Arc Marine

GIS for a Blue Planet

Dawn J. Wright
Michael J. Blongewicz
Patrick N. Halpin
Joe Breman

Foreword by Jane Lubchenco

ESRI Press
REDLANDS, CALIFORNIA

ESRI Press, 380 New York Street, Redlands, California 92373-8100

 First edition 2007
10 09 08 07   1 2 3 4 5 6 7 8 9 10

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data
Arc marine : GIS for a blue planet / Dawn J. Wright ... [et al.] ; foreword by Jane Lubchenco. -- 1st ed.
p.   cm.
Includes bibliographical references.
ISBN 978-1-58948-017-9 (pbk. : alk. paper)
1. Oceanography--Geographic information systems. I. Wright, Dawn J., 1961–
GC38.5.A73 2008
551.460285—dc22   2007000708

Ask for ESRI Press titles at your local bookstore or order by calling 1-800-447-9778. You can also shop online at www.esri.com/esripress. Outside the United States, contact your local ESRI distributor.

ESRI Press titles are distributed to the trade by the following:

In North America:
Ingram Publisher Services
Toll-free telephone: (800) 648-3104
Toll-free fax: (800) 838-1149
E-mail: customerservice@ingrampublisherservices.com

In the United Kingdom, Europe, and the Middle East:
Transatlantic Publishers Group Ltd.
Telephone: 44 20 7373 2515
Fax: 44 20 7244 1018
E-mail: richard@tpgltd.co.uk

Cover and interior design by Savitri Brant

Contents

9

chapter nine

Epilogue

This book has provided a detailed presentation of the Arc Marine data model: its main objectives and intended uses, various aspects of its design, descriptions of the main features and objects, thematic groups and classes, and practical ways to implement the model with data. The previous chapters guide users through the implementation of marine GIS projects with Arc Marine and show them standard ways to describe data and develop GIS tools to consistently and effectively solve ocean and coastal problems. This epilogue closes the book with some views on what the future may hold for this and related ESRI-supported data models.

Basic GIS tasks involving the Internet include searching for appropriate data through spatial data clearinghouses, adding datasets to ArcCatalog for analysis, mapping from the Geography Network and endless other sources, and accessing metadata about datasets. Recent developments in technology include the ability to make maps and data available on the Internet via map servers and the addition of decision support tools to Web GIS sites. This can be done either by downloading an extension to the desktop or the more difficult coding of the analysis functions directly into Web GIS sites, to be used interactively. The Ocean Biogeographic Information System-Spatial Ecological Analysis of Megavertebrate Populations (OBIS-SEAMAP) described in chapter 4 is an excellent example of marine data acquisition. Interactive mapping is one of many key site features (examples may be found at http://marinecoastalgis.net). These are based on commercial solutions such as ArcIMS, or open source solutions such as Minnesota MapServer, PostGIS, and GRASS GIS.

With the recent explosion of Internet mapping sites and data clearinghouses on the Web, the relationship of data models to generic Web services and to project-specific Internet map servers such as OBIS-SEAMAP will certainly increase in importance, as will the emergence of capabilities for streaming data from these sites directly into the Arc Marine structure. Trends in this arena include serving more and more real-time data (often with automated sensors, sensor networks, and wireless technology), the continued popularity of open source, and the continued quest to add more analysis functions to Web GIS (going "beyond mapping"). The coupling of desktop applications or on the server with the Web (e.g., the ESRI ArcWeb, the Microsoft Web programming environment .NET, or the Sun J2EE) will become more commonplace. Another example is the ArcGIS Server that complements ArcGIS desktop by allowing GIS analysts to author maps, globes, and geoprocessing tasks on their desktops and publish them to ArcGIS Server using integrated tools (http://www.esri.com/software/arcgis/arcgisserver/).

Arc Marine will increase the interoperability of tools and data for marine applications by providing standardized data structures for Internet-based, Web-services processes (for a related review see Wright and Halpin 2005). Providing marine data as Web services using Open Geospatial Consortium protocols is increasingly popular. The technology for enabling Internet-based automation centers on the use of Extensible Markup Language (XML), which provides the tag-control encoding for data transfer. The Open Geospatial Consortium has issued spatially explicit specifications for image (Web map service, WMS), vector (Web feature service, WFS), and raster (Web coverage service, WCS) Web services. The request and response communications rely on XML encoding. More specifically, Geographic Markup Language (GML) handles vector representations.

The need for standardized data models also increases as more applications rely on standardized Web-services data to integrate this data into scientific workflows. Developers must anticipate the data structures clients will likely use. Initial development will likely include tools or scripts using the ArcGIS ModelBuilder workflow to harvest data directly from Web services to help offload computational processes through remote grid services.

With the recent rise of Google Earth and the ESRI ArcGIS Explorer as GIS visualization applications, developers have quickly adopted the simplified XML-based representation of spatial objects using Keyhole Markup Language (KML) (e.g., Pilouk and Fine 2006). The wrapping of Web services has more generically gained momentum with the implementation of Simple Object Access Protocol (SOAP), which enables a common set of programming interfaces to Web services. ArcGIS Server will also have a large impact on data and services that researchers will be able to share via the Web. ArcGIS Server combines mapping, visualization, geoprocessing, and data management in one product, while supporting customization using .NET and Java programming (including asynchronous JavaScript and XML, also known as AJAX), and Open Geospatial Consortium WMS, WFS, GML, and Standard Query Language application programming interfaces.

How are these various Internet protocols and Web services related to Arc Marine? The tools for data harvest and grid services can also take advantage of the most sensible relational database storage for the data, that is, the Arc Marine data model. The tools based on the harvested data may more automatically configure the mechanics of the local storage. For instance, one may specify into a tool the bounding box and species of interest, which is then harvested from an OBIS Web service. The tool parses the XML data response into the sensible Arc Marine geodatabase, allowing for subsequent processes to take advantage of the Arc Marine format.

Semantics and ontology will also become critical for marine Internet GIS applications. These applications are the key to successful discovery of data beyond just searching the metadata. Semantics are captured by associating formal terms and descriptions (e.g., “shoreline” versus “coastline”) and making cross-disciplinary connections between them to attach well-defined meaning to data and to other Web resources. This greatly increases the quality of data retrieval or integration based on meaning instead of on mere keywords (Berners-Lee et al. 2001). Ontology is briefly defined as the formalization of concepts and terms used in a practice or discipline (for background see Gruber 1993; Mark et al. 2003). Ontologies can thus provide the semantic aspects of metadata, including lists of terms with definitions, more complex relationships between terms, rules governing those relationships, and potential values for each term. Closely related is the area of semantic interoperability and the semantic Web (Egenhofer 2002). Despite ontologies, words may still mean different things to different people within an interdisciplinary community. How does one, for example, search effectively through shared databases based on the words in the metadata (e.g., coastline versus shoreline, seabed versus seafloor, engineering versus ecological resilience, coastal wetland buffering versus GIS buffering).

In a hypothetical marine Internet GIS scenario, a keyword search for “shoreline” in a data portal may return hundreds of datasets, but a search for “coastline” will return none. Users and developers will need to incorporate innovative changes to metadata catalogs to more effectively search among the existing portal datasets (for an example in ocean and coastal management, see Eleveld et al. 2003). The language of data models may provide the key.

Data portals have been criticized as providing data descriptions only at the syntactic level (i.e., explicit, machine usable), making it difficult for users and providers to interpret or represent the applicable constraints of data, including the related inputs and outputs of analyses or decisions (e.g., Cabral et al. 2004). Compared to a syntactic means, a semantic approach provides higher quality and more relevant information for improved decision-making (Helly et al. 1999; Sheth 1999; Cabral et al. 2004). Semantics deal with meanings of terms that may not be machine usable at the outset. For instance, we know what the term “sea lion” means, but a computer may not initially “know” this and cannot infer additional meaning, such as a “sea lion” is a kind of “marine mammal” and automatically point a user to other datasets that might be related.

Equally important will be the development of multiple spatial and terminological ontologies to define meanings and formal descriptions (Egenhofer 2002; Goodchild 2003). One may think of an ontology at various levels, from a simple catalog (i.e., a list of terms), to a glossary (a definition of those terms), to a thesaurus (the terms and definitions, but with hierarchical relationships between terms and synonyms). All of these are ontologies. A more formal and desirable ontology would be a listing of terms with definitions, more complex relationships between the term, rules governing those relationships, and potential values for each term (in other words, a data model!). Building the necessary tools to define, verify, and deliver these ontologies is a significant research challenge. Researchers must also understand the gaps and inconsistencies in ontologies and handle changes in the material represented by ontologies in ways that go beyond simple versioning (e.g., Fonseca et al. 2002; Cushing et al. 2005). These approaches will be greatly informed by the feature class glossaries of data models (in this case, the feature, object, and relationship classes of Arc Marine and the Common Marine Data Types that they build on).

To implement an effective semantic Web resource, a dataset's ontology should include a vocabulary drawn from its metadata, ultimately revealing which datasets are interoperable. Again, the ArcGIS Marine data model and its Common Marine Data Types may be a natural conceptual framework for identifying important metadata elements unique to different marine datasets for future large distributed data archives and cyberinfrastructures. Ontologies can act as registration mechanisms for vocabularies and as a means of mapping vocabularies to each other using defined relations. Consider the possibility of using relations such as "shoreline same as coastline" or "SST same as sea surface temperature" or "seafloor same as seabed" to map vocabularies. If that were possible, the results could be stored in a collected ontology and used to translate between covocabularies and generate other inferences about the relationships between the different vocabularies and their terms. The benefits of this approach include the following:

- Better and more complete discovery and filtering of data
- Clearer and more precise and computable characterization of data
- Contextualization of information, so that it is provided in the right format, place, and language
- Semantic value, where human users and also computerized inference engines and harvesters can make better use of information, leading to the next item in the list
- Better display of search results, where terms can be substituted if they are equivalent
- Integration into additional tools for data portals, which will then immediately be working with more appropriate datasets

These exciting challenges and developments are being considered now in the context of large ocean observatories with scores of (1) cabled or moored platforms, (2) mobile autonomous systems, and (3) remote-sensing platforms (e.g., the Global Ocean Observing System, GOOS, http://www.ioc-goos.org/, and the Integrated Ocean Observing System, IOOS, http://www.ocean.us/ioos_system). Related efforts such as OBIS, Ocean Research Interactive Observatory Networks (ORION), and more recently the Marine Metadata Interoperability (MMI) project consider scores of critical issues and possible solutions concerning marine data management. Here again, the enterprise solution approach of ESRI-supported data models may make an important contribution. MMI bears especially close watch as it seeks to engage and inform the ocean science community in the creation of interoperable, metadata-centric data systems by (a) providing guidance and reference documentation on properly using and developing metadata, controlled vocabulary, and ontology solutions for the ocean science community; (b) encouraging community involvement in the development and evaluation of those documents; and (c) using test-bed activities (including Arc Marine) to demonstrate cross-platform, cross-disciplinary, interoperable distributed data systems (Bermudez et al. 2005; http://marinemetadata.org/).

While we have described Arc Marine mainly as an isolated resource in this book, another emerging trend is the linking and integrating of two or more data models, where a user combines feature classes from one model with feature classes of another. For example, one could envision the interleaving of Arc Marine with Arc Hydro to study processes from a coastal watershed to an estuary and out into the pelagic ocean. Arc Marine and the Climate and Weather data model might combine to look at air-sea interactions, tracking hurricanes and the like. Arc Marine and the IHO S-57 data model could work together in understanding scientific and resource management applications in the context of navigation issues.

Finally, data models will likely play a larger role in university and professional workshop instruction, but not as initially assumed. Just as there is now a distinction between learning GIS and learning with GIS (Thompson and Buttenfield 1997; Hall-Wallace et al. 2002), we will likely see the use of data models as tools for teaching not just GIS concepts but scientific concepts in a host of disciplines. For example, courses arising throughout the United States use Arc Hydro as a means for teaching students about water resources (water quality, availability, flooding, the natural environment, and management of water resources and surface and groundwater hydrology). Because Arc Marine has been as much about marine science as it has about GIS, some may find the book suitable as a supporting textbook for courses in marine resource management, marine geography, and marine remote sensing. At the other end of the spectrum, the ESRI-supported data models will always provide an effective avenue for teaching students about the advanced features of ArcGIS.

In this and other undertakings, a final note is that this book need not and should not stand alone. There are important resources on the Web (http://dusk.geo.orst.edu/djl/arcgis and the Marine link at http://support.esri.com/datamodels). These include several Arc Marine schemas from the case studies, geodatabases already populated with data from the case studies, the detailed Arc Marine poster, tools and scripts, animations, the Arc Marine tutorial, and various background documents and Microsoft PowerPoint files.

References

Bermudez, L., J. Graybeal, N. Galbraith, and D. J. Wright. 2005. Engaging a community towards marine cyberinfrastructure: Lessons learned from the Marine Metadata Interoperability Initiative. *Eos, Transactions of the American Geophysical Union* 86(52), Fall Meet. Suppl., Abstract IN44A–08.

Berners-Lee, T., J. Hendler, and O. Lassila. 2001. The semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American* 284:34–43.

Cabral, L., J. Domingue, E. Motta, T. Payne, and F. Hakimpour. 2004. Approaches to semantic Web services: An overview and comparisons. In *The semantic Web: Research and applications: First European Semantic Web Symposium, Lecture Notes in Computer Science*, ed. Bussler, C., J. Davies, and D. Fensel, vol. 3053, Sep. 2004, 225–39.

Cushing, J., T. Wilson, L. Delcambre, E. Hovy, et al. 2005. *Preliminary report of eco-informatics & decision making: Defining research objectives for digital government for ecology,* Workshop on Biodiversity and Ecosystem Informatics, NSF, NASA, USGS-NBII, Olympia, Wash., http://www.evergreen.edu/bdei/home.php.

Egenhofer, M. J. 2002. Toward the semantic geospatial Web. In *Proceedings of the Tenth ACM International Symposium on Advances in Geographic Information Systems,* McLean, Va.

Eleveld, M. A., W. B. H. Schrimpf, and A. G. Siegert. 2003. User requirements and information definition for a virtual coastal and marine data warehouse. *Ocean and Coastal Management* 46:487–505.

Fonseca, F. T., M. J. Egenhofer, P. Agouris, and G. Camara. 2002. Using ontologies for integrated geographic information systems. *Transactions in GIS* 6(3): 231–57.

Goodchild, M. F. 2003. The nature and value of geographic information. In *Foundations of geographic information science,* ed. M. Duckham, M. F. Goodchild, and M. F. Worboys, 19–32. New York: Taylor & Francis.

Gruber, T. R. 1993. A translation approach to portable ontologies. *Knowledge Acquisition* 5(2): 199–220.

Hall-Wallace, M. K., C. S. Walker, T. C. Wallace, and R. F. Butler. 2002. Geographic information systems help teach introductory earth science. *EOS, Transactions, American Geophysical Union* 83(31): 333, 339–40.

Helly, J., T. T. Elvins, D. Sutton, and D. Martinez. 1999. A method for interoperable digital libraries and data repositories. *Future Generation Computer Systems* 16:21–8.

Mark, D. M., M. Egenhofer, S. Hirtle, and B. Smith. 2003. *Ontological foundations for geographic information science,* UCGIS long-term research challenge white paper, Leesburg, Va., University Consortium for Geographic Information Science, http://www.ucgis.org.

Pilouk, M., and Y. Fine. 2006. Best practices for developing with ArcGlobe, *Proceedings of the ESRI Developer Summit 2006.* http://gis.esri.com/library/userconf/devsummit06/index.html.

Sheth, A. 1999. Changing focus on interoperability in information systems: From system, syntax, structure to semantics. In *Interoperating geographic information systems,* ed. M. F. Goodchild, M. J. Egenhofer, R. Fegeas, R., and C. A. Kottman, 5-30. New York: Kluwer.

Thompson, D., and B. Buttenfield. 1997. *Learning with GIS, learning about GIS.* UCGIS education priority paper. Washington, D.C.: University Consortium for Geographic Information Science, http://www.ucgis.org/priorities/education/priorities/learning.htm.

Wright, D. J., and P. N. Halpin. 2005. Spatial reasoning for Terra Incognita: Progress and grand challenges of marine GIS. In *Place matters: Geospatial tools for marine science, conservation and management in the Pacific Northwest*, ed. D. J. Wright and A. J. Scholz, 273-87. Corvallis, Ore.: Oregon State University Press.