INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

SMR.478 - 30

THIRD AUTUMN COURSE ON MATHEMATICAL ECOLOGY

(29 October - 16 November 1990)

"Personal Computers in Ecosystem Modelling"

A. A. VOINOV
Academy of Sciences of the U.S.S.R.
Institute of Atmospheric Physics
109 017 Moscow
U.S.S.R.

These are preliminary lecture notes, intended only for distribution to participants.

PERSONAL COMPUTERS IN ECOSYSTEM MODELING

A.A.Voinov
(Inst.of Atmospheric Physics, USSR Academy of Sciences)

ABSTRACT

Mathematical modeling in ecology is an essentially interdisciplinary subject and it calls for joint efforts of specialists in the fields of mathematics,computer science and those representing natural sciences - biologists, ecologists, chemists.Not always is it easy to achieve a mutual understanding between different branches of science. With the advent of personal computers mathematical and systems reasoning started to penetrate directly into ecological media, stimulating interdisciplinary communication. An outlook of various applications of PCs in different aspects of ecological studies is presented. Special attention is payed to systems of automatic modeling and programing.

In contrast with the mainframes or even minicomputers, personal computers can be operated by the user directly with no intermediate links provided by programers, operators etc. As a result very quickly they became fairly widespread in very different branches of science. Their application by biologists, ecologists and other natural scientists, at first was generally confined to data storage and processing, which is quite natural since actually these are the main tasks in the fields dealing with experimental studies. On the other hand, it can be noted that the first user-friendly packages supplied to the market were data bases (dBase, etc.) and spreadsheets (Lotus, SuperCalc, etc.). They were quickly absorbed by the public and served as efficient propagators of PCs in natural sciences.

After data is accumulated and stored, obviously it needs to be processed. The methods of data processing traditional for experimental sciences come from statistics. Some simple statistical methods have been already included into the spreadsheet programs. Further on an abundance of different statistical packages flooded the market serving very special tastes and preferences of natural scientists. Starting from very simple tutorial-like packages (SYSTAT), more complicated ones (Abstat, Statgraphics) and up till the sophisticated and powerful systems (NTSYS, BMDP, etc.) all these products are again characterized by user-friendly and fool-proof interface.

Data processing is a useful step in ecological modeling and application of PCs helped a lot at the preliminary stages of modeling concerned with storage analysis and sorting of data. However the modeling stage itself is still hardly perceived by ecologists with no special background in system analysis and programing. There is a number of packages actually made of one or several models with a user-friendly shell (LAKE, RAMASa, etc.), which can be operated by laymen in programing. But most of the models realized on PCs are simply transfered from larger computers and are far from being within the capacity of an average ecologist.

On the other hand any successful modeling project needs joint efforts of specialists. While ecologists are assumed to provide conceptual models describing the system structure and processes, identifying the parameters, mathematicians and computer scientists formalize the qualitative schemes and definitions, provide computer codes and run the resulting models. Next all of them together analyze the results and improve the model. Their should be close interaction and good mutual understanding within such a multidisciplinary team, which is not always easily achieved.

Note that simulation modeling, should be considered just as another means of representation of experimental information for formulation and verification of hypotheses about operating principles for ecosystems and their separate units. The only great advantage of simulation modeling in contrast to other methods of data handling (statistics, etc.) is that models incorporate not only the quantitative information obtained from experiments, but also the conceptual qualitative insights of researchers, which they formulate when creating the conceptual model of the ecosystem. From this point of view the efforts of non-mathematicians seem to be much more "artistic", while the systems part of the investigation is rather routine: more or less standard formalizations, programing, calibration, etc.

Quite naturally there appears the idea of automatization of the most standard stages of modeling procedure. If the formalization and programing stages are automatized, nothing stops naturalists from using and creating simulation models by themselves. This actually gives a new means of interdisciplinary communication: through a system of computer formalized languages, which would be equally understandable to representatives of different disciplines, and could interpret knowledge from various branches of science in terms of some common formalism.

Actually it turns out that perception of methods of simulation modeling by poorly mathematized natural sciences - biology, chemistry, ecology - is very much impeded by a lack of an interface, which could, just as in the case of statistics, provide workers ignorant of mathematics and programing a means to apply mathematical methods of analysis, allowing them to build and modify mathematical models all by themselves.

A number of systems of automatic programing have appeared lately. But only a few of them provide the user with a friendly interface. It is most natural to classify the known systems by arranging them along some axis which lies between the computer codes at its one extreme and some narrow problem-oriented languages or realizations of concrete models at the other (Fig.1).

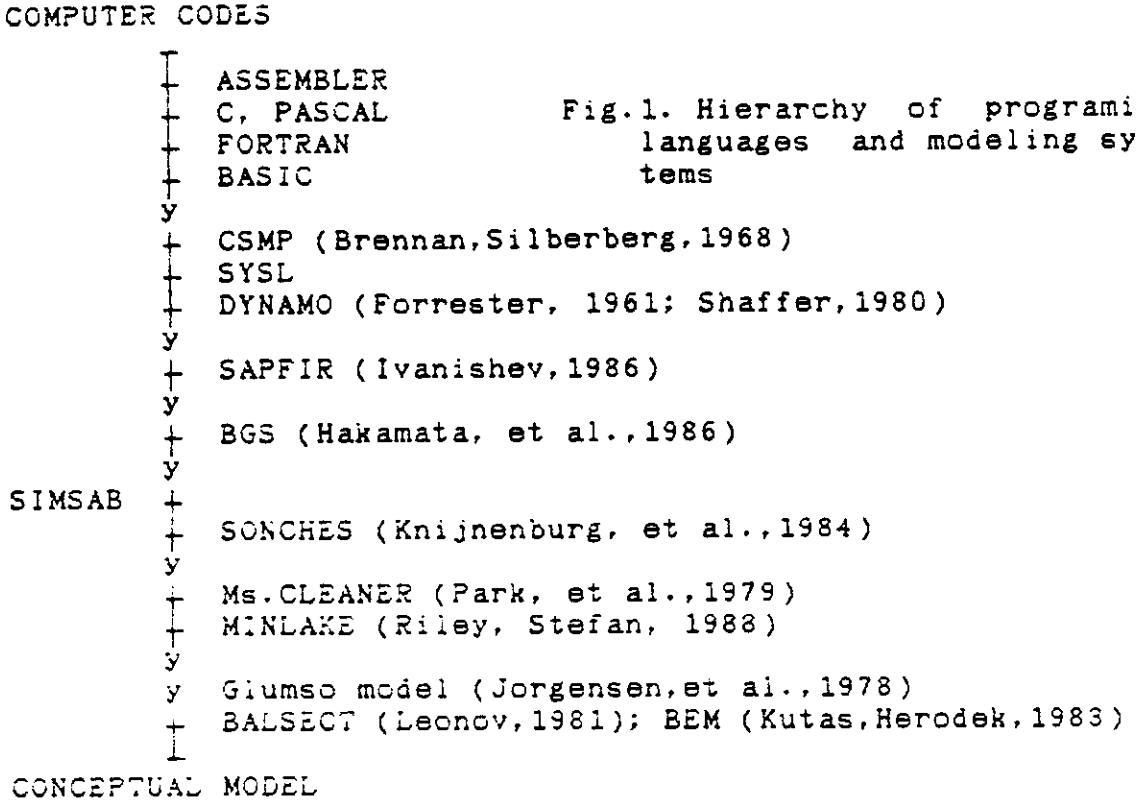

Fig.1. Hierarchy of programing languages and modeling systems

Any attempt to simplify interaction between man and computer results in losses in generality. And it is a very delicate problem to find the optimal balance between simplification of formulations and generality of the language. Since there is a continuous spectrum of possible solutions to this problem and since the delicate balance is mostly a matter of user's taste it is only natural that more and more modeling languages and systems appear.

In this paper we consider an example of one such system of automatic modeling - the Simulation Modeling System for Aquatic Bodies (SIMSAB). The Modeling system is actually an attempt to formalize the procedure of modeling, which is still rather an art than a science. SIMSAB can be placed somewhere in the middle of the axis in Fig.1. However this is not quite precise, since the system possesses features of both computer and problem-oriented languages. On the one hand it can be treated as an expansion of FORTRAN, because any kind of FORTRAN constructions are valid in SIMSAB. On the other hand, if FORTRAN is set aside, the present version is quite narrowly oriented on aquatic modeling, and this makes it closer to the other end of the axis.

Being most loosely linked to postulates of a certain school of mathematical modeling, the modeling interface is flexible enough to cover models of different kinds. Those principles which make up the invariant part of the system are clearly identified and open for criticism.

Since one of the goals of SIMSAB is to bring methods of simulation modeling into the biological media,naturally it is accomplished by a number of auxiliary programs which give an overview of some basic principles of simulation modeling in general and describe basic SIMSAB postulates in particular, allowing even a layman to master this system. SIMSAB-helper provides the user with all necessary information when working with the system.

Suppose we have a conceptual model. This means that knowing the purposes of the model, we have bounded our ecosystem in space and time, we have identified the variables, which are the ecosystem components of interest provided with experimental time series, and we have defined the links between these ecosystem components. Besides, we have determined the interactions between the ecosystem and its environment by identifying the forcing functions which specify the time-variable effects that are independent of the ecosystem. These are the functions that affect the ecosystem with no feedback from it. Our conceptual model also includes the spatial representation of the ecosystem, which means that we have decided, whether the ecosystem is spatially homo- or heterogeneous and we have chosen how to represent its spatial distribution. And finally we assume at least a qualitative description of the interactions between the variables. All this is actually what we call a conceptual model.

The stage of formalization of the conceptual model assumes that various qualitative processes are quantitatively interpreted in mathematical terms. All the variety of interactions in an ecosystem can be classed into three levels of complexity. Firstly it is the functions, which define some simple interactions explicitly determined by some internal or external factors in the ecosystem. An example can be presented by the temperature limitation function, which determines the effect of temperature upon the rate of some processes, or the so-called trophic functions, which define the uptake rate depending upon the predator and prey concentrations.

Secondly it is the functionals, which may depend upon several processes or flows in the ecosystem and define several flows at a time. An example of such a functional is the formalization of the Liebig's principle of limiting factors. In the model this functional is to determine the synchronized flows of limiting nutrients into an organism, depending upon the potentially available flows of separate nutrients.

And finally we can distinguish processes, meaning large-scale transformations in the water-body, such as the formation of wind-induced currents (the hydro-dynamics), or the material exchange between the water-body and the watershed or among different parts of the water-body (mixing).

In terms of this classification any interaction between ecosystem components can be presented as a superposition of some functions and functionals while various spatial or temporal irregularities may be described as some kind of processes. It is most common to present an ecosystem model in terms of equations, which right-hand sides actually contain the necessary functions (see for instance Jorgensen, et al., 1978; DiToro et al.,1980 and many others). Such presentations seem to be quite natural and self-explanatory and the SIMSAB flow-language actually follows the lines of this kind of formalism. Instead of a detailed formalization when all functions should be further described as it is usually done and all parameters specified, SIMSAB provides a means of qualitative description in terms of functions, functionals and processes, so that it is enough for the user just to know the qualitative effect of various terms, and he may not bother about the mathematical representation of the appropriate functions. The necessary functions will be substituted automatically and the needed parameters will be searched in a special parameter's library and inserted into the appropriate function.

In fact there is a number of functions which prove to be rather adequate and go over from one model to another. The light limitation functions of Steele (1962) and DiToro (DiToro, et al.,1971) are one such example. Another example is the variety of trophic functions, starting from the Volterra function (1927), the functions from the theory of enzyme kinetics and their analogies from population dynamics presented by variations of the Mono and Michaelis-Menten functions. Likewise Jorgensen (1980) presents a list of temperature limitation functions encountered in models.

Practically, it turns out that we can make a list of functions, which would cover most of the interactions used in models. As a result we find out that a limited number of functions can represent practically all the necessary ecological interactions. The same can be said about functionals and processes. However naturally to keep in pace with ecological developments the resulting system should be open for modifications and expansions.

Spatial structuring.

Let us now look at the spatial representation assumed by the conceptual model. Generally speaking, spatially heterogeneous ecosystems can be modeled by systems of equations in partial derivatives. However since such models need very much CPU time for model runs and require extremely extensive experimental backup, there is only a few examples of full scale case studies based on models of this type (e.g. Krapivin,1980; Venice LAgoon ...).

Much more wide-spread are the so called compartmental models, which assume that the whole water body can be split into a number of ecologically homogeneous segments (compartments) (Chen,Orlob,1975). Within each of the segments ecosystem components are presented by spatially averaged variables (numbers, concentrations). It should be noted that such a structuring is quite natural for a practical worker, who usually deals with data measured on a fixed network of stations. In this case it is only natural to think that each of the stations (especially if it produces data significantly different form those observed at the adjacent ones) describes an ecologically uniform region and thus represents a certain segment of the model.

In SIMSAB the segmental presentation of spatial irregularities is assumed. However within horizontally homogeneous segments vertically distributed models can be constructed. In fact in aquatic ecosystems usually the spatial scale of irregularities in the horizontal dimension by far exceeds that in the vertical one. In special cases the distributed dimension can be turned in order to model 1-D horizontal ecosystems (e.g. rivers). In order to bring together the representations of different segments within the framework of a single model, ecosystem components modeled by variables in segments are to be recalculated according to the intra-segmental material exchange provided by advective flows (currents induced by wind, outflows, inflows, sedimentation, etc.) and dispersive fluxes (turbulent and molecular diffusion). All these flows can be either input into the model as forcing functions, or calculated within special hydrodynamic models, which are a part of the system.

Temporal discreteness of processes.

After considering the model organization in space, let us look at its temporal organization. In SIMSAB it is assumed that a sequence of events can be distinguished on the time axis, each event standing for some processes realized in the ecosystem. These processes are assumed to be realized instantly, and no processes occur in between. Then the functioning of an ecosystem can be described as follows: at time, say, t boundary segments receive a nutrient load; next, for instance, at time t+dt concentrations of various components are altered due to material transfer along trophic chains or in hydrobiochemical cycles (ecological block); then at time t+3dt the wind induced currents are formed, determining the water exchange between segments (hydrodynamic block); next at time, say, t+7dt the whole water body is mixed up by water exchange and diffusive fluxes between segments, correspondingly changing the concentrations of ecological components; then at time t+10dt again some nutrient load enters the water body, and so on. Note that, generally speaking, the contents of each of the events is arbitrary.

General structure of SIMSAB.

The resulting structure of SIMSAB is presented in Fig.2. Within a segment material transformations are described by a system of ordinary differential equations. In order to formulate them one should specify their right-hand sides. This is performed in terms of a special flow language which operates with various functions and functionals contained in the SIMSAB library of functions.

The syntax details of this language can be found elsewhere (Akhremenkov,1988), here we just outline the main points:

- Q[X,Y] stands for the flow from variable X to Y;
- QIN[X] and QOUT[X] stand for the inflow into X from without of the system and outflow from X to without, respectively;
- Q[*,X] and Q[X,*] stand for the sums of all flows entering X and leaving X, respectively;
- $ denotes a SIMSAB library function, for instance, $FT2(T,X), $RS(X,Y), etc.;
- #VAR:NAME.EXT is a model parameter: VAR (optional) stands for the variable name, NAME is the parameter name and EXT (optional) is an extension, which can be another variable name for parameters in binary interactions, or any other identifier;
- if a SIMSAB function, returning several values, is used, then these values will be assigned to the flows in curly brackets {..} in the left-hand side.

As for the rest, the basic FORTRAN syntax rules are adopted ("C" starts the comments, entries start from the 6th position, etc.).

The resulting differential equations of the model which will be composed by the system itself have the form

$$dXi/dt = \sum_{j=1}^{n} (Q[Xj,Xi]-Q[Xi,Xj]) + QIN[Xi] - QOUT[Xi],$$

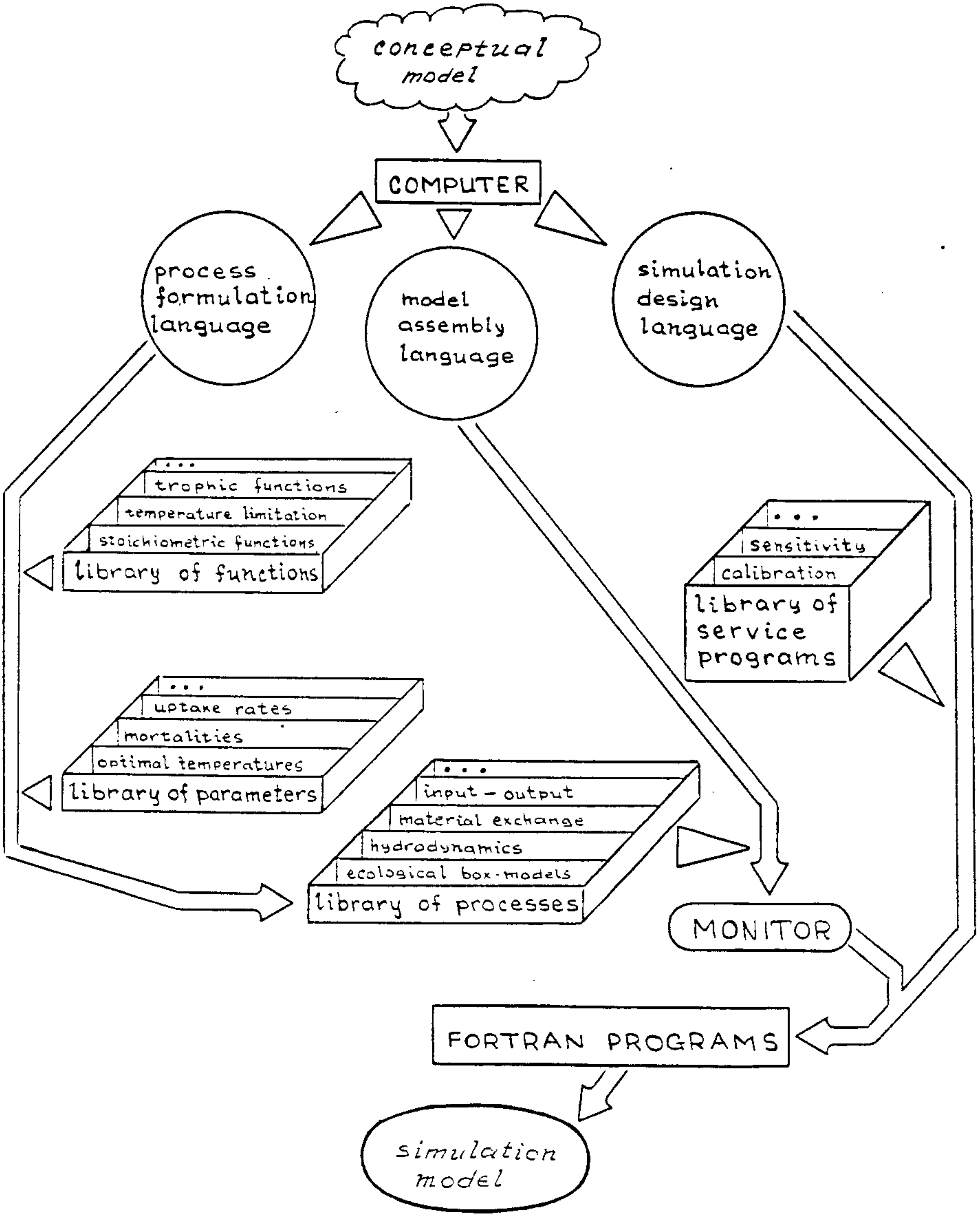
conceptual model
COMPUTER
process formulation language
model assembly language
simulation design language
. . .
trophic functions
temperature limitation
stoichiometric functions
library of functions
. . .
uptake rates
mortalities
optimal temperatures
library of parameters
. . .
input – output
material exchange
hydrodynamics
ecological box-models
library of processes
. . .
sensitivity
calibration
library of service programs
MONITOR
FORTRAN PROGRAMS
simulation model

Fig. 2

$i=1,\ldots,n$, where n is the number of variables.

SIMSAB allows one to build models operating with qualitative, conceptual ideas about ecosystem material transformations. This means that restricting oneself to the functions from SIMSAB library, one can describe the processes in the ecosystem as a superposition of various elementary processes, which are already mathematically formalized and programmed.

Generally speaking, the right-hand sides can be formulated in a file as a user-written FORTRAN program. If you apply means of SIMSAB, then the right-hand sides are calculated as sums of flows, specifying the material transfers from one ecosystem component into another. These flows may be described either in terms of SIMSAB, using its library of functions, or as fragments of a FORTRAN-program. In any case, SIMSAB results in a number of FORTRAN subroutines, calculating the right-hand sides of systems of differential equations, presenting various segments of the water body. The user who does not know any FORTRAN is restricted in his model formulations by the SIMSAB library of functions, which incidentally includes all of the most widely used functions.

The formulations of the SIMSAB flow-language are automatically processed to produce a FORTRAN program. All the parameters are automatically inserted in the appropriate places. In order to do this the system makes use of a specially structured file storing the information about various parameters encountered in ecosystem modeling of water bodies. This file is the library of parameters.

The SIMSAB-monitor brings together all the processes, which are subroutines, including those which compute the wind-induced currents, the material exchange between segments and within a segment due to the flows between different ecological components (variables) defined by the differential (difference) equations, and those subroutines which provide input of data and explication of results. In the monitor we actually define a number of time loops which control the sequence of various events presented by the model. There is also a number of standard functions used by the monitor.

SIMSAB assumes that some of the monitor functions can be specially predefined by the user. The formulations of the right-hand sides is actually the special definition of those monitor functions, that solve systems of ordinary differential equations ($RUNGE, $ADAMS or $EULER). Likewise the formation of a file containing the matrix of the depths of the water body and the wind and flow scenarios according to a special pattern can be considered as the predefinition of the hydrological monitor function ($HYDRO), which calculates the patterns of wind-induced currents and material exchange between segments.

Just as in the case of the right-hand sides, the monitor can include any of the user-prepared FORTRAN fragments and subroutines. In what follows an example is presented to show how models are formalized in SIMSAB.

```
MODEL A
VAR A,Z,D,N,P;
FORC T;
PARAM IRCUT.PRM;
C  uptake of nutrients by phytoplankton with limiting by N
C  or P: $LIM1 returns two values, two flows are defined  :
     (Q[N,A],Q[P,A])=$FT2(T,A)*$LIM1(A,N,P,$R2(N,A),$R2(P,A))
C  uptake of carbon described as an external flow :
     QIN[A]=Q[N,A]/#A:C.N*#A:C.C
C  grazing of zooplankton on phytoplankton :
     Q[A,Z]=$FT2(T,Z) * $R2(A,Z)
C  metabolic losses and mortality :
     Q[A,D]=Q[*,A]*#A:MB + A*#A:MOR
     Q[Z,D]=Q[*,Z]*#Z:MB + Z*#Z:MOR
C  decomposition of detritus into nutrients :
     (Q[D,N],Q[D,P],QOUT[D])=$FT4(T,D)*$ST1(D,N,P,C,$DES(D))
END

MONITOR MOD A;
INITIAL
C   define the initial values of variables :
      $READ(0.5, 0.1, 2, 0.01, 0.005) VAR
C   period of simulation is 60 days :
DYNAMIC TEND=60;
ST=0,7;
C   input of temperature starting from time 0
C   and every 7 days :
      $READ(TEMP) T
ST=0,0.5;
C every 12 hours (0.5 days) :
C   output of variables as graphs on the monitor:
      $GRAPH A(0,2), Z(0,2)
```

```
C   compute the system of equations by Adams method :
       $ADAMS A,VAR,FORC,5
END

DATA
TEMP:  3, 8.5, 13, 18, 14, 17, 13, 11, 20;
```

A model thus formulated is automatically translated into FORTRAN. Thus complicated models of high dimensionality can be analyzed on big computers, while the assembly of programs and analysis of simple models is performed on personal computers.

As against mathematics, natural sciences are much more qualitative in their methodologies and speculations. It is the strict reasoning of mathematics, which is hardly conceivable by naturalists and hinders the mutual efforts in ecological studies. However if we assume certain postulates, the qualitative reasoning can be at least partly automatically translated into quantitative formulations good for computer or mathematical analysis. This is actually the first goal of SIMSAB. The second one is to secure modellers themselves from technical errors in programs and to free them from some routine programing works.

The second computer revolution, that started with the advent of the personal computer, brought the computing and information processing facilities directly to the natural and humanitarian scientist. Special modeling packages can bring together ecological and mathematical studies in an interactive mode to obtain new insights into principles of ecosystem operation and more adequate representations of ecosystem dynamics.