Introduction to Biological Models

Lecturer: Glenn Flierl

May 11, 2011

1 Introduction

The ocean is inhabited by innumerable individuals of many different genera and species, each having its own developmental/physiological state and each being immersed in its own environment. The organisms move, both because of water flow and because of their own swimming or buoyancy, and interact with their environment by gathering resources which they need and by excreting waste products. The assimilated material can be used for maintenance, growth, or reproduction. Finally, the organisms can die either from natural causes or because of attacks by another organism.

Furthermore, the processes just described must generally be regarded as stochastic. For example, the probability of a predator capturing a prey item will depend on multiple factors, each with its own probability: finding a prey item in range, the choice to attack, success in the attack, competition against others. Such a description suggests an "agent-based" or "individual-based" model (IBM), with each agent carrying information about its position, its species, its physiological state, etc. Individuals can grow, reproduce, and die. Certainly, we can build small versions of such models, but the number of individuals is necessarily limited (compared to nearly 20,000 copepods per cubic meter or to phytoplankton densities on the order of 10^8 per cubic meter). However, such experiments may indeed give insight into the way in which the local, stochastic interactions translate into terms representing, for example, grazing rates in terms of average densities.

Or we could take the alternative view of attempting to predict the probability distribution for biomass in a continuous space, in this case using something like weight and species as our variables. The latter is, of course, discrete, yet different organisms can be genetically or, more importantly, functionally quite close to others. If we choose a species ordering such that the maximum growth rate varies smoothly, we may expect that other terms entering the dynamics such as the losses by predation will also fall on a fairly smooth curve. Certainly on any diagram such as figure (1) the gaps will be so small as to be negligible, and viewing the ordinate as a continuous variable is not unreasonable. We can then consider the ways that the processes described above alter the biomass distribution in this space.

As in the individual-based model, the number of variables we would have to consider is still unmanageably large. Furthermore, for each (w, sp), we need to specify the sources, the sink, and the transfer rates, including possible nonlinear dependence on the local biomass $b(w, s|\mathbf{x}, t)$ and the density at the source/sink $b(w, s|\mathbf{x}, t)$.

Any attempt to construct such a model or the IBM model, will inevitably point out how little we know about most of the species inhabiting the ocean. We retreat to dealing
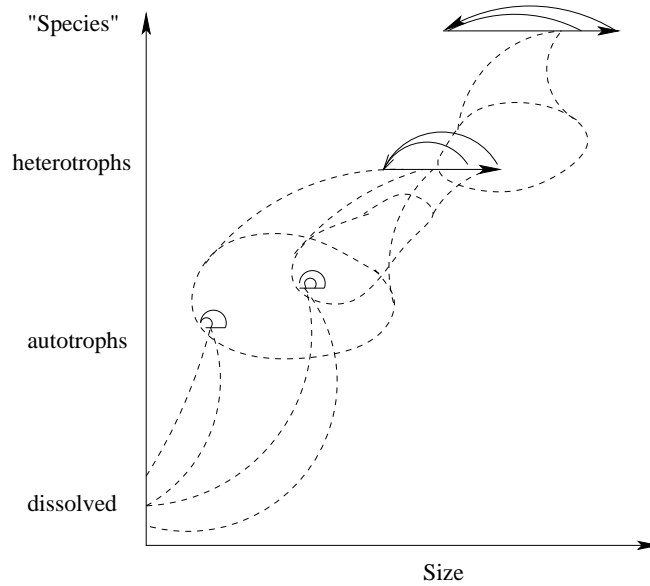
15

Figure 1: The range of sizes for various species

with variables such as P which can be viewed as the integral over some range $(s1- > s2)$ of species and over weight. We then presume that transfers into and out of the resulting black-boxes can be represented as functions just of the integrated values and attempt to parameterize those.

2 Basic biological models

Exponential growth

Suppose each individual (or pair) reproduces in a time order τ, then the time evolution of the population would be described as

$$b(t + \delta t) = [1 + g\delta t]b(t), \quad g = ln2/\tau \tag{1}$$

$$\frac{\partial b}{\partial t} = gb \ \Rightarrow \ b(t) = b_\circ exp(gt). \tag{2}$$

where b is the population and g is its growth rate. This model predicts an unrealistic exponential growth of the population (see figure (2)). Thus, some limiting factors should exist.

Resource limitation

For example, one could write a model with the growth rate depending on a resource R which itself evolves with time:

$$\frac{\partial b}{\partial t} = (gR - d)b \tag{3}$$

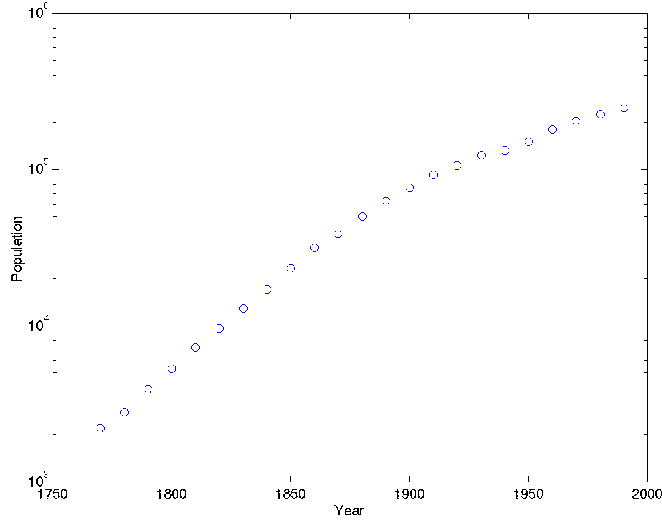$$\frac{\partial R}{\partial t} = -gRb - \lambda(R - R_\circ) \tag{4}$$

Figure 2: USA population as a function of time showing exponential growth at initial stage, but a decay of a growth rate at later stages

Here, the population evolves with a growth rate gR and a death rate d. The resource decreases due to consumption by population and is restored to $R_\circ$ with a time scale $1/\lambda$. This model allows for a steady state solution with $\bar{R} = d/g, \bar{b} = -\lambda(d/g - R_\circ)/d$. However, the growth rate is unconstrained with increase of the resource; a more realistic growth model would have to limit it at $R \to \infty$. One possibility is to use the form $g\frac{R}{R+k}$.

Predator-pray

Assume that zooplankton (Z) grows by consuming phytoplankton (P) and phytoplankton grows due to source of nutrients (N) assumed to be fixed; both have natural death rates d_P, d_Z

$$\frac{\partial P}{\partial t} = gPZ - \mu NP - d_P P \tag{5}$$

$$\frac{\partial Z}{\partial t} = \alpha gPZ - d_Z Z \tag{6}$$

This model (the Lotka-Volterra equations) gives a solution with cycles that have a conserved quantity $H = \alpha gP - d_Z lnP + \mu NZ + d_P lnZ$. The quadratic NPZ model assumes $N = N_T - P - Z$, where N_T is a total amount of nutrients. This model has three steady states if $N_T > (d_Z/\alpha g) + (d_P/\mu)$, with a P, Z non-zero point being a stable attractor.

NPZD

The NPZ model assumes all dead organisms or excreted material is immediately remineralized to usable nutrient. In contrast, the NPZD model assumes that dead organisms and unassimilated phytoplankton would contribute to a detrital pool that eventually be-
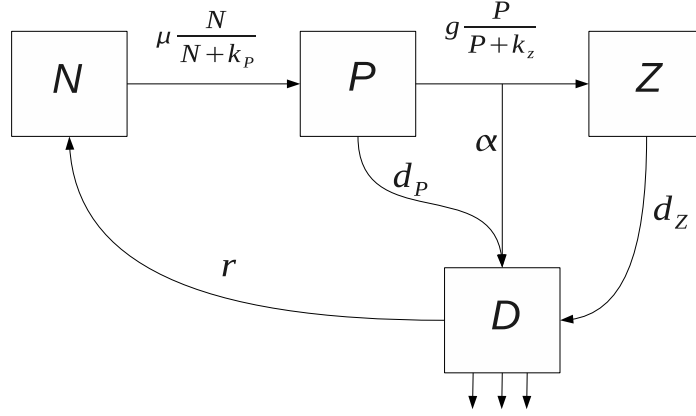
17

Figure 3: Schematic representation of NPZD model showing the fluxes of biomass and the parameters controling them

comes a source of nutrients. A schematic representation of this model is shown in figure (3), and a typical version of the corresponding equations (including sinking of detritus) would be:

$$\frac{D}{Dt}P = \mu\frac{N}{N+k_P}P - g\frac{P}{P+k_Z}Z - d_P P + \nabla\cdot(\kappa\nabla P) \tag{7}$$

$$\frac{D}{Dt}Z = \alpha g\frac{P}{P+k_Z}Z - d_Z Z + \nabla\cdot(\kappa\nabla Z) \tag{8}$$

$$\frac{D}{Dt}D + \frac{\partial}{\partial z}w_s D = (1-\alpha)g\frac{P}{P+k_Z}Z + d_P P + d_Z Z - rD + \nabla\cdot(\kappa\nabla D) \tag{9}$$

$$\frac{D}{Dt}N = rD - \mu\frac{N}{N+k_P}P + \nabla\cdot(\kappa\nabla N) \tag{10}$$

$$\tag{11}$$

where $\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla$ denotes the material derivative of the advected scalar.

In general case the system equations have the form

$$\frac{D}{Dt}b_i + \nabla\cdot(\mathbf{u}_i b_i) = B_i(\mathbf{x},t|\mathbf{b}) + \nabla\cdot((\kappa+\kappa_i)\nabla b_i) \tag{12}$$

Here, $\mathbf{u}_i$ and κ_i represent the motion relative to the fluid. The function $B_i(\mathbf{x},t|\mathbf{b})$ represents interactions between different parts of the system. Some times a condition $\sum B_i(\mathbf{x},t|\mathbf{b}) = 0$ is used, which assumes that the total biomass is conserved; this reduces the order of the system.

Such biological models require a large number of parameters. For example, in the NPZD model with 4 variables, the parameters are: $\mu, k_P, k_Z, g, d_P, d_Z, \alpha, r$ in addition to the

biological movement term w_s and physical parameters such as κ; together, they determine the fundamental behavior of the system. Thus, in addition to the mathematical question of the behavior of the equations for different parameter values, modellers need to decide on numbers which seem to give "realistic" solutions. Since the variables are not generally directly observed, comparisons with data tend to be pretty qualitative in an case.

Controlled (i.e., laboratory) experiments can be a guid for parameter choicese; however the behavior of the organisms might be affected by the laboratory constraints so that things like predation rates may be quite different. Another way is to observe the biosystem in the nature, but in reality the system is much more complicated than the model, and this can also introduce large uncertainties.

3 Stochastic Dynamics

An alternative method for modeling organism's population dynamics is through the use of probability density functions for the population size. The probability density functions will be written as $\mathcal{P}(n,t)$ which represents the probability of having a population of size n at time t. In case of multiple interacting populations, the evolution of the population probability is based on the equation

$$\mathcal{P}(\mathbf{z}, t + \delta t) = \sum \mathcal{P}_T(\delta \mathbf{z} | \mathbf{z} - \delta \mathbf{z}) \mathcal{P}(\mathbf{z} - \delta \mathbf{z}, t) \tag{13}$$

where $\mathbf{z}$ is a vector representing the population of the various species considered in the model and $\mathcal{P}_T$ is the transition probability. This equation is essentially an application of the law of total probability. The probability of a specific population at a future time step is the sum of all possible previous populations times the probability of transitioning to the specified population during the given time step. For one of the species, the transition probability can be modeled as

$$\mathcal{P}_T = \begin{cases} nd_n\delta t & \delta n = -1 \\ 1 - ng_n\delta t - nd_n\delta t & \delta n = 0 \\ ng_n\delta t & \delta n = 1 \end{cases}$$

where g_n and d_n are the growth and death rates *per capita* depending on the population size n.

To make this more tractable, let us consider a case where only one population is modeled. For simplification we will change the notation from $\mathcal{P}(n,t)$ to be $\mathcal{P}_n(t)$. Equation 13 becomes

$$\begin{aligned} \mathcal{P}_n(t + \delta t) &= \mathcal{P}_{n-1}(t)(n-1)g_{n-1}\delta t + \mathcal{P}_n(1 - ng_n\delta t - nd_n\delta t) + \mathcal{P}_{n+1}(n+1)d_{n+1}\delta t \\ \frac{\mathcal{P}_n(t+\delta t) - \mathcal{P}_n}{\delta t} &= \mathcal{P}_{n-1}(t)(n-1)g_{n-1} - \mathcal{P}_n n(g_n + d_n) + \mathcal{P}_{n+1}(n+1)d_{n+1} \\ \frac{d}{dt}\mathcal{P}_n &= \mathcal{P}_{n-1}(n-1)g_{n-1} - \mathcal{P}_n n1(g_n + d_n) + \mathcal{P}_{n+1}(n+1)d_{n+1} \end{aligned} \tag{14}$$

From this the first few equations can be calculated to give a foundation for insight

$$\begin{aligned} \frac{d}{dt}\mathcal{P}_0 &= \mathcal{P}_1 d_1 \\ \frac{d}{dt}\mathcal{P}_1 &= -\mathcal{P}_1(g_1 + d_1) + 2\mathcal{P}_2 d_2 \\ \frac{d}{dt}\mathcal{P}_2 &= \mathcal{P}_1 g_1 - 2\mathcal{P}_2(g_2 + d_2) + 3\mathcal{P}_3 d_3 \\ \frac{d}{dt}\mathcal{P}_3 &= 2\mathcal{P}_2 g_2 - 3\mathcal{P}_3(g_3 + d_3) + 4\mathcal{P}_4 d_4 \end{aligned} \tag{15}$$

Now we will look at the expectation of certain properties and will define the functional form

$$< f_m >= \sum_n f_m \mathcal{P}_n v$$

Clearly the sum of the probabilities at a given time is constant so

$$\frac{d}{dt} < 1 >= 0.$$

A more relevant calculation is the understanding of the time rate of change of the expected population change which can be shown to be

$$\frac{d}{dt} < n >=< n(g_n - d_n) > .$$

If however, the death and growth rates are independent of the population size than the constant can be pulled outside of the expectation calculation, so

$$\frac{d}{dt} < n >= (g_n - d_n) < n >$$

which has a similar form to the deterministic population model. Additionally it can be shown that given

$$\frac{d}{dt} < n^2 >= 2 < n^2(g_n - d_n) > + < n(g_n + d_n) >$$

and $\sigma^2 =< n^2 > - < n >^2$, it is possible to calculate the long term trend of the ratio between the standard deviation of the population (σ) and the expected population size with growth and death rates independent of population size

$$\lim_{t \to \infty} \frac{\sigma}{< n >} = \sqrt{\frac{\sigma_0^2 + < n >_0 \frac{g+d}{g-d}}{< n >_0^2}}$$

where σ_0 and $< n >_0$ are the initial standard deviations and the expected population size. If we compare a set of realizations, we would find the range of variation increases with time just as the mean size is increasing.

For populations which can double in a day, the exponential model cannot apply for long. We can examine the logistic model with $g_n = g_0(1 - n/n_0)$, in which case the distribution can reach a nearly state. It cannot be completely steady, since a constant value of $\mathcal{P}_1$ implies the extinction probability, $\mathcal{P}_0$, is continuing to grow.

3.1 Fokker-Planck Expansion

when the number of individuals is very large, $\mathcal{P}(n, t)$ can be treated as a smooth function of n, regarded as a real number. If we start again with the Equation 13, the following Taylor expansion can be performed to approximate the dependence on the vector of population sizes

$$\begin{aligned}
\mathcal{P}(\mathbf{z}, t) + \delta t \frac{\partial}{\partial t} \mathcal{P}(\mathbf{z}, t) &= \sum \mathcal{P}_T(\delta \mathbf{z} | \mathbf{z}, t) \mathcal{P}(\mathbf{z}, t) - \frac{\partial}{\partial z_i} \sum \delta z_i \mathcal{P}_T(\delta \mathbf{z} | \mathbf{z}, t) \mathcal{P}(\mathbf{z}, t) \\
&+ \frac{1}{2} \frac{\partial^2}{\partial z_i \partial z_j} \sum \delta z_i \delta z_j \mathcal{P}_T(\delta \mathbf{z} | \mathbf{z}, t) \mathcal{P}(\mathbf{z}, t)
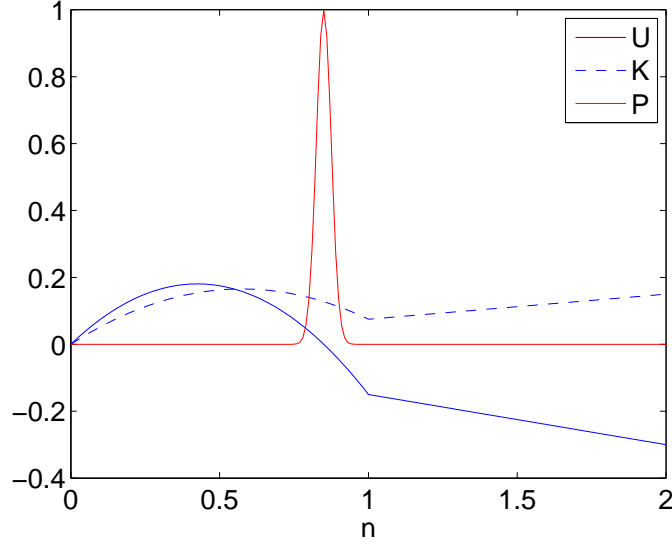\end{aligned}$$

20

Figure 4: Plot of the U and K for a linearly deacreasing growth rate and a constant death rate and the resulting population density distribution.

where $\mathcal{P}(\mathbf{z}, t)$ is still the population probability function dependent on the coordinate $\mathbf{z}$ and time, t, and all δ's represent a differential step in the provided coordinate. This Taylor expansion can be simplified to

$$\frac{\partial}{\partial t}\mathcal{P}(\mathbf{z}, t) = -\frac{\partial}{\partial z_i}U_i\mathcal{P} + \frac{\partial^2}{\partial z_i \partial z_j}K_{ij}\mathcal{P} \tag{16}$$

where the substitutions

$$U_i(\mathbf{z}, t) = \sum \frac{\delta z_i}{\delta t}\mathcal{P}_T\mathbf{z}|\mathbf{z}, t), \qquad K_{ij}(\mathbf{z}, t) = \sum \frac{\delta z_i \delta z_j}{2\delta t}\mathcal{P}_T(\delta\mathbf{z}|\mathbf{z}, t)$$

have been made. For the case like 14 considering birth and death rates in one species, $U = n(g_n - d_n)$ and $K = \frac{1}{2}n(g_n + d_n)$. An example growth rate that can be used is $g_n = g_0(1 - n/n_0)$ which corresponds to a linearly decreasing growth rate for populations under a maximum population, n_0. With a death rate at a fixed constant the function of U and K can be determined and are presented in Figure 4. The resulting population probability distribution function is a narrow Gaussian centered at $g_n = d_n$.