# ELECTRONIC SEISMOLOGIST

July/August 2004

Thomas J. Owens
Department of Geological Sciences
University of South Carolina
Columbia, SC 29208
Ph: 803-777-4530
Fx: 803-777-0906
Em: owens@sc.edu

## THE FRUITS OF FISSURES RIPEN

It has been 7 years since the original FISSURES Workshop (Malone, 1997). FISSURES, the ES looked up, stands for "Framework for Integration of Scientific Software for University Research in Earth Sciences". FISSURES was the blueprint for a revolution in seismological software. As luck would have it, back in 1997, no one (with the possible exception of the ES) was in the mood for a revolution. So it was that the revolution turned into a quiet effort by the IRIS Data Management System to create a framework for a new generation of data request mechanisms. FISSURES, an acronym tainted by the rantings of a small band of revolutionaries who will remain nameless lest one of them lose a cushy columnist position with SRL, was reincarnated as DHI (Data Handling Interface; Ahern, 2001a,b). And, good things began to happen. Today, there 5 standalone applications that access data from DHI-enabled data centers (**www.iris.edu/DHI**). Two (GEE and VASE) are data-viewers, targeted at different audiences, that actually do quite a bit more than view data. Two (JEvalResp and JPlotResp) are the DHI-equivalents of popular utilities to access and view instrument responses. MATLAB users can now draw data directly into their applications via DHI using the FMI (FISSURES-MATLAB Interface). And, the ES's personal favorite, SOD allows data to land on your desktop and get funneled directly to your own processing codes automatically after an earthquake of interest to you. SOD is the application that the ES envisioned when the FISSURES was being formulated … something that would eliminate the need to make individual data requests when individual earthquakes occur as well as the awkward task of manually ftp-ing and then organizing data from SEED files. If someone knows what kind of data they are looking for, they should be able to tell the IRIS DMC or other data centers just once and automatically get that data when it arrives. Right? Well, thanks to a dedicated group of programmers in various locations around the globe, that simple need is now close to being a reality. The ES is grossly biased in his excitement about the potential that SOD has to streamline our data handling grunt work. Give it a try! Maybe the ES is right this time.

# SOD – STANDING ORDER FOR DATA

Thomas J. Owens, H. Philip Crotwell, Charles Groves, and Phillip Oliver-Paul
Department of Geological Sciences
University of South Carolina
Columbia, SC 29208
803-777-4530
803-777-0906
owens@seis.sc.edu

## Introduction

Access to digital seismological data from around the world has become increasingly rapid and easy with the expansion of data holdings at international, national, and regional data centers. Nonetheless, the actual task of getting the data from the data center to your own machine and into a form that is ready for analysis has always been a task that required considerable human interaction. That task just became much easier when accessing data from data centers that support the IRIS/FISSURES Data Handling Interface (DHI) protocols (Ahern, 2001a,b). With the release of SOD2.0 (Standing Order for Data; http://www.seis.sc.edu/SOD), seismologists can now configure complex data requests that not only reach back into existing archives, but also remain in effect into the future. For example, a seismologist involved in a PASSCAL experiment can configure SOD to request data from permanent stations in his/her study region for all significant earthquakes that occur during the experiment. SOD will monitor global seismicity and request data for appropriate event-station pairs that will be delivered in SAC format to the seismologist's home machine for the duration of the experiment. Users can configure SOD to undertake a number of simple pre-processing tasks on the arriving data to prepare it for further processing or even build full analysis codes for inclusion in the SOD processing scheme.

SOD is the core application behind two long-term automated processing efforts just getting underway at our institution. First, in conjunction with IRIS Education and Outreach and the DLESE Program Center, we are using SOD to automatically analyze global seismograms in order to take only the highest quality seismograms and (eventually) populate a website for K-16 education purposes. Second, we have developed a SOD module that calculates receiver functions and estimates bulk crustal properties for use on EarthScope data. Thus, we have tested SOD extensively for our particular interests. We hope others will give it a try and let us know how it works (and what it needs) for other applications.

## The Roots of SOD

SOD, like a growing number of clients, accesses distributed data centers using the IRIS FISSURES/DHI servers (http://www.iris.washington.edu/DHI/). Currently, three servers of varying sizes offer DHI-based access to their data holdings. The IRIS DMC (www.iris.washington.edu) is the largest. The Northern California Earthquake Data Center (quake.geo.berkeley.edu) has begun offering this access and the South Carolina

Earth Physics Project (SCEPP; www.seis.sc.edu/scepp) was the original DHI-enabled data center offering data from high-school based instruments throughout South Carolina. Each of these centers offers waveform information, event information, and station information through DHI servers and can be seamlessly accessed through DHI clients such as SOD. At this point, all of these clients are Java-based.

SOD itself is configured via an XML file in which the user can specify a number of criteria for selecting desired events and stations. Internally, SOD is organized in what we call "arms". Event-based information is gathered from DHI Event Servers and flows through the Event Arm. Similarly, Station-based information is gathered from DHI Network Servers and flows through the Network Arm. Decisions about the suitability of events or stations are made in "subsetters". For example, if only large earthquakes are of interest, then a Magnitude Range subsetter would be used. In XML, the subsetter in the event arm would look like this:

```
<magnitudeRange>
    <magType>mb</magType>
    <min>5.5</min>
</magnitudeRange>
```

More complex subsetters can be constructed using AND, OR, NOT, and XOR logicals.

Subsetters that require only event information are included in the Event Arm and subsetters that require only station information are included in the Network Arm. The Waveform Arm is where qualified events and stations are combined. Subsetters that apply criteria that require information about both the event and the station, such as distance range criteria, reside in the Waveform Arm.

In addition to the event-station based search criteria, SOD has a number of configuration parameters that define your preferences on the longevity of the run, the type of files to output, and the type of status information to maintain. These parameters and a history of the events-station pairs received to date are kept in a local database so that SOD can be restarted after a local system failure, if necessary. For example, you can request that existing events be re-opened periodically in case new data has arrived at a data center after the original request. Data latency in global real-time systems is normally small, but network outages occasionally delay data for hours or days. In addition, event notifications that trigger SOD to request waveforms may actually arrive before surface waves can propagate to remote stations. In either of these cases, it might be desirable to re-open events to check for new data. The XML for this looks like:

```
<property>
     <name>sod.start.ReopenEvents</name>
     <value>true</value>
</property>
```

Other tags in the Property section define the preferred interval for checking for new data from existing events.

One configurable property of SOD is the option of creating HTML status pages to monitor your run as it harvests and processes seismograms. By enabling this feature and placing the status files where they are accessible to a web server, SOD runs can be easily

monitored. The main page of a SOD run (Figure 1a) just summarizes the events recovered to date with some general information about the run. The Events page (Figure 1b) provides a detailed, clickable, listing of the individual events with some information about how much data is available, pending, as well as how much data failed at least one of the criteria set for the run. For each event, a summary of the recovered stations is available (Figure 2a) and for each station for which there is data, a summary and quick plot of the data is linked to this list (Figure 2b).

**A GUI for SOD**

When SOD v1.0 was completed in June of 2002, it could go back in time in a data archive and request data with capabilities similar to the popular WEED utility. However, at that time, the persistence was not built in to allow SOD to perform as a true Standing Order for Data. In addition, no GUI was available, which some felt would decrease the potential user base. Now, with the release of SOD 2.0, both of those limitations have been addressed. SOD is a very flexible, highly configurable package. At this point, you can only get access to all of the features of SOD if you are willing to edit the SOD XML using a simple text editor. However, we have developed a simple GUI that uses templates to allow users to change predefined fields and subsetters. For example, we have created an XML file for SOD that simulates the features of WEED (Figure 3). By running the SOD GUI with the WEED configuration file, users can change all of the search parameters that they could change in the traditional WEED application, but they are limited to using only those features, not the whole SOD toolkit. This approach allows for GUI-based manipulation of the SOD configuration, but the changes are limited to the parameters defined in a particular XML file. Nonetheless, this template-based method can meet the needs of a large percentage of seismologists. The SOD GUI is continuing to evolve and may reach full-functionality in the near future. For now, if you need more configurability, dig into the XML.

**Output from SOD**

You can request several different types of output from SOD. Most users will likely want to request data delivered to their machine in SAC format. In the Waveform Arm, this is done in XML as follows:

```
<saveSeismogramToFile>
    <fileType>sac</fileType>
    <dataDirectory>SOD_Data</dataDirectory>
    <eventDirLabel>Event_<originTime>yyyy_DDD_HH_mm_ss</originTime>
    </eventDirLabel>
</saveSeismogramToFile>
```

The above XML generates a directory called SOD_Data and subdirectories for each individual event with a prefix of Event_ and a full name based on the origin time of the event. Optionally, SOD can output MSEED files that contain compressed data, but are currently readable by far fewer processing packages. Finally, if you are requesting

restricted data (such as PASSCAL datasets that are still proprietary), or you really want SEED data for some reason, or you are simply just not ready to break from your long held processing routine, you can elect to have SOD generate good old-fashioned BREQFAST files:

```
<breqFastAvailableData>
      <dataDirectory>XJ97_breqfast</dataDirectory>
      <label>
         Event_
         <originTime>yyyy_DDD_HH_mm_ss</originTime>
      </label>
      <name>Thomas J Owens</name>
      <inst>Univ. of South Carolina</inst>
      <mail>Dept. of Geol. Sci., USC, 29208</mail>
      <email>owens@seis.sc.edu</email>
      <phone>803-777-4530</phone>
      <fax>none</fax>
      <media>Electronic</media>
      <altmedia1>Electronic</altmedia1>
      <altmedia2>Electronic</altmedia2>
      <quality>b</quality>
</breqFastAvailableData>
```

In this case, the directory X97_breqfast will be filled with BREQFAST requests with names like Event_1997_230_23_13_00.breqfast that can then be emailed to IRIS. Its very '90s, but it is necessary in some cases.

By default when outputting SAC or MSEED format files, SOD also generates a DSML (DataSet Markup Language) file that allows the seismograms received to be easily viewable in the Global Earthquake Explorer (GEE: http://www.seis.sc.edu/gee).

**Preprocessing and Automated Processing in SOD**

While outputting seismograms as SAC files is useful, it is very common for there to be additional preprocessing steps that are always applied before the data is analyzed. SOD is capable of handling many of these, applying them to the data before it is saved to a file. The processing section of SOD acts as a pipeline, with the output seismograms from one processor being the input to the next. SOD has a small but growing number of these processors implemented including cutting, removing the mean and the trend, filtering, applying the response gain as well as an external processor that allows the user to create new custom processors. There is also a fork processor that allows more than one processing sequence to be applied to the data. It is conceivable that some automated processing systems would not even save the data, but instead save just the results, knowing that the data can be retrieved easily later if needed.

Here is the XML for a processing sequence that would be at the end of the Waveform Arm. This example first cuts to a window around the predicted P arrival, removes the mean and trend, applies a taper and then a band pass filter, all before saving the seismograms to files.

```
<phaseCut>
       <beginPhase>P</beginPhase>
       <beginOffset>
           <unit>SECOND</unit>
           <value>-120</value>
       </beginOffset>
       <endPhase>P</endPhase>
       <endOffset>
           <unit>SECOND</unit>
           <value>360</value>
       </endOffset>
 </phaseCut>
 <rMean/>
 <rTrend/>
 <taper/>

 <filter>

        <lowFreqCorner>
        <value>50</value><unit>SECOND</unit>
        </lowFreqCorner>

        <highFreqCorner>
        <value>5</value><unit>HERTZ</unit>
        </highFreqCorner>

        <numPoles>2</numPoles>
        <filterType>NONCAUSAL</filterType>
   </filter>
   <saveSeismogramToFile>
       <fileType>sac</fileType>
       <dataDirectory>SOD_Data</dataDirectory>
        <eventDirLabel>Event_
             <originTime>yyyy_DDD_HH_mm_ss</originTime>
        </eventDirLabel>
   </saveSeismogramToFile>
```

At this point, all processors directly interacting with SOD must be written in Java. However, it is easy to envision a processor that simply generated a valid input file for an existing code and then launched that code within SOD. Future plans for enhancements to SOD include creation of Tcl, Python, and perhaps even SAC macro script processors to facilitate execution of legacy scripts within SOD.

## Downloading and using SOD

SOD can be found at http://www.seis.sc.edu/sod. SOD will run on any platform that has a recent version of Java (Windows, Mac OSX, Sun Solaris, Linux). On the SOD web page, there is a tutorial that explains the XML configuration options in more detail, links to detailed documentation, and download/install instructions. Enjoy!

## References

Ahern, T (2001a). Data Handling Infrastructure at the IRIS DMC. IRIS DMC Newsletter, Vol3-1, 3. (http://www.iris.edu/news/newsletter/vol3no1/page3.htm)

Ahern, T (2001b). What happened to FISSURES? -or- Exactly what is the Data Handling Interface?, Vol3-3, 3. (http://www.iris.edu/news/newsletter/vol3no3/page3.htm)

Malone, S. (1997). The Electronic Seismologist goes to FISSURES, *Seis. Res. Letters,* 68, 489-492.

## Figure Captions

**Figure 1.** A. (top) SOD Summary Status Page. B. (bottom) SOD Event Summary Page.

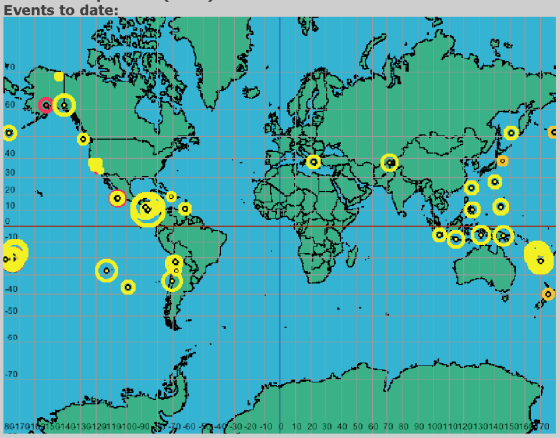
**Figure 2.** A. (top) SOD Event Page. Blue triangles on the map are stations for which SOD has recovered data. Station list appears below the map. B. (Bottom) SOD Station-Event page showing location and distance information and images of the recovered seismograms.

**Figure 3.** Screen shot of the Event tab on the SOD GUI Editor for the WEED.xml configuration file illustrating configurable fields.

REV Test
Configuration
Events
Networks
About
Errors
Wily

## REV Test Information

**Started on:** 4/30/2004 21:13 GMT
**Amount of SAC data written:** 1358228132 bytes
**Last event received:** TONGA ISLANDS
**Number waveforms processed:** 706
**Waveforms per hour(WPH):** 13.02925524564783
**Events to date:**



Page last generated:
5/3/2004 3:24:42.618 GMT

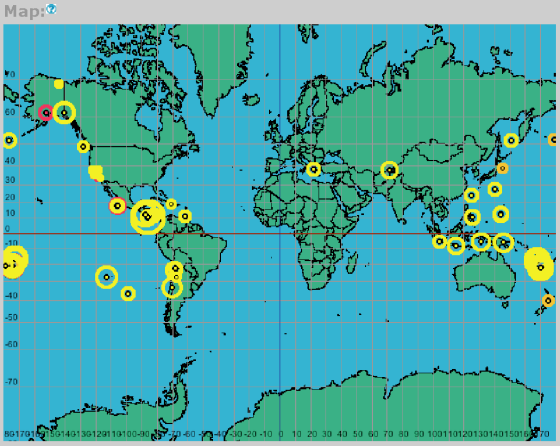
REV Test
Configuration
Events
Networks
About
Errors
Wily

## Events

**Status:**

Waiting for the wave form queue to process some events before getting more events

**Map:**



**Successful:**

| Region | Time (GMT) | Magnitude | Depth | # Processed Channels | # Rejected Channels | # Retry Channels |
|---|---|---|---|---|---|---|
| TONGA ISLANDS | 05/02/2004 05:11:30 | 5.6 M | 74.9 km | 72 | 34 | 222 |
| OFF COAST OF JALISCO, MEXICO | 05/02/2004 04:19:27 | 4.9 M | 10.0 km | 0 | 0 | 0 |
| SOUTHERN ALASKA | 05/01/2004 14:24:20 | 4.7 M | 170.0 km | 21 | 326 | 84 |
| CENTRAL CALIFORNIA | 05/01/2004 02:52:58 | 3.6 M | 3.6 km | 3 | 478 | 18 |
| SOUTHERN CALIFORNIA | 04/30/2004 18:17:17 | 3.2 M | 12.4 km | 6 | 484 | 18 |
| AFGHANISTAN- | 04/30/ | | 56.4 | | | |

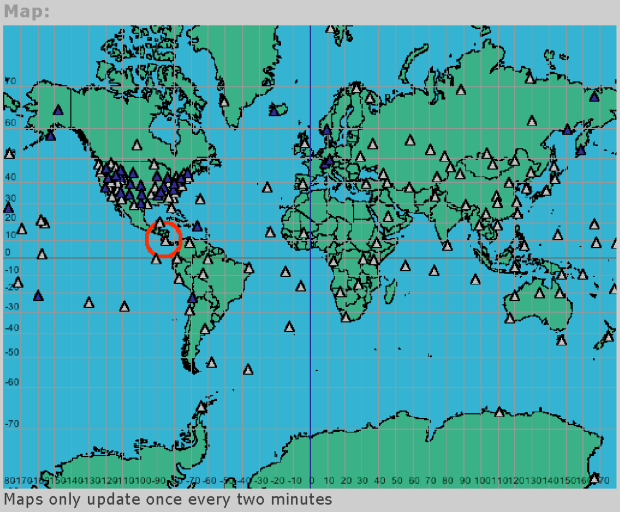
**SRL-Electronic Seismologist, Jul/Aug 2004, Owens et al, Figure 1**

Events > OFF COAST OF COSTA RICA 04/29/2004 00:57:25

REV Test
Configuration
Events
Networks
About
Errors
Wily

## Station information for this event

### Map:



Maps only update once every two minutes

### Stations:

| Id | Begin Time | End Time | # Successful Channels | # Retry Channels | # Reject Channels |
|---|---|---|---|---|---|
| II.BFO | 05/29/1996 | - | 3 | 0 | 0 |
| II.BORG | 07/31/1994 | - | 3 | 0 | 0 |
| II.KDAK | 06/09/1997 | - | 3 | 0 | 0 |
| II.PFO | 10/24/1986 | - | 3 | 0 | 0 |
| IU.ANMO | 11/19/2002 | - | 3 | 0 | 0 |
| IU.BILL | 05/18/1997 | - | 3 | 0 | 0 |
| IU.COLA | 06/14/1996 | - | 3 | 0 | 0 |
| IU.GRFO | 01/26/1994 | - | 3 | 0 | 0 |
| IU.HKT | 07/11/1995 | - | 3 | 0 | 0 |
| IU.HRV | 01/01/1988 | - | 3 | 0 | 0 |

09/17/

Events > OFF COAST OF COSTA RICA 04/28/2004 20:57:25 -0400 > IU.HKT-Hockley, Texas

REV Test
Configuration
Events
Networks
About
Errors
Wily

## Seismograms for this Event

### Event Info

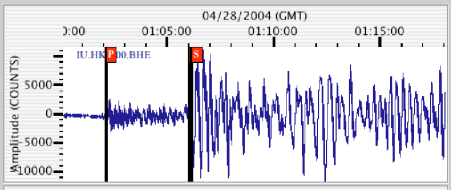
| | |
|---|---|
| Region | OFF COAST OF COSTA RICA |
| Time | 04/28/2004 20:57:25 -0400 |
| Magnitude | 6.1 M, |
| Depth | 40.6 km |
| Latitude | 10.69 |
| Longitude | -86.02 |
| Distance | 21.247 deg |
| Azimuth | 335.899 deg |
| Back Azimuth | 152.452 deg |

### Station Info

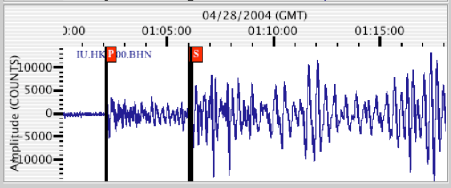
| | |
|---|---|
| Id | HKT |
| Name | Hockley, Texas |
| Latitude | 29.9618 |
| Longitude | -95.8384 |
| Begin Time | 07/10/1995 20:00:00 -0400 |
| End Time | 12/31/2498 19:00:00 -0500 |

### Seismograms

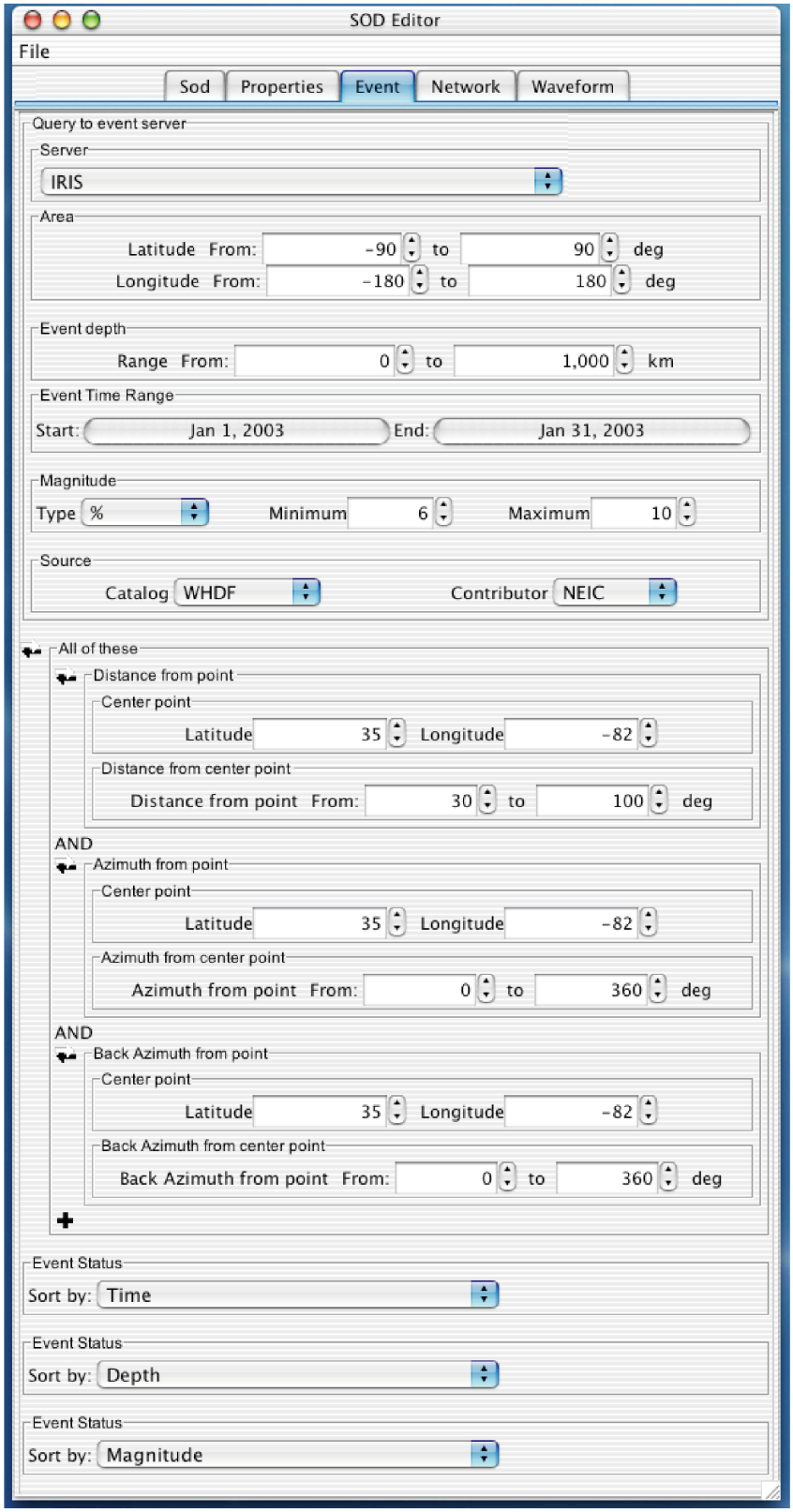
IU.HKT.00.BHE
43.17257
AD 2004.04.29
00:57:29.810
GMT




IU.HKT.00.BHN




**SRL-Electronic Seismologist, Jul/Aug 2004, Owens et al, Figure 2**




**SRL-Electronic Seismologist, Jul/Aug 2004, Owens et al, Figure 3**