## Introduction to Biological models

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 (presumably depending on the level of satiation of the predator and the perceived nature of the prey)
- success in the attack (a function of the condition of the predator and of the prey)
- competition against others

Such a description suggests an "agent-based" or IBM with each agent carrying information about its position, its species, its physiological state, etc. Organisms 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 observed during Globec [Incze, 19xx] , 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.

Once again, 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 this, 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 $(w, s p)$, we need to specify the sources, the sink, and the transfer rates, including possible nonlinear dependence on the local biomass $b(w, s \mid \mathbf{x}, t)$ and the density at the source/ $\operatorname{sink} b\left(w^{\prime}, s^{\prime} \mid \mathbf{x}, t\right)$.

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 with variables such as $P$ which can be viewed as the integral over some range $(s 1 \rightarrow s 2)$ 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.

## Basics

## Exponential growth

Suppose each individual (or pair) reproduces in a time order $\tau$

$$
\begin{gathered}
b(t+\delta t)=[1+g \delta t] b(t) \quad, \quad g=\ln 2 / \tau \\
\frac{\partial}{\partial t} b=g b \quad \Rightarrow \quad b(t)=b_{0} \exp (g t)
\end{gathered}
$$

As Malthus pointed out this implies the need for limitations.
Stochastic version
What are the effects of stochasticity on the dynamics? We shall look at a few problems to decide whether the expected values satisfy differential equations like the ones we've been using.

## Birth-death

Consider a vector $\mathcal{P}_{n}(t)$ giving the probability that there are $n$ organisms at time $t$; we would like to predict $\mathcal{P}_{n}(t+d t)$ or, in the limit as $d t \rightarrow 0, \frac{\partial}{\partial t} \mathcal{P}$. Assuming that a given organism can double, can die, or can survive during time $d t$ leads to
$\mathcal{P}_{n}(t+d t)=\mathcal{P}_{n-1}(n-1) g_{n-1} \delta t+\mathcal{P}_{n+1}(n+1) d_{n+1} \delta t+\mathcal{P}_{n}(t)\left(1-n g_{n} \delta t-n d_{n} \delta t\right)+\mathcal{O}\left(\delta t^{2}\right)$
I.e., we can move from having $n-1$ organisms to $n$ if one of the $n-1$ individuals divides (the probability of two dividing in time $\delta t$ is another factor of $\delta t$ smaller and will be neglected.) From this, we derive the "master equation":

$$
\begin{equation*}
\frac{\partial}{\partial t} \mathcal{P}_{n}=(n-1) g_{n-1} \mathcal{P}_{n-1}+(n+1) d_{n+1} \mathcal{P}_{n+1}-n\left(g_{n}+d_{n}\right) \mathcal{P}_{n} \tag{1}
\end{equation*}
$$

where we have allowed for the possibility that the per capita birth and death rates may be density dependent; if not, they are simply given by $g$ and $d$, respectively.

The first few equations help us gain insight into the character of the solutions

$$
\begin{aligned}
& \frac{\partial}{\partial t} \mathcal{P}_{0}(t)=d_{1} \mathcal{P}_{1}(t) \\
& \frac{\partial}{\partial t} \mathcal{P}_{1}(t)=-\left(d_{1}+g_{1}\right) \mathcal{P}_{1}(t)+2 d_{2} \mathcal{P}_{2}(t) \\
& \frac{\partial}{\partial t} \mathcal{P}_{2}(t)=-2\left(d_{2}+g_{2}\right) \mathcal{P}_{2}(t)+g_{1} \mathcal{P}_{1}(t)+3 d_{3} \mathcal{P}_{3}(t) \\
& \frac{\partial}{\partial t} \mathcal{P}_{3}(t)=-3\left(d_{3}+g_{3}\right) \mathcal{P}_{3}(t)+2 g_{2} \mathcal{P}_{2}(t)+4 d_{4} \mathcal{P}_{4}(t)
\end{aligned}
$$

The probability of extinction, $\mathcal{P}_{0}$, increases monotonically; whether it is bounded by some value less than one will depend on how rapidly $\mathcal{P}_{1}$ decreases with time. Thus, we cannot
expect to find a truly steady solution to the master equation. ${ }^{\dagger}$ Populations either grow indefinitely or become extinct; however, as we shall see, the extinction time for large populations can be so long that we find a quasi-steady solution. We shall return to this issue when we discuss density-dependent dynamics.

In general, we are less concerned with the detailed probability distribution than with its moments

$$
\left\langle n^{m}\right\rangle=\sum_{n} n^{m} \mathcal{P}_{n}
$$

in particular, the mean $\langle n\rangle$ and the variance $\sigma^{2}=\left\langle n^{2}\right\rangle-\langle n\rangle^{2}$; these are quantities which may be measurable and which can be related to the deterministic density $b$.

From the form of (1), we can show that

$$
\frac{\partial}{\partial t} \sum \mathcal{P}_{n}=-\sum n g_{n} \mathcal{P}_{n}-\sum n d_{n} \mathcal{P}_{n}+\sum(n-1) g_{n-1} \mathcal{P}_{n-1}+\sum(n+1) d_{n+1} \mathcal{P}_{n+1}
$$

so that

$$
\frac{\partial}{\partial t} \sum_{n} \mathcal{P}_{n}(t)=0
$$

implying the $\left\langle n^{0}\right\rangle$ is preserved at its initial value 1 - as required for a probability distribution. The first moment equation is

$$
\frac{\partial}{\partial t}\langle n\rangle=\sum n\left(g_{n}-d_{n}\right) \mathcal{P}_{n}
$$

When the birth and death rates are constant, this reduces to

$$
\frac{\partial}{\partial t}\langle n\rangle=(g-d)\langle n\rangle
$$

and the mean satisfies the same exponential rule $\langle n\rangle=\langle n\rangle_{0} \exp ([g-d] t)$ as the deterministic case.

The second moment (with constant rates) satisfies

$$
\frac{\partial}{\partial t}\left\langle n^{2}\right\rangle=2(g-d)\left\langle n^{2}\right\rangle+(g+d)\langle n\rangle
$$

substituting $\left\langle n^{2}\right\rangle=\sigma^{2}+\langle n\rangle^{2}$ and using the dynamical equation for $\langle n\rangle$ results in the same equation for the variance. The solution is

$$
\sigma^{2}(t)=\left(\sigma^{2}(0)+\langle n\rangle_{0} \frac{g+d}{g-d}\right) e^{2(g-d) t}-\langle n\rangle_{0} \frac{g+d}{g-d} e^{(g-d) t}
$$

$\dagger$ Unless a single individual is immortal $-d_{1}=0!$

For large times and a growing population, the first term dominates, and the ratio of the standard deviation to the mean becomes constant

$$
\lim _{t \rightarrow \infty}\left(\frac{\sigma}{\langle n\rangle}\right)=\sqrt{\frac{\sigma^{2}(0)+\langle n\rangle_{0} \frac{g+d}{g-d}}{\langle n\rangle_{0}^{2}}}
$$

implying that the error bars on a plot of $\log \langle n\rangle$ versus $t$ would be uniform.
The extinction probability (for cases with $g \geq d$ so that the mean is not decaying) is

$$
\lim _{t \rightarrow \infty} p(0, t)=\sum_{n} \mathcal{P}_{n}(0)\left(\frac{d}{g}\right)^{n}
$$

If we start with exactly $N$ organisms, the extinction probability will be $(d / g)^{N}$. The example shows histograms $\mathrm{g}=0.4, \mathrm{~d}=0.38$, the means, and $\log$ mean .

## Resource limitation

## Competition/ logistic

One form of limitation arises from direct competition for scarce resources, leading to a decrease in the growth rate $g$ as $b$ increases:

- The nutrient available per organism is inversely proportional to the population density - Avail $=R / b$.
- The organisms require a minimum amount of nutrient to offset respiration - Avail > Resp. This sets the carrying capacity $b_{c}=R /$ Resp. The growth/ reproduction rate increases as the available nutrient increases.
- The growth rate has a maximum value.

These suggest a growth rate vs. available nutrient like

$$
\begin{aligned}
\text { growth rate } & =\text { Max rate } \frac{\text { Avail }- \text { Resp }}{\text { Avail }} \\
& =g_{m} \frac{1 / b-1 / b_{c}}{1 / b} \\
& = \begin{cases}g_{m}\left(1-\frac{b}{b_{c}}\right) & b<b_{c} \\
0 & b>b_{c}\end{cases}
\end{aligned}
$$


we can rewrite in terms of $1 / b$ and solve analytically

$$
\begin{equation*}
b=\frac{b_{c}}{\frac{g_{m}}{g_{m}-d}\left[1+e^{-\left(g_{m}-d\right) t}\right]+\frac{b_{c}}{b(0)} e^{-\left(g_{m}-d\right) t}} \tag{2}
\end{equation*}
$$

The solution has a sigmoid shape. For $b(0) \ll b_{c}$ and short times, we have

$$
b=b(0) \exp \left(\left[g_{m}-d\right] t\right)
$$

For long times $b$ limits to $b_{c} g_{m} /\left(g_{m}-d\right)$, the carrying capacity, whether $b$ starts above or below $b_{c}$. The solutions show the sigmoid shape; this also applies to the case with a cutoff when $b>b_{c}$. Seasonal cycles in the forcing lead to periods of exponential growth, limitation, and decay.

## Resource depletion

For each organism, take the resource uptake as proportional to the concentration of resource $(N)$. The population biomass is proportional to the number of individuals; therefore the rate of change of biomass is proportional to the product $N P$.

$$
\begin{aligned}
\frac{\partial}{\partial t} P & =\left[\mu N-d_{P}\right] P \\
\frac{\partial}{\partial t} N & =-\mu N P+\lambda\left(N_{T}-N\right)
\end{aligned}
$$

The example with $m u=1, \mathrm{dp}=0.2, \mathrm{l} a=0.1, \mathrm{nt}=5$ shows the sigmoid shape, but altered by the variations in $N_{T}$.
$P$ no longer grows without bounds; instead the system limits to $\bar{N}=d_{P} / \mu, \bar{P}=$ $\lambda\left(N_{T}-\bar{N}\right) / \mu \bar{N}$. Special cases in which the equation for $P$ resembles the logistic equation are:

- closed - dead PP are remineralized immediately

$$
\frac{\partial}{\partial t} N=-\mu N P+d_{p} P
$$

which implies $P+N=N_{T}$ a constant. Then

$$
\frac{\partial}{\partial t} P=\left[\mu N_{T}-d_{p}-\mu P\right] P
$$

- chemostat $d_{P}=\lambda$ (washout of nutrients and PP)

$$
\frac{\partial}{\partial t}(P+N)=-\lambda(P+N)+\lambda N_{T} \quad \Rightarrow \quad P+N \rightarrow N_{T}
$$

If we start with $P=0$ and $N=N_{T}$ and seed it with a small population $P_{0}$, then

$$
P+N=N_{T}+P_{0} e^{-\lambda t}
$$

and

$$
\frac{\partial}{\partial t} P=\left[\mu N_{T}+\mu P_{0} e^{-\lambda t}-\mu P-d_{p}\right] P
$$

If $P_{0} \ll N_{T}, P$ will satisfy the logistic equation.

## Stochastic version

If the birth/death processes are stochastic, we can again write

$$
\frac{\partial}{\partial t} \mathbf{p}=\mathbf{M} \mathbf{p}
$$

with the matrix $\mathbf{M}$ being

$$
\mathbf{M}=\left(\begin{array}{cccccc}
0 & d_{1} & 0 & 0 & 0 & \cdots \\
0 & -d_{1}-g_{1} & d_{2} & 0 & 0 & \\
0 & g_{1} & -d_{2}-g_{2} & d_{3} & 0 & \\
0 & 0 & g_{2} & -d_{3}-g_{3} & d_{4} & \\
0 & 0 & 0 & g_{3} & -d_{4}-g_{4} &
\end{array}\right)
$$

Previously $g_{n}=n g$ and $d_{n}=n d$; now, we assume a logistic form (based on the same food availability arguments) $g_{n}=g\left(n-n^{2} / n_{0}\right)$ [or zero for $n>n_{0}$ ], $d_{n}=n d$. We still have

$$
\sum_{n} M_{n m}=0 \Rightarrow \frac{\partial}{\partial t} \sum_{n} \mathcal{P}_{n}(t)=0
$$

but now

$$
\sum_{n} n M_{n m}=(g-d) m-\frac{g}{n_{0}} m^{2}
$$

so that

$$
\frac{\partial}{\partial t}\langle n\rangle=(g-d)\langle n\rangle-\frac{g}{n_{0}}\left\langle n^{2}\right\rangle
$$

If $\sigma \ll\langle n\rangle,\left\langle n^{2}\right\rangle \simeq\langle n\rangle^{2}$ and we recover the logistic form; in general, however, they will not agree, since the variance $\left\langle n^{2}\right\rangle-\langle n\rangle^{2}$ is not zero. If we assume we do know $\sigma^{2}$, we can find the mean; the answers agree well for $n_{0}=400, g=0.2, d=0.15$ case (if $\beta=n_{0}(g-d) / g$ then the first estimate would be $\langle n\rangle=\beta$, the second $\langle n\rangle=$ $\beta-\sigma^{2} / \beta$, and the third $0.5 \beta+0.5 \sqrt{\beta^{2}-4 \sigma^{2}}$ giving $100,96.258$, and 96.106 compared to the numerically calculated value 96.104 ). They remain close for $n_{0}=200$ where extinction is happening at a still-noticeable rate (50, 45.782, 45.349 compared to 45.384 ). The numerics is $n 0=200, \mathrm{~g}=0.2, \mathrm{~d}=0.15$, means extinction prob $n 0=400, \mathrm{~g}=0.2, \mathrm{~d}=0.15$ means extinction prob illustrate the effects of varying the carrying capacity.

However, we do not know apriori the variance. The equation for the second moment is

$$
\frac{\partial}{\partial t}\left\langle n^{2}\right\rangle=2(g-d)\left\langle n^{2}\right\rangle-2 \frac{g}{n_{0}}\left\langle n^{3}\right\rangle+(g+d)\langle n\rangle-\frac{g}{n_{0}}\left\langle n^{2}\right\rangle
$$

We now run into the first version of a closure problem: to calculate $\langle n\rangle$, we need $\left\langle n^{2}\right\rangle$; to calculate $\left\langle n^{2}\right\rangle$ we need $\left\langle n^{3}\right\rangle$; etc. If we assume the first and second moment equations reach a steady state and write $n=\langle n\rangle+n^{\prime}$, we have

$$
\left[2(g-d)-4 \frac{g}{n_{0}}\langle n\rangle\right] \sigma^{2}=-2 d\langle n\rangle+2 \frac{g}{n_{0}}\left\langle n^{\prime 3}\right\rangle
$$

If we assume the final state has zero skewness (is symmetric), $\left\langle n^{\prime 3}\right\rangle=0$, we can solve the resulting equations for the mean and variance. The solutions give $\langle n\rangle=46.514$ and 96.794 for the two cases above - certainly in the right direction, but quantitatively off. Shapes mean relative error extinction ratg

Given that the use of the logistic equation for the mean seems to have very small errors when $n_{0}$ (or $\langle n\rangle$ ) is large, why should we be concerned any further? The problem lies in the presumption, inherent in this model, that the biota and their resources are all well-mixed; otherwise, extinction can occur in small subregions (which can, by themselves, support fewer organisms) and replenishment by mixing may overcome such losses.

We follow a procedure like that of Hernández-Garcia and López: For each organsim, we calculate an estimate of neighbor density $R$ (using xx.xx) and then adjust the birth probability to be $g\left(1-R / \rho_{0}\right) \delta t$; the probability of dying remains $d \delta t$. Births occur at the location of the parent; in between time steps the organisms move with random flights characterized by the parameters $\kappa$ and $r$. We show the densities compared to the logistic value $\rho_{0}(g-d) / g$. For large diffusivities the competition is essentially global; for small densities, it is local, and we find small clumps of organisms which die out while new
ones form as individuals move out of the competitive range and begin to reproduce more effectively. forming a new clump. ${ }^{\dagger}$

## Functional forms

These are often expressed in terms of Holling's types I-IV. QNP example Saturating example of an NP model with

$$
\frac{\partial}{\partial t} P=\mu \frac{N}{N+N_{1 / 2}} P-d_{p} P \quad, \quad N=1-P
$$

shows how the population saturates at high light. The equilibrium values are

$$
N=N_{1 / 2} \frac{d_{p}}{\mu-d_{p}} \quad, \quad P=1-N_{1 / 2} \frac{d_{p}}{\mu-d_{p}}
$$

This Michaelis-Menton or Monod form arises from two-step processes (cf. Follows and Williams)

Enzyme kinetics:

$$
E+N \underset{k_{-1}}{\stackrel{k_{1}}{\rightleftarrows}} C \xrightarrow{k_{2}} E+R
$$

Prey capture and assimilation:

$$
P \xrightarrow{k_{1}} g u t \xrightarrow{k_{2}} Z \text { biomass }
$$

E.g., we have a system in which the middle step has "active" states which can take up resources and "occupied" states which are handling the resource and turning it into biomass (c.f. Follows and Williams). The rate at which the resource is taken up depends on the product of the resource density $R$ and the density of active states $A$. When uptake occurs, the state becomes occupied in handling it; that takes a time of order $1 / \lambda$ at which point the resource has been converted to biomass and the state flips back to active.

$$
\begin{aligned}
\frac{\partial}{\partial t} A & =-m A R+\lambda O \\
\frac{\partial}{\partial t} O & =m A R-\lambda O
\end{aligned}
$$

These imply $A+O=T$ a constant so that

$$
\frac{\partial}{\partial t} O=m R T-(m R+\lambda) O
$$

If $R$ is varying slowly, this equation shows that $O$ will be in local equilibrium

$$
\lambda O=\lambda T \frac{R}{R+\lambda}
$$

But this is just the growth term in the biomass; thus we recover

$$
\frac{\partial}{\partial t} P=\mu \frac{N}{N+N_{1 / 2}} P-\ldots
$$

† Unlike Hernández-Garcia and López (20xx), but consistent with the results of Birch and Young (20xx), we do not find regular patterns for the parameters used.

Foraging: Suppose an individual can take food from a volume $\nu$ in a time $\tau$. Assume that the resources are randomly distributed at each time with an average number density $R / m_{R}$. If $R$ is distributed randomly, the probability of feeding is

$$
1-\mathcal{P}(0)=1-\exp \left(-\nu R / m_{R}\right)
$$

If a fraction $a$ of the food is assimilated, the individual incorporates mass at a rate

$$
\begin{gathered}
\frac{a m_{R}}{\tau}\left[1-\exp \left(-\nu R / m_{R}\right)\right] \\
\frac{D}{D t} b=a \frac{m_{R}}{\tau m_{b}} b\left[1-\exp \left(-\frac{\nu}{m_{R}} R\right)\right]
\end{gathered}
$$

which has a typical saturating shape.

## Predator-prey

Mortality - mostly getting eaten (viruses though). If the density of prey or the foraging volume is small, the growth rate for the predator is just proportional to prey density.

$$
\begin{aligned}
\frac{\partial}{\partial t} P & =\left[\mu N-g Z-d_{p}\right] P \\
\frac{\partial}{\partial t} Z & =\left[a g P-d_{Z}\right] Z
\end{aligned}
$$

Lotka-Volterra fixes $N$; the system cycles with

$$
H=a g P-d_{Z} \ln P+g Z-\left(\mu N-d_{p}\right) \ln Z
$$

conserved. Implies steady point is formally stable. LV trajectories vs. QNPZ trajectories show the neutral vs. stable behavior.

The Quadratic NPZ model replaces $N$ by $N_{T}-P-Z$. This has three steady states if $N_{T}>\left(d_{Z} / a g\right)+\left(d_{P} / \mu\right)$, with the $P, Z$ non-zero point being a stable attractor.
NPZD

$$
\begin{aligned}
\frac{D}{D t} P & =\mu \frac{N}{N+N_{1 / 2}} P-g Z \frac{P}{P+P_{1 / 2}}-d_{P} P+\nabla \kappa \nabla P \\
\frac{D}{D t} Z & =a g \frac{P}{P+P_{1 / 2}} Z-d_{Z} Z+\nabla \kappa \nabla Z \\
\frac{D}{D t} D+\frac{\partial}{\partial z}\left(w_{s} D\right) & =(1-a) g \frac{P}{P+P_{1 / 2}} Z+d_{P} P+d_{Z} Z \\
& -\lambda D+\nabla \kappa \nabla D \\
\frac{D}{D t} N & =\lambda D-\mu \frac{N}{N+N_{1 / 2}} P+\nabla \kappa \nabla N
\end{aligned}
$$

- Parameters: $\mu, N_{1 / 2}, g, P_{1 / 2}, d_{P}, a, d_{Z}, w_{s}, \lambda, N_{T}$.
- Complications: different types of $P, Z$, even $D$. Different nutrients (law of minimum, Redfield ratios Nitr $=16$ Phos), prey preference, mixotrophy, ...

