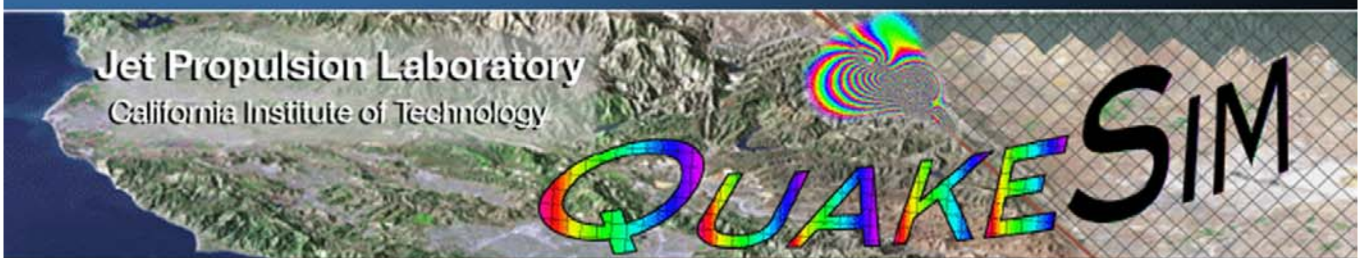
2003 Annual Report



Numerical Simulations For Active Tectonic Processes: Increasing Interoperability And Performance

Andrea Donnellan
Principal Investigator

2003 Annual Report Completed

Numerical Simulations For Active Tectonic Processes: Increasing Interoperability And Performance

The QuakeSim Project

JPL Task Order: 10650

Milestone C – Administration due date: 8/30/2003

Second Annual Report Delivered.

Executive Summary

In order to develop an active tectonics simulation and analysis environment for the geophysics community, the QuakeSim project made rapid progress in 2003 in two major areas: the development of the QuakeSim Portal, a web-based environment for using data and interoperating tools, and high-end computing implementation of two major simulation tools. These tools are PARK, a simulator for earthquake initiation and GeoFEST, which computes permanent deformation from earthquakes in realistic settings.

First, we demonstrated a web-services problem-solving environment prototype that links together diverse earthquake science applications on distributed computers. We can now use the prototype to build a model with faults and layers from the fault database, automatically generate a finite element mesh, solve for crustal deformation and produce a full color animation of the result integrated with remote sensing data. This portal environment is rapidly expanding to include many more applications and tools.

We then made major performance improvements to the simulation codes. The capabilities of the PARK code have increased from 15,000 fault elements and 500 time steps in just under eight hours (one processor) to 150,000 fault elements and 5,000 time steps in the same amount of time on 256 processors. This was demonstrated on the Chapman computer at Ames. Integration with the parallel fast multipole method has resulted in a one hundredfold performance improvement in solving problems of this size.

GeoFEST has had similar success. Its demonstrated capabilities have gone from solving a 55,000 finite element problem in just under fourteen hours (one

processor) to 1.4 million elements in 2.8 hours (64 processors on the Thunderhead Linux computer at GSFC). Parallel implementation includes integration with the Pyramid adaptive mesh tools developed by the Computational Technologies (CT) project. GeoFEST has been ported to eight different parallel machines, including Linux, SGI and Apple. Ongoing development with Pyramid will allow use of near optimal meshes for high quality regional simulations.

QuakeSim is working with the CT visualization team to display surface deformation with topography and LandSat images. Work has also begun to show the underground stress field due to earthquakes and tectonic forces.

Team

Andrea Donnellan:
Principal Investigator
Jet Propulsion Laboratory
Mail Stop 183-335
4800 Oak Grove Drive
Pasadena, CA 91109-8099
donnellan@jpl.nasa.gov
818-354-4737

Geoffrey Fox:
Information Architect

Community Grid Computing Laboratory
Indiana University
501 N. Morton, Suite 224
Bloomington, IN 47404-3730
gcf@indiana.edu
812-856-7977

Terry Tullis:
Fast Multipole Methods
Brown University
Box 1846, Brown University
Providence, RI 02912-1846
Terry_Tullis@Brown.edu
401-863-3829

John Rundle:
Pattern Recognizers
Center for Computational Science and Engineering
U. C. Davis
Davis, CA 95616
rundle@geology.ucdavis.edu
530-752-6416

Greg Lyzenga:
Inversion and Assimilation Codes

Jet Propulsion Laboratory
Mail Stop 126-347
4800 Oak Grove Drive
Pasadena, CA 91109-8099
gerg.lyzenga@jpl.nasa.gov
818-354-6920

Jay Parker:
Overall Software Engineer
Jet Propulsion Laboratory
Mail Stop 238-600
4800 Oak Grove Drive
Pasadena, CA 91109-8099
Jay.W.Parker@jpl.nasa.gov
818-354-6790

Marlon Pierce:
Code Interoperability Software Engineer

Community Grid Computing Lab
Indiana University
501 N. Morton, Suite 224
Bloomington, IN 47404-3730
marpierc@indiana.edu
812-856-1212

Dennis McLeod:
Database Interoperability
Professor
Computer Science Department
University of Southern California
Los Angeles, CA 90089-0781
mcleod@usc.edu
213-740-4504

Lisa Grant:
Fault Database Architect
University of California, Irvine
Environmental Analysis and Design
Irvine, CA 92697-7070
lgrant@uci.edu
949-824-5491

Robert Granat:
Pattern Recognizers

Jet Propulsion Laboratory
Mail Stop 126-347
4800 Oak Grove Drive
Pasadena, CA 91109-8099
robert.granat@jpl.nasa.gov
818-393-5353

Objective

Project Objective

The full objective over this three-year program is to produce a system to fully model earthquake-related data. Components of this system include:

- A database system for handling both real and simulated data
- Fully three-dimensional finite element code (FEM) with an adaptive mesh generator capable of running on workstations and supercomputers for carrying out earthquake simulations
- Inversion algorithms and assimilation codes for constraining the models and simulations with data
- A collaborative portal (object broker) for allowing seamless communication between codes, reference models, and data
- Visualization codes for interpretation of data and models
- Pattern recognizers capable of running on workstations and supercomputers for analyzing data and simulations

In order to develop a solid Earth science framework for understanding and studying of active tectonic and earthquake processes, this task develops simulation and analysis tools to study the physics of earthquakes using state-of-the-art modeling, data manipulation, and pattern recognition technologies. We develop clearly defined accessible data formats and code protocols as inputs to the simulations. These are adapted to high-performance computers because the solid Earth system is extremely complex and nonlinear resulting in computationally intensive problems with millions of unknowns. With these tools it will be possible to construct the more complex models and simulations necessary to develop hazard assessment systems critical for reducing future losses from major earthquakes.

Project details and documentation are available at the QuakeSim main web page at http://quakesim.jpl.nasa.gov.

Year 2 Objectives

Our objectives in the second year were twofold: to build a working prototype of the QuakeSim problem solving environment, and to implement and measure performance and scalability of high end computing simulations of earthquakes using the GeoFEST and PARK codes.

The Year 2 goal is a prototype environment that demonstrates the fault database, mesh-generation, GeoFEST simulation and RIVA visualization for solving a typical crustal deformation problem. This was successfully completed and documented in Milestone I *Interoperability Prototype* at http://quakesim.jpl.nasa.gov/milestones.html.

QuakeSim is developing three high-end computing simulation tools: GeoFEST, PARK and Virtual California. The Year 2 objective here is to demonstrate parallel scaling and efficiency for the GeoFEST finite element code and the PARK boundary element code for unstable slip. Both these codes were enhanced by integration with high performance computing tools: GeoFEST with Pyramid adaptive meshing, and PARK with a fast multipole library. This code improvement was achieved and documented in Milestone F *First Code Improvement.*

Approach for Year 2

Interoperability/Portal

Our approach in Year 2 is to build a three-tiered architecture system. The tiers are:

- A portal user interface layer that manages client components.
- A service tier that provides general services (job submission, file transfer, database access, etc.) that can be deployed to multiple host computers.
- Backend resource, including databases and earthquake modeling codes.

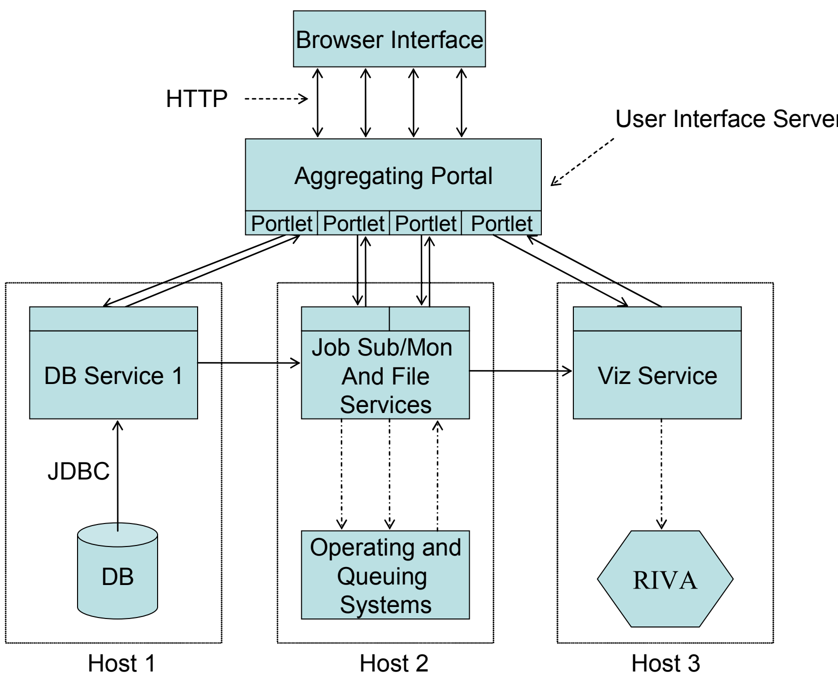


Figure1. Portal and service architecture.

These tiers are illustrated in Figure 1. The user interacts with the system through the Web Browser interface (top). The web browser connects to the aggregating portal, running on the User Interface Server (http://complexity.ucs.indiana.edu in the testbed). The "Aggregating Portal" is so termed because it collects and manages dynamically generated web pages (in JSP) that may be developed independently of the portal and run on separate servers. The components responsible for managing particular web site connections are known as portlets. The aggregating portal can be used to customize the display, control the arrangement of portlet components, manage user accounts, set access control restrictions, etc. See Figure 2 for an example of the user interface.

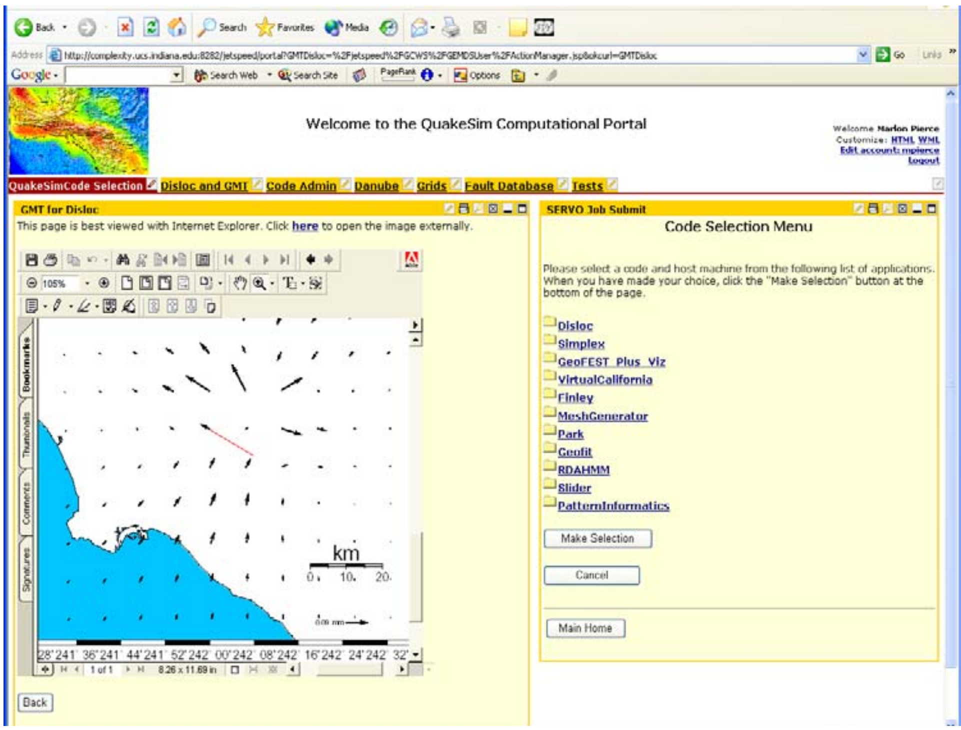


Figure 2. The QuakeSim portal is based on portlet component. This figure shows the browser interface. The screen shot displays two user interfaces (a visualization of Disloc output on the left, and a the code selection interface on the right). These are independent web pages pulled into the aggregating portal. Tabs across the top can be used to navigate to other portlet component displays.

The portlet components are responsible for loading and managing web pages that serve as clients to remotely running Web services, shown in the lower figure. Figure 1 shows a DB service running on host 1 (http://infogroup.usc.edu in the testbed), job submission and file management services in the center (typically running on danube.ucs.indiana.edu in the testbed) and visualization services (such as RIVA, running on the host jabba.jpl.nasa.gov). We use Web Services to describe the remote services and invoke their capabilities. Unless otherwise indicated, solid arrow connections in the figure indicate SOAP over HTTP connections. Broken dash-dot arrows indicate unspecified connections—these are local invocations in our testbed but we are testing ssh/scp connections. We may also use Grid connections (GRAM and GridFTP) to access these applications. Database connections between the Database service and the actual database are handled by JDBC (Java Database Connectivity), a standard technique.

The QuakeSim portal effort has been one of the pioneering efforts in building Computing Portals out of reusable portlet components. Pierce and Fox of the QuakeSim team collaborate with other portal developers following the portlet component approach through the Open Grid Computing Environments consortium (OGCE: Argonne National Laboratory, Indiana University, the University of Michigan, the National Center for Supercomputing Applications, and the Texas Advanced Computing Center). This project has been funded by the NSF National Middleware Initiative (Pierce, PI) to develop general software releases for portals and to end the isolation and custom solutions that have plagued earlier portal efforts. Pierce and Fox's involvement with both the QuakeSim project and the OGCE will ensure that the QuakeSim portal will benefit from the larger community of portal development: we may extend the QuakeSim portal to use capabilities developed by other groups and may share the capabilities developed by the QuakeSim portal with the portal-building community.

Fault Database

Our approach in year 2 was to provide a searchable fault database with annotated earthquake fault records from publications. The database team designed the fields that constituted the database records and provided a web-based interface that enables the submitting and accessing of those records. A small group of geologists/paleoseismologists searched the literature and collected annotated records of Southern California earthquake faults to begin population of the fault database. This is detailed in the Milestone I documentation.

Code Improvement

Two of the most complex codes were targeted for improved performance, chiefly through design changes that make them efficient, high-performance parallel codes. These codes are PARK, a boundary-element based code for studying unstable slip at the Parkfield segment of the San Andreas fault, and GeoFEST, a fully three-dimensional finite element code to model active tectonics and earthquake processes. Together with an adaptive mesh generator that constructs a mesh based on geometric and mechanical properties of the crustal structure, the GeoFEST system makes it possible to efficiently model time-dependent deformation of interacting fault systems embedded in a heterogeneous Earth structure.

GeoFEST and PARK were re-written with standard Message Passing Interface (MPI) with careful attention to low communication overhead and load balance among processors. In addition, the PARK code has been re-written to use a fast multipole library for enhanced performance and GeoFEST has been modified to use the Pyramid adaptive mesh refinement library. These improvements were tested on increasingly large problems on NASA computing testbeds to ensure efficiency in scaling.

More details on performance and scaling may be found in the Technology Accomplishments section below, and in the Milestone F documentation at http://quakesim.jpl.nasa.gov/milestones.html.

Scientific Accomplishments

On-Demand Science Capability Demonstration – San Simeon Earthquake

Simplex, which is implemented in the QuakeSim portal, was used to investigate the geometry of the fault ruptured by the December 22, 2003, magnitude 6.5, San Simeon earthquake. The Simplex portal tool enables scientists to rapidly find fault parameters based on surface deformation and a nonlinear least-squares optimal Okada solution in an elastic half space. Determination of the elastic rupture geometry is necessary to investigate immediate time-dependent deformation processes. Additionally, the rupture geometry is important for determining stress transfer to adjacent faults and understanding triggered seismicity. In future use, rapidly available information about the rupture parameters could be used to optimally deploy campaign GPS receivers and select InSAR mission targets, as well as assess increased seismic hazard.

The preliminary report for the San Simeon earthquake provided details about the earthquake[1]. The report was available December 24, 2003. The San Simeon main shock was located 11 km NE of San Simeon and caused damage and casualties in Paso Robles, 39 km to the ESE. The focal mechanism derived from regional seismic networks is nearly pure thrust motion on a fault striking ESE-WNW. The estimated dip slip is 50 cm. The main shock was followed by more than 897 aftershocks greater than magnitude 1.8 in the first 48 hours. The aftershock region extends 30 km in length, with a surface projection width of 10 km. The aftershock depths rarely exceed 8 km.

The epicenter of the San Simeon earthquake is located in a region with many mapped active faults, however the fault that ruptured had not been determined at the time of the preliminary report. Additionally, no surface rupture had been identified. The preliminary report described two possible fault planes, one

[1] http://www.cisn.org/special/evt.03.12.22/prelimreport.html

dipping NE and the other dipping SW. The aftershock distribution and finite-fault models suggest a NE dipping fault.

Using Simplex, the two possible fault dip directions were tested. Co-seismic displacements for 12 local stations of the Southern California Integrated GPS Network (SCIGN) provided solution constraints. Professor Tom Herring of MIT, Chairman of SCIGN, provided these preliminary displacement estimates. Initial fault guesses were located near the epicenter and parameters consistent with the reported magnitude and focal mechanism were chosen. For both the SW and NE dipping fault initial guesses, the fault orientation and geometry were allowed to vary. The SW dipping fault consistently reached lower residual values than the NE dipping fault for tests with modifications to the parameters in the input file. The residual fit to the best solution is shown in Figure **3** below. This fault dips 46 degrees and strikes approximately ESE. The fault reaches a depth of 8 km and does not break the surface. The dip slip magnitude in the minimized Simplex result is 50 cm. Due to the lack of sensors close to the fault, and absence of sensors SE of the epicenter, it is difficult to accurately constrain the fault location.

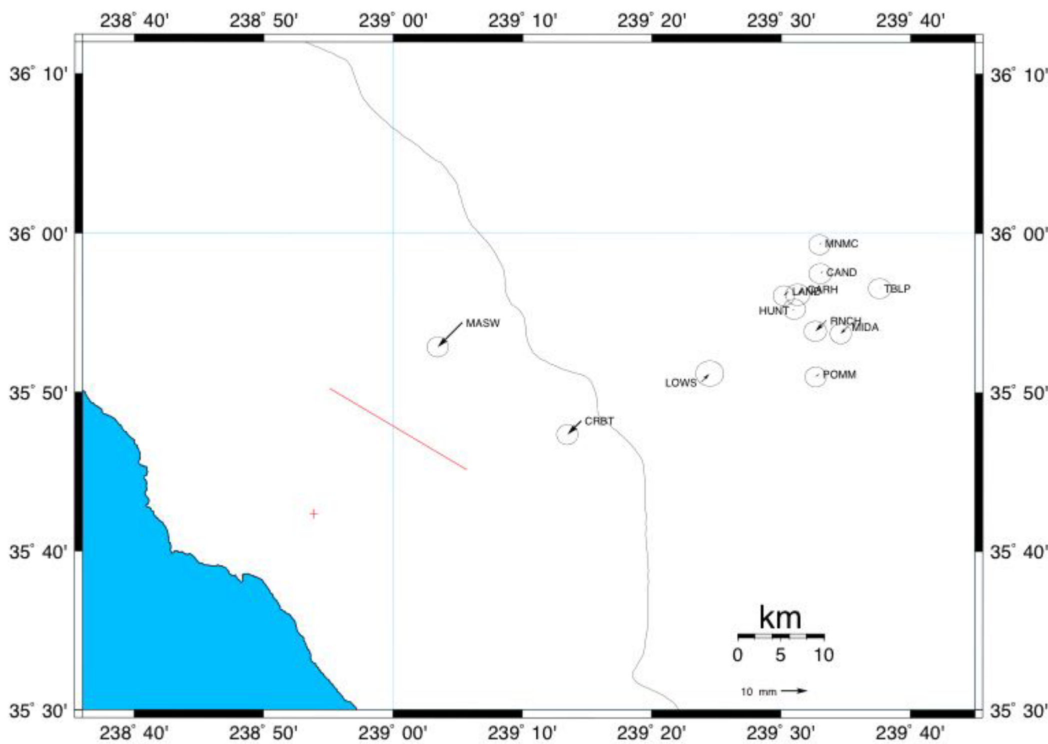


Figure 3. Residual displacements calculated by Simplex for the fault best-fit to San Simeon co-seismic GPS displacements. Vectors show the displacement residuals at the contributing stations, the red line is the surface trace of the best-fit fault, and the red plus marks the main shock epicenter.

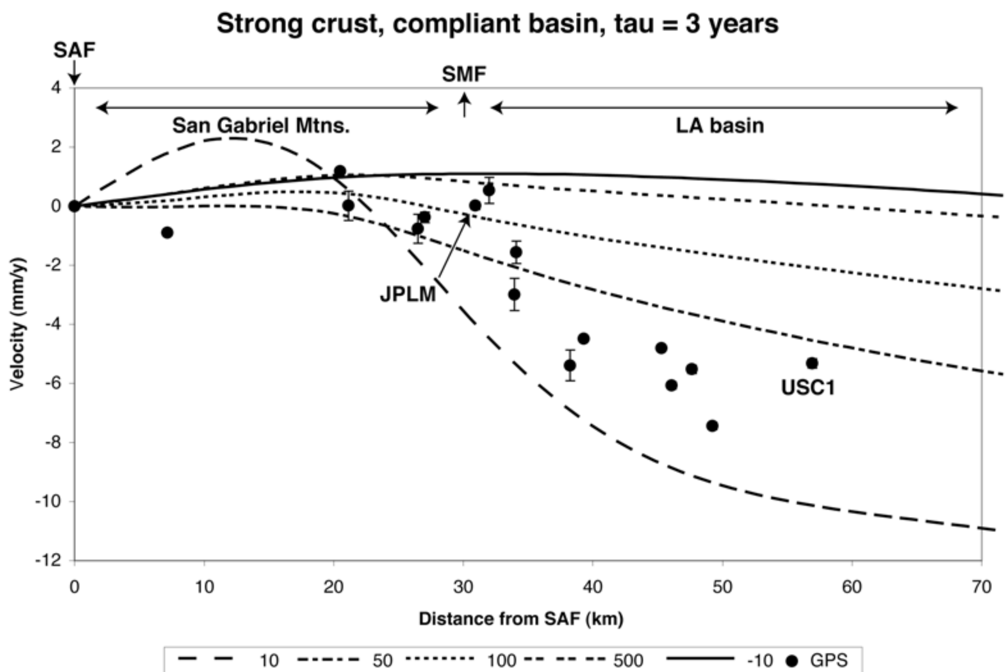
The preliminary results described above demonstrate that science capabilities are enhanced by use of the portal. Additionally, these capabilities are available to all of the community using the QuakeSim portal. While in this demonstration alternate fault dips were investigated, other studies shortly after the earthquake could have included testing each of the known faults in the area near the rupture.

New Implications for Los Angeles Basin Surface Deformation

GPS data shows an anomalous band of compression in the northern part of the Los Angeles basin. Scientists at UC Davis used QuakeSim tools to investigate both elastic and viscoelastic models of the basin to test what set of model parameters best reproduced the observed GPS data. They used SIMPLEX for elastic models and GeoFEST for viscoelastic models.

A GeoFEST model was constructed using geologic boundary conditions, recurring fault slip and a layered viscoelastic crustal structure. The top of the model is a free surface, which produces surface velocities that can be compared to observations from SCIGN measurements.

The models that yielded the best results featured a vertically strong crust, a single fault, and a low-rigidity, anelastically deforming basin abutting the rigid San Gabriel Mountains block (see Figure 4 below). These models produced a horizontal velocity gradient that matched the observed gradient across the LA basin both in magnitude and shape. While the match is not perfect, the models suggest that the sedimentary basin may be controlling the deformation in northern metropolitan Los Angeles, rather than faults south of the Sierra Madre fault as some have claimed.



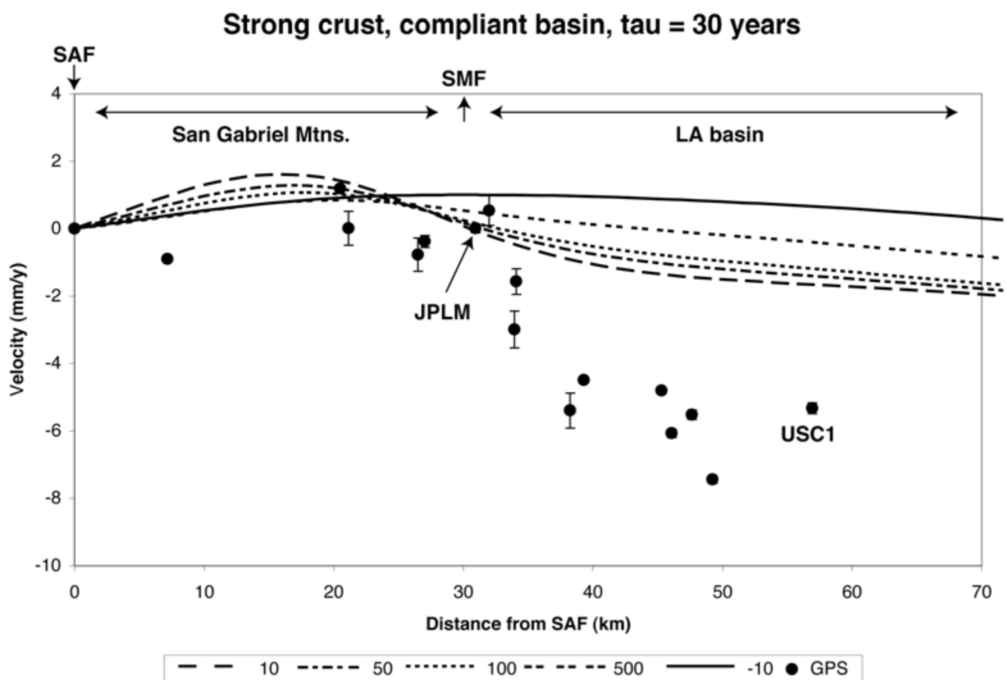


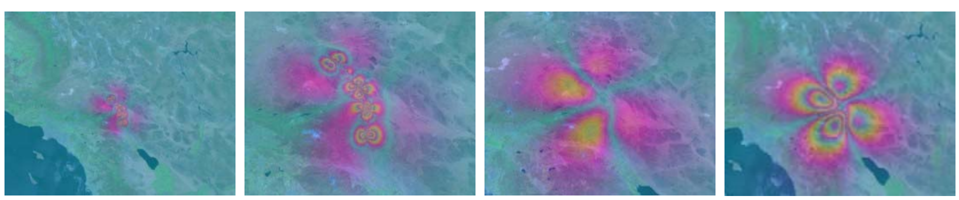
Figure 4. Comparison of strong crust/weak basin models with GPS residual velocities. The figures show horizontal velocities at the following times: 10 years after an earthquake event, 20 years after an earthquake event, 50 years after an earthquake event, 100 years after an earthquake event, 500 years after an earthquake event, and 10 years prior to the next scheduled earthquake event. The N 40º E component of the residual GPS velocity field is plotted to compare the model results to the observed data. SAF and SMF mark the locations of the San Andreas and Sierra Madre faults. JPLM and USC1 mark GPS stations at the Jet Propulsion Laboratory (at the edge of the basin) and the University of Southern California (near the middle of the basin).

Significance of Code Improvements to Geophysics

Achieving the First Code Improvement Milestone is significant because it opens up the way to probe significant earthquake physics. For the first time it presents to the scientific community fast parallel codes that allow creating simulations of the entire earthquake cycle on a fault in a 3D model that uses the most accurate description of fault friction, rate and state friction, and the quasi-dynamic radiation damping approximation to full elastodynamics. We now have the potential for greatly increasing the number of elements that can be included in the model over what could be done in the past.

With the PARK model, enough elements can now be used that is it possible to represent a reasonably sized fault with elements that are small enough that they can properly represent the behavior of a continuum. Larger numbers of elements also allow occurrence of earthquakes with a large range of sizes in the simulation. It will now be possible to simulate small earthquakes occurring in isolation and ones that cascade or grow into larger ones. It is currently not understood what causes small earthquakes to grow into large ones or stop at small events. Hence, these new simulations should be key to understanding earthquake rupture processes.

Comparing the PARK model with field data will also shed light on whether patterns of microseismicity can be used to predict earthquakes. Because of the new ability to model a wide range of earthquake dimensions, a single simulation can include thousands of tiny earthquakes and how they reflect the changes that lead up to massive events.



(a) Instant Uplift (b) Close-up of Uplift (c) Relaxation at 100 Yrs (d) Postseismic at 500 Yrs

Figure 5 (a-d). Landers event simulation with GeoFEST, showing post-seismic earthquake cycle phases.

GeoFEST users may now move beyond single fault analysis. As shown in Figure 5, regional analysis with multiple interacting faults is now within reach. The ability to solve domains with millions of elements implies we can simulate regions with multiple faults, such as the Los Angeles basin. Interactions among slipping faults and possible emergent structures from these nonlinear interactions represents a substantial advance in forecasting earthquake risk. While many kinds of simulation codes are emerging in the earthquake community, the GeoFEST finite element code approaches the greatest degree of flexibility in including the effects of realistic material diversity and faulting in the Earth.

Technology Accomplishments

Interoperability/Portal

We have the following accomplishments during the 2003 period of development:

- Completed core development of the portal system based on the Jakarta Jetspeed. We extended this to more easily support externally running web interfaces: we developed portlets to support HTTP sessions, SSL connections, HTTP GET/POST messages, and File Transfers.
- Continued development of several core services: Database access and querying, remote command execution, remote file management, user session management, and services to support sequences of closely related jobs. Many of these services were started in the previous year. During the current period we focused on enhanced performance of the session management services and the support for job sequencing. The latter was required to complete Milestone I.
- Completed integration of the Fault Database into the portal/service architecture. Software modules can directly access the Fault Database, allowing fault entries to be used directly by simulation codes such as GeoFEST.
- Integrated Disloc, Simplex, and GeoFEST into the portal. The Mesh Generation suite (Apollo and Akira) were also integrated into the portal and coupled to GeoFEST. Preliminary work on Virtual California, PARK, and Finley (from the University of Queensland) was begun.
- Coupled GeoFEST with the visualization code RIVA to generate movie visualizations of the Northridge fault.
- Demonstrated our portal and service system's coupling of resources at USC, JPL, and Indiana University to combine mesh generation, GeoFEST simulations, and RIVA simulations (Milestone I).
- Integrated Generic Mapping Tool (GMT) image visualization for Disloc and Simplex. A sample map for Disloc is shown in Figure 2.

Database Middleware

Four categories of database functions performed by the middleware have been implemented: insertion, selection, deletion, and modification. This middleware is portable since JAVA-based web technologies are cross-platform. There are three divisions of customized browser-based user interfaces that have been implemented: Fault Database, Layer Database, and California Geological Survey Database.

In addition to the browser-based user interfaces, customized web-service stubs are implemented by using the Web Service Description Language (WSDL) interface and the JAVA programming language. These stubs are used to retrieve data from different databases to perform further operations (e.g, simulations).

Code Improvement

Code	Machine Wallclock Time	Processors	Date	Elements	Time Steps
PARK Milestone E (7/30/02)	*Chapman (AMES)* 7.888 Hours	1	September 18, 2002	15,000	500
PARK Milestone F	*Chapman* (AMES) 7.879 Hours	256	August 15, 2003	150,000	5,000
GeoFEST Milestone E (7/30/02)	Solaris workstation (JPL) 13.7 Hours	1	July 30, 2002	55,369	1,000
GeoFEST Milestone F	*Thunderhead* (GSFC) 2.8 Hours	64	September 1, 2003	1,400,198	1,000

For Milestone F, we demonstrated use of the new parallel versions of PARK and GeoFEST to solve problems that would have taken 1000 times (PARK) and 400 times (GeoFEST) longer on the original baseline machines with no algorithmic improvement (based on extrapolation). For PARK, the improvement relies on integration with a parallel fast multipole library and good parallel coding practices. For GeoFEST the improvement is due to introduction of an efficient parallel iterative solver, integration with the PYRAMID library for domain decomposition, and porting to a computing cluster with faster processors than the baseline computer.

For PARK, problem preparation for scaling analysis is straightforward, and we performed 36 scaling runs that fill a grid with four problems sizes, each done with differing numbers of processors. The wallclock time for each of these 100 time step runs is shown in Figure 6. The scaling for a single processor follows the expected N logN curve, which displays dramatic improvement over the N^2 scaling of the previous full-interaction algorithm (close to a factor of 100 for the largest problem attempted). Additional speedup is obtained by applying more processors for the largest problem, running 50 times faster on 256 processors. Further improvements in efficiency are expected in the coming months. The PARK code implementation is the first simulation of an earthquake fault using the fast multipole technique, using a library originally developed for astrophysical, gravitationally-interacting bodies.

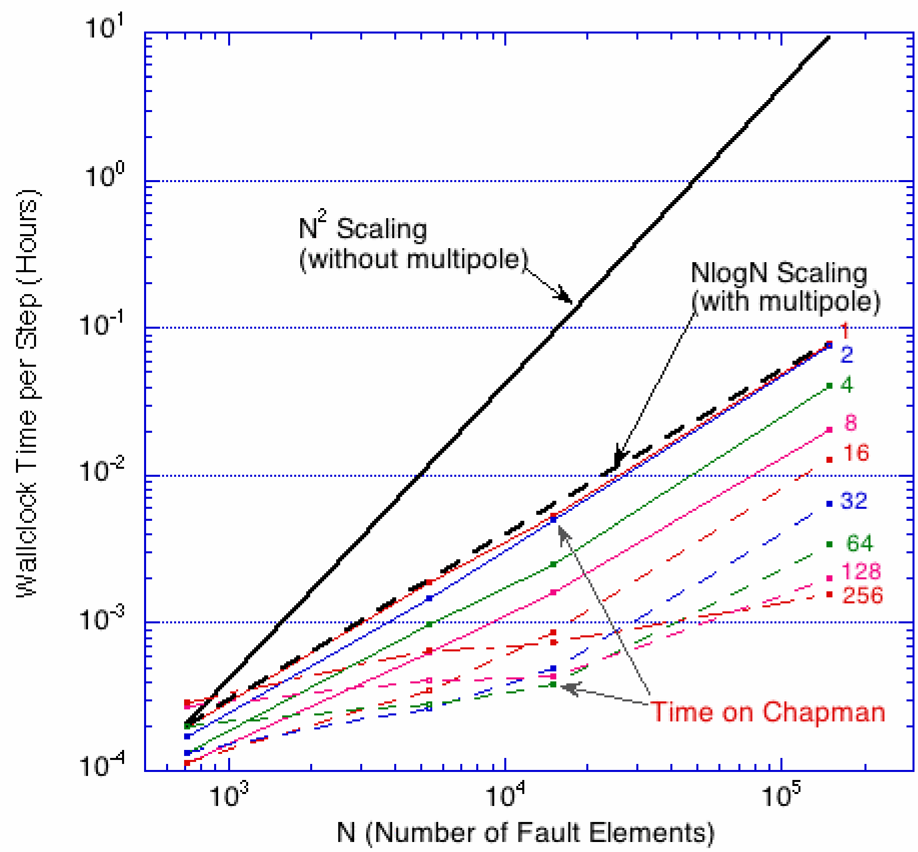


Figure 6. Time savings resulting from using fast multipole method and parallel computing. Black solid line shows scaling of time required for PARK per time step using traditional N^2 full interaction method, while black dashed line shows the expected N log N scaling for the multipole method on one processor. Colored curves show measured times with varying problem size, and indicate parallel improvement found with additional processors (colored numbers indicate processors used, 1 to 256). Scaling for small number of processors follow the expected N logN trend, while scaling for larger sets of processors appears anomalously favorable due to relatively high overhead for small problems.

For GeoFEST, the problem sizes are somewhat irregular due to generation of unstructured meshes. But the operation count is easy to estimate, and we have been able to demonstrate excellent scaling. Problems with > 4,000 elements per processor see negligible parallel overhead, and even smaller problems show substantial speed up with many processors. More detail on GeoFEST scaling is available in the Milestone F report.

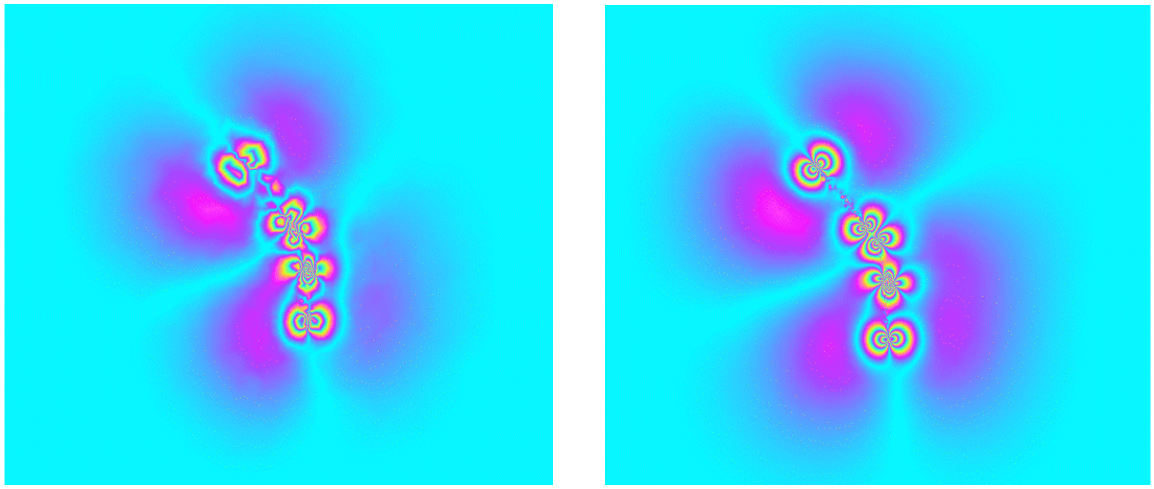
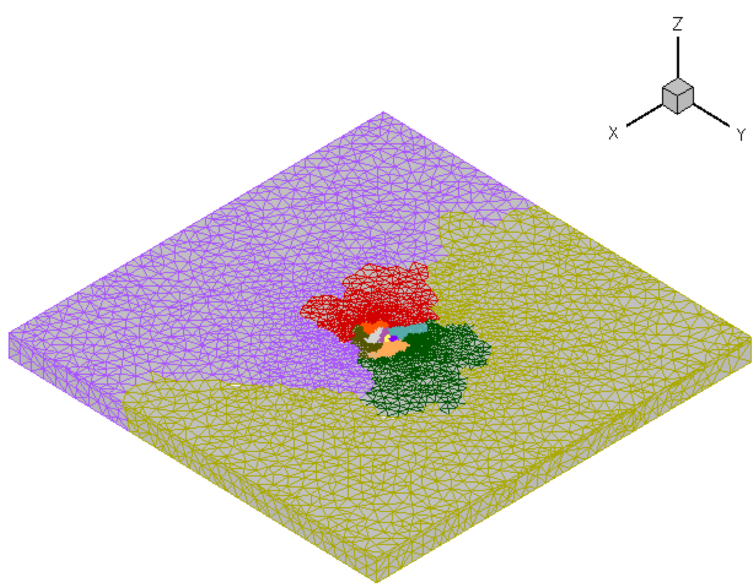


Figure 7. Models of Landers earthquake deformation at two resolutions. These images show the accuracy improvement going from 82,000 finite elements on four processors (left image) to 1.4 million finite elements on 64 processors (right image).

The main Milestone F case has 1.4 million elements on 64 processors (about 23,000 elements per processor). Figure 7 shows the surface elastic uplift for the Landers fault case for two meshes of different sampling density, for 4 processors (82,000 elements) and for 64 processors (1.4 million elements). The visual quality of the solution is dramatically improved, highlighting the need for meshes with millions of elements and parallel computers. Figure 8 shows the domain and mesh for the Milestone F 1.4 million element problem, with colors indicating domain decomposition among processors. The decomposition was made using the PYRAMID library. Figure 9 shows the surface nodes in the central surface area near the faults. This quality of mesh and GeoFEST solutions are easy to obtain using the QuakeSim portal.

Figure 8. Finite element mesh LandersGap64 for GeoFEST milestone F improvement code improvement problem. Colors indicate partitioning among processors (limited to 16 processors in this image for clarity, actually 64 processors were used). Partitions cluster near domain center due to the high mesh density that is used near the faults.



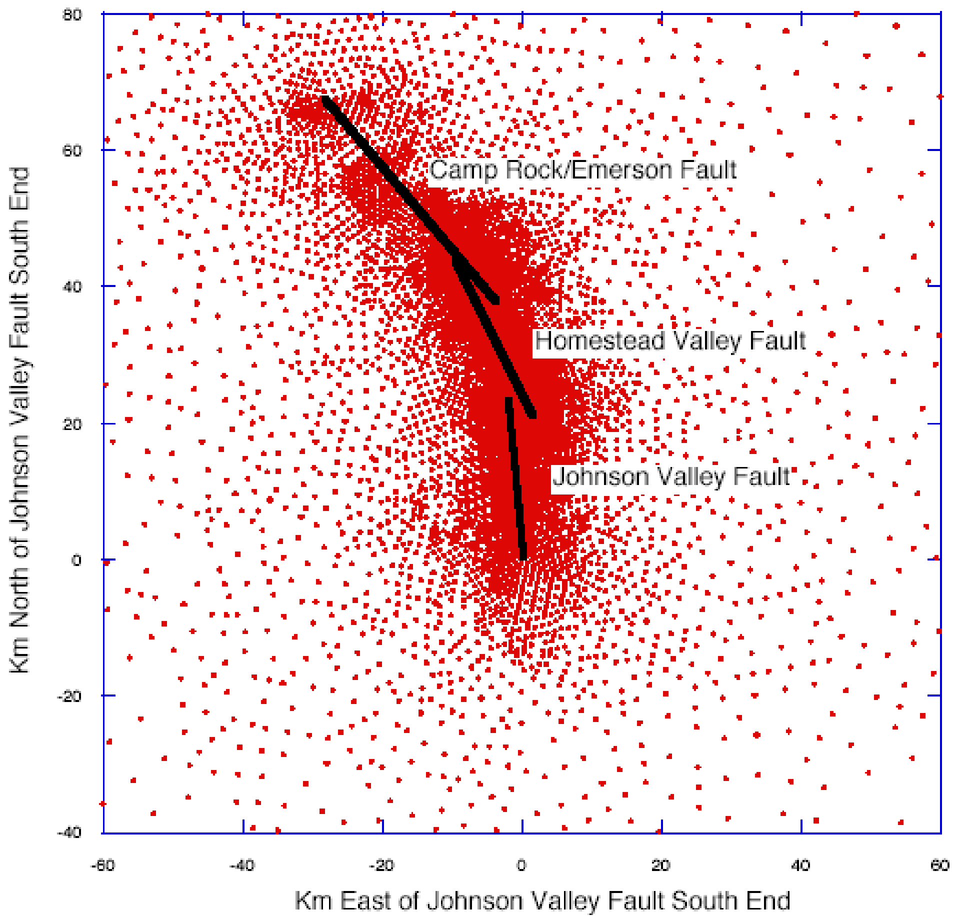


Figure 9. Vertical fault segments (10 km deep) and surface nodes for LandersGap64 mesh, center region. Node spacing is near 1 km close to the faults, gradually relaxing to 35 km at the boundaries beyond view. This mesh was generated automatically with the mesh generation tools integrated into the portal.

GeoFEST has been integrated with the PYRAMID library, creating a versatile system for parallel simulations via domain decomposition. PYRAMID is a system developed by the Computational Technologies project for parallel unstructured adaptive mesh refinement (AMR). AMR is an advanced computational technology applicable to many simulation problems, enabling highly accurate solutions to unstructured mesh problems in reduced time. For complex domains such as faulted plate boundary regions, AMR allows dynamic focusing of computational resources where they are most needed, avoiding wasted resources in domain regions where solutions are smooth and uninteresting. PYRAMID has advanced capabilities for modeling large-scale structure, simulation and visualization of results over extensive scales of length and resolution where highly accurate fitting to arbitrary geometries is required. Ongoing PYRAMID research seeks to validate advanced automated quality control for mesh elements. QuakeSim is a major partner in this work through this integration and validation of PYRAMID with the GeoFEST code.

In Year 2, the PYRAMID team worked closely with QuakeSim authors to instrument the GeoFEST code with PYRAMID library commands. To do this they developed an innovative C-adaptor library to bridge the GeoFEST C code with the FORTRAN 90 PYRAMID library. Work to date has brought into GeoFEST PYRAMID functions for parallel I/O, parallel mesh partitioning, and support for communication of boundary data among processors. Optimization of key communication functions was added after initial assessment of performance. Porting to a variety of parallel computing platforms ensures the future utility of this combination of GeoFEST and PYRAMID in future developments.

In the coming year, the full AMR capability of PYRAMID will be used by QuakeSim. Key features that will be added are implementation and verification of a strain-based indicator of mesh adequacy and assessment of strategies for cycling snapshot solutions from GeoFEST with a limited number of PYRAMID refinements. Using AMR, we expect to use GeoFEST for long-term simulations, with refined meshes that make near-optimal use of over 16 million finite elements. In this way users will be able to model exquisite details of complex regional active tectonics.

GeoFEST has been successfully run on the following computers:

- NASA GSFC "*thunderhead*" 512-processor Intel Pentium 4 Xeon Cluster
- NASA ARC "*altix*" 512-processor Intel Itanium II SGI Altix
- NASA JPL “*orion*” 64-processor Itanium II SGI Altix
- NASA JPL “*newyork”* 64-processor Apple G4 Xserve Cluster
- NASA JPL “*pluto*” 64-processor Intel Pentium 3 Cluster
- NASA JPL “*losangeles”* 32-processor Intel Pentium 4 Cluster
- UCLA “*uclapic”* 4-processor Apple G5 Tower Cluster
- NASA JPL “*leo”* 64-processor HP Itanium 2 Cluster
- Virginia Tech “*X”* 2200-processor Apple G5 Tower Cluster

Status/Plans

In the final year of the project, we plan to expand the prototype portal to the full QuakeSim portal, with over 13 component applications demonstrating interoperability through a web services framework. These will include generation of GeoFEST simulations by selection of fault database faults, automatic adaptative mesh generation, and visualization of 3D deformation and stress.

Further improvements will boost performance of the major simulations. These simulations are:

- PARK, which will demonstrate rupture generation for 50,000 time steps for a fault with 400,000 elements,
- Virtual California which will show 10,000 time steps of interaction of 700 fault segments representing the west coast tectonic system, and
- GeoFEST, which will demonstrate 1000 time steps for a regional model with 16 million finite elements. For GeoFEST, PYRAMID will be used to refine the mesh where needed due to high strain.

A "wizard" system will be demonstrated, enabling outside users to integrate their own applications into the QuakeSim framework, resulting in an expanding community tool. Finally, the QuakeSim tools will form the basis for a five-year earthquake forecast for events larger than magnitude 5 for southern California.

All software will be publicly posted on the World Wide Web. Publications are planned for open literature journals and the work will be documented in a final report.

To get user feedback and increase awareness of QuakeSim in the scientific and computational communities, PI Andrea Donnellan has proposed a QuakeSimintroductory tutorial at the 4th ACES (APEC Cooperation for Earthquake Simulation) workshop in Beijing, China, July 2004.

Although the QuakeSim project had difficulty meeting milestone dates in 2003 (mainly due to long lead time for JPL execution of university subcontracts), the work is now proceeding smoothly toward expected completion of all the agreed milestones by the end of 2004.

Points of Contact

Andrea Donnellan:
Principal Investigator

Jet Propulsion Laboratory
Mail Stop 183-335
4800 Oak Grove Drive
Pasadena, CA 91109-8099
donnellan@jpl.nasa.gov
818-354-4737

Michele Judd:
Technical Task Manager

Jet Propulsion Laboratory
Mail Stop 183-335
4800 Oak Grove Drive
Pasadena, CA 91109-8099
michele.judd@jpl.nasa.gov
818-354-4994

Publications

Berman, F., Fox G., and Hey, A., eds., *Grid Computing: Making the Global Infrastructure a Reality*, Wiley Series in Communications Networking & Distributed Systems, 2003.

Chen, A., Donnellan, A., McLeod, D., Fox, G., Parker, J., Rundle, J., Grant, L., Pierce, M., Gould, M., Chung, S., and Gao, S., Interoperability and Semantics for Heterogeneous Earthquake Science Data, *International Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data*, Sanibel Island, FL, October 2003.

Donnellan, A., Fox, G., Rundle, J., Tullis, T., McLeod, D., and Grant, L., Numerical Simulations for Active Tectonic Processes: Increasing Interoperability and Performance (invited), *Summary Symposium on the Solid Earth Simulator Project*, Tokyo, Japan, 2003.

Donnellan, A., Granat, R., and Rundle, J., Detecting Features in Seismic and Geodetic Data (invited), *35th Symposium on the Interface: Computing Science and Statistics*, Salt Lake City, UT, 2003.

Fox, G. and Walker, D., *e-Science Gap Analysis*, http://grids.ucs.indiana.edu/ptliupages/publications/GapAnalysis30June03v2.pdf.

Fox, G., Pallickara, S., Pierce, M., and Walker, D., Towards Dependable Grid and Web Services, *ACM Ubiquity*, v. 4, n. 25, 2003.

Gould, M., Grant, L., Donnellan, A., McLeod, D., and Chen, Y-A., The QuakeSim Fault Database for California, *Geological Society of America Abstracts with Programs*, v. 35, n. 6, 2003.

Gould, M., Grant, L., Donnellan, A., McLeod, D., and Chen, Y-A., The Quakesim Fault Database for California, *SCEC annual meeting*, Oxnard, CA, 2003.

Holliday, J., Rundle, J.B., Tiampo, K.F., Donnellan, A., Klein, W., and Turcotte, D.L., Evaluation and Development of the Pattern Informatics Method of Earthquake Forecasting, *Fall AGU meeting*, San Francisco, CA, 2003.

Parker, J.W., Donnellan, A., Lyzenga, G., Rundle, J.B., and Tullis, T., Performance Modeling Codes for the QuakeSim Problem Solving Environment, Computational Science, *Proceedings of the International Conference on Computational Science (Part III)*, Springer-Verlag, Berlin, p. 845-862, 2003.

Pierce, M., Youn, C., and Fox, G., Interacting Data Services for Distributed Earthquake Modeling, *ACES Workshop at the International Conference on Computational Science*, Australia, 2003.

Rundle, J.B., Rundle, P.B., Donnellan, A., and Li, P., Computing and Visualizing the Complex Dynamics of Earthquake Fault Systems: Towards Ensemble Earthquake Forecasting, *Fall AGU meeting*, San Francisco, CA, 2003.

Rundle, J.B, Kellogg, L.H., Tiampo, K.F., Klein, W. and Rundle, P.B., Integrating NASA Space Geodetic Observations with Numerical Simulations of a Changing Earth, in *Proceedings of the 3rd ACES Workshop, Maui HI*, GOPRINT, Brisbane, Australia, p. 129-136, 2003.

Rundle, J.B., Turcotte, D.L., Sammis, C., Klein, W. and Shcherbakov, R. Statistical Physics Approach to Understanding the Multiscale Dynamics of Earthquake Fault Systems (invited), *Review of Geophysics and Space Physics*, v. 41, n. 4, DOI 10.1029/2003RG000135, 2003.

Rundle, J.B., Klein, W., Tiampo, K.F., Donnellan, A., and Fox, G., Strategies for the Detection and Analysis of Space-Time Patterns of Earthquakes on Complex Fault Systems, *Proceedings of the International Conference on Computational Science (Part III)*, Springer-Verlag, Berlin, p. 827-836, 2003.

Rundle, J.B., Rundle, P.B., Donnellan, A. and Fox, G., Statistical Mechanics Approaches to the Modeling of Nonlinear Earthquake Physics, *Conference proceedings of the International Union of Geodesy and Geophysics,* Sapporo, Japan, 2003.

Rundle, J.B., Scaling and Critical Phenomena in Nature (book review of Hergarten, S., *Self-Organized Criticality in Earth Systems*, Springer-Verlag, 2002; ISBN 3540434526), *Computing in Science and Engineering (CiSE)*, v. 5, n. 5, p. 80-81, 2003.

Rundle, J.B., Klein, W., Tiampo, K.F., Donnellan, A., and Fox, G.C., Strategies for the Detection and Analysis of Space-time Patterns in Complex Driven Threshold Systems, *Proceedings of the International Conference on Computational Science (Part III)*, Springer-Verlag, Berlin, p. 827-844, 2003.

Tiampo, K.F., Rundle, J.B., McGinnis, S., Klein, W. and Ben-Zion, Y., Using Eigenpattern Analysis to Constrain Numerical Fault Models – Application to Southern California, in *Proceedings of the 3rd ACES Workshop, Maui HI*, GOPRINT, Brisbane, Australia, p. 67-74, 2003.

Tiampo, K.F., Rundle, J.B., Klein, W., Martins, J.S.S., and Ferguson, C.D., Ergodic Dynamics in a Natural Threshold System, *Physics Review Letters*, v. 91, n. 238501, p. 1-4, 2003.

Tiampo, K.F., JB Rundle, W Klein and JSS Martins, Ergodicity in Natural Fault Systems, in *Proceedings of the 3rd ACES Workshop, Maui HI*, GOPRINT, Brisbane, Australia, p. 75-82, 2003.

Tullis, T. E., Earthquake Models Using Rate And State Friction And Fast Multipoles, *Geophysical Research Abstracts* (European Geophysical Society), v. 5, p. 13211, 2003.

Youn, C., Pierce, M.E., and Fox, G.C., Building Problem Solving Environments with Application Web Service Toolkits, *Special Issue of FGCS Magazine*, 2003.

Youn, C., Pierce, M., and Fox, G., Developing Secure Web Services for Computational Portals, in *Proceedings of 2003 International Conference on Security and Management,* Las Vegas, NV, 2003.

Youn, C., Pierce, M., and Fox, G., Building Problem Solving Environments with Application Web Service Toolkits, *International Conference on Computational Science*, Australia, 2003.

Media References

Articles and Press Releases

Cole, K., "Tiny California Town Is The Focus Of Geologist's Effort To Predict Quakes", George Street Journal, July 11, 2003. Included interviews with Terry Tullis and Andrea Donnellan. http://www.brown.edu/Administration/George_Street_Journal/vol27/27GSJ31f.html

Chamot, J., "Earthquake Warning Tools" *Geotimes.Org* October 2003. http://www.geotimes.org/oct03/technology.html

McGraw, D., "Taking a Crack at Predicting Quakes" *PRISM* pp. 30-33, November 2003.

Oskin, B., "Quake-Simulation Tool in Works" *Pasadena Star-News* May 23, 2003.

Wilson, J., "Predicting Earthquakes from Space" *Popular Mechanics* November 10, 2003.

JPL Press Release, "New NASA Computer Models May Lead to Quake Forecast System", May 22, 2003. http://www.jpl.nasa.gov/releases/2003/74.cfm

"NASA Computer Models May Lead To Quake Forecasts" *SpaceFlightNow.com* May 24, 2003. http://spaceflightnow.com/news/n0305/24quake/

"NASA Cites Progress In Earthquake Research" *SpaceFlightNow.com* December 4, 2003. http://spaceflightnow.com/news/n0312/04earthquake/

"Progress, Promise In Space-based Earthquake Research" *ScienceDaily.com* December 8, 2003. http://www.sciencedaily.com/releases/2003/12/031205051954.htm

"Space-Based Research Helps Predict Earthquakes" *MIT News* December 12, 2003. http://web.mit.edu/newsoffice/nr/2003/earthquakes.html

"Leading Computational Earth Scientists Meet in Brisbane" *University of Queensland News Online* June 5, 2003. http://www.uq.edu.au/news/index.phtml?article=4524

Live Interviews

Donnellan, A.: Live interviews on QuakeSim Computational Technologies project, including CNN Headline News and top story feature on Next@CNN, May 2003.

Donnellan, A., Rundle, J.: NASA Earth Science Update, Washington D.C., December 2003

Donnellan, A.: Live interview with Los Angeles based radio station, KFWB, on InSAR and computing, December 2003.

Donnellan - Selected Talks and Briefings

Space Technology and Solid Earth Science Briefings - 2003

- Chip Groat, Director of the US Geological Survey, January 2003.
- Michele Burkett, Staff Assistant, House Committee on Appropriations, January 2003.
- Congressman Dreier, February 2003;
- Congressman Rhorabacher, February 2003;
- Mike O'Brien and Joe Wood, NASA HQ, Code I, August 2003.
- Amy Kaminski, OMB examiner for NASA Code S, August 2003.
- Margaret Leinen, NSF Deputy Director for Geosciences, October 2003
- Congressional Committee on Science Research, November 2003.
- John Marburger, President's science advisor, December 2003.

Outreach - 2003

- The Use of Space Technology for Studying Earthquakes, Von Karman Public Lecture, JPL and Pasadena City College, January 2003.
- The Use of Space Technology to Study Earthquakes, Adventurer's Club, Los Angeles Chapter, April 2003.
- Talk on earthquakes and space technology to 3–5 graders at St. Elizabeth's School in Altadena, California, April 2003.
- Talk on *Plate Tectonics and You: A Cool View of Earthquakes*, Happy Valley School, Summer Science Program to high school students, July 2003.
- *Using Space Technology to Understand Earthquakes*, IEEE Women in Engineering, Ventura Chapter, October 2003.

Other Presentations

Geoffrey Fox Applications as Web (Grid) Services and Related Issues Argonne National Laboratory November 10 2003.

Marlon Pierce Metadata and the Semantic Web Access Grid Presentation for PET August 29 2003.

Geoffrey Fox, David Walker e-Science Technology/Middleware (Grid, Cyberinfrastructure) Gap Analysis and OMII SEAG Meeting DTI June 20 2003

Geoffrey Fox iSERVOGrid Architecture Working Group Third ACES Working Group Meeting Brisbane June 5 2003.

Geoffrey Fox and Marlon Pierce Grid Technology Implications for ACES and SERVOGrid Third ACES Working Group Meeting Brisbane June 5 2003

Marlon Pierce, Choonhan Youn, and Geoffrey Fox Interacting Data Services for Distributed Earthquake Modeling ICCS Melbourne Australia June 4 2003

Geoffrey Fox Some Future Semantic Grid Activities CrisisGrid and ServoGrid presentation at GGF7 Tokyo Semantic Grid Meeting March 6 2003.

G.Fox, D. Gannon, M. Pierce, M. Thomas Overview of Grid Computing Environments presentation at GGF7 Tokyo Portal Architecture Workshop March 6 2003 summarizing document of same name

Academic Training

***Fall 2003 CSUN Geology 464** (Applied Geophysics) course: Professor Gerald Simila and his geophysics students used the QUAKESIM portal to run DISLOC to model the 1994 Northridge earthquake. Dr. Simila has a NASA Faculty Fellowship Program (Summer 2003) at JPL with Dr. Andrea Donnellan. His project involved using the QUAKESIM portal to test the usability and development of the fault modeling programs (DISLOC and SIMPLEX) and associated graphics.

STUDENTS and POSTDOCS

Jordan Van Allsburg – University of California, Davis, Graduate Student, Department of Physics, graduate student working on percolation analysis foundations for statistical earthquake data analysis.

Galip Aydin – Indiana University, graduate student developing the GML schemas to describe seismicity and GPS data. Implemented services for extracting legacy data from online repositories. Building services and user interfaces to extract sections of the data.

Teresa Baker – Massachusetts Institute of Technology, BS undergraduate student, used Simplex, Disloc, and GeoFEST to model deformation associated with the Northridge earthquake and adjacent Ventura basin prior to matriculation in May 2003. Teresa is currently working on the QuakeSim project at JPL.

Anne Yun-An Chen – University of Southern California, PhD graduate student is undertaking development of the fault and federated databases.

James Holiday – University of California, Davis, Graduate Student, Department of Physics, developing and testing new Pattern Informatics methodologies.

Shan Gao – University of Southern California, PhD graduate student is developing the interoperability aspects of the fault databases.

Maggi Glasscoe – University of California, Davis, MS graduate student used geoFEST to model observed GPS deformation and coupling between the crust and the mantle in the Los Angeles basin prior to matriculation in December 2003. Maggi has taken a full-time position at JPL and currently supports the QuakeSim project at JPL.

Miryha Gould – University of California, Irvine, PhD graduate student is developing the geological aspects of the fault database.

Dr. Gleb Morein – University of California, Davis, Postdoctoral Fellow developing parallel slider code for high performance applications. He is also converting Virtual California to parallel/high performance format.

Paul Rundle – Undergraduate Student, Harvey Mudd College, working on earthquake fault model data bases and static data assimilation for Virtual California.

Ahmet Sayar – Indiana University, graduates student developing authoring tools to create scripts for coupling services needed by the QuakeSim portal.

Dr. Robert Shcherbakov – University of California, Davis, Postdoctoral Fellow Analysis of seismicity and geodetic databases for Pattern Informatics earthquake forecasting.

Sang Soo Sung – University of Southern California, PhD graduate student is assisting on the development of the federated databases.

Choonhan Youn - Syracuse University. Youn defended his Ph.D. thesis in Nov 2003. Dissertation work was based in part on work performed as part of the NASA CT project.