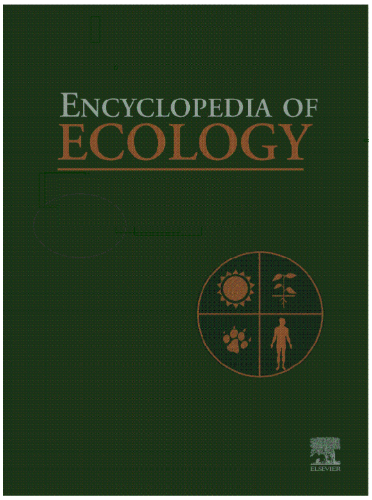

M J Dreyfus-Leon and M Scardi. Application of Ecological Informatics. In Sven Erik Jørgensen and Brian D. Fath (Editor-in-Chief), Ecological Informatics. Vol. [1] of Encyclopedia of Ecology, 5 vols. pp. [222-227] Oxford: Elsevier.

Application of Ecological Informatics

M J Dreyfus-Leon, Universidad Autonoma de Baja California, Ensenada, Mexico

M Scardi, University of Rome 'Tor Vergata', Rome, Italy

Introduction

Ecological Systems

Ecological systems are open systems characterized by a great number of interactions within and between levels of organization and by complex exchanges with other neighboring systems. Their inherent complexity makes their study, prediction, and management very difficult.

The mathematical modeling and statistical tools that have been traditionally used in ecological research allowed significant advances in ecological knowledge, but they were mainly aimed at a reductionistic approach, which can only be successful in case very simple systems are studied.

Real ecosystems, however, are always very complex (and more complex than they appear) in their structure and dynamics. The combination of many parallel and/or sequential nonlinear interactions often induce unexpected responses, which sometimes reveal chaotic dynamics, making prediction of ecosystem behavior impossible.

Another aspect in ecology that we are dealing with is nonequilibrium systems. Many models that have been used are based on equilibrium, therefore making weak assumptions about reality. Modeling is about simplifying to get a tight description of a structure and its response to a certain stimulus or its dynamics in space or time. It is almost certain that we will never get a 'perfect model' unless we are able to reproduce the system itself. Any simplification will stay short in the model representation ability.

Another problem facing researchers is that our knowledge in one ecological system is not completely transferable to another ecosystem; it is not reducible as classical physical systems are. This again is the burden of all the above.

While our understanding of ecosystem functioning is only partial, the amount of available ecological data keeps growing, and it grows much faster than our ability to turn new data into new insights into ecological processes.

Ecological Informatics

In this scenario, 'ecological informatics' can be regarded as an extremely promising research field, which has the potential to help bridge the gap between data and knowledge. As many emerging disciplines, 'ecological informatics' is still ill-defined, and several different definitions can be found. Most of them, however, agree regarding 'ecological informatics' as a combination of several research fields. It can be summarized as the application of the latest computationally intensive tools to ecological research and the development of novel computational methods inspired by biological and ecological systems.

The purposes of 'ecological informatics' are multiple, but in most cases they involve the development of modeling, data mining, data management, visualization, expert systems, or similar applications in ecological research. Computational techniques such as neural networks (see Artificial Neural Networks), cellular automata (CAs) (see Cellular Automata), or evolutionary algorithms (see Evolutionary Algorithms) are the basis for many successful 'ecological informatics' applications, but any computationally intensive method or information technology may play a role in supporting new applications.

Artificial neural networks (ANNs) have been extensively applied to ecological sciences through supervised and unsupervised learning models, and the number of applications has been growing exponentially during the last decade. Multilayer perceptrons (MLPs) (see Multilayer Perceptron) trained with the backpropagation (BP) algorithm are the most popular neural networks in ecological applications and they have been applied to a number of empirical modeling problems. While MLPs are very effective as generalized regression tools, 'self-organizing maps' (SOMs) (see Animal Defense Strategies) may be successfully applied to ordination and classification of ecological data (e.g., in indirect gradient analysis).

Although very popular, neural networks are not the only tools upon which 'ecological informatics' relies. For instance, CAs, although among the earliest artificial life models, are still applied in ecology, and they are certainly also part of 'ecological informatics'. CA shows that complex behavior and self-replicating patterns may be obtained from simple rules, when they are applied iteratively. CAs have been applied in many ecological studies, especially when population dynamics or landscape ecology is involved.

Individual-based models (IBMs) are another typical application that can be regarded as a member of the 'ecological informatics' family. They represent plants or animals as individual entities that are programmed to react to environmental stimuli, including interactions with other individual entities. The discrete nature of individual entities in IBMs leads to nonequilibrium systems, and their properties and behavior must be carefully defined in order to obtain useful simulations.

Evolutionary algorithms are certainly the methods that were more directly inspired to biological systems among those in the 'ecological informatics' toolbox. In fact, 'genetic algorithms' (GAs) (see Evolutionary Algorithms) exploit the analogy with biological evolution to solve complex optimization problems. However, the application of evolutionary algorithms should not be regarded as a mere tool for problem solving, because it also stimulated new insight into ecological problems, especially in combination with IBMs.

The list of methods that can be applied in an 'ecological informatics' framework is virtually endless, and it overlaps with the ones of other disciplines, for instance, 'ecological modeling' or 'bioinformatics'. Therefore, a couple of examples are probably more useful than theoretical definitions or comprehensive lists of methods in showing how ecological value can be additionally obtained from the application of appropriate 'ecological informatics' techniques (see Ecological Informatics: Overview).

Examples

Example 1

A very straightforward example of an 'ecological informatics' application is the Fish-based Decision Support System (FIDESS): a 'decision support system' (DSS) that has been recently developed in Italy is based on artificial intelligence and aims at assisting environmental management policies.

The need for such a DSS stemmed from the European Water Framework Directive (WFD), which set a very ambitious goal for all the member states, that is, improving the quality of all the superficial water bodies by 2015 up to a level that can be considered as 'good'. Obviously, in order to enforce the WFD policies accordingly, appropriate evaluation methods are required. The WFD clearly states that the key criterion is the 'ecological status', that is, an expression of the quality of the structure and functioning of aquatic ecosystems associated with surface waters, which is mainly based on biotic 'quality elements'. Fish fauna plays a major role among the latter, not only because fish species are effective biological indicators of environmental quality in aquatic ecosystems, but also because of their iconic value.

The majority of the available assessment methods based on fish have been developed during the last two decades, and they are mostly inspired by the seminal work by J. Karr, who developed a multimetric index (the 'index of biotic integrity', IBI), which combines 12 attributes of the fish assemblage that are supposed to respond to environmental disturbance (i.e., metrics) into a single score. This approach is inherently flexible, and therefore it has been adapted to a number of countries and river basins, not only in North America, but also in Europe and other continents.

Although multimetric biotic indices have become commonplace tools in environmental management, they are not optimized from a computational point of view and therefore even the most successful ones often fail, providing evaluations that are not consistent with other ecological evidences. This limited capability is not surprising, as no evaluation method can be simple, general, and accurate at the same time. Multimetric indices are certainly simple, so they have to give up generality in order to be accurate, and in fact the most successful ones are usually aimed at a single river basin or at a single, very homogeneous ecoregion. Basically, multimetric biotic indices usually rely upon a sound ecological rationale, but they exploit the available information in a suboptimal way.

In order to be both general and accurate, methods for evaluating the ecological status must be more complex than multimetric indices in the way they process the available information. 'Ecological informatics' is certainly the appropriate conceptual and methodological framework for developing such an optimized method.

Therefore, a DSS based on an ANN was trained to associate fuzzy expert judgments to environmental and fish assemblage data. This solution was based on the assumption that complex biotic relationships that link fish assemblage composition to environmental conditions can be embedded into an ANN and that such an ANN can be trained to mimic the way human experts issue their judgments.

In fact, expert judgment, although inherently subjective, is the key for any environmental assessment method, from the selection of relevant metrics to the discretization of the scoring scales of multimetric indices. The same subjectivity affects the evaluation of the ecological status, which cannot be univocally defined, and it is mostly based on the personal interpretation of natural phenomenologies. In spite of the lack of objective criteria, ecologists usually agree in ranking sites according to their ecological status, because they share a common rationale.

FIDESS is still under development, as more information (fish assemblage data, environmental data, and multiple expert judgements) is needed to fully train the ANN with respect to a full spectrum of ecoregional conditions, and at present it is optimized for central Italy.

In spite of theoretical problems related with the so-called 'curse of dimensionality' and thanks to the strong biotic relationships that implicitly constrained the learning phase, a few hundreds records allowed to properly train a very complex ANN. (The curse of dimensionality refers to the exponential increase in volume caused by the addition of new dimensions to an *n*-dimensional space. In machine learning applications it usually hinders the solution of problems involving a limited number of patterns in a high-dimensional feature space.) This ANN is a 59-25-5 MLP, which has 27 abiotic and 32 biotic inputs. Among the latter, several hydromorphological attributes as well as some chemicophysical ones are considered, while presence/absence of 30 species, plus overall and juveniles-only species richness were included as descriptors of the fish assemblage. The ANN has five outputs, which correspond to fuzzy membership estimates relative to each one of the ecological status classes that are defined by the European WFD (and that are considered in the human expert judgments). The ANN outputs can be regarded as memberships as they sum up to one thanks to a softmax activation function in the output nodes.

The training of the ANN-based DSS is performed not only using data directly obtained from sampling, but also 'virtual' records. Basically, during expert judgment elicitation, human experts are also asked to point out which changes in biotic and/or abiotic would affect their evaluation, or to explain how their evaluation would change in case different (but likely) environmental and faunistic properties were observed. In this way alternate scenarios can be easily simulated and new expert judgments can be associated to each 'virtual' record, thus widening the knowledge base upon which FIDESS is built.

Even though at its present development stage FIDESS can be regarded as a very early alpha release of the final tool, it has been tested using an independent data set ($n = 69$). A confusion matrix, that is, a 5×5 contingency table, was obtained by cross-tabulating human expert judgment against FIDESS classification, showing a very good agreement: two out of three cases were correctly classified after defuzzification, while the worst-case error was within a single quality class. A typical measure for interobserver agreement, the weighted Kappa statistics, confirmed that the deviation of the FIDESS classification from a random agreement with expert judgment was highly significant.

Although computationally intensive, an ANN-based DSS cannot be regarded as a paradigm for 'ecological informatics'. An essential component in this light is the 'Graphical User Interface' (GUI) that was wrapped around the ANN to provide a user-friendly and interactive access to FIDESS (**Figure 1**). The GUI makes the ANN – that is, the unnecessary complexity – absolutely transparent to users who are free to interact with FIDESS. As soon as they modify the input data, changes in classification in real time can be observed. Although it is trivial if compared to the ecological and computational background of FIDESS, the GUI is not a secondary feature. On the contrary, it plays a major role in the acceptance of FIDESS. In fact, while most users are familiar with multimetric and other biotic indices, they do not feel comfortable with an ANN, which is perceived as a rather obscure 'black box'. Interacting with FIDESS in real time, thanks to a user-friendly GUI, for example, by moving sliders, helps users to learn how FIDESS reacts to changes in biotic and abiotic variables and to understand that FIDESS just mimics their own way of reasoning. The relationships between user's input, ANN, and FIDESS outputs are summarized in **Figure 2**.

In conclusion, this combination of a typical artificial intelligence technique, a smart knowledge elicitation procedure, and a very user-friendly and interactive GUI can be regarded as a good example of what 'ecological informatics' is all about: combining available methods, data, knowledge, and software into new, viable solutions for ecological problems.

Example 2*

This example deals with a representation of animal behavior and learning. The agent or artificial animal is generated and attempts to cope with the features of the world with its limited knowledge. The objective is to adapt to the problem presented to it through learning, gradually modifying its behavior related to movement in the space (**Figure 3**).

The agent used in the example carries an ANN (three layers: input layer where data enter the networks, hidden layer where they are classified, and output layer where decisions are made) that must learn by reinforcement (punishment and reward) the best strategy to get as much profit as possible from the world in terms of food. In order to catch its food, it must be in the same pixel as its prey (prey do not move). Unfortunately as all animals, it is not perfect. It has limited knowledge and capacities. What it can do is make a decision each unit of time upon three possibilities: keep moving straight, turn slightly to the left, or turn slightly to the right. These decisions are relative to its current direction.

As input it has only memory consisting of knowledge as well as knowledge of failure of the last three decisions. It lacks any sensorial capacity or knowledge of its location in the space.

The toroidal world is a square area on the top of the screen (**Figure 3**); below it, a performance histogram will appear (**Figure 4**), describing by time intervals the catch that this predator has achieved through the actual animation run.

---

* Demonstration program will be available through the first author by an email request.

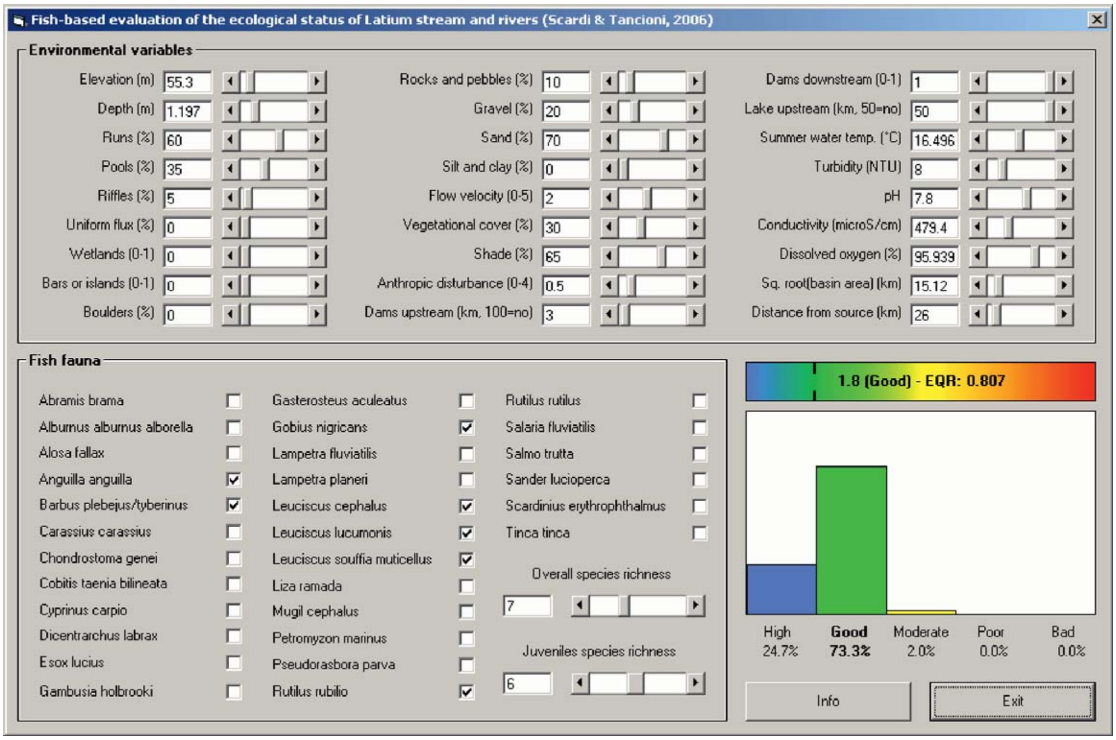

Figure 1 The Fish-based Decision Support System (FIDESS). Quantitative data can be modified by moving the sliders, while the classification results (shown in the lower-left part of dialog) change in real time. The very user-friendly GUI played a fundamental role in the acceptance of the method among ecologists and fish biologists who were not familiar with the underlying computational methods but are used to apply simpler biotic indices.

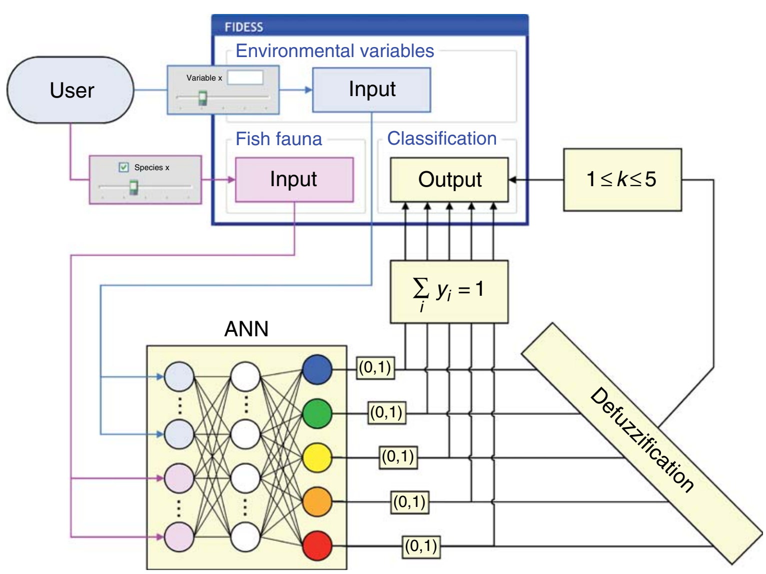

Figure 2 The GUI of FIDESS makes the underlying ANN completely transparent to the user. Quantitative data are entered in input fields, while binary data are entered by means of check boxes. Sliders are also available for quantitative variables, thus providing immediate feedback to the user, who can easily compare FIDESS actual responses with its expected behavior. Both fuzzy and crisp (i.e., defuzzified) classifications are provided in output.

No matter which decision this agent takes, there is a possibility of failure, and no matter how fitted it is, there is a possibility of being in a deserted area. With its limited information, it cannot avoid being starved at some intervals, but as learning proceeds it changes its behavior, starting from a random walk and finishing with sinusoidal movement when it has encountered prey or with a more straightforward movement when it has not. This improves

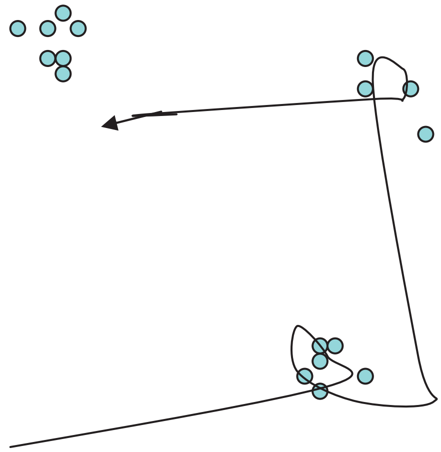

Figure 3 Schematic representation of the toroidal world with patches of food and an animal track.

the chances of catching prey when it is close to a patch (Figure 4). This type of strategy has been described in predators when prey is found in clusters. A sinusoidal movement in a situation close to a patch will increase the probability of keeping close to the patch and a straight movement in the opposite case will increase the probability of escaping from an empty area. This behavior has been described from small predators like insects all the way to humans (fishing vessels).

This example not only shows the main feature of an ANN, its learning capability, but also a way of classifying situations and relating them in this case to actions. This artificial animal is forecasting and taking the best action course that will lead it to its prey.

Ecological informatics methods have great potential and are not limited to forecasting or classifying. In this example not only was this achieved, but also a representation of learning or animal behavior, and an IBM was developed at the same time.

Goals

Application

'Ecological informatics' can be successfully applied to any complex ecological problem, but it may be really effective in case data are more abundant or reliable than theoretical knowledge.

In ecological modeling, for instance, methods which stem from 'ecological informatics' should be applied (but not exclusively) when many variables are involved in the system being modeled, when some of those variables are not precisely accounted for, when they are categorical or nominal, or when nonlinear effects and/or interaction between variables are suspected to occur.

In general, 'ecological informatics' can play a relevant role when there is not enough theory to explain the dynamics of a system or the relationships between its components. This is the case, for instance, in most studies based on a bottom-up approach.

Finally, 'ecological informatics' provides several methods that are particularly useful in empirical modeling applications, that is, when one or more variables whose measurements are expensive and time consuming, information can be accurately estimated on the basis of other variables, which are cheaper and easier to measure. A typical application of this approach is in remote sensing and in the calibration of instrumental measures.

The Future

'Ecological informatics' methods are not restricted or limited to the main purpose they were developed for, they can achieve more. It all depends on the imagination of the researcher and possibilities increased by hybridizing related methods in ecological informatics/modeling.

Where will this bring us – how much closer to our understanding of ecological systems? This is something that will be answered probably by the next generation of researchers, not by the ones who developed or first applied them in ecology but by those born in the age of personal computers, with higher interrelation to computers, who will find new approaches with a new way of looking at nature and machines.

'Ecological informatics' presents a new option or approach on modeling ecological systems. This approach has been growing in the last decade but much has yet to be accomplished. It is not the purpose that ecoinformatics methods displace traditional methods, but they present

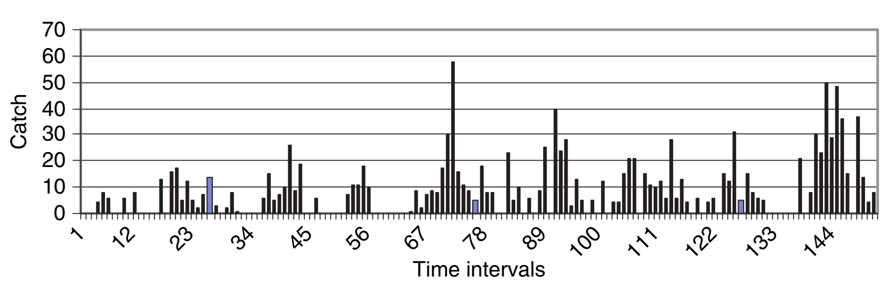

Figure 4 Prey catch by the artificial animal from starting time with a random walk and through learning stage.

another point of view in getting insight into future ecological systems.

In this scenario, 'ecological informatics' has a potential for growing as fast in ecological applications as bioinformatics grew in studies at cell or individual scale. It is not predictable, however, to what an extent and how 'ecological informatics' will evolve during the next decades. It will be certainly influenced by advances in computer science, but only our ability to deal with increasingly complex ecological problems will foster 'ecological informatics' as an independent discipline.

In fact, when new technologies or new methods are developed, their application to existing disciplines is usually regarded as a spinoff that may define a subdiscipline, and this is the present state of 'ecological informatics'. However, when a subdiscipline gains enough momentum as to become widely accepted by nonspecialists, it eventually flows back into the mainstream discipline, thus broadening its scope. We hope this will be the destiny of 'ecological informatics'.

See also: Abundance Biomass Comparison Method; Adaptive Agents; Animal Defense Strategies; Artificial Neural Networks; Cellular Automata; Ecological Complexity; Ecological Informatics: Overview; Ecosystem Health Indicators; Empirical Models; Evolutionary Algorithms; Individual-Based Models; Learning; Multilayer Perceptron; Orientation, Navigation, and Searching; River Models.

Further Reading

Dreyfus M (1999) Individual-based modelling of fishermen search behaviour with neural networks and reinforcement learning. *Ecological Modelling* 120: 287–297.

Fielding AH (ed.) (1999) *Machine Learning Methods for Ecological Applications*, 280pp. New York: Kluwer (ISBN-0412841908).

Lek S, Scardi M, Verdonschot PFM, and Descy J-P (eds.) (2005) *Modelling Community Structure in Freshwater Ecosystems*, 518pp. London: Springer (ISBN-3540239405).

Recknagel F (ed.) (2003) *Ecological Informatics: Understanding Ecology by Biologically Inspired Computation*, 425pp. London: Springer (ISBN-3540434550).

Applied Ecology

A Georges, L J Hone, and R H Norris, Institute for Applied Ecology, Canberra, ACT, Australia

Applied Ecology

The science of ecology involves the study of interactions between organisms and their environment, both biotic and abiotic, with particular focus on those interactions that determine their distribution and abundance. Applied ecology is the science of the application of ecology to contemporary problems in managing our biological resources. It includes scientific study of the effects of humans on the interactions between organisms and their environment, but excludes human ecology.

Applied ecology has two broad themes. The utilitarian theme concerns the interests of humans in their food, shelter, welfare, and health, that is, the material services the natural environment provides. Such ecosystem services, once compromised, can be very expensive to replace despite our technological advances. How do we bring ecology to bear in maintaining and improving these ecosystem services where they currently exist, in restoring or replacing them if they have been lost, or in mitigating the impact if those services are under threat? A second theme concerns nonconsumptive values of the biota, for recreation, tourism, psychological well-being, or simply because humans have an ethical responsibility as custodians of the natural environment and the species it contains. How do we bring ecology to bear in conserving these important nonconsumptive values?

These two broad themes overlap, since the nonconsumptive values of the environment are connected through biodiversity to the services healthy environments deliver. Naturally biodiverse systems are typically more resilient to human-induced perturbation than are systems that are highly modified, structurally simplified or