

A GIS-based tool for storage, selection and visualization of time series 4D marine datasets

Ricardo N. Fernandes · Vasilis D. Valavanis

© Springer Science+Business Media B.V. 2008

Abstract We present an initial attempt for the development of a distributed and scalable GIS-like tool for the storage, selection and visualization of 4D marine datasets. The aim of this tool is to standardize the variety of data available for the water column and support non-technical marine biologists in manipulating 4D marine datasets. The tool is developed towards answering specific environmental and biological questions regarding ocean processes and essential fish habitat mapping in three dimensions.

Keywords Geographic information systems (GIS) · Marine datasets · Database · 3D graphics · 4D visualization

Introduction

Extensive amounts of accumulated and diverse marine datasets need to be commonly organized, accessed and queried by marine scientists (Su et al., 2000). The result of these queries should be displayed using sophisticated techniques, such as 3D graphics

and animations (Chen et al., 2004). Three-dimensional visualizations can contain more information at once and are therefore more suitable for presenting large datasets (Schonhage et al., 2000). However, taking into account another important factor, time (as the fourth dimension) is still one of the challenges in the field. Thus, there is a need for a powerful, yet simple to use, GIS-based tool that stores time series of marine datasets, performs analytical GIS functions (i.e. data selection using certain criteria defined by the user) and visualizes the results in 3D animations. These analytical functions may be used to answer specific questions, such as the spatiotemporal relations between different environmental variables (e.g. 3D mapping of upwelling, fronts, gyres and essential fish habitats).

In many disciplines, large data collections are increasingly emerging as one of the most important resources. In domains as diverse as global climate change, high energy physics and computational genomics, the data volume already reaches terabytes and will soon approach petabytes. User groups with data access and mining needs are often large and almost always geographically distributed (Moore et al., 1998). The accumulated large amount of marine datasets, coming from diverse sources such as remote sensing or electronic sensors, must be efficiently stored to be queried by marine scientists according to specific criteria (e.g. where are the areas that simultaneously have a temperature between 10 and 15°C, salinity greater than 35 psu and chlorophyll greater

Guest editor: V. D. Valavanis
Essential Fish Habitat Mapping in the Mediterranean

R. N. Fernandes (✉) · V. D. Valavanis
Marine GIS Laboratory, Institute of Marine Biological Resources, Hellenic Centre for Marine Research, Thalassocosmos, 71003 Heraklion Crete, Greece
e-mail: ricardo@her.hcmr.gr

than 0.003 mg/m³) and must display the result of these queries in 3D graphics.

Recently, several 3D visualization tools have been developed for terrestrial (Jasnoch et al., 2001; Dunbar, 2003; Grunwald & Barak, 2003) and marine applications (Su & Sheng, 1999; Campbell, 2002; McCann, 2002). Most of these tools visualize specific data formats, although the current stream of applications makes use of latest technologies in data–metadata organization, such as the eXtensible Markup Language (XML). The GIS-based tool described here follows this current stream of applications focusing on the development of a ‘light’ multi-platform tool to be easily used by marine biologists. The tool aims to answer specific biological questions with the capability to store, select and visualize time series 4D (latitude, longitude, depth and time) marine datasets in a distributed and heterogeneous environment (i.e. with different types of equipment or platform, accessing data stored in one or more remote servers). Here, the architecture, the various features of the application as well as the tools used to develop it are described and its functionality is explained.

Characteristics

The tool has been implemented as a client/server architecture, since this is a versatile, message-based and modular infrastructure which assures usability, flexibility, interoperability and scalability (Kinshuk et al., 2001). The main characteristics of this 4D tool are:

A. Distributed. Since the GIS-based tool may have many users (often referred to as clients), it was built to allow many clients to query datasets stored in one or more remote servers, although it can operate in a centralized way, as well (both the client and the server residing in the same machine). Query results are stored in the user’s computer and may be displayed in 3D graphics.

B. Heterogeneous. Since users work with different types of hardware and operating systems, the tool was implemented in Java programming language to support this heterogeneity by being able to be easily transferable and run on any platform.

C. Simple. The user interface is one of the most important parts of any tool, because it is the connection between the tool and the user (Preece

et al., 1994). Therefore, the tool’s interface, based on the graphical user interface (GUI) paradigm, was conceived to be very simple and intuitive to be used by non-technical users. After just a few trials, users start to familiarize themselves with the tool’s interface and its functionalities.

D. Scalable. Since diverse datasets accumulate over time, the tool is able to easily absorb additional data. With the current design of the tool, it is possible to have up to 12 servers capable of storing and querying data simultaneously (i.e. in a parallel way), allowing users to add more servers according to their current needs and the existing volume of data.

E. High-performance. Since a large amount of marine datasets is accessed and queried by many users across different sites and is displayed on local screens as 3D graphics, high computing power is required. Consequently, the tool had been implemented in a way that ensures high performance and usability through the use of Java 3D because of its optimization towards real-time 3D rendering and ability to take advantage of native graphics hardware.

To build all the features mentioned above, the following components have been employed: under a distributed and heterogeneous environment, Java was the natural choice as the programming language, since it is simple, distributed, portable and multithreaded (Gosling, 1995), and thus very suited to implement the logical layer of the tool. Swing and Java 3D were used to develop the user interface, i.e. the presentation layer of the tool. For the persistence layer, MySQL was chosen because of its well proven scalability, flexibility and high performance. The connection layer, between the client and the server, was assured by Java Database Connectivity (JDBC), the industry standard for connectivity between the Java programming language and a wide range of databases. All components are open source and available in the public domain.

Functionality

The tool’s functionality has been divided into four steps (Fig. 1), each performing a well-defined task. Such modularization of the functionality helps the user utilizing the tool, leading him step-by-step through the analysis process.

A. Transformation/standardization. Usually, marine datasets are collected from heterogeneous sources

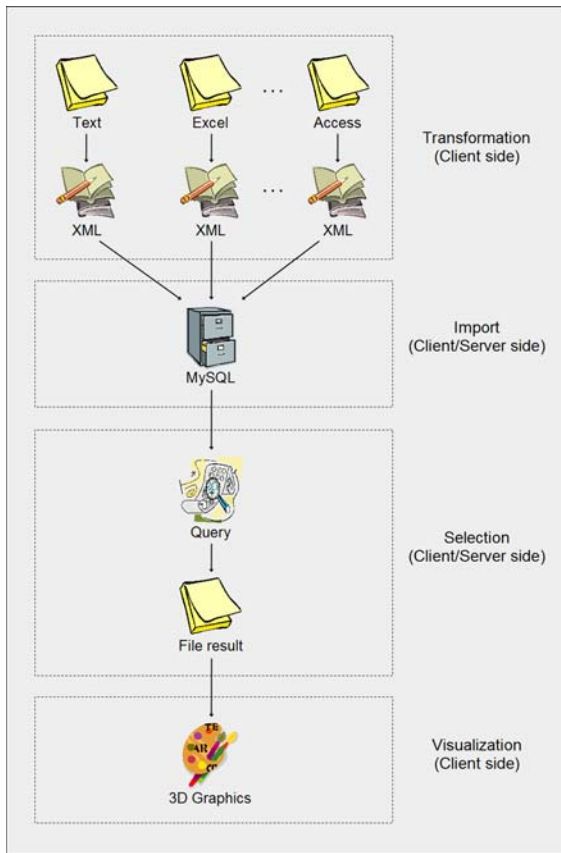


Fig. 1 The data flow through the tool's four-stage functionality

(e.g. remote sensing or electronic sensors) and represented in a variety of different formats. Therefore, the first step is to represent these marine datasets in a standardized format using XML. One of the major benefits of using XML is based on its capability to represent not only data itself but also the nature and structure of that data (metadata). This combination of data and metadata in the same file means XML has considerable advantages (e.g. data quality control or data formatting errors) over traditional data structures (Good, 2005). Marine datasets in any format may be transformed and imported into the system, as long as they contain information about the latitude, longitude, depth, the value of the attribute measured (e.g. salinity, temperature, etc.) and the date of the measurements. For each existing marine dataset, the equivalent in XML is generated with the help of tools especially developed for this purpose, some of them freely available in the public domain. Such tools are easy to use and guide the user

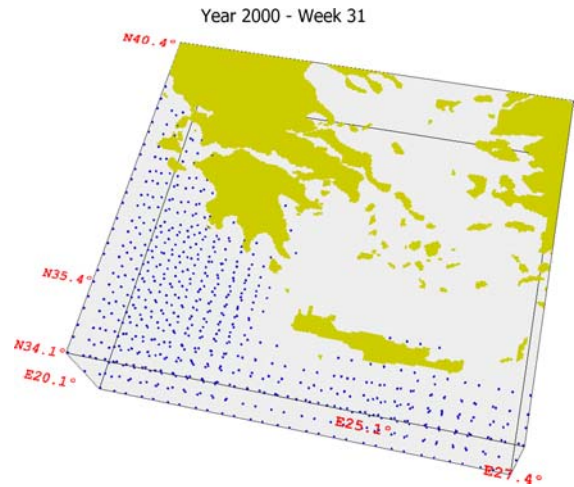


Fig. 2 Example of a 3D graphics showing results from a query of multiple variables

during the transformation process without requiring in-depth knowledge about XML.

B. Import. After the datasets have been transformed into XML, they are imported into the underlying MySQL database. The import process is automated and the user just needs to select the relevant XML files that will be checked for their validity, in terms of structure and range of values, through the use of an XML schema definition (XSD).

C. Selection. Having populated the database, the user may perform queries and apply criteria in order to select the data subsets of interest (e.g. where are the areas where a specific fish species was captured with a weight greater than 1.5 kg and the temperature of the water was below 17°C, during 2006) and save the selection results in his computer.

D. Visualization. Finally, after having performed a certain query, the user may visualize the query result in 3D graphics (Fig. 2). The user can manipulate the 3D display by standard visualization capabilities (e.g. rotation, pan and zoom in/out) but also perform more advanced operations, like navigating through time when displaying the graphics or create an animation (movie) from the sequence of graphics.

Conclusion

This tool merges expertise from computer and marine sciences (i.e. bioinformatics) in order to facilitate the understanding of the ever-changing marine world

through the manipulation and visualization of a high variety of 4D marine datasets (latitude, longitude, depth and time). Targeting users in the marine biology domain, the system aims to facilitate biological questions, such as where optimum living environmental conditions for a species are located or where certain productivity-enhancing ocean processes occur. After a period of evaluation and feedback by end-users, their suggestions will be implemented towards the improvement of the tool. In addition, future work will include the enabling of the tool for web-based usage (as a Java applet) or the embedding of the Web Feature Service standard of the Open Geospatial Consortium. The tool is freely distributed through the following website: <http://arch.her.hcmr.gr>.

Acknowledgements The development of this GIS-based tool was supported by the Sixth Framework Programme of the European Union through the Marie Curie Action ECOSUMMER. We thank the two referees that greatly improved earlier versions with constructive comments.

References

- Campbell, B., P. Collins, H. Hadaway, N. Hedley & M. Storer, 2002. Web3D in ocean science learning environments: Virtual big beef creek. In Wagner, M. G., K. S. Candan & M. T. Beitler (eds), *Proceedings of the Seventh International Conference on 3D Web Technology*, Tempe Arizona: 85–91.
- Chen, S. C., X. Wang, N. Rishé & M. A. Weiss, 2004. A web-based spatial data access system using semantic R-trees. *Information Sciences* 167: 41–61.
- Dunbar, M., 2003. 3D visualization for the analysis of forest cover change. *Proceedings of the ISPRS Commission IV Joint Workshop on Challenges in Geospatial Analysis, Integration and Visualization*, Stuttgart, Germany.
- Good, J., 2005. The benefits and practicalities of using extensible markup language (XML) for the interfacing and control of object-oriented simulations. In Zerger, A. & R. M. Argent (eds), *MODSIM 2005 International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand: 655–661.
- Gosling, J., 1995. Java: An Overview. Sun Microsystems. <http://research.sun.com/features/tenyears/volcd/papers/7Gosling.pdf>. Accessed 12-04-2008.
- Grunwald, S. & P. Barak, 2003. 3D geographic reconstruction and visualization techniques applied to land resource management. *Transactions in GIS* 7: 231–241.
- Jasnoch, U., V. Coors & U. Kretschmer, 2001. Virtual tourist guide. *GIM International* 15: 39–41.
- Kinshuk, H. H., N. Albi, A. Patel & C. Jesshope, 2001. Client-server architecture based integrated system for education at a distance. In Ruokamo, H., O. Nykanen, S. Pohjolainen & P. Hietala (eds), *Intelligent Computer and Communications Technology—Learning in On-line Communities*. Proceedings of the Tenth International PEG Conference, Tampere University of Technology: 57–61.
- McCann, M., 2002. Creating 3D oceanographic data visualizations for the web. In Wagner, M. G., K. S. Candan & M. T. Beitler (eds), *Proceedings of the Seventh International Conference on 3D Web Technology*, Tempe Arizona: 179–184.
- Moore, R., C. Baru, R. Marciano, A. Rajasekar & M. Wan, 1998. Data-intensive computing. In Foster, I. & C. Kesselman (eds), *The Grid: Blueprint for a Future Computing Infrastructure*. Morgan Kaufmann Publishers, San Francisco: 105–129.
- Preece, J., Y. Rogers, H. Sharp, D. Benyon, S. Holland & T. Carey, 1994. *Human-computer Interaction*. Addison-Wesley, Wokingham, UK.
- Schonhage, B., A. van Ballegooij & A. Eliens, 2000. 3D gadgets for business process visualization—A case study. *Proceedings of the Fifth Symposium on Virtual Reality Modeling Language (Web3D-VRML)*. Monterey California: 131–138.
- Su, S. & Y. Sheng, 1999. Visualizing upwelling at Monterey bay in an integrated environment of GIS and scientific visualization. *Marine Geodesy* 22: 93–103.
- Su, Y., J. Slottow & A. Mozes, 2000. Distributing proprietary geographic data on the world wide web: UCLA GIS database and map server. *Computers and Geosciences* 26: 741–749.