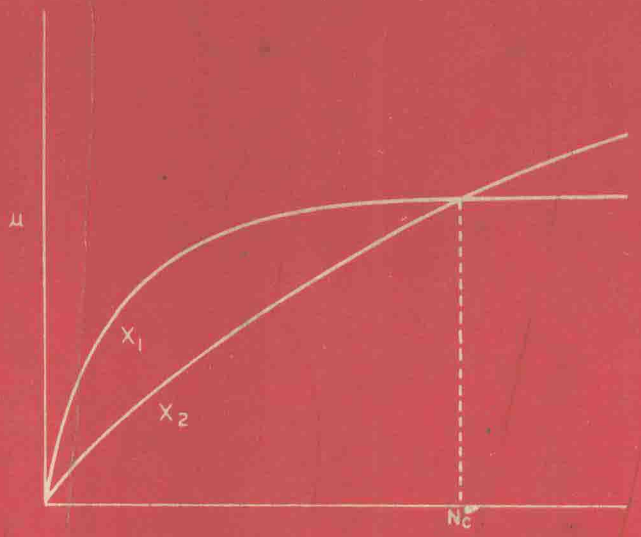
BASIC Microcomputer Models in Biology



James D. Spain

BASIC Microcomputer Models in Biology

James D. Spain
Michigan Technological University

1982

Addison-Wesley Publishing Company
Advanced Book Program/World Science Division
Reading, Massachusetts

London . Amsterdam . Don Mills, Ontario . Sydney . Tokyo

This book was prepared in camera-ready form by Phyllis and Brian Winkel (Way With Words), Terre Haute, Indiana, on Ohio Scientific Instruments Challenger III microcomputer system using WordStar software and NEC Spinwriter printer using Math/Times Roman font.

Library of Congress Cataloging in Publication Data

Spain, James D.
BASIC microcomputer models in biology.

Bibliography: p.
Includes index.
1. Biology—Mathematical models. 2. Biology—Data processing. 3. Basic (Computer program language)
I. Title.
QH323.5.S644 574'.0724 81-19132
ISBN 0-201-10678-7 AACR2

ABCDEFGHIJ-AL-8987654321

Manufactured in the United States of America

PREFACE

As far as I am aware, this is the first general introductory text on the subject of computer modeling for life scientists. On numerous occasions I have tried to justify why I should consider myself qualified to write a book of this nature. Such a book is clearly needed to introduce the student to the exciting and rapidly developing field of computer modeling, and yet, no general text seemed to be forthcoming. Eventually I decided that even though my credentials as a mathematician are essentially nonexistent, I could draw enough from my experiences teaching quantitative courses in biology and chemistry to write an introductory text on modeling. The exercises in the text are drawn from ten years experience developing instructional simulation models and teaching introductory courses on biological simulation techniques. The mathematics employed is generally quite simple, and hence understandable to most students of life sciences. The emphasis is on applications rather than sophisticated mathematical methodology. Few people realize the exciting things that can be done on the digital computer with relatively simple numerical methods. Even unsophisticated models can produce realistic and instructive simulation output when implemented in this way.

Neither understanding of nor appreciation for computer modeling can occur without actually programming the models and interacting with the computer. For this reason, this book consists of a series of exercises which are integrated with the descriptive material. Thus, it may seem more like a laboratory manual than a text. The computer is, in fact, a powerful laboratory tool ideally suited for developing and testing biological concepts. As students complete the exercises, they begin to see each system from a new perspective. Real understanding of a system often comes more from the process of model development than from examination of the simulation data the model produces. Students are always more enthusiastic about developing their own models than they are about working with 'canned' programs which someone else has prepared. The modeling process serves to reinforce the basic understanding of the quantitative principles of biology. Thus, the text might be described as an introduction to the numerical approach to biomathematics.

Biology is a highly integrated science and mathematics is one of the chief integrating forces. As an example, the mathematical concepts of chemical kinetics employed in biochemistry are often used in modeling population dynamics. Compartment modeling and transport processes find applications at all levels of biological organization from the sub-cellular level to the ecosystem level. For this reason, the material has been organized cumulatively. Each chapter builds in one way or another on preceding chapters. My students are encouraged to complete at least one or two representative exercises from each chapter. A diversity of material is provided so that most students can satisfy their particular interests. Experience with simple models provides them with the background to appreciate the potential role of computer modeling in understanding complex biological systems such as large ecosystem models, world models, and complex models dealing with human physiology.

The text is divided into three parts. The first deals with simple equations that model single biological systems or system components. Individual chapters illustrate the various ways in which simple model equations are derived, and used in the computer to generate simulation data. This approach gives the students time to learn computer programming techniques while doing simple examples. Those sections or exercises essential in understanding the other parts have been marked with a double asterisk (**).

The second part is concerned with deterministic models of multicomponent systems. These models generally employ numerical techniques to solve multiple equations with interdependent variables. This part is organized according to specific areas of application and/or modeling approaches within a specific area of application. To a limited extent, one may pick and choose according to his or her field of interest. However, there are certain techniques and concepts with broad applicability which should be understood irrespective of field of interest. As before, these essential concepts have been marked with the double asterisk.

The third part deals with the effects of random processes on biological systems. Its chapters are devoted to topics such as sampling processes, random walks, and queueing. The effect of discrete events occurring randomly in time is emphasized. The examples and exercises employ a direct approach called the Monte Carlo technique.

The book was originally designed to employ BASIC programming language in conjunction with time-sharing terminals. This approach was not totally successful for several reasons. Students without previous computer course work often dropped the course as a result of the initial trauma of interacting with a major computer system. For example, they experienced difficulties with account numbers, run identification numbers, passwords, lost programs and long turnaround time. The problems were often compounded by a communication breakdown between the computer center, the instructor, and the student. Other problems included crowding of the terminal room, and the difficulty of contacting the instructor when questions arose about programming exercises.

When microcomputers became available, it was evident that many of these problems could be solved by setting up a small independent computer lab in the department. Subsequent experience has proven this approach to be correct. Many more students now complete the course, and the quality of work has been enhanced because of the improved graphics capability and quick turnaround time provided by the microcomputer. Thus, the present text has been modified extensively to take advantage of the exceptional capabilities of the microcomputer. However, the general aspects of the original book have been retained so that it may be employed with time-sharing systems if these are preferred. For more detailed discription of the development of the simulation course at Michigan Tech, see Spain (1981).

References have been provided throughout the text to give access to the more important literature on the subject of modeling. However, the reader is reminded that this is an introductory text and it is not possible to include a rigorous literature review on each subject. Models developed in conjunction with my own course work have been included without any citation or source indicated. These deal with common biological systems, so it is likely that similar models are described elsewhere in the literature. I apologize in advance to those who see their pet ideas published here without proper citation. When such oversights occur, please inform me so they can be properly referenced in future editions.

I am heavily indebted to Dr. Edwin T. Williams of the Department of Chemistry and Chemical Engineeering for thoughtfully and critically reviewing early drafts of the text and for making many suggestions which have been incorporated. I would also like to thank my colleagues in the Biological Sciences Department for their comments and criticisms. Special thanks go to Dr. Janice Glime, Dr. Kenneth Kramm, and Dr. Robert Keen, especially for their help on the sections dealing with ecology, and to Dr. Martin Auer for his review of the sections on limiting factors for growth. Thanks also go to Dr. James Horton at the Biology Department, California State College in Bakersfield, and to Mrs. Lois Young for their editorial assistance.

Special gratitude goes to Dr. Brian Winkel, my close friend and colleague at the Division of Mathematics, Rose-Hulman Institute of Technology, for his comments, criticisms, and encouragement during the final stages of manuscript development, and to Phyllis Winkel, for the hours spent patiently transcribing various versions into the word-processor. Finally, thanks go to the many Michigan Tech students who contributed comments and suggestions during the development of this course material.

I believe that the objectives outlined in this preface are very desirable for an introductory course in modeling for biologists. I have tried to make this text meet those objectives and fill what many believe is a need in the curriculum of life science majors.

The reader should recognize that this is not intended to be the 'last word' in computer modeling. Rather it represents a first attempt to organize the general principles of computer modeling into a single integrated body of information. I hope that others will be stimulated to build on my beginning and ultimately produce a text which is fully representative of this exciting new field.

James D. Spain

Utility programs and sample exercises designed to complement the text are available from Addison-Welsey. Programs on mini-floppy disk or cassette for either TRS-80* or Apple-II microcomputers include CURFIT, POLYFIT, POISSON, GRAPH, TOLLGATE, NORMAL and sample exercises from the book. Inquiries and orders should be directed to the publisher.**

---

*TRS-80 is a trademark of Radio Shack, a division of Tandy Corporation.
**Apple II is a trademark of Apple Computer, Inc.

CONTENTS

CHAPTER 9

CHAPTER 10

CHAPTER 11

CHAPTER 12

CHAPTER 13

CHAPTER 14

CHAPTER 20

CHAPTER 21

CHAPTER 22

APPENDIX 1

APPENDIX 2

APPENDIX 3

APPENDIX 4

APPENDIX 5

APPENDIX 6

APPENDIX 7

INTRODUCTION

THE ROLE OF COMPUTER MODELING AND SIMULATION IN BIOLOGY

Future life scientists may be characterized more by their ability to use the computer for data analysis and simulation than by their ability to use the traditional microscope. Biology is potentially the most mathematical of the sciences. This results from the fact that living systems involve a complex interaction of chemical and physical processes all of which are capable of being described in mathematical terms. These systems which are much more complex than any devised by the mind of man have for the most part resisted mathematical analysis by the classical methods so successfully employed by physicists and chemists. Only recently, with the advent of the digital computer, have biologists been able to use numerical methods to deal with these multi-component systems. The resulting new interest in biomathematics has been further stimulated by the availability of the personal computer. Suddenly biologists find that they can easily bring into their office or lab a tool capable of a wide range of data analysis techniques. Statistical methods that previously required hours to accomplish using the calculator can now be completed in minutes. Thus, many techniques, such as cluster analysis and multiple regression, should soon become routinely employed by most knowledgable biologists and a necessary component of undergraduate biology instruction.

Computer modeling and simulation is another area of biomathematics which is now accessible to all biologists as a result of the microcomputer revolution. Simulations previously employed by the few who happened to have the necessary equipment and expertise are now available to anyone willing to take the time to learn the few simple programming tricks that are needed to use BASIC as a simulation language. The objective of this book is to provide an organized discussion of biological simulation techniques which may be implemented on the microcomputer.

James D. Spain, BASIC Microcomputer Models in Biology ISBN 0-201-10678-7

Conceptual Models

Models and other analogies have always played an important role in the scientific thought process. The use of models is so ingrained in our thinking that we are often unconscious of the important distinction between models and real systems. A model may be defined as any representation of a real system. It may deal with either the structure or function of the system. It may use words, diagrams, mathematical notation, or physical structures to represent the system. It is often synonymous with concept, hypothesis, or analogy. As no model can totally represent the real system in every detail, it must always involve varying degrees of simplification. Using this broad definition, it becomes evident that essentially all science deals with the formation, examination, and improvement of conceptual models about our universe. At this point, it would be valuable to review some of the conceptual models that have contributed to our present ideas about living systems.

The atom as a unit of elemental structure and the molecule as a unit of chemical reactivity were for a very long time simply conceptual models. Only in the last few decades have we been able to observe directly a few very large molecules using electron microscopy. X-ray diffraction allows one to construct physical models of molecules based upon the diffraction patterns they cast. It is still an indirect technique for observing that which we may never be able to observe directly. When Linus Pauling (1977) worked out the α-helix structure now found to be present to some degree in most protein polypeptide chains, he experimented with various molecular configurations using paper models until he arrived at one which had repeat distances that were consistent with his x-ray data. Another classical example is the double helix structure of deoxyribose nucleic acid (DNA) ,first proposed by Watson and Crick, based upon a molecular model which fitted many of the known properties of DNA (Watson, 1968). Like most good models, it triggered a burst of experimental activity which has continued to the present.

What is often referred to as the typical animal cell or the typical plant cell is actually a diagrammatic model of the cell based upon a composite of many observations of many kinds of cells using a variety of observation techniques (Hardin, 1966). The three-dimensional structure of most organs within the body is based upon the serial analysis of hundreds of two-dimensional tissue sections using the technique of stereology (Elias and Pauly, 1966). The resulting three-dimensional organ models are often essential to the understanding of their function under normal and abnormal conditions.

The conceptual model of the gene has gone through a long history of evolution (Glass, 1963). The blending theory of heredity gave way to the concept of hereditary particles described as 'beads on a string.' Subsequently, the gene has been considered as that portion of chromosomal DNA which codes for a single polypeptide chain. Despite the great advances made in this field, it is still dominated by conceptual models based largely upon the observation of genetic effects. Each new model has led to a new set of questions, which in turn have led to a new and better understanding of the nature of the gene.

In ecology, the food chain and the food pyramid are important conceptual models which have been used to explain the flow of energy and cycling of materials within the ecosystem. The theory of evolution is based on an enormous amount of evidence, but still must be considered a conceptual model.

In biochemistry, we deal with a variety of conceptual models about enzyme action, including enzyme attachment to the substrate and enzyme responses to changes in temperature and pH. The metabolic pathways that we call glycolysis, Krebs' cycle, and the carbon cycle of photosynthesis are in reality conceptual models which, like other models, may require modification to be consistent with new findings. Each of these models has led to great spurts in experimental activity because of the questions raised. We depend on these models to make decisions about medical treatment, drug action, treatment of poisons, and nutrition of both plants and animals.

It is clear that biology has profited greatly by translating real systems into models of various types. Some of the earliest models involved careful observation and drawing in great detail the morphology of systems under investigation. This technique still provides a valuable means of forcing oneself to see details about a structure which would otherwise go unnoticed. Those who have taken the time to accurately draw some detailed biological structure on paper soon realize that they had never really examined the subject before. New morphological details suddenly appear, and interesting relationships become evident. This illustrates that the greatest benefit of a model often comes as a direct result of the thought process involved in model development. Clearly, modeling has played an important role in the understanding of most biological systems.

Life science is involved in more than just description and understanding of natural systems. Very early in history, man decided to be a manager of his environment, and one of the objectives of biology is to gain sufficient information about complex living systems to manipulate them for our benefit. Management decisions are made on the basis of the conceptual models as perceived by the manager. Whether we are talking about wildlife management or medicine, it is obvious that improvement of our conceptual models will lead to better management.

Conceptual models, by themselves, are generally lacking in rigor. They can be imprecise and interpreted differently by different people. To circumvent this disadvantage, a method has been developed for translating the conceptual model into a form which is more subject to precise description, evaluation, and validation. This form of the conceptual model is called the mathematical model.

The Mathematical Model

Mathematical models deal with the rate of change of systems of cells, organisms, populations, or molecules with time and the extent to which such systems are effected by light, pH, temperature, or other environmental factors. A mathematical model may be as simple as a single equation relating one variable to another, or it may involve the interaction of many equations having several mutually dependent variables. The latter will be referred to as a multicomponent system model. Simple mathematical models may be obtained in various ways, but the two main approaches involve either theoretical derivation, or empirical derivation through statistical analysis.

An equation or group of equations by themselves may not contribute much to the understanding of a particular phenomenon. For this reason, it is usually necessary to solve the equation for some representative values of the independent variable (for example, time, pH, or temperature) and to present the resulting information in the form of a graph. Because of the amount of data involved, this is best done by implementing the mathematical model on a computer, especially one with graphics display capability.

Computer Simulation

Simulation in its simplest form involves implementing of a mathematical model on the computer to produce simulation data. In this way, the output of the mathematical model may be readily compared with experimental data from the real system in order to evaluate the model. Because of the complex interrelationships involved, simulation is particularly essential to the understanding of the multicomponent system model.

The process of developing a simulation model forces the investigator to describe the system in simple terms. When working with the model in this way, the investigator must take into account details about the system which might otherwise go unnoticed. The objective of any kind of model is to provide a means for obtaining new insights into the operation of a system. Simulation assists in this regard by permitting experimental interaction with the model to produce verifiable responses.

There are two basic approaches to simulation using mathematical models. One involves the use of the analog computer to solve the mathematical equations by reducing them to electrical analogs such as resistors, capacitors, and amplifiers. The output of the analog computer is a time varying voltage which is recorded on either an X-Y plotter or a cathode ray tube. It is especially useful for simulation models which involve the integration of complex differential equations such as those describing growth or energy flow. However, not all systems lend themselves to the analog approach, and programming requires considerable understanding of the electronics principles involved. The other basic approach involves the simulation of a system by numerical methods employing the digital computer. This text is concerned exclusively with the latter technique.

The general approach to digital computer simulation usually involves the following steps. First, the system must be analyzed in order to determine the basic components required for the development of a conceptual model. Often this analysis results in the construction of a block diagram of the system. Next, the key variables are defined, and each is expressed in the form of a simple functional relationship with the other variables in the system. Equations are then derived establishing the actual mathematical relationship between the variables. This derivation is done either empirically, through the use of statistical methods such as curve fitting, or analytically, by deriving the equations from theoretical considerations. Next, the mathematical expressions are programmed into the computer, and various rate constants or coefficients are assigned. The simulation is then allowed to run and produce a set of simulation data. The simulation data are then compared with experimentally obtained data to validate the performance of the model. Usually, the model requires some modification at this point in order for it to simulate the real system more accurately. After sufficient verification and validation, the model may be used to perform experiments in much the same manner as one performs experiments on the real system.

In many ways, simulation may be likened to a game that one plays with the computer. The computer keeps track of all the rules of the game, follows through the play in correct sequence, and provides any random numbers that might be necessary to simulate the effects of chance. Persons working with the simulation would have various decisions to make just as they would in playing computerized football, where, depending on conditions, they would select certain strategies to optimize the chances of scoring. In the same way, an individual working with a population growth simulation would have an opportunity to decide initial population levels, growth rates, whether there is predation or not, and how predation relates to changes in population density. The simulation would be run for a while so that its behavior might be observed. Subsequently, one may wish to alter conditions in an attempt to stabilize the system or to observe the effects of different growth rates. The outcome of such a "game" will depend on the assumptions made in designing the simulation and on the initial parameters selected by the simulator.

Just as computerized football can only approach the real game in terms of complexity, the whims of chance, and multiplicity of decisions to be made, simulations can only approach the real system to varying degrees depending upon the complexity of the computer model employed. On the other hand, it is possible to simulate systems which would be almost impossible to investigate experimentally because of the magnitude of time and/or space involved. For example, investigation of a real predator-prey system could involve population estimates taken over a 10-50 year period and a 10-1000 square mile area. Even population experiments in small closed systems may require intense study over many weeks to provide meaningful results. A simulation of these same systems could be carried out in seconds on the computer.

<u>The Relationship of Modeling and Simulation to the Research Process</u>

Some of the concepts presented above have been summarized in a diagram relating modeling and simulation to the overall process of research and management. (See Figure 1.)

It all begins, of course, with the Real System. It must be understood that any biological system must always remain to some extent a "black box." No matter how much information we have about a particular biological entity we must always remain on the outside looking in. For this reason, the Real System is distinguished from all other components of the diagram by representing it as a circle. All of the boxes represent forms of information that in one way or another are derived from the Real System. As such, they are reflections or perceptions of the Real System. Arrows represent processes by which this information is obtained and manipulated.