The Fault Activity Database and its Evolution into a Fault Information System

P.I.: Thomas Jordan

Annual Report by Sue Perry 2004

SCEC year 2004 saw completion of the first phase of development in the SCEC Fault Activity Database (SCEC-FAD) and the SCEC Fault Information System (FIS). The next phases will be funded under new grants awarded by the USGS/EHRP-External and the California Earthquake Authority, and will be part of an effort to create a statewide earthquake rupture forecasting tool with programmatic inputs from fault databases at the USGS and SCEC. For this project, a new database is being fashioned, the Reference Geologic Fault Parameters Database (RefGF).

The FIS has been designed to streamline interdisciplinary research, with a goal to provide scientists with 'one-stop,' internet access to a suite of fault-related data, models, and visualization tools. The FIS is meant to complement, not replace, the databases and user interfaces of other groups. Few of its datasets will be housed at, or maintained by, SCEC. The FIS will never copy the datasets of others (because that creates corrupted and confusing data sets) nor supplant the web pages of other organizations (because that is a wasteful duplication of effort). Instead, the FIS enables users to discover and locate the myriad of available products, and understand how the datasets interrelate.

Currently, the FIS (http://www.scec.org/FIS) provides users with browser and/or programmatic access to four resources:

1) The SCEC-FAD contains numeric data from published observational studies, and currently holds over 300 publications on 110 faults. Every datum is keyed to the fault, the reference, and the study site. Users may query by fault, by parameters (Fig 1), and by reference. Currently the SCEC-FAD contains a complete record of slip rate data for every southern California fault in the USGS National Quaternary Fault and Fold database (NQFFD), and has been designed to complement the text-based fault compilations of that comprehensive resource (Fig 2). The NQFFD summarizes what is known about a fault. The SCEC-FAD enables users to easily download and query the published numeric data. The SCEC-FAD now contains slip rate data for about 85% of faults in the SCEC Community Fault Model (CFM). Almost all of the remaining 15% lack published activity data, and are not even scheduled for compilation by the NQFFD. This year, FIS manager Sue Perry completed data entry for slip rate data, and designed and implemented browser queries based on references (Fig 3).

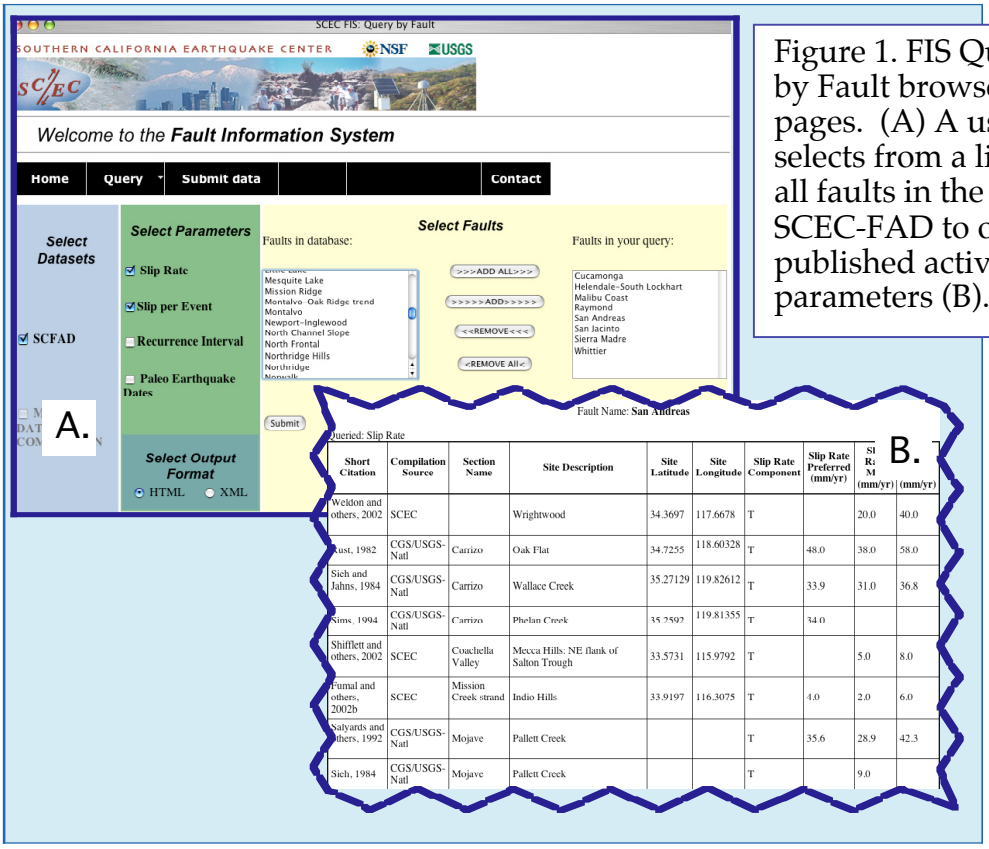

Figure 1. FIS Query by Fault browser pages. (A) A user selects from a list of all faults in the SCEC-FAD to obtain published activity parameters (B).

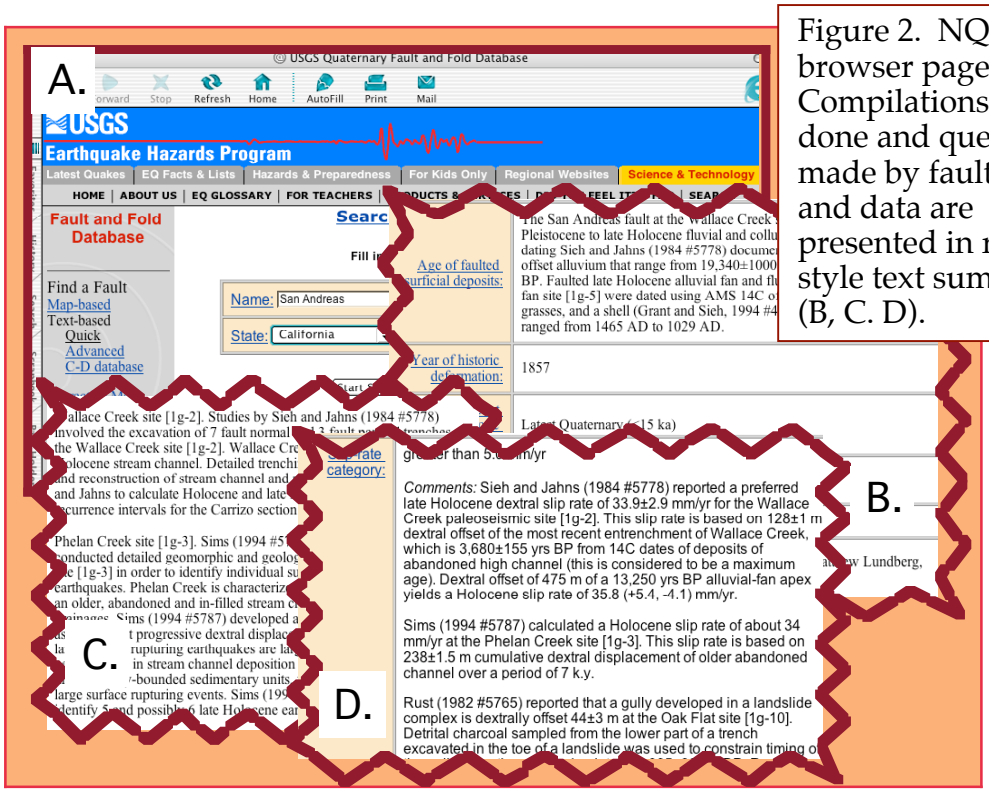

Figure 2. NQFFD browser pages. Compilations are done and queries are made by fault (A), and data are presented in review-style text summaries (B, C. D).

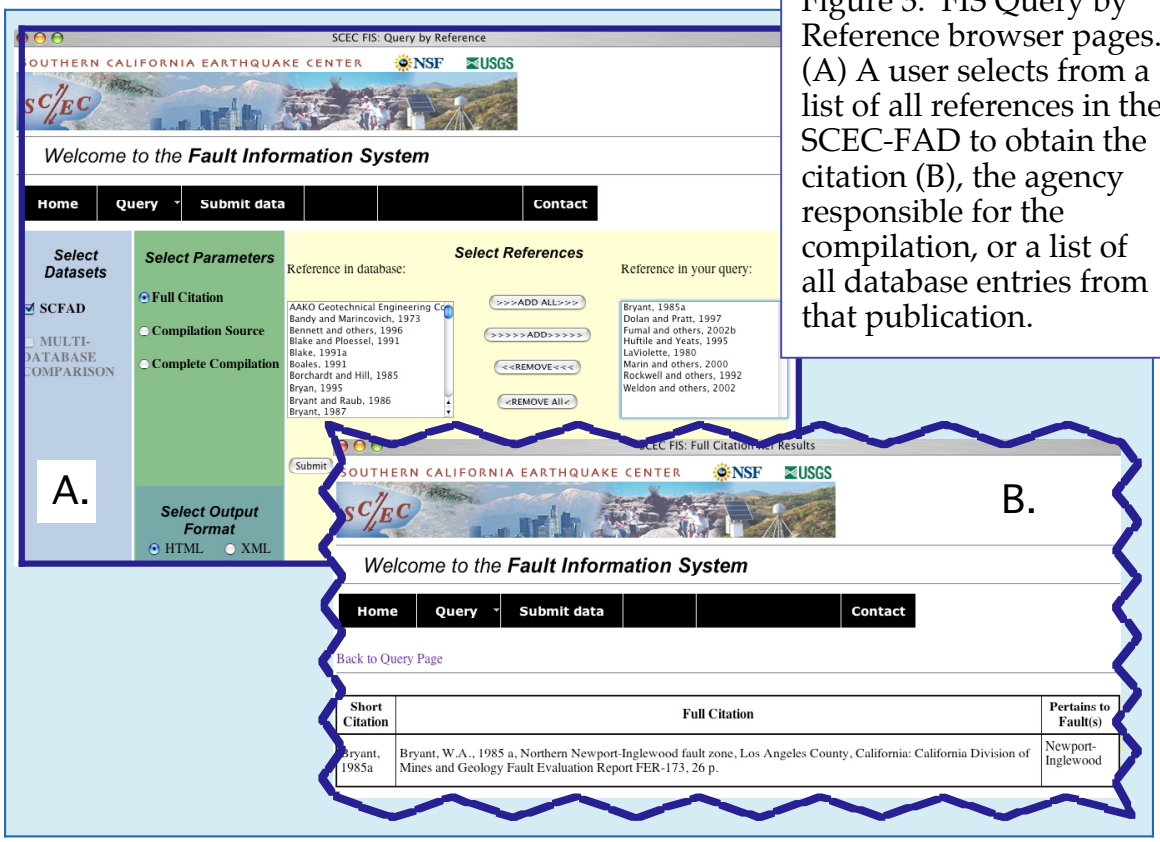

Figure 3. FIS Query by Reference browser pages. (A) A user selects from a list of all references in the SCEC-FAD to obtain the citation (B), the agency responsible for the compilation, or a list of all database entries from that publication.

2) A Web-based data contribution form by which scientists may submit new findings to the database manager for inclusion in the SCEC-FAD. The database manager also uses the form to review and edit submissions, then commit them to the SCEC-FAD.

3) A database to create a "look-up table" that cross-correlates 2-D and 3-D fault models and data across several important datasets (Table I): the SCEC-FAD, the NQFFD, the 1996 and 2002 input parameters to the National Seismic Hazard Maps (PSHA params), the CFM v. 1.2, and the 1994 Jennings fault map. During this SCEC year, Perry designed and implemented this database and manually populated it.

4) An executable distribution of LA3D, the open-source, object oriented, interactive software package created by the SCEC/USEIT (Undergraduate Studies in Earthquake Information Technology) interns (Fig 4). LA3D was first distributed as a free fault viewer to scientists evaluating CFM v. 2. LA3D users may add and view a great variety of geoscience and cultural datasets, including the Community Block Model, earthquake catalogs, and hazard maps like the SCEC Phase III map. During this SCEC year, Perry oversaw the USEIT interns as they embarked into a brave new world of software distribution and support (Fig 5), which has proved orders of magnitude more difficult than software

development as an undergraduate research project. She also served as a liaison with a growing number of SCEC scientists who want to add their data sets to LA3D.

Table I. Sample of Variations in Fault Names and IDs

FAD Fault Name	FAD Section Name	FAD ID	NQFFD ID	NQFFD Section Name	NQFFD Section ID	NQFFD Fault Name	CFM Name	CFM ID
Chino		109	126	Chino	b	Elsinore	chino	13
Elsinore	Coyote Mountain	25	126	Coyote Mountain	f	Elsinore	coyote_mountain_elsinore	18
Glen Ivy		25	126	Glen Ivy	c	Elsinore	glen_ivy	12
Elsinore	Julian	25	126	Julian	e	Elsinore	julian_elsinore	21
Earthquake Valley		23	126	Julian	e	Elsinore	earthquake_valley_elsinore	30
Laguna Salada		117	126	Laguna Salada	g	Elsinore	laguna_salada	32
Elsinore	Temecula	25	126	Temecula	d	Elsinore	temecula_elsinore	20
Whittier		101	126	Whittier	a	Elsinore	whittier	76

Perry also worked on two projects that have been set aside while the USGS-Golden revamps its databases. Early in the year, she designed and implemented a database housing the CGS/USGS 1996 and 2002 PSHA parameters, the input values to the National Seismic Hazard Maps. These important collections of consensus values were previously only available in html tables and Adobe *.pdf* files. However, the USGS-Golden is now in the process of databasing these parameters - which is preferable to housing them in a SCEC database, since the USGS–Golden is responsible for maintaining and updating these datasets. Once the new database is complete, the FIS will access the PSHA parameters from Golden. In addition, the USGS-Golden is changing database management systems for the NQFFD. After the NQFFD is migrated from Filemaker Pro to Oracle, Perry will resume work to access NQFFD text commentaries programmatically during FIS queries.

Finally, Perry collaborated closely with Ned Field and Glenn Biasi to develop a conceptual data model for the next generation of fault database, the Reference Geologic Fault Parameters Database (RefGF). Schema for the SCEC-FAD and the NQFFD pre-date recent paleoseismic advances in the representations of uncertainties in event dates and displacements. The RefGF is being engineered to accommodate these more sophisticated representations, such as probability density functions. In 2005, the conceptual data model will be used to create a logical and then physical data model, and database contributions will be made using a GUI developed by Perry. After that, the RefGF should be ready to become a component in the FIS.

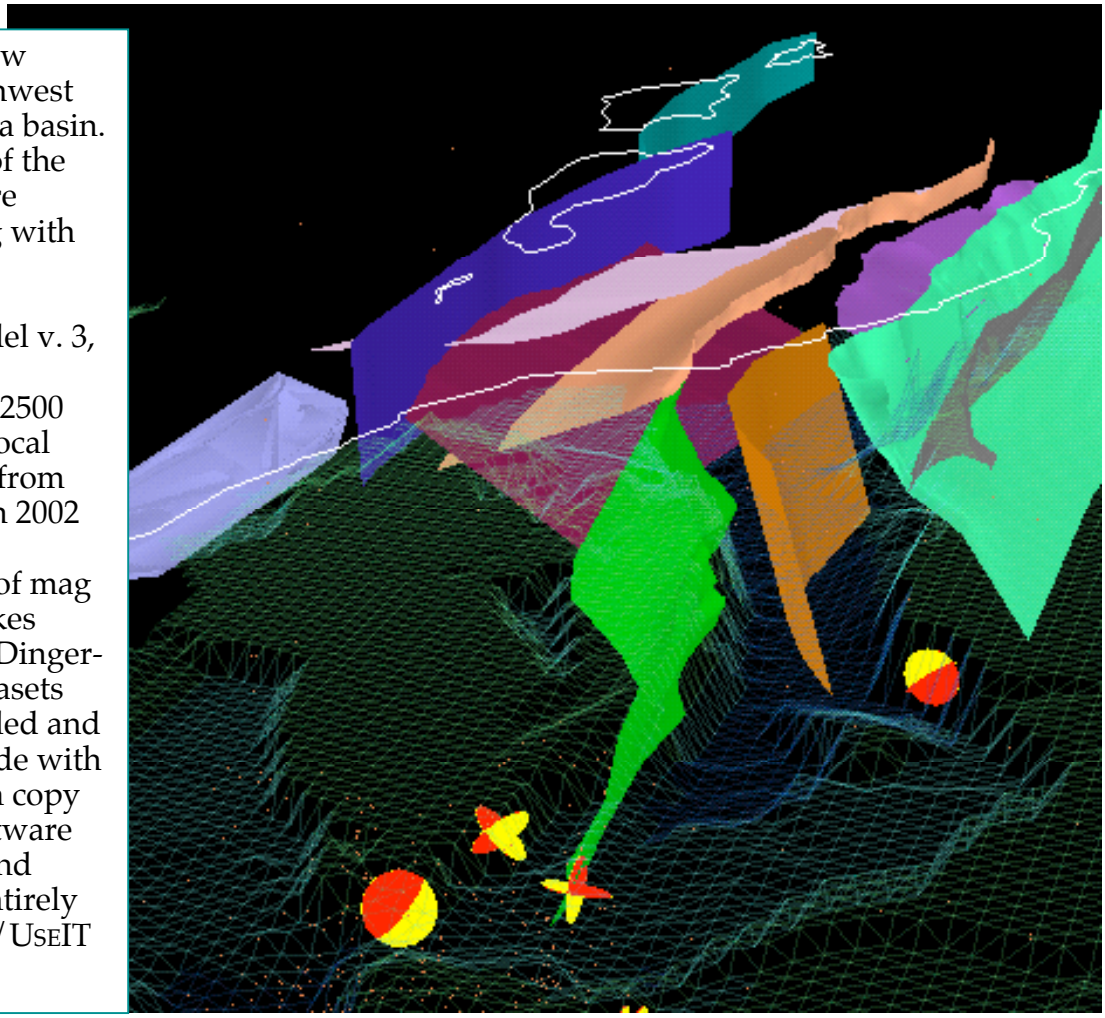

Figure 4. View looking southwest at the Ventura basin. Some faults of the CFM v. 1.2 are shown, along with the SCEC Community Velocity Model v. 3, (basin depth contoured at 2500 m/s), some focal mechanisms from the Hauksson 2002 catalog, and hypocenters of mag 3-4 earthquakes relocated by Dinger-Shearer. Datasets were assembled and this view made with a distribution copy of LA3D, software engineered and developed entirely by the SCEC/USEIT interns.

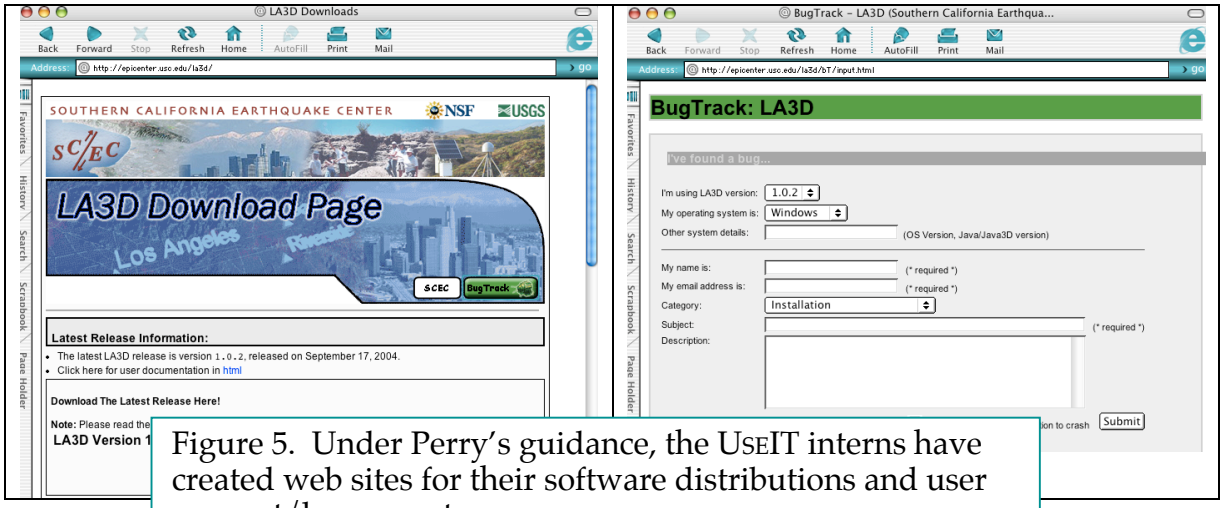

Figure 5. Under Perry's guidance, the USEIT interns have created web sites for their software distributions and user support/bug reports.