

**Mathematical and computational challenges
in population biology and ecosystem science**

Simon A. Levin*, Bryan Grenfell, Alan Hastings and Alan Perelson

**S. A. Levin, Department of Ecology and Evolutionary Biology
Princeton University, Princeton, NJ 08544-1003**

**B. Grenfell, Zoology Department, Cambridge University
Downing Street, Cambridge, CB2 3EJ, UK**

**A. Hastings, Division of Environmental Studies,
Institute for Theoretical Dynamics and Center for Population Biology
University of California, Davis, CA 95616**

**A. Perelson, Theoretical Biology & Biophysics
Los Alamos National Laboratory, Los Alamos, NM 87545**

***To whom correspondence should be addressed**

Introduction:

Mathematical and computational approaches to biological questions, a marginal activity a short time ago, are now recognized as providing some of the most powerful tools in learning about Nature, guiding empirical work and providing a framework for synthesis and analysis (1, 2). In some areas of biology, such as molecular biology, the advent has been recent but dramatic, for example as an adjunct to the analysis of nucleic acid sequences or the structural analysis of macromolecules. In population biology, in contrast, the marriage between mathematical and empirical approaches has a century-long history, rich in tradition and in the insights it has provided. Statistics and stochastic processes, for example, derive their origins from biological questions, as in Galton's invention of the method of genetic correlations and Fisher's creation of the analysis of variance to study problems in agriculture (1). Branching processes were developed to describe genealogical histories; and even such classical subjects as dynamical systems theory have been enriched by contact with problems in population biology (see for example 3, 4).

In recent years, the nature of the game has changed, primarily due to the availability of high speed computation. Whereas classical approaches to population biology, as indeed classical approaches to other problems in biology, emphasized deterministic systems of low dimensionality and thereby swept as much stochasticity and heterogeneity as possible under the rug, new techniques and the availability of more powerful computers have led to the development of highly detailed models in which a wide variety of components and mechanisms can be incorporated. In a model of animal grouping, every animal can be tracked; in a forest model, every tree; in an epidemiological model, every individual in the population. Since models of this sort may provide an unjustified sense of verisimilitude, it is important to recognize them for what they are: imitations of reality that represent at best individual realizations of complex processes in which stochasticity,

contingency and nonlinearity underlie a diversity of possible outcomes. Individual simulations cannot be taken as more than representative of this diversity, but repeated simulations can provide statistical ensembles that contain robust kernels of truth. The problem becomes one of the central ones in science: determining what is signal and what is noise by understanding what detail at the level of individual units is essential to understanding more macroscopic regularities.

The issues raised in the previous paragraphs cut across population biology and ecosystem science, from the immune system to the biosphere. At each level, one observes dynamics that emerge from the collective behaviors of individual units. The challenge then is to develop mechanistic models that begin from what is understood (or hypothesized) about the interactions of the individual units, and to use computation and analysis to explain emergent behavior in terms of the statistical mechanics of ensembles of such units. In the following sections, this challenge is examined for a range of scientific problems. Many of the ideas are explicated in more detail in Levin (1), and represent conclusions derived more recently in Kollman *et al.* (5). The areas discussed range across a spectrum of problems in population biology, from the populations of *B*-cells and *T*-cells in the immune system, to the variety of genotypes within a population, to the diversity of populations in the biosphere. Though the nature of the biological problems differs, the similarity is what stands out: An individual organism is a biosphere in miniature, with competition, exploitation, mutualism, succession and nutrient cycling, and providing the stage for evolutionary changes on the small scale, including selfish and cooperative behaviors. Though the subdisciplines that are highlighted have their individual cultures and dynamics, the commonality of the mathematical and computational challenges can foster positive feedbacks that would otherwise not occur.

2. Ecology.

The characterization of ecological interactions provides one of the most venerable of venues for mathematical biology, dating back certainly to Volterra's consideration of the fluctuations of the Adriatic fisheries, if not before. The challenges facing us today, for example, in the consideration of global change and the loss of biodiversity, and in achieving a sustainable future (6), elevate the complexities to new levels.

General circulation models are providing detailed information on likely scenarios of climate change and the global fluxes of key elements such as carbon and nitrogen. The resolution of such models, typically, is at the scale of hundreds of kilometers; how then can we assess likely effects on natural and managed systems, where the scales of interest are typically on the order of meters or even centimeters? Even more difficult, how can we extrapolate from the level of effects on individual plants and animals to changes in the distribution of individuals over longer time scales and broader space scales, and hence in community level patterns and the fluxes of nutrients? Individual-based models, such as the forest growth simulators JABOWA (7), FORET (8) and SORTIE (9), provide a point of departure; but the amount of detail in such models cannot be supported in terms of what we can measure and parameterize. The result is that such models produce cartoons that may look like Nature, but represent no real systems. However, they do represent powerful experimental tools, which become more valuable when used to produce exhaustive simulations that allow exploration of parameter space and model structures, and permit adequate representation of the full statistical ensemble of possible realizations associated with the many stochastic elements. Development of extensive sets of outputs from multiple runs forms the basis for extracting essential and more robust features that can be compared with data, and that can provide the meat for simplification (10, 11). Simplification techniques may include familiar tools such as renormalization or moment closure (12), in approximations that present more interpretable representations of pattern