# STOCHASTIC MODELS IN ANIMAL POPULATION ECOLOGY

DOUGLAS G. CHAPMAN
UNIVERSITY OF WASHINGTON

## 1. Introduction

A fundamental problem of animal ecology is the distribution and abundance of animals, the title that Andrewartha and Birch [1] chose for their important textbook on the subject.

At the outset Andrewartha and Birch specify that an animal ecologist needs to be a careful naturalist, an able experimenter and have "a sound working knowledge of statistical methods." This order is in fact suggestive of the evolutionary development of ecology—beginning with description, moving on to experiments mostly though not entirely in the laboratory and now turning more and more to statistical methods to analyze experimental results and to evaluate field data.

Statistics might be thought to include the development of mathematical models. Andrewartha and Birch presumably do not—they seem more hesitant about such models and stress the possibilities of their being misused. On the other hand, one should not interpret their warnings as an injunction against the use of mathematical models for they open their second chapter with a quotation from Plato to wit "He who would be truly initiated should pass from the concrete to the abstract, from the individual to the universal."

An even more pessimistic view of the use of mathematical models particularly in the social sciences but also in the biological sciences is expressed by Hajnal [18]. For some contrary views reference is made to some of the discussants of this paper (particularly Skellam and Scott). Scott particularly points out the distinction between interpolation and structural models.

Nevertheless, structural models are the abstraction which may lead us to a more complete understanding of our universe. Even the pessimists who point to the large gap between the simplicity of most mathematical models and the complexity of the real world must admit that this is the long run solution to the problem of understanding. But the simpler models of the present time are a necessary stage in the building up of more complicated models. Furthermore, the models that we are constructing are of great value in clarifying basic concepts and assumptions for a structural model can be constructed only when there is a clear understanding of the fundamental elements and a clear formu-

This research began with the support of National Science Foundation Project No. G-19645.

lation of elementary definitions. Finally, such models are of value even when they are incomplete or wrong in providing insight into new areas to explore new experiments to conduct (see, for example, Neyman Park, and Scott [30], p. 71).

While structural models do not have to be mathematical, since mathematics is the language of science it is reasonable to suppose that most models ultimately will be mathematical. In passing, however, we note many important studies in population dynamics are being made which, while not constructing a fundamental mathematical model structure, do essentially involve models. Studies of Tinbergen [36], [37] published posthumously and the large study on the spruce budworm under the leadership of Morris [29] are examples.

Such mathematical models can be deterministic or stochastic. Since random fluctuations are inherent in animal populations and in their environment, it is desirable to incorporate chance variation as an integral component of the model. Yet the earliest mathematical models were deterministic rather than stochastic for the equally obvious reason that one must learn to walk before one can run.

Such deterministic models, while obviously oversimplified, have nevertheless played some role in ecological research and particularly have been a stimulus for both abstract and experimental work. In some situations these deterministic models give first order answers that are useful; in others the models do not agree with reality. In either case we wish to move on to improved models. Such improvements may lead to more complex deterministic models or to stochastic models. We must at some stage examine stochastic models to provide error bounds on the deterministic models, to provide answers where the deterministic models fail, and above all to shed light on the true underlying nature of the natural phenomenon being studied. This is true even though in some areas the possible deterministic models are far from completed and there remain formidable statistical problems in the measurement of the population parameters. This is the case for most large exploited populations in fisheries and wildlife work.

Limitations of space make it necessary to impose some severe limitations upon the range of topics covered. Thus, the stochastic models of genetics and of evolutionary theory certainly are relevant to the theory of the distribution and abundance of animal populations. Also, the enumeration of an animal population and the determination of estimates of the population parameters, both of which are basic to understanding of the population, often require statistical models. Both of these topics are excluded here as is the use of information theory in the study of diversity.

## 2. Population growth and interaction

The mathematical models that were developed in this area originally by Lotka and Volterra are the most well known. (See D'Ancona [13] for references and an exposition.) The basic idea is to regard the population size $N(t)$ as a continuous differentiable function of time and express $dN/dt$ as some function of $N$ and of $t$. When two or more interacting populations are involved, the differential equations will be functions of the several populations and even the deterministic models can rapidly become too difficult to solve in closed form. However, it is usually possible to study the equilibrium situation and additional information could be obtained from numerical solutions, if desired, and if enough information is available to yield estimates of the parameters involved in the equations.

Three such equations have been studied particularly as models for population interaction in the large:

(i) single population (logistic)

$$\frac{1}{N}\frac{dN}{dt} = \epsilon - \delta N; \tag{2.1}$$

(ii) two populations (competition)

$$\frac{1}{N_i}\frac{dN}{dt}i = \epsilon_i - \delta_{i1}N_1 - \delta_{i2}N_2, \qquad i = 1, 2; \tag{2.2}$$

(iii) two populations: prey predator

$$\begin{aligned} \frac{1}{N_1}\frac{dN_1}{dt} &= (\epsilon_1 - \delta_{11}N_1 - \delta_{22}N_2), \\ \frac{1}{N_2}\frac{dN_2}{dt} &= \left(\epsilon_2 - \delta_2\frac{N_2}{N_1}\right). \end{aligned} \tag{2.3}$$

Here $N_1$ denotes the prey, $N_2$ the predators. The $\epsilon$ and $\delta$ are parameters which of course play different roles in the several models. Equation (2.1) leads to simple logistic growth. Equation (2.2) has in general, when the matrix $(\delta_{ij})$ is nonsingular, four limiting possibilities. For some parameter values species 1 always persists; for another set species 2, or there may be a stable equilibrium with both species persisting or there may be an unstable equilibrium in which only one species survives but the survivor depends on the initial state of the system.

A number of differential equations have been proposed as a basis for prey-predator models. Leslie [24] argues that the model represented by (2.3) is preferable since it allows for intraspecific competition. In general it has a stationary state, with both species persisting, which is approached by a series of damped oscillations.

The first major attempt to provide stochastic versions of these deterministic models was by Feller [14]. It was immediately evident that the stochastic models are far more intractable than the deterministic models. It is thus appropriate to distinguish between three different types of attack: analytic, approximate, and Monte Carlo.

2.1. *Analytic solutions.* The obvious analytic approach is to replace the differential growth equations by transition probabilities which may depend on population size and time. For example, D. G. Kendall [20], [21] solved the single population process where the transition probabilities are linear functions of population size but completely general functions of time. Thus, the model has the following transitions and associated transition probabilities in the interval $(t, t + dt)$

$$
\begin{array}{lll}
 & n_t \rightarrow n_t + 1 & \lambda(t) n_t \, dt \\
(2.4) & n_t \rightarrow n_t & 1 - [\lambda(t) + \mu(t)] n_t \, dt, \\
 & n_t \rightarrow n_t - 1 & \mu(t) n_t \, dt.
\end{array}
$$

Here $n_t$ denotes the (random) population size at time $t$ and $\lambda(t)$, $\mu(t)$ are the birth and death rates. Defining

$$
P_n(t) = P_n[n_t = n], \qquad \Phi(z, t) = \sum_{n=0}^{\infty} P_n(t) z^n, \tag{2.5}
$$

it can be shown that $P_n(t)$ satisfies the difference differential equation

$$
\begin{aligned}
\frac{\partial P_n(t)}{\partial t} &= (n + 1)\mu P_{n+1}(t) + (n - 1)P_{n-1}(t) - n(\lambda + \mu)P_n(t), \quad n \geqq 1, \\
\frac{\partial P_0(t)}{\partial t} &= \mu P_1(t),
\end{aligned} \tag{2.6}
$$

while $\Phi(z, t)$ must satisfy the partial differential equation

$$
\frac{\partial \Phi}{\partial t} = (z - 1)(\lambda z - \mu) \frac{\partial \Phi}{\partial z} \tag{2.7}
$$

together with the boundary condition $\Phi(z, 0) = z$. It is remarked that the dependence of $\lambda$ and $\mu$ on $t$ is suppressed. The distribution derived from this is a geometric series with a modified first term. Furthermore, if the initial number is $N_0$, then

$$
E(n_t) = N_0 e^{-\rho(t)}, \qquad \mathrm{Var}\,(n_t) = N_0 e^{-2\rho t} \int_0^t e^{\rho(\tau)} [\lambda(\tau) + \mu(\tau)] \, d\tau, \tag{2.8}
$$

where

$$
\rho(t) = \int_0^t [\mu(\tau) - \lambda(\tau)] \, d\tau. \tag{2.9}
$$

It is seen that if $\mu$, $t$ are constant the population explodes or becomes extinct. Perhaps the most interesting biological result to be drawn from this model is for the case $\lambda = \mu$ (constant). Such equality could be expected in a population in equilibrium with its environment. For such a population $E(n_t) = N_0$ but $\mathrm{Var}\,(n_t) = 2\lambda N_0 t$, that is, while the population average remains constant, the fluctuations about the ceiling level increase. More precisely the standard deviation of $n_t$ is proportional to $t^{1/2}$.

Of course, in any actual population it is also to be presumed that if these increasing statistical fluctuations tend to move $n_t$ sufficiently far from $N_0$, then whatever density dependent factors are present will begin to play a role.

While this might appear to be of little biological interest, in fact it could represent a reasonable approximation over a limited range of population values. In particular, where density independent factors are more important than density dependent factors, this could serve as a starting model. The birth and death processes could be sinusoidal functions of time reflecting seasonal variations. To provide additional realism it would be desirable to add to the model large external random "shocks." Despite limited biological application to date, this solution has been an important tool in the Monte Carlo approach to stochastic population models.

This method has been used in a number of other similar or somewhat more complex situations. For example, it was pointed out in the discussion to Kendall's paper by Prendiville [33] that a solution could be found for single population growth of a "logistic like" model. Thus, it is reasonable to ask that in any single population model involving density dependence that the general relative birth and death probabilities $\lambda(n_t)/n_t$ and $\mu(n_t)/n_t$ be monotone decreasing and increasing functions, respectively. The strict equivalent to the logistic would arise if they are also linear functions. However, they may also be assumed to be hyperbolic, namely,

$$\frac{\lambda(n_t)}{n_t} = \begin{cases} \alpha\left(\frac{M_2}{n_t} - 1\right), & n_t \leqq M_2 \\ 0, & \text{otherwise,} \end{cases} \tag{2.10}$$
$$\frac{\mu(n_t)}{n_t} = \begin{cases} \beta\left(\frac{1 - M_1}{n_t}\right), & n_t \geqq M_1 \\ 0, & \text{otherwise,} \end{cases}$$

where $M_1$, $M_2$, $\alpha$, $\beta$ are parameters. A published solution can be found in Takashima [35]. In particular,

$$E(n_t) = \frac{1}{\alpha + \beta} \{\alpha M_2 + \beta M_1 - [(\alpha M_2 + \beta M_1) - N_0(\alpha + \beta)]e^{-(\alpha+\beta)t}\} \tag{2.11}$$

with limiting value $(\alpha M_2 + \beta M_1)/(\alpha + \beta)$. The interesting thing about this model is that the corresponding deterministic differential equation

$$\frac{dN}{dt} = (\alpha M_2 + \beta M_1) - (\alpha + \beta)N \tag{2.12}$$

has exactly the same solution. Furthermore, this function which is variously known as the Mitscherlich law or the Bertanlanffy growth equation has been widely used in recent years as a model in fisheries (see, for example, Beverton and Holt [9]). The estimation of the parameters of the law (in the form $y = \alpha + \beta\rho^x$) has been much studied (see Tomlinson and Abramson [38] for a recent useful set of tables with references to earlier publications) and several high speed computer programs are also available for this purpose. The estimation of $\alpha$, $\beta$, $M_1$, $M_2$ for the stochastic model has not been studied.

In most other exact formulations either the model is greatly simplified, for example, Weiss [40], or only partial results are obtained. Weiss deals with a competition process where each encounter between members of the competing species results in death of one or the other with probabilities $\sigma/(1 + \sigma)$, $1/(1 + \sigma)$, respectively. No births are assumed so that the process is immediately seen to be a slight modification of the ruin of the gambler or of random walk with absorbing boundaries. The exact probabilities of extinction and of the expected numbers of the surviving species are explictly determined which permits a comparison between deterministic and stochastic models. Denote by $p_i$ the probability that species $i$ survives. The $p_i$ are functions of $N_1$ and $N_2$, the initial population sizes, as well as of $\sigma$. Define the region in the $(N_1, N_2)$ plane where $p_i$ differ from zero and unity as the zone of "stochastic indeterminacy." As expected this zone decreases as $N_1$, $N_2$ increase.

The work of Mode [28] provides an interesting study of a prey-predator system in which the stochastic model is investigated analytically and some results obtained through general theory of recurrent processes.

2.2. *Approximate methods.* As pointed out, the stochastic equivalent of the logistic model would imply transition probabilities as follows

$$\begin{aligned} &n_t \to n_t + 1 && \lambda(n_t)\,dt, \\ &n_t \to n_t && 1 - \{\lambda(n_t) + \mu(n_t)\}\,dt, \\ &n_t \to n_t - 1 && \mu(n_t)\,dt, \end{aligned} \tag{2.13}$$

where

$$\begin{aligned} \lambda(n_t) &= \begin{cases} \alpha(M_2 - n_t)n_t, & n_t < M_2, \\ 0, & \text{otherwise,} \end{cases} \\ \mu(n_t) &= \begin{cases} \beta(n_t - M_1)n_t, & n_t > M_1, \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \tag{2.14}$$

where $0 \leqq M_1 < M_2$ are parameters of the model. For this model Feller [13] showed that $E(n_t)$ is smaller than $N_t$ obtained as a solution of the differential equation (2.1), that is, stochastic variations would lead to a smaller average population size than suggested by the deterministic model. Kendall obtained a heuristic solution for $t \to \infty$ and showed that

$$P\{n_\infty = M_1 + m\} = \frac{C}{M_1 + m} \binom{M_2 - M_1}{m} \alpha^m \beta^{(M_2 - M_1) - m}, \tag{2.15}$$

where $C$ is adjusted so that

$$\sum_{n=M_1}^{M_2} P[n_\infty = n] = 1. \tag{2.16}$$

It is seen that

$$E(n_\infty) = \frac{1}{E^*\left[\dfrac{1}{M_1 + k}\right]}, \tag{2.17}$$

where $E^*$ is expectation with respect to the $B(k; M_2 - M_1, \alpha/(\alpha + \beta))$ distribution. For $M_1 = 1$ this is approximately $M_2\alpha/(\alpha + \beta)$. Bartlett [5] has developed different approximation methods to throw light on the model; one of these is now discussed.

Assume that $\lambda(0) \neq 0$ so that the process cannot become permanently extinct at $n = 0$ and further that the probability distribution of $n_t$ tends to a limiting distribution. The first assumption is, of course, a biological impossibility unless births include immigration, but if the time to extinction is large, this may be regarded as a reasonable approximation.

For general $\lambda(n)$, $\mu(n)$ the differential equation generalizing (2.6) is

$$\frac{\partial P_n(t)}{\partial t} = \mu_{n+1}P_{n+1}(t) + \lambda_{n-1}P_{n-1}(t) - (\lambda_n + \mu_n)P_n(t), \tag{2.18}$$

so that the limiting distribution must satisfy

$$\mu_{n+1}P_{n+1} + \lambda_{n-1}P_{n-1} = (\lambda_n + \mu_n)P_n, \tag{2.19}$$

which is easily seen to have the solution

$$P_n = \frac{\lambda_{n-1}}{\mu_n} P_{n-1}, \tag{2.20}$$

where $P_n = \lim_{t\to\infty} P_n(t)$. Further, let $m$ be the solution of the equation $\lambda_{n-1} = \mu_n$ and expand $\lambda_{n-1}/\mu_n$ about $m$, namely,

$$\frac{\lambda_{n-1}}{\mu_n} = 1 + (n - m)\left[\frac{\partial}{\partial n}\left(\frac{\lambda_{n-1}}{\mu_n}\right)\right]_{n=m}, \tag{2.21}$$

where $n$ is now treated as a continuous variable. Denote the partial derivative as $-\sigma^{-2}$. Then, to a further approximation

$$P_n \sim \left[1 - \frac{(n - m)}{\sigma^2}\right]P_{n-1}, \tag{2.22}$$

or

$$\frac{\partial \log P_n}{\partial n} \sim -\frac{n - m}{\sigma^2}, \tag{2.23}$$

whence, we have

$$P_n \sim C'e^{-(n-m)^2/2\sigma^2}; \tag{2.24}$$

that is, to this approximation, $n$ is normal with mean and variance as indicated.

This and other approximations due to Bartlett [5], [6] throw light on the behavior of the process in the vicinity of the stationary state. Additional methods to yield approximate solutions to the evolutionary stages of the process and/or bounds on this stage are desirable. One possibility is that used by Gani [15] in a study of the interaction of phages and bacteria. He solves the usual partial differential equation for the probability generating function not for the fully stochastic situation but with some random variables replaced by their expectations. This equation is solvable in closed form to yield the required probability distribution.

2.3. *Monte Carlo or simulation methods.* Approximate methods are most valuable and more such need to be developed. There is a tendency to disparage them unless some bounds can be placed upon the errors of such approximations. Insight into such approximations may be obtained from Monte Carlo studies or simulation on high speed computers. When no exact or approximate solution is available the computer may represent the only way to gain additional theoretical information. Bartlett [4] and Leslie [25] were the first to demonstrate that much could be accomplished with such an approach particularly when these studies are guided by approximate or partial solutions or the solutions of the corresponding deterministic models.

There are several different ways of performing the Monte Carlo studies. If the transition probabilities can be written down in explicit form, then the computer can be programmed to generate a sample of the actual ecological process—at each step a random number determines which transition is made. Such transition probabilities can be functions, both of the past and present states of the process and also of time. Such a simulation is exemplified by Barnett [3]. This paper is valuable as a source of helpful procedures to facilitate these simulations.

A second method of Monte Carlo similation is to replace the continuous time process by a discrete time process assuming that the birth and death rates are constant over each time interval. The solution obtained by Kendall (equations (2.8) above) then yields

$$\begin{aligned} E(n_{t+1}|n_t) &= n_t e^{-\lambda_t - \mu_t}, \\ \operatorname{Var}(n_{t+1}|n_t) &= n_t \left(\frac{\lambda_t + \mu_t}{\lambda_t - \mu_t}\right)(e^{\lambda_t - \mu_t} - 1)e^{\lambda_t - \mu_t}, \qquad \lambda_t \neq \mu_t, \\ &= 2\lambda_t n_t, \qquad \lambda_t = \mu_t. \end{aligned} \tag{2.25}$$

If $n_{t+1}$ given $n_t$ is assumed to be normal with mean and variance as specified, it is easy to generate an artificial series representing the development of the process. The $\lambda_t$ and $\mu_t$ change from step to step according to their dependence on time and on the sizes of the population or populations involved. This method has been explored extensively by Leslie [25], [26] and Leslie and Gower [27]. Leslie noted that the stochastic equivalent of any of these deterministic models is not well determined and he introduced the idea of studying two extreme cases, namely, birth rate constant, so that only the death rate is a function of the various variables and conversely death rate constant but birth rate variable.

From the simulation work done to date the following conclusions may be drawn, subject to the qualification that the Monte Carlo experiments have been quite limited in number.

(a) The stochastic population model with birth and death rates which are linear functions of population size can be fitted satisfactorily by a logistic curve.

(b) The parameters of the stationary distribution of the population resulting from such a model can be well estimated by the approximate formulae due to Bartlett, that is, assuming the birth and death rate functions are known.

(c) The problem of estimating the four parameters $\alpha$, $\beta$, $M_1$, $M_2$ of (2.21) from an observed process is largely open. If one of $\lambda(n)$ or $\mu(n)$ is constant, then there are only three parameters, which may be expressed as functions of the three parameters of the logistic; but even here whether satisfactory estimates can be obtained from a single observed process is not yet known. In the more difficult case where neither $\lambda(n)$ nor $\mu(n)$ may be assumed to be constant it is presumably also necessary to utilize additional information such as the variance of $n$ about its mean (in the limiting stationary distribution). Of course an alternative approach to this problem is to obtain independent estimates of $\lambda(n)$, $\mu(n)$ from other biological observations—if these are available.

(d) Empirical probability functions (for the probability of survival of either species) have been determined for a competition model for a region of the $n_1(0)$, $n_2(0)$ plane for some parameter values that could be associated with the unstable stationary state in the deterministic model.

(e) The behavior of the competition model has been studied in the region of the stationary state (for the stable equilibrium case). Leslie and Gower ([27], p. 327) conclude from realizations that "It appears then, that, in a system of two competing species with a stable stationary state, the number of individuals over a relatively long period of time settles down to a type of distribution which is approximately normal in form, but with a degree of variation which may be greater than that expected for small deviations about the stable state. This greater degree of variation about the equilibrium level can only lead to an increased chance of random extinction of one or other of the two species."

Some of Leslie and Gower's artificial realizations also show as one would expect that near the borderline of the region in the $N_1(0)$, $N_2(0)$ plane where according to the deterministic model only species 1 should survive, there is a positive probability that in the stochastic case not species 1 will survive but rather species 2.

As yet insufficient studies have been attempted on prey-predator models, though Bartlett [4] did obtain some simulations of a model slightly simpler than that corresponding to (2.3). In all cases the predator or prey became extinct which means that the model is inadequate. Bartlett considered the effect of a time lag due to the time between birth and maturity though obviously there may be other factors which in nature serve to prevent extinction (see for example Bailey, Nicholson, and Williams [2].)

One other comparison of deterministic and stochastic models by Monte Carlo methods needs to be mentioned briefly—that due to Larkin [23] who studied the effect of a fishery on competing populations. The fishery alters the equilibrium as is to be expected—and may in fact change the situation so that one species becomes extinct. If the fishery is such that both species can continue to exist, Larkin's results suggest that there is little difference between the deterministic and stochastic models for populations in the hundreds or greater.

2.4. *Tribolium models.* The extensive and detailed series of experiments carried out at the University of Chicago by Park and his students on flour beetles (*Tribolium confusum and Tribolium castaneum*) deserve special note for they have stimulated much of this general research as well as special studies. In these *Tribolium* experiments a small number of adults of one species alone or of both species together are introduced into a vial of flour. At periodic intervals a census is taken of the population and the flour renewed. In some sense this represents an artificial aspect since in more "natural" circumstances the beetles would, as they exhaust their local supply of flour have to move to new sources.

Diffusion is a problem even within the controlled experiments but would become more so for "wild" populations.

In these single species experiments the population grows rapidly at first but eventually oscillates about some ceiling; in the two species experiments one or other species dies out. The winner depends on temperature, humidity and initial numbers but even with these conditions held fixed the winner may vary. Since the *Tribolium* has four stages in its life history, two active and two passive, and since the active stages are cannibalistic, that is, adults or larvae feed on eggs or pupae of either species, it is seen that the model is of a prey-predator system as well as involving a competition for food.

Numerous deterministic models have been tried to fit this behavior—some references are to be found in Leslie [26]. The two most ambitious attempts to set up a stochastic model for these experimental results are due to Neyman, Park, and Scott [30] and to Leslie [26] (also see Barnett [3]).

The Neyman, Park, Scott study as well as that of several of their students involve setting up probability models for several aspects of the dynamic process involved. These involve the egg laying rate which appears to vary with age (of the beetle) and with population size. Even more complex is the cannibalism factor since even in a simplified model with two stages ("active" and "passive") rather than the true four, there are four coefficients of "voracity" (the rate of consumption of passives of species $i$ by actives of species $j$ with $i, j = 1, 2$). Further insight into the cannibalism seems to involve the nature of the diffusion of the beetles both in regard to egg laying and searching for eggs. In view of all these complexities Neyman, Park, and Scott do not attempt to do more than obtain some qualitative agreements between their mathematical theory and the experimental results.

Leslie, on the other hand, attacks the problem in more classical terms using the birth and death model extended however to two stages, immature and mature which have different probabilities of death. He also introduces the obvious time lag between generations—he chooses the unit of time for his discrete time model as the time from birth to egg laying. This simplifies the model equations but creates difficulties in the estimation of the basic parameters, because the observations were not made in such time units. Then using necessarily rather crude estimates, he generates artificial realizations of the model to compare with the experimental results. The method of generating these is similar to that discussed for the simpler models. While there are some satisfactory agreements, Leslie concludes "it seems quite evident from the results presented in this paper, that this simple two age-class model does not provide an adequate description of these *Tribolium* populations." However, in a broad sense the simple model does yield qualitative agreement since models are generated with unstable stationary states. As such the analysis points the way to possibilities of better agreement when better parameter estimates are available and with somewhat more complex models. The availability of such better estimates may depend on the setting up of experiments with the requirements of the model in mind.

## 3. Dispersion and migration

As stated, even in such apparently simple experiment as with the *Tribolium*, the movement of the animals plays a role in the population growth and interaction (for a study on this see Cox and Smith [12]). Thus, no growth and interaction model of natural populations will be complete without bringing in movement also.

Naturalists have accumulated a large amount of qualitative information on the large scale and very impressive movement of animals—for example, the semiannual migration of most birds and of many mammals, the spawning migration of many fish, and so forth. The phenomenon of dispersion or random movement does not lend itself to such easy observation—experiments must be designed to measure these movements and such experiments often involve wide scale marking programs. Because of inadequacies of the placement or recovery of the marks it has been customary to avoid any quantitative analysis of movement in marking experiments.

While marking is the obvious tool that is used to identify subgroups of a population—though morphometric measurements, bloodtyping and population composition are also useful—it is possible to measure the initial spread of an animal without such refinements. This has been particularly true for a number of entomological studies which have given rise to a number of quantitative models of dispersion. Most of those are of the empirical curve fitting type. For a review of these as well as more sophisticated models see Gooch [16].

These more sophisticated models are for the most part based on the theory of diffusion and the theory of random walks. The first of such studies in ecology is due to Karl Pearson [32]. That such models can answer specific ecological questions has been demonstrated by Skellam [34].

Skellam shows that if the distribution in space of the progeny of an individual at $(\xi, \eta)$ in the $(x, y)$ plane is of the form $F(x - \xi, y - \eta)$ and if the initial starting point of the first generation was $(0, 0)$, then $\varphi_n$, the characteristic function of the distribution of the $n$th generation is $\varphi^n$ where

$$\varphi(t, u) = \iint e^{itx+uy}\, dF(x, y). \tag{3.1}$$

From this it immediately follows that if $F$ is the circular normal distribution with common variance $a^2$, then the radial distribution of the $n$th generation has the density

$$f_n(r, a^2) = \frac{2r}{na^2} e^{-r^2/na^2}, \qquad 0 \leqq r \leqq \infty. \tag{3.2}$$

Skellam uses this and an exponential (deterministic) growth law to try to answer a question on the mechanism of the spread of oaks in Great Britain in the post glacial period.

It is of interest to attack this problem somewhat differently—the spread of the population can be regarded not as a function of the distribution of all of the surviving progeny in the $n$th generation, but of the distance of the most distant of the progeny from its parent. We are thus led to extreme value theory. To simplify slightly and perhaps not too inexactly for the distribution of oaks in Great Britain consider a linear model. Denote by $S$ the most northward location of the $k$th generation. Clearly, it can be argued that if the progeny of any generation with parent located at $(\xi, \eta)$ have distribution function $F(x - \xi, y - \eta)$ where $\int x\, dF(x) = 0$, $\int x^2\, dF(x) = a^2$, then

$$S = \sum_{i=1}^{k} \max_{1 \leq j \leq n_i} X_{ij}, \tag{3.3}$$

where $n_i$ is the number of progeny in generation $i$.

It is clear that for large $k$, $S$ is approximately normally distributed but its mean will depend very strongly on the underlying distribution of the $X_{ij}$ and in fact on the tails of this distribution. For example, if

$$E\left(\max_{1 \leq j \leq n} X_{ij}\right) = g_n, \qquad \sigma^2\left(\max_{1 \leq j \leq n} X_{ij}\right) = \tau_n, \qquad \Phi(Z_\alpha) = 1 - \alpha, \tag{3.4}$$

then the equation to be solved to assure that $S = \mathfrak{Z}$ with probability $1 - \alpha$ in $k$ generations with $n$ progeny per generation is

$$\mathfrak{Z} = a[Z_\alpha(k\tau_n)^{1/2} + k_n]. \tag{3.5}$$

Suppose that $X_{ij}$ have a logistic distribution; then $g_n = \ln n + \gamma + o(1/n)$ where $\gamma = 0.57722$ and $\tau_n \to \pi^2/6$. For $k = 300$, the figure used by Skellam, $n = 1000$, $1 - \alpha = 0.997$ $\mathfrak{Z} = 600$ miles, $a = 0.27$, somewhat less than the value calculated by Skellam. However, it should be noted that Skellam used a factor of 9,000,000 for the number of progeny per parent and also assumed a normal spatial distribution. With a normal distribution and with $n = 1000$ $a = 0.62$ which is closer to the value that Skellam calculated. However, while it may not be reasonable that the root mean square of daughter oaks about their parent exceeds 1/2 mile, on the other hand, it is probable that the distribution of the location of progeny is a mixture of at least two distributions—those due to wind dispersal and those carried by animals. The latter may represent only a small fraction of the total, but for these $a^2$ might well be of the order of magnitude indicated and it is seen that the spread of oaks can be explained without resort to the assumption that pockets of oaks survived in Britain throughout the glacial period.

3.1. *Diffusion process.* The diffusion process in the plane can be obtained as the limit of a two dimensional random walk, though of course, a more rigorous approach involves the Kolmogorov forwards and backwards equations. If $N(x, y, t)$ denote the density of animals at $(x, y)$ at time $t$, and if the animals move subject to two processes, diffusion with parameter $D$ and drift with parameters $c_x$, $c_y$, then $N(x, y, t)$ satisfies the differential equation

$$\frac{\partial N}{\partial t} = -c_x \frac{\partial f}{\partial x} - c_y \frac{\partial f}{\partial y} + D\nabla^2 N \tag{3.6}$$

where $\nabla^2 = \partial/\partial x^2 + \partial/\partial y^2$ as usual.

This is the equation derived by Skellam [34] (with $c_x = c_y = 0$) and by Jones [19] (with $c_x = 0$) from first principles.

Consider now a population $N_0$ (for example, tagged fish placed at the origin at time $t = 0$ so that the boundary conditions are $N(0, 0, 0) = N_0$, $N(x, y, 0) = 0$, for all $x, y \neq 0$).

The solution to (3.2) is well known, namely,

$$N = \frac{N_0}{2\pi Dt} \exp\left\{-\frac{1}{4Dt}\left[(x - c_x t)^2 t(y - c_y t)^2\right]\right\}. \tag{3.7}$$

If these marked or tagged fish are subject to a uniform catching process then the distribution of recaptures in time is given by (3.7).

The maximum likelihood estimates of $c_x$, $c_y$, $D$ based on recaptures at $(x_i y_i)$ at times $t_i$ are given by

$$\begin{gathered} \hat{c}_x = \frac{\sum x_i}{\sum t_i}, \qquad \hat{c}_y = \frac{\sum y_i}{\sum t_i}, \\ \hat{D} = \frac{1}{4n} \sum_{i=1}^{n} \frac{1}{t_i}\left[(x_i - c_x t_i)^2 + (y_i - c_y t_i)^2\right]. \end{gathered} \tag{3.8}$$

The asymptotic variances of these estimates are found in the usual manner. In fact, much of standard normal inference theory is applicable. Unfortunately, the true situation may be much more complicated than this because the sampling is not uniform over time and space. Models that permit such additional complications are not available yet.

In his study cited earlier Skellam [34] combines the diffusion model of dispersion with deterministic growth models (exponential and logistic) to obtain a variety of differential equations for both one and two dimensional situations. As is to be expected the boundary conditions play a significant role. It is apparent that if full stochastic models are to be considered that incorporate both processes, Monte Carlo methods must presently be used.

## 4. Conclusion

That the models considered so far are much too simple is fully agreed—yet it is also true that even these simple models are often insolvable by present analytic techniques. Still the simple models have experimental value. For example, the models of Cole [11] that attempted to explain population cycles as purely random fluctuations, while not generally accepted today as the sole explanation of cycles, did pave the way for much more critical thinking on this important topic and, furthermore, demonstrated the need for much better observational data than had been often available before.

Similarly it may well be true that most animal dispersion fails to conform to the diffusion models which have been so useful in physical theories. Nevertheless, their study will provide better data than is now available.

That much more can be done with Monte Carlo or simulation techniques is obvious. In particular it is being used to answer specific management questions, for example, in fisheries exploitation or in insect, pest control (Watt [39]).

One difficulty with the simulation processes used to date is the introduction of the stochastic element. Thus, many of those studied may be unrealistic because only internal variation is permitted; yet in fact this may be the smaller component of variation. The larger component may be the random shocks of the environment. However, the superimposition of this external variance should cause no difficulty in the simulation process.

On the other hand, it should not be believed that simulation can provide theoretical solutions to models in the large without being accompanied at least in part by some analytic solutions—perhaps through approximate methods, formal solutions of the corresponding deterministic models and so forth. However, simulation and Monte Carlo methods should provide a valuable tool to gain insight into models in the small, particularly when these are tied to specific experimental or observational models. Moreover, they can provide guidance to the proper information needed in such experiments or field studies.

## REFERENCES

[1] H. G. Andrewartha and L. C. Birch, *The Distribution and Abundance of Animals*, Chicago, University of Chicago Press, 1954.

[2] V. A. Bailey, A. J. Nicholson, and E. J. Williams, "Interactions between hosts and parasites when some host individuals are more difficult to find than others," *J. Theoret. Biol.*, Vol. 3 (1962), pp. 1–18.

[3] V. D. Barnett, "The Monte Carlo solution of a competing species problem," *Biometrics*, Vol. 18 (1962), pp. 76–103.

[4] M. S. Bartlett, "On theoretical models for competitive and predatory biological systems," *Biometrika*, Vol. 44 (1957), pp. 27–42.

[5] ———, *Stochastic Population Models in Ecology and Epidemiology*, London, Methuen, 1960, 90 pp.

[6] ———, "Some stochastic models in ecology and epidemiology," *Contributions to Probability and Statistics; Essays in Honor of Harold Hotelling*, Stanford, Stanford University Press, 1960, pp. 89–96.

[7] ———, "Monte Carlo studies in ecology and epidemiology," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1961, Vol. 4, pp. 39–56.

[8] M. S. Bartlett, J. C. Gower, and P. H. Leslie, "A comparison of theoretical and empirical results for some stochastic population models," *Biometrika*, Vol. 47 (1960), pp. 1–12.

[9] R. J. H. Beverton and S. J. Holt, "On the dynamics of expected fish populations," *Fishery Investigations, Series II*, Vol. 19, London, Her Majesty's Stationary Office, 1957.

[10] C. L. Chiang, "Competition and other interactions between species," *Statistics and Mathematics in Biology*, Ames, Iowa State College Press, 1954.

[11] L. C. Cole, "Some features of random population cycles," *J. Wildlife Management*, Vol. 18 (1954), pp. 2–24.

[12] D. R. Cox and W. L. Smith, "On the distribution of *Tribolium Confusum* in a container," *Biometrika*, Vol. 44 (1957), pp. 393–404.

[13] U. D'Ancona, "The Struggle for Existence," *Bibliotheca Biotheoretica*, Vol. 6 (edited by E. J. Brill), Leiden, Netherlands, 1954.

[14] W. Feller, "Die Grundlagen der Volterraschen Theorie des Kampfes um Dasein in wahrscheinlichkeitstheoretischer Behandlung," *Acta Biotheoretica*, Vol. 5 (1939), pp.11–40.

[15] J. Gani, "Stochastic phage attachment to bacteria," *Biometrics*, Vol. 21 (1965), pp. 134–139.

[16] B. C. Gooch, "Critique of the literature dealing with mathematical aspects of dispersal and migration," unpublished report of AEC Project No. 123, University of California at Davis, 1965.

[17] E. J. Gumbel, *Statistics of Extremes*, New York, Columbia University Press, 1958, 375 pp.

[18] J. Hajnal, "Mathematical models in demography," *Cold Spring Harbor Symp. Quant. Biol.*, Vol. 22 (1957), pp. 97–103.

[19] R. Jones, "A method of analysis of some tagged haddock returns," *J. Cons. Perm. Int. Explor. Mir.*, Vol. 25 (1959), pp. 58–72.

[20] D. G. Kendall, "On generalized birth-and-death process," *Ann. Math. Statist.*, Vol. 19 (1948), pp. 1–15.

[21] ———, "Stochastic processes and population growth," *J. Roy. Statist. Soc. Ser. B*, Vol. 11 (1949), pp. 230–282.

[22] ———, "Les processus stochastiques de croissance en biologie," *Annls. Inst. Henri Poincaré*, Vol. 13 (1952), pp. 43–108.

[23] P. A. Larkin, "Interspecific competition and exploitation," *J. Fish. Res. Bd. of Canad.*, Vol. 20 (1963), pp. 647–678.

[24] P. H. Leslie, "Some further notes on the use of matrices in population mathematics," *Biometrika*, Vol. 35 (1948), pp. 213–245.

[25] ———, "A stochastic model for studying the properties of certain biological systems by numerical methods," *Biometrika*, Vol. 45 (1958), pp. 16–31.

[26] ———, "A stochastic model for two competing species of *Tribolium* and its application to some experimental data," *Biometrika*, Vol. 49 (1962), pp. 1–26.

[27] P. H. Leslie and J. C. Gower, "The properties of a stochastic model for two competing species," *Biometrika*, Vol. 45 (1958), pp. 316–330.

[28] C. J. Mode, "A stochastic model of the dynamics of host-pathogen systems with mutations," *Bull. Math. Biophys.*, Vol. 26 (1964), pp. 205–233.

[29] R. F. Morris (ed.), "The dynamics of epidemic spruce budworm populations," *Mem. Ent. Soc. Canad.*, Vol. 31 (1963), pp. 1–332.

[30] J. Neyman, T. Park, and E. L. Scott, "Struggle for existence. The *Tribolium* model: biological and statistical aspects," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1956, Vol. 4, pp. 41–79.

[31] G. J. Paulik and J. W. Greenough, Jr., "Management analysis for a salmon resource system," *System Analysis in Ecology* (edited by Kenneth Watt), New York, Academic Press, 1966, pp. 215–250.

[32] K. Pearson and J. Blakeman, "Mathematical contributions to the theory of evolution. XV. A mathematical theory of random migration," *Drap. Co. Mem. Biom. Ser.*, Vol. 3 (1906), pp. 1–54.

[33] B. J. Prendiville, "Discussion to the Symposium on Stochastic Process," *J. Roy. Statist. Soc.*, Vol. 11 (1949), p. 273.

[34] J. G. Skellam, "Random dispersal in theoretical population," *Biometrika*, Vol. 38 (1951), pp. 196–218.

[35] M. Takashima, "Note on evolutionary processes," *Bull. Math. Statist.*, Vol. 7 (1956), pp. 18–24.

[36] L. TINBERGEN, "The dynamics of insect and bird populations in pine woods. The natural control of insects in pine woods. I. Factors influencing the intensity of predation by songbirds," *Arch. Neerl. Zool.*, Vol. 13 (1960), pp. 265–343.

[37] L. TINBERGEN and H. KLOMP, "The dynamics of insect and bird populations in pine woods. The natural control of insects in pine woods. II. Conditions for damping of Nicholson oscillations in parasite-host systems," *Arch. Neerl. Zool.*, Vol. 13 (1960), pp. 344–379.

[38] P. R. TOMLINSON and N. J. ABRAMSON, "Fitting a von Bertalanffy growth curve by least squares," *Fish Bull.*, California Department of Fish and Game, No. 116 (1961), pp. 1–69.

[39] K. E. F. WATT, "Mathematical models for use in insect pest control," *Canad. Ent. Supp.*, Vol. 19 (1961), pp. 5–62.

[40] G. H. WEISS, "Comparison of a deterministic and a stochastic model for interaction between antagonistic species," *Biometrics*, Vol. 19 (1963), pp. 595–602.