AAPG International Conference & Exhibition '94
August 21-24, 1994, Kuala Lumpur, Malaysia

A microcomputer intelligent database system for Indonesia marine geosciences data

SIGIT SUKMONO[1], DJOKO SANTOSO[1] AND BASRIE M. GANIE[2]

[1]Bandung Institute of Technology
Bandung, Indonesia
[2]BPP Teknologi
Jakarta, Indonesia

Abstract: The Agency for Assessment and Application of Technology — Indonesia (BPP Teknologi) has collected extensive marine geosciences data in Indonesia's maritime area by using the Baruna Jaya I, II and III research vessels. Types of data include: geological/geophysical data (coring, sampling, dredging, bathymetry, heat flow, magnetic, seismic, etc.), oceanography data (water temperature, salinity, sound velocity, chemical concentrations, etc.) and hydrography/meteorology data (air temperature, pressure, humidity, sea current and direction, etc.). A microcomputer database system has been developed using Prolog language for managing the data and it provides the integrated capabilities of database management system and expert system. The database system provides easy data entry, storage and retrieval in which the user could query the data based on the coordinate ranges, types of survey, name of survey, name of vessel and survey location. The output can be displayed in map format or text format and can be imported by some commercial softwares. The expert system provides intelligent assistance in searching specific geological, geophysical, oceanographical and hydrographical/meteorological information. The Prolog declarative expression of the database itself makes the data accessible for any logic or expert system programming. This can be built to suite specific purposes. The system is especially useful for institutions involved with marine data (i.e. oil company, oceanographic institutions, etc.). It is designed as an economical and practical microcomputer based system. The advantages of using Prolog language and Intelligent Database System concept is discussed.

INTRODUCTION

Data represents one of an organization's three most important assets, the other two are money and people. Therefore it shall be managed properly by utilizing fully the computer technology. It is widely predicted that knowledge information processing will be a major area of computer database applications in the 1990's, where problem solving and logic inference will play the central role. Along this line, the foundation of the work is to combine concepts from three areas of computer science research: Artificial Intelligence (AI), Logic Programming, Data Base Management System (DBMS) and Associate Processing. From AI and Logic Programming, the concept of Expert System (ES) is used and from DBMS the concept of a relational database is utilized. These ideas are married to form an intelligent relational database system, yielding a novel and powerful collection of facilities in the system-user interface, and giving great benefits in reduced cost and increased flexibility of implementation and application.

The Agency for Assessment and Application of Technology Indonesia (BPP Teknologi) has collected extensive marine geosciences data in Indonesia's maritime area by using the Baruna Jaya I, II and III research vessels. A microcomputer intelligent database system has been developed to manage the data.

Intelligent Database System (IDS), also known as "Expert Database System" or "Knowledge-based Management System", is a computer database system which integrates the technology of Relation Database Management System (RDMS) and Expert System (ES). The RDMS here provides basic data management operations including data insertion, data retrieval and data modification. The ES here can be built to act as human expert within one particular field of knowledge. The principal power of an ES is derived from the knowledge the system embodies rather than from searching algorithms and specific reasoning methods.

In an IDS, the ES modules can be incorporated in the front or/and rear-end of the database system (Fig. 1). This enable the system to offer not only direct retrieval of facts or groups of facts from the database, but also rules of inferences and hierarchy of inferences to make deductions and low-level interpretation (Fig. 2). Deductions can be based on rules of inference, where the user is supplied with information not stored explicitly in the database, but implied by the data. This is the aspect that justifies the term "intelligent".

The computer languages commonly used for developing IDS are PROLOG or LISP. They are declarative computer languages which are usually used for developing artificial intelligence, expert system, database system and it's combination. The popularity of the IDS is boosted by the announcement that this system is selected as the fundamental concept of the Japanese Fifth Generation Computer Project (Ishizuka and Moto-Oka, 1984). Research by USGS also revealed that this system is very promising for handling geological data (Brodie and Mylopoulos, 1986).

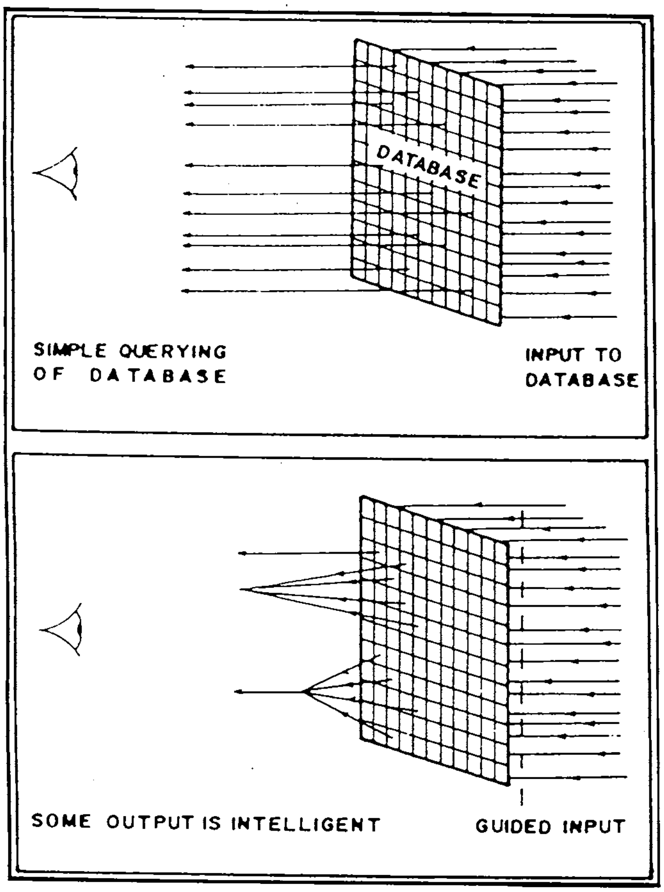

Figure 1. Difference between IDS and conventional database system.

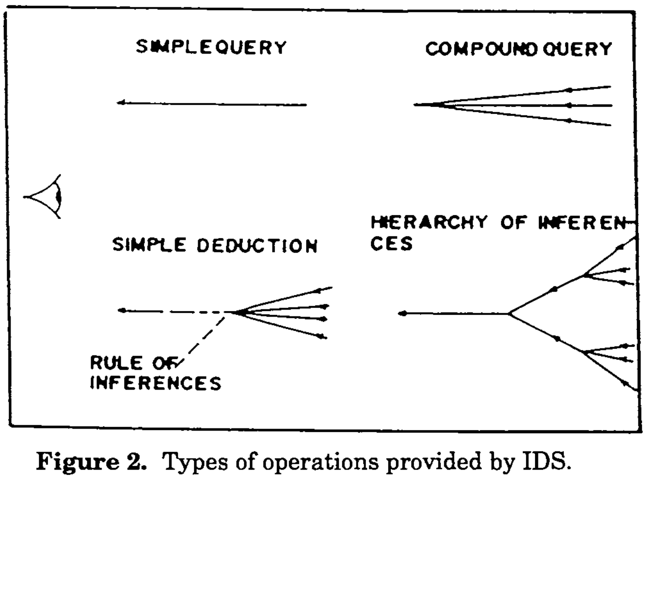

Figure 2. Types of operations provided by IDS.

STRUCTURE OF THE SYSTEM

The software here is developed using Prolog ("PROgramming in LOGic"), a declarative computer language commonly used in artificial intelligence programming. The Prolog program consists of a series of facts and rules which are entered in the internal database. The rules are the equivalent of procedures or subroutines in conventional language and the facts are equivalent to data items. Because the user only specifies the data items and the relationship between rule and facts, Prolog is called non-procedural that is the user need not specify any logarithm to solve the problem but only their relationships: the Prolog system determines the algorithm at run time.

The relational model is implied in the RDMS. In the meantime, an ES module is built only in the front-end of the RDMS to give intelligent assistance to the users in performing various searching procedures. We plan to add one more ES module in the rear-end to do subsurface condition prediction based on available coring and dredging data using cross-correlation and cluster analysis methods such as used before for geotechnical bore data (Sukmono *et al.*, 1993). The managed data is the data collected during marine surveys done by BPP Teknologi since 1989 using Baruna Jaya I, II, III research vessels (RV) (Tables 1 and 2).

The output can be displayed in map format or text format. It can also be transferred into ASCII file and imported by commercial packages. In the meantime, the data is still in Indonesian. We plan to translate it to English gradually. Figures 3–8 show some samples of outputs.

To ensure that the software can be installed widely, it is designed to be installable on IBM PC microcomputer and it's compatible. The ES modules will also give very user-friendly guidances so the system can be operated by people with little computer knowledge.

DISCUSSION

Familiar programming languages such as Basic, Pascal, and C have contributed significantly to numeric aspects of geology, including geological modeling, geophysical modeling and geostatistics.

Table 1. Data collected by Baruna Jaya I (Oceanography RV) and Baruna Jaya III (Multifunction RV).

1.	Name of survey.
2.	Name of vessel.
3.	Type of vessel.
4.	Date of survey.
5.	Location of survey.
6.	Departure port, midway port, arrival port
7.	Name, institution, and country of survey leader.
8.	Type of survey.
9.	Objective of survey.
10.	Results of survey (summary).
11.	Name, institution, and country of participants.
12.	Lines of survey.
13.	Oceanography station coordinate (for Baruna Jaya I only).
14.	Oceanography measurement data (for Baruna Jaya I only): measurement depth, water temperature, salinity, sigma T, sound velocity, chemical constituent and it's concentration (O_2, PO_4, NO_3, $SiOH_4$).
15.	Coring data (for Baruna Jaya III).
16.	Dredging data (for Baruna Jaya III).
17.	Heatflow data (for Baruna Jaya III).
18.	Bathymetry data (for Baruna Jaya III).
19.	Magnetic data (for Baruna Jaya III).

Table 2. Data collected by Baruna Jaya II.

1.	Name of survey.
2.	Name of vessel.
3.	Hydrography: date of observation, reference time, spheroid, projection, scale, survey area, base point, dangerous point, etc.
4.	Oceanography: date of observation, reference time, reference position, tide and current observation.
5.	Meteorology: date of observation, reference time, reference position, observation interval, air temperature, humidity, air pressure, rainfall, cloud condition, wind direction, wind velocity.
6.	Maritime geography: date of observation, survey area, survey method, reference.
7.	Summary.

The user can perform various searching procedures for specific data type and output it on the screen or printer using the database retrieval program. Specific searching procedures provided by the program are:

1. Based on the name of survey
2. Based on the name of vessel
3. Based on the location of survey
4. Based on the date of survey
5. Based on the type survey
6. Based on the coordinate range

These language can be characterized as "procedural", reflecting their emphasis on highly structured, fixed programs operating sequentially on variable (usually numeric) data.

Languages of the other major class, including Prolog and LISP, are adapted more to dealing with non-numeric data and their properties and relationships — types of information characteristic in the geological field. Because in using these languages, the distinction between program and data is often blurred, and relations rather than sequential procedures are emphasized, they are characterized as "declarative" languages.

Our previous studies showed that the application of declarative computer language and IDS concept for managing geological data offers three prominent advantages (Sukmono and Santoso, 1992):

a. Capability in manage and process qualitative information (which is very common in geology).
b. It's flexibility in ES and RDMS software design to suite specific requirements.
c. The managed data is ready for any logic or/and expert system programming.

Analysis of our developed IDS for Indonesia Marine Geosciences Data shows that there are some more specific advantages in using declarative languages and IDS concept compared to procedural programming:

a. It allows declarative expression of database, queries and facts which are very suitable for managing qualitative information which is common in marine geosciences data.
b. It allows a single language to be used for database, queries and programs, or in particular there is no need for a separate query and host programming language.
c. It provides a structure which is so similar to natural language, hence the expression of the database and the knowledge-base is more understandable.
d. It provides an expressive environment for data modeling, since the use of database statements allows a single general statement to replace

NAME OF SURVEY	/ Nama Survei	: MINA ZEE 2–90	VESSEL NAME	/ Name Kapal	: BARUNA JAYA I
VESSEL TYPE	/ Jenis Kapal	: K.S. Oseanografi	DATE OF SURVEY	/ Waktu Survei	: 4/7/1990 s/d 30/7/1990
DEPARTURE PORT	/ Pelab. Berangkat	: Tg. Priok, Jakarta	MID-WAY PORT	/ Pelab. Singgah	: Teluk Bayur, Padang
LAST PORT	/ Pelab. Kembali	: Tg. Priok, Jakarta	LEADER	/ Pemimpin Survei	: Edi Mulyadi Amin (BPPI, Indonesia
DATA STORAGE	/ Tempat Data	:			
SURVEY LOCATION	/ Lokasi Survey	: 1. LAUTAN HINDIA, Barat P. Enggano			

TYPE OF SURVEY / Jenis Survei:

Pengamatan oseanografi fisika: Suhu, salinitas dan densitas air laut dengan menggunakan CTD SUBER.
Pengamatan oseanografi kimia: Oksigen, posfat, nitrat dan silikat.
Pengamatan boimassa ikan dengan metoda akustik.
Pengamatan biomassa plankton.

SURVEY OBJECTIVES / Maksud dan Tujuan Survei:

1. Memperoleh data potensi sumberdaya ikan pelagis serta penyebarannya secara vertikal maupun horizontal.
2. Memperoleh komposisi jenis sumberdaya perikanan pelagis terutama yang bernilai ekonomis tinggi, baik sebagai komoditas ekspor maupun untuk pemehuhan kebutuhan dalam negeri dalam meningkatkan kebutuhan gizi masyarakat.
3. Memperoleh informasi dasar untuk penentuan alat tangkap ikan yang sesuai bagi pengusahaan sumberdaya tersebut.
4. Untuk mendapatkan informasi tentang kondisi oseanografi yang meliputi: suhu, salinitas, densitas, kandungan oksigen serta kandungan zat-zat harta seperti posfat, nitrat dan silikat.

PRELIMINARY RESULTS / Kesimpulan Sementara:

1. Kelompok ikan yang ditemukan pada kedalaman lebih dari 100 m; kelompok besar berada di sekitar selatan Selat Sunda sedangkan kelompok kecil di selatan P. Enggano.
2. Dari hasil penangkapan ikan, panjang ikan Cakalang yang tertangkap 48–56 cm, dengan berat 2, 3–2, 7 kg, ikan Madihiang panjang 31–43 cm dengan berat 0, 6–1, 3 kg, cumi cumi panjang mantel 11–20 cm dengan berat 40–400 gram.

DATA OF SURVEY : IRJA '92
OCEAMOGRAPHIC STATION MEASUREMENT DATA

Station No.	: 1	Date/Time	: 5 7 1992/1938
Depth	: 979.00	Wind/Direction	: 13.8/235
Latitude	: 130 24 59	Air Temp. Dry Bulb	: 27.30°C
Longitude	: 0 –19 –58	Air Temp. Wet Bulb	: 27.10°C
Air Pressure	: 1011.00		

Depth (m)	Temp. (°C)	Sal. PSU	Sigma T	Sound V (m/s)	0_2 (ml/l)	PO_4 µg A/l	NO_3 µg A/l	$SiOH_4$ µg A/l
0	28.58	34.177	21.59	1541.70	4.48	0.17	0.29	5.65
10	28.54	34.182	21.61	1541.80	0.00	0.00	0.00	0.00
20	28.42	34.201	21.66	1541.70	0.00	0.00	0.00	0.00
25	28.39	34.200	21.67	1541.80	4.42	0.41	0.29	17.83
30	28.36	34.204	21.68	1541.80	0.00	0.00	0.00	0.00
40	28.35	34.207	21.68	1541.90	0.00	0.00	0.00	0.00
50	28.35	34.210	21.69	1542.10	4.39	0.58	0.51	24.78
60	28.36	34.221	21.69	1542.30	0.00	0.00	0.00	0.00
70	28.10	34.286	21.83	1542.20	0.00	0.00	0.00	0.00
75	27.93	34.344	21.94	1541.60	3.51	0.88	2.00	24.35

DATA OF SURVEY : SALAWATI (HASARAM '93)
CORING DESCRIPTION DATA

Sta. No.	Depth (m)	Sample Length (m)	Core Description
1	972.30	2.12	Clay, greenish-grey, soft, large foram.
2	2019.00	0.99	Limestones, very hard, highly deformed.
3	665.70	2.25	–
4	565.30	0.00	–
5	56.60	2.09	–
6	730.60	1.33	Sand, fine-coarse, grey, large foram.
7	742.60	0.19	Limestone, coralline, very hard
8	663.60	2.05	Mud, grey
9	802.50	1.89	Clay, very soft, grey, green, small foram
10	46.14	2.38	Clay, very soft, grey, mollusca shell

Figure 3. Example of text format output.

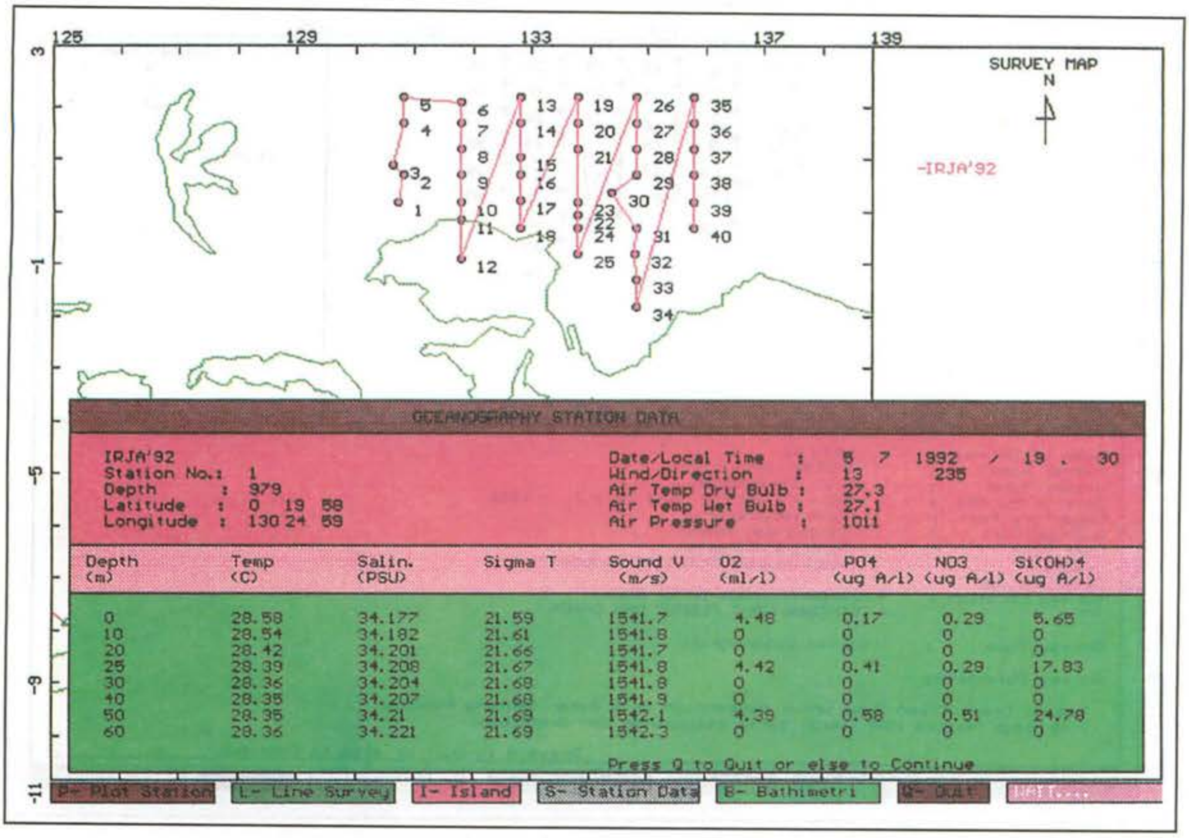

Figure 4. Example of map display for oceanography data.

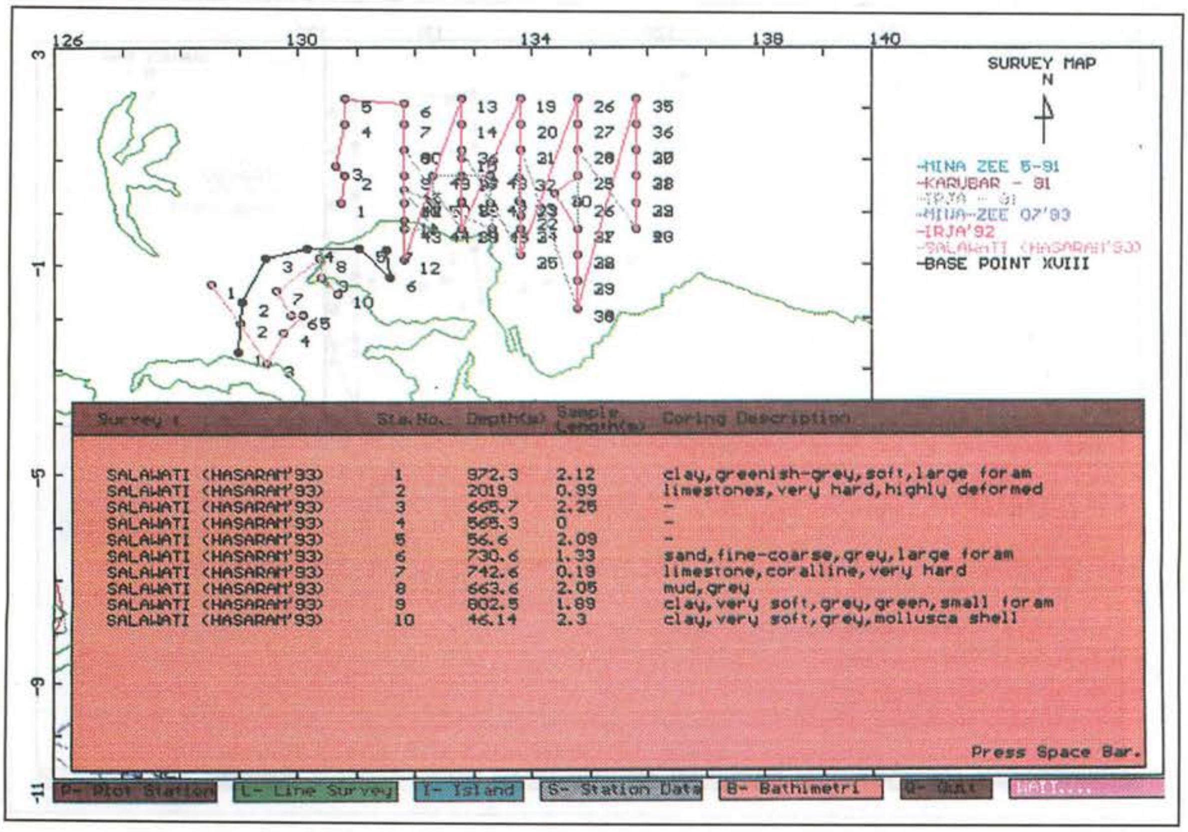

Figure 5. Example of map display for coring data.

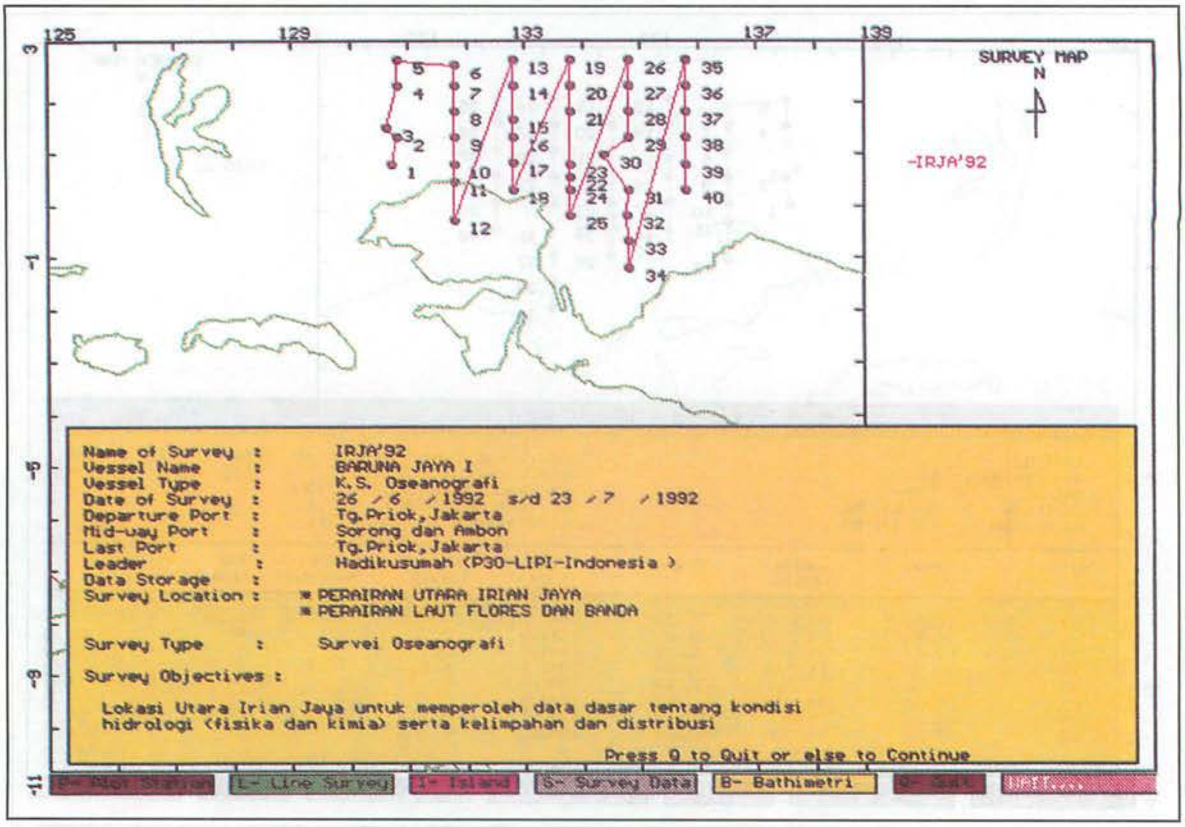

Figure 6. Example of map display for basic survey data.

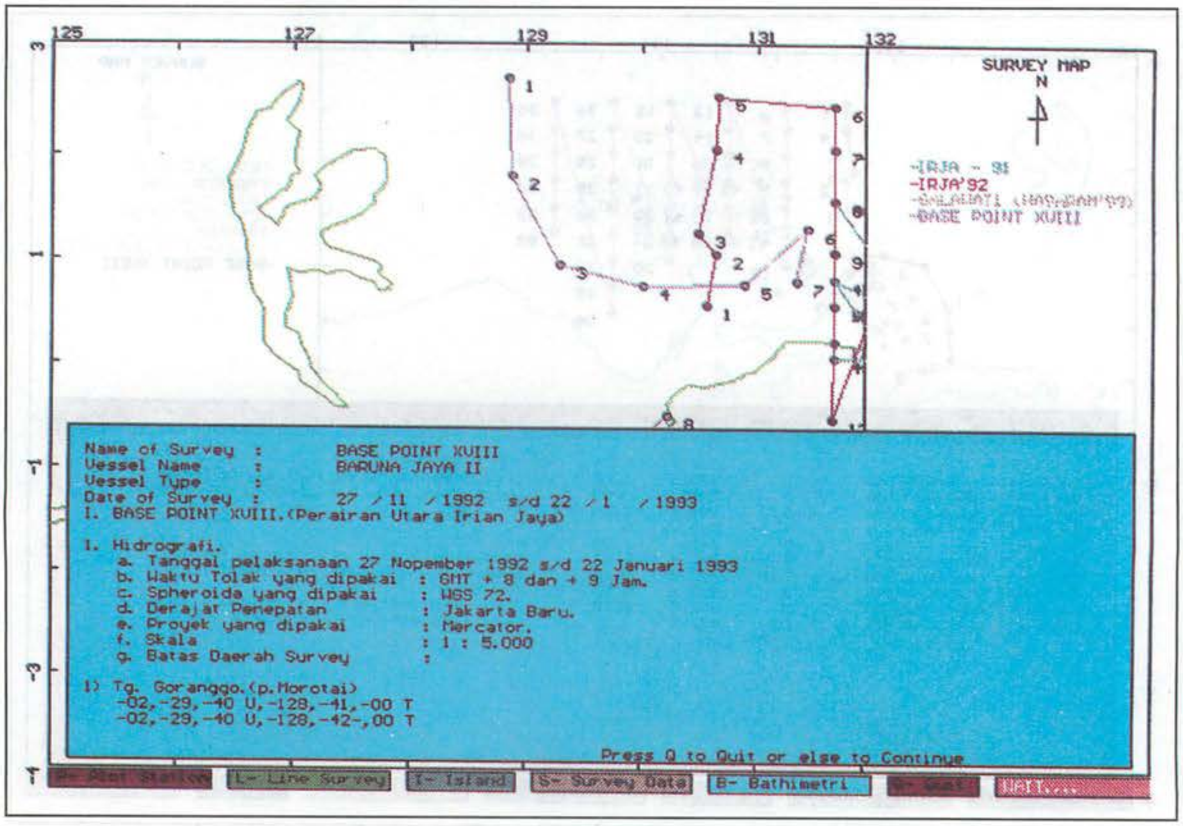

Figure 7. Example of map display for maritime geography data.

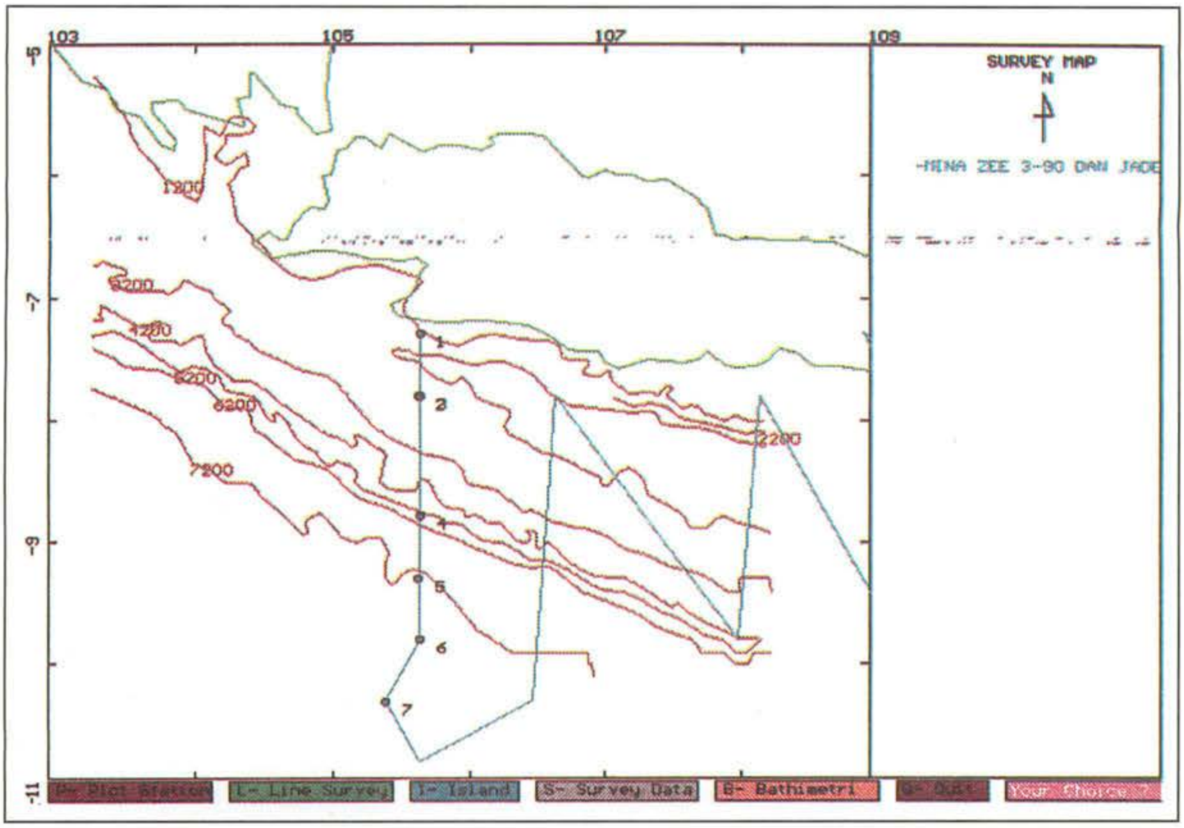

Figure 8. Example of bathymetry map display.

many facts.

e. The combined capability of a RDMS and an ES can be obtained. The RDMS serve as the fundamental basis for the IDS and the data are accessible for any ES or logic programming which can be built to suite specific requirements of the users such as doing the interpretation, diagnosis, prediction, instruction, monitoring, planning and design.

CONCLUSION

Marine geosciences data is one of the most important data in earth science; therefore proper management of these data by utilizing fully the computer technology is very important. Comparing to the conventional database the usage of IDS for managing marine geosciences data gives several prominent advantages in the system structure, system capability and system flexibility. Therefore further studies on the application of IDS for managing earth data, especially marine geosciences data, should be further encouraged.

REFERENCES

BRODIE, M.L. AND MYLOPOULOS, J., 1986. *On Knowledge Based Management System.* Springer-Verlag, New York.

ISHIZUKA, M. AND MOTO-OKA, T., 1984. Overview of Expert System in Japan. *In Expert System: State of the Art Report 12, 7.* Pergamon Infotech Ltd., 63–69.

SUKMONO, S. AND SANTOSO, D., 1992. Utilizing Intelligent Database System Concept for Managing Geological Data (An Introduction). *Proceeding of XXI IAGI Meeting,* 341–344.

SUKMONO, S., NUTALAYA, P. AND SANTOSO, D., 1993. An Intelligent Database System for Geotechnical Bore Data. *Proc. of 11th SEA Geotechnical Conference, Singapore,* 211–215.

Manuscript received 2 October 1994