Understanding Earthquake Fault Systems using QuakeSim Analysis and Data Assimilation Tools

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

Lisa Grant
University of California, Irvine
Irvine, CA 92697

Abstract—We are using the QuakeSim environment to model interacting fault systems. One goal of QuakeSim is to prepare for the large volumes of data that spaceborne missions such as DESDynI will produce.. QuakeSim has the ability to ingest distributed heterogenous data in the form of InSAR, GPS, seismicity, and fault data into various earthquake modeling applications, automating the analysis when possible. Virtual California simulates interacting faults in California. We can compare output from long time-history Virtual California runs with the current state of strain and the strain history in California. In addition to spaceborne data we will begin assimilating data from UAVSAR airborne flights over the San Francisco Bay Area, the Transverse Ranges, and the Salton Trough. Results of the models are important for understanding future earthquake risk and for providing decision support following earthquakes. Improved models require this sensor web of different data sources, and a modeling environment for understanding the combined data.[1,2]

TABLE OF CONTENTS

1. INTRODUCTION 1
2. FEDERATED DATABASE 2
3. MODELING AND SIMULATION APPLICATIONS 3
4. FORECASTING 4
5. CONCLUSIONS 6
6. ACKNOWLEDGEMENTS 7
REFERENCES 7
BIOGRAPHIES 8

1

[1] 978-1-4244-2622-5/09/$25.00 ©2009 IEEE.
[2] IEEEAC paper #1100, Version 3, Updated December 10, 2008

1. INTRODUCTION

Traditional earthquake science is done by domain experts in individual, disjointed efforts and has typically been ground-based. As an example, crustal deformation observations conventionally come from triangulation, trilateration, or strain meters. More recently GPS networks have been used to measure crustal deformation. InSAR measurements provide detailed images of surface displacements. Large volumes of systematic measurements will be made with the advent of UAVSAR and DESDynI, one of the recommended Decadal Survey missions. Crustal deformation modeling can include analytic solutions, finite element models, or elastic and viscoelastic models. Fault mechanics includes knowledge of fault frictional properties, fault nucleation, and rupture propagation. Integrating these models and the various datasets will enhance our understanding of earthquake processes.

The advent of web infrastructure and applications allows the integration of various scales of models to improve understanding of earthquake processes and mechanics. Data constrain the models; hence data sources also need to be integrated into the computational infrastructure. Developing prototype systems now using UAVSAR, will prepare us for routinely analyzing data from DESDynI. We use a service architecture exploiting a mix of Grid and Web 2.0/Cloud technologies to support computational and data systems. Descriptions of datasets and their inter-relationships are essential to support effective interfaces for

Table 1. Application descriptions, data sources, and how they are compared.

Application	Purpose	Data Source	Compared To
Virtual California	Interacting fault model	Faults, friction	Earthquakes
GeoFEST	Finite element deformation model	Faults	GPS and InSAR surface deformation
Disloc	Surface displacements from fault motion	Faults	GPS and InSAR surface deformation
Simplex	Inversion for fault motion from deformation data	InSAR and GPS surface deformation	Fault data constrain model
RIPI	Seismic pattern analysis	Seismicity	Earthquake faults
RDAHMM	Time series analysis	GPS	Earthquake, aquifer sources

users, including data discovery, delivery, visualization, and integration. The QuakeSim federated database QuakeTables, interconnects data sources providing information that is then accessed by the modeling applications. The increased volumes of data provide an opportunity to carry out pattern analysis and study earthquakes from a nonlinear dynamical perspective.

QuakeSim is a project to integrate both real-time and archival sensor data with high-performance computing applications for data mining and assimilation to improve earthquake forecasts. The multi-scale nature of earthquakes requires integrating many data types and models to fully simulate and understand the earthquake process. QuakeSim focuses on modeling the interseismic process through various boundary element, finite element, and analytic applications, which run on various platforms ranging from desktops to high-end computers. Making these data available to modelers is leading to significant improvements in earthquake forecast quality and thereby mitigating the danger from this natural hazard.

Modeling earthquake fault systems enables us to understand how earthquake faults interact, how stress is transferred between faults, and how an earthquake on one fault will delay or advance the likelihood of an earthquake on other faults. QuakeSim incorporates several approaches to modeling these systems (Table 1). Virtual California is a boundary element software application that models interactions between vertical strike-slip faults. GeoFEST is finite element software that computes the stress and strain field from arbitrary faults within a mesh. GeoFEST can model faults of any geometry heterogeneous elastic or viscoelastic crustal material properties. Analytic tools Disloc and Simplex calculate surface deformation or fault parameters in an elastic half space. Pattern analysis algorithms have proved useful for analyzing time series (GPS) and spatial (seismicity) data. The latter looks promising for earthquake forecasting.

2. FEDERATED DATABASE

Modeling interacting fault systems and surface deformation due to fault slip requires specifications of faults as input to the models. Through a web-services based portal, QuakeSim provides global access to reference models of faults and fault data as well as the analysis tools. Federation of the data records and provides portal-based access to diverse data types. Currently, QuakeTables houses a comprehensive set of paleoseismic and fault data that can be ingested into QuakeSim applications. The database includes processed InSAR data and is being populated with processed California interferograms, courtesy of Paul Lundgren, as they become available.

QuakeSim applications are used to model fault activity such as rate of strain accumulation or offset related to earthquakes over a finite fault segment. 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. It is a challenge to convert data, particularly those collected and reported by a variety of means into standard data for modeling applications. For many faults there are multiple interpretations.

The purpose of QuakeTables is to standardize data for QuakeSim modelers and allow the modeler to further refine interpretations about faults [1]. We have developed ontologies for the different types of data, such as paleoseismic fault data, and InSAR data. QuakeTables does not house one single, self-consistent, fault model for California. Rather than defining a single reference model

for faults or other data QuakeTables houses published data sources. Multiple different interpretations and data sets allows the user to access a self-consistent set of faults for their model and to be able to trace a fault segment recorded in the database back to the original reference. This allows the modeler to test various fault interpretations and leave the option for uploading a new fault interpretation based on refined analysis. An important capability of the 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.

3. Modeling and Simulation Applications

Disloc and Simplex

Disloc and Simplex are both based on Okada's [2] formulation for calculating surface displacements from dislocations in an elastic half-space. Disloc is a forward model for calculating surface deformation from prescribed slip on an arbitrary number of faults. Simplex is an inverse algorithm residual minimization algorithm using disloc to calculate fault motions and errors based on surface deformation measurements [3].

Simplex is limited to homogeneous elastic models, and so we are developing a correction step, wherein the Simplex estimate of fault slip automatically becomes the basis of a suite of detailed finite element simulation including known materials variation. This suite of results will indicate the best first-order correction to the fault parameters. The goal is a broadly adaptable integrated system of precision surface deformation monitoring, combined with a modeling system that incorporates processes at multiple scales. This will allow definition of a baseline model of regional and global deformation processes, to be continuously compared with sensor observations for automatic early detection of unusual events. Although this feature is not yet part of QuakeSim, it illustrates the goal of interfacing modeling tools of different scales and with different features, to improve the overall model of fault system deformation and interactions.

Virtual California (VC)

[4-6] is a topologically realistic boundary element numerical simulation of earthquakes occurring on the fault systems of California. It includes all the major strike-slip faults in California and has now been extended to depth-dependent boundary elements and dipping faults. VC is an efficient, parallel object-oriented C++ numerical code that runs on NASA computers Columbia and JPL computer COSMOS using MPI-II protocols. VC can also compute associated surface displacements through time. From the related simulation, InSAR interferograms can be computed for use in analyzing and interpreting signals from real interferograms obtained via radar satellites such as DESDynI. This process of comparing simulated interferograms with real interferograms will allow us to study questions relating to physics of earthquakes such as: s1) Precursory failure process of major earthquakes on complex fault systems; 2) Timing and statistics of major earthquakes on complex fault systems; and 3) Origin of space-time correlations between major earthquakes. Another major application of VC simulations lies in earthquake forecasting. In weather forecasting, current and past observational data are routinely assimilated into numerical simulations to produce ensemble forecasts of future events in a process termed "model steering." We have developed a similar approach (Van Aalsburg et al., 2008) that is motivated by analyses of previous 30-year forecasts of the Working Group on California Earthquake Probabilities (Field, 2008). By systematically comparing simulation to observed data the variability of paleoseismic and historic data, a series of spatial probability density functions (PDFs) can be computed that describe the probable locations of future large earthquakes. These forecasts yield fault-based locations for the next earthquake, as well as most probable locations for earthquakes during the next 30 years.

Virtual California produces a very large synthetic seismic record. Systematic observations over a long time period provide an opportunity to observe emergent behavior and fault interactions in the system. Examining 40,000 years of VC data and calculating the correlation of events on an "initiating" fault segment with subsequent events on a second fault element produces a correlation score matrix that shows the relative amount of correlation between events on two elements (Figure 1). Analysis has been expanded to include processing of data at different time window lengths. Examining correlation score matrices at different time window lengths can provide insight into the relationships between faults at different time scales. The events on the Eastern California Shear Zone typically precede, but do not follow events on the southern San Andreas fault.

GeoFEST

GeoFEST is a finite element program for modeling fault motion in a mechanically realist crust. It includes nonlinear rheology, slip on curved and wrinkled faults, and production of images of Coulomb Stress [7]. Anomalies can be detected and assessed with less approximation in the modeling. Stress transfer can be estimated for hazard assessment.

GeoFEST is being used to investigate the effects of large earthquakes over time. The magnitude 7.7 [8] 1906 San Francisco earthquake produced a rupture of 2–8 m rupturing the northern third of the San Andreas fault [9, 10]. It has

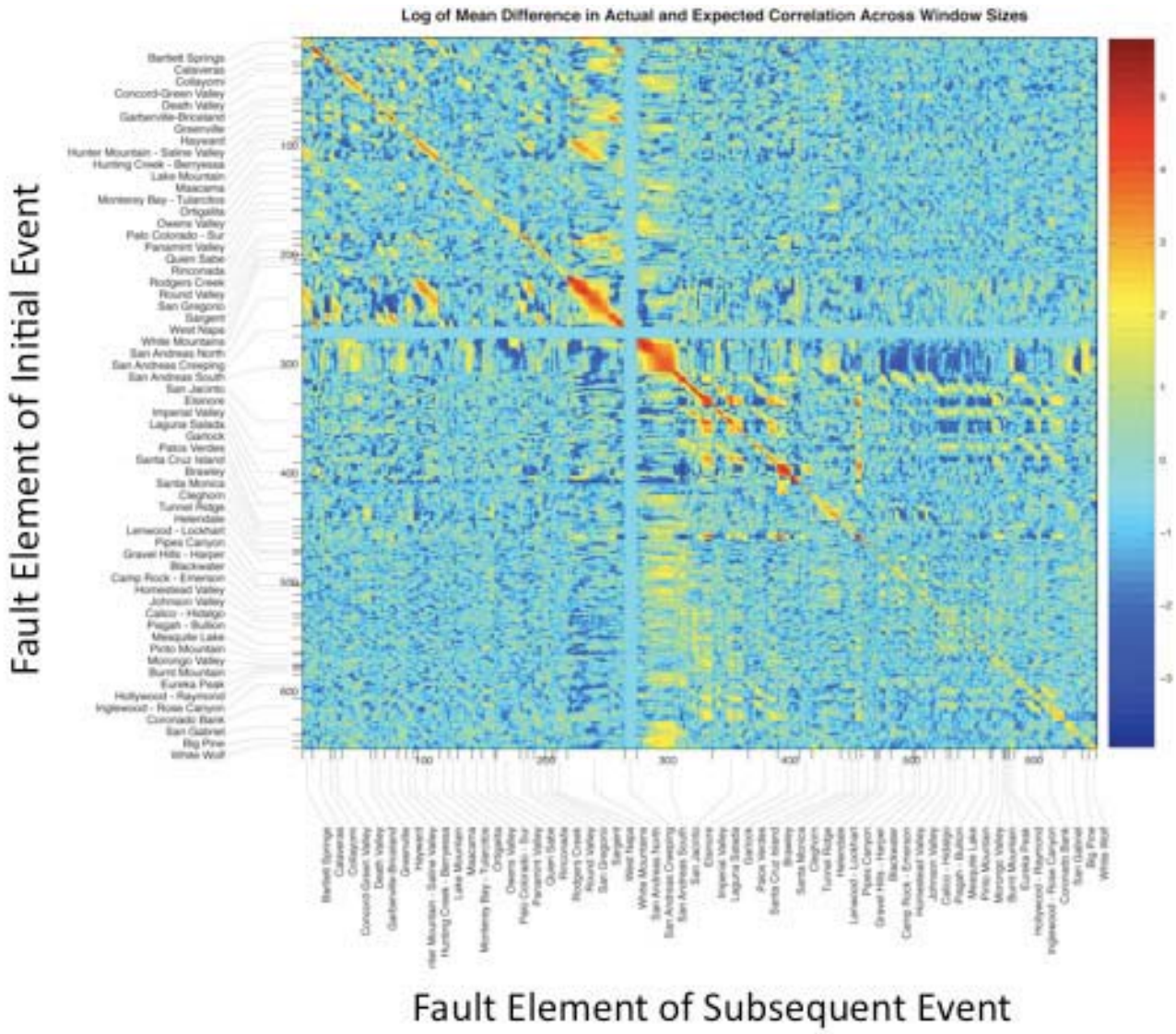

Figure 1. Log of correlation score matrix 400 yr time window. Raw score matrix is saturated at values around zero, so a log function is applied to highlight interesting features. Analysis above features 59 faults (639 elements) and does not include the "creeping" section of the San Andreas fault because of computational considerations.

been questioned whether there is still a detectable strain signature in the present-day geodetic data. Studies to date with GeoFEST point to the importance of developing realistic models of crustal deformation (Figure 2). More complex (realistic) models show postseismic effects of the 1906 earthquake of 1 mm/yr rather than the 2–5 mm/yr for the less complex models (Figure 3).

4. Forecasting

The Pattern Informatics (PI), Relative Intensity (RI), and RIPI methods use online seismicity catalogs to generate space-time forecast maps. Systematic and ongoing real-time tests of these forecasts are posted on the NASA/JPL quakesim web site. A real-time test of the original method was published in Rundle et al., 2002 [11], and a test of an updated and modified method as published in Holliday et al. 2007 [12]. The great majority of the earthquakes have occurred on or near a colored anomaly, or "hotspot" on that map. The location of the recent M5.4 Chino Hills earthquake was successfully forecast by both maps, but the more recent M 5.1 Ludlow earthquake (Figure 4). At present we update versions of a forecast map on the QuakeSim web site as publications occur. Eventually we will publish continuously updated forecasts.

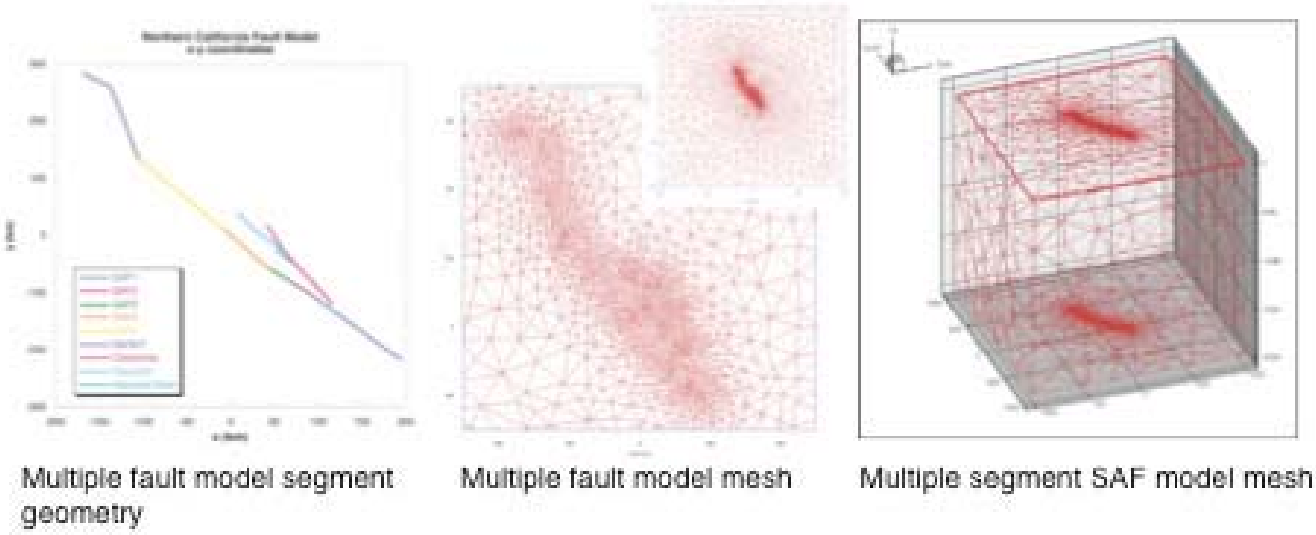

Figure 2. Model geometry for complex 1906 earthquake postseismic deformation models. QuakeSim portal tools were used to generate initial and subsequent meshes for input into GeoFEST viscoelastic simulation software.

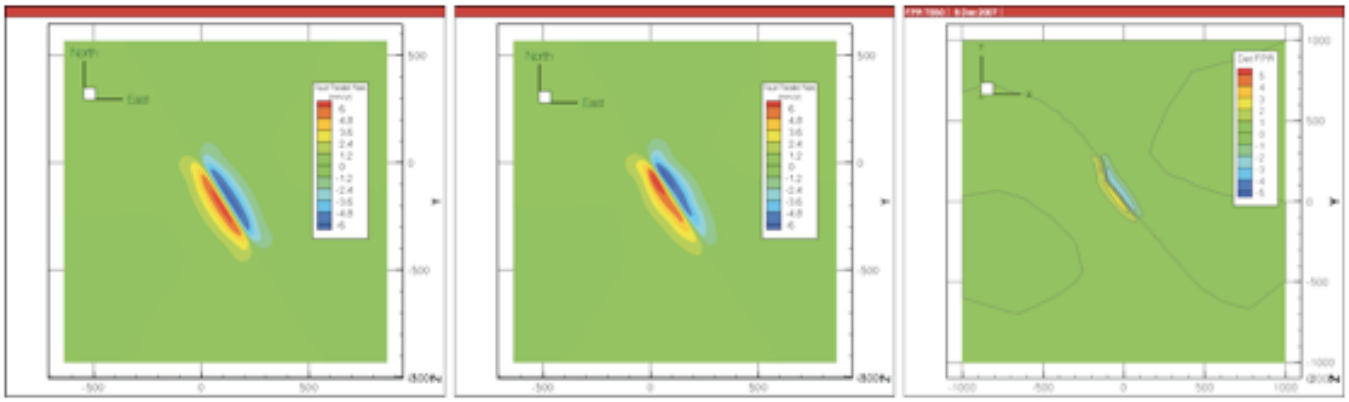

Figure 3. Fault parallel displacement for a: left panel: uniform slip model, middle panel: variable slip on the fault, and right panel: multiple segment San Andreas fault with plate boundary conditions. Residual velocities become smaller as the model becomes more complex.

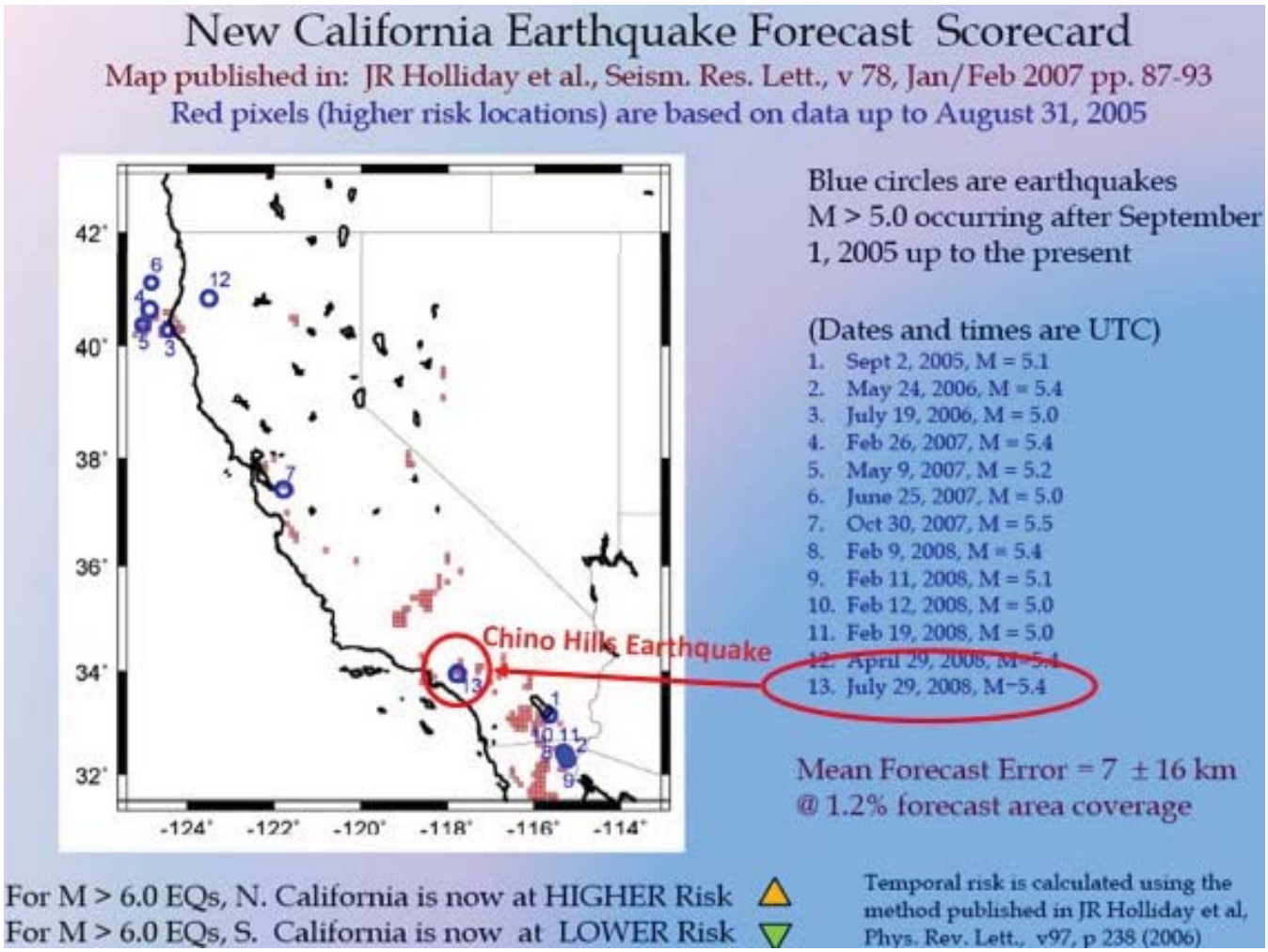

Figure 4. Recent successful forecast of the M 5.4 Chino Hills earthquake of July 29, 2008.

In the last twelve months eight earthquakes above magnitude 5 have occurred in identified hotspots (Figure 4). The identified hotspots make up only 1.2% of the total map area of the forecast or the state of California. Approximately half the total boxes have at least 1 M>3 earthquake in them. Therefore for the total active area, it would be forecast area of 2.4%. The approach is to minimize the forecast area, which is essentially the false alarm rate, while still detecting all the large earthquakes (maximizing the hit rate). The mean forecast error is the average distance that a M>5 earthquake occurs from a 11 km pixel box boundary. Most of this error is due to offshore earthquake # 6, which is about 50 km off the nearest red pixel. Without that earthquake, the error would be less than 5.5 km, half of one pixel box size. Additional earthquakes will allow us to optimize the forecast area.

Earthquake forecasts and models with time frames of years rather than decades and much smaller spatial scales on the order of 10 km rather than 100 km scales are now possible. The models are based on patterns of seismicity at present. Similar pattern recognition techniques will be used on the radar interferograms from DESDynI when they become available. The methodology requires systematic data, which DESDynI will produce. Current InSAR missions do no produce systematic crustal deformation measurements.

5. Conclusions

QuakeSim model applications are showing the importance of modeling the complexity and detail of fault systems. Faults interact with each other, and the order in which earthquakes occur impacts the order of subsequent earthquakes. Developing models to match observed data requires realistic models. Efficiency is gained by the access of the data by models such as of QuakeTables through QuakeSim. Results from the Pattern Informatics method and Virtual California indicate that seismicity and crustal deformation produce detectable patterns that can be used to infer earthquake hazard. Continued modeling will indicate patterns of crustal deformation that may be key to improved earthquake forecasting. Developing these prototypes now is important for being prepared to maximize the use of InSAR crustal deformation data from DESDynI and to develop a routine operational data analysis system. UAVSAR will be

a valuable testbed for developing a prototype system for DESDynI, but will also be important for filling in data gaps in response to earthquakes or other events.

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. Grant, L.B., et al., *A Web Services-Based Universal Approach to Heterogeneous Fault Databases.* Computing in Science and Engineering Special Issue on Multi-Physics Modeling, 2005. 7(4): p. 51-57.

2. Okada, Y., *Surface deformation due to shear and tensile faults in a half-space.* Bulletin of the Seimological Society of America, 1983. 75(4): p. 1135-1154.

3. Lyzenga, G.A., W.R. Panero, and A. Donnellan, *The Influence of Anelastic Surface Layers on Postseismic Thrust Fault Deformation.* Journal of Geophysical Research, 2000. 105: p. 3151-4157.

4. Rundle, P.B., et al., *Virtual California: Fault model, frictional parameters, applications,.* Pure and Applied Geophysics, 2006. 163(1819-1846).

5. Rundle, J.B., et al., *A simulation-based approach to forecasting the next great San Francisco earthquake.* Proceedings of the National Academy of Sciences, 2005. 102: p. 15,363-15,367.

6. Van Aalsburg, J., et al., *A feasibility study of data assimilation in numerical simulations of earthquake fault sytems.* Physics of the Earth and Planetary Interiors, 2007. 163(149-162).

7. Parker, J., et al., *Geophysical Finite Element Simulation tool (GeoFEST): algorithms and validation for quasistatic regional faulted crust problems.* Pure and Applied Geophysics, 2008. 165: p. 497-521.

8. Wald, D.J., H. Kanamori, and D.V. Helmberger, *Source Study of the 1906 San Francisco Earthquake.* Bulletin of the Seimological Society of America, 1993. 83: p. 981-1019.

9. Ellsworth, W.L., *Earthquake History 1769-1989*, in *The San Andreas Fault System*, R.E. Wallace, Editor. 1990, U.S. Geological Survey Professional Paper. p. 152-187.

10. Lawson, A.C., *The California Earthquake of April 18, 1906*, in *Report of the State Earthquake Investigation Commission*. 1908, Carnegie Institution of Washington.

11. Rundle, J.B., et al., *Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and forecasting.* Proceedings of the National Academy of Sciences, 2002. 99: p. 2514-2521.

12. Holliday, J.R., et al., *A RELM Earthquake Forecast Based on Pattern Informatics.* Seismological Research Letters, 2007. 78: p. 87-93.

BIOGRAPHIES

Andrea Donnellan is the 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.

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.