Netherlands Journal of Sea Research
33 (3/4): 247-260 (1995)

A STRUCTURE AND METHODOLOGY FOR MARINE ECOSYSTEM MODELLING

J.C. BLACKFORD and P.J. RADFORD
Plymouth Marine Laboratory, Prospect Place, West Hoe, Plymouth PL1 3DH, UK

ABSTRACT

Comprehensive ecosystem models involve aspects from each of the biological, chemical and physical sciences. Consequently such models, for example the European Regional Seas Ecosystem Model (ERSEM), are often based on computer code written by a number of different research groups each with a well-defined area of responsibility. To construct a meaningful integrated model capable of combining the expertise from each of these disciplines requires a formalized ecosystem structure that provides a template for model development. Assumptions made at this stage will to a large extent determine the qualities of the resulting model. Essential too is a system of quality management that assures a correct and consistent end product. Such a system, however, requires the flexibility to allow the various contributors the freedom to pursue individual and novel representations of ecosystem function, if the model is to be ground breaking. The system adopted by the ERSEM group, which should have a wider applicability than marine ecosystem modelling, is described here. Finally, the methodology used to gauge the relative merits of alternative ecosystem representations is described.

1. INTRODUCTION

The purpose of this paper is to explain the organizational structure and modelling techniques that are implemented to facilitate the construction of the European Regional Seas Ecosystem Model (ERSEM). Much of the background material explained here will be of great benefit to readers in understanding the interrelation with the other papers about ERSEM.

The ERSEM project utilizes the skills of nine participating institutes (Fig. 1). The main objective of this group is to produce a generic shelf sea ecosystem model which will dynamically simulate the cycling of organic carbon, nutrients and oxygen over a seasonal cycle initially applied to the North Sea. Its activities are organized by a representative committee of project partners. Regular workshops throughout the project ensure a coordinated approach and cross fertilization of ideas.

Ecosystem function is divided into a number of 'building blocks' with each group in the project responsible for modelling one particular aspect of the system. Computationally each building block is expressed as a submodel/module. In such a scheme the interfaces between each element become the essential glue that binds the model together. With each group simultaneously working on their particular module and regularly submitting improvements for consideration, there is a need not only to formalize the procedure by which new modules are judged and assimilated, but also a requirement to ensure that each group is working from the same developmental base or standard model at any one time.

The concept of the standard model, which comprises the best co-functioning set of submodels/modules, is key to the project's progress. Without this coherent basis, the ability to join the various modules would become seriously compromised and comparisons between sites invalid. The standard model is regularly updated, approximately twice yearly, and incorporates all validated module improvements submitted to Plymouth Marine Laboratory, the partner charged with overall model development.

The initial standard model within ERSEM was based on the Ems model (Baretta & Ruardij, 1988). This was applied to the ERSEM compartments of the North Sea with forcing functions and a physical transport model relevant to the area under question.

2. LOGICAL STRUCTURE OF ERSEM

For the purpose of the ERSEM model the (North Sea) ecosystem is currently divided into ten boxes based on the ICES divisions of which five are considered as stratified (Fig. 2). Both the pelagic system and the benthic system are modelled. Light and temperature are supplied as forcing functions and the model is designed to be driven by water transport terms derived from an independent three-dimensional hydrodynamic model constructed by Backhaus (1985) predicting the dynamics for the years 1988 and 1989.

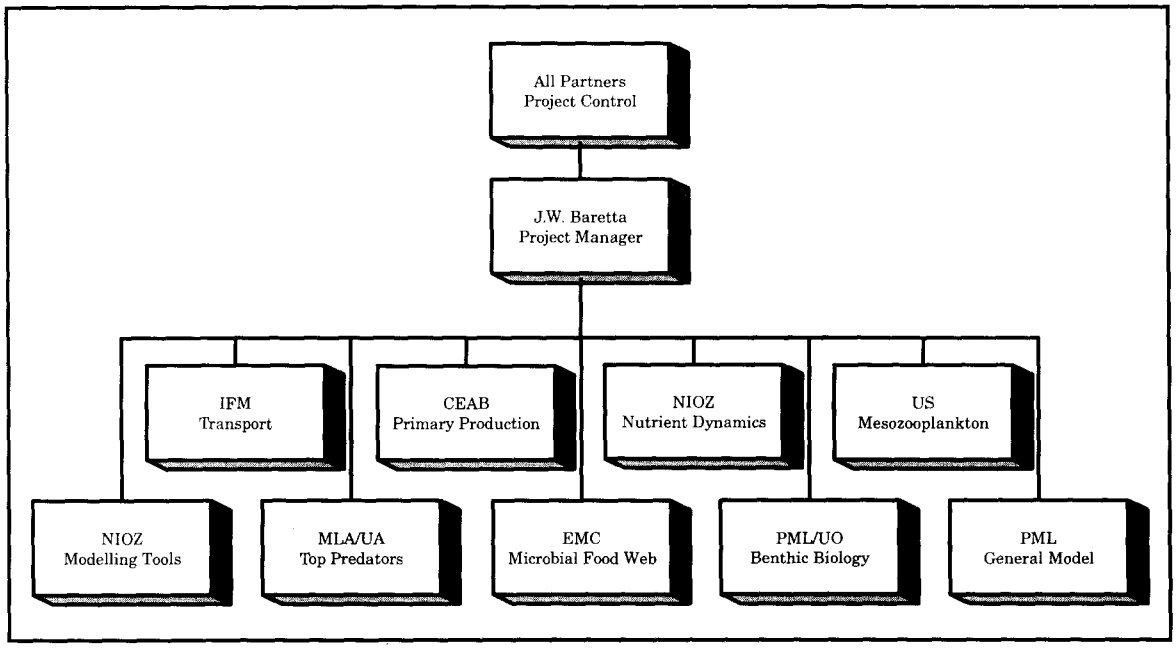

Fig. 1. ERSEM project organization. Project participants: CEAB (Centro d'Estudis Avançats de Blanes); IFM (Institut fur Meereskunde Hamburg); MLA (Marine Laboratory Aberdeen); NIOZ (Nederlands Instituut voor Onderzoek der Zee); PML (Plymouth Marine Laboratory); UA (University of Aberdeen); UO (Universität Oldenburg); US (University of Strathclyde); EMC (Ecological Modelling Centre).

ERSEM represents the ecosystem as a group of parallel ordinary differential equations (ODEs), solved as an open ended recursive system using continuous system simulation techniques. This method of representation, which can be traced back to the work of Forrester (1961), was introduced into the field of ecology by De Wit & Goudriaan (1974) and applied to marine systems by Radford (1979) and Baretta & Ruardij (1988) to estuarine and near-shore aquatic ecosystems. The essence of the approach is first to define a group of variables which together represent the state of the whole system. The strength of this approach is that the rates of all the individual processes over any small time interval may then be defined simply as functions of the state of the system at that time as expressed by the instantaneous values of the state variables modified by external forcing. This enables the large number of individual processes to be conveniently subdivided into groups which may form independent modules linked only by the state variable matrix. This feature of the method facilitates the modular approach that has been adopted for ERSEM.

The material composition of the state variables and of the transfers which interlink them may be any convenient tracer (*e.g.* energy). A comprehensive representation of the marine ecosystem requires the inclusion of the parallel cycling of carbon, nitrogen, phosphorus, and silicon, wherever each is present. For some states the C:N:P:Si ratio is considered as fixed and only the carbon content is explicitly modelled. However, variable carbon:nutrient ratios are considered as important aspects of other processes and states. Thus several of the fifty or so state variable types and their related groups of processes comprise some or all of these four elements as appropriate, making a total of approximately seventy individual state variables in the current model.

The choice of state variables is governed by a desire to keep the model as simple as possible without omitting any component which would exert a large influence on the balance of energy flow. Broad functional groups, based on trophic level, were selected which are subcategorized into states according to well-established procedures, often based on size classes or feeding method. This assists the mapping of the model against existing data. However, the exact nature of the set of state variables developed with the modelling experience. Whilst the functional groups have tended to be constant their constituent states have changed. For example there has been redefinition of feeding types in the zooplanktonic and zoobenthic populations and extensions to age-structured populations in the higher trophic levels.

Each module is required to model the processes mediated by one or more state variables. Fig. 1 and Table 1 detail the grouping of state variable types into modules and the project partner with the primary responsibility. Only those processes actuated by the particular state variable are included in that module.

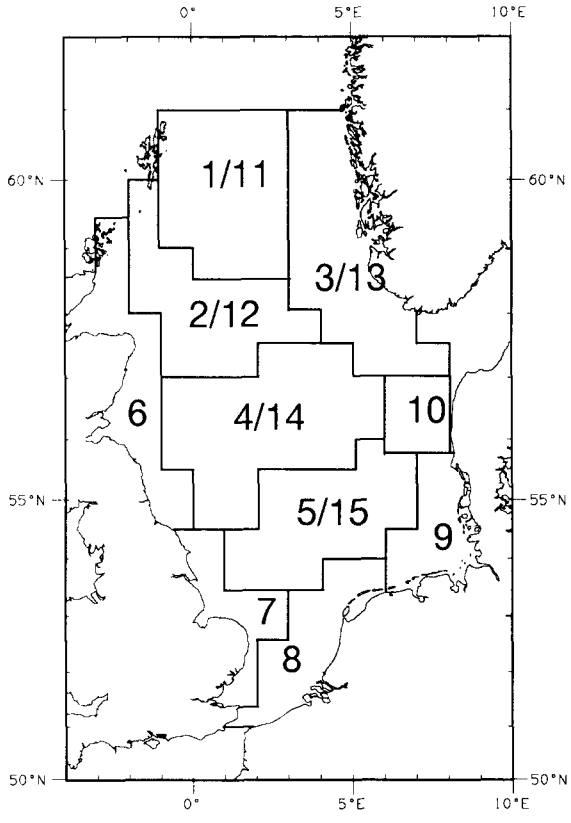

Fig. 2. ERSEM North Sea box structure.

For example the description of predation by a higher trophic level belongs to the module describing the predator rather than the prey. In some senses therefore the construction of the model can be seen as hierarchical. As a simplification higher trophic levels may be stripped off and replaced by constant mortalities without necessitating changes in the modules describing the lower trophic levels.

Fig. 3 gives a simplified representation of the model, as it has evolved towards the end of the three-year period of research. Each box in the figure represents one module and lists the constituent state variable types of that group, with the exception that the pelagic bacteria are considered along with microzooplankton as the microbial loop submodel and benthic bacteria have been combined for convenience with the zoobenthos. The lines joining the boxes represent each and every process included in the model, although of course the processes must be individually specified for each constituent state variable. Additionally the processes for each nutrient component may be modelled either explicitly (for states with variable nutrient ratios) or implicitly (for fixed nutrient ratios).

The grouping of state variables introduces the concept of the standard organism (Baretta & Ruardij, 1988). Biological state variables can be thought of as essentially three types, producer, consumer and decomposer, each of which may be defined as having a standard set of processes (Ebenhöh & Kohlmeier, 1993). Whilst during the period of the ERSEM project there has been no onus on modellers to conform to a precise template for modelling the bio-states, the concept of a standard set of processes helps greatly in assuring the completeness of modules. It is important, however, to allow for variations and additional processes specific to particular state variables. The age-structured representation of the higher trophic levels provides an example. The standard processes associated with each type are listed in Table 2.

3. QUALITY ASSURANCE

In addition to the necessary decisions which relate to the structure and component parts of the model itself, several other standardizations and procedures are needed to provide a coherent structure through which the modelling effort can progress whilst ensuring a quality result. These can be grouped under the categories of: computing environment, modelling tools, code conventions, coding practices, acceptance criteria version control and documentation.

3.1. COMPUTING ENVIRONMENT

The prerequisite for intergroup model homogeneity is to ensure that all partners use the same hardware and software environment in which to develop the model. Although computer environmental differences should only produce tiny variations in quantitative behaviour, modern developments in mathematics show that there are circumstances in which such dif-

TABLE 1
The allocation of modules and their active constituent state variables amongst the ERSEM project partners.

partner	*submodel/module*	*state variable types*
IFM	transport	-
CEAB	primary production	phytoplankton (P1, P2)
EMC	microbial food web	bacteria (B1), microzooplankton (Z5, Z6)
US	mesozooplankton	mesozooplankton (Z3, Z4)
MLA/UA	top predators	fish (FD, FP, FV)
NIOZ	benthic nutrients	benthic nutrients (M1-M5, M11, M14, K1-K6, K11, K14)
UO/PML	benthic biology	zoobenthos (Y1-Y5), heterotrophic bacteria (H1, H2)

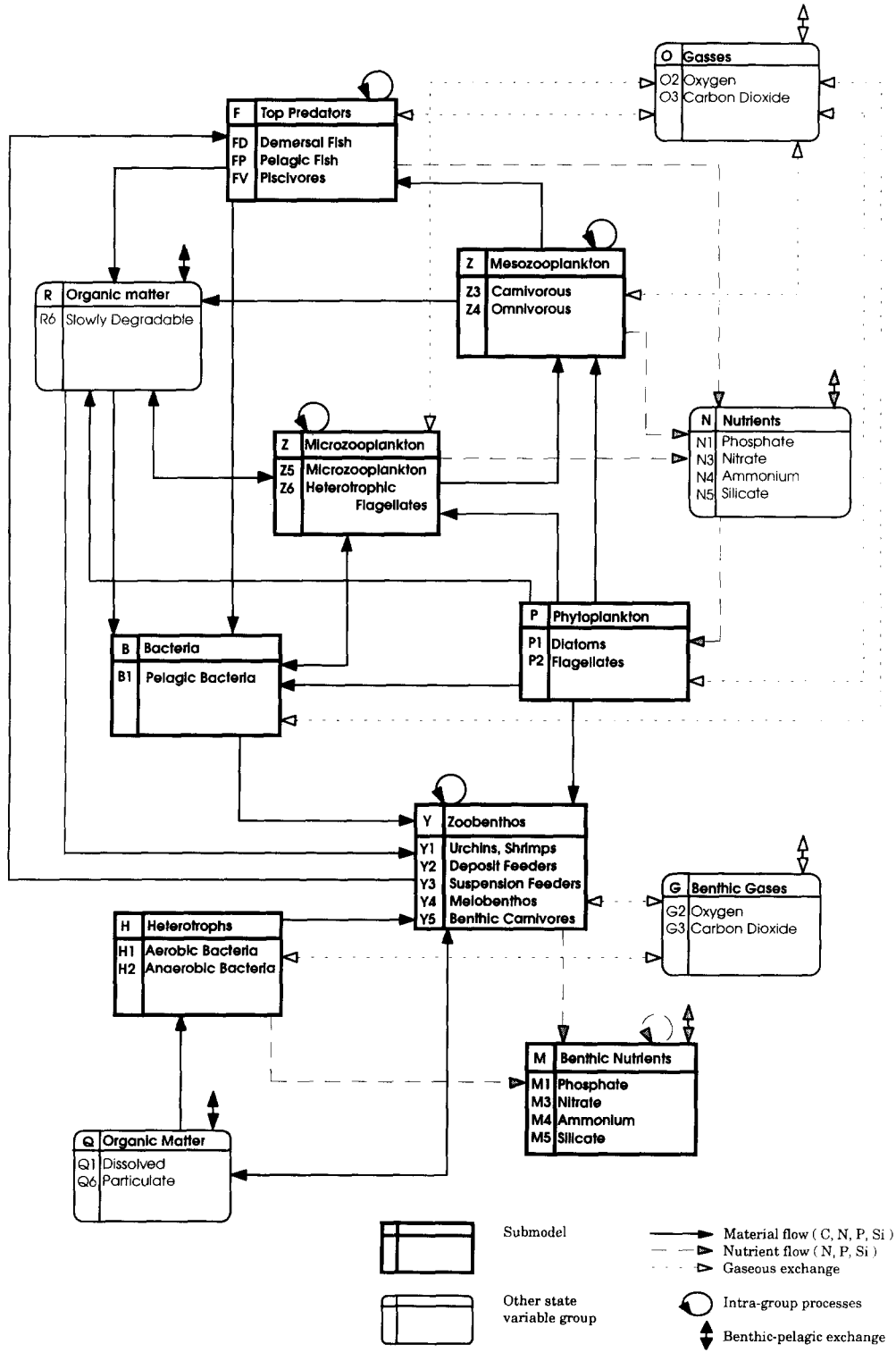

Fig. 3. Simplified ERSEM process flow diagram.

ferences can be qualitatively important. To this end each of the nine participating laboratories purchased Sun Sparc workstations (Sun Microsystems Ltd.) and associated peripherals at the outset of the ERSEM project.

Commercial software requirements are limited to Sun Fortran (versions 1.3 and 1.4), GKS (versions 3.0 or 4.1) and either the Sunview or Openwindows operating system interfaces. Although changes to the Fortran version should not numerically alter the simulation output, model simulations comparing the old and new software must be made to establish this as fact. In each case the software is chosen for its proven reliability rather than for any innovative features.

All participating institutes in the ERSEM project are interconnected via the Joint Academic Network (JANET) in the UK and the European Academic Research Network (EARN) for the continental partners. This allows distribution of new standard model versions and data electronically and encourages a quick and easy flow of communication amongst project participants.

3.2. MODELLING TOOLS

Much more crucial to model behaviour are the mathematical algorithms used to solve the differential equations. To this end a modelling tool, SESAME (Software Environment for the Simulation and Analysis of Marine Ecosystems) has been developed for the ERSEM project. The SESAME package (Ruardij *et al.*, 1995) runs under Unix and has been optimized for use on Sun RISC processors. It is written in Fortran-77 using VMS extensions, having evolved from BAHBOE (Baretta & Ruardij, 1988), developed on and for use with Norsk-Data hardware. The SESAME simulation tool is designed to solve systems of parallel ordinary differential equations, within a framework which facilitates participation of multiple partners in the construction and use of large ecosystem models. A full description of SESAME is given in Ruardij *et al.*, 1995, but the broad features which contribute to harmonization and quality assurance are also reviewed in this chapter.

The philosophy of SESAME revolves around the concept of a standard model consisting of a group of associated Fortran routines which together describe an ecosystem. The interfaces between the modules—the set of state variables—are automatically packaged into a Fortran structure called a common block and via these made available to all parts of the model code. Removing the necessity for the modellers to explicitly deal with the module interfaces reduces the error potential considerably.

Many of the more mundane aspects of programming, such as initialization of variables, compilation, parameter input and data output are also performed by SESAME, thus removing many of the more common sources of computational error.

TABLE 2
The standard processes associated with each of the three types of organism considered by the ERSEM project (Ebenhöh & Kohlmeier, 1993).

type of organism	*associated standard process*
consumer	uptake of (several) food sources
	excretion
	respiration (basal, activity and stress)
	mortality (not due to predation)
	nutrient release
decomposer	respiration
	mortality
	nutrient uptake and release
	degradation and assimilation of detritus and dissolved matter
	production of dissolved organic matter
primary producer	assimilation
	respiration
	nutrient uptake
	exudation
	lysis
	sedimentation

Each participating site holds a copy of the current standard model. Partners wishing to modify any particular subprogram use one of the menu driven facilities of the SESAME package to copy the module file and associated parameter file to a different subdirectory. Here they are free to adapt the module as they wish without compromising any part of the standard model. At compilation (the generation of an executable program from the source code) SESAME automatically integrates the adapted module with the remainder of the standard model thereby forming a full ecosystem model. In this way the performance of revised modules may be tested in the context of the whole ecosystem model.

Other important features of SESAME which contribute to harmonization of approach include options which allow for model simulation, the generation of tabular and graphical output, validation against field data and comparison with previous model versions. Thus the treatment of results is standardized and intercomparable. These facilities free the user to concentrate upon defining the dynamics of the ecological system.

3.3. CODE CONVENTIONS

The first essential feature of the conventions relates to the units of mass adopted throughout the model. The standard units, ensuring compatibility with most measured data sets are used. It is important to keep in mind the distinction between those states that occupy a volume and those that occupy a surface. Hence carbon state variables are represented in

mg $C \cdot m^{-3}$ in the pelagic and in benthic pore water region of the model but as mg $C \cdot m^{-2}$ in the benthic layer. Similarly nutrients and gases are expressed in $mmol \cdot m^{-3}$ in the pelagic and in pore water but as $mmol \cdot m^{-2}$ in the benthic layer.

Certain invariate code conventions are required by the SESAME package and need to be adopted if the model is to be internally consistent. Experience has shown it to be also desirable to introduce other conventions to encourage mutual understanding between the various partners' subprograms. The conventions adopted are set out in detail in Appendix 1. Control over conventions is centralized, with requests for alterations or extensions being channelled through the one partner taking overall responsibility for model development, Plymouth Marine Laboratory.

Compulsory conventions cover only the state variable names, which provide the unifying elements of the modules and certain global variables principally those describing environmental factors. The representation of state variables indicates the trophic level, specific type, composition and geographical box, so that for example *Z3c*(7) can be translated as zooplanktonic carnivores (*Z3*) measured in mg $C \cdot m^{-3}$ (*c*) in box 7.

The SESAME package automatically creates the source terms in which the calculated rates of change are stored by each module ready for integration by the SESAME package. These source terms consist of the state variable name preceded by the character '*S*'. Hence *SZ3c* is the source term for carnivorous mesozooplankton carbon content.

It is necessary to ensure that within each module two source equations follow each flux (or mass transfer) definition. For example to ensure that the process of uptake by omnivorous zooplankton (*Z4*) of diatoms (*P1*) is integrated within SESAME the following source equations must be written.

Flux calculation	*fP1Z4c* = ...
Source equation	*SP1c(I)* = *SP1c(I)* - *fP1Z4c*
Source equation	*SZ4c(I)* = *SZ4c(I)* + *fP1Z4c*

Experience with the GEMBASE (Radford & Joint, 1980) and Ems model (Baretta & Ruardij, 1988) confirmed the value of the introduction of a range of code conventions which together, although not essential to the integrity of the model, act as an aide-mémoire and serve to give a unifying feel to the program as a whole. Rather than creating the restrictive straight jacket to programming that one might expect they have proved to bring about a freedom of mutual understanding amongst the group of partners.

Fluxes of material from one state variable to another are indicated by the character '*f*' followed by the state variable name of the source, the destination and the composition code. (*e.g.* primary production of pelagic diatoms in carbon units is named *fO3P1c* where *O3* is the name for carbon dioxide). If the flux destination is in different units from its source then a letter *j* (rather than *f*) is used to emphasize that a unit conversion must be made in one of the source terms. The fluxes represent the end value of the process calculations and especially during the model development stage can be of as much interest to the modellers as the state variables. A unifying convention for fluxes enables these values to be readily identified.

Auxiliary variables, which are precursors to flux definitions and to some extent parameters commonly fall into a number of different categories, each of which can be signified by a particular character. Further characters specify more precisely the nature of the variable indicating, the process it represents, a description of that process, and its relative magnitude. The addition of a state variable name associates the process with the relevant component of the ecosystem and often provides all that is needed to complete a unique acronym. For example *rutZ3n* represents the total (*t*) rate (*r*) of uptake (*u*) by carnivorous mesozooplankton (*Z3*) in terms of nitrogen (*n*).

Parameters or constants are signified by the $ character so that for example, *sumZ3*$ represents the (fixed) maximum (*m*) specific rate (*s*) of uptake (*u*) of carnivorous zooplankton (*Z3*).

A simple look-up table (Appendix 1, Table 6) has been provided to assist partners in the building of auxiliary and parameter acronyms and this can also be used in the interpretation of Fortran names created by the other partners. Although initially requiring some extra effort, it has been found that modellers rapidly become conversant with this system. The use of such closely defined acronyms has proved of great value in reading the computer code as well as adding rigour to the activity of programming.

3.4. CODING PRACTICES

One further simple but effective practice has been adopted as standard. Including the 'implicit none' statement in all subroutines of the model ensures that all identifiers (parameters and variables) must be explicitly defined either globally via the SESAME common block system or within that particular subroutine. This protects against the accidental mistyping of an identifier, the presence of which would be flagged as undeclared at model compilation. In addition the standard protocols regarding presentation and layout of Fortran are adhered to, as much as is practical.

3.5. ACCEPTANCE CRITERIA

When revised modules are presented for possible inclusion in the next standard model it is necessary to assess not only the merits of the new code (see Section 5) but also its integrity (Section 4). To facilitate this, in addition to the code files containing the prospective modifications, two further items are required

with each new module; a sample set of model output generated by the new module run in conjunction with the current standard model and a documentation detailing the scope of and reasoning behind the modifications, with a qualitative description of how the submodel performance has improved. This allows the verification of results and intent.

3.6. VERSION CONTROL

Extensive use of a UNIX utility, the Source Code Control system (SCCS) is made in the process of assembling new verified modules into the next standard version. SCCS monitors changes made to a file or set of files and maintains a record of all such changes. The program controller may re-instate any previous version of any module file into the working directory. This system is especially useful when a number of new modules which tested individually against the present standard perform reasonably, but in combination interact in such a way as to cause unacceptable results. SCCS simplifies the task of identifying the best combination of modules and the source of conflict.

New standard model versions are released with a sample set of results thus enabling all partners to check the model as received works as intended.

3.7. DOCUMENTATION

Adequate documentation of ERSEM software is required at all levels of the project. The SESAME (Ruardij *et al.*, 1990) manual describes the structure of the types of models which can be formulated within its framework, as well as all the facilities which it provides. This gives an essential common reference for all partners as they develop their individual modules within the context of the standard model.

Detailed descriptions of these new modules are produced at each stage of module development and the final versions are published internally (Radford, 1993). At a lower level the Fortran code itself is liberally annotated with comments to explain line by line how ecosystem concepts had been translated into computer code.

4. QUALITY CONTROL

Quality control represents the active efforts undertaken to assess the integrity of the model code. All but the most trivial code modifications must be subject to a rigorous verification process that ensures that the code correctly represents the intended biological, chemical or physical interactions. This is separate from the task of assessing the relative merits of one model version over another.

An initial check on new modules is made by attempting to duplicate the results generated by the submitting partner by running the updated code against the reference standard model held at the code coordinating centre. Although not a rigorous check this indicates not only that the new module has been received as intended, but further that the standard model held by the submitting partner has not diverged from the agreed standard. Reference then to the documentation ensures that the modified code faithfully represents the interactions as intended by the submitting partner.

The final stage is to ensure that the new sections of the code maintain coherency with the rest of the model. A number of points need to be checked; that the variables used are correctly defined, initialized and passed between subroutines; that the naming conventions are adhered to; that comments in the code are useful; that obsolete code is removed. Most importantly a check is made that each flux defined (*i.e.* a transfer of material between two state variables) is correctly included in the source terms, (*i.e.* is added to the source term of the recipient state variable and subtracted from the donor) hence ensuring mass balance. Here it is important to ensure the correct signs have been used and that, if the units of the two state variables differ, (as in the case of benthic, m^{-2} and pelagic, m^{-3} variables) the correct conversion factor is applied.

To assist this task a number of pieces of software have been developed that analyse the structure of a submodel/module. Using these it is possible to: trace the usage of any given identifier or group of identifier, be it parameter, state variable, source term, flux term or other variable, through any submodel; list all identifiers present in a module; list all source equations and dissect the modular structure of code. These utilities are based on the standardization of code conventions described above.

5. EVALUATION OF MODEL PERFORMANCE

In any given ecosystem model it is often seen that small numerical variations in parameters or initial conditions can make the difference between a reasonably behaving model and nonsense. Whilst the complexity of the interactions in ERSEM to some extent may be expected to protect against such occurrences, this complexity makes it rather difficult to predict the effect of a particular change to the model formulation.

Another factor that increases sensitivity is a lack of differentiation in some of the trophic levels; for example there are currently only two types of primary producers supporting the biological food chain. This reflects the decisions made at the outset of the project to restrict the initial scheme to the major trophic groups and feeding types.

Once a submodel/module has been accepted as a correct representation of the intended processes and assuming that it improves the performance against measured data of the particular groups it represents,

the ramifications on the other parts of the model must be analysed. This presents problems for two reasons. There are a large number of state variables (currently about 70) in each of fifteen boxes that can be assessed and for many variables especially in the northern and offshore regions of the North Sea there is a lack of data against which to validate the model output.

The first stage in developing a set of validation data that can be trusted is to recognize the variety of problems that present themselves. A distinction can be made between state variables that can be instrumentally measured such as nutrients and those that cannot, for example 'carnivorous mesozooplankton' (one of the ERSEM state variable categories). For the former there generally exist reasonable data sets or at worst the potential to obtain such data within the constraints of time and budget. For the latter the situation is much more problematic. It is necessary to define what, for example, carnivorous mesozooplankton comprises in terms of particular species. Then with access to data for relevant species it is necessary to consider the effects of sampling technique and analysis, adjust the data to take account of missing species, age or size classes and to provide some conversion from species counts to the units of the model, in this case mg $C \cdot m^{-3}$. All of these elements contribute to a large degree of uncertainty about the eventual validation data.

Processes that lead to short-term and spatially restricted events such as the spring bloom can also be difficult to validate as data sampling programs often tend to miss this type of phenomenon. To validate the ERSEM code we require for any given period a value applicable to a whole geographical region as represented by one of the 15 boxes. However, the existence of marked gradients across the boxes coupled with the haphazard distribution in space and time of sampling sites and the variation in shape and size of the ICES boxes makes it probable that the process of interpolating data to produce average box values will have as a consequence a larger margin of error than desirable.

Within the ERSEM group the approach to data use is centralized, with PML providing the data banking facilities. Most data are distributed in the format required by the SESAME package for model validation. Initially the construction of a formal database system for the project was considered but rejected for two reasons. Firstly, as most data were already held in a database there was no desire to duplicate this effort. Secondly, the specialist data processing facilities required by the project would not easily be achieved using commercial database packages. Instead it was decided to develop in house software explicitly tailored to the ERSEM requirements.

There is a conflict between the practicalities of data management and the requirement to screen each data profile and period aggregate to ensure its validity. Coastal effects and seasonal storm events can often produce very misleading readings. The software developed will:

- integrate or average depth profiles;
- sort data for either the ERSEM boxes or for any other specified (rectangular) area;
- aggregate data in space either by simple averaging or interpolation;
- aggregate data for any specified time period;
- store details of intermediate stages in the calculations, allowing screening of data as required;
- produce error margins as either standard deviations or quantiles;
- allow rejection of data points based on specified criteria.

Assessing the effect of new submodels is problematic, given that changes to model results tend to be subtle, especially for state variables other than those directly dealt with by the changed code. It is very difficult to envisage a single objective criterion which could be applied to the overall model result which could then be used to provide a simple comparison with the current standard version. Thus in assessing the relative merits of one model formulation against another, given the uncertainty of the validation data, subtle changes that make no serious divergence from the expected results are excluded from consideration. Less subtle changes can to some extent be quantified using standard goodness of fit formulations (another facility of the SESAME package (Stroo, 1986)). When faced with difficulty in distinguishing two model formulations, more weight is placed on components that drive the system, such as nutrients or primary production and components that have detailed validation data sets available. Finally one must also consider the merits of a new module that models a particular process in more detail and realism than previously, whilst not improving the model output. Qualitative decisions are to some extent unavoidable when assessing this type of modelling project.

In practice assessing the relative merits of module revisions has not provided any major uncertainties regarding the relative merits of alternative modules, within the first phase of the ERSEM project. However, as the project continues it can be expected that such problems will arise with greater frequency.

6. CONCLUSION

The methodology of project organization and quality management described here is equally applicable to a wide range of multi-programmer ecosystem modelling projects. The essential ingredients can be summarized as:

- compatible hardware and software environments;
- a logically structured standard model that divides the modelling effort into several conceptually inde-

pendent units or modules;
- modelling tool that standardizes the mathematical algorithms used, automates inter-module communication, automates module integration and provides standardised treatment of results;
- set of coding conventions ensuring a common interface between modules;
- protocols for submitting and testing new modules;
- quality control system that establishes the correctness of the model elements.

The success of any system of quality control can ultimately only be judged by its record. During the three years of the ERSEM project it is believed that no major mistakes have gone persistently unnoticed and hence little time has been wasted working with erroneous models. Of course many errors have cropped up, the vast majority minor, but these have been quickly noticed and corrected. It is also very important to note that informal internal review, in the context of a project like ERSEM which involves a high degree of co-operation, contributes significantly to overall quality assurance.

Acknowledgements.—The ERSEM I project receives financial support from the Marine Science and Technology programme (MAST I) of the European Union (contract number CT90-0021).

7. REFERENCES

Backhaus, J.O., 1985. A Three dimensional model for the simulation of shelf sea dynamics.—Dt. hydrog. Z. **38:** 167-262.

Baretta, J.W. & P. Ruardij, 1988. Modelling the ecosystem of the Ems Estuary. In: J.W. Baretta & P. Ruardij. Tidal flat estuaries. Ecol. Studies **71**. Springer-Verlag, Heidelberg: 1-353.

De Wit, C.T. & J. Goudriaan, 1974. Simulation of ecological processes. Centre for agricultural publishing and documentation, Wageningen: 1-159.

Ebenhöh, W. & C. Kohlmeier, 1993. Standard organisms in the ERSEM model. Internal Publication. University of Oldenburg, Germany: 1-33.

Forrester, J.W., 1961. Industrial Dynamics. M.I.T. Press, Cambridge, MA: 1-464.

Radford, P.J., 1979. Some aspects of an estuarine ecosystem model, GEMBASE. In: S.E. Jørgensen. State of the art ecological modelling. Proc. Int. Soc. Ecosystem Modelling Conf., Copenhagen: 301-322.

——, 1993. The modules. A description of each of the ERSEM I modules as at the end of MAST I. Internal Publication. Plymouth Marine Laboratory, Plymouth: 1-250.

Radford, P.J. & I.R. Joint, 1980. The application of an ecosystem model to the Bristol Channel and Severn Estuary. Institute of Water Pollution Control Annual Conference. Water Pollution Control **79:** 244-254.

Ruardij, P., J.W. Baretta, & J.G. Baretta-Bekker, 1990. SESAME, a Software Environment for the Simulation and Analysis of Marine Ecosystems. Manuals. publ. Netherlands Institute for Sea Research, Texel.

——, 1995. SESAME, a Software Environment for Simulation and Analysis of Marine Ecosystems.—Neth. J. Sea Res. **33:** 261-270.

Stroo, D., 1986. A method for validation. Netherlands Institute for Sea Research, interne verslagen EON 1986-1: 1-35.

APPENDIX 1

STANDARD SPECIFICATION OF CODE CONVENTIONS FOR ERSEM

To facilitate the modular structure of submodels and to assist mutual validation of computer code, certain conventions are recommended. Some are considered as essential but others allow a little latitude. This section explains the philosophy of the nomenclature to be adopted and specifies the agreed standards.

1.1. UNITS TO BE USED FOR STATE VARIABLES

TABLE 1
Units as used in ERSEM.

composition	*code*	*pelagic*	*pore water*	*benthic*
carbon	c	mg C·m^{-3}	mg C·m^{-3}	mg C·m^{-2}
nitrogen	n	mmol·m^{-3}	mmol·m^{-3}	mmol·m^{-2}
phosphate	p	mmol·m^{-3}	mmol·m^{-3}	mmol·m^{-2}
silicate	s	mmol·m^{-3}	mmol·m^{-3}	mmol·m^{-2}
oxygen	o	mmol·m^{-3}	mmol·m^{-3}	mmol·m^{-2}
reduction equivalent	e	-	-	mmol·m^{-2}
depth	m	m	m	m

1.2. STATE VARIABLES

Modellers are limited to using a subset of the complete list of state variables. Additional state variables must be agreed by all partners and implemented through Plymouth Marine Laboratory.

The initial letter of the acronym indicates the trophic group and has been chosen where possible to be reminiscent of the group name. This character must be in upper case.

P phytoplankton
N nutrients
FD demersal fish

A following numeric code distinguishes between the various state variable types in the group.

N1 phosphate
N3 nitrate

Finally a single character code represents the composition and hence unit (see Table 1) of the state variable.

Table 2 shows all state variables used in the ERSEM model.
So for example:
P1c Pelagic diatoms expressed in carbon
P1n Pelagic diatoms expressed in nitrogen
N3n Nitrogen

TABLE 2

The ERSEM state variables. The state variable identifier together with one of the composition codes completes the state variable name.

PELAGIC

identifier	*comp. code*	*state variable*
N		**Nutrients**
N1	p	phosphate
N3	n	nitrate
N4	n	ammonium
N5	s	silicate
R		**Dead organics**
R1[1]	cnps	dissolved labile
R6	cnps	slowly degradable
P		**Phytoplankton**
P1	cnps	diatoms
P2	cnp	flagellates
B		**Bacteria**
B1	cnp	bacteria
Z		**Zooplankton**
Z3	c	carnivorous mesozooplankton
Z4	c	omnivorous mesozooplankton
Z5	cnp	microzooplankton
Z6	cnp	heterotrophic flagellates
F		**Fish cohorts** [2]
FP	c	pelagic planktivores
FD	c	demersal
FV	c	voracious piscivores, pelagic
O		**Gases**
O2	o	oxygen
O3	c	carbon dioxide
X		**Miscellaneous**
X1	x	salinity (PSU)
X2	x	arbitory tracer (no unit)
X3	x	proportion of box contents derived from the Rhine and Meuse (no unit)

[1] Not explicitly implemented in current model

[2] Fish populations are modelled as age-structured cohorts.

BENTHIC

identifier	*comp. code*	*state variable*
K		**Nutrients**
K1	p	phosphate
K3	n	nitrate
K4	n	ammonium
K5	s	silicate
K11	p	anaerobic phosphate
K14	n	anaerobic ammonium
K6	e	reduction equivalent
Q		**Dead organics**
Q1	cnp	dissolved labile
Q6	cnps	slowly degradable
H		**Heterotrophs**
H1	c	microbacteria aerobic
H2	c	microbacteria anaerobic
Y		**Zoobenthos**
Y1	c	urchins / shrimps
Y2	c	deposit feeders
Y3	c	suspension feeders
Y4	c	meiobenthos
Y5	c	benthic carnivores
G		**Gases**
G2	o	oxygen
G3	c	carbon dioxide
D		**Dimensions of layers**
D1	m	aerobic layer thickness
D2	m	depth of anoxic surface
D6	m	penetration of Q6c
D7	m	penetration of Q6n
D8	m	penetration of Q6p
D9	m	penetration of Q6s

PORE WATER

identifier	*comp. code*	*state variable*
M		**Nutrients**
M1	p	phosphate
M3	n	nitrate
M4	n	ammonium
M5	s	silicate
M11	p	anaerobic layer phosphate
M14	n	anaerobic layer ammonium

1.3. FLUXES

A flux represents a transfer of material per unit of time from one state variable to another. The standard flux nomenclature is the lower case letter *f* followed by the state variable acronyms of the source and destination in that order. They must end with the composition code of the state variables.
For example:

fO3P1c carbon flux from carbon dioxide to diatoms (primary production)
fP1Z4c carbon flux from diatoms to omnivorous mesozooplankton (grazing)
fN1P1p phosphate uptake by diatoms

However in some cases a flux occurs between a state variable measured per unit volume and one measured per unit area. In such cases a conversion factor (depth) is a necessary component in the source terms (see below). These fluxes are indicated by an initial letter '*j*', which serves as an aide-mémoire. For example:

jP1Y3c carbon flux from diatoms to suspension feeders (grazing)

Finally in some cases an initial letter '*g*' is used in place of the '*f*' to indicate a net rather than gross flux.

1.4. STATE VARIABLE INTEGRANDS OR SOURCE TERMS

All fluxes affecting a particular state variable must be summed into the source term specific to that state so that the SESAME package can integrate the calculated rates of change at the end of each time step. These source terms are denoted by the capital letter '*S*' followed by the state variable name. Hence *SP1c(I)* represents the source term for pelagic diatoms. For example, the source equations resulting from the calculation of the flux *fP1Z4c* must be written as:

$SP1c(I) = SP1c(I) - fP1Z4c$
$SZ4c(I) = SZ4c(I) + fP1Z4c$

and in the case where a flux, (say *jP1Y3c* - calculated per m^2) is between a volume state and a surface state:

$SP1c(I) = SP1c(I) - (jP1Y3c \,/\, \text{depth}(I))$
$SY3c(I) = SY3c(I) + jP1Y3c$

1.5. ENVIRONMENTAL FACTORS

All environmental factors start with the letter E followed by two upper case characters

TABLE 3
Environment factors.

model name	*meaning*	*unit*
ETA	temperature in the air	°C
ETW	temperature in the water	°C
ETB	temperature in the bottom	°C
EIR	irradiation	$W \cdot m^{-2}$
EQU	mass transport in horizontal direction x and/or y	$m \cdot s^{-1}$
EQX	mass transport in horizontal x-direction	$m \cdot s^{-1}$
EQY	mass transport in horizontal y-direction	$m \cdot s^{-1}$
EQZ	mass transport in vertical z-direction	$m \cdot s^{-1}$
EDU	diffusion coeff. in horizontal direction x and/or y	-
EDX	diffusion coeff. in x-direction	$m \cdot s^{-1}$
EDY	diffusion coeff. in y-direction	$m \cdot s^{-1}$
EDZ	diffusion coeff. in z-direction	$m \cdot s^{-1}$
ESS	suspended sediment	$mg \cdot m^{-3}$

1.6. AUXILIARY VARIABLES

Auxiliary variables are defined as precursors to flux definitions. They may be classified into different types.

TABLE 4
Auxiliary variable types.

identifier	*type*
c	concentrations
e	environmental variable
p	proportions (dimensionless)
q	quotient (not dimensionless)
q10	Q10 value of a process
r	rates
s	specific rates
t	temperatures
u	unit conversions
v	volume rates (*e.g.* filtration rates)
x	all others

One of these lower case characters must be the initial letter of the auxiliary variable name.

Some auxiliary variables represent biological processes and often these may be qualified by commonly used adjectives which describe the type and magnitude of the process. These characters will follow the auxiliary variable identifiers.

TABLE 5
Process descriptions and magnitudes.

category	*code*	*description*
process	d	death (mortality)
	e	excretion/exudation
	r	respiration
	u	uptake/feeding
description	a	active (routine)
	s	standard (basal, rest)
	o	other
	t	total
magnitude	g	gross
	h	half saturation
	l	lowest (minimum)
	m	maximum
	n	net

If more than one of these categories are used they should appear in the order process, description, magnitude.

The '-' sign may be used to link (L) separate parts of the auxiliary variable acronym to indicate relationships. It has two meanings:

a. available for

b. going from 'A' to 'B' (indicating direction of flow).

For example:

sraZ3c specific rate (*s*) of active (*a*) respiration (*r*) of carnivorous mesozooplankton in terms of carbon.

rumZ5 rate (*r*) of maximum (*m*) uptake (*u*) of microzooplankton.

Exceptions to these auxiliary variable conventions must be clearly documented.

Rules for building auxiliary variable names.

1. The first character must be one of the identifiers (Table 4).
2. This may be followed by one or more of the process/description/magnitude characters (Table 5).
3. If appropriate, groups of characters from Tables 4 and 5 may be linked by the "-" character to indicate the preposition "for" or "to".
4. One or two state variable acronyms may follow, finishing with the composition code (Table 2).
5. Parameters must end with a "$".

TABLE 6
(Summary of rules for building acronyms for auxiliary variables and parameters.)

identifier		*process*		*description*		*magnitude*		*to/for*	*state*	*comp.*	*param.*
c	concentrations	d	death	a	active	g	gross	'-'	P1	c	$
e	environmental	e	excretion	s	standard	h	half saturation		Z3	n	
p	proportions	r	respiration	o	other	l	lowest		.	p	
q	quotients	u	uptake	t	total	m	maximum		.	s	
q10	Q10 values					n	net		etc.	o	
r	rates									e	
s	specific rates										
t	temperatures										
u	unit conversions										
v	volume rates										
x	all others										

1.7. PARAMETERS

Fixed value parameters follow the same rules as auxiliary variables but are identified by the addition of the '$' character. For example:

srsZ5$ specific rate (*s*) of basal or standard (*s*) respiration (*r*) for microzooplankton.

1.8. TRANSPORT TERMS

Water transport terms are indicated by a three-letter code which may prefix any state variable code.

w initial letter

The second letter represents the direction or type of transport.

b horizontal transport across sea boundaries
d diffusion mediated transport
h horizontal transport
r riverine transport
s vertical transport other than by water transport
t total transport into or out of a compartment
v vertical transport between adjacent layers

The third letter represents inputs to or outputs from a geographical compartment (box).

i input
o output

This concept can be expanded to define transport of any material dissolved or suspended in water for example:

wbiN1p(1) transport of phosphate into box 1 across an external boundary.
wtiPTc(6) total transport of pelagic carbon into box 6.