A Unified Source Model for the 1906 San Francisco Earthquake

by Seok Goo Song, Gregory C. Beroza, and Paul Segall

Abstract  We reconcile two previously discordant source models of the 1906 San Francisco earthquake and obtain a model that satisfies both triangulation and seismic data by allowing the rupture velocity to exceed the shear-wave velocity. Employing a projection method to remove the dependence on initial station positions allowed us to make use of a more stable triangulation network, including nonrepeated angle observations along the northern San Andreas fault. This strengthens the case for significant slip over the entire northern segment of the San Andreas fault from San Juan Bautista to Cape Mendocino during the 1906 earthquake. We also found that the teleseismic body-wave data can be reconciled with the geodetically derived slip model by allowing supershear rupture. This resolves a longstanding conflict between the two previous slip models (geodetic and seismic) of this earthquake. Supershear rupture has long been recognized as a theoretical possibility for strike-slip faulting, and it has been observed in several recent large strike-slip earthquakes, which raises the prospect that it might be typical for such events. Supershear rupture leads to substantially different strong ground motion, and as a result, may need to be taken into account when developing ground motion prediction relations for large strike-slip earthquakes. Our final slip model has a seismic moment of $7.9 \times 10^{20}$ N m, which corresponds to a moment magnitude of $M_w 7.9$.

Online Material:  A digital version of geodetic displacements and slip distribution.

Introduction

The 1906 San Francisco earthquake ruptured the northern segment of the San Andreas fault at the dawn of the twentieth century and is perhaps the single most important earthquake in the history of earthquake science. Despite the importance of the earthquake, the two most recently published source models, one based on geodetic data (Thatcher et al., 1997) and the other based on seismic data (Wald et al., 1993), differ substantially from one another, particularly in the total rupture length. The geodetic slip model maps slip from San Juan Bautista to Cape Mendocino (~500 km rupture length), whereas the seismic model finds almost no slip north of Point Arena (~300 km rupture length). We find that these two models can be reconciled if the rupture velocity exceeded the shear-wave velocity of the Earth’s crust north of San Francisco.

While it is often assumed that earthquake rupture velocity does not exceed the Rayleigh wave velocity, theoretical studies indicate that in-plane rupture can propagate at infrasonic speeds, that is, between the $S$-wave and the $P$-wave velocities (Burridge, 1973; Andrews, 1976). Recent large strike-slip earthquakes—the 1999 Izmit, Turkey, the 2001 Kunlunshan, Tibet, and the 2002 Denali events—have all exhibited characteristics of supershear rupture (Bouchon et al., 2001; Bouchon and Vallée, 2003; Dunham and Archuleta, 2004; Ellsworth et al., 2004). Recent laboratory experiments (Rosakis et al., 1999; Xia et al., 2004) confirm and extend previous theoretical work on supershear rupture propagation. Thus, it seems plausible that supershear rupture could have occurred during the 1906 San Francisco earthquake.

Geodetic Analysis

The differences between existing source models are primarily north of Point Arena, where the San Andreas fault runs offshore, rendering direct observation of surface rupture impossible. There were 4.9 m of slip measured at Alder Creek, the northernmost observation of unambiguous faulting in the 1906 earthquake (Lawson, 1908). The same report found offset at Seal Cove, farther to the north, where the fault comes on shore again, but raised the possibility that this offset might have been due to landslides rather than tectonic faulting.

Only repeated triangulation observations were used to estimate slip in the previous geodetic study (Thatcher et al., 1997), resulting in a weakly connected network (Thatcher et al., 1997), resulting in a weakly connected network, especially north of Point Arena, and hence large uncertainties in the inferred displacements. In this study, we utilize all of the available triangulation measurements in the northern region,
employing a projection method to remove the dependence on initial station positions (Yu and Segall, 1996). This allows us to strengthen this part of the network using nonrepeated observations (Fig. 1). It can be shown that the Yu and Segall (1996) method reduces to the standard approach using only repeated angle measurements when all of the measurements are repeated before and after the earthquake (see Appendix). Our data set in this region contains 60 pre-1906 and 172 post-1906 angle observations, respectively, compared with the 37 angle changes used in the previous study. We used the same data set, consisting of repeated angles, in the southern region (see Fig. 2).

The geodetic displacements along the entire rupture trace of the 1906 earthquake were estimated (Fig. 2 and 3, Table 1 in the electronic edition of BSSA) using a model coordinate solution (Segall and Matthews, 1988) to constrain rigid body motions and scale changes. The result shows a displacement field characteristic of the coseismic faulting as far north as Cape Mendocino (Fig. 2a). Two stations near the fault trace immediately south of Cape Mendocino show large displacements parallel to the changing local strike of the fault, which strongly supports fault slip as opposed to land sliding in this area. The magnitude of these displacement vectors indicates that the amount of fault slip that caused them is substantial.

We estimate the coseismic slip distribution by a linear inversion of the triangulation data (Fig. 4 and Table 2 in the electronic edition of BSSA) using the surface trace from the new 3D geologic model constructed by Jachens et al. (2006). A homogeneous elastic half-space was assumed in the forward calculation. An appropriate level of smoothing was determined by cross validation (Matthews and Segall, 1993) and used in the inversion. The idea behind cross validation is that a good model should predict data not used in the estimation. The appropriate smoothing is determined by testing what level of smoothing generates a model that best predicts the unused data. Our model slip varies only in the horizontal direction along the fault trace as shown in Figure 2. Each value along the fault indicates averaged slip on a 10-km-long and 12-km-wide (deep) vertical fault patch. The fault is assumed to be extended vertically with 90° dip. We note that the average slip on each patch trades off with the assumed vertical extent of the fault (12 km in this study) with the result that, for example, models with greater depth extent but smaller averaged slip can fit the data equally well. But the integral of slip in the vertical direction should not change much irrespective of the assumed rupture width within a reasonable range of the seismogenic zone in California. Because of our inability to obtain the original data records, the slip distribution south of Point Arena is constrained only by the same repeated angle observations used in the previous study (Thatcher et al., 1997). The slip north of Point Arena is significantly improved by the use of nonrepeated angle measurements. Our slip on a ~500-km-long rupture successfully fits the triangulation data, both confirming and refining the previous geodetic analysis (Thatcher et al., 1997).

Seismic Analysis

With the mapped surface slip and geodetic data both consistent with the longer fault rupture, the short rupture length inferred from the seismic data stands out. For long-strike-slip events like the 1906 San Francisco earthquake,
the duration of observed waveforms is related to the ratio of rupture length to rupture velocity; however, the duration of the observed teleseismic waves and the fixed sub-Rayleigh rupture velocity assumed previously (Wald et al., 1993) favors a shorter fault rupture. Our hypothesis is that by allowing more flexibility in the rupture velocity, including the potential for supershear rupture, we might fit all data with a single model.

We apply a Bayesian inversion approach coupled with a Monte Carlo sampling method (the Metropolis algorithm) (Mosegaard and Tarantola, 2002; Metropolis et al., 1953). The posterior distribution of the model (slip and rupture velocity) is proportional to the product of a prior distribution and a likelihood function. The likelihood function contains only the seismic data and the geodetic inversion results obtained above were used in the prior distribution in order to stabilize the slip in the inversion, which we believe was relatively well resolved in the linear geodetic inversion. The prior for the rupture velocity is a Gaussian distribution centered at the previously used sub-Rayleigh velocity (2.7 km/sec) with a 0.5 km/sec standard deviation. By assuming subshear rupture velocity in the prior, we ensure that the method will only find supershear rupture if the data require it.

We use waveforms from the 1984 $M$ 6.2 Morgan Hill earthquake as empirical Green’s functions to calculate teleseismic waveforms (Wald et al., 1993). Both the Morgan Hill and the 1906 San Francisco earthquake are vertical strike-slip events and share approximately the same strike ($\sim$N35°W), particularly in the central portion of the 1906 rupture. But the San Andreas fault bends in the northern and southern portion of the study area as shown in Figure 2. The average strike of these segments is about 10°–15° different from that of the Morgan Hill earthquake, such that modest error is introduced by using the same Green’s function for the entire fault trace. The Green’s function is time lagged with elapsed rupture time along the fault trace from the hypocenter, and a linear summation of the time-lagged Green’s function weighted by slip provides synthetic waveforms for the 1906 earthquake.

As discussed in detail by Wald et al. (1993), the age and quality of the seismic data limit the resolution of source characteristics for the 1906 earthquake. The uncertain instrument response, limited accuracy of the available Green’s func-
tions, particularly when applied to the northernmost end of
the fault, and complex wave propagation effects for the $SV$
component, all contribute to the data residuals. A waveform
inversion without the analysis of very-long-period data
(> 40 sec) may underestimate the rupture area of large earth-
quakes as observed in the Sumatra earthquake (Stein and
Okal, 2005). Despite the limited data coverage quality, for-
ward modeling indicates that the duration and amplitude of
the teleseismic waves can constrain the overall duration of
the rupture, and hence the average rupture velocity when
combined with the fault length determined from the geodetic
data.

Supershear Rupture and Combined Slip Model

Sensitivity tests indicate that, because of the geometry of
the problem, rupture north of the hypocenter is primarily
constrained by the two European seismograms (Gottingen
and Uppsala), while rupture to the south is primarily
constrained by the Puerto Rico data. Thus, we first tried a
two-segment rupture velocity model split at the hypocenter
and solved for a single rupture velocity on each of these two
segments. Several locations have been suggested for the
hypocenter of the earthquake by analyzing local and tele-
seismic observations (Reid, 1910; Bolt, 1968; Lomax,
2005). We used the latest estimate determined by Lomax
(2005), which is located about 3 km west of the San Fran-
cisco zoo. We find that it takes about 85 and 52 sec, respec-
tively, for the rupture to propagate along the northern
(330 km long) and southern (150 km) segment from the
hypocenter, which indicates that the rupture travels to the
north at an average speed of $3.9 \pm 0.006$ km/sec, exceeding
the average shear wave velocity of the Earth’s crust north
of San Francisco, and to the south at $2.9 \pm 0.1$ km/sec, re-
spectively. The standard errors of the rupture velocity esti-
mates are quite small, due to the fact that neighboring
points in seismic waveforms are highly correlated, a fact
not accounted for in the inversion. In an attempt to localize
the rupture velocity, we divided the rupture into five seg-
ments (three segments north of the hypocenter, 110 km
for each; two segments to the south, 70 and 80 km for each).
While it is possible that the data could resolve such spatial
variations in rupture velocity, the fact that the total rupture

Figure 3. As for Figure 2, the estimated model coordinate displacement fields (blue arrows) with 95% confidence ellipses in four local
networks: (a) Point Arena, (b) Fort Ross, (c) Tomales Bay, (d) Colma.
durations north and south of the hypocenter are about the same in both the two- and five-segment models (Fig. 5) suggests that while supershear rupture to the north of the hypocenter is required to fit the data, it may be difficult to localize it further.

Figures 6 and 7 show the slip and waveform comparisons, respectively, obtained from the two-segment model. A digitized version of the slip estimates are given in Table 2 in the electronic edition of BSSA. The slip in the northern segment is smoother than that obtained solely from geodetic data (Fig. 4), but there is significant slip in the northern region of the fault. This confirms that the long rupture length (~500 km) is compatible with the seismic data, although the amount of slip is somewhat smaller than the geodetically preferred value. The synthesized waveform envelopes capture the duration and amplitude of the seismograms, and reasonable waveform fits are achieved at the European stations even though the objective function is defined using the waveform envelopes (Fig. 7). The polarity of the synthetic waveform (SV component) at the PTR station is reversed. This can occur while fitting the waveform envelope because it does not preserve polarity information. Inaccurate arrival time alignment, an inaccurate Green's function, reversed polarity on the instrument, and delayed rupture propagation to the southeast are all possible explanations of this mismatch. Because the fit at PTR is primarily controlled by the slip and rupture propagation to the southeast of the hypocenter, it does not affect the inference that the rupture extended a total length of 500 km or that rupture to the north of the hypocenter was supershear.

Our final slip model has an average slip of 4.3 meters and a seismic moment of $7.9 \times 10^{20}$ N m, which corresponds to a moment magnitude of $M_w 7.9$. Although the final slip model was obtained by a joint inversion of the geodetic and seismic data, the final static slip distribution is primarily constrained by the triangulation data. Because the triangulation survey data used in this study span an interval as long as 40 yr, our slip estimates include postseismic and interseismic, as well as coseismic, deformation. Because the long-
Figure 6. Final slip models with 2 sigma errors obtained from both geodetic and seismic data compared with two previous slip models. Significant slip is observed in the northern region of the fault, although the actual amount of slip is somewhat smaller than the previous geodetic slip (Thatcher et al., 1997). The rest is the same as Figure 4.

(a) Waveform Envelopes

(b) Waveforms

Figure 7. Comparison of observed (solid) and calculated (dashed) waveforms and their envelopes. Data are direct S phases band-pass filtered between 0.01 and 0.1 Hz. Maximum amplitudes are given on the left of each trace in mm at the WWSSN LP seismograph (top, observed; bottom, calculated). The calculated waveform envelopes successfully reproduce most of the amplitude and duration information of the observed waveform envelopes. Some stations also achieve good waveform fitting, although our objective function is based on the waveform envelopes (Gottingen, Germany, GOT; Uppsala, Sweden, UPP; Puerto Rico, PTR; W.I.; Kobe, Japan; KOB; Osaka, Japan, OSK).
term aseismic strain accumulates in the opposite direction of the deformation caused by earthquakes (i.e., right-lateral strain accumulation between earthquakes versus right-lateral strain release during the earthquake), our estimate of $M_w 7.9$ should be considered a lower bound on the size of the 1906 earthquake in that sense. However, afterslip at seismogenic depths, which is very difficult to constrain given the data available, could bias the estimated magnitude to higher values.

Discussion and Implications

The long rupture length is also strongly supported by seismic intensity data from the 1906 earthquake (Boatwright and Bundock, 2005). The intensity map clearly shows the severely damaged area (intensity VII or larger) extending north to the Mendocino triple junction and rules out the possibility that the slip on the fault north of Point Arena occurred aseismically. Moreover, a recent examination of the northernmost San Andreas fault near Shelter Cove concluded that slip mapped there in 1906 was likely tectonic, extensive, and located on the main trace of the San Andreas fault (Prentice et al., 1999). A simpler analysis of the waveform data that does not directly model the spatial variation of slip also supports supershear rupture north of the hypocenter (Fig. 8). Deconvolution of the empirical Green’s functions from the 1906 seismograms yields an estimate of the apparent source duration and, given the fault length, the average rupture velocity. Using the European stations and employing positivity, smoothness, and moment-minimization constraints in a

Figure 8. Total rupture duration of the earthquake inferred from the deconvolution of observed waveforms. (a) Apparent source time functions for two European stations (GOT and UPP) obtained from deconvolving the empirical Green's functions of the Morgan Hill earthquake from the observed waveforms. (b) Stacked version of the all source time functions except the SH component of the Uppsala station. (c) Expected duration of waveforms at each station as a function of the rupture velocity north of the hypocenter. The rupture velocity south of the hypocenter was fixed as 2.9 km/sec. The expected duration of waveforms were calculated from the maximum difference of arrival times from each fault patch. The duration of the apparent source time functions are well defined at each component except for the Uppsala SH, and is quite well resolved in the stacked source time function (b). The apparent rupture duration at the two European stations is about 75 sec, which is equivalent to a 3.4 km/sec average rupture velocity north of the hypocenter. This independent measure of the total rupture duration of the earthquake supports the high rupture velocity, although the estimated velocity is somewhat less than in the kinematic slip inversion.
time-domain deconvolution of the empirical Green’s function event from the 1906 mainshock, we find a duration of \(\sim 75\) sec, which corresponds to an average rupture velocity of \(\sim 3.4\) km/sec north of the hypocenter. This should be regarded as a lower bound because the smoothness constraint on the source time function deconvolution results in a longer duration, and hence lower inferred rupture velocity.

Both the intensity data and the deconvolution support our long-rupture-length slip model with supershear rupture. If the rupture velocity in this earthquake was supershear, as we have suggested, then it has important implications for seismic hazard. First it demonstrates that slip models derived from geodetic and seismic observations are compatible, which is relevant for northern California because it provides a unified slip model to be used in recurrence models for future earthquakes. More generally, because the nature of strong ground motion for earthquakes that undergo supershear rupture is profoundly different from those in which the rupture is subshear (Aagaard and Heaton, 2004), it will be necessary to account for this when predicting both the level and the variability of strong ground motion in future large strike-slip earthquakes.

Acknowledgments

We thank David Wald for providing us with his digitized teleseismic data and Wayne Thatcher for sharing his notes on the triangulation data. We also thank the National Geodetic Survey for helping us extract triangulation data from their archives. We greatly benefited from the advice, criticism, and assistance of David Wald, Jack Boatwright, and Brad Aagaard throughout the study, and we thank two reviewers, David Wald and Pengcheng Liu, for their insightful comments, which improved the manuscript. This research was supported by the U.S. Geological Survey (USGS), Department of Interior, under USGS Award Number 06HQGR0025. The views and conclusions contained in this document are ours and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. government.

References


Appendix

Here we show that the denuisancing procedure of Yu and Segall (1996) reduces to the standard method, using the differences between repeated angles, when the pre- and post-earthquake networks are identical. That is, when all angles are repeated. Yu and Segall (1996) write the equations relating angle measurements \(d\theta\) to the coordinate corrections at the initial epoch \(dx_1\) and displacements \(u\) as

\[
\begin{bmatrix}
  \frac{d\theta_1}{d\theta_2}
\end{bmatrix} = \begin{bmatrix}
  A_1 \\
  A_2
\end{bmatrix} \begin{bmatrix}
  dx_1 \\
  0
\end{bmatrix} + \begin{bmatrix}
  0
\end{bmatrix} u, \quad (A1)
\]
where the subscripts 1 and 2 refer to the initial and final surveys. The displacements are assumed to arise from fault slip \( s \) in an elastic medium so that \( u = Gs \), and (A1) becomes

\[
\begin{bmatrix}
\frac{d\theta_1}{dx_1}
\end{bmatrix} = \begin{bmatrix}
A_1 \\
A_2
\end{bmatrix} \begin{bmatrix}
x_1 \\
A_2 G
\end{bmatrix} + \begin{bmatrix}
0 \\
A_2 G
\end{bmatrix} s. \tag{A2}
\]

If the network geometry is identical in both surveys, then \( A_1 = A_2 \equiv A \) and subtracting the first set of equations from the second leads directly to

\[
d\theta_2 - d\theta_1 = AG. \tag{A3}
\]

In this case, the changes in angles are directly related to the fault slip, with no need to consider the corrections to the initial coordinates. For least-squares estimation, (A3) leads to normal equations

\[
G^T A^T (d\theta_2 - d\theta_1) = G^T A^T AG. \tag{A4}
\]

We now consider the denuisancing procedure. Equation (A2) can be written more compactly as

\[
d\theta = \Psi dx + \Omega s. \tag{A5}
\]

Yu and Segall (1996) define a projection operator \( Q \) that annihilates the dependence on \( dx \)

\[
Q \equiv I - \Psi \Psi^\dagger, \tag{A6}
\]

where \( I \) is an identity matrix and \( \Psi^\dagger \) is the generalized inverse of \( \Psi \). Premultiplying (A5) by \( Q \) leads to

\[
Q d\theta = Q \Omega s, \tag{A7}
\]

since

\[
Q\Psi = (I - \Psi \Psi^\dagger) \Psi = 0. \tag{A8}
\]

A least-squares estimate of fault slip then follows from the normal equations

\[
\Omega^T Q d\theta = \Omega^T Q \Omega s, \tag{A9}
\]

since \( Q \) is both symmetric and idempotent.

If the network geometry is repeated, then

\[
\Psi = \begin{bmatrix}
A \\
A
\end{bmatrix}. \tag{A10}
\]

The matrix \( A \) has singular value decomposition (SVD) given by \( A = U_p S_p V_p^T \), where \( p \) is the number of nonzero singular values. It follows then that \( \Psi \) has an SVD with the same singular vectors spanning the model space \( V \) and repeated data space singular vectors given by

\[
\begin{bmatrix}
\frac{1}{\sqrt{2}} U_p \\
\frac{1}{\sqrt{2}} U_p
\end{bmatrix}, \tag{A11}
\]

where the factor of \( 1/\sqrt{2} \) arises due to normalization. Thus, in this case, the projection operator \( Q \) is

\[
Q = \begin{bmatrix}
I - \frac{1}{2} H & -\frac{1}{2} H \\
-\frac{1}{2} H & I - \frac{1}{2} H
\end{bmatrix}, \tag{A12}
\]

where \( H \equiv AA^\dagger = U_p U_p^T \). Substituting (A12) into (A9) leads to

\[
G^T A^T \left[ \left( I - \frac{1}{2} H \right) d\theta_2 - \frac{1}{2} H d\theta_1 \right] = G^T A^T \left( I - \frac{1}{2} H \right) AG.
\]

\[
\frac{1}{2} G^T A^T (d\theta_2 - d\theta_1) = \frac{1}{2} G^T A^T AG, \tag{A13}
\]

since \( A^T H = (HA)^T = (AA^\dagger) A^T = A^T \). Comparing (A13) to (A4), we see that both methods lead to the same result when all angles are repeated. The denuisancing procedure, however, allows all measurements to be used when the angles are not all repeated.
Queries

1. Please provide names of all editors. “Et al.” is not allowed in the reference list.
2. Is this correct? Figure 2a? The original manuscript was a little confusing at this point.
3. The word “above” here must be changed to a descriptive term, as the journal’s style does not allow for references to “above” and “below” information within the article. Please suggest an alternative to “above.” (Suggestion: “...following the aforementioned steps...”)
4. Please define PTR, if possible.
5. Please carefully check all math in the Appendix. Is everything correctly styled? Keep in mind that journal style requires bold, Roman font styles for all matrices, vectors, tensors, and moment magnitudes. Almost all other variables should be styled as nonbold italic characters.
6. Please define WWSSN and LP.
7. What is “W.I.”?