DYNAMICAL MODELS OF ECOSYSTEMS AND EPIDEMICS

by

S.A. Levin,
Kirk Moloney,
Linda Buttel,
and
Carlos Castillo-Chavez

BU-1015-M

February 1989
Dynamical Models of Ecosystems and Epidemics

by

Simon A. Levin, Kirk Moloney, Linda Buttel and Carlos Castillo-Chavez

(BU-1015-M)

ABSTRACT

This paper discusses problems in ecology and epidemiology that represent challenging and important scientific problems. The emphasis is on ecology and epidemiology of infectious diseases, and the use of supercomputers as a complementary and novel tool to experimentation and classical approaches. A brief discussion to address the concept of a fundamental scale, a necessary step in the development of a landscape model, is illustrated through the use of a model for the serpentine grassland at Jasper Ridge, CA. Further issues related to the proper level of aggregation and scale are addressed through the myxoma-rabbit system and the human immunodeficiency virus-man system.
INTRODUCTION

Taking the charge of the Molokai conference literally, this paper discusses problems in ecology and epidemiology that represent grand scientific challenges. The areas that we will discuss are of fundamental scientific importance; they are also ones where the potential for supercomputing to have major impact is tremendous, but not yet realized.

Supercomputing represents a novel tool that complements experimentation and classical theoretical approaches. It can transform the way ecologists and epidemiologists even frame their questions, freeing them from constraints that have been taken for granted. Such liberation will take time, but the issues are too important to delay initiating the process.

In this paper, we join ecology and epidemiology—more precisely, the epidemiology of infectious diseases—because they represent closely related subjects. Ecology deals with the relationships between organisms and their environment, which includes other such organisms and other species; among the interspecific interactions of concern are competition, predation, and parasitism. The dynamics of infectious diseases represent a special case of the latter: an interaction between parasite or pathogen and host, where the particular infectious agent may be a virus, a bacterium, or other organism. Viewing the dynamics of infectious diseases within an ecological as well as an epidemiological context has provided a new perspective on disease spread, control, and
evolution (Anderson and May 1979, Levin 1983), and has given new impetus to modeling efforts.

Given the parallels discussed in the last paragraph, it is not surprising that efforts at modeling ecological and epidemiological interactions, though historically they developed separately, have followed quite similar paths. The classical fare in both fields has involved dynamical systems of low dimensionality, in large part in deference to what could be done analytically. In ecology, statistical mechanical theories of interacting species have been developed, but have had no real impact on the subject because the necessary simplifying assumptions cannot be justified biologically. Ecological systems are not composed of ensembles of interchangeable particles, and show tremendous heterogeneity and complexity across a range of scales. Analytical approaches have not been capable of dealing with the full range of complexity, although of course they remain critical elements in an integrated approach to studying complex ecosystems. The supercomputer complements these methods, and opens up dramatic new possibilities for understanding the structure and dynamics of ecosystems and epidemics.

ECOLOGICAL SYSTEMS

In the early part of the twentieth century, two views of ecosystems were advanced. One, due to Gleason (1926), argued that plant communities were very loose assemblages, that species were distributed individualistically, and that stochastic factors were of major importance. The contrary view (Clements 1936) argued that the vegetation at a particular location would tend to an asymptotic state, the climax, determined by local climatic and edaphic (soil) conditions and consisting of a single or a small number of competitively dominant species. The latter point of view was also represented in the mathematical theories of Lotka (1925) and Volterra (1931), whose approach was developed to describe animal communities, but was applied more generally. These models, which are taught in introductory ecology courses and continue to provide grist for the mills of mathematicians today, consist of systems of differential equations representing the dynamics of interacting
species, treated as if chemicals in solution. The emphasis is almost always on the asymptotic behavior, which usually involves convergence to a singular point or a limit cycle. The general characteristics of such approaches involve determinism, predictability, homogeneity, and equilibrium, all of which are features that cause the ecologist some concern.

The most influential applications of these equations have been to pairs of interacting species; attempts to expand beyond this have led to some interesting mathematical excursions, but to very little biological success. Given the uncertainty with which the forms of the equations, much less the parameters, are known, the classical problems of analytical complexity, sensitivity to parameters and initial conditions, and error propagation have seriously restricted our ability to use such formalisms to predict the dynamics of ecosystems.

Equally problematical is the fact that the choice of the ecological species as the fundamental unit to use in modelling is arbitrary. In some cases, one must break the species into age classes, developmental stages, genetic types, or social groups. In others, it is appropriate to lump species together into groups that perform similar functions. Indeed, it may be appropriate to do both at the same time, lumping together the juvenile stages of a group of species that show similar behavior at that life stage, and creating different groupings at adult stages. Indeed, any choice of aggregate variable is to some extent arbitrary, suppressing heterogeneity within a group in an attempt to organize individuals into aggregates for which some characteristic behavior can be assumed.

One of the most intriguing aspects of a population's structure is its spatial distribution. In much the same way that a population can be subdivided into age groups, it can be subdivided into local subpopulations; and it is this patchiness of populations, coupled with environmental variability, that accounts for the coexistence of large numbers of species that otherwise could not exist together.
For continuously distributed species, as for example oceanic plankton, continuum descriptions are appropriate; and a considerable literature exists based on the resultant coupled parabolic partial differential equations. Such approaches can be very instructive in treating such problems as the spread of invading species, including those that cause epidemics, and there is a beautiful and applicable mathematical literature on the asymptotic speed of propagation that emerges from such descriptions (Kolmogorov et al. 1936, Bramson 1983, Andow et al., ms.). Nonetheless, such approaches are not easily extended to the consideration of localized random disturbances, which are common features of ecosystems, and for which percolation theory and approaches involving interacting particle systems (Liggett 1985, Zallen 1983, Durrett 1988) may be more appropriate. Furthermore, most ecologists are unable to take data over a continuum of spatial scales simultaneously, and tend to select study sites of fixed (and small) size for reasons that have more to do with logistics and finances than with science.

One of the most exciting challenges facing ecological theorists is to address these and other questions by developing methods for relating processes taking place on different spatial, temporal, and organizational scales. A proper theory of scaling for ecosystems should stimulate experimentation and data collection that address explicitly the manifestation of patterns on multiple scales, and provide the basis for extrapolation from data collected on one scale to patterns on broader or narrower scales. Such a theory also would address two related and emerging problems of substantial applied importance. The rapidly accumulating base of information arising from remote sensing is restricted in resolution by the minimum pixel size that can be used; theories that can interpret that information in terms of processes played out on finer scales are a fundamental need. Furthermore, as attention turns to global climate change and its implications for Earth's ecosystems, various modeling groups have seen the attractiveness of coupling global climate models with models of the dynamics of ecosystems; such approaches would allow one to address not only the effects of a changing climate on ecosystems, but also the
reciprocal effects that altered ecosystems (for example due to deforestation) must have on climate change. Yet the typical scales of resolution in global climate models are orders of magnitude coarser that those used in ecosystem models; and even with massively parallel processing, the potential for refining the grid of the climate models is severely limited. Thus the alternative approach must be developed, of scaling up ecosystem models to provide believable models at the landscape level. This is the ultimate goal of our approach.

A necessary step in developing a landscape model is to define a fundamental scale, that of the primary building block of the larger model. Processes occurring at smaller scales are represented by their average effect on processes at the fundamental scale. For example, one of the systems we are modeling is a serpentine grassland at Jasper Ridge, California, and the fundamental scale we have chosen is a 10-cm by 10-cm area. (We will use the Jasper Ridge model in the following discussion to illustrate our approach.) Of course, the choice of any scale is to some degree arbitrary ignoring heterogeneity within a cell in order to retain a fundamental unit sufficiently large that its dynamics are statistically reasonably predictable. We also are investigating models that account for dynamics occurring at a smaller scale—tracking mortality, reproduction, and competition as a function of spacing at length-scales on the order of millimeters or centimeters in an attempt to find an optimal degree of resolution. In this paper, we ignore the finer details and model demographic processes as a function of population densities within a 10-cm by 10-cm neighborhood.

One of the factors influencing the choice of fundamental scale of the Jasper Ridge model is the size-scale of disturbances affecting the grassland, since it has been established that species distributions at Jasper Ridge are controlled to a large extent by disturbance regime (Hobbs and Mooney 1985); approximately 10% of the landscape is disturbed per year by the actions of small mammals and insects. In order to capture the landscape dynamics of the grassland, we choose the fundamental scale of the model (100 cm²) an
order of magnitude smaller than the size of a typical disturbance (1100 cm² on average; Hobbs and Mooney 1985).

Once the scale of the smallest model unit is set, the second step in our approach is to build a landscape level model by constructing a $10^4$ cell grid incorporating dispersal, competition, and spatially correlated extrinsic factors. The dynamics portrayed at the landscape level are generated by copies of a basic demographic model running in each fundamental unit, or grid cell, of the landscape (Fig. 1). The model in its simplest form can be run efficiently on a microcomputer; but as greater detail is added, it is necessary to use the facilities of the Cornell National Supercomputer Facility (CNSF). This is particularly true when we wish to explore parameter space in detail. Because of these considerations, the model is being coded in FORTRAN; and the same version, with slight modification, can be run on an IBM clone, the CNSF, or a Macintosh. Spatial and temporal statistics applied to model output are more computationally intensive than the model runs themselves and must, in general, be conducted on the CNSF. The structure of the model is designed to capitalize on the parallel capabilities of the CNSF.

Currently, the Jasper Ridge model is built around four plant species, chosen to represent demographic types characteristic of the plant community. Connections between the species occur through competitive interactions, which, in the current version, act to modify seed production as a function of the densities of the four species within each model cell. This represents a major regulatory feature of the demographic model, as it introduces a density dependent dynamic regulating the production of seeds by individual plants. As density within a cell increases, seed production by individual plants declines. The dynamics are such that each species, in the absence of disturbance, would reach an equilibrium population density (or limit cycle) over time (Fig. 2a); in the presence of random disturbances, the stochastic process is stationary and approximately ergodic (depending on boundary conditions).
Landscape level processes in the Jasper Ridge model are produced, in part, through mechanisms that act to couple dynamics among cells across the landscape. These include disturbance, dispersal, and competition. For example, the dispersal of seeds among cells allows recolonization of a disturbed site, or invasion by an introduced species, and acts to connect population dynamics across the grid. However, such coupling is not sufficient to maintain landscape level heterogeneity. Otherwise, synchronization across cells occurs because of the self-damping of the demographic model within each cell; thus, without disturbance the model approaches a stable point or limit cycle at the level of the landscape through time (Fig 2b). Without an outside forcing function, the landscape model simply becomes a large scale realization of the model constructed at the fundamental scale, with dispersal serving simply to homogenize the system across space.

Spatial heterogeneity is introduced by disturbing sites at random within the landscape; Monte Carlo techniques are employed to determine the location of disturbances (Fig. 3). Disturbances are allowed to occur over two distinct time periods in the yearly cycle depicted by the model. The impacts of introducing disturbances are manifold and depend in part on time of occurrence (Fig. 2c) relative to the seasonal dynamics of species. Disturbances can affect local population structure by removing plants from the population before they reproduce (Disturbance 1 in Fig. 1), and can have a large impact on the local abundance of a species, especially since all of the plants being modeled are annuals. In contrast, a disturbance occurring after seed set (Disturbance 2 in Fig. 1) can have either little or no effect on the local population or a large effect, depending on whether the seed bank is destroyed. Differences among species in time of flowering and in dispersal ability under a disturbance regime can act to create substantial differences in distributional patterns observed across the landscape. Another major impact of disturbance is to create a different physical environment within affected cells through a change in moisture and nutrient availabilities (Koide et al. 1985, Mielke 1977). This is modeled as a change in parameter values for rates of germination, survival, and reproduction of plants growing on cells.
affected by disturbances. Again, environmental differences associated with disturbances act to create pattern across the landscape.

Another source of landscape heterogeneity can be introduced by building environmental gradients into the model. This would be similar to modeling a landscape over a range of altitudes or range of soil types, where population processes change in response to a changing environment. Parameters describing demographic processes along a static gradient in the model would then be a function of location as well as of disturbance regime. This provides a mechanism for modeling over several community types and for developing a model at larger landscape scales.

Model outputs are analyzed by evaluating a variety of spatial and temporal pattern descriptors, correlograms and semivariograms, power spectra, fractal dimensions, and nested analyses of variance (see for example, Fig. 4). These may be used to compare model outputs with data taken from the natural system, and used to refine the model. The model itself can then be used as an experimental tool on the computer, on which the parameters of various key processes can be varied to study their importance in defining pattern on various scales. With Richard Durrett, we have begun the consideration of percolation model approximations to our detailed model, in the hope of gaining understanding of the importance of the basic processes. Finally, analogous versions of the model are being implemented for forest systems.

Clearly, it is possible to make a model more and more detailed by adding complexity. However, adding complexity to any model can be dangerous, adding to the dynamical complexity as well. Furthermore, with increase in resolution comes increasing computational needs, as well as problems of parameter estimation. Developing models at larger fundamental scales is one way to simplify the process, both in terms of computation and comprehension; our goal is to develop these descriptions by examining the emergent properties of systems structural at finer scales, and to compare the output of the aggregated
behaviors with observed features of real systems. This is an iterative process, which can be repeated several times.

EPIDEMIOLOGICAL SYSTEMS

Although Daniel Bernoulli presented a mathematical model for the effects of smallpox vaccination as early as 1760, current mathematical models of the spread of epidemics trace back to work of En'ko in 1889 (see Dietz 1988), and to the notion, advanced by Hamer (1906), that the critical process in the dynamics of disease was in the contact rate between susceptibles and infectives (see also Ross 1911, Brownlee 1907, McKendrick 1912). The contact rate describes the essential nonlinearity in what otherwise is basically a linear system. The classical assumption is that the rate of new infections is proportional to the number of infectives and to the number of susceptibles, an assumption that is easy to motivate from first principles when densities are low. It is recognized that this simple relationship is likely to break down at higher densities, but it remains a convenient starting point for most models in use today. A related problem, particularly important in the consideration of populations that fluctuate in size, is whether contact rates should be functions of the actual numbers of susceptibles and infectives in a particular area, or only on their relative proportions with respect to the total population, including recovered, immune, and latent individuals. It is not the purpose of this paper to enter into that debate, which of course can only be resolved with regard to particular diseases and under particular conditions.

Ross (1911), in studying the dynamics of malaria, and Kermack and McKendrick (1927) produced two of the classic papers of the mathematical epidemiological literature. The contribution of Kermack and McKendrick was the demonstration of the existence of a threshold value for the number of susceptibles, such that if fewer were present in a population, a disease could not take hold. Such a threshold is important for a variety of reasons, most importantly because it provides a target for vaccination strategies. If the susceptible population is reduced below the threshold value by removal techniques such as
vaccination, then the disease can be controlled. Of course, computation of the threshold can be problematical and highly model dependent, and particular interest today focuses on how that threshold is to be computed when the population is broken into distinct subgroups. This is one of the foci of our own work.

Another objective involves understanding the periodic and more complicated dynamics observed in the dynamics of many diseases (Hethcote and Levin 1989; Schaffer et al. 1985). Soper (1929) began the investigation of the causes of periodic behavior; but the list of contributory causes continues to expand (Hethcote and Levin 1989).

The availability of supercomputers has expanded greatly our capacity to investigate these issues. In particular, parallel processing techniques are especially suited to the problem of nonhomogeneous mixing, which is a consequence of the fact that populations are subdivided into different risk groups. Within groups, the rates of interaction are typically very different than they are among groups; and furthermore, intergroup interactions are not likely to be homogeneous. For example, in the dynamics of diseases such as influenza, age is a primary consideration, since different age groups have different contact rates, different susceptibilities, and different risks of mortality. For sexually transmitted disease such as AIDS, the adult population is key; but subdivision is necessary with regard to sexual preference and behavior. For AIDS, of course, drug users also represent a risk category that needs to be considered explicitly. In the remainder of this paper, we will discuss briefly models of three viral diseases: myxomatosis, a disease of rabbits; influenza; and AIDS.

Myxomatosis is a disease of rabbits, and has been used as a biological method of controlling the spread of rabbits in Australia and in Europe (Fenner and Ratcliffe 1965). The early success of the viral control soon was dissipated, however, as evolutionary pressures drove the virus towards lower levels of virulence, and the rabbit to increased resistance. Differential equation models of this system captured many of the principal features, but were unable to represent the full complexity: the multiplicity of grades of the
virus, the geographic and seasonal variability, the etiology of the disease, and the
demography of the host population. As such, still unresolved are explanations of the
coeexistence of viral grades in the field, and projections of prospects for future control of the
rabbit population.

Detailed simulations (Dwyer et al. 1989) on the CNSF have demonstrated that even
the interaction between single grades and homogeneous hosts can show complicated
dynamics, including periodic and chaotic behavior. In particular, the natural oscillations
that are characteristic of any host parasite system interact with the seasonal forcing that
derives from host demographic patterns to produce a panoply of dynamic behaviors as
critical parameters are varied. Further investigations have elucidated the interactions among
strains, and the coevolution between virus and host.

Viral evolution takes on a new dimension in the annual ravages of influenza, whose
rapid changes, facilitated by a particularly labile molecular structure, foil host coevolution
and result in the sporadic reappearance of strains long disappeared. Although the molecular
shift and drift involving the surface antigens of the virus are well understood, such
knowledge is insufficient to explain the interepidemic periods. Such periods are
determined by epidemiological characteristics, in particular the time necessary for a
sufficient susceptible population to build back up after an epidemic or pandemic. The
threshold behavior mentioned earlier as a general property of epidemics is key to an
understanding of when new outbreaks can occur. A separate but related problem involves
an understanding of how the seemingly dormant strains of the virus are maintained for the
decades that can pass before the susceptible population reaches threshold.

Our investigations of influenza (Castilla-Chavez et al. 1988, 1989a, Liu and Levin
1989) on the CNSF have incorporated details of the age structure of the host population,
including age specific contact rates, mortalities, and fecundities. It is well appreciated that
the human population does not mix homogeneously; for example, school children tend to
mix with school children, and have a particularly high contact rate and probability of
transmission of the disease. Non-homogeneous mixing of different age groups does not appear capable of leading to sustained oscillations, but can produce oscillations that are very slowly damped. A separate phenomenon that produces similar behavior is the cocirculation of distinct strains of the virus, which show partial but incomplete cross-immunity with one another. When a full age-structured model is considered in a population in which such partially incompatible viral strains are cocirculating, the two weakly damped oscillators are mutually excitatory; the result is that sustained oscillations are maintained. Analytic models have provided deep insights into this problem, but a full understanding could be obtained only when those analyses were carried out in concert with detailed simulations on the CNSF.

Finally, and most recently, we have turned our attention to models of the AIDS epidemic. AIDS, as a primarily sexually transmitted disease, is a disease for which the assumption of homogeneous mixing is a total failure. Detailed models (Castillo-Chavez et al. 1987, 1989b,c,d,e,f) have been developed that subdivide the population into risk groups based on sexual behavior or other characteristics, and show how to compute the threshold for such multi-group models. However, what are needed now are models that incorporate the rapidly accumulating information on human sexual behaviors, and that are tailored to different geographical regions. For these detailed models, the supercomputer is essential.

CONCLUSION

We have presented here only the briefest introduction to the application of the supercomputer to emerging problems in ecology and epidemiology. In the short space of this paper, it is impossible to develop these investigations in detail; for that purpose, the reader is directed to the referenced works. Suffice it to say that the epidemiological and ecological problems facing us in the next decade represent some of the grand challenges for society; this applied importance, and the models' rich and fascinating mathematical behaviors, place them in the elite group of Grand Challenges for Supercomputing.
ACKNOWLEDGMENTS

Support for this research was provided by NSF grant BSR-8806202 and McIntire-Stennis grant NYC-183550 to SAL.

This is also Ecosystems Research Center Publication ERC-201, and was supported in part by Cooperative Agreement CR812685 between the U.S. Environmental Protection Agency and the Ecosystems Research Center at Cornell University. The work and conclusions published herein represent the views of the authors, and do not necessarily represent the opinions, policies, or recommendations of the Environmental Protection Agency.

The research discussed in this paper was conducted using the Cornell National Supercomputer Facility, a resource of the Center for Theory and Simulation in Science and Engineering (Cornell Theory Center), which receives major funding from the National Science Foundation and IBM Corporation, with additional support from New York State and members of the Corporate Research Institute.

REFERENCES


FIGURE LEGENDS

FIGURE 1. General structure of the Jasper Ridge serpentine grassland model as it runs within individual cells of the model grid. The same model structure is used for each species with connections between species coming through competitive interactions during the flowering stage. Connections among cells of the model grid occur through dispersal.

FIGURE 2. Mean abundance at each time step for *Lashthenia californica*, one of the four species in the Jasper Ridge model, averaged over all cells of the same disturbance type. (A) The model run as one cell for 100 generations with no disturbances. (B) The model run as a 100 x 100 grid of cells for 100 generations with no disturbances. (C) The model run as 100 x 100 grid of cells for 100 generations with two disturbance types, April and July, affecting approximately 10% of the model cells during each time step.

FIGURE 3. A subsection of the grid for the Jasper Ridge model showing randomly distributed disturbances of two types.

FIGURE 4. Examples of descriptors of model output. (A) Semivariogram, representing the mean squared variance of the abundance of four species as a function of spatial disturbance. (B) Power spectral analysis of the spatial distribution of four species and disturbance regime. (C) Spatial variance of four species as a function of size of basic sampling unit.
Initialize

Generation Counter

Disturbance 1

Flower

Disperse?

Dispersal Pool

Seed Bank

Germinate

Grow

Die?

Census

Disturbance 2
Model Grid with Disturbances

- No disturbance
- April disturbance
- July disturbance
- April and July disturbance
Spatial Spectra for Jasper Ridge Model

at Generation 80

Power Spectrum

Frequency (Cycles/Cell)
VARIANCE VS SCALE

\[ y = 5.5002 - 0.53935x \quad R^2 = 0.948 \]

\[ y = 3.9567 - 0.53931x \quad R^2 = 0.949 \]

\[ y = 1.9327 - 0.37770x \quad R^2 = 0.979 \]

\[ y = -0.51989 - 0.48109x \quad R^2 = 0.960 \]