# Collaboration to Populate a SCEC Community Paleoseismic Model


Principal Investigators:

Thomas Rockwell, San Diego State University, San Diego

Glenn Biasi, U.S. Geological Survey, Pasadena


Proposal Category: Data Gathering and Products

SCEC Science Priorities:

P5c, "patterns of earthquake occurrence, …paleoseismic observations"

P5d: "test supercycles"

P1: "How are faults loaded across temporal and spatial scales?"


**Abstract**

The project objective is to develop and construct a paleoseismic database that will eventually contain all dated paleoearthquakes in California. The first-year effort focused on two objectives: 1) design of the database and input structure, and 2) initial population of the database with unpublished data from the southern San Jacinto fault. This database will serve the large community that is interested in long-term patterns of seismicity in California, and potentially as a test to computer simulations of long term seismicity. The data structure has been constructed in XML for flexibility and ready incorporation into a relational database. The text-based XML format has been used as a preliminary input structure also, and most southern San Jacinto paleoseismic sites have been entered. The structure includes site name and location, timing and uncertainty of past events (or cumulative displacement over multiple events), estimate of separation between events, with uncertainty, references or locations of primary data, trench logs, raw radiocarbon data and, if available, the OxCal models and associated PDFs of the paleo-event ages.


**Project Objectives**

The objective of this project is to develop and populate a community paleoseismic database that can be compared to physics-based computer simulations of seismicity in California. This proposal addresses the SCEC 5 science plan by developing a complete database of currently available paleoseismic data to "Assess the limitations of long-term earthquake rupture forecasts by combining patterns of earthquake occurrence and strain accumulation with neotectonic and paleoseismic observations of the last millennium" (priority P5.c.), and "Test the hypothesis that “seismic supercycles’’ seen in earthquake simulators actually exist in nature and explore the implications for earthquake predictability.” (priority P5.d.). The development and vetting of a Community Paleoseismic Model (vetted database) will facilitate both of these priorities, as well as illuminate areas and faults where additional work is critical. It will also address priority P1.a which states “Refine the geological slip rates on faults in southern California…”, as we will include the full range of slip rate uncertainties and test where some data are incompatible with the paleoseismic record or other slip rate data.

## Methodology

Task 1 was to develop a template for data input and begin to populate it with published and unpublished paleoseismic data from the southern San Andreas fault system. Figure 1 shows the form of the template. Task 2 was to initially populate the database with data from sites in the southern San Andreas system and test the entry system (debug).

## Results – Work Accomplished in Year 1

To date, we have completed a first cut template that incorporates all of the proposed items to compile, including site name and location, paleoseismic event timing and uncertainty, displacement in past events (or cumulative displacement over multiple events), estimate of separation between events, with uncertainty, references or locations of primary data, trench logs, and raw radiocarbon data. We have begun populating the database by filling in a template for each site, with the initial focus on many of the unpublished sites along the southern San Jacinto fault (many are in theses from SDSU and would be otherwise difficult to find). Data have been compiled from the majority of paleoseismic sites in the southern San Andreas system, and many have been incorporated using the XML format. We have also drafted the interface through which data will be accessed by the community (Figure 2), although this remains to be up and running.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="cpm_db.xsl"?>
<!-- v1:  add NEvents, CompleteAfter, CompleteBefore, RecurInterval,
RIMethod, CoV, DatingModel -->
<!--      add example of alternate event pdf                     -->
<!--      add References block and support for unpublished stuff -->
<!--      Add Logs block for logs -->
<!--      Add event mean date -->

<PaleoSites>
<PaleoSite> <!-- start of a new site -->
   <SiteName> Superstition Mountain North Shoreline </SiteName>
<!-- in this file, this oddly demarked bit is a comment -->
   <SiteCode> TP </SiteCode>
   <Lat> 32.997316 </Lat>
   <Lon> -115.943607 </Lon>
   <Fault> Superstition Mountain </Fault>
   <!-- Add later - Section, Subsection (UCERF3 sense), Orientation
(from CFM), Slip Rate, Slip type -->
   <MREDate> 1715 </MREDate>
   <MREDateUncertainty> 10 </MREDateUncertainty>

   <EventDates>

        <Event>

            <EvName> TP1 </EvName>  <!-- start with letter, unique for
site, no spaces -->

            <EvMeanDate> 1715 </EvMeanDate>  <!-- nominal mean event
date -->

            <EvDatePDF> TP1.txt </EvDatePDF>   <!-- plan external file
named TP1.txt, strict GPB format -->

            <EvDatePDFRebinned> TP1Rebinned.txt </EvDatePDFRebinned>
<!-- use GPB code for this -->
                 <!-- idea is a Date PDF, resampled on (say) 10,000
small parts -->

            <EvConfidence> 80 </EvConfidence>
                 <!-- scale 1-100 evidence Confidence, where
available, Scharer et al. 2007 Rank -->

            <EvLogs>
                <EvLog>TPLog3, x= 13.2, z=1.5</EvLog>
                   <!-- unique log labels, e.g. TPLog3, with
stationing point -->
                   <!-- each log listed in Log section -->
                <EvLog>TPLog1, x= 4.2, z=1.25</EvLog>  <!-- if

desired, second location -->
                <EvLog>TPLog5, x= 5.1, z=0.25</EvLog>   <!-- if
desired, third location -->
            </EvLogs>

            <EvComment> optional, short, no reserved characters </
EvComment>

        </Event>

   </EventDates>

   <EventDisplacements>
       <EvDisplacement>
           <EvName> TP1 </EvName>   <!-- example name by event when
single event displ is known -->
           <EvDispl>2 </EvDispl>
           <EvMin> 0.5 </EvMin>
           <EvMax> 2.5 </EvMax>
       </EvDisplacement>

    <TimeSeparations>
       <EvPairs>
           <EvNames> TP1 TP2 </EvNames>
           <TSeparation> 50 </TSeparation>   <!-- years, 50 if a soil
formed, say -->
           <TSepMin>     20 </TSepMin>
           <TSepMax>    100 </TSepMax>
       </EvPairs>

   </TimeSeparations>

    <EventHistories>
        <EventHistory>
           <EvHistoryNum> 1 </EvHistoryNum>
           <EvNames> TP5 TP4 TP3 TP2 TP1 </EvNames>  <!-- all events
considered good -->
           <NEvents> 5 </NEvents>  <!-- Number of events; convenience
entry -->
           <CompleteAfter> 950 </CompleteAfter>  <!-- event chronology
complete after this.  will predate oldest event -->
           <CompleteBefore> 2017 </CompleteBefore>  <!-- event
chronology up to this.  most often date of entry -->
           <RecurInterval> 222 </RecurInterval>  <!-- nominal
recurrence interval -->
           <RIMethod> PoissonWithOpenInterval </RIMethod>
           <CoV> 0.65 </CoV>  <!-- arithmetic CoV -->
           <DatingModel> TP5Events.oxcal </DatingModel>  <!-- OxCal or
other -->
        </EventHistory>

    </EventHistories>

    <References>
        <Ref>
           <RefType> Journal </RefType>
           <SCECid> 99999 </SCECid>
           <Citation> Fumal, T., M. J. Rymer, and G. G. Seitz (2002).
Timing of large earthquakes since A.D. 800 on the Mission Creek strand
of the San Andreas fault zone at Thousand Palms Oasis, near Palm
Springs, California, Bull. Seismol. Soc. Am., 92, 2841-2860.
           </Citation>
    </References>
    <Logs>
        <Log>
            <LogName> TPLog1 </LogName>
            <LogFileName> TPLog1.jpg </LogFileName>
     </Logs>

</PaleoSite> <!-- end of site -->
</PaleoSites>  <!-- ends xml file -->
```

**Figure 1.** *XML template for entry of paleoseismic data.*

## Preliminary Issues to be Solved

The template for inputting data is bulky and prone to minor errors, and as we have not actually tested it in an on-line form, it is uncertain as to whether it will work, as promised.  This issue is exacerbated by a lack of an on-line GUI through which data can be entered more directly. We have been discussing this issue with other SCEC scientists (Ned Field, Tom Jordan) in search of the best path forward. The volume of data (currently over 100 gb of files) along with the time-consumptive nature of current data input requires a slight change in direction to make this project a success.  Development of a GUI and decisions as to the best way to host and store data are a primary foci of current work, in addition to data entry.

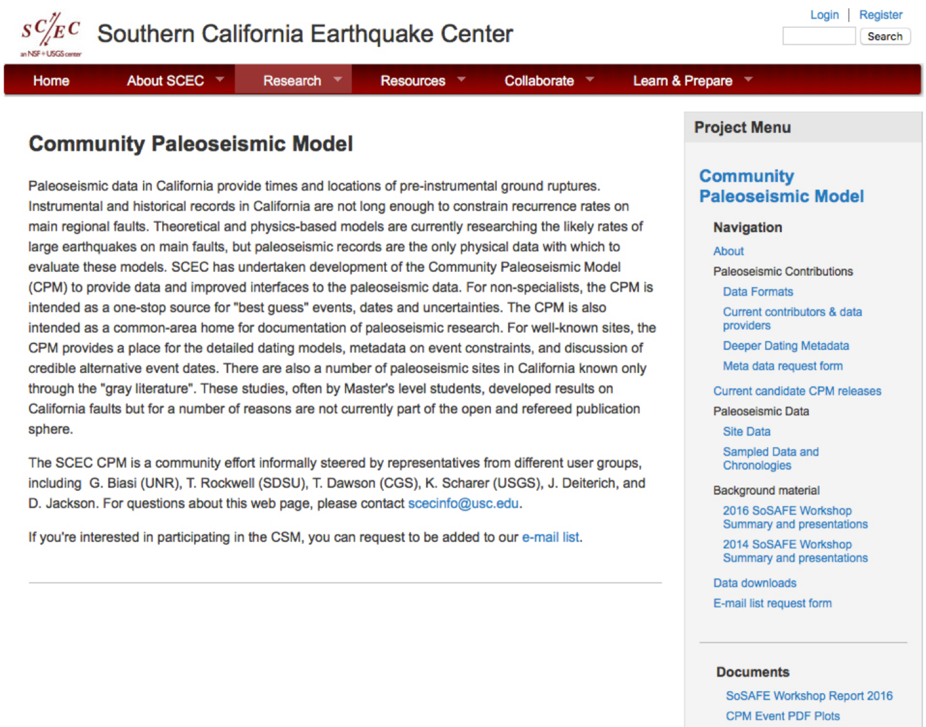


**Figure 2.** *Mock-up of the web site that will host the paleoseismic data.*