GeoModeler: Tightly linking spatially-explicit models and data with a GIS for analysis and geovisualization

T. C. Vance
NOAA/NMFS/AFSC
7600 Sand Point Way NE
Seattle, WA 98115 USA
001.206.526.6767
tiffany.c.vance@noaa.gov

N. Merati1, S. M. Mesick2 and
C. W. Moore1
1 NOAA/OAR/PMEL
2 NOAA/NESDIS/NCCDC
nazila.merati@noaa.gov
sharon.mesick@noaa.gov
christopher.moore@noaa

D. J. Wright
Department of Geosciences
Oregon State University
Corvallis, OR, 97331
dawn@dusk.geo.orst.edu

ABSTRACT
Numerical simulation models provide a way to understand and predict the behavior of natural and human systems. Ideally spatially-explicit models would be easily linked to a geographic information system (GIS) for data analysis and visualization. In the past, these two have not been well integrated for scientific uses. This lack of true integration hinders the ability of scientists and managers to create interactive, GIS-based models for research and policy planning. However, GIS packages are starting to expose code and objects to allow closer coupling of core GIS functionality and analytical/modeling tools.

In creating GeoModeler, we have provided a prototype of how one might integrate a GIS with oceanographic and decision-support models. Through the use of Java-based application programming interfaces and connectors, a GIS is directly linked with the Regional Ocean Modeling System (ROMS) and with the Method of Splitting Tsunami (MOST) model. Scientists and managers are able to use a graphical interface to display datasets, select the data to be used in a scenario, set the weights for factors in the model and run the model. The results are returned to the GIS for display and spatial analysis. Three-dimensional visualizations are created using elements of the Visualization Toolkit (VTK) and OpenGL. The project creates a framework for linking to other types of back-end models written in a variety of programming languages.

Categories and Subject Descriptors
J.2 PHYSICAL SCIENCES AND ENGINEERING; I.3.2 Graphics Systems

General Terms
Design, Languages, Management

Keywords
GIS, oceanography, geovisualization, scientific models, 3-D

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference'04, Month 1–2, 2004, City, State, Country.
Copyright 2004 ACM 1-58113-000-0/00/0004...$5.00.

1. INTRODUCTION

1.1 Background

Models provide a way to understand and predict the behavior of natural and human systems. In simplifying these systems, models provide a window to greater understanding. Chorley and Haggett [7] argue that models can take a number of forms - they can be a theory, a hypothesis, a role, a relation or a synthesis; they can include translations in space to create a spatial model, and translations in time to create a temporal model [7]; citing [14]. More recent models attempt to deal with both types of translations in order to create spatio-temporal models. As such, models, while simplifying reality, can still involve complex mathematical and analytical exercises. While some of these calculations are possible within a geographic information system (GIS), others require greater computational abilities than are currently available in a GIS. This requirement for greater computational capabilities produces the need to integrate the graphical and analytical abilities of a GIS with the computational abilities of various programming languages and the processing power of computer servers.

Models can be created within a GIS, created entirely separately using scripting or advanced programming languages, or ideally they might be created using combined techniques yielding the best of both worlds. Models can be implemented within a GIS in a number of ways. They can be loosely coupled, with the GIS used to prepare data for use in a separate computational model. The GIS can be used just to visualize the model output. The model can be implemented using the functionality of the GIS, for example in calculating hillslopes and drainage patterns for an hydrological model. Finally, the model and the GIS can be tightly computationaly coupled, with the GIS used both for the input and for visualization of the output.

GIS-centered models include modeling of seafloor habitat [9]; [13]; [24], models of three-dimensional oceanographic data [24], models of the spread of disease [8] and models of the aesthetics of landscapes [12]. Examples of hydrological models created within a GIS, taking advantage of the native analytical functions of the GIS, include those built for vector data using the ArcHydro Toolset [3], or for raster data, such as TauDEM [15].

Examples of stand-alone models that could then be visualized in a GIS include weather models such as the Community Climate System Model (CCSM), Rapid Update Cycle (RUC) and
ECMWF Re-Analysis models [18]. Tools such as RAMAS GIS use a GIS to organize data for input into a stand-alone habitat model [1]. The MODFLOW groundwater flow model can be visualized with the GRASS GIS [4]; [5]; [6]. Use of a GIS for model parameter setup and the display of results is seen in the watershed management modeling of the ArcView nonpoint source pollution modeling (AVNPSM) system [11].

Application of GIS to oceanographic models has been limited by a number of factors including: an inability to easily handle datasets such as three-dimensional model output and common data formats such as netCDF [28], the lack of volumetric three-dimensional analytical capabilities, and a perception that there is a steep learning curve for their use (e.g. [20]; [21]; [22]; [23]). However, there has been a wide acceptance of GIS by emergency managers and others who routinely use GIS-based decision support and crisis response systems, with integrated scenario modeling and spatial analyses capabilities, to support rapid responses in emergency situations.

GIS-model linkages are a part of the national research agenda of the University Consortium for Geographic Information Science [2]; [19]. In this research agenda, these linkages are seen as a way to introduce dynamic modeling to the GIS community and to integrate computationally intensive applications within a GIS.

1.2 Motivation and Aims
A GIS has great abilities to organize, analyze and display geospatial information. It can provide powerful tools for the integration of data and for the creation of new data products. Ideally, data could be exchanged directly and seamlessly between a model and a GIS. In this instance, the user would choose datasets, define model structures and select parameters for a scenario or model run within the GIS user interface. The model itself would combine spatial analytical tools from the GIS world with scientific modeling capabilities from the theoretical realm. Use of high end processors for the models would create an almost real-time interaction between the model back-end and the GIS front-end. Users would be able to describe a scenario, generate results and re-run the scenario with altered parameters in a timely and efficient manner. The results would be enhanced by the automatic generation of maps and geospatial displays.

GeoModeler addresses a number of needs that can be met by enhancing a standard GIS and coupling it with spatially-explicit models. These include:

1. The need to analyze and display three-dimensional oceanographic data. In contrast to the regularly spaced output of most models, the data are usually collected along transects and may involve individual samples or measurements taken from the surface to the seafloor. The data are dense along tracklines, sparse between tracklines and very dense with depth. Analysis and visualization requires reading in standard three-dimensional data formats such as netCDF and tools for three-dimensional interpolation. Analyses include calculating parameters such as the depth to which the water column is uniformly mixed (the mixed-layer or thermocline depth) and depth integrated temperature or chlorophyll concentrations. For display, the data and analyses need to be combined with baseline data including bathymetry or coastline data.

2. The need to easily set up the inputs for complicated ocean circulation models and creative tools to display the output from the model. Models have numerous parameters and compiler specifications and the typical technique of creating a parameter file is nonintuitive for new users. The output of the models is voluminous and hard to visualize. Being able to quickly create three-dimensional visualizations of one or more model outputs and to view these visualizations interactively makes understanding the output easier. Rapid visualization makes it easier to see errors in model setup and to compute corrected model runs.

3. The need to combine the output of model runs with socio-economic data, such as census data, to calculate affected populations and other emergency management scenarios. While a model may illustrate a physical process, the process can have direct impacts upon human activities. An emergency manager may be much more interested in the impact of an event on the population than the extent or timing of an event.

4. The need to create visualizations and analyses of the intersection of planes and paths with three- and four-dimensional data. Animals follow linear three-dimensional paths. Intersecting a plane with a volume is a good way to visualize structures. Phenomena in the oceans and the atmosphere are inherently three- and four-dimensional. Models provide output in three spatial dimensions and a fourth temporal dimension. Standard GIS packages do not handle truly three-dimensional spatial features such as volumes and convex hulls. There are standard data formats, such as netCDF, for storing four-dimensional data and there are now tools for reading netCDF into standard packages such as ArcGIS. While these data can be read in and animations of two-dimensional slices of the data can be created, analyses are limited. For research it would be useful to be able to do analyses such as the intersection of the three-dimensional line with a volume or the intersection of a volume with a volume to create a new volume.

As GeoModeler is a prototype, some of these needs have been met, and other will require further enhancements and developments.

2. SOFTWARE COMPONENTS
GIS packages are starting to expose software code and objects to allow closer coupling of core GIS functionality and analytical/modeling tools. The GeoModeler application protypes the direct integration of a GIS and modeling capabilities in support of research, management, and decision making. Through the use of Java-based application programming interfaces (APIs) and connectors, a GIS front-end is directly linked with models. Scientists and managers are provided with a GIS-based graphical interface to display datasets, select the data to be used in a scenario, set the weights for factors in the model and run the model. The results are returned to the GIS-based application for display and spatial analysis. The project creates a Java-based framework for back-end models written in a variety of programming languages. The analytical and visualization tools provide more than just a simple coupling of a GIS to a model. GeoModeler provides spatial analytical functions but does not provide temporal functions beyond making animations. It does not provide time series analysis tools such as Fourier analysis or filtering.
The GeoModeler application employs the Java3D API [27], which is designed as a high-level, platform independent 3-D graphics programming API, and is amenable to very high performance implementations across a range of platforms. To optimize rendering, Java3D implementations are layered to take advantage of the native, low-level graphics API available on a given system. In particular, Java3D API implementations are available that utilize OpenGL, Direct3D, and QuickDraw3D. This means that Java3D rendering will be accelerated across the same wide range of systems that are supported by these low-level APIs.

The GeoModeler application also makes use of a second 3-D API called the Visualization Toolkit (VTK) [31]. VTK is a cross-platform 3-D application programming interface built upon, and independent of, the native rendering library (OpenGL, etc). It exposes Java bindings (as well as Tcl and Python). It is written in C++ and includes similar scene-graph, lighting models, and graphic primitives as Java3D. VTK performs Boolean operations on 3-D volumes (intersection and union), volume rendering, filtering (including convolution, FFT, Gaussian, Sobel filters, permutation, and high- and low-pass Butterworth filters), and gradient and divergence calculations. The VTK data model allows for fast topology traversal, making these filters very fast, and allows for rapid mesh decimation. VTK also offers powerful 3-D probe "widgets" that allow easy interaction with the data, and has methods to utilize parallel architecture through the Message Passing Interface (MPI).

ArcGIS Engine and implementations of ArcObjects [25] are used in a GeoModeler component creating a tsunami inundation decision support tool with communication to an ArcGIS-based front-end. The front-end provides both setup - allowing the user to specify the datasets to be used, the weights for elements of the model and the outputs desired, and for display - showing the results of the model run in a map or other spatial output. ArcGIS Engine allows for a Java API while ArcObjects, written in C++, are used for analysis. With the use of the Java connector to ArcIMS, the results may also be displayed in an ArcIMS or other map server. Analyses include intersections, unions, and buffering.

ArcEngine is an ESRI developer product for creating and deploying ArcGIS solutions. It is a simple API-neutral cross-platform development environment for ArcObjects - the C++ component technology framework used to build ArcGIS. ArcObjects are the core of the ArcGIS functionality and include tools such: intersect; proximity – buffer or point distance; surface analysis - aspect, hillshade, or slope; and data conversion - shapefile, coverage or DEM to geodatabase. ArcEngine's object library makes full GIS functionality available though fine and coarse-grained components that can be implemented in Java and other environments. Using ArcEngine, solutions can be built and deployed to users without requiring the ArcGIS Desktop applications (ArcMap, ArcCatalog) to be present on the same machine. It supports all the standard development environments, including Java, and C++, and all the major operating systems. In addition, some of the functionality available in the ArcGIS extensions can be embedded. This product is a developer kit as well as a deployment package for ArcObjects technology.

3. RESULTS AND CASE STUDIES
GeoModeler provides a prototype of an integrated system implementing GIS-based analysis, compute intensive scientific models and geovisualization. The system illustrates the strength of using Java as a linking mechanism and the integration of a number of open-source utilities and analysis tools. It takes advantage of advances in GIS software including APIs to ArcObjects and direct reading of netCDF files. A number of test cases have been created to illustrate the tools available in GeoModeler. These include applications in physical oceanography, tsunami inundation and propagation, and marine mammal studies.

3.1 Displaying and Analyzing Oceanographic Data
Analyzing oceanographic data requires the ability to ingest and manipulate three-dimensional data. A standard GIS can read three-dimensional data, but cannot go beyond displaying two-dimensional slices through the data or 2.5 –dimensional surfaces. In GeoModeler we have created tools to read netCDF files of model output and shapefiles of oceanographic sample data (x and y with multiple z values). GeoModeler is not limited to simply displaying visualizations. We are also developing analytical tools for oceanographic data and integrating them with the GUI. The tools are intended to be familiar both to GIS users and to users of scientific graphics packages. GeoModeler builds upon an earlier tool called OceanGIS [21]. OceanGIS was initially designed to allow 3-D oceanographic calculations on in situ data, and to overlay the results of the calculations on base data such as bathymetry. Tools were developed to perform oceanographic analyses such as calculating basic water properties given by conductivity-temperature-depth (CTD) measurements, such as Figure 1. Tools for analyzing conductivity-temperature-depth data from the Bering Sea, Alaska mixed-layer depth, geostrophic velocity and dynamic height. These tools are integrated into GeoModeler. The GeoModeler interface and display use graphical objects to provide functionality related to the type of data being displayed. As a data layer is added, the relevant tools for analyses are exposed for use. In the example shown in Figure 1, the addition of a layer of CTD data
causes tools for calculating mixed layer depth and other appropriate oceanographic parameters to become available. Data can be read directly from an OPeNDAP server by the Java code in GeoModeler. Data can also be read from shapefiles and directly from netCDF files.

GeoModeler has met the challenge of making analytical calculations on three-dimensional data. It can read in standard data formats and can integrate multiple formats in a single display. Open-source and proprietary tools have been used for data ingestion. Further work is needed to integrate algorithms for three-dimensional interpolation and to expand the types of calculations that can be made, e.g. depth integrated temperature. Other needed enhancements are the inclusion of a variety of surface interpolators. As these types of data are typical of model outputs, GeoModeler has made it easier to see and understand these types of complicated and extensive datasets.

### 3.2 Setting Up the Parameters for Complicated Ocean Models

While models are a very powerful tool, they are also hard to set up and run. Changing parameters and rerunning a model can be a tedious and non-intuitive process. For a new user of a model, the process is especially daunting and can be a barrier to non-modelers using models. A graphical interface for setting parameters and displaying results can make using a model much easier. The Visualization Toolkit (VTK) provides a way to create a visual interface. The initial implementation of GeoModeler creates a VTK interface to the Regional Ocean Modeling System (ROMS) [30]. ROMS is a widely used ocean circulation model. It is a free-surface, terrain-following, primitive equations ocean model that includes accurate and efficient physical and numerical algorithms. It has been expanded for a diverse range of applications (e.g. [10]) and adapted to several coupled models for biogeochemical, bio-optical, sediment, and sea ice applications.

The test application, written in Java, uses an interface that allows the user to modify any of the C-preprocessor directives that ROMS uses in its build-script to enable various physical and numerical options (Figure 2). The initialization file can also be modified to reflect changes in timestep, tiling, and initial conditions. The model is configured, compiled and launched through the GeoModeler interface, and model results are output to an OPeNDAP (Open-source Project for a Network Data Access Protocol) server directory, allowing either viewing in GeoModeler or sharing results with remote colleagues.

The GeoModeler ROMS Data Reader is a GUI class that reads netCDF output of ROMS model data, and allows 3-D renderings to be created and animated in a geo-referenced framework. Since GeoModeler utilizes the GeoTools [26] library, shapefiles of data using standard projections can be rendered simultaneously (as opposed to simply overlaying). ROMS model output is then loaded through the ROMS Data Reader, and the user selects the variable of interest, contour levels, color maps, etc. and can animate the resulting rendering (Figure 3).

Ocean circulation models are not the only type of model used in marine research. Models of the propagation and height of runup of tsunamis are used both by researchers and by emergency managers. As with the ROMS model, integrating a graphical,
GIS-based interface with the model makes use of the model easier.

A second test model for GeoModeler is the implementation of an interface to launch tsunami models and allow the integration of results into a GIS framework. The initial results for this implementation are shown in Figure 4. The model output renderer/animated reads results of the Method of Splitting Tsunami (MOST) model, and surface height is rendered at each of 1440 timesteps. This model run includes "runup", or the adjusting of the model boundary condition to simulate inundation [16]. Rendering this surface over high-resolution topography, overlaid with an aerial photograph produces a fairly realistic view of the event. A complete description of this event can be found in Titov and Synolakis [17].

The MOST model parameters include a series of uniform-sized seismic faults, called “unit sources”, shown in Figure 5. The user loads a shapefile of seismic fault data, shown as yellow rectangles, along the coastal trenches. These sources are initially set using an interface that allows the user to set the magnitude for each source, as well as the vertical distance the fault moves (the slip). An initial condition for running the inundation model is built up from a linear combination of model runs for each unit source. This initial condition is passed to the inundation model, with default parameters pre-set to give a rapid estimate of inundation in time for emergency managers to view results well before the approaching wave strikes populated areas.

GeoModeler has shown that it is possible to use a GIS-based interface to set up model parameters. The resulting application has used VTK and Java3D to create a powerful tool for visualizing model output. Future work includes getting all the parameters for MOST model runs into the interface so the GeoModeler tool can be used for research. GeoModeler will also be used as an interface to other ROMS models and will also be used to display particle tracking models tied to ROMS and other ocean circulation models. Linking models directly with a GIS has improved the ability of modelers to visualize their results and the ability of non-modelers to use these types of models.

### 3.3 Combining the Output of Model Runs with Socio-Economic Data

For model output to be fully useful for emergency managers, it needs to be integrated with related socio-economic and infrastructure data. These types of data are easily available in GIS formats. The GeoModeler framework has proven useful for tsunami modeling, in both the propagation phase and the inundation phase. The MOST model described above is an example of a model of the propagation phase, demonstrating that we can deliver estimates to emergency managers rapidly. However, emergency managers still require tools integrating critical infrastructure, evacuation routings and at risk populations and hazards for decision support system applications. A prototype application was developed using ArcEngine to integrate inundation model output scenarios with critical GIS layers. It is a way of visualizing results and performing common analysis functions that can be found in a GIS, but are difficult to run in standard modeling software packages.

The Tsunami GIS application allows users to create inundation scenarios for pre-calculated near-shore and off-shore sources for a tsunami event in a selected region. The final product is a map of inundated areas and estimates of the affected population in the inundation zone. The tsunami height modeling application shown in Figure 6 is an example of modeling the inundation phase for the city of Seaside, Oregon.

The ArcObjects development environment lets us build a stand-alone GIS application that has the look and feel of a standard ArcGIS desktop, but with tsunami-specific menus added to the standard ArcGIS functionality and editing capabilities. The geoprocessing API operates in the back-end to merge the gridded inundation result with the census information for the area of study to create a new polygon of affected areas. Census data are used to calculate estimates of populations at risk including the elderly or handicapped. Critical infrastructure – hospitals, schools, bridges and emergency centers within the inundation zone can be...
highlighted and standard geoprocessing functions such as proximity analysis can be run to determine mitigation strategies. Users are able to print maps and reports and export map images to be used in operation manuals and reports.

ArcEngine’s framework allows users to add their own data sets (e.g. inundation grids; evacuation routes; infrastructure) as well as links to live data feeds and servers to add current and derived data products on the fly. The ability to customize the application using the ArcObjects modules will allow developers to implement additional models and algorithms as they are developed. These will include adding elevation data and distance from the shoreline to the calculation of populations at risk.

GeoModeler has made it easier to combine the results of models with related socio-economic data to support emergency managers. Use of the ArcObjects exposed through ArcEngine has added GIS analysis functions such as buffering and merging of data layers. Further work is needed to implement other types of GIS analyses. Linking models with socio-economic provides a tool to plan for impacts and mitigation during emergencies. The ability to run the tool and change parameters such as the height of the tsunami supports scenario testing.

3.4 Visualizations and Analyses of the Intersection of Planes and Paths with Three- and Four-Dimensional Data

With the inherent three-dimensional nature of oceanographic or atmospheric model data, GIS spatial analysis tools fall short. Simple calculations such as buffering have a natural extension in three dimensions, and calculating volumes enclosed by isosurfaces and fluxes of properties through an area is common for the geo-scientist. There is a pressing need to extend the standard statistical and spatial analysis tools to three dimensions. The Visualization Toolkit (VTK) has some functionality for determining whether a point in three-space falls inside or outside a closed region, and its data representation allowing for rapid topology traversal makes these calculations efficient.

We are extending this functionality to include calculating the intersection of a line and a volume, a plane and a volume, as well as the intersection of two volumes.

Figure 7 shows the intersection of the path followed by a diving marine mammal with a volume defined by the sea surface and a mixed-layer depth surface as described in Section 3.1. The illustration shows the technique simply as an example, but applications of this technology to marine mammal studies include determining the sensitivity of foraging depth to water temperature and the response of animals to physical and biological oceanographic parameters.

Allowing GIS applications to spatially analyze enclosed volumes is of particular importance in the field of marine hydrothermal vent studies. Hydrothermal vent effluent forms a buoyant plume that rises from the vent caldera and forms a plume with a sub-surface maximum chemical concentration.

Isosurfaces of this neutrally buoyant plume then form completely enclosed volumes that can be probed by intersecting this volume with a plane and contouring concentration at the intersection (Figure 8). User interaction allows the plane to be rotated and translated about all 3 spatial axes, with intersection contours updating in real time. This probing technique can be quite powerful, and can be accomplished as the vent plume evolves in time, providing both an animation of the isosurface, as well as an animation of the probed concentration.

GeoModeler has met the challenge of visualizing arbitrary slices through three-dimensional data through the use of VTK functions. Tools for calculating the intersection of a path in three-dimensions with a volume and volume on volume intersections are being created. Linking models directly with a GIS has improved the ability of a scientist to visually interrogate model results and to create animations to highlight temporal features in the output form a model. All of these tools have made it easier to comprehend and explain the complicated spatio-temporal features seen in both data collected by instruments and data created by models.
4. CONCLUSIONS

GeoModeler provides a fully functional prototype of closely linking scientific models and a GIS. GeoModeler provides more than simple coupling tools to these models. It provides analytical tools specifically adapted to the type of model, and type of model output, being analysed. Future improvements to GeoModeler will integrate model outputs by storing them in a database within the GIS. With this integration, the models will become simply another analysis tool available within GeoModeler. At the same time, the analytical functions available will be expanded and enhanced. While our current emphasis is on spatial analysis tools, future developments may include temporal analyses.

The features of GeoModeler have been applied to a variety of models. The ease with which these models have been implemented suggests that the integration of further models should be straightforward. The ease with which modelers have been able to use new GIS-based functionality has made them more open to implementing GeoModeler. The enhanced visualization capabilities and the ability to easily include other datasets such as GIS-based socioeconomic data will enhance the results of the existing scientific models.

5. ACKNOWLEDGMENTS

Funding for this project was provided by the High Performance Computing and Communications (HPCC) project of the NOAA Office of the Chief Information Officer. For more details please see http://wwwweb.mnw.noaa.gov/hpcc/nwg/. Additional funding was provided by the Pacific Region Integrated Data Enterprise (http://apdre.soest.hawaii.edu/PRIDE/). Additional support for this research was provided by the NOAA/NMFS/AFSC Recruitment Processes program. This publication was supported by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement #NA17RJ1232, PMEL contribution 3084. This research is contribution EcoFOCI-0637 to NOAA's Fisheries-Oceanography Coordinated Investigations. The paper was greatly improved by suggestions provided by May Yuan and two anonymous reviewers. The views expressed herein are those of the author(s) and do not necessarily reflect the views of NOAA or any of its sub-agencies. Mention of software products does not imply endorsement of these products.

6. REFERENCES


Software and data formats:


