MULTIMODELING: NEW APPROACHES FOR LINKING ECOLOGICAL MODELS

Louis J. Gross

The Institute for Environmental Modeling

University of Tennessee

Knoxville, TN 37996-1610

and

Donald L. DeAngelis

Biological Resources Division, USGS and

Department of Biology

University of Miami

Coral Gables, FL 33124

Dr. Louis J. Gross

Department of Ecology and Evolutionary Biology

University of Tennessee Knoxville, TN 37996-1610

Phone: (423)974-4295; fax: (423)974-3067,

E-mail: gross@tiem.utk.edu 20 pages, 1 figure

Abstract: Computational ecology is an emerging multidisciplinary field, with the objective of combining 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. One ongoing effort in the area, the ATLSS (Across Trophic Level System Simulation) project, involves analyzing the ecological impacts of hydrologic planning across the Everglades region of South Florida. ATLSS is structured as a multimodel, a mixture of modeling approaches based upon the inherent temporal resolution 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 resolutions has been a major challenge, requiring a variety of spatial interpolation methods, and careful design of model interfaces and linkages with remote sensing data and GIS. Such multimodeling efforts are an essential tool to link together disparate data sets, aid monitoring programs, and produce relative assessments of the ecological impacts at landscape-levels of spatially-explicit management plans.

Key Words: computational ecology, multimodeling, Everglades, across trophic level, spatially-explicit modeling

INTRODUCTION:

The Everglades region of South Florida presents one of the major natural system management challenges facing the United States. 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 (Davis and Ogden 1994). 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 eight billion dollars or more over the next several decades (U.S. Army Corps of Engineers 1999). 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 (Helly et al. 1995). 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 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), such as 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 Sacramento River Basin, 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.

MULTIMODELING AND 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 (DeAngelis et al. 1998). Just as the Atlas of mythology bore the world on his shoulders, we expect the methodologies we are developing to 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.

The application of mathematical models in most fields involves determining whether the key features of the problem under consideration require a discrete or continuous formulation, for example populations with discrete life stages as compared to those with overlapping generations. A second dichotomy concerns whether stochastic factors should be included or ignored, that is if unpredictable abiotic environmental factors must be explicitly modeled or whether the averages of such variations are sufficient to describe and analyze the situation. Once basic modeling issues such as these are decided, this specifies a single modeling approach, such as a matrix approach for structured populations in discrete time, an ordinary differential equation system for discretely-structured populations in continuous time, or a partial differential equation for a continuously structured population in continuous time. It is very atypical for the modeling framework used for a particular problem to utilize more than one modeling approach. Rather, as the modeling process proceeds, and additional complexities of the system under study are being considered, sometimes a new approach is taken. Thus, a simple population model for abundance dynamics could be phrased as a single differential equation, but when age structure is added, the model formulation would then involve a system of differential equations.

Unlike the situation described above, ATLSS relies upon a variety of different mathematical approaches, a process known as multimodeling. This mixture of approaches is based upon the inherent temporal and spatial resolution and 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 (e.g. land-use) factors. The approaches currently involved include spatially-explicit indices, compartment models, differential equations for structured populations and communities, and individual-based models. Linking models which operate at very different spatial and temporal extents has been a major challenge, requiring a variety of spatial interpolation methods (Luh et al. 1997), and careful design of model interfaces (Duke-Sylvester and Gross 1999). 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++.

Figure 1 illustrates the variety of dynamic approaches utilized in ATLSS relative to the spatial and temporal resolutions of the trophic levels involved. The hierarchical pattern in Fig. 1 for the components of ATLSS arises from the association of different model types with different species or groups of species in the ecosystem that have different typical spatial extents appropriate for the variables in the models. The spatial extent indicated ranges from the finest resolution available, currently 28.5 m from Landsat image information, to regional extents on the orders of hundreds of kilometers. The horizontal range of one of the boxes indicates the range of spatial extents over which variables for that type of model may reasonably be considered as acting. Models that follow individuals necessarily require localized information on individuals in order to appropriately assess the interactions between individuals which only interact locally. Although the variables in these individual-based models may be averaged over larger spatial extents, these averaged variables cannot then be used to assess the basic interactions within the model. Other

approaches have variables, such as those indicated as index models, for which the underlying variables can be estimated at a wide range of spatial extents without affecting the basic assumptions of the model. Models such as those for structured populations have variables that may be defined at an intermediate range of spatial extents, over which it may be assumed that within this extent the spatial variability of environmental conditions may be reasonably averaged.

The vertical axis of Fig. 1 illustrates the complexity of the associated models, using the number of system variables as a measure of such complexity. Alternatively this may be viewed as indicating the amount of aggregation being utilized in the description of the component of the biotic system of interest. Some submodels, here indicated as compartment models typically applied to lower trophic levels including periphyton and zooplankton, can for practical purposes be well-described by simple coupled differential equations. These describe interactions between functional groups, defined here as groups of species that perform similar roles (on the same trophic level) with forcing functions representing the effects of abiotic influences. Thus a reasonable description of a zooplankton community can be reduced to a single or a few state variables. Other components of the biotic community may have a more complex structure that requires a structured population or mixed-species community description. Additionally, this structure may directly impact higher trophic components. For example, fish survival and growth are typically size dependent, implying that a single state variable for a fish population may be inappropriate for many situations. The number of system variables needed to describe a structured population is higher than those needed when simply following population abun-

dances or densities, thus leading to a lower level of aggregation.

Higher trophic level organisms are generally much more behaviorally complex and are able to display a repertoire of responses to environmental conditions and other organisms that cannot be captured in a small number of state variables. Each individual may behave differently based upon both its current condition and its current environment. For such organisms, an individual-based approach may be indispensable to determine how the individual actions affect population-level phenomena. 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. When considering populations with hundreds to thousands of individuals, the total number of state variables can easily be in the tens of thousands.

Although temporal resolution also has great importance in this modeling scheme, it is not as straightforward as spatial extent. In part this is due to the fact that all populations are made up of individuals that carry out actions on short time periods. The relevant temporal resolution of importance here is that associated with the system variables of interest. Thus if one is dealing with zooplankton density rather than actions of an individual zooplanktor, the appropriate temporal resolution is not the seconds upon which individuals move, but rather the many day time period over which significant changes in plankton density occurs at the spatial extent chosen for the model. In Fig. 1, the relevant temporal resolution should be viewed as the upper bounds on which a temporal description of the model makes sense. At the level of individuals, the relevant temporal resolution is set by the unit of time on which they can execute actions that affect their survival and

reproduction. Thus, for individual birds, weekly time steps are inappropriate since within a week an individual bird's actions can have great impact on its survival or growth. For fish populations however, in which the relevant state variables are size or age distributions, an upper bound on temporal effects may be weeks or longer.

THE ATLSS MODELS

The ATLSS hierarchy starts with models which translate coarse resolution hydrologic information to a finer resolution appropriate for biotic components that operate at spatial extents much smaller than the resolution of the main hydrologic model. This main hydrologic model, a product of the South Florida Management District, has been utilized to produce all of the alternative plans for water management and typically provides a 30-year plan of daily water depths across approximately 1700 spatial locations, each of which represents a 2-mile by 2-mile grid cell. A full description of the methods used to produce a high resolution hydrology is available on the ATLSS Project Home Page (ATLSS). This relies upon vegetation maps and the associated limitations on hydroperiod associated with these, to characterize a 28.5 meter resolution topography chosen to preserve the volumes of water derived from the 2-mile resolution hydrology model.

The ATLSS hierarchy next includes Spatially-Explicit Species Index (SESI) models which make use of the spatially-explicit, within-year dynamics of hydrology to compare the relative potential for breeding and/or foraging across the landscape (Curnutt et al. 1999). These models produce yearly comparisons of the spatial extent of the index across the landscape. SESI models are viewed as approximations which are useful in

coarse evaluations of scenarios and are an aid in interpreting the more detailed models. SESI models have been constructed and applied during the Restudy to the Cape Sable Seaside Sparrow, the Snail Kite, Short- and Long-Legged Wading Birds, and White-tailed Deer.

The compartment models involve biotic components for which it is reasonable to model variation 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. This is particularly appropriate for the primary producers (e.g., periphyton, aquatic macrophytes, terrestrial macrophytes), and meso- and macroinvertebrate consumers and detritivores, which can be combined into a few main functional groups. In it's current state of development for the Restudy, these models had not been adequately developed to link them to hydrology, and so they were not applied. However, from the point of view of restoring the higher consumer populations, the importance of this lower trophic level web is the food resources it provides to support these higher trophic levels and validation for the fish models has indicated that accounting for the temporal variability in these components is quite critical (Gaff 1999).

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 short distance 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. Detailed

models have been developed and applied in the Restudy to estimate responses of functional groups of fish across the freshwater landscape (Gaff et al. 1999). These models consider the size distribution of large and small fish as important to the basic food chain which supports wading birds. It has been applied to assess the spatial and temporal distribution of availability of fish prey for wading birds.

Individual-based models are employed to represent populations of top predators or other large-bodied species. Individuals of these species may move over large areas, with movements over short periods of time spanning areas of thousands of spatial unit cells. Such models are also used for organisms of particular interest, such as endangered species, for which detailed behavioral information is available and there is particular interest in the detailed response of the population to alternative plans. 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. ATLSS models of this type include ones for the Cape Sable Seaside Sparrow, the Snail Kite, the White-tailed Deer, the Florida Panther and various Wading Bird Species. The models include great mechanistic detail, and their outputs may be compared to the wide variety of organism distribution data available, including that from radio collared individuals. An advantage of these more detailed models is that they link each individual animal to specific environmental conditions on the landscape. These conditions (e.g., water depth, food availability) can change dramatically through time and

from one location to another, and determine when and where particular species will be able to survive and reproduce.

HOW ATLSS WAS APPLIED DURING THE RESTUDY

The U.S. Army Corps of Engineers was charged in 1996 by the U.S. Congress to produce a restoration plan for the Everglades region of South Florida by July 1, 1999 (see U.S. Army Corps of Engineers 1999 for details). The efforts to produce such a plan involved a massive effort on the part of numerous federal and state agencies. We cannot here begin to describe the full procedures for the development of the final plan presented to Congress, but do want to indicate the procedures used to apply ATLSS. The essential goal was to provide a rational, scientific basis for developing relative rankings of the biotic impacts of the proposed hydrologic scenarios as input to the planning process.

A key feature of the ATLSS application is the reliance on what we call relative assessment. By this we mean that we do not claim that the variety of models within ATLSS can produce quantitatively accurate forecasts of population responses across the Everglades over the 30-year plan. There are simply too many uncertainties in the models and the data used to estimate the parameters to attach any great confidence to the exact spatial variation in population estimates provided by the models. Rather, what we provided during the Restudy was a ranking of how various alternative plans caused the models to respond, relative to a base plan chosen by the agencies involved. The procedure was that once a hydrologic plan was developed, the ATLSS team (and this was a team effort involving eight individuals) was given one week to carry out all the various model runs, evaluate

them, provide written comparisons of the results of each model to the base scenario and post a variety of the model results and our summaries to the ATLSS Home Page. There would follow several days of conversations with various individuals from various agencies involved in assessing the scenarios (the Alternative Evaluation Team) who would then make a set of recommendations to the Alternative Design Team (ADT). It was then the responsibility of the ADT to revise the hydrologic plan, run a new hydrology model based upon the design changes, and provide a new plan. The turnaround time between hydrologic plans was typically three weeks, and the ATLSS team carried out over 15 complete evaluations of plans during a 14 month period.

THE FUTURE

The focus of ATLSS to date has been on the freshwater systems, with emphasis on the intermediate and upper trophic levels. The ATLSS structure was purposely formulated to provide for extension to estuarine and near-shore dynamic models once physical system models for these regions are completed. This would involve the construction of a variety of additional models for the biotic components. A further effort at lower trophic levels in the freshwater regions is needed to account for the impact of hydrologic plans on vegetation change and associated nutrient fluxes. Closely linked with these would be models for the effects of major disturbances to the system, including fire and hurricanes. Finally, ecotoxicological models coupled to transport models for toxicants such as mercury may be readily incorporated into the biotic components already constructed.

In a more general context of future developments, 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 regional extents 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. We argue based upon our experience with the Restudy 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 we 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 to address problems of spatially-explicit control (Hof and Bevers 1998). Management that occurs at landscape extent (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.

ACKNOWLEDGEMENTS:

The efforts described here have been supported by the U.S. Geological Survey, Biological

Resources Division through Cooperative Agreement #1445-CA09-95-0094 with the University of Tennessee. The statements, findings, conclusions, recommendations, and other data in this report are solely those of the authors and do not necessarily reflect the views of the U.S. Geological Survey. A number of individuals have contributed greatly to the empirical work which underlies the ATLSS project, notably Sonny Bass, John Chick, John Curnutt, William Loftus, and Joel Trexler. Several individuals have contributed greatly to the ongoing computer modeling for this project, notably Jane Comiskey, Scott Duke-Sylvester, Holly Gaff, Mark Palmer, Michael Peek, and Rene' Salinas. We acknowledge the assistance of all these individuals in bringing the ATLSS project to its current stage of development.

LITERATURE CITED

ATLSS: Across Trophic Level System Simulation. 1999. http://www.atlss.org (assessed 23 June 1999).

Curnutt, J. L., J. Comiskey, M. P. Nott and L. J. Gross. 1999. Developing and using landscape-based spatially-explicit species index models for natural system planning. Ecological Applications (in review). Davis, S. M. amd J. C. Ogden (editors). 1994. Everglades: the system and its restoration. St. Lucie Press, Delray Beach, Florida, USA.

DeAngelis, D. A., L. J. Gross, M. A. Huston, W. F. Wolff, D. M. Fleming, E. J. Comiskey, S. M. Sylvester. 1998. Landscape modeling for Everglades ecosystem restoration. Ecosystems 1:64-75.

Duke-Sylvester, S. and L. J. Gross. 1999. Integrating spatial data into an agent-based mod-

eling system: ideas and lessons from the development of the across trophic level system simulation (ATLSS). In R. Gimblett (ed.) GIS and Agent-Based Modeling. University of Arizona Press (in review).

Gaff, H. D. 1999. Spatial Heterogeneity in Ecological Models: Two Case Studies. Unpublished Ph.D. dissertation. University of Tennessee, Knoxville.

Gaff, H. D., D. L. DeAngelis, L. J. Gross, R. Salinas, and M. Shorrosh. 1999. A dynamic landscape model for fish in the Everglades and its application to restoration. Ecological Modelling (in review)

Helly. J., T. Case, F. Davis, S. Levin, and W. Michener (editors). 1995. The State of Computational Ecology. http://www.sdsc.edu/compeco_workshop/report/report.html (assessed 23 June 1999).

Hof, J. and M. Bevers. 1998. Spatial Optimization for Managed Ecosystems. Columbia University Press, NY.

Luh, H.-K., C. Abbott, M. Berry, E.J. Comiskey, J. Dempsey, and L. J. Gross. 1997. Parallelization in a spatially-explicit individual-based model (I) Spatial data interpolation. Computers and Geosciences 23:293-304

U.S. Army Corps of Engineers. 1999. Central and South Florida Comprehensive Review Study. Final Integrated Feasibility Report and Programmatic Environmental Impact Statement. http://www.restudy.org/overview.pdf (assessed 23 June 1999).

FIGURE LEGENDS

Figure 1 A general approach to across trophic level modeling is illustrated with four different modeling approaches, each with a range of spatial extents upon which the variables utilized within the models may be viewed as acting. Individual-based models are the most localized, and involve the largest numbers of variables, with associated low levels of aggregation. Index models, typically produce a single index value per location and may be constructed at a wide variety of spatial extents, depending upon the level of spatial resolution available for the input data and the model assumptions. Compartment models and structured population/community models are intermediate in the spatial extents appropriate for their variables. 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. The hatching illustrates the temporal resolution appropriate for these models. In the ATLSS case, the index models operate on yearly time steps while the individual-based models have time steps less than a day.