COMPUTATIONAL ECOLOGY: ENVIRONMENTAL PROBLEM SOLVING FOR THE NEW MILLENNIUM - submitted to SIAM News 9/97

Louis J. Gross
The Institute for Environmental Modeling
Departments of Ecology and Evolutionary Biology
and Mathematics
University of Tennessee
Knoxville, TN 37996-1610
gross@tiem.utk.edu
http://www.tiem.utk.edu/~gross/

INTRODUCTION:

The Everglades region of South Florida presents one of the major natural system management challenges facing the U.S. With its assortment of alligators, crocodiles, manatees, panthers, large mixed flocks of wading birds, highly diverse sub-tropical flora, and sea of sawgrass, the ecosystem is one of the most unique in the country. The region is also perhaps the largest human-controlled system on the planet, in that the major environmental factor influencing the region is water, and water flows are managed on a daily basis -- subject to the vagaries of rainfall -- by a massive system of locks, pumps, canals and levees constructed over the past century. The changes brought about by such control have led to extensive modifications of historical patterns and magnitudes of flow, causing large declines in many native species, extensive changes in nutrient cycling and vegetation across South Florida, and great increases in pollutants such as mercury. Constrained by the conflicting demands of agriculture, urban human populations, and wildlife for control of water resources, and the varying agendas of hosts of government agencies and non-governmental organizations, there is now an ongoing effort to plan for major changes to the system with expenditure estimates of five billion dollars or more over the next decade. Carrying out such planning, particularly as it impacts upon the natural systems of the region, provides one of the major challenges to the new field of computational ecology.

COMPUTATIONAL ECOLOGY AND REGIONAL ASSESSMENT:

Computational ecology is an emerging multidisciplinary field, similar in concept to the cell and molecular emphasis of bioinformatics, which applies modern computational methodology to key problems at higher levels of biological organization. The goal of computational ecology is to combine realistic models of ecological systems with the often large data sets available to aid in analyzing these systems, utilizing techniques of modern computational science to manage the data, visualize model behavior, and statistically examine the complex dynamics which arises. Success in applying this to regions such as the Everglades requires diverse expertise in areas such as field biology, complex systems theory, computational science, remote sensing, and mathematical modeling, as well as the development of mechanisms to link approaches that are at the forefront of research in many of these fields. The field presents exciting challenges for applied mathematicians desirous of having a real impact on applied environmental problems.

The vast majority of theory in ecology has started from very simple differential equations in which a single variable represents population densities; solutions of these are analyzed mathematically, and may be compared to abundance estimates from field or lab observations. Although these models have had a great influence on theory in ecology, their aggregated form is particularly difficult to relate to observational biology. Their application to complex natural systems with spatially and temporally varying environmental factors which force the system leads to models which are not analytically tractable, and must be investigated numerically. The study of equilibrium and stability behavior typical in mathematical ecology is of little relevance for natural systems with strong abiotic forcing.

One of the most important functions of the environmental sciences today is to analyze the impacts of human actions on ecosystems and to provide management recommendations in order to ameliorate these impacts. In all parts of the world ecosystems are affected by the shrinkage and dissection of natural areas, disruptions of natural cycles, and the input of pollutants. Ecological assessment refers to the determination of the impacts of various anthropogenic influences on a natural system. Common components of such an assessment would be:

* Changes in population densities of species considered "important" species for either cultural or economic reasons, including endangered species such as the manatee and Florida panther.
* Changes in biodiversity of the system, particularly introductions of non-native species and associated issues of hybridization with species currently present, an example being the rapid spread of invading melaleuca trees in the Everglades.
* Changes in community structure (which may not necessarily be associated with biodiversity changes), such as the shift from sawgrass to cattail dominated wetlands.
* Effects of pollutant inputs, an example being the high mercury levels in fish throughout the Everglades.
* Direct effects of human actions on the system (e.g. hunting, deforestation, sewage/waste disposal), such as the effect of poaching on the Florida panther
* Indirect effects of human actions (e.g. habitat fragmentation, soil erosion, salinity changes), suchas the increase in salinity in Florida Bay due to reduced fresh water flows through Everglades National Park.

The spatial extent of the effects of anthropogenic impacts range from very local (tens of meters) to regional, on the order of hundreds of kilometers, and therefore require assessments that can span these scales as well. Regional problems involve controls and dynamics changing over time periods far longer than the one-to-five-year periods typical of academic research projects. In addition to Everglades restoration, projects involving extensive computational components are ongoing in the Columbia River basin, the San Diego Bay area, and the Southern Appalachians, to name just a few of the efforts in the U.S. Dealing with such regional problems has led naturally to the use of remote sensing data and geographic information systems (GIS). Still very much in their infancy are methods to couple GIS to information on tracking of animal movements (now feasible for large numbers of organisms due to the availability of inexpensive and highly reliable radio tags), site specific studies of animal behavior, and dynamic models. Yet there are immediate needs for rigorous scientific methods to couple available data to realistic ecological models to assess the potential impacts of alternative management scenarios.

MODELING THE EVERGLADES:

It is exactly this need for assessment of the long-term (e.g. over 30 years) impacts of alternative hydrologic scenarios on biotic components of the natural systems in South Florida which has led a large group of collaborators to develop an approach we call Across Trophic Level System Simulation, or ATLSS. Just as the Atlas of mythology bore the world on his shoulders, we hope the methodologies we are developing will provide a firm basis for bearing the weight of ecological impact analysis for many natural systems across the planet. Key to this is our use of a multimodeling methodology (Figure 1) in which we argue that the use of any single modeling approach is inappropriate for problems spanning a wide variety of temporal, spatial and organismal scales.

ATLSS uses a mixture of approaches, based upon the inherent temporal scales and spatial extent of various trophic components, linked together by spatially-explicit information on underlying environmental (e.g. water, soil structure, etc.), biotic (e.g.vegetation), and anthropogenic factors (e.g. land-use). The approaches currently involved include static spatially-explicit indices, compartment analysis, differential equations for structured populations and communities, and individual-based models. Linking models which operate at very different spatial and temporal scales has been a major challenge, requiring a variety of spatial interpolation methods, and careful design of model interfaces. Due to the modular nature of the project and the desire for long-term flexibility in what submodels may be included, we chose from the start an object-oriented design, utilizing C++.

The hierarchical spatial pattern shown in Figure 1 for the dynamic components of ATLSS arises from the association of different model types with different species or groups of species in the ecosystem that have different-sized typical spatial extents over which they are active. The compartment models deal with variables representing spatially localized biota, mainly the biomasses of lower trophic levels such as algae. These biomass variables only interact locally. Therefore, one can represent these variables across a landscape by means of many local uncoupled spatial unit cell models, in which the cell size is chosen small enough to represent a tract relatively homogeneous in substrate and elevation, the extent of which might be several hundreds of meters in a relatively flat landscape such as the Everglades.

The age- and size-structured population and community models represent intermediate trophic levels, such as fish, macroinvertebrates, and small non-flying vertebrates. These may undergo small-scale movements in response to changes in water levels, so their spatial domains of interaction are larger, on the order of up to a square kilometer, thus encompassing many unit cells, coupled now to allow population movements.
Individual-based models are employed to represent populations of top predators or other large-bodied species, such as wading birds and panthers. Individuals of these species may

move over large areas, with movements over short periods of time spanning areas of thousands of spatial unit cells. These individual-based models, an ecological form of agent-based models, are rule-based approaches which can track the growth, movement, and reproduction of many thousands of individuals across the landscape. Adequate description of a single individual can easily require a number of variables detailing its current age, size, location, physiological status, and other information pertinent to its future actions. Associated with the spatial extent of the different model components are temporal scales and the vertical axis of Figure 1 shows the range in number of state variables or level of aggregation necessary to describe the system.

Figure 2 illustrates our vision of how ATLSS will continue to develop, though at present not all modules of the multimodel are complete. Despite this, ATLSS is being used now to evaluate alternative hydrologic scenarios, and is in fact the only approach available to assess the spatially explicit dynamic responses of fish, vegetation, wading birds, deer and panther and compare scenarios. From October 1997 through May 1998, every 3 weeks a new scenario is being provided by the South Florida Water Management District and U.S. Army Corps of Engineers. Each hydrologic scenario consists of a file of daily predictions of water depth over a 31-year time period, across a region of approximately 80 miles by 130 miles with a resolution of 2 miles by 2 miles. This is the hydrology input to ATLSS, which then produces maps of species densities, time series of spatially-averaged population abundances and size-class structure, and maps of breeding potential for several species. Interested readers can view the current set of evaluations along with more details on ATLSS at the Web site listed at the end of this article. The ATLSS models involved are computationally intensive, taking several days to run on fast UNIX workstations, though we have made progress on developing parallelization methods for some components, particularly the individual-based models.

THE FUTURE:

Computational ecology allows us to investigate far more realistic ecological models than we might otherwise be able to explore. Its practitioners are driven by the need to improve our predictive capabilities -- to more accurately assess the future impact of human actions on natural systems. The difficulty of manipulating and replicating experiments at landscape scales raises issues about experimental design, the regularity and longevity of sampling, and the integration and storage of data. Associated with this are the need for methods to archive complex data structures arising from spatially explicit dynamic models, which we have only begun to address by utilizing object-oriented databases. The field also has the capability to suggest appropriate long-term monitoring plans for natural systems to aid adaptive management.

A central issue in computational ecology is the need to link dynamic processes that operate across differing spatial regions and at different rates. How do we link natural and anthropogenic forces that influence the demand for biological resources with the dynamics of those resources? How much averaging and smoothing of very high resolution biological data must be done to match the lower resolution of geophysical data while preserving the predictive capabilities of the approach for the underlying natural systems? All this is clearly tied in with both what it is we wish to predict, as well as the accuracy desired for such a prediction to be useful. As always for such complex models, issues of model validation and error propagation are difficult to address. I would argue that for many purposes, what we wish to produce are relative assessments of different scenarios, rather than exact quantitative predictions for any particular scenario. Under the assumption that errors in parameter estimates and functional forms within such complex models do not interact differentially with changes in scenarios (a reasonable assumption for scenarios produced by methods external to the models being used to assess them), it is reasonable to assume that errors propagate similarly in model runs on different scenarios. This may justify the use of highly complex models for relative assessments, though I have had little success mathematically proving such assertions.

In addition to further development of multimodeling methods, and computationally efficient approaches to carry them out, computational ecology offers the opportunity to combine spatially explicit ecological models with economic and social impact models. Associated with this is the need for applied mathematicians to address problems of spatially-explicit control. Management that occurs at landscape scale (e.g. forest harvesting, water flow management, conservation preserve design, etc.) is not an all-or-nothing affair that occurs uniformly in space. Rather, realistic management scenarios must take account of spatial heterogeneity in underlying resources, as well as how such heterogeneity interacts with management through time (local ecological succession for example). Given that there are a variety of potential criteria which affect the system management, so that the underlying non-spatial issue may be viewed as a multiple criteria optimization problem, how should the "control" of the system be applied spatially in order to carry out the optimization? This is a little-developed area of applied mathematics, particularly in systems in which there are stochastic factors which interact with the management scheme. Yet such control problems are at the heart of much of applied ecology today.

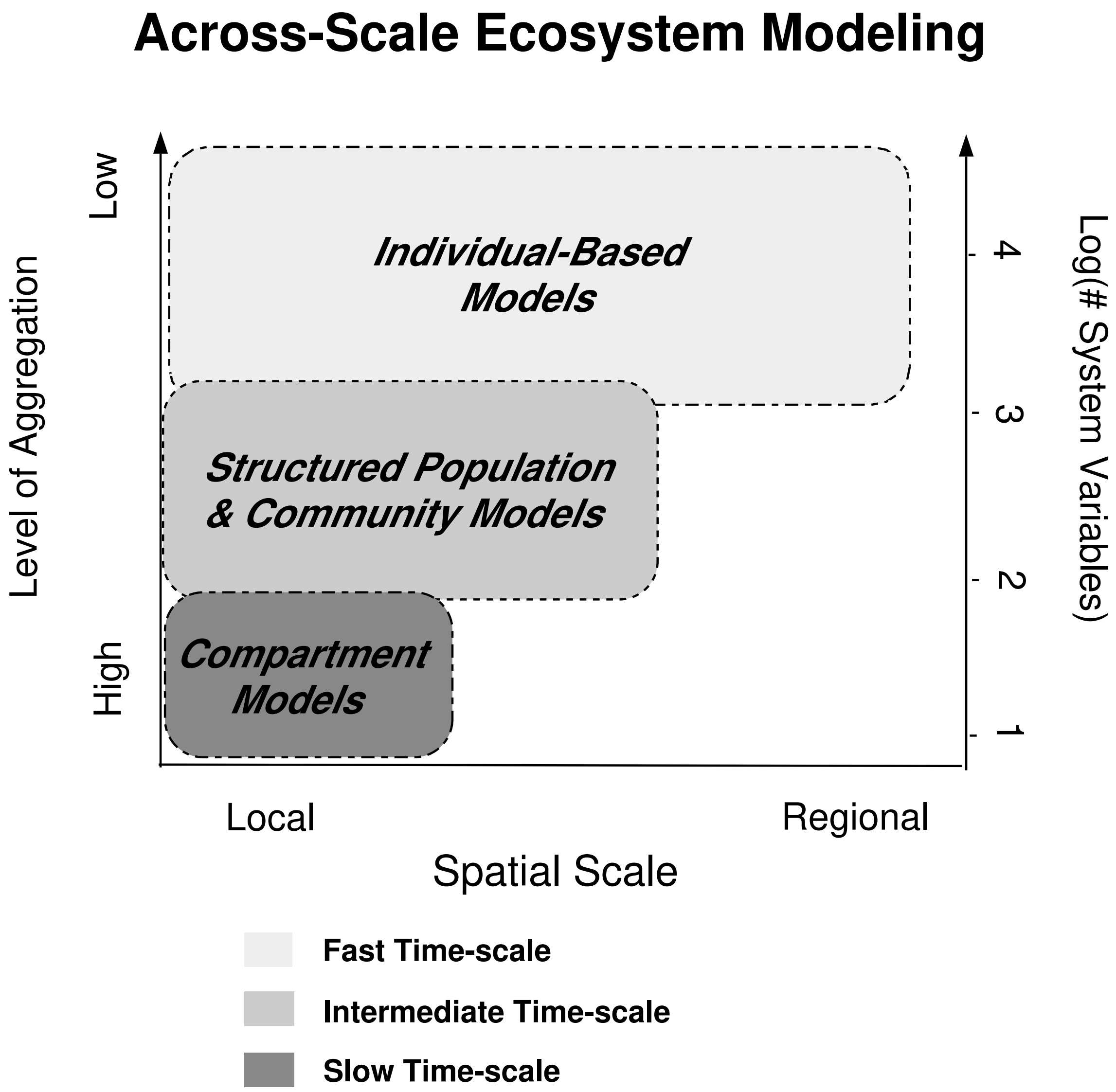
For further reading:
A Workshop Report on Computational Ecology is available at
http://www.sdsc.edu/compeco_workshop/report/report.html
The ATLSS Home Page contains detailed information on the ATLSS project at
http://www.tiem.utk.edu/~gross/atlss_www/atlss_frame.html

ACKNOWLEDGEMENTS:
I appreciate the support of the U.S. Geological Survey, Biological Resources Division (Cooperative Agreement #1445-CA09-95-0094), the National Science Foundation (BIR-9 318160), and the Environmental Protection Agency (R825157-01-0). The statements in this report are solely those of the author and do not necessarily reflect the views of any of these agencies. A number of my colleagues have been influential in the development of my thoughts in connection with this article, and I particularly thank Jane Comiskey, Donald DeAngelis, Thomas Hallam, Michael Huston, Marilyn Kallet and Scott Sylvester.

FIGURE LEGENDS:

Figure 1 A general approach to multimodeling for ecosystems is illustrated with three different modeling approaches, each with a range of spatial scales upon which they operate. Compartmental models are the most localized, tracking changes in components on a highly local scale. These compartment models have 10-100 state variables, and operate on a relatively slow time scale. Structured population and community models take account of the details of size, age or physiological state structure within a population or community by breaking down populations and communities into discrete classes. Within a localized grid cell such models can include 100-1000 state variables and operate on time scales that may be faster than the compartment models. Individual-based models, which track the behavior of each individual within a population, operate across wide spatial ranges due to individual organism movement abilities. These models can easily have1000 or more state variables, with temporal resolution as fineas that associated with significant changes in individual behavior.

Figure 2. The plan for the components of ATLSS, illustrating the variety of trophic components being addressed using each modeling approach. Although several of these component models are still under development, ATLSS is currently being applied to analyze the relative impacts on biotic systems in South Florida of alternative hydrologic scenarios.


Across-Scale Ecosystem Modeling
Low
High
Level of Aggregation
Individual-Based Models
Structured Population & Community Models
Compartment Models
4
3
2
1
Log(# System Variables)
Local
Regional
Spatial Scale
Fast Time-scale
Intermediate Time-scale
Slow Time-scale

GENERAL STRUCTURE OF ATLSS EVERGLADES MODELING STUDY