

review article

Thresholds and breakpoints in ecosystems with a multiplicity of stable states

Robert M. May*

Theory and observation indicate that natural multi-species assemblies of plants and animals are likely to possess several different equilibrium points. This review discusses how alternate stable states can arise in simple 1- and 2-species systems, and applies these ideas to grazing systems, to insect pests, and to some human host-parasite systems.

In all but the most trivial areas of enquiry, there arise questions about the extent to which events are shaped by predictable natural laws as against the accidents of initial conditions and perturbations. Is the human story largely a deterministic tale of civilisations marching to Toynbee's tune, three and a half beats to disintegration, or did the hinge of history turn on the length of Cleopatra's nose? Such questions of the relative roles of chance and necessity¹ are fundamental in modern cosmology^{2,3}, in the foundations of statistical mechanics^{4,5}, and in evolutionary biology¹ and ecology, even though they may arise in less blatant and romantic fashion than the 'what ifs' of history and the social sciences.

Viewing the grand sweep of evolution, we can see many examples where the taxonomic details of the plant or animal that occupies a given niche at a given time and place depend on historical accident, but where the niches themselves, and the broad patterns of community organisation, are remarkably constant⁶⁻⁹.

Taking a much narrower and more local view, it is interesting to consider a particular assembly of species, with specified interactions among them, and to ask questions about the dynamics of the system. Is the dynamical behaviour described by the multi-dimensional generalisation of a single valley (a global attractor)? Or is the dynamical landscape pockmarked with many different valleys, separated by hills and watersheds? If the former, the system has a unique stable state, to which it will tend (like a marble seeking the bottom of a cup) from all initial conditions, and following any disturbance. If the latter, the state into which the system settles depends on the initial conditions; the system may return to this state following small perturbations, but large disturbances are likely to carry it into some new region of the dynamical landscape (so that the system behaves rather like the ball in a pin-ball machine). If there is a unique stable state, historical effects are unimportant; if there are many alternative locally stable states, historical accidents can be of overriding significance. Obviously, questions of this kind are very important in the understanding and management of ecosystems.

A large body of empirical observations shows that many natural communities have a multiplicity of stable states. Sutherland¹⁰ has demonstrated that for the fouling community (a complex assembly of hydroids, tunicates, bryozoans, sponges and associated species) at Beaufort, North Carolina, the order of larval recruitment determines the way the community develops. Reviewing other work on the marine rocky intertidal (see also the review by Levin¹¹), on coral reefs, on freshwater lakes, and on terrestrial plant communities, Sutherland concludes that community structure can often "be explained only by referring to specific historical events" and therefore that "multiple stable points are an undeni-

able reality." A similar conclusion emerges from Connell's and Slatyer's¹² survey of mechanisms of succession in natural communities.

The view that complicated ecosystems possess many alternative stable states is also supported by theoretical studies of mathematical models that caricature such systems. From the growing number of possible examples, I mention only two, chosen from opposite ends of the spectrum. Austin and Cook¹³ have made computer studies of a system in which 94 species (embracing plants, herbivores and carnivores) are linked together by interactions that aim to be relatively realistic; the system has many equilibrium points, and is easily transferred from one to another. Case and Gilpin¹⁴ have explored a relatively abstract system, in which the coefficients in the interaction matrix for a n -species Lotka-Volterra model are assigned random values; if n is at all large, the system typically collapses to one or other of a variety of simpler systems with fewer species, and this final steady state depends on the initial population values. The notion of 'resilience' has been introduced by Holling¹⁵ in an attempt to characterise the degree to which a system can endure perturbations without collapsing or being carried into some new and qualitatively different state. Theoretical ideas about resilience, along with interpretive reviews of diverse other meanings that can be attached to 'stability' in an ecological context, are the subject of many recent papers¹⁶⁻²².

It is thus clear that real ecosystems possess multiple stable states, as do plausible mathematical models. Unfortunately, the complications inherent in multi-species systems almost invariably preclude any quantitative confrontation between theory and data. For multi-species communities, the empirical observations remain largely anecdotal, and the theory remains largely metaphorical.

For simple 1- and 2-species situations it is, however, beginning to become possible to put theory and observation together, to gain insight into the workings of systems with more than one stable state. This review is a synthesis of several examples of this kind. I begin with grazing ecosystems (and then, more generally, any harvested crop or animal population), go on to insect pest systems (particularly the spruce budworm), and conclude with some human host-parasite systems.

Grazing ecosystems

Consider^{23,24} a population of herbivores, maintained at a constant density H , and sustained by vegetation whose biomass is V .

Our interest is centred on the dynamics of the vegetation biomass. Following Noy-Meir²³, suppose that in the absence of grazing the growth rate of the vegetation as a function of V is $G(V)$, and that the herbivores consume the vegetation at a net rate $C(V)$ (corresponding to a per capita consumption at a rate $c(V)$: $C(V) = Hc(V)$). Then the rate at which V changes is given by

*Biology Department, Princeton University, Princeton, New Jersey 08540

$$\frac{dV}{dt} = G(V) - C(V) = G(V) - Hc(V) \quad (1)$$

The vegetation biomass will thus tend to an equilibrium level where the natural growth rate exactly balances the loss rate due to grazing; that is, to a value of V such that $G(V) = C(V)$. Noy-Meir shows how these equilibrium V -values can be found graphically, for various assumptions about the way $G(V)$ and $C(V)$ depend on V .

Fig. 1 illustrates one plausible situation. Here, as shown by the solid curve, $G(V)$ is positive even when $V = 0$ (corresponding to some constant background contribution to the vegetation growth rate, for example from the sprouting of seeds blown in from other areas). At first $G(V)$ increases as V increases, but as V continues to increase $G(V)$ decreases as shading and competition for nutrients becomes important, until $G(V) = 0$ at $V = K$; K is the maximum stable biomass of ungrazed vegetation. The general shape of the per capita consumption function $c(V)$ is a saturation curve: at low V , the herbivore intake is limited by forage availability, so that $c(V)$ increases with increasing V ; at high V , $c(V)$ saturates to some constant determined by the animal's intake capacity or digestion rate. It is useful to christen V_0 as the characteristic value of V at which the consumption function saturates. The dashed curves for total consumption rate $C(V) = Hc(V)$ in Fig. 1 are for a $c(V)$ that increases linearly with V at small V , corresponding to a herbivore that searches randomly (Holling's "Type II" or "invertebrate" predator search pattern²⁵).

We see from Fig. 1 that for relatively small values of H there is a single equilibrium value for V , given by the point A where $G(V)$ and $C(V)$ curves intersect; this value is slightly less than K . For relatively large values of H , there is again a unique equilibrium value for V (corresponding to the point E), now at a low value of V . For intermediate values of H , the $G(V)$ and $C(V)$ curves intersect at three points. The points B and D correspond to locally stable V -values, whose domains of attraction are divided by the unstable equilibrium point corresponding to C; B and D are the bottoms of adjacent valleys, and C marks the watershed between them.

The essential feature that leads to two alternative stable states for intermediate H values in Fig. 1 is the assumption that $c(V)$ saturates to a constant for values of V significantly below the ungrazed equilibrium. That is, defining

$$\alpha = V_0/K \quad (2)$$

Fig. 1 is for α significantly less than unity. If α exceeds, or is of the

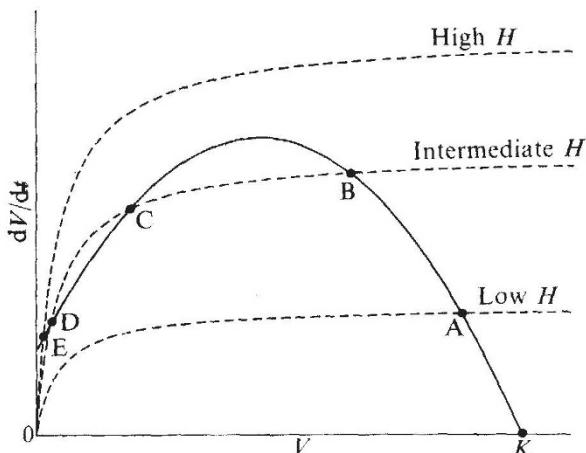


Fig. 1 The rate of change of vegetation biomass, dV/dt , is shown as a function of V . The solid curve is the natural, ungrazed vegetation growth rate (which here is finite even for $V = 0$). The dashed curves are loss rates due to grazing (of Type II pattern) at high, intermediate and low herbivore densities, H . Where the solid curve lies above the dashed one, the net growth rate is positive; where the solid curve lies below the dashed one, the net growth rate is negative; the points of intersection of the curves correspond to possible equilibrium points. For further discussion, see the text.

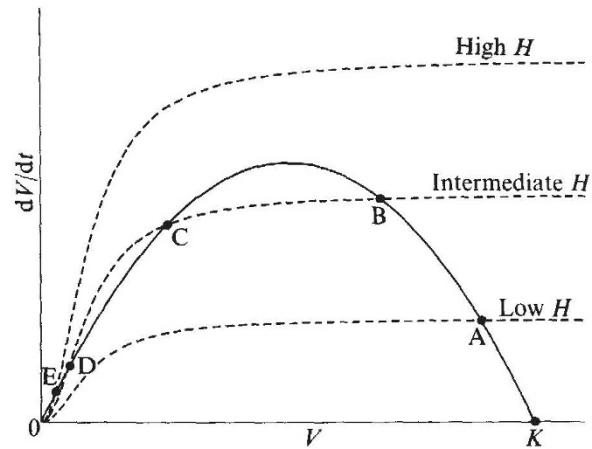


Fig. 2 As for Fig. 1, except that here the natural growth rate (solid curve) is linearly proportional to V at small V , and the loss rate due to grazing (dashed curves) is of Type III. Specifically, this figure corresponds to equations (3) and (4); the α of equation (2) is here $\alpha = 0.1$, and 'high, intermediate and low H ' are represented by $\gamma = 0.35, 0.22$ and 0.10 , respectively. The basic features shown by the figure are, however, generic.

order of unity, there is always only one stable state in the figure corresponding to Fig. 1.

Fig. 2 illustrates another plausible circumstance, which differs from Fig. 1 in details but not in essentials. Here, as V increases, $G(V)$ increases from zero to some maximum value and then decreases back to zero at $V = K$. The per capita consumption function $c(V)$ is now for herbivores whose foraging efficiency increases faster than linearly with V at low V values, but which again saturates to a constant for V above some characteristic value V_0 (Holling's "Type III" or "vertebrate" predator consumption function²³, which is also typical of many invertebrates²⁶⁻²⁸). Provided that V_0 is significantly less than K (that is, α is small), we can again distinguish three domains of dynamical behaviour for the vegetation biomass: for small H there is a unique equilibrium value of V , slightly below K (the point A); for large H there is a unique equilibrium at a low V value (the point E); for intermediate H there are two alternative stable states (the points B and D), divided by an unstable point (C).

Specifically, Fig. 2 illustrates the situation where $G(V)$ is given by the familiar logistic equation, $G(V) = rV(1 - V/K)$, and where $C(V)$ is the "Type III" consumption function $C(V) = \beta HV^2/(V_0 + V^2)$. Then equation (1) has the particular form

$$\frac{dV}{dt} = rV\left(1 - \frac{V}{K}\right) - \frac{\beta HV^2}{V_0 + V^2} \quad (3)$$

This equation may equivalently be written in dimensionless form, by using the earlier definition of α , and introducing the rescaled variables $X = V/K$, $\tau = rt$ and $\gamma = \beta H/rK$

$$\frac{dX}{d\tau} = X(1 - X) - \gamma X^2/(x^2 + X^2) \quad (4)$$

This equation can exhibit (D. Ludwig, D. Jones and C. S. Holling, to be published) two alternative stable states if, and only if, $\alpha < 1/3 \sqrt{3}$.

Figure 3 is derived from the situation depicted in Fig. 2, and shows the stable equilibrium value(s) of the vegetation V , as a function of the herbivore density H . For low H , the vegetation biomass tends to settle to a unique steady value, slightly below K . As H increases beyond a threshold value (T_1), a second stable state for V appears, in a discontinuous fashion. As H continues to increase, there occurs a second threshold (T_2), beyond which there is again only one stable equilibrium for V (the original low- H state having disappeared, discontinuously). For stocking densities in the intermediate region, $T_1 < H < T_2$, the vegetation will tend

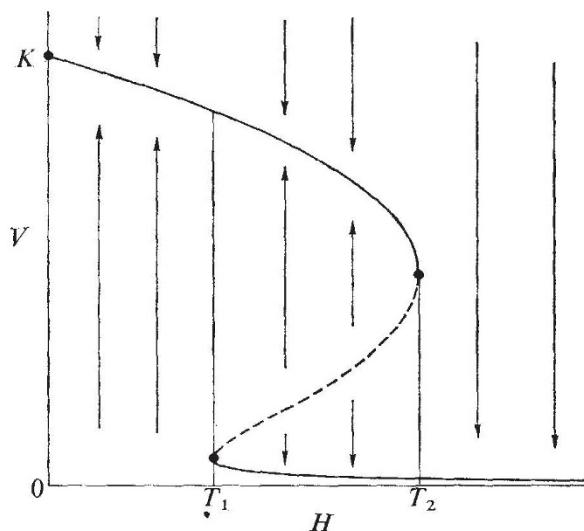


Fig. 3 The equilibrium values of the vegetation biomass, V , are shown as a function of the stocking density, H . For a fixed value of H below the lower threshold at T_1 , or above the upper threshold at T_2 , there is a unique equilibrium value of V ; any initial V value will move to this equilibrium, as indicated by the arrows. For H between T_1 and T_2 , there are two alternative equilibria for V : as shown by the arrows, the system will move to the upper or lower equilibrium, depending on whether the initial value of V lies above or below the dashed 'breakpoint' curve. (This curve is constructed from Fig. 2 and equations (3) and (4), with $\alpha = 0.1$.)

toward either the upper or the lower equilibrium value, depending on the initial value of V : for initial values lying above the dashed line in Fig. 3, V will move toward the upper equilibrium; initial values below the dashed line will move to the lower equilibrium. Borrowing an epidemiological term²⁹, we may call the dashed line the locus of the "breakpoint" values of V .

Although Fig. 3 was constructed from the specific equation (3), its important features are generic to any grazing ecosystem where the dynamics of the vegetation biomass is described by something like Figs 1 or 2. Even in the specially simple case where the $G(V)$ of Fig. 2, with $G(0) = 0$, is combined with the $C(V)$ of Fig. 1, with a linear dependence on V for small V , we still obtain Fig. 3, except that now the lower equilibrium value (occurring for $H > T_1$) is $V = 0$.

If it is assumed that gross animal production, P_G , is linearly proportional to total consumption, $C(V)$, then the equilibrium level of P_G as a function of stocking density H can be read off from Fig. 3. This is done in Fig. 4. Fig. 4 has all the properties just adumbrated for Fig. 3, and is similarly generic.

Noy-Meir first points out some management morals implicit in all this, and then discusses the extent to which such theoretical insights accord with known facts.

If one has a grazing system capable of manifesting the discontinuities illustrated in Figs 3 and 4, then the vegetation will tend always to recover to a high level following an environmental disturbance only if H is kept below some threshold density, T_1 . But, as can be seen from Fig. 4, this usually implies unacceptably low gross productivity. Conversely, for H above T_1 , there is the danger that an environmental fluctuation will carry V below the breakpoint value (the dashed line), whereupon the system will move into the alternative steady state, with dramatically lower values of V and P_G . The closer the system is pushed toward the point of maximum productivity (at $H = T_2$), the more likely is this discontinuous collapse. And of course if the stocking rate for maximum productivity is misjudged, so that H is pushed beyond T_2 , then the system must collapse to the lower equilibrium state. As Noy-Meir^{2,3} sums it up, "a discontinuously stable system is highly labile at the stocking rate which allows the highest productivity. Such a grazing system can be maintained at or near this maximum production rate only by very frequent, almost constant, adjust-

ments of stock density in response to fluctuations in vegetation. A somewhat lower stocking rate will ensure more stable, though on the average somewhat lower, production."

Noy-Meir reviews two classes of empirical evidence.

For intensive, or pasture, systems he pulls together data on the growth curves as a function of green biomass, $G(V)$, for ryegrass-clover in New Zealand^{30,31} and for clover in Australia³², and on the consumption curves for sheep, $C(V)$, for *Phalaris*-clover³³, for *Phalaris*-annuals-clover³⁴, and for ryegrass-clover³⁵, all in Australia. He concludes that the sheep tested by Arnold and by Willoughby would be capable of discontinuous stability properties in October (down under) in pastures of ryegrass, ryegrass-clover, and possibly in clover. There is also direct evidence for two alternate stable states in an experiment done by Morley³⁶, which involves sheep at three different stocking rates in two grazing systems (continuous and with rotation), with three replicates of each.

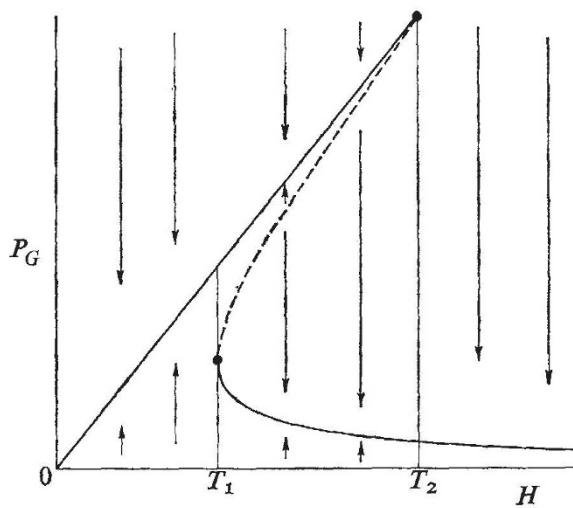
For extensive, or range, systems, indirect evidence can be culled from traditional management practices. Noy-Meir lists the conventional distinction between 'safe' and 'maximum' stocking rates, the notion of range readiness (allowing the vegetation to build to some pre-determined level before introducing animals), and the observation that productivity and animal condition may remain high even though the range is on the verge of collapse. This 'conventional wisdom' is readily justified by Figs 3 and 4.

Throughout this discussion, the herbivore density H has been treated as a constant. In natural (as opposed to managed) situations, H itself will be a dynamical variable, whose rate of change depends on V and H . The consequent pair of differential equations for $H(t)$ and $V(t)$ have been discussed by many people: see the review by Caughley²⁴. The system may have a unique stable point, or two (or even more) alternate stable points, or the populations may oscillate in a stable limit cycle. This subject will be pursued further, when we come to the budworm model, below.

Harvesting animal populations

In grazing ecosystems, we considered the dynamical behaviour of a population (the 'vegetation', $V(t)$) which was being harvested by herbivores. Much of the discussion can be translated to apply to other systems where a plant or animal population is harvested, either directly or indirectly, by man. In particular, fisheries provide an example where one is interested in the dynamics of a fish population, and where the net population growth rate involves a natural growth term, $G(V)$, and a loss rate due to harvesting, $C(V)$, which in general depends both on the fishing effort and on the fish

Fig. 4 Gross animal productivity, P_G , is shown as a function of H . This curve is derived from Fig. 3, under the assumption that P_G is proportional to the total consumption, $C(V)$, of equation (1). The main features are as in Fig. 3; for further discussion, see the text.



population density. The same is true of whaling industries, and of commercial foresteries.

Brauer and Sanchez³⁷ have considered a fish or other population that is harvested to give a constant yield. Their analysis corresponds to Fig. 2 in the limit $\alpha \rightarrow 0$; that is, they have the $G(V)$ curve of Fig. 2, but their consumption or harvesting curves are purely horizontal lines, already saturated at $V = 0$. The resulting equilibrium fish population (V) as a function of yield (H) is a limiting version of Fig. 3, with T_1 at the origin, and the lower equilibrium curve collapsed to the H axis ($V = 0$).

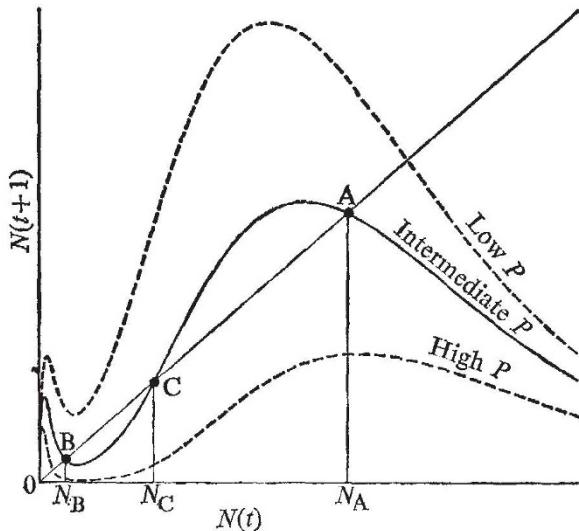


Fig. 5 This figure is a difference equation version of Fig. 3, for animals with non-overlapping generations. The curves relating the population magnitudes in successive generations are shown for various levels of predation. P . Possible equilibrium points occur where the curves intersect the 45° line. (Specifically, the difference equation used here is $N(t+1) = N(t) \exp[r(1 - N - PN/(x^2 + N^2))]$, with $x = 0.1$, $r = 4$, and $P = 0.35, 0.22, 0.15$.)

A more realistic harvesting curve will acknowledge that, at low fish population values, the effort needed to keep the yield constant is prohibitively high. Consequently, a realistic curve for attempted 'constant yield' harvesting will be more like those in Fig. 1 and 2, and much of the discussion of the system's dynamics (ref. 38 and J. R. Beddington, to be published) parallels that by Noy-Meir. Beddington and May's³⁸ study is focused on the response of the system to environmental fluctuations, and to this end employs differential equations with stochastic coefficients. Nevertheless, the qualitative conclusions of this relatively sophisticated study (especially the caution against trying to maximise sustained yield under a strategy of constant quotas, which is roughly equivalent to operating around T_2 in Fig. 4) are laid bare by the simple and deterministic graphical analysis outlined above.

Other recent studies of harvested systems which can have alternate stable states are due to Goh³⁹ and to Huberman⁴⁰ (who uses equation (3)). Clark⁴¹ has given a fine review.

Insect pests: general remarks

We now alter our viewpoint, to think of $V(t)$ more generally as a 'prey' population. Its dynamical behaviour depends on the trophic level above it (the 'predators', which were the herbivores H in the grazing systems) and on the trophic level below it (the resources, which were the light and nutrients determining K in the grazing systems). At the risk of confusing the reader, I shall symbolise this change of viewpoint by using $N(t)$, to replace $V(t)$, for the prey population, and by using P , to replace H , for the predators.

In dealing with insect pests, our attention has shifted up one rung in the trophic ladder. Instead of the vegetation $V(t)$ which

depends on a variable H and a constant K , we have a population $N(t)$ of herbivorous insects which depends on predators P and vegetational resources K , either of which may vary.

For a given value of the environmental carrying capacity (K) and for a specified density of predators (P , formerly H), the considerations that determine the dynamics of the prey population (N , formerly V) are likely to be as in Fig. 2. That is, the growth rate of the prey population may be typified as logistic, and the predators' attack rate will have the general form shown in Fig. 2 both for vertebrate²⁵ and for many invertebrate²⁶⁻²⁸ predators.

One final complication must be disposed of. For the majority of temperate zone insects, population growth is a seasonal (often annual), rather than a continuous, affair. Thus the continuous rate of population change, dN/dt , should usually be replaced by discrete changes, $N(t+1) - N(t)$, at time intervals one unit (often one year) apart. The relation between the population values in successive generations will still typically be determined by Fig. 2, but with $N(t+1) - N(t)$ replacing dN/dt on the y -axis. It is more traditional, in this case, to plot the curve that relates $N(t+1)$ to $N(t)$. This is done in Fig. 5. The 45° line represents population values that are unchanging from one generation to the next, and therefore the points where the curves intersect this line are possible equilibrium states. Where the solid curve lies above the dashed curve in Fig. 2, the net growth rate is positive, and the curve lies above the 45° line in Fig. 5. Conversely, where the solid curve lies below the dashed one in Fig. 2, the curve lies below the 45° line in Fig. 5.

When predation is relatively unimportant (low H in Fig. 2, low P in Fig. 5), there is a unique equilibrium point, at a population value slightly below K ; the population level is set primarily by resources. When predation is relatively important (high H in Fig. 2, high P in Fig. 5), there is again a unique equilibrium point, this time at a low population level; the population is predator-controlled. But for intermediate levels of predation there are, as before, two attracting equilibrium points (A and B in Fig. 5), divided by a repelling point (C in Fig. 5).

The consequences of this intermediate level of predation, with its two alternative equilibria for the pest population, have been noted by Takahashi⁴² and Watt⁴³. They observe that such pests may usually fluctuate at low numbers, around the lower equilibrium point at N_B . But if fluctuations in the number of predators, or of pests, or of the carrying capacity, happen to carry pest numbers above N_C in any one year, then the population will explode towards the upper equilibrium point at N_A . A subsequent crash is likely (the complications that make the dynamics of difference equations more unsteady than those of differential equations are discussed in the conclusion), with the population returning to the neighbourhood of N_B . The pest population may thus exhibit a periodic or episodic pattern of "outbreaks", followed by relatively long intervals at low densities.

Southwood⁴⁴⁻⁴⁶ has recently given a more quantitative discussion, and he and others have incorporated the practical conclusions into a morphology of strategies appropriate for the control of various kinds of insect pests⁴⁷⁻⁵⁰.

For example, Southwood and Comins⁴⁴ have interpreted Clark's thorough studies of the eucalyptus psyllid (a plant louse) *Cardiaspina albitextura* in the light of Fig. 5. Using Clark's data for egg-to-adult survival for 29 generations, along with information about the depression of natality by direct crowding and by host plant deterioration, they estimate the shape of the $N(t+1)$ -versus- $N(t)$ curve. They find the curve to be of 'intermediate P ' type, and assign to N_B , N_C and N_A values corresponding to 2.8, 22 and 107 eggs per shoot, respectively. Not only does this give a qualitative explanation of the observed episodic outbreaks, but it is in remarkable quantitative agreement: Clark's field data give an average of 3.0 eggs per shoot at low densities, suggest a "rapid increase to outbreak level" beyond 10-15 e/s, and have a mean for 12 high density populations of 110 e/s (with a range 23-280).

Sufficient data have also been published on the European spruce sawfly, *Diprion hercyniae*, in New Brunswick for a tentative analysis of this kind. Southwood⁴⁷ suggests the present situation is one of 'intermediate P ', with the sawfly controlled by a com-

bination of virus disease and parasitoids: life-table estimates and field observations agree on a value of N_c around 0.2 larvae m^{-2} ; fluctuations to higher densities result in runaway to around $N_A \sim 1$ larvae m^{-2} , which is economically acceptable. Before the advent of the virus disease, the sawfly in New Brunswick seems to have been^{4,7} in a "low P" situation, with outbreaks to densities around 20 larvae m^{-2} followed by crashes.

Insect pests: the spruce budworm

One of the best studied of all insect pests is the spruce budworm, *Choristoneura fumiferana*, in Canada. This forest defoliator erupts at approximately 40-year intervals, causing much damage in northern coniferous forests, and economic stress in the lumber industry. The massive amount of data gathered by Morris^{5,1} and his associates has been subjected to much analysis by Holling and his collaborators^{1,5,2} (and D. Ludwig, D. Jones and C. S. Holling, to be published) at UBC and at the International Institute for Applied Systems Analysis (IIASA). One of the most pleasing features is the way the models have become progressively simpler as the basic mechanisms have become better understood. Thus the earlier "systems models" have given way to Ludwig *et al.*'s^{5,3} 3-component model (budworms; average leaf area per tree; energy reserve determining the condition of trees and foliage); the present review gives a crude 2-component model (budworms; leaf area) that retains the essentials.

Consider first the dynamics of the budworm population, $N(t)$. Yet again, this is plausibly described by Fig. 2 (with N replacing V), with the net population growth rate counterpoised between a natural growth term and losses due to predators. The general features described by Fig. 2 may be typified by the specific equation (3), which now reads

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) - \frac{\beta PN^2}{N_0^2 + N^2} \quad (5)$$

Here r is the intrinsic growth rate for budworms, and K is the carrying capacity, which depends on the average leaf area per tree, S ; we write $K = \kappa S$. N_0 is the characteristic budworm population at which the predator attack rate saturates (to the constant level β per predator). This characteristic value for predator switching will usually depend on budworm density per unit leaf area rather than on absolute budworm numbers, so that we may write $N_0 = \eta S$. Finally, for a given value of average leaf area S , it is convenient to use the rescaled variables $X = N/\kappa S$, $\tau = rt$ and $\gamma = \beta P/r\kappa S$ to rewrite equation (5) in dimensionless form as

$$\frac{dX}{d\tau} = X(1 - X) - \gamma X^2 / (\alpha^2 + X^2) \quad (6)$$

Here α is formally defined as $\alpha = \eta/\kappa$, but biologically it retains the meaning it had in equations (2) and (4): for any given value of S , α is the ratio between the budworm density that saturates the predator attack capacity and the maximum budworm density that the vegetation can sustain. Equations (4) and (6) are identical.

As before, it is clear, either from the general Fig. 2 (which is drawn for small α) or from the specific equation (6) (provided $\alpha < 1/3$, 3), that the budworm system may have two alternative stable states. If we fix the resource level S , and plot the stable budworm density N as a function of the number of predators P , we get exactly Fig. 3 (with N replacing V , and P replacing H).

It is more interesting, however, to fix the predator level P and consider how the equilibrium value(s) of N varies with S . This can be done graphically from Fig. 2, or algebraically from equation (5). The result is Fig. 6. A simpler way to arrive at Fig. 6 from Fig. 3 is to notice, from equation (6), that the equilibrium value(s) of N/S depends only on γ , once α is fixed. But $\gamma = \beta P/r\kappa S$, so that S plays a part exactly like $1/P$: small P is equivalent to large S , and vice versa, whence Fig. 6 (with S for the x axis) follows from Fig. 3 (with P or H for the x axis). Moreover this makes biological sense. For budworms, the good life is few predators or lots of food; low P or high S . The bad life is the converse.

The main features of Fig. 6 need no elaboration. At low values of

S , there is a unique, and low, budworm density; the system is under predator control. At T_2 a second stable state appears discontinuously, and, for $T_2 < S < T_1$, N will tend to one or other of the two stable states, depending on which side of the 'breakpoint' locus it starts from. For S above T_1 there is again a unique stable state, in which the predators play little part.

This discussion of the budworm dynamics is not sufficient for a full elucidation of the way the interactive system of budworms and forest foliage behaves. We need also to consider how the average leaf area per tree, S , depends on N .

The essentials of such a discussion can be carried out^{5,3}, independent of any further details, by noting that the characteristic time scales for changes in N and in S are very different. The time scale for budworm population growth is months, or even weeks. The basic time scale for change in the average leaf area per tree depends on average branch area, and thence on the time scale for tree growth, which is typically measured in decades. Thus N changes on a relatively fast time scale, S on a slow one.

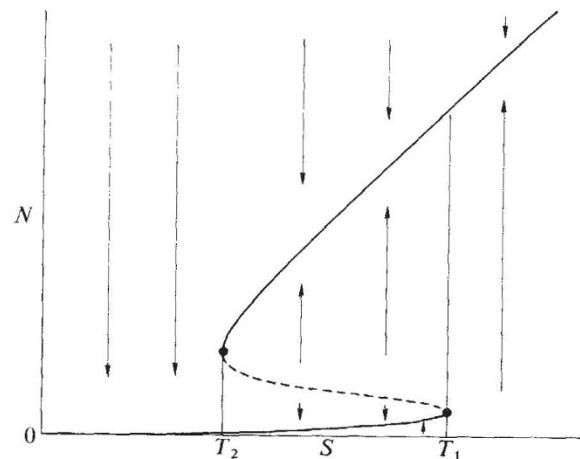


Fig. 6 The equilibrium budworm population, N , shown as a function of the average leaf area per tree, S . (This figure is based on equation (6), with $\alpha = 0.06$.) The threshold and breakpoint features are similar to those in Fig. 3, and are discussed more fully in the text.

Returning to Fig. 6, we further suppose that budworm populations along the lower, 'predator-controlled' equilibrium curve do not have a significant impact on the foliage, so that S undergoes slow natural growth in this regime. Conversely, we assume that the large budworm populations along the upper equilibrium curve have a significant adverse effect on S , decreasing it. The qualitative nature of the system's dynamical behaviour now follows. Starting from any low values of N and S , the system will move rapidly (on the fast time scale) on to the lower budworm equilibrium curve in Fig. 6. S will now slowly increase, with N always quickly adjusted to the current S value, and the system will move (on the slow time scale) to the right along the lower equilibrium curve. When the system arrives at the point T_1 , no further continuous change is possible, and N jumps (on the fast time scale) to the upper equilibrium branch. Now S decreases because of the uncontrolled depredations of the budworms, and the system moves (fairly fast) to the left on the upper equilibrium curve. Finally, when the point T_2 is reached, N must fall (again on the fast time scale) back to the lower equilibrium branch, and the cycle begins again.

The upshot is a cyclic pattern of budworm explosions and crashes, as shown in Fig. 7. The system spends most of the cycle at low budworm densities, traversing the lower equilibrium curve from T_2 to T_1 . The cycle length is thus set by the slow time scale, and is of the order of decades.

A more explicitly mathematical treatment can be given crudely by assuming that the rate of change of S is given by a natural logistic growth term (with an intrinsic growth rate ρ), offset by

losses linearly proportional to the budworm density

$$\frac{dS}{dt} = \rho S(1 - S/S_{max}) - cN. \quad (7)$$

Under the assumptions discussed above, the locus of equilibrium values of S (the points where $dS/dt = 0$) are as indicated by the dashed line in Fig. 7. The pair of equations (5) and (7) then lead to a stable limit cycle, as shown in Fig. 7. The limit $\rho \ll r$ corresponds to the foregoing distinction between slow and fast time scales, and leads to a limit cycle with the almost discontinuous character that was discussed above.

The work of Ludwig *et al.*⁵² is more detailed (involving three

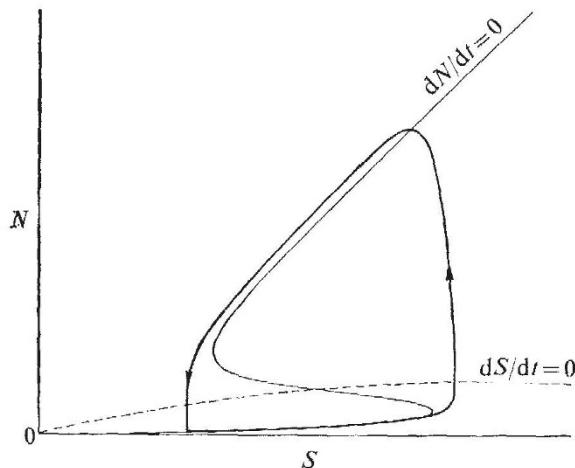


Fig. 7 The (solid, $dN/dt = 0$) equilibrium curve for budworms, N , as a function of foliage, S , is as in Fig. 6. The (dashed, $dS/dt = 0$) equilibrium curve for S as a function of N has the form of equation (7). Under these circumstances, the system will settle to oscillate in a stable limit cycle, as indicated. The figure is for $\rho = 0.01r$, so that the budworm growth time scale is fast compared with the time scale for tree growth; as explained more fully in the text, this explains the main features of the cycle.

differential equations), and is compared with field data. They obtain numerical agreement with the observed 40-year cycle.

This model has implications for the management of budworm populations. Notice first that a program aimed at stopping budworm outbreaks has a propensity to hold the population perpetually poised on the brink of explosion, around the threshold point T_1 on the lower equilibrium curve in Fig. 6. More quantitatively, the effects of using insecticide against budworms can be mimicked by adding an extra mortality term, $-sN$, to the right hand side of equation (5). If we could maintain $s > r$ for many years, the budworms could in principle be eradicated. In the more likely event that $r > s$, the new version of equation (5) can again be brought into the dimensionless form of equation (6), with α replaced by $\alpha' = (r/r - s)\alpha$. We recall that the kinkiness in Figs 3, 4 and 6, and the ensuing alternative stable states, depends on α being small. Thus use of insecticides can increase the effective value of α to such an extent as to straighten out the kink in Fig. 6 (that is, can increase α' to a value $\alpha' > 1/(3\sqrt{3})$), leading to a unique stable state. The price, however, is daunting: not only are we committed to endless insecticide use, but the new equilibrium is necessarily established at a budworm density somewhere intermediate between the original upper and lower equilibrium levels. Such a budworm density represents a perpetual low-level outbreak.

It should be added that this discussion has ignored climatic fluctuations. One argument in favour of trying to 'hold the line' with insecticides is that one may thereby buy time, waiting for a year when unfavourable weather will carry the population back down to very low values.

Human host-parasite systems

Similar threshold and breakpoint phenomena occur in the transmission dynamics of many infections of humans and other

animals, particularly when the transmission cycle involves intermediate vectors.

Malaria is probably the best-known example exhibiting a threshold^{54,55}. Suppose that a single malarious mosquito is introduced into a closed and previously malaria-free community of people and mosquitoes. Each person bitten by this mosquito may, with a probability that can be estimated, become victim to the disease. These infected people may in turn be bitten by uninfected mosquitoes, which thereby (again with some probability factor, and after the elapse of a latent period of around 10 days) are recruited as malaria vectors. The key question is whether the original infectious mosquito produces, on the average, more than one subsequently infectious mosquito, or not. If it does, the system is 'above threshold', and the fraction of the human and of the mosquito populations that have malaria at any one time will grow until an equilibrium is reached, at which new infections are balanced against recoveries and deaths. Conversely, if it does not, the system is 'below threshold', and the introduced infection cannot be maintained.

This discussion makes clear the underlying nature of the threshold relation for malaria. The relation can be written explicitly in terms of the number of mosquitoes, biting rates, and recovery and death rates of mosquitoes and people: as reviewed by Macdonald^{54,55} and Conway⁵⁶, this threshold formula has many implications for management.

The main fact, however, is that for simple models of the malaria transmission process there is always a unique stable state: for mosquito densities above threshold, there is some endemic level of infection; below threshold the level is zero. For many helminthic infections the situation is made more complicated by the parasites' having a sexual stage in the human host. Thus, for example, an adult female schistosome will produce eggs only if she has a mate, which may be unlikely if the mean worm load per person is less than, or of the order of, unity.

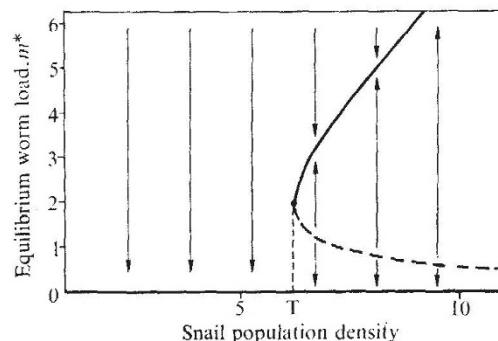


Fig. 8 For a simple model of the transmission dynamics of schistosomiasis, the equilibrium mean worm load per person, m^* , is shown as a function of the snail population density (or other relevant transmission factor). The threshold and breakpoint phenomena are as discussed in the text. (From Bradley & May⁵⁹.)

As first stressed by Macdonald^{29,55}, the result is that for schistosomiasis and other parasitic infections with a sexual stage in the primary host there is, as before, a threshold transmission value which depends on the population density of the snails or other intermediate vector, on the rate of egg production in the primary host, and on other transmission parameters. Below threshold, there is only one equilibrium state, namely no infection. Above threshold, there are two alternative equilibrium states: if the initial mean worm load per person, m , is sufficiently high, the infection cycle is maintained; but if m is initially too small, the number of mated females (which will scale as m^2) may be insufficient for the system to 'take off' and attain its endemic equilibrium level. This situation is illustrated in Fig. 8, which corresponds to simple models for schistosomiasis^{29,56-58}. If the snail density (or other relevant transmission factor) is below the threshold value T , the

equilibrium level of adult schistosomes is zero. Above threshold, the mean worm load per person moves either to the upper equilibrium curve, or to zero, depending on whether its initial value lies above or below the "breakpoint" line.

Figure 8 is for a closed system, with no immigration of infectious snails or people from outside. If a small amount of such immigration is included (I. Nasell, to be published) the lower equilibrium curve in Fig. 8 must shift up to some small, but finite, level of infection, and a figure similar to Fig. 6 is obtained (with m replacing N , and snail density or the like replacing S).

The existence of a breakpoint holds important implications for control of the system^{29,55-59}. For diseases such as malaria, where there is always a unique equilibrium state, the only control strategy is to reduce the mosquito numbers and/or the biting rate and other parameters in such a combination as to take the system below threshold. The system must, moreover, subsequently be kept below threshold. But infections whose dynamics are as described by Fig. 8 can in principle be eradicated by temporarily displacing the mean parasite load below the breakpoint, thus taking advantage of the two alternative equilibrium states. This control strategy has the advantage that the basic transmission parameters do not have to be modified to, and maintained at, sub-threshold values. (If the infection dynamics of snails operates on a fast time scale, and of humans on a slow time scale, the locus of the breakpoint depends only on the level of infection among humans; more generally, the breakpoint will depend on both snail and human infection levels⁵⁶⁻⁵⁸. It also depends on whether the adult parasites are distributed among humans in random or clumped fashion^{58,59}.)

Current estimates are that about 200 million people have schistosomiasis, and about 300 million suffer from the various filarial infections (which include 'river blindness' and elephantiasis). But empirical information on the overall transmission dynamics of schistosomiasis and filariasis is scant, so that theoretical notions about threshold and breakpoints remain untested in this context.

Conclusion

This review has drawn together a variety of mathematical models for ecological systems, each of which is dealt with individually in the scattered literature. The models are united by the common theme of possessing a regime of dynamical behaviour in which there are two alternative stable states, so that continuous variation in a control variable can produce discontinuous effects. Thus smooth changes in stocking rates can cause discontinuous changes in the grazed vegetation; continuous changes in harvesting rates can cause discontinuous collapse in fisheries; continuous changes in environmental parameters or foliage growth or predation rates can lead to discontinuous outbreaks of insect pests; continuous changes in snail or dipteran population densities can cause discontinuous appearance or disappearance of helminthic infections.

These models can often be tied to empirical data, whence they yield specific insights into the management of the system. (Some of these insights can be recast in the language of catastrophe theory, but this is usually *post hoc* window-dressing. It is too often forgotten that catastrophe theory is, strictly speaking, a local theory; we want global descriptions of the dynamics.)

The review has been largely confined to the dynamical behaviour of a single population, albeit as a function of other populations of resources and predators. As the number of interacting dynamical variables, and of equations, increases things get more complicated. Consider, for example, a predator-prey system where the predator's dynamics depends on the prey population according to the backward-bending curve of Fig. 6 (N versus S), and where the prey's dynamics in turn depends on the predator population according to Fig. 3 (V versus H). These two curves—one for the predator equilibrium, the other for the prey equilibrium—are in principle capable of intersecting at 9 points, corresponding to no fewer than four alternative stable states (separated by four saddle points and one central hilltop), each with its own domain of attraction. As the dimensionality of the system increases to encompass more and more species, the dynamical

landscape can begin to look like the surface of the moon, and any detailed comparison with field or laboratory data becomes very difficult.

A different and additional complication is that systems with discrete generations are described by difference equations, which can exhibit kaleidoscopic dynamics. For such systems, of which Fig. 5 is an example, increasingly severe nonlinearities can make the dynamical behaviour range from a stable point, through a bifurcating hierarchy of stable cycles, into a regime which is in many ways indistinguishable from random noise⁶⁰⁻⁶². Similarly, patterns of stable points or cycles giving way to chaos can be found in systems described by differential equations with time lags, or even in simple first-order ordinary differential equations once 3 or more species are interacting^{60,62,63}.

These manifold complications can lead one to take a gloomy view of the possibility of making predictions about multi-species systems^{60,62}. The one kindly light amidst this encircling gloom is that many complex communities are, arguably, made up mainly of loosely coupled lower-order systems^{17,64,65}, to which the simple models reviewed here are pertinent.

I thank Henry S. Horn for helpful conversations, and the US NSF for support (grant DEB 75-10464).

- ¹ Monod, J. *Chance and Necessity: an Essay on the Natural Philosophy of Modern Biology* (Knopf, New York, 1971).
- ² Weinberg, S. *The First Three Minutes: a Modern View of the Origin of the Universe* (Basic Books, New York, 1977).
- ³ Hawking, S. W. *Sci. Am.* **236**, 34-40 (1977).
- ⁴ Layzer, D. *Sci. Am.* **233**, 56-69 (1975).
- ⁵ Blatt, J. M. *Theor. Phys.* **22**, 745-756 (1959).
- ⁶ Simpson, G. G. *The Major Features of Evolution* (Columbia University Press, New York, 1953).
- ⁷ Raup, D. M. *Am. Sci.* **65**, 50-57 (1977).
- ⁸ MacArthur, R. H. *Geographical Ecology* (Harper and Row, New York, 1972).
- ⁹ May, R. M. in *Progress in Theoretical Biology* (eds Rosen, R. & Snell, F.), 1-50 (Academic, New York, 1974).
- ¹⁰ Sutherland, J. P. *Am. Nat.* **108**, 859-873 (1974).
- ¹¹ Levin, S. A. *Adv. Ecol. Syst.* **7**, 287-310 (1976).
- ¹² Connell, J. H. & Slatyer, R. O. *Nat.* (in the press).
- ¹³ Austin, M. P. & Cook, B. G. *J. theor. Biol.* **45**, 435-458 (1974).
- ¹⁴ Case, T. J. & Gilpin, M. E. *Nature* **261**, 40-42 (1976).
- ¹⁵ Holling, C. S. *Adv. ecol. Syst.* **4**, 1-23 (1973).
- ¹⁶ Lewontin, R. C. in *Diversity and Stability in Ecological Systems* 13-24 (US National Bureau of Standards, Springfield, Virginia, 1969).
- ¹⁷ May, R. M. *Stability and Complexity in Model Ecosystems* 2nd ed. (Princeton University Press, Princeton, 1974).
- ¹⁸ Orians, G. H. in *Unifying Concepts in Ecology* (eds van Dobben, W. H. & Lowe-McConnell, R. H.), 139-150 (W. Junk, The Hague, 1975).
- ¹⁹ Margalef, R. in *Unifying Concepts in Ecology* (op. cit.) 151-160.
- ²⁰ May, R. M. in *Unifying Concepts in Ecology* (op. cit.) 161-168.
- ²¹ Whittaker, R. H. in *Unifying Concepts in Ecology* (op. cit.) 169-181.
- ²² Grimm, H. R. *HASA Publication RR-76-005* (2361 Laxenburg, Austria, 1976).
- ²³ Noy-Meir, I. *J. Ecol.* **63**, 459-481 (1975).
- ²⁴ Caughey, G. in *Theoretical Ecology: Principles and Applications* (ed. May, R. M.) 94-113 (Blackwell, Oxford, 1976).
- ²⁵ Holling, C. S. *Mem. ent. Soc. Can.* **45**, 3-60 (1965).
- ²⁶ Hassell, M. P., Lawton, J. H. & Beddington, J. R. *J. Anim. Ecol.* **46**, 249-262 (1977).
- ²⁷ Murdoch, W. W. & Oaten, A. *Adv. ecol. Res.* **9**, 1-131 (1975).
- ²⁸ Hassell, M. P. *Arthropod Predator-Prey Systems* (Princeton University Press, Princeton, in the press).
- ²⁹ Macdonald, G. *Trans. R. Soc. trop. Med. Hyg.* **59**, 489-506 (1965).
- ³⁰ Brougham, R. W. *Aust. J. agric. Res.* **6**, 804-812 (1955).
- ³¹ Brougham, R. W. *N. Z. J. Sci. Technol.* **A38**, 78-87 (1956).
- ³² Davidson, J. L. & Donald, C. M. *Aust. J. agric. Res.* **9**, 53-72 (1958).
- ³³ Willoughby, W. M. *Aust. J. agric. Res.* **10**, 248-268 (1959).
- ³⁴ Arnold, G. W. in *Grazing in Terrestrial and Marine Environments* (ed. Crisp, D. J.) 133-154 (Blackwell, Oxford, 1963).
- ³⁵ Allden, W. G. *Proc. Aust. Soc. Anim. Prod.* **4**, 163-166 (1962).
- ³⁶ Morley, F. H. W. *Proc. N. Z. Soc. Anim. Prod.* **26**, 8-21 (1966).
- ³⁷ Brauer, F. & Sanchez, D. A. *Theor. Pop. Biol.* **12**, 31 (1975).
- ³⁸ Beddington, J. R. & May, R. M. *Science* **197**, 463 (1977).
- ³⁹ Goh, B. S. *Math. Biosci.* **33**, 359 (1977).
- ⁴⁰ Huberman, G. *Theor. Pop. Biol.* (in the press).
- ⁴¹ Clark, C. W. *Mathematical Biogeconomics* (Wiley, New York, 1976).
- ⁴² Takahashi, F. *Rep. Pop. Ecol.* **6**, 29-36 (1964).
- ⁴³ Watt, K. E. F. *Ecology and Resource Management* (McGraw-Hill, New York, 1968).
- ⁴⁴ Southwood, T. R. E. & Comins, H. N. J. *Anim. Ecol.* **45**, 949-966 (1976).
- ⁴⁵ Southwood, T. R. E. in *Proceedings of XI International Congress of Entomology* (ed. Packer, J. S.) 36-51 (Entomological Society of America, College Park, Maryland, 1977).
- ⁴⁶ Southwood, T. R. E. *Am. Sci.* **65**, 30-39 (1977).
- ⁴⁷ Southwood, T. R. E. in *The Origins of Pest, Parasite, Disease and Weed Problems* (eds Cherrett, J. M. & Sagar, G. R.) (Blackwell, Oxford, 1977).
- ⁴⁸ May, R. M. *Nature* **264**, 211-212 (1976).
- ⁴⁹ Conway, G. R. in *Theoretical Ecology: Principles and Applications* (op. cit.) 257-281.
- ⁵⁰ Conway, G. R. *Nature* (in the press).
- ⁵¹ Morris, R. Mem. ent. Soc. Can. **31**, 1-332 (1963).
- ⁵² Holling, C. S., Dantzig, G., Baskerville, C., Jones, D. & Clark, W. *Proc. Int. (Canadian) Conference on Applied Systems Analysis*, 1975 (in the press).
- ⁵³ Levandowsky, M. & White, B. S. *Evol. Biol.* (in the press).
- ⁵⁴ Macdonald, G. *The Epidemiology and Control of Malaria* (Oxford University Press, London, 1957).
- ⁵⁵ Macdonald, G. *Dynamics of Tropical Disease* (Oxford University Press, London, 1973).
- ⁵⁶ Nasell, I. & Hirsch, W. M. *Comm. pure appl. Math.* **26**, 395-453 (1973).
- ⁵⁷ Nasell, I. *Theor. Pop. Biol.* **10**, 133-144 (1976).
- ⁵⁸ May, R. M. *Math. Biosci.* (in the press).
- ⁵⁹ Bradley, D. J. & May, R. M. *Trans. R. Soc. trop. Med. Hyg.* (in the press).
- ⁶⁰ May, R. M. & Oster, G. F. *Am. Nat.* **110**, 573-599 (1976).
- ⁶¹ May, R. M. *Nature* **261**, 459-467 (1976).
- ⁶² Guckenheimer, J., Oster, G. F. & Ipakchi, A. *Theor. Pop. Biol.* (in the press).
- ⁶³ Rossler, O. E. Z. *Naturforsch.* **31a**, 1664-1670 (1976).
- ⁶⁴ Gilbert, L. E. in *Coevolution of Animals and Plants* (eds Gilbert, L. E. & Raven, P. H.) (Texas University Press, Austin, 1975).
- ⁶⁵ Lawton, J. H. *Nature* **264**, 138-139 (1976).