SDSC: Envisioning the Future
Research Advances 2007

Zooming in on Earth's Surface, and Below

by Paul Tooby

GEON cyberinfrastructure for the geosciences democratizes access to state-of-the-art tools and data

Geoscientists who study the subterranean world must be ingenious, finding indirect ways to "see" into the earth and unravel the mysteries of the past events that have given rise to the complex 3-D structure of today's earth. A key tool researchers rely on is to gather the most detailed possible topography data. These sharper views hold vital clues to both geological hazards and mineral resources that lie hidden below.

Today, modern laser scans of the earth's surface known as Light Detection And Ranging (LiDAR), one of the hottest tools for precision topography, can capture elevation points every few feet across vast areas. But the flood of data points that sharpens the view also adds up to multi-terabyte (one thousand gigabytes) data sets that present daunting computational challenges for handling, processing, and analyzing, effectively keeping valuable LiDAR topography out of reach for most geoscientists.

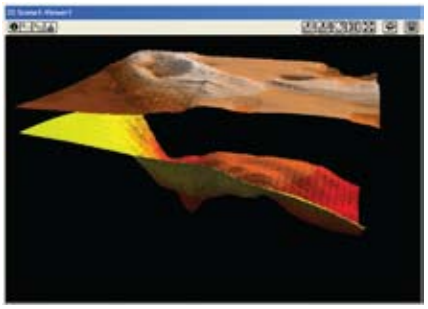
To solve this problem, the GEON "Cyberinfrastructure for the Geosciences" project, with the information technology component based at the San Diego Supercomputer Center (SDSC), is providing broad community access to this key topographic data.

Professor Ramón Arrowsmith, a GEON principal investigator from Arizona State University, and graduate student Chris Crosby have worked with GEON colleagues at SDSC to develop the GEON LiDAR Workflow, which lets users access "raw" LiDAR data and process it on-the-fly in response to user requests through GEON's Web portal, giving geoscientists much sharper views of earth's surface than they could previously get.

The Kepler Scientific Workflow-based system provides access to a growing collection of LiDAR datasets that already contains 7.7 billion data points totaling three terabytes for key geological areas such as the northern San Andreas Fault in California. The data, which is hosted at SDSC, is searched on a 128 gigabyte "fat node" of SDSC's DataStar system. So far, some 100 users have submitted more than 1,500 jobs and processed 27.5 billion data points.

Cyberinfrastructure for the Geosciences

Resources such as GEON's LiDAR Workflow represent one of the most exciting promises of today's advancing information technologies, fulfilling the vision of "science at your fingertips." The pioneering GEON initiative, a collaboration between Earth Science and Computer Science researchers, is providing



Seamless Access to Synthetic Seismograms
Through the GEON Portal users reach the Web interface to the SYNSEIS synthetic seismogram tool for modeling 2-D and 3-D seismograms of earthquake ground motion. Seamlessly accessing diverse input data sets such as the subsurface model shown and powerful SDSC supercomputers behind the scenes, users build complete simulation models, submit jobs, and monitor and receive job output. Image: Dogan Seber, SDSC/UCSD.

the geoscience community a cyberenvironment that integrates services, tools, applications, high performance computing resources, and diverse data collections, all in a unified portal.

"GEON is bringing the power of data cyberinfrastructure to the geoscience community," said SDSC's Chaitan Baru, a GEON principal investigator and information technology research lead. "Users span universities nationwide as well as the U.S. Geological Survey, industry geophysicists, state agencies, and even a high school student in Wisconsin who used our LiDAR topography system for his science fair project."

A collaboration among sixteen institutions and an array of partner projects and agencies, including the USGS, GEON is funded by the National Science Foundation (NSF) Information Technology Research program. GEON is also a Science Gateway in the NSF TeraGrid.

Seamless Synthetic Seismograms

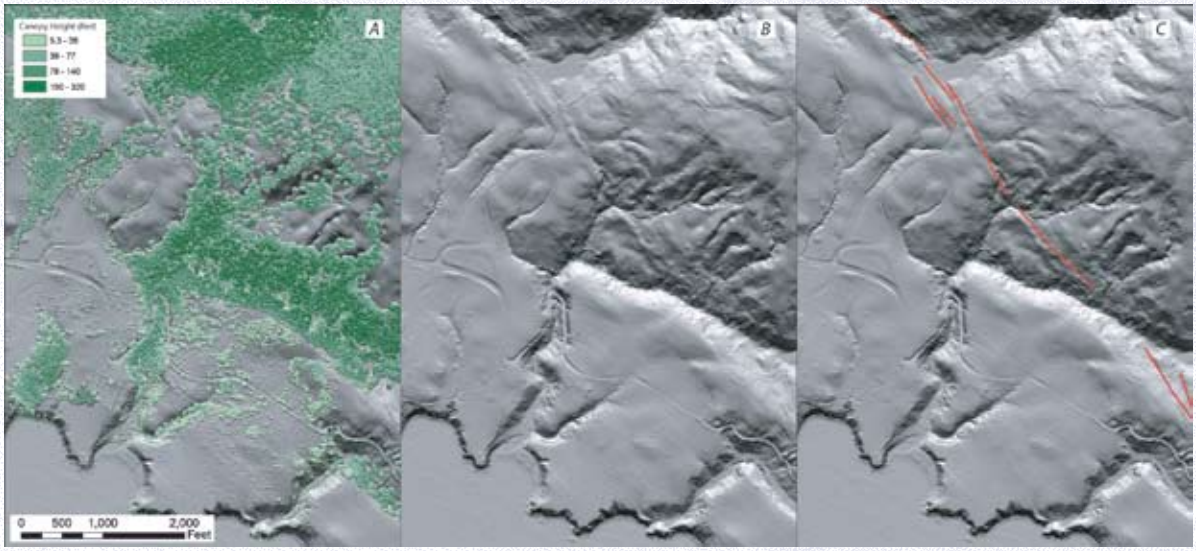
In addition to providing access to key data such as the LiDAR topography, as part of GEON's cyberenvironment the researchers have also developed an easy-to-use Web interface to a modeling environment that lets users compute seismograms, both two and three-dimensional regional seismic waveforms.

SYNSEIS, for SYNthetic SEISmogram generation tool, gives users, from novice students to experienced researchers, access to a sophisticated finite-difference simulation code, E3D, developed at the Lawrence Livermore National Laboratory, running behind the Web interface on powerful TeraGrid resources.

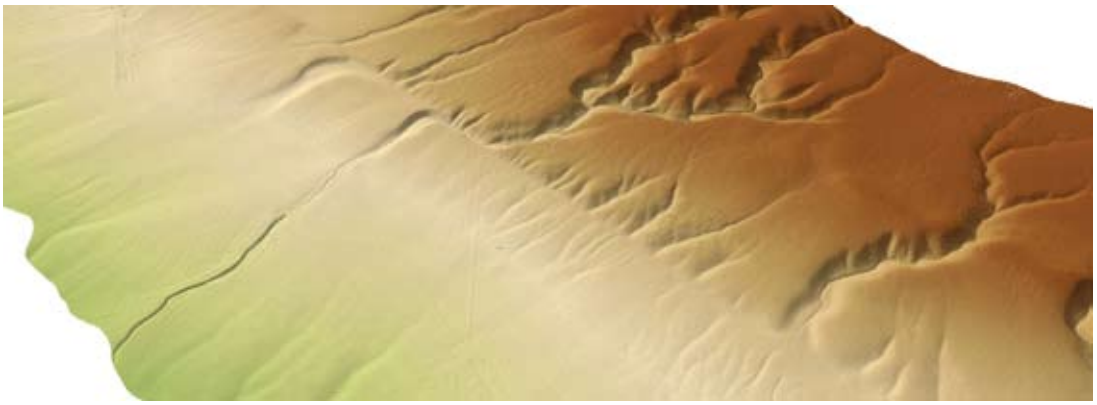
For example, researchers in the large-scale EarthScope project, who need to rapidly incorporate data from hundreds of seismic stations into their models, can interactively set their study region, access seismic event and station locations, extract earthquake waveforms on the fly, and compute a synthetic seismogram using built-in tools—and do this far more quickly than previously possible.

Seeing how well the model agrees with data gives new insights into the structure of the hidden subsurface world below. The system can also be used in reverse, starting with a subsurface structure model and generating an artificial earthquake in order to see what the seismogram will look like. That is, the system can model the resulting ground motion from the virtual earthquake at various locations.

"It's amazing to geoscientists that they can run this powerful tool on national-scale supercomputing resources at SDSC right from their desktop through a simple Web interface," said Dogan Seber, a GEON PI and director of SDSC's Geoinformatics Lab, who is leading development of the system. "The GEON portal hides the complexities of the code, account management, CPU allocation, and many other things, bringing this resource within reach for a much wider range of scientists and students."



High-Resolution Topography for Geoscience
LiDAR laser scans generate digital topography models ten times more accurate than before. These hillshade images of the Northern San Andreas Fault were produced with the GEON LiDAR Workflow. A. topography and overlying vegetation, darker green indicates taller canopy height; B. bare earth terrain (vegetation digitally removed) for the same area; C. bare-earth terrain showing the San Andreas fault in red. Image: Christopher Crosby, ASU/GEON.



Perspective View of the San Andreas Fault
This view, looking north, was created from detailed LiDAR topography data processed with the GEON LiDAR Workflow. The San Andreas Fault is the northwest-southeast linear feature. Note how Wallace Creek (upper left) has been offset more than 120 meters by repeated earthquake slip along the fault. Farther northwest is an older channel abandoned 3,700 years ago. Image: Christopher Crosby, ASU/GEON.

A Model for Other Disciplines

In addition to providing advanced resources to geoscientists such as the LiDAR and SYNSEIS projects and the Paleointegration Project, which provides seamless searching and ontology-based integration of extensive fossil and sedimentary rock databases and paleomapping tools, GEON's portal is helping educate a new generation of cyberinfrastructure-savvy researchers. Ongoing workshops address a range of areas from ontologies and data integration to visualization in the geosciences, and GEON researchers were influential in launching the new Geological Society of America-sponsored Geoinformatics meeting, now held annually.

For the past four summers, geoscientists have gathered at SDSC to attend the first-of-its-kind Cyberinfrastructure Summer Institute for Geoscientists, which introduces researchers to commonly-used as well as emergent information technology tools.

Because GEON's core cyberinfrastructure is generic and standards-based, it is proving useful in a range of other disciplines. More than a half dozen other projects are adopting GEON cyberinfrastructure, including the EarthScope Data Portal in the geosciences, the Archaeoinformatics archaeology consortium, the Tropical Ecological Assessment and Monitoring (TEAM) network, the Chesapeake Bay Environmental Observatory project, the Network for Earthquake Engineering Simulation (NEESit) in earthquake engineering, and the NEON ecology testbed.

Participants

A full list of GEON participants is available on the GEON website. LiDAR: Ashraf Memon and Viswanath Nandigam, SDSC/UCSD; Christopher Crosby, Newton Alex, Josh Coyan Jeff Conner, and Gilead Wurman, ASU; SYNSEIS: Tim Kaiser, Choonhan Youn, and Cindy Santini, SDSC/UCSD; Paleointegration Project: David Rowley, U. Chicago; Judith Parrish, U. Idaho; David Weishampel, JHU; Chris Scotese, UT Arlington; John Alroy, UCSB

Project Leaders

GEON: Chaitan Baru and Dogan Seber, SDSC/UCSD; LiDAR: Ramón Arrowsmith, ASU; SYNSEIS: Dogan Seber, SDSC/UCSD; Paleointegration Project: Allister Rees, U. Arizona

Related Links

GEON Website http://www.geongrid.org

GEON Portal https://portal.geongrid.org