

QuakeSim: Efficient Modeling of Sensor Web Data in a Web Services Environment

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

Geoffrey Fox and Marlon Pierce
Community Grids Laboratory, Indiana University
501 N. Morton, Suite 224
Bloomington, IN 47404

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

Lisa Grant
University of California, Irvine
Irvine, CA 92697

Walter Brooks
NASA Ames Research Center
Moffett Field, CA 94035

Abstract—QuakeSim is a project to develop a modeling environment for studying earthquake processes using a web services environment. In order to model interseismic processes multiple data types must be ingested including spaceborne GPS and InSAR data, geological fault data, and seismicity data. QuakeSim federates data from these multiple sources and integrates the databases with modeling applications. Because the models are complex and compute intensive we are using the Columbia computer located at NASA Ames to integrate and run software programs to improve our understanding of the solid Earth and earthquake processes. The complementary software programs are used to simulate interacting earthquake fault systems, model nucleation and slip on faults, and calculate run-up and inundation from tsunamis generated by offshore earthquakes. QuakeSim also applies pattern recognition techniques to real and simulated data to elucidate subtle features in the processes.^{1,2}

TABLE OF CONTENTS

1. INTRODUCTION	1
2. DISTRIBUTED COMPUTING INFRASTRUCTURE	3
3. QUAKE TABLES DATABASE.....	4
4. APPLICATIONS	6
5. CONCLUSIONS.....	9
6. ACKNOWLEDGEMENTS.....	9
REFERENCES.....	9
BIOGRAPHIES.....	10

1. INTRODUCTION

We are expanding the development of our QuakeSim Web Services environment to integrate both real-time and archival sensor data with high-performance computing applications for data mining and assimilation. The goal of this work is to substantially improve earthquake forecasts, which will ultimately lead to mitigation of damage from this natural hazard. We are federating sensor data sources, with a focus on InSAR and GPS data, for an improved modeling environment for forecasting earthquakes.

¹ 1-4244-1488-1/08/\$25.00 ©2008 IEEE.

² IEEEAC paper #1224, Version 5, Updated November 29, 2007



Figure 1. Example of the QuakeSim portal. Users can access the QuakeTables database and ingest the data into various modeling and visualization applications.

Improved earthquake forecasting is dependent on measurement of surface deformation as well as analysis of geological and seismological data. Space-borne technologies, in the form of continuous GPS networks and InSAR satellites, are the key contributors to measuring surface deformation. These disparate measurements form a complex sensor web in which data must be integrated into comprehensive multi-scale models. In order to account for the complexity of modeled fault systems, investigations must be carried out on high-performance computers.

The QuakeSim applications GeoFEST and Virtual California are the key applications for modeling fault systems. GeoFEST and its supporting software provide a suite of tools for creating and refining finite element meshes that can be used to calculate very detailed seismic deformations associated with individual faults and small fault systems. At a large scale, Virtual California has been developed to perform simulations of large interacting fault systems (i.e. the entire western U.S.) over hundreds of years. Both applications have been ported to and are being optimized for running in parallel on Columbia. Both applications are key interpreting surface deformation data from planned InSAR missions for understanding underlying fault properties and interaction.

Work here will lay the groundwork for NASA's planned Dynamics, Ecosystem Structure, and Dynamics of Ice (DESDynI) mission, with a target launch date of 2014. The mission will produce a minimum of 650 GB of data per day. If data downlink bandwidth limitations are overcome raw

data production will be greater than 1 TB/day. The sheer volumes of data will require routine automated data processing on supercomputers. Data and products must be transported to and from the supercomputing resources and distributed for further processing and analysis. QuakeSim, is intended to establish infrastructure for the upcoming DESDynI mission as well as other potential missions.

We are building upon our Grid of Grids approach, which includes the development of extensive Geographical Information System based Data Grid services. We are extending our earlier approach of integrating the Data Grid components with improved Execution Grid services to interact with high-end computing resources. Our first targets for deploying these services are the Columbia computer at NASA Ames and the Cosmos computer cluster at JPL.

One of the key issues in this project is the utilization and representation of streaming InSAR and GPS data. A federated ontology is being developed to semantically represent GPS data and manage its availability via real-time streaming data services. As part of the development process, multiple GPS data representations and data access services are currently being examined. An Application Programming Interface (API) is used to interface users with the data represented by the ontology. This API is part of the QuakeSim portal. Scientists are able to perform queries at different levels of abstraction and run simulations of the data obtained from QuakeSim. The communication between QuakeTables and the federated ontology is managed through a middleware API.

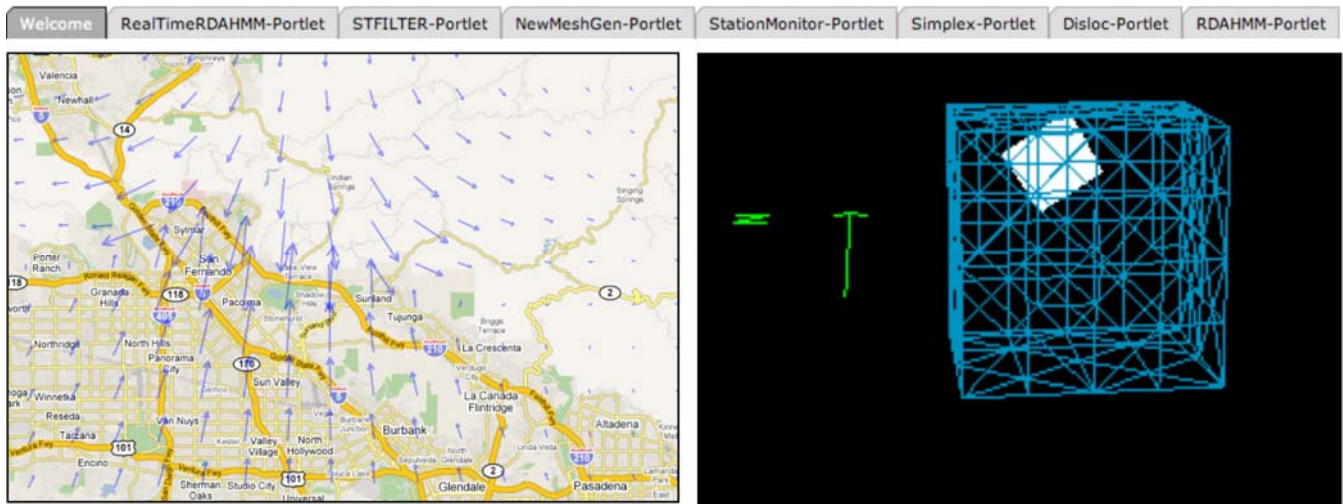


Figure 2. Examples from the portal version 2.0 showing improved richness of interfaces, including the available QuakeSim applications, Google Maps output of surface deformation from the Disloc application, and an example mesh with embedded fault generated from the portlet.

2. DISTRIBUTED COMPUTING INFRASTRUCTURE

QuakeSim's current distributed computing infrastructure consists of Web services interacting with a clients in a component-based Web portal (Figure 1). The Web services provide access to data (particularly fault models) and application codes through well-defined programming interfaces (expressed in WSDL). The QuakeSim portal is a graphical user interface that provides the following capabilities:

- (1) Allows the user to couple databases with simulation codes. This is typically done in the input file creation process, in which users select desired fault models that will be used in the simulation from our fault database. GPS data sources are also supported.
- (2) Assists users with setting up the complicated input files used by the codes.
- (3) Allows the user to track the progress of running jobs.
- (4) Allows the user to do simple plotting to inspect results, such as finite element meshes and calculated surface stresses.
- (5) Allows the user to create and manage archives of jobs by storing metadata (all parameters used, times submitted, simple text descriptions) generated by the user's interactions with the portal. This allows the user to know exactly how a particular results was

obtained and to quickly modify and resubmit it if desired.

- (6) Allows the user to download output files created by a particular run.

We recently released version 2.0 of the QuakeSim portal. We use the JSR 168 portlet-compliant GridSphere container, which is a popular product in the science gateway community. GridSphere enables developers to quickly develop and package third-party portlet web applications that can be run and administered within the GridSphere portlet container. It is used by the Open Grid Computing Environment Project, the Scripps GPS Explorer portal, and many TeraGrid Science Gateways. All portlets are developed using Java Server Faces (JSF). QuakeSim 2.0 has improved the richness of interfaces (Figure 2). We make use of Google Maps, YUI JavaScript Libraries, and BFO Plotting libraries for meshes.

The service architecture we have adopted is useful for running relatively small simulation problems but will need major enhancements to interact securely with the batch schedulers used by Columbia and other high-end supercomputers. Rather than developing this from scratch, we are integrating our approach with the Globus toolkit and services. The classic grid is Globus used by the National Science Foundation TerraGrid, and Open Science Grid (<http://www.globus.org>). The Globus Toolkit is an open source software toolkit for building Grid systems and applications. The "Grid" allows people to share computing power, applications, and databases across boundaries without sacrificing local autonomy. Globus provides the following relevant to QuakeSim:

InterpId	Author1	Publication	Title	Fault_id	FaultName	SegmentName	SegmentId
26	Wald, D. J.	Bulletin of the Seismological Society of America	The Slip History of the 1994 Northridge, California, Earthquake Determined from Strong-Motion, Teleseismic, GPS, and Leveling Data	75	Northridge2	Northridge2	1

Figure 3. Example output on a search from the QuakeTables database.

- A secure remote execution and job management service (GRAM) that has bindings to several queuing systems (PBS, LSF, LoadLeveler, etc);
- Remote file management and file transfer (GridFTP);
- Information services (MDS);
- A single sign-on security environment (GSI) that enables limited delegation (useful, for example in GridFTP third-party file transfers); and
- A client programming API (the Java COG Kit) for its services. The Java COG has been used by the QuakeSim portal and related projects to provide access to the NSF TeraGrid.

Adopting Globus will provide several important features missing from the current command-line based system. The COG provides a rich client development environment that allows us to build graphical user interfaces as well as command line tools. The COG also supports the creation of graph-based workflows for chaining together several operations. The GRAM service supports multiple scheduling/queuing systems and provides an API for programmatically creating batch scripts that is independent of the queuing system. GridFTP supports third-party transfers in addition to uploads and downloads. This allows us to directly transfer files between two backend computers from a portal server. Globus provides optional information services (MDS) that can be used to access machine information that can be displayed to the user or used internally to assist with job submission decisions. Globus services can be used to set up cross-realm authentication. For example, services running at NASA JPL can be configured to accept user credentials signed by the NASA ARC Certificate Authority. We can take advantage of numerous external projects (such as Kepler and Condor-G) for workflow composition and high throughput computing.

3. QUAKE TABLES DATABASE

The QuakeTables database is part of the QuakeSim environment (Figure 3). Currently, QuakeTables houses paleoseismic and fault data that can be ingested into QuakeSim applications. We are expanding the database to include GPS velocities and interferograms processed from Synthetic Aperture Radar data. It is a challenge to convert data, particularly those collected and reported by a variety of means into standard data for modeling applications.

In QuakeSim applications we model fault activity such as rate of strain accumulation or offset related to earthquakes over a finite fault segment. Therefore, the modeler is interested in the general fault characteristics, such as geometry and average rate of slip with an associated uncertainty. Paleoseismic data and results are typically reported in scientific publications and there is no standard format or method for this reporting. Typically a geologist digs a trench across a fault and looks for disrupted layers and carbon samples within these disrupted layers. The samples are carbon dated, and ultimately the geologist publishes a paper with information on a particular earthquake rupture, or sequence of ruptures for a single point on a fault. Alternatively, there may be measurements or models to estimate fault parameters as a result of the occurrence of an earthquake. In order to ingest this information into a model then, judgment must be exercised as to how to extrapolate this information along the length of a fault segment. We have expended considerable effort in combing through the literature and other existing databases, online or off, to include as much information as possible about the faults in California in the QuakeTables database.

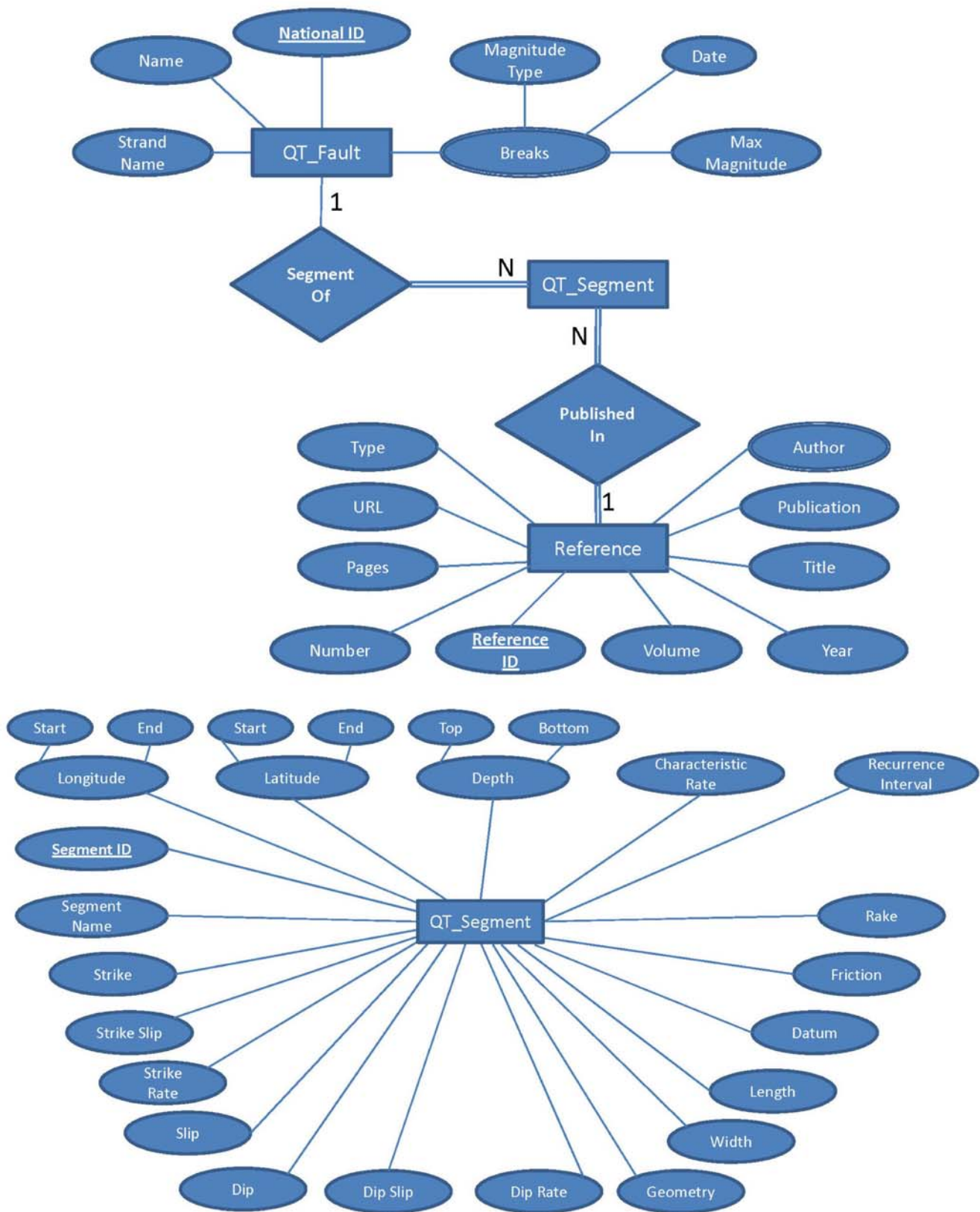


Figure 4. Extended Entity Relationship (EER) for the QuakeTables fault database. Consistency in the data is maintained and uncertainties are allowed. The representation includes data items and types, and geophysical definitions (e.g. width is in km).

For many faults there are multiple interpretations. The purpose of QuakeTables is to standardize data for modelers and allow the modeler to further refine interpretations about faults. As such, then, QuakeTables does not house one single, self-consistent, fault model for California. Rather, it houses the many different interpretations, which can be many even for a single earthquake. It is therefore important for the user to be able to access a self-consistent set of faults for their model and to be able to trace the fault segment recorded in the database back to the original reference.

Another issue is that different applications may use parameters that are reported in different ways. For example, slip on a fault can be reported in Cartesian or polar coordinates. As a result, we have also created mathematical relationships between fault data items to ensure the consistency and semantic integrity of the data. The QuakeTables fault database also includes entries for including uncertainties on the data. (Figure 4). The current design of QuakeTables allows for rectangular faults, which is consistent with the modeling applications.

One important requirement for the new QuakeTables design is its capability to store data from different data sources and keep it in its original format along with any calculated or derived datasets based on this original set. This feature was implemented using two different dataset representations within QuakeTables. The first is DataSet, which are the original datasets by authors in their own format. These sets are stored in dynamic tables to preserve their original format. This type of dataset could also be snapshots of specific data that people want to preserve in a specific format. For example, we find that in carrying out pattern recognition of seismicity, the seismic catalog is occasionally updated and earthquakes are inserted, removed, or their magnitude or location is changed. The previous catalogue is no longer available, and these changes can impact our results. Hence we want to store all versions of the "standard" seismic catalogue. The other data representation is QTSet, which is a dataset that is derived from 'DataSet' and conforms to the QuakeTables format that is used by simulation programs. Each QTSet is linked to its original DataSet, and a DataSet could have multiple QTSets. Since DataSets are originally public domain, QTSets could be set to public or private to users or groups of users.

4. APPLICATIONS

Our QuakeSim applications include traditional high performance software as well as data analysis and assimilation codes. The high-performance modeling applications include GeoFEST [1], a finite element model that simulates stresses associated with earthquake faults, Virtual California [2], which simulates large, interacting fault systems, and PARK [3], which simulates complete earthquake cycles and earthquake interaction. The portal also contains Disloc, which models surface deformation from faults within an elastic half-space, and Simplex, which is an

inversion application, which finds the optical dislocation model of fault slip from GPS and InSAR deformation data [4]. Analysis methods include Pattern Informatics [5], which examines seismic archives to forecast geographic regions of future high probability for intense earthquakes, and RDAHMM [6], a time series analysis application that can be used to determine state changes in instrument signals (such as generated by Global Positioning System arrays). The portal also has a mesh generation tool and tool to filter GPS time series data. We expand on some of the applications here.

Virtual California

Virtual California (VC) is a numerical simulation program for studying the system-level dynamics of the vertical strike-slip fault configuration in California [7,8]. The majority of plate boundary deformation in California is accommodated by slip (i.e. earthquakes) on the strike-slip faults included in the Virtual California models (Figure 5).

Virtual California uses topologically realistic networks of independent fault segments that are mediated by elastic interactions. Virtual California is a backslip model, inasmuch as the plate tectonic stress increases are produced by means of applying a negative (backslip) velocity to each segment whose magnitude is that of the long-term rate of slip on the segment. Since "positive slip" reduces the stress on a fault segment, "negative slip" due to the backslip increases the stress. On each time step, all faults are checked to determine whether the shear stress has reached the failure threshold. Once at least one segment reaches the threshold, the "long time steps" stop, and "short (failure) time steps" (a.k.a. Monte Carlo Sweeps, or mcs) begin. An mcs begins with a check of each site to determine whether it has failed, followed by a parallel updating of each segment. An update of a segment consists of increasing the sudden seismic slip on each segment so that the stress of the segment, considered in isolation, drops to a residual value, plus or minus a random overshoot/undershoot. The elastic stress on all segments is then recalculated, and another mcs is carried out. This iterative process repeats until all segments are below the failure threshold, at which time the mcs time steps cease and the long plate tectonic time steps begin again.

Virtual California also includes a stress-dependent precursory slip, or stress leakage of the type that has been observed in laboratory experiments by [9] and [10]. The physics of this process is that as the stress on a segment increases, a small amount of stable sliding occurs that is proportional to the level of the stress above the residual. Lab experiments and field data suggest that the frictional parameter alpha [8] is of the order of a few percent. Alpha is defined as the fraction of aseismic slip relative to total slip. Therefore, it may be possible to detect precursory signals before earthquakes using InSAR data from missions such as DESDynI. Virtual California simulations enable testing for precursory signals. Hence a focus is to analyze the magnitude and spatial distribution any precursory slip in the simulations.

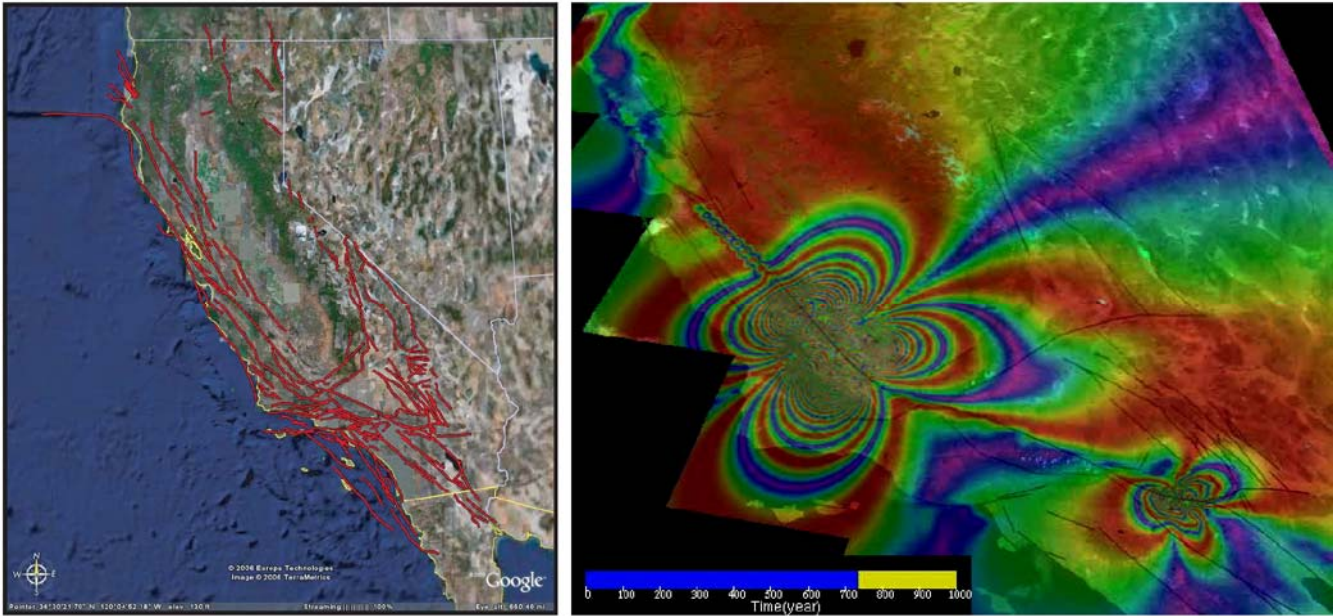


Figure 5. Example fault model used by Virtual California (left panel), which is also included in the QuakeTables database. Output converted to InSAR fringes for a given time step (right panel). The fringes represent surface deformation for an earthquake. In this time step two earthquakes have occurred.

GeoFEST

GeoFEST uses stress-displacement finite elements to model stress and flow in a realistic model of the Earth's crust and upper mantle in complex regions such as southern California, including the Los Angeles Basin. The model includes stress and strain due to the elastic response to an earthquake event in the region of the slipping fault, the time-dependent viscoelastic relaxation, and the net effects from a series of earthquakes. The physical domain may be two- or three-dimensional and may contain heterogeneous materials and an arbitrary network of faults. Finite element modeling in three dimensions allows faithful modeling of complex faulting geometry, inhomogeneous materials, realistic viscous flow, and a wide variety of fault slip models and boundary conditions. Because finite elements conform to (nearly) any surface geometry and support wide variations in mesh density, solutions may be made arbitrarily accurate with high computational efficiency.

GeoFEST runs in the high-performance domain of message-passing parallel computer systems [11] including the Columbia system at NASA Ames and the COSMOS system at JPL, among others. It includes the functions of the PYRAMID parallel adaptive mesh refinement library [12]. Source code is available with a no-fee license from Open Channel and it runs within the QuakeSim web-based problem-solving environment [13]. All documentation and links to Open Channel and the portal can be found at <http://quakesim.org>.

The primary quantity computed by GeoFEST is the displacement at each point in a domain. The stress tensor is

also computed as a necessary byproduct. The computational domain represents a region of the earth's crust and possibly underlying mantle. It is typically a square or rectangular domain in map view, with a flat upper free surface and constant depth, but the domain may deviate from this. The only requirement is that it be a bounded 3D domain with appropriate surface boundary conditions to render the problem well defined. These boundary conditions may be specified as surface tractions and/or displacements, which are usually specified on all surfaces and at times on interior surfaces such as faults. Free surfaces have zero surface traction by definition. Faults are interior surfaces, and may have associated dislocation increments at set times. The solid domain may contain layers or other distributions of material with associated rheological properties.

Currently supported materials are isotropic, Newtonian elastic, Newtonian viscoelastic, and non-Newtonian power-law viscosity. Elastostatic solutions are supported, such as computing the displacements and stresses immediately caused by a specified slip distribution on a fault or finding the interior displacement and stress distribution due to a surface traction or displacement. These solutions are not time-dependent. Viscoelastic solutions, which are time dependent, are also supported, in which the material flows and relaxes in response to imposed stress, such as an earthquake event. One may compute the viscoelastic response to a single event, or to multiple events in a sequence. The sequence may be user-specified. Location-specific body forces are supported.

Click on a station symbol for more information.

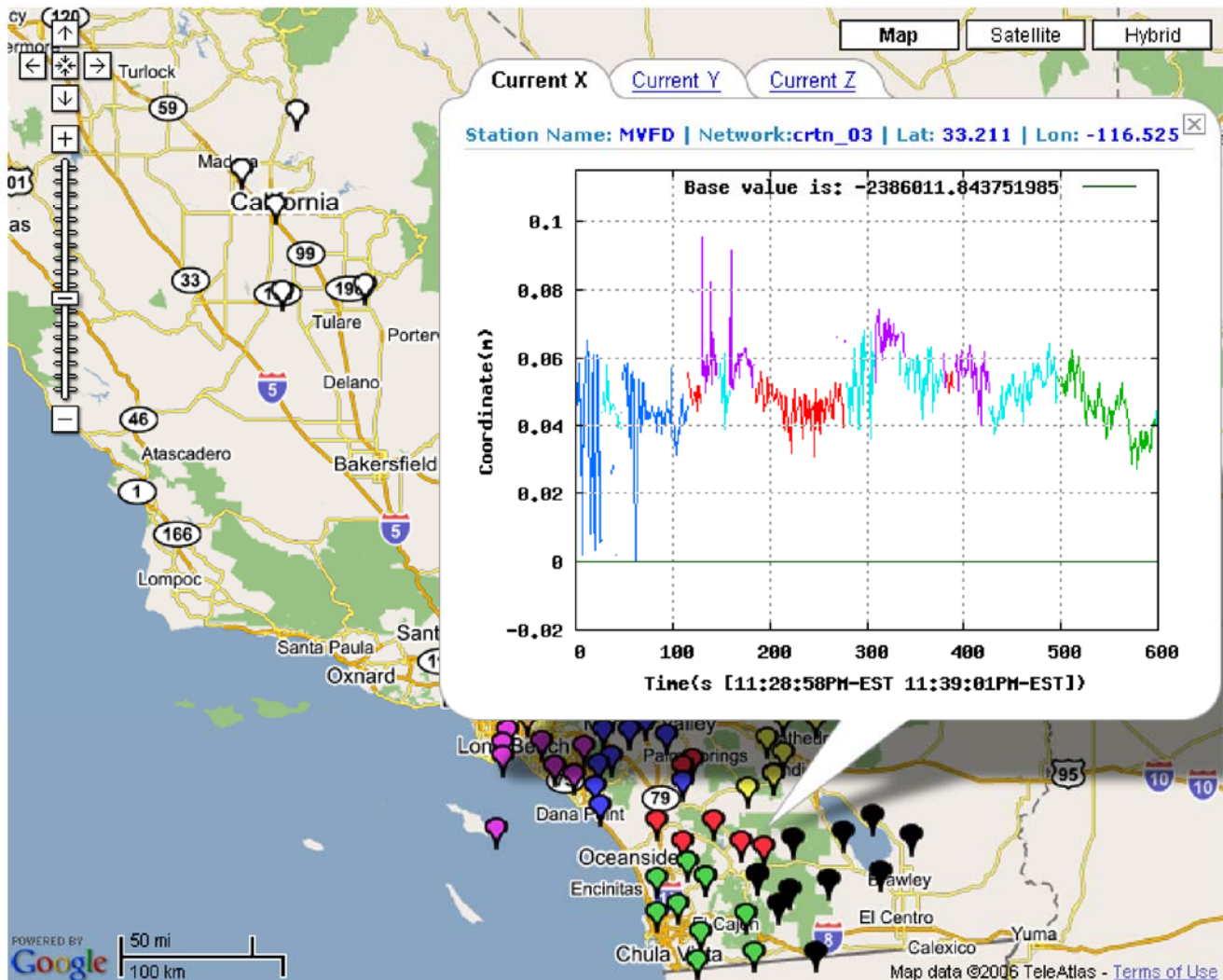


Figure 6. Google maps interface through the QuakeSim portal showing classified GPS time series data using the RDAHMM services.

Boundary conditions and solutions apply to a finite-element discretized approximation to this domain. The domain is defined internally as a mesh of space-filling tetrahedral or hexahedral elements, with three components of displacement at each mesh node constituting the solution. Stress is computed for each element, and is element-wise constant for the current linear tetrahedral element type. Surface nodes carry special boundary conditions such as tractions or specified displacements. Nodes on faults are special split-nodes that define screw or tensile dislocation on the fault without perturbing the mesh geometry. Temporal evolution is by discrete time steps using an implicit solution technique, allowing large time steps without numerical instability.

RDAHMM

RDAHMM, or Regularized Deterministic Hidden Markov Model, carries out time series analysis and mode detection in GPS and other signals. Examples of signals that

RDAHMM can detect are ground subsidence from withdrawal of water from aquifers and earthquake co-seismic and post-seismic signals.

We have integrated the processing of GPS position time series data into the QuakeSim portal. By wrapping the RDAHMM time series analysis software as a web service filter, it is seamlessly integrated into work and data processing flows. Raw GPS data (1Hz) are converted to RYO (real-time) format and made available through a data server. Then data are passed through a series of filters that perform format conversion and station separation. Message passing is handled through NaradaBrokering. Finally, data are passed to the RDAHMM analysis application.

We have implemented an interface through which the RDAHMM software can be applied to archived daily GPS solutions to perform time series segmentation. Segmentation results are provided both graphically and

through numerical descriptions of segmentations and fitted models, which are available for download. In addition, we have implemented a proof-of-concept Google maps interface to RDAHMM analysis of real-time streaming GPS data. The segmentation analysis is performed on the last ten minutes of real-time data, and then displayed graphically upon mouse-over in the Google maps interface (Figure 5).

5. CONCLUSIONS

Through the QuakeSim project we are developing tools for carrying out comprehensive simulations of earthquake fault interactions. Our focus to date has been on California but the tools are applicable to any region in the world and through collaborations with the University of Queensland, Australia we will be extending the database and models to cover that regions of the world. We have recently released the QuakeSim 2.0 portal, based on GridSphere, the popular science gateway product. We are working to extend our distributed computing environment by interfacing with the Columbia computer at NASA Ames, and the Cosmos computer at JPL. Our QuakeTables fault database is designed to allow modelers to make use of paleoseismic data and the associated uncertainties. We have expanded QuakeSim to access GPS time series data and are developing the ontologies to access GPS velocity data, and synthetic aperture radar interferograms (InSAR data).

Current earthquake risk estimation is based on static models inferred from past earthquake activity as determined through paleoseismology and historical earthquakes. Earthquake fault systems are continuously changing state based on deformation of the Earth's crust and mantle as well as strain release and transfer from earthquakes. It is therefore important to develop time-dependent models for earthquake forecasting. Current earthquake hazard maps have an outlook of decades [14,15]. Our goal is to improve earthquake forecasting by, in effect, migrating from static hazard maps to dynamically changing earthquake forecasts based on the current state of the system.

6. ACKNOWLEDGEMENTS

We'd like to thank the many other QuakeSim contributors. These include JPL staff members Charles Norton for the adaptive mesh refinement and porting and optimization of software to the Columbia and Cosmos systems, and Margaret Glasscoe for testing the portal using science applications and developing the web pages, and Harout Nazerian for developing the web pages and testing the portal. Galip Aydin, John Youl Choi, and Zhigang Qi, at Indiana University have also contributed to development of the QuakeSim portal. Gleb Morein at UC Davis has worked on the development of Virtual California including porting to the Columbia computer system. Lorena Medina corrected and validated the QuakeTables database at UC Irvine. Terry Tullis at Brown University and Nick Beeler from the USGS develop and maintain the PARK application. This work was

carried out at the Jet Propulsion Laboratory, California Institute of Technology under contract with NASA from the Earth Science and Technology Office, and at University of Southern California, Indiana University, NASA Ames, University of California Davis, and University of California Irvine, and Brown University under subcontract with the Jet Propulsion Laboratory.

REFERENCES

- [1] Parker, Jay, Andrea Donnellan, Gregory Lyzenga, John B. Rundle, Terry Tullis: Performance Modeling Codes for the QuakeSim Problem Solving Environment. International Conference on Computational Science: 855-862, 2003.
- [2] Rundle, J.B., K.F. Tiampo, W. Klein and J.S.S. Martins, Self-organization in leaky threshold systems: The influence of near mean field dynamics and its implications for earthquakes, neurobiology and forecasting, *Proc. Nat. Acad. Sci. USA*, **99**, Supplement 1, 2514-2521, (2002).
- [3] Tullis, T. E., Earthquake models using rate and state friction and fast multipoles, *Geophysical Research Abstracts (European Geophysical Society)*, **5**, 13211, 2003.
- [4] Lyzenga, G.A., W.R. Panero, A. Donnellan, The Influence of Anelastic Surface Layers on Postseismic Thrust Fault Deformation, *J. Geophys. Res.*, **105**, 3151-3157, 2000.
- [5] Tiampo, K.F., Rundle, J.B., McGinnis, S., and Klein, W. Pattern dynamics and forecast methods in seismically active regions, *Pure and Applied Geophysics*, **159**, 2002.
- [6] Granat, R. A., *Regularized Deterministic Annealing EM for Hidden Markov Models*, Ph.D. Thesis, University of California, Los Angeles, 2004.
- [7] Rundle JB, Klein W, Tiampo K, et al., Linear pattern dynamics in nonlinear threshold systems, *Phys. Rev. E*, **61**, 2418-2431, 2000.
- [8] Rundle P.B., J.B. Rundle, K.F. Tiampo, et al., Nonlinear network dynamics on earthquake fault systems, *Phys. Rev. Lett.*, **87**, Art. No. 148501, 2001.
- [9] Tullis, T.E., Rock friction and its implications for earthquake prediction examined via models of Parkfield earthquakes, *Proc. Nat. Acad. Sci.*, **93**, 3803-3810, 1996.
- [10] Karner SL, Marone C, Frictional restrengthening in simulated fault gouge: Effect of shear load perturbations, *J. Geophys. Res.*, **106**, 19319-19337, 2001.

- [11] Parker J., Donnellan A., Lyzenga G., et al., Performance modeling codes for the QuakeSim problem solving environment, *Lect. Notes In Computer Sci.*, 2659: 855-862 2003.
- [12] Norton, C., G. Lyzenga, J. Parker, E. Tisdale, Developing Parallel Active Tectonics Simulations Using GeoFEST and the PYRAMID Adaptive Mesh Refinement Library, *Eos Trans. AGU*, 85, Fall Meet. Suppl., Abstract SF43A-0785, 2004.
- [13] Pierce M, Youn C, Fox G, Interacting data services for distributed earthquake modeling, *Lect. Notes In Computer Sci.*, 2659: 863-872 2003.
- [14] Blanpied, M., Working Group on California Earthquake Probabilities, Earthquake Probabilities in the San Francisco Bay Region: 2002–2031, *U.S. Geol. Survey Open File Report*, 30-214, 2003.
- [15] Petersen, M.D., W.A. Bryant, C.H. Cramer, T. Cao, M. Reichle, A.D. Frankel, J.J. Lienkaemper, P.A. McCrory, and D.P. Schwartz, Probabilistic Seismic Hazard Assessment of for the State of California, *US. Geol. Survey Open File Report*, 96-706, 1996, revised 2003.

BIOGRAPHIES



Andrea Donnellan is the DESDynI Science Lead and QuakeSim principal investigator at NASA's Jet Propulsion Laboratory and is a research professor at the University of Southern California. Donnellan uses GPS and InSAR satellite technology coupled with high performance computer models to study earthquakes, plate tectonics, and the corresponding movements of the earth's crust. She has been a geophysicist at JPL since 1993. She received a B.S. from the Ohio State University in 1986, with a geology major and mathematics minor. She received her M.S. and Ph.D. in geophysics from Caltech's Seismological Laboratory in 1988 and 1991 respectively. Donnellan received an M.S. in Computer Science from the University of Southern California in 2003. She held a National Research Council Postdoctoral Fellowship at NASA Goddard Space Flight Center. Donnellan was a Visiting Associate at the Seismological Laboratory at Caltech from 1995 to 1996. In 1996 Donnellan received the Presidential Early Career Award for Scientists and Engineers, in 2003 the Women in Aerospace Award for Outstanding Achievement, and in 2006 she was the MUSES of the California Science Center Foundation Woman of the Year.



Jay Parker is a Senior Scientist in the Geodynamics and Space Geodesy group of the Jet Propulsion Laboratory, a NASA center administered by the California Institute of Technology. His graduate research used computer simulations to explain mesospheric ionization response to solar flares and dynamic instabilities. Dr Parker's research subjects at the Jet Propulsion Laboratory include a variety of topics in remote sensing analysis and modeling. These include supercomputing algorithms for electromagnetic scattering and radiation, satellite geodesy and finite element simulation for earthquake-related deformation, and ocean sensing through GPS signal reflection. He is currently the software engineer and co-investigator for the QuakeSim project, which has developed a solid Earth science framework including a variety of simulation and analysis tools. He also develops the SEASCRAPE software system for high-fidelity simulation and parametric retrieval of atmospheric infrared spectrometry at Remote Sensing Analysis Systems, Inc. of Altadena CA. Dr. Parker is a member of the American Geophysical Union and co-chair of the Data Understanding and Assimilation working group of the APEC Cooperation for Earthquake Simulation.



Robert Granat is a senior member of the technical staff in the Data Understanding Systems Group at NASA's Jet Propulsion Laboratory. He received a B.S. in engineering and applied science from Caltech in 1996, an M.S. in electrical engineering in 1998 from University of California, Los Angeles, and a Ph.D. in electrical engineering with an emphasis in signal processing from UCLA. He works on the application of hidden Markov models to the study earthquakes and spacecraft systems.



Geoffrey Fox received a Ph.D. in Theoretical Physics from Cambridge University and is now professor of Computer Science, Informatics, and Physics at Indiana University where he is director of the Community Grids Laboratory. He previously held positions at Caltech, Syracuse University and Florida State University. He has published over 550 papers in physics and computer science and currently works in applying computer science to Defense, Earthquake and Ice-sheet Science and Chemical Informatics. He is involved in several projects to enhance the capabilities of Minority Serving Institutions.



Marlon Pierce is assistant director of the Community Grids Laboratory and leads the portal and services development for the QuakeSim project. He has a Ph. D. in computational condensed matter physics from Florida State University (1998). His research interest is the application of distributed computing techniques and technologies to problems in physical sciences.



John Rundle is the Director of the Computational Science and Engineering Center at the University of California at Davis, and is Director of the California Institute for Hazard Research of the University of California. He was educated at Princeton University (BSE, 1972) and the University of California-Los Angeles (MS, 1973, PhD, 1976). His research is focused on understanding the dynamics of earthquakes through numerical simulations; pattern analysis of complex systems; dynamics of driven nonlinear Earth systems; and adaptation in general complex systems. He has published over 200 papers in the peer-reviewed literature.



Dennis McLeod is currently Professor of Computer Science at the University of Southern California, and Director of the Semantic Information Research Laboratory. He received his Ph.D., M.S., and B.S. degrees in Computer Science and Electrical Engineering from MIT. Dr. McLeod has published widely in the areas of data and knowledge base systems, federated databases, database models and design, and ontologies. His current research focuses on dynamic ontologies, user-customized information access, database semantic heterogeneity resolution and interoperation; personalized information management environments; information management environments for geoscience and homeland security information, crisis management decision support systems, and privacy and trust in information systems.



Rami Al-Ghanmi is a PhD student in the Computer Science Department at the University of Southern California. He received his MS in Computer Science from the USC (2006) and BS in Computer Engineering from King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia (2002). He is currently working with Professor Dennis McLeod at the Semantic Information Research Group at USC. His research

interests are Semantic Web Services and Ontology-based Federation of scientific data.



Lisa Grant is an associate professor of public health at University of California, Irvine. She studies environmental problems from a geologic perspective with emphasis on natural hazards. The primary objective of her research program is to identify active faults and to quantify their potential for generating large earthquakes by discovering their history of earthquake production over geologic time intervals. The record of previous large earthquakes on active faults is one of the best indicators of future earthquake activity. The results of her research on earthquake occurrence patterns are applied for earthquake forecasting, land-use planning, building design, risk assessment, disaster preparedness planning, and public education about the earthquake threat. Grant received a B.S. in environmental Earth science from Stanford in 1985, an M.S. in both environmental engineering and science and geology in 1989 and 1990 respectively and a Ph.D. in geology and geophysics from Caltech in 1993. She serves on the board of directors of the Southern California Earthquake Center, and is an associate director of the California Hazards Research Institute.



Walter Brooks has worked at NASA Ames Research center since 1977. As chief of the NASA Advanced Supercomputing (NAS) Division, he oversaw the full range of high-performance computing efforts within the division. Brooks has led groups that simulated and designed space science missions such as Infrared Astronomical Satellite (IRAS), Space InfraRed Telescope Facility (SIRTF) and Stratospheric Observatory for Infrared Astronomy (SOFIA). In 1993 he was selected to lead the Space Station redesign management team. Awarded the NASA Outstanding Leadership medal for his work in the area of high-end computing, Dr. Brooks has played an important role in the high-end computing aspects of the Return-to-Flight effort. He was a member of the High End Computing Revitalization Task Force that produced an important report in 2003 that outlines a set of the organization's key findings and recommendations to advance the state of high-end computing in the U.S. He is past president of the SGI user group, a Sloan Fellow, an AIAA Associate Fellow, and author of more than 50 publications. Brooks received a doctorate in physics from Stevens Institute of Technology, performing research for his thesis at Brookhaven National Laboratory. He earned a master's degree from the Stanford Graduate School of Business in 1991.