MONTE CARLO STUDIES IN ECOLOGY AND EPIDEMIOLOGY

M. S. BARTLETT
UNIVERSITY OF MANCHESTER

1. Introduction

"Monte Carlo" investigations, that is (at least in the present context), the simulation of real phenomena, or idealized models of them, involving a random or probabilistic element in their structure, by the deliberate use of "random" (or pseudorandom) numbers, have already played an important role in many applications of stochastic models and processes, both by way of background material in understanding qualitatively some of the properties of such models, and more quantitatively, in the study of particular problems that are not amenable to complete mathematical solution. The advent of electronic computers has especially facilitated the extension of such investigations to include more realistic conditions not easily incorporated into tractable mathematical models; and in physics, for example, a growing use of Monte Carlo studies is evident. However, it is in biology that some of the most important developments are being realized, partly because of the often complex situations to be studied, but also because no genuine progress can be made until biologists themselves accept some of the consequences and predictions to be deduced from theoretical models. Such acceptance is often more convincing to them if the properties of the models are demonstrated numerically in typical cases than if presented in abstract mathematical terms; often, as already noted, the latter results are unavailable anyway.

In the study of biological populations changing over time, two types of process, one involving discrete units of time (or generations), and the other continuous time, are conveniently separated. In the first type, the probability $f(\mathbf{n}_r)$ of the vector number $\mathbf{n}$ (a scalar if only one type of individual is involved) at time r is given in the model in terms of the number $\mathbf{n}_{r-1}$ at time $r-1$; the model is for simplicity assumed Markovian in $\mathbf{n}_r$, though not necessarily homogeneous as regards the time, and the Monte Carlo study consists of generating one or more random series of the vector $\mathbf{n}$ as the time r increases. In the second type, the process (again assumed Markovian) is usually specified by infinitesimal transition probabilities $g(\mathbf{n}_t, \mathbf{m})\,dt$ for $\mathbf{n}_t$ changing to $\mathbf{m}(\neq \mathbf{n}_t)$. While this process is sometimes conveniently approximated by an appropriate process in discrete time steps δt, a precise realization (first suggested, I believe, by D. G. Kendall [14])

is obtained by noting (a) that the random time interval T before the next transition occurs is exponential with mean $1/G(\mathbf{n}_t)$, where

$$G(\mathbf{n}_t) \equiv \sum_{\mathbf{m}} g(\mathbf{n}_t, \mathbf{m}), \tag{1}$$

and (b) that the relative odds of the different possible transitions, when the transition occurs at time $t + T$, are proportional to $g(\mathbf{n}_t, \mathbf{m})$. Note that (b) still holds if the process is not time-homogeneous, but the random time interval T is then no longer strictly exponential, having a distribution

$$P\{T > u\} = \exp\left[-\int_t^{t+u} G(\mathbf{n}_t; \tau)\, d\tau\right], \tag{2}$$

where the transition probabilities $g(\mathbf{n}_t, \mathbf{m}; t)\, dt$ and their sum $G(\mathbf{n}_t; t)\, dt$, now explicitly vary with the time. The vector formalism covers a population specification over a discrete spatial lattice, but not more general specifications over a continuous spatial region.

One of the most quantitatively precise fields of mathematical biology is in genetics, and in particular in the evolutionary study of genetic populations; it is therefore interesting to see the recent development of Monte Carlo studies in this general domain (see [17] and [11]), for example, to check approximate conclusions from somewhat complicated mathematics. In the fields of ecology and epidemiology, both quantitative, and more qualitative or illustrative, studies have proved important. Thus in epidemiology, some of the earliest Monte Carlo demonstrations of the development of an epidemic appear to be those of Reed and Frost (compare Bailey [1]). The purpose of the present paper is to report the current position of Monte Carlo studies with which I have been associated; some of these are in ecology, but the most extensive so far are in epidemiology, and the latter field will therefore be discussed first.

2. General purpose and outline of the epidemiological studies

My own epidemiological investigations have been largely concerned with a stochastic theory for a certain class of recurrent epidemic, typified in measles, and defined and developed in detail in my last Berkeley paper [4]. In that paper I illustrated with artificial series the contrasting behavior of (i) a small community in which the level of susceptibles dropped low enough after each epidemic to ensure extinction of infection until this could be replenished from outside ([4], figure 6); (ii) either a larger community, or a disease with different parametric values, for which epidemics were likely to recur for a considerable time before fade-out of infection ([4], figure 10). The first type of behavior had also previously been illustrated [2] by a model representing measles incidence in a boarding school. The possibility of recurrence of epidemics depends of course on the renewal of susceptibles. The extinction of infection, however, is associated with the stochastic behavior in a closed population of a level of susceptibles

above or below the "threshold" value; and logically this should be studied first. Some empirical studies on this problem were reported by D. G. Kendall [15] at the last Berkeley Symposium (compare also Horiuchi and Sugiyama [13]).

Later Monte Carlo work of my own has concentrated on studying more quantitatively, with particular reference to measles, the critical community size required to maintain recurrent epidemics in a community closed to further infection from outside. The more qualitative early results have been supplemented by further (i) investigation of actual measles statistics [6], [7]; (ii) theoretical though approximate calculations [8]; (iii) Monte Carlo results, some of which have been provisionally referred to in previous papers, but which are more fully reported in the present paper.

The basic epidemiological model is defined in [3] or [4], but is repeated here for completeness. In its continuous-time (C) formulation it is defined by the infinitesimal transition probabilities:

Type of transition	*Transition probability*
$S \to S + 1, I \to I$	$\nu\, dt$
$S \to S, \quad I \to I - 1$	$\mu I\, dt$
$S \to S - 1, I \to I + 1$	$\lambda IS\, dt$

S and I here denote numbers of susceptibles and infective persons, respectively. The analogous discrete-time (D) formulation is

$$S_r = S_{r-1} + m - I_r, \tag{3}$$

where I_r is a binomial random variable with probability per "trial" $1 - (1 - p)^{I_{r-1}}$, and number of trials, S_r. The use of (C) was started for comparison with theoretical results, but the greater rapidity of (D), together with its probable closer resemblance to the true situation in measles (with its fairly well-defined incubation period), resulted in a later switch to (D). Two other complications have also been introduced. The first was the spread of actual communities over an area, so that diffusion (d) of infection in space had to be allowed for. The Monte Carlo results with this variant of the model are hardly feasible except on an electronic computer, and the series obtained are of considerable interest in illustrating the extent to which different regions in the model remain in phase. The model (Cd), which was described in detail in [6], incorporated a 6×6 grid or lattice of cells, diffusion of infectives taking place between adjacent cells. In the analogous discrete-time version (Dd), the spread of infection was more conveniently achieved by allowing infection across the common frontier of adjacent cells. The precise details are given in the Appendix.

However, with regard to the problem of time to fade-out of infection, the effect of spatial extension of the community appears less important than I thought at first; and a considerable number of series (D) ignoring spatial spread was obtained "manually." This last set was finally extended to cover some allowance for seasonal variation in infectiousness (Ds), the coefficient of infectiousness λ being modified to $\lambda[1 + 0.1 \cos(2\pi t/52)]$, where t is in weeks.

3. Further discussion of results

It is of course impossible to present here all the individual series in detail, these including the figures for each of the 36 cells in the electronic computer results. Some of them are depicted graphically, including the total notifications both for the continuous-time series as far as this was taken (figure 1) and for

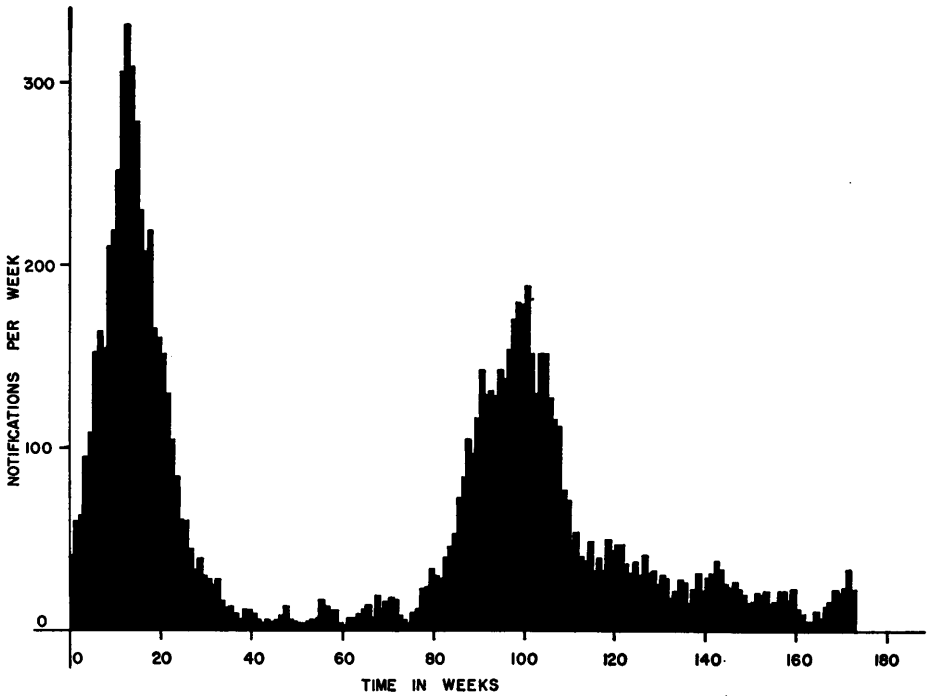

FIGURE 1

Continuous-time series
(total notifications per week, with $m = 108$).

one of the discrete-time series up to extinction (figure 2). The separate notifications for the four quarters of the grid in the case of those latter series are shown in figures 3 and 4 for two series obtained for two-thirds the original population size, that is, $m = (2/3) \times 108 = 72$. They both terminated after two epidemics, and the random variation not only between the individual quarters but also between the two series is well depicted by these two figures. The times to fade-out, which are of course very variable for the individual series, are summarized in figure 5. The series obtained manually were extended to rather small populations to try to obtain a reasonable range of population sizes, though for very small community sizes with rapid fade-out, the times to extinction are more strongly conditioned by the epidemic cycle and the starting conditions used. The empirical line

$$\log_{10} T = 2.1 + 0.005m \tag{4}$$

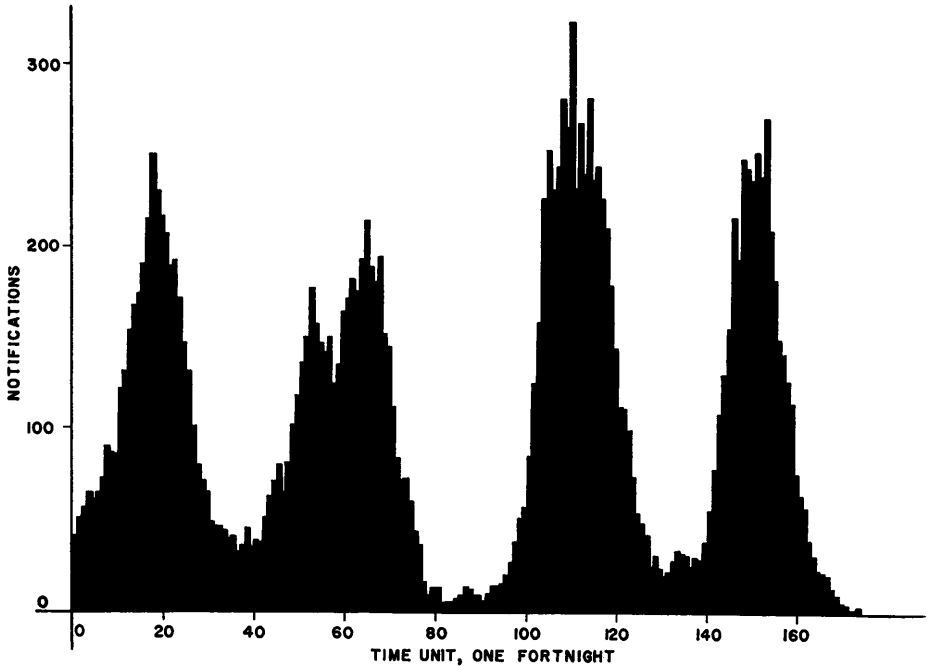

FIGURE 2

Discrete-time series
(total notifications per fortnight, with $m = 108$).

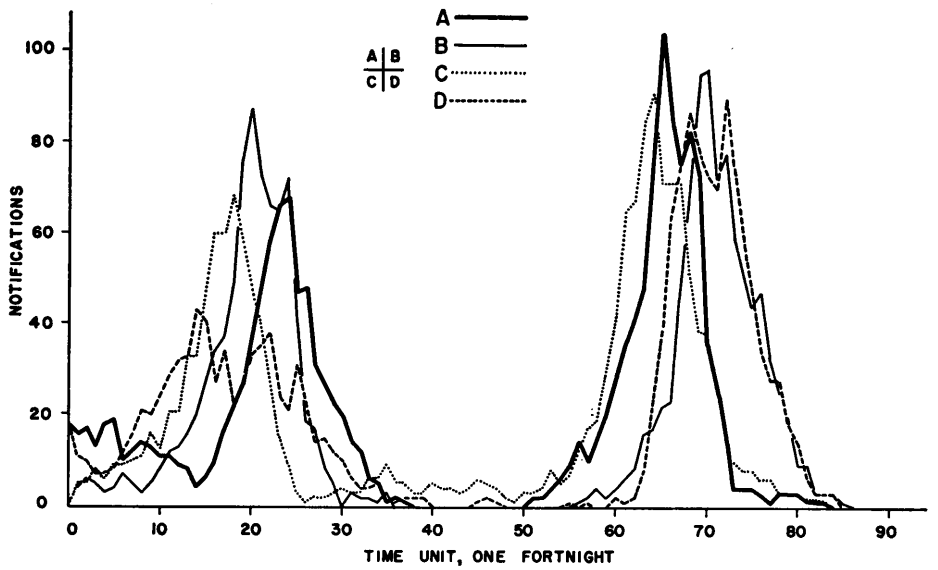

FIGURE 3

Discrete-time series
(notifications for separate quarters of the grid, with $m = 72$).

(shown on figure 5) was suggested [7] as a rough summary of all the results without seasonal variation. These last results with seasonal variation in λ are useful in providing a check on the theoretical calculation of the resulting seasonal variation in notifications based on the deterministic form of the model [this having been used to check that the 0.1 cos $(2\pi t/52)$ component should be sufficient to ensure a realized seasonal variation comparable to that observed for real measles series]. The effect of the seasonal term on average notifications during the year is shown in figure 6, and is of the order anticipated. Its effect on fade-out is less clear-cut; the suggestion, especially for $m = 27$ and 54, is that times to

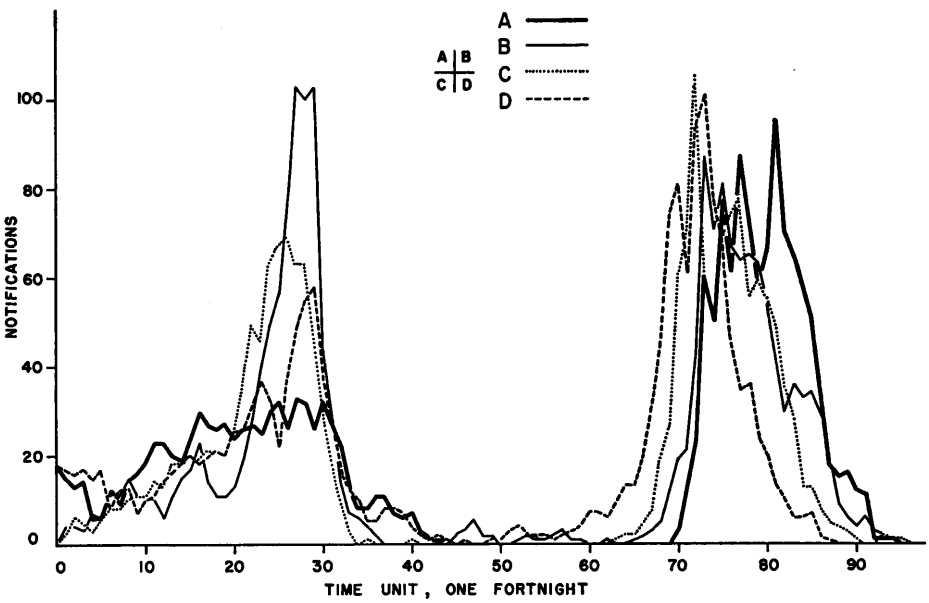

FIGURE 4

Discrete-time series
(notifications for separate quarters of the grid, with $m = 72$).

fade-out are reduced on the average, though the variability between points for the same value of m unfortunately makes the exact quantitative effect still rather doubtful, and a further set, say for the seasonal coefficient doubled, would be useful. This variability is rather strikingly illustrated by one of the series for $m = 54$, shown for interest in figure 7 (compare figure 10 of [4]) and lasting for 981 fortnight units or about 38 years.

The final figure depicted in this set, figure 8, is one of the series with seasonal variation. In spite of the average seasonal effect shown in figure 6, it is doubtful whether the individual series with and without seasonal variation could easily be separated by eye.

The full series should of course be of use for various further statistical analyses

(compare the remarks on the ecological Monte Carlo series in the next section). Some of these, which could be concerned with investigation of realized marginal distributions, say of the number of infectives, are not at first sight directly relevant to the fade-out problem, but it will be remembered that one theoretical difficulty with this problem is that no very accurate distributions are yet available theoretically, and further information from the series should, when properly

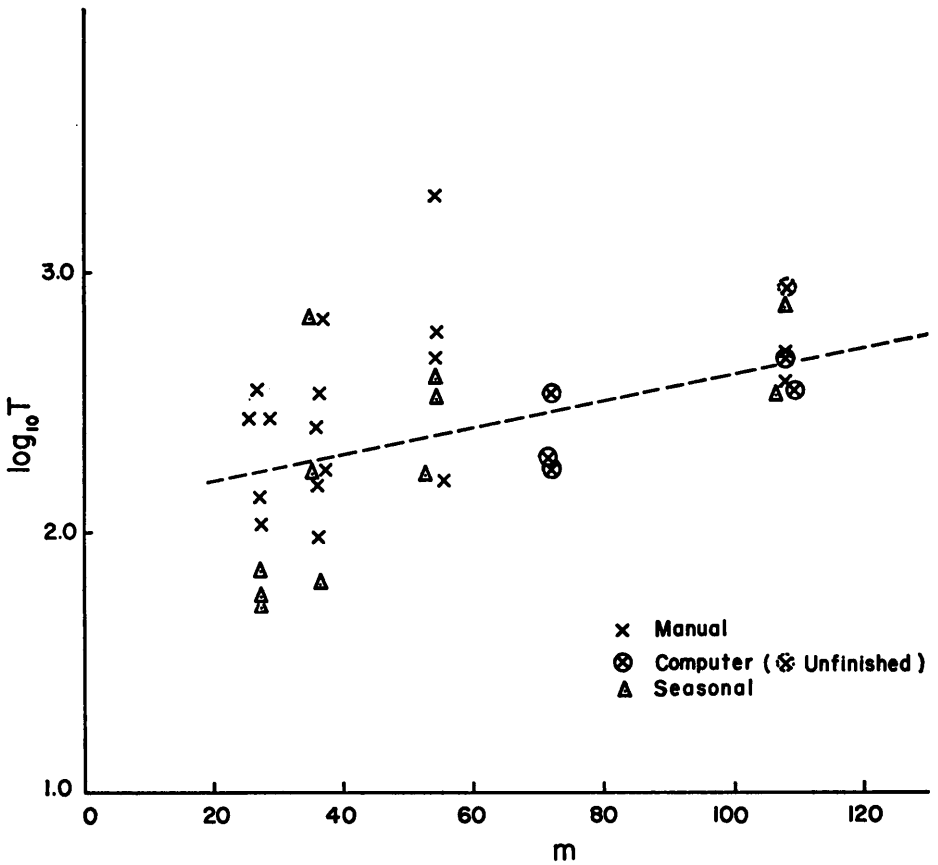

FIGURE 5

Times to extinction (T) plotted against m.

assessed, lead to more accurate estimates of extinction times than do direct observations on fade-out.

An indication of this kind of information is given in figure 9, which shows the marginal distribution of new notifications each fortnight for the long series of figure 7. The skew character of the distribution is of course not surprising, but stresses the inappropriateness of any crude normal approximation for this dis-

tribution. The observed mean of the distribution is 53.4. The standard deviation is 34.7, which is rather less than the theoretical first approximation 43.1 calculated from the formula $[m(1 + n/m)]^{1/2}$, with $m = 54$, $n = 1800$, for the continuous-time model; but how far this is accounted for by the use of the discrete-time model and also by the distribution coming from a series with a

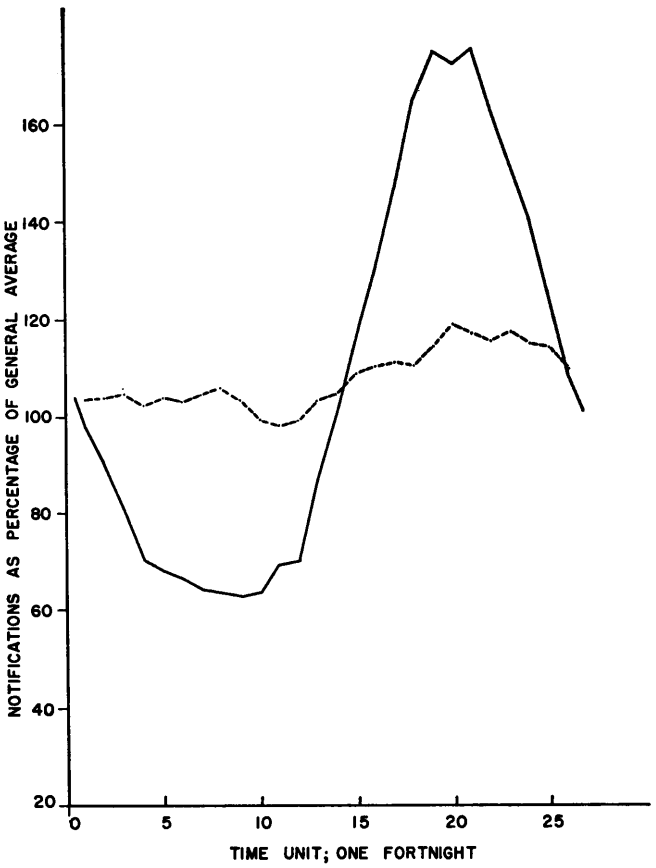

FIGURE 6

Seasonal variation in notifications over the year
(continuous line for series with 10% variation in λ,
dotted line for series with λ constant).

rather anomalously long lifetime (and hence probably with a somewhat undersize variance) has not yet been estimated. It is, however, noted that the overall notifications for the first of the electronic computer discrete-time series (which includes of course the transmission of infection from cell to cell) gave an observed mean 103.9 and standard deviation 57.1, the latter for this series being much closer to the value 60.9 obtained from the formula quoted above. An interesting problem that has arisen here is the *theoretical* relation between variance (condi-

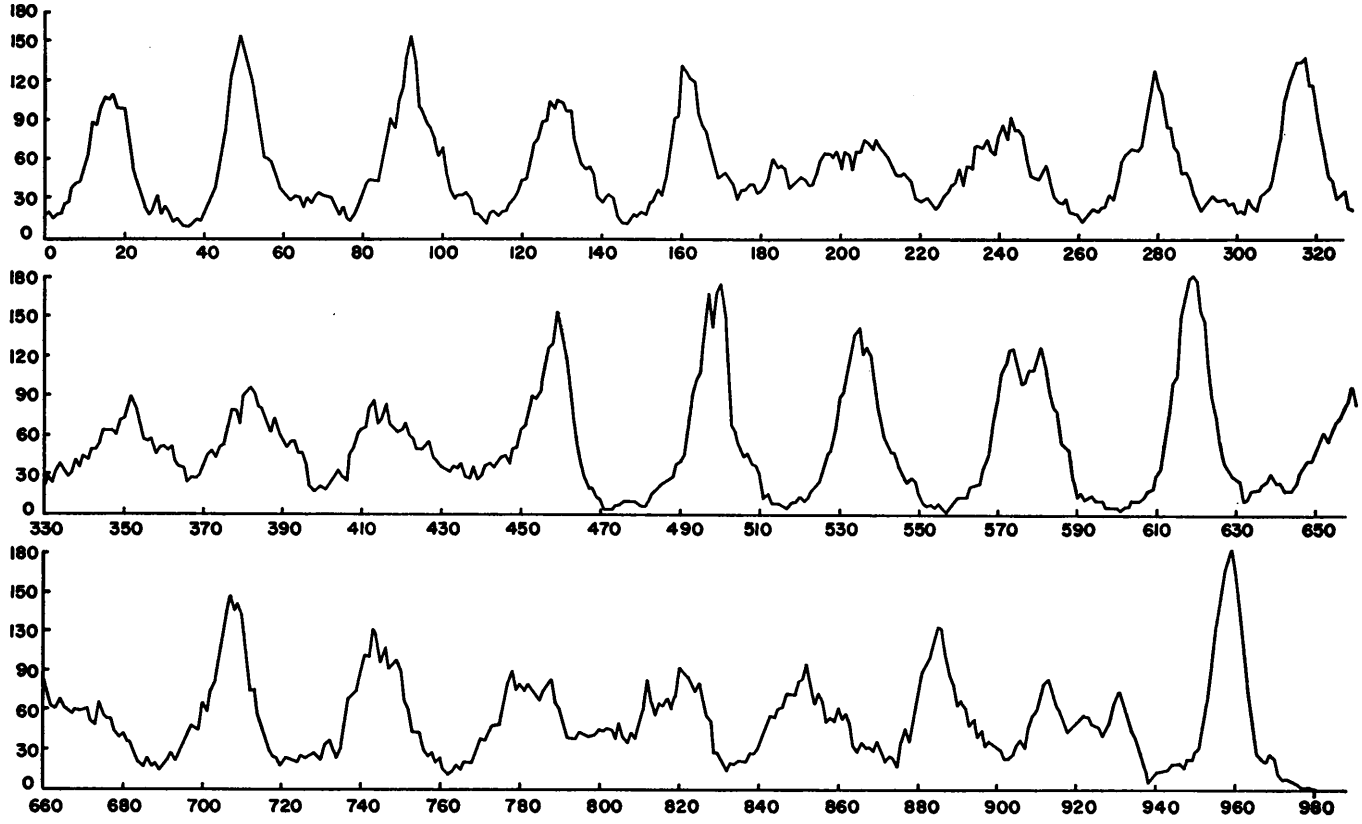

FIGURE 7

Longest series (obtained manually, with $m = 54$).

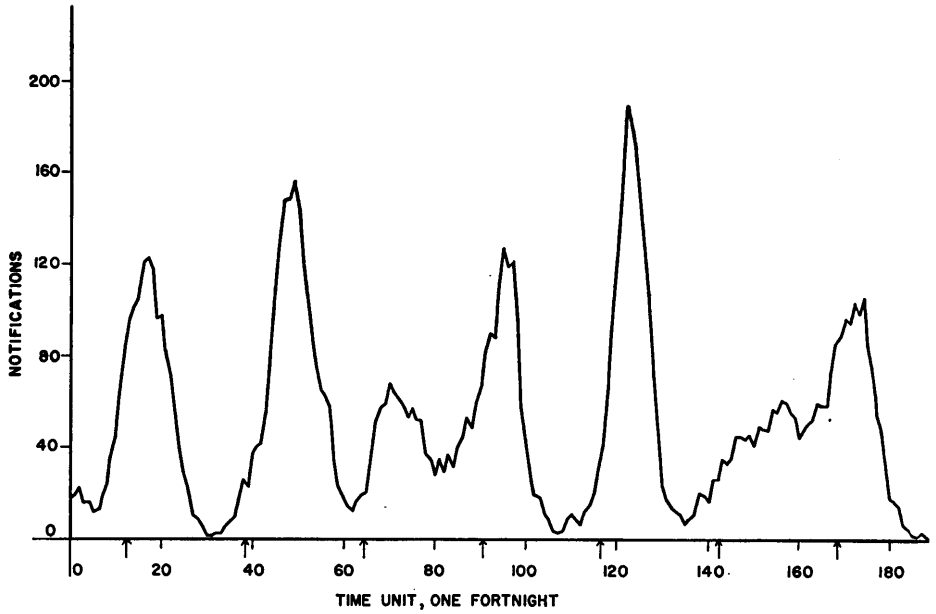

FIGURE 8

Example of series with 10% seasonal variation in λ (with $m = 54$). The arrows indicate times of maximum infectivity.

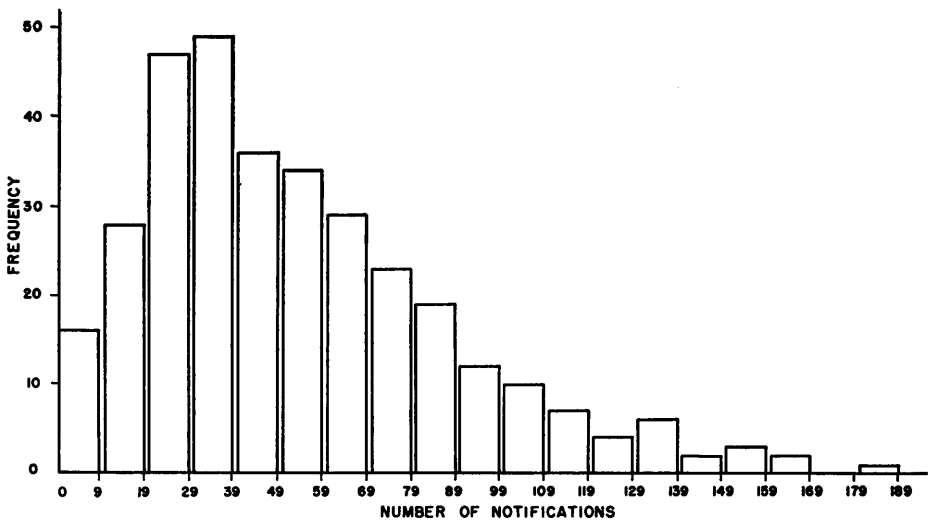

FIGURE 9

Marginal distribution of notifications for series in figure 7.

tional on observed lifetime) and "lifetime" for these "quasi-stationary" series, as I have called them.

4. Ecological studies involving two species

Among Monte Carlo studies of an ecological character must be included the various empirical investigations by plant ecologists of the statistical distributions, Poisson or otherwise, obtained from various artificial procedures such as throwing counters and the like. However, the theoretical problems in this field are usually "static" in character, and problems in animal ecology are closer to the epidemiological ones discussed above. In [5] I gave Monte Carlo realizations of simple stochastic models for (i) host and parasite, and (ii) competition between two species. Leslie and Gower [16] took up the latter problem, making use of a discrete-time version proposed by Leslie, and generating series exemplifying the situations both of a stable equilibrium point and of an unstable one. Where "stable" distributions were generated, the results could under appropriate conditions be quite well predicted by approximate theoretical formulas [10], and were bivariate analogues of the series generated for one species alone. The latter series were tested somewhat more precisely in the paper [10] just referred to, making use of second-order approximations for such quantities as the mean and skewness of the distribution of population size.

The points in common between some of these ecological situations and the epidemiological problem of sections 2 and 3 should be noticed. Thus in [5] I showed that the simple host-parasite (or prey-predator) relation was "neutral" as regards stability, but could be made more, or less, stable by introducing immigration, or age lag to maturity, respectively. The epidemiological model defined above is stable (deterministically), but this does not prevent, with the full stochastic version of the model, extinction of infectives in small communities. For larger communities what I have termed a quasi-stationary series can be generated, the "lifetime" of which can be very long. This situation is analogous to the series generated for the stochastic logistic model, for which the passage time to zero once the population size has reached the neighborhood of its equilibrium value can be effectively infinite for a small enough variance; on the other hand, for conditions conducive to a large dispersion, the extinction probability can become appreciable.

Monte Carlo results obtained for the competition problem where the deterministic equilibrium point does not exist or is unstable seem more useful, for no approximating theory is available. In the simplest case of a continuous-time model with transition probabilities:

Type of transition	*Transition probability*
$N \to N + 1, M \to M$	$A(N, M)\, dt$
$N \to N - 1, M \to M$	$B(N, M)\, dt$
$M \to M + 1, N \to N$	$C(N, M)\, dt$
$M \to M - 1, N \to N$	$D(N, M)\, dt$

where N and M represent the numbers of the two species, it is straightforward to write down the equation for the extinction probability $p_t(n, m)$, say for the first species, for initial numbers $N_0 = n, M_0 = m$; and hence to have the equation satisfied by its limiting value $p(n, m)$, namely

$$(5) \quad [A(n, m)\Delta_n - B(n, m)\Delta_n E_n^{-1} + C(n, m)\Delta_m - D(n, m)\Delta_m E_m^{-1}]p(n, m) = 0,$$

with $p(0, m) = 1, p(n, 0) = 0$, where $n, m > 0$ and $\Delta_n p(n, m) \equiv p(n + 1, m) - p(n, m)$, $E_n^{-1}p(n, m) = p(n - 1, m)$, and so on. However, it seems in practice still most convenient to obtain solutions by Monte Carlo means. This was first done systematically (using an electronic computer) by Leslie and Gower [16] for a discrete-time model, in some particular cases which were analogues of a continuous-time model with functions A, B, C, D of the form

$$(6) \quad \begin{aligned} A(N, M) &= \alpha_1 N - \alpha_{11}N^2 - \alpha_{12}NM, \\ B(N, M) &= \beta_1 N + \beta_{11}N^2 + \beta_{12}NM, \end{aligned}$$

and so on. The "birth" probability $A(N, M)$ is defined as zero when the above expression becomes negative, and this ensures a "ceiling" to the numbers N, M (and hence no mathematical complications due to escape of N or M to infinity).

The extent to which a model of this last type is representative of the ecological situation described in, say, [18] when two *Tribolium* species "compete," has been discussed in [5] and [9], and some comments about this are also made below. For the moment I shall merely define a model by the functions

$$(7) \quad \begin{aligned} A(N, M) &= 0.11N - 0.0007N^2 - 0.001NM, \\ B(N, M) &= 0.01N, \\ C(N, M) &= 0.08M - 0.0007NM - 0.0007M^2, \\ D(N, M) &= 0.005M, \end{aligned}$$

with the remark that this model, while grossly oversimplified for the purpose, was considered relevant to the *Tribolium* competition problem (at 29°C, 70% humidity) with N representing *Tribolium castaneum* and M *Tribolium confusum*. It has been studied by V. D. Barnett, to whom I am very grateful for the summary (table I) of his results on the estimation of the extinction probabilities $p(n, m)$. These were obtained by analogous methods to those of Leslie and Gower for the discrete-time model, a large number of stochastic paths on the N, M graph being run off for varying initial numbers n, m. (It should be recalled that for these stochastic paths the random time intervals between successive events are not needed.) It was found that over the limited range of initial conditions shown in table I the theoretical formula

$$(8) \quad \log \frac{p}{1 - p} = -1.71477 \log n + 1.00638 \log m - 2.01481 - 0.349612\, n + 0.139616\, m$$

fitted fairly well when the calculated probabilities were compared with the actual extinction relative frequencies, although some systematic discrepancies can

TABLE I

EXTINCTION PROBABILITIES FOR SPECIES N

m \ n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
24	.985	.934	.832	.681	.507	.347	.223	.139	.085	.052	.032	.020	.012	.008
	31/31	31/31	28/32	29/47	26/38	10/35	9/34	2/25	1/35	1/33	0/17	1/17	0/20	0/20
23	.982	.922	.805	.641	.462	.307	.193	.118	.072	.044	.027	.016	.010	.006
	38/38	40/41	32/40	25/44	23/40	12/41	8/34	2/25	0/32	1/35	2/24	1/21	0/20	0/24
22	.979	.907	.775	.597	.416	.269	.166	.101	.061	.037	.022	.014	.008	.005
	45/45	43/45	27/45	21/48	17/49	14/52	6/33	1/34	0/32	1/34	2/23	0/15	0/23	0/20
21	.974	.890	.741	.552	.372	.234	.142	.085	.051	.031	.018	.011	.007	.004
	50/50	47/49	32/51	22/46	18/50	16/46	5/32	2/45	1/37	1/34	1/26	0/23	0/30	0/24
20	.969	.871	.703	.505	.329	.202	.120	.071	.042	.025	.015	.009	.006	.004
	60/60	59/66	50/77	24/60	19/54	11/44	4/37	3/51	0/36	0/32	1/22	1/28	0/27	0/22
19	.963	.847	.661	.457	.288	.173	.102	.060	.035	.021	.013	.008	.005	.003
	67/69	51/62	52/85	32/64	17/51	9/41	5/41	4/52	1/37	0/34	0/22	1/29	0/18	0/18
18	.955	.821	.617	.409	.250	.147	.085	.050	.029	.017	.011	.006	.004	.002
	71/73	54/64	54/96	31/67	13/61	10/50	5/50	4/52	1/48	0/26	0/23	1/26	1/23	/025
17	.946	.790	.569	.363	.215	.124	.071	.041	.024	.014	.009	.005	.003	.002
	68/72	62/74	51/100	27/66	12/64	10/56	4/55	2/36	1/31	0/26	0/24	1/24	1/22	0/22
16	.935	.754	.519	.318	.183	.104	.059	.034	.020	.012	.007	.004	.003	.002
	70/74	68/83	46/101	31/81	16/82	10/58	2/53	1/31	1/36	0/29	0/29	1/33	0/15	0/19
15	.921	.715	.468	.275	.154	.086	.049	.028	.016	.010	.006	.004	.002	.001
	75/80	69/92	45/104	32/98	12/79	6/51	2/52	1/33	1/33	1/30	1/33	1/24	0/21	0/22
14	.904	.670	.417	.235	.129	.071	.040	.023	.013	.008	.005	.003	.002	.001
	74/82	68/99	44/112	24/93	6/77	5/46	3/44	1/28	0/28	0/24	0/23	0/19	0/25	0/20
13	.884	.621	.366	.199	.107	.058	.032	.018	.011	.006	.004	.002	.001	
	89/105	83/124	44/124	23/96	8/66	4/46	0/45	1/32	0/37	0/29	0/31	0/28	0/23	
12	.860	.568	.316	.166	.088	.047	.026	.015	.009	.005	.003	.002		
	97/116	85/136	50/160	19/107	4/63	4/57	2/50	1/31	0/36	0/33	0/37	0/33		
11	.830	.512	.269	.137	.071	.038	.021	.012	.007	.004	.002			
	103/128	44/82	55/174	14/98	4/66	3/57	0/50	0/42	0/44	0/33	0/27			
10	.794	.453	.226	.112	.057	.030	.017	.009	.005	.003				
	109/143	34/84	19/71	11/89	2/63	0/47	0/47	0/36	0/32	0/24				
9	.751	.393	.186	.089	.045	.024	.013	.007	.004					
	69/98	27/71	15/59	8/102	0/48	0/47	0/54	0/36	0/27					
8	.700	.334	.150	.071	.035	.018	.010	.006						
	69/108	23/67	13/68	2/45	0/38	0/34	0/40	0/28						
7	.639	.276	.118	.055	.027	.014	.008							
	40/69	18/65	8/67	3/48	0/41	0/41	0/44							
6	.569	.221	.091	.041	.020	.011								
	41/76	24/96	5/82	3/70	0/51	0/43								
5	.488	.170	.067	.030	.015									
	45/93	28/142	4/101	3/74	1/47									
4	.399	.125	.048	.021										
	44/103	4/46	0/26	0/25										
3	.302	.085	.032											
	34/90	2/44	0/22											
2	.200	.051												
	15/62	0/30												
1	.098													
	4/39													

be noticed in the regions of small p and $1 - p$. When considering these discrepancies, it should be remembered that each stochastic path contributes to several "initial" starting points, the accumulated frequencies for which are thus not independent. One interesting feature of the use of the Manchester electronic computer to obtain these results was the possibility of photographing typical paths from an oscilloscope, two such paths being reproduced as accurately as possible in figure 10. These two paths were both for the initial conditions $n = 2$, $m = 12$, which (with $p \sim 0.568$) represented a point from which the outcome was very uncertain. For $n = 4$, $m = 4$, on the other hand, the proba-

bility of survival of N (representing *Tribolium castaneum*) was near unity ($1 - p \sim 0.979$).

Before comparing such results with those reported by Professor Park for *Tribolium*, two points must be stressed. The first is that the above model has telescoped all the stages of the flour-beetle into one, whereas the development from egg to larva to pupa to mature adult (which eats some of the eggs) is a process over time which should (and obviously does from the records) lead to oscillatory waves in the population growth. Such oscillations, especially at the

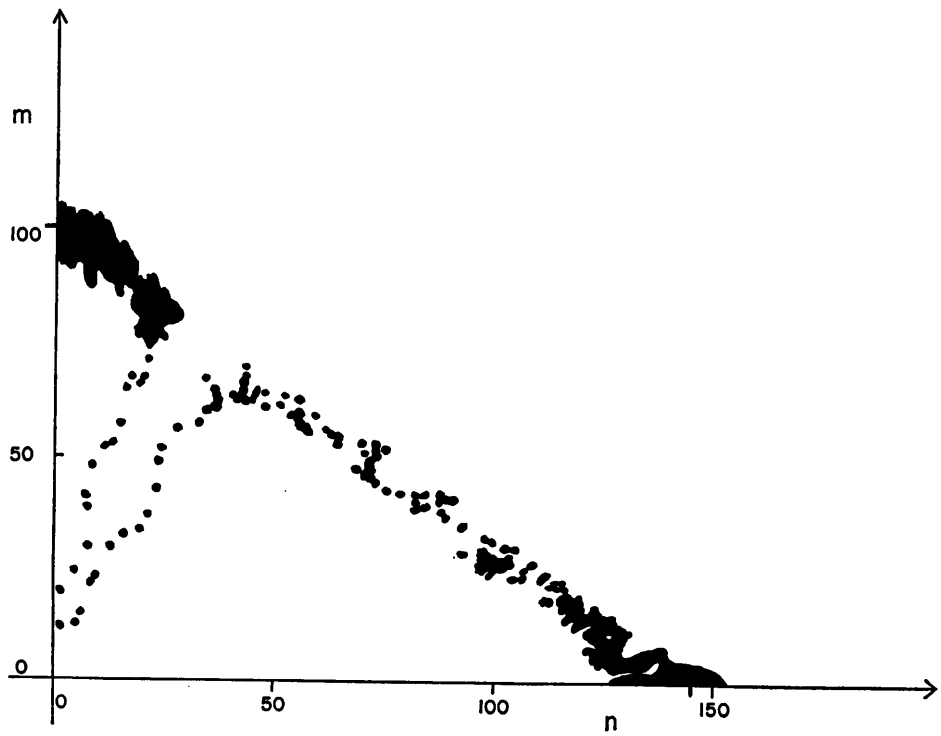

FIGURE 10

Two paths from the same starting point $n = 2$, $m = 12$
(based on photographs).
Path to right 2497 steps; path to left 1142 steps.

initial stages of growth of the population, must be very relevant to the extinction phenomenon when two species are put together, and should consequently be incorporated, for example, by some appropriate age lag to maturity, into any model aiming at quantitative agreement.

The second point is that the results already published by Park (for example, an estimated probability $1 - p(4, 4) = 0.86$ at 29°C, 70% humidity) are in process of revision, as Professor Park has subsequently isolated different strains of each species with differing behavior, and any final comparison must wait

until the results of his further investigations are available and properly assessed.

Before concluding this section I will note one further property of some ecological models possibly relevant to the first point above of the effect of age lag to maturity. Nicholson [19] reported a very regularly oscillating population of sheep blowfly (*Lucilia cuprina*), attributable to the dependence of the production of eggs on the density of the adult population. Suppose we consider a very simple deterministic model with two stages of growth with numbers P (say eggs) and A (adults), using discrete time to emphasize the age lag to maturity (the time unit being taken to correspond to the maturing of P to A). Let

$$\begin{aligned} P_r &= (\lambda - \delta A_{r-1})A_{r-1}, \\ A_r &= A_{r-1}(1 - \epsilon) + P_{r-1}, \end{aligned} \tag{9}$$

where in the first equation we take $P_r = 0$ if the expression on the right becomes negative. The "equilibrium" values of P and A are

$$P_e = \frac{(\lambda - \epsilon)\epsilon}{\delta}, \qquad A_e = \frac{\lambda - \epsilon}{\delta}; \tag{10}$$

but if we put $P = P_e(1 + p)$, $A = A_e(1 + a)$, we find for small a and p

$$p_r \fallingdotseq -\frac{\lambda - 2\epsilon}{\epsilon} a_{r-1}, \qquad a_r = a_{r-1}(1 - \epsilon) + \epsilon p_{r-1}, \tag{11}$$

whence

$$a_r \fallingdotseq a_{r-1}(1 - \epsilon) - (\lambda - 2\epsilon)a_{r-2}. \tag{12}$$

This recurrence relation gives instability if $\lambda - 2\epsilon > 1$. With regard to models for *Tribolium*, note that this conclusion is unaltered if in (9) we insert a "cannibalism" term $-\mu A_{r-1}P_{r-1}$ as well as, or instead of, the direct density-dependence term $-\delta A_{r-1}^2$. But owing to the restriction of P_r to zero if A_{r-1} increases above λ/δ, then A_{r+1} decreases by the second equation. This leads to a *self-regulating* cyclic mechanism even with the deterministic model, and stochastic fluctuations with such a model would be unimportant. This conclusion does not depend on the *abrupt* change in the behavior of the production of eggs, when $A_r = \lambda/\delta$, as the first equation may be replaced if desired by

$$P_r = \lambda A_{r-1}e^{-\eta A_{r-1}}. \tag{13}$$

5. Concluding and summarizing remarks

While a general approach to Monte Carlo studies in biology has been made by way of introduction, the specific results discussed have tended to center on two problems: (i) that of a model for recurrent epidemics, with particular reference to fade-out of infection for measles, (ii) that of competition between species in ecological models, with particular, though not very quantitative, reference to Park's *Tribolium* data.

The series obtained should still yield a good deal more information than has yet been extracted, at least in the direction of further summarizing of empirical

distributional and other results. This emphasizes the very valuable secondary function of investigations of this kind, that even when they are still rather provisional in regard to a specific purpose, such as fade-out times for measles or extinction probabilities for *Tribolium confusum* vs. *Tribolium castaneum*, they also yield useful artificial series as background material, the results from which are available for comparison with theory if or when the latter becomes available.

I am much indebted to the following persons for help in obtaining the various Monte Carlo series: J. C. Gower, for the continuous-time spatial epidemic model; E. Kerr, for the discrete-time spatial epidemic model; Christine Caley, for the discrete-time epidemic series obtained "manually"; and V. D. Barnett, for the competition between species model.

APPENDIX

In [6] the coefficients used for a continuous-time model with a 6×6 grid were (scheme II):

$\lambda = 0.005$ (rate of infection per week per infected person per susceptible person in one cell);

$\nu = 1.5$ (rate of entry per week of new susceptibles into one cell);

$\mu = 0.5$ (rate of removal or recovery of infectives per week per infected person);

$\epsilon = 0.125$ (rate of migration per week per infected person in one cell to any one neighboring cell with common boundary).

This gave $n = 36\mu/\lambda = 3{,}600$, $m = 36\nu/\mu = 108$.

In the corresponding discrete-time model the unit of time was a fortnight, and the effective number J of infectives in any interior cell (with an obvious adjustment for boundary cells) available for infecting the susceptibles in that cell was assumed to be

$$J = \frac{1}{2}I + \frac{1}{8}[I^{(a)} + I^{(b)} + I^{(c)} + I^{(d)}] \tag{14}$$

where I was the actual infectives in the cell, and $I^{(a)}$, $I^{(b)}$, $I^{(c)}$, and $I^{(d)}$ the numbers in the four relevant adjacent cells.

To drop the population size down to two-thirds its previous value, the rate of entry of susceptibles per cell per fortnight was dropped from 3 to 2, the probability p increased from 0.01 to 0.015, and the formula for J modified to

$$J = \frac{1}{4}I + \frac{3}{16}[I^{(a)} + I^{(b)} + I^{(c)} + I^{(d)}]. \tag{15}$$

One series was attempted with a change in the extent of cross-infection, the expression for J (for the larger population size) being

$$J = \frac{3}{4}I + \frac{1}{16}[I^{(a)} + I^{(b)} + I^{(c)} + I^{(d)}]. \tag{16}$$

Unfortunately, owing to a machine fault this series had to be curtailed.

It should be noted that the basis of comparability of cross-infection used in the continuous- and discrete-time models is that diffused infectives in the continuous-time model will on average only be available in the cell they have migrated to for half the total (small) time-interval under consideration.

REFERENCES

[1] N. T. J. BAILEY, *The Mathematical Theory of Epidemics*, London, Griffin, 1957.

[2] M. S. BARTLETT, "Stochastic processes or the statistics of change," *Appl. Statist.*, Vol. 2 (1953), pp. 44–61.

[3] ———, *An Introduction to Stochastic Processes*, Cambridge, Cambridge University Press, 1955.

[4] ———, "Determinisitic and stochastic models for recurrent epidemics," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1956, Vol. 4, pp. 81–109.

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

[6] ———, "Measles periodicity and community sizes," *J. Roy. Statist. Soc., Ser. A*, Vol. 120 (1957), pp. 48–70.

[7] ———, "The critical community size for measles in the United States," *J. Roy. Statist. Soc., Ser. A*, Vol. 123 (1960), pp. 37–44.

[8] ———, "Some stochastic models in ecology and epidemiology," *Hotelling Festschrift*, Stanford, Stanford University Press, 1960.

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

[10] 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. 1–11.

[11] W. F. BODMER, "Discrete stochastic processes in population genetics," *J. Roy. Statist. Soc., Ser. B*, Vol. 22 (1960), pp. 218–244.

[12] P. GREIG-SMITH, *Quantitative Plant Ecology*, London, Butterworths, 1957.

[13] K. HORIUCHI and H. SUGIYAMA, "On the importance of the Monte Carlo approach in the research of epidemiology," *Osaka City Med. J.*, Vol. 4 (1957), pp. 59–62.

[14] D. G. KENDALL, "An artificial realisation of a simple 'birth-and-death' process," *J. Roy. Statist. Soc., Ser. B*, Vol. 12 (1950), pp. 116–119.

[15] ———, "Deterministic and stochastic epidemics in closed populations," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1956, Vol. 4, pp. 149–165.

[16] P. H. LESLIE and J. C. GOWER, "The properties of a stochastic model for two competing species," *Biometrika*, Vol. 45 (1958), pp. 316-330.

[17] H. LEVENE and T. DOBZHANSKY, "New evidence of heterosis in naturally occurring inversion heterozygotes in *Drosophila Pseudo-obscura*," *Heredity*, Vol. 12 (1958), pp. 37–49.

[18] 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.

[19] A. J. NICHOLSON, "An outline of the dynamics of animal populations," *Austral. J. Zool.*, Vol. 2 (1954), pp. 9–65.