Reprinted from:
ESTUARINE RESEARCH, VOL. I
Chemistry, Biology and
the Estuarine System

THE *A POSTERIORI* ASPECTS OF ESTUARINE MODELING[1]

Robert E. Ulanowicz
David A. Flemer
Donald R. Heinle
Curtis D. Mobley[2]

ABSTRACT

This exercise is the application of an analytical method for systematically modeling ecosystems data to observations made on a naturally eutrophic, mesohaline planktonic microcosm. The theory and experimental design are briefly outlined and the particular steps in the acutal modeling process follow. Then there is a discussion as to how the whole endeavor can be refined to culminate in models with predictive capabilities.

INTRODUCTION

The type of marine ecosystem most useful to man on a per-unit basis is the mesohaline, or estuarine system. For this reason, there has been quite a strong demand from management and commercial sources for models that can predict the time behavior of these baseline systems, and the evolution of perturbations upon them. Predictive models however are elusive, and any new philosophical approach which would hasten the development of models well founded upon observation would be most welcome.

"Good data are the precursors of good models" (12), and one would expect that any models with good precision and prediction would be of an *a posteriori*

1. Contribution No. 532, Natural Resources Institute, University of Maryland, Solomons, Maryland 20688.

2. Chesapeake Biological Laboratory, Natural Resources Institute, University of Maryland, Solomons, Maryland 20688.

12. Steeman-Nielsen, E.
1957 The chlorophyll content and the light utilization in communities of plankton algae and terrestrial higher plants. **Physiol. Plantarum,** 10: 1009-1021.

13. Steeman-Nielsen, E. and Hansen, V. K.
1959 Measurement with the carbon 14 technique of rates of respiration in natural population of phytoplankton. **Deep-Sea Res.,** 5: 222-233.

14. Teal, J. M.
1962 Energy flow in the salt marsh ecosystem of Georgia. **Ecology,** 43: 614-624.

15. Teal, J. M. and Kanwisher, J.
1961 Gas exchange in a Georgia salt marsh. **Limnol. Oceanogr.,** 6: 388-399.

16. Teal, J. M. and Wieser, W.
1966 The distribution and ecology of nematodes in a Georgia salt marsh. **Limnol. Oceanogr.,** 11(2): 217-222.

17. Wiegert, R. G.
1973 A general ecological model and its use in simulating algal-fly energetics in a thermal spring community. In **Insects: studies in population management,** Vol. 1, (eds. Geier, P. W., Clark, L. R., Anderson, D. J., and Nix, H. A.) Occasional Papers, Ecol. Soc. of Austr., Canberra.

18. Wiegert, R. G.
1974 Simulation modelling of the algae-fly components of a thermal ecosystem: effects of spatial heterogeneity, time delays and model condensation. In **Systems analysis in Ecology,** Vol. 3., (ed. Patten, B. C.) Academic Press, New York. (In press.)

19. Wolf, P., Shanholtzer, S. and Reimold, R.
Population estimates for **Uca pugnax** on the Duplin estuary marsh. **Crustaceana** (In press.)

nature in the sense of Watt (16). Generally, such models are accepted to be ones in which data are regressed to determine the parameters of a given model. But *a posteriori* models can be more precisely defined. To be specific, it is useful to consider Dale (3) and Ross' (13) chronology of the problem-solving process: (A) the lexical phase, or the delimination of the entities or parts; (B) the parsing phase, or the choice of relationships between entities of interest; (C) modeling, or the specification of the mechanism by which these relationships take place; and (D) the analysis or validation of the model. Regression usually occurs at step (D), after the first three phases have been initially executed on *a priori* grounds. In a truly *a posteriori* approach, regression should take place before steps (B) and (C) and should strongly influence the subsequent steps.

Bellman (1), in treating regression as the most significant step in the modeling process, referred to a procedure with such emphasis as solving the "inverse problem" in ecological modeling. It remained, however, for Mobley (9, 10) to couple regression with statistical hypothesis testing and construct an algorithm whereby regression is actually antecedent to parsing and modeling.

From the foregoing, it might seem as if the development of *a posteriori* models has languished because of the lack of an adequate theoretical framework. In reality, however, a dearth of good data on multi-species population dynamics has probably been the rate-limiting step. One can readily cite the two-species competition data of Gause (4) and Park (11), but it is not easy to find equally adequate information on the time dynamics of all the compartments involved in the biomass cycle of an ecosystem. That this vacuum exists is not surprising, given the complexities of most real systems, the manifold of exchanges of any particular open system with its neighbors, and the unpredictability of the intrinsic variables that drive the system.

In planning the data collection, the authors tried to surmount the latter two difficulties by enclosing a system and subjecting it to nearly constant salinity, temperature, and incident light. In an effort to assemble populations large enough to reduce sampling noise, the system was limited to planktonic species.

The two sections immediately following outline the theory and experimental methods in order to provide background for the example analysis, in which the modeling process is demonstrated.

THEORY

The inverse problem, as viewed by the mathematician, is as follows: Given observations on several populations, N_i at l separate times, how does one construct a model that will fit the data as closely as possible, and yet include only those interactions which the data indicate are significant?

Since the object of this approach is to make a few *a priori* assumptions as possible about the significant interactions, it will be assumed that all possible

interactions of a given class are initially pertinent. For the sake of mathematical tractability, the class of models under consideration will be of the form

$$\frac{d\,Ni}{dt} = \sum_{p=1}^{k} b_p f_p \tag{1}$$

where the b_p are the parameters which multiply the arbitrary functions f_p. The f_p model the various phenomena which influence the change in the population of species i and may be any single-valued function of any population, time, or extrinsic variable, e.g.

$$f_p = f_p(N_1, N_2, \ldots, N_n, t, \text{etc.}) \tag{2}$$

In addition, any phenomenon may be represented in equation 1 by more than one f_p, where the different f_p are alternative mathematical representations of the same effect.

This general model must first be regressed to fit the data. Subsequently, each f_p must be examined for its contribution to the fit, and those terms with minimal contribution are then systematically dropped from the model. This systematic exclusion is accomplished with the help of statistical hypothesis testing, and effects both parsing and modeling (since alternative mathematical expressions may be evaluated against one another). The full mathematical exposition of the regresseion and hypothesis testing is to be found in Mobley (9, 10); only a brief outline of these methods is given below.

The method of regression used is specially adapted to the overall analysis and is fundamentally different from standard linear regression. To demonstrate what is involved, it is useful to write (1) in full.

$$\begin{aligned}
\left.\frac{d\,Ni}{d\,t}\right|_{t_1} &= b_1 f_1(t_1) + b_2 f_2(t_1) + \ldots + b_k f_k(t_1) + e_1 \\
\left.\frac{d\,Ni}{d\,t}\right|_{t_2} &= b_1 f_1(t_2) + b_2 f_2(t_2) + \ldots + b_k f_k(t_2) = e_2 \\
&\vdots \\
\left.\frac{d\,Ni}{d\,t}\right|_{tl} &= b_1 f_1(tl) + b_2 f_2(t_2) + \ldots + b_k f_k(tl) + el
\end{aligned} \tag{3}$$

The left-hand sides in equation (3) are the observed rates of change as calculated from the series of observed populations by any one of a number of standard techniques. The first k terms on the right-hand side yield the rate of change at t_i as predicted by the hypothesized interaction f_j. Finally, the e_i are the differences between the observed and predicted rates of change; i.e., the "error" to be minimized.

Equations (3) can consisely be written in matrixvector form as:

$$\bar{Y} = \bar{\bar{X}}\,\bar{B} + \bar{E} \tag{4}$$

where:

$$Y_j = \frac{dN_i}{dt}(\text{at } t = t_j) \tag{4a}$$

$$X_{ij} = f_j(t_i) \tag{4b}$$

$$B_j = b_j \tag{4c}$$

$$E_j = e_j \tag{4d}$$

In general Y is a vector in n-dimensional Euclidean space, whereas the column vectors of matrix $\bar{\bar{X}}$ span a subspace of dimension m, V_m ($m \leq n$). The special case $n = 3$, $m = 2$ is illustrated graphically in Figure 1. All possible vectors $\bar{\bar{X}}\bar{B}$ will lie in subspace V_m (the plane defined by $\bar{X}_1$, $\bar{X}_2$ in Figure 1). It is readily seen from Figure 1 that the magnitude of the error vector $\bar{E}$ will be minimal when $\bar{E} = \bar{E}^* \perp V_m$. Or equivalently,

$$\bar{\bar{X}}^t\bar{E} = \bar{0} \tag{5}$$

applying (2.6) to (2.5) one arrives at the familiar solution

$$\bar{B}^* = [\bar{\bar{X}}^t\bar{\bar{X}}]^{-1} \quad [\bar{\bar{X}}^t\bar{Y}] \tag{6}$$

Standard linear regression theory is an implementation of equation 6.

The execution of (6) is difficult when the dimension of $\bar{\bar{X}}$ becomes large and it is impossible when the columns of $\bar{\bar{X}}^t\bar{\bar{X}}$ are linearly dependent. To circumvent these difficulties, Mobley works instead with an algorithm that minimizes the magniture of $\bar{E}$ by, equivalently, maximizing the value of $\cos^2\theta$. This method begins with the assumption $\bar{B} = \bar{0}$ and determines which element of $\bar{B}$ can be changed by what amount δ such that the resulting increment in $\cos^2\theta$ is the maximum obtainable by changing only one element of $\bar{B}$. The element of $\bar{B}$ thus located is incremented by the appropriate value δ, and the method is reiterated with the revised B until the calculated increment in $\cos^2\theta$ is negligible. The vector $\bar{B}$ must converge to $\bar{B}^*$ since $\cos^2\theta$ is bounded from above by unity.

The above iterative procedure is useful in more than just obviating the numerical difficulties that accompany equation 6. In practice it often happens that the required fit is achieved without ever changing certain elements of $\bar{B}$

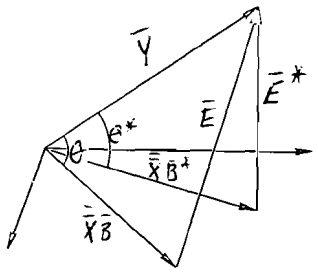

Figure 1. Illustration of $\bar{Y} = \bar{\bar{X}}\bar{B} + \bar{E}$.

from their initial values of zero. The corresponding f_p may be dropped from further consideration, thus effecting a preliminary contribution to the parsing.

Having fit the general model to the data, it is necessary to ascertain which of the many f_p may be neglected without significantly affecting the ability of the resultant system to mimic the observations. For the sake of brevity, Ω will denote the set of assumptions that the data can be represented by the general model, and that the error vector of the fit, E*, has as elements normally distributed, statistically independent, random variables with means of zero and identical, constant, albeit unknown, variances. A measure of the inability of the full model to perfectly represent the data will be designated by

$$\mathfrak{J}_{\Omega} = \bar{E}^t\bar{E}. \tag{7}$$

H will denote the hypothesis under test; namely, that any given combination of q elements of B may be set equal to zero (i.e., H denotes the hypothesis that the interactions modeled by the corresponding f_p are insignificant). Finally, the error-sum-of-squares of the resultant reduced model, $W(W = \Omega \cap H)$, will be labeled $\mathfrak{J}_w$.

One sees intuitively that H should be rejected if $(\mathfrak{J}_w - \mathfrak{J}_{\Omega})$ becomes large. To be more quantitative, it is possible to employ the likelihood-ratio statistic

$$F = \frac{(\mathfrak{J}_w - \mathfrak{J}_{\Omega}) / q}{\mathfrak{J}_{\Omega} / (n-k)} \tag{8}$$

where, under W, F is a central F-variable with q and (n-k) degrees of freedom. For any chosen level of significance, say .05, one can determine a value, $F_{.05}$, such that if $F > F_{.05}$, the hypothesis H may be rejected with 95% confidence.

On the other hand, if $F < F_{.05}$, H is accepted, which implies that the reduced model can provide as adequate a mathematical description of the observed population dynamics as can the general model. It is possible, however, that this accepted hypothesis is in reality false. It is useful then to compute the

power of the test; that is, the probability that if H is false, it would have been correctly rejected. If H is supposed false, then Ω (but not W) is assumed to hold and the F in (8) becomes a non-central F-variable, F'. The power of the test, β, is given by

$$\beta = \text{Prob}\,(F' > F_{.05}) \tag{9}$$

One observes that as the confidence level is raised to insure that a valid reduced model is not falsely rejected, it becomes more probable that a spurious reduced model will be accepted as a good representation of the data (i.e., the higher the confidential level, the lower the power). Selection of a significance level commensurate with the noise of the data is one of the few magnitudes the modeler must estimate.

By systematic iteration of the apparatus outlined above, the full model can be pruned until it consists only of terms which contribute significantly to the model's ability to mimic the data. An example of the application of this procedure to planktonic data is presented in Section 4, and the biological significance of both the modeling process and the ensuing model is discussed in Section 5.

METHODS

The microcosms under study were contained in 757-liter polyethlyene cylinders sheltered in the Chesapeake Biological Laboratory pier house. Twelve hours of daylight were simulated by a 500 W General Electric quartz-iodide, wide-flood lamp placed 1.5 m above each cylinder. Lighting was attenuated to a satisfactory level at the surface of 0.18 cal m^{-2} min^{-1} with layers of window screening. Twilighting was provided each cylinder by a single 20 W cool-white fluorescent light which preceded and followed the intense lighting by ½ hour. Temperature varied diurnally by about 2°C around a mean of 22°C. The tanks were stirred four times a day by large polyethylene propellers driven at 24 rpm for intervals of one hour.

To fill the tanks initially, Bay water from a depth of 0.5 m was pumped through a 28 μ plankton net into a reservoir and thence simultaneously fed to the three experimental chambers. Natural populations of rotifers, protozans, and algae were thereby introduced. The final salinity of the systems was 9 °/oo. The copepod *Eurytemora affinis* (Poppe) was added to the system from mass cultures grown by the techniques of Heinle (5). When predators were desired, the selective planktivore, *Menidia menidia* (Linnaeus) larvae, were added.

The systems were allowed to evolve in the batch mode for a period of 15 days. They were sampled at the same time each day by dipping water from the surface with a bucket after the stirrers had run for 10 to 20 minutes. Each

sample was analyzed for the following: total seston, particulate carbon, particulate nitrogen, particulate carbohydrate, total phosphorous, dissolved organic phosphorous, dissolved inorganic phosphorous, ammonia, nitrates, nitrites, chlorophyll *a*, primary productivity (C^{14} uptake), algal taxonomy and cell count, herbivore biomass, salinity, temperature, and dissolved oxygen. A number of the above variables were dominated by sampling and analytical noise. Discernible trends appeared, however, on the abiotic level in the dissolved nitrogen ($NO_3 + NO_2$), on the producer level as active chlorophyll *a*, and on the second trophic level in the total herbivore biomass. For brevity, only the analyses of these quantities germane to the model derived in Section 4 are referenced below.

Part of the water samples collected were filtered through GF/C glass fiber filters. The filtrate was analyzed for nitrite and nitrate by the procedure of Strickland and Parsons (14). Chlorophyll *a* was extracted from the filter pads with a 90% acetone solution and the resulting concentration was measured on a Turner fluorometer after Yentsch and Menzel (18) and Holm-Hansen (7). Calibration determination of active chlorophyll *a* followed the procedure of Lorenzen (8) and Yentsch (17).

Herbivore biomass was estimated by measuring necessary dimension of all the herbivores found in the material retained from a one-liter sample by 63 μ pore size net. Previously established length-weight relationships were used for the copepods (6). Dimensions of representative samples of rotifers and protozoans were measured, their volumes thereby estimated, and dry weights arrived at by assuming a water content of 80%. With few exceptions, copepods dominated the herbivore biomass.

EXAMPLE ANALYSIS

Data taken during two weeks of observation of a microcosm are displayed in Table 1. The compartments to be modeled are those previously identified–nitrate plus nitrite (species 2), active chlorophyll *a* (species 3), and herbivore biomass (species 1). A predator was also present, but its biomass remained unchanged during the course of the experiment and is deleted from further analysis. Data were taken for 13 days, although each component was not measured on each day. Since the modeling algorithm requires the populations to be measured simultaneously, linear interpolation between adjacent measurements is used to provide missing values.

To minimize the *a priori* assumptions imposed upon the mathematical form of the model, a generalized first-order, quadratic, ordinary differential equation is employed as the "full" model for each species:

TABLE 1.

The populations of the three ecosystem components modeled. Component 1 is total herbivore biomass in μg dry wt./liter, Component 2 is dissolved nitrogen in μg at N/liter, and Component 3 is active chlorophyll a in μg/liter. An asterisk denotes values obtained by interpolation.

TIME IN DAYS	COMPONENT 1	COMPONENT 2	COMPONENT 3
0	0	8.	5.4
1	3.	8.	4.7
2	7.	7.	7.8
3	5.	4.	8.0
4	10.	5.5*	9.5
5	40.	7.	8.5
6	55.*	6.	10.1
7	70.	7.*	11.8*
8	73	8.	13.5
9	56.	7.*	13.0
10	67.*	6.	10.1
12	89.	4.*	8.6*
13	69.*	3.	7.8

$$\frac{d\,Ni}{d\,t} = A_i + \sum_{j=1}^{3} B_{ij}N_j + \sum_{j=1}^{3} C_{ij}N_iN_j + \sum_{\substack{j=1 \\ j\neq i}}^{3} D_{ij}N_j^2 \qquad (10)$$

Referring to (1), the correspondences $b_1 = A_i$, $f_1 = 1$, $b_2 = B_{i1}$, $f_2 = N_i$, etc. are obvious. Equation (10) may be viewed as generalized Lotka-Volterra interactions or may be regarded as the result of expanding the f_p as Taylor series and retaining only the linear and lowest order non-linear terms.

It is important to remember that when using equation 10, it is not the populations that are being modeled, but rather their derivatives. These derivatives are not directly measurable and must be estimated by some suitable scheme. For the present example, the derivative at a given time t_i is estimated by averaging the slopes of the straight-line segments connecting the population at t_i to the two adjacent populations at t_{i-1} and t_{i+1}.

Once the three final single-component models have been obtained, the equations must be simultaneously integrated to yield the predicted populations. Only a small change in a population derivative can sometimes lead to a large

change in the population, especially after long integration times. Thus it often occurs that a restricted model which by acceptance of the associated hypothesis is equivalent to the full model, yields upon integration a population significantly different from the population obtained from integration of the full model. Since one goal of our analysis is a model with predictive power, there is imposed the additional requirement that a derivative model, when integrated, closely mimic the data.

To proceed with the actual modeling, equation 10 is applied to each component in turn. The iterative algorithm for solving the "inverse problem" yields the sets of least-squares parameters for the three general models as tabulated in Table 2. The predicted populations obtained by simultaneous integration of these general models are exhibited in Figure 2. To gain some idea as to which terms of the general model best represent the dominant ecosystem

TABLE 2

The values of the least-squares parameters for the general models of equation (10).

COMPONENT 1		COMPONENT 2		COMPONENT 3	
A_1	$= -4.518 \cdot 10^{1}$	A_2	$= 9.967 \cdot 10^{-4}$	A_3	$= 2.484$
B_{11}	$= -2.859 \cdot 10^{-1}$	B_{21}	$= -3.068 \cdot 10^{-2}$	B_{31}	$= -6.048 \cdot 10^{-2}$
B_{12}	$= 2.636 \cdot 10^{1}$	B_{22}	$= 8.386 \cdot 10^{-3}$	B_{32}	$= -1.553 \cdot 10^{-1}$
B_{13}	$= 0$	B_{23}	$= 1.264 \cdot 10^{-1}$	B_{33}	$= -4.891 \cdot 10^{-2}$
C_{11}	$= -1.811 \cdot 10^{-3}$	C_{21}	$= 7.897 \cdot 10^{-3}$	C_{31}	$= 3.910 \cdot 10^{-3}$
C_{12}	$= 1.756 \cdot 10^{-1}$	C_{22}	$= -9.478 \cdot 10^{-3}$	C_{32}	$= 2.158 \cdot 10^{-1}$
C_{13}	$= -4.630 \cdot 10^{-2}$	C_{23}	$= -9.988 \cdot 10^{-3}$	C_{33}	$= -9.443 \cdot 10^{-2}$
D_{12}	$= -2.502$	D_{21}	$= -1.342 \cdot 10^{-4}$	D_{31}	$= 2.053 \cdot 10^{-4}$
D_{13}	$= -2.036 \cdot 10^{-1}$	D_{23}	$= -9.999 \cdot 10^{-3}$	D_{32}	$= -1.088 \cdot 10^{-1}$

interactions, it is useful to test the series of hypotheses resulting when the model parameters are deleted, one at a time, from the general model. For brevity, only one component, say, the herbivore biomass, is discussed. The series of hypotheses are

$$H_1: A_1 = 0$$
$$H_2: B_{11} = 0$$
$$\vdots$$
$$H_9: D_{13} = 0$$

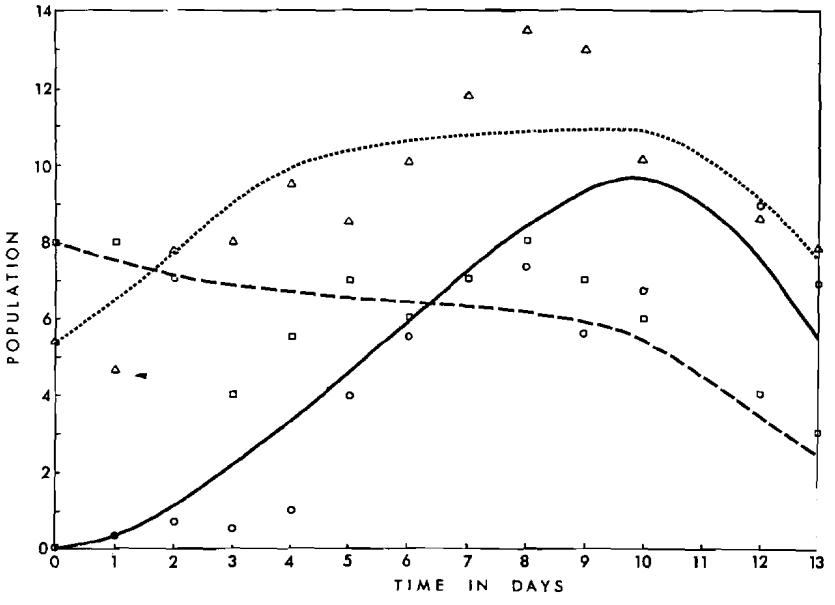

Figure 2. The observed populations and the populations predicted by the general models of equation (10). o and —— are the observed and predicted total herbivore biomasses, respectively, in μg dry wt./liter times 10^{-1}; □ and – – –are the observed and predicted dissolved nitrogen concentration in μg at N/liter; and Δ and - - - - are the observed and predicted active chlorophyll *a* concentrations in μg/liter.

The results of the tests of $H_1, \ldots, H_9$ when made at the .25 significance level are shown in Table 3.

Table 3 shows that H_1 is rejected (with a .75 probability that the rejection is correct), H_2 is accepted (with a .71 probability that if H_2 were really false it would have been correctly rejected), and so on. These results immediately suggest testing the hypothesis

$$H_{10}: B_{11} = B_{13} = C_{11} = C_{13} = D_{13} = 0$$

And indeed, H_{10} is accepted with a power of .59. A further series of tests can determine whether any more parameters can be deleted from the general model:

$$H_{11}: A_1 = B_{11} = B_{13} = C_{11} = C_{13} = D_{13} = 0$$

$$\vdots$$

$$H_{14}: B_{11} = B_{13} = C_{11} = C_{13} = D_{12} = D_{13} = 0$$

TABLE 3

Results of the first hypothesis tests on Component 1. In the case of acceptance, the power of the test is given. The significance level was .25.

HYPOTHESIS	REJECTION OR POWER
H_1: $A_1 = 0$	rejected
H_2: $B_{11} = 0$	.71
H_3: $B_{12} = 0$	rejected
H_4: $B_{13} = 0$	.65
H_5: $C_{11} = 0$	.73
H_6: $C_{12} = 0$	rejected
H_7: $C_{13} = 0$	.77
H_8: $D_{12} = 0$	rejected
H_9: $D_{13} = 0$	.63

H_{11}, H_{12}, and H_{14} are all rejected. However, H_{13}, which attempts to delete C_{12} in addition to the parameters already deleted by acceptance of H_{10}, is accepted with a power of .63. This may seem a surprising result since H_6: $C_{12} = 0$ was rejected. It must be remembered that the degrees of freedom for the two F-tests are different; thus, the possibility of such a result. It is found, however, that upon integration, the populations predicted by the model resulting from H_{13} deviate considerably (about 50%) from the observed populations, and for this (subjective) reason the model is excluded from further consideration. Thus, the final model for component has the structure

$$\frac{d\,Ni}{d\,t} = A_1 + B_{12}N_2 + C_{12}N_1N_2 + D_{12}N_2^2 \qquad (11)$$

It is to be noted that five of the nine terms of the general model have been found to play no significant role in the modeling. Each of the remaining terms does contribute significantly to the integrated model's ability to mimic the observed data.

The analysis proceeds similarly for the other two ecosystem components, the results being

TABLE 4.

The values of the least-squares parameters for the final component models of equations (11) and (12).

COMPONENT 1	COMPONENT 2	COMPONENT 3
$A_1 = -1.171 \cdot 10^{2}$	$B_{23} = 1.489 \cdot 10^{-2}$	$A_3 = 4.623$
$B_{12} = 4.404 \cdot 10^{1}$	$C_{21} = 5.286 \cdot 10^{-3}$	$B_{31} = -1.158 \cdot 10^{-2}$
$C_{12} = -9.998 \cdot 10^{-3}$	$C_{23} = 1.498 \cdot 10^{-2}$	$B_{32} = -1.378 \cdot 10^{-2}$
$D_{12} = 3.647$	$D_{21} = -3.418 \cdot 10^{-4}$	$C_{32} = 2.223 \cdot 10^{-1}$
		$C_{33} = -9.034 \cdot 10^{-2}$

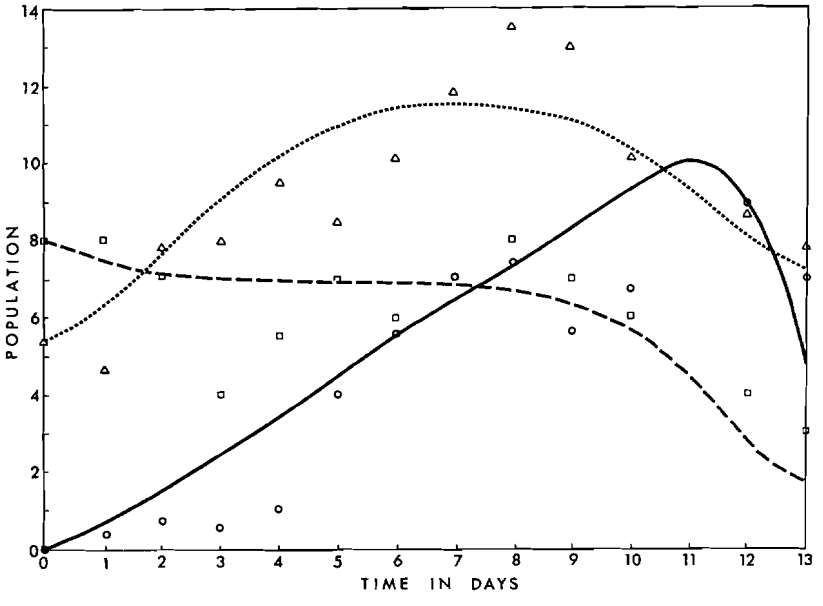

Figure 3. The observed populations and the populations predicted by the final models of equations (11) and (12). o and —— are the observed and predicted total herbivore biomasses, respectively, in μg dry wt./liter times 10^{-1}; □ and - - - are the observed and predicted dissolved nitrogen concentrations in μg at N/liter; and Δ and ------ are the observed and predicted active chlorophyll *a* concentrations in μg/liter.

$$\frac{d\,N_2}{d\,t} = B_{23}N_3 + C_{21}N_2N_1 + C_{23}N_2N_3 + D_{21}N_1^2$$

$$\frac{d\,N_3}{d\,t} = A_3 + B_{31}N_1 + B_{32}N_2 + C_{32}N_3N_2 + C_{33}N_3^2 \qquad (12)$$

The sets of least-squares parameters for equations (11) and (12) are displayed in Table 4, and the populations obtained from equations 11 and 12 simultaneously integrated with these parameter values are shown in Figure 3.

DISCUSSION AND FUTURE WORK

Having derived a suitable mathematical representation of observed population interactions, it is tempting to proceed with a bioligical interpretation of each term in the model. Such analysis must proceed with caution engendered by the following observations. When the matrix $\bar{\bar{X}}^t\bar{\bar{X}}$ is non-singular (as it is in this instance) the vector $\bar{B}^*$ of least-squares parameters is unique. The approach to this vector is, however, (a) non-unique, and (b) at times, mathematically very stiff. As an example of the latter difficulty, iteration of the regression algorithm may proceed until the increment of $\cos^2\theta$ is, say, 10^{-6}. Continuation of the iteration until the $\cos^2\theta$ is incremented by only 10^{-12} may result in substantial changes in certain elements of B. Generally speaking, the noisier the data, the greater such difficulties become.

Looking at equation 11, it can be said with confidence that A_1 is necessary to mimic the data and, as such, has physical and/or biological significance (in this case the loss of copepod biomass to the planktivores and sediments). To say, however, that such loss is 117.5 μg dry wt./ l /day, as indicated by the last regression, is stretching the point, in view of the noise in the given data.

Equation (11) reveals the somewhat surprising independence of the copepods to the phytoplankton standing crop. Of course, this is not to say that the herbivores can do without their food source. It reveals, however, that under the conditions to which the microcosm was subjected, the pertinent observable kinetics occur between copepods and nitrogen. Evidently, phytoplankters are in a condition of minimum response to nutrients and grazing pressure. Caperon (2) suggests that such minimal responses are characteristic of eutrophic ecosystems, and the thread of evidence presented here indicates the possibility that grazers may be of prime importance in assessing the stability of natural eutrophic mesohaline waters to additional nutrient loading.

When strong interactions are present (viz., the N_1 - N_2 coupling above), it is not surprising that all terms including one or the other of the populations will be kept as significant in the final reduced model. This is only an indication that the actual relation is not adequately described by any single term. If the reduced

model is viewed as a Taylor approximation of the actual functionalities, then clues to the actual function are at least provided.

Thus, at this time, only semi-quantitative information is yielded by the modeling exercise. However, the pathway toward a model with quantitative predictability is clear and will direct the thrust of future work at this Laboratory. An evident necessity is a reduction in the noise of the data, or at least programmed redundancy to offset such noise. This will be accomplished by enlarging the microcosm and sampling it more frequently. An extension of the period of measurement and an increased effort at replication will also add to the confidence level of the final model.

Work also is being initiated on the construction of an algorithm to assess the confidence interval of each individual parameter calculated from the data. Also, a sensitivity analysis after Ulanowicz (15) is planned to help elucidate the temporal sequence of interations as they occur in the derived model.

In conclusion, the authors feel that this work constitutes a beginning toward the systematic derivation of empirical models capable of quantitative application to management problems. In addition, the information gathered during the process outlined above should engender a number of hypotheses regarding the mechanism of important interactions (*a priori* models).

ACKNOWLEDGMENTS

The work described herein was supported by the National Science Foundation's Committee on Research Applied to National Needs (RANN), Grant Number GI-29906, to the Chesapeake Research Consortium (CRC). Computer time was provided by the United States Army Corps of Engineers' Contract DACW31-70-C-0077 to the University of Maryland Computer Science Center. The authors wish to thank Ms. Shelley Sulkin and Mr. Rogers Huff for their excellent support in obtaining the data, and Ms. Margaret Roper for typing the manuscript.

REFERENCES

1. Bellman, R., Kagiwada, H., and Kalaba, R.
   1966 Inverse problems in ecology. **Jour. Theoretical Biol.,** 11: 164.
2. Caperon, J., Cattell, S. A., and Krasnick, G.
   1961 Phytoplankton kinetics in a subtropical estuary: eutrophication. **Limnol. Oceanogr.,** 16: 599.
3. Dale, M. B.
   1970 Systems analysis and ecology. **Ecology,** 51(1): 1.
4. Gause, G. F.
   1934 **The Struggle for Existence.** Williams and Wilkins, Baltimore.
5. Heinle, D. R.
   1969 Culture of calanoid copepods in synthetic seawater. **Jour. Fish.**

Res. Bd. Canada, 26: 150.

6. Heinle, D. R.
   1969 Effects of temperature on the population dynamics of estuarine copepods. Doctoral dissertation, University of Maryland, College Park, Maryland.
7. Holm-Hansen, O., Lorenzen, C. J., Holmes, R. W., and Strickland, J. D. H.
   1965 Fluorometric determination of chlorophyll. **Jour. Conseil, Perm. Intern. Explor. Mer.,** 30: 3.
8. Lorenzen, C. J.
   1967 Determination of chlorophyll and pheo-pigments: spectro-photometric equations. **Limnol. Oceanogr.,** 12: 343.
9. Mobley, C. D.
   1972 Statistical hypothesis testing as applied to the inverse problem in ecosystems analysis. N.R.I. reference no. 72-84, Natural Resources Institute, Solomons, Maryland.
10. Mobley, C. D.
   1973 A Systematic Approach to Ecosystems Analysis. (In preparation.)
11. Park, T.
   1948 Experimental studies of interspecies competition. **Ecol. Monogr.,** 18: 265.
12. Patten, B. C.
   1968 Mathematical models of plankton production. **Int. Revue ges. Hydrobiol.,** 53: 357.
13. Ross, D. T.
   1967 The AED approach to generalized computer-aided design. **Proc. Ass. Comput. Mach. National Meeting,** 367.
14. Strickland, J. D. H., and Parsons, T. R.
   1965 A manual of sea-water analysis. **Fish. Res. Bd. Canada Bull. No. 125.** Ottawa.
15. Ulanowicz, R. E.
   1971 Transport through hemoglobin solutions: a sensitivity analysis. **Jour. Theoretical Biol.,** 32: 185.
16. Watt, K. E. F.
   1962 Use of mathematics in population ecology. **Ann. Rev. Ent.,** 7: 243.
17. Yentsch, C. S.
   1965 In **Primary Productivity in Aquatic Environments.** (ed. Goldman, C. R.) Mem. 1st Ital. Idrobiol., 18. Univ. California Press, Berkeley.
18. Yentsch, C. S., and Menzel, D. W.
   1963 A method for the determination of phytoplankton chlorophyll and phaeophytin by fluorescence. **Deep-Sea Res.,** 10: 221.