

portant at this stage among sparse populations. The survival rate of dense populations of the female pupae varied primarily in response to ichneumonids and disease. Variation in the survival rate of the instar I - III larvae was probably primarily a function of variation in the dispersion rate of the newly hatched larvae.

- An index of disease incidence among instar IV - VI larvae was a curvilinear function of insect density, a linear function of precipitation during June, a linear function of the percent of swamp white oak in the overstory, and a function of an interaction between density and precipitation. Several other variables (the apparent survival rate of instar I - III larvae; the proportion of the female pupae killed by ichneumonids; and the number of eggs deposited per adult female) were curvilinear functions of insect density.
- A generation model was developed that describes density at the beginning of a second generation as a function of the environmental variables associated with the above mortality-causing factors. This model was tested against an independent body of data with some success.

This information, with more complete details and interpretive discussion, has now been published.²

A study was begun on sparse, stable gypsy moth populations in 1965, and continued in 1966 and 1967. This work began on the assumptions that: (1) variation in the rate of vertebrate predation is mainly determined by the location of the larvae within the environment when the predators are foraging for food; and (2) larval distribution in sparse populations is mainly determined by the distribution and quality of resting places, which the insect uses during daylight. Some of our combined 1965 and 1966 results are summarized in tables 5 and 6. These results support the above assumptions.

An approximate mean value life table is shown for the 1965-66 data (table 7).

² Campbell, R. W. THE ANALYSIS OF NUMERICAL CHANGE IN GYPSY MOTH POPULATIONS. Forest Sci. Monogr. 15, 33 pp. 1967.

Table 5.—Relationship between number of bark flaps per 0.01-acre plot and number of adult gypsy moths produced

Number of bark flaps per plot	Number of plots observed	Total adults	Adults per plot
0	280	6	.02
1	77	4	.05
2 - 3	70	6	.09
4 and over	56	15	.27

Table 6.—Relationship between pupal location, pupal sex, and mortality rate; and between pupal sex and mortality rate from vertebrate predation

Pupal location	Number found		Mortality rate		Mortality from vertebrates	
	♂♂	♀♀	♂♂	♀♀	♂♂	♀♀
Bark flaps	64	27	.78	.85	.20	.41
Other	71	85	.87	.95	.56	.86

Table 7.—Life table typical of sparse gypsy moth populations in northeastern Connecticut

x	lx	dx	dx	100qx
Age interval	Number alive at beginning of x	Factor responsible for dx	Number dying during x	dx as percent of lx
Eggs	550	Parasites	82.5	15
		Other	82.5	15
		Total	165.0	30
Instars I-III	385	Dispersion, etc.	142.5	37
Instars IV-VI	242.5	Deer mice	48.5	20
		Parasites and Disease	12.1	5
		Other	167.3	69
		Total	227.9	94
Pre-pupae	14.6	Predators, etc.	2.9	20
Pupae	11.7	Vertebrate predators	9.8	84
		Other	0.5	4
		Total	10.3	88
Adults	1.4	Sex(S:R=30:70)	1.0	70
Adult ♀♀	0.4	—	—	—
Generation	—	—	549.6	99.93

METHODS OF DEVELOPING LARGE-SCALE SYSTEMS MODELS

by K. E. F. WATT, Department of Zoology, University of
California, Davis, California.

WHEN A MATHEMATICAL model for a complex process is constructed, instead of merely selecting an empirical formula useful for interpolation—as in regression analysis—we try to write one or more differential, difference, or difference-differential equations based on insight into the mechanics of the process under study. Before discussing the methods for building large models, the differences between these three kinds of equations will be explained.

Differential Equations

Differential equations arise whenever a rate of change of some variable with respect to another can be expressed in terms of a continuous set of variate values of one or more other variables. Consider, for example, the simple attack equation (*Watt 1959*) in which

N_A represents numbers attacked,

N_0 represents numbers vulnerable to attack,

P represents the numbers of attackers,

a , b and K are constants, and all numbers have been measured in the same universe,

$$\frac{dN_A}{dN_0} = aP^{1-b} (PK - N_A)$$

This is a differential equation, because the variables P and N_A are assumed to be able to vary continuously. The notion of differential equations can in fact be generalized considerably to include accelerations and higher-order derivatives, squares of derivatives or higher degrees than this, and partial differential equations, in which there are several independent variables, as in

$$\frac{\delta^2 Z}{\delta x^2} + \frac{\delta^2 Z}{\delta y^2} = x^2 + y$$

However, the common element in all differential equations is the notion that the variables can take any value on a continuous scale, including non-integer values. A moment's reflection indicates that such equations do not give an accurate picture in many biological systems, unless we handle them in a special way. Consider an animal laying eggs. It lays 8 eggs, 1 egg, 233 eggs, or 119,236 eggs, but never 6.237 eggs. Consider an area of ground, divided up into squares 10 meters by 10 meters. If there are 9 mice living in one such square, emigration from the square will occur one mouse at a time, not 1.23 mice at a time. Many biological processes have this character, in which variables can take only discrete values. We conclude that difference equations are often more appropriate than differential equations, or at least differential equations must often be treated as difference equations.

Difference Equations

Difference equations are used wherever the variables take only discrete values. For example, the differential equation for exponential growth is

$$\frac{dY}{dt} = rY$$

In contradistinction, the difference equation for exponential growth is

$$\frac{Y_{k+1} - Y_k}{t_{k+1} - t_k} = \frac{\Delta Y}{\Delta t} = rY_k$$

Equation (1) states that at any instant of time, the rate of growth in Y will be proportional to the value of Y at that instant. Equation (2) states that the difference in magnitude of Y from time t_k to time t_{k+1} will be r times the value of Y at t_k . The essential difference is that Y and t are assumed to vary only by discrete (i.e., steplike) values. A book by Goldberg (1958) is an excellent introduction to this field for behavioral scientists.

Difference-Differential Equations

Suppose, now, that we wished our dependent variable to be a derivative, as in differential calculus, but our independent variables are to be treated as discrete values at some prior point in

time. Here the exponential growth equation would take the form

$$\frac{dY}{dt} = rY_{t-\tau} \quad (3)$$

where τ is the period of time lag.

No matter what type of equation is most suitable for describing a complex system, the problem faced in all cases is that of determining the particular structure for the model which gives the most realistic description of the system.

In the use of large complex problems with many interaction terms, the most convenient way to approach model construction is by trying to split the problem into bits, so we can deal with the bits individually. For example, suppose we know that Y , the volume of timber grown in a forest per unit time, is a function of 15 other variables, so that we could write

$$Y = f(X_1, X_2, \dots, X_{15}). \quad (4)$$

The first step is to determine how to split (4). Do we have

$$Y = \frac{f(X_1, X_2, X_3)}{f(X_4)} + f(X_5, \dots, X_{15}),$$

or $Y = [f(X_1, X_2, X_3)] [f(X_4, X_5, X_6)] [f(X_7, \dots, X_{15})]$.

How to combine such component terms, or sub-models, to make a large model will often be apparent from the following two probability theorems.

The addition theorem.—The probability that one out of m events, any two of which are mutually exclusive, occurs is equal to the sum of the probabilities of the occurrence of each event separately.

The multiplication theorem.—The probability that two stochastically independent events occur together is equal to the product of the probabilities of the occurrence of each event separately.

Given that the researcher has determined how to split his problem into pieces, how does he write equations for a sub-model? Suppose that the whole model is

$$Y = A \cdot B \cdot C \cdot D, \text{ and the sub-model we are concerned with is} \\ A = f(X_1, X_2, X_3). \quad (5)$$

Table 1.—Result from sorting and tabulating data to reveal form of functional relationship

Number of cards in sub-deck	A		X ₁		X ₂		X ₃	
	Total	Mean	Total	Mean	Total	Mean	Total	Mean
5	—	—	—	2	—	2	—	40
5	—	—	—	2	—	2	—	80
5	—	—	—	2	—	2	—	120
5	—	—	—	2	—	4	—	40
5	—	—	—	2	—	4	—	80
5	—	—	—	2	—	4	—	120
5	—	—	—	2	—	8	—	40
—	—	—	—	2	—	8	—	80
—	—	—	—	2	—	8	—	120

and so on

To determine the particular form of (5), the analyst can sort the data, using cards, to separate out the effects of X₁, X₂, X₃, by

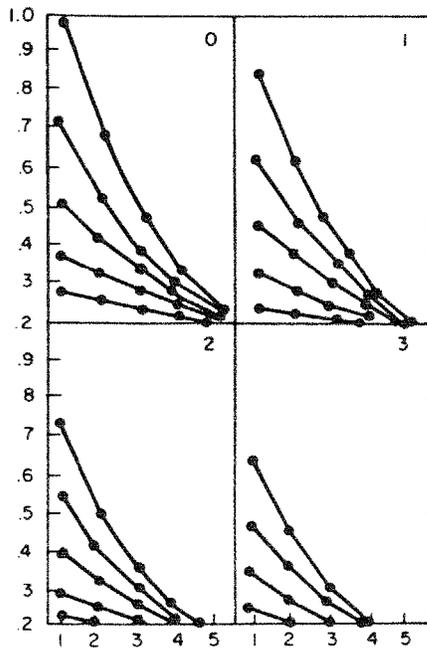


Figure 1.—Plot of row means from table as in Table 1 to determine form of functional relationship $A = f(X_1, X_2, X_3)$. In this figure, A is plotted on the Y-axis; X₁ is plotted on the X-axis; different values of X₂ are represented by the different lines in each panel, and different values of X₃ are represented by the four different panels (from Watt 1967).

obtaining tabulations as in table 1. After making such a table, he will plot families of graphs, as in figure 1. The order in which he then develops sub-sub-models to express the effects of X_1 , X_2 , X_3 on A will depend on the proportion of the variance in A accounted for by each of the X 's. One should begin by modelling the effect of that independent variable for which the graph of A on X has the steepest slope and least scatter.

The researcher will then make various assumptions about the mode of operation of, say, X_3 on A , on the basis of his understanding of the phenomenon, and check these by examining the graphs. It is best to check assumptions systematically by asking oneself a series of questions, the answers to which form a logical branched tree. For example, using only a small list of questions, we form a logical tree as in figure 2. The following list of questions used in figure 2 could be expanded to make a tree from which all known equations could be derived.

1. Is dA/dX_3 proportional to X ?
2. Is dA/dX_3 proportional to A ? (i.e. are we dealing with some form of compound growth law?)
3. Is dA/dX_3 inversely proportional to X_3 ?
4. Does dA/dX_3 approach zero as A approaches some upper asymptote A_{max} ?
5. Does dA/dX_3 approach infinity as X_3 approaches some lower limit $X_3 \text{ min.}$?

Integral forms of some of the most commonly encountered equations are given below.

A—32	$Y = a + bX$
A—28	$Y = a + b \ln X$
A—24	$\ln Y = a + bX$, or $Y_x = Y_0 e^{bx}$
A—16	$Y = a + bX - cX^2$
A—20	$Y = aX^b$
A—22	$Y = \frac{Y_{max}}{1 + e^{-bx}}$ the logistic
A—30	$Y = Y_{max} (1 + e^{-bx})$

Once an elementary equation has been chosen to describe the effect of X_3 on A , this equation is integrated (using tables of integrals, if necessary), and transformed into a form suitable for testing. Graphical testing of the validity of the equation is done

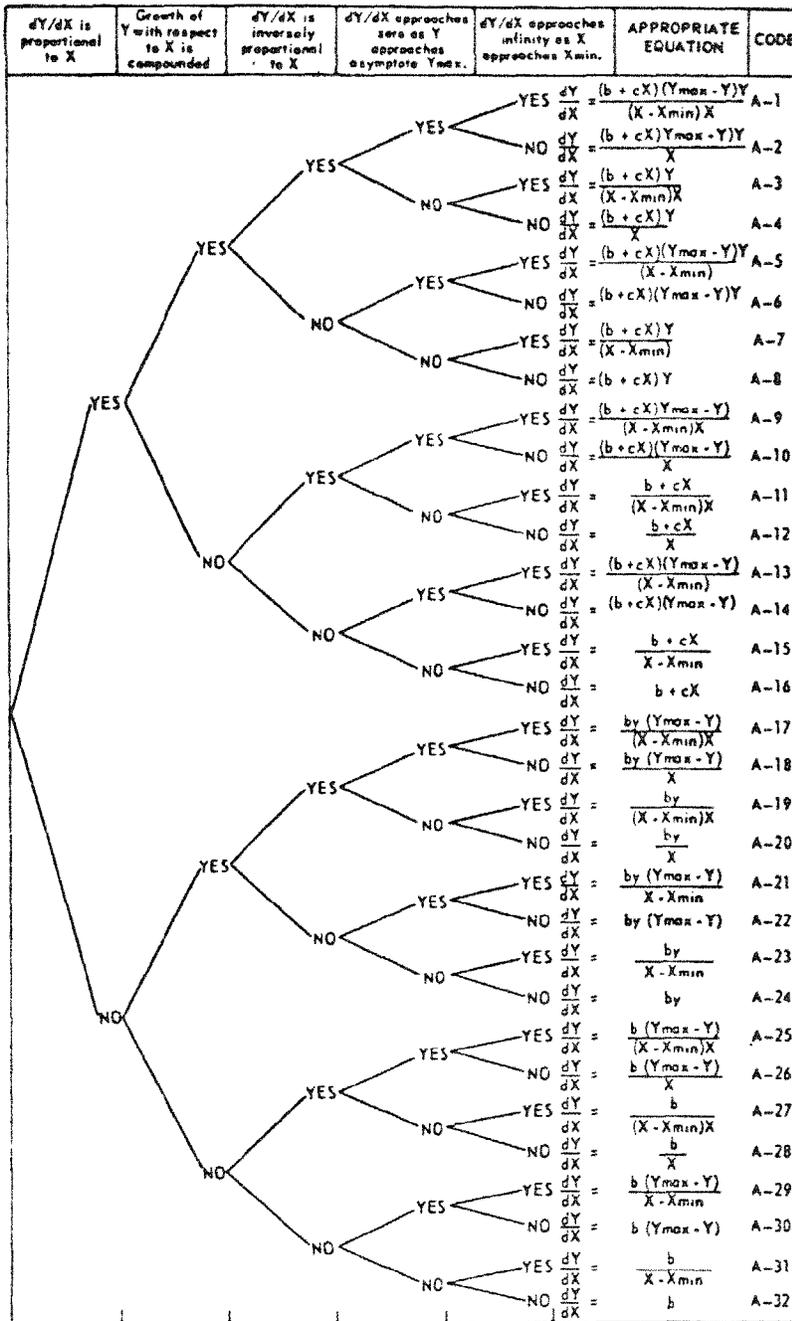


Figure 2.—Logical branching tree for obtaining appropriate differential equation to describe a set of data (Watt, 1961).

as follows. Suppose we decide that $A - 30$ describes the effect of X_3 on A , while X_1 and X_2 are held constant.

That is

$$\frac{dA}{dX_3} = b (A_{\max} - A),$$

and integrating,

$$- \ln(A_{\max} - A) = bX_3 - \ln A_{\max}$$

and since we wish first to see if b is indeed constant, this can be rearranged to yield

$$\ln \left[\frac{A_{\max}}{A_{\max} - A} \right] = bX_3 \quad (6)$$

Plotting the transformed data on semi-logarithmic graph paper will yield a straight line if (6) in fact describes the relation between A and X_3 , X_1 and X_2 being held constant. If a straight line is not obtained, some other elementary equation will be needed in place of $A - 30$. If there is rectilinearity, we proceed to the next step. This consists of plotting the data in terms of some other independent variable, say X_2 .

We plot

$$Z = \ln \left(\frac{A_{\max}}{A_{\max} - A} \right) \text{ against } \frac{X_2}{X_3}$$

If plotting this transformation against X_2 yields a straight line parallel to the $X_2 -$ axis, X_2 has no significant effect on A . Otherwise, we know that b is not a constant, but in fact is some function of X_2 . If so, we decide for X_2 the appropriate form of the function, and test this guess by plotting the appropriate transformation to see if we get a straight line. For example, if we decide that

$$\frac{dZ}{dX_2} = \frac{cZ}{X_2},$$

then

$$Z = gX_2^c,$$

which we test by plotting Z against X_2 on log-log graph paper.

If we get a straight line, we know that

$$\ln \left[\ln \left(\frac{A_{\max}}{A_{\max} - A} \right) \frac{X_3}{X_2} \right] = \ln g + c \ln X_2$$

In this case, we can proceed to determine the effect of X_1 on A . We do this plotting

$$\ln \left(\frac{A_{\max}}{A_{\max} - A} \right) \frac{1}{X_3 X_2^c} \text{ against } X_1.$$

When this step is complete, we solve the resultant form for A . The other sub-models are handled in the same way, and finally the whole model is put together.

Now that the basic process of cyclical model building and testing has been explained, there are a large number of questions on detail that need to be considered.

First, where we have made the postulate that A is a function of X_1 , X_2 and X_3 , how do we determine the order in which we will test the significance of X_1 , X_2 and X_3 ? In fact, we proceed as in stepwise multiple regression, and examine the effects of the independent variables starting with the most important and proceeding to the least. The reason for proceeding in this sequence has to do with the logic of statistical testing (by leaving the least significant variables until last we have stronger error terms with which to test the significance of the first variables removed, using the variance-ratio, or F test).

Second, how do we decide the fundamental structure of the whole system model into which we incorporate terms for sub-models, into which we have in turn built sub-sub-models? One way of splitting up a model for a system is to structure it in terms of time. Consider, for example, an equation to account for the change in numbers of an insect population from one generation to another. First, we need to define a number of symbols.

Let	N_t	Represent the density of adult insects present immediately prior to oviposition in year t .
	N_{t+1}	The density of adult insects present at the corresponding time in year $t + 1$.
	$T_{t:t+1}$	N_{t+1}/N_t , i.e. the trend index of the population from t to $t + 1$.
	P_t	The proportion of N_t consisting of females that oviposit at t .
	F_t	The mean fecundity of the $N_t P_t$ females.

S_E	The proportion of the eggs surviving to eclosion.
$S_I, S_{II}, \dots, S_{VI}$	The proportion surviving of first instar larvae, second instar larvae . . . sixth instar larvae.
S_P, S_A	The proportion surviving of pupae and adults, respectively. Adult populations are measured just before oviposition in year $t + 1$.

We have, then, the model

$$T_{t:t+1} = \frac{N_{t+1}}{N_t} = P_t F_t S_E S_I \dots S_{VI} C_p S_A.$$

This equation structures our system in a common-sense sequence, and is called a model. It is built up out of terms such as

$$S_E = f(X_1, X_2, X_3, X_4, \dots, X_n)$$

which are called sub-models, and these in turn may be built up out of sub-sub-models, which for example relate the variate value of X_4 to factors which govern X_4 .

Another matter needing consideration is the list of five questions which produced the 32 (2^5) equations in figure 2. The list is, in fact, by no means complete, but was merely presented to show that most of the simple, commonly encountered differential equations can be derived from a simple logical branching tree. Biologists will be able to think of a variety of other questions which might be asked to expand the logical tree enormously. However, it is much easier to think of a differential equation than to solve it, and the biologist will need a powerful aid if he does much of this type of model-building. Apart from the usual handbooks, the most useful such aid is the compendium of methods of solution by Murphy (1960).

A logical branching tree is not the only way to arrive at the equation which best describes a particular process (Turner, Monroe and Lucas 1961; Turner, Monroe and Homer 1963). Another approach is to make use of general underlying equations from which a great variety of commonly-used models may be derived as special cases. Using such formulas, we can employ curve-fitting techniques to determine the values of various parameters that

give best fit of the general equation to a particular body of data. The particular values the parameters take indicate which of the special cases we are dealing with. To illustrate, consider the differential equation used by Turner and his associates, in which n and ϵ are variables, δ , a , and γ are parameters.

$$\frac{dn}{d\epsilon} = \frac{\delta(n-a)}{\epsilon-\gamma\delta}$$

This equation integrates to produce

$$n = a + \beta (\epsilon - \gamma\delta)^\delta \quad (7)$$

The constant α determines which special case we have represented in our data. If β is a positive integer, we have a polynomial process, which is non-linear if δ exceeds one. As δ approaches either positive or negative infinity, (7) approaches the exponential model

$$n = a + \beta e^{-\epsilon/\gamma}$$

The following values of $2/\delta$ produce the given simple models

$2/\delta$	model
-4	inverse square root law
-3	inverse two-thirds law
-2	rectangular hyperbole
-1	inverse square law
0	exponential
1	parabola
2	straight line

When $2/\delta$ is negative, the curves have two asymptotes, when $2/\delta$ is zero they have one, and when $2/\delta$ is positive they have no asymptotes.

Another such basic equation has been proposed by Grosenbaugh (1965).

$$\text{Consider } Y = H + A \left[e^{(N^2-1)U} - NU \right]^{NM+1} \quad (8)$$

where U represents a 1- or 2-parameter function such as

$$U = E^{-B(X-G)} \quad (9)$$

Equation (8) with variant (9) inserted takes the form of almost all well-known mathematical functions for two variables,

depending on the values assigned to N and M . The idea behind Grosenbaugh's research is to facilitate the computations for iterative regression by developing tables of the partial derivatives used for each of the elementary functions corresponding to a particular N and M pair of values. This suggests a large-scale computer routine for doing iterative regression, in which the computer starts with only the raw data. In the first part of the program, the computer determines the values of N and M and the form of U . Then, using these three pieces of information, it selects from memory the appropriate derivatives for calculating the values in the set of equations. The second part of the program is the iterative regression routine.

It should be noted that both the logical branching tree method and the general equation method suggest means of programming a computer so that it decides which model best describes a given body of data.

Model-building techniques of the type we have described can obviously lead to extremely complicated models. The question naturally arises as to how one can perform logically valid statistical or other tests on the model, to determine how accurately it describes reality. One approach is to test the whole model statistically piece by piece (sub-model or sub-sub-models separately), using iterative regression where necessary because parameters enter the sub-models nonlinearly. However, this approach has a logical weakness, in that we lose a degree of freedom every time a new parameter enters an equation, and our error terms in statistical tests lose strength. This means, that as our models become more and more complex, it becomes progressively more difficult to demonstrate statistically that they do not in fact describe the data. In other words, ANY sufficiently complicated model is a highly flexible interpolatory formula, even though its structure bears no relationship to the structure of the process we are trying to describe.

The reader can satisfy himself on this point by noting in tables of F -values, that the smaller the number of numerator degrees of freedom, the higher the F -value must be to produce significance at a given probability level. To circumvent this degrees-of-freedom

impasse, the following procedure is suggested. Obtain data from one set of surveys, or experiments on components of whole systems, then build a systems model. The model is then tested by seeing how it can predict the outcome from another set of surveys or experiments on whole systems, or of a different type. No degrees of freedom are lost in such a test. For example, we might build a systems model describing a historical process, using data from experiments on particular components of the whole process. Then having used these sub-models to build the whole model, the model is tested against data on the history of the whole system. Such an approach has been used by Holling (1966) and Watt (1955).

This concludes a general discussion on methods of developing models. A typical situation that arises subsequent to following the steps outlined is for the researchers to find that their systems model only accounts for 40 per cent of the variance in the system under study. It often develops that a major proportion of the entire model-building effort is expended on efforts to refine the initial systems model. A number of procedures are available to help with this step.

1. Using the model, obtain values of "Y calculated" for each case on which data were obtained. Then make a large graph in which "Y calculated" values are plotted against "Y observed." Every point on the graph should be labeled as to the characteristics of the sample that produced the datum. When all points are plotted, we study the graph to observe if particular sets of sample values deviate consistently above or below the 45° line on which all points would lie in an error-free situation, and where we have an accurate model. Such sets of values may provide clues about factors omitted from the analysis which should have been measured.

2. Another method for discovering structural weakness in the model is to plot "Y observed - Y calculated" against each of the independent variables in turn. Systematic departures from a straight line parallel to the X- axis indicate that a term in the model does not mimic nature accurately.

3. A third source of difficulty in a systems model occurs if one

or more variables have been incorrectly measured, defined, or coded. Suppose we know from field experience that torrential downpours or high-velocity wind gusts can drive delicate insects off their food plants onto the ground where they may starve or be drowned. It is not adequate to express the force of such factors in standard meteorological units, such as inches of rain collected in a rain gauge or miles of wind per hour moving past a stationary point. Rather, we should record and enter into our equations the biologically relevant variables: maximum drop velocity within a 24-hour period or maximum gust velocity.

4. If some of the independent variables that entered into multiple regression analyses were of the form X_1X_2 , then computer output may reveal that the cross-product or other interaction terms are needed in the systems model.

The Experimental Components Approach

The preceding discussion outlined mathematical techniques by which one could develop a model for a system on which a large body of data had already been collected. However, such a procedure is only useful as a "macro" approach. In order to obtain the kind of insight into the mechanics of a system that will allow us to manipulate it profitably, in many cases more detailed information about the quantitative nature of processes will be required than those available from field studies. Further, the "macro" approach, typical of field studies, limits us to the ranges of variate-values provided by nature. Hence, in many cases a "micro" approach is called for in addition to the "macro" approach. Holling (1961, 1965) has presented an elaborate exposition of the logic and methodology of the "micro" approach, which he calls experimental components analysis. The following discussion is based in part on his publications.

The essence of the experimental components approach is that the processes of experimentation and mathematical model construction are conducted as two interlocking parts of an integrated program. A systems model grows out of a sequence of steps which proceed as follows.

The process which we wish to study (e.g. predation, parasitism,

dispersal, reproduction) is conceived of as comprising a set of component processes, or constituent fragments. We determine by experiment and observation what these fragments are for the process in question, then through *a priori* considerations sort them into two groups, basic and subsidiary. A basic component of a process is a constituent factor that invariably operates where that process occurs. For example, prey density and predator density must always operate where predation occurs. On the other hand, predator speed is a subsidiary component, because it is not relevant in the case of an ambush predator, or a filter-feeder, or a Portuguese man of war. In short, we assume that for any process, such as predation, there is a basic model that explains the underlying processes common to all species pairs of predators and prey. We further assume that the great diversity of different predator-prey processes is caused by the additions to this basic model that occur in various situations as subsidiary components. Hence, a sequence of experiments is conducted, beginning with experiments on a small group of basic components of a process. The experiments are designed to analyze the operation of these components in sufficient depth that a mathematical model can be built to describe the operation of the components. The model is so chosen as to meet two desiderates: it must give a statistically satisfactory fit to the data and it must incorporate real insight into the phenomenon.

Additional experiments are conducted until all the basic components of the process have been analyzed, modelled, and incorporated into a systems model. Then the whole process is repeated for the various subsidiary components that can be found in the various situations where the type of process under study occurs. Finally, a systems model is constructed which can be used to simulate the process on a computer. Any particular instance of the process (representing a particular variant of the model) can be simulated, merely by informing the computer, by means of control and parameter cards, which terms are to be added to the basic model, and what values the relevant parameters are to take.

The preceding is a very general description of what is in practice a highly complex experimental and mathematical pro-

cedure. In order to give any real insight into the kind of reasoning used in the experimental components approach, it is necessary to describe a particular application. Holling's program on predation will be explained so as to bring out the sequence of steps in his reasoning.

Holling first decided, on the basis of his own observations, experiments on many predator-prey systems, and study of the literature that five groups of variables affect predation:

- density of the prey species
- density of the predator species
- characteristics of the prey species
- characteristics of the predator species
- characteristics of the environment.

The first two variables must operate in every predator-prey situation, and hence are basic variables. There are situations in which the last three groups of variables do not affect predation; therefore they are subsidiary. The two basic factors can each have their effect through numerous causal pathways. Thus, there are four components of the response to prey density by predators, each of which can be subdivided as follows:

- (1) Searching rate —predator speed relative to prey speed
 - the maximum distance from a predator at which it notices and attacks a prey, and the effect of hunger on this distance.
 - the proportion of attacks that result in successful capture.
- (2) The time predators are exposed to prey
 - time predators spend in non-feeding activities.
 - time predators spend in feeding activities.
- (3) Time spent handling each prey
 - time spent in pursuit for prey
 - time spent eating
 - digestive pause, while predator is not hungry enough to attack
- (4) Hunger
 - rate of digestion and assimilation
 - maximum food capacity of the gut

Experiments are designed to obtain the form of, and parameter

values in mathematical equations for each of these components and sub-components. As such equations are developed, they are incorporated into a systems model which is carefully checked against experimental data on the whole system. Finally, the systems model, when complete, serves as the basis for computer simulation and optimization studies. These give rise to new insight about the most important criteria for evaluating biological control agents, and the most effective parameter values for these criteria.

A similar components analysis can be conducted for any type of complex process in the ecology-behavior domain. It should be emphasized that the key concept in this whole approach is the intimate reciprocal feedback between the experimental program and computer analysis and simulation, with the model arising from one type of data and being subjected to test against other data at each stage in its development. Whenever the simulation studies predict a result not corroborated by experiment, the model clearly needs re-examination.

Literature Cited

- Goldberg, S.
1958. INTRODUCTION TO DIFFERENCE EQUATIONS. N. Y. Sci. Editions, John Wiley and Sons.
- Grosenbaugh, L. R.
1965. GENERALIZATION AND REPARAMETERIZATION OF SOME SIGMOID AND OTHER NONLINEAR FUNCTIONS. *Biometrics* 21:708-714.
- Holling, C. S.
1961. PRINCIPLES OF INSECT PREDATION. *Annu. Rev. Entomol.* 6:163-182.
- Holling, C. S.
1965. THE FUNCTIONAL RESPONSE OF PREDATORS TO PREY DENSITY AND ITS ROLE IN MIMICRY AND POPULATION REGULATION. *Entomol. Soc. Canada Mem.* 45.
- Holling, C. S.
1966. THE FUNCTIONAL RESPONSE OF INVERTEBRATE PREDATORS TO PREY DENSITY. *Entomol. Soc. Canada Mem.* 48.
- Murphy, G. M.
1960. ORDINARY DIFFERENTIAL EQUATIONS AND THEIR SOLUTIONS. Princeton, N. J. D. Van Nostrand.
- Turner, M. E., R. J. Monroe, H. J. Lucas, Jr.
1961. GENERALIZED ASYMPTOTIC REGRESSION AND NON-LINEAR PATH ANALYSIS. *Biometrics* 17:120-143.
- Turner, M. E., R. J. Monroe, and L. D. Homer.
1963. GENERALIZED KINETIC REGRESSION ANALYSIS: HYPERGEOMETRIC KINETICS. *Biometrics* 19:406-428.
- Watt, K. E. F.
1955. STUDIES ON POPULATION PRODUCTIVITY. I. THREE APPROACHES TO THE OPTIMUM YIELD PROBLEM IN POPULATIONS OF *TRIBOLIUM CONFUSUM*. *Ecol. Mongr.* 25:269-290.
- Watt, K. E. F.
1959. A MATHEMATICAL MODEL FOR THE EFFECT OF DENSITIES OF ATTACKED AND ATTACKING SPECIES ON THE NUMBER ATTACKED. *Canad. Entomol.* 91:129-144.
- Watt, K. E. F.
1967. ECOLOGY AND RESOURCE MANAGEMENT. McGraw-Hill Book Co., N. Y.

DYNAMIC MODELS FOR POPULATION SYSTEMS

by D. GORDON MOTT, Principal Ecologist, Forest Insect and Disease Laboratory, Northeastern Forest Experiment Station, Forest Service, U. S. Department of Agriculture, Hamden, Connecticut.

STUDENTS of forest insect population dynamics have been remarkably slow in adopting those mathematical methods that have been widely developed in the study of other dynamic systems. I think that most of us have been seduced by statistical methods and, as a result, have been intensively engaged in making careful measurements and analyses without having given any rigorous thought to the logical structure of the system we are concerned with. We have been guided by our intuition in choosing system parameters for study.

In this paper I will discuss the mathematical structure of some forms of dynamic system models. There are two objectives to the discussion. First, I hope to show that the construction of a mathematical system model will provide all the relevant information about the system—whether our interests are in understanding the biological structure, in applying regulation, or in predicting system behavior. Second, I hope to make clear that the construction of a system model will give the best possible guide to system dynamics research by focusing effort on the measurement of important system parameters rather than relying upon intuition to serve as is generally the case at present.

System Models

It is convenient to use as a point of departure, the population change model:

$$N_{t+1} = \prod_{i=1}^h S_i \cdot P \cdot F \cdot N_t \quad (1)$$

where N_t is number at time t , S_i is survival during age interval i , P is proportion of adults that are female, and F is fecundity.

It is interesting that this equation represents the most advanced form of our thinking about population system dynamics today. The interest arises because it is so patently incomplete, as will be seen shortly.

Equation (1) is a *difference equation* model for a population system. In simplest form, where each member of a population produces in a unit of time k individuals, and a proportion m of the individuals present at time t dies in a unit of time, we find

$$N_{t+1} = kN_t - mN_t \quad (2)$$

If we knew the values of k and m and the initial value of N , we would be able to calculate the number at any time thereafter. That is, from (2) it would be possible to obtain a numerical solution for N as a function of time. In fact, this is our objective in constructing a system model: we want to know how the system will behave over time so that we can regulate it, harvest from it, and so forth. Now, a numerical solution might be perfectly acceptable, but in this case it is possible also to obtain an analytic solution:

$$N = f(t)$$

where $f(t)$ is a function of time defined on integral values of t . To find f in this case, we proceed:

$$\text{Given (2): } N_{t+1} = kN_t - mN_t, N_{t=0} = N_0$$

$$\text{Then: } N_{t+1} = (k-m) N_t \quad (2a)$$

$$\text{or: } N_{t+1} - (k-m) N_t = 0$$

Assume that $N = Cx^t$, such that if C and x were known, a solution could be found. To find C and x , assume this solution and substitute in (2a):

$$Cx^{t+1} - (k-m) Cx^t = 0$$

$$Cx^t [x - (k-m)] = 0$$

$$\text{Since: } Cx^t \neq 0$$

$$\text{Then: } x - (k-m) = 0 \text{ and } x = k-m$$

To find C :

$$\text{Since: } N_{t=0} = N_0$$

$$\text{Then in: } N = Cx^t$$

$$N_0 = Cx^0$$

$$\text{or } C = N_0$$

Therefore: $N = N_0 (k-m)^t$ is the desired solution (figure 1).

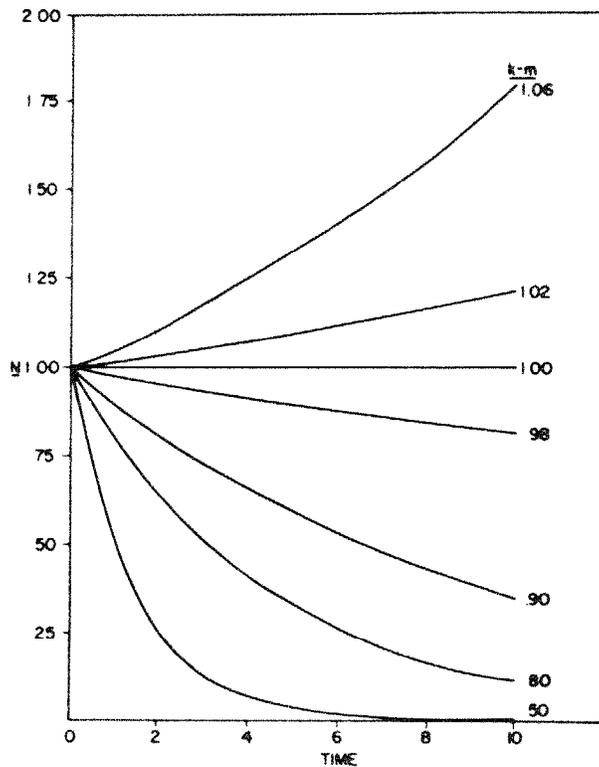


Figure 1.—Solution of the difference equation $N_{t+1} - (k-m)N_t = 0$, for given values of $(k-m)$.

It would be unusual to encounter such a simple system. Mortality usually depends upon other system elements such as predators, parasites, and disease organisms. Thus, if P represents, say, predator abundance, and in a unit of time each predator kills h of organism N , we would obtain:

$$N_{t+1} = kN_t - hP_t \quad (3)$$

However, in order to obtain a numerical solution in this case, we must know the value of P at each time t , P_t . Clearly, the system model is incomplete unless we know how P changes over time. This is exactly why the model in (1) is so deficient: it says nothing whatever about the many elements other than N present in the system. Thus, the substantial effort currently devoted to field studies with the objective of measuring the components of

(1), which ignore in large part the other system components, will produce seriously incomplete knowledge about the system.

A complete system model in the present case would consist of equation (3) and an equation for P. Suppose that during a unit of time a proportion r of P dies, and that each N permitted s progeny of P to be produced (s would be a complex parameter that included the probability of encounter between N and P and the reproductive rate of P). The following set of equations would then describe the system:

$$\left. \begin{aligned} N_{t+1} &= kN_t - hP_t & N_{t=0} &= N_0 \\ P_{t+1} &= sN_t - rP_t & P_{t=0} &= P_0 \end{aligned} \right\} \quad - \quad (3a)$$

Here, too, it is possible to obtain a numerical and an analytic solution for N and P over time.

We have now progressed to the point where the structure of a dynamic system model can be appreciated. It consists of these fundamental elements:

a. equations that state how each variable in the system changes *over time* — dynamic equations that incorporate:

b. the interrelationships among system elements described by static equations.

This deserves some amplification. A great deal of biological research has been devoted to obtaining the material in (b) above. The relationships between the number of attacked hosts and host density, parasite/predator density, spatial distribution of each organism, and so forth are fairly well known in some circumstances. Relationships between reproductive capacity and food quality and quantity, temperature (or "heat units"), and so forth are similarly known. In fact, the description of such relationships seems to be the main objective of many population studies, and both field studies and behavioral, genetic, and physiological studies in the laboratory have produced a large catalog of relationships of this kind. However, it is important to realize that the study of system dynamics depends critically upon such relationships being incorporated into mathematical descriptions of the way in which systems change *over time*. Thus, static relationships must be of suitable form to be incorporated

into dynamic equations and, in order for the dynamic model to be constructed for any system, all such relevant relationships must be known. It follows that early in the establishment of system studies thought must be devoted to the form of the system model.

The difference equation models discussed above are particularly simple; they are linear models. In general, system models will be non-linear. The distinction lies in the form of the equations. For any model in which variables occur only to the first power, no products of variables occur, and variables do not appear as exponents, the model is said to be linear. Analytic solutions to linear models can be found. In general, this is not possible for non-linear models. However, numerical or graphical solutions can always be obtained.

As an example of a non-linear difference equation model, consider the following much simplified case. Suppose we are dealing with an insect like the white pine weevil (W) which attacks white pine leaders (P). Suppose that the number of leaders that will be attacked in any season depends upon the number of weevils present. That is, as the weevil population increases, intraspecific competition intensifies for attack sites and the increase in the number of attacks per weevil declines. A simple function to describe (approximately) this sort of curve would be:

$$A_t = W_t \left\{ 1 - \exp \left(- b \frac{P_t}{W_t} \right) \right\} \quad (4)$$

where A is the number of leaders attacked in a given year and b is a constant. Suppose, furthermore, that 2 years are necessary before an attacked leader again enters the leader population. Then, the following difference equation could be constructed for leader population dynamics:

$$\begin{aligned} P_{t+1} &= P_t - A_t + A_{t-1} \\ &= P_t - W_t \left\{ 1 - \exp \left(- b \frac{P_t}{W_t} \right) \right\} + W_{t-1} \left\{ 1 - \exp \left(- b \frac{P_{t-1}}{W_{t-1}} \right) \right\} \quad (5) \end{aligned}$$

In order to construct an equation for weevil population dynamics, let us suppose that each attack in a season produces

a constant number, k , weevils in the next season. Thus:

$$W_{t+1} = kW_t \left(1 - e^{-b \frac{P_t}{W_t}}\right) \quad (5a)$$

Thus, (5) and (5a) together with initial values constitute a complete system model. It is non-linear since the variable W appears in the exponent of both functions. Some solutions for this model are shown in figure 2.

Besides the non-linearity, an additional element characteristic of biological systems has been introduced—a time-lag in the form of an effect in the present season of events that took place in a previous season. That is, leader numbers have changed partly because of attacks in the present year and partly because of the intensity of attack in previous years. Time lags introduce a particularly interesting feature to such systems; they can account by themselves for oscillation in an otherwise non-oscillatory system. In effect, the time lag changes the "order" of the equation

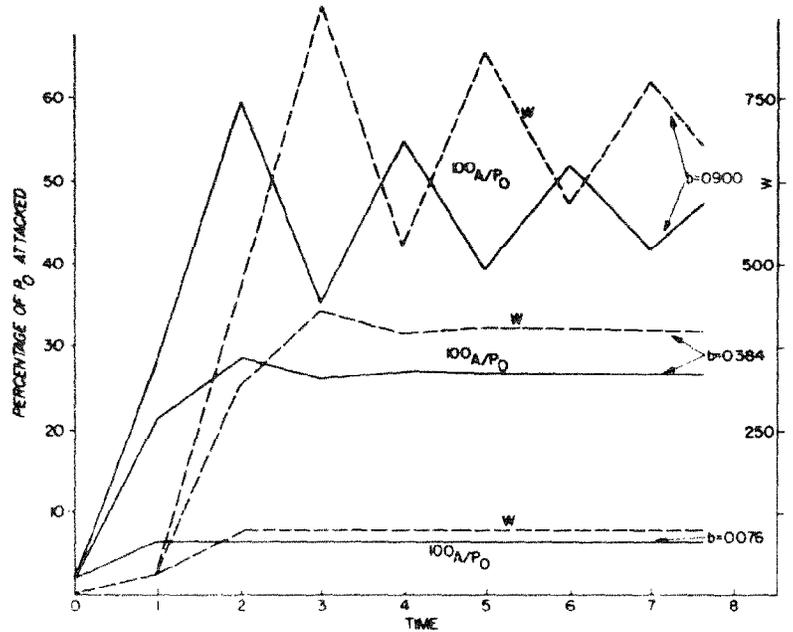


Figure 2.—Solutions of equations 5 and 5a showing behavior of the system over time.

from first to second. As a simple example of this phenomenon, consider the following general single variable system model:

$$Y_{t+1} + aY_t + bY_{t-1} = 0$$

Assume, as in equation 2a, that: $Y = Cx^t$

$$Cx^{t+1} + aCx^t + bCx^{t-1} = 0$$

$$Cx^{t-1}(x^2 + ax + b) = 0$$

$$x^2 + ax + b = 0$$

$$\text{and } x = -\frac{a}{2} \pm \frac{\sqrt{a^2 - 4b}}{2}$$

Since there are two possible values for x , assume that the solution is a linear combination of the two, that is:

$$Y = C_1 \left\{ -\frac{a + \sqrt{a^2 - 4b}}{2} \right\}^t + C_2 \left\{ -\frac{a - \sqrt{a^2 - 4b}}{2} \right\}^t$$

Since $Y(0) = Y_0$ and $Y(-1) = Y_{-1}$, the initial conditions, we can obtain solutions for C_1 and C_2 by solving the two simultaneous linear equations obtained from substituting $t = 0$ and $t = -1$.

There are several possible results depending upon the values of the coefficients a and b . If, for example, the two values of x are complex conjugates, the solution will oscillate.

Now, let us consider a second mathematical form for constructing system models—the *differential equation*. This is the most widely utilized mathematical form in the sciences. Using equation (2), $N_{t+1} = kN_t - mN_t$, we can consider the rate at which N changes per unit time. By definition, the average rate of change consists of the difference between final and initial values of N , divided by the time period over which the change took place. That is,

$$\frac{\Delta N}{\Delta t} = \frac{N_{t+\Delta t} - N_t}{\Delta t}$$

where ΔN is the increment in N and Δt the increment in t . Suppose we desired a representation for change in N that yielded intermediate values between integer values of t , either because we knew the system behaved in a continuous fashion (or nearly so, as in a bacterial population in a flask) or because there was some advantage in having such a continuous function (some of

the advantages will be apparent below). Since our objective is to obtain a continuous function, we assume that the changes in N are taking place continuously—that is, for any Δt , no matter how small, there will be some change in N . Formally, we are searching for the derivative of N —the “instantaneous rate of change” defined as

$$\text{Limit}_{\Delta t \rightarrow 0} \left\{ \frac{N_{t+\Delta t} - N_t}{\Delta t} \right\} = \frac{dN}{dt}$$

In the case of equation (2), we have shown the solution

$$N = N_0(k-m)^t$$

$$\text{or, if } k - m = r, \quad N = N_0 r^t$$

Let us now calculate the instantaneous rate of change in N :

$$\begin{aligned} \frac{dN}{dt} &= \text{Limit}_{\Delta t \rightarrow 0} \left\{ \frac{N_0 r^{t+\Delta t} - N_0 r^t}{\Delta t} \right\} \\ &= \text{Limit}_{\Delta t \rightarrow 0} \left\{ (N_0 r^t) \frac{(r^{\Delta t} - 1)}{\Delta t} \right\} \\ &= \left\{ \text{Limit}_{\Delta t \rightarrow 0} (N_0 r^t) \right\} \left\{ \text{Limit}_{\Delta t \rightarrow 0} \frac{r^{\Delta t} - 1}{\Delta t} \right\} \end{aligned}$$

The value of the first factor is obviously $N_0 r^t$, which in fact is N . In order to calculate the second limit, we resort to L'Hospital's rule and obtain:

$$\text{Limit}_{\Delta t \rightarrow 0} \left\{ \frac{r^{\Delta t} - 1}{\Delta t} \right\} = \text{Limit}_{\Delta t \rightarrow 0} \left\{ \frac{r^{\Delta t} \ln r}{1} \right\} = \ln r$$

Thus, we discover that

$$\frac{dN}{dt} = N \ln r \quad (7)$$

This is a differential equation in N . A solution to this equation would be obtained by integration and would yield

$$N = N_0 \exp(\ln r \cdot t) \quad (7a)$$

This function represents a smooth curve that passes through all the points generated by the difference equation (2).

Thus, differential equations can be viewed as approximations for difference equations, and in fact the converse also applies.

As demonstrated, the construction procedure for differential equation models for population systems is similar to that for

difference equations. A rate equation is constructed for each variable in the system. In general, these will be non-linear and involve such things as time lags and variable coefficients. One of the simplest and earliest such models is that of Lotka-Volterra. Where H and P are a host and parasite population, this model states that

$$\left. \begin{aligned} \frac{dH}{dt} &= (b - dP)H \\ \frac{dP}{dt} &= (rH - m)P \end{aligned} \right\} \quad (8)$$

This is a simple non-linear model, for which a good approximate solution exists, yielding an oscillation in time of P and H (figure 3).

This and associated models have received a great deal of attention—far more than they warrant, perhaps. Even casual speculation about the model for a predator-prey system quickly leads to a much more complex set of functions. For example, both to illustrate this point and to illuminate further the construction of a dynamic system model in differential equation form, let us construct a hypothetical model.

To begin, consider that we are dealing with a closed system in which we have neither immigration nor emigration, so that all changes must occur through births and deaths. Next, in regard

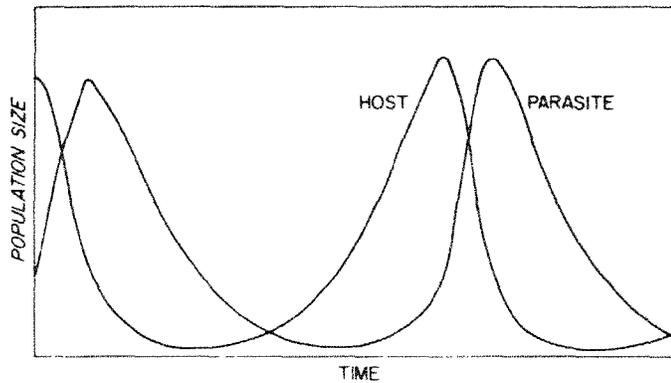


Figure 3.—Behavior of the Lotka-Volterra predator-prey model.

to host reproduction, it would be most unusual to discover that reproductive rates were constant. A more reasonable speculation would be that reproduction behaved as in the logistic equation:

$$B = \text{Birth Rate} = \frac{bH(K-H)}{K} \quad (9)$$

where H is population number, b the maximum reproductive rate, and K some constant representing the "carrying capacity" of the environment. Or, we might invoke a more complex model like that proposed by Watt (1960) as a general fecundity model. In any event, we would probably speculate that reproduction depended in some way upon density. Furthermore, it might be reasonable that it depended also upon temperature. In an environment in which temperature varied, temperature in turn could be considered a function of time, so that we might write

$$\begin{aligned} B &= f(t) \frac{(K-H)}{K} \\ &= f(t)H - \frac{f(t) H^2}{K} \end{aligned} \quad (10)$$

Here, we have a non-linear differential equation with variable coefficients. Furthermore, it might be reasonable to suppose that reproductive rates now depend upon the density of the host, H , some time ago ($t - \tau$) since it takes time (τ units) for animals to mature. Thus, we might write:

$$B = f(t) H_{t-\tau} \left\{ \frac{K-H_t}{K} \right\} \quad (11)$$

It would also be reasonable to speculate that K depended upon the food supply so that $K = g(F)$, and we would have to enter a third rate equation for F . Let us avoid this complication now.

Now, consider the death rate of the host. In the Lotka-Volterra model it is assumed to be a linear function of P . However, again because of the large amount of data available on the relationships between predators and prey, we know that this is almost certainly not the case. Rather, very likely as predator population increases relative to that of the host, predator efficiency decreases. We might utilize a function such as:

$$D = \text{Death Rate} = H(1 - \exp(-cP)) \quad (12)$$

or any other suitable function, such as from Holling (1966) or Watt (1959). Then, we could write:

$$\frac{dH}{dt} = f(t) H_t \left[\frac{K - H_t}{K} - H_t \right] \left\{ 1 - \exp(-cP) \right\} \quad (13)$$

Consider now predator dynamics. Let us assume that predator reproduction again behaved like the logistic except that in this case K depended upon H and that $K = h(H)$. Furthermore, we might presume that all predators of a certain age, a , died. We would then write:

$$\frac{dP}{dt} = bP_t - \frac{bP_t^2}{h(H_t)} - P_{t-a} \quad (14)$$

Equations (13) and (14) make up a model for this rather simple system. It is obvious that the nature of biological material takes us very rapidly into rather complex models. We have neglected many kinds of phenomena that might be considered. For example, variation in the constants would result from evolution in the system; it is known that reproductive and mortality constants depend upon oscillatory genetic changes that result from differential mortality over time. Also, probably few systems are really closed, and almost any system can be enlarged as variables are added to include the whole world. However, ultimately, the addition of more variables will have little effect on the system, and its dynamics will depend upon the effects of a few most important variables. The problem of discerning relative importance of variables obviously depends upon obtaining a knowledge of the system—and thus upon constructing its model.

Obtaining A Solution for the Model

Let us now turn to the matter of studying system behavior, given that its model is available. The easiest way to accomplish this is to obtain an analytic solution and study its form. In several cases considered above this has been done, and in any linear case an analytic solution can be found. Since in many real systems it is possible to find a linear model that is a good approximation to the real non-linear case over some operating range, it is worth

acquiring the rather elementary mathematics to obtain such solutions. Furthermore, since most population system models will be of first-order (in contrast to some physiological models in which the application of mechanical principles yield higher order equations) the very limited mathematics can be stated in a few paragraphs.

Given a linear system:

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n \\ \frac{dx_2}{dt} &= a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n \\ &\vdots \\ \frac{dx_n}{dt} &= a_{n1} x_1 + a_{n2} x_2 + \dots + a_{nn} x_n \end{aligned}$$

With given initial conditions $x_i(0)$, the system model can be written in matrix form as:

$$\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

The solution to the set of equations consists of a set of functions:

$x_i = C_{1i} \exp(\lambda_{1i}t) + C_{2i} \exp(\lambda_{2i}t) + \dots + C_{ni} \exp(\lambda_{ni}t)$ (15)
 where the C_{ij} depend upon initial conditions, and the λ are the eigenvalues of the a_{ij} matrix (providing the eigenvalues are distinct; otherwise the result is only slightly more complicated.)

The eigenvalues of the matrix are obtained by solving the polynomial equation that results from taking the determinant of the a_{ij} matrix minus λ times the identity matrix and setting it equal to zero. To illustrate:

Given:

$$\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \begin{array}{l} x_1(0) = x_{10} \\ x_2(0) = x_{20} \end{array}$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{bmatrix} = 0$$

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{21} a_{12} = 0$$

$$\text{or } \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11} a_{22} - a_{21} a_{12}) = 0 \quad (16)$$

This is a quadratic equation in λ of the form:

$$\lambda^2 + b\lambda + c = 0$$

that has roots:

$$\lambda_1 = \frac{-b}{2} + \frac{\sqrt{b^2 - 4c}}{2}$$

$$\lambda_2 = \frac{-b}{2} - \frac{\sqrt{b^2 - 4c}}{2}$$

Systems of larger size pose no particular difficulty. In order to interpret the results, it is necessary to be able to change complex roots to trigonometric form (which implies some oscillation in the solution), but in general the solutions are straightforward. Depending upon the form of the non-linearity in the model, some analytic procedures will yield solutions in the form of series. There are also graphical (isocline) methods, but they are limited to second order systems.

However, one does not really have to know any mathematics in order to solve any of these models; solutions can be obtained on either an analog or digital computer. Thus, where the biologist 15 years ago had to turn to a trained mathematician (if he could find one, if he could communicate with him, and if he could understand the answers given), today, simply by learning to program a computer, a task of about the same complexity as learning first year high-school algebra, he can perform more complete analyses of most mathematical problems.

Consider, for example, the matter of programming for an analog computer. This device can add, multiply, and integrate with respect to time, and perform the inverse operations. Given a model such as that of Lotka-Volterra, one simply supplies to the integrators (a particular component on the machine) the values of the derivatives (in the form of electrical signals, the voltage of which is proportional to the variable) and obtains the values of the variables. Since the values of the derivatives consist of combinations of the variables, the input signals are obtained from the output. Figure 4 shows a circuit for solving this model. Since the electronic analog computer is a continuous device, it solves differential equations directly.

It is difficult to simulate time lags on the analog computer, and it also is difficult to develop some of the more complex static functions. The digital computer can easily be programmed to provide these elements, and it is also possible to integrate numerically to any desired degree of accuracy. A number of numerical integration methods have been developed—Euler, Runge-Kutta, and Adams methods, to name a few—and they are easily programmed if they are not readily available as subroutines.

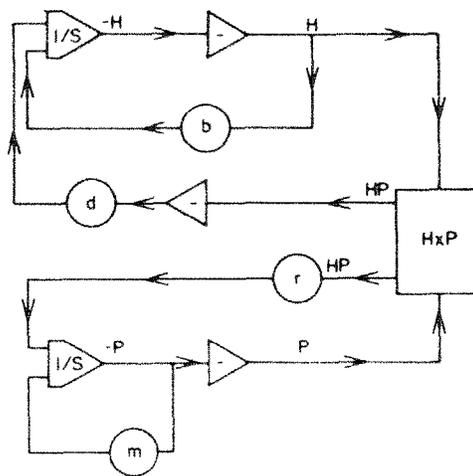


Figure 4.—Analog computer circuit for solving the Lotka-Volterra equations.