COMPUTER BASED DEVELOPMENT OF LARGE SCALE ECOLOGICAL MODELS

PROBLEMS AND PROSPECTS

Efraim Halfon
Canada Centre for INland Waters
Burlington, Ontario
Canada L7R 4A6

Modelling an ecosystem requires knowledge of the system, obtained with experiments, and its abstraction within a mathematical framework. Systems methods can be used effectively in the latter phase of the modelling process. Indeed, when coupled with experimental work, these mathematical methods can help in the development of an ecologically realistic mathematical model. These models formalize hypotheses proposed to describe the ecosystem structure and explain its behavior. For example, a forest is not only made of trees, but also of shrubs, fungi and other flora. Animals contribute to modify and influence the forest, and chemical nutrients cycle between flora, fauna and soil. A mathematical model of this ecosystem would try to conceptualize relationships among these domains. If this model is set in dynamical terms, e.g., a set of differential equations, then the model behavior should be similar to the forest behavior.

In theory, the problem of modelling an ecosystem is similar to any other modelling problem. In practice this is not so: one of the greatest difficulties is that an ecosystem is not designed by man. Therefore, it takes much experimental work on the real system to obtain the information necessary to model it. Some man-made systems are also large scale and need special approaches to study them, for example, electrical networks. Even so, the engineers have an advantage over ecologists, they know the physical laws that regulate the system and can express them in mathematical terms. Ecologists do not know these laws yet. Thus, when quantification is necessary, intuition and approximations must be used and this process involves errors.

The purpose of this paper is to analyze the problems that arise in the development of an ecosystem model, and during its verification and validation. Some solutions devised by ecologists are presented. System theory and empirical modelling interact to create a model that is compatible with current knowledge.

The first step in model development is the selection of a hypothesis. In most cases the goal is to have a model that has an ecologically realistic structure and and which is able to predict the behavior of an ecosystem under normal conditions and under stress, usually man-made. For this purpose ecologists describe the system They are interested in with systems of first order ordinary or partial difference or differential equations; other mathematical formulations are more rarely used. The state space approach is the most modern and appropriate because it allows description of both observable and unobservable variables. The models are memoryless and nonanticipatory and the state of the system is predicted using information on the present state and inputs to the system. These state variables are called compartments, thus the model is a compartmental model.

How do we chose these compartments? If we want to model a forest we have several possibilities. A compartment can be a tree, or trees from the same species, or many species or even the whole forest. A compartment can also be the calcium in the vegetation or in the soil. If we want to model a lake, a compartment can be a species of algae, or all the algae in the lake. Should the algae be expressed as weight or as chlorophyll or as phosphorus? The answer depends on the modelling goal and the experimental information available. It also depends on the ability of the modeller to apply system theoretic concepts which can help him in making this choice. Seldom until now have these concepts been applied to construct a model that is simple enough to be simulated in a computer, but complex enough to include our ecological insight. The problem of integrating system theory and ecological theory to establish the best choice of state variables is analyzed in the paper.

A second topic of interest is system identification. System identification is a field of system theory that can formally be defined as the process of determining coefficients, parameters, and structure of a mathematical model in such a way that it describes a physical process in accordance with some predetermined criteria. In the development of ecosystem models the solution of the identification problem is important because it permits objective determination of the existence of interaction among compartments. The hypothesis here is that information about the structure of systems is implicitly included in the data and can be utilized for developing models in areas such as ecology, where knowledge regarding the structure is not available. For example, are some algae eaten by some animals or not? What are the relations between fish and algae? These are questions that regard the model structures and that can be answered with the appropriate use of system identification methods.

A third topic concerns verification and validation. This topic has received less theoretical attention from ecologists and in current modelling exercises the solution has mostly been empirical and associated with particular problems. The model has been considered verified it if is behaving as expected. For example algae should grow in spring and summer and soluable phosphorus present in the water should decrease because algae use it. More sophisticated methods, such as those introduced by Mirham, are not usually used. Model validation has received more attention because it is more closely related to the problem of prediction. The problem here is that observations are noisy and data can seldom be collected more often than once a week or once a month. We are far from the engineers' capability of sampling every few milliseconds. Thus, given the biological variability and the limited amount of data, most ecological models can be considered validated: model parameters can be modified to have a reasonable fit to the data. Two or several models can be validated on the same set of data but when used for prediction they can give very different answers. Finally, time is not the only important dimension, space must be considered too. If a large city is located near the shore of a lake, then there are local problems of pollution. The model must be able to differentiate between parts of the lake and behave accordingly. Some solutions to this problem will be discussed.

From the examples above, it is evident that the modelling of ecosystems is becoming more and more a theoretical and methodological problem. Ecologists are now contributing to the development of general systems theory as biologists, e.g., Von Bertalanffy, did decades ago.

☆ ☆ ☆

STRUCTURE CHARACTERIZATION FOR SYSTEM MODELLING IN UNCERTAIN ENVIRONMENTS

J. Bens
Belgian National Science Research Foundation

G.C. Vansteenkiste
Universities of Ghent and Brussels

Systems from uncertain environments are characterized by uncertainties as to their geometry, their field parameters and their internal characteristics; moreover, the system observations are of relatively low quality. It is shown how system modelling is not likely to be surpassed by any other method in handling such problem measurements in an adequate way and in deciding on their control.

Three fundamental stages can be recognized in system modelling: the selection of the type of model, the identification of the parameters and the validation of the model. It is shown what particularities system modelling in uncertain environments brings in to the three stages.

When developing a model for some new phenomenon, one usually starts with that model structure that is expected to be the best suited to the problem. Using experimental data the parameters are identified and the model is a posteriori compared to the data. If there is no good agreement, another model structure is chosen. Using this procedure no indications will be given and the search will have no directions. Another approach to the model building problem will be presented in this paper, where the emphasis lies on structure characterisation instead of parameter identification. The structure is indeed more fundamental to the understanding of what is happening rather than the parameter values.

The most crucial part of the method is one of feature extraction. The choice among the different models will be guided on the basis of characteristic expressions. These features must be insensitive to noise and have good discriminating properties. It is explained how this can be achieved. When the features have been computed from the data, they are used to classify the system. Here most of the techniques of pattern recognition can be used. Very useful is the method of discriminant functions. In this way, the vast amount of experimental data, and the corresponding features, is reduced to a few very significant numbers. These numbers are the input to a classifying algorithm. This algorithm will, with a predetermined reliability, classify the data as best satisfying one of the proposed models. After classification it should not be too hard to obtain the values of the parameters in the model.

The second stage, parameter identification is an old problem. Modern filter theory has improved parameter estimation quite a bit, Kalman filtering allows to extract the parameters with great accuracy from noise measurement.

The third stage, the validation of the model, is often neglected. Very few criteria exist to say whether a model is a valid representation for a real world phenomenon or not.