane. The rule of the permease is that neither a proton by itself nor a galactoside molecule by itself can traverse the membrane. In effect, the galactoside permease is a machine that normally functions to pump galactosides from outside to inside powered by the protonmotive force; i.e., by the excess of protons pumped outside the cell by an electrogenic process of proton extrusion (chemiosmosis). But that was not known in the middle 1950s and the kinetics appeared to Monod and his colleagues as follows:

$$S_0 + P \overset{k_1}{\underset{k_2}{\Leftrightarrow}} SP \overset{k_3}{\Rightarrow} P + S_i. \tag{12}$$
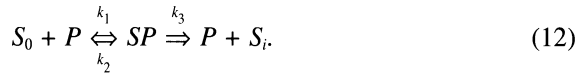
This is very closely the case if the protonmotive force is constant. This mechanism translates into mathematics as

$$d[SP]/dt = k_1[S_0][P] - (k_2 + k_3)[SP]. \tag{13}$$

In the steady state (technically a quasi-steady state) where the rate of entry is balanced by outflow, the two terms on the righthand side are equal and this expression for the rate of entry can be set to zero, yielding

$$K_{en} = [S_0][P]/[SP], \tag{14}$$

where $K_{en} = (k_2 + k_3)/k_1$, with the subscript "en" designating the entry process. By calling the total amount of permease, p; i.e., $p = [P] + [SP]$, we can then add $[S_0]$ to both sides of the second form of the equation, and write

$$K_{en} + [S_0] = [S_0]([P]/[SP] + 1) = [S_0]([P] + [SP])/[SP], \tag{15}$$
$$= [S_0]p/[SP], \tag{16}$$

or

$$[SP] = [S_0]p/(K_{en} + [S_0]) \tag{17}$$

The rate of the entry process is equal to the rate of dissociation of the complex to yield the internal galactoside as a product, therefore,

$$v = k_3[SP] = k_3p\ [S_0]/([S_0] + K_{en}) \tag{18}$$

Then by calling k_3p as V_{en}, again with the subscript designating the entry process, we obtain

$$v = k_3[SP] = V_{en}[S_0]/([S_0] + K_{en}) \tag{19}$$

In this way, we have rederived the classical 1913 Michaelis-Menten equation in the 1924 Briggs and Haldane elaboration in the context of bacterial uptake as proposed in the 1956 derivation by Monod (Rickenberg et al. 1956). Again it can be written with the microbiologist's symbols as

$$\mu = \mu_{max}S/(K + S). \tag{20}$$

The point of the 1913, 1924, and 1956 derivations of the basic law of enzyme and microbial growth kinetics is that useful results are obtained even though we have no idea of actual concentrations of enzymes or membrane-bound components. This is in spite of the fact that we have used square brackets, such as [S_i], to indicate concentrations as is the convention of chemists. Actually, in many cases today, we can know the concentrations for soluble enzymes of known molecular weight and known weight concentrations, but this is still actually a presumptuous symbol to be used for a material embedded in the cytoplasmic membrane. That, however, was the important point, our ignorance does not matter because we end up with an expression in which all the quantities are measurable, the concentration of substrates and rates of reaction, so that the useful and significant quantities, V_{en} and K_{en}, can be extracted from well-designed experiments.

2.2. Elementary Derivation of Hyperbolic Dependency of Growth on Substrate

Let me give one more derivation and unfortunately one more symbol: we have used X for the collective biomass. But the X concept does not relate to the number of cells and their surface; so let us chose B for the number of bacteria, it is both the product and the catalyst. So we can write the quasi-chemical process as:

$$B + S \Leftrightarrow SB \Leftrightarrow B + B. \tag{21}$$
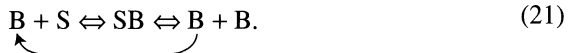
Thus 1 B becomes 2 Bs, and S is converted to B (with a conversion factor of Y).

This scheme suggests that a bacterium is a catalyst that binds to the substrate and converts it into a product, just as an enzyme does, only the product is another bacteria. We assert that the many reactions of intermediary metabolism are irrelevant; we will need a strong defense to maintain these theses, not for low substrate concentration, but for high concentrations. Thus we assert with confidence that a bacterium will be limited in finding substrate at low concentrations and with less confidence be limited by the rate it can convert it into the newly created product bacterium at high concentrations. Then our argument is just the same as in the Michaelis-Menten formulation for enzymes. Now, however, the possibility of a connection between growth and the uptake capability is clear. But at this point, the reader should be confused about the whole cell being the equivalent of an enzyme and at the same time each permease molecule in the cytoplasmic membrane being an enzyme.

3. Limitation of the Hyperbolic Model

The formulation $\mu = \mu_{max}S/(K + S)$ should strictly apply only if a single irreversible transport system exhibiting saturation kinetics limits growth at all substrate concentrations. The Monod equation would not apply when multiple transport systems function; then a composite with several Michaelis-Menten constants would be needed. Nor would it apply for a single transport process that exhibited kinetics more complex that those of a one-way carrier model. Finally, and most commonly, it would not apply if other secondary processes besides transport limit growth at high substrate. From this point of view and because of other restrictions to be mentioned soon, it is perhaps surprising that hyperbolic kinetics fit so many sets of growth rate data as well as they do.

Questions about the applicability of simple hyperbolic kinetics to bacterial growth were raised by Powell (1967); Koch (1972); and Dabes et al. (1973). These workers derived alternate models applicable when the maximum transport capability for the growth resource under study exceeds the cell's needs for maximal growth in that environment. Powell (1967) and Dabes et al. (1973) performed statistical tests showing that most sets of experimental data did not fit the Monod model well. The approach in all these studies was to assume that the specific growth rate is a hyperbolic function of the internal concentration of the nutrient, but that as an alternative kinetic model for the transport of the nutrilite into and out of the cell and that transport was of the passive diffusion type.

3.1. Evolution and Uptake Systems

The extensive use of the Monod hyperbolic simple model has been particularly noteworthy in the fields of microbial ecology and in biotechnology. A hyperbolic relationship for growth/transport has been used in many thousands of publications describing complex ecosystems or considering the design and operation of fermentation systems. It has been particularly useful to dissect multiple processes. Although it is satisfactory and useful in many cases, it is theoretically and actually wrong in almost all cases!

It is right for transport in the restricted case in which proton extrusion (or other source) is never limiting, the permease is always limiting, and the permease molecules do not complete with each other on the surface of the cell. (It also requires that the diffusion up to the cell, that the surface to volume ratio, and that the concentration of permease per unit of cell surface are not factors. These effects are considered later.) But imagine that it was once right in an earlier eon; how would evolutionary forces respond and modify the process? Whenever the permease's substrate was limiting, selection would favor better or more numerous permeases (Koch, 1972). Eventually the transport capability would become so great that the entry rate would be magnified and that entry at low substrate
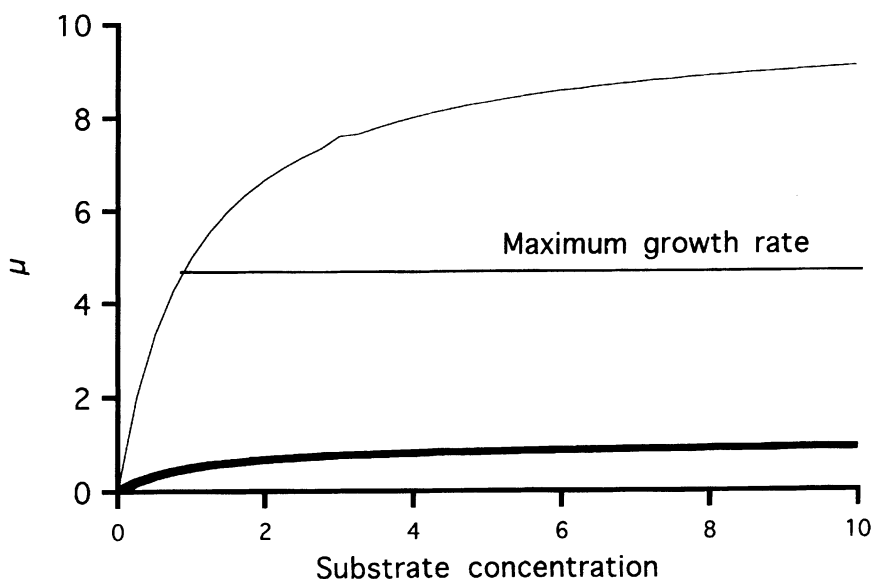
Figure 4.2 Evolution of an uptake process. The thick line describes a hyperbolic Monod-type dependency on substrate concentration. During times of stress selection would act to increase the quality and/or quantity of permease systems in the membrane, as shown by the upper curve. However, during times when substrate is abundant, growth has an upper bound limited by other cellular processes. This will result in a discontinuity or a near discontinuity as the substrate is increased to the point where uptake yields a rate satisfies the cell's needs. This is the basis of the Blackman law.

concentrations would be very steeply proportional to substrate concentration (compare the original lower hyperbola to the upper curve corresponding to increased transport ability in Fig. 4.2). Then even at a quite low concentration, the substrate would be taken into the cell at the rate that more than suffices for all the cell's needs (Koch 1967, 1972, 1996). The cell would need ways to shut off the uptake or let the excess pass out of the cells.

4. The Blackman (1905) Model and the Best (1954) Model

4.1. Diffusion Enzyme Model

These ideas fit into a model devised by Blackman (1905) even before the concepts of enzyme kinetics had been formulated. For Blackman limitation, uptake is characterized by a second-order rate constant for uptake (first order in

the substrate); i.e., k_1 in the preceding formulation. But then there is a discontinuity and uptake is squared off when $v = V$, where V refers to the maximum ability of the organism to grow given an unlimited internal supply; i.e., there is no longer a supply-sided limitation, but a limitation on the demand side (see Fig. 4.2). Note that V must be less than V_{en}, but otherwise is unrelated to it. The equations of the two branches of uptake are then

$$v = k_1 S, \quad S \leq V/k_1: \quad \text{Blackman, low S,} \tag{22}$$
$$v = V, \quad S > V/k_1: \quad \text{Blackman, high S.} \tag{23}$$

The first branch is the substrate-limited case (supply-side limitation) and the second branch the growth is limited by the cell's internal physiology (demand-side limitation). Of course, we do not know *a priori* how far evolutionary processes for any given case in any given organism have gone, so therefore we must resort to refined experiments to establish the kinetics in particular cases.

But before we consider experimental results, we first consider an intermediate formulation by Jay Boyd Best (1955), which was developed a year before the "Monod permease" equation (Rickenberg et al. 1956). It considers an enzymatic process in tandem following a diffusion process. I'll not go through the algebra, which is also given in Koch and Coffman (1970) and Koch (1985), but the Best solution is

$$v = V(S + K + J)\{1 - [1 - 4SJ/(S + K + J)^2]^{1/2}\}/2J. \tag{24}$$

Equation (24) is called the diffusion:enzyme model as well as the Best equation. This is more complex than the "Monod permease" uptake equation but much simpler than models considered next. The extra parameter is J, which is defined as $J = V_{max}/(AP)$, where A is the surface area of the bacteria and P is the permeability constant, i.e., the diffusion constant in the membrane divided by the thickness of the membrane. When J is very small the equation reduces to the Monod equation and when J is very large it becomes the two limbs of the Blackman equation (see Fig. 4.3). This same equation was later independently derived by Powell (1967), by Koch and Coffman (1970), and Koch (1972) and by Dabes et al. (1973) and applied to microbial growth and by Zimmermann and Rosselet (1977) for a pharmacological application. Sample curves for the gradation between the extreme Blackman and Monod models are shown in Fig. 4.3.

4.2. Alternative Couplings of Uptake with Consumption

The two cases considered so far have a natural coupling of uptake to consumption. In the Monod case, uptake is linked to consumption because the hypothesis
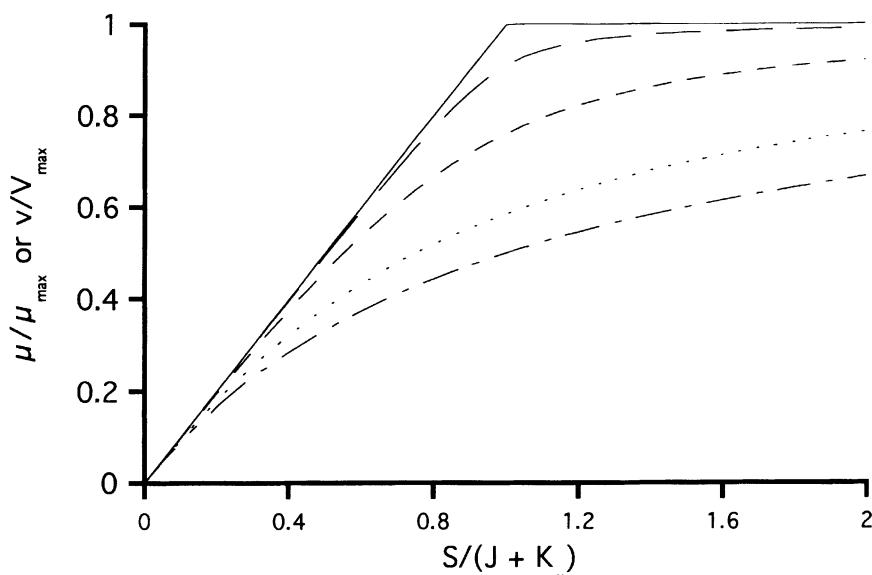
Figure 4.3 Dependence of uptake rate on concentration for the Best equation. The specific growth rate is shown for a number of cases for growth in which a pure diffusion process through the wall is followed by irreversible hyperbolic consumption for growth. The Best equation (Eq. [24]) via its parameter J, ranges from the Monod case when J is small (lowest lines) to the two branches of the Blackman case when J is large (uppermost line).

is that uptake system at the cell surface always limits growth and that all molecules taken up by the cell will be converted into cytoplasm no matter how many enter. If entry is by passive reversible diffusion, as in the Best case, there is a natural coupling between uptake and growth. If uptake should exceed need, then the internal concentration of substrate, waiting to be used, would increase and this would increase efflux and slow net diffusion. This is probably the reason that this model best fits the experimental data. But then it is inconsistent with the fact that uptake is a process that is catalyzed by an enzyme-like permease in many cases.

Logic therefore compels us to assume that an additional element is needed to do this. I proposed in accordance with ideas of Cohen and Monod (1957) an enzyme, leak:enzyme model in which uptake had the hyperbolic kinetics of a permease followed by irreversible consumption for growth, but the system was self-regulated because there was also a passive leakage, or efflux, through the cytoplasmic membrane. Thus when the internal concentration is high, substrate would go back to the outside. We see that this does not fit the data quite as well as the diffusion:enzyme model of Best.

Other types of regulation are imaginable and their kinetics are illustrated in Figures 4.4, 4.5, 4.6, and 4.7. They involve more parameters that the four of the Best or the enzyme, leak:enzyme models. These are all variants on the reversible permease process preceding an enzymatic consumption process. In the steady-state net entry and consumption are equal. The rate of a reversible enzymatic process is given by

$$v = [(V_{en}/K_{en})S_0 - (V_{ef}/K_{ef})S]/[1 + S_0/K_{en} + S/K_{ef}], \tag{25}$$

Where "en" stands for entry and "ef" stands for efflux, and S is the internal concentration. In terms of consumption by an irreversible process, S is given by the inversion of the standard Michaelis/Menten expression as
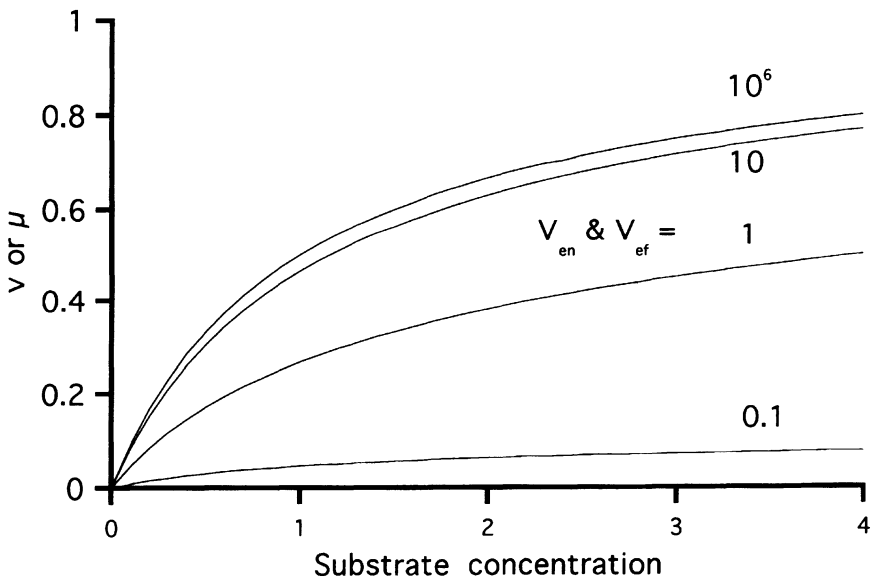
Figure 4.4 Simulated facilitated diffusion. The rate of entry (and hence the specific growth rate) dependence for a model of passive reversible carrier preceding irreversible consumption is shown; results calculated from Eq. (27). When K_{en} and K_{ef} are chosen large, the equation simplifies to the Best case Eq. (24). Curves for high and low membrane capacity are shown. For a reversible catalyzed, but non-energy-requiring transport through the cytoplasmic layer, V_{en} and V_{ef} were varied together, while K_{en} and K_{ef} were kept constant at 1. Here V and K for consumption were both fixed at 1. It can be seen that whether transport limits growth or consumption limits uptake and growth, the curves are nearly hyperbolic and differ only slightly from the Monod model.
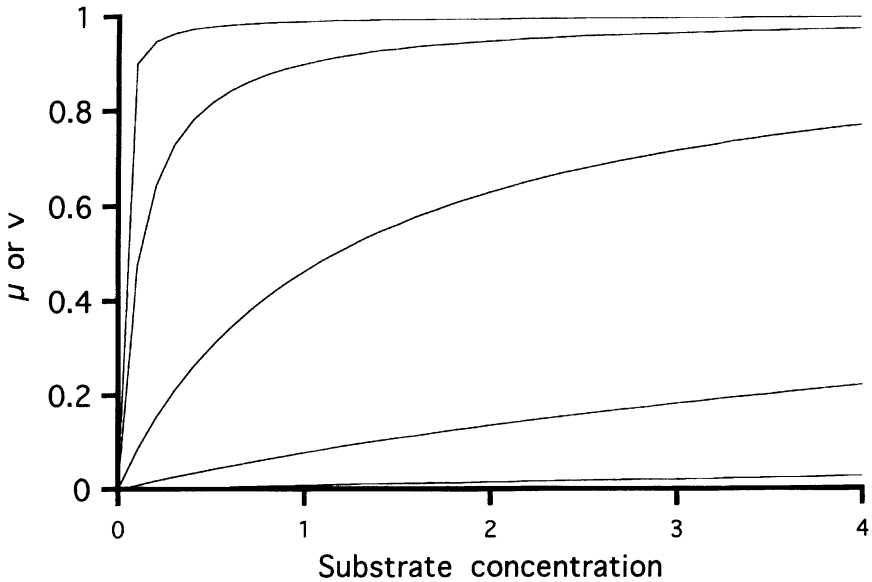
Figure 4.5 Simulated active transport. Uptake rates or growth rate dependencies for a type of energy coupled of transport kinetics proceeding to an irreversible hyperbolic consumption process are shown. This corresponds to the transition between the Monod case and the Blackman case during evolutionary course in the development of an uptake system. As the rate of entry increases the hyperbolic (Monod type) kinetics changes to the case where entry can be rapid so that the internal substrate concentration is so high as to achieve the maximum specific growth rate, giving nearly the two-phase (Blackman-type) growth kinetics. The figure shows from lowest to highest the effect of increasing V_{en} from much less than V to greater than V.

$$S = vK/[V - v]. \tag{26}$$

Between the two expressions S can be eliminated to give

$$(-K_{ef}K_{en} + KK_{en} - K_{ef}S_0)v^2 \tag{27}$$
$$+ (VK_{en}K_{ef} + V_{ef}KK_{en} + V_{en}K_{ef}S_0 + V_{ef}S_0)v - V_{en}VK_{ef}S_0 = 0$$

This quadratic expression is reprinted from Koch (1972) because there were typographical mistakes in the previous publication. Figures 4.4 to 4.7 were calculated for appropriate values of the parameters by a program that solved the quadratic equation and selected between the two roots, based on the requirement that the value of S be positive and that the value of y could not exceed the smaller of V and V_{en}.
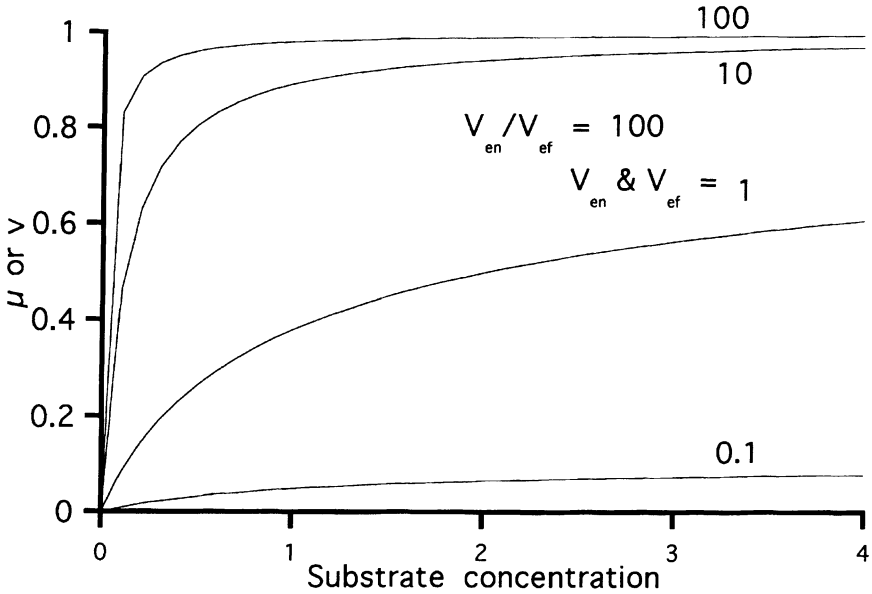
Figure 4.6 Simulated variable amounts of active transport. By maintaining the ratio of entry rate to efflux rate at 100 and the Michaelis constants for entry and efflux equal, active transport is simulated (see Fig. 4.5). As the level of the permease system is varied, the curves vary from hyperbolic at 0.1 to nearly the Blackman case at 100.

Figure 4.4 shows this reversible carrier-catalyzed process preceding irreversible hyperbolic uptake with the constants chosen so that the kinetics of entry would be exactly those of efflux; i.e., $V_{en} = V_{ef}$ and $K_{en} = K_{ef}$, but the ratio of the transport process to the consumption varies from curve to curve. This is the case of facilitated diffusion. It is apparent that no matter whether transport or consumption limits growth the specific growth rate is almost hyperbolic in external substrate concentration. The half-saturating concentration varies from that of the K_m of the consumption process at low and high values of the transport capability upward by a factor of less than 2 at intermediate values of the transport capability. Figure 4.5 shows the effect of energy coupling, simulated by increasing the entry rate alone; similar curves (not shown) result if either V_{ef} or K_{ef} progressively decrease. Figure 4.6 shows the effect of varying the transport capability relative to the consumption process in an energized system. Finally Figure 4.7 shows the effect of product inhibition on an irreversible permease system.

It would be possible when still better data are available to fit these formulations to the experimental results and analyze the problem further; however, since the number of independent parameters is six in Figures 4.4 through 4.7, it is doubtful that further analysis of growth curves alone could possibly distinguish the cases.
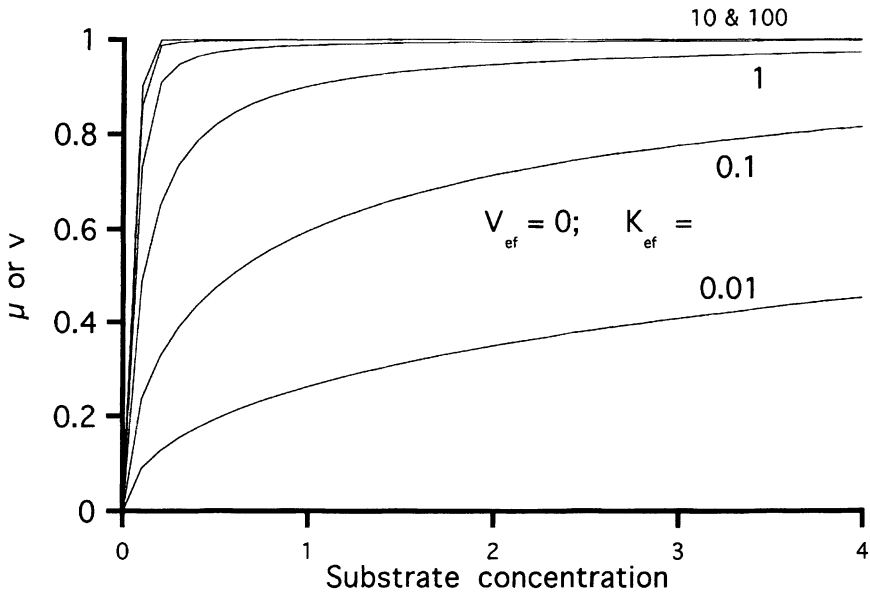
Figure 4.7 Simulated product inhibition. Reaction products must necessarily inhibit the enzyme that produces them. By setting $V_{ef} = 0$, an active transport system was simulated, and by varying K_{ef} various degrees of product inhibition were produced; $V_{en} = 10$; $K_{en} = 1$; $V = 1$; and $K = 1$ were the other parameters, as in most of the cases in Figs. 4.4 to 4.7.

The results of actual experiments of the most extensive and accurate experimental measurements of growth where independent rates of growth are presented as a function of residual glucose concentration are shown in Figures 4.8 and 4.9. These data were fitted to a variety of possible models. Note that in the analysis (shown in Fig. 4.8) the fitting was marginal for a model with only four independent parameters. Data much more precise would be needed to make more elaborate fits.

   An extended treatment of the coupling of consumption with uptake has recently appeared (Koch, 1997).

5. Still More Complication: The Phosphotransferase System

   Although lactose is accumulated inside of permease-containing cells through the use of energy generated directly from chemiosmosis in the form of a protonmotive force or a proton electrochemical gradient, some of the other small molecules are accumulated by the cell with energy provided directly from high energy phosphate bonds. The most energy-consuming case is that involved in the phos-
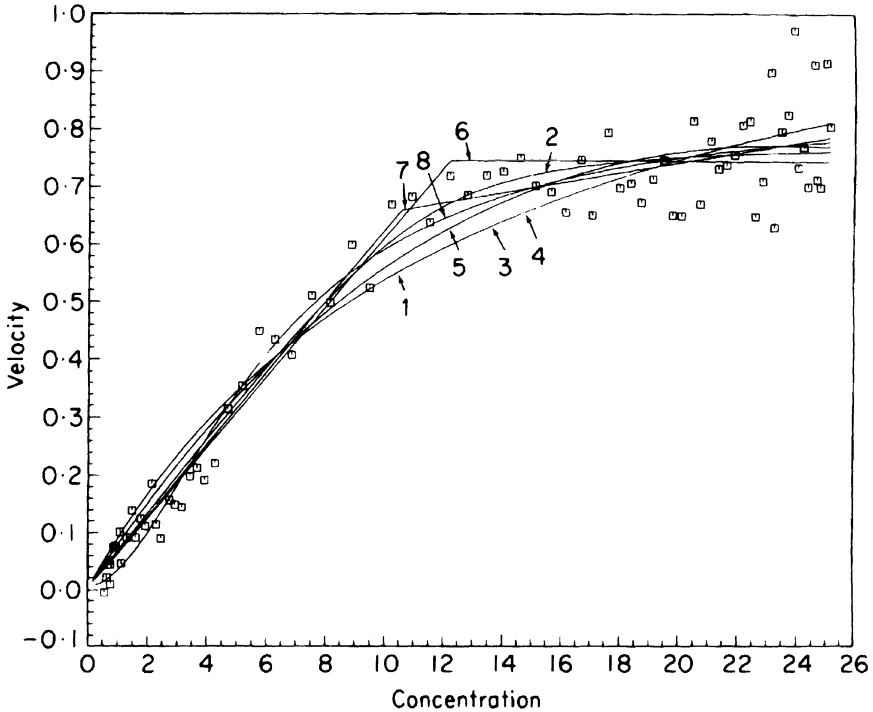
Figure 4.8 Glucose consumption by *E. coli.* Experimental data first published in Koch and Wang (1982) are reproduced. In this experiment, growth was measured by computer regression of absorbance measurements in a double-beam spectrophotometer. The glucose concentration was calculated for each point by subtracting from the original concentration the amount consumed in forming the bacteria that were then present. Curves for the best fits to the Blackman and Monod models and the intermediary Best model that could encompass them both are shown. The models are: 1, Michaelis-Menten; 2, diffusion:enzyme (this is the Best equation); 3, enzyme, leak:enzyme; 4, concomitant enzyme (positive); 5, 2 commitant enzymes; 6, bilinear (Blackman mode); 7, trilinear through the origin; 8, Hill equation (allosteric case).

photransferase system (PTS) (see Postma et al. 1992, Medow et al. 1990). It starts with phospho-enol-pyruvate (PEP). The free energy of hydrolysis of this compound is due to the fact that after the phosphate is removed the resultant enol pyruvate tautomerizes to the more stable keto form, thus pulling the reaction. The negative free energy of hydrolysis of PEP is greater than that of ATP. In fact, there is an important pathway that generates ATP from PEP. For the PTS transport process, PEP's phosphate is transferred first to Enzyme I (a soluble protein) and then to HPR (a small heat-stable protein), which in turn phosphorylates one of a number of different Enzyme II species. Each of these is specific
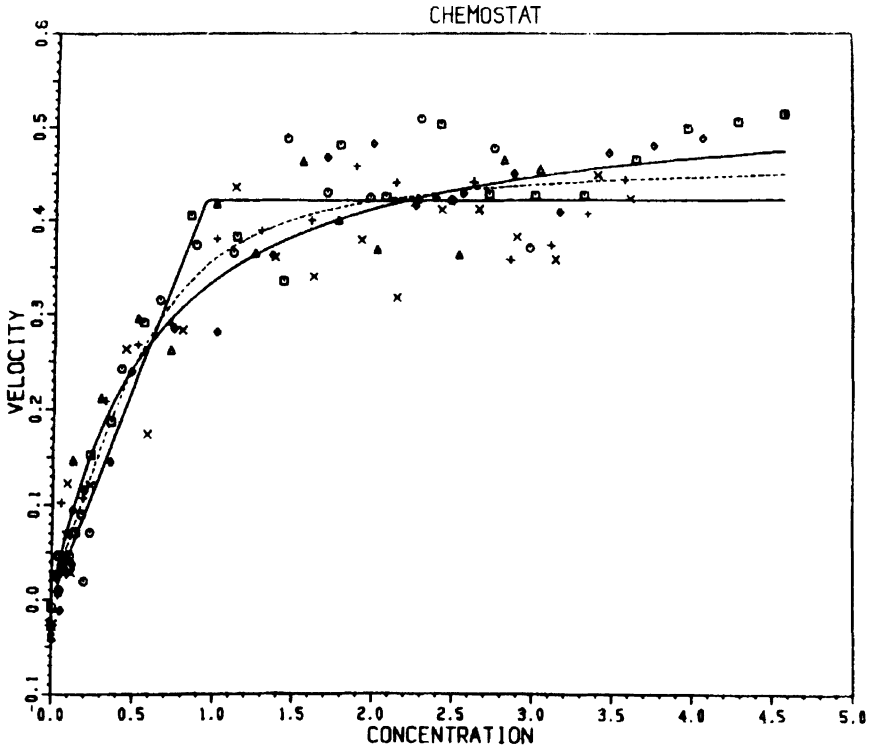
Figure 4.9 Glucose consumption after selection by chemostat growth. After extensive selection for growth in the chemostat under glucose limitation a new strain (probably an *ompR*) mutant was selected for that led to only the formation of *ompF*. Its ability to consume glucose at very low concentrations was enhanced (see text). Reproduced from Koch and Wang (1982).

for a different individual sugar. These proteins are membrane bound and carry out "group translocation" and "vectorially phosphorylate" sugars. These terms mean that the sugar is not only transported across the cytoplasmic membrane into the cell, it is also esterified to become, for example, glucose-6-phosphate. As such it is immediately available to enter into the metabolism of the cell via the Embden-Myerhof-Parnas lactic acid fermentation pathway and into several other pathways as well.

There is still one more component of the PTS and from the point of view of this article it is this one that is most relevant. This is Enzyme III^{Glu}, a soluble protein which is specific for glucose. When it is phosphorylated as part of the chain, it directly inhibits the galactoside and other permeases and thus prevents accumulation of galactosides and other inducing sugars. This is the phenomenon

of inducer exclusion and is the mechanistic explanation of diauxie because this inhibition prevents cells growing in a mixture of lactose and glucose from acquiring the lactose needed to induce the transcription of the *lac* operon to produce the galactoside permease to become further induced. (Another part of the story has to do with the role of the β-galactosidase in the induction process; see Koch 1996a). The β-galactosidase is necessary to convert some of the lactose into a compound that is actually the inducer). The kinetics of the PTS system at low substrate *in vivo* were hard to establish and various estimates have been published (see Koch and Wang 1982). This discrepancy in the literature is part of the reason for assembling and using the apparatus described in Section 7.

6. Still More Complications: The Kinetic Contribution of Porins and Passage through the Outer Membrane

The outer membrane of gram negative organisms is organized so that one leaflet has a hydrophilic covering, whereas the phospholipids in the other leaflet are hydrophobic. Seemingly this sandwich should prevent small molecules whether they are hydrophobic or hydrophilic from penetrating. However, to selectively allow the entry of certain compounds, bacterial cells have devised porins that allow water soluble molecules with molecular weights of up to 700 to pass through. In *E. coli* there are two porin proteins regulated by *ompR*, which controls production of the products of genes *ompF* and *ompC*. The pore produced by the latter is somewhat smaller than the former.

The simple fact of an additional barrier is important because it impedes transport even though the molecular weight of glucose (180) is much smaller than the sieving size of the porins. Also the switch between the two porin types has profound influence on the kinetics at low substrate. These two effects will be analyzed in Section 11.

7. The Experimental Measurement of Glucose Consumption

In the late 1970s Koch and Wang (1982) carried out experiments to actually measure the growth kinetics at very low limiting carbon source concentrations. To do this we employed a device previously built (Wang and Koch 1978). This was a long-path-length (10-cm) flowthrough cuvette maintained at constant temperature with water flowing through its jacket as well as in the walls of the spectrophotometer chamber. The apparatus was arranged so that the growing culture was circulated through the cuvette into a small reservoir by an air lift system in such a way that air bubbles did not enter the optical path. The spectrophotometer was a highly precise double-beam Cary 16 spectrophotometer that

was very stable. The absorbance signal was sampled more than once a second and sent to a minicomputer. The system was so sensitive, precise, and accurate that in only 200 sec, it could calculate the growth rate to better than one percent. For the actual experiments, a culture of a derivative of ML30; i.e., ML308, was grown and subcultured in a minimal (mineral salts) medium in 0.02% glucose. This is indeed a low concentration of carbon source and the 10-cm-light-path cuvette was needed to achieve adequate sensitivity, but even a forty fold dilution of this was made into medium containing no carbon source at all. After this dilution the final glucose concentration equivalent was only 27 μm, was so low that the growth rate progressively decreased. We could show that the yield of bacteria per unit amount of glucose was the same as at higher glucose concentrations so we could use the turbidity data to calculate the residual glucose concentration. Our actual data for one case for glucose consumption by *E. coli* ML 308 are shown in Figs. 4.8 and 4.9 (Koch and Wang 1982; Koch 1982ab). The experimental data was fitted to a variety of models which are discussed later.

8. Selection of a Mutant Growing More Avidly at Low Glucose Concentrations

An important facet to the problem of growth at low substrate is the diffusion of substrate through the medium up to the cell. There is, consequently, a maximum rate that cells could grow at low substrate concentration even if they had infinitely great and avid transport capabilities. If every molecule of glucose colliding with a bacterium was taken up and consumed, then no faster growth would be possible unless the culture medium was super efficiently stirred. It could be shown (Koch 1971) that the cell's own motility would not aid significantly and that very strong shear forces would be needed to increase uptake sufficiently. The question then is How close to the diffusion limit do, or can, cells operate? For this reason, the strain used in Figs. 4.8 and 4.9 was grown in a chemostat with a dilution rate of $D = 0.06$ hour^{-1} (doubling time = 11.56 hour) for 34 days. In the chemostat the actual concentration of the nutrient in the growth medium is extremely small due to the consumption. As a result a mutant with a more avid uptake system will take over the culture. The results of growth measurement on this culture are shown in Fig. 4.9. It can be seen that the initial slope is greater than in the unselected culture. The interpretation of these findings is presented in Section 10.

9. The Data Fitting: The Role of Models

The word "models" has quite different connotations to workers in different fields. On one hand, a model is a mathematical formalism that may describe the

actual data reliably and, on the other, it is an abstraction for a real physical or biological situation. Both types are needed. Thus faced with the data of which a sample is shown in Figure 4.8, there are two types of approach that can be made. We could say that the specific growth rate versus substrate concentration could be modelled as a polynomial and that the statistical job is to try to fit, by the method of least squares, the data with a series of polynomials each extending to higher and higher powers. Then it is by the use of statistical criteria that a decision can be made as to which is the smallest polynomial that reflects the data sufficiently accurately. In the same vein, we could try other mathematical equations and again choose the one that performed the best statistically.

Alternatively, we can frame biological models such as the Monod, the Blackman, and the Best, but not those used to produce Figures 4.4 through 4.7 because they have too many parameters. The diffusion:enzyme or Best formulation is an example in which the underlying mechanism of action was imagined and translated into equations. Its fit to the data is shown by the curve marked 2 in Figure 4.8 and by the dashed line in Figure 4.9. However, this general approach is limited: (1) because we can not imagine all the relevant details, although in fact, many of them are irrelevant to the understanding of the general process; (2) because generating a mathematical description may be difficult, exceed our mathematical capability, or it may be impossible to write the model down in mathematical terms; and (3) because we may not be able to execute the statistical fitting of the available data to the model if they have too many parameters.

Perhaps the most important constraint in this type of model is that the sample of data does not cover the range of dependent variables sufficiently. Thus, in the sample depicted in Figure 4.8, the data do not show whether or not there is a lower limit to the glucose concentration at which growth stops. We tacitly assumed that growth stopped when the glucose concentration became zero and supported the assumption in a variety of ways. However, a threshold has been observed with a number of oligotrophs (see Button 1985) and even with *E. coli* in slow chemostat culture (Koch and Coffman 1970). Evidently important biology is related to this possibility.

A major problem with the first modeling procedure is that the mathematical form is generally chosen both for ease of fitting and appropriateness to the sample of data at hand and usually independent of more general properties of the experimental system. Thus, for example, on the high substrate end of the growth curve of the experiment shown in Figure 4.8 the glucose concentration is less than 27 μM, and although a power series fits this limited range of data well, it is known from other experiments that at 100 or a 1000 times larger glucose concentration the growth rate would be nearly the same. If that fact is added as a constraint then no power series is appropriate because none reach a positive constant value at a very large value of the independent variable. This is a sufficient reason to reject the power series fitting for this type of problem.

Actually to extend the argument further, at very high concentrations of sugars osmotic effects would become evident and growth would be inhibited and then none of the models considered in this paper would be counted acceptable.

So I argue that for *my purposes*, and probably those of many microbiologists, the data need to be interpretable in terms of some biological mechanisms (entry into cells and to cell metabolism) but not others (thresholds, osmotic effects). In this way, I would justify using the simple Monod formulation to interpret many ecological studies, with the recognition that it is only approximate and that there are many various ways to generate curves, as shown above, that are nearly hyperbolas. I also point out that deviations can be important under certain conditions, especially when the problem at hand is to delve deeper into the cell's physiology.

What should we do? We wish to study the substrate dependence of growth and wish to isolate and interpret the dominant features of the process. We should admit to the possibilities of a threshold concentration of nutrient for growth and that high concentrations may be inhibitory and work between these limits. Doing this does include the idea that at low concentrations, the limiting process is the organism finding the substrate and is therefore basically a second order-reaction, and that at high substrate, the growth rate should saturate because finding the substrate is no longer the limiting step, whereas processing the substrate or other cellular constraints are. In the paper in which the data were treated statistically (Koch 1982a), nonsaturable models were discarded with the justification just given even though they fitted the sample of data better than other models. Then in a comparison of available "saturable" models, the Blackman was picked as second best and the Best model as the most satisfactory choice. It turned out that the *J* parameter was of a size that made the fit nearer to that of the Blackman formulation and farther away from the Monod formulation.

10. The Statistical Fitting

Our actual data for two cases of glucose consumption by *E. coli* ML 308 are shown in Figures 4.8 and 4.9 (Koch and Wang 1982; Koch 1982ab). In the late 1970s the data were analyzed by a subroutine ZXSSQ of an early version of the numerical package IMSL. The fitting used a modified Levenberg-Marquardt algorithm to find the minimum of the sum of squares. It was allowed to terminate when all parameters agreed to four significant figures on two successive iterations. Although the fitting was very much more cumbersome than it would be with newer software these early fittings should be reliable. Table 4.1 lists the fit of the data to 11 models for the data for the original strain of ML308. Five of these models were nonsaturable and of these the quartic and bilinear fits gave R^2 values of 0.960, indicating a very good fit, better than any of the saturable models. But

Table 4.1 Parameters to a variety of models from a least-squares fit of the data of a single run in the double-beam spectrometer linked to a computer

	RSS	R^2
Nonsaturable		
Linear, Equation (9) of Koch (1982a)		
0.1145 (α_0) 0.0308 (α_1)	0.9303	0.860 30
Quadratic, Equation (10) of Koch (1982a)		
−0.0130 (α_0) 0.0738 (α_1) −0.0017 (α_2)	0.3404	0.948 20
Cubic, Equation (11) of Koch (1982a)		
−0.0592 (α) 0.1029 (α_1) −0.0048 (a_2) 0.000 077 (a_3)	0.2954	0.954 44
Quartic, Equation (12) of Koch (1982a)		
−0.0034 (α_0) 0.0511 (α_1) 0.0050 (α_2)		
−0.000 524 (α_3) 0.000 012 (α_4)	0.2542	0.960 24
Bilinear, Equation (15) of Koch (1982a)		
−0.0133 (α_0) 0.0642 (a_1) 0.5625 (a_0') 0.0075 (a_1')	0.2564	0.960 46
Saturable		
1. Michaelis-Menten, Equation (7)		
1.2250 (μ_{max}) 12.9675 (K)	0.3829	0.942 51
2. Diffusion: enzyme, Equation (24), the Best equation		
0.7794 (V) 0.3965 (K) 11.4789 (J)	0.2850	0.956 63
3. Enzyme, leak:enzyme, Equation (5) of Koch (1982a)		
12.739 (V) 14.2969 (K_{en}) 37.640 (V_{en}) 342.31 (Kk)	0.382 85	0.940 96
5. Two concomitant enzymes, Equation (8) of Koch (1982a)		
363.91 (V_1) 109.31 (K_1) −381.08 (V_2) 117.25 (K_2)	0.3316	0.948 85
6. Bilinear (Blackman), Equations (22) and (23)		
0.0609 (a_1) 0.7403 (a_0')	0.3058	0.954 07
8. Hill equation, Equation (17) of Koch (1982a)		
0.8392 (V) 26.1968 (K) 1.7850 (n)	0.285 24	0.956 60

Seventy-eight pairs of growth rate versus dry weight concentration were obtained from a single growth experiment with strain ML 308 depleting a very low level of glucose, designated as run VI in the original work (Koch 1982a) and the data are shown in Figure 4.8. The turbidity values converted to dry weight were subtracted from 30.8 to give the residual glucose concentration in terms of bacterial dry weight. This was multiplied by 0.9026 to convert to micromolar concentrations. This yielded the data analyzed. The curves in the figure are identified with numbers that correspond to the fitted equations shown above in bold face. Lines 4 and 7 in Figure 4.8 are for models that can be rejected on other grounds (see Koch 1982a). None of the nonsaturable entries, whose parameters are given in the top part of this table are presented in Figure 4.8. The error of the individual estimates of growth rate derived from 300 turbidity measurements over a 200-sec interval was ± 0.03 to 0.08/hour.

as noted, we rejected them because they did not predict zero growth rate at zero substrate nor did they predict a substrate independent growth rate at very high substrate concentration. The other six models did have these properties and the Best equation with three parameters being statistically more acceptable than the two-concomitant enzymes and the Best diffusion:enzyme model with four parameters each. The Best equation is superior to the Monod and Blackman. The value of J (11.4789, in this example) fitted from such data, because of its meaning of $V/(AP)$, can be useful in dissecting the value of the three factors that make it up. More extensive use of the "diffusion:enzyme" equation of Best, will be invaluable to the next generation of microbial physiologists and microbial ecologists.

11. Diffusion Limitation and Effect of Multiple Layers

To enter the cell, a small molecule, such as glucose, needs to face a number of diffusion problems. If dq/dt is the rate of transport for each of the diffusion steps in the steady state, then diffusion up to the cell given by

$$dq/dt = 4\pi RD(S_0 - C_1), \tag{28}$$

where D is the diffusion constant of the substance in the aqueous environment; S_0, as earlier, is the concentration in the bulk phase of the growth medium; and C_1 is the concentration at the outer surface of the cell. At best, C_1 would be zero for an infinitely avid cell. This formula assumes that the cell is spherical with a radius R (see Koch 1971, 1990, 1996b).

For the next step, diffusion through the porins follows a simple form of Fick's law:

$$dq/dt = (A\pi fD/x)(C_1 - C_2), \tag{29}$$

where C_2 is the concentration just inside the outer membrane, A is the collective area of pores in the entire membrane, and x is the thickness of the outer membrane. Here D/x is the permeability constant, P, such as used in the given Best formulation. The Renkins factor, f, corrects for the chance that the molecule hits the wall of the pore and bounces back and also the resistance to flow due to the fact that fluid must move in accommodation to the passage of the molecule through the porin.

For the third step, diffusion through the periplasmic space would follow an even simpler case of Fick's diffusion law:

$$dq/dt = (4\pi RD/y)(C_2 - C_3), \tag{30}$$

where y is the thickness of the aqueous path through the periplasm and we neglect the effects of objects in the path such as the murein layer and periplasmic proteins. Here C_3 is the concentration immediately outside of the cytoplasmic membrane.

For the fourth step (i.e., transport through the membrane), we need to choose an equation from the many given in an earlier section. We also need to consider that the concentration, C_3, may vary from the immediate vicinity of a permease or PTS transport site (see Koch 1990) to other points equally close to the cytoplasmic membrane but farther from the source of entry. After transport through the cytoplasmic membrane, the consumption process can start to take place. We can easily take this to be due to an irreversible enzymatic action. After this step, it is assumed that consumption will be complete even though much biochemistry remains to be done, and we assume that these steps have no regulatory effect on any of this "front end" series of processes.

In Koch and Wang (1982) these various terms were evaluated from data of our experiments and those in the literature. The intermediate concentrations, C_1, C_2, and C_3, can be eliminated by combining the equations, which results in the conclusion that the resistance to flow of a series of processes is the resistance to flow of a combined process whose resistance is the combination of those of the component parts, mathematically related like the resistance of a set of linked electrical resistors when they are connected in parallel. It was shown in Koch and Wang (1982) that diffusion through the periplasmic space would contribute the least restriction to flow and that the most significant term is the diffusion through the porins.

The latter is an interesting and surprising result. Thus even though glucose can readily pass through the porins, at low external glucose concentrations with glucose entry limiting growth, the porins in the outer membrane slow the uptake very significantly. The porin limitation was less after the strain was selected for growth under chemostat conditions. Analysis of these data showed, however, that if one could remove the outer membrane of such organisms without impairing them, then indeed diffusion up to the cytoplasmic membrane would be limiting. This implies that the transport through the cytoplasmic layer with these chemostat-adapted cells was so effective that essentially every glucose molecule hitting the cytoplasmic membrane anywhere would be taken up. It could be that more or better permease molecules could exit, but there would never be a biological situation that would select for them.

12. The Effect of the Variation of the Surface Area to Volume during the Cell Cycle

A very thin rod-shaped organism that grows by elongation and binary fission is about the only case where a microbe maintains a virtually constant surface

area to volume (*S/V*) ratio as it grows and divides. Neglecting the ends, the *S/V* is $2\pi rl/\pi r^2 l$ or $2/r$, where r is the radius and l is the length. At the other extreme is the hypothetical case of a spherical organism that grows bigger and suddenly (instantly) divides into two spherical cells with half the volume, the *S/V* ratio is $4\pi r^2(4/3)/\pi r^3$ or $(3/r)$ at birth. Just before division, the radius, r, has increased by $2^{1/3}$, or 1.2599-fold, and the surface to volume ratio is lower by the same factor. Variations within this 26% range may or may not be important biologically.

   Clearly much more important is the fact that r in these cases appear in the denominator of these expressions which shows that with small objects of the same general shape the *S/V* will be higher. This, of course, is the key point in microbiology: Microbes are small so that transport can be simple and still be adequate. This is their essential strategy; it also leads to the generality that small organisms can grow faster (von Bertalanffy 1968; Kooijman 1993).

   No doubt as evolution has tuned organisms to their chosen niches, being able to grow as fast as possible is the key issue for many growth strategies under many circumstances. Does the growth rate vary during the cell cycle because of systematic variations in the *S/V* for a typical gram-negative rod-shaped organism? From Figure 4.10 it can be seen that for a culture of *E. coli* growing in minimal
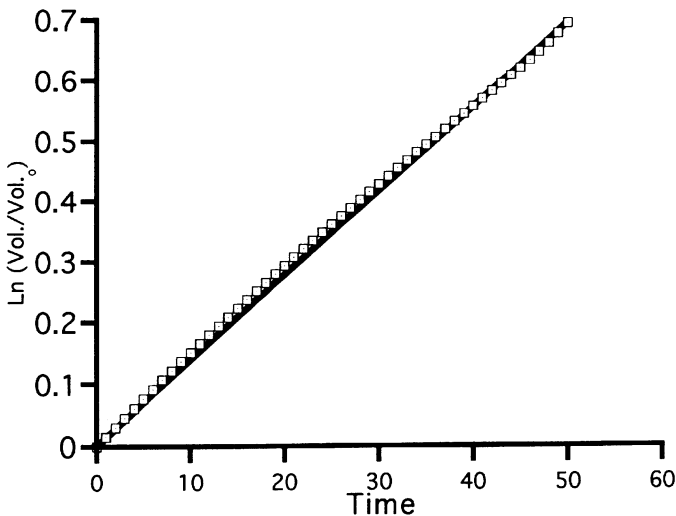
Figure 4.10 Surface-limited growth through the cell cycle. The open squares show the rate of growth based on the assumptions that the concentration of uptake sites on a unit area of the cell envelope is constant and that the cell is a cylinder of constant radius with hemispherical poles. The dimensions for a particular strain and growth conditions were taken from Woldringh et al. (1987). The solid line corresponds to exponential growth. It can be seen that the deviations caused by surface area are very small.

medium with an adequate level of nutrient, changes in the aspect ratio are almost without effect. They would be even less for very thin or long rods, and a little more evident for coccal type growth. I had hoped that this effect would be more profound and might lead to an understanding of the experiments of Kepes and Kepes (1981) in which cultures propagated in a special machine in a way so that at the end of each nominal doubling time, half the culture was removed and replaced with fresh medium containing a limiting amount of phosphate. If this S/V effect were of sufficient magnitude, it could have led to a synchronizing effect because cells in an unfavorable stage of growth would not have been as efficient as some others in scavenging the phosphate. Unfortunately, the calculations showed that the changes in the length to width or the S/V ratio were inadequate to account for the synchrony. (A good explanation of how this cycling technique functions to synchronize a growing culture has not been forthcoming, and because the technique is not sufficiently reproducible, its use has been discontinued.)

13. Grave Omissions

This chapter on the Monod model is very incomplete. First, there are the myriad of uses of the Monod equation in microbial ecology that have not been mentioned at all. Such a listing would include several of the chapters in this book. Moreover, the formulation of fermentation processes commonly uses Monod's concept. There is a vast body of work carried out by the school started by Frederickson in the Chemical Engineering Department at Minnesota. His colleagues there and his students now at a number of universities and even their students continue on in this tradition. The English school started at Porton by Herbert led to many studies of continuous culture. The Czechoslovakian school was important up until 1968 and the Dutch school still remains vital.

14. Conclusions

A detailed look at the mechanism of transport and consumption of even a single limiting substance whose metabolism is reasonably well studied showed that there are many facets that modify the growth process. A simple hyperbola describes uptake quite generally as a first approximation. Many times this will be approximately true, not because of the kind of explanations that Monod made, but because of a more general explanation; namely, that in a plot of specific growth rate versus substrate concentration the graph must go through the origin, be initially linear, and become horizontal at large substrate concentrations. The only question is just how the two branches join. Probably either the Monod

model, or better, the Blackman model will continue to be of great use in microbial ecology only because the exact shape in the intermediary region is usually unimportant. On the other hand, further analysis of the exact dependency backed up with measurements of other physiological variables and processes will be important in obtaining a deeper understanding of bacterial physiology. It can be expected that with more precise data, the Best model will come to more common usage as will detailed modeling of transport and its linkage to cell metabolism (Koch 1997).

References

Best, J. B. 1955. The inference of intracellular properties from observed kinetic data. *J. Cell Comp. Physiol.* 46:1–27.

Blackman, F. F. 1905. Optima and limiting factors. *Ann. Botany.* 19:281–295.

Briggs, G. E. and J. B. S. Haldane. 1925. A note on the kinetics of enzyme action. *Biochem. J.* 19:338–339.

Button, D. K. 1985. Kinetics of nutrient-limited transport and microbial growth. *Microbiol. Rev.* 49:270–297.

Cohen, G. N., and J. Monod. 1957. Bacterial permeases. *Bacteriol. Rev.* 21:164–194.

Dabes, J. N., R. K. Finn, and C. R. Wilke. 1973. Equations of substrate-limited growth: The case for Blackman kinetics. *Biotechnol. Bioeng.* 15:1159–1177.

Jacob, F., and J. Monod. 1961. Genetic regulatory mechanisms in the synthesis of protein. *J. Mol. Biol.* 3:318–356.

Judson, H. F. 1996. *The Eighth Day of Creation: The makers of the revolution in Biology.* expanded edition. Simon and Schuster, New York.

Kepes, F., and A. Kepes. 1981. Long-lasting synchrony of the division of enteric bacteria. *Bioche. Biophys. Res. Commun.* 99:761–767.

Koch, A. L. 1967. Kinetics of permease catalyzed transport. *J. Theor. Biol.* 14:103–130.

Koch, A. L. 1971. The adaptive responses of *Escherichia coli* to a feast and famine existence. *Adv. Microb. Physiol.* 6:147–217.

Koch, A. L. 1972. Deviations from hyperbolic dependency of transport processes. *J. Theor. Biol.* 36:23–40.

Koch, A. L. 1982a. Multistep kinetics: Choice of models for growth of bacteria. *J. Theor. Biol.* 98:401–417.

Koch, A. L. 1982b. Diffusion limit and bacterial growth. In V. Krumphanzl, B. Sikyta, and Z. Vanek (eds.), *Overproduction of Microbial Products,* pp. 571–580. Academic Press, London.

Koch, A. L. 1985. The macroeconomics of bacterial growth. In M. M. Fletcher and G. D. Floodgate (eds.), *Bacteria in Their Natural Environment,* pp. 1–42. The Society for General Microbiology, London.

Koch, A. L. 1990. Diffusion: The crucial process in many stages of the biology of bacteria. *Adv. Microb. Ecol.* 11:37–70.

Koch, A. L. 1993a. Biomass growth rate during the cell cycle. *CRC Crit. Rev.* 19:17–42.

Koch, A. L. 1993b. The growth law of *Bacillus subtilis. Antonie van Leeuwenhoek* 63:45–53.

Koch, A. L. 1996. What size should a bacterium be? A question of scale. *Annu. Rev. Microbiol.*, 50:317–348.

Koch, A. L. 1997. The microbial physiology and ecology of slow growth. *Microbiol. and Mol. Biol. Rev.*, September.

Koch, A. L., and R. Coffman. 1970. Diffusion, permeation, or enzyme limitation: A probe for the kinetics of enzyme induction. *Biotech. Bioeng.* XII:651–677.

Koch, A. L., and C. H. Wang. 1982. How close to the theoretical diffusion limit do bacterial uptake systems function? *Arch. Microbiol.* 131:36–42.

Kooijman, S. A. L. M. *Dynamic Energy Budgets in Biological Systems.* Cambridge University Press, Cambridge.

Medow, N. D., D. K. Fox, and S. Roseman. 1990. The bacterial phosphoenol pyruvate:glycose phosphotransferase system. *Ann. Rev. Biochem.* 59:497–542.

Michaelis, L., and Menten, M. M. 1913. Die Kinetik der Invertinwirkung. *Biochem. Z.* 49:333–369.

Monod, J. 1942. *Recherches sur las croissance des culture bacteriennes.* Herman et Cie, Paris.

Monod, J. 1949. The growth of bacterial cultures. *Ann. Rev. Microbiol.* 3:371–394.

Monod, J., and A. Audureau. 1946. Mutation et adaptation enzymatique chez *Escherichia coli-mutabile. Ann. Inst. Pasteur* 72:868–867.

Monod, J., G. Cohen-Bazire, and M. Cohn. 1951. Sur la biosynthèse de la β-galactosidase (lactase) chez *Escherichia coli:* la spécificité de l'induction. *Biochim. Biophys. Acta* 7:585–599.

Pearl, R., and L. J. Reed. 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. *Proc. Natl. Acad. Sci. USA* 6:275–288.

Postma, P. W., G. J. G. Ruijter, J. van der Vlog, and K. van Dam. 1992. Control of carbohydrate metabolism in enteric bacteria: qualitative and quantitative aspects. In E. Quagliariello and F. Palmieri (eds.), *Molecular Mechanisms of Transport,* pp. 97–105. Elsevier, New York.

Powell, E. O. 1967. Growth rate of microorganisms as a function of substrate concentration. In E. O. Powell, C. G. T. Evans, R. E. Strange, and D. W. Tempest (eds.), *Microbial Physiology and Continuous Culture,* pp. 34–55. Her Majesty's Stationery Office, London.

Rickenberg, H. V., G. N. Cohen, G. Buttin, and J. Monod. 1956. La galactoside permease d'*Escherichia coli. Ann. Inst. Pasteur* 91:829–857.

Ricklefs, R. E. 1990. *Ecology,* 3rd ed. W. H. Freeman, New York.

Robinson, J. A., and J. M. Tiedje. 1983. Nonlinear estimation of Monod growth kinetic parameters from a single substrate depletion curve. *Appl. Environ. Microbiol.* 45:1453–1458.

Simkins, S., and M. Alexander 1985. Non-linear estimation of the parameters of Monod kinetics that best describe mineralization of several substrate concentrations by dissimilar bacterial densities. *Appl. Environ. Microbiol.* 50:816–824.

von Bertalanffy, L. 1968. *General Systems Theory.* George Braziller, New York.

Verhulst, P. F. 1838. Notice sur la loi que la population suit dans son accroissement. *Corr. Mat. et Phys.* 10:113–121. Reprinted in *Readings in Ecology* (ed. E. J. Kormondy), pp. 64–66. Prentice Hall, Englewood Cliffs, N.J., 1956.

Wang, C. H., and A. L. Koch. 1978. Constancy of growth on simple and complex media. *J. Bacteriol.* 136:969–975.

Woldringh, C. L., P. Huls, E. Pas, G. H. Brakenhoff, and N. Nanninga. 1987. Topography of peptidoglycan synthesis during elongation and polar cap formation in a cell division mutant of *Escherichia coli* MC43100. *J. Gen. Microbiol.* 133:575–586.

Zimmermann, W., and A. Rosselet. 1977. Function of the outer membrane of *Escherichia coli* as a permeability barrier to beta-lactam antibiotics. *Antimicrob. Agents Chemotherapy* 12:368–372.

5

Using Transport Model Interpretations of Tracer Tests to Study Microbial Processes in Groundwater

Richard L. Smith and Stephen P. Garabedian

1. Introduction

It has long been known that microorganisms affect the geochemistry of groundwater. But despite this recognition, little detailed information is available regarding the rates and the factors controlling microbial processes in groundwater. Part of the reason stems from the relatively inaccessible nature of most groundwater and the difficulties encountered in obtaining representative samples of groundwater and subsurface sediments. At the same time, most groundwater systems are nutrient poor or oligotrophic environments in which the resident microorganisms are severely stressed and often nearly inactive. These populations are functioning so slowly that many types of activity measurements designed to assess microbial processes in more productive environments are ineffective for groundwater. However, because groundwater is by far the largest reservoir of freshwater in the world (Freeze and Cherry 1979), our lack of knowledge about groundwater microorganisms and their processes represents a significant void in the study of microbial ecology and in our ability to predict the outcome when these reserves are compromised by contamination.

Compared to groundwater microbiology, which is a fledgling discipline, groundwater hydrology is a mature science (e.g. see Freeze and Cherry 1979). A great deal is known about the flow of water through an aquifer and the importance of many of the factors affecting it. A variety of approaches, techniques, and models have been developed to facilitate hydrologic studies of groundwater and to test theoretical principles. Each of these aspects of groundwater hydrology has a direct bearing on groundwater microbiology because groundwater hydrology defines the environment within which groundwater microorganisms live. Microbial processes in any

habitat must be delineated within the context of the microbes' environment. Therefore, a great deal of the technology developed by and for groundwater hydrology can be aptly applied to studies of groundwater microbiology.

This chapter presents an overview of the adaptation of tracer-test technology, which was originally developed to study groundwater transport and movement, to also study microbial processes *in situ*. This approach is achievable because groundwater and tracer constituents will move through an aquifer. Yet at the same time, it is considerably complicated by this same condition. In addition to microbial processes, a variety of physical and chemical processes will also affect the transport of a tracer in groundwater. Fortunately, many of these factors are sufficiently characterized to the extent that conservative transport can be successfully simulated by groundwater transport models. These models can also be modified to include reaction terms for microbial processes. Because they deal with groundwater flow, the models describe certain aspects of a groundwater microorganism's environment, even though that was not the original intention when the models were developed. The modified models have actually become essential tools in the interpretation of tracer-test results for determining microbial processes. We review the development and the features of these models within the context of the groundwater environment, and then present two examples demonstrating how the models can be used to assess specific microbial processes in the subsurface.

2. The Groundwater Environment

2.1. Groundwater Occurrence

Although groundwater systems are basically aquatic environments from a microbial standpoint, there are several features of these systems that make them different from surface-water environments. These features include the total absence of light, generally low but constant temperatures, and the predominance of mineral surfaces. Subsurface earth materials generally have pores and fractures that can contain air, water, or other fluids (Freeze and Cherry 1979). The subsurface environment can be separated into two basic zones, the vadose (unsaturated) zone in which water is under less than atmospheric pressure, and the phreatic (groundwater) zone in which the water has pressure greater than atmospheric (Fig. 5.1). These zones are very distinct because water will only enter a well if it is at a pressure greater than atmospheric pressure. For most of the vadose zone, the earth materials are unsaturated—that is, the soil or parent materials have pores that contain both air and water—whereas in the phreatic zone pores contain only water. The importance of this distinction will become more apparent as the discussion leads to those situations in which oxygen in a groundwater system
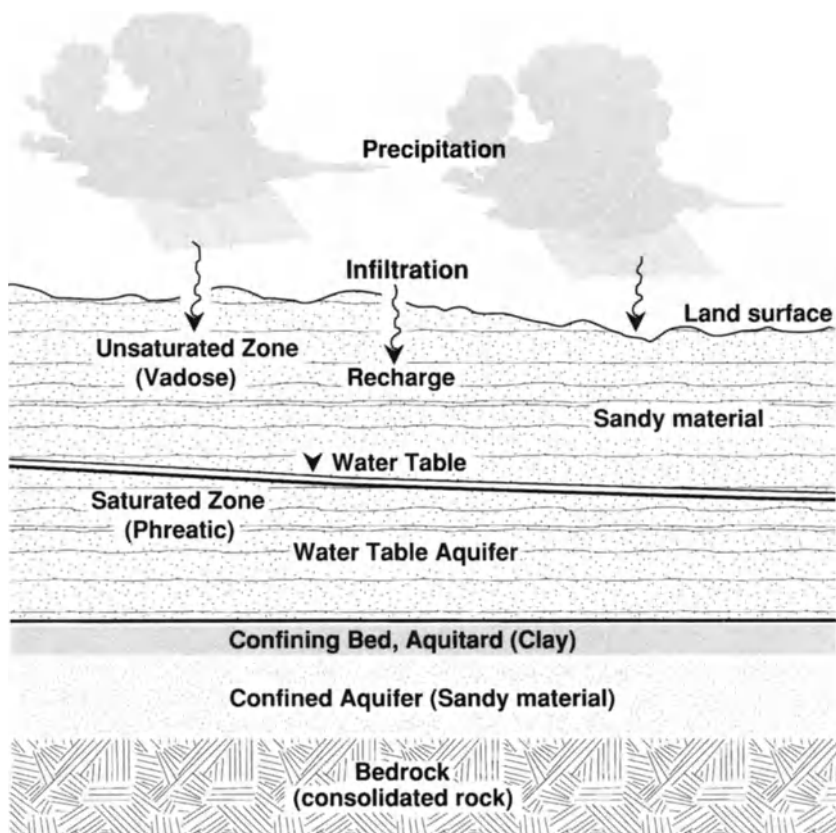
Figure 5.1 Generalized diagram of the subsurface.

becomes depleted and the redox condition can become reduced. Another major difference between the two zones is that the movement of water in the vadose zone is primarily downward; whereas the movement of water in the phreatic zone is primarily lateral. For example, some major aquifer systems have travel paths of tens to hundreds of kilometers.

Within the phreatic zone, water can be contained within aquifers or aquitards (Fig. 5.1). Aquifers are defined in an operational manner by the amount of water that can be extracted; when this amount is a useful quantity the groundwater system is considered an aquifer. Aquitards, however, may contain significant amounts of water, but the yield from them is very small. Aquifers can be classified as either unconfined or confined. In the case of an unconfined aquifer, the top of the aquifer is in direct contact with an unsaturated zone. In a confined aquifer,

the top of the aquifer is in direct contact with a unit that transmits water less readily. This confining bed (i.e., an aquitard) is typically composed of fine-grained silts and clays.

Aquifers can be found in several different geologic environments (Freeze and Cherry 1979). These include loose, unconsolidated sediments composed of gravel, sand, and some silt. Consolidated sedimentary rocks, such as sandstones and limestones, will also transmit water. Both consolidated and unconsolidated sedimentary rocks will typically transmit water through interstitial pores, that is, through pores that are between the individual mineral grains. The size of these pores will vary depending on the size of the mineral grains, their shape, and degree of sorting (or size variance); ranges of pore sizes are typically from micrometers to millimeters. In addition, sedimentary rocks may often have secondary pores, which are developed by the dissolution of minerals.

Other transmissive geologic materials include volcanic rocks (e.g., basaltic lava), intrusive igneous rocks (e.g., granite), and metamorphic rocks (e.g., schist and gneiss). Igneous and metamorphic rocks usually transmit water through fractures; these are developed in joints and faults, and typically are of micrometer to millimeter aperture size. Sandstones and limestones will also transmit significant quantities of water through fractures, and in the case of limestones these fractures can become larger due to dissolution of the limestone. It is interesting to note that the lay term for "underground river," although not applicable to aquifers composed of small pores and fractures, does accurately describe the situation in some limestone aquifers, where large cave systems can develop and be the primary means of transmitting water in these areas.

2.2. Groundwater Movement

Groundwater moves through granular or fractured porous media from areas of high hydraulic head to areas of lower head. Hydraulic head is a potential energy term which includes the potential energy of both elevation and pressure. This combined term is the elevation to which the water rises in a well, and is used because it is easily measured in the field.

The volumetric flow rate of water (Q) moving through a porous media has been found (Darcy 1856) to be linearly related to the change in hydraulic head over distance (dh/dx), which is the hydraulic gradient; the cross-sectional area of the porous media (A); and a proportioning factor defined as the hydraulic conductivity (K). Darcy's equation,

$$Q = -K\left(\frac{dh}{dx}\right)A, \tag{1}$$

is the equation used to describe fluid movement in the subsurface. The negative sign in this equation is necessary to indicate that fluid flows from higher to lower hydraulic head. Hydraulic conductivity (K) is a factor that generally is found from laboratory tests of porous media or from field tests using measurements of head around a pumped well.

Although hydraulic conductivity is defined in the same units as velocity (length/time), it does not represent the velocity of water in the porous media. To quantify the average velocity of water (v) in the porous media, a modified version of Darcy's equation (Freeze and Cherry 1979) is used which recognizes that fluid flow occurs only in the pores,

$$v = -\frac{K}{n}\left(\frac{dh}{dx}\right),\qquad(2)$$

where n is the porosity of the material. The porosity is defined as a fraction, and is the volume of voids in a porous media relative to the total volume. A representative example of velocity in a sandy material could involve a hydraulic conductivity value (K) of 100 m/day, a hydraulic gradient (dh/dx) of 1 m per 1000 m (0.001), and a porosity (n) of one third (0.33). This would result in a velocity (v) of 0.3 m/day. Although porosity will typically vary from 20% to 50% for most porous media, hydraulic conductivity can vary over several orders of magnitude, and velocities can have an equally large range. Most groundwater systems have velocities in the range of millimeters to meters per day.

2.3. Groundwater Transport

Although water moves through the porous media it will carry along with it solutes, gases, and in some cases colloidal-sized particles. Solutes and gases that do not interact with the solids (which can include well-crystallized minerals, amorphous solids, and organic material) and do not react in solution will generally travel (advect) with the same velocity as the water. Advective transport is the most important mechanism for solute movement in groundwater systems and will expose solutes to different regions of the aquifer over time.

As solutes move through porous media they also are dispersed. This dispersion occurs as the solutes travel through a number of routes in the porous media, some of which are faster and some slower than the average rate. The result of this range of velocities is the spreading of solute molecules, and if there is a chemical gradient, a dilution of the solute concentration.

During transport, solutes can be added or removed from groundwater, depending on the type of reactions that occur. For instance, if the solute undergoes a reaction with the solid surfaces, such as adsorption, it will be retarded relative

to the rate of transport of a conservative solute. Solutes that generally are conservative include anions like chloride and bromide; whereas cations like sodium and calcium will be adsorbed to mineral surfaces. In addition, precipitation-dissolution reactions are possible that deposit solid precipitates within the porous media or to dissolve solids, thereby adding solutes to the groundwater. There are also many reactions in groundwater systems that are microbially mediated that affect solute concentrations.

2.4. Groundwater Microbiology

Our present understanding of the diversity and ecology of groundwater microorganisms has been the subject of several reviews (Ghiorse and Wilson 1988; Harvey and Widdowson 1992; Chapelle 1993) and is not reiterated here. However, there are a few key points that need to be emphasized that are particularly relevant to the use of tracer tests for *in situ* activity measurements.

Because an aquifer is an aquatic environment that is dominated by solid surfaces, it should stand to reason that a considerable fraction of the microorganisms in an aquifer would be attached to surfaces (Ghiorse and Wilson 1988). For example, on average, greater than 98% of the bacteria in a pristine portion of a sand and gravel aquifer in Massachusetts were attached (Harvey et al. 1984). Generally, bacterial abundances in groundwater range from 10^5 to 10^8/g of sediment (Harvey and Widdowson 1992). However, the distribution is not uniform. In the Massachusetts aquifer, 98% of the adherent bacteria were attached to particles that were less than 60 μm in size, although these silt and clay-sized particles represented only about 2% (based on mass) of the total sediment particulates. Although a minor component, free-living bacteria are present in the groundwater and some are capable of traveling relatively long distances via groundwater flow (Harvey et al. 1989). The remainder are immobilized by straining or sorption after moving only short distances (Harvey 1991). An idealized scheme of the microbial community in a groundwater pore is shown in Figure 5.2. From the perspective of the microorganism, the rate of groundwater flow through this pore can actually be rather fast. The groundwater flow rate in the Massachusetts aquifer is 0.4 m/day (Garabedian et al. 1991), which is 4.6 mm/sec. The travel time across 100 μm, the approximate dimension of Figure 5.2, would be 22 sec.

In most groundwaters, nutrient availability is extremely sparse. Dissolved organic carbon concentrations are usually less than 1 mg/L in pristine groundwater, much of which may be refractory compounds (Thurman 1985). Dissolved nitrogen and phosphorus concentrations are often even lower or absent (Ghiorse and Wilson 1988). The result is that groundwater microorganisms are nutrient-limited and grow very slowly. Growth rates can be several orders of magnitude lower than those found in surface sediments (Thorn and Ventullo 1988). In situations
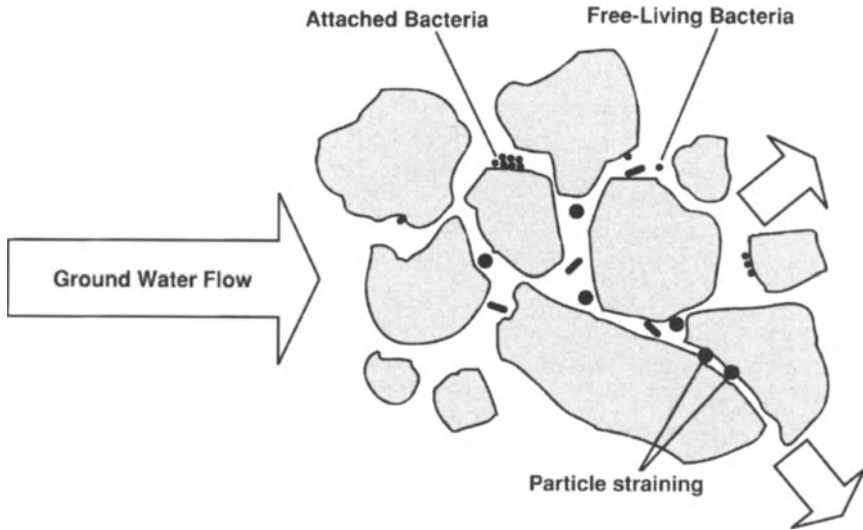
Figure 5.2 Schematic presentation of groundwater microorganisms in an aquifer pore. (Modified after Harvey (1991)).

where the groundwater is contaminated, growth rates can increase, but most are still low compared to surface-water habitats (Harvey and George 1987). Other general physiological indicators such as DNA content, total adenylates, and adenylate energy charge indicate that groundwater microorganisms are stressed and struggling to survive, even in contaminated aquifers (Metge et al. 1993).

Often an aquifer is viewed as a homogeneous environment, in which the physical, chemical, and microbiological conditions are considered to be uniform throughout. Usually this is an erroneous assumption. Geochemical gradients are common in groundwater, occurring both along and perpendicular to groundwater flowpaths (Chapelle 1993; Lyngkilde and Christensen 1993; Smith, Harvey, and LeBlanc 1991). The presence of these gradients can have important implications to the microbial processes occurring within an aquifer as they are often accompanied (or more appropriately caused) by sequential changes in the predominant terminal electron acceptor and consequently the redox potential. In many cases vertical gradients can be relatively steep (Fig. 5.3) and maintained for considerable horizontal transport distances (Smith, Harvey and LeBlanc 1991). This is due to a tendency of groundwater not to mix (i.e., disperse) well vertically when flow is horizontal. This condition is expressed as a small vertical dispersion coefficient. When examining microbial processes in groundwater, chemical gradients should be delineated at an appropriate spatial resolution and interpretation of microbiological results should be within the context of that spatial scale.
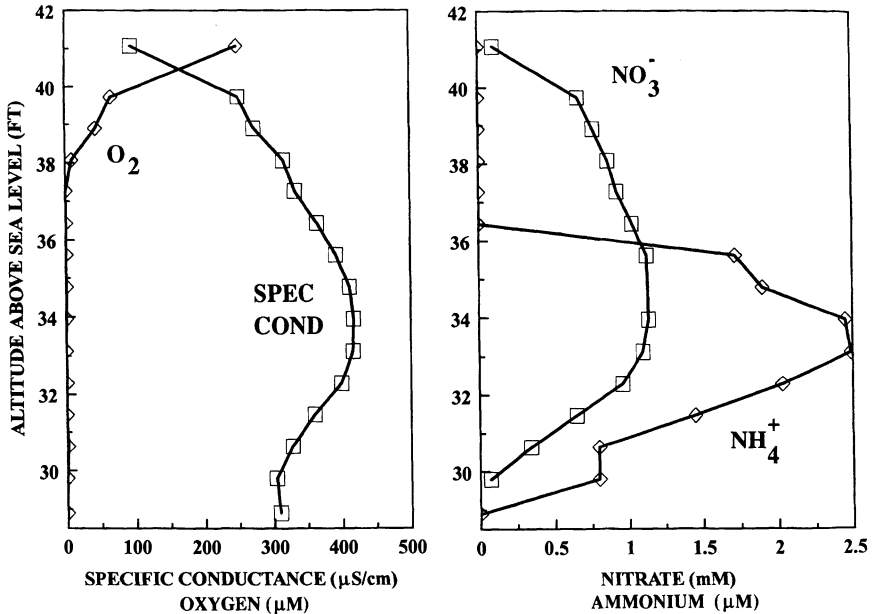
Figure 5.3 Vertical profile of dissolved constituents in a groundwater contaminant plume located on Cape Cod, MA, USA.

3. Measuring Microbial Processes in an Aquifer

The most common approach for measuring microbial processes in groundwater is to conduct batch incubations with aquifer material obtained using coring procedures (Aelion and Bradley 1991; Chapelle and Lovley 1990; Madsen et al. 1991; Phelps et al. 1989; Smith and Duff 1988; Wilson and Wilson 1985). Although the results of these kinds of incubations can give instructive insight into the metabolic potential of the groundwater microbial community, they should not be viewed as representing actual *in situ* rates of activity. The drilling procedure to obtain the cores is a disruptive process, fraught with the potential for contamination from the surface and geologic material above the horizon of the sample, and from drilling fluids. Once the core has been obtained, the incubation conditions usually do not reflect the hydrologic and geochemical conditions present within the aquifer. In addition to the uncertainty that these disruptions represent, but virtually totally ignored by groundwater microbiologists, is the heterogeneity that occurs within the matrix of an aquifer. This heterogeneity may require a large number of cores to obtain activity estimates that would be appropriate on an aquifer scale, or on any scale larger than 1 to 2 cm.

A more effective approach for assessing groundwater microbial processes on a habitat-relevant basis would be to conduct the activity assay directly within the aquifer. This approach recognizes the unique physical nature of an aquifer by avoiding attempts to reproduce it in the laboratory and it circumvents the problems associated with trying to collect representative samples. However, despite the conceptual appeal, there are imposing logistical barriers inherent in this approach that must be overcome. To date, only a handful of studies have successfully done so, several of which have been associated with enhanced activity due to bioremediative treatment. For example, at Moffett Field, California, the *in situ* biodegradation of chlorinated aliphatic compounds was characterized using induced populations of methanotrophic bacteria (Roberts et al. 1990; Semprini et al. 1990), whereas Hutchins et al. (1991) examined degradation of a JP-4 jet fuel spill by supplementing groundwater with nitrate and nutrients. Several studies that have characterized microbial processes *in situ* have been conducted at the Borden site in Ontario, Canada, including studies to determine the aerobic degradation rates of bromoform, dichlorobenzene, and hexachloroethane (Roberts et al. 1986) and BTEX (benzene, toluene, ethylbenzene, and xylene) degradation with both oxygen and nitrate as electron acceptors (Barbaro et al. 1992; Barker et al. 1987). Trudell et al. (1986) used a specially designated drive point that could first inject and then withdraw tracers in a nitrate-contaminated aquifer to measure rates of denitrification, and Gillham et al. (1990) reported a microcosm-like device that could be used *in situ* to measure microbial processes.

4. Tracer-Test Technology

Because tracer tests can and have been used for assessing groundwater microbial processes *in situ,* a brief overview on the technology is in order. Tracer tests have been used extensively in groundwater investigations, particularly for hydraulic research purposes (Davis et al. 1990; Domenico and Schwarz 1990). Tracers have been used to define travel paths and times through both granular and fractured-rock systems. Tracers used to quantify the direction and rate of water movement are usually conservative, that is they don't react with other solutes or with the solid surfaces. Typically anionic solutes are used for this purpose, including chloride and bromide; these solutes will usually move at the same rate as the water molecules. Another tracer used for the same purpose is tritiated water. Although tritium is nonconservative, its decay can be readily predicted, and is ideal in the sense that the water molecule itself is marked. A major problem with the use of tritium is the reluctance of regulatory agencies to allow release into the environment, and generally tritium's use is limited to laboratory experiments.

In addition to travel directions and rates, tracer tests are also used to characterize

and quantify the rates of reactions occurring in the subsurface. Examples of tracers used for these purposes include cations that electrostatically adsorb on to solid surfaces (e.g. cation exchange on clays). In this case the total mass of the cation remains conservative in that the cation does not degrade and there will be stable concentrations of the cation on the solid and in solution at equilibrium conditions. If, however, the equilibrium is disturbed, for example by a decreased concentration in solution, there will be an adjustment of the cation concentration on the solid until a new equilibrium is reached. An example of a different, consumptive reaction is the degradation of an organic compound (e.g. benzene) by microbial uptake. Here the compound may undergo both adsorption to solid organic matter in the aquifer and be consumed by microbially mediated reactions, decreasing its total mass. Other tracers have included colloidal-sized particles used to quantify the transport and reaction rates of groundwater particles; these tracers include both organic (bacteria, viruses, and synthetic microspheres) and inorganic (iron precipitates) particles.

There are two basic types of tracer tests used in groundwater systems, forced- and natural-gradient experiments (Domenico and Schwartz 1990). Forced-gradient tests involve the injection of water, the withdrawal of water, or both. The intent is to control the direction and rate of water movement in the aquifer. One type of forced-gradient test is the radial injection (Fig. 5.4A) experiment in which the flow of groundwater and tracer moves radially outward from the injection well. In this type of experiment only small volumes of water and tracer are extracted as they pass by the sampling well; therefore the radial flow pattern induced by the injection well is not disturbed by the sampling. A second design is the pumping experiment, in which the sampling well is now a large capacity pumping well, and a smaller capacity well is used to inject a small amount of the tracer (Fig. 5.4A). A third design is the doublet test in which the pumping and injection rates are usually the same (Fig. 5.4A). Advantages of these types of tests include the short duration of the experiment and the simplicity of the design, which requires only two wells. Disadvantages include the continuously varying velocity of the water as it moves away from (or toward) the injection (or pumping) well, and in the case of the radial injection or pumping test, the need to supply or dispose of large quantities of water some distance from the experiment.

A natural-gradient experiment involves injecting the tracers into the aquifer over a short period of time, therefore limiting the disturbance to the flow system, and allowing the solute to move at the rate of natural groundwater flow in the aquifer (Fig. 5.4B) (Garabedian et al. 1991; LeBlanc et al. 1991). This type of test has the advantage that it is conducted under the hydrologic and chemical conditions that are similar to those that would affect a contaminant or solute of interest in systems that are not undergoing pumping or other hydrologic stresses. Another advantage of this design is the time of exposure of the solute to the
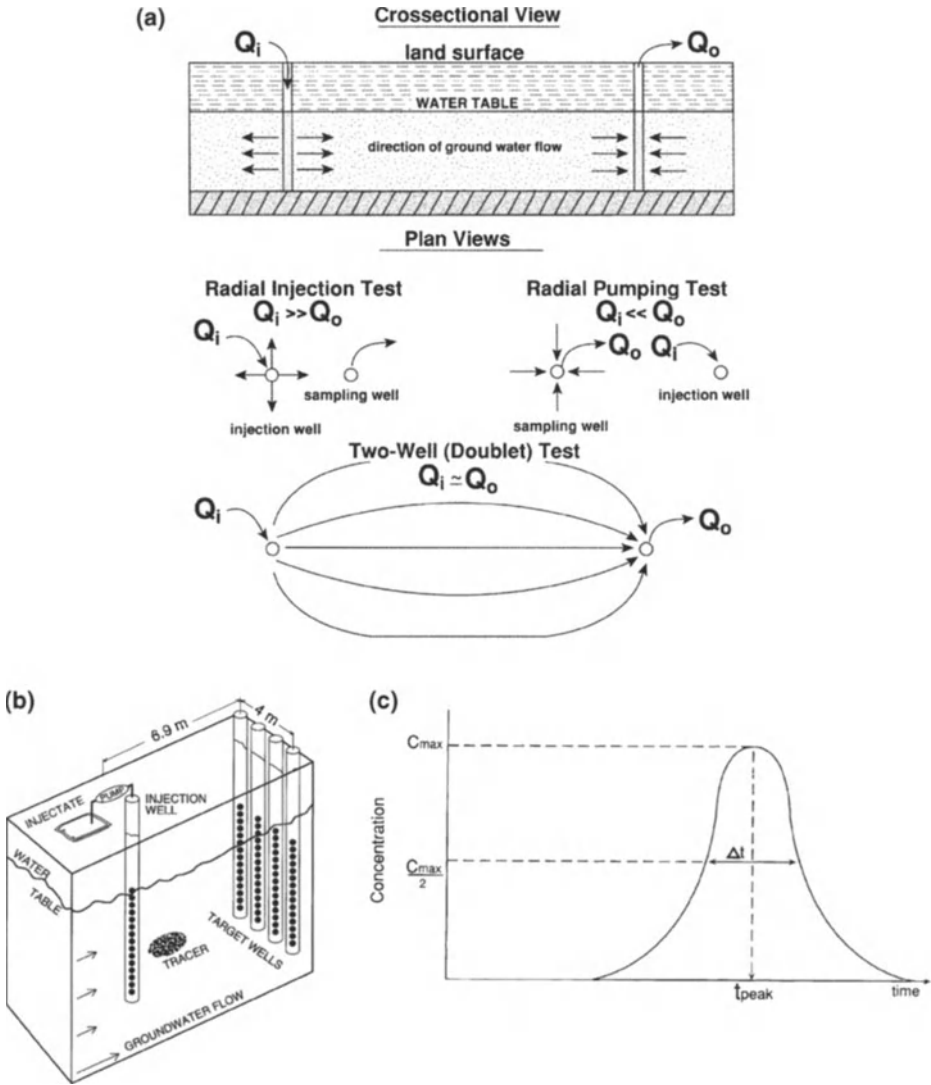
Figure 5.4 Types of A) forced-gradient and B) natural-gradient tracer tests used in ground-water studies and C) an example of a tracer test breakthrough curve. (Part B reprinted with permission from Smith, Howes and Garabedian (1991)).

sediment. Many of the reactions that occur in the subsurface are noninstantaneous; the time of exposure to the sediment is important and the rate and degree of degradation is dependent on the advection rate of the solute through the aquifer. Natural-gradient experiments enable reactions to occur at rates expected under natural (i.e., unstressed) conditions. A major disadvantage of natural-gradient experiments is that they can take weeks to months to complete because the flow of groundwater can be quite slow and it may take a long period of time to achieve the desired transport distances. Generally, the type of tracer test that is chosen will be determined by the objectives of the test, and the hypotheses that need to be tested.

5. Transport-Process Models

Mathematical models can be used to test various hypotheses and interpretations of the microbial processes occurring in the subsurface. The basic equation used for natural-gradient tracer tests is a modified version of the advection-dispersion equation. This one-dimensional equation is developed by assuming that both porosity and velocity of the groundwater system are constant with respect to both space and time (Domenico and Schwartz 1990; Freeze and Cherry 1979), and dispersion is assumed to be proportional to the concentration gradient and only occurs in the longitudinal direction. These assumptions result in the following equation:

$$\frac{\partial c}{\partial t} = D\,\frac{\partial^2 c}{\partial x^2} - v\,\frac{\partial c}{\partial x} + \frac{r}{n}\,, \tag{3}$$

where

c = concentration of solute in solution.
t = time,
D = dispersion coefficient = $\alpha_L v$,
α_L = longitudinal dispersivity,
x = spatial coordinate along the direction of flow,
v = velocity,
r = solute mass consumed or produced per unit volume per unit time,
n = porosity.

In the case of a conservative solute, one that does not undergo reaction in solution or with the solid surfaces, the reaction rate (r) would be zero. However, in situations where reactions do occur, various forms are possible (Domenico

and Schwartz 1990). One of the simplest would be the case of a zero-order reaction, in which r would be a constant production or consumption rate. For this case, the mass of the solute would increase or decrease at a constant rate over time. If the rate was negative, the initial mass was finite, and there were no other sources, the solute mass would eventually drop to zero. In the case of a first-order reaction, the reaction rate (r) is equal to the concentration of the solute (c) times a reaction constant (k): $r = kc$. Here the reaction rate will increase or decrease over time as the mass is produced or consumed by the reaction. An example of a first-order reaction is the decay of a radioactive isotope. Another form of the advection-dispersion equation can be developed, which includes the Michaelis-Menten reaction term; this term is often used to describe consumption of a biologically active solute:

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} - v \frac{\partial c}{\partial x} - \frac{V_{max}c}{n(c + K_m)}, \tag{4}$$

where V_{max} is the maximum rate of solute degradation and K_m is the Michaelis-Menten constant. This equation has the effect of a first-order reaction rate at low concentrations ($c \ll K_m$; $r = V_{max}c/K_m$) and a zero-order rate at high concentrations ($c \gg K_m$; $r = V_{max}$).

Deriving a solution to the preceding transport equation requires both boundary conditions and an initial condition for the solute concentration. In the case of an infinite porous media and a very thin slug of conservative solute ($r = 0$) injected into the porous media, the solution is

$$c(x,t) = \frac{M}{n\sqrt{4\pi Dt}} \exp\left[-\frac{(x - vt)^2}{4Dt}\right], \tag{5}$$

where M is the mass of injected solute.

The preceding solution can be used to identify the velocity and dispersivity of a conservative solute during an experiment. After the tracer injection and subsequent movement through the aquifer, the tracer concentration is sampled at a downgradient point, producing a breakthrough curve (Fig. 5.4C). The tracer concentration at this point will have been affected by dispersion and any reactions, reducing the tracer concentration relative to the initial concentration. The velocity is found by dividing the distance of the sampling point from the injection of the solute (x_1) by the time to the peak concentration ($v = x_1/t_{peak}$; see Fig. 5.4C). The dispersivity (α_L) is found by measuring the time (Δt) in which the concentration of the solute is greater than one half the peak concentration ($c(t) > c_{max}/2$; see Fig. 5.4C) (Smith, Howes and Garabedian 1991), and deriving a relation between Δt and α_L using the solution presented in equation 5:

$$\alpha_L = \frac{x_1(\Delta t/t_{peak})^2}{16\ \ln\ (2)}\ ,$$

(6)

where Δt is the duration of the breakthrough curve when $c(t) > 0.5c_{max}$ and c_{max} is the peak concentration and t_{peak} is the time to peak concentration.

To obtain a general solution for equation 4 it is necessary to use numerical methods because of the nonlinear nature of the Michaelis-Menten term (Smith, Howes, and Garabedian 1991). The partial differential equation is replaced with an algebraic approximation, the area of interest (i.e. the solution domain) is discretized into subregions (e.g. blocks) over which the approximations are valid, and time is divided into steps. An example of a numerical approach is the finite-difference approximation with a Crank-Nicholson time-stepping scheme (Huyakorn and Pinder 1983). After the approximations are introduced into equation 4, the equation is reduced to an algebraic form,

$$
\begin{aligned}
&c(x - \Delta x,\ t + \Delta t) \left[\frac{-D}{2(\Delta x)^2} + \frac{-v}{4(\Delta x)} \right] \\
&+ c(x,\ t + \Delta t) \left\{ \frac{1}{\Delta t} + \frac{D}{(\Delta x)^2} + \frac{V_{max}}{2n[K_m + c(x,\ t + \Delta t)]} \right\} \\
&+ c(x + \Delta x,\ t + \Delta t) \left[\frac{-D}{2(\Delta x)^2} + \frac{v}{4(\Delta x)} \right] = c(x - \Delta x,t) \left[\frac{D}{2(\Delta x)^2} + \frac{v}{4(\Delta x)} \right] \\
&+ c(x,t) \left[\frac{1}{\Delta t} + \frac{-D}{(\Delta x)^2} + \frac{-V_{max}}{2n(K_m + c(x,t))} \right] \\
&+ c(x + \Delta x,t) \left[\frac{D}{2(\Delta x)^2} + \frac{-v}{4(\Delta x)} \right]
\end{aligned}
$$

(7)

where Δx is the block step and Δt is the time step. One of these equations is applied to each of the blocks in the model domain, resulting in a large set of simultaneous equations to solve, typically by using a digital computer. Note that all of the concentration values on the left-hand side of equation 7, ($c(x - \Delta x, t + \Delta t)$, $c(x, t + \Delta t)$, $c(x + \Delta x, t + \Delta t)$), are at the new time step ($t + \Delta t$) and are unknown; those on the right-hand side, ($c(x - \Delta x,t)$, $c(x,t)$, $c(x + \Delta x,t)$) are at the present time level and are known.

The above algebraic equation is nonlinear because concentration appears in the denominator in the Michaelis-Menten term on the left-hand side. Therefore, an iterative scheme is employed to solve the equation and advance the solution to the next time step. The concentration value used in the denominator of the Michaelis-Menten term is held at a value from the previous iteration step. The fixed value is then updated to the new concentration approximation with every iterative step until the concentration on both sides of the equation converge and

the solution is stable. After convergence, the newly calculated concentration values are then used as the present time values and another set are calculated for the next time step, marching onward in time. (See Smith, Howes, and Garabedian (1991) for additional details.)

6. Assessing Methane Oxidation

Now we will illustrate two specific examples of using the coupled tracer test-model approach to assess microbial processes in an unconfined sand and gravel aquifer located on Cape Cod, Massachusetts. The aquifer has been continuously contaminated with treated sewage effluent since 1934, resulting in a contaminant plume that is more than 5 km long and nearly 1 km wide (LeBlanc 1984). Although it is rather dilute, the contamination has altered the composition and the distribution of the microbial community within the aquifer (Harvey and Barber 1992; Harvey and George 1987; Harvey et al. 1984; Metge et al. 1993). The plume is characterized by vertical geochemical gradients, which are maintained even after transport downgradient of several kilometers (Smith, Harvey, and LeBlanc, 1991). The predominant terminal electron-accepting processes of the microbial community change accordingly across these gradients as the redox potential changes. The result is a vertical zonation in microbial metabolism, similar to what is found in surface sedimentary environments, even though the overwhelming flux or movement of solutes in the aquifer is in the horizontal, downgradient direction.

More than 50 volatile organic compounds, many of them halogenated, have been detected in the Cape Cod contaminant plume (Barber 1988; Barber et al. 1988). Methane oxidation has been identified as a process that could be exploited to bioremediate groundwater contaminated with halogenated aliphatic compounds (Roberts et al. 1989; Semprini et al. 1990; Wilson and Wilson 1985). *In situ* treatment would involve enriching for methane-oxidizing bacteria in the subsurface by injecting methane into a known area of contamination (Semprini et al. 1990). The overall applicability of this approach will depend on the presence of native populations of methane-oxidizing bacteria in both contaminated and uncontaminated groundwater environments and their capacity to respond to the added methane. The results of a tracer test at the Cape Cod site designed to study the transport of dissolved gases suggested that when methane was added as a tracer it readily disappeared, even though it was not generally present in the aquifer. Therefore, because the *in situ* activity of this "methane-consuming" or methylotrophic process appeared to be relevant to the ecology of methane-oxidizing, groundwater bacteria and to their potential for bioremediation, the process was selected as a focal point for tracer test-modeling studies (Smith, Howes, and Garabedian 1991).
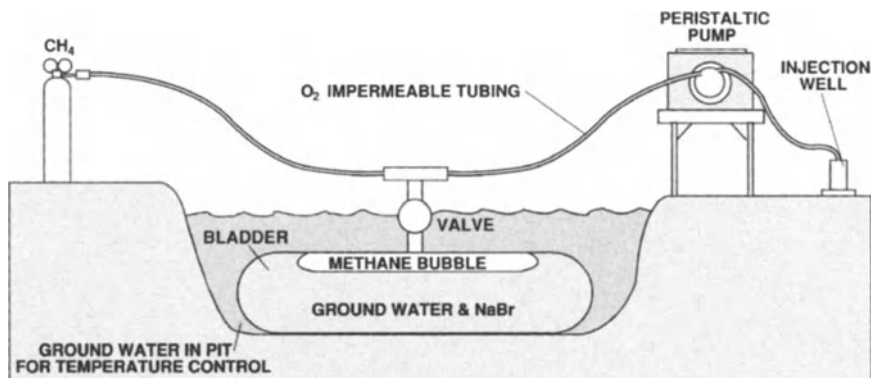
INJECTATE PREPARATION

● FLUSH BLADDER WITH N_2, THEN VENT
● ADD NaBr
● FILL WITH GROUND WATER
● ADD CH_4 , EQUILIBRATE WITH MIXING
● VENT GAS
● PUMP SOLUTION INTO GROUND

Figure 5.5 Diagram of the technique used to prepare the injection solution for the methane oxidation tracer test.

The basic approach for measuring methane oxidation *in situ* was to add the substrate, methane, to the aquifer, without altering the geochemistry of the zone of interest. This was accomplished by dissolving methane gas and NaBr (used as a conservative tracer) in 100 L of groundwater in a gas-impermeable bladder (Fig. 5.5) and then injecting the tracer solution into the ground through one or more ports of a multilevel sampling well. The use of the multilevel sampler as the injection well enables the selection of a specific vertical interval as the target zone for the activity measurement. The tracer cloud is then transported downgradient by natural groundwater flow, where it is intercepted by rows of multilevel samplers (Fig. 5.4B).

Tracer tests to measure methane oxidation were conducted in a well oxygenated, uncontaminated portion of the Cape Cod aquifer and in an anoxic nitrate-containing zone in which denitrification was the predominant electron-accepting process (Smith, Howes, and Duff 1991). Breakthrough curves of the injectate constituents clearly demonstrate that there was a significant loss of methane within the transport interval relative to a conservative species such as bromide (Fig. 5.6A). When normalized to the injectate concentration, methane concentrations in the tracer cloud were significantly attenuated at both the contaminated and uncontaminated
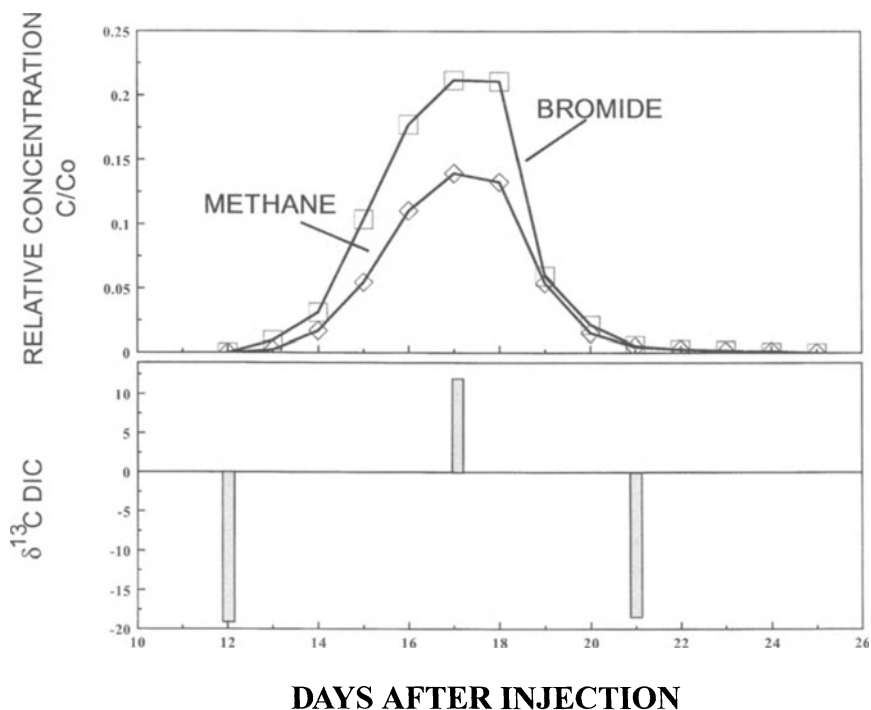
Figure 5.6 Breakthrough curves of A) bromide and methane and B) $\delta^{13}C$ values for dissolved inorganic carbon (DIC) for a natural-gradient tracer test conducted at the Cape Cod, MA site using $^{13}CH_4$ as the tracer. (Reprinted with permission from Smith, Howes, and Garabedian, (1991)).

sites, yet the time to peak concentration and the spread of the methane break-through curves were identical to the bromide curves (Fig. 5.6A). Methane oxidation was identified as the process responsible for the methane loss in replicate tracer tests using ^{13}C-enriched methane (Fig. 5.6B). The ^{13}C enrichment in the dissolved inorganic carbon pool could only have resulted from the oxidation of the added ^{13}C-enriched methane.

The amount of methane consumed within the transport interval can be calculated by taking the difference between the dimensionless concentrations (C/C_o) of bromide and methane for a given sample date and multiplying it by the injectate methane concentration. The amount of methane consumed is related to the methane concentration in the tracer cloud as it passed by the well fence (Fig. 5.7). Although the shape of this curve might suggest Michaelis-Menten saturation kinetics, it is actually an empiricial relationship and is site specific. That is because the curve is relating the total amount of substrate consumed within the entire transport interval with the substrate concentration at the end of the interval.
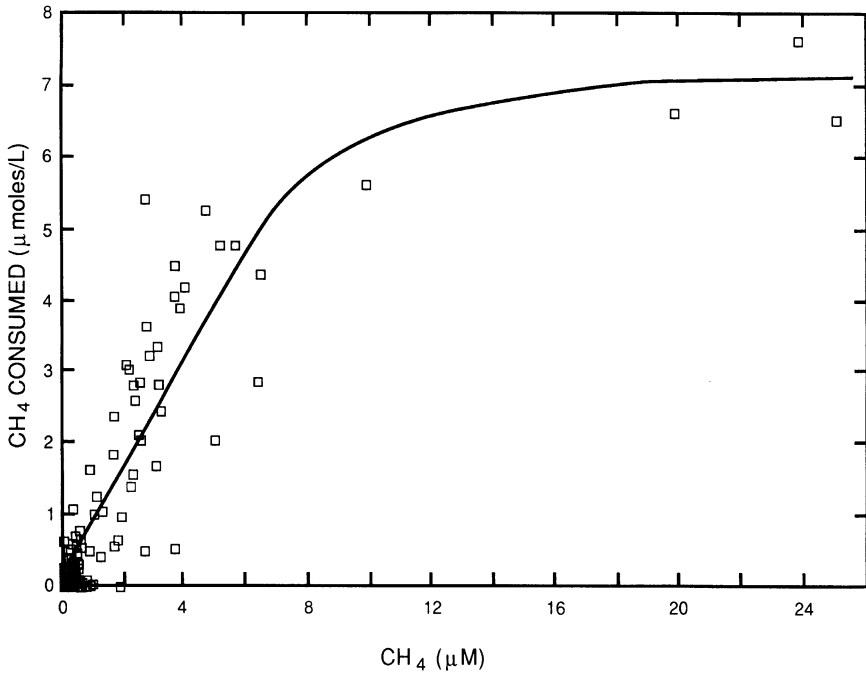
Figure 5.7 Relationship between the amount of methane consumed within a transport interval and the measured methane concentration in the tracer cloud at the end of the interval. The plot is a composite of data from several tracer tests conducted at one site. (Reprinted with permission from Smith, Howes and Garabedian (1991)).

In this type of test, substrate concentrations within the tracer cloud are being continually diluted during transport via dispersion. Therefore, if the process of interest is governed by Michaelis-Menten kinetics, the rate of the process is probably also changing as the substrate concentration changes. The only exception would be when the substrate concentration was sufficiently high to maintain a zero-order reaction throughout the transport interval. In that case the rate of the process would be constant during transport and a V_{max} can be calculated. For the methane oxidation tracer tests at the contaminated site, the amount of methane consumed was independent of methane concentrations $> 15 \, \mu M$ (Fig. 5.7). The average V_{max} for methane oxidation in samples in which the methane concentration exceeded this value* was $0.16 \, \mu moles$ (L aquifer $\times$ day)$^{-1}$.

The transport-microbial reaction model (Equation [4]) does not have a unique

*The value reported by Smith, Howes, and Garabedian (1991) was 0.4 mM/day (i.e., per liter of groundwater). Normalizing for aquifer porosity (39%) gives units that are better suited for comparison to other aquifers.

solution for both K_m and V_{max}. However, for any given V_{max}, there is a unique K_m. Therefore, if an independent estimate for one can be obtained, best-fit, trial and error simulations of the model can then be made to estimate the other. This is demonstrated with the methane tracer test data in Figure 5.8 using the V_{max} value of 0.16 μmoles (L aquifer × day)$^{-1}$, that was derived from Figure 5.7. The model adequately simulates the methane breakthrough curve. This indicates that in addition to capably predicting the physical factors that are affecting methane transport, the model is also able to successfully incorporate the effect that methane oxidation has on methane transport. The K_m value can be chosen either by visually selecting the simulation that best matches the peak of the breakthrough curve or by calculating the root mean square error for a statistical best fit (Smith, Howes, and Garabedian 1991). The statistical approach gives equal emphasis on the entire breakthrough curve, including the leading and trailing edges, while the visual approach places more emphasis on the central portion of the tracer cloud. The K_m for methane oxidation in the Cape Cod aquifer was determined to be 9.0 and 6.0 μM, at the contaminated and uncontaminated sites, respectively, using a visually determined best fit.

The results of the methane oxidation tracer tests emphasize that the best approach for such tests is to carefully select the total mass and the concentration of the substrate to be tested. The injection concentration should be high enough to result in zero-order kinetics for the center of the tracer cloud through the transport interval so that several samples can be collected from this region of the cloud. This enables a determination of V_{max} that is independent of the model. On the other hand, excessive additions of substrate to the aquifer will likely result in growth of the subsurface populations, which will interfere with determining the *in situ* rates of the process. Although growth kinetics could be incorporated into the transport model, it would be more straightforward to minimize growth by limiting the amount of substrate added and to minimize the contact time that the substrate has with any given volume of aquifer, and, therefore, with its attached resident microorganisms. For example, in the methane tests, the longest contact time, which was at the sampling well located farthest downgradient, was 16 days, but for only 4 days did the concentration exceed the calculated K_m for methane oxidation.

7. Assessing Denitrification

Inorganic nitrogen compounds are usually an important aspect of sewage-derived contamination, and the groundwater plume at the Cape Cod site is no exception (LeBlanc 1984). Nitrate is a common constituent of the contaminant plume, reaching concentrations of up to 2 mM. At the same time, the contamination has resulted in an increased electron acceptor demand, one which exceeds
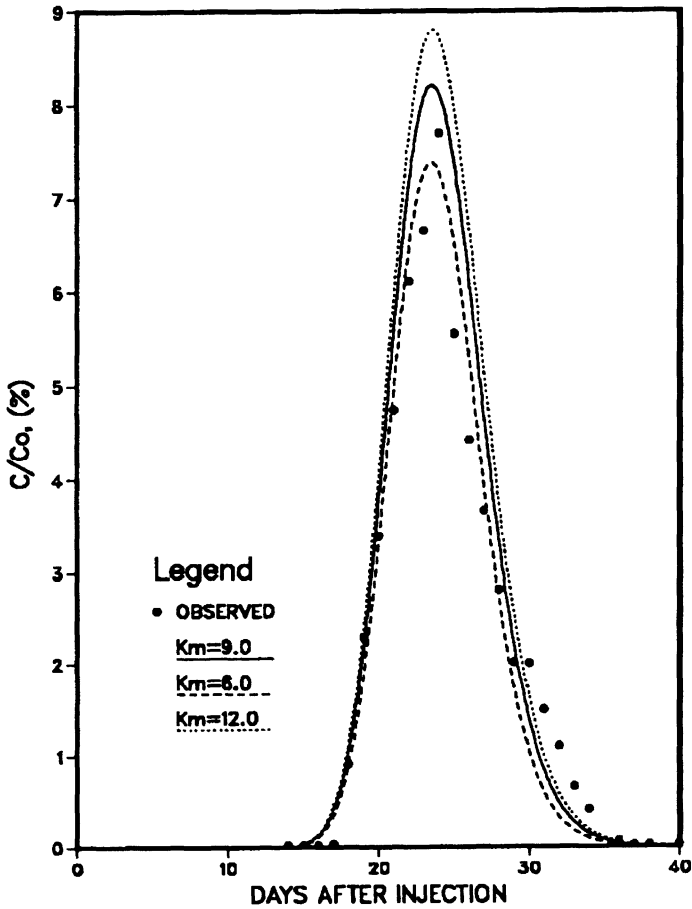
Figure 5.8 Model simulations as compared to the observed relative methane concentration for a tracer test measuring methane oxidation in a sand and gravel aquifer. The figure demonstrates the sensitivity of the model to different values for K_m (μM) for methane oxidation. Values used for the simulation were: v, 0.28 m day^{-1}; α_L, 0.06 m; V_{max}, 0.16 μmoles (L aquifer $\times$ day)$^{-1}$. (Reprinted with permission from Smith, Howes and Garabedian (1991)).

the dissolved oxygen supply in most locations. Consequently, denitrification is a predominant terminal electron-accepting process in much of the contaminant plume (Smith and Duff 1988; Smith, Howes, and Duff 1991). The factors that control denitrification and the extent and rate at which it is occurring are, therefore, important facets of the total microbial community metabolism within the contaminant plume.

Small-scale natural-gradient tracer tests were conducted within the contaminated aquifer to determine the *in situ* rates of denitrication at a site in which the process was known to be occurring (Smith et al. 1996). Denitrifying activity was measured with the acetylene blockage technique (Balderston et al. 1976; Yoshinari et al. 1977). This technique exploits the sensitivity of the terminal enzyme of denitrification, nitrous oxide reductase, to acetylene. In the presence of acetylene nitrous oxide is the end product of denitrification, rather than N_2. Tracer tests were conducted in anoxic ground water containing high concentrations of nitrate by adding acetylene and NaBr as tracers (Table 5.1).

The overall nature of this tracer test is quite different from the methane oxidation tracer tests. In this case, the substrate for the process of interest is present throughout the transport interval at high concentrations, the process itself is occurring continuously, while the tracer cloud contains an inhibitor for the process rather than the substrate. Activity is subsequently quantified by the accumulation of a 'product' within the tracer cloud above the background concentration. Both the inhibitor and the 'product' will be diluted by dispersion during downgradient transport, but the substrate will not be diluted.

The injectate solution for the acetylene block tracer test was prepared as depicted in Figure 5.5, except that acetylene was substituted for methane. The breakthrough curves for the conservative tracer (bromide) and acetylene in the sand and gravel aquifer are very similar (Figure 5.9). This is a clear indication that the acetylene was transported in a conservative manner, both hydrologically and biologically. There was also an increase in nitrous oxide concentration coincident with the tracer cloud (Fig. 5.9). This nitrous oxide peak confirms that denitrification was actively occurring within the transport interval. However, the

Table 5.1 Acetylene block tracer test

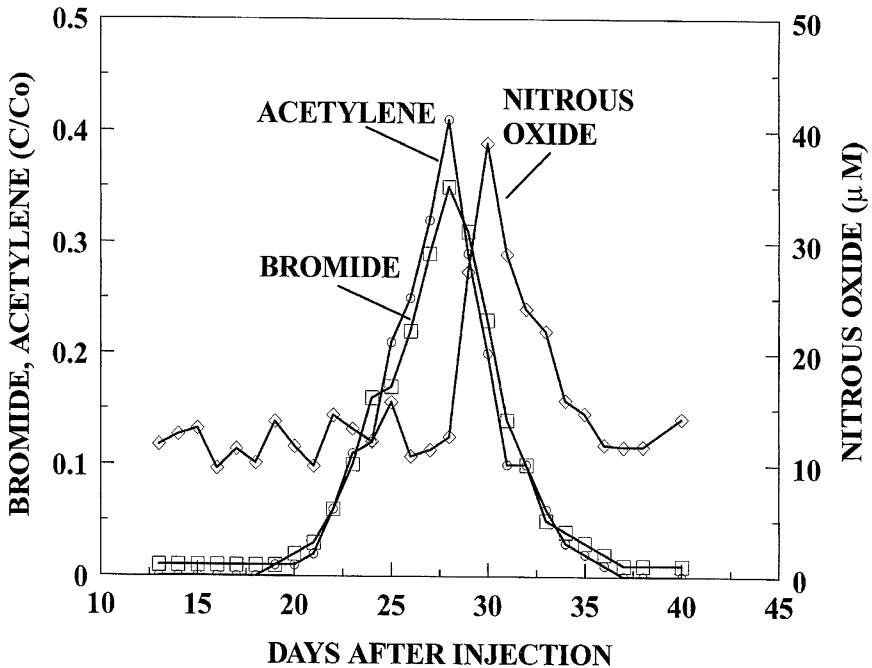| Groundwater geochemistry	
Specific conductance	416 µS/cm
Oxygen	<2 µM
Nitrate	1.8 mM
Nitrous oxide	6.6 µM
Tracer constituents	
Acetylene	6.1 mM
Sodium bromide	1.0 mM

Figure 5.9 Breakthrough curves of bromide, acetylene and nitrous oxide for an acetylene block tracer test that measured denitrification in nitrate-contaminated ground water. (Reprinted with permission from Smith, Garabedian and Brooks (1996)).

peak nitrous oxide concentration occurred 2 days after and was several days narrower than the acetylene and bromide peaks.

This apparent lag or delay of the nitrous oxide peak is readily explainable given the context of this tracer test. The nitrate concentration present in the zone in which the test was conducted is more than sufficient to allow denitrification to function as a zero order reaction (relative to nitrate) (Firestone 1982). The effect this has within the tracer cloud is depicted in Figure 5.10. As the tracer cloud passes by a fixed point, such as a sampling well, the acetylene concentration in the leading edge will increase to a level that will result in a complete inhibition of nitrous oxide reductase. Then nitrous oxide accumulation would occur at a constant, zero-order rate as the remainder of the tracer cloud passes by, even after acetylene concentrations peak and begin to decline. As a result, nitrous oxide concentrations will continue to increase throughout the time of exposure to the acetylene cloud, with the peak nitrous oxide value occurring well after the center of the acetylene cloud has passed. Nitrous oxide production should drop off only at the trailing edge of the tracer cloud when acetylene falls below the amount needed for a complete inhibition.
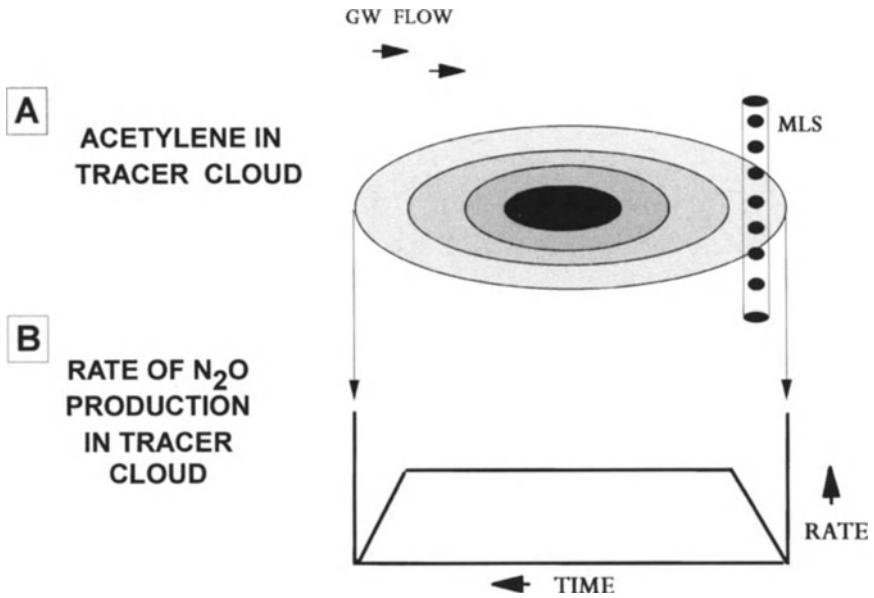
Figure 5.10 Schematic of the rate of nitrous oxide production within a tracer cloud of varying acetylene concentration when denitrification is zero order due to high nitrate concentrations.

The model used to simulate the acetylene block tracer test was somewhat different from the model used for the methane oxidation determination. A zero-order rate constant was used to calculate nitrous oxide production within the tracer cloud. However, at low acetylene concentrations the rate constant was linearly scaled to account for those situations in which the acetylene inhibition was not 100%. The equations used were

$$N_p = 0, \qquad\qquad C_a \le C_{min}, \qquad\qquad (8a)$$

$$N_p = K_{max} \left(\frac{C_a - C_{min}}{C_{max} - C_{min}} \right), \quad C_{min} \le C_a \le C_{max}, \qquad (8b)$$

$$N_p = K_{max} \qquad\qquad C_{max} \le C_a, \qquad\qquad (8c)$$

where

N_p = nitrous oxide production rate,
K_{max} = maximum nitrous oxide production rate,

C_a = acetylene concentration,
C_{min} = minimum acetylene concentration for inhibition to occur,
C_{max} = acetylene concentration at which K_{max} is achieved.

Initial values for C_{min} and C_{max} were selected as 0 and 10 μM, respectively, based on the data reported by Yoshinari et al. (1977).

This initial version of the model matched the nitrous oxide breakthrough curves, but with limited success (Fig. 5.11). It simulated the delayed arrival time in the nitrous oxide peak relative to the acetylene and bromide peaks, however, it did not aptly simulate the narrowness of the nitrous oxide peak. The simulated peak can be narrowed by altering C_{min} and C_{max}, although the model is not very sensitive to changes in these parameters. Unrealistically high values (>> 1000 μM for C_{min}) are needed to match the narrowness of the observed data, yet increasing these values resulted in progressively earlier arrival times for the simulated nitrous oxide peak, no longer matching the observed arrival time. So there appeared to be an additional factor(s) affecting denitrification *in situ* or the function of the acetylene block *in situ*.

Incubation experiments with aquifer core material routinely exhibit a 3- to 4-day lag period when denitrification was assayed using the acetylene blockage technique (Smith et al., 1996). This lag was usually attributed to disturbances from the core collection process and (or) to bottle effects during the incubation. However, when the acetylene block transport model was modified to include a 4-day lag term, the simulated nitrous oxide breakthrough curve closely matched the tracer test data for both arrival time and peak shape (Fig. 5.12). Adding a lag term functioned to decrease the amount of nitrous oxide production computed by the model, and also narrowed the simulated breakthrough curve. It also delayed the arrival time of the peak, and this had the effect of countering the effect of increasing C_{min}. With the time lag in place, the calibrated best-fit value for C_{min} was 500 μM, which is not an unreasonably high value. This modified model suggests that ground-water denitrification is much less sensitive to acetylene than are pure cultures or surface soils and that the time lag in response to the presence of acetylene is a general phenomenon for denitrification at this groundwater site.

The *in situ* denitrification rate that yielded the best-fit simulation for the modified version of the model was 1.5 μmoles N_2O (L aquifer × day)$^{-1}$. This rate is up to 26-fold less than equivalent rate measurements using laboratory incubations with core material obtained from the tracer test location and depth (Smith et al. 1996). Bottle or flask activity measurements are well known for overestimating *in situ* rates with surface sediments and soils. Tracer tests are also inherently more capable of integrating the natural heterogeneity that occurs along an aquifer flow path, an integration that would require the collection of up to several hundred cores for laboratory incubations.
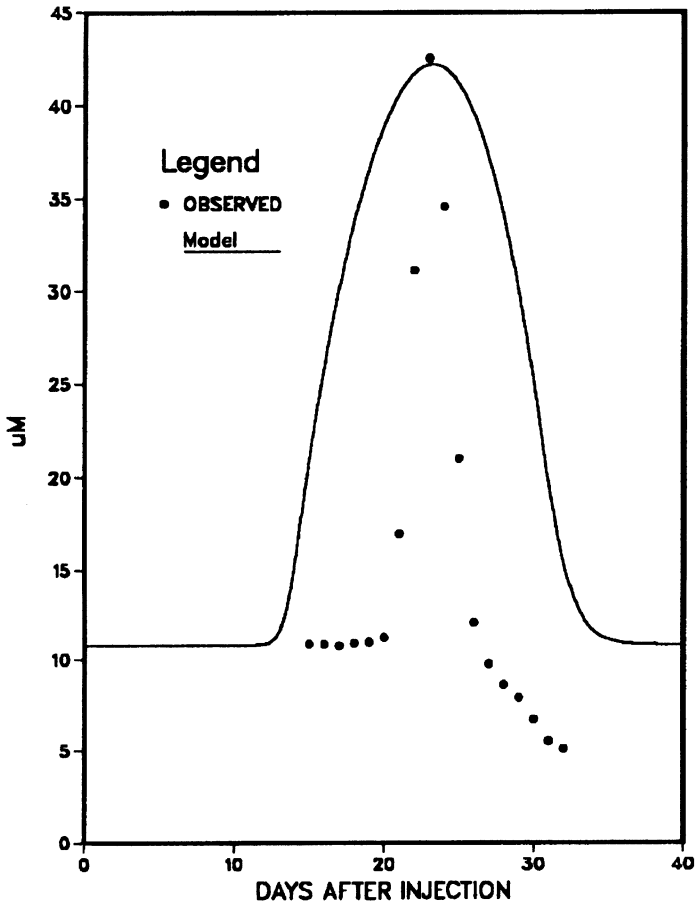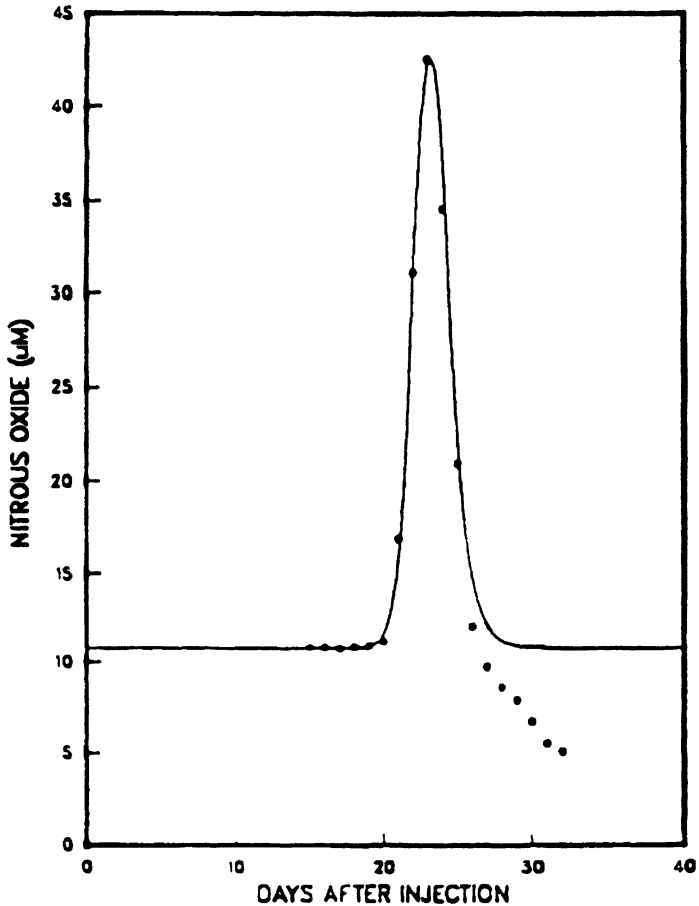
Figure 5.11 Model simulation as compared to the observed nitrous oxide concentration for an acetylene block tracer test to measure denitrification. Values used for the simulation were: v, 0.45 m day^{-1}; α_L, 0.05 m; c_{min}, 0 μM; c_{max}, 10 μM; rate = K_{max}, 0.6 μmoles N_2O produced (L aquifer × day)$^{-1}$. (Reprinted with permission from Smith, Garabedian and Brooks (1996)).

Figure 5.12 Model simulation as compared to the observed nitrous oxide concentrations for an acetylene block tracer test to measure denitrification. The model used in Figure 5.11 was modified to include a 4 day lag term for the commencement of nitrous oxide production after exposure to an acetylene concentration that exceeded C_{min}. Values used for the simulation were: v, 0.45 m day^{-1}; α_L, 0.05 m; c_{min}, 500 μM; c_{max}, 501 μM; rate = K_{max}, 1.51 μmoles N_2O produced (L aquifer × day)$^{-1}$. (Reprinted with permission from Smith, Garabedian and Brooks (1996)).

8. Future Applications and Limitations

Natural-gradient tracer tests offer a unique opportunity to assess microbial processes occurring in groundwater within the *in situ* environment. Valid interpretation of the tracer-test results requires the application of a model that accounts for advection and dispersion during transport and also includes microbial processes. The models, then, are an integral part of the overall approach and must be considered when designing these types of tests.

Many different types of microbial processes could be examined using natural-gradient tracer tests. An example of a bacterial transport experiment is reported by Harvey and Garabedian (1991), in which a model based on colloid-filtration theory was applied to a natural-gradient tracer test using stained indigenous bacteria. Particularly apropos would be the predominant terminal electron-accepting processes, which would be indicative of total microbial metabolism in the aquifer. Tracer tests could also be used to examine short-term responses to perturbations such as changes in redox, electron donor or electron acceptor supply, metabolism of individual substrates of interest, or the effect of various types contamination. In some situations, such as with the methane oxidation studies at the Cape Cod site, these tests could also be used to gauge the potential for bioremediation. These tests do not necessarily have to be slug injections, they could also potentially be continuous injections for extended periods of time, which might be better suited for bioremediative studies.

The tracer-test model approach does have certain drawbacks. First, it requires a good understanding about the hydrology of a given aquifer. Obviously, one must have reasonable confidence regarding the direction and rate of travel that a tracer cloud will take to intercept it downgradient. These tests are also rather time-consuming and can be initially expensive with regard to well construction and installation. Logistical considerations will probably limit these tests to relatively shallow depths (perhaps 100 ft below land surface). And as with most field experiments, suitable controls and replication will be subjected to the vagaries of a changeable environment. There will also be limitations on the types of compounds that can be added to groundwater; toxic or radioactive substrates may be inappropriate for drinking water aquifers. However, the recent improvement in analytical sensitivities suggests that using tracers labeled with stable isotopes could be a practical alternative.

In general, this tracer-test transport model approach is an interactive, interdisciplinary combination of microbial ecology, groundwater hydrology, and numerical simulation. There are many situations for which the approach will be a valuable tool for examining the activities and the ecology of groundwater microorganisms directly within their native environment.

References

Aelion, C. M., and P. M. Bradley. 1991. Aerobic biodegradation potential of subsurface microorganisms from a jet fuel-contaminated aquifer. *Appl. Environ. Microbiol.* 57: 57–63.

Balderston, W. L., B. Sherr, and W. J. Payne. 1976. Blockage by acetylene of nitrous oxide reduction in *Pseudomonas perfectomarinus. Appl. Environ. Microbiol.* 31:504–508.

Barbaro, J. R., J. F. Barker, L. A. Lemon, and C. I. Mayfield. 1992. Biotransformation of BTEX under anaerobic denitrifying conditions: Field and laboratory observations. *J. Contam. Hydrol.* 11:245–272.

Barber, L. B. 1988. Dichlorobenzene in ground water. Evidence for long-term persistence. *Ground Water* 26:696–702.

Barber, L. B., E. M. Thurman, M. P. Schroeder, and D. R. LeBlanc. 1988. Long-term fate of organic micropollutants in sewage-contaminated groundwater. *Environ. Sci. Technol.* 22:205–211.

Barker, J. F., G. C. Patrick, and D. Major. 1987. Natural attenuation of aromatic hydrocarbons in a shallow sand aquifer. *Ground Water Monit. Rev.* pp. 64–71. Vol. 7 #3.

Chapelle, F. H. 1993. *Ground-Water Microbiology and Geochemistry.* John Wiley & Sons, New York.

Chapelle, F. H., and D. R. Lovley. 1990. Rates of microbial metabolism in deep coastal plain aquifers. *Appl. Environ. Microbiol.* 56:1865–1874.

Darcy, H. 1856. *Les fontaines publiques de la ville de Dijon.* Victor Dalmont, Paris.

Davis, S. N., G. M. Thompson, H. W. Bentley, and G. Stiles. 1990. Ground water tracers: A short review. *Ground Water* 18:14–23.

Domenico, P. A., and F. W. Schwartz. 1990. *Physical and Chemical Hydrogeology.* John Wiley and Sons, New York.

Firestone, M. K. 1982. Biological denitrification. In F. J. Stevenson (ed.), *Nitrogen in Agricultural Soils,* pp. 289–326. American Society of Agronomy, Madison, WI.

Freeze, R. A., and J. A. Cherry. 1979. *Groundwater.* Prentice-Hall, Inc., Englewood Cliffs, NJ.

Garabedian, S. P., D. R. LeBlanc, L. W. Gelhar, and M. A. Celia. 1991. Large-scale natural-gradient tracer test in sand and gravel. Cape Cod, Massachusetts. 2. Analysis of spatial moments for a nonreactive tracer. *Water Resour. Res.* 27:911–924.

Ghiorse, W. C., and J. T. Wilson. 1988. Microbiol ecology of the terrestrial subsurface. *Adv. Appl. Microbiol.* 33:107–172.

Gillham, R. W., R. C. Starr, and D. J. Miller. 1990. A device for *in situ* determination of geochemical transport parameters. 2. Biochemical reactions. *Ground Water* 28:858–862.

Harvey, R. W. 1991. Parameters involved in modeling movement of bacteria in ground water. In C. J. Hurst (ed.), *Modeling the Environmental Fate of Microorganisms,* pp. 89–114, American Society for Microbiology, Washington, D.C.

Harvey, R. W., and L. B. Barber. 1992. Associations of free-living bacteria and dissolved organic compounds in a plume of contaminated ground water. *J. Contam. Hydrol.* 9:91–103.

Harvey, R. W., and S. P. Garabedian. 1991. Use of colloid filtration theory in modeling movement of bacteria through a contaminated aquifer. *Environ. Sci. Technol.* 25:178–185.

Harvey, R. W., and L. H. George. 1987. Growth determinations for unattached bacteria in a contaminated aquifer. *Appl. Environ. Microbiol.* 53:2992–2996.

Harvey, R. W., L. H. George, R. L. Smith, and D. R. LeBlanc. 1989. Transport of microspheres and indigenous bacteria through a sandy aquifer: Results of natural and forced-gradient tracer tests. *Environ. Sci. Technol.* 23:51–56.

Harvey, R. W., R. L. Smith, and L. George. 1984. Effect of organic contamination upon microbial distribution and heterotrophic uptake in a Cape Cod, Mass., aquifer. *Appl. Environ. Microbiol.* 48:1197–1202.

Harvey, R. W., and M. A. Widdowson. 1992. Microbial distributions, activities, and movement in the terrestrial subsurface: Experimental and theoretical studies. In R. J. Wagenet, P. Baveye, and B. A. Stewart (eds.), *Interacting Processes in Soil Science,* pp. 185–225. Lewis Publishers, Ann Arbor, MI.

Hutchins, S. R., W. C. Downs, J. T. Wilson, G. B. Smith, D. A. Kovacs, D. D. Fine, R. H. Douglas, and D. J. Hendrix. 1991. Effect of nitrate addition on biorestoration of fuel-contaminated aquifer: Field demonstration. *Ground Water* 29:571–581.

Huyakorn, P. S., and G. F. Pinder. 1983. *Computational Methods in Subsurface Flow.* Academic Press, New York.

LeBlanc, D. R. 1984. Sewage plume in a sand and gravel aquifer, Cape Cod, Massachusetts. *U.S. Geological Survey Water Supply Paper 2218.* U.S. Government Printing Office, Washington, D.C.

LeBlanc, D. R., S. P. Garabedian, K. M. Hess, L. W. Gelhar, R. D. Quadri, K. G. Stollenwerk, and W. W. Wood. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 1. Experimental design and observed tracer movement. *Water Resour. Res.* 27:895–910.

Lyngkilde, J., and T. H. Christensen. 1993. Redox zones of a landfill leachate pollution plume (Vejen, Denmark). *J. Contam. Hydrol.* 10:273–289.

Madsen, E. L., J. L. Sinclair, and W. C. Ghiorse. 1991. *In situ* biodegradation: Microbiological patterns in a contaminated aquifer. *Science* 252:830–833.

Metge, D. W., M. H. Brooks, R. L. Smith, and R. W. Harvey. 1993. Effect of treated-sewage contamination upon bacterial energy charge, adenine nucleotides, and DNA content in a sandy aquifer on Cape Cod. *Appl. Environ. Microbiol.* 59:2304–2310.

Phelps, T. J., C. B. Fliermans, T. R. Garland, S. M. Pfiffner, and D. C. White. 1989. Methods for recovery of deep terrestrial subsurface sediments for microbiological studies. *J. Microbiol. Meth.* 9:267–279.

Roberts, P. V., G. D. Hopkins, D. M. Mackay, and L. Semprini. 1990. A field evaluation

of *in-situ* biodegradation of chlorinated ethenes: Part 1, Methology and field site characterization. *Ground Water* 28:591–604.

Roberts, P. V., M. N. Golz, and D. M. Mackay. 1986. A natural gradient experiment on solute transport in a sand aquifer. 3. Retardation estimates and mass balances for organic solutes. *Water Resour. Res.* 22:2047–2058.

Roberts, P. V., L. Semprini, G. D. Hopkins, D. Grbic-Galic, P. L. McCarty, and M. Reinhard. 1989. *In-situ* aquifer restoration of chlorinated aliphatics by methanotrophic bacteria. *EPA Res. Dev.* 600:1–7.

Semprini, L., P. V. Roberts, G. D. Hopkins, and P. L. McCarty. 1990. A field evaluation of *in-situ* biodegradation of chlorinated ethenes: Part 2, Results of biostimulation and biotransformation experiments. *Ground Water* 28:715–727.

Smith, R. L., S. P. Garbedian, and M. H. Brooks. 1996. Comparison of denitrification activity measurements in ground water using cores and natural gradient tracer tests. *Environ. Sci. Technol.* 30:34–48.

Smith, R. L., and J. H. Duff. 1988. Denitrification in a sand and gravel aquifer. *Appl. Environ. Microbiol.* 54:1071–1078.

Smith, R. L., R. W. Harvey, and D. R. LeBlanc. 1991. Importance of closely spaced vertical sampling in delineating chemical and microbiological gradients in ground water studies. *J. Contam. Hydrol.* 7:285–300.

Smith, R. L., B. L. Howes, and J. H. Duff. 1991. Denitrification in nitrate-contaminated ground water: Occurrence in steep vertical geochemical gradients. *Geochim. Cosmochim. Acta* 55:1815–1825.

Smith, R. L., B. L. Howes, and S. P. Garabedian. 1991. *In situ* measurement of methane oxidation in ground water by using natural-gradient tracer tests. *Appl. Environ. Microbiol.* 57:1997–2004.

Thorn, P. M. and R. M. Ventullo. 1988. Measurement of bacterial growth rates in subsurface sediments using the incorporation of tritiated thymidine into DNA. *Microb. Ecol.* 16:3–16.

Thurman, E. M. 1985. *Organic Geochemistry of Natural Waters.* Nijhoff-Junk, Publishers, Boston, MA.

Trudell, M. R., R. W. Gillham, and J. A. Cherry. 1986. An *in-situ* study of the occurrence and rate of denitrification in a shallow unconfined sand squifer. *J. Hydrol.* 83:251–268.

Wilson, J. T., and B. H. Wilson. 1985. Biotransformation of trichloroethylene in soil. *Appl. Environ. Microbiol.* 49:242–243.

Yoshinari, T., R. Hynes, and R. Knowles. 1977. Acetylene inhibition of nitrous oxide reduction and measurement of denitrification and nitrogen fixation in soil. *Soil Biol. Biochem.* 9:177–183.

6

Modeling of Pesticide Biodegradation in Soil

Daniel R. Shelton, Michael A. Doherty, Timothy B. Parkin, and Joseph A. Robinson

1. Introduction

Pesticides are an integral part of modern agriculture. A significant proportion of the increase in agricultural productivity since World War II is directly attributable to effectively controlling weed, insect, and fungal pests with herbicides, insecticides, and fungicides, respectively. The need for pesticides is underscored by a total U.S. production in 1990 of approximately 1.1 billion lb (Anonymous 1990). However, the intentional release of large quantities of synthetic pesticides has given rise to serious concerns over the potential for adverse human health and environmental effects. Largely as a consequence of Rachel Carson's book *Silent Spring,* the majority of organochlorine insecticides were banned (e.g., DDT) in the 1970s because of their volatility, persistence, and tendency to bioaccumulate in the food chain. In recent years, there has been renewed concern over the fate of soil-applied herbicides and insecticides because of the potential for runoff and/or leaching through the soil profile, resulting in the contamination of surface and ground waters. Various strategies have been proposed for minimizing contamination; however, the only permanent and environmentally benign solution is biodegradation. Consequently, the use of pesticides that are rapidly and extensively biodegraded to CO_2, H_2O, NH_4^+, etc. is desirable.

Rates of biodegradation that are too rapid, however, can lead to losses of efficacy. In the last 10 years, losses of efficacy have been reported for both soil-applied herbicides, such as S-Ethyl dipropyl thiocarbonate (EPTC) (Wilson 1984; Bean et al. 1988; Moorman 1988; Tal et al., 1989), and insecticides, such as isophenphos (Racke and Coats 1987) and carbofuran (Harris et al. 1984; Chapman et al. 1986; Camper et al. 1987; Turco and Konopka 1990). This topic was the

focus of a recent symposium (Racke and Coats 1990). In each instance, the loss of efficacy has been demonstrated to be due to enhanced rates of microbial degradation, which result in soil pesticide concentrations below the lethal threshold required for effective pest control. Therefore, successful pest management strategies using biodegradable soil-applied pesticides will, in large measure, be dependent on the ability to reasonably predict soil pesticide concentrations when pest control is most critical (typically the first 4 to 8 weeks after application and planting) to ensure that losses of efficacy do not occur. Consequently, a knowledge of the kinetics of pesticide biodegradation in soils is necessary to assess both the potential for environmental contamination due to persistence and the losses of efficacy due to accelerated biodegradation.

2. Modeling

Rates of pesticide degradation in soil have traditionally been described by first-order kinetics; Rao and Davidson (1988) have compiled first-order degradation rate constants for a wide range of pesticides. First-order kinetics are to be expected when pesticides are cometabolized, i.e., not utilized as carbon and/or nitrogen sources for growth (Horvath 1972). The pesticide-degradation literature indicates that the majority of currently used pesticides are degraded via cometabolic mechanisms, e.g., oxidative, reductive, and/or hydrolytic transformations; although new strains of microorganisms (or microbial consortia) are periodically isolated with the ability to utilize previously "persistent" pesticides as growth substrates (e.g., atrazine; Mandelbaum et al. 1993; Radosevich et al. 1995). Actually, the more appropriate kinetic expression is pseudo first-order kinetics, where biomass is assumed to remain constant with time. Consequently, depending on the population densities of pesticide-degrading strains, rate constants (half-lives) may vary significantly from one site to another, resulting in the spatial variability generally observed in pesticide dissipation studies.

The primary advantage of assuming first-order kinetics is the ease of data summarization since a single rate constant is obtained. A single rate constant is particularly useful where the goal is to compare and contrast the relative persistence of pesticides, or in pesticide fate models where biodegradation is not the primary focus. For example, models designed to estimate rates of leaching of pesticides to groundwater (e.g., pesticide root zone model (PRZM); Carsel et al. 1984) assume first-order kinetics. Of course, modeling of the effect(s) of environmental variables on rates of pesticide degradation are simplified if first-order kinetics are assumed. It has long been recognized that environmental parameters that affect overall rates of soil microbial respiration and/or activity

(e.g. moisture, temperature, pH) also affect rates of pesticide biodegradation (Lichtenstein and Schulz 1964; Getzin 1968; Walker 1978). Walker (1974) has proposed a pesticide biodegradation model that accounts for temperature effects using the Ahhrenius equation and moisture effects using the empirical equation: $H = a\text{MC}^{-b}$, where H is half-life, MC is moisture content, and a and b are constants, in conjunction with a first-order biodegradation rate constant. For those pesticides that are degraded via cometabolic mechanisms and are relatively nonpersistent, the application of pseudo-first-order kinetics, adjusted for temperature/moisture effects, may result in reasonably accurate estimates of pesticide dissipation.

Deviations from simple first-order kinetics, however, are frequently observed in soils studies. Increasing rates of degradation may be observed due to utilization of pesticides as growth substrates by soil microorganisms ("accelerated" or "enhanced" degradation). Conversely, decreasing rates of biodegradation may also be observed due to sorption to soil particles (decreased bioavailability). It has long been recognized that pesticides in soil partition between soluble and sorbed phases (K_d) and that this partitioning can effect availability to soil microorganisms, as well as exposure to the target pests (Osgerby 1973; Pignatello 1989; Weber et al. 1993). Furthermore, the sorption of pesticides to soil particles can continue over a sustained period of time resulting in both lower bioavailability and decreased extraction efficiencies, i.e. "bound" residues (Hamaker and Goring 1976; Karickhoff 1980; McCall and Agin 1985). The topic of "bound" pesticide residues was reviewed (Kaufman et al. 1976), although there was no apparent consensus as to the precise fate of such residues.

Based on the observation that rates of pesticide dissipation in soil (over a period of several months) were biphasic, Hamaker and Goring (1976) proposed the use of a two-compartment model. Pesticides were proposed to partition between "labile" (i.e., soluble) and "unavailable" (i.e., sorbed) pools, with first-order rate constants (k_1/k_{-1}) describing the rates of transfer between the two pools (i.e., sorption/desorption). Rates of "decomposition" (i.e., degradation) were described by a third first-order rate constant (k), based on the assumption that only pesticides in the soluble pool were available for degradation. Consequently, the more rapid initial rate of decomposition was a reflection of the fraction of pesticide that was readily available for degradation (soluble), whereas the slower rate was a reflection of the fraction that was not readily available (sorbed). This modeling approach has proved to be useful in assessing the long-term environmental fate of relatively persistent pesticides, and has been used successfully by investigators to predict the potential for groundwater contamination and/or carryover between growing seasons (Helling et al. 1994). This modeling approach is not very helpful, however, in assessing the potential for losses of efficacy where short-term rates of biodegradation are of primary importance.

3. More Elaborate Models

Recently, more sophisticated models that include terms for convection, dispersion, and retardation have been used by soil physicists to describe rates of leaching/transport of organic contaminants through subsurface soils (van Genuchten and Wagenet 1989; Brusseau et al. 1991; Gamerdinger et al. 1991). These models allow for the distribution of pesticides into one of three compartments: (1) pesticides in aqueous or soil solution, (2) pesticides sorbed to soil surfaces but that are in rapid equilibrium with the soil solution (readily available for biodegradation), and (3) pesticides that have diffused into soil aggregates and/or organic matter particles and are slowly available for biodegradation. The earlier model, as mentioned, for this is the Hamaker-Goring model. The model assumes that pesticides are initial in soil solution and that degradation and sorption to soil particles occurs simultaneously. Sorbed pesticides are assumed to be unavailable for degradation, but can be degraded after they desorb. Although conceptually similar to the Hamaker-Goring model, these newer models are designed to predict the fate or movement of pesticides over a much shorter time period: hours to days as opposed to weeks to months. The leaching models have proved to be particularly useful for describing the short-term sorptive behavior of pesticides (albeit under homogenous conditions) and in distinguishing rates of sorption/diffusion from rates of biodegradation. However, to the extent that these models typically assume first-order biodegradation, due either to the assumption that pesticides are not utilized as growth substrates or that concentrations are too low to support microbial growth ($<1\mu g/mL$), their applicability in assessing the potential for losses of efficacy is limited.

As previously stated, losses of pesticide efficacy are due to utilization of pesticides as growth substrates, i.e., carbon and/or nitrogen sources. In those instances, the assumption of first-order biodegradation kinetics may not be valid; rather, depending on initial substrate and biomass concentrations, sigmoidal kinetics may be observed. In one of the first studies to demonstrate pesticide biodegradability, Audus (1951) observed sigmoidal kinetics of 2,4-D degradation as a result of microbial growth. Sigmoidal kinetics have been observed in other aqueous and soils studies, but due to the difficulties in summarizing nonlinear, sigmoidal data, deviations from first-order kinetics have typically been ignored or, alternatively, data have been summarized by including a lag period followed by first-order kinetics. With the advent of the digital computer and appropriate nonlinear regression software, however, analysis of sigmoidal data has been greatly simplified.

Recently, a variety of models have been proposed for fitting and summarizing sigmoidal mineralization i.e. conversion to (CO_2) data, or substrate disappearance data. One approach to summarizing sigmoidal mineralization data has been to

utilize nonmechanistic models with the simple objective of reducing the dimensionality of the data to facilitate statistical comparisons. Parkin et al. (1991) used a four-parameter logistic equation (the Morgan-Mercer Flodin (MMF) model), developed as a general model for characterizing the nutritional response of organisms (Morgan et al. 1975), to describe the sigmoidal kinetics of CO_2 production resulting from hydrolysis of the insecticide carbofuran. The MMF model makes no assumptions regarding the nature of microbial growth or pesticide metabolism; the summary parameters produced by the model are DT-50% (time required to achieve 50% of asymptotic dissipation) and maximum rate. This model was successfully used to assess spatial and temporal variability in rates of carbofuran biodegradation in agricultural soils (Parkin and Shelton 1992), and to model the effects of moisture and temperature on rates of biodegradation (Parkin and Shelton 1994). Such an approach may be preferable when treatment comparisons, or treatment modeling, are the primary goal and goodness-of-fit is the critical criterion.

Alternatively, mechanistic models have been proposed for summarizing sigmoidal rates of biodegradation of organic compounds, including pesticides. Models derived from Monod kinetics (see later) have been used successfully to summarize mineralization data from aqueous incubations (Simkins and Alexander 1985). However, these models have not proven to be applicable to soils, presumably, due to sorption and diffusional constraints (Scow et al. 1986). Scow et al. (1986) proposed a two-compartment model (analogous to the Hamaker and Goring model) in which the substrate was assumed to be distributed between two separate pools. This model was generally superior in describing rates of mineralization from soil, however, Monod kinetics were not incorporated into the model; rather, first-order kinetics were used to describe both the rates of transfer between the two pools and the rates of biodegradation. Brunner and Focht (1984) proposed a three-half-order kinetic model for summarizing mineralization data. This model makes no assumptions regarding the reaction order, i.e., it is not assumed that the compound (pesticide) of interest is the primary growth substrate or that biological factors, as opposed to physical and chemical factors, are rate limiting. Consequently, the three-half-order model can be used to summarize first-order, linear, or sigmoidal kinetic data. Despite attempts to provide mechanistic interpretations for rate constants, however, all of these models are essentially exercises in curve fitting and provide little predictive value. In addition, mineralization models are limited by the fact that rates of parent compound disappearance (i.e., losses of efficacy) are not necessarily highly correlated with rates of mineralization. Consequently, it is difficult to ascertain which factors (biological versus physical and chemical) are controlling rates of biodegradation. Consequently, there is a need for more information-intensive models that, with appropriate databases, can be used to predict rates of enhanced pesticide biodegradation.

Rates of bacterial growth have traditionally been described using Monod kinet-

ics (Monod 1949). If the fraction of substrate utilized for cell growth (Y) is assumed to be constant over time, then rates of substrate depletion can be described according to the equation

$$\frac{dS}{dt} = \frac{\mu_{max}SX}{(K_s + S)Y}, \tag{1}$$

where μ_{max} is the maximum growth rate, K_s is the half-saturation growth constant, s is substrate concentration, and X is biomass concentration. Figure 6.1 illustrates the shape of substrate depletion curves (aqueous culture) as a function of different
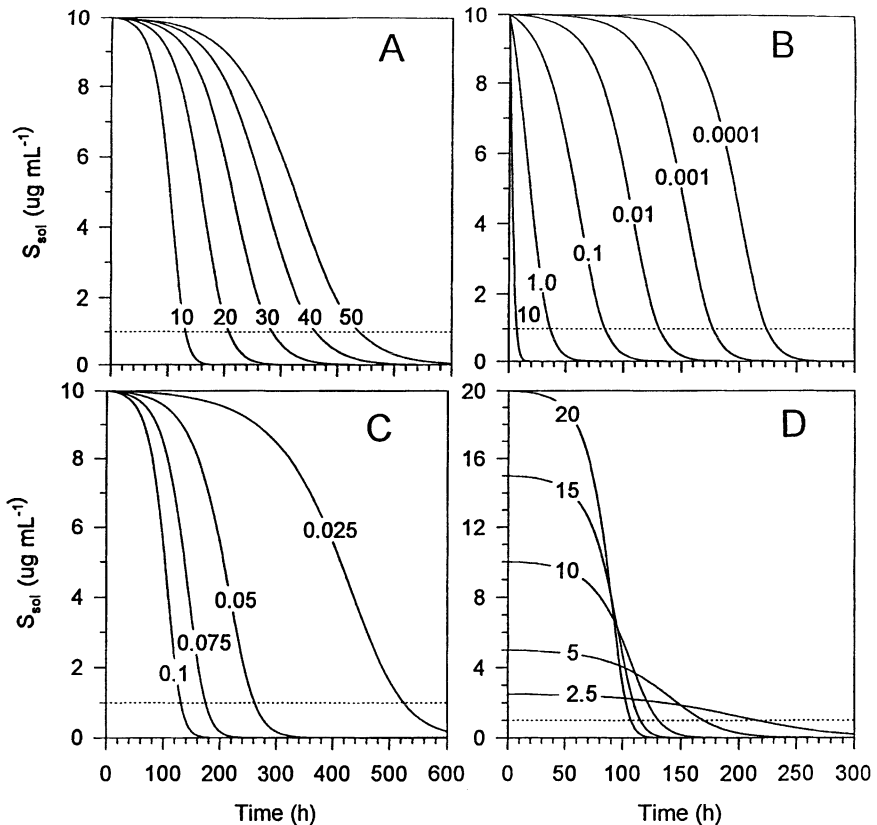
Figure 6.1 Rates of substrate/pesticide and biodegradation in aqueous culture as a function of Monod growth constants: (A) effect of K_s, (B) effect of X_0, (C) effect of μ_{max}, and (D) effect of S_0. Default values were $K_s = 10$ µg/mL, $X_0 = 0.01$ µg/g dry soil, $\mu_{max} = 0.1$/hour, $Y = 0.25$, and $S_0 = 10$ µg/g dry soil.

combinations of K_s, X_0, μ_{max}, and S_0 (simulated data; from Shelton and Doherty 1997b). For purposes of comparison, 1 µg/mL was assumed to represent a hypothetical loss of efficacy. Clearly, differences in growth constants or initial conditions can have dramatic effects on the time required to achieve substrate depletion as well as the shape of the curve (Simkins and Alexander, 1984; Shelton and Doherty 1997b).

Application of Monod kinetics was limited primarily to chemostat studies prior to 1983, when Robinson and Tiedje (1983) demonstrated the potential for estimation of Monod constants from analysis of substrate depletion data using nonlinear regression techniques. The application of this technique to describing the metabolism of organic substrates in aqueous culture has since been demonstrated by others (Simkins and Alexander 1984; Alexander and Scow 1989). The application of microbial growth kinetics to soils is problematic, however, due to the lack of quantitative knowledge regarding boundary conditions and factors controlling microbial processes.

Pesticide concentrations in soil are typically assumed to be in the range of 0.5 to 2 µg/g soil, which is generally assumed to be too low to support microbial growth. This concentration is estimated from dividing the typical application rate for many herbicides (1 to 4 lb/acre) by the approximate weight of mineral soil (plough depth) per acre (~ 2 million lb). However, pesticide concentrations in the surface layer of soil, or in the vicinity of granules, may be significantly higher. For example, the uniform application of a herbicide to the soil surface at a rate of ~ 3 lb/acre (e.g., atrazine) would result in concentrations in the top centimeter of soil of 30 µg/g soil. In the case of carbofuran granules, which are applied in the seed furrow at the time of planting, Shelton et al. (1993) determined that the average carbofuran concentration in soil solution within 1 cm of granules after a heavy rainfall was 24 µg/mL. However, carbofuran concentrations in the immediate vicinity of granules in moist soil (in the absence of rainfall) were calculated to be significantly higher (up to 400 µg/mL, depending on the distance from the granule) due to relatively rapid rates of granular dissolution but slow rates of molecular diffusion away from granules. Consequently, pesticide concentrations in soil shortly after application may be sufficient to support microbial growth, depending on initial biomass levels (X_0) and K_s values.

4. Effect of Microbial Numbers

The extent of microbial growth (i.e., sigmoidicity of the dissipation curve) is, in large measure, dependent on the population density of pesticide degraders at the time of pesticide application. The maximum theoretical biomass of pesticide-degrading bacteria in soils can be estimated from the initial pesticide/substrate concentration (S_0) × yield (Y), where Y is defined as the weight of biomass produced/weight of pesticide consumed. For example, assuming an initial S_0 of

10 µg/g soil and a Y of 0.25 (e.g., 2,4-D), the maximum biomass will be 2.5 µg/g soil; assuming a typical bacterium mass of ~ 0.1 pg (1×10^{-13}), the maximum population density would be 2×10^7 bacteria/g soil. The actual population density of pesticide degraders in soils will also depend on other factors such as mortality rates (i.e., cell death, predation, parasitism) as well as proliferation with alternative growth substrates. Consequently, rates of biodegradation (and losses of efficacy) will depend not only on kinetic growth constants for specific pesticides, but also on mortality rates and growth rates with other substrates throughout the year prior to pesticide application. Actual quantification of population densities of specific pesticide degraders in soil is frequently difficult. Plate counting techniques are reasonably precise (typical CVs = 25%), however, they typically underestimate population densities due to problems with extracting cells from the soil matrix. In addition, this technique is often precluded by limited pesticide solubilities and/or inhibitory effects of parent or product(s) at higher concentrations. An alternative approach is the most probable number (MPN) technique, however, this technique is imprecise, with 95% confidence intervals of approximately one order of magnitude.

Values for Y can vary depending on the extent of metabolism and on the metabolic pathway; in particular, whether the pesticide is utilized as a carbon or nitrogen source. For example, carbofuran-degrading bacteria have been isolated that are able to utilize (1) the aromatic ring as a carbon source, (2) the methylamine moiety as a carbon source, and/or (3) the methylamine moiety as a nitrogen source (Karns et al. 1986; Chaudry and Ali 1988; Ramanand et al. 1988). In each instance, Y varies as a function of the metabolic pathway. Based on pure culture studies with strain WMIII (Karns et al. 1986), utilization of methylamine as a sole carbon source resulted in a Y value of 0.225 µg biomass µg^{-1} methylamine, or 0.03 µg biomass µg^{-1} carbofuran (Shelton and Parkin, unpublished data). However, utilization of methylamine as a sole nitrogen source, in the presence of alternate carbon substrates, would result in an approximate 20-fold increase in biomass (assuming a biomass C:N ratio of 5:1).

There is little information on the range of μ_{max} and K_s values for pesticide-degrading microorganisms. To our knowledge, the only systematic assessment of Monod constants was conducted by Greer et al. (1992), who estimated μ_{max} and K_s values for seven strains of 2,4-D-degrading bacteria in aqueous culture. They observed μ_{max} values ranging from 0.1 hour^{-1} to 1.0 hour^{-1} and K_s values ranging from ~ 2 µg/mL to 30 µg/mL. Whether the variability and/or absolute values observed in this study are applicable to other pesticide-degrading organisms is unclear.

5. Role of Sorption

As previously discussed, sorption plays a major role in limiting bioavailability. Consequently, models designed to estimate enhanced rates of biodegradation in

soil must also account for the effects of sorption on microbial growth. Several studies indicate that rates of microbial growth (μ) are strictly dependent on soluble or soil solution substrate concentrations (Apajalahti and Salkinoja-Salonen 1984; Ogram et al., 1985; Gordon and Millero, 1985; Speitel et al. 1988; Robinson et al. 1990; Shelton and Parkin 1991; Greer and Shelton, 1992). One study appears to indicate that some bacteria are capable of directly utilizing sorbed substrates (Guerin and Boyd 1992); however, the most straightforward interpretation of this data is the production of biosurfactants by one of the strains. Consequently, pesticide residues that are sorbed to soil surfaces and particles must first desorb before biodegradation can occur. Depending on the relative rates of sorption/desorption, a significant fraction of those pesticide residues sorbed to soil surfaces may be rapidly desorbed, hence readily available for biodegradation. In comparison, those pesticide residues that have diffused into soil aggregates, or into organic matter particles, will be slowly available, or "unavailable," for biodegradation since reverse diffusion/desorption must first occur. The size of the "unavailable" pool depends primarily on the relative quantity and size of organic matter particles, as well as the contact time between soil and pesticide prior to microbial degradation. It should be noted that the "unavailable" pool is not necessarily synonymous with "bound residues," since "bound residues" is operationally defined based on a specific extraction procedure.

Scow has recently described a model for studying the effect(s) of sorption and diffusion on rates of microbial degradation (Scow and Hutson 1992; Scow and Alexander 1992). The diffusion-sorption-biodegradation (DSB) model contains terms for sorption to particulate surfaces (linear sorption isotherm), diffusion within a polymeric matrix (radial Fickian diffusion), and biodegradation (including Monod kinetics). The model has been used to describe rates of substrate consumption (usually from mineralization data) in defined soil/polymeric bead slurries. The DSB model has been particularly useful in elucidating the effects of particulate size, sorptivity, and porosity, as well as different Monod growth constants, on rates of substrate degradation. However, to the extent that it describes pesticide behavior in well-defined aqueous systems, results are not directly applicable to soil.

Recently, Shelton, and Doherty (1997a) have proposed a model to describe enhanced rates of pesticide biodegradation in soil, with emphasis on rates of biodegradation in the first 4 to 8 weeks after pesticide application. The compartmentalized model, with the relevant differential equations, is shown in Figure 6.2. The model is based on the following assumptions: (1) pesticides are used as growth substrates (carbon and/or nitrogen sources) by soil microorganisms; (2) rates of biodegradation can be described by Monod kinetics, where the pesticide is the rate limiting nutrient and yield (Y) is constant; (3) rates of microbial growth (μ) are strictly dependent on soil solution concentrations (soluble pool); (4) there is a rapid rate of sorption/desorption of pesticides to and from soil
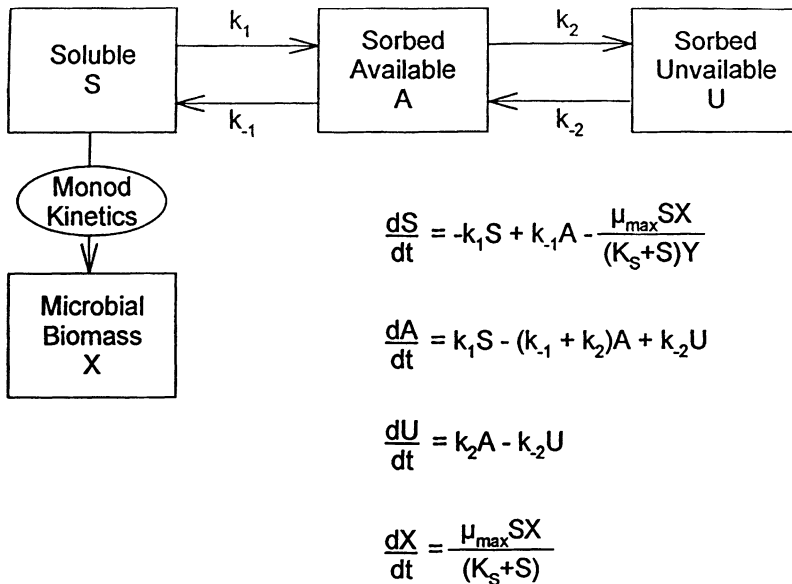
Figure 6.2 Four-compartment model with relevant differential equations describing substrate depletion (soluble compartment), microbial growth (biomass), sorption to soil surfaces (sorbed-available compartment), and diffusion into the soil matrix (sorbed-unavailable compartment).

surfaces (sorbed-available pool); (5) there is a time-dependent diffusion of pesticides to less accessible sites such as aggregates and/or organic matter particles (sorbed-unavailable); and (6) rates of sorption/desorption and diffusion can be described by first-order kinetics. Although conceptually similar to previous models (Brusseau et al. 1991; Gamerdinger et al. 1991; Scow and Hutson 1992; Estrada et al. 1993), the proposed model differs in that it attempts to describe pesticide behavior in actual field moist surface soils. Pesticides in the soluble, sorbed-available, and sorbed-unavailable pools are estimated using a hydraulic pressure technique that allows for independent estimates of soluble versus sorbed pesticide concentrations in soil over a wide range of moisture contents (Shelton and Parkin 1991; Greer and Shelton 1992). This technique assumes that pesticides are homogeneously distributed throughout the soil solution and that extraction efficiencies remain constant with time.

   Here K_{d1} (k_1/k_{-1}) is operationally defined as the ratio of pesticide sorbed to soil surfaces (sorbed-available), shortly after application before significant diffusion into aggregates/particles has occurred (< 3 hours) versus pesticide in soil solution (soluble). On the other hand, K_{d2} (k_2/k_{-2}) is operationally defined as the time-

dependent diffusion/sorption of pesticides to "internal sites" (McCall and Agin 1985) within the organic matter/aggregate matrix (sorbed-unavailable). In addition, the mass of pesticide in the sorbed-unavailable pool can be calculated by subtracting sorbed-available from sorbed-total, where sorbed-available = $S_{sol} \times K_{dl}$. Although k_1 and k_{-1} are defined in terms of rates of sorption/desorption, they also undoubtedly reflect the rate of diffusion through the soil solution to and from the sorptive sites, which under some circumstances (particularly low moisture conditions), may be the rate limiting process (Shelton and Parkin 1991). Similarly, although k_2 and k_{-2} are defined in terms of first-order rate constants, they reflect the much more complicated process of intraorganic matter diffusion, where the rate of diffusion is a function of K_{dl}, amount and relative size distribution of organic matter particles (ratio of internal to external sorption sites), and organic matter matrix size in conjunction with the pesticide molecular weight (Wu and Gschwend 1986).

The ability of the model to accurately estimate sorption, diffusion, and microbial growth rate constants is illustrated in Figures 6.3 and 6.4 (from Shelton and Doherty 1997a) using 2,4-D degradation data from Greer and Shelton (1992). High and low organic matter soils, amended with 100 µg 2,4-D g^{-1} soil, were inoculated with a 2,4-D-degrading *Alcaligenes* (strain MI), and soluble versus sorbed concentrations of 2,4-D monitored with time. High r^2 values indicate that estimates of k_1/k_{-1}, k_2/k_{-2}, μ_{max}, and K_s generated by the model accurately described the data. Estimates of K_s and μ_{max} were generally comparable with previous pure culture determinations ($\mu_{max} = 0.19$ hour^{-1}, $K_s = 2.7$ µg/mL; Greer et al. 1992). These observations suggest that rates of microbial growth in soil are consistent with Monod kinetics when adjusted for soil solution (soluble) pesticide concentrations.

Potential limitations of this model are that (1) the pesticide may not be the growth-limiting nutrient, or the primary growth substrate; (2) it does not account for cell mortality; and (3) it does not account for temperature and moisture effects on microbial growth. Although inorganic nutrients are not likely to be limiting in agricultural soils, the utilization of alternative substrates could result in inaccurate estimates of biomass. Depending on microbial growth rates versus mortality rates (e.g., parasitism, predation, cell death), mortality could be an important variable. Fortunately, the Monod growth equation can be easily modified to account for mortality, pending the availability of reliable data. Accounting for moisture and temperature effects could be more difficult. Although models have been proposed to account for moisture and temperature effects, assuming either first-order (Walker 1974) or sigmoidal kinetics (Parkin and Shelton 1994), neither of these models explicitly described the effects of moisture and/or temperature on rates of sorption, diffusion, or microbial growth constants (μ_{max} and K_s). Ultimately, however, in conjunction with information on rates of pesticide dissolution from various formulations, moisture and temperature effects, rates of microbial mortal-
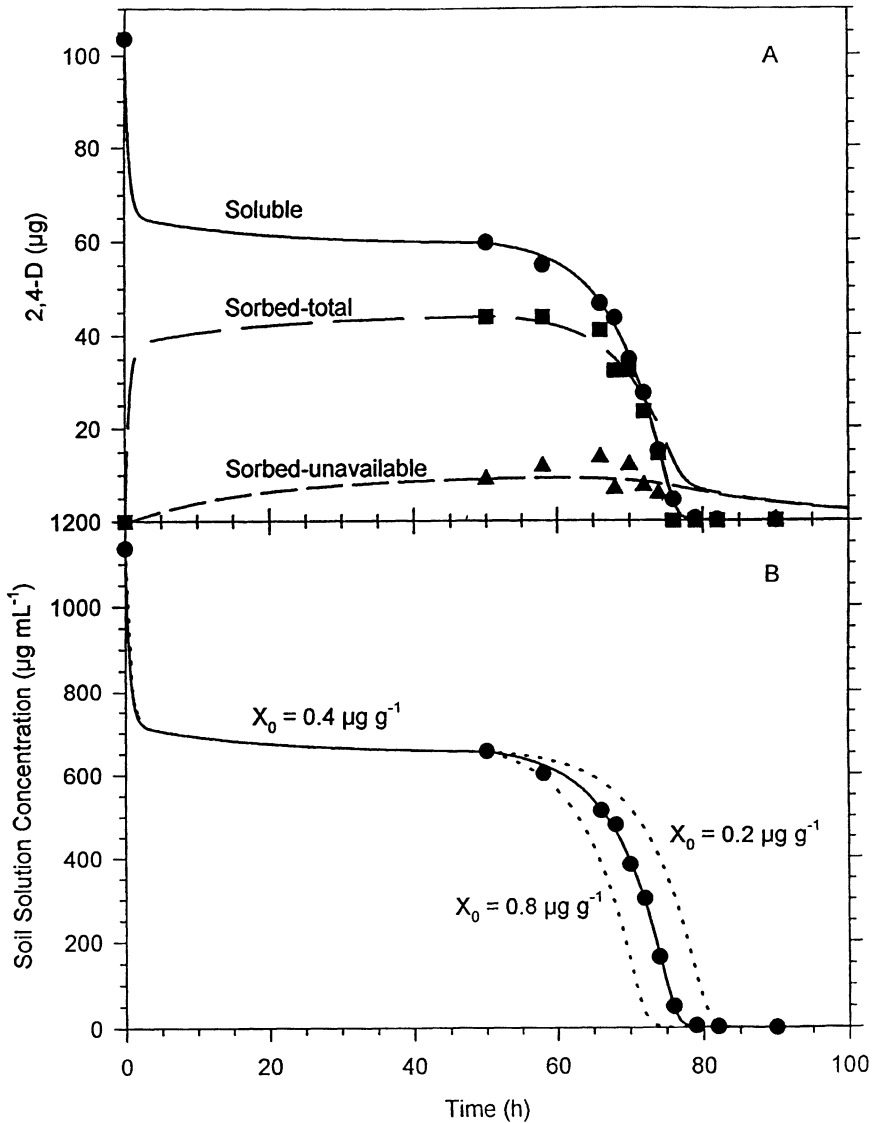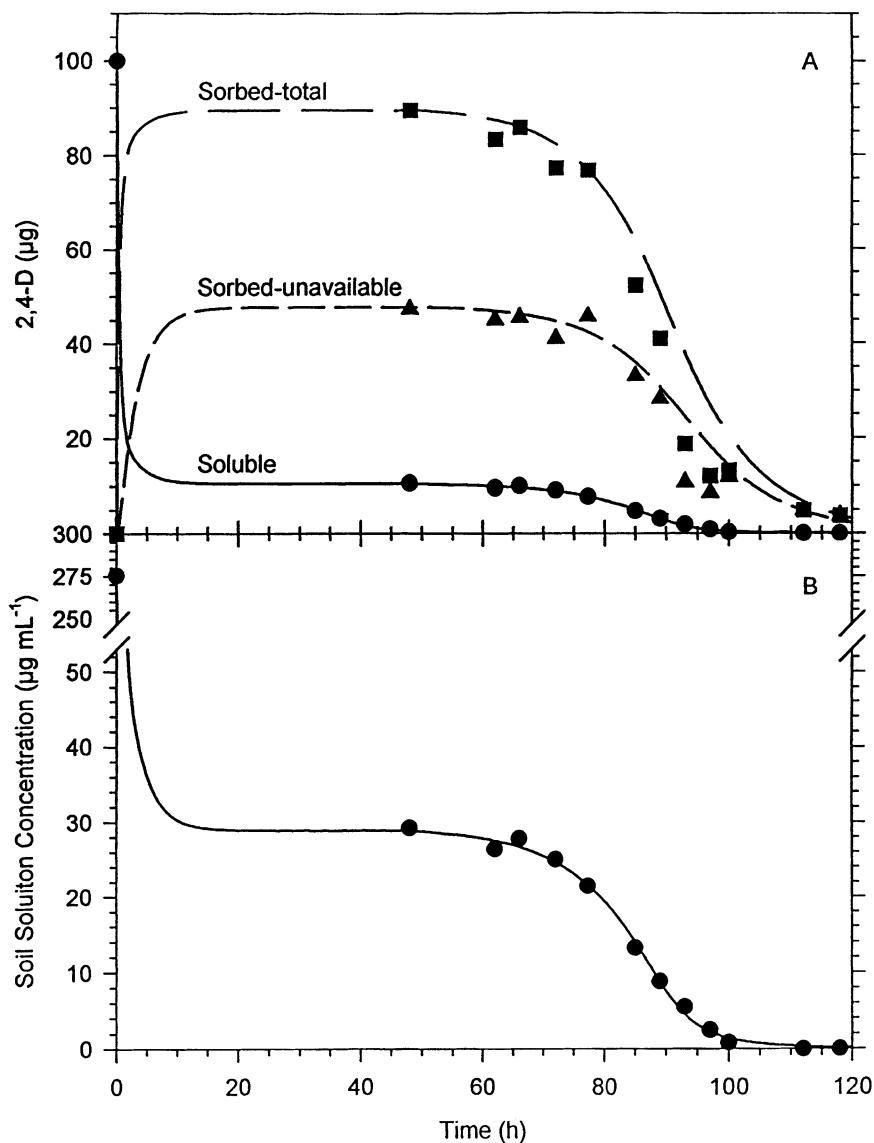
Figure 6.3 Soluble, sorbed-total, sorbed-unavailable (A), and soil solution (B) 2,4-D (~ 100 µg/g soil) in low organic matter soil inoculated with MI. Symbols represent measured or calculated data; lines represent model estimates. Model estimates for k_1/k_{-1}, k_2/k_{-2}, X_0, μ_{max}, and K_s were 0.726/12.8, 0.015/0.055, 0.4, 0.17, and 41.5, respectively ($r^2 = 0.9994$).

Figure 6.4 Soluble, sorbed-total, sorbed-unavailable (A), and soil solution (B) 2,4-D (~ 100 μg/g soil) in high organic matter soil inoculated with MI. Symbols represent measured or calculated data; lines represent model estimates. Model estimates for k_1/k_{-1}, k_2/k_{-2}, X_0, μ_{max}, and K_s were 2.75/1.21, 0.314/0.275, 0.2, 0.13, and 5.9, respectively ($r^2 = 0.9999$).

ity, and rates of pest mortality as a function of time and pesticide concentration, it should be possible to develop a predictive model for estimating potential losses of efficacy. Such a predictive model is necessary if biodegradable, soil-applied pesticides are to be extensively used in agricultural production.

6. Summary

A variety of models have been proposed over the past 30 years for estimating rates of pesticide biodegradation in soil. All of the models have some utility depending on the nature of the metabolic pathway, the time frame of degradation, and the purpose(s) for which the model is to be used. The simplest approach is to assume first-order kinetics, where the rate of biodegradation is dependent only on pesticide concentration. The primary advantage of this model is that a single rate constant is generated, which is particularly useful where the goal is to compare/contrast the long-term persistence of pesticides. There are severe limitations, however; the model assumes that pesticides are co-metabolized (not used as a growth substrate), and it does not account for the effects of sorption on bioavailability or temperature and moisture on microbial activity. Modifications to the basic first-order model have been proposed to account for sorption or temperature/moisture effects on rates of biodegradation. A first-order model has been described which incorporates the Ahhrenius equation and the empirical equation, $H = a\mathrm{MC}^{-b}$, to account for temperature and moisture effects, respectively. The long-term effects of sorption on rates of biodegradation have been modeled by partitioning pesticides into soluble versus sorbed pools, and assuming that only pesticides in the soluble pool are bioavailable. This model incorporates first-order rate constants to describe the rates of transfer between the two pools (i.e., sorption/desorption) and the rate of biodegradation. Recently, more sophisticated models have been described that include rate terms for sorption of pesticides to soil surfaces (rapid process) and subsequent diffusion into soil aggregates and/or organic matter particles (slower process). These models have been particularly useful in describing the short-term bioavailability of pesticides. None of the above models, however, can account for utilization of pesticides as microbial growth substrates (i.e., sigmoidal kinetics). Recently, models have been proposed that combine pesticide sorption/diffusion with microbial growth (Monod) kinetics. These models are significantly more complex since nonlinear regression techniques are required to describe rates of biodegradation. However, to the extent that readily biodegradable pesticides are to be extensively used in agricultural production, biodegradation models are needed which can accurately account for microbial activity/growth as a function of bioavailability and soil parameters such as temperature and moisture.

References

Alexander, M., and K. M. Scow. 1989. Kinetics of biodegradation in soil. In B. L. Sawhney, K. Brown (eds.), *Reactions and Movement of Organic Chemicals in Soils,* pp. 243–269. SSSA Special Publication No. 22. Madison, WI.

Anonymous. 1990. *National Agricultural Chemicals Association Industry Profile.* Compiled by Ernst and Young, Washington, DC.

Apajalahti, J. H. A., and M. S. Salkinoja-Salonen. 1984. Absorption of pentachlorophenol (PCP) by bark chips and its role in microbial PCP degradation. *Microb. Ecol.* 10:359–367.

Audus, L. J. 1951. The biological detoxification of hormone herbicides in soil. *Plant and Soil* 3:170–192.

Ball, W. P., and P. V. Roberts. 1991. Diffusive rate limitations in the sorption of organic chemicals. In R. A. Baker (ed.), *Organic Substances and Sediments in Water: Vol. 2, Processes and Analytical,* pp. 273–310. Lewis Publishers, Chelsea, MI.

Bean, B. W., F. W. Roeth, A. R. Martin, and R. G. Wilson. 1988. Influence of prior pesticide treatments on EPTC and butylate degradation. *Weed Sci.* 36:70–77.

Brunner, W., and D. D. Focht. 1984. Deterministic three-half-order kinetic model for microbial degradation of added carbon substrates in soil. *Appl. Environ. Microbiol.* 47:167–172.

Brusseau, M. L., R. E. Jessup, and P. S. C. Rao. 1991. Nonequilibrium sorption of organic chemicals: Elucidation of rate-limiting processes. *Environ. Sci. Technol.* 25:134–142.

Camper, N. D., M. N. Fleming, and H. D. Skipper. 1987. Biodegradation of carbofuran in pretreated and nonpretreated soils. *Bull. Environ. Contam. Toxicol.* 39:571–578.

Carsel, R. F., C. N. Smith, L. A. Mulkey, and J. D. Jowise. 1984. *User's Manual for the Pesticide RootZone Model (PRZM): Release 1.* U.S. Environmental Protection Agency. EPA-600/3-109. Athens, GA.

Chapman, R. A., C. R. Harris, and C. Harris. 1986. Observations on the effect of soil type, treatment intensity, insecticide formulation, temperature and moisture on the adaptation and subsequent activity of biological agents associated with carbofuran degradation in soil. *J. Environ. Sci. Health.* B21:125–141.

Chaudry, G. R., and A. N. Ali. 1988. Bacterial metabolism of carbofuran. *Appl. Environ. Microbiol.* 54:1414–1419.

Estrella, M. R., M. L. Brusseau, R. S. Maier, I. L. Pepper, P. J. Wierenga, and R. M. Miller. 1993. Biodegradation, sorption, and transport of 2,4-dichlorphenoxyacetic acid in saturated and unsaturated soil. *Appl. Environ. Microbiol.* 59:4266–4273.

Gamerdinger, A. M., A. T. Lemley, and R. J. Wagnet. 1991. Nonequilibrium sorption and degradation of three 2-chloro-*s*-triazine herbicides in soil-water systems. *J. Environ. Qual.* 20:815–822.

Getzin, L. W. 1968. Persistence of diazinon and zinophos in soil: effects of autoclaving, temperature, moisture, and acidity. *J. Econ. Entomol.* 61:1560–1565.

Gordon, A. S., and F. J. Millero. 1985. Adsorption mediated decrease in the biodegradation rate of organic compounds. *Microb. Ecol.* 11:289–298.

Greer, L. E., J. A. Robinson, and D. R. Shelton. 1992. Kinetic comparison of seven strains of 2,4-dichlorophenoxyacetic acid-degrading bacteria. *Appl. Environ. Microbiol.* 58:1027–1030.

Greer, L. E., and D. R. Shelton. 1992. Effect of inoculant strain and organic matter content on kinetics of 2,4-dichlorophenoxyacetic acid degradation in soil. *Appl. Environ. Microbiol.* 58:1459–1465.

Guerin, W. F., and S. A. Boyd. 1992. Differential bioavailability of soil-sorbed naphthalene to two bacterial species. *Appl. Environ. Microbiol.* 58:1142–1152.

Hamaker, J. W., and C. A. I. Goring. 1976. Turnover of pesticide residues in soil. In D. D. Kaufman, C. G. Still, G. D. Paulson, and S. K. Bandal (eds.), *Bound and Conjugated Pesticide Residues,* pp. 219–243. ACS Symposium Series 29. Washington, DC.

Harris, R., R. A. Chapman, C. Harris, and C. M. Tu. 1984. Biodegradation of pesticides in soil: Rapid reduction of carbonate degrading factors after carbofuran treatment. *J. Environ. Sci.* B19:1–11.

Helling, C. S., B. F. Engelke, and M. A. Doherty. 1994. DDT dissipation in hawaiian *in-situ* soil columns. *J. Environ. Sci. Health.* B29(1):103–119.

Horvath, R. S. 1972. Microbial co-metabolism and degradation of chloroaromatics in nature. *Bacteriol. Rev.* 36:146–155.

Karickhoff, S. W. 1980. Sorption kinetics of hydrophobic pollutants in natural sediments. In R. A. Baker (ed.), *Contaminants and Sediments,* pp. 193–205. Ann Arbor Science, Ann Arbor, MI.

Karns, J. S., W. W. Mulbry, J. O. Nelson, and P. C. Kearney. 1986. Metabolism of carbofuran by a pure bacterial culture. *Pest. Biochem. Physiol.* 25:211–217.

Kaufman, D. D., C. G. Still, G. D. Paulson, and S. K. Bandal. 1976. *Bound and Conjugated Pesticide Residues.* ACS Symposium Series 29. Washington, DC.

Lichtenstein, E. P., and K. R. Schulz. 1964. The effects of moisture and microorganisms on the persistence and metabolism of some organophosphorus insecticides in soils with special emphasis on parathion. *J. Econ. Entomol.* 57:618–627.

Mandelbaum, R. T., P. W. Wackett, and D. L. Allan. 1993. Mineralization of the s-triazine ring of atrazine by stable bacterial mixed cultures. *Appl. Environ. Microbiol.* 59:1695–1701.

McCall, P. J., and G. L. Agin. 1985. Desorption kinetics of picloram as affected by residence time in the soil. *Environ. Toxicol. Chem.* 4:37–44.

Monod, L. J. 1949. The growth of bacterial cultures. *Ann. Rev. Microbiol.* 3:371–394.

Moorman, T. B. 1988. Populations of EPTC-degrading microorganisms in soils with accelerated rates of EPTC degradation. *Weed Sci.* 36:96–101.

Morgan, P. H., L. P. Mercer, and N. W. Flodin. 1975. General model for nutritional responses of higher organisms. *Proc. Nat. Acad. Sci.* 72:4327–4331.

Ogram, A. V., R. E. Jessup, L. T. Ou, and P. S. C. Rao. 1985. Effects of sorption on

biological degradation rates of 2,4-dichlorophenoxyacetic acid in soils. *Appl. Environ. Microbiol.* 49:582–587.

Osgerby, J. M. 1973. Processes affecting herbicide action in soil. *Pest. Sci.* 4:247–258.

Parkin, T. B., and D. R. Shelton. 1992. Spatial and temporal variability of carbofuran degradation in soil. *J. Environ. Qual.* 21:672–678.

Parkin, T. B., and D. R. Shelton. 1994. Modeling environmental effects on enhanced carbofuran degradation in soil. *Pest. Sci.* 40:163–168.

Parkin, T. B., D. R. Shelton, and J. A. Robinson. 1991. Evaluation of methods for characterizing carbofuran hydrolysis in soil. *J. Environ. Qual.* 20:763–769.

Pignatello, J. J. 1989. Sorption dynamics of organic compounds in soils and sediments. In B. L. Sawhney and K. Brown (eds), *Reactions and Movement of Organic Chemicals in Soils,* pp. 45–80. SSSA Special Publication No. 22. Madison, WI.

Racke, K. D., and J. R. Coats. 1987. Enhanced degradation of isofenphos by soil microorganisms. *J. Agric. Food Chem.* 35:94–99.

Racke, K. D., and J. R. Coats. 1990. *Enhanced Biodegradation of Pesticides in the Environment.* American Chemical Society, ACS Symposium Series, 426. Washington, DC.

Radosevich, M., S. J. Traina, Y.-L. Hao, and O. H. Touvinen. 1995. Degradation and mineralization of atrazine by a soil bacterial isolate. *Appl. Environ. Microbiol.* 61:297–302.

Ramanand, K., M. Sharmila, and N. Sethunathan. 1988. Mineralization of carbofuran by a soil bacterium. *Appl. Environ. Microbiol.* 54:2129–2133.

Rao, P. S. C., and J. M. Davidson. 1988. Estimation of pesticide retention and transformation parameters required in nonpoint source pollution models. In M. R. Overcash and J. M. Davidson (eds.), *Environmental Impact of Nonpoint Source Pollution,* pp. 23–67. Ann Arbor Science Publishers, Ann Arbor, MI.

Robinson, K. G., W. S. Farmer, and J. T. Novak. 1990. Availability of sorbed toluene in soils for biodegradation by acclimated bacteria. *Water Res.* 24:345–350.

Robinson, J. A., and J. M. Tiedje. 1993. Nonlinear estimation of Monod growth kinetic parameters from a single substrate depletion curve. *Appl. Environ. Microbiol.* 45:1453–1458.

Scow, K. M., and M. Alexander. 1992. Effect of diffusion on the kinetics of biodegradation: Experimental results with synthetic aggregates. *Soil Sci. Soc. Am. J.* 56:128–134.

Scow, K. M., and J. Hutson. 1992. Effect of diffusion and sorption on the kinetics of biodegradation: Theoretical considerations. *Soil Sci. Soc. Am. J.* 56:119–127.

Scow, K. M., S. Simkins, and M. Alexander. 1986. Kinetics of mineralization of organic compounds at low concentrations in soil. *Appl. Environ. Microbiol.* 51:1028–1035.

Shelton, D. R., and M. A. Doherty. (1997a). A model for estimating enhanced rates of pesticide biodegradation in soil. Submitted to *Soil Sci. Soc. Am. J.* July/August.

Shelton, D. R., and M. A. Doherty. (1997b). Estimating losses of efficacy due to enhanced

rates of biodegradation in soil: Model simulations. Submitted to *Soil Sci. Soc. Am. J.* July/August.

Shelton, D. R., and T. B. Parkin. 1991. Effect of moisture on sorption and biodegradation of carbofuran in soil. *J. Agric. Food Chem.* 39:2063–2068.

Shelton, D. R., A. M. Sadeghi, and A. R. Insensee. 1993. Estimation of granular carbofuran dissolution rates in soil. *J. Agric. Food Chem.* 41:1134–1138.

Simkins, S., and M. Alexander. 1984. Models for mineralization kinetics with the variables of substrate concentration and population density. *Appl. Environ. Microbiol.* 47:167–172.

Simkins, S., and M. Alexander. 1985. Nonlinear estimation of the parameters of monod kinetics that best describe mineralization of several substrate concentrations by dissimilar bacterial densities. *Appl. Environ. Microbiol.* 50:816–824.

Speitel, G. E., C. J. Lu, M. Turakhia, and X.-J. Sho. 1988. Biodegradation of trace concentrations of substituted phenols in granular activated carbon columns. *Environ. Sci. Technol.* 23:68–74.

Tal, A., B. Rubin, J. Katan, and N. Aharonson. 1989. Fate of ^{14}C-EPTC in a soil exhibiting accelerated degradation of carbamathioate herbicides and its control. *Weed Sci.* 37:434–439.

Turco, R. F., and A. Konopka. 1990. Biodegradation of carbofuran in enhanced and non-enhanced soils. *Soil Biol. Biochem.* 22:195–201.

van Genuchten, M. T., and R. J. Wagenet. 1989. Two-site/two region models for pesticide transport and degradation: theoretical development and analytical solutions. *Soil Sci. Soc. Am. J.* 53:1303–1310.

Walker, A. 1974. A simulation model for prediction of herbicide persistence. *J. Environ. Qual.* 3:396–401.

Walker, A. 1978. The degradation of methazole in soil. I. effects of soil type, soil temperature and soil moisture content. *Pest. Sci.* 9:326–332.

Weber, J. B., J. A. Best, and J. U. Gonese. 1993. Bioavailability and bioactivity of sorbed organic chemicals. In D. M. Linn (ed.), *Sorption and Degradation of Pesticides and Organic Chemicals in Soil,* pp. 153–196. SSSA Special Publication No. 32. Madison, WI.

Wilson, R. G. 1984. Accelerated degradation of thiocarbamide herbicides in soil with prior thiocarbamide herbicide exposure. *Weed Sci.* 32:264–268.

Wu, S.-C., and P. M. Gschwend. 1986. Sorption kinetics of hydrophobic organic compounds to natural sediments and soils. *Environ. Sci. Technol.* 20:717–725.

7

Modeling Nitrogen Transformations in Soil

David D. Myrold

1. Introduction

Studies of N cycling in soils are inherently concerned with measuring process rates and examining the environmental factors that affect these rates. Consequently, models and modeling always have been a part of such studies.

Nitrogen cycle models come in an almost infinite variety. Therefore it is useful to sort these models into meaningful and useful categories. This could be done in numerous ways: by N cycle process or the number of N transformations (complexity), theoretical versus empirical, static versus dynamic, deterministic versus stochastic, etc. Although I use these criteria to some degree, I attempt to categorize N cycle models by their use. Models of N transformations can be used to (1) generate data, e.g., calculate process rates or pool sizes; (2) gain further understanding of how the N cycle is regulated and functions; and (3) make predictions. My approach is to discuss a few models in each category. The models were selected to represent the range of models that have been applied to N transformations.

2. Using Models to Calculate Data

Studies of the N cycle invariably involve measuring pool sizes and/or rates of N cycle transformations (Legg and Meisinger, 1982). Doing so, in most cases, requires at least a conceptual model of the N cycle and often a mathematical model. Examples of applications of models for determining pool sizes or rates include: N budgets, kinetic expressions, and compartmental models.

2.1. Nitrogen Budgets

Perhaps the first used and conceptually simplest type of model used in studies of N cycling was the construction of a N budget, which are also known as mass balance or input-output models. A N budget is basically an accounting approach that sums all inputs and subtracts all outputs from a system to determine the net gain or loss of N to the system. The system itself is often treated as a "black box" and transformations of N within the system are ignored. In many cases where the soil is the system of interest, the soil N is assumed to be in equilibrium, i.e., there is no net gain or loss of soil N, and the difference between known N inputs and outputs can be used to identify unknown, or unmeasured processes. This was often how denitrification was estimated, for example, prior to the introduction of ^{15}N and acetylene block methods.

Nitrogen budgets have been applied at scales ranging from a few grams of soil in a laboratory flask to the globe (Hauck and Tanji, 1982). Their most common application, however, has probably been at the field-plot scale. Fertilizer recovery experiments in agricultural systems are a typical example. The sensitivity of the mass balance approach can be enhanced by using ^{15}N-labeled materials.

The recently reported "sandbox" study by Bormann et al. (1993) is a good example of a N mass balance. The sandboxes, consisting of reconstituted soil horizons, were isolated on all four sides and capped on the bottom to measure N contained in water escaping the rooting zone. Each sandbox was vegetated with a different tree species or left unvegetated as a control. Over a period of 3 to 5 years, N inputs in precipitation, N outputs in leachate, and changes in N stored in the soil and vegetation were measured. This is summarized by the following mass balance equation:

$$N_{unexplained} = \Delta N_{storage} + N_{precipitation} - N_{drainage} \tag{1}$$

Precipitation inputs were relatively small with low variability (5.2 ± 0.5 kg N ha^{-1} yr^{-1}) and were largely balanced by small but more variable drainage outputs (0.3 to 7.0 kg N ha^{-1} yr^{-1}, depending on treatment). Changes in N stored in vegetation and soils were typically an order of magnitude greater and had much larger relative variability, despite the great care taken in this experiment. The high variability in vegetation and soil N storage is often the major weakness in constructing N budgets (Legg and Meisinger, 1982). The most interesting result from Bormann et al. (1993) is the relatively large amount of unexplained N (Fig. 7.1). Without vegetation, there an unexplained loss of 94 ± 75 kg N ha^{-1} yr^{-1}, which is far in excess of what might be expected for denitrification and volatilization losses. The sandbox containing black alder, a symbiotic N$_2$ fixing species, gained 320 ± 70 kg N ha^{-1} yr^{-1}, which is on the upper end of previous estimates of N$_2$ fixation in alder strands (Hibbs and Cromack, 1990). More surprising was
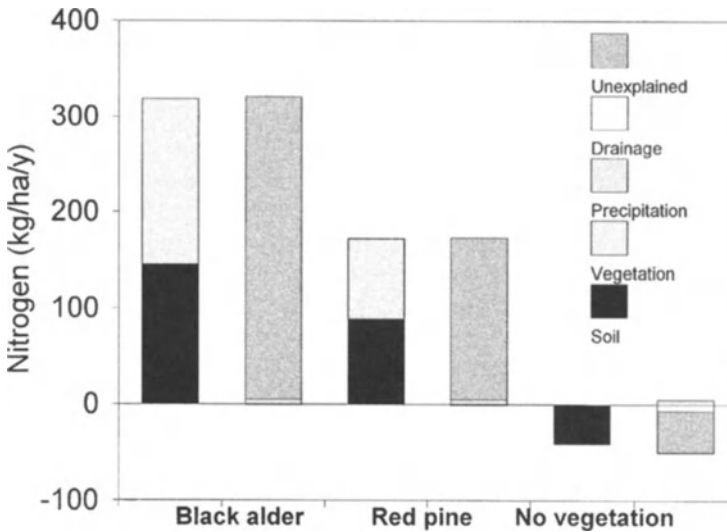
Figure 7.1 Nitrogen balance for three of the "sandbox" systems of Bormann et al. (1993). The left bar of each system shows the gains or losses of N within the soil and vegetation within the plant-soil system. The right bar represents measured inputs and outputs and also shows the unexplained gain or loss of N assuming that each sandbox is a closed system.

the large gain in N (140 ± 90 kg N ha^{-1} yr^{-1}) in the red pine sandbox, which was presumably from nonsymbiotic and/or root-associated N$_2$ fixation, a process not thought to add more than a few kg N ha^{-1} yr^{-1} to most soils. These results illustrate that, even with the relatively large errors that accumulate when using the mass balance approach, N budgets can be important in learning more about how N is cycled in soils.

2.2. Kinetic Expressions

There has been a long standing interest in the kinetics of N transformations in soil. Of these transformations, the kinetics of N mineralization have been studied most extensively. Stanford and Smith (1972) can probably be credited with starting the trend of describing net N mineralization using first-order kinetics: $N = N_0(1 - e^{-kt})$, where N is the amount of inorganic N (NH$_4^+$ + NO$_3^-$) at a given time, N_0 is the amount of potentially mineralizable N, k is the first-order rate constant, and t is time. An example of this approach is shown for some simulated data (Fig. 7.2).

Although Broadbent (1986) found that a power function ($N = At^{1/2}$, where A is an arbitrary constant) fit the data of Stanford and Smith (1972) better than the
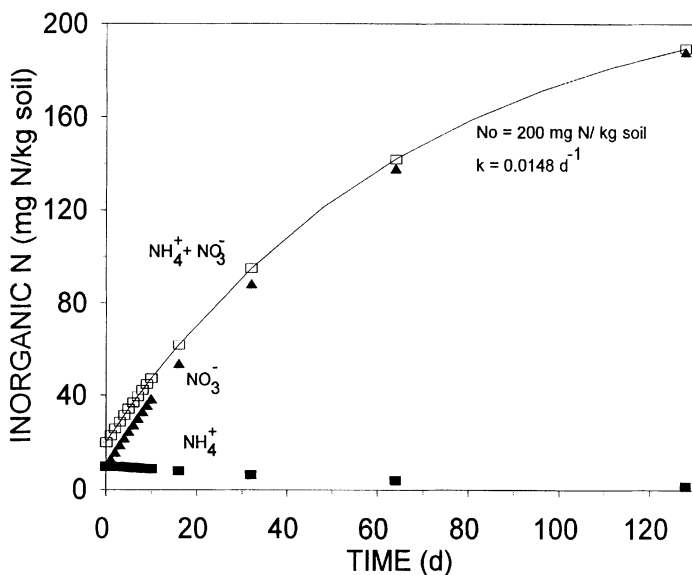
Figure 7.2 The net change in soil inorganic N pools from a simulation using the model shown in Fig. 3. The solid line is the best-fit curve based on a first-order N mineralization model (Stanford and Smith 1972). This model was used to estimate the available organic N pool (200 mg N kg^{-1} soil) and first-order rate constant (0.0148 day^{-1}).

first-order kinetic model, the first-order model is still preferred because of the potential utility of N_0 and the intuitively satisfying assumption that the reaction rate should depend on substrate concentration. Subsequently, many additional models have been proposed, some completely empirical, others with a greater mechanistic basis (Juma et al. 1984; Deans et al. 1986; Bonde and Lindberg 1988; Ellert and Bettany 1988; White and Marinakis 1991). The overriding conclusion from these studies is that no single model gives the best fit to the data for all soils or under all conditions. The best guidance is to use the simplest model that fits the data well and to avoid simplistic mechanistic interpretations.

2.3. Compartmental Models

The distinction between kinetic models and compartmental models is somewhat arbitrary but for this discussion compartmental models will be defined as having more than one compartment (N pool) and/or more than one process rate. Typically a compartmental model is graphically depicted as a set of boxes, the compartments

or N pools, connected by arrows, the N transformations (Fig. 7.3). This pictorial representation of a compartmental model is written mathematically as a series of first-order differential equations, which represent the changes in N pool sizes over time.

The simpler compartmental models, i.e., those with relatively few pools and rates, can be solved analytically. Some early examples of these, and ones that are still used widely today, are the models of Kirkham and Bartholomew (1954, 1955). By using $^{15}NH_4^+$ as a tracer, gross rates of N mineralization and N immobilization can be calculated given certain assumptions about the kinetic order of the reaction and whether immobilized N can be mineralized.

The compartmental model shown in Fig. 7.3 was used to create a simulated dataset. For the simulation, all reactions were assumed to be first-order, remineralization of immobilized N was allowed, and the NH_4^+ pool was labelled with $^{15}NH_4^+$ at zero time. Initial pool sizes were 200 mg N kg^{-1} for organic N and 10 mg N kg^{-1} for NH_4^+ and NO_3^- (Figs. 7.2 and 7.4). First-order rate constants of 0.025 day^{-1} for N mineralization, 0.2 day^{-1} for N immobilization, and 0.3 day^{-1} for nitrification were used. This simulation resulted in a gradual decrease in NH_4^+ concentration and a continuous, curvilinear increase in NO_3^- (Fig. 7.2). The ^{15}N in the NH_4^+ pool was rapidly diluted (Fig. 7.4) whereas the atom % ^{15}N of the NO_3^- pool increased rapidly, reached a maximum at 3 days, and declined thereafter. It is interesting to note that the calculated rate constant for net N mineralization (0.0148 day^{-1}, Fig. 7.2) was much smaller than the gross N mineralization rate constant, however, the N_0 estimate of 199 mg N kg^{-1} agreed well with the organic N pool size.

The first 10 days of simulated data were analyzed by the method of Kirkham and Bartholomew (1955) to determine daily gross N mineralization rates. The actual gross N mineralization rate, based on the organic N pool size and first-order rate constant decreased gradually from 5 to about 4.5 mg N kg^{-1} day^{-1} over this time period (Fig. 7.5). The gross N mineralization rate calculated by assuming that remineralization occurred agreed quite well with the actual rates, just slightly overestimating the rate for each daily interval. When remineralization was excluded, gross N mineralization rates were underestimated and after 2 days this
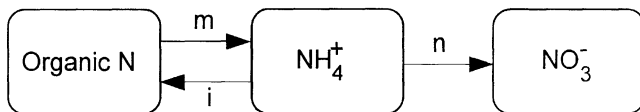
Figure 7.3 An example of a simple compartmental model for the soil N cycle. There are three pool of N and three transformation processes: N mineralization (*m*), N immobilization (*i*), and autotrophic nitrification (*n*).
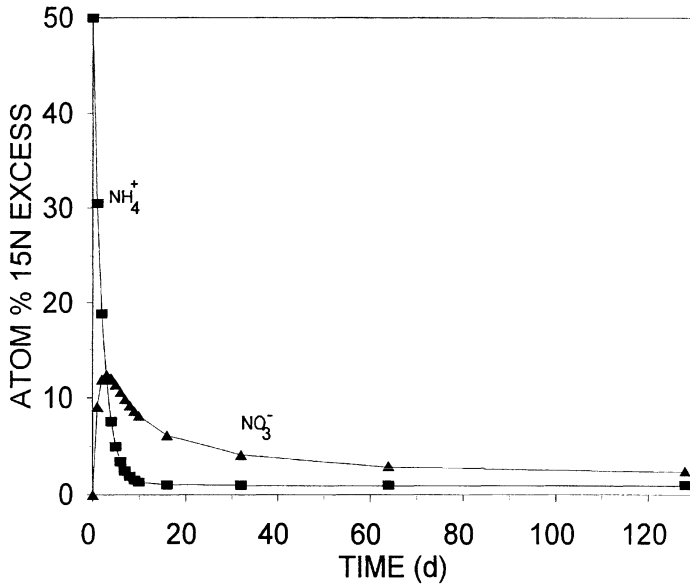
Figure 7.4 The temporal dynamics of the ^{15}N labeling of inorganic N pools based on a simulation using the compartmental model shown in Figure 7.3 At zero time, $^{15}NH_4^+$ was added to the soil so that the NH_4^+ pool was labeled at 50 atom % ^{15}N excess. Figure 7.2 shows the changes in NH_4^+ and NO_3^- concentrations for this simulation.

underestimation became large. Thus for short time periods following the addition of $^{15}NH_4^+$ and for short time intervals, it is probably reasonable to assume zero-order kinetics (pool sizes do not change appreciably) and no remineralization.

The analytical approach of Kirkham and Bartholomew has subsequently been extended to include autotrophic nitrification (Nishio et al., 1985) and NO_3^- immobilization (Wessel and Tietema 1992). As model complexity increases, by adding more pools or transformations, it is necessary to use numerical solutions to the resulting differential equations (Myrold and Tiedje 1986; Nason and Myrold 1991).

3. Using Models to Understand N Cycle Transformations and Their Regulation

Some of the models described previously, such as mass balance and compartmental models, have given insights into the N cycle, for example, by identifying important processes. Greater understanding, however, has come from models that have more fully incorporated other components of the soil ecosystem. By
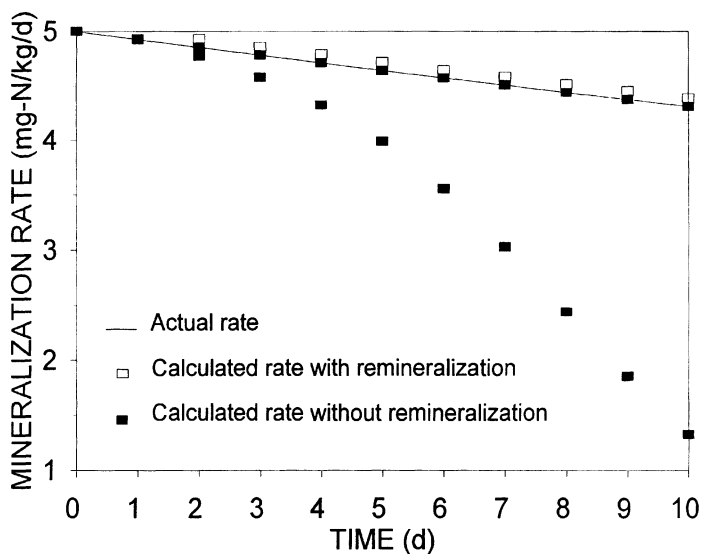
Figure 7.5 Actual and estimated gross N mineralization rates from the simulated data shown in Figures 7.2 and 7.4. Gross N mineralization rates were calculated using the formulas of Kirkham and Bartholomew (1955).

adding other important biological processes, incorporating population dynamics, or including important physical or chemical factors, such as aeration, pH, etc., more realistic models have been created. Three examples are given to illustrate the range of models that have been developed to explore the N cycle and its functioning in soil.

3.1. Regulation of N Mineralization-Immobilization

It has been long recognized that net N mineralization, the balance between N immobilization and mineralization, is tightly coupled with the C cycle. As a rule of thumb, the C:N ratio of a material undergoing decomposition can be used to predict if net N mineralization or immobilization will take place. For many natural substrates a C:N ratio of 25 or less results in net N mineralization, however, more information about the chemical composition of the substrate, the nature of the decomposing population, and soil conditions is required to determine the time course of N mineralization and immobilization. For this purpose a number of models of varying complexity have been developed (Juma and Paul 1981; McGill et al., 1991; Molina et al., 1983; van Veen et al. 1984; Bosatta and Berendse 1984; Grant et al. 1993). As an example, the development and use of the NCSOIL model of Molina et al. (1983) is reviewed.

NCSOIL divides soil organic matter into three pools: biomass, humads, and humus. The biomass and humads fractions are thought to be part of the actively cycling organic matter, whereas the humus is relatively unreactive or passive. Biomass and humads fractions are further divided into two pools each, one labile and one resistant, each of which decomposes independently according to first-order kinetics. The submodels for C and N are quite similar (Fig. 7.6) except that CO_2, the product of C mineralization, is not immobilized into organic matter. They are linked by the C:N ratios of the pools and efficiency factors for C utilization. (NCSOIL is coded in FORTRAN. It is available from the authors on request).

Various modifications of the original NCSOIL model have been used to successfully simulate N and C dynamics over time periods of weeks to years (Molina et al. 1983; Nicolardot et al. 1994; Nicolardot and Molina 1994). Of particular interest are studies that used NCSOIL and experimental data to examine whether microorganisms preferentially assimilated inorganic or organic N. The original
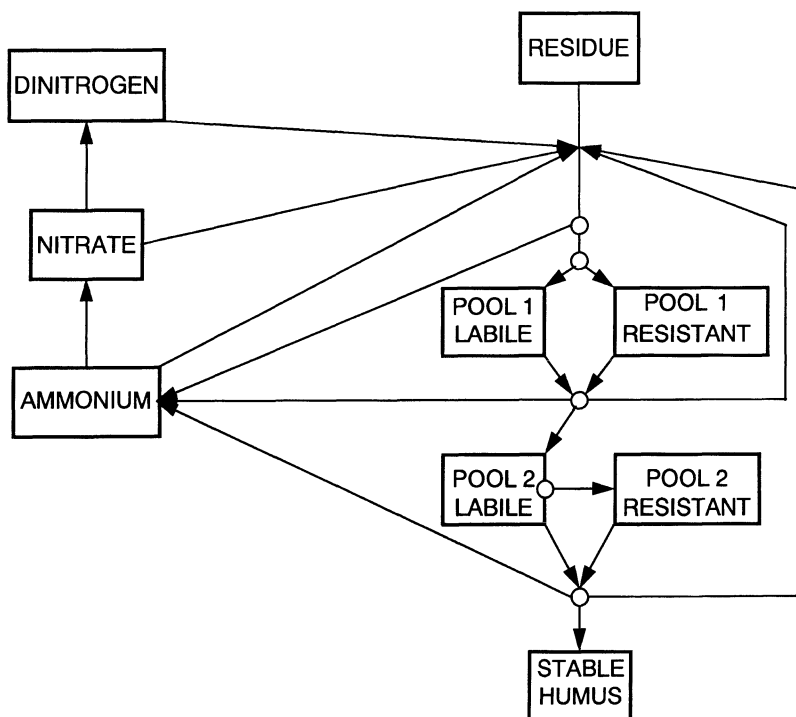
Figure 7.6 Diagram of the N submodel of the NCSOIL model. Redrawn from Molina et al. (1983).

NCSOIL model allowed for organic N to be directly assimilated by the microbial biomass. A modification, known as the mineralization-immobilization turnover model, allowed only inorganic N to be assimilated into microbial biomass. Incubation studies with $^{15}NH_4^+$ generally demonstrated that the mineralization-immobilization turnover model provided the better fit to the data (Fig. 7.7) (Hadas et al. 1987, 1992a, 1992b). When ^{15}N-labeled alanine was used (Hadas et al. 1992b), however, the direct immobilization model more accurately simulated the ^{15}N dynamics of the system. Thus, soil microorganisms apparently will assimilate small organic N molecules when they are available. Under natural soil conditions,
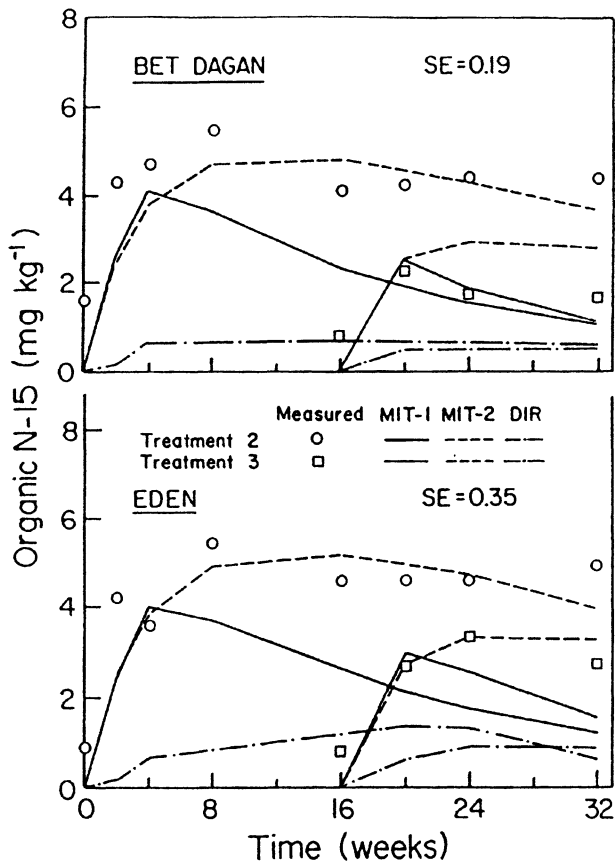
Figure 7.7 Changes in ^{15}N excess of soil organic N for two soils incubated with $^{15}NH_4^+$. Three versions of NCSOIL were fit to the data: MIT-1 (inorganic N assimilation, with NH_4^+ preferred), MIT-2 (inorganic N assimilation of NH_4^+ and NO_3^-, and DIR (direct assimilation of soluble organic N). Data from Hadas et al. 1992b).

however, small soluble organic N compounds may not be present at concentrations sufficient for their assimilation to be a major factor in N immobilization.

3.2. Population Dynamics of Nitrifying Bacteria

Because of the unique autotrophic character of nitrifying bacteria and the importance of nitrification and NO_3^- in determining the availability of N in soils, nitrifiers and their activity have been extensively researched for more than 50 years (Lees and Quastel 1946). Packed soil columns, with either a closed flow (perfusion; Lees and Quastel 1946) or continuous flow (Macura and Kunc, 1965) of nutrients, have often been used to study nitrification in the laboratory. Such column systems with a constant flow rate have been modeled in a variety of ways and experimental data from such systems frequently was fitted well by these models (McLaren, 1970, 1971; Bazin and Saunders 1973; Saunders and Bazin 1973; Ardakani et al. 1974). Subsequently, further refinements have been made that more closely mimic actual soil conditions where diffusion is often the dominant transport process.

Darrah and others (Darrah et al. 1983, 1985, 1986a, 1986b, 1986c, 1986d) incorporated the diffusion of substrates and products into a model of nitrification in soil. Their model quite accurately predicted temporal and spatial changes in NH_4^+ and NO_3^- profiles (Fig. 7.8). Predicted nitrifier biomass was found to closely parallel the spatial and temporal trends of NO_3^-, however, actual measurements of nitrifier populations were not made (Darrah et al. 1985). Subsequent sensitivity analysis of the model suggested that parameters associated with nitrifier growth and activity had the greatest impact on modeled predictions (Darrah et al. 1986d). Maximum specific growth rate and growth yield were the most important parameters, whereas the model was quite insensitive to changes in the substrate affinity constant, probably because it was much smaller than typical soil substrate concentrations. This example demonstrates how models can be used to identify the most important biological parameters affecting an N cycling process.

3.3. Environmental Controls of Denitrification

Denitrification is extremely variable in time and space. This, along with its importance to N cycling, has resulted in significant study and the generation of numerous models. Conceptually, Groffman et al. (1988) formulated a model that suggested that the environmental variables important in regulating denitrification differed as the spatial and temporal scales changed (Fig. 7.9). At the finer scale of centimeters and hours, the supply of oxygen (or creation of anaerobic microsites), C, and NO_3^- are the dominant variables.

Much of the modeling of denitrification at the microsite level has used soil aggregates as their physical basis. Greenwood (1962) modeled soil aggregates
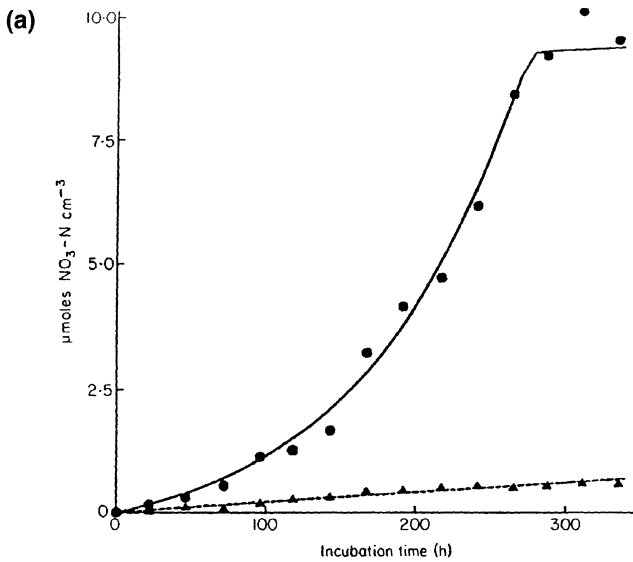
(a) Nitrate formation following NH₄Cl addition in the incubation study; experimental
data +N (●) and −N (▲) and NO₃⁻ predicted +N (———) and −N (---).
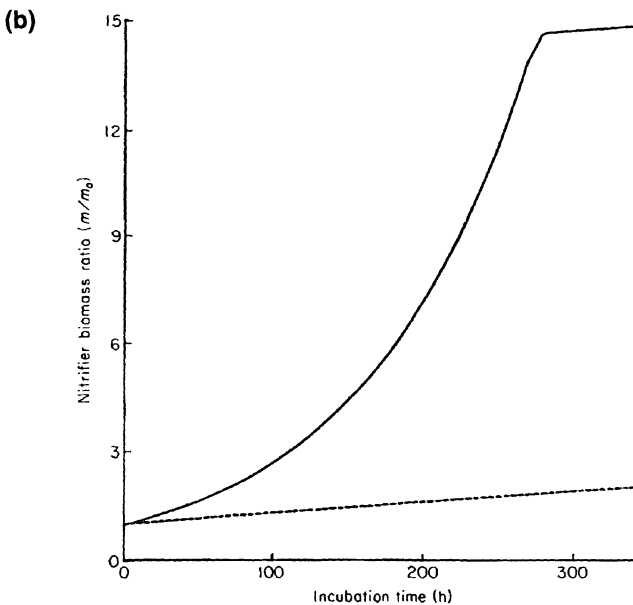
(b) Predicted biomass of nitrifiers in the incubation experiment expressed as a ratio of
the initial biomass: +N (———), −N (---).

Figure 7.8 Experimental and modeled data from Darrah et al. (1985) based on a reaction-diffusion model of nitrification. (a) Temporal changes in measured and modeled NO_3^- concentrations. (b) Modeled nitrifier population dynamics relative to initial nitrifier biomass.

SCALE OF REGULATING FACTORS
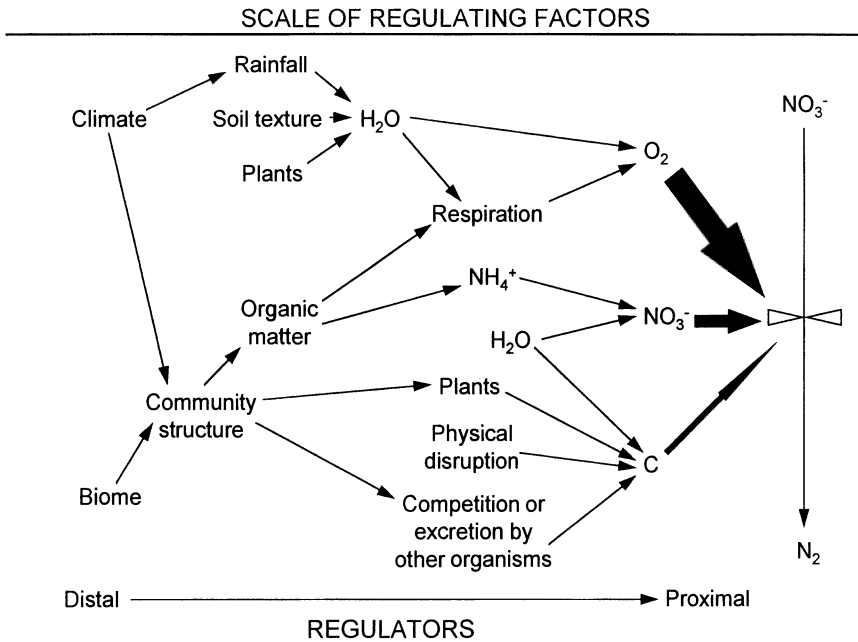
Figure 7.9 The effect of spatial scale on the factors that regulate denitrification in soil ecosystems. Based on the concepts of Groffman et al. (1988).

as spheres with the assumptions of steady-state conditions, uniform distribution of microorganisms, a constant diffusion coefficient, and a constant respiration rate. This enabled him to calculate the anaerobic volume within a given aggregate. Experimental work with nitrification as an indicator of aerobic activity and denitrification as an index of anaerobic activity showed that the ratio of denitrification to nitrification was linearly related to the fraction of the aggregate volume that was anaerobic. This work on individual aggregates has been further confirmed by direct measurement of oxygen profiles and denitrification activity in native soil aggregates by Sexstone et al. (1985; Table 7.1).

   Aggregates come in a variety of sizes, of course, so extending this aggregate model to native soils requires information about the size distribution of aggregates. Smith (1980) developed a model based on a lognormal distribution of aggregates sizes and was able to calculate the anaerobic volume of a soil based on data about aggregate size distribution. Further refinements have been made by including the transport of NO_3^- (Myrold and Tiedje 1985; Arah and Smith 1989). Myrold and Tiedje (1985) suggested that in a well aerated soil, NO_3^- could be limiting to denitrification because only large aggregates contained anaerobic centers. This conclusion was later extended to show that the optimum aggregate size for

Table 7.1 Relationship between the size of the anaerobic zone within water-saturated soil aggregates as measured with an O_2 microelectrode and denitrification rates

Aggregate Radius (mm)	Measured Anaerobic Radius (mm)	Calculated Anaerobic Radius (mm)	Respiration Rate (mm^3-CO_2 cm^{-1} $hour^{-1}$)	Denitrification Rate (ng–N g^{-1} $hour^{-1}$)
9.0	0.3	0.0	1.0	0.03
18.0	14.0	13.6	2.2	1.70

The soil aggregates came from a silt loam soil of an uncultivated native prairie. Data taken from Sexstone et al. (1985).

denitrification was a function of both oxygen and NO_3^- transport (Arah and Smith 1989).

4. Using Models to Make Predictions about N Cycling

Over the past 20 years numerous models have been developed to describe the biological, chemical, and physical fate of N in terrestrial ecosystems and several symposia and books have been devoted to them (Beck and Frissel 1973; Nielsen and MacDonald 1978; Frissel and van Veen 1981; Groot et al. 1991). These ecosystem-scale models normally contain several submodels, which themselves may contain more detailed components (Fig. 7.10). Their size and complexity has led to the proliferation of these models, which vary in the types of processes included and the detail in which these processes are considered. Many of these models have been compared and reviewed by others (Frissel and van Veen 1982; Rao et al. 1982; Addiscott and Wagenet, 1985), with the reviews of Tanji (1981) and de Willigen (de Willegen and Neeteson 1985; de Willegan 1991) being particularly detailed and informative. Instead of dwelling on such comparisons, however, I present one model as an example of this class of ecosystem-level models that have been used for making predictions about N cycling.

LEACHN, the N version of LEACHM (Leaching Estimation And CHemistry Model), is a model developed at and available from Cornell University (Hutson and Wagenet 1991, 1992). LEACHN is a deterministic model that includes separate components for heat flow, water flow, solute transport, evapotranspiration, N transformation, and plant growth. Potential evapotranspiration and canopy and root growth are simulated with daily time steps, whereas other processes use smaller time steps.

Fluxes are modeled mechanistically. Convective heat transport is based on Fourier's Law (which is the equivalent of Fick's diffusion laws, but apply for
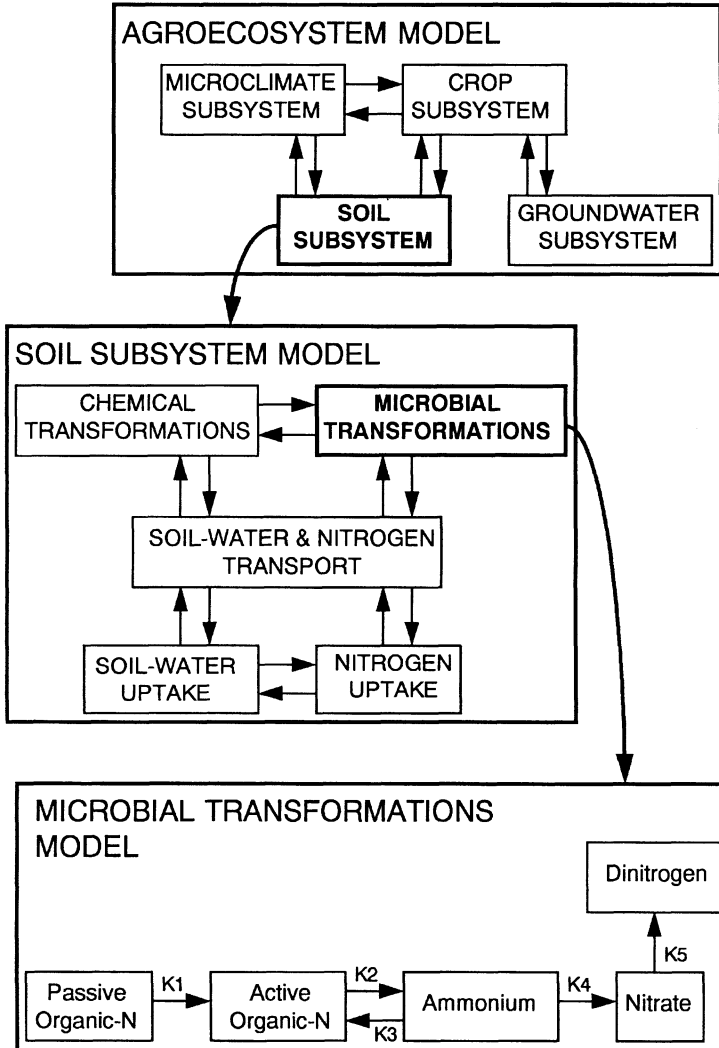
Figure 7.10 An example of hierarchical modeling used in the construction of complex predictive models. Based on the diagram of Rao et al. (1982).

diffusion of heat), unsaturated water flow is modeled by the Richard's equation (which is a second-order partial differential equation for water flow), and solute transport is represented by the convective-dispersion equation. All transport processes are modeled using finite difference approximations.

Plant growth is based on empirical equations that simulate a logistic accumulation of biomass. As such, there is no feedback between the plant growth and important soil properties, such as water and N.

The N transformation subroutine in LEACHN is based on the concepts and equations of Johnsson et al. (1987; Fig. 7.11). The organic N pool is divided into three components: litter, feces, and humus, the latter pool being relatively recalcitrant. Mineralization from each organic pool is modeled as a first-order kinetic process. Nitrification rates are scaled proportionally to the NO_3^-/NH_4^+ ratio. Michaelis-Menten kinetics, with NO_3^- as the single substrate, are used to model denitrification. All these biological processes are adjusted for temperature using a Q_{10} function, and water content, using a scaling factor between 0 and 1 for the optimum water content for each process. The water content adjustment empirically accounts for aeration effects. In addition, ammonia volatilization is a first-order function of NH_4^+ concentration.

An almost infinite variety of simulations can be run with models such as LEACHN because of the large number of input variables that can be altered. For example, one could examine the effects of cover crops versus winter fallow, different fertilization regimes or irrigation schedules, and changes in microbial
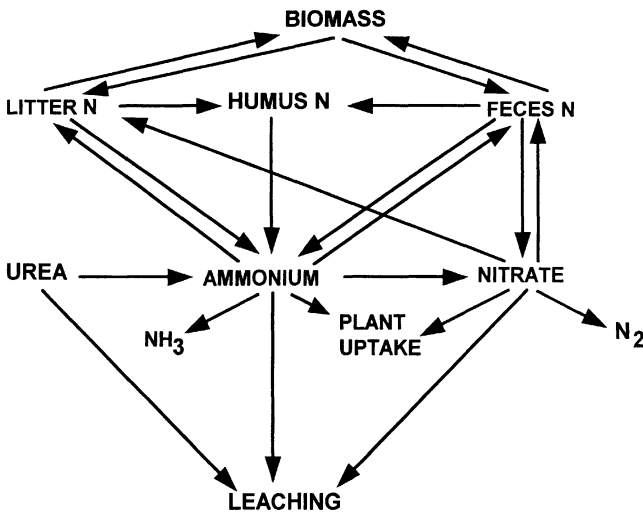
Figure 7.11 The N submodel of LEACHM (Hutson and Wagenet 1991; Johnsson et al. 1987).

transformation rate constants. Hutson and Wagenet (1991) briefly examined the latter situation and found that soil NO_3^- contents were more sensitive to changes in N mineralization and denitrification rate constants than to variations in ammonia volatilization rates.

5. Summary

Whether implicitly, in our conceptual models of how N is transformed in soil systems, or explicitly, in mathematical models of varying complexity, models are a useful tool for studying N transformations in soil. They enable us to calculate gross and net rates of N cycle processes, help to identify the more important N transformations in a given soil system, provide a means to study regulatory mechanisms, and may even have meaningful predictive capabilities. This brief overview has touched on only a few examples of such uses and many additional modeling approaches and applications can be found by those interested in quantitatively studying the soil N cycle.

References

Addiscott, T. M., and R. J. Wagenet. 1985. Concepts of solute leaching in soils: A review of modeling approaches. *J. Soil Sci.* 36:411–424.

Arah, J. R. M., and K. A. Smith. 1989. Steady-state denitrification in aggregated soils: A mathematical model. *J. Soil Sci.* 40:139–149.

Ardakani, M. S., J. T. Rehbock, and A. D. McLaren. 1974. Oxidation of ammonium to nitrate in a soil column. *Soil Sci. Soc. Am. Proc.* 38:96–99.

Bazin, M. J., and P. T. Saunders. 1973. Dynamics of nitrification in a continuous flow system. *Soil Biol. Biochem.* 5:531–543.

Beck, J., and M. J. Frissel. 1973. *Simulation of Nitrogen Behaviour in Soils.* Pudoc, Wageningen, The Netherlands.

Bonde, T. A., and T. Lindberg. 1988. Nitrogen mineralization kinetics in soil during long-term aerobic laboratory incubations: A case study. *J. Environ. Qual.* 17:414–417.

Bormann, B. T., F. H. Bormann, W. B. Bowden, R. S. Pierce, S. P. Hamburg, D. Wang, M. C. Snyder, C. Y. Li, and R. C. Ingersoll. 1993. Rapid N_2 fixation in pines, alder, and locust: Evidence from the sandbox ecosystem study. *Ecology* 74:583–598.

Bosatta, E., and F. Berendse. 1984. Energy or nutrient regulation of decomposition: implications for the mineralization-immobilization response to perturbations. *Soil Biol. Biochem.* 16:63–67.

Broadbent, F. E. 1986. Empirical modeling of soil nitrogen mineralization. *Soil Sci.* 141:208–213.

Darrah, P. R., P. H. Nye, and R. E. White. 1983. Diffusion of NH_4^+ and NO_3^- mineralized from organic N in soil. *J. Soil Sci.* 34:693–707.

Darrah, P. R., P. H. Nye, and R. E. White. 1986a. Simultaneous nitrification and diffusion in soil. II. The effects of the addition of ammonium sulphate. *J. Soil Sci.* 37:53–58.

Darrah, P. R., P. H. Nye, and R. E. White. 1986b. Simultaneous nitrification and diffusion in soil. V. The effects of pH change, following the addition of ammonium sulphate, on the activity of nitrifiers. *J. Soil Sci.* 37:479–484.

Darrah, P. R., R. E. White, and P. H. Nye. 1985. Simultaneous nitrification and diffusion in soil. I. The effects of the addition of a low level of ammonium chloride. *J. Soil Sci.* 36:281–292.

Darrah, P. R., R. E. White, and P. H. Nye. 1986c. Simultaneous nitrification and diffusion in soil. II. The effects at levels of ammonium chloride which inhibit nitrification. *J. Soil Sci.* 37:41–52.

Darrah, P. R., R. E. White, and P. H. Nye. 1986d. Simultaneous nitrification and diffusion in soil. IV. Simplification and sensitivity analysis of the simulation model for ammonium chloride. *J. Soil Sci.* 37:469–478.

Deans, J. R., J. A. E. Molina, and C. E. Clapp. 1986. Models for predicting potentially mineralizable nitrogen and decomposition rate constants. *Soil Sci. Soc. Am. J.* 59:323–326.

de Willigen, P. 1991. Nitrogen turnover in the soil-crop system; comparison of fourteen simulation models. *Fert. Res.* 27:141–149.

de Willigen, P., and J. J. Neeteson. 1985. Comparisons of six simulation models for the nitrogen cycle in the soil. *Fert. Res.* 8:157–171.

Ellert, B. H., and J. R. Bettany. 1988. Comparison of kinetic models for describing net sulfur and nitrogen mineralization. *Soil Sci. Soc. Am. J.* 52:1692–1702.

Frissel, M. J., and J. A. van Veen (eds.). 1981. *Simulation of Nitrogen Behaviour of Soil-Plant Systems.* Pudoc, Wageningen, The Netherlands.

Frissel, M. J., and J. A. van Veen. 1982. A review of models for investigating the behaviour of nitrogen in soil. *Phil. Trans. R. Soc. Lond.,* B296:342–349.

Grant, R. F., N. G. Juma, and W. B. McGill. 1993. Simulation of carbon and nitrogen transformations in soil: Mineralization. *Soil Biol. Biochem.* 25:1317–1329.

Greenwood, D. J. 1962. Nitrification and nitrate dissimilation in soil. II. Effect of oxygen concentration. *Plant Soil* 17:378–391.

Groffman, P. M., J. M. Tiedje, G. P. Robertson, and S. Christensen. 1988. Denitrification at different temporal and geographical scales: proximal and distal controls. In J. R. Wilson (ed.), *Advances in Nitrogen Cycling in Agricultural Ecosystems,* pp. 174–192. CAB International, Wallingford, UK.

Groot, J. J. R., P. De Willigen, and E. L. J. Verberne. 1991. *Nitrogen Turnover in the Soil-Crop System.* Kluwer Academic Publishers, Dordrecht, The Netherlands.

Hadas, A., J. A. E. Molina, S. Feigenbaum, and C. E. Clapp. 1987. Simulation of nitrogen-15 immobilization by the model NCSOIL. *Soil Sci. Soc. Am. J.* 51:102–106.

Hadas, A., J. A. E. Molina, S. Feigenbaum, and C. E. Clapp. 1992a. Factors affecting nitrogen immobilization in soil as estimated by simulation models. *Soil Sci. Soc. Am. J.* 56:1481–1486.

Hadas, A., M. Sofer, J. A. E. Molina, P. Barak, and C. E. Clapp. 1992b. Assimilation of nitrogen by soil microbial population: NH_4^+ versus organic N. *Soil Biol. Biochem.* 24:137–143.

Hauck, R. D., and K. K. Tanji. 1982. Soil nitrogen budgets. In F. J. Stevenson (ed.), *Nitrogen in Agricultural Soils,* pp. 891–925. American Society of Agronomy, Madison, WI.

Hibbs, D. E., and K. Cromack, Jr. 1990. Actinorhizal plants in Pacific Northwest forests. In C. R. Schwintzer and J. D. Tjepkema (eds.), *The Biology of Frankia and Actinorhizal Plants,* pp. 343–363. Academic Press, San Diego, CA.

Hutson, J. L., and R. J. Wagenet. 1991. Simulating nitrogen dynamics in soils using a deterministic model. *Soil Use Manage.* 7:74–78.

Hutson, J. L., and R. J. Wagenet. 1992. LEACHM: *Leaching Estimation and Chemistry Model—A Process-Based Model of Water and Solute Movement, Transformations, Plant Uptake and Chemical Reactions in the Unsaturated Zone, Version 3.* Cornell University, Ithaca, NY.

Johnsson, H., L. Bergström, P.-E. Jansson, and K. Paustian. 1987. Simulated nitrogen dynamics and losses in a layered agricultural soil. *Agric. Ecosys. Environ.* 18:333–356.

Juma, N. G., and E. A. Paul. 1981. Use of tracers and computer simulation techniques to assess mineralization and immobilization of soil nitrogen. In M. J. Frissel and J. A. van Veen (eds.), *Simulation of Nitrogen Behaviour of Soil-Plant Systems,* pp. 145–154. Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands.

Juma, N. G., E. A. Paul, and B. Mary. 1984. Kinetic analysis of net nitrogen mineralization in soil. *Soil Sci. Soc. Am. J.* 48:753–757.

Kirkham, D., and W. V. Bartholomew. 1954. Equations for following nutrient transformations in soil, utilizing tracer data. *Soil Sci. Soc. Am. Proc.* 18:33–34.

Kirkham, D., and W. V. Bartholomew. 1955. Equations for following nutrient transformations in soil, utilizing tracer data, II. *Soil Sci. Soc. Am. Proc.* 19:189–192.

Lees, H., and J. H. Quastel. 1946. The site of soil nitrification. *Biochem. J.* 40:815–823.

Legg, J. O., and J. J. Meisinger. 1982. Soil nitrogen budgets. In F. J. Stevenson (ed.), *Nitrogen in Agricultural Soils,* pp. 503–566. American Society of Agronomy, Madison, WI.

Macura, J., and F. Kunc. 1965. Continuous flow method in microbiology: V. Nitrification. *Folia Microbiol.* 10:125–135.

McGill, W. B., H. W. Hunt, R. G. Woodmansee, and J. O. Reuss. 1991. Phoenix, a model of the dynamics of carbon and nitrogen in grasslands soils. In F. E. Clark and T. Rosswall (eds.), *Terrestrial Nitrogen Cycles. Processes, Ecosystem Strategies and Management Inputs, Ecol. Bull.* Swedish Natural Science Research Council; Stockholm Ecological bulletins (*Stockholm*) 33:49–115.

McLaren, A. D. 1970. Temporal and vectorial reactions of nitrogen in soil: a review. *Can. J. Soil Sci.* 50:97–109.

McLaren, A. D. 1971. Kinetics of nitrification in soil: growth of the nitrifiers. *Soil Sci. Soc. Am. Proc.* 35:91–95.

Molina, J. A. E., C. E. Clapp, M. J. Shaffer, F. W. Chichester, and W. E. Larson. 1983. NCSOIL, a model of nitrogen and carbon transformations in soil: Description, calibration, and behavior. *Soil Sci. Soc. Am. J.* 47:85–91.

Myrold, D. D., and J. M. Tiedje. 1985. Diffusional constraints on denitrification in soil. *Soil Sci. Soc. Am. J.* 49:651–657.

Myrold, D. D., and J. M. Tiedje. 1986. Simultaneous estimation of several nitrogen cycle rates using ^{15}N: Theory and application. *Soil Biol. Biochem.* 18:559–568.

Nason, G. E., and D. D. Myrold. 1991. ^{15}N in soil research: appropriate application of rate estimation procedures. *Agric. Ecosys. Environ.* 34:427–441.

Nicolardot, B., and J. A. E. Molina. 1994. C and N fluxes between pools of soil organic matter: Model calibration with long-term field experimental data. *Soil Biol. Biochem.* 26:245–251.

Nicolardot, B., J. A. E. Molina, and M. R. Allard. 1994. C and N fluxes between pools of soil organic matter: model calibration with long-term incubation data. *Soil Biol. Biochem.* 26:235–243.

Nielsen, D. R., and J. G. MacDonald (eds.). 1978. *Nitrogen in the Environment, Vols. 1 & 2.* Academic Press, New York.

Nishio, T., T. Kanamori, and T. Fujimoto. 1985. Nitrogen transformations in an aerobic soil as determined by a ^{15}NH$_4^+$ dilution technique. *Soil Biol. Biochem.* 17:149–154.

Rao, P. S. C., R. E. Jessup, and A. G. Hornsby. 1982. Simulation of nitrogen in agroecosystems: Criteria for model selection and use. *Plant Soil* 67:35–43.

Saunders, P. T., and M. J. Bazin. 1973. Nonsteady state studies of nitrification in soil: Theoretical considerations. *Soil Biol. Biochem.* 5:545–557.

Sexstone, A. J., N. P. Revsbech, T. B. Parkin, and J. M. Tiedje. 1985. Direct measurement of oxygen profiles and denitrification rates in soil aggregates. *Soil Sci. Soc. Am. J.* 49:645–651.

Smith, K. A. 1980. A model of the extent of anaerobic zones in aggregated soils and its potential application to estimates of denitrification. *J. Soil Sci.* 31:263–277.

Stanford, G., and S. J. Smith. 1972. Nitrogen mineralization potentials of soils. *Soil Sci. Soc. Am. Proc.* 36:465–472.

Tanji, K. K. 1981. Modeling of the soil nitrogen cycle. In F. J. Stevenson (ed.), *Nitrogen in Agricultural Soils,* pp. 721–772. American Society of Agronomy, Madison, WI.

van Veen, J. A., J. N. Ladd, and M. J. Frissel. 1984. Modelling C & N turnover through the microbial biomass in soil. *Plant Soil* 76:257–274.

Wessel, W. W., and A. Tietema. 1992. Calculating gross N transformation rates of ^{15}N pool dilution experiments with acid forest soil litter: Analytical and numerical approaches. *Soil Biol. Biochem.* 24:931–942.

White, C. S. and Y. D. Marinakis. 1991. A flexible model for quantitative comparisons of nitrogen mineralization patterns. *Biol. Fertil. Soils* 11:239–244.

8

Construction and Analysis of Static, Structured Models of Nitrogen Cycling in Coastal Ecosystems

Robert R. Christian, Mariachiara Naldi, and Pierluigi Viaroli

1. Introduction

Microbial processes often dominate biogeochemical cycling within ecosystems. Therefore, microbial ecology interfaces with ecosystem ecology as appropriately as plant or animal ecology does. However, it is difficult for many microbiologists to interpret ecosystem-level phenomena. Often field measurements are made separately on individual components of an ecosystem. Thus one measures a single process, such as phytoplankton primary productivity, or a single assemblage of organisms, such as benthic (i.e., referring to sediments) microalgae. The interactions of components are inferred from observations of patterns of several variables over time or space. Or one conducts experiments using a limited number of variables where inferences involve only direct interactions of a few components. It is more difficult to examine indirect influences of one component on another or to simultaneously evaluate numerous interactions. We have found that constructing and analyzing simple mathematical models can be useful in greater synthesis of microbiological and ecological information (Christian et al. 1986; Christian and Wetzel 1991). Modeling provides a mechanism by which one can extend inferences to include numerous components and both direct and indirect interactions simultaneously.

In ecosystem ecology, mathematical models often are "structured," containing a number of compartments (also called state variables). The first-year biology student is aware of the general, structured model of the food chain with compartments of primary producers, herbivores, etc. Compartments generally categorize ecological units (e.g., species populations, communities, trophic levels, nutrient pools), which represent matter, energy, or individuals. Compartments may be

connected to each other and to the environment of the system. These connections, or interactions, represent flows of matter, energy, individuals, or information.

There may be several objectives for employing models in ecological research. These objectives are met at different steps in the process: before, during, and after data collection. In many instances for biologists (as opposed to mathematicians), the most important objective of modeling is to aid field or laboratory research or to give insight to new directions for research. Thus, it is valuable to construct models early in the research program. These models may be crude because of the lack of information. But they can be improved as research proceeds. The recognition of ways for improvement may direct certain avenues of research. In this way, one can view modeling as a "synergistic" partner to normal ecological research (Christian and Wetzel 1991).

The steps of ecological modeling begin with establishing a purpose. What specific question or questions concerning the system can be addressed by the construction and analysis of a model? For example, one may ask whether the import of nutrients or their recycling most controls primary productivity. Next, it is important to define the system precisely: the compartments, imports, exports, interactions, and the boundary conditions in both space and time. Again, for example, if one wishes to define the temporal boundaries of a system as a "season," one must decide if the traditional seasons (between soltice and equinox) are ecologically relevant or if some other interval is more appropriate. This process of definition can be rewarding in itself, as it forces a formality of thought about the ecological system of concern. There should be interplay between what has been, is being, or will be measured and the characteristics of the model. To be useful and efficient, the definitions should reflect the research program. In this regard, a model at the species level may be of little value if information is obtained only at the community level.

Once each compartment, interaction, and connection with the environment is defined qualitatively, quantification is the next step. If simulation through time is desired, one must also define the interactions through specific equations. These equations can take a variety of forms, linear or nonlinear, based on empirical relationships or perceived mechanisms. Simulation models are discussed throughout this book in other chapters. Overviews of ecosystem-level simulation models can be found in such works as those by Shugart and O'Niell (1979) and Odum (1983).

In the current effort, multicompartmental models are static; that is, they are time-averaged and represent a "snapshot" of the ecosystem. Analyses of these are directly on the structure of the models (i.e., how are compartments connected) and magnitudes of the averaged fluxes. One benefit of construction and analysis of static models is in the avoidance of the need for the assumptions necessary in developing equations for simulation. The detriment is the preclusion of the predictive powers of simulation (Mann et al. 1989). For the static models, values

of standing stocks and fluxes are derived for one sampling or may be averaged in time and space. The sources of these values may be varied and include (1) direct measures, (2) interpolations and extrapolations from related studies, (3) general literature and ecological principles, and (4) results of mass balance calculations of other compartments. Thus the degree of reliability can vary among values, depending on their sources. Some numbers may be well estimated, whereas others may be little more than guesses. This exercise helps determine what is being measured in the context of total system's structure and function. The lack of a particular value may direct a change in research plan. Also, the balancing of inputs and outputs of compartments may help define the limits of values for which there is little information. For example, knowing the productivity of an animal species population sets a minimum estimate on its rate of ingestion. General inferences from similar species concerning respiration, excretion, and egestion improves the estimate. In fact, merely designing a model can provide much useful information without formal analysis.

The formal analyses of models described here come under the general heading of network analysis. Network analysis is actually several descrete analyses that employ a variety of subdisciplines of mathematics and associated algorithms. The analyses are designed to evaluate the qualitative (the topology of the connections among compartments or "map of the system") and quantitative (related to the amounts of each flux) structure of the system. Overviews of ecological network analysis and its applications are found in Patten (1985), Ulanowicz (1986), and Wulff et al. (1989). Issues that can be addressed include (1) the relative importance of imports to recycled material in the system, (2) the role of each compartment and interaction in recycling, (3) the importance of each compartment to others, (4) the trophic structure of the community, (5) the distribution of positive feedback loops in the system and their importance to flow, and (6) global attributes of growth and development of the system. These various system characteristics can be used for comparisons within the system, among times, and among systems. Hence, network analyses provide powerful tools for the interpretation of both nutrient cycling and energy flow.

In this chapter, we illustrate our use of static multicompartmental models and network analysis in the understanding of nitrogen cycling within coastal ecosystems. Compartments are states of nitrogen within these systems (e.g., phytoplankton N). Some commonly measured and designated forms include dissolved N (DN), dissolved inorganic N (DIN), dissolved organic N (DON), and particulate organic N (PON). Connections represent transformations of nitrogen from one state to another (e.g., uptake of DIN by phytoplankton) or transfers from one location to another (e.g., import of DIN into a coastal lagoon). We have analyzed models from the Neuse River Estuary in North Carolina (Christian et al. 1992) and from rice fields and a coastal lagoon in the Ebro River delta of Spain (Forès and Christian 1993; Forès et al. 1994). Comparative ecosystem

ecology has been attempted through comparison of analyses of these three systems with each other, with another lagoon in the Ebro River delta and with the Sacca di Goro in Italy (Christian et al. 1993, 1996). We concentrate our discussion for the present chapter on the latter system. We have refined our model through three generations of improved information acquisition and incorporation. We describe this evolution to address the following issues: (1) establishment of purpose, (2) iterative nature of the process and reliability of information, (3) sensitivity of analysis results to model changes, and (4) benefits and limitations of such models and their analysis.

2. Methods

2.1. Study Site

The Sacca di Goro is a coastal lagoon within the Po River delta, Italy at approximately 44°47′–44°50′ N and 12°15′–12°20′ E (Fig. 8.1). It is 26 km²
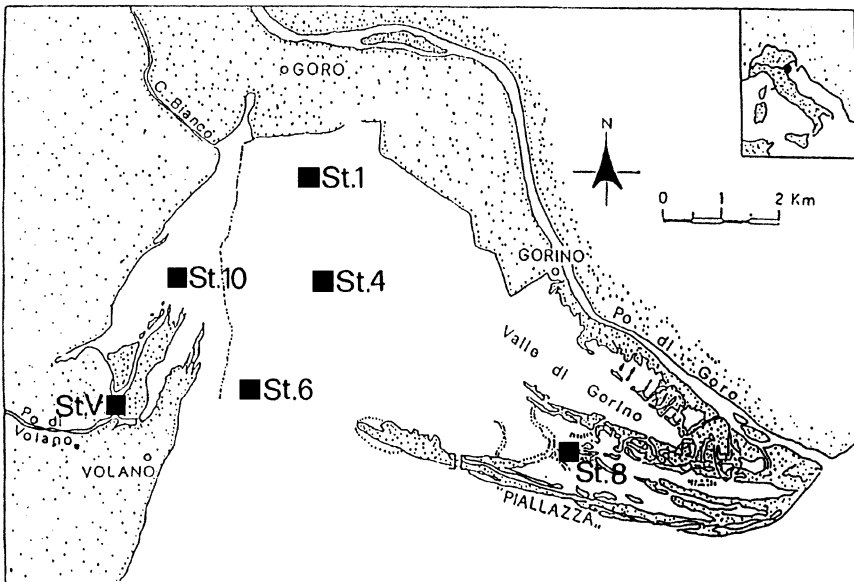
Figure 8.1 Map of the Sacca di Goro, Italy. The coastal lagoon receives fresh water from the Po di Goro, the Po di Volano, and local pumping stations and connects with the northern Adriatic Sea (see insert). Field research involves sampling stations shown. Models of nitrogen cycling are from the Valle di Gorino in the vicinity of station 8.

with an average depth of 1.5 m. The bottom is flat, and sediments are composed of typical alluvial mud with a high clay and silt content, whereas sandy mud prevails in the eastern region. The eastern region, the Valle di Gorino, accounts for half the area with a shallow depth of 0.8 m. Fresh water enters the lagoon from the Po di Goro, the southernmost distributary of the Po River; Po di Volano, a polluted canal that drains the Provinces of Ferrara and Modena; and from local pumping stations releasing highly polluted water directly into the lagoon. Saline water enters from the northern Adriatic Sea through a 2-km-wide mouth and at times through a man-made ditch into the Valle di Gorino. The lagoon is the site of numerous enterprises, including fin and shell fisheries, tourism, a bird sanctuary and mariculture. However, the incoming waters have high concentrations of nutrients; macroalgae, especially *Ulva rigida* and *Gracilaria* spp., grow in large abundances; and in summer there are periods of dystrophy. Dystrophy is the trophic state in which the primary supply of organic matter and energy is from preformed material and not active primary production. Heterotrophy dominates. In the Sacca di Goro this results from the decomposition of macroalgae and is associated with extended anoxia. The Valle di Gorino is the site of greatest macroalgal biomass and most severe dystrophy (Viaroli et al. 1992). Station 8 is taken to be representative of the Valle di Gorino (Fig. 8.1).

The Sacca di Goro has been the site of continuous study since 1987 by researchers at the Universities of Ferrara and Parma. This research stems from recognition of the high degree of cultural eutrophication of the lagoon and associated dystrophy. Much of the initial work was on either community structure (e.g., Pugnetti 1990; Ferrari et al. 1993) or water quality (e.g., Colombo 1989). More recent studies have focused on the processes of nutrient cycling and their relationship to the growth and decomposition of macroalgae (e.g., Viaroli 1992; Viaroli et al., 1992; 1993 a,b,c). The results of these studies provide a body of knowledge sufficient enough to allow us to construct reasonable models of the nitrogen cycle within the Valle di Gorino to be used in network analysis.

2.2. Nitrogen Cycle

Nitrogen is the nutrient element that has been considered the most likely to be limiting in coastal marine systems (Ryther and Dunstan 1971; Day et al. 1989). The cycle is complex, involving numerous oxidations and reductions of the element. These oxidations and reductions are involved in both the assimilation of organic matter and catabolic metabolisms. Some processes occur only in aerobic environments, some only in anaerobic environments, and some in both. Details of the cycle are beyond the scope of this chapter. We assume the reader to be versed in the general processes of the nitrogen cycle. If background information is required, refer to Day et al. (1989) for descriptions of estuarine nitrogen cycling. In fact, the models presented here do not explicitly include all of the

transformations, having been constructed in the context of available information from the field studies in the Sacca di Goro. Christian et al. (1992) developed two models of nitrogen cycling within the Neuse River Estuary which included a larger suite of transformations. See this work for a more detailed network analysis of nitrogen cycling.

2.3. Basis for Model Construction and Analysis

Our modeling approach involves designating individual compartments that represent standing stocks of nitrogen and flows of nitrogen between compartments or between a compartment and outside the system. These latter flows are imports and two forms of outputs: exports and dissipations resulting from respiration. This separation of outputs has been useful in the analysis of energy flow and carbon cycling models as done by Ulanowicz (1986). The energetic ground state for carbon is carbon dioxide, resulting from respiration. In published models of carbon, as a surrogate for energy, the carbon dioxide pool is considered outside the system, connected to the compartments via autotrophy and respiration (e.g., Baird and Ulanowicz 1989; Baird et al. 1991). In nitrogen cycles we may consider molecular nitrogen as having a similar position outside the system, although it is not the energetic ground state. Its connections with the system are through nitrogen fixation and denitrification (a dissipation in the sense of Ulanowicz [1986]). No distinction has been made between molecular nitrogen and other gaseous endproducts of denitrification (e.g., nitrous oxide). All flow rates (i.e., fluxes) have positive values with the units mmole N m^{-2} day^{-1}. Standing stocks are as mmoles N m^{-2}.

We have constructed three generations of models for the Valle di Gorino. Our initial model (called VDG1A) was developed as a form of introduction to the research results from the Sacca di Goro for Christian during a stay in Italy in 1991. The most complete data set available at that time was for April 1989, which served as a rough temporal setting for this first attempt. Based in part on the results of the first model's construction and analysis, a second generation was developed using a better data set, with much, but not all, field information collected during 1991. Two separate time periods were modeled: (1) the productive period of early summer (VDG2B) and (2) the dystrophic period in July (VDG2C). These were used in initial comparisons of cycles among coastal ecosystems (Christian et al. 1993). A third generation of models was based on a more complete analysis of the data from the 1991 studies of the Sacca di Goro. Three periods were modeled: (1) spring 1991 (VDG3A), (2) early summer 1991 (VDG3B), and (3) the dystrophic period of July (VDG3C). This generation is being used for a more complete and specific evaluation of the effects of macroalgal growth and dystrophy on nitrogen cycling. The reliability of information has increased greatly during the evolution of the models. This reliability is in the

form of greater specificity in both the time and location of data collection to the desired boundary conditions of the models. Details of data sources and model construction are given in the next section.

We used ecological network analyses to interpret the nature of nitrogen cycling in our models and by inference in the field. Calculations were made on an IBM-PC compatible computer using the MS-DOS software package "NETWRK4: A Package of Computer Algorithms to Analyze Ecological Flow Networks" by R. E. Ulanowicz (University of Maryland, Center for Environmental and Estuarine Studies, Chesapeake Biological Laboratory, Solomons, MD 20688-0038, USA, 1987). The package contains several programs in FORTRAN to aid in network analysis. The two used for our analyses were DATAIN used to convert data into SCOR input format, usable for the network analysis in the program NETWRK. (Details of the software's use are given in the documentation provided by Ulanowicz 1987.) NETWRK computes a variety of analyses under the headings of Structure Analysis, Lindeman Trophic Aggregations, Biogeochemical Cycle Analysis, Information Indices and Connectance Indices. Of these we have found the Structure Analysis and Biogeochemical Cycle Analysis to be the most informative for our nitrogen cycling models. The other analyses are more appropriate when a developed food web is represented.

The mathematical derivations of the various algorithms within the network analysis have been presented several times by others. Patten (1985), Ulanowicz (1986), and Kay et al. (1989) give perhaps the most comprehensive discussions of most to all of the algorithms. Individual analyses have been described by their original authors, and we cite these references where appropriate. Rather than reiterate the mathematics, we provide what we hope are simple explanations of each algorithm designed to interest and inform microbiologists. Further, the jargon associated with network analysis can be formidable. Where we could, we have tried to use terms that we consider of more common usage to people with minimum modeling experience. In many cases, however, new vocabulary is necessary.

Structure analysis involves a series of matrix algebra manipulations resulting in matrices and vectors designed to quantify both direct and indirect effects and fates of imports and interactions (Hannon 1973; Finn 1976; Patten et al. 1976). The NETWRK4 software organizes separate vectors of fluxes involving imports, exports, dissipations and a matrix of interactive flows (exchange matrix). Each flux or coefficient within a vector or matrix is represented as f_{ij}, where i refers to a donor, x_i; j refers to a recipient, x_j; and O refers to outside the system. "Throughput" of each compartment is used for the analyses and is equal to the sum of all of each compartment's outputs or inputs. When a compartment is in steady state, these two sums are equivalent. When a compartment is not in steady state, throughput is the larger of the two sums. Finn (1976, 1980) relates this

latter definition to the sequestering of materials from or release to the system with growth and decline, respectively.

   Two matrices that examine the interaction of compartments are calculated first in the Structure Analysis section. A "total contribution matrix" is used to evaluate the fraction of a compartment's throughput that contributes to another compartment's throughput. A "total dependency matrix" is calculated to evaluate the fraction of a compartment's throughput that resided at some point in another compartment. These are derived from the exchange matrix of each interaction or exchange among compartments. In the total contribution matrix, each interaction f_{ij} is divided by the throughput of the donating compartment x_i; then the resulting matrix $\mathbf{F}$ is subtracted from the identity matrix $\mathbf{I}$; and the inverse of the difference is computed $(\mathbf{I} - \mathbf{F})^{-1}$. For the total dependency matrix, each interaction f_{ij} is divided by the throughput of the recipient compartment x_j; then the resulting matrix $\mathbf{G}$ is subtracted from the identity matrix $\mathbf{I}$; and the inverse of the difference is computed $(\mathbf{I} - \mathbf{G})^{-1}$. Coefficients along diagonals are further manipulated by subtracting each from 1 and dividing the difference by the original coefficient to evaluate recycling (Ulanowicz 1986). The coefficients for either matrix represent connections between compartments that may be either direct or indirect. That is, no direct connection is necessary for positive dependence or a contribution to occur. For example, nitrogen released as ammonium during the decomposition of detritus may be taken up by phytoplankton. The connection between phytoplankton and detritus is indirect and mediated by an ammonium pool. These matrices compute relative dependencies and contributions of both the direct connections, such as that between phytoplankton and ammonium and the indirect connections, such as between detritus and phytoplankton.

   Another analysis, called input environ analysis, computes the contributions of each import to other flows within and from the system through a series of vectors and matrices. Each import is considered separately. In our studies, the coefficients in each resultant vector and matrix represent the number of times or the probability that an atom of imported N passes along a flow path between entering and leaving the system. This analysis begins with the matrix $\mathbf{F}$ as described above. It is transposed $(\mathbf{F}^T)$ and subtracted from the identity matrix, and the inverse is taken of the difference $(\mathbf{I} - \mathbf{F}^T)^{-1}$. Each column represents a vector $(\boldsymbol{T}_i)$ of throughputs of compartments for unit inputs to the compartment represented by the column (i.e., column i refers to unit input to x_i). To construct an exchange matrix, matrix $\mathbf{F}$ is muliplied in separate subsequent analyses by each vector $\boldsymbol{T}_i$ for which there is an input f_{0i}. The unit fate of the material to export and respiration are calculated by mass balance after accounting for internal flows. Thus through structure analysis, the influence of both flows entering the system and those within the system (i.e., loading versus recycling of nutrients) can be considered.

   Biogeochemical cycle analysis employs graph theory to evaluate the character-

istics of cycles or positive feedback loops within the system (Ulanowicz 1983, 1986). The terminology for graph theory is somewhat new to most biologists. For our purposes an arc is a flux or flow between two compartments. A cycle is a series of arcs that in combination pass material from a compartment, through one or more unique compartments, and returns material to the original one. This is in fact a positive feedback loop. The arc within a cycle that has the smallest flux is the weak arc. And a group of cycles sharing the same weak arc is a nexus. One can infer that the weak arc is potentially the controlling flow within a cycle and that cycles in a nexus have a common control. For example, a cycle may exist as the following: nitrogen is passed from an ammonium pool to phytoplankton and then with death the nitrogen is passed to detritus. Mineralization of the detritus returns nitrogen to the ammonium pool. If the smallest flux is the assimilation of ammonium by phytoplankton, this is the weak arc. If there is also a cycle from ammonium to phytoplankton to detritus to sediments and back to ammonium, and if the smallest flux of that cycle is also the assimilation by phytoplankton; the two cycles form a nexus. The cycling of nitrogen through both cycles may be controlled by their common weak arc. If assimilation is increased, then one might infer that both cycles would have a greater flow.

NETWRK4 possesses an algorithm that searches for and identifies all cycles. The program then identifies the smallest weak arc and its cycle or cycles (nexus). If only one cycle is involved, the amount of flow associated with this weak arc is subtracted from all arcs in the cycle; thus breaking the cycle. If more than one cycle share the weak arc, the amount in the weak arc is divided equally among the cycles, and subtraction occurs again to break the cycles at the weak arc. This continues with ascending values of weak arcs until all are broken. As cycles may have different lengths (i.e., number of arcs per cycle) and quantities of flow, the system is characterized by not only the number of cycles but also the distribution of flow according to cycle length. Finally, as the total flow associated with these cycles represents cycled flow it can be compared to flows of the total system (i.e., total system throughput equals the sum of all inputs, outputs and interactions). The latter comparison (i.e., cycled flow as a fraction of total system throughput) is commonly called the Finn Cycling Index (Finn 1976, 1980).

3. Model Development

3.1. First Generation

In this chapter, we explain not only the results of the modeling exercise but also the processes of developing a model. The original use of modeling for this exercise was to aid one of us (RRC) in organizing information on nitrogen cycling

within the Sacca di Goro (Fig. 8.1) during his stay in Italy in spring 1991. Having surveyed the literature on the lagoon, we constructed the first conceptual model (VDG1A), as shown in Figure 8.2. This model has four compartments containing nitrogen: seston (SES), *Ulva rigida* (ULV), sediments (SED), and dissolved inorganic nitrogen (DIN). These compartments are connected by 10 flows of nitrogen representing different processes (e.g., feeding by heterotrophs of one community on another, deposition, benthic release of DIN). Seston is defined as particulate matter in the water column and includes nitrogen in both living and nonliving entities. Seston and DIN are imported and exported with water currents, nitrogen in sediments is exported through burial, and nitrogen "dissipation" occurs through denitrification in sediments. The degree of aggregation of compartments reflects the information available at the time. Both process and standing stock information were collected at the hierarchical levels represented. For example, nitrogen exchange rates between sediments and the water column were available (Barbanti et al. 1992b, in press at the time), but information on cycling within sediments was not. Flows of nitrogen were computed based on both available information and reasonable conjecture of ecological interaction. Thus the gross primary productivity of seston was estimated from field data, but feeding by
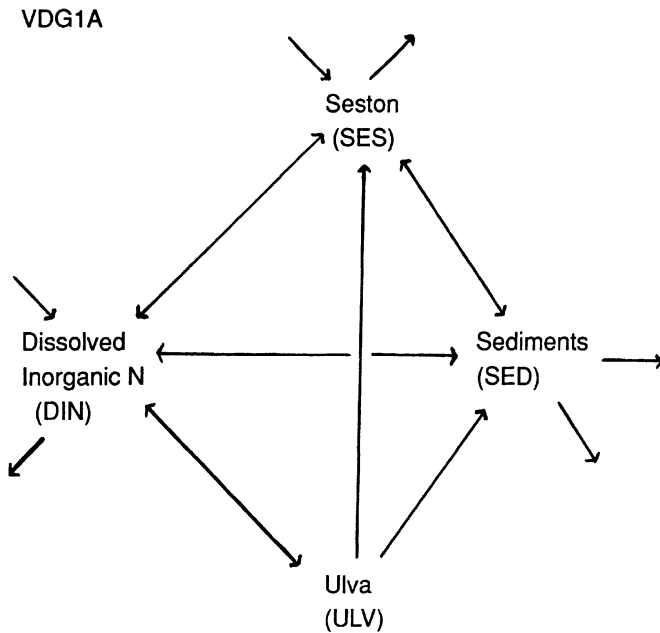
Figure 8.2 First-generation model of nitrogen cycling in the Valle di Gorino, for the period April 1989 (VDG1A).

sestonic, heterotrophic organisms (animals and bacteria) on *Ulva* was not. Also, April 1989 was chosen solely because most pertinent information was available for this time. Therefore, values were derived in a variety of ways: from the literature directly, from unpublished data, from conversions of related literature values, from mass balance calculations of compartments, and from educated guesses based on general ecological principles or knowledge of similar systems.

A total of 19 values (4 standing stocks and 15 flows) were needed for this first, rather small and simplistic model. One important consideration in obtaining such numbers is documentation. In most cases, the available information is not exactly in the form needed for the model. Available information on units, times, locations, species, or elements may be different from those required for the model. Many decisions may be made in converting this information. These decisions must be documented at least for the researcher if not for the readers of any publication. This documentation is equivalent to that needed to identify data in a field or laboratory study (i.e., you must be able to construct the history of each value). Our documentation for this small model is over 20 pages. A summary of the designation and meaning of each compartment and flow, its value, and a reference or brief explanation are given in Table 8.1.

The mere construction of such a model can provide information even before any formal analysis. Based on the comparison of the various standing stocks and rates estimated for the first model, the following observations were made for the prescribed time and location. The availability of DIN during the growth of *Ulva* appears to come largely from loading rather than recycling. This hypothesis is based on two observations. First, estimated respiration rates and sediment release cannot sustain reasonable rates of macroalgal growth. Second, based on concentration patterns, the Valle di Gorino appears to be a sink for DIN such that it draws N from the Po di Goro, intruding saline waters from the Adriatic Sea, and neighboring regions of the lagoon. A result of the low recycling rates, varied sources of loaded nitrogen, and *in situ* concentration patterns is that nitrogen was probably the nutrient limiting growth of autotrophs at this time.

Other observations were also made. During April 1989, phytoplankton gross primary productivity was comparable to that of *Ulva*, and primary productivity by benthic microalgae was minimal. Also, as a general consideration, denitrification may cause significant losses of nitrogen from the system. However, the major source of nitrogen for this process is not necessarily nitrate in the water column. Organic nitrogen reaching sediments may be mineralized to ammonium, this may be nitrified to nitrate in the sediments, and then the nitrate is denitrified. Finally, trophic and migratory interactions between compartments are likely to be small but may be important qualitative links between compartments. Little is known of the rates of these interactions, and research on these interactions may provide a novel link between ecological studies of material and energy processing and community structure. Thus before formal analysis is even conducted, these

Table 8.1 Summary of information for model of nitrogen cycling in the Valle di Gorino during April 1989 (VDG1A)

Description	Value	Reference Explanation
State variables	(mmole N m^{-2})	
Seston (SES)	46	10, 12
Ulva (ULV)	136	4, 9, 12
Sediments (SED)	1000	7
Dissolved inorganic N (DIN)	15.2	1
Imports	(mmole N m^{-2} day^{-1})	
Seston	3.3	1, 13
DIN	21	11
Exports		
Seston	3.3	11
Burial in Sediments	2.2	3, 7
DIN	0	11
Dissipations		
Denitrification from SED	13.3	7, 11
Interactions		
Seston feeding on *Ulva*	1	13
Feeding on benthos and migration	1	13
Seston: Gross primary production	12.4	2, 12
Ulva: Gross primary production	7	4, 9
To sediments from seston	10	11
To sediments from *Ulva*	1	13
Benthic uptake of DIN	5.6	7
Seston: community respiration	1	2, 12
Ulva: Community respiration	1	9, 12
Benthic release of DIN	2.2	2

1, Aquater, 1990; 2, Barbanti et al. 1992a; 3, Barbanti et al. 1992b; 4, Fujita et al. 1989; 5, Pugnetti 1990; 6, Pugnetti et al. 1992; 7, Viaroli et al. 1991; 8, Viaroli et al. 1992; 9, Viaroli and Naldi 1992; 10, unpublished data; 11, mass balance; 12, conversions of elements needed; 13, educated guess.

various observations lead to numerous hypotheses that may provide fruitful avenues for future research.

3.2. Second Generation

The research on the Sacca di Goro was part of a program funded by the European Economic Community to study eutrophication at the mouths of major rivers in the Mediterranean basin (Po, Ebro, and Rhone Rivers). A second genera-

tion of models was designed to compare nitrogen cycles among several of the ecosystems studied and with Christian's previous work on the Neuse River Estuary (Christian et al. 1993). Two models (VDG2B and VDG2C) were constructed to represent June and July, a period of high productivity and one of dystrophy, respectively. Structural changes from the first generation included the following: (1) phytoplankton nitrogen was separated as a compartment from the rest of the seston, (2) *Gracilaria* spp. was added as a compartment for July, (3) net primary productivity was estimated (as opposed to gross primary productivity), and (4) new flows associated with this compartmentalization were defined. New information had become available from the 1991 field season, and as much of this information as possible was incorporated into the second generation models (Table 8.2 and Fig. 8.3).

The second-generation models are intermediate in reliability between the first- and third-generation models. As such, we do not describe them in the detail devoted to the others. We do allude to some results of their analysis in comparison with the other models. Comparative results ranked the Valle di Gorino as the most eutrophic of the systems studied. Nitrogen recycling was found to be the least, and we proposed that macroalgae were responsible for the "short circuiting" of the cycle (Christian et al. 1993).

3.3. Third Generation

By the autumn of 1993, data analysis of the intensive sampling that took place in 1991 was largely complete. Inferences from the first- and second-generation models could now be evaluated with the more reliable data. Therefore, we constructed three models to follow nitrogen cycling in 1991 from the spring (VDG3A), through early summer (VDG3B), and into the dystrophic period of July (VDG3C). Data from six sampling dates from March through July 1991 were used in conjunction with information from laboratory studies using macroalgae from the site. These data provided more specificity to desired time intervals than one normally has in these exercises.

Third generation model structure is summarized in Figure 8.4, and values are given in Table 8.3. The number and designation of compartments are largely as used for the second generation model (VDG2C). *Gracilaria* was found to be present at each represented period during the 1991 field season and not just in July. Dissolved nitrogen now includes both inorganic and organic forms. We estimated fluxes for 3 imports, 5 exports, 1 dissipation, and 13 interactions for each period model and used the following rules in estimating standing stocks and fluxes. When data were directly measured during the period, they were used as an initial estimate. Averages from at least two dates were preferred to single measures. The most satisfactorially estimated fluxes were imports, phytoplankton and macroalgal uptakes of dissolved nitrogen, and sediment exchanges. Next we

Table 8.2 Summary of values for initial models of nitrogen cycling in the Valle di Gorino for 1991 during early summer (VDG2B) and dystrophy in July (VDG2C)

Description	Phase 2 VDG2B	Phase 3 VDG2C
State variables	(mmole N m^{-2})	
Phytoplankton (PHY)	0.7	0.6
Ulva (ULV)	687.9	73.8
Gracilaria (GRA)	0	66.8
Seston (SES)	6.4	11.8
Sediments (SED)	1131	990
Dissolved inorganic N (DIN)	1.5	2
Imports	(mmole N m^{-2} day^{-1})	
Phytoplankton	1.3	1.6
Seston	3.5	3.1
DIN	22.3	6.9
Exports		
Phytoplankton	0.1	0.1
Ulva	6.3	0.7
Seston	0.8	1.2
DIN	0.2	0.2
Dissipations		
Denitrification from SED	49.7	31.2
Interactions		
PHY to SES	1.0	1.4
PHY to SED	0.2	0.15
SES feeding on ULV	4.5	0.4
ULV to SED	12.7	1.1
SES feeding on GRA		0.6
GRA to SED		1.6
SES to SED	0.1	0.1
SES mineralization	0.3	0.6
SED release of DN	0.1	4.5
ULV: net productivity	149.9	11.8
GRA: net productivity		5.9
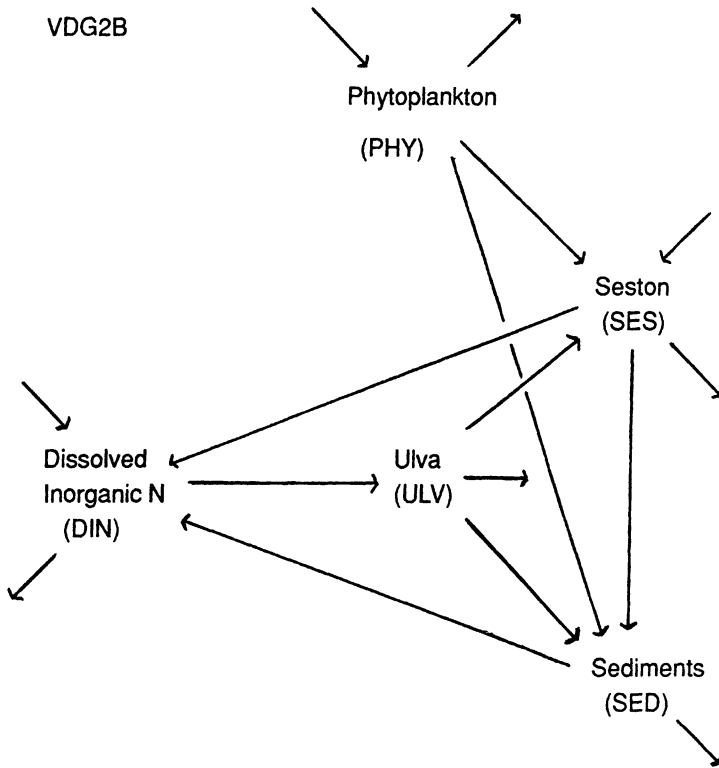
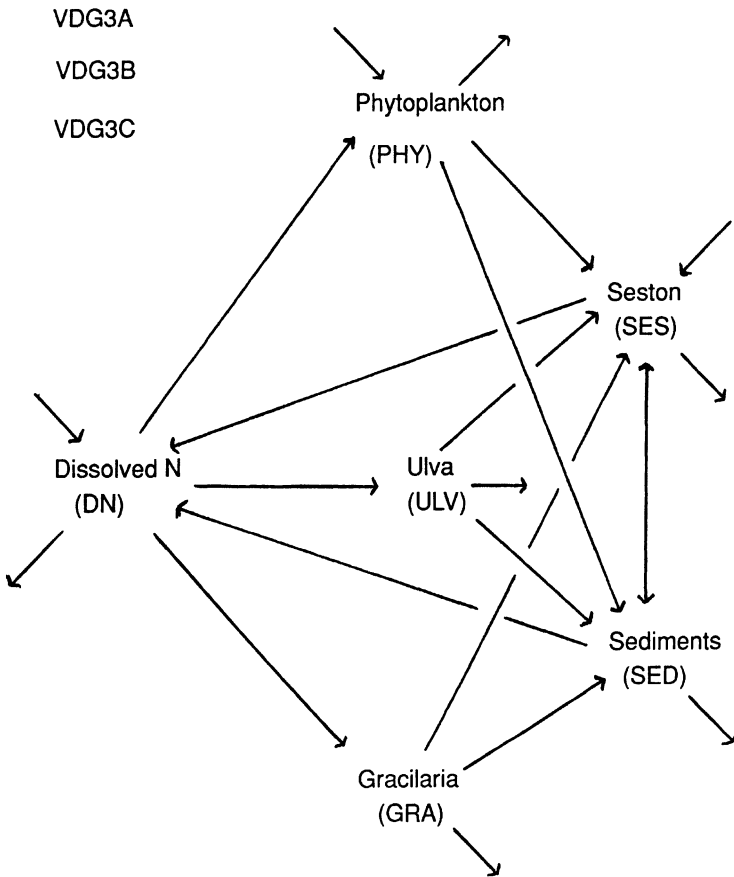Figure 8.3 Second-generation models of nitrogen cycling in the Valle di Gorino, during 1991 for early summer (VDG2B) and July (VDG2C). *Continued*

balanced inputs and outputs for all individual compartments except the macroalgae. This involved estimating some fluxes for which we had little or no direct measurement. For example, we had little information on the fates of macroalgal nitrogen during periods of biomass decrease. If minimal information was available in deciding the distribution of nitrogen fluxes among two or more paths, equal fluxes were assumed initially. Inequality was then instituted as more information was incorporated. Further, we modified fluxes that we had previously estimated to achieve mass balance. This was always done by minimizing change to fluxes for which we felt we had the best estimates.

A central question was always asked when a change was proposed or an estimate made based on little direct data: Does this make sense based on what is known about the ecology of the system? Thus a degree of subjectivity was introduced. Models at the ecosystem level invariably contain such subjectivity. Both the temporal and spatial scales, as well as system's complexity, preclude accurate and precise measurement of all necessary values. This condition has
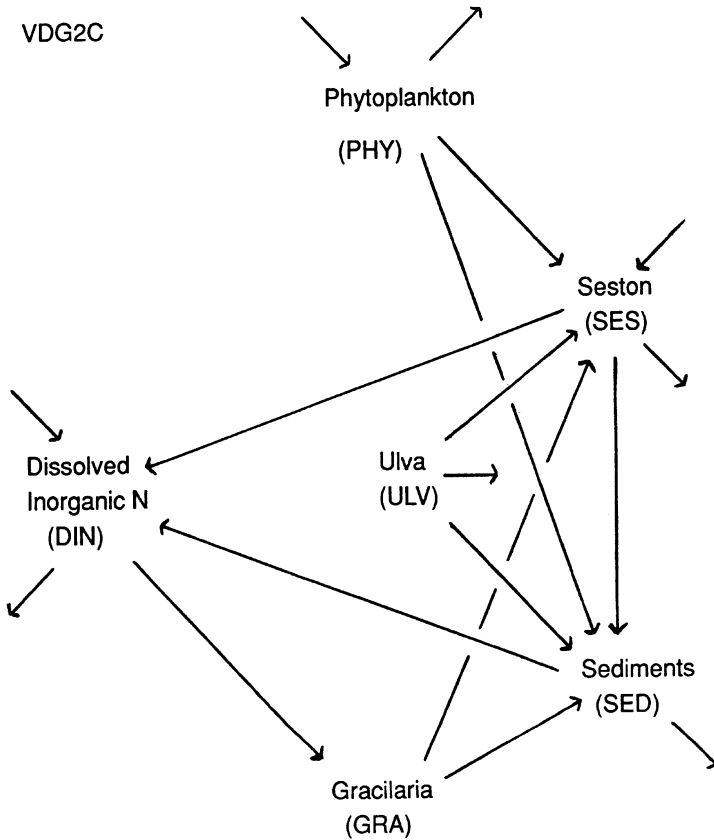
Figure 8.3 *Continued*

been responsible for many to reject ecosystem-level modeling as a worthwhile pursuit (Mann et al. 1989). We contend that modeling can be valuable over a variety of levels of accuracy and precision of information. After all, science progresses through increasing the reliability of hypotheses as more information becomes available. Models represent hypotheses or composites of hypotheses. The important point is to remember that the model and the results of its analysis must be interpreted in the context of its perceived reliability.

4. Analysis Results and Interpretation

In this section we present and interprete those network analyses that we consider most useful to biogeochemistry. The reader is encouraged to peruse papers on the analyses of food webs that use a broader range of network analyses (e.g.,

Figure 8.4 Third-generation models of nitrogen cycling in the Valle di Gorino, during 1991 for spring (VDG3A), early summer (VDG3B), and July (VDG3C).

Baird and Ulanowicz 1989, 1993; Christensen and Pauly 1993). We begin by providing an in depth description of the analysis of the first generation model. We do this to take advantage of its relative simplicity, acknowledging its lack of reliability as a descriptor of the Sacca di Goro. We then use the other model generations for selected comparisons within a generation across periods and across generations.

4.1. First-Generation Structure Analysis

The total dependency matrix of VDG1A is shown in Table 8.4. Each coefficient represents the fraction of nitrogen input to the recipient compartment (designated

Table 8.3 Summary of values for models of nitrogen cycling in the Valle di Gorino for 1991 during spring (VDG3A), early summer (VDG3B), and dystrophy in July (VDG3C)

Description	Phase 1 VDG3A	Phase 2 VDG3B	Phase 3 VDG3C
State variables	(mmole N m^{-2})		
Phytoplankton (PHY)	0.5	0.5	0.4
Ulva (ULV)	47.3	314.4	237.2
Gracilaria (GRA)	124.9	80.5	91.8
Seston (SES)	1.0	2.5	3.8
Sediments (SED)	1429.0	1429.0	1357.0
Dissolved nitrogen (DN)	65.8	57.2	42.3
Imports	(mmole N m^{-2} day^{-1})		
Phytoplankton	1.6	1.8	0.6
Seston	4.5	4.4	1.7
Dissolved nitrogen	26.8	31.1	9.7
Exports			
Phytoplankton	0.1	0.1	0.1
Ulva	0.3	2.1	2.7
Gracilaria	9.9	1.2	0.6
Seston	0.3	0.5	0.3
Dissolved Nitrogen	20.8	9.6	3.8
Dissipations			
Denitrification from SED	4.2	14.8	15.7
Interactions			
SES feeding on PHY	0.9	1.3	0.1
PHY to SED	1.0	1.0	0.7
SES feeding on ULV	0.3	2.1	10.2
ULV to SED	0.3	2.1	10.2
SES feeding on GRA	1.6	1.2	0.6
GRA to SED	1.7	7.5	0.6
SES to SED	2.0	5.0	7.6
SES to mineralization	5.1	3.6	4.8
SES feeding on SED	0.1	0.1	0.1
SED release of DN	0.7	0.7	3.3
PHY: net productivity	0.4	0.6	0.3
ULV: net productivity	2.7	18.2	7.3
GRA: net productivity	8.7	7.0	6.4

by columns) that once resided in the donor compartment (designated by rows). For example, 57% of the nitrogen flowing into sediments (SED) was once in seston (SES). This is largely a direct interaction potentially involving filter feeding by benthic fauna, sedimentation, and migration of organisms to the benthos. Also, 9% of the flow to *Ulva* (ULV) was once in seston even though no direct interaction was included in the model. The only direct source of nitrogen for *Ulva* was from DIN which thus supplied 100% of the nitrogen needs. The 9% from seston must have first passed through the DIN pool. In fact some also passed through the sediments before reaching the DIN pool. Therefore this matrix can be used to evaluate the extended sources of material to a compartment. If this were a model of energy flow through a food web the extended source could be considered the "extended diet", a concept effectively used by Baird and Ulanowicz (1989) for trophic interactions in Chesapeake Bay, USA.

From the total dependency matrix (Table 8.4) we see that, as expected, DIN was the most important donor of nitrogen to all other compartments. Nitrogen was recycled through the system from each compartment. *Ulva* was the least important source of nitrogen, whereas seston and sediments acted similarly as sources to both *Ulva* and DIN. Coefficients along the diagonal represent the fraction of complete recycling, indicating how much of the flow to a compartment once previously resided in that compartment. These values ranged from 5% to 14% giving the first indication of relatively little recycling.

The reverse perspective of interaction is seen through the total contribution matrix (Table 8.5). Here each coefficient represents the relative amount of output from a compartment (represented by row) that enters another (represented by column), directly or indirectly. The diagonal coefficients are the same in both value and meaning as those for the total dependency matrix. In this view, for example, 60% of the outputs from both seston and DIN passed into sediments, whereas 32% of *Ulva* output entered sediments. Less than 15% of sediment release of nitrogen (by all processes) entered any other single compartment by any combination of internal interactions. Denitrification and burial were major

Table 8.4 Total dependency matrix of nitrogen cycling in the Valle di Gorino during April 1989 (VDG1A)

Donor	Recipient			
	SES	ULV	SED	DIN
SES	0.10	0.09	0.57	0.09
ULV	0.10	0.05	0.12	0.05
SED	0.13	0.10	0.10	0.10
DIN	0.80	1.00	0.78	0.14

Coefficients represent fraction of recipient's throughput that once resided in donor.

Table 8.5 Total contribution matrix of nitrogen cycling in the Valle di Gorino during April 1989 (VDG1A)

Donor	Recipient			
	SES	ULV	SED	DIN
SES	0.10	0.04	0.60	0.13
ULV	0.24	0.05	0.32	0.18
SED	0.12	0.04	0.10	0.12
DIN	0.59	0.31	0.60	0.14

Coefficients represent fraction of donor's throughput that passes into recipient.

outputs from the sediments, and their magnitude ensured relatively low contributions to other compartments. Thus, this matrix extends our understanding of the fates of materials leaving a compartment.

A third type of structure analysis (input environs analysis) evaluates the fate of each import through the system. Coefficients of calculated vectors and matrices represent the number of atoms of nitrogen associated with a flow for every atom of nitrogen entering through a particular path. In Table 8.6 we show two vectors and the exchange matrix summarizing the contributions from the sole import of 1 atom of seston nitrogen. Twenty-one percent left as seston, and 8% was buried

Table 8.6 Input environs analysis of imported seston in the Valle di Gorino during April 1989 (VDG1A)

	Export from			
	SES	ULV	SED	DIN
	0.21	0.00	0.08	0.00
	Dissipation from			
	SES	ULV	SED	DIN
	0.00	0.00	0.48	0.00
Exchange Matrix				
		Recipient		
Donor	SES	ULV	SED	DIN
SES	0.00	0.00	0.63	0.06
ULV	0.01	0.00	0.01	0.01
SED	0.04	0.00	0.00	0.08
DIN	0.07	0.04	0.03	0.00

Coefficients represent the number of atoms of nitrogen within a flux per atom of seston nitrogen imported as seston.

in the sediments ("export from" vector). The equivalent to nearly half of the nitrogen entering as seston was dissipated through denitrification ("dissipation from" vector). The contribution of seston nitrogen to denitrification occurred largely through the considerable flow of seston (SES) to sediments (SED) as seen in the exchange matrix (0.63). Presumably, within the sediments organic N from seston was first mineralized and nitrified. Thus, using this analysis one can trace the fate of material as it passes through and out of a system.

4.2. Biogeochemical Cycle Analysis

In this analysis, the nature of the pathways of nitrogen cycling are examined. Within the first-generation model were 10 cycles or positive feedback loops (Table 8.7). Four cycles involved only 2 arcs (2 fluxes and, therefore, 2 compartments), 4 cycles involved 3 arcs, and 2 involved all 4 compartments and 4 arcs. Seven cycles involved seston, 5 involved *Ulva*, 7 involved sediments, and all but 1 involved DIN. This pattern of occurrence fits the degree to which each compartment was an extended source of nitrogen to others described in the previous section: DIN > seston = sediment > *Ulva*.

As discussed, the sources for information for this model were varied, and the accuracy of some of the values is suspect. We wished to represent five fluxes, for which we had virtually no information, as positive but relatively small numbers. These were fluxes involving trophic interactions for the most part, and such interactions were not directly studied. The value 1 mmole N m^{-2} day^{-1} was used for each. Unfortunately, the use of the same small number so many times compromised our ability to group cycles into nexuses (cycles sharing the same

Table 8.7 Biogeochemical cycles of nitrogen in the Valle di Gorino during April 1989 (VDG1A)

Cycle	Weak Arc	Weak Arc Value (mmole N m^{-2} day^{-1})
SES-SED-SES	SED-SES	1.0
SES-DIN-SES	SES-DIN	1.0
SES-DIN-SED-SES	SES-DIN, SED-SES	1.0, 1.0
SES-DIN-ULV-SES	SES-DIN, ULV-SES	1.0, 1.0
SES-DIN-ULV-SED-SES	SES-DIN, ULV-SED,	1.0, 1.0
	SED-SES	1.0
SES-SED-DIN-ULV-SES	ULV-SES	1.0
SED-DIN-ULV-SED	ULV-SED	1.0
DIN-ULV-DIN	ULV-DIN	1.0
SES-SED-DIN-SES	SED-DIN	2.2
SED-DIN-SED	SED-DIN	2.2

weak arc or smallest flux). The value, 1, served as the weak arc for all but two cycles. However, this value occurred more than once in two cycles. Although the grouping into nexuses helps to delineate the potential for shared controlling processes among cycles, the duplication of weak arc values makes this analysis less tenable for this particular model. It is evident, however, that remineralization pathways into the DIN pool (SES-DIN and SED-DIN) may be important limits to the cycling of nitrogen.

We also considered the distribution of flows through cycles of different lengths. This distribution can be normalized to total systems throughput of nitrogen in the system. Thus, 7.2% of total systems throughput (85.3 mmole N m^{-2} day^{-1}) passed through cycles possessing a length of 2 (i.e., between two compartments). The percentages of flow through cycles of length 3 and 4 were 3.9% and 4.7%, respectively. Information of this sort has been used to evaluate the degree of stress on systems represented as food webs. Stress may be represented as "changes in the flows of energy in a system, disappearance of previously existing flows, or acceleration of repair work" (Lugo 1978). For food webs it is postulated that stress results in more flow in shorter cycles than in unperturbed systems (Wulff et al. 1989; Baird et al. 1991). However, we have not found this to be true for biogeochemical cycle-based models (Christian et al. 1996).

By summing the normalized flows per cycle length (given in the previous paragraph), we obtained the total percentage of throughput associated with recycling equal to an FCI of 15.8%. This is quite small compared to the N recycling portrayed in models of other coastal systems, as studied by Christian et al. (1993, 1996) and Baird et al. (1995). Although differences in model structure may contribute to differences in FCI, the Sacca di Goro, as depicted in VDG1A, has features which would minimize FCI: (1) a relatively short residence time of water, (2) a positive growth of both micro- and macroalgae, and (3) minimal mineralization and release into the DIN pool.

4.3. Across-Period Analysis

We consider that the most valuable use of network analysis is in comparisons: within system across time (Baird and Ulanowicz 1989; Christian et al. 1992; Forès and Christian, 1993; Baird et al. 1995), and across systems (Baird et al., 1991; Baird and Ulanowicz, 1993; Christensen and Pauly 1993; Christian et al. 1993, 1996). Here we compare models of three periods in 1991: VDG3A for spring, VDG3B for early summer, and VDG3C for the July dystrophic period. Two system-level indices are total systems throughput and Finn cycling index (FCI). The former remained relatively constant through time as 94, 116, and 87 mmole N m^{-2} day^{-1} for the respective periods. In contrast FCI was low during the first two periods (9.5% and 11.4%, respectively, for spring and early summer) and increased to 29.2% during the dystrophic period. During the periods of low

FCI, macroalgal growth sequestered nitrogen and restricted cycling. The increase in cycling would be expected as *Ulva* decomposed and organic nitrogen was released and/or mineralized.

The total dependency matrices of the three models were used to examine the relative importances of compartments to DN dynamics (Fig. 8.5). Again recycling from the macroalgae was low except during dystrophy, when *Ulva* was the major source for other compartments. All recycling to DN from macroalgae was indirect either through seston or sediments. Included in Figure 8.5 is the recycling of DN associated with all.intermediates. Again the value is low overall, but is greatest during the dystrophic period. Although direct relationships are easily made without network analysis, such conclusions concerning indirect relationships are not readily made by alternative means.

Input analysis was also used to compare models (Fig. 8.6). We evaluated the fate of imported DN. Again we emphasize indirect relationships, this time associated with sediments. The largest fractions of imported DN to pass to sediments occurred during early summer from *Gracilaria* and during July from
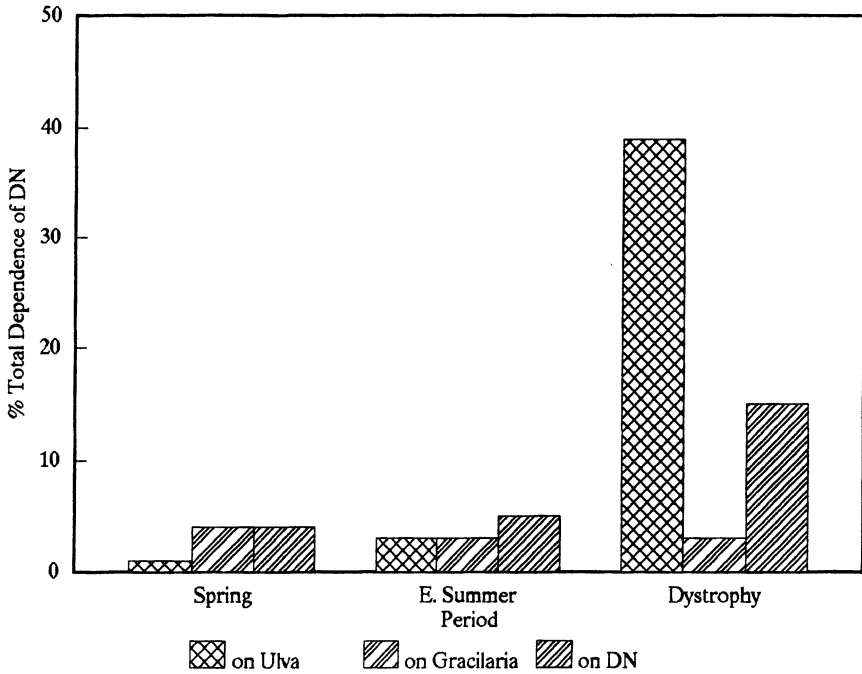
Figure 8.5 Total dependency of dissolved nitrogen (DN) on selected compartments by period in third-generation models.
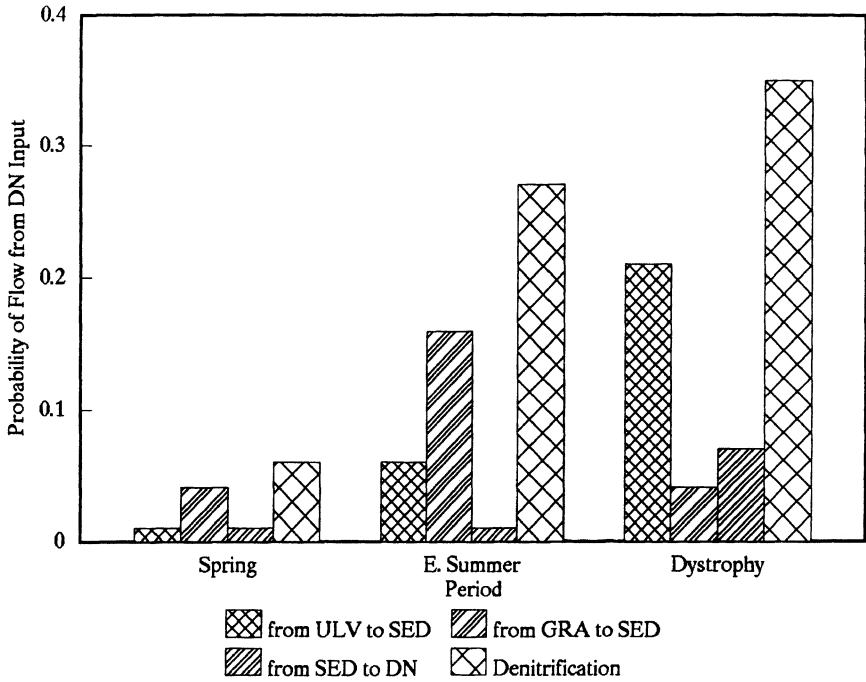
Figure 8.6 Probability of nitrogen flow from the import of dissolved nitrogen (DN) through selected pathways by period in third generation models. Abbreviations of compartments through which flow occurs are as shown in Figure 8.4 and defined in text.

Ulva. Dystrophy in July was largely in response to *Ulva* death and decomposition. *Gracilaria* actually grew during this period. Subsequent return of nitrogen from sediments to DN occurred with greatest probably during dystrophy. Loss of imported DN from the sediments via denitrification occurred with greater probability during both summer periods than during spring.

Each model had the same 13 cycles. One was 2 arcs in length, six had 3 arcs, and the remainder had 4 arcs. Those with lengths of three arcs were associated with the most cycled flow (Fig. 8.7). This large contribution of 3 arc cycles to cycled flow was primarily the result of one or more cycles involving the flow from DN to a macroalga to either seston or sediments and back to DN. The macroalga involved in these important cycles changed from *Gracilaria* in spring to *Ulva* during both summer periods. All cycles but one involved DN, 11 involved sediments, and 10 involved seston. Each primary producer was involved in four cycles. Thus, although macroalgae were involved in a limited number of cycles, their quantative contribution to cycling was important. Again, we find that network
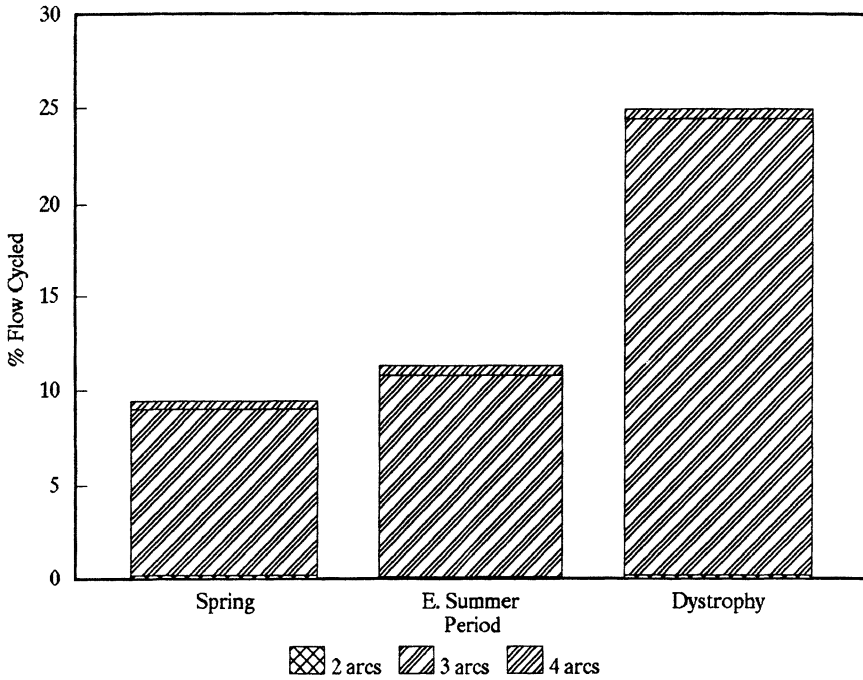
Figure 8.7 Distribution of cycled fluxes according to the length of cycles by period in third-generation models. The number of arcs refers to the number of flows associated with cycles. For example "3 arcs" includes all cycles that involve three flows connecting three compartments.

analysis provides a unique view of model structure and inferentially nutrient cycling.

4.4. Intergenerational Comparisons

As previously indicated, we consider modeling an iterative process, as is science in general. Better information can be used to produce more reliable models, and each generation of model can promote the search for better information. In this section we provide examples of how analysis results compare among models of different generations. Model structure changed with increased numbers of compartments and flows with successive generations (Table 8.8). As a result, the most direct comparisons can be made with system-level attributes. Total systems throughputs were most similar for generations 1 and 3 in spite of the doubling of number of flows in the latter. The second-generation models had the highest (VDG2B) and lowest (VDG2C) values. The high value was largely the

Table 8.8 Comparison of selected traits of three generations of models and their network analyses concerning nitrogen cycling in the Valle di Gorino

Index	VDG1A	VDG2B	VDG2C	VDG3A	VDG3B	VDG3C
Compartments	4	5	6	6	6	6
Number of flows	10	16	19	21	21	21
Total systems throughput (mmole N m^{-2} day^{-1})	85	221	68	94	116	87
Dependency (as percentage)						
DN on ULV	5	<1	1	1	3	39
DN on GRA	NA	NA	4	4	3	3
DN on SES	9	<1	6	17	11	36
DN on SED	10	<1	34	2	2	20
DN on DN	14	<1	3	4	5	15
Input of DN (as fraction of atom per 1 atom of DN input)						
to PHY	0.57	0.00	0.00	0.01	0.02	0.02
to ULV	0.32	1.00	0.57	0.09	0.54	0.48
to GRA	NA	NA	0.45	0.28	0.21	0.42
from ULV to SED	0.05	0.11	0.08	0.01	0.06	0.21
from GRA to SED	NA	NA	0.12	0.04	0.16	0.04
from SED to DN	0.08	<0.01	0.03	0.01	0.01	0.07
denitrification	0.48	0.11	0.18	0.06	0.27	0.35
Biogeochemical cycles						
Number	10	3	6	13	13	13
Number of nexuses	6	3	4	4	3	4
FCI	15.8	0.34	13.8	9.5	11.4	29.2

NA refers to not applicable.
Season designations are A for spring, B for early summer, and C for dystrophy in July.

result of a large estimate of macroalgal primary productivity during early summer compared to VDG3B. This estimate was reduced with better data analysis for the third generation. The low value resulted from several small fluxes throughout when compared to the third generation model of dystrophy (VDG3C). As one might expect, greater information stabilizes "mean" values, in this case, the values used to represent fluxes for set time intervals.

Lower FCI values and DN recycling for early summer and July were also found for the second generation when compared to the respective third-generation models (Fig. 8.8). This is to be expected as the second-generation models had a reduced number of cycles; increased denitrification; and, during early summer, high rate of sequestering of nitrogen by macroalgae compared to the third genera-
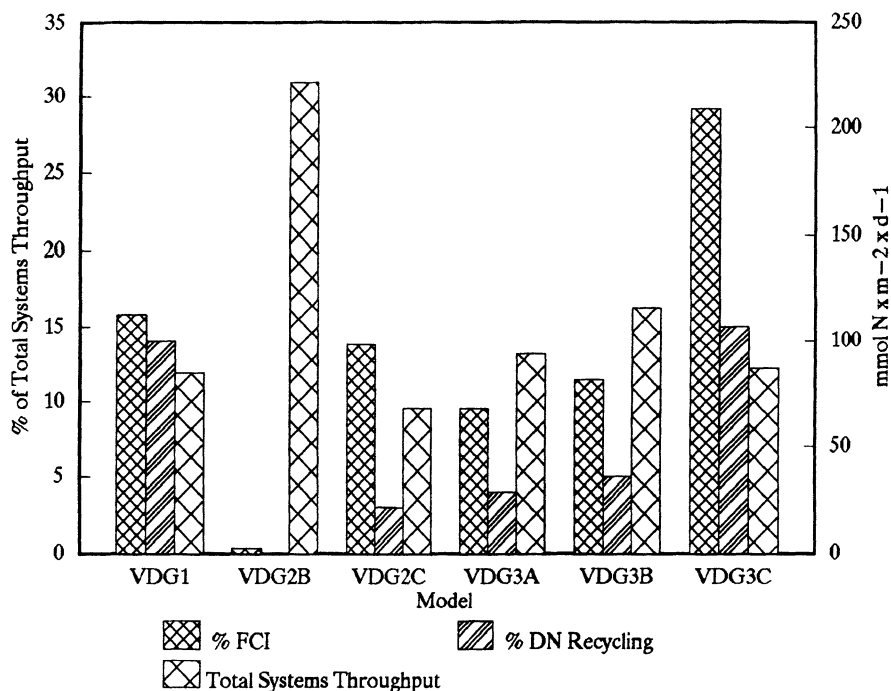
Figure 8.8 Comparison of models by generation for total systems throughput (mmole N m^{-2} day^{-1}), DN recycling (percentage total systems throughput) and Finn cycling index (FIC, percentage total systems throughput). Model designations are given in text.

tion. The rates of denitrification used for the second generation were much greater than those used by us for other systems (Christian et al., 1993), had produced considerable imbalance in sediment nitrogen, and were based on a technique that might promote overestimation (El-Habr and Golterman 1990). Denitrification rates are particularly hard to estimate (Seitzinger 1988). In the third-generation models, we used measured estimates as maximum limits. Nitrate reduction rates were measured and considered twice the rates of denitrification (Hattori 1983). Rates were then modified by comparing these with mass balance requirements. Again, by putting a rate in context of other flows, one may be forced to revaluate its accuracy, as we have done.

Overall, however, by comparison to similar models of other coastal ecosystems, the intergenerational differences are small. Total systems throughput values for all VDG models were large compared to two lagoons in the Ebro River delta (Christian et al. 1993, 1996; Forès et al. 1994) and in the range of highly productive ricefields (Forès and Christian 1993) and an eutrophic estuary (Christian et al. 1992). Similarly, FCI values for all generations were low compared to those

for other systems, and interperiod trends were as expected for both the second- and third-generation models.

Quantitative precision and accuracy should increase with greater reliability of information used for model construction. For our models, the general trends that we have inferred have not been modified significantly and may confirm our perceptions of the system. However, changes in trends over generations may also provide valuable information to the scientist and force him or her to understand why a change occurred and to reevaluate the assumptions associated with that change.

5. Conclusions and Subsequent Directions

Three generations of static, multicompartmental or structured models of nitrogen cycling of the Valle di Gorino were developed and subjected to network analysis to formalize understanding of the system. Several inferences concerning cycling were made from the analyses:

1. Recycling represented a small percentage of total nitrogen flow within this system with maximum, but still small, recycling occurring during dystrophy.
2. Imports of DN are largely sequestered within macroalgae during their growth, and the relative contribution of each macroalgal species to recycling changes over time.
3. The decomposition of *Ulva* during July's dystrophy dominates nitrogen cycling.
4. Denitrification may cause a significant loss of N, but the major source of nitrate within the sediments is probably from the deposition and decomposition of organic matter to sediments followed by nitrification and not from nitrate uptake from the water column.

The first model, constructed in 1991 with values largely based on conjecture, was developed as a crude tool to address the ongoing research. Consistent research was lacking on such areas as external loading, benthic fluxes and nutrient uptake by macroalgae over the growing and dystrophic periods. Early field studies did reveal regular trends in nutrient concentrations in the water column, potential for nitrogen limitation of primary production and nitrogen pulses conforming to the life cycle of *U. rigida* (Viaroli et al. 1992; Viaroli and Naldi 1993). Early experience in modeling did provide more precise hypotheses about the key role of *Ulva* in the functioning of the lagoon; by storing and immobilizing nitrogen in the spring and by allowing release during dystrophy. Subsequently we improved our research program to consider the following: (1) trends in hydrological and hydrochemical variables and in external loading; (2) seasonal succession of

Table 8.9 Nitrogen content in *Ulva* and the water column within the Valle di Gorino during field studies from April through August 1992

Date	*Ulva*-N (mmole N m⁻²)	Total N in Water (mmole N m⁻²)	%DIN	%DON of Total N in Water	%PON
9 Apr.	381.0	30.2	1.8	75.7	22.5
5 May	635.4	50.4	2.7	70.0	27.3
9 Jun.	374.7	42.2	3.7	84.2	12.1
7 Jul.	162.8	32.4	2.9	87.7	9.4
21 Jul.	0.0	60.7	3.9	70.2	25.9
28 Jul.	0.0	90.0	10.5	52.2	37.3
4 Aug.	0.0	54.5	11.9	48.0	40.1

macroalgae and phytoplankton; (3) production and respiration of both macroalgal and plankton communities; (4) changes in the elemental composition of macroalgal biomass; and (5) decomposition of macroalgal detritus and its effects on sediment oxygen demand, water-sediment nitrate reduction, and regeneration of ammonium and soluble reactive phosphorus (Viaroli et al., 1993 a,b). This second series of studies was syneristic with modeling by being in part directed by model construction needs and analysis results and by providing information for better models.

There was significant agreement between the improved model output and experimental and field results. In the models, nitrogen cycling was least dependent on *Ulva* during the two growing phases (Fig. 8.5). The same interpretation was made of our 1992 field data (Table 8.9). *Ulva* is seen as a sink during active growth. When *Ulva* undergoes decomposition, increased nitrogen cycling occurs through both organic dissolved and particulate (seston) forms, but DIN remineralization seems relatively negligible. Direct remineralization of *Ulva* nitrogen is small, and much DIN is released after further processing. This was evidenced in both field and laboratory results (Viaroli and Naldi 1993; Viaroli et al. 1993a) and in the length of quantitatively important cycles within the models of dystrophy (Fig. 8.7).

Modeling helped us develop more appropriate experimentation in the field. In 1993, we tested the hypotheses involving the sequestering of nitrogen by *Ulva* during growth and subsequent ineffective direct recycling and mineralization to DIN during dystrophy. Three experiments sere conducted with light and dark benthic chambers and bell jars, providing different combinations of plankton, sediment, and *Ulva*: one in the Sacca di Goro and two in the lagoon Etang du Prevost, France (Viaroli et al. 1993b,c). Here we summarize results from Etang du Prevost. Over a 3-day period in the dark 14% of the initial nitrogen in *Ulva* was accounted for as total dissolved nitrogen (TDN), but only 6% was as ammonium. In the light TDN accounted for only 8% of the initial macroalgal

nitrogen, and ammonium was 1%. These results can be interpreted to confirm the hypothesis that released nitrogen during decomposition included considerable organic material.

The general results of field and laboratory studies, modeling, and network analysis are in agreement with a conceptual model of macroalgal growth and decomposition. Nutrient cycling shifts from a spring growing phase through early summer to a decomposition phase (Fig. 8.9). If water renewal is slow, as observed in the Valle di Gorino, the rapid changes from the growing to stationary to decomposition phases may produce an oxygen imbalance which in turn causes increased reducing conditions and anoxia. The rapid and strong anoxic decomposition of *Ulva* does not release all the energy of the decaying detritus; part of this energy accumulates in reduced organic compounds and part as sulfides. Sulfide production decouples carbon cycling and energy flow, and organic matter decomposition is incomplete.

Nitrogen seems to be one of the most important factors in the "bottom-up" or resource availability control of primary production in the Valle di Gorino. During
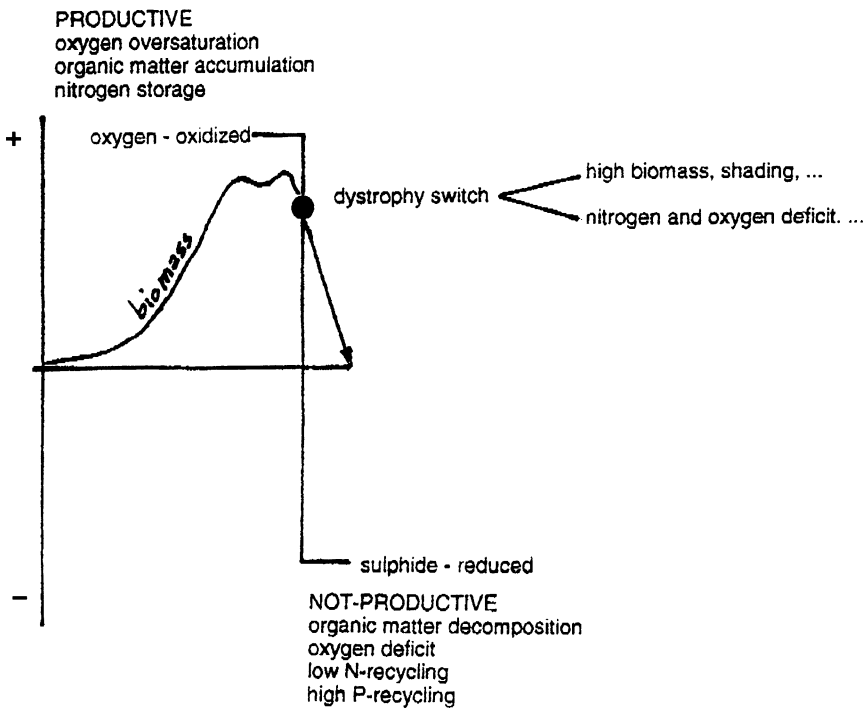
Figure 8.9 Conceptual model representing mechanisms and processes involved in the dystrophic crisis.

the growing phases, nitrophilous macroalgae take up most of the available DIN and make it more limiting to other primary producers. *Ulva* and *Gracilaria* dominate with a depression of the planktonic food web. The dystrophic event seems to start inside and beneath the macroalgal mats and overlaps the processes taking place in the surface sediments: inorganic nutrient release, potential nitrate losses due to nitrate reduction, and imbalance in organic matter decay and nutrient regeneration. The trophic structure is dominated by a loop that includes macroalgae and associated microbial communities. Incomplete decomposition fosters continued nitrogen limitation.

The temporal uncoupling between growth and decomposition of macroalgae with incomplete recycling requires further study. We are considering the extent to which the life cycle of *Ulva* and intensity of its production in the Valle di Gorino actually provide negative feedback for the macroalgal growth. The question remains open as it is difficult to distinguish between the effects of external factors, such as loading, from internal ecological processes. Currently, we are focusing on field studies concerning the decoupling of primary production and decomposition processes as having key responsibility for system control. Although static models have been most helpful to this point; in matters that involve important temporal changes in the mechanisms of control, dynamic modeling may be needed for future evaluations. The interface between dynamic models and network analysis remains an interesting direction for the future.

Acknowledgments

References

Aquater. 1990. *Piano di Risanamento della Sacca di Goro, Relazione finale.* Amministrazione Provinciale di Ferrara, Ferrara, Italy.

Baird, D., and R. E. Ulanowicz. 1989. The seasonal dynamics of the Chesapeake Bay. *Ecol. Monogr.* 59:329–364.

Baird, D., and R. E. Ulanowicz. 1993. Comparative study on the trophic structure, cycling and ecosystem properties of four tidal estuaries. *Mar. Ecol. Prog. Ser.* 99:221–237.

Baird, D., J. M. McGlade, and R. E. Ulanowicz. 1991. The comparative ecology of six marine ecosystems. *Philos. Trans. R. Soc. Lond. B* 333:15–29.

Baird, D., R. E. Ulanowicz, and W. R. Boynton. 1995. Seasonal dynamics in Chesapeake Bay: A network approach. *Estuar. Coast. Shelf Sci.* 41:137–162.

Barbanti, A., V. U. Ceccherelli, F. Frascari, G. Rosso, and G. Reggiani. 1992a. Nutrient release from sediments and role of bioturbation in the Goro lagoon (Italy). *Sci. Total Environ., Suppl.* 1992:475–487.

Barbanti, A., V. U. Ceccherelli, F. Frascari, G. Reggiani, and G. Rosso. 1992b. Nutrient regeneration processes in bottom sediments in a Po delta lagoon (Italy) and the role of bioturbation in determining the fluxes at the sediment-water interface. *Hydrobiologia* 228:1–21.

Christensen, V., and D. Pauly (eds.). 1993. *Trophic models of aquatic ecosystems.* International Center for Living Aquatic Resources Management, Manila, Philippines.

Christian, R. R., J. N. Boyer, D. W. Stanley, and W. M. Rizzo. 1992. Network analysis of nitrogen cycling in an estuary. In C. Hurst (ed.), *Modeling the Metabolic and Physiologic Activities of Microorganisms*, pp. 217–247. Wiley, New York.

Christian, R. R., E. Forès, F. Comín, P. Viaroli, and I. Ferrari. 1993. Comparative network analysis of nitrogen cycling in several eutrophic coastal ecosystems. In R. Guerrero and C. Pedros-Alio (eds.), *Trends in Microbial Ecology*, pp. 449–452. Spanish Soc. For Microbiology, Barcelona, Spain.

Christian, R. R., E. Forès, F. Comín, P. Viaroli, M. Naldi, and I. Ferrari. 1996. Nitrogen cycling networks of coastal ecosystems: influence of trophic status and primary producer form. *Ecol. Model.* 87:111–129.

Christian, R. R., and R. L. Wetzel. 1991. Synergism between research and simulation models of estuarine microbial food webs. *Microb. Ecol.* 22:111–125.

Christian, R. R., R. L. Wetzel, S. M. Harlan, and D. W. Stanley. 1986. Growth and decomposition in aquatic microbial systems: alternate approaches in simple models. In F. Megusar and M. Gantar (eds.), *Perspectives in Microbial Ecology*, pp. 38–45. Slovene Society for Microbiology, Ljubljana, Yugoslavia.

Colombo, G. 1989. *Variazioni spaziali e temporali delle caratteristiche fisico-chimiche delle acque e delle biomassa fitoplanctonica. Relazione finale (Settembre 1987– Dicembre 1988),* Convenzione tra Amministrazione Provinciale di Ferrara e l'Universita' di Ferrara, Istituto di Zoologia, Ferrara, Italy.

Day, J. W., Jr., C. A. S. Hall, W. M. Kemp, and A. Yanez-Arancibia. 1989. *Estuarine Ecology.* John Wiley & Sons, New York.

El-Habr, H., and H. L. Golterman. 1990. *In vitro* and *in situ* studies on nitrate disappearance in water-sediment systems of the Camargue (Southern France). *Hydrobiologia* 192:223–232.

Ferrari, I., V. U. Ceccherelli, M. Naldi, and P. Viaroli. 1993. Planktonic and benthic communities in a shallow-water dystrophic lagoon. *Verh. Internat. Verein. Limnol.* 25:1043–1047.

Finn, J. T. 1976. Measures of ecosystem structure and function derived from analysis of flow. *J. Theor. Biol.* 56:363–380.

Finn, J. T. 1980. Flow analysis of models of the Hubbard Brook Ecosystem. *Ecology* 61:562–571.

Forès, E., and R. R. Christian. 1993. Network analysis of nitrogen cycling in temperate, wetland ricefields. *Oikos* 67:299–308.

Forès, E., R. R. Christian, F. A. Comín, and M. Menendez. 1994. Network analysis on nitrogen cycling in a coastal lagoon. *Mar. Ecol. Progr. Ser.* 106:283–290.

Fujita, R. M., P. A. Wheeler, and R. L. Edwards. 1989. Assessment of macroalgal nitrogen limitation in a seasonal upwelling region. *Mar. Ecol. Prog. Ser.* 53:293–303.

Hannon, B. 1973. The structure of ecosystems. *J. Theor. Biol.* 41:535–546.

Hattori, A. 1983. Denitrification and dissimilatory nitrate reduction. In E. J. Carpenter and D. G. Capone (eds.), *Nitrogen in the Marine Environment*, pp. 191–232. Academic Press, New York.

Kay, J. J., L. A. Graham, and R. E. Ulanowicz. 1989. A detailed guide to network analysis. In F. Wulff, J. G. Field, and K. H. Mann (eds.), *Network Analysis in Marine Ecology: Methods and Applications*, pp. 15–61. Springer-Verlag, Berlin.

Lugo, A. E. 1978. Stress and ecosystems. In J. H. Thorp and J. W. Gibbons (eds.). *Energy and Environmental Stress in Aquatic Systems*, pp. 62–101. DOE Symposium Series (CONF-771114). National Technical Information Service, Springfield, VA.

Mann, K. H., J. G. Field, and F. Wulff. 1989. Network analysis in marine ecology: an assessment. In F. Wulff, J. G. Field, and K. H. Mann (eds.), *Network Analysis in Marine Ecology: Methods and Applications*, pp. 259–282. Springer-Verlag, Berlin.

Odum, H. T. 1983. *Systems Ecology: an Introduction.* Academic Press, New York.

Patten, B. C. 1985. Energy cycling in the ecosystem. *Ecol. Model.* 28:1–71.

Patten, B. C., R. W. Bosserman, J. T. Finn, and W. G. Cale. 1976. Propagation of cause in ecosystems. In B. C. Patten (ed.), *Systems Analysis and Simulation in Ecology, Vol. 4*, pp. 457–579. Academic Press, New York.

Pugnetti, A. 1990. *Struttura e dinamica dei popolamenti planctonici di una laguna del Delta del Po (Sacca di Goro).* Tesi di dottorato in Scienza dell'Ambiente, Universita' di Parma, Parma, Italy.

Pugnetti, A., P. Viaroli, and I. Ferrari. 1992. Processes leading to dystrophy in a Po River delta lagoon (Sacca di Goro): phytoplankton-macroalgae interaction. *Sci. Tot. Environ., Suppl.* 1992:445–456.

Ryther, J. H., and W. Dunstan. 1971. Nitrogen, phosphorus, and eutrophication in the coastal marine environment. *Science* 171:1008–1013.

Seitzinger, S. 1988. Denitrification in freshwater and coastal marine ecosystems: ecological and geochemical significance. *Limnol. Oceanogr.* 33:702–724.

Shugart, H., and R. V. O'Niell (eds.). 1979. *Systems Ecology: Benchmark Papers in Ecology.* Dowder, Hutchinson and Ross, Inc., Stronsburg, PA.

Ulanowicz, R. E. 1983. Identifying the structure of cycling in ecosystems. *Math. Biosci.* 65:219–237.

Ulanowicz, R. E. 1986. *Growth and Development: Ecosystems Phenomenology.* Springer-Verlag, New York.

Ulanowicz, R. E. 1987. *NETWRK4: A Package of Computer Algorithms to Analyze Ecological Flow Networks.* University of Maryland, Chesapeake Biological Laboratory, Solomons.

Viaroli, P. 1992. Eutrophication of the Po delta lagoons: Evolution and prospects for restoration. In M. Finlayson, T. Hollis, and T. Davis (eds.), *Managing Mediterranean Wetlands and Their Birds*, pp. 159–164. IWRB Special Publication No. 20. Parma.

Viaroli, P., M. Bartoli, C. Bondavalli, M. Cattadori, G. Giordani, M. Naldi, I. Ferrari, and I. Fumagalli. 1993c. Macroalgae and macrophytes, oxygen and nutrient fluxes at the community level. In F. Rodriguez-Valera and P. Caumette (eds.), *CLEAN, Coastal Lagoon, Eutrophication and Anaerobic Processes*, pp. 255–279. Progress Report to Commission of the European Communities Environmental Research Programme, Bordeaux Univ., France.

Viaroli, P., M. Bartoli, C. Bondavalli, M. Naldi, I. Ferrari, E. Fano, L. G. Gatti, and V. Gaiani. 1993b. Macroalgae growth and decomposition and nutrient cycling in the Sacca di Goro (Po River Delta). In F. Rodriguez-Valera and P. Caumette (eds.), *CLEAN, Coastal Lagoon Eutrophication and Anaerobic Processes*, pp. 281–307. Progress Report to Commission of the European Communities Environmental Research Programme, Bordeaux Univ., France.

Viaroli, P., M. Cavalca, I. Fumagalli, V. Gaiani, and V. Zaccaria. 1991. Nitrogen losses in a shallow eutrophic coastal lagoon (Sacca di Goro): preliminary results. *SItE Atti 12* (Italian Society of Ecology—Proc. of 4th Congress, in Italian), pp. 149–153.

Viaroli, P., and M. Naldi. 1992. Research on nitrogen and phosphorus cycles in a coastal dystrophic lagoon (Sacca di Goro, Delta del Po). *SItE Atti 15* (Italian Society of Ecology—Proc. of 5th Congress, in Italian), pp. 95–115.

Viaroli, P., M. Naldi, R. R. Christian, and I. Fumagalli. 1993a. The role of macroalgae and detritus in the nutrient cycles in a shallow-water dystrophic lagoon. *Verh. Internat. Verein. Limnol.* 25:1048–1051.

Viaroli, P., A. Pugnetti, and I. Ferrari. 1992. *Ulva rigida* growth and decomposition processes and related effects on nitrogen and phosphorus cycles in a coastal lagoon (Sacca di Goro, Po River Delta). In G. Colombo, I. Ferrari, V. U. Ceccherelli, and R. Rossi (eds.), *Marine Eutrophication and Population Dynamics*, pp. 77–84. Olsen & Olsen, Fredensborg, Denmark.

Wulff, F., J. G. Field, and K. H. Mann (eds.). 1989. *Network Analysis in Marine Ecology: Methods and Applications.* Springer-Verlag, Berlin.

9

A Modeling Approach to Elucidating the Distribution and Rates of Microbially Catalyzed Redox Reactions in Anoxic Groundwater

Derek R. Lovley and Francis H. Chapelle

1. Introduction

Microbially catalyzed redox reactions have an important influence on the chemical composition of many groundwaters. For example, carbon dioxide production during the oxidation of organic matter drives carbonate and silicate mineral dissolution in pristine carbonate aquifers and thus has a significant impact on water quality and secondary porosity (Chapelle 1993). Microbial reduction of Fe(III) to Fe(II) generates undesirably high concentrations of dissolved iron in aquifers (Lovley et al. 1990; Chapelle and Lovley 1992) and microbial sulfate reduction and methane production result in the accumulation of sulfide and methane (Thorstenson et al. 1979). The degradation of organic contaminants in polluted aquifers is a major mechanism for attenuating the transport of contaminants (Salanitro 1993; Lyngkilde and Christensen 1992; Baedecker et al. 1993). Thus, to better understand existing groundwater quality, and to predict the effect of perturbations on groundwater quality, it is necessary to have information on the types of microbial processes taking place in the subsurface and on the rates of these processes.

One way to study microbial process in aquifers is to sample the aquifer sediments and measure the rates of the processes of interest. However, this technique has several limitations. One is that it is generally expensive to sample the sediments of all but the shallowest of aquifers. Most measurements of microbial metabolism in incubated aquifer sediments are also expensive, labor intensive, and often technically difficult. Furthermore, as discussed later, activities measured in aquifer sediments incubated in the laboratory may bear little relationship to *in situ* activities.

We have attempted to develop techniques in which the types of microbial processes and their rates could be discerned through analyses of groundwater chemistry. Wells for sampling groundwater are often already in place in aquifers under investigation. It also is often easier to justify the cost of drilling for well installation than for sediment collection because once a well is installed, the groundwater can be sampled repeatedly over long periods of time, whereas sediments can only be collected at the time of drilling. Furthermore, chemical analyses of groundwater are generally simple and relatively inexpensive. Here we demonstrate how mathematical modeling of microbial metabolism and geochemistry have been used to elucidate the distribution and rates of microbial processes in groundwater.

2. Use of H_2 Concentrations to Predict Terminal Electron-Accepting Processes in Anoxic Groundwater

2.1. Model for H_2 Concentrations in Anoxic Sedimentary Environments

H_2 is an important intermediate in the decomposition of organic matter in many anoxic ecosystems. H_2 is produced from the fermentation of organic matter and in methanogenic systems, and is also generated in the oxidation of short-chain fatty acids. The principal sinks for H_2 in anoxic sedimentary environments are microorganisms that couple the oxidation of H_2 to the reduction of inorganic electron acceptors such as nitrate, Mn(IV), Fe(III), sulfate, and carbon dioxide. It is the competition for H_2 (as well as other substrates, most notably acetate) among the various H_2-consuming populations that, accounts for the often observed segregation of anaerobic processes in sedimentary environments (reviewed in Lovley and Chapelle 1995). When Fe(III) is readily available, Fe(III)-reducing microorganisms are able to maintain H_2 concentrations at levels too low for sulfate reducers or methanogens to metabolize. When Fe(III) is no longer available but sulfate is not limiting, sulfate reducers maintain H_2 too low for methanogens to metabolize. As sulfate levels become limiting sulfate reducers are no longer able to maintain H_2 below the minimum threshold necessary for methanogenesis and methane production become the predominant terminal electron accepting process.

A large number of studies (reviewed in Lovley and Goodwin 1988) on H_2 uptake by either natural assemblages or pure cultures of sulfate reducers and methanogens have indicated that, as long as the H_2 does not fall below a minimum threshold necessary for H_2 uptake, H_2 uptake can be modeled with Michaelis-Menten kinetics:

$$v = \frac{V_{max} \times [H_2] \times B}{K + [H_2]},\tag{1}$$

where v is the velocity of H_2 uptake, V_{max} is the maximum rate of H_2 uptake per unit of biomass, $[H_2]$ is the concentration of dissolved H_2, K is the H_2 concentration at which v is one half V_{max}, and B is the amount biomass of the H_2-consuming population.

The growth rate of a given H_2-consuming population can be modeled as

$$\frac{\partial B}{\partial t} = (v \times Y) - (b \times B),\tag{2}$$

where $\partial B/\partial t$ is change in the biomass of the H_2-consuming population over time; Y is the amount of biomass produced per H_2-consumed; and b is the cell decay coefficient, which includes all sources of loss of cellular material such as maintenance energy requirements, cell lysis or death, and grazing by predators.

It is assumed that in many anoxic sedimentary environments, H_2 production and consumption approach steady conditions such that the rates of H_2 production and consumption are equal and there is no net change in the size of the H_2-consuming population ($\partial B/\partial t = 0$). When $\partial B/\partial t$ is replaced with 0, Equations (1) and (2) can be combined and solved for $[H_2]$ with the following result:

$$H_2 = \frac{K}{(V_{max} \times Y/b) - 1} .\tag{3}$$

According to this model, under steady-state conditions, the concentration of dissolved H_2 in anoxic groundwater should be dependent solely on the physiological characteristics of the microorganisms consuming the H_2.

As previously discussed (Lovley and Goodwin 1988), V_{max} and b are expected to remain relatively constant, regardless of the microbial population consuming the H_2. However, the K and Y parameters are different for different types of H_2-consuming microorganisms. In general, the higher the potential energy yield from the H_2 oxidation the lower the K and the higher the Y (Lovley and Goodwin 1988). The potential energy yield for H_2 oxidation with the various electron acceptors in anoxic environments is: nitrate > Mn(IV) > Fe(III) > sulfate > methane and thus the expected pattern of H_2 concentrations in different types of sediments is nitrate-reducing < Mn(IV)-reducing < Fe(III)-reducing < sulfate-reducing < methanogenic (Lovley and Goodwin 1988). This pattern has consistently been observed in a wide variety of anoxic aquatic sediments (Lovley and Goodwin 1988) and ground waters ((Lovley et al. 1994; Vroblesky and Chapelle

1994; Chapelle et al. 1995) (Fig. 9.1). Remarkably, not only is the pattern of H_2 concentrations the same in different types of sediments, but also, the absolute values of dissolved H_2 for a given microbial process all fall within a narrow range. The constancy of H_2 concentrations for a given process in such a wide diversity of environments suggests that microorganisms with similar physiological characteristics for H_2 uptake are widespread.

It was previously suggested that further evidence supporting the hypothesis that the H_2-consuming microorganisms control the steady-state H_2 concentrations in sediments was that if the K, Y, and V_{max} values for pure cultures of H_2-consuming sulfate-reducers or methanogens were put into Equation (3), along with a b value obtained for H_2-consuming methanogens living in sediments, then the calculated steady-state H_2 concentrations for sulfate reduction and methanogenesis were comparable to the H_2 concentrations measured in sediments in which sulfate reduction or methanogenesis predominated (Lovley and Goodwin 1988). However, the Y values that were used in the previous calculations were 100-fold higher than the actual values because there was a printer's error in the
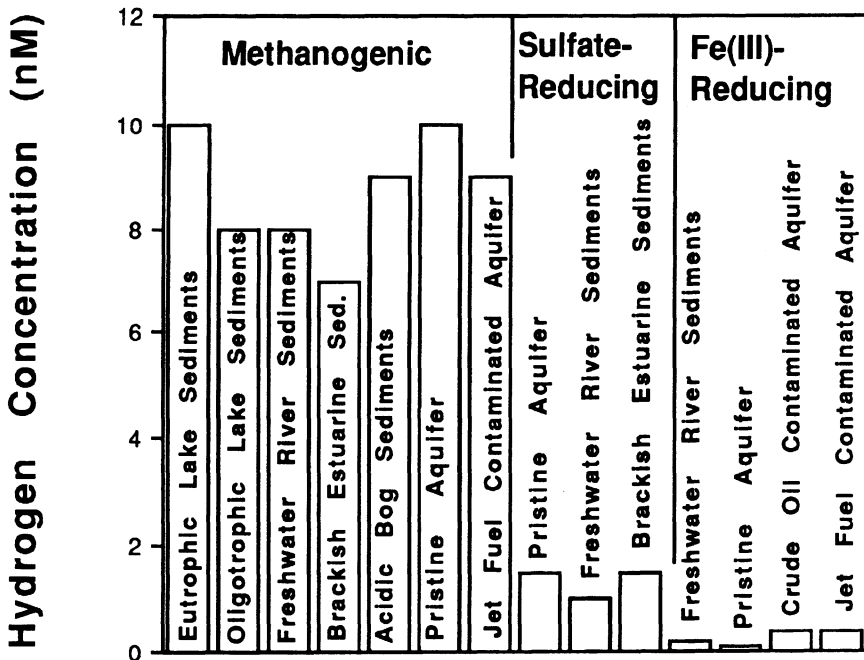
Figure 9.1 H_2 concentrations observed in aquatic sediments (Lovley and Goodwin 1988) and groundwater (Lovley et al. 1994) in which the stated terminal electron accepting processes predominated.

table that the Y values were taken from (Robinson and Tiedje 1984). Thus, using the correct Y values, the calculated H_2 values would be ~ 100-fold higher than those observed in sediments.

A subsequent study has indicated that the K estimates for that were used in the original calculations were also probably wrong (Giraldo-Gomes et al. 1992). When the K for methanogens was estimated by a novel technique designed to avoid mass transfer limitations, the estimated K was two orders of magnitude lower than the previous K estimates. This lower K value would then compensate for the error in the Y values in the original calculations and again give calculated H_2 concentration close to what has been observed in methanogenic aquatic sediments. However, the large difference in estimated K values with different methods emphasizes that extreme caution is required in trying to extrapolate from pure culture data to sediments. Calculations based on culture data may not be suitable to prove or disprove models of sedimentary metabolism.

2.2. Application of H_2 Model to Anoxic Ground Water

This simple model for H_2 metabolism not only explains the pattern of dissolved H_2 that is observed in anoxic sedimentary environments, but also provides a technique to determine which microbially catalyzed redox reactions predominate in groundwater. Historically, redox reactions in anoxic groundwater have been evaluated in terms of equilibrium thermodynamics and the master variable pe (Lindberg and Runnells 1984; Hostettler 1984; Fish 1993, and references therein). The hope was that with measurements of redox potential with electrodes or through calculations of redox potential from an analysis of groundwater chemistry it would be possible to define which redox reactions predominated in a given groundwater. This approach has proven not to be useful in practice primarily because groundwaters are rarely if ever at redox equilibrium and because redox electrodes do not respond to most of the important redox couples in groundwater (Lindberg and Runnells 1984; Hostettler 1984; Fish 1993, and references therein). However, with H_2 measurements it has been possible to reliably predict which anaerobic processes predominate at a given site.

An example of this was in a study of microbial metabolism in a shallow, petroleum-contaminated aquifer in Hanahan, South Carolina (Lovley et al. 1994). A large area of the aquifer downgradient from the source of contamination was anoxic. Methane production and Fe(III) reduction were important anaerobic processes in the anoxic zone. However, it was impossible to localize the zones of Fe(III) reduction and methane production with standard geochemical analyses of the groundwater because there were high concentrations of the products of Fe(III) reduction (Fe(II)) and the methane production (CH_4) even in zones where Fe(III) reduction or methane production were not taking place. In a study at two sites within the aquifer, measurements of dissolved H_2 concentrations in the

groundwater predicted that Fe(III) reduction predominated at one site, whereas methane production was the terminal electron accepting process at the other. Measurements of microbial processes in laboratory incubations of sediments from the two aquifers confirmed that the dissolved H_2 concentrations had accurately predicted the distribution of the microbial processes.

H_2 measurements can also be helpful in elucidating the distribution of anaerobic processes in deep pristine aquifers. In many such aquifers, there is an orderly succession of anaerobic processes that can be inferred through monitoring changes in the concentrations of substrates and/or products along the groundwater flowpath (see next section). However, a limitation of the geochemical modeling approach is that many of the reactants and products of important redox reactions reactions are insoluble (i.e. particulate organic matter, Fe(III) and Mn(IV) oxides, sulfate minerals, sulfide, and Fe(II) and Mn(II) minerals) or may diffuse in from the neighboring confining beds (i.e., sulfate and dissolved organic matter) (Chapelle and Lovley 1990, 1992; McMahon and Chapelle 1991; Plummer et al. 1990). Thus, numerous assumptions about which compounds are being consumed and produced are often required.

An example of this was a study of an anoxic reach of the Black Creek aquifer in South Carolina (Chapelle and McMahon 1991; McMahon and Chapelle 1991a). The accumulation of dissolved inorganic carbon along the ground water flow path indicated that organic carbon was being oxidized to carbon dioxide. However, there was not sufficient loss of potential electron acceptors or accumulation of reduced products to account for the amount of carbon dioxide produced. H_2 measurements predicted that the organic matter oxidation was coupled to sulfate reduction. Subsequent analysis of sediment cores demonstrated that sulfate to support sulfate reduction was diffusing into the aquifer from the surrounding confining beds and that the sulfide that was produced from sulfate reduction was precipitating as iron sulfide. Again, H_2 measurements accurately predicted the anaerobic process in an aquifer where standard geochemical analysis of ground water could not.

3. Estimating Rates of Microbial Processes with Geochemical Modeling

The limitation of the H_2 technique is that it can only specify which microbial processes predominate in a given section of an aquifer, it can not be used to determine the rates of microbial processes. However, it is often possible to estimate the rates of microbial processes in aquifers based solely on changes in water-chemistry changes along groundwater flowpaths. The basic equation describing water-chemistry changes as groundwater moves along a one-dimensional flowpath in an aquifer system may be written (Konikow 1977):

$$D\ \partial^2 C/\partial x^2 - v\ \partial C/\partial x - \text{CHEM}/E = \partial C/\partial t, \qquad (4)$$

where (when L is the flowpath length, T is time, and M is mass of the solute):

D = coefficient of hydrodynamic dispersion (L^2T^{-1}),
C = the concentration of a solute (ML^{-3}),
x = distance in the direction of groundwater flow (L),
v = velocity of groundwater flow (LT^{-1});
CHEM = chemical/biological reaction source (+) or sink (−)
of solute C per unit volume of aquifer ($ML^{-3}T^{-1}$),
E = aquifer porosity (dimensionless).

Equation (4) is a general statement of the principle of the conservation of mass and is applicable to all aquifer systems. The first term on the left side of Equation (4) incorporates the effects of hydrodynamic dispersion of a solute as it moves with flowing ground water. The second term on the left side of Equation (4) accounts for the effects of connective mixing on solute concentrations. The CHEM term is the summation of all chemical and/or biological reactions that serve either as a source or a sink of the solute and, depending on the scale of the system and types of reactions under consideration, can be formulated in a variety of ways, as illustrated later.

3.1. Rate of Organic Matter Oxidation in a Regional Aquifer System

If groundwater flow and solute transport is considered on a regional scale (flowpath lengths on the order of tens of kilometers), the influence of hydrodynamic dispersion (term 1) and convective mixing (term 2) on overall solute concentrations are often negligible and Equation (4) can be simplified as

$$\text{CHEM} = \partial C/\partial t. \qquad (5)$$

Mathematically rigorous methods (Plummer et al. 1983) have been developed to identify operative chemical/biological processes in groundwater systems and to quantify amounts of mass transfer into or out of solution. These methods are collectively known as geochemical modeling. Geochemical modeling is based on the principle of mass conservation in which changes in water chemistry along a flowpath can be explained as:

$$\text{Initial water composition + Reactants} \qquad (6)$$
$$= \text{Final water composition + Products.}$$

This principle can be mathematically expressed as (Plummer et al. 1983)

$$\left(\sum_{p=1}^{p} \infty_p b_{p,k} = \Delta m_{\text{Tot},k} \right), \ k = 1, j, \tag{7}$$

Which states that the change in total moles of a particular element k, $(\Delta m_{\text{Tot},k})$ is equal to the sum of all the sources and sinks for element k. This relationship holds for 1 through j elements. The sources and sinks may include dissolution, precipitation, microbial degradation, gas transfer, and so forth, for P phases (i.e., minerals or gases) along the flowpath where ∞ is the number of moles reacting and b is the stoichiometric coefficient for the element in the pth phase. In practice, $\Delta_{\text{Tot},k}$ values are derived directly from groundwater chemistry data using equations of the form

$$\Delta_{\text{Tot},k} = m_k \ (\text{initial water}) - m_k \ (\text{final water}). \tag{8}$$

An example of an application of this approach to microbial processes was the attempt to estimate rates of organic matter oxidation in three aquifers in a regional groundwater system in the Atlantic Coastal Plain of South Carolina (Chapelle and Lovley 1990). In this system, changes in groundwater chemistry were dominated by microbial CO_2 production which drove the dissolution of magnesium-bearing carbonate shell material in the aquifers:

$$CO_2 + Ca_xMg_{1-x}CO_3 + H_2O \longrightarrow Ca_x^{2+} + Mg_{1-x}^{2+} + 2HCO_3^-. \tag{9}$$
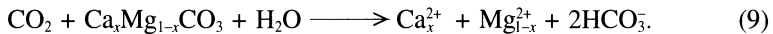
The calcium and magnesium liberated by this reaction could then exchange with sorbed sodium:

$$Ca^{2+} + Na_2(\text{clay}) \longrightarrow Ca(\text{clay}) + 2Na^+, \tag{10}$$
$$Mg^{2+} + Na_2(\text{clay}) \longrightarrow Mg(\text{clay}) + 2Na^+. \tag{11}$$
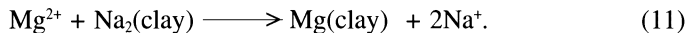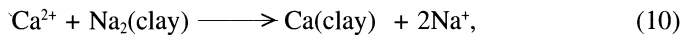
An added consideration was the fact that residual seawater present in the aquifer mixed with the incoming freshwater and served to increase concentrations of sodium and chloride.

From this information, mass balance equations that account for the observed water chemistry changes along aquifer flowpaths could be formulated. For example, the change in dissolved inorganic carbon (ΔM_C) is equal to the sum of carbon dioxide produced by microbial metabolism (m_{CO_2}) and the carbonate shell material (m_{sm}) dissolved:

$$\Delta M_C = m_{CO_2} + m_{sm} \tag{12}$$

The carbonate shell material contained 98 mol % calcium and 2 mol% magnesium. Thus, the change in dissolved calcium along each flowpath segment (ΔM_{Ca}) was equal to $0.98m_{sm}$ minus the amount of calcium that was removed from solution by sorption onto clays ($m_{Ca(clay)}$):

$$\Delta M_{Ca} = 0.98m_{sm} - m_{Ca(clay)}, \tag{13}$$

and the change in dissolved magnesium (ΔM_{Mg}) was

$$\Delta M_{Mg} = 0.98m_{sm} - m_{Mg(clay)}. \tag{14}$$

The change in sodium (ΔM_{Na}) along each flowpath segment was set as the amount of sodium resulting from calcium-sodium ($2m_{Ca(clay)}$) and magnesium-sodium ($2m_{Mg(clay)}$) exchange as well as sodium coming from the residual seawater (m_{sw}):

$$\Delta M_{Na} = m_{sw} + 2m_{Ca(clay)} + 2m_{Mg(clay)}. \tag{15}$$

The amount of chloride from the seawater (ΔM_{Cl}) was represented as:

$$\Delta M_{Cl} = m_{sw}. \tag{16}$$

These five mass balance equations (Equations [12] to [16]), with five unknowns, could be solved algebraically to determine the mass of carbon dioxide produced by microbial metabolism:

$$m_{CO_2} = \Delta M_C - \Delta M_{Ca} - \Delta M_{Mg} - (\Delta M_{Na} - \Delta M_{Cl})/2. \tag{17}$$

However, computer programs designed to formulate and solve these kinds of mass-balance problems in geochemical modeling are presently available (Parkhurst et al. 1982) and make the procedure much easier.

Once the time for groundwater to travel between two points along a groundwater flowpath is known then a rate of carbon dioxide production can be calculated. The approximate velocity of groundwater flow is given by Darcy's equation:

$$v = K(dh/dL)/E, \tag{18}$$

where

$$v = \text{average ground-water velocity (LT}^{-1}\text{)},$$
$$K = \text{hydraulic conductivity (LT}^{-1}\text{)},$$
$$dh/dl = \text{hydraulic gradient (dimensionless)},$$
$$E = \text{is aquifer porosity (dimensionless)}.$$

Rates of microbial carbon dioxide production (R-CO$_2$) could be calculated as:

$$R - CO_2 = (m_{CO_2} \times v)/L. \tag{19}$$

Using this method, overall rates of microbial carbon dioxide production in the Black Creek, Middendorf, and Cape Fear aquifers were estimated to be on the order of 10^{-5} mmoles of CO$_2$ per liter of groundwater per year. In contrast, depending on the incubation technique, laboratory incubations of sediment material from these aquifers yielded estimated rates of organic matter decomposition that were 10^2 to 10^6 times faster (Chapelle and Lovley 1990). Considerations of the age of the aquifer sediments and the amount and microbial availability of organic mater in the sediments and groundwater clearly indicate that the high estimates from laboratory incubations can not be correct and that the geochemical modeling technique must more closely approximate *in situ* rates of organic matter decomposition (Chapelle and Lovley 1990). The only other environment known to have such low rates of microbial metabolism is the cold bottom waters of deep ocean basins (Williams and Carlucci 1976).

3.2. Compound-Specific Rates of Contaminant Biodegradation

Geochemical modeling can also be used to estimate the rates of microbial decomposition of specific organic compounds in groundwater. In the simplest case, where degradation of the organic can be assumed to follow first-order kinetics, the CHEM term in Equation (4) can be replaced by:

$$CHEM/E = k_d C, \tag{20}$$

where k_d is the first-order rate constant (T^{-1}). Most studies on degradation of organic compounds are concerned with the local scale, where groundwater flowpaths are considered over distances of tens or hundreds of meters. In these instances, the effects of hydrodynamic dispersion and convective mixing are relatively important in solute transport and they must be included in Equation (4). Thus, the equation for the first-order decomposition of an organic compound on the local level can be expressed as

$$D\, \partial^2 C/\partial x^2 - v\, \partial C/\partial x - k_d C = \partial C/\partial t. \tag{21}$$

Equation (21) can be simplified for the special case where concentration changes approach steady-state conditions. At steady state, Equation (20) simplifies to an ordinary differential equation,

$$D \, d^2C/dx^2 - v \, dC/dx - k_d C = 0, \tag{22}$$

that is easily solved given appropriate boundary conditions. For boundary conditions of $C = C_0$ at $x = 0$, and $C \longrightarrow 0$ as $x \longrightarrow \infty$, the solution of Equation (22) is

$$C(x) = C_0 \exp^{\{[v - (v^2 + 4Dk_d)^{1/2}]/2D\}x} . \tag{23}$$

If a hydrologic system has been well-characterized, so that reasonable estimates can be made for the coefficient of hydrodynamic dispersion (D) and groundwater velocity (v), and if water-chemistry data are available to describe concentrations of C in time and space, then estimates of k_d can be obtained by curve-fitting calculated concentration changes with measured concentration changes. As an example, this method can be used to estimate rates of toluene degradation in the contaminated shallow aquifer located in Hanahan, South Carolina, that was mentioned earlier in connection with H_2 measurements. The aquifer was polluted by gasoline and jet fuel spills in the 1960s, and groundwater contaminated with benzene, toluene, and xylene has emanated from the site ever since. Because contaminated groundwater has been transported from the site for such a long period of time, the steady-state assumption is a reasonable approximation. Extensive hydrologic testing, including tracer tests estimates of hydrodynamic dispersion, indicate that D and v are approximately 0.038 ft²/day and 1 ft/day. Using these parameters, it is evident that, allowing for inevitable variability in the data, apparent biodegradation rate constants for toluene degradation are bracketed by k_d values of about 0.01 and 0.005 day^{-1} (Fig. 9.2).

Studies in which the rates of oxidation of [ring-^{14}C]-toluene to $^{14}CO_2$ were monitored in laboratory incubations of sediments from this site indicated that under the sulfate-reducing conditions that predominate in the zone modeled, the k_d values for toluene oxidation were on the order of 0.009 day^{-1} (F. H. Chapelle and P. M. Bradley, unpublished data). Thus, in contrast to the deep subsurface studies in which there was poor agreement between estimates from laboratory incubations and geochemical modeling, laboratory estimates and geochemical modeling gave similar rate estimates in this shallow, contaminated aquifer.

4. Conclusions

These studies have demonstrated how relatively simple modeling of microbial processes can aid in evaluating the geochemistry of groundwater and how modeling of geochemistry can help in elucidating microbial processes. The combined approach of using dissolved H_2 concentrations to define which microbially cata-
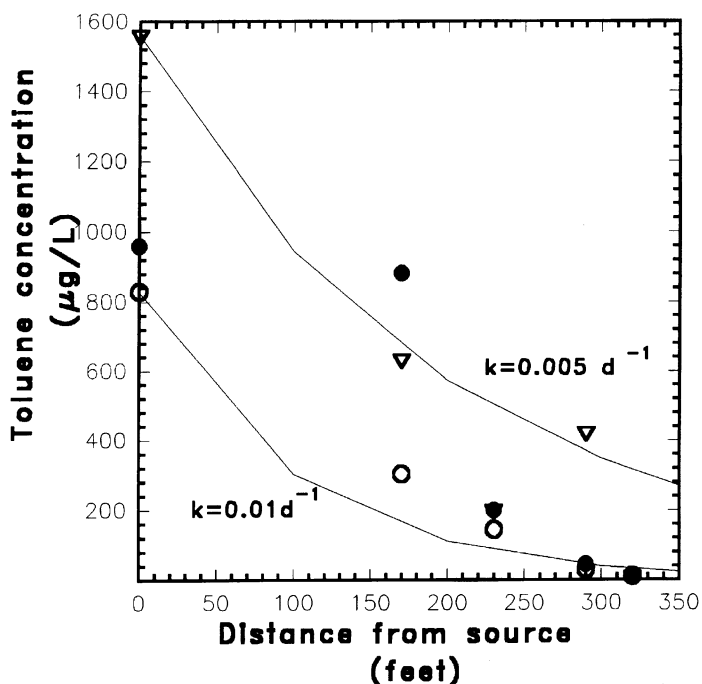
Figure 9.2 Concentrations of toluene near the Hanahan, South Carolina, site measured in July 1992 (○), November 1992 (△), and June 1993 (●) superimposed on solutions of equation 23 for k_d values of 0.01 and 0.005 day^{-1} and boundary conditions of 830 and 1560 µg/L, respectively.

lyzed redox process are taking place in subsurface environments and using geo-chemical modeling to estimate the rates of the microbial processes permits detailed examination of microbial metabolism without sampling the aquifer sediments. Since sampling aquifer sediments is generally much more costly and requires much more technical expertise than sampling groundwater and since, at least in some instances, results from laboratory incubations of aquifer sediments fail to reflect *in situ* conditions, whenever possible, a modeling approach to the study of microbial processes in groundwater should be considered.

References

Baedecker, M. J., I. M. Cozzarelli, D. I. Siegel, P. C. Bennett, and R. P. Eganhouse. 1993. Crude oil in a shallow sand and gravel aquifer: 3. Biogeochemical reactions and mass balance modeling in anoxic ground water. *Appl. Geochem.* 8:569–586.

Chapelle, F. H. 1993. *Ground-water microbiology and geochemistry.* John Wiley & Sons, New York.

Chapelle, F. H., and D. R. Lovley. 1990. Rates of microbial metabolism in deep coastal plain aquifers. *Appl. Environ. Microbiol.* 56:1865–1874.

Chapelle, F. H., and D. R. Lovley. 1992. Competitive exclusion of sulfate reduction by Fe(III)-reducing bacteria: A mechanism for producing discrete zones of high-iron ground water. *Ground Water* 30:29–36.

Chapelle, F. H., and P. B. McMahon. 1991. Geochemistry of dissolved inorganic carbon in a coastal-plain aquifer: 1. Sulfate from confining beds as an oxidant in microbial CO_2 production. *J. Hydrol.* 127:85–108.

Chapelle, F. H., et al. 1995. Deducing the distribution of terminal electron-accepting processes in hydrologically diverse ground water systems. *Water Resour. Res.* 31: 359–371.

Fish, W. 1993. Sub-surface redox chemistry: A comparison of equilibrium and reaction-based approaches. In H. E. Allen, E. M. Perdue, and D. S. Brown (eds.), *Metals in Groundwater,* pp. 73–101. Lewis Publishers, Ann Arbor, MI.

Giraldo-Gomez, E., S. Goodwin, and M. S. Switzenbaum. 1992. Influence of mass transfer limitations on determination of the half saturation constant for hydrogen uptake in a mixed culture CH_4-producing enrichment. *Biotech. Bioeng.* 40:768–776.

Hostettler, J. D. 1984. Electrode electrons, aqueous electrons, and redox potentials in natural waters. *Am. J. Sci.* 284:734–759.

Konikow, L. F., and D. B. Grove. 1977. Derivation of equations describing solute transport and dispersion in ground water. *U.S. Geo. Surv. Water Res. Invest.* 77:19–30.

Lindberg, R. D., and D. D. Runnells. 1984. Ground water redox reactions: an analysis of equilibrium state applied to Eh measurements and geochemical modeling. *Science* 225:925–927.

Lovley, D. R., and F. H. Chapelle. 1995. Deep subsurface. *Microbial. Processes. Rev. Geophsy.* 33:365–381.

Lovley, D. R., F. H. Chapelle, and E. J. P. Phillips. 1990. Fe(III)-reducing bacteria in deeply buried sediments of the Atlantic Coastal Plain. *Geology* 18:954–957.

Lovley, D. R., F. H. Chapelle, and J. C. Woodward. 1994. Use of dissolved H_2 concentrations to determine the distribution of microbially catalyzed redox reactions in anoxic ground water. *Environ. Sci. Technol.* 28:1025–1210.

Lovley, D. R., and S. Goodwin. 1988. Hydrogen concentrations as an indicator of the predominant terminal electron accepting reactions in aquatic sediments. *Geochim. Cosmochim. Acta* 52:2993–3003.

Lyngkilde, J., and T. H. Christensen. 1992. Fate of organic contaminants in the redox zones of a landfill leachate pollution plume (Vejen, Denmark). *J. Contamin. Hydrol.* 10:291–307.

McMahon, P. B., and F. H. Chapelle. 1991a. Geochemistry of dissolved inorganic carbon in a coastal plain aquifer. 2. Modeling carbon sources, sinks, and $\partial^{13}C$ evolution. *J. Hydrol.* 127:109–135.

McMahon, P. B., and F. H. Chapelle. 1991b. Microbial production of organic acids in aquitard sediments and its role in aquifer geochemistry. *Nature* 349:233–235.

Parkhurst, D. L., L. N. Plummer, and D. C. Thorstenson. BALANCE—A Computer Program for Calculating Mass Transfer for Geochemical Reactions in Groundwater. U. S. Geological Survey Water Resources Investigation Report 82-14. 1982. U. S. Geological Survey, Reston, VA.

Plummer, L. N., J. F. Busby, R. W. Lee, and B. B. Hanshaw. 1990. Geochemical modeling of the Madison Aquifer in parts of Montana, Wyoming, and South Dakota. *Water Resour. Res.* 26:1981–2014.

Plummer, L. N., D. L. Parkhurst, and D. C. Thorstenson. 1983. Development of reaction models for ground-water systems. *Geochim. Cosmochim. Acta* 47:665–686.

Robinson, J. A., and J. M. Tiedje. 1984. Competition between sulfate-reducing and methanogenic bacteria for H_2 under resting and growing conditions. *Arch. Microbiol.* 137:26–32.

Salanitro, J. P. 1993. The role of bioattenuation in the management of aromatic hydrocarbon plumes in aquifers. *Ground Water Monitor. Remed.* 13:150–161.

Thorstenson, D. C., D. W. Fisher, and M. G. Croft. 1979. The geochemistry of the Fox Hills-Basal Hell Creek aquifer in southwestern North Dakota and northwestern South Dakota. *Water Resour. Res.* 15:1479–1498.

Vroblesky, D. A., and F. H. Chapelle. 1994. Temporal and spatial changes in terminal electron accepting processes in a petroleum hydrocarbon contaminated aquifer and the significance for contaminant biodegradation. *Water Resour. Res.* 30:1561–1570.

Williams, P. M., and A. F. Carlucci. 1976. Bacterial utilization of organic matter in the deep sea. *Nature* 262:810–811.

10

From the Ground Up: The Development and Demonstrated Utility of the Ruminal Ecosystem Model

R. L. Baldwin and K. C. Donovan

1. Introduction

Our understanding of rumen digestion and microbial ecology has developed in several stages. The first was recognition in the 1940s of the fact that the rumen microbes play a major role in the digestive process, producing volatile fatty acids, and that they are a major source of nutrients to the ruminant animal (Hungate 1966). This recognition was followed by studies that led to the isolation, taxonomic identification, and enumeration of microbes present in the rumen (Bryant 1959). These studies were followed by studies of pure cultures of rumen microbes in which their physiological properties and metabolism were defined. In addition, interactions among rumen microbes were evaluated. This area of inquiry is not complete, even though activity has been sporadic and limited during recent years. The final phase of our evaluation of the rumen as an ecosystem requires quantitative and dynamic analyses of the extent to which microbes express their characteristics in mixed cultures and in vivo. To achieve this goal we must understand microbial interdependence and competition in the rumen (Russell and Hespell 1981). Progress toward this goal has been limited and may require several decades for adequate resolution. With the exception of the model of Russell and Allen (1984) of the behavior of *Streptococcus bovis* and *Ruminococcus albus* in coculture, no dynamic, mechanistic models of interactions among rumen microbes have been published. This may well be because of a lack of interest in modeling the system, a lack of required data, a lack of opportunities to undertake the coupled modeling and experimental analyses required to make progress, or all of these. Whatever the case, this limitation clearly limits the range of models of ruminant digestion to those that consider the rumen microbes in aggregate or as

functional groups. Discussions of models of rumen function in this chapter reflect this reality.

2. Balance Models of Rumen Digestion

Most rumen microbiologists and ruminant nutritionists have used balance models in teaching and writing (Wolin 1960; Hungate 1966). Examples of balance models frequently used in the past are presented in Table 10.1. These simplified models emphasize the requirement for maintenance of redox balance in anaerobic systems; illustrate the varied metabolic functions of rumen microbes and interactions among microbes; and, in part, demonstrate relationships among fermentation patterns and ATP yields. Although these simple balance models illustrate underlying metabolic considerations, they are misleading. For example, the high estimates of carbohydrate energy lost as methane energy in examples II and III (Table 10.1) are inflated because carbohydrate and reducing equivalents used in support of microbial growth are not considered. This observation led to the view that more complex models must be used in quantitative evaluations of ruminant digestion. Such models must accommodate the fermentation of the several classes of carbohydrates by different microbes; fermentation of nutrients other than carbohydrates; the use of carbon, nitrogen, and reducing equivalents for microbial cell growth; and passage to and digestion of nutrients in the lower gut. This need led Reichl and Baldwin (1975, 1976) to develop balance models of ruminant digestion.

Two balance models were constructed: an elemental balance model and a metabolic pathways balance model. These were, initially, constructed to provide convenience and accuracy in calculating rumen balance relationships from complete and partial fermentation data. The stoichiometric relationship developed in these models have been used in many subsequent models. The Reichl and Baldwin (1975, 1976) models were designed to aid investigators in evaluations of experimental data for internal consistency in terms of essential balance relationships. Fermentation data had to be input to these models. These were not predictive models.

Murphy et al. (1982) developed an analytical model to estimate proportions of fermentation products formed from specific chemical components of ruminant diets based on the molar proportions of acetate, propionate, butyrate, and valerate produced during fermentation of various diets. A premise in this study was that the eventual development of predictive models of rumen fermentation will require the availability of estimates of the proportions of products formed from specific chemical entities in the diet. Two sets of data were compiled from the literature. The first set of data was from 20 different publications on 137 animals fed forage

Table 10.1 Three early rumen balance models

I. Glucose conversion to acetate and propionate:

1.5 glucose + 3ADP → 3 pyruvate + $6H^+$ + $6e^-$ + 3ATP

1 pyruvate + ADP → CO_2 + acetate + $2e^-$ + $2H^+$ + ATP

2 pyruvate + $8H^+$ + $8e^-$ + [2ADP] → 2 propionate [+ 2ATP][a]

1.5 glucose + 6ADP → CO_2 + acetate + 2 propionate + 6ATP
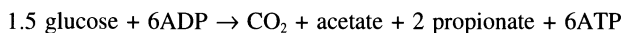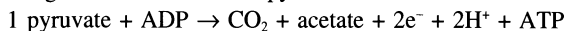
$$\text{Efficiency} = \frac{2 \text{ propionate} + 1 \text{ acetate}}{1.5 \text{ glucose}} * 100 = 94\%$$

Energy loss as heat = 6.0% (does not exclude energy trapped in ATP)

II. Glucose conversion to acetate and methane

Organism A

polysaccharide → 1.5 glucose-P

1.5 glucose-P + 4.5ADP → 3 pyruvate + $6H^+$ + $6e^-$ + 4.5ATP

3 pyruvate + $6H^+$ + $6e^-$ → 3 lactate

Organism B

3 lactate + 3ADP → 3 acetate $3CO_2$ + $6H_2$ + 3ATP

Organism C

$6H_2$ + $1.5CO_2$ + 1.5ADP → $1.5CH_4$ + $3H_2O$ + 1.5ADP

1.5 glucose units + 9ADP → 3 acetate + 1.5 CO_2 + $1.5CH_4$ + $3H_2O$
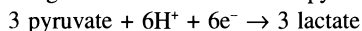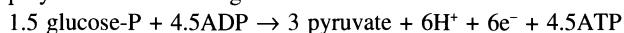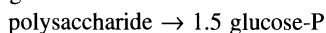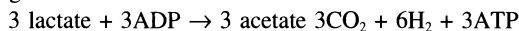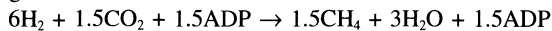
$$\text{Efficiency} = \frac{3 \text{ acetate}}{1.5 \text{ glucose}} * 100 = 61\%$$

Methane loss = 30.5%   Heat loss = 8.5%

III. Standard rumen fermentation[b]

100 glucose units → 107 acetate + 37 propionate + 28 butyrate + 59 methane + heat

Net efficiency of fermentation = 75%

Methane loss = 18%

Net ATP yield[a] (depending on microbes considered) = 3.65 − 4.58

Heat loss = 7%

[a]ATP yields can vary depending on microbe and metabolic pathway (Baldwin et al. 1970).
[b]Based on Hungate (1966).

diets. The second set of data was from studies of 48, mostly concentrate, diets fed to 374 experimental animals.

An arbitrary set of initial parameter values defining proportions of each substrate converted to each fermentation product were input to initiate the iterative, parameter estimation process. In sequential solutions, the parameter values were adjusted systematically using the simplex and steepest descent algorithms to minimize the weighted sums of squares of residual errors of prediction. Weightings of experimental data were based on the assumptions that measurements on a number of animals are more likely to be representative than are those based upon one animal. Standard errors of the final parameter values were estimated using the jackknife procedure.

Estimated fermentation parameters for both the roughage (R) and concentrate (C) data sets are presented in Table 10.2. Evaluations of the uniqueness of parameters deduced using complex analytical models is difficult. Qualitative evaluations based on whether or not the parameter estimates are compatible with known fermentation patterns and diet-induced shifts in the relative numbers of major rumen microbes were satisfactory.

A second set of evaluations involved comparisons of estimated parameter values with data derived from in vitro incubations of rumen microbes with soluble ^{14}C-labeled substrates. These evaluations also indicated that the deduced parameter values reflect reality. Additional evaluations, which are not repeated here, suggested that parameter values deduced for acetate and propionate formation from the several substrates were unique and correct. Parameter values defining protein (amino acid) fermentation and butyrate fermentation were found to be suspect.

Russell et al. (1992) described the rumen element of the Cornell Net Carbohydrate and Protein System (CNCPS; Search:Agriculture, 1990) for evaluating cattle diets and predicting energy and amino acids available to the ruminant animal. In this system, inputs required to predict microbial yields are feed carbohydrate, which is divided into nonstructural carbohydrate (NSC) that includes sugars, pectin, and starch; structural carbohydrate (SC), which includes hemicellulose, cellulose, and lignin; and feed protein. Feed crude protein is partitioned into four fractions: nonprotein nitrogen; rapidly degraded soluble true protein; slowly degraded true protein; and unavailable protein. Relative amounts of each component of each feed degraded in or passed from the rumen are calculated from values tabulated for each feed ingredient. All nutrients in each feed ingredient are either digested in or passed from the rumen.

Net microbial growth yields are dependent on the amount of NSC and SC digested in the rumen; a theoretical maximum growth yield, Y_g, which is defined as maximum grams of bacteria formed per gram of carbohydrate fermented in the rumen in the absence of maintenance; a maintenance requirement; and the input of hemicellulose plus cellulose and lignin (neutral detergent fiber [NDF])

from forage in the diet. Thus, the Cornell system for microbial growth accommo-
dates variation due to composition of feed ingredients and their digestibilities in
the rumen, as well as microbial growth yields dependent on relative rates of
fermentation, growth, maintenance energy expenditures, and availabilities of
amino acids and peptides (Search:Agriculture, 1990). In our view, such accommo-
dations are essential to adequate predictions of amino acid availabilities to rumi-
nant animals. Thus, these criteria are emphasized in the evaluations of dynamic
models in the next section of this chapter.

3. Dynamic Models of Ruminant Digestion

   Most dynamic models discussed in this section were constructed to evaluate
current data and concepts regarding ruminant digestion and metabolism and
identify critical research needs. As a result, direct comparisons with the static
models discussed in the previous section are not always relevant, because these
models were often constructed with different goals, including prediction, in mind.
In many cases, the development of dynamic models emphasized the concept that
model evolution should proceed from the simple to the complex, with increasing
complexity forced by failures of the model. The examples used herein were
selected to illustrate this concept and to represent the evolution of dynamic
models of rumen digestion supported by an interplay between experimental and
modeling research.

4. Early Dynamic Models

   The first detailed mechanistic and dynamic model of ruminant digestion we
are familiar with was published by Baldwin et al. (1970). The stated modeling
objectives were to illustrate the utility and flexibility of using mathematical
models in support of research. Important results of the study were an updating
of stoichiometric relationships for fermentation and microbial growth and recogni-
tion of the fact that a number of microbial interactions have significant impacts
on quantitative aspects of ruminant digestion. A number of provisions in the
model were quite similar to those in some current models.
   Three composite microbial groups were defined: a cellulolytic group (A),
which encompassed known fermentation and growth properties of *Ruminococcus
flavefaciens*, *R. albus,* and *Bacteroides* (now *Fibrobacter*) *succinogenes*, plus
methanogenic species; an amylolytic complex (B); and a complex (C), which
included species with weak cellulolytic and very active saccharolytic activities
such as those associated with *Bacteroides ruminicola* and several *Butyrivibrio*
spp. Group A microbes utilized ammonia as a nitrogen source for growth, whereas

groups B and C used ammonia and amino acids. Limitations in nitrogen availability could disrupt microbial growth. Group C microbes compete for hydrolytic products from holocellulose. This accommodated the presence of high numbers of group C microbes on high-fiber diets. Based on the view that rumen microbes utilize and ferment chemicals rather than conventionally defined fiber, protein, and nitrogen-free extract, description of dietary inputs to the model in chemical terms was considered necessary. Potentially digestible holocellulose, soluble carbohydrate, starch, protein, and nonprotein nitrogen expressed as urea were chosen to describe dietary input. Based on application of similar reasoning to the animal, rates of production of individual volatile fatty acids (VFA) were traced. Insoluble substrates were separated into large (slow passage rate) and small (readily passed) particle pools. Equations in the model were mass action in form and parameter values required for solution of the model were calculated from steady-state (continuous feeding) data. A number of additional factors such as rates of protein and starch solubilization were considered but were not incorporated because of a lack of data. The already mentioned provisions are quite compatible with those discussed in the previous section with regard to the Cornell system. The major difference is that dynamic data can be used to challenge the dynamic model but not the static Cornell model. Failures of the dynamic model to stimulate dynamic data provided insights regarding types of data required to formulate subsequent models to better simulate reality.

Mazinov and Nolan (1976) and Nolan (1975) developed and illustrated the use of a dynamic, analytical model of nitrogen metabolism in sheep. Their basic objective was to enable the interpretation and integration of ^{15}N isotope dilution experiments, in which tracers were introduced at different sites. The submodel of nitrogen metabolism in the reticulo-rumen, omasum, and abomasum was premised on the view that quantitative assessments of the following are necessary (Nolan 1975): the quantities of soluble and insoluble protein and nonprotein N from the diet and of endogenous urea and other endogenous N compounds, and their contribution to the total pool of N compounds available for fermentation in the rumen; the extent to which these N compounds are degraded to simpler compounds (i.e., peptides, amino acids, and ammonia) or pass undegraded from the rumen; the extent to which peptides, amino acids, ammonia, and nucleic acids are assimilated by bacteria and protozoa, and the net rate of efflux of microbial N to the small intestine; the quantities of soluble N compounds absorbed through the walls of the forestomach, or their efflux in water; and the quantities of N recycled through pools within the rumen itself (e.g., as a result of ingestion of other microorganisms by protozoa, and lysis of bacteria).

Further studies indicated that nitrogen exchange required consideration of transactions in the lower digestive tract and in the animal. Critical elements of the model included separation of diet inputs into protein and Non Protein nitrogen (NPN); explicit consideration of amino acid as well as ammonia use for microbial

growth; nitrogen digestion and secretion into the gastrointestinal tract (GIT); urea recycling to the GIT; turnover of extracellular and intracellular amino acids, as well as body protein; and N excretion in urine and feces. Equations used in the model were essentially mass action in form with lags to allow the fitting of experimental data. A very important aspect of this work was that the experimental approach of administering ^{15}N compounds at multiple sites enabled "identification" of components of the model and computations of flux relationships among pools.

Baldwin et al. (1977) presented a research model of ruminant digestion for use in evaluating factors that influence nutritive value. Specific objectives were development of a quantitative and dynamic model of the ruminant digestive process based on currently accepted and defensible concepts; identification of specific aspects of ruminant digestion where current generally accepted concepts and/or data are inadequate; and development of a model that could be used to test hypotheses regarding factors affecting nutritive value. Baldwin et al. (1977) stated the belief that a model suitable for these purposes should be comprised of mechanistic rather than empirical equations and satisfy the following criteria: applicability to all ruminant diets and feeding regimens with no changes in basic functional forms (equations), rate constants, or parameter values other than those required to describe feed input and the physiological status of the animal; accommodation of current concepts regarding physical, chemical, biochemical, physiological, and microbiological factors that affect nutritive value and interactions amongst these; manageability in terms of size, i.e., fulfill the preceding objectives of model development without incorporating excessive detail; and provide outputs in a form comparable with a wide range of experimental data. Their model incorporated the following concepts: (1) the relative rates of fermentation in and passage from the rumen are central determinants of extent of digestion in the rumen, partition of digestion between the stomach and intestine, and amounts and patterns of nutrients absorbed by the animal; (2) both the physical and chemical properties of feedstuffs influence rates and patterns of rumen availability, utilization, and passage of nutrients; (3) microbes fulfill an essential, integrative, and variable role in ruminant digestion; (4) interactions among animals and their digestive processes strongly influence digestion rates and patterns; (5) estimates of most rate constants and other essential numerical inputs can be obtained or deduced from experimental data with animals fed frequently (so as to simulate a steady-state system); and (6) stoichiometric relationships in rumen fermentation can be calculated based on available knowledge of metabolic pathways. The bases for these concepts are self-evident and have been discussed in many comprehensive reviews (Hungate 1966; Phillipson 1970; McDonald and Warner 1975; Russell and Hespell 1981; Baldwin and Allison 1983; Dobson and Dobson 1988).

The first three concepts simply suggest that rumen microbial functions should be emphasized in the model. The fourth concept requires that animal functions such as rumination, salivation, etc. be represented. The fifth concept essentially

restricts the data set used in model development to continuously fed sheep (actually fed only one diet, alfalfa, at one intake). This allows all other diets, intakes, and feeding patterns to be used in model evaluations. The final concept accepts stoichiometric coefficients from rumen balance models discussed in the previous section.

The basic block diagrams used to describe the digestion of soluble nutrients, insoluble, rapidly available nutrients, and insoluble, particulate material varied in complexity.

Several concepts regarding microbial growth were introduced into this model. The equation for microbial growth was given as the commonly known function of ATP availability and yield of microbial cells/mole ATP (YATP). Additionally, the concept of ATP use for microbial maintenance (ATPM) was introduced. Also, dependence of microbial growth rates and yields on amino acid availability derived from Maeng and Baldwin (1976) and Maeng et al. (1976) were incorporated. Third, an equation describing a limitation of microbial growth by low ammonia concentrations was formulated, based on the data of Hogan (1975).

Equations for nutrient transactions in the Baldwin et al. (1977) model were all mass action in nature. Most parameter values incorporated into the model were estimated using a static model based on data from studies of 40-kg sheep fed 910 g/day of an alfalfa diet in equal amounts at hourly intervals. Although available data were not adequate to fully parameterize the model (overparameterized), input:output data were good and considered to be an adequate reference base for simulations of the steady-state condition used in initial stages of model evaluation (debugging etc.). Numerical inputs descriptive of stoichiometric relationships of fermentation and microbial growth were similar to those reported by Baldwin et al. (1970), Reichl and Baldwin (1975), and Koong et al. (1975). Given the constraints inherent to the use of a very specific data set to parameterize the model and the almost exclusive use of mechanistic elements, evaluations of the model and the concepts that comprise it were quite promising (Baldwin et al. 1977; Murphy et al. 1986). However, a number of specific failures of the model, including an inability to simulate the digestion of poor (straw) and very high (concentrate) quality diets, helped identify several critical limitations in then current data and concepts. These included water dynamics including salivation; factors involved in particle size reduction and passage; effects of diet and pH on patterns of volatile fatty acids (VFA) formed from the several chemical constituents in the diet; and nitrogen metabolism in the rumen and exchange between the rumen and the animal (Argyle and Baldwin, 1988).

Several improvements were incorporated into the model by Murphy et al. (1986). These were undertaken to correct specific omissions and weaknesses in the original model and were based, largely, on experimental data collected in response to conceptual and data problems identified in the course of the first modeling analysis.

A major problem with the initial model was correct simulation of rumen

ammonia concentrations across diets and feeding frequencies. Rates of insoluble protein hydrolysis and degradation were the most sensitive variable affecting rumen ammonia nitrogen. Representation of this was improved by explicit recognition of large and small particle pools for insoluble protein and addition of a provision that only microbes closely associated with small particles hydrolyze insoluble protein. This allowed accommodation of effects of rumen fill, passage, and rumination and the concept of colonization. These changes improved model behavior markedly. However, effects of feeding frequency were still not simulated adequately. For example, predictions of rumen ammonia nitrogen in animals fed once daily differed by as much as 50% from experimental estimates at the same time of day. If the rate constant for ammonia absorption from the rumen deduced from continuous feeding data were adjusted, once-a-day feeding can be adequately simulated. However, this type of change was not compatible with concept 5. These observations imply that factor(s) other than those recognized, at that time, influence protein hydrolysis and ammonia absorption.

In construction of the first model, data needed for estimation of a number of rate constants were very limited and relevant data sets available were often quite variable. A premise in modeling is that experimental data are correct. When data available are limited and variable, the modeler is tempted to select the datum that works best and can never be certain whether he or she has resisted this temptation. Mechanistic models are complex in explicitly expressing causal as compared to empirical relationships. If the computation of numerical inputs— rate constants and pool sizes—is not firmly based on experimental data, the model can become nonunique; i.e., when a model is complex with many equations and numerical inputs, many combinations of numerical inputs can produce satisfactory outputs. Adequate experimental data are required to prevent this. When confidence regarding uniqueness is low, the model and evaluations and conclusions derived therefrom are suspect. The very complete study of Ulyatt et al. (1984) helped satisfy this need during evolution of the Murphy et al. (1986) revision of the rumen model. All numerical inputs to the model were evaluated carefully based on these data, and adjustments, where required, were made. One could then believe that when the model behaved well, the result reflected adequacy of the conceptual frame on which it was based. Conversely, when the model failed, this could be taken as an indication that the conceptual frame was inadequate rather than a problem with nonunique parameter values. As noted, concepts regarding the dynamics of nitrogen metabolism in the rumen required improvement. Other aspects that required further study were water dynamics, which are poorly represented in most current models; particle size reduction through chewing; rumination and other factors; and factors influencing passage rates across species.

At the same time that Baldwin et al. (1977) were working on the discussed model, Black et al. (1981) were also developing a dynamic model of ruminant digestion. Their objective was to integrate current concepts of processes occurring

in the rumen in equation forms compatible with their existing model of growth (Black et al. 1981) and enable the prediction of growth and production in the whole animal. This model, although more aggregated than that of Baldwin et al. (1977), captured many of the same key concepts, including nitrogen recycling from the animal to the rumen via saliva, microbial maintenance requirements, use of amino acids and ammonia as sources of nitrogen for microbial growth, independent consideration of soluble and insoluble substrates, etc. Major differences were uses of empirically deduced equations, estimates of potentially degradable protein and α and β hexose polymers, and rates of hydrolysis as inputs to the model.

Evaluations of the model were very good (Beever et al. 1981) and certainly better than those obtained with the original Baldwin et al. (1977) model, whether the comparisons were between experimental data and model predictions or of systematic errors of prediction. This should be the expected result, in that the Baldwin et al. (1977) model was designed as a research model, and one of the imposed requisites was that the model apply to all ruminant diets and feeding regimes with no changes in rate constants or parameter values (to test the adequacy of current data and concepts). The Black et al. (1981) model was designed, however, as a predictive model in which estimates of potential degradability and fractional degradation rates were defined on the basis of empirical or experimental estimates made on the specific feedstuffs to be simulated. Despite their differences in objectives and form, these two models provided valuable insights regarding the types of data and additional inquiry needed to advance our understanding of quantitative and dynamic aspects of ruminant digestive function.

5. Current Dynamic Models

Three recent models were selected as representative of current mechanistic and dynamic models of ruminant digestion. The first model was published by Danfaer (1990). This was selected because the objective was to develop a research model that could eventually evolve into a model for practical use in animal agriculture. The second model evolved from the Baldwin et al. (1987) study and was initially devised to formulate nutrient inputs to a lactating cow model; it has since been improved in a number of regards. The final model selected was that of Dijkstra et al. (1992) and Neal et al. (1992). These models share a number of common attributes and in this illustrate a convergence that appears to be taking place and may indicate future directions in the evolution of dynamic models of ruminant digestion. To a significant extent, the three models trace their origins to the analysis of France et al. (1982), entitled *A Mathematical Model of the Rumen*. All three models have been described in detail (Danfaer 1990; Dijkstra et al. 1992; Baldwin 1995) and thus, are available for use and comparison by

all interested workers. Evaluation criteria differed in the three cases, but all of these are relevant and reveal important limitations in our current understanding of digestion in the ruminant.

A very important departure from previous approaches evident in all three of these models is that most of the equations used exhibit saturation or Michaelis-Menten type kinetics, rather than the first-order, mass action arguments used previously. There are a number of good reasons for the departure. First, almost all metabolic systems are saturable such that when substrate availability is high, the system approaches a maximum reaction velocity. Also, as nutrient availabilities vary over a physiological range, patterns of nutrient use can vary in a fashion not adequately simulated using mass action kinetic arguments. Another important reason is that equations that exhibit saturation kinetics are more stable or robust in their behavior when numerical methods are used to solve large models. A major disadvantage arises from the fact that the number of parameters that must be defined essentially doubles. As with early models, a static model of flux rates observed in a reference, steady-state system was utilized heavily. Such data are adequate and appropriate for parameterizing mass action equations. However, these data are not adequate for Michaelis-Menten type equations. To parameterize such equations an assumption or an additional data fact is required for each equation. Common assumptions that can often be defended in general terms are that the V_{max} is very high relative to observed rates and need not be specified accurately (solutions insensitive to this parameter) or that k_s and concentration of substrate are about equal under most conditions. However, such assumptions cannot always be defended and thus invoked. We are not pointing this out as a criticism of the newer models, although all three of the models contain parameter estimates that can be challenged on these grounds, but rather to emphasize that further evolution of these models toward the goal of general applicability will require generation of kinetic data in addition to steady-state data. In presentation of the three models, discussions of parameterization and types of data required for further evolution are included.

A very troubling area of parameter estimation is detailing feed inputs to the models. In all three models, these inputs are based on chemical and physical properties of the feedstuffs. There are notable differences in aggregation of inputs among the models. The most detailed feed composition inputs (16) are required by the Baldwin et al. (1987) model. A drawback of this approach is that such complete analyses of diets are rarely available. This problem identifies one area of needed research—more extensive, detailed analyses of feedstuffs. Dijkstra et al. (1992) also noted that data on diet composition affects the accuracy of model simulations.

The models of Danfaer (1990) and Dijkstra et al. (1992) utilize feed fractions that are composites of several chemical entities. For example, Danfaer (1990) combines organic acids, pectin, lactate, and starch into a fermentable sugars pool,

and hemicellulose and cellulose into a fermentable cell-wall carbohydrate pool. Dijkstra et al. (1992) combine hemicellulose, cellulose, and lignin into a neutral detergent fiber pool. A problem with aggregation of chemicals differing in elemental composition and heats of combustion per gram is that computation of carbon and redox balances and accounting for energetic relationships is not feasible.

Additionally, Danfaer (1990) and Dijkstra et al. (1992) define fermentable/degradable and unfermentable/undegradable fractions of protein, starch, and NDF. These inputs are set based on in vitro fermentation rates and in vivo degradation rates of feeds in nylon bags. By utilizing such data to set the input data, consideration of important biological processes that contribute significantly to (explainable) variance in the digestion process is avoided. In this sense, the models become partially self-fulfilling and dependent on the accuracy of data used to define inputs on digestibility or degradability.

It is very important to recognize in this context that utilization of inputs that can be adjusted to accommodate variance partially compromises the modelers' goal of identifying areas where current data on the chemical and physical properties of feeds are not adequate to define the reactions and relationships occurring in ruminant digestion. Also, although it is recognized that deficiencies exist in all three models with respect to predicting nutrient outflow from the rumen and, ultimately, nutrient availability to the animal, these deficiencies give emphasis and direction as to where research and modeling efforts should be concentrated.

Dijkstra et al. (1992) identified several areas in earlier models "as being insufficiently represented." These included "aspects relating to microbial recycling, microbial substrate preference, uncoupled microbial fermentation, effect of pH on microbial activity and VFA and ammonia absorption," and "variation in microbial chemical composition, as influenced by microbial species and nutrient availability."

The concept of recycling microbial matter was captured by Dijkstra et al. (1992) by including model elements defining protozoa as a subfraction of the amylolytic microbial pool. Inclusion of protozoa in the model allows for microbial matter recycling within the rumen by protozoa engulfment. It also provides for equations describing the retention and lysis of protozoa, lactate metabolism by protozoa, and the effects of pH on protozoal growth.

Substrate preferences of the highly diverse and variable population in the rumen continue to pose a major problem to modelers of the rumen fermentation process. As noted in the introduction to this chapter, the many factors that define interactions among the individual microbial species in the rumen have, probably, not all been identified and, certainly, have not been characterized in the quantitative fashion required for the formulation of models of the rumen ecology. This problem has forced modelers of rumen digestion to the use of highly aggregated representations of rumen fermentation. Baldwin et al. (1970) defined three distinct rumen microbe groups based on substrate specificity. These workers recognized

that this decision stretched the available data resulting in overparameterization of the model and simplified their representation to two microbial groups. One of these groups was identified as those fermenting readily available carbohydrates and the other as those that grow in association with fibrous particles (Baldwin et al. 1977). This approach was modified by adoption of the concept that the microbes in each group vary with the amount of concentrate (starch) in the diet and the deduction of differing stoichiometric coefficients for carbohydrates in roughage and concentrate diets, as already discussed and in Table 10.2 (Murphy et al. 1982, 1986). Similarly, Baldwin et al. (1987), Dijkstra et al. (1992), and Baldwin (1995) adopted the same approach—two microbial groups with differences in fermentation patterns based on diet composition. The strategies in formulation of equation forms differ but the underlying conceptual frame is held in common.

Effects of pH on cellulose and hemicellulose hydrolysis were implemented by Argyle and Baldwin (1988) by modifying the rate constant for these reactions in a linear manner below pH 6.2. The authors recognized that the response form is likely sigmoidal in form, but insufficient data were available at that time to parameterize a nonlinear equation. In addition to pH effects on microbial action of degradable fiber, Dijkstra et al. (1992) also implemented pH effects on protozoa and absorption of VFA and ammonia by modifying the maximal velocity of the respective equations as a function of pH minimum and period of time below a set pH.

Table 10.2 Estimated rumen fermentation parameters[a]

Substrate	Group	Acetate A(I) ± SE	Propionate B(I) + SE	Butyrate C(I) ± SE	Valerate D(I) ± SE
Soluble carbohydrate[b]	R[c]	0.69 ± 0.06	0.20 ± 0.01	0.11 ± 0.05	0.00 ± 0.00
	C[d]	0.45 ± 0.03	0.21 ± 0.04	0.30 ± 0.04	0.04 ± 0.00
Starch	R	0.59 ± 0.04	0.14 ± 0.02	0.20 ± 0.05	0.06 ± 0.00
	C	0.40 ± 0.01	0.03 ± 0.00	0.20 ± 0.00	0.10 ± 0.01
Hemicellulose	R	0.57 ± 0.06	0.18 ± 0.03	0.21 ± 0.05	0.05 ± 0.00
	C	0.56 ± 0.01	0.26 ± 0.01	0.11 ± 0.01	0.07 ± 0.00
Cellulose	R	0.66 ± 0.10	0.09 ± 0.00	0.23 ± 0.09	0.03 ± 0.00
	C	0.79 ± 0.00	0.06 ± 0.00	0.06 ± 0.00	0.09 ± 0.00
Protein	R	0.45 ± 0.07	0.30 ± 0.04	0.18 ± 0.03	0.07 ± 0.00
	C	0.36 ± 0.03	0.37 ± 0.01	0.20 ± 0.03	0.07 ± 0.00

[a]Adapted from Murphy et al. (1982).
[b]Soluble carbohydrate plus organic acids and pectin.
[c]Roughage diets. Values are mean ± standard errors (10 jackknife runs).
[d]Concentrate diets. Values are mean ± standard errors (8 jackknife runs).
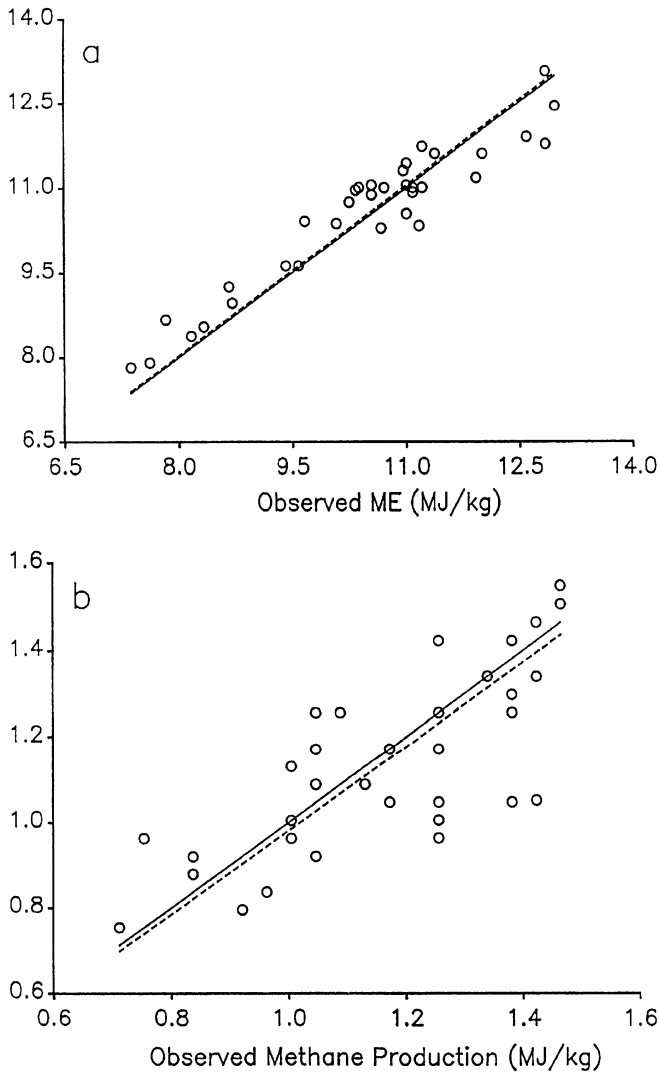
Figure 10.1 Comparisons of predicted versus observed estimates of (a) metabolizable energy and (b) methane production. Regression equations for best fit lines (—) were $y = a + bx$ with r^2 values for ME and CH_4 of 0.85 and 0.60, respectively. Regression equations for lines of equality (– –) were $y = bx$. The 34 experimental estimates were reported by Armsby and Fries (1915), Armstrong (1964), Armstrong et al. (1958), Armstrong et al. (1964), Blaxter (1967), Blaxter and Clapperton (1965), Colovos et al. (1970), (1979), Coppock et al. (1964), Flatt et al. (1967), Graham et al. (1958), Moe et al. (1972, 1973a,b), Moe and Tyrrell (1976, 1979a,b), Tyrrell and Moe (1972), and Tyrrell et al. (1976) for diets including high- and low-quality legume, corn silage, corn meal, soybean meal, and high- and low-quality grass hays.

It has long been recognized that the proportion of polysaccharides in microbes is variable (Hespell and Bryant 1979; Storm and Ørskov 1983; Czerkawski 1986). This concept was captured in Dijkstra et al. (1992) by allowing storage polysaccharides to vary as a function of the concentration of hexose available to amylolytic microbes. This is an important concept and represents an advance.

These models of ruminant digestion (Danfaer, 1990; Dijkstra et al. 1992; Neal et al. 1992; Baldwin 1995) are the current state of the art. They have been built on biological concepts laid out in earlier generations of mathematical models, and have been further refined both biologically and mathematically. Some aspects of ruminant digestion have been aggregated or removed dependent on their significance to overall digestion, a determination that was made as part of the modeling exercises. Other aspects of ruminant digestion have systematically been added to the models to improve model behavior and predictions. The overall behavior of the models is considered good. As an example of this efficacy, results of model estimates of observed and predicted values of metabolizable energy (ME) and methane (CH_4) using Baldwin's (1995) model are presented in Figures 10.1a and 10.1b. Differences between observed and predicted values may be due to experimental or animal variance, errors in data evaluation of diet composition as inputs to the model, or model errors. In view of these potential sources of errors, the r^2 values for ME and CH_4, 0.85 and 0.60, respectively, indicate acceptable agreement and a lack of systematic error of predictions. It must be recognized that several areas exist where model predictions are poor. These include predicting cellulose and hemicellulose digestion on high-concentrate diets and predicting the yield of VFA. Errors in estimates of butyrate production present a particularly difficult problem, as discussed.

Current models are more heavily parameterized than earlier ones due, in large part, to the adoption of Michaelis-Menten type equation forms. Many assumptions have been made in estimating the parameters for these equations. Making assumptions and testing them is an integral part of the modeling/experimental research cycle. The fact that assumptions are made should not be forgotten, however. In the words of Czerkawski (1986), "the danger lies within, that others will not realize that assumptions have been made."

References

Argyle, J. L., and R. L. Baldwin. 1988. Modeling of rumen water kinetics and effects of rumen pH changes. *J. Dairy Sci.* 71:1178–1188.

Armsby, H. P., and J. A. Fries. 1915. Net energy values of feeding stuffs for cattle. *J. Agric. Res.* 3:435–491.

Armstrong, D. G. 1964. Evaluation of artificially dried grass as a source of energy for sheep. II. The energy value of cocksfoot, timothy and two strains of rye grass at varying stages of maturity. *J. Agric. Sci.* 62:399–416.

Armstrong, D. G., K. L. Blaxter, N. McC. Graham, and F. W. Wainman. 1958. The utilization of the energy of two mixtures of steam-volatile fatty acids by fattening sheep. *Br. J. Nutr.* 12:177–188.

Armstrong, D. G., K. L. Blaxter, and R. Waite. 1964. The evaluation of artificially dried grass as a source of energy for sheep. III. The prediction of nutritive value from chemical and biological measurements. *J. Agric. Sci.* 62:417–424.

Baldwin, R. L. 1995. *Modeling Ruminant Digestion and Metabolism.* Chapman & Hall Ltd., UK.

Baldwin, R. L., and M. J. Allison. 1983. Rumen metabolism. *J. Anim. Sci.* 57:461–477.

Baldwin, R. L., H. L. Lucas, and R. Cabrera. 1970. Energetic relationships in the formation and utilization of fermentation end-products. In A. T. Phillipson (ed.), *Physiology of Digestion and Metabolism in the Ruminant*, pp. 319–334. Oriel Press, Newcastle Upon Tyne.

Baldwin, R. L., L. J. Koong, and M. J. Ulyatt. 1977. A dynamic model of ruminant digestion for evaluation of factors affecting nutritive value. *Agric. Syst.* 2:255–288.

Baldwin, R. L., J. H. M. Thornley, and D. E. Beever. 1987. Metabolism of the lactating cow II. Digestive elements of a mechanistic model. *J. Dairy Res.* 54:107–131.

Beever, D. E., Black, J. L., and Faichney, G. J. 1981. Simulation of the effects of rumen function on the flow of nutrients from the stomach of sheep: Part 2. Assessment of computer predictions. *Agric. Syst.* 6:221–241.

Black, J. L., D. E. Beever, G. J. Faichney, B. R. Howarth, and N. M Graham. 1980. Simulation of the effects of rumen function on the flow of nutrients from the stomach of sheep: Part I. Description of a computer program. *Agric. Syst.* 6:195–219.

Blaxter, K. L. 1967. *The Energy Metabolism of Ruminants*, 2nd ed. Hutchinson and Co. Ltd., London.

Blaxter, K. L., and J. L. Clapperton. 1965. Prediction of the amount of methane produced by ruminants. *Br. J. Nutr.* 19:511–522.

Bryant, M. P. 1959. Bacterial species of the rumen. *Bacteriol. Rev.* 23:125–153.

Colovos, N. F., J. B. Holter, R. M. Koes, W. E. Urban, Jr., and H. A. Davis. 1970. Digestibility, nutritive value and intake of ensiled corn plant (*Zea Mays*) in cattle and sheep. *J. Anim. Sci.* 30:819–824.

Colovos, N. F., H. A. Keener, J. R. Prescott, and A. E. Teeri. 1979. The nutritive value of timothy hay at different stages of maturity as compared with second cutting clover hay. *J. Dairy Sci.* 32:659–664.

Coppock, C. E., W. P. Flatt, L. A. Moore, and W. E. Stewart. 1964. Effect of hay to grain ratio on utilization of metabolizable energy for milk production by dairy cows. *J. Dairy Sci.* 47:1330–1338.

Czerkawski, J. W. 1986. *An Introduction to Rumen Studies.* Pergamon Press Ltd., Oxford, U.K.

Danfaer, A. 1990. A dynamic model of nutrient digestion and metabolism in lactating dairy cows. Ph.D. Dissertation, National Institute of Animal Science, Foulurm, Denmark.

Dijkstra, J., H. D. S. C. Neal, D. E. Beever, and J. France. 1992. Simulation of nutrient digestion, absorption and outflow in the rumen: Model description. *J. Nutr.* 122:2239–2256.

Dobson, A., and M. J. Dobson. (eds.). 1988. *Aspects of Digestive Physiology in Ruminants.* Comstock Publishing Associates, Cornell, University Press, Ithaca, NY.

Flatt, W. P., P. W. Moe, L. A. Moore, A. W. Munson, and T. Cooper. 1967. Energy utilization by high producing dairy cows. II. Summary of energy balance experiments with lactating holstein cows. *Eur. Assoc. Anim. Prod. Publ.* 12:235–251.

France, J., J. H. M. Thornley, and D. E. Beever. 1982. A mathematical model of the rumen. *J. Agric. Sci.* 99:343–353.

Graham, N. McC., K. L. Blaxter, and D. G. Armstrong. 1958. Analytical and other techniques used in respiration calorimetry and their errors. *Eur. Assoc. Anim. Prod. Publ.* 8:157–163.

Hespell, R. B., and M. P. Bryant. 1979. Efficiency of rumen microbial growth: Influence of some theoretical and experimental factors on Y_{ATP}. *J. Anim. Sci.* 49:1640–1659.

Hogan, J. P. 1975. Symposium: Protein and amino acid nutrition in the high producing cow. Quantitative aspects of nitrogen utilization in ruminants. *J. Dairy Sci.* 58:1164–1177.

Hungate, R. E. (ed.). 1966. *The Rumen and Its Microbes.* Academic Press, New York and London.

Koong, L. J., R. L. Baldwin, M. J. Ulyatt, and T. J. Charlesworth. 1975. Iterative computation of metabolic flux and stoichiometric parameters for alternate pathways in rumen fermentation. *Comp. Program. Biomed.* 4:209–213.

Maeng, W. J., and R. L. Baldwin. 1976. Factors influencing rumen microbial growth rates and yields: Effect of amino acid additions to a purified diet with nitrogen from urea. *J. Dairy Sci.* 59:648–655.

Maeng, W. J., C. J. Van Nevel, R. L. Baldwin, and J. G. Morris. 1976. Rumen microbial growth rates and yields: Effect of amino acids and protein. *J. Dairy Sci.* 59:68–79.

Mazanov, A., and J. V. Nolan. 1976. Simulation of the dynamics of nitrogen metabolism in sheep. *Br. J. Nutr.* 35:149–174.

McDonald, I. W., and A. C. I. Warner (eds.). 1975. *Digestion and Metabolism in the Ruminant.* The University of New England Publishing Unit, Armidale, N.S.W., Australia.

Moe, P. W., W. P. Flatt, and H. F. Tyrrell. 1972. Net energy value of feeds for lactation. *J. Dairy Sci.* 55:945–958.

Moe, P. W., and H. F. Tyrrell. 1976. Effect of feed intake and physical form on energy value of corn in timothy hay for lactating cows. *J. Dairy Sci.* 60:752–758.

Moe, P. W., and H. F. Tyrrell. 1979a. Effect of endosperm type on incremental energy value of corn grain for dairy cows. *J. Dairy Sci.* 62:447–454.

Moe, P. W., and H. F. Tyrrell. 1979b. Methane production in dairy cows. *Eur. Assoc. Anim. Prod. Publ.* 26:59–62.

Moe, P. W., H. F. Tyrrell, and N. W. Hooven, Jr. 1973a. Energy balance measurements with corn meal and ground oats for lactating dairy cows. *J. Dairy Sci.* 56:1149–1153.

Moe, P. W., H. F. Tyrrell, and N. W. Hooven, Jr. 1973b. Physical form and energy value of corn grain. *J. Dairy Sci.* 56:1298–1304.

Murphy, M. R., R. L. Baldwin, and L. J. Koong. 1982. Estimate of stoichiometric parameters for rumen fermentation of roughage and concentrate diets. *J. Anim. Sci.* 55(2): 411–421.

Murphy, M. R., R. L. Baldwin, and M. J. Ulyatt. 1986. An update of a dynamic model of ruminant digestion. *J. Anim. Sci.* 62:1412–1422.

Neal, H. D. S. C., J. Dijkstra, and M. Gill. 1992. Simulation of nutrient digestion, absorption and outflow in the rumen: Model evaluation. *J. Nutr.* 122:2257–2272.

Nolan, J. V. 1975. Quantitative models of nitrogen metabolism in sheep. In I. W. Mcdonald and A. C. I. Warner (eds.), *Digestion and Metabolism in the Ruminant. Proceedings of the IV International Symposium on Ruminant Physiology.* The University of New England, Armidale, N. S. W. Australia. pp. 417–427.

Phillipson, A. T. (ed.). 1970. *Physiology of Digestion and Metabolism in the Ruminant.* Oriel Press, Newcastle upon Tyne.

Reichl, J. R., and R. L. Baldwin. 1975. Rumen modeling: Rumen input-output balance models. *J. Dairy Sci.* 58:879–890.

Reichl, J. R., and R. L. Baldwin. 1976. A rumen linear programming model for evaluation of concepts of rumen microbial function. *J. Dairy Sci.* 59:439–454.

Russell, J. B., and M. S. Allen, (eds.). 1984. *Gastrointestinal Microecology.* American Society for Microbiology, Washington, DC.

Russell, J. B., and R. B. Hespell. 1981. Microbial rumen fermentation. *J. Dairy Sci.* 64:1153–1169.

Russell, J. B., J. D. O'Conner, D. G. Fox, P. J. Van Soest, and C. J. Sniffen. 1992. A net carbohydrate and protein system for evaluating cattle diets: I. Ruminal fermentation. *J. Anim. Sci.* 70:3551–3561.

Search:Agriculture. 1990. *The Cornell Net Carbohydrate and Protein System for Evaluating Cattle Diets.* Cornell University Agriculture Experiment Station, Number 34, Ithaca, NY.

Storm, E., and E. R. Ørskov. 1983. The nutritive value of rumen micro-organisms in ruminants. I. Large-scale isolation and chemical composition of rumen micro-organisms. *Br. J. Nutr.* 50:463–470.

Tyrrell, H. F., and P. W. Moe. 1972. Net energy value for lactating of a high and low concentration ration containing corn silage. *J. Dairy Sci.* 55:1106–1112.

Tyrrell, H. F., C. K. Reynolds, and P. W. Moe. 1976. Effect of basal ration consumed upon utilization of acetic acid for lipogenesis by mature cattle. *Eur. Assoc. Anim. Prod. Publ.* 19:57–60.

Ulyatt, M. J., G. C. Waghorn, A. John, C. S. W. Reid, and J. Munro. 1984. Effect of intake and feeding frequency behavior and quantitative aspects of digestion in sheep fed chaffed lucerne hay. *J. Agric. Res. (Camb.)* 102:645–657.

Wolin, M. J. 1960. A theoretical rumen fermentation balance. *J. Dairy Sci.* 43:1452–1459.

11

Mathematical Models of
Bacterial Chemotaxis

Roseanne M. Ford and Peter T. Cummings

1. Introduction

Many bacterial species exhibit chemotactic behavior, the ability to bias their otherwise random motion in the direction toward increasing concentrations of nutrients (referred to as attractants) or away from increasing concentrations of metabolites or compounds toxic to the bacteria, which may be indicators of unfavorable conditions (referred to as repellents). Chemotaxis can provide a competitive advantage for bacteria because in their natural habitats they are continually exposed to changing environmental conditions, and their survival depends on their capacity to respond favorably to adverse circumstances. Because their small size (1 to 2 µm) and simple structure limits their ability to modify their surroundings, they respond either by migration to a more desirable location or by adaptation of their internal metabolic processes (Macnab 1980). Adaptation occurs naturally through genetic modification, but is relatively slow. Chemotactic bacteria can clearly respond much more quickly by moving to a more favorable environment. Chemotaxis has many practical applications and is known to play important roles in nitrogen fixation in plants, the pathogenesis of disease, and the bioremediation of contaminated aquifers. This last case is of particular interest in our research group because it has been shown that bacteria are capable of degrading many toxic organic materials—including halogenated hydrocarbons via anaerobic degradation (Bouwer 1992; Harvey 1991)—and additionally respond chemotactically to these compounds. We are pursuing a long-range research program aimed at understanding the role of bacterial motility and chemotaxis in *in situ* bioremediation processes. The objectives are to quantitatively measure bacterial migration at the macroscopic level (both in the presence and absence

of one or more attractant and/or repellent species), understand the basis for the macroscopic behavior by measuring and analyzing the motion of individual bacteria, develop mathematical models for bacterial migration based on microscopic and macroscopic level information, and use the model to predict bacterial migration in natural processes, with particular emphasis on *in situ* bioremediation processes.

To interpret experimental data on bacterial migration in terms of fundamental transport coefficients and to predict the behavior of bacterial populations in response to changing environmental conditions, it is necessary to develop and solve mathematical models for bacterial migration. Several such models have been proposed in the literature and can be categorized into three broad classes: (1) population-level models, which attempt to describe the bacterial migration at a purely macroscopic level by convective diffusion equations and so are partial differential equations for the bacterial density $c(r,t)$ at the point r at time t; (2) subpopulation-level models, which are based on balance equations for the density of bacteria $n(r,\hat{s},t)$ at the point r at time t moving in direction $\hat{s}$ (a unit vector), and hence require a model for the way in which bacteria change the direction of their motion; and (3) microscopic level models that are based on individual cell dynamics, which can be described by stochastic equations of motion.

In this chapter, we review several mathematical models for bacterial chemotaxis. We particularly focus on our own modeling efforts at the microscopic and subpopulation levels and the utility of these models for analyzing experiments, visualizing bacterial migration, and predicting bacterial migration in natural environments (including exposure to multiple gradients). The chapter is divided up as follows: In the remainder of this Introduction, the known microscopic dynamics of bacteria both in the absence and presence of chemical stimuli (attractants or repellents) is described in Section 1.1, the experimental assay developed by one of us (Ford and Lauffenburger 1991; Ford et al. 1991) for observing bacterial migration at the macroscopic level is outlined in Section 1.2, and the philosophy and utility of our two modeling approaches (cellular dynamics simulation and subpopulation models) is discussed in Section 1.3. In Section 2, we present Alt's subpopulation model (Alt, 1980), which is the most general subpopulation level description of bacterial migration considered in the published literature. We also describe a reduced form of Alt's equations suitable for systems in which the stimulus gradients occur in one spatial direction only (Frymier et al. 1994), and compare the reduced form of Alt's model with a one-dimensional model proposed by Rivero et al. (1989). Cellular dynamics (Frymier et al. 1993), a technique for studying large populations of bacteria by performing simulations of large numbers of individual bacteria based on the known single-cell dynamics, is presented in Section 3. In Section 4, the solutions of the reduced form of Alt's equations for a one-dimensional stimulus gradient and of the Rivero et al. model are compared to cellular dynamics simulations and to experimental data. In Section 5, we

apply the reduced form of Alt's equations to the case of bacteria responding simultaneously to two attractants to demonstrate modeling a more relevant situation in natural microbial processes. We propose three simple models for the tumbling probability associated with the swimming behavior in such a case and solve the resulting balance equations by using finite element methods. We conclude in Section 6 with some remarks about the necessity and utility of mathematical modeling in understanding bacterial migration.

1.1. Mechanism of Bacterial Chemotaxis

The swimming behavior exhibited by individual bacteria at the microscopic level consists of a series of relatively straight line "runs" followed by "tumbles," which usually result in a change in orientation (Berg and Brown 1972; Macnab and Koshland 1972; Spudich and Koshland 1975). Typical swimming speeds are 20 to 60 μm/sec (Macnab and Aizawa 1984) with runs lasting up to several seconds and tumbles taking about 0.1 sec. In an isotropic medium, the swimming behavior resembles a three-dimensional random walk analogous to molecular diffusion. In molecular diffusion, however, a change in direction results from a collision with another molecule, whereas for bacteria a change in direction results from a reversal of flagellar rotation from counterclockwise to clockwise. In the presence of an attractant gradient, a bacterium is able to decrease its tumbling frequency (increase its run time) if it is moving in a favorable direction and thereby bias its random walk to achieve net migration toward higher attractant concentrations. Bacteria respond to repellents by extending their run lengths when swimming toward lower concentrations.

The migration behavior of bacteria in the absence and presence of an attractant is illustrated in Figure 11.1, which contains visualizations of cellular dynamics simulations (see Section 3) of bacteria in an isotropic medium (left column of photographs) and in the presence of a linear attractant gradient (right column). Straight line runs, interrupted by tumbles and changes in direction, are evident in all the photographs. In the absence of an attractant, the overall motion of the population is uniform (without bias) so the bacteria spread out in a uniform sphere. For the bacteria in the presence of an attractant, it is clear that the runs in the direction of increasing attractant concentration are on average longer than those in the opposite direction with the result that the population spreads out nonuniformly with a bias toward increasing attractant concentration.

As bacteria move about exploring their surroundings they monitor changes in chemical concentrations through special proteins called receptors located near the cell surface. These receptors, like enzymes, have specific binding sites to which only a narrow range of structurally similar chemical substrates can bind. It is the change in the number of bound receptors over time that provides information to the cell regarding chemical gradients (Macnab and Koshland 1972). This

Random Motility   Chemotaxis
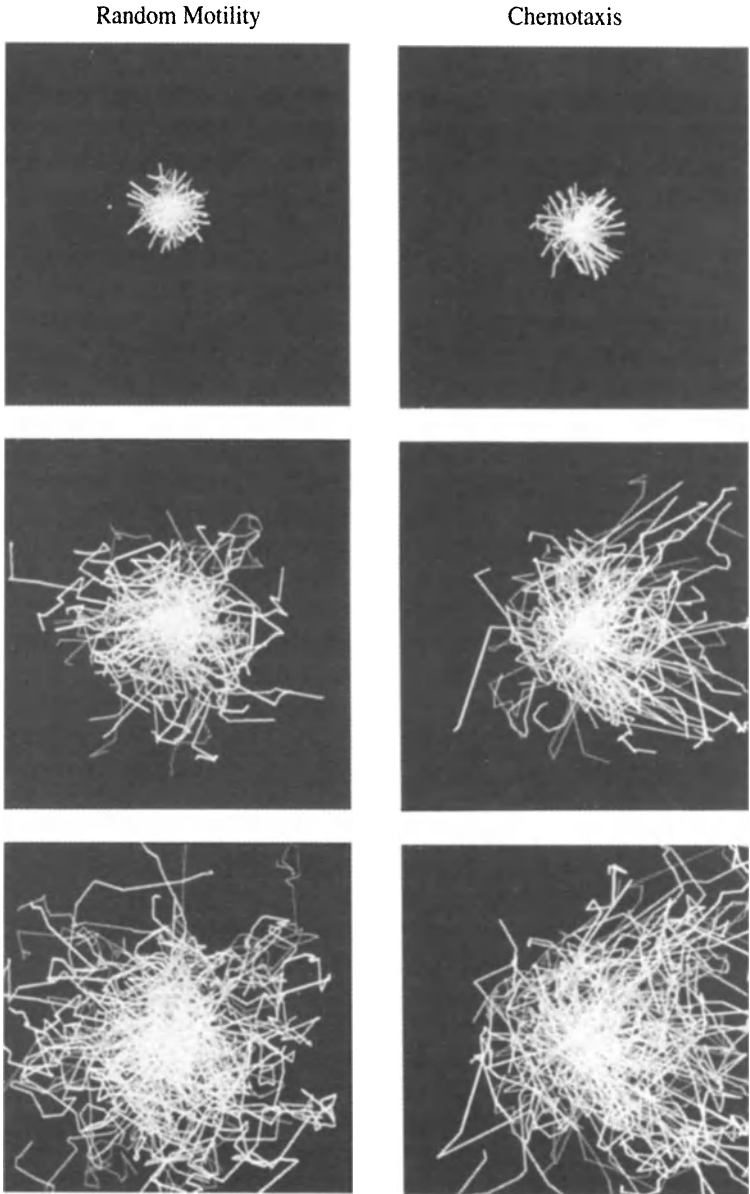
Figure 11.1 Cellular dynamics (CD) simulations of 100 bacteria (originating from the center of the photograph) shown as traces of bacterial trails. In the left-hand series, no attractant gradient is present and the bacteria disperse essentially uniformly in all directions as time increases from top to bottom. In the right-hand series, a linear attractant gradient that increases from left to right generates a chemotactic response. Bacterial run lengths to the right are on average longer than those to the left resulting in an overall bias of bacterial motion toward the right.

231

interaction generates a biochemical signal that controls the flagellar rotation and hence the "decision" whether to tumble or to continue swimming. In the analysis of our data, a simple mechanism that assumes equilibrium binding of attractant to the receptors and a direct correspondence between the number of bound receptors and the chemotactic signal adequately describes experimentally observed chemotactic responses to single stimuli.

1.2. Stopped-Flow Diffusion Chamber

The stopped-flow diffusion chamber (SFDC) assay shown schematically in Figure 11.2 is advantageous for the study of bacterial chemotaxis because it
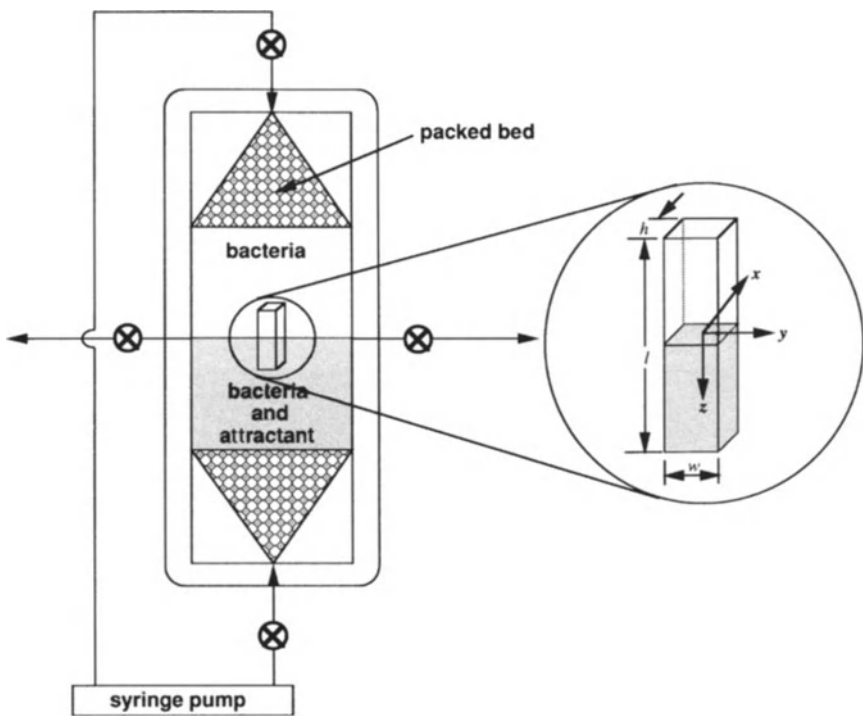
Figure 11.2 Diagram of the SFDC assay (left). Two bacterial suspensions differing only in the concentration of an attractant are contacted by impinging flow to create an initial step change in attractant concentration at the center of the chamber and a uniform distribution of bacteria. The walls of the chamber are formed by two microscope slides, which serve as optical windows. The approximate dimensions of the chamber are $4 \times 2 \times 0.2$ cm. On the right is an exploded view of the simulation box, which is referred to in Section 3 of the text.

provides well-characterized chemical gradients (Ford et al. 1991; Ford and Lauf-fenburger 1991). The fluid flow behavior within the chamber involves the contact of two suspensions by impinging flow. The impinging flow is created by fluid flow in equal volumetric amounts into the top and bottom of the SFDC with the fluid exiting at the two side ports, as illustrated in Figure 11.2. One critical aspect is that while the fluid is flowing no mixing occurs between the two suspensions and a step change in stimulus concentration can be maintained to establish well-defined initial conditions for solving the mathematical model equations. After establishing the initial conditions, flow is stopped, and the chemical stimulus in the bottom half begins to diffuse into the upper half, setting up a transient gradient in the stimulus concentration. As bacteria respond to this gradient, a band of high cell density forms and moves downward toward higher stimulus concentra-tions. The changing bacterial distribution is observed through light scattering under dark-field illumination and recorded with a charge-coupled device (CCD) video camera. Digitized images at several time points over a 10-min interval are generated using a frame grabber board connected to a microcomputer with image processing software. Bacterial density profiles are computed from gray level measurements over a rectangular region in each image. The bacterial transport coefficients for random motility and chemotaxis, as defined in the Rivero-Tran-quillo-Buettner-Lauffenburger (RTBL) mathematical model presented in Section 2.2.4, are determined from the observed bacterial distributions. Note that the experiments in the SFDC are performed to minimize the number of factors that need to be taken into account in modeling bacterial movement. Bacterial growth and decay are negligible because the experimental observation times (6 to 12 min) are short compared to the expected bacterial doubling times at the suboptimal growth conditions in the SFDC assay. The attractant species studied are nonmetab-olizable analogs of metabolizable attractants (such as fucose and α-methylaspar-tate as analogs for galactose and aspartate) so that consumption of the attractant and the production of metabolites (which can act as repellents) can be ignored. Hence, in the mathematical models described in the remainder of this chapter, growth, death, attractant consumption, and metabolite production are all assumed to be negligible and are not discussed further.

1.3. Relationships between Computer Simulation, Mathematical Modeling, and Experiments

To characterize the migration behavior of chemotactic bacteria we use three different research tools—experiments, mathematical modeling, and computer simulation. The relationship between the modeling approaches and experimenta-tion is shown in diagram form in Figure 11.3. The experimental system to which we apply our model is the stopped-flow diffusion chamber. We have used both the RTBL model (Section 2.2.4) and the reduced form of Alt's equations (Section
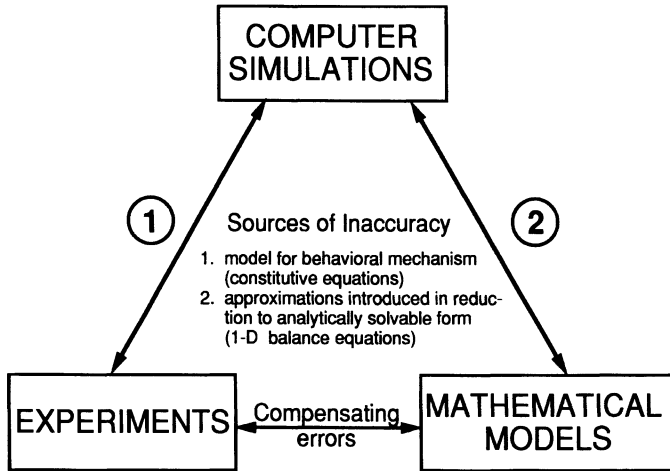
Figure 11.3 Schematic illustration of the relationship between the three approaches (experiments, mathematical models, and computer simulations) we use in characterizing bacterial migration.

2.2.5) to analyze our experimental data in terms of macroscopic transport properties. The solution of these model equations requires approximations that may introduce errors into the theoretical bacterial density profiles. The computer simulation methodology we have developed provides an independent check of the error introduced in the numerical solution of the model equations because it is based on an exact solution of the rigorous equations of motion for a single bacterium. A comparison of the simulation results to the experimental data indicates the accuracy of the physical model used to describe bacterial detection of and response to chemical gradients. Without the independent check afforded by the simulations it is possible that errors associated with approximations in the mathematical model could be compensated for by an incorrect physical model. For example, the restriction of bacterial motion to one-dimension in the RTBL model underpredicts the bacterial response, but the agreement with experiments is still very good. One explanation is that the physical model for the swimming behavior away from the gradient is incorrect and results in an overprediction of the chemotactic response as we showed in Frymier et al. (1994) and is described in Section 4.3. If the computer simulations provide an exact prediction of the behavior for a given physical model, one might then ask what the utility is of the mathematical model. The answer lies in recognizing that the simulations require significant computational resources to yield statistically meaningful results and are therefore not practical to use routinely as a modeling tool. Once a mathematical model is developed that is consistent with and agrees with the

simulations, and that is computationally efficient, characterization of experimental data can be performed on a routine basis using the mathematical model. As we demonstrate in this chapter, the reduced form of Alt's equations, derived by Ford and Cummings (1992) and solved using finite element techniques (see Section 2.2.5), is just such a mathematical model.

2. Population Balance Models

To mathematically describe the motion of chemotactic bacteria, two sets of equations must be formulated. The first are the cell balance equations that describe the evolution in time and space of the cell density. For bacterial migration in the absence of growth or death of the bacteria, these constitute the statement that cell number is conserved and are exact. The second set of equations that need to be formulated are the constitutive equations that relate the quantities in the conservation equation to properties of the cells and their swimming motion. It is in the constitutive equations that the approximations are introduced that define the physical model. One can draw a parallel with the development of a mathematical description of a chemical process: one begins with the exact conservation laws (such as conservation of mass, momentum, and/or energy) and supplements these with constitutive relations (such as equations of state, Newton's law of viscosity, and Fick's law of diffusion) to obtain a closed mathematical problem. In the mathematical modeling of the motion of chemotactic bacteria, several authors (Alt 1980; Othmer et al. 1988; Patlack 1953; Segel 1977; Stroock 1974; Rivero et al. 1989) have developed cell balance equations of varying degrees of rigor. The most general and rigorous are the three-dimensional cell balance equations derived by Alt (1980) that take into account explicitly the position, direction, and elapsed run time of a bacterial subpopulation. A simplified three-dimensional cell balance equation, known as Stroock's equation (1974), can be derived from Alt's equations when one assumes that the probability that a bacterium tumbles is independent of run time. Segel's (1970) equations are one-dimensional phenomenological equations that do not explicitly include the dependence on run time. One can also attempt to describe the bacterial migration at the macroscopic level via a convective diffusion equation. This is the basis for the model of Keller and Segel (1971). Each of these models (composed of balance equations and constitutive relations) and the relationship between them is briefly described here. A mathematically more rigorous derivation of the relationships between the models was published by Ford and Cummings (1992), to which the interested reader is referred.

2.1. Models Based on Population Balance Equations

The time evolution of the cell density $c(\mathbf{r},t)$ is given by the conservation equation

$$\frac{\partial c}{\partial t} = -\nabla_r \cdot \boldsymbol{J}_c , \tag{1}$$

where $\boldsymbol{J}_c$ is the cell flux. This is an exact statement in the absence of cell death or reproduction. In their original phenomenological model to describe population behavior, Keller and Segel (1971) proposed that the bacterial flux include one term to describe diffusion-like or random motility behavior and a second term for convection-like or chemotactic motion,

$$\boldsymbol{J}_c(\boldsymbol{r},t) = -\mu\nabla_r c(\boldsymbol{r},t) + \boldsymbol{V}_c c(\boldsymbol{r},t), \tag{2}$$

where μ is the random motility coefficient, analogous to a molecular diffusion coefficient, and $\boldsymbol{V}_c$ is the chemotactic velocity. The fundamental transport properties, μ and $\boldsymbol{V}_c$, are in general both functions of the stimulus concentration $a(\boldsymbol{r},t)$. Keller and Segel considered the case of bacteria responding to an attractant whose concentration varies only in the z direction, so that the chemotactic velocity has only a single nonzero component, V_{cz}, for which they proposed the constitutive relation,

$$V_{cz} = \chi(a)\, \frac{\partial a}{\partial z} , \tag{3}$$

where

$$\chi(a) = \frac{\Delta}{a} .$$

This particular form was chosen because it generated a traveling wave solution that could describe the traveling bands experimentally observed by Adler (1966, 1969). For an exponential attractant gradient this relationship predicts a constant chemotactic velocity. However, experiments performed in an exponential gradient did not yield a constant chemotactic velocity as expected; the velocity varied by a factor of eight over the concentration range studied (Dahlquist et al. 1972). Lapidus and Schiller (1976) proposed a different functional form,

$$\chi(a) = \delta\frac{k}{(k + a)^2} , \tag{4}$$

to qualitatively satisfy the behavior of a dose-response curve in the capillary assay, as reported by Mesibov et al. (1973). The chemotactic response was proportional to the chemotactic mobility, δ, and had a maximum sensitivity at a

particular concentration, k, which they let equal the initial attractant concentration. Segel (1977) later derived the following expression,

$$\chi(a) = \chi_M \frac{a^{n-1} K^n}{(K^n + a^n)^2} \tag{5}$$

based on experimental observations of individual cell behavior by Brown and Berg (1974). The Hill cooperativity number, n, was included to match the sensitivity of response observed by Dalquist et al. (1972) over a wide range of attractant concentrations. For $n = 1$, the Lapidus and Schiller (1976) form is recovered with the sensitivity constant now defined as the dissociation constant, K, and the chemotactic mobility as the constant χ_M. As we shall see, some of the relationships just given can be derived from more microscopic considerations described in the discussion of models based on subpopulation balance equations described next.

2.2. Models Based on Subpopulation Balance Equations

In this section, we describe several models for bacterial migration based on subpopulation balance equations. The relationships between the models, and the assumptions needed to derive one from another, are represented schematically in Figure 11.4.

2.2.1. ALT'S BALANCE EQUATIONS

In Alt's cell balance equations, the cells are assumed to have piecewise linear paths (runs) where the mean speed, v, depends on position and time—i.e., $v = v(r,t)$. The quantity $\sigma(r,s,\tau,t)$ is the number of cells per unit volume at point r moving in direction $\hat{s}$ (a unit vector) at time t with run time τ. The probability
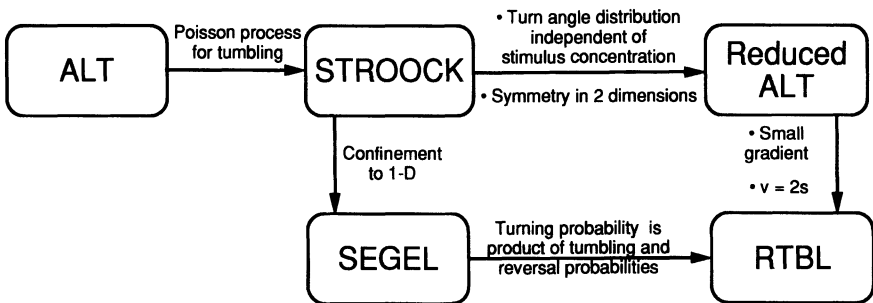
Figure 11.4 Relationships between population and subpopulation cell balance equations in three dimensions and one dimension (shown in boxes) and the assumptions required to derive one from the other (shown as lines).

per unit time that a cell moving in direction $\hat{s}$ at r at time t with run time τ (counted from the beginning of the run) tumbles at r at time t is given by $\beta(r,s,\tau,t)$. If a cell tumbles at r at time t after a run in direction $\hat{s}'$, the probability that after tumbling it chooses the direction $\hat{s}$, as its new direction is $k(r,\hat{s}',t;\hat{s})$, where k is referred to as the turning probability. Alt's fundamental equations (1980) are then

$$\frac{\partial\sigma(r,\hat{s},\tau,t)}{\partial t} = -\frac{\partial\sigma(r,\hat{s},\tau,t)}{\partial\tau} - \hat{s}\cdot\nabla_r[v(r,t)\sigma(r,\hat{s},\tau,t)] - \beta(r,\hat{s},\tau,t)\sigma(r,\hat{s},\tau,t), \quad (6)$$

for $\tau > 0$ and for $\tau = 0$, we have

$$\sigma(r,\hat{s},0,t) = \int_0^\infty\!\!\int \beta(r,\hat{s},\tau,t)\sigma(r,\hat{s}',\tau,t)k(r,\hat{s}',t;\hat{s})\ d\hat{s}'\ d\tau. \tag{7}$$

Physically, Equation (6) states that the rate of change in the population of cells at r at time t moving in direction $\hat{s}$ with run time τ is given by a term that takes into account the change in the run time of the cell population, an efflux term which accounts for the net motion of the cells away from the point r and a loss term due to cells tumbling (with probability β, so density $\beta\sigma$). The subscript r on the divergence operator is included to emphasize that the divergence is with respect to the spatial coordinate r. Note that implicit in this cell balance equation is the assumption that the swimming motion of the cells is piecewise linear and that changes in direction are assumed to occur only through the tumbling mechanism (taken into account by the $\beta\sigma$ term on the right hand side of Equation [6]). Equation (7) states that the way in which one obtains an initial (i.e., run time = 0) population of cells moving in direction $\hat{s}$ is by considering cell populations that were moving in another direction $\hat{s}'$ which tumbled at time t with run time τ, $\beta(r,\hat{s},\tau,t)\sigma(r,\hat{s},\tau,t)$, multiplying by the probability that the cell after tumbling moves in the direction $\hat{s}$, given by $k(r,\hat{s}',t;\hat{s})$. One needs to add up over all such populations-hence, the integration over all directions $\hat{s}'$ and all run times τ.

2.2.2. STROOCK'S BALANCE EQUATION

The general form of Alt's equations can be simplified considerably by applying an important constitutive relationship. If one assumes that the tumbling probability, $\beta(r,\hat{s},\tau,t)$, is independent of run time τ, Alt's equations can be integrated (Ford and Cummings 1992) to give the single equation

$$\frac{\partial n(r,\hat{s},t)}{\partial t} = -\hat{s}\cdot\nabla_r[v(r,t)n(r,\hat{s},t)] - \beta(r,\hat{s},t)n(r,\hat{s},t) \tag{8}$$
$$+ \int\beta(r,\hat{s}',t)n(r,\hat{s}',t)k(r,\hat{s}',t;\hat{s})\ d\hat{s}',$$

where

$$n(\boldsymbol{r},\hat{s},t) = \int_0^{\infty} \sigma(\boldsymbol{r},\hat{s},\tau,t) \, d\tau$$

is the number density of cells per unit volume at point $\boldsymbol{r}$ moving in direction $\hat{s}$ at time t. The assumption that $\beta(\boldsymbol{r},\hat{s},\tau,t)$ is independent of τ is equivalent to assuming that the distribution of run times is governed by a Poisson distribution. This has been shown experimentally to be the case for *Escherichia coli* by Berg and Brown (Berg and Brown 1972; Brown and Berg 1974). For comparison, Stroock's three-dimensional equation for the transport of chemotactic cell populations is given by (Stroock 1974)

$$\frac{\partial n(\boldsymbol{r},\hat{s},t)}{\partial t} = -\hat{s} \cdot \nabla_r[v(\boldsymbol{r},t)n(\boldsymbol{r},\hat{s},t)] - \int p(\boldsymbol{r},\hat{s},\hat{s}',t)n(\boldsymbol{r},\hat{s},t) \, d\hat{s}' \tag{9}$$
$$+ \int p(\boldsymbol{r},\hat{s}',\hat{s},t)n(\boldsymbol{r},\hat{s}',t) \, d\hat{s}',$$

where $p(\boldsymbol{r},\hat{s}',\hat{s},t)$ is the probability density for a cell to change direction at point $\boldsymbol{r}$ from direction $\hat{s}'$ to direction $\hat{s}$. If one makes the identifications

$$p(\boldsymbol{r},\hat{s}',\hat{s},t) = \beta(\boldsymbol{r},\hat{s}',t)k(\boldsymbol{r},\hat{s}',t;\hat{s}) \tag{10}$$
$$\beta(\boldsymbol{r},\hat{s},t) = \int p(\boldsymbol{r},\hat{s},\hat{s}',t) \, d\hat{s}' ,$$

then Stroock's equation is equivalent to Alt's equations for bacteria whose running phase is governed by a Poisson distribution. Furthermore, if one recognizes that

$$c(\boldsymbol{r},t) = \int n(\boldsymbol{r},\hat{s},t) \, d\hat{s} \tag{11}$$

and

$$\boldsymbol{J}_c = v \int n(\boldsymbol{r},\hat{s},t)\hat{s} \, d\hat{s} \tag{12}$$

integration of either Equation (8) or Equation (9) over $\hat{s}$ yields the macroscopic conservation equation, Equation (1). Multiplication of either Equation (8) or Equation (9) by $\hat{s}$ and integration of $\hat{s}$ should lead to an expression for the bacterial flux. However, except for very simple formulations of the tumbling probability and turning probability, there is no obvious simplification of the resulting flux equation to a form easily identified with Equation (2). For a more complete discussion, see Ford and Cummings (1992).

2.2.3. SEGEL BALANCE EQUATIONS

Segel (1977) considered the motion of bacteria in one dimension (z) only and wrote down cell balance equations accordingly. If we define $n^+(z,t)$ as the density of cells at point z at time t moving in the positive z direction and $n^-(z,t)$ as the density of cells at point z at time t moving in the negative z direction, then Segel's equations are

$$\frac{\partial n^+}{\partial t} = -\frac{\partial}{\partial z}(vn^+) + p^-n^- - p^+n^+ \tag{13}$$

$$\frac{\partial n^-}{\partial t} = -\frac{\partial}{\partial z}(vn^-) + p^+n^+ - p^-n^-.$$

In these equations, $p^+ = p^+(z,t)$ is the probability per unit time that a cell moving in the positive z direction becomes a cell moving in the negative z direction (by tumbling), and $p^- = p^-(z,t)$ is the probability per unit time that a cell moving in the negative z direction becomes a cell moving in the positive z direction. The one-dimensional version of Equation (1) is obtained by adding together Equations (13) (Rivero et al. 1989) and recognizing that

$$c(z,t) = n^+(z,t) + n^-(z,t), \tag{14}$$

and

$$J_{cz}(z,t) = v[n^+(z,t) - n^-(z,t)]. \tag{15}$$

Comparing the one-dimensional cell density conservation equation thus obtained with the Keller and Segel equations, Equations (1) and (2), after some manipulation (involving neglect of the short time dynamics of the cell flux) we find that (Rivero et al. 1989)

$$\mu = \frac{v^2}{p^+ + p^-}, \quad V_{cz} = \frac{v(p^- - p^+)}{p^+ + p^-}, \tag{16}$$

thus enabling the macroscopic transport properties (random motility and chemotactic velocity) to be related to microscopic quantities.

Ford and Cummings (1992) recently showed that the Segel equations could be obtained from Alt's equations by assuming a Poisson process for tumbling and assuming that the bacterial motion was confined to one dimension. The latter assumption is unrealistic (one exception may be bacteria confined to capillaries whose diameter is of the same order of magnitude as the effective cell diameter),

and in fact the intended application of the Segel equation is to bulk systems (bacteria moving freely in three-dimensional space) subject to one-dimensional attractant gradients. In this case, physically we have symmetry in the x and y directions and it is a natural question to ask whether Alt's equations applied to bacteria subject to a one-dimensional gradient can be reduced to Segel's equations by integration over x and y. Ford and Cummings (1992) recently showed that only in the limit of infinitesimally small gradients could the Segel equations be obtained rigorously by integrating Alt's equations over x and y.

2.2.4. RTBL MODEL

The model of Rivero et al. (1989), hereafter referred to as the RTBL model, follows from the use of Equations (13) with p^+ and p^- given by

$$p^\pm = p_t^\pm p_r, \ p_t^\pm \ = \frac{1}{\langle \tau^\pm \rangle} \ , \ p_r = (1 - \psi)/2. \tag{17}$$

In this equation, p_r is the probability that a cell reverses direction after tumbling which, on the basis of experimental observations (Macnab 1980), is independent of the presence of an attractant or repellent, $p_t^\pm = p_t^\pm(z,t)$ is the probability of tumbling, $\langle \tau^\pm \rangle$ is the mean run time for a cell moving in the $\pm z$ direction, and ψ is the directional persistence. Note that the inverse relation between the tumbling probability and mean run length reflects the Poisson distribution of run times. A positive value for the directional persistence implies that a bacterium after tumbling is more likely to continue in the same direction as it was going before tumbling than to reverse direction. This is consistent with experimental observations (Berg and Brown 1972) in which the average angle between trajectories before and after tumbling for *E. coli* was found to be around 62°, compared to 90° which would be the average for random reorientations.

Berg and Brown (1972) also studied the effect of a chemical stimulus on individual cell paths and observed that the mean run times, $\langle \tau \rangle$, increased exponentially with the change in the number of receptor-attractant complexes, N_b, over mean run times measured in the absence of a chemical gradient, $\langle \tau_0 \rangle$. Based on this observation, the RTBL model for mean run time is given by

$$\langle \tau \rangle = \langle \tau_0 \rangle \exp \left[\nu \frac{DN_b}{Dt} \right] = \langle \tau_0 \rangle \exp \left[\nu \frac{\partial N_b}{\partial t} + \nu v \hat{s} \cdot \nabla_r N_b \right], \tag{18}$$

where ν is the differential tumbling frequency and D/Dt is the material (or substantial) derivative. The material derivative in Equation (18) is necessary to account for both temporal and spatial changes in attractant concentration sensed

by cells swimming at speed v (Berg and Brown 1972; Macnab and Koshland 1972; Spudich and Koshland 1975). The quantity N_b is related to the chemical attractant concentration, a, by the attractant/receptor binding equilibrium

$$N_b = \frac{N_T a}{K_d + a},$$

(19)

where K_d is the dissociation constant for attractant/receptor binding and N_T is the total number of receptors. For substitution into the one-dimensional Segel equations through Equation (17), Rivero et al. (1989) made use of a simplified version of Equation (18) given by

$$\langle \tau^{\pm} \rangle = \langle \tau_0 \rangle \exp \left[v \frac{\partial N_b}{\partial t} + v v \frac{\partial N_b}{\partial z} \right].$$

(20)

Using these constitutive relationships (Equations [17] to [20], Rivero et al. (1989) derived expressions for the random motility coefficient and the chemotactic velocity in terms of microscopic quantities as follows:

$$\mu_0 = \frac{v^2}{p_t(1 - \psi)}, \quad V_{cz} = v \tanh \left[\frac{\chi_0}{v} \frac{K_d}{(K_d + a)^2} \frac{\partial_a}{\partial z} \right],$$

(21)

where μ_0 is the random motility coefficient in the absence of an attractant gradient and χ_0 is referred to as the chemotactic sensitivity coefficient and is given by

$$\chi_0 = v v^2 N_T .$$

(22)

In the presence of an attractant gradient, the random motility μ exhibits a weak dependence on the gradient but for the conditions associated with the SFDC experiment it is adequately characterized by its gradient-independent value μ_0 (Ford and Lauffenburger 1991).

Thus, bacterial migration can be characterized by two fundamental transport coefficients, μ_0 and χ_0. The stopped flow diffusion chamber assay provides a method for measuring these coefficients for given bacterial species and attractants. This method is described in Section 4.3.

The macroscopic form of the RTBL model, which follows from ignoring some of the short term dynamics in the cell flux (Rivero et al. 1989), is given by

$$\frac{\partial c}{\partial t} = \mu_0 \frac{\partial^2 c}{\partial z^2} - \frac{\partial}{\partial z} (V_{cz} c),$$

(23)

where μ_0 and V_{cz} are given by Equation (21). The numerical solution of Equation (23) can be obtained by finite differences using a predictor-corrector method (Ford and Lauffenburger, 1991).

2.2.5. MODEL BASED ON REDUCED FORM OF ALT'S EQUATIONS

The RTBL model is based on the Segel subpopulation balance equations (Segel 1977) that can be derived from Alt's equations if the bacterial motion is confined to one dimension (z, in our notation). In reality, bacteria are free to move in three dimensions but symmetry exists in two of the three coordinate directions (x and y in our case). Therefore, the use of the RTBL model to describe bacterial migration has the inherent approximation of one-dimensional bacterial motion. To develop a model that did not involve this approximation, Ford and Cummings integrated Equation (8) over the x and y directions to obtain a reduced equation of the form

$$\frac{\partial n_z(z,\theta,t)}{\partial t} = -\frac{\partial v(z,t)n_z(z,\theta,t)}{\partial z} - \beta(z,\theta,t)n_z(z,\theta,t) \qquad (24)$$

$$+ \int_0^\pi \beta(z,\theta',t)n_z(z,\theta',t)K(\theta',\theta)\ sin\ \theta'\ d\theta',$$

where θ is the angle made by $\hat{s}$ with the z axis,

$$n_z(z,\theta,t) = \int_0^{2\pi} \int \int n(r,\hat{s},t)\ dx\ dy\ d\phi, \qquad (25)$$

and

$$K(\theta',\theta) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} k(\hat{s}',\hat{s})\ d\phi'\ d\phi. \qquad (26)$$

The relationship between $c(z,t)$ and $n_z(z,\theta,t)$ is given by

$$c(z,t) = \int_0^\pi n_z(z,\theta,t)sin\ \theta\ d\theta. \qquad (27)$$

Equation (24) was solved numerically by Frymier et al. (1994) using Galerkin's method on finite elements (Allaire 1985) for integration over the spatial variables z and θ. We briefly describe some aspects of this because the solution of Equation (24) represents an unusual application of the finite element technique. This is because it is an integropartial differential equation, in contrast to the typical

application to purely partial differential equations. A weighted implicit/explicit finite difference method was used for the integration over time. With the definition

$$F(z,\theta,t) = -s_z \frac{\partial v(z,t)n_z(z,\theta,t)}{\partial z} - \beta(z,\theta,t)n_z(z,\theta,t) \tag{28}$$
$$+ \int_0^\pi \beta(z,\theta',t)n_z(z,\theta',t)K(\theta',\theta)\sin\theta'\ d\theta',$$

and first order differencing in time, Equation (24) gives

$$\frac{n_z(z,\theta,t_m) - n_z(z,\theta,t_{m-1})}{\Delta t} = \omega F_m + (1-\omega)F_{m-1}, \tag{29}$$

where the subscript m indicates that F is evaluated at the current time step $t = t_m = t_0 + m\Delta t$ and the subscript $m-1$ indicates that F is evaluated at the previous time step. The weighting parameter ω can be varied from 1 (resulting in a fully implicit method) to 0 (resulting in a fully explicit method). The $z - \theta$ space was divided into 2000 rectangular elements. A bilinear function was used to approximate $n_z(z,\theta,t)$ on the elements, that is

$$n_z(z,\theta,t)^{(e)} = \gamma_1^{(e)} + \gamma_2^{(e)}z + \gamma_3^{(e)}\theta + \gamma_4^{(e)}z\theta, \tag{30}$$

where the superscript (e) denotes the element e. The constants γ_i, $i = 1, \ldots, 4$, are related to the (unknown) nodal values of the dependent variable $n_z(z,\theta,t)$. The interpolation function for $n_z(z,\theta,t)$ can also be approximated by the more convenient form

$$n_z(z,\theta,t)^{(e)} = G_i n_{z(i)} + G_j n_{z(j)} + G_k n_{z(k)} + G_l n_{z(l)}, \tag{31}$$

where the constants $n_{z(\alpha)}$, $\alpha = i, j, k, l$ are the values of $n_z(z,\theta,t)$ at the "nodes" of the element e and G_α, $\alpha = i, j, k, l$ are called the shape functions of e. The shape functions are related to γ_i, $i = 1, \ldots, 4$ and are constructed so that $n_z(z,\theta,t)$ varies linearly along the sides of the element and takes on its nodal values at the nodes $i, j, k,$ and l. For each element, the integration over θ' indicated in Equation (24) was calculated using a two-point Gaussian integration that, because of the small variation of the integrand over the width of an element in θ space $(\pi/10)$, was found to be sufficiently accurate. The values of θ and θ' for which the direction change is needed are therefore known at the beginning of the solution method, enabling $K(\theta',\theta)$ to be calculated at the outset of the solution procedure and tabulated so that it need not be recalculated at each time step. A typical finite element solution to the balance equation for one-dimensional attractant gradients requires about 40 CPU minutes on an IBM RS/6000 Powerstation 320.

Numerical solutions of Equation (24) using the finite element method are presented in Section 4.

3. Cellular Dynamics Simulation

The cellular dynamics (CD) simulation methodology was developed by us as a method for studying the collective transport properties of populations of bacteria essentially by simulating the dynamics of large populations of individual bacteria (Frymier et al. 1993). It thus shares considerable common philosophical ground with molecular dynamics simulations methods used to predict the many-body thermophysical properties of liquids (Allen and Tildesly 1987). In essence, the stochastic differential equation that describes the dynamics of an individual bacterium is solved for 10^4 to 10^5 bacteria in a geometry appropriate to a small subvolume ($1 \times 1 \times 8$ mm) of the SFDC located at its center with the long axis parallel to the coordinate direction (z) in which there is an attractant gradient, as shown in Figure 11.2.

Computer simulations of the motion of individual bacteria have been reported by several researchers. Berg (1988, 1993) performed simulations to illustrate the random walks generated by single cells assuming Poisson statistics for the tumbling probability and a simple approximation to the bias in the turn angle distribution. Bornbusch and coworkers (1984, 1986) investigated the effect of limiting the turning field size of a cell on its ability to locate an attractant source. Tankersley and Conner (1990) performed simulations of single-cell migration to illustrate how the differences between various cell types in the mechanisms used to move toward an attractant source resulted in markedly different patterns of migration. Unlike these computer simulations, we simulate a large population of bacteria (20,000 or more cells) and incorporate the experimentally measured turn angle distribution into the mechanism for reorienting the cells after tumbling.

The logic diagram in Figure 11.5 summarizes the cellular dynamics simulation methodology described below. The simulation box is first initialized. At the start of the simulation loop, each bacterium decides whether or not to tumble (according to Equations [36] and [37]). Based on whether or not it tumbles, either a new direction is chosen or the bacterium continues its run in the same direction. Finally the bacterium moves to a new position based on its swimming speed and direction vector using an Euler integration of the equation of motion. We now describe the implementation of these steps.

The simulation "box" of size $h \times w \times l$ is given an initial uniform random distribution of bacteria, both in terms of the spatial distribution of the bacteria as well as the direction of motion for each bacterium. The initial concentration $c(r,t = 0)$ (see Fig. 11.2) is given by
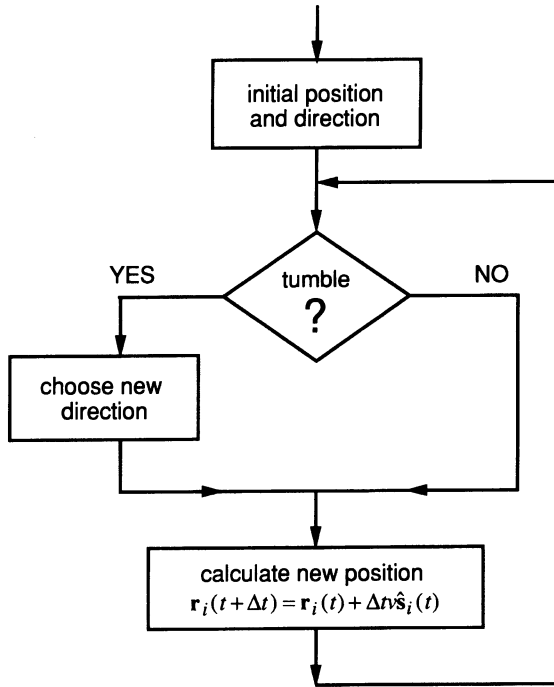
Figure 11.5 The motion of bacteria is simulated by performing three calculations for each individual bacterium at each time step. First, based on a probability that is dependent on the local attractant gradient, it is stochastically determined whether or not the cell tumbles. Second, if the cell tumbles, then its new direction is chosen using an experimentally measured turn angle distribution. If it does not tumble, its direction is maintained. Third, the cell's new position is computed based on its direction vector, swimming speed, and the time step.

$$(32)$$

$$c(\mathbf{r}, t = 0) = c_0, \quad \begin{aligned} -\frac{h}{2} &\leq x \leq \frac{h}{2} \\ -\frac{w}{2} &\leq y \leq \frac{w}{2} \\ -\frac{l}{2} &\leq z \leq \frac{l}{2} \end{aligned}.$$

An initial bacterial concentration of $c_0 = 2 \times 10^7$ cells/cm^3 was used in the experiments of Ford et al. (1991); however, for the simulation results presented here, we used an initial concentration of bacteria of $c_0 = 2 \times 10^6$ cells/cm^3 to reduce the calculations necessary at each time step by an order of magnitude.

The bacteria are assigned initial direction vectors at random. A step change along the z axis at time $t = 0$ in the concentration of an attractant, a, exists such that

$$a(z,t = 0) = 0, \; -l/2 \leq z \leq 0, \tag{33}$$
$$a(z,t = 0) = a_0, \; 0 \leq z \leq -l/2,$$

$$a(z,t) = \frac{a_0}{2} \left[1 + \text{erf} \left(\frac{z}{\sqrt{4Dt}} \right) \right], \; t > 0, \; -l/2 \leq z \leq l/2, \tag{34}$$

where a_0 is the initial attractant concentration introduced into one half of the chamber, D is the diffusion coefficient of the attractant, z is the position along the z axis, and t is time. The solution given above for $t > 0$ can be found in Crank (1979). (Strictly speaking, Equation [34] is the solution of the diffusion equation for $l \rightarrow \infty$. For practical purposes, over the course of a simulation, the attractant concentration goes to the limiting value a_0 at $z = \pm l/2$, so that Equation [34] is valid for our application.) For $t > 0$, the motion of each bacterium is modeled as an independent, three-dimensional, biased random walk. Our assumption that the motion of each bacterium can be considered independent of the motion of other bacteria is justified by noting that the cell densities used in the SFDC experiments are low enough that intercellular distances are at least an order of magnitude greater than cellular diameters. However, at much higher cell densities (by an order of magnitude or more), hydrodynamic interactions among swimming bacteria may be significant (Guell et al. 1988), in which case our simulation technique would require additional terms in the equations of motion to account for hydrodynamic interactions between the bacteria.

The position r of the bacterium i at time $t + \Delta t$, where Δt is the time step in the simulation, is given by

$$r_i(t + \Delta t) = r_i(t) + v\hat{s}_i\Delta t, \tag{35}$$

where v is the three-dimensional swimming speed of the bacterium and $\hat{s}_i$ is the unit vector in the direction of motion of the bacterium. The simulation box employs periodic boundary conditions (Allen and Tildesly 1987). If the cell moves out of the box, it is repositioned so that it reenters at the opposite wall, retaining a constant number of cells in the simulation box. The length of the simulation box in the direction of the attractant gradient (the z direction) is 0.8 cm. This is sufficiently large to ensure that the attractant concentration is uniform at the ends of the box ($z = \pm 0.4$ cm), at which points the cell density returns to the bulk density c_0.

The unit direction vectors for the bacteria evolve in time according to

$$\hat{s}_i(t + \Delta t) = \lambda_i \hat{s}_i(t) + (1 - \lambda_i)\hat{s}_i', \tag{36}$$

where

$$\lambda_i = 1 \text{ if } \rho_i \geq p_t\Delta t \tag{37}$$
$$= 0 \text{ if } \rho_i < p_t\Delta t.$$

In Equation (37), ρ_i is a random number chosen at time $t + \Delta t$ from a uniform distribution on $[0,1]$ so that a bacterium has probability $p_t\Delta t$ of tumbling and probability $(1 - p_t\Delta t)$ of continuing its run. Thus, the tumbling probability in a zero gradient is constant and consistent with a Poisson distribution. The time step used in the simulation is $\Delta t = 0.1$ sec, the time corresponding to the experimentally observed duration of a tumble. Therefore the tumbling process is effectively instantaneous in our simulation. Details on the method of choosing $\hat{s}_i'$ in Equation (36) using the experimentally determined turn angle probability are provided in the appendix of Frymier et al. (1994).

Figure 11.1 shows the paths of 100 bacteria generated by applying the simulation algorithm in the absence of a chemical gradient (left column) and in the presence of a linear attractant gradient (right column). Note that periodic boundary conditions were not implemented in this sample trace. The cell swimming speed (22 μm/sec) and zero gradient tumbling probability (0.17 sec^{-1}) are the same as for the simulations reported in Section 4, where results of the cellular dynamics simulation of the SFDC assay are compared with the solution of Equation (24), the reduced form of the Alt equations.

4. Comparison of Modeling Approaches

In this section, we compare the finite element method (FEM) solution of the reduced Alt (RA) model, Equation (24), with results obtained from cellular dynamics (CD) simulations (Section 4.1), with the one-dimensional RTBL model (Section 4.2) and with experimental data for *E. coli* responding to fucose and α-methylaspartate in the SFDC assay (Section 4.3).

Before doing so, it is instructive to summarize the relationship between the three approaches being used to describe bacterial migration in this paper, RTBL, the FEM solution to the balance equation for one-dimensional attractant gradients and CD simulations. All three approaches have the same conceptual basis: that tumbling in bacteria is governed by a stochastic Poisson process and is related to the attractant gradient by the experimentally derived relationship Equation (18). The CD simulation method is, in essence, a brute force Monte Carlo method for solving Alt's equations, Equations (6) and (7). The FEM solution is for the

RA equation derived by Ford and Cummings (1992) for the case of a one-dimensional attractant gradient by exploiting symmetry in two of three coordinate directions. Thus, CD and FEM are alternative methods for solving the same model—that is, the same equations for the same physical situation—and thus should yield equivalent results. Note that the CD method involves the solution of the dynamical equations for individual cells and thus, unlike the FEM solution, is unable to take advantage of the overall symmetry of the population as it responds to a one-dimensional attractant gradient. This is one sense in which the FEM solution of the RA model is more efficient. By contrast to the model underlying the CD simulation and the RA model, the RTBL model involves an additional physical assumption, namely that the bacteria are confined to one dimension in their motion (Ford and Cummings 1992).

Hence, the purpose of these comparisons is to validate the FEM solution of the RA model by comparing it to CD simulations, evaluate the accuracy of the RTBL model compared to the more rigorous RA model, and to evaluate the accuracy of the physical model embodied in both CD simulations and the RA model by comparing the RA theoretical predictions against experimental data.

4.1. Comparison of Reduced Alt Model to Cellular Dynamics Simulation

We begin with a comparison of the bacterial density profiles generated by the RA model, CD simulation and RTBL model for the set of conditions shown in Table 11.1. The attractant concentration, diffusivity and dissociation constant (a_0, D, and K_d) and bacterial properties (β_0, v, and p_r in Table 11.1 correspond to those in SFDC experiments measuring the response of *E. coli* bacteria to a gradient of fucose reported by Ford et al. (1991). Note that the RTBL quantities for cell speed v and chemotactic sensitivity coefficient χ_0 in Table 11.1 are one-dimensional quantities and are related to the corresponding three-dimensional quantities used in the CD simulation and RA model calculations according to the relationships derived by Ford and Cummings (Ford and Cummings 1992):

$$v^{1D} = \frac{v}{2},$$

(38)

which leads to the following relationship between χ_0^{1D} and χ_0 (Ford and Cummings 1992):

$$\chi_0^{1D} = \frac{\chi_0}{4}.$$

(39)

These relationships were verified by Frymier et al. (1992) as being essential to obtain consistency between the one-dimensional RTBL and three-dimensional CD descriptions. The value $\chi_0 = 3.5 \times 10^{-4}$ cm^2/sec (equivalent to $\chi_0^{1D} = 0.88 \times 10^{-4}$ cm^2/sec) corresponds to the experimentally measured value for this system. The value of $\chi_0 = 105 \times 10^{-4}$ cm^2/sec is significantly higher, but is consistent with values measured experimentally for bacteria cultured at low growth rates (Mercer et al. 1993).

In Figures 11.6 and 11.7, the FEM solution of the RA model is compared to the CD simulation and to the solution of the RTBL model for the conditions given in Table 11.1 over a period corresponding to 6 min of elapsed time. Equation
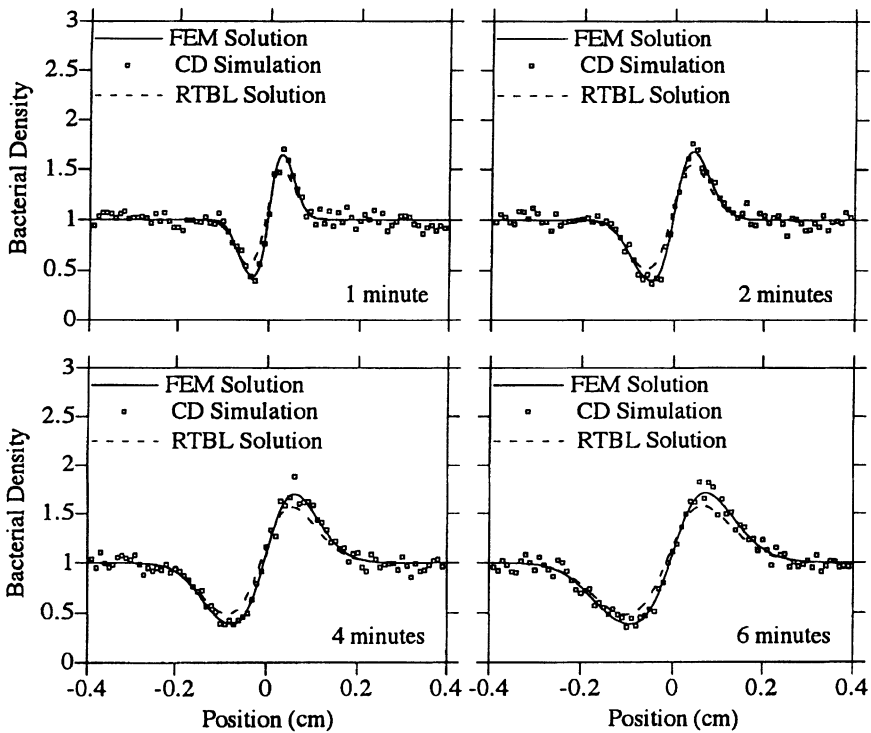
Figure 11.6 Comparison of the finite element solution of the reduced form of Alt's equations for a one-dimensional gradient (solid line), CD simulation (squares), and the RTBL model (dashed line) for $\chi_0 = 3.5 \times 10^{-4}$ cm^2/sec. The dimensionless bacterial density c/c_0 is plotted as a function of the position along the SFDC for times of 1, 2, 4, and 6 min. The position $z = 0$ corresponds to the position of the initial step change in the attractant concentration at $t = 0$ with a fucose concentration of 0.2 mM initially in the bottom of the SFDC ($0 < z \leq 4$ cm in the graphs).
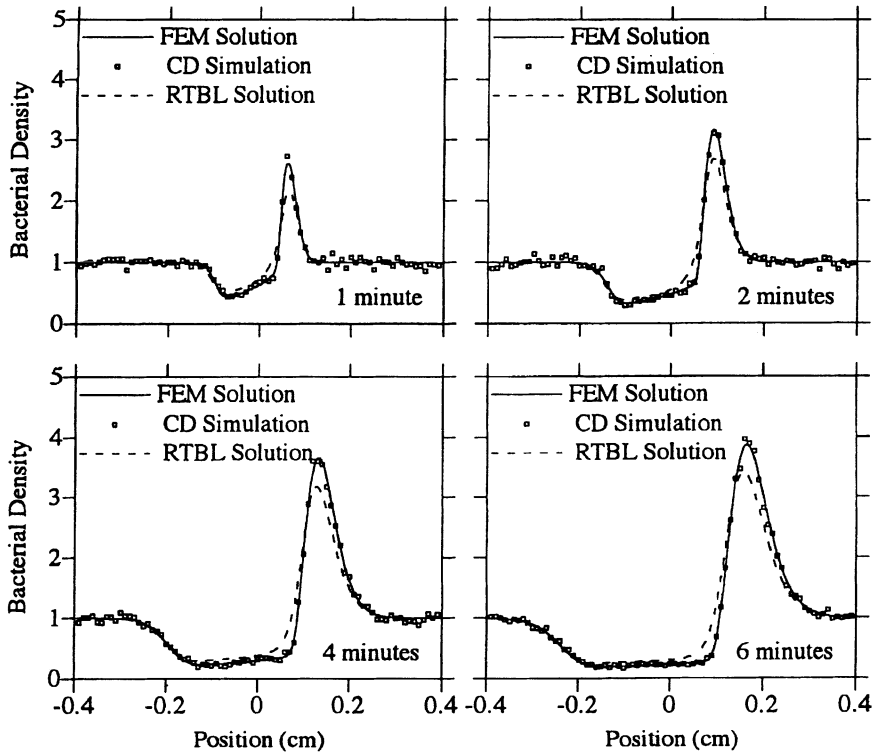
Figure 11.7 Comparison of the finite element solution of the reduced form of Alt's equations for a one-dimensional gradient (solid line), CD simulation (squares), and the RTBL model (dashed line) for a large value of $\chi_0 = 105 \times 10^{-4}$ cm^2/sec.

(27) corresponds to the bacterial density that is observable experimentally. This density is plotted in dimensionless form as a function of position. The dimensionless bacterial density is defined as c/c_0, where c_0 is the initial uniform bacterial density.

The model for the tumbling probability used in the FEM solution and the CD simulations is Equation (18) and in the RTBL model is Equation (20), both with the time derivative $\partial a/\partial t$ term neglected, an assumption validated by Frymier et al. (1994). This model for the tumbling probability is consistent with that used in previously reported numerical studies (Ford and Lauffenburger 1991).

The stochastic differential equation solved in the CD embodies the same individual cell dynamics as are assumed in Alt's equations and the RA model. Hence, from the mathematical point of view, the density profiles obtained from the CD simulations should be the same as those obtained from the FEM solution

Table 11.1 Summary of constants used in the CD simulation and in solving the RA and RTBL equations for comparisons shown in Figures 11.6 and 11.7.

Quantity	β_0 (sec^{-1})	v (μm/sec)	a_0 (mM)	χ_0 ($\times 10^4$ cm^2/sec)	D ($\times 10^6$ cm^2/sec)	K_d (mM)	p_r
RA and CD	0.17	22.	0.2	3.5, 105	6.9	0.08	—
RTBL	0.17	11.	0.2	0.88, 26	6.9	0.08	0.32

for the RA model provided the same boundary conditions (i.e., SFDC geometry), operating conditions and sensing mechanisms are used in both cases. It is clear from Figures 11.6 and 11.7 that the FEM solutions of the RA model are completely consistent with the CD simulations as expected. The advantage of the FEM solution to Equation (24) over CD simulation is that the numerical computation involved is considerably less. This is because the error in CD simulations is $O(N^{-1/2})$, whereas the computation time is $O(N)$ where N is the number of bacteria simulated. Obtaining results from CD simulation with errors similar to those of FEM solution of the RA model (around 1%) would be computationally expensive. For example, with 20,000 bacteria the noise in CD appears to be around 5%, so reducing this to 1% would require 25 times this number (or 500,000) bacteria and would require 26 hours of CPU time on a IBM RS/6000 Powerstation 320 for a simulation corresponding to 6 min of real time. For comparison, the FEM solution of the RA model requires 39 min of CPU time on an IBM RS/6000 Powerstation 320 for a calculation corresponding to 6 min of real time. The advantage of the CD method is that it enables visualization of the individual bacterial motion as well as the collective motion of the population.

4.2. Comparison of Reduced Alt Model to RTBL Model

From Figures 11.6 and 11.7, it is clear that there is a small but discernible, quantitative difference between the predictions of the RTBL model and the solution of the three-dimensional models (RA model and CD simulations). In view of the relative speed of the numerical solution of the RTBL equation (1 min of CPU time on an IBM RS/6000 Powerstation 320 for a calculation corresponding to 6 min of real time compared to 39 min for the analogous FEM solution of the RA model), it might be tempting to regard the difference as inconsequential for practical purposes and compensated for by the more rapid computational speed of RTBL. However, as Frymier et al. (1994) show, the small differences in the density profiles in the RTBL model lead to significant overestimates in the values of the chemotactic sensitivity coefficient χ_0 inferred by comparison with experimental data, with overestimates ranging from 30% for $\chi_0 = 3.5 \times 10^{-4}$ cm/sec to 230% for $\chi_0 = 105 \times 10^{-4}$ cm^2/sec. These overestimates

are with respect to values obtained from the RA model and CD simulations, which are assumed to be physically correct.

4.3. Comparison of Reduced Alt Model to Experimental Results

Having established the correctness of the FEM solution of the RA model by comparison to CD simulation, it is clear that the FEM solution provides a convenient and computationally inexpensive route to calculating the bacterial migration profiles for bacteria subject to one-dimensional attractant gradients. It is thus the appropriate vehicle for comparison with experimental measurements, particularly for determining values of transport coefficients.

The model for the tumbling probability used in the previous two sections is

$$\beta(r,\hat{s},t) = \beta_0 \exp(-\epsilon), \tag{40}$$

where $\beta = 1/\langle\tau\rangle$, β_0 is the tumbling probability in the absence of an attractant gradient and

$$\epsilon = \frac{\chi_0}{v} \frac{K_d}{(K_d + a)^2} \hat{s} \cdot \nabla a. \tag{41}$$

This implies that bacteria decrease their tumbling frequency when moving toward an attractant and increase their tumbling frequency when moving away from an attractant. Depending on the swimming behavior of the particular bacterial species, an alternate form of Equation (40) may be appropriate, which is given by

$$\begin{aligned}\beta(r,\hat{s},t) &= \beta_0 \exp(-\epsilon) \quad \epsilon > 0 \\ &= \beta_0 \quad\quad\quad \epsilon < 0.\end{aligned} \tag{42}$$

That is, for some bacterial species (which include *E. coli*) the probability of tumbling does not increase when the bacteria are moving in the direction of a decreasing attractant gradient but simply returns to β_0, the basal tumbling frequency in the absence of a gradient (Berg and Brown 1972).

In order to compare the RA model with experiment, it is necessary to determine the fundamental transport parameters, the random motility μ_0 and the chemotactic sensitivity coefficient χ_0. We now describe how this is done.

First, a random motility experiment is performed in the SFDC in which, under flow conditions, the lower half of the chamber contains only motility buffer and the upper half contains buffer and bacteria. Hence, there is a step change in bacterial density at the midpoint of the SFDC. Once flow is stopped, the bacteria exhibit net migration into the lower half of the chamber where initially no bacteria

were present. Bacterial density is measured by light scattering to determine $N(t)$, the number of bacteria entering the lower half of the SFDC. Experimentally, it is observed that $N(t) \propto t^{1/2}$. From the mathematical modeling point of view, all the models described in Section 2 predict that the macroscopic diffusion equation for the bacterial density,

$$\frac{\partial c}{\partial t} = \mu_0 \frac{\partial^2 c}{\partial z^2}, \tag{43}$$

will hold for this random motility experiment. The solution of this equation for the initial conditions of the random motility experiment also predicts $N(t) \propto t^{1/2}$ with a slope related to μ_0. Figure 11.8a shows the experimental result for $N(t)$ in the case of *E. coli*. The random motility is determined by the value of μ_0 which leads to the best fit of the experimental data. The random motility is used to determine the tumbling probability in the absence of an attractant gradient β_0 based on the rigorous connection between μ_0, β_0, the cell swimming speed v, and directional persistence $\psi = \langle \cos \varphi \rangle$ (where φ is the angle between successive runs) given by (Lovely and Dahlquist 1975)

$$\mu_0 = \frac{v^2}{3\beta_0(1 - \psi)}. \tag{44}$$

For *E. coli*, we use the values of $v = 22$ μm/sec and $\psi = 0.32$ measured by Berg and Brown (1972) to obtain β_0 from the experimentally determined random motility.

Knowledge of β_0 is necessary for the next step, determination of the chemotactic sensitivity coefficient χ_0. Under flow conditions at the beginning of a chemotaxis experiment, the lower half of the SFDC contains motility buffer, bacteria, and attractant, whereas the upper half of the SFDC contains buffer and an equal concentration of bacteria but no attractant. Hence, there is a uniform density of bacteria in the SFDC and a step change in the concentration of the attractant. Once flow is stopped, the attractant diffuses upward according to Fickian diffusion and the bacteria exhibit net migration into the lower half of the SFDC chamber in response to the attractant gradient. The density of the bacteria is again measured by light scattering, and $N(t)$, the number of bacteria entering the lower half of the SFDC, obtained. Experimentally, it is found that $N(t) \propto t^{1/2}$, and this is also found to be the case from the solution of the RA model with a slope that increases monotonically with χ_0. The inferred value of χ_0 from the experiment is that which yields the same theoretical slope of $N(t) \propto t^{1/2}$ as that given by a linear least-squares regression of the experimental data. This is illustrated in Figure 11.8b, where experimental data on *E. coli* responding to fucose is used to determine χ_0.

Frymier et al. (1994) investigated the degree to which the use of Equation
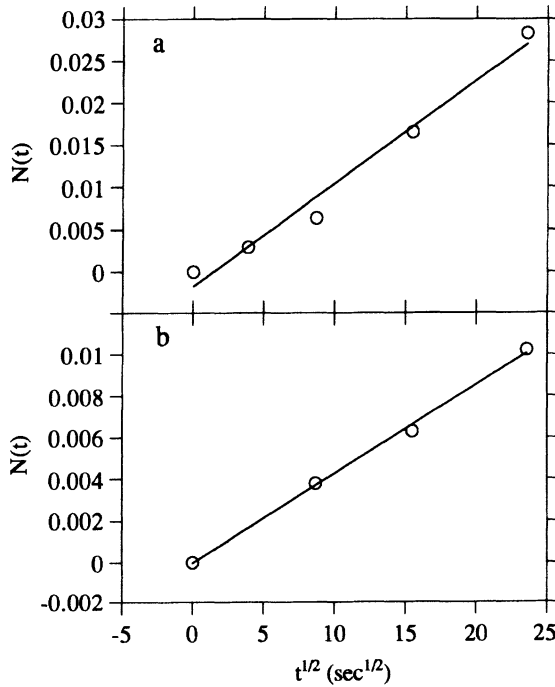
Figure 11.8   (a) Graph of $N(t)$, the number of bacteria entering the bottom half of the SFDC, determined experimentally as the area under the bacterial density profile, plotted as a function of $t^{1/2}$ from a random motility experiment. A linear regression analysis of the data produced the solid line shown in the graph with a slope of 1.22×10^{-3} cm/sec$^{1/2}$ and a correlation coefficient r of 0.98952. The resulting random motility coefficient is $\mu_0 = (4.7 \pm 0.8) \times 10^{-6}$ cm^2/sec. (b) Graph of the area of the bacterial density peak as a function of $t^{1/2}$. The data are from a chemotaxis experiment with an initial attractant concentration of 0.1 mM fucose. A linear regression analysis of the data produced the solid line shown in the graph with a slope of 4.29×10^{-4} cm/sec$^{1/2}$ and a correlation coefficient r^2 of 0.99754. The chemotactic sensitivity coefficient χ_0 is determined by matching the experimental slope to theoretical predictions over a range of χ_0 values. For this set of data, $\chi_0 = (3.9 \pm 0.1) \times 10^{-5}$ cm^2/sec.

(40) for the tumbling probability, rather than the more correct Equation (42), affected the experimental determination of χ_0. Specifically, for *E. coli* responding to α-methylaspartate, they found that Equation (40) for the tumbling probability leads to a value of $(1.9 \pm 0.1) \times 10^{-4}$ cm^2/sec, whereas the use of the more correct Equation (42) (bacteria moving against the attractant gradient return to the basal tumbling frequency) yields $(4.1 \pm 0.2) \times 10^{-4}$ cm^2/sec (Strauss et al. 1995). Therefore, determination of χ_0 from experimental data requires *a priori* knowledge of whether Equation (40) or Equation (42) applies for the particular bacteria under investigation. We note that the 100% increase in χ_0 obtained using Equation (42) rather than Equation (40) is not a general result but is consistent with the small values of χ_0 involved. It can be shown that in the limit of small ϵ, in which case the exponentials in the tumbling frequency equations can be linearized, this factor of two naturally arises using the perturbative expansion of Ford and Cummings (1992).

Figures 11.9 and 11.10 show the experimental data for *E. coli* responding to fucose and α-methylaspartate in SFDC experiments (Ford et al. 1991; Strauss 1992; Strauss et al. 1995) conducted under the conditions given in Table 11.2. Also shown in Table 11.2 are the random motility and chemotactic sensitivity coefficients measured in the SFDC. The theoretically obtained bacterial density profiles from the RA model are also shown in the figures. The agreement between theory and experiment is very good, which suggests that the major assumptions in the development of the RA model are reasonable.

5. Application to Multiple Stimuli

In this section, we extend the reduced Alt model to include experiments involving two chemical attractant gradients. We propose three models for describing chemotactic signals generated from two different gradients into the constitutive equation for the tumbling probability. By comparing the bacterial density profiles from the theoretical predictions with experimental data from the SFDC assay for two chemical stimuli we can infer some information about how bacteria process multiple signal inputs from their surroundings to produce a single output response.

Although the biochemical pathway for communicating concentration gradient information to the flagellar motors—in particular, signal integration and regulation—is not completely understood, the primary steps in the pathway are well known from studies of nonchemotactic mutants of *E. coli* and *Salmonella typhimurium*. Recent reviews of the signal transduction mechanism include those by Stock et al. (1989) and Eisenbach (1991). Transmembrane proteins convey information from the environment across the cytoplasmic membrane to the interior of the cell. Located on the periplasmic (exterior) side of the transducer proteins
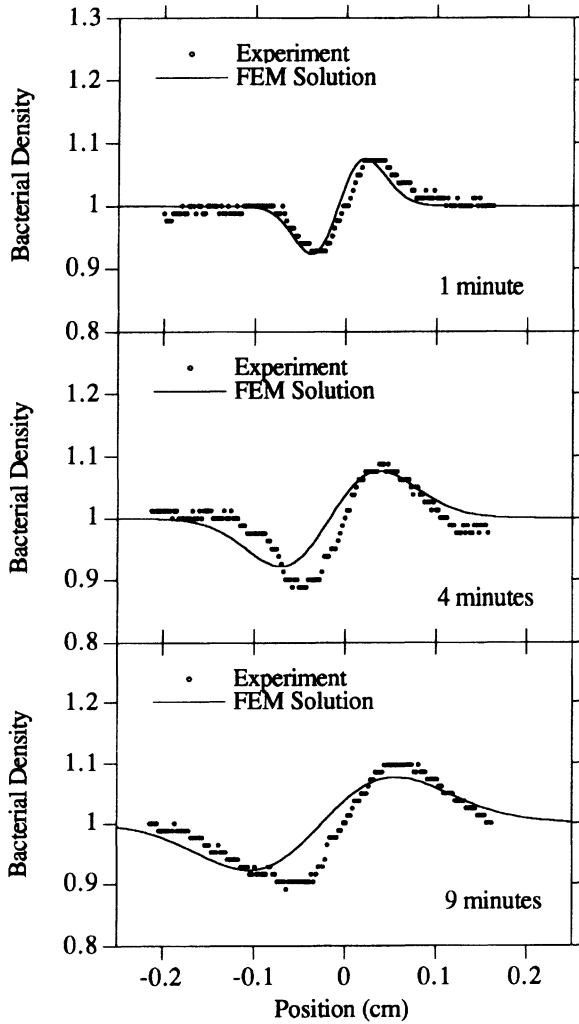
Figure 11.9 Bacterial density profiles from the SFDC assay for chemotaxis experiments (circles) with an initial attractant concentration of 0.1 mM fucose at times of 1, 4, and 9 min. Fucose was initially present in the bottom of the chamber, which corresponds to positions greater than zero on the plot. Theoretical profiles (FEM solution) are superimposed with $\chi_0 = 3.9 \times 10^{-5}$ cm^2/sec (solid line) for comparison.
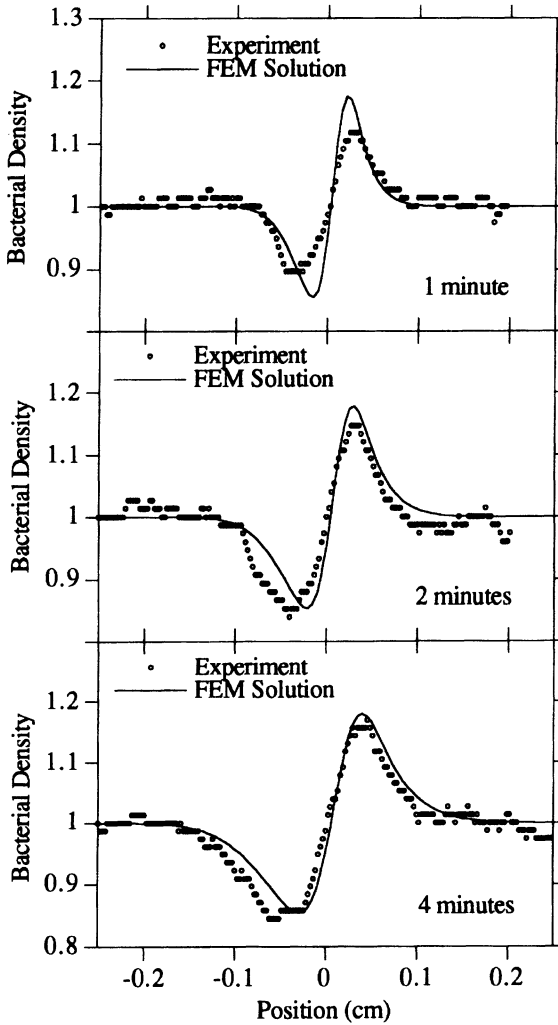
Figure 11.10 Bacterial density profiles from the SFDC assay for chemotaxis experiments (circles) with an initial attractant concentration of 0.01 mM α-methylaspartate at times of 1, 2, and 4 min. Methylaspartate was initially present in the bottom of the chamber, which corresponds to positions to the right of zero on the plot. Theoretical (FEM solution) profiles are superimposed with $\chi_0 = 4.1 \times 10^{-5}$ cm^2/sec (solid line) for comparison.

Table 11.2 Conditions for experimental measurement of the response of *E. coli* to fucose (Ford et al. 1991) and α-methylaspartate (Strauss 1992)

Quantity	v (μm/sec)	a_0 (mM)	μ_0 ($\times 10^7$ cm²/sec)	D ($\times 10^6$ cm²/sec)	K_d (mM)	χ_0 ($\times 10^5$ cm²/sec)
Fucose	22	0.1	47 ± 8	6.9	0.08	3.9 ± 0.1
α-Methyl-aspartate	22	0.01	8.8 ± 3.8	8.6	0.125	41 ± 2

are specific binding sites for chemoattractants and repellents. The binding event on the periplasmic portion of the transducer induces a structural change on the cytoplasmic (interior) side, which initiates a cascade of phosphorylation reactions that control the flagellar rotation and ultimately the tumbling behavior.

Our experimental system, *E. coli*/fucuse/α-methylaspartate, represents two attractants that utilize different transmembrane proteins for signal transduction. Fucose binds to the galactose-binding protein, which then binds to the Trg transmembrane protein to initiate the chemotactic signal. The attractant α-methyl-aspartate is a nonmetabolizable structural analog to the amino acid aspartate and binds directly to the Tar transmembrane protein.

5.1. Tumbling Probability Models

The general equation describing the tumbling probability, $\beta(z,\theta,t)$, in the presence of a one-dimensional attractant gradient for a single attractant is given by Equation (42) and Equation (41) in Section 4.3. We consider three hypothetical models for the tumbling probability $\beta(z,\theta,t)$ in the presence of multiple stimuli: (1) response only to the attractant present with the highest chemotactic sensitivity χ_0; (2) response to the maximum individual chemotactic signal ϵ at any one time; and (3) response proportional to simple additivity of individual chemotactic signals, $\sum \epsilon_i$. These models can be compared with mixture rules used in thermodynamics to predict the thermophysical properties of fluid mixtures based on measurements of the properties for pure substances (Sandler 1977). In our experimental system, the first model, "high sensitivity (HS)," would use the tumbling probability equation for a single attractant Equation (42) with parameter values for α-methylaspartate, since the chemotactic sensitivity of α-methylaspartate is significantly higher than that for fucose. This model implies that the chemotactic response is controlled by a single attractant, which is preferred by the bacteria and all other attractants are ignored. The second model, "maximum signal (MS)," is based on the hypothesis that bacteria respond to the strongest chemotactic signal, which combines both the chemotactic sensitivity and the strength of the gradient. This differs from the HS model because bacteria will respond to a

fucose signal in the presence of methylaspartate if the fucose gradient is large enough to compensate for a lower chemotactic sensitivity coefficient. In terms of the tumbling probability, this model is defined by

$$\beta(z,\theta,t) = \begin{cases} \beta_0\exp[-\max(\epsilon_i)], & \max(\epsilon_i) > 0 \\ \beta_0, \max(\epsilon_i) < 0, \end{cases} \tag{45}$$

where

$$\epsilon_i = \frac{\chi_{0,i}}{v} \frac{K_{d,i}}{(K_{d,i} + a_i)^2} \hat{s} \cdot \nabla a_i . \tag{46}$$

The third model, "simple additivity (SA)," suggests that the responses are additive as suggested by Rubik and Koshland (1978) for *S. typhimurium* ST171/aspartate/serine and *S. typhimurium* ST171/aspartate/ribose. The tumbling probability equation for this model is written as

$$\beta(z,\theta,t) = \begin{cases} \beta_0\exp\left[\sum_i(-\epsilon_i)\right], & \sum_i(-\epsilon_i) > 0 \\ \beta_0, & \sum_i(-\epsilon_i) < 0. \end{cases} \tag{47}$$

5.2. Experimental Observations

Our experimentally observed responses of *E. coli* to multiple gradients of fucose and α-methylaspartate are shown in Figure 11.11 for the co-directional configuration and the antidirectional configuration. A codirectional configuration refers to fucose and α-methylaspartate both in the bottom half of the SFDC and antidirectional refers to fucose only in the bottom and α-methylaspartate in the top. If we compare the gradients relative to the respective dissociation constants that correspond to the concentrations evoking the strongest response, fucose is greater by a factor of 5. However, bacterial bands were observed to move toward α-methylaspartate and away from fucose in the antidirectional configuration indicating that the greater chemotactic sensitivity of α-methylaspartate more than compensated for a larger gradient. From qualitative visual observations during the experiments it also appeared that the magnitude of the chemotactic response of the bacteria in the presence of both attractants was much greater than the single stimulus responses to α-methylaspartate and fucose. Nevertheless, it is difficult to compare directly because the random motility of the population used for the experiment to measure the response to α-methylaspartate alone was significantly less than that for the multiple-stimuli experiment.
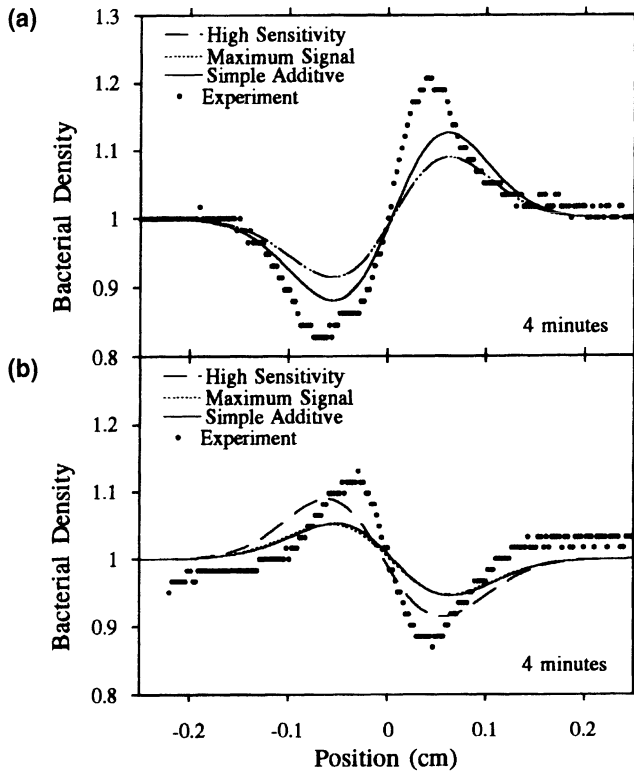
Figure 11.11 (a) Bacterial density profiles for multiple stimulus experiments (circles) with an initial attractant concentration of 0.01 mM α-methylaspartate and 0.032 mM of fucose compared to theoretical predictions based on the HS (long-dashed line), MS (short-dashed line), and SA (solid line) models for the codirectional configuration. The HS and MS model predictions are identical and appear as a line with a combination of long and short dashes. (b) Bacterial density profiles for multiple stimulus experiments compared to theoretical predictions for the antidirectional configuration.

The one direct experimental comparison that can be made is between the results for the co- and antidirectional configurations of the two attractants. For these cases, the fucose concentrations and random motility coefficients are the same. According to the HS model, bacteria would only respond to the attractant with the larger chemotactic sensitivity coefficient, in our case α-methylaspartate. The presence of the less-preferred attractant, fucose, would be ignored. We would then expect the bacterial density profiles to be mirror images of each other because the α-methylaspartate gradient is identical except it is initiated from

opposite sides of the chamber. It is clear from Figure 11.11 that the profiles differ significantly and therefore, we conclude that the HS model does not accurately model the multiple gradient response.

Because the fucose concentrations were different between the single stimulus and multiple stimuli experiments and the random motility coefficients were different for each of the experiments, it would not be possible to properly evaluate the physical models from the experimental profiles alone. However, because we have a mathematical model that describes the physical model, we can generate theoretical predictions for any set of experimental conditions thereby permitting us to evaluate the validity of the physical model.

5.3. Comparison between Theoretical Models and Experiments

We compare theoretical predictions of the bacterial density profiles for the three different tumbling probability models to each other and to the experimental data for the codirectional and antidirectional configurations in Figure 11.11. The comparisons are shown at 4 min because the differences between the models and the experimental data are more pronounced at the later times. The same qualitative trends were observed at 1 and 2 min.

In assessing the MS model, we note that for the codirectional gradients, the α-methylaspartate signal (as defined by Equation [41]) will always be stronger because of the 10-fold larger chemotactic sensitivity coefficient although the fucose gradient is initially larger in magnitude by a factor of three. Therefore, in a codirectional configuration we expect multiple gradient response to be very similar to the single stimulus response to α-methylaspartate (and to the HS model). As expected, the MS model is identical to the HS model for this configuration and underpredicts the response to multiple gradients. For the antidirectional configuration, bacteria moving against the α-methylaspartate gradient would respond to the fucose stimulus instead of returning to their basal tumbling frequency. This would increase the run lengths for bacteria moving away from the α-methylaspartate and tend to reduce the number of bacteria crossing the interface and hence reduce the peak area of the bacterial density profile as compared to the HS model and α-methylaspartate alone. So, for the antidirectional configuration the HS model agrees more closely with the experimental data than the MS model, but both significantly underpredict the observed response.

The SA model predicts a response to the codirectional gradient equivalent to the sum of the individual responses to fucose and α-methylaspartate at the same initial concentrations. In comparing the models, we see a larger response for the SA model due to the inclusion of a response to fucose, which was ignored in the other two models. Even with this increase accounting for a response to fucose, the theoretical prediction still falls short of the experimental data. For the antidirectional configuration, the response should still be toward the α-methyl-

aspartate side, but with a reduction in peak area related to the fucose peak area for a single stimulus experiment at 0.032 mM. For this configuration, the HS model more closely represents the experimental data, although based on previous arguments the HS model is clearly incorrect.

5.4. Evaluation of Proposed Models

The work of Tsang et al. (1973) and Adler and Tso (1974) with attractants and repellents eliminated the dominance of an attractant over a repellent and vice versa. The mere presence of an attractant was not sufficient to overcome the response to a repellent; a gradient in the attractant was required. Our results for two attractants are also consistent with the elimination of the HS model. Although this model provides the closest fit to the experimental data for the antidirectional case, it is clear from the experimental data that the presence of the fucose causes a diminished response to the α-methylaspartate in the antidirectional configuration in comparison to the response in the codirectional configuration. The height of the peak drops from 1.22 to 1.14 in dimensionless units and the area under the peak representing the number of bacteria that responded to α-methylaspartate is significantly decreased.

Previous work (Tsang et al. 1973; Adler and Tso 1974; Pfeffer 1988) with attractants and repellents indicated that the decision to respond to either the attractant or repellent was influenced by the size of the gradient as well as the type of chemoattractant. Our proposed MS model therefore incorporated a measure of the gradient along with the chemotactic sensitivity as the input for the decision to which chemoattractant they should respond. This model, like the HS model processes only one input, but uses a different criterion for selection. Comparison to experimental data for the multiple attractants shows that this model is no better than the HS model in representing the data. The apparent conclusion is that bacteria do process both inputs using a common mechanism.

The simplest common mechanism for signal processing is represented by the SA model. It comes closest to what we might expect physically for a receptor-mediated response. It adds together the signals initiated by all the bound receptors. The other two models require the bacteria to somehow distinguish, compare and evaluate several stimuli and then switch off the signal from one or more of them which would be a more complex biochemical process to carry out via a receptor-mediated response. Algebraic additivity of the responses was reported by the previous researchers considering responses to attractants and repellents (Tsang et al. 1973; Adler and Tso 1974; Rubik and Koshland 1978). We have represented this in our model by addition of the signal ϵ from each of the stimuli as measured in single stimulus experiments. Even this model underestimates the observed experimental data for both co- and antidirectional configurations.

Rubik and Koshland (1978) proposed a response regulator model to account

for their observations that seems consistent with the most recent literature regarding the signal transduction mechanism with the level of phosphorylated CheY protein acting as the regulator. They suggested that such a model would predict different additive properties which could be studied using mutants or multiple stimuli. Their studies with mutants revealed that the bacterial system was capable not only of exact additivity, but also nonadditive behaviors such as desensitization and potentiation observed in higher order organisms. One explanation of our experimental data, which is consistent with Rubik and Koshland's hypothesis, is that the signal from the α-methylaspartate is amplified during the signal-processing step. This would explain the result for both configurations. A more complex model than any of those presented here is clearly required to account for signal being amplified within the cell during processing. To develop such a model, we need to understand more about the relationship between the number of bound receptors and the switch activator for the flagellar rotation to incorporate a signal amplification model.

6. Concluding Remarks

In this chapter we have illustrated the necessity and utility of mathematical modeling for characterizing bacterial motility and chemotaxis and for understanding the underlying mechanisms governing this behavior. A cell balance equation derived by reducing Alt's equations for the case of symmetry in two of three dimensions was used with constitutive equations for the tumbling probability to evaluate experimental data from the SFDC assay to determine the macroscopic transport coefficients, random motility, and chemotactic sensitivity. This demonstrates that mathematical models can be applied to specific experimental conditions in which other complicating factors can be minimized or eliminated to determine values of fundamental transport properties of bacteria. In the SFDC, we observe the formation of traveling bands in response to an attractant gradient. The observed bacterial density profiles can be interpreted within the context of a mathematical model to provide macroscopic transport properties.

Once the values for these properties have been determined in independent experiments, they can be used as parameter values in the mathematical model to predict bacterial density profiles for different experimental conditions. We determined μ_0 and χ_0 from total numbers of cells accumulating on one side of the chamber or within a band, then we used these values to predict the bacterial density profiles for comparison with experimental data and found good agreement for the systems we studied, namely, *E. coli* with fucose and α-methylaspartate as attractants.

Good agreement with the experimental profiles suggests that the physical models we have used adequately describe the bacterial swimming behavior so

long as the approximations used in solving the mathematical model are valid. Computer simulations are a tool that can be used to check these approximations to avoid the possibility of compensating errors leading to good agreement. For example, the approximation that confines bacteria to one-dimensional swimming used in the application of the RTBL model to evaluate SFDC assay data yields lower bacterial density profiles when compared with the cellular dynamics simulations as shown in the upper two graphs of Figure 11.12. This model requires a higher value for χ_0 to match the simulation results. However, the physical model also used in the RTBL model, which assumes bacteria decrease run lengths when swimming down the attractant gradient, has the opposite effect (see the lower
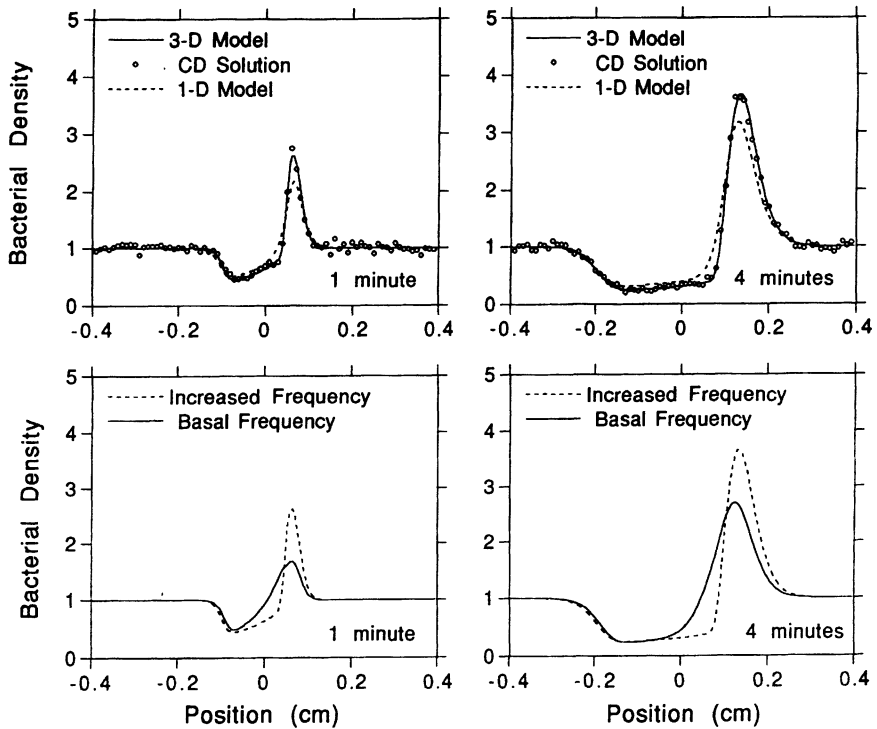
Figure 11.12 Illustration of compensating errors in RTBL model. Comparison of bacterial density profiles generated from the reduced form of Alt's equations, CD simulations and RTBL model (top figures) to show that approximations in RTBL underpredict the bacterial density profile. Solutions of the reduced form of Alt's equations (bottom figures) for different physical models of the bacterial swimming behavior away from increasing attractant concentrations show differences in the bacterial density profiles. RTBL used the increased frequency model which overpredicts the bacterial density. Good agreement with experimental data is still possible because the errors are compensating for each other.

two graphs in Fig. 11.12). The bacterial density profiles are larger than the corresponding ones from the simulations and require a lower value of χ_0 to match the results. Thus, compensating errors can lead to good agreement with experimental data and even provide reasonable values for the transport properties. Computer simulations are necessary to provide an independent method to check for such problems.

Being confident that the RA model yielded reliable results for the cases with single stimuli in a one-dimensional gradient, we extended the model to multiple gradients to investigate the signal-processing mechanism in bacteria by comparing theoretical predictions based on a tumbling probability model to experimentally observed responses for codirectional and antidirectional configurations of fucose and α-methylaspartate gradients. None of the simple models which we proposed matched the experimental data particularly well, suggesting that the bacteria do have a processing mechanism that weights signal inputs differently in the presence of secondary inputs. The development of a model for such a mechanism is part of our ongoing research.

The current emphasis in our modeling work is in several areas. First, we are extending our consideration of stimuli to include repellents, both alone and in the presence of attractants. We have developed a more detailed model of the cellular response that considers additional elements of the signal transduction mechanism based on the work of Bray et al. (1993), and preliminary results indicate that this model works well in predicting the response of *E. coli* to attractant/repellent systems. Second, all of our modeling to date has been for bacteria swimming in a bulk aqueous medium. We are now performing experimental studies and developing mathematical models for bacterial migration within porous media, both in the presence and absence of attractant species. This research moves us toward our ultimate goal of a mathematical description of *in situ* bioremediative processes in terms of fundamental cellular processes.

Acknowledgments

expressed in this material are those of the authors and do not necessarily reflect the views of the IBM Corporation.

References

Adler, J. 1966. Chemotaxis in bacteria. *Science* 153:708–716.

Adler, J. 1969. Chemoreceptors in bacteria. *Science* 166:1588.

Adler, J., and W.-W. Tso. 1974. Decision-making in bacteria: chemotactic response of *Escherichia coli* to conflicting stimuli. Science 184:1292–1294.

Allaire, P. E. 1985. *Basics of the Finite Element Method.* Wm. C. Brown, Dubuque, IA.

Allen, M. P., and Tildesly, D. J. 1987. *Computer Simulations of Liquids.* Oxford University Press, New York.

Alt, W. 1980. Biased random walk models for chemotaxis and related diffusion approximations. *J. Math. Biol.* 9:147–177.

Berg, H. C. 1988. A physicist looks at bacterial chemotaxis. *Cold Springs Harbor Symp. Quantitative Biol.* 53:1.

Berg, H. C. 1993. *Random Walks in Biology.* Princeton University Press, Princeton, New Jersey.

Berg, H. C., and Brown, D. A. 1972. Chemotaxis in *Escherichia coli* analyzed by three-dimensional tracking. *Nature* 239:500–504.

Bornbusch, A. H. 1984. Turning field size and its effects upon computer simulated klinotactic orientation. *J. Theor. Biol.* 107:151.

Bornbusch, A. H., and Conner, W. E. 1986. Effects of self-steered turn size and turn bias upon simulated chemoklinotactic behavior. *J. Theor. Biol.* 122:7.

Bouwer, E. J. 1992. Bioremediation of organic contaminants in the subsurface. In R. Mitchell (ed.), *Environmental Microbiology,* pp. 287–318. Wiley-Liss, New York.

Bray, D., R. B. Bourret, and M. I. Simon. 1993. Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. *Mol. Biol. Cell* 4:469–482.

Brown, D. A., and H. C. Berg. 1974. Temporal stimulation of chemotaxis in *E. coli. Proc. Natl. Acad. Sci. USA* 71:1388–1392.

Crank, J. 1979. *The Mathematics of Diffusion.* Clarendon Press, Oxford, England.

Dahlquist, F. W., P. Lovely, and J. D. E. Koshland. 1972. Quantitative analysis of bacterial migration in chemotaxis. *Nature* 236:120–123.

Eisenbach, M. 1991. Signal transduction in bacterial chemotaxis. In J. L. Spudich and B. H. Satir (eds.), *Sensory Receptors and Signal Transduction,* pp. 137–208. Wiley-Liss, Inc., New York.

Ford, R. M., and P. T. Cummings. 1992. On the relationship between cell balance equations for chemotactic cell populations. *SIAM J. Appl. Math.* 52:1426–1441.

Ford, R. M., and D. A. Lauffenburger. 1991. Measurement of bacterial random motility and chemotaxis coefficients, II. Application of single-cell-based mathematical model. *Biotechnol. Bioeng.* 37:661–672.

Ford, R. M., J. A. Quinn, B. R. Phillips, and D. A. Lauffenburger. 1991. Measurement of bacterial random motility and chemotaxis coefficients. I. Stopped-flow diffusion chamber assay. *Biotechnol. Bioeng.* 37:647–660.

Frymier, P. D., R. M. Ford, and P. T. Cummings. 1993. Cellular dynamics simulation of bacterial chemotaxis. *Chem. Eng. Sci.* 48:687–699.

Frymier, P. D., R. M. Ford, and P. T. Cummings. 1994. Analysis of bacterial migration. I. Numerical solution of balance equation for one-dimensional attractant gradients. *AIChE J.* 40:704–715.

Guell, D. C., H. Brenner, R. B. Frankel, and H. Hartman. 1988. Hydrodynamic forces and band formation in swimming magnetotactic bacteria. *J. Theor. Biol.* 135:525.

Harvey, R. W. 1991. Parameters involved in modeling movement of bacteria in ground-water. In J. C. Hurst (ed.), *Modeling the Environmental Fate of Microorganisms,* pp. 89–114. American Society for Microbiology, Washington, DC.

Keller, E. F., and L. A. Segel. 1971. Model for chemotaxis. *J. Theor. Biol.* 30:225.

Lapidus, I. R., and R. Schiller. 1976. A model for the chemotactic response of a bacterial population. *Biophys. J.* 16:779–789.

Lovely, P. S., and F. W. Dahlquist. 1975. Statistical measures of bacterial motility and chemotaxis. *J. Theor. Biol.* 50:477.

Macnab, R. M. 1980. Sensing the environment: Bacterial chemotaxis. In R. F. Goldberger (ed.), *Biological Regulation and Development,* pp. 377–412, Plenum, New York.

Macnab, R. M., and S. I. Aizawa. 1984. Bacterial motility and the bacterial flagellar motor. *Annu. Rev. Biophys. Bioeng.* 13:51–83.

Macnab, R. M., and D. E. Koshland. 1972. The gradient-sensing mechanism in bacterial chemotaxis. *Proc. Natl. Acad. Sci. USA* 69:2509–2512.

Mercer, J. R., R. M. Ford, J. L. Stitz, and C. Bradbeer. 1993. Growth rate effects on fundamental transport properties of bacterial populations. *Biotechnol. Bioeng.,* 42:1277–1286.

Mesibov, R., G. Ordal, and J. Adler. 1973. The range of attractant concentrations for bacterial chemotaxis and the threshold and size of response over this range. *J. Gen. Physiol.* 23:203.

Othmer, H., S. Dunbar, and W. Alt. 1988. Models of dispersal in biological systems. *J. Math. Biol.* 26:263.

Patlack, C. S. 1953. Random walk with persistence and external bias. *Bull. Math. Biophysics.* 15:311.

Pfeffer, W. 1988. Über chemotaktische Bewegungen von Bakterien, Flagellaten und Volvocineen. *Untersuch. Bot. Inst. Tübingen* 2:582–663.

Rivero, M. A., R. T. Tranquillo, H. M. Buettner, and D. A. Lauffenburger. 1989. Transport

models for chemotactic cell populations based on individual cell behavior. *Chem. Eng. Sci.* 44:2881–2897.

Rubik, B. A., and D. E. Koshland. 1978. Potentiation, desensitization and inversion of response in bacterial sensing of chemical stimuli. *Proc. Natl. Acad. Sci. USA* 75:2820–2824.

Sandler, S. I. 1977. *Chemical and Engineering Thermodynamics.* Wiley, New York.

Segel, L. A. 1977. A theoretical study of receptor mechanisms in bacterial chemotaxis. *SIAM J. Appl. Math.* 32:653–665.

Spudich, J. L., and J. D. E. Koshland. 1975. Quantitation of the sensory response in bacterial chemotaxis. *Proc. Natl. Acad. Sci. USA* 72:710–713.

Stock, J. B., A. J. Ninfa, and A. M. Stock. 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* 53:450–490.

Strauss, I. 1992. Bacterial chemotaxis in the presence of multiple stimuli. M. S. thesis, University of Virginia.

Strauss, I., P. D. Frymier, C. M. Hahn, and R. M. Ford. 1995. Analysis of bacterial migration. II. Studies of multiple attractant gradients. *AIChE J.* 41:402–414.

Stroock, D. W. 1974. Some stochastic processes which arise from a model of the motion of a bacterium. *Z. Wahrsch. Ver. Geb.* 28:305–315.

Tankersley, R. A. and W. E. Conner. 1990. Not-so-random-walks-computer simulations of chemo-orientation behavior. *BioScience* 40:392.

Tsang, N., R. Macnab, and J. D. E. Koshland. 1973. Common mechanism for repellents and attractants in bacterial chemotaxis. *Science* 181:60.

Index