

# Computerizing Marine Biota: A Rational Approach

V.S. CHAVAN, D. CHANDRAMOHAN AND A. H. PARULEKAR

*Biological Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004, India.*

## SUMMARY

Data on marine biota while being extensive are also patchy and scattered; thus making retrieval and dissemination of information time consuming. This emphasises the need for computerizing information on marine biota with the objective to collate various aspects like taxonomy, biology, ecology, biogeography, pollution and the bibliography. It would help in knowing endangered and extinct forms, extent and effect of pollution, systematics, biodiversity, environmental status and would help in the development and management of marine living resources for sustainable utilization. A methodology, towards computerizing various aspects of marine biota, is described and discussed here.

## Introduction

Approximately, over  $50 \times 10^6$  species are projected to exist in the world, though only,  $1.5 \times 10^6$  species have so far been documented (Swaminathan, 1991). Most of the unknown and undescribed life forms are expected to be from the marine environment (Ray, 1988; Swaminathan, 1991).

Available data on marine life is scattered and not maintained properly, causing delay in the retrieval. Similarly, there is no proper collation of the impact of different pollutants to marine life forms. Research, management and conservation activities in the marine environments are always effective with adequate, appropriate and rapid accessibility to desired information (Nair, 1987; Duellmen, 1992). These could be done with computerization of required data and is being practiced (popularly known as Bioinformatics) elsewhere.

## Marine Bioinformatics in Indian Context

India has a coastline of about 7000 km. with the exclusive economic zone (EEZ) of over  $201 \times 10^6$  hectares and is one of the richest in the world in natural resources (Das and George, 1991; Khosho, 1991) and

has a century-old tradition of marine biological research).

Computerisation of data on marine biota has much importance especially for the developing country like India because there is a need for readily accessible taxonomic information and expertise in order to overcome constraints in development, management and conservation of marine living resources (Parulekar, 1990; Schalk, 1992).

## Global Status of Databases on Marine Biota

The idea of a computer based taxonomic, biogeographic and ecological data bank, is of recent origin. (Schalk, 1992). With the inception of Oceanographic Data Centre (NODC) at Washington, DC in early 70's listing of the marine flora and fauna increased. NODC's aim is to list the marine biota and to develop a numerical taxonomic code for it. In 1989, the Institute of Marine Research, Norway, first released a similar database with taxonomical, biological and ecological information on 350 marine protists. In 1991, International Expertise Centre for Taxonomic Identification (ETI) was established at the University of Amsterdam, with the objective to develop description database on marine biota.

Among the third world countries, India is first to initiate building database related to marine biota. In mid 1990, Distributed Information Sub Centre (DISC) was established at the National Institute of Oceanography (NIO), Goa, India with funding under the Biotechnology Information System (BTIS) of the Department of Biotechnology (DBT) of the Government of India. The major aim and objective of the NIO DISC is to develop database on flora and fauna of the marine and estuarine ecosystems which would be known as NIO's Database on Marine Life of India (NIODMLI).

### NIODMLI

NIODMLI would initially cover the taxonomic, biological, ecological, pollutional and economical information (Fig. 1) along with the sketch or photograph and distribution map of the species occurring in the marine and brackish water environment. Information for the database will be obtained from the expertise available in India and through literature survey. The database would also list major Indian and international

publications on the species along with the contact addresses of the experts on the subject.

In order to identify and address each species and to facilitate easy retrieval, numerical taxonomic code (NIO CODE) has been developed. The code is a hierarchical system of numerical code of upto 12 digits used to represent the organisms to the level of species, subspecies or variety (Table 1). For example, an economically important algal species (*Sargassum aquistifolium*) is coded by a 10 digit number-151004102, whereas *giesbrechti* variety of the copepod *Undinula vulgaris* is coded by a 12 digit number 611801040101 (Table 2). The code links the Linneaus system of biological nomenclature to a numerical scheme that facilitates modern methods of data storage and retrieval.

Fourth Generation Language (4GL) Database Management Software (DBMS), "FOCUS" operating under "UNIX" environment is used. FOCUS software facilitates local and global searching which would be of great help in retrieving data in multivariate

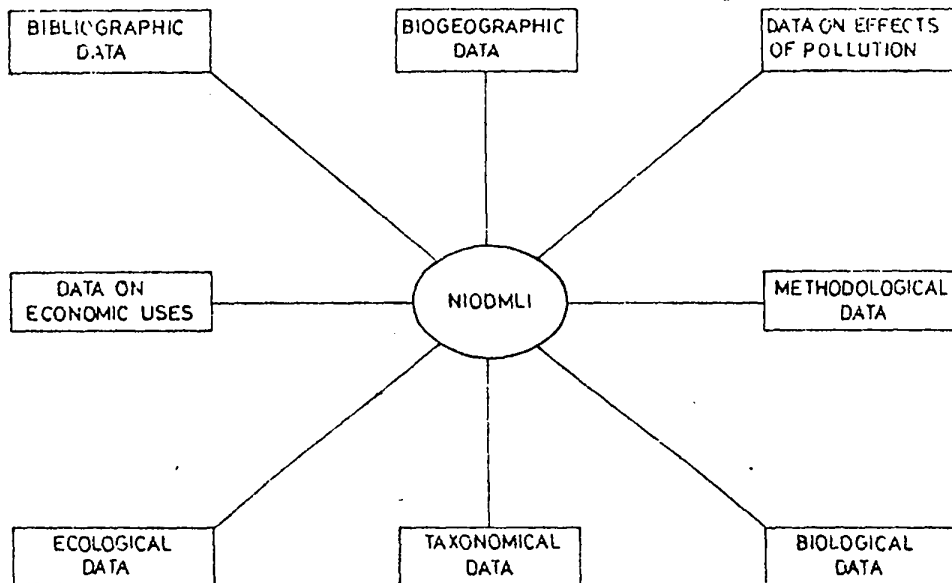


Fig. 1. Variety of Data accessible through NIODMLI

**Table 1: Structure of Numerical Taxonomic Code (NIO CODE) Developed by Distributed Information Sub Centre (NIO DISC)**

02 Digits	Sub Kingdom, Phylum, Sub Phylum, Class, Superorder, Order.
04 Digits	Superclass, Class, Sub Class, Superorder Order, Sub Order, Infraorder, Superfamily.
06 Digits	Class, Order, Sub Order, Infraorder, Superfamily, Family, Sub Family, Infracfamily.
08 Digits	Genus.
10 Digits	Species.
12 Digits	Subspecies, Variety.

**Table 2: Code for *Sargassum angustifolium* an economically important algae and an variety of copepod *Undinula vulgaris* Var. *giesbrechti*, whereas abbreviations used in the parenthesis indicate the taxonomic level of classification.**

15	Phaeophyta (P)
1501	Phaeophyceae (C)
1510	Fucales (O)
151004	Sargassaceae (F)
15100401	Sargassum (G)
1510040102	<i>Sargassum angustifolium</i>
55	Arthropoda (P)
61	Crustacea (C)
6117	Copepoda (O)
6118	Calanoidae (SO)
611801	Calanidae (F)
61180104	Undinula (G)
6118010401	<i>Undinula vulgaris</i>
611801040101	<i>Undinula vulgaris</i> var. <i>giesbrechti</i>

Abbreviations:	P	:	Phylum,	C	:	Class,
	SC	:	Sub Class,	Order		
	SO	:	Sub Order,	F	:	Family,
	G	:	Genus, var.	:		Variety.

combinations. This database could be retrieved by employing online and offline methods.

The outline of the numerical taxonomic code (NIO CODE) will be made available to user as a key to locate sections of the code that covers organisms of interest. NIO DISC will help users to contact respective expert(s) in order to get more detailed information on the desired species.

## Advantages of NIODMLI

1. Create record of existing, endangered and extinct species.
2. Build basic and applied data, related to marine organisms.
3. Help to trace the origin, evolution and distribution of the species.
4. A better tool for comparative study of the individual, population and community in relation to environment.
5. Present sharp and clear picture about the marine biodiversity.
6. Highlight the threatened life forms and environmentally sensitive areas.
7. Help in planning for the sustainable resource utilisation, environmental conservation and biosphere reserves.
8. Identify the species and areas for commercial applications.
9. Help taxonomists, systematists and provide media to the marine scientists to interact with each other.
10. List the basic literature on individual marine organisms.

## Acknowledgement

Authors are grateful to Dr. B.N. Desai, Director, National Institute of Oceanography, Dona Paula, Goa 403 004, India for unstinted support. Thanks to Department of Biotechnology, Government of India for funding the Distributed Information Sub Centre (DISC) at the National Institute of Oceanography, Goa. Authors are thankful to Dr. T.G. Jagtap and Mr. P.D. Kunte, our colleagues for their constructive criticism and valuable suggestions and to Mr. V.K. Kulkarni for his assistance.

## References

1. Das, P. and J. George. (1991). Biodiversity and fish improvements. *Biology Education* 8(1): 47-62.

2. Dudleyan, W.E. (1992). Addressing the biodiversity crisis: the neotropical biological diversity programme. *ASC Newsletter* 20(1): 85-88.
3. Khoshoo, T.N. (1991). In situ conservation of biological diversity. *Biology Education* 8 (1): 20-26.
4. Nair, N.C. (1987). Priority of rare plant data. *Biosphere Reserve: Proceedings of the first National Symposium, Udagamandalam, 24-26 Sept. 1986*, Ministry of Environment and Forest, Govt. of India, New Delhi. pp.147-154.
5. Parulka: A.H. (1990). Taxonomy in biomedical resources of the Exclusive Economic Zone of India. *Taxonomy In Environment and Biology: 49-57* (Zoology Survey of India, Calcutta (India)).
6. Ray, C.G. (1988). Ecological Diversity in Coastal Zones and Oceans. *Biodiversity* ed. E.O. Wilson, pp. 36-50, National Academy Press, Washington, DC.
7. Schalk, P.H. (1992). Systematics Goes Multimedia: Computerising Biological Knowledge. *Experience Centre for Taxonomic Identification, Amsterdam, The Netherlands* (Private Circulation).
8. Swaminathan, M.S. (1991). Biological Diversity and Biological Productivity. *Biology Education* 8(i): 04-19.