Hybrid Models: Lake St. Lucia

Ecological Modelling, 46 (1989) 107–119
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

A RULE-BASED ECOLOGICAL MODEL FOR THE MANAGEMENT OF AN ESTUARINE LAKE

A.M. STARFIELD [1,2], B.P. FARM [2] and R.H. TAYLOR [3]

[1] *Department of Computational and Applied Mathematics, University of the Witwatersrand, P.O. Wits 2050, Johannesburg (South Africa)*
[2] *Department of Ecology and Behavioral Biology, University of Minnesota, Minneapolis, MN 55455 (U.S.A.)*
[3] *Natal Parks Board, P.O. Box 62, St Lucia Estuary, 3936 (South Africa)*

ABSTRACT

Starfield, A.M., Farm, B.P. and Taylor, R.H., 1989. A rule-based ecological model for the management of an estuarine lake. *Ecol. Modelling*, 46: 107–119.

It is often difficult to build conventional dynamic models for an ecological system because the relationships between the abiotic and biotic components of the system are understood only in a rough, qualitative rather than a detailed, quantitative sense. This paper shows how a rule-based model can be formulated and used in this type of situation.

The purpose of the model is to assist managers of a large, shallow coastal lake connected to the sea by a narrow estuary. Fluctuations in the salinity of the lake have a marked effect on the biotic components and managers are interested in the likely effects of alternative strategies for ameliorating the salinity.

The paper shows how available information lends itself to a representation in terms of rules that indicate how important biotic components change (on a crude scale of 1 to 5) depending on the prior state of the system and current water conditions. The model was incomplete at the time of writing, but at least one important (and unexpected) result has already emerged: the abundance of underwater plant biomass is sensitive to the rate of change of salinity rather than the salinity level per se.

Several consequences of this type of modeling are noted: it draws on the experience of both scientists and non-scientists, provides a consistent, logical basis for discussion, improves communication between field biologists and managers, lends itself to an adaptive approach, and can provide assessments of the quality of each simulation.

It is suggested that this approach is pertinent whenever the effects of abiotic events dominate mutual interactions between the biotic components of a system.

CONTEXT OF THE MODEL

Lake St Lucia (28° 15′S 32°30′E) is a saline coastal lake situated at the southern end of the Mocambique coastal plain (see Fig. 1). It is 60 km long,

0304-3800/89/$03.50

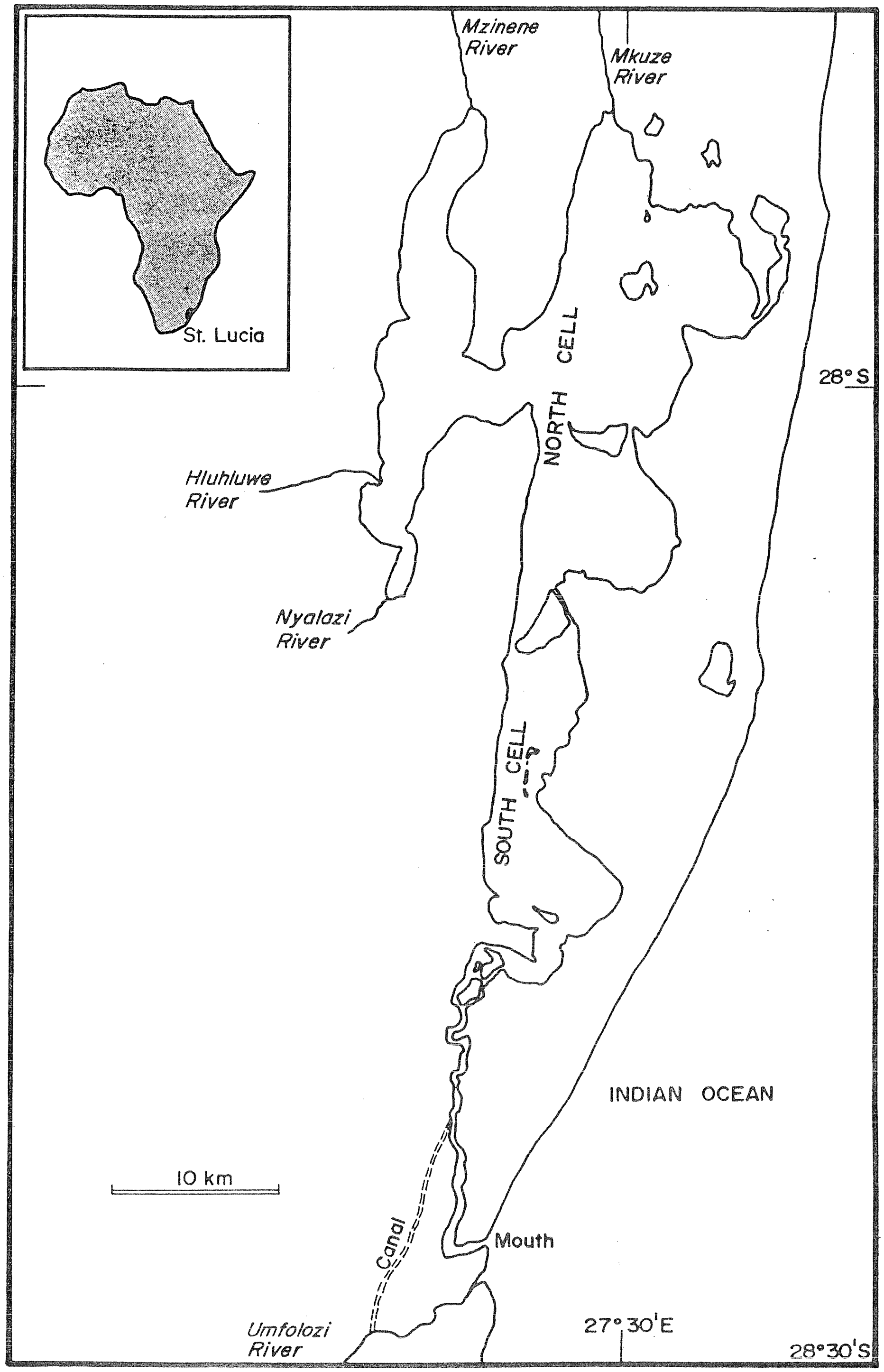

Fig. 1. Location of Lake St Lucia, the position of the estuary mouth and canal, and the division of the lake into two cells.

3–5 km wide, and has a surface area of 350 km^2. It is an extremely shallow system, most of it being less than 1 m deep.

The lake is connected to the Indian Ocean by a 21 km long narrow channel and has a mouth which may be wide open, constricted, or closed.

Since it has a high surface–area to volume ratio, it loses a considerable proportion of its water to evaporation. The lake level generally fluctuates about the mean sea level (MSL) mark. When it is above MSL, there is a net outflow of water from the lake to the ocean; when below MSL, there is a net inflow of seawater, and salt enters the system. During seasons of normal rainfall most of the freshwater entering the system is from direct precipitation and from the rivers which enter the northern extremities of the lake. A salinity gradient develops, ranging from freshwater in the north to sea-water in the south. During drought conditions, no freshwater enters the system from the rivers and water losses through evaporation concentrate the salt. Under these conditions the salinity gradient is reversed; the water with the lowest salinity is the seawater near the mouth, and during extreme conditions the salinity of the water farthest from the mouth can be concentrated to more than 3 times that of seawater.

The St Lucia ecosystem is predominantly detritus-based. The source of the detritus is reeds and submerged plants during low and moderate salinities, and phytoplankton during periods of high salinity. During low salinities the main trophic pathway is short – consisting mainly of the water plant *Potamogeton* and herbivorous birds such as ducks and coots. When salinity is closer to that of seawater, the main trophic pathways include the benthic fauna, fishes and piscivorous birds. During high salinities, much of the benthic fauna dies off, fish move out, and the dominant fauna are zooplankton which are fed upon by as many as 60 000 flamingos.

Human activities in the catchments and the alteration of the mouth have increased the sensitivity of the lake to climatic fluctuations, and the salinity changes in St Lucia are now considered to be more extreme than in the past. Changes of salinity result in different biota, and it is clear that catchment land uses or any other activities which change the hydrology of the lake can have undesirable effects. As a counter measure, the hydrology of the system can be manipulated. For example, the flow of water through the estuary mouth can be controlled by dredging, or a recently built canal can be used to divert fresh water from a nearby river.

A prerequisite for catchment regulation and hydrological manipulation is a sound understanding of physical changes in the lake and the associated biological responses. Managers wish to be able to predict, within reasonable limits, the biological consequences of any actions they may be considering.

Hutchison and Midgley (1978) developed a complex hydrological model of St Lucia which simulates the water gains and losses and the changes in salinity and lake level. This model was developed to evaluate the hydrological effects of various management options; it was never intended, however, to show the biotic outcome of the options tested.

Although there has been a significant amount of biological research conducted at St Lucia, most of it is in the form of short-term studies on specific components. There has been no coordination between studies and there is little synthesis to explain how St Lucia functions as a holistic unit; i.e., as an ecosystem. It would be attractive to try to build a model similar to Kremer and Nixon's (1978) model of Naragansett Bay, but in this case there was neither the quantitative knowledge nor the resources to develop a complex compartmental model. However, there is some knowledge of what will happen to the various components of the lake under different sets of circumstances. For example, while it may not be possible to say that 'at x salinity, phytoplankton p will be consumed with y efficiency by predator q', it may be known that if salinity rises above a certain level then the submerged plants will die back and after a short delay the numbers of ducks will drop. Observations such as this are currently used in an informal and uncoordinated fashion to influence management decisions.

Starfield and Bleloch (1986) describe (but do not implement) the idea of a rule-based qualitative predictive model. We wondered whether this approach could be used to coordinate the existing knowledge of the system in order to predict the biological state of St Lucia, in a coarse and qualitative manner, from a given history of physical conditions (e.g. salinity and lake level). The long-term objective is to combine this qualitative biological model with the existing hydrological model to make the best use of current understanding in guiding management decisions.

In this paper, we begin by outlining the differences between a conventional system model and a qualitative rule-based model. We then go on to describe parts of a rule-based model which is currently being developed for Lake St Lucia, and end by discussing what we have learnt from this exercise.

RULE-BASED MODEL

In both conventional and rule-based models, the objective is to simulate how different biotic components of the system respond over time to changes in both the abiotic components and other biotic components. In a conventional model, each component is represented by a real number variable. The conventional model uses mathematical operations in the form of difference or differential equations to calculate the amount of change in a given variable over some period of time. Examples of conventional system models are the Lotka–Volterra equations (at the theoretical level) and simulation models such as that of Kremer and Nixon (1978).

Suppose, however, that we decide to represent both the physical (abiotic) and biotic components using a small set of discrete states rather than real number variables. We now wish to predict how each variable changes from state to state over a period of time. Instead of mathematical operators and differential equations, we use logical operators in the form of IF–THEN rules to predict these changes. A set of such rules makes up a qualitative predictive model, indicating how the system would change for each combination of states. Note that in conventional models time may be either continuous or discrete; rule-based models are event driven and therefore discrete.

In the case of St Lucia, the relevant physical states are past and present water and salinity levels. As mentioned in the introduction, there is a salinity gradient in the lake. For the sake of simplicity we represent this by dividing the lake into two uniform cells (see Fig. 1). The biological states represent abundance (in discrete intervals) of functionally equivalent groups (e.g. piscivorous birds) within each cell. The model consists of rules which update these states every 3 months.

Table 1 shows how the abiotic states have been defined. Notice that the

TABLE 1

Representation of biotic and abiotic components in the model

Biotic states

State	Biomass range represented (percent of the maximum observed historically)
1	0– 5
2	6–12
3	13–25
4	26–50
5	> 50

Salinity states		Lake level states	
State	Salinity level (ppt)	State	Lake level
1	0– 4	1	low: below Mean Sea Level (MSL)
2	5–12	2	medium: 0–20 cm above MSL
3	13–25	3	high: more than 20 cm above MSL
4	26–45		
5	46–65		
6	> 65		

Seasons

Quarter	Actual time of year
1	summer (January, February, March)
2	fall (April, May, June)
3	winter (July, August, September)
4	spring (October, November, December)

ppt, parts per thousand.

resolution (as reflected in the intervals) has been carefully considered. For example, the first salinity interval (0–4 ppt) represents the range that fresh-water organisms can tolerate, the second interval (5–12 ppt) is optimal for the water plant *Potamogeton pectinatus*, and so on. For truly estuarine organisms, the stenohaline (narrow tolerance) species can survive only in state 4 (25–45 ppt), while the euryhaline (wide tolerance) species will tolerate states 3 through 5. The resolution also matches understanding: there are only three states for water level because there is no detailed knowledge of how water level influences the biotic dynamics. Note that the seasons are skewed to match available hydrological data.

The biotic components (plant and animal) are each allowed five states on a logarithmic scale, as shown in Table 1. This level of discrimination was thought to be adequate, while the choice of a logarithmic scale reflects the exponential part of a population growth curve. (The system is disturbed so often that one seldom sees sigmoidal growth.) The biotic components included in the model are the detritus, phytoplankton, zooplankton, the benthos, reeds, three underwater plants (*Potamogeton pectinatus, Ruppia cirrhosa* and *Zostera capensis*), fish, ducks (representing all herbivorous birds), flamingos, pelicans (representing all piscivorous birds) and humans. This choice of components was guided by the purpose of the model . Thus, for example, crocodiles and hippopotamuses, which do not react to the short-term lake dynamics, are omitted. The human component is included because management actions are related to recreation. There is also a bias towards birds because they are a good indicator of the state of the ecosystem.

As an example of how the IF–THEN rules are constructed, consider the rules relating to *Phragmites* reeds. The local biologist knows that the salinity tolerance of these plants is as follows:

Salinity (ppt)	Plant status
< 25	optimal, there is net growth
26–45	suboptimal, plants maintain levels with no net growth or reduction
> 46	reeds die back slowly

In addition, it is known that extreme salinities (> 65 ppt) will quickly kill the above ground portion of the reeds. If the rootstock is not flooded by the highly saline water, it becomes dormant and will grow again once conditions are less saline. If, however, lake levels are high the rootstock will be flooded and killed by the salt. After such an occurrence reeds will have to recolonize and hence regrowth will be slow. Finally, the plants are observed to grow most quickly in the spring and summer months.

With only this much information, an attempt can be made to build rules about the reed growth. Using the first conditions, with low or medium lake levels, the first rule is developed:

RULE 1:
"IF the salinity level is high, and lake level is low or moderate
THEN reeds die back within three months."

To develop the second rule, which takes into account both high salinities and high water levels, further information is needed about the rate of recolonisation after the rootstock has been killed. With a bit of prodding the biologist estimates that after the rootstock has been killed it will take at least a year for the reeds to reestablish and (conditions being favourable) reach the next level, state 2. In other words, after extremely high salinity (> 65 ppt) and high lake level, the reeds will remain in state 1 for at least a year. A rule can be formulated almost directly from this clarified observation:

RULE 2:
"IF the salinity level is high
AND the lake level is high,
THEN the reeds will remain at level 1 for at least a year."

Similarly, four more rules can be found within the above observations.

RULE 3:
"IF the salinity level is between 45–65 ppt,
THEN the reeds will die back slowly (drop 1 level each quarter)."

RULE 4:
"IF the salinity level is between 25–45 ppt,
THEN the reeds will remain at their present level."

RULE 5:
"IF the salinity level is below 25 ppt
AND the season is either spring or summer,
THEN the reeds will grow very quickly (increase 2 levels)."

RULE 6:
"IF the salinity level is below 25 ppt
AND the season is either autumn or winter,
THEN the reeds will grow well (increase 1 level)."

These rules have been written rather loosely in English. In the syntax required by the program we have developed, they would actually be written:

SET: REEDS.
RULE 1: IF salinity = 6 & (level = 1 or level = 2) THEN reeds: zero.
RULE 2: IF salinity = 6 & level = 3 THEN reeds: future4.
RULE 3: IF salinity = 5 THEN reeds: down1.
RULE 4: IF salinity = 4 THEN reeds: same.
RULE 5: IF salinity < 4 & (quarter1 OR quarter4) THEN reeds: up2.
RULE 6: IF salinity < 4 & (quarter2 OR quarter3) THEN reeds: up1.

This syntax is easy to learn: "reeds: down1", for example, instructs the program to set the reed abundance for the next period at one level below its current level; "reeds: future4" sets the level at the lowest abundance for the next four time periods irrespective of conditions during that year.

Similar rule sets can be developed for other biotic components of the system. The reed example illustrates, however, how the structure of the rules and the way in which we have represented the states provide a framework for the biologist, one he can use to make his accumulated knowledge explicit to some degree of accuracy. Without knowing all the mechanisms and interactions involved, and without detailed quantitative information about the system, we see that it is nevertheless possible to create a small model of a portion of the system.

SAMPLE OUTPUT

As an example of another portion of the model, consider the three main submerged plants in the lake: *Potamogeton pectinatus*, *Ruppia cirrhosa* and *Zostera capensis*. Growth and decay of these plants does not depend on water level, but each has a different salinity tolerance and different response time. Growth is generally fastest in fall and winter when the water is calm and clear. Under unfavorable salinity conditions, die-back is also seasonal and is faster in spring and summer (as a result of rougher wave action and increased turbidity).

Figures 2(A) and (B) show the results from two different simulations using rules based on the above information. In both simulations salinity levels were first increased then decreased; in case (A) these changes were relatively slow (over a period of 15 years), while in case (B) they were considerably faster (over a period of 7 years). Both figures show the response of all three submerged plants to these salinity changes. Both show the *Potamogeton* flourishing at low salinities, while *Ruppia* and *Zostera* respond best at higher salinities.

These responses are as expected. What is really interesting is a comparison of the two figures, which reveals that longer periods of stability in the salinity are needed to attain maximum plant biomass: the faster changes

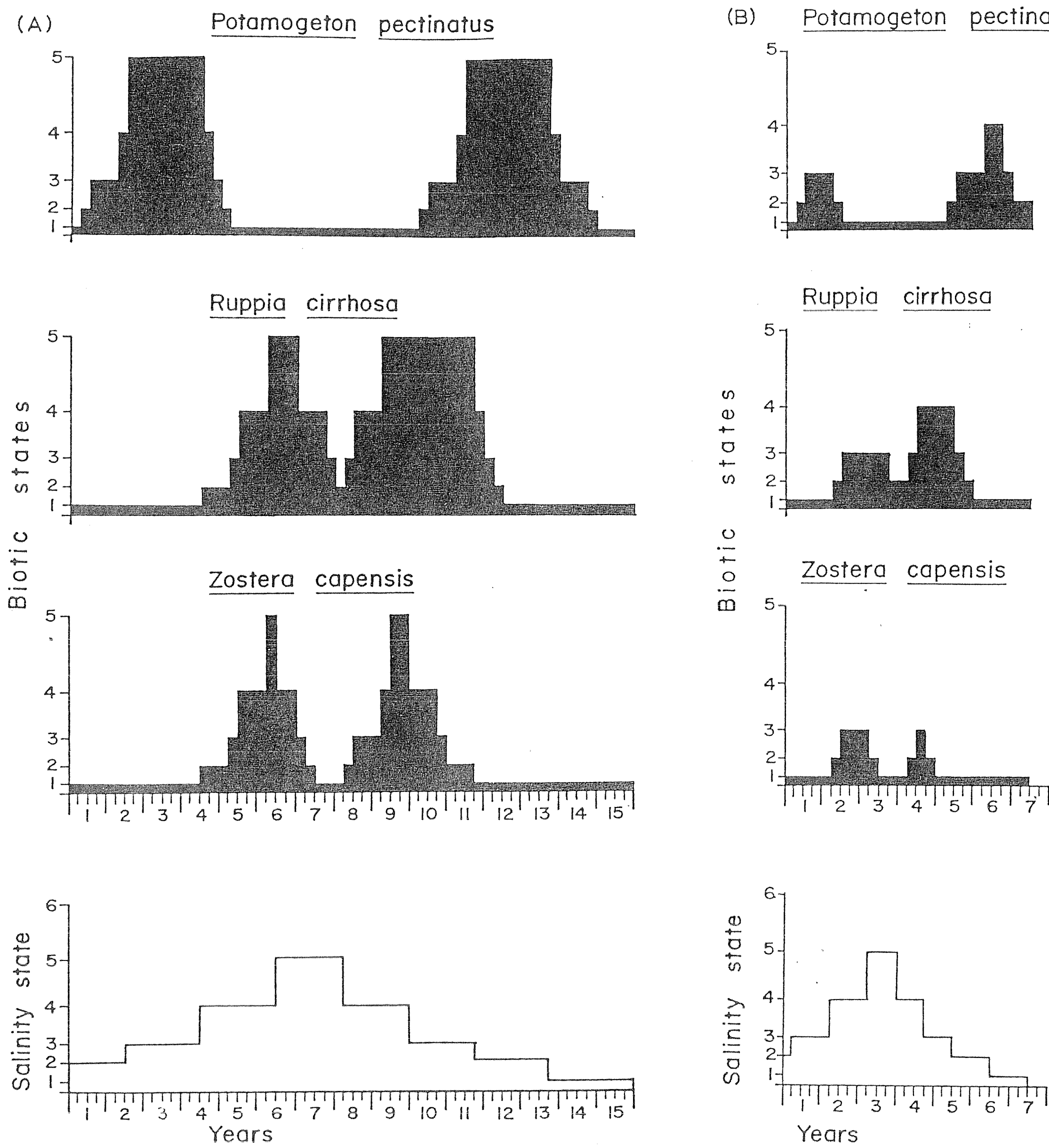

Fig. 2. Simulated response of three underwater plants to: (A) relatively slow salinity changes, and (B) faster salinity changes.

lead to noticeably lower levels of submerged plants in the system as a whole. This in turn will affect all higher trophic components when the rules for those components are added to the simulation. There will also be indirect effects because decayed plants contribute to the detritus component.

IMPLEMENTATION, TESTING AND FUTURE USE

The model has been implemented on a personal computer using a modified version of an expert system shell (Starfield et al., 1985; Starfield and Louw, 1986). The rules, as indicated previously, are grouped in sets where each set updates the state of a specific biotic component. As presently implemented, the program addresses the rules in each set sequentially, starting with the first rule and exiting the set when a rule is satisfied, i.e. only one rule in each set can be fired. This has both advantages and disadvantages. The main advantage is that one does not have to worry about the consistency of the rules provided that their order is carefully planned. The disadvantages are that one does have to be wary of the order (which makes the rule sets difficult to update) and that the number of rules in a set can expand considerably when feedback is included in the model, i.e. when the state of a variable is affected by variables above as well as below it in the food chain. The advantages and disadvantages of a program that addresses every rule in a set, and can implement more than one rule per set per time set, are currently being explored.

Two other features of the modified expert system shell have still to be implemented in the model. The first is an explanation feature which allows for explanatory text to be provided for each rule, while the second is a tracing feature which lists, in order, all the rules fired during a simulation. As explained in the discussion section below, both these features are considered essential. A third feature which will be added is a confidence rating for each rule (perhaps as a scale of 1 to 5 where 1 represents a guess and 5 represents a rule that has been validated). This can then be used in conjunction with the trace feature to provide measures of confidence for the logical pathways associated with different simulations.

The rule set is currently incomplete in the sense that the rules for some variables (e.g. the detritus) are almost purely guesswork, while the rules for others (e.g. fish) are still too sketchy. The rules have been constructed on the basis of accumulated expertise without direct reference to historical records. Intermittent records do in fact exist for the past 20 years (e.g. duck counts) and the intention is to compare the model with this data set. This cannot be considered as a validation of the model, since the same data set in a sense constitutes the accumulated expertise, but it will provide a stringent testing, and no doubt updating, of the rules.

This leads to the question of how the model will be used in the future. At this stage it should be thought of as a learning tool which at any time reflects the best interpretation of current understanding. For it to be adaptive and effective it is essential that it be used on a regular basis, that its predictions be compared with what actually happens, and that its rules be modified in the light of discrepancies. As the model develops, so confidence in it will increase and its role in guiding management decisions will expand.

The fact that the model runs on a personal computer, accessible to managers, scientists and field personnel, is crucial. The intention is to link this model to the existing hydrological model (the way in which the abiotic components have been defined makes this easy to do) but the hydrological model would first have to be implemented on a personal computer too. One way in which the linked models could then be used would be to refer back to past conditions (e.g. a period of severe drought and high salinity) and simulate what might have happened if alternative management strategies had been applied at that time (e.g. if the estuary mouth had been closed). Even with the current set of biological rules, such an exercise could provide the basis for healthy argument about the merits of alternative management actions.

DISCUSSION

Despite the fact that the St Lucia model is incomplete, it is possible to draw some conclusions about this type of modeling. We begin by listing some of the perceived advantages of rule-based modeling.

(1) An important advantage is that field biologists respond positively to a rule-based model. It provides a format which frequently corresponds with their state of knowledge of the system. This allows them to structure their knowledge, which in turn imposes discipline on their approach to understanding the system dynamics.

This is an important point. One might have a philosophical preference for conventional system models, but our experience shows that biologists who were uncomfortable with conventional models were enthused by the rule-based approach, and were eager to contribute to and argue about the rules. Their response to this form of modeling was "It fits what I know."

Managers, too, respond enthusiastically to this type of modeling. The model promotes discussion and facilitates communication between scientists and management. It also facilitates interactions between resource managers and the public in that it provides a reasoned, consistent explanation for management policy.

(2) Qualitative modeling also fits well with the type of information presently available. It is often difficult to prevent a conventional system model from becoming too detailed. The way in which the variables are defined in a qualitative model, and the way in which the rules are structured, both make it easier to build consistently at a fixed level of resolution.

(3) Quality control is always difficult in a system model. In conventional models one is forced to invent plausible relationships and, similarly, in rule-based models one is forced to invent plausible rules. However, it is relatively simple in a rule-based model to keep track of the flow of logic and the rules that have been fired. In this way, one can evaluate the quality of the conclusions reached. The workings of a rule-based model are available to the user; the workings of a conventional model are often hidden from the user.

(4) Compartmental models have proved useful for describing systems where interactions between the components are crucial. Rule-based models are likely to prove more useful for describing systems that are dominated by abiotic changes. The biotic changes in Lake St Lucia, for example, are driven by climatic fluctuations.

(5) There is often a gulf between natural resource managers and scientists because the former have to make immediate decisions while the latter are unwilling to express an opinion until they have 'completed' their research. Rule-based models can bridge that gulf, since they encapsulate the current state of knowledge. That knowledge may be incomplete or even inconsistent, but making it explicit and dynamic provides a solid basis for, at least, consistent arguments. The model ensures that everybody has the same mental picture. Moreover, a rule-based model has the ability to grow with the available knowledge. It is more flexible than conventional models and easier to update because:

- the rules which constitute the model are structured as a data file rather than a computer program;
- their syntax is user-friendly;
- the trace feature makes the workings of the model easy to follow and helps to explore the consequences of rule changes.

(6) As with any form of modeling, the process of building the model forces people to think logically, carefully, and consistently and helps to identify areas of poor understanding or insufficient information. For example, the St Lucia exercise highlighted the need for more information about detritus.

Some of the deficiencies or unanswered questions with this form of modeling are:

(1) Although we have indicated that the model is flexible and easy to update, there is still a need to develop updating procedures that maintain the integrity of the model.

(2) A sensitivity analysis is an important part of any modeling exercise and such an analysis can often be made routinely on a conventional model. There is a need to develop a parallel approach for rule-based models.

(3) Quantitative models have the advantage in that they provide an audit of what goes in and what comes out of a system. For example, Kremer and Nixon (1978) used their compartmental model to show that the pre-bloom winter phytoplankton biomass in Naragansett Bay was too small to sustain observed zooplankton levels. Rule-based models are incapable of drawing such conclusions, and it might prove useful to ask how one could superimpose a rough audit on a rule-based model. It does not follow that rule-based models cannot produce insightful results; the St Lucia model showed, for

example, how periods of stable salinity were essential for the build-up of large quantities of certain components.

(4) We have used a production rule representation because we were familiar with our expert system shell and because the biologists were comfortable with the IF–THEN rule syntax. We have not asked whether there are better artificial intelligence tools for this type of modeling.

The St Lucia model looks promising but can only be regarded as successful once it has been integrated into the management decision-making process. This will take some years to evaluate. In the meanwhile, it should be apparent that rule-based models have a role to play: they provide a means for accessing useful information which would be unsuitable for a conventional model.

ACKNOWLEDGEMENTS

The authors would like to thank A. L. Bleloch, M. Sears and K. Alman for their useful comments, and Glenn Radde for presenting the paper on their behalf.

REFERENCES

Hutchison, I.P.G. and Midgley, D.C., 1978. Modelling the water and salt balance in a shallow lake. Ecol. Modelling, 4: 211–235.

Kremer, J.N. and Nixon, S.W., 1978. A Coastal Marine Ecosystem: Simulation and Analysis. Springer, Berlin.

Starfield, A.M. and Bleloch, A.L., 1986. Building Models for Conservation and Wildlife Management. McGraw-Hill, New York, pp. 231–232.

Starfield, A.M. and Louw, N.J., 1986. Small expert systems: as perceived by a scientist with a computer rather than a computer scientist. S. Afr. J. Sci., 82: 552–555.

Starfield, A.M., Adams, S.R. and Bleloch, A.L., 1985. A small expert system shell and its applications. In: Proc. 4th Int. Phoenix Conf. Computers and Communications. IEEE Computer Sociery Press, Silver Spring, MD, pp. 262–267.