Quakesim Computational Environment

Andrea Donnellan, Jay Parker, Margaret Glasscoe, Robert Granat
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
818-354-4737
Andrea.Donnellan@jpl.nasa.gov

John Rundle
Department of Physics, University of California, Davis
One Shields Avenue
Davis, CA, 95616

Dennis McLeod, and Rami Al-Ghanmi
University of Southern California
Mail Code 0781, 3651 Trousdale Parkway
Los Angeles, CA 90089

Marlon Pierce and Geoffrey Fox
Indiana University
2719 East 10th Street
Bloomington IN 47408

Lisa Grant
University of California, Irvine
Irvine, CA 92697

Abstract— QuakeSim is a computational environment that integrates multiple heterogeneous data sets, crustal deformation modeling tools, and pattern recognition techniques for studying earthquake processes and forecasting their behavior. Recent developments in QuakeSim include improved mapping and visualization tools for exploring and selecting data, enhancement to model applications, addition of UAVSAR data to the QuakeTables database, and improved pattern analysis methods. The recent magnitude 7.2 El-Mayor/Cucapah earthquake in Baja California that occurred in April 4, 2010 has provided a useful testbed and science environment. RDAHMM, disloc, simplex, Virtual California, and RIPI have been used to analyze motions that occurred as a result of the earthquake as well as address the implications for future earthquakes in southern California. Analysis of Virtual California simulations suggests that the long strike-slip faults in southern California could rupture following a Baja type event. GPS time series as analyzed through RDAHMM showed state changes on the San Andreas, Elsinore, and San Jacinto faults in conjunction with the earthquake. The locations of these state changes correlated to creep events observed on these faults. A northward propagation of state changes several weeks after the earthquake culminated in the M 5.4 Collins Valley earthquake near the Coyote Creek Segment of the San Jacinto Fault. The RIPI forecasting methodology was used to target regions of higher hazard in California, which included southern California north of the rupture termination. Currently Virtual California handles vertical strike-slip faults. Combining Virtual California with GeoFEST will allow for studying interacting faults of any orientation. We are populating the QuakeTables database with the best estimates of the faults in California and are adding UAVSAR data as they become available.

TABLE OF CONTENTS

1. INTRODUCTION

QuakeSim is a project to integrate computational infrastructure, remote sensing data, pattern analysis, and model applications into a seamless environment. The data in the project are housed in the QuakeTables database or are accessed remotely via web services. Model applications are available through a web portal and results from pattern analysis are continuously displayed and updated on the QuakeSim web page. The web page is being transitioned to Drupal to fully utilize a content management system. The QuakeTables database and model applications are accessed through the web page, but are developed independently from any content management system for greater flexibility. The focus has been to make the site much more user-friendly, cohesive, and to more directly integrate QuakeSim applications into a single homogeneous interface. Where possible we have moved away from requiring logging in to access the portal, though the log in area allows for users to save, access, and copy projects for better science analysis.

2. QuakeTables

The QuakeTables database houses fault data from various studies, and InSAR data from from UAVSAR and spaceborne missions. We have worked with the UAVSAR team to improve the data format for UAVSAR as well as present the data in a form useful to modelers. All of the UAVSAR data are available in a map browsable interface and all products are available for download. We have added timestamping to the files to that they can be viewed in Google Earth in timeframes for which the data were collected.

QuakeTables uses a reference fault specification called "QuakeSim Format" as its reference format for fault data. When a new fault dataset is incorporated into QuakeTables, a conversion scheme from and to QuakeTables is set as part of the metadata for this new dataset. The QuakeTables team has also continued participating in the Southern California Earthquake Center (SCEC) Earthquake Simulators Comparison project in an advising role. Our function has been in discussing tools and formats available for the project to share and present information.

3. Portal, Mapping and Visualization Tools

Several enhancements have been made to the portal based on lessons learned from science analysis, assessment of user access, and response to recent earthquakes. In order to improve ease of access and encourage use of QuakeSim tools we have begun to move applications outside of a login to the portal. We developed "anonymous" versions of the RDAHMM and Disloc web interfaces that can be used in the new quakesim.org web site. These do not require users to create accounts. The time series and pattern analysis tools have also been moved outside of the login area.

Using QuakeSim for science analysis has resulted in several improvements to the interface. The most recent Uniform California Earthquake Rupture Fault 2.4 set is now available to the crustal deformation applications. Users can browse the database and select faults to model or can draw faults directly on a map for modeling. The inversion software, Simplex, now accesses UNAVCO GPS velocity solutions. We learned that GPS velocity solutions are put out in various time frames and in various reference frames. The QuakeSim applications need to access as many of these solutions as possible, because different solutions are impacted by various earthquakes, post-seismic motions, or are easier to interpret in one reference frame versus another.

Several revisions to both Simplex and Disloc Web user interfaces and associated plotting tools to improve their usability. A key enhancement is in the plotting service. Disloc produces interferogram by calculating elevation angle parameters for each pixel from UAVSAR RPI metadata file (airplane heading and altitude) if it is available. As a result there is a better match with UAVSAR fringe interferograms. Work has also been initiated to automatically show the appropriate number of vectors on the map.

The recent Tohoku-Oki earthquake highlighted the need for low-latency GPS position time series for RDAHMM analysis. As a result we added UNAVCO data sets to the RDAHMM analysis pipeline and updated the user interfaces to display this information. We are also in discussion with other GPS analysis centers about obtaining more rapid solutions. These centers recognize the value of RDAHMM for highlighting state changes in GPS stations. The volume of data an solutions makes it too difficult to analyze the data manually and RDAHMM provides an automated approach for doing this.

3. Implementing Interacting Fault Models

Interacting faults have been implemented into GeoFEST. GeoFEST has a robust implementation of split node faults, which rather than responding to stress, move by a commanded amount. Previous ideas for stress-driven faults, both in GeoFEST and Pylith, involve definition of a novel type of element that is designed to respond to ambient stress through internal forces and/or rheological yielding. Such approaches involve the building and testing of significant new physics and coding. The process was simplified through the use of existing functions. Iterative slip adjustment is performed on split node faults in the model, in order to minimize elastic strain energy. The fault senses failure conditions and triggers slip initiation on the fault(s).

The central idea is to have the user identify nominal or "characteristic" fault patches and slip amounts. When a failure stress criterion is reached on a fault patch, the nominal slip is provisionally applied to it. Then an iteration is performed in which the slip amount is varied in an attempt to minimize global strain energy. Ideally, within a few iterations, the "best" slip for that particular event and time will be found, at which point regular time stepping is resumed until such time as another patch is ripe to fail. Relatively minor high-level coding was necessary to get the code to function in this way, and it adds the ability for the code to internally evolve a realistic history of slip intervals and amounts without pre-programming by the user.

4. Recent Earthquakes

The M 7.2 El-Mayor/Cucapah earthquake that occurred in Mexico on April 4, 2010 was well instrumented with continuous GPS stations in California. Large offsets were observed at the GPS stations as a result of deformation from the earthquake providing information about the co-seismic fault slip as well as fault slip from large aftershocks. The

April 11, 2011 M 9.0 Tohoku-Oki earthquake in Japan was also well instrumented and large enough that 30 minute GPS position time series could be analyzed for information. Information can also be obtained from the position time series at each station in both earthquakes.

We want to understand the details of ground displacement as measured by GPS**.** RDAHMM uses statistical, data-driven approaches rather than basing results on physical models. This avoids bias due to models and can be completely automated, which is important as the number of continuously operating GPS stations grows. RDAHMM does not answer science questions directly, but allows for identification of interesting behavior**.** We can define "interesting" in various ways.

RDAHMM uses Hidden Markov models (HMMs), which allow us to segment GPS time series into discrete modes in order to extract information. Each mode corresponds to an underlying "hidden state" of a model and is described by the statistics of its member observations. This method can be done in the absence of labeled training data or other human supervision and is entirely a data-driven approach.

Fitting a hidden Markov model in the absence of a priori information using the standard expectation-maximization (EM) method is a difficult problem due to the presence of numerous local maxima in the objective function. This problem is addressed through the use of the regularized deterministic annealing EM (RDAEM) algorithm. The algorithm produces stable, high-quality model fits and has been implemented in the QuakeSim RDAHMM software package.

Daily position time series were analyzed fro the JPL GIPSY GPS solutions in southern California. Users can explore data sets and analysis results on both the micro-scale (individual station) and macro-scale (whole network). Users can focus in on or scroll through particular spatial or temporal time windows and observe dynamic behavior by created movies that display the system state. Analysis of the data for several months before and after the El Mayor-Cucapah earthquake suggest that a creep event may have occurred starting about 1 week before the event. Stations change state before the earthquake in a northward progression at a rate of about 15 km per day.

There are no to few state changes in the Salton Trough to Eastern California Shear Zone region for several months before the M 7.2 El Mayor-Cucapah earthquake that occurred on April 4, 2010. Seven days before the event a station just north of the Salton Sea and near the San Andreas fault changed state. A creep event was observed in the geologic and GPS data associated with the April 4, 2010 M 7.2 earthquake and is seen in this time series plot, but after the station changed state.

Five days before the event a station about 30 km north of the San Andreas station changes state. The station changes state frequently before the earthquake and several months after the earthquake. It does not change state for several months following the earthquake, perhaps representing a creep stress smoothing. Creep events can be observed in latitude (north) on the middle plot associated with the event, and starting around June 10, 2010 just before a M 5.7 aftershock occurred on the northern end of the rupture. Similar southward creep suggesting right-lateral fault slip is observed at a station in the Eastern California Shear Zone. This station changes state many times before the April 4, 2010 earthquake and June 15, 2010 aftershock, but does not change state following these events.

Only one station, that has a large offset and was near the mainshock ruptpure, shows a state change the day of the earthquake. The event occurred at 22:40 UTC and the analysis is done in UTC time, so the earthquake happened at the end of that day. This station had a large offset and indicates a state change the day of the event. Several station changed state the next day indicating a direct relation to the mainshock. The time series shows a rapid decay in longitude (west) immediately following the earthquake.

Analysis of time series data around the 2010 El-Mayor/Cucapah earthquake shows that GPS stations change state near the rupture as expected. GPS stations also change state farther from the rupture, but near other faults, such as the San Andreas and San Jacinto faults. The GPS stations that change state associated with the earthquake tend to be in close proximity to faults that exhibited creep events associated with the earthquake in Baja.

30 minute solutions for the Tohoku-Oki earthquake were produced for several weeks before the earthquake and for a few days after the earthquake by the JPL/Caltech ARIA team using Japan Geonet GPS data. Initiation of the rupture is indicated in the 0530 – 0600 solution. State changes propagate away from the rupture with a timescale of about 1.5 hours indicating a fast postseismic process, but one that is much slower than the rupture propagation. Following the earthquake there are no systematic state changes until about two days after the earthquake when GPS stations west of the triple junction between the North American, Pacific, and Phillipine Sea plates frequently change state. This region will be analyzed more closely for a geophysical explanation.

5. The National Level Exercise

During the week of May 16th an earthquake was simulated for the Central United States. QuakeSim tools were used to generate deformation products for this National Level Exercise for the simulated earthquakes in the Central US. Output was used to calculate slope changes and as input to E-DECIDER for construction of PESH files for HAZUS. PESH files represent Potential Earth Science Hazard. HAZUS is FEMA's methodology for estimating potential losses from disasters. All products were ingested into ESRI's Common Operational Picture (COP).

Disloc was also used to illustrate the size of the New Madrid earthquake sequence in 1811-1812. The Central United States experienced a series of large earthquakes in late 1811 and early 1812. Simulations of slip on the faults show how the Earth's crust deformed as a result of these earthquakes. These initial disloc models will be used as input cases for using the interacting fault model implementation in GeoFEST. The visualization tools can be used to show how varying radar observation parameters change observables from the events.

6. CONCLUSIONS

QuakeSim has undergone numerous improvements in the last year. This has been driven largely by the tools maturing to the point that they can be used for science analysis. This analysis provides an efficient means for indicating where further improvements can be made in the model applications as well as interfaces. Users are sometimes reticent to create and use logins for conducting analysis and as a result we are making more of the QuakeSim tools publically available. This has the drawback of users not being able to save, modify, or reuse projects, but allows for rapid model development and analysis.

The recent earthquakes have served to provide real scenarios for use of QuakeSim and have spurred many improvements in the interfaces. As advanced information system technology tools are developed it is important to engage with end users to ensure utility of the tools and optimize their capability.

BIOGRAPHY

Andrea Donnellan *is a geophysicist at NASA's Jet Propulsion Laboratory, and a research professor at the University of Southern California. Andrea studies earthquakes and crustal deformation by integrating satellite technology with high performance computer models. She is Principal Investigator of NASA's QuakeSim project, as well as supercomputing, earthquake modeling, and UAVSAR projects. Donnellan has also been Deputy Manager of the JPL's Science Division, Pre-Project Scientist of a mission to study natural hazards, ice sheets, and ecosystems, and NASA's Applied Sciences Program Area Co-Lead for Natural Disasters. She has conducted field studies in California, in Antarctica, on the Altiplano of Bolivia, in Mongolia, and on Variegated Glacier in Alaska. She has been a geophysicist at JPL since 1993. She received a bachelor's degree from the Ohio State University in 1986, with a geology major and mathematics minor. She received her master's and Ph.D. in geophysics from Caltech's Seismological Laboratory in 1988 and 1991 respectively and held a National Research Council postdoctoral fellowship at NASA's Goddard Space Flight Center. Donnellan received an M.S. in Computer Science from the University of Southern California in 2003.*

Table 2. Summary of Style Requirements

Paper and Size: Paper size/type 8 1/2" x 11” White Number of pages 6–20, (Or longer if required by the nature of the material. Invited papers are individually determined.) Margins: Top and bottom margins 3/4" Left and right margins 3/4" Columns: Number of columns 2 Space between 1/4" Title: Title typeface Bold 20 pt Times Roman upper & lower case, text centered on the full-page width. Initial caps on all words except articles. Keep less than 100 characters Author: Typeface 10 pt Times Roman upper and lower case centered on the full-page width. Include affiliation, address, phone number, and e-mail. Do not include degrees or titles except military rank. Text: • Typeface: Times Roman 10 pt • Line to line spacing single • Space after paragraph 10 pt • Paragraph indent None • Justification Left & right Acronyms: Define all acronyms on first usage. Page numbers: Bottom center of every page. Footnotes: • Font size 8 point Copyright Notice Footnote: Include a copyright notice as a footnote on the first page (government employees see text above): 978-1-4244-3888-4/10/$25.00 ©2010 IEEE. Paper number Footnote: Put your paper number in a footnote on the first page. Clearly mark a numeric version number so that we review the latest one. *Can’t Find a Rule for Formatting?* Where this document is silent on a formatting question, it is because it is not important or is the writer’s option. When in doubt, make a choice that makes your document most readable.	Headings: • Spacing before major or subheading: Double space • Spacing after major or subheading 1 1/2 space • Major headingsCenter, use 12 point Small Caps, Bold • Subheadings Italic, flush left, separate line, same size as text • Subsubheadings: Italic, run into paragraph with em dash Equations: • Where? ....Centered • Equation numbers: In parentheses, flush with right of column. Include special fonts with the paper. Figure and Table Titles: • Where? Centered directly below figure • Where? Centered directly above table • Font? Times Roman10 point bold • Scanned Image 300 dpi • Image format 300 dpi JPEG References: • Where? End of paper • Font Times Roman 10 point • References numbers: In square brackets [ ] • Style As shown in examples Biography: Brief biography and photo of each author. Electronic copy: • Submit electronic copies both for the review and the final paper. No hard copies. • Remove passwords from paper. • Scan for viruses. Document properties: Put your paper’s title and author name in Word “Properties” FAQ on formatting see: www.aeroconf.org/2008_web/Frequently%20Asked%20Questions.doc Filled-out, signed IEEE copyright form must be submitted, by November 1, 2009. Submission is electronic, submit at the same link used to submit papers.