

OPEN SOURCE ARCHITECTURE FOR WEB-BASED OCEANOGRAPHIC DATA SERVICES

R Venkat Shesu, TVS Udaya Bhaskar*, E Pattabhi Rama Rao, R Devender, and T Hemasundar Rao

Indian National Centre for Ocean Information Services (INCOIS), Ministry of Earth Science, Pragathinagar (BO), Nizampet (SO), Hyderabad, Andhra Pradesh, India - 500 090
*Emails: venkat@incois.gov.in; *uday@incois.gov.in; pattabhi@incois.gov.in; devender@incois.gov.in; hemasundar.t@incois.gov.in*

ABSTRACT

A GIS for ocean data applications named "Ocean Data and Information Systems (ODIS)" was designed and developed. The system is based on the University of Minnesota MapServer, an open source platform for publishing spatial data and interactive mapping applications to the web with MySQL as the backend database server. This paper discusses some of the details of the storage and organization of oceanographic data, methods employed for visualization of parameter plots, and mapping of the data. ODIS is conceived to be an end-to-end system comprising acquisition of data from a variety of heterogeneous ocean platforms, processing, integration, quality control, and web-based dissemination to users for operational and research activities. ODIS provides efficient data management and potential mapping and visualization functions for oceanographic data.

Key words: MapServer, Web GIS, Oceanographic data, Visualization, Open-source

1 INTRODUCTION

The oceans play an integral role in many of the Earth's systems including climate and weather. Continuous monitoring of oceans is vital for sustainable exploitation and utilization of ocean resources. In view of this, a large amount of surface and sub-surface data is being continuously collected. Advancements in the technology and progress of ocean observing systems have enabled an increase in ocean monitoring activities. Many ocean monitoring activities of the Indian Ocean (e. g., the Bay of Bengal Monsoon Experiment (BoB-MEX), the Arabian Sea Monsoon Experiment (AR-MEX), etc.) have been planned and successfully executed. The Indian National Centre for Ocean Information Services (INCOIS) is the nodal agency for the central repository of ocean data and services in India. It receives voluminous amounts of data from most of the ocean observation systems both in real time as well as offline. Various *in situ* platforms, including moored buoys, drifting buoys, expendable bathythermographs (XBT), automatic weather stations (AWS), current meter arrays, and many others, contribute their data to INCOIS (see Figure 1). The heterogeneous data acquired vary according to their nature of acquisition (sampling interval, sensor type, media, formats, fixed/mobile or drifting platforms) and their parameter properties (time series, profile, spectral, dynamic and static variability). In addition, some of the parameters are measured directly (temperature) while some are derived (salinity from conductivity) from the observed data. All these diverse datasets need to be quality controlled, organized, and disseminated in real-time to data users via the web and email. Researchers from various scientific organizations and students from universities rely heavily on INCOIS for obtaining data to fulfill their research and academic needs, and INCOIS has been successful in supplying data as per their requirements. However, the diversified and heterogeneous nature of the ocean data has caused problems in handling them in a unified way under single window.

In order to overcome this difficulty, the Ocean Data and Information System (ODIS) was conceived at INCOIS to provide more unified, reliable, and efficient management of the dynamic ocean data along with their analysis and application by means of a backend database. ODIS is a unified system that is capable of archiving, visualizing, and sharing heterogeneous data from a single point. Thus the ODIS serves the needs of the user community by delivering data and preliminary information services at lower costs, manpower and time-scales than has been associated with ocean state monitoring and management previously. Initially, the traditional method of data handling

(e.g., flat files) was used in INCOIS. Over a period of time with a huge inflow of data from various sources, the traditional method proved to be difficult to handle. Conventional methods have difficulties with comprehensive management, analysis, and application of the dynamic ocean data, owing to the spatial and temporal complexity of the data (ZHAO et al., 2009).

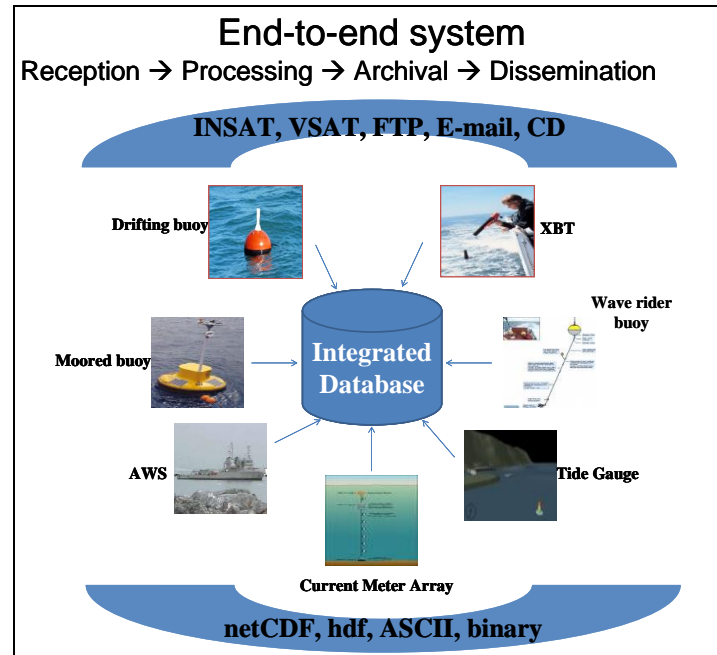


Figure 1. The flow of heterogeneous *in situ* data (Eularian and Lagrangian types) from different observational networks into the centralized database established at INCOIS

New technology and scientific developments have been changing the traditional use of the geographic information system (GIS) (Brovelli, et al., 2003; Goodchild, et al., 2004). Oceanographers have over the years developed their own specialized system, GIS software, for the analysis and display of their data without even knowing that they were in fact using the GIS techniques (Goldsmith, 2009). The use of GIS tools in oceanography has been gaining wider acceptance recently. GIS provides powerful functions to support the comprehensive management of spatio-temporal data. Introducing GIS technology and methods to oceanographic data is a step towards the comprehensive management of dynamic and complex atmospheric and oceanographic data and information (Wang et al., 1999). It is also the major goal envisaged for ODIS. Our aim is for ODIS to become one of a kind in ensuring delivery of heterogeneous data sets, mainly from research organizations and university students, to the oceanographic community. The rest of the paper is organized as follows. Various components of ODIS are described in Section 2. The steps involved in building it and the functionalities associated with it are described in Section 3. The interaction between database and map server is described in Section 4, and visualization and web-GIS functionalities are described in Section 5. Some of the results are summarized in Section 6.

2 COMPONENTS OF ODIS

ODIS adopts a client-server architecture that runs on a single computer or is distributed onto multiple computers. The whole ODIS system is composed of five major components in a Java based programming environment, running on a Linux platform (Figure 2). These components are:

- UMN MapServer
- Apache HTTP Server + Apache Tomcat
- MySQL
- Web Browser

- ChartDirector™

The University of Minnesota (UMN) MapServer is an open source platform that serves the purpose of displaying and querying dynamic data spatially (Mapserver Team, 2013). It supports several Open Geospatial Consortium (OGC) web specifications, including WMS (Web Map Service), non-transactional WFS (Web Feature Service), and GML (Geography Markup Language) (OGC Inc., 2007). The UMN MapServer is used in this context to manage queries and display data from heterogeneous ocean observation systems deployed in the Indian Ocean under various programs. Entire data sets obtained from these observations are loaded into the MySQL data base. MySQL is a multi-threaded, multi-user relational database management system (RDBMS). Although MySQL evolved as an alternative to expensive proprietary software, it is capable and stable in handling high-scale needs as well. The Apache HTTP Server is primarily used to serve both static and dynamic contents on the World Wide Web (WWW). The Apache Tomcat is a J2EE container used for server side data computations. These serve both the dynamic and static content pertaining to the ocean observations via the ODIS. ChartDirector™ is the only proprietary software component that is being used in our ODIS. It serves the purpose of on-the-fly generation of plots to visualize the query sent by the user.

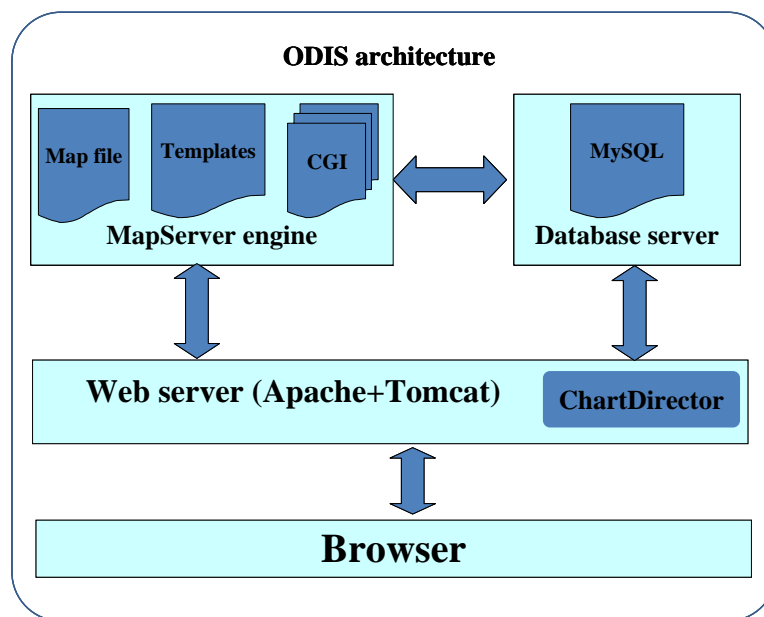


Figure 2. The three level ODIS Architecture for processing the user query and generating the on-the-fly output in the form of data or plots or other related information

All the components of ODIS are loosely coupled for greater flexibility. The open source (OS) modules can be updated or replaced with higher versions/advanced flavors. For example, MySQL can be replaced with PostgreSQL without the need for major changes or disturbing other components of ODIS. This loosely coupled system gives the flexibility for experimental changes to the ODIS system, such as when a new feature needs to be added or an existing one deleted. Although the MapServer and a web browser are enough to implement a simple web GIS, the role of MySQL is to store and organize live processed data into database tables in simple vector form, which can be directly loaded by the GIS engine.

3 BUILDING STEPS

In line with users' requirements for *in situ* data, ODIS was realized in two phases: (i) the development of an integrated data base management system and (ii) inclusion of a web GIS interface.

3.1 Development of the integrated data base management system

The database is configured and automated according to the data flow into and out of the system. In flow and out flow of data into the MySQL database is shown in Figure 3. The main processes involved in the database management system are: Data Acquisition, Data Processing, Quality Checks, and Data Archival. Furthermore, the system disseminates data to users by email or pushes data to an FTP (see Figure 3). The best way to manage *in situ* information is by implementing the relational data management system, which contains both data and metadata. MySQL A database is updated on an operational basis (ships and buoys reporting at synoptic hours) and the updated data is immediately available to users. This operational updating also helps in monitoring (ship tracking and moorings drifting from their position) through ODIS. A variety of applications for input, query, retrieval, visual quality control (VQC), and reporting functions are incorporated into the database. This is designed as a flexible data management system, capable of extending to new platforms when the need arises. It is also based on client/server architecture. MySQL is located on a server PC while other applications are located in different client PCs. A replica of the database is maintained in a backup database server, by means of replication data from the Master MySQL database. This master database is replicated to the Slave MySQL database. Raw data acquired from different media are pushed to the backup FTP Server.

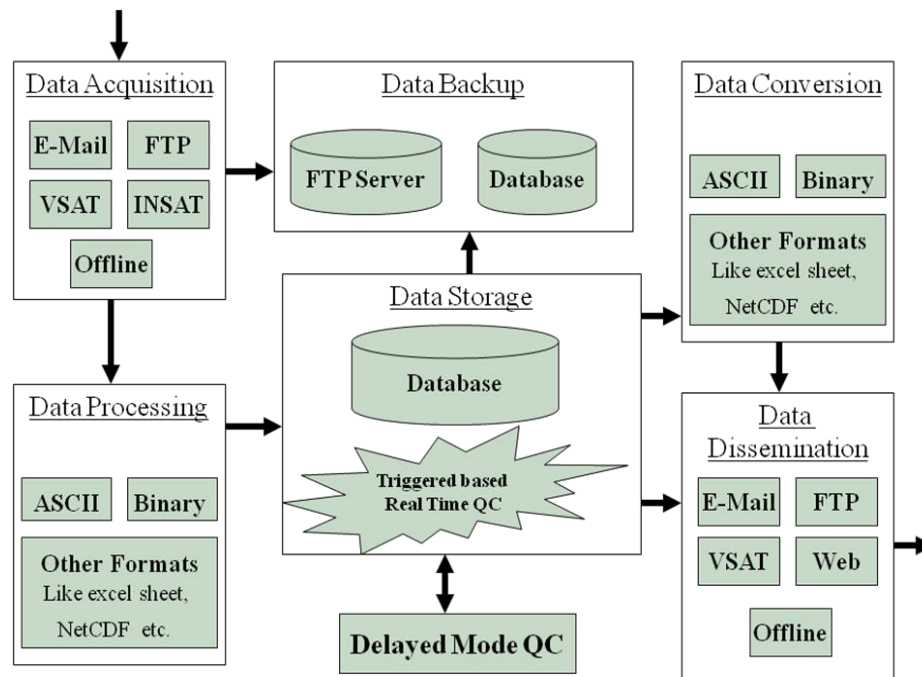


Figure 3. Flow of *in situ* data from various sources into the ODIS system established at INCOIS

It is pertinent to discuss the quality control (QC) of the *in situ* data being pushed into the database. Two types of QC measures are followed based on the type of data set. For data received in real time, quality control is based on a trigger set in the database. For the most part, the data are checked for their range and possible presence of spikes, and a stuck value test is done using this trigger based QC. For instance, data from moored buoys, Automatic Weather Stations, drifting buoys, and wave rider buoys are all quality controlled in this way. Apart from this, these data sets are also quality controlled in a delayed mode (Jayaram et al., 2009). The data that are received in near real time or delayed mode are received either with QC flags assigned by the agency responsible for collection of the data or are quality controlled at INCOIS. Data falling in this category are XBTs, current meters, etc. XBTs are quality controlled based on procedures laid down in the CSIRO cook book (Bailey et al., 1994). Once appropriate QC flags are assigned, the data are pushed into the database. In any case, internationally approved standard QC procedures are used for assigning quality flags to the data.

3.2 The web GIS phase

In order to visualize hidden relations between parameters, long term trends in oceanographic observational data graphs and maps are often used. The map-making function is implemented in ODIS using 1D graphs and 2D distribution maps. These functions have a uniform and easy-to-use user interface. They are developed using the UMN MapServer. The UMN MapServer consists of three important elements:

- Mapfile
- Template files
- CGI program

The Mapfile is the heart of MapServer. It defines the relationships among objects, points MapServer to where data are located, and defines how things are to be drawn. There is a built-in object oriented scripting language in MapServer that defines how to create and use the maps and their layers. The OGR connection (from the OGR library) is used to connect the database in MySQL and web GIS.

Template files are used to define the look of a MapServer CGI application interface and to display the results of the query. They guide the presentation of results, either as a text or as map, to the user. Common HTML page tags are provided by the MapServer and can thus be viewed in a web browser. The CGI program provides a standard protocol for the interfacing browser and web server. It reads and processes the map file settings, the template file, and the user defined variables and returns the processed outputs as maps, variables, or values. Query results are shown in the template files. Every CGI output is a temporary image or value updated at each CGI work session.

Usually the plots displayed through the web GIS are either 1D or 2D plots. A one-dimensional graph is used to represent how a parameter changes as a function of another variable (time, depth, etc.). For instance, to make a 1D graph in ODIS, a user chooses a specific moored buoy, parameter, and starting and ending time. A line plot of the parameter in respect to time is made using the ChartDirector and displayed on-the-fly to the user. Figure 4 shows a line plot of the air pressure on moored buoy AD03. Figure 5 shows the profile plots (temperature against depth) of temperature obtained from an XBT. In ODIS, the 2D distribution maps are currently restricted to the display of data density. For instance, a user can choose an in situ platform(s) and get the data distribution of the parameter between the chosen time periods (Figure 6).

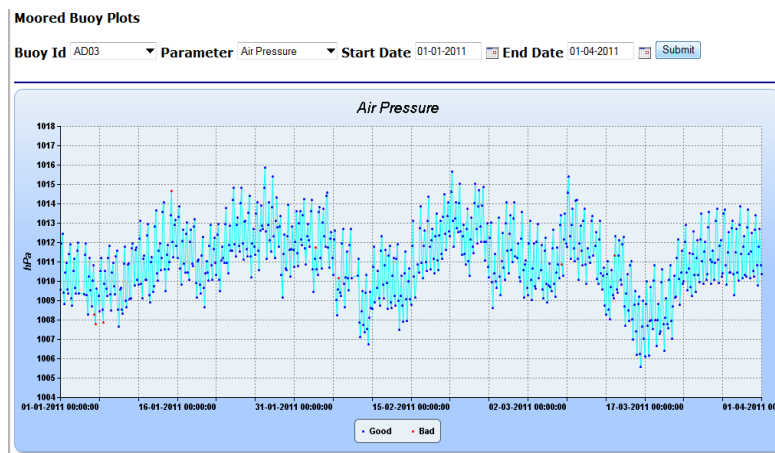


Figure 4. Sample time series plot of air pressure from a moored buoy within a specified period based on a query

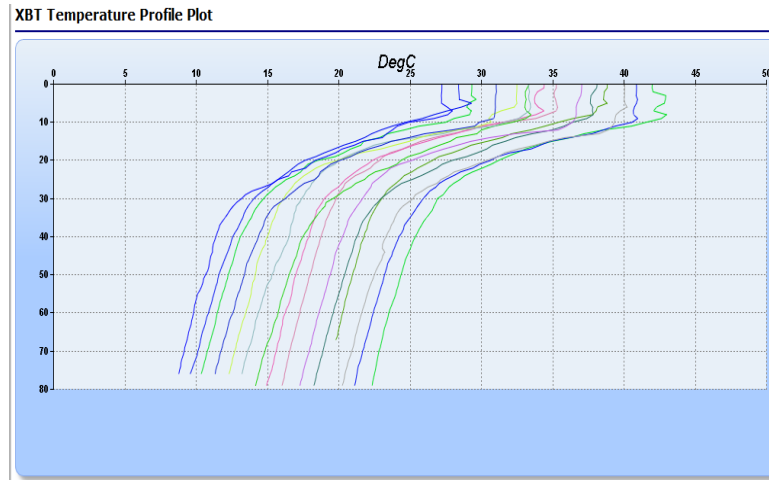


Figure 5. Sample plots of temperature obtained from XBTs within the region of interest of the user

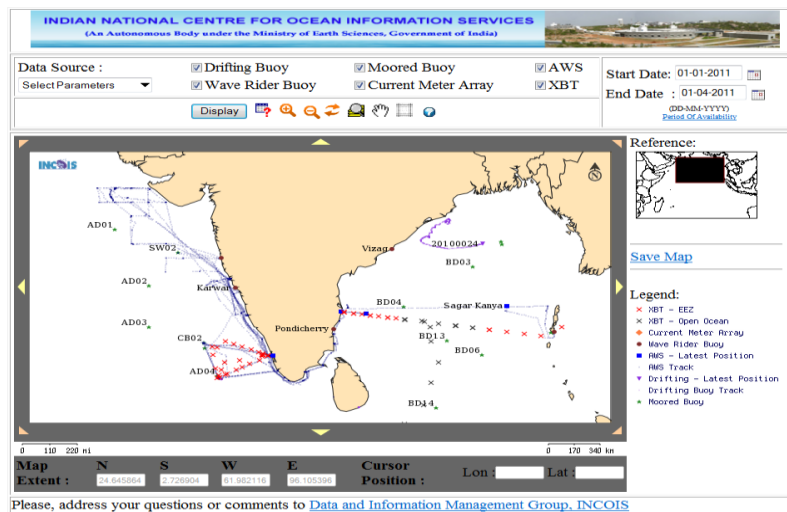


Figure 6. Snapshot of the main web GIS interface through which the user can visualize, query, and extract the information of choice through the main interface

4 MySQL AND ITS INTERACTION WITH MAPSERVER AND VISUALIZATION

The geo-referenced data from different sources are organized and stored in MySQL tables. For loading these data from the database, the necessary query parameters required in a mapfile Layer Object are:

- Connection parameters, in particular the name of the database containing the table to be loaded, user name, and password
- Name table and its geometry column
- Filter description for a SQL query, WHERE clause

This is made possible using OGR/GDAL. The OGR Simple Features Library is a C++ open source library (with command-line tools) providing read (and sometimes write) access to a variety of vector file formats including Shapefiles and MapInfo mid/mif and TAB formats. OGR is actually part of the GDAL library so some references point to GDAL. Figure 7 highlights a sample code base for the extracted information from a specific moored buoy selected by the user.

```

: : : : :
CONNECTION "<OGRVRTDataSource>
  <OGRVRTLayer name='Insitu'>
    <SrcDataSource>
      ODBC:user/password@DSN
    </SrcDataSource>
    <SrcSQL>
      select buoy_id,longitude,latitude
      from insitu_table where longitude between 20 and 120 and
      latitude between -30 and 30
    </SrcSQL>
    <GeometryField encoding='PointFromColumns' x='longitude'
y='latitude' />
    <GeometryType>
      wkbPoint
    </GeometryType>
  </OGRVRTLayer>
</OGRVRTDataSource>"
CONNECTIONTYPE OGR
: : : : :

```

Figure 7. Sample OGR syntax specification used in ODIS for a MapServer interaction with MySQL

To realize on-the-fly visualization, ChartDirector was used here. ChartDirector, a professional chart component for web applications, is involved in presenting the data plots for resulting data layers. The Java Server Page (JSP) program is used for accessing the time series as well as profile datasets from the MySQL database and to display the charts using ChartDirector.

5 WEB GIS STRUCTURE AND FUNCTIONALITIES

The big advantage of using GIS in oceanography, or any other field for that matter, is that the user can dynamically interact through the web GIS interface to obtain the desired result. Through ODIS, users can interact with the GIS engine via the Apache HTTP Server for Map viewing or the Apache Tomcat with ChartDirector for viewing plots resulting from processing the query. Figure 8 represents the work flow of the web interface:

1. Map viewing or querying parameters setting
2. MySQL data loading setting
3. MySQL geo-referenced data return
4. Map views or query results displaying
5. Chart viewing or querying parameters setting
6. Chart views or query results displaying

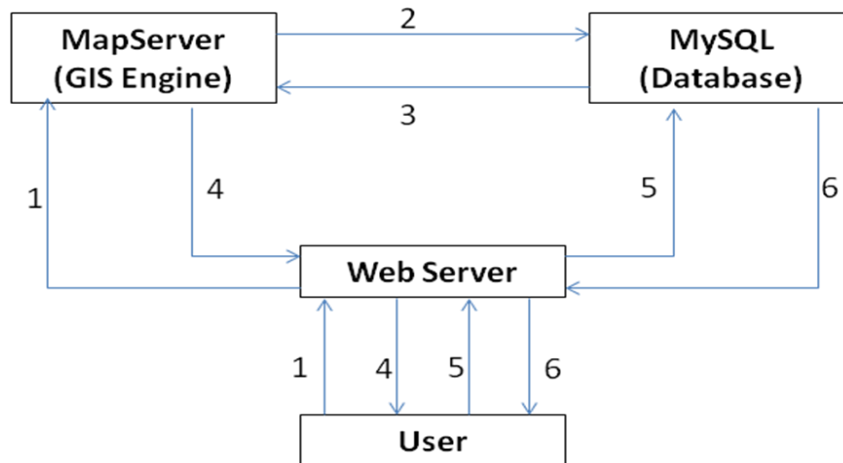


Figure 8. Structure of the adopted web GIS framework. Arrows show the flow of information requested by the user to the MapServer and database and the resultant output back to the user

A user can work with proper MapServer template files and in particular can use the functionalities provided within the MapServer. The functionalities are:

- Data source (layer) selection
- Map browsing (pan, zoom in, zoom out, predefined views)
- Map queries (multiple queries)

Query functionalities are especially useful for obtaining the availability of data from different sources (ocean observations platforms) for the selected region. A user can get the information regarding all the available platforms in the region of interest based on what the web interface allows other functionalities. These web GIS functionalities give the user the power to narrow down the search, view, and filter further until satisfied with the output. These download only the necessary data, which is very helpful as most university students are handicapped by only having access to small amounts of bandwidth.

6 SUMMARY AND CONCLUSION

The use of geographical information systems (GIS) is gaining wide spread acceptance in oceanography and Earth science. INCOIS, the central repository for heterogeneous data sets, receives a voluminous amount of data from various in situ platforms. The diversified and heterogeneous nature of ocean data has caused problems with their handling in a unified way under a single window. In order to overcome this difficulty, INCOIS conceived the Ocean Data and Information System (ODIS). This system provides unified, reliable, and efficient management, analysis, and application of dynamic ocean data through the use of a backend database.

ODIS has been successfully deployed and running at INCOIS since July 2010, clearing many of the problems faced with the old system of handling flat files (which was in place prior to the establishment of ODIS). As a stand-alone desktop GIS system built on open tools, ODIS has achieved effective data management and is providing powerful mapping and visualization tools for oceanographic data. The system has met its performance requirements and the goals of the ocean user community (students and researchers) excellently.

As expected, ODIS is a good example of the large application potential of a GIS to the comprehensive management and visualization of dynamic and complex oceanographic data and information. The integration of real-time in situ data into a web GIS platform delivers a dynamic tool for users ranging from novice to expert in querying the latest scientific datasets obtained from various heterogeneous data sets. The software architectural tools used in the ODIS offer a unique method of data exchange with ever increasing amounts of data and user demand.

However, this system is not perfect as it is built using open source tools. As the name suggests, open source tools are made available by a group of professionals, typically created as a collaborative effort. Some open source tools offer features or performance benefits that surpass the features present in similar commercial software. Some special maps and functionalities presented with commercial software need to be worked out and implemented. As this system is driven greatly by the users, it is always open to future improvements in system and functional up-gradation and implementation of extensions.

7 ACKNOWLEDGEMENTS

The authors thank the Director of the Indian National Centre for Ocean Information Services (INCOIS) for his encouragement. The free MapServer software provided by the University of Minnesota, USA is greatly appreciated. The authors are indebted to colleagues M. Raghavendra, R. S. Mahendra, and P. C. Mohanty for their help in providing MapServer frame work, software development, valuable suggestions, and comments. This is INCOIS contribution number 151.

8 REFERENCES

- Bailey, R., Gronell, A., Phillips, H., Tanner, E., & Meyers, G. (1994) Quality control cookbook for XBT data, Version 1.1. *CSIRO Marine Laboratories Report 221*, pp 55.
- Brovelli, M. A. & Magni, D. (2003) An architectural web GIS application based on MapServer and PostGIS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXIV*, Part 5/W12.
- Goldsmith, R. (2000) Some applications and challenges in extending GIS to oceanographic research. Proceedings of *the 20th Annual ESRI User Conference*, San Diego, CA, Paper 390.
- Goodchild, M.F., Johnston, D.M., Maguire, D.J., & Noronha, V.T. (2004) Distributed and mobile computing. In: McMaster, R.B. & Usery, E.L. (Eds.) *A Research Agenda for Geographic Information Science*, Boca Raton: CRC Press, pp. 257–286.
- Jayaram, Ch., Venkat Shesu, R., Pattabhi Rama Rao, E., Udaya Bhaskar, T.V.S., & Tiwary, S. (2011) Delayed mode quality control procedures for moored buoys in the Indian Seas. DMG-TR-2011-03, Technical Report, May 2011.
- MySQL AB (2008-2009) Sun Microsystems, Inc., *MySQL 5.0 Reference Manual, 2009*. Retrieved from the World Wide Web, August 20, 2013: <http://dev.mysql.com/doc/refman/5.0/en/index.html>
- Open Geospatial Consortium Inc. (2007) OpenGIS® Geography Markup Language (GML) Encoding Standard, OGC 07-036, Version 3.2.1, pp 437.
- The MapServer Team (2013) MapServer: open source web mapping. MapServer documentation, Release 6.2.1, pp 760.
- Wang, H. & Zhu Z. (1999) The progress of ocean geographic information systems at home and abroad. *Remote Sensing Technology and Application 14*(3), pp 49 - 55.
- ZHAO, Q., FAN, H., & LAI, J. (2009) Application of GIS to supporting Atmospheric and Oceanographic Data Management and Visualization. *Geo-spatial Information Science 12*(1), pp 50 - 55.

(Article history: Received 31 January 2013, Accepted 19 August 2013, Available online 28 August 2013)