Data Files for Ground-Motion Simulations of the 1906 San Francisco Earthquake and Scenario Earthquakes on the Northern San Andreas Fault

Data Series 413
Data Files for Ground-Motion Simulations of the 1906 San Francisco Earthquake and Scenario Earthquakes on the Northern San Andreas Fault

By Brad Aagaard, Michael Barall, Thomas M. Brocher, David Dolenc, Douglas Dreger, Robert W. Graves, Stephen Harmsen, Stephen Hartzell, Shawn Larsen, Kathleen McCandless, Stefan Nilsson, N. Anders Petersson, Arthur Rodgers, Bjorn Sjogreen, and Mary Lou Zoback

Data Series 413

U.S. Department of the Interior
U.S. Geological Survey
## Contents

1906 San Francisco Earthquake Modelers and Scenarios .............................................................. 1
1989 Loma Prieta Earthquake Modelers and Scenarios .............................................................. 1
Organization of the Data Files .................................................................................................. 5
File Formats ............................................................................................................................. 7
References Cited .......................................................................................................................... 7

### Appendix—Methodology Details

- Problem Statement .............................................................................................................. 16
- Validation ............................................................................................................................. 16
- Wave-Propagation Codes ................................................................................................... 16
  - Material Properties and Attenuation .............................................................................. 16
  - Topography and Water ................................................................................................. 17
- Slip-Time Function ............................................................................................................ 17
- 1906 San Francisco Source Models .................................................................................... 17
- 1989 Loma Prieta Source Models ....................................................................................... 21
- Modified Mercalli Intensity .............................................................................................. 21

### Figures

1. Model domains for the 1906 San Francisco earthquake simulations ........................................ 3
2. Model domains and epicenter for the 1989 Loma Prieta earthquake simulations ....................... 4
3. Directory tree for 1906 San Francisco earthquake simulation results (left) and 1989 Loma Prieta earthquake simulation results (right) .......................................................... 10
4. Geodetic displacement files for 1906 San Francisco earthquake simulation ............................... 10
5. Peak ground velocity files for 1906 San Francisco earthquake simulation ............................... 11
6. Peak ground velocity files for 1989 Loma Prieta earthquake simulations ............................... 11
7. Geodetic displacement file format ....................................................................................... 11
8. Peak ground velocity file format. A peak ground velocity file contains the peak ground velocity at each point on a dense grid of points ............................................................... 12
9. Rupture trace file format .................................................................................................... 12
10. Waveform site list file format ............................................................................................ 12
11. Waveform file format ....................................................................................................... 13
12. Loma Prieta station file format ........................................................................................ 13
13. Hypocenter locations for the 1906 San Francisco earthquake simulations ............................. 18
14. Source models for the 1906 San Francisco earthquake simulations ....................................... 19
15. Rise times for the 1906 San Francisco earthquake simulations ............................................ 20
16. Source models for the 1989 Loma Prieta earthquake simulations ........................................ 20
**Tables**

1. Wave propagation codes for the 1906 San Francisco earthquake simulations ..................2
2. Scenarios for the 1906 San Francisco earthquake simulations ........................................5
3. Coordinates of the earthquake scenario hypocenters, all of which are at an elevation of -10 km ..................................................................................................................5
4. Propagation codes for the 1989 Loma Prieta earthquake simulations ..............................6
Data Files for Ground-Motion Simulations of the 1906 San Francisco Earthquake and Scenario Earthquakes on the Northern San Andreas Fault

By Brad Aagaard, Michael Barall, Thomas M. Brocher, David Dolenc, Douglas Dreger, Robert W. Graves, Stephen Harmsen, Stephen Hartzell, Shawn Larsen, Kathleen McCandless, Stefan Nilsson, N. Anders Petersson, Arthur Rodgers, Bjorn Sjogreen, and Mary Lou Zoback

This data set contains results from ground-motion simulations of the 1906 San Francisco earthquake, seven hypothetical earthquakes on the northern San Andreas Fault, and the 1989 Loma Prieta earthquake. The bulk of the data consists of synthetic velocity time-histories. Peak ground velocity on a 1/60th degree grid and geodetic displacements from the simulations are also included. Details of the ground-motion simulations and analysis of the results are discussed in Aagaard and others (2008a,b).

1906 San Francisco Earthquake Modelers and Scenarios

Five groups participated in modeling the 1906 San Francisco earthquake (1) Aagaard, (2) Graves, (3) Harmsen and others (Harmsen and Hartzell), (4) Larsen and others (Larsen, Dreger, and Dolenc), and (5) Petersson and others (Petersson, Rodgers, McCandless, Nilsson, and Sjogreen). Each group employed a different modeling code to solve the elastic wave equation (table 1 and fig. 1).

The earthquake scenarios include two models of the 1906 earthquake (the Song scenario and the SongMod scenario) and seven models of similar-sized hypothetical earthquakes (SongModHypoN, SongModHypoC, SongModHypoS, RandomHypo06, RandomHypoN, RandomHypoC, and RandomHypoS) on the northern San Andreas Fault (table 2).

The Song scenario is based on the source model by Song and others (2008). The Song scenario specifies uniform slip on fault patches measuring 10 km along-strike and 12 km down-dip and, therefore, contains slip variations only at long scales. The hypocenter is a few kilometers offshore from San Francisco at a depth of 10 km.

The SongMod scenario is similar to the Song scenario, but it uses a source model that has been modified by adding shorter-scale variations in slip and rupture speed. This is the preferred source model for the 1906 ground motion simulations. The additional variations in slip are randomly phased and passed through a high-pass spatial filter whose amplitude falls off as the inverse square of the wavenumber for large wavenumbers. The rupture speed is modified so that the rupture propagates more quickly through regions with larger slip. The SongModHypoN, SongModHypoC, and SongModHypoS scenarios are hypothetical earthquakes that have the same slip distribution as the SongMod scenario but different hypocenters (N, north; C, center; S, south; table 3).

The RandomHypo06, RandomHypoN, RandomHypoC, and RandomHypoS scenarios use a randomly generated slip distribution. The slip distribution was produced by generating a slip distribution with random phase at all wavelengths and passing it through the wavenumber squared low-pass spatial filter mentioned above, while constraining the average slip to be the same as in the SongMod model. These four random slip scenarios are all the same, except for the location of the hypocenter. Additional information is available in the Appendix.

1989 Loma Prieta Earthquake Modelers and Scenarios

Four groups participated in modeling the 1989 Loma Prieta earthquake (1) Aagaard, (2) Dolenc and others (Dolenc, Dreger, and Larsen), (3) Graves, and (4) Harmsen and others (Harmsen and Hartzell). Each uses a different elastic wave propagation code (table 4 and fig. 2). There are two scenarios for the 1989 earthquake:

- The Beroza scenario is based on the Beroza (1991) source model.
- The Wald scenario is based on the Wald and others (1991) source model.

Additional information is available in the Appendix.
## 2 Data Files for Ground-Motion Simulations of the 1906 San Francisco Earthquake and Scenario Earthquakes

### Table 1. Wave propagation codes for the 1906 San Francisco earthquake simulations.

<table>
<thead>
<tr>
<th>Model domain</th>
<th>Aagaard</th>
<th>Graves</th>
<th>Harmsen and others</th>
<th>Larsen and others</th>
<th>Petersson and others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length, in km</td>
<td>250</td>
<td>555</td>
<td>128</td>
<td>630</td>
<td>550</td>
</tr>
<tr>
<td>Width, in km</td>
<td>110</td>
<td>162</td>
<td>52</td>
<td>320</td>
<td>200</td>
</tr>
<tr>
<td>Max. depth, in km</td>
<td>40</td>
<td>45</td>
<td>31</td>
<td>55</td>
<td>40</td>
</tr>
<tr>
<td>NW corner(1)</td>
<td>-123.7083, 38.2832</td>
<td>-125.5000, 40.2000</td>
<td>-122.7313, 37.6313</td>
<td>-126.2037, 39.6589</td>
<td>-125.4990, 40.0440</td>
</tr>
<tr>
<td>NE corner</td>
<td>-122.7002, 38.8944</td>
<td>-123.9482, 41.0505</td>
<td>-122.3116, 37.9610</td>
<td>-123.2940, 41.3485</td>
<td>-123.6000, 41.1000</td>
</tr>
<tr>
<td>SE corner</td>
<td>-121.0208, 37.0716</td>
<td>-120.2911, 36.9623</td>
<td>-121.2978, 37.1331</td>
<td>-119.0100, 36.7700</td>
<td>-119.9590, 37.1030</td>
</tr>
<tr>
<td>SW corner</td>
<td>-122.0180, 36.4894</td>
<td>-121.8528, 36.1118</td>
<td>-121.7173, 36.8069</td>
<td>-121.8819, 35.0804</td>
<td>-121.8060, 36.0470</td>
</tr>
<tr>
<td>Projection</td>
<td>none (3-D Earth)</td>
<td>spheroidal</td>
<td>ellipsoidal trans. Mercator</td>
<td>spheroidal</td>
<td>spheroidal</td>
</tr>
<tr>
<td>Discretization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>unstructured finite element</td>
<td>staggered-grid finite element</td>
<td>staggered-grid finite difference</td>
<td>staggered-grid finite difference</td>
<td>node centered finite difference</td>
</tr>
<tr>
<td>Space accuracy</td>
<td>2nd order</td>
<td>4th order</td>
<td>4th order</td>
<td>4th order</td>
<td>2nd order</td>
</tr>
<tr>
<td>Time accuracy</td>
<td>2nd order</td>
<td>2nd order</td>
<td>2nd order</td>
<td>2nd order</td>
<td>2nd order</td>
</tr>
<tr>
<td>Resolution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Element size, in m</td>
<td>variable</td>
<td>150</td>
<td>50-150</td>
<td>100</td>
<td>125</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>T &gt; 2.0 s</td>
<td>T &gt; 1.0 s</td>
<td>T &gt; 1.0 s</td>
<td>T &gt; 1.0 s</td>
<td>T &gt; 1.0 s</td>
</tr>
<tr>
<td>Min. Vs, in m/s</td>
<td>700</td>
<td>760</td>
<td>330</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>Features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topography(3)</td>
<td>yes</td>
<td>bulldozed</td>
<td>squashed</td>
<td>yes</td>
<td>bulldozed</td>
</tr>
<tr>
<td>Water</td>
<td>air filled</td>
<td>sediment filled</td>
<td>excluded</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>Mat. properties</td>
<td>USGS 05.1.0</td>
<td>USGS 05.1.0</td>
<td>USGS 05.1.0</td>
<td>USGS 05.1.0</td>
<td>USGS 05.1.0</td>
</tr>
<tr>
<td>Attenuation(4)</td>
<td>none</td>
<td>Graves</td>
<td>Graves</td>
<td>USGS 05.1.0</td>
<td>none</td>
</tr>
<tr>
<td>Earthquake source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>offset in mesh</td>
<td>point sources</td>
<td>point sources</td>
<td>point sources</td>
<td>point sources</td>
</tr>
<tr>
<td>Number of point sources</td>
<td>N/A</td>
<td>6,084</td>
<td>1,200–1,400</td>
<td>6,084</td>
<td>12,313</td>
</tr>
<tr>
<td>Fault surface</td>
<td>3D geol. model</td>
<td>3D geol. model</td>
<td>3D geol. model</td>
<td>3D geol. model</td>
<td>3D geol. model</td>
</tr>
</tbody>
</table>

1Domain corners are given as WGS84 longitude and latitude.
2Graves employs a hybrid method which combines a deterministic model for periods $T > 1.0$ s with a stochastic model for periods $1.0$ s $> T > 0.1$ s.
3For topography, “bulldozed” means that all material above some elevation is removed, and voids below that elevation are filled with generic material. “Squashed” means that the geologic model is vertically distorted to produce a planar top surface. Squashing also eliminates bodies of water.
4For material properties and attenuation, “USGS 05.1.0” refers to the U.S. Geological Survey Bay Area Velocity Model 05.1.0, and “Graves” refers to a simple attenuation model developed by Graves. See the Appendix for more information.
Figure 1. Model domains for the 1906 San Francisco earthquake simulations. Solid lines are the bounding boxes of the domains used by the five modeling groups. Dotted lines are the bounding boxes of the US Geological Survey Bay Area Velocity Model 05.1.0. The thick red curve is the portion of the San Andreas Fault that ruptured in 1906. Thin red curves are the surface traces of major faults in the region.
Figure 2. Model domains and epicenter for the 1989 Loma Prieta earthquake simulations. Large rectangles are the bounding boxes of the domains used by the four modeling groups. The small red rectangle is the projection of the fault plane in the Wald model, and the star is the epicenter. Thin red curves are the surface traces of major faults in the region.
Table 2. Scenarios for the 1906 San Francisco earthquake simulations. The table lists the names of the nine scenarios, along with the slip distribution and hypocenter location used for each scenario.

<table>
<thead>
<tr>
<th>Scenario name</th>
<th>Slip distribution and rise time</th>
<th>Hypocenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Song</td>
<td>Song</td>
<td>1906</td>
</tr>
<tr>
<td>SongMod</td>
<td>SongMod</td>
<td>1906</td>
</tr>
<tr>
<td>SongModHypoN</td>
<td>SongMod</td>
<td>Rockport</td>
</tr>
<tr>
<td>SongModHypoC</td>
<td>SongMod</td>
<td>Bodega Bay</td>
</tr>
<tr>
<td>SongModHypoS</td>
<td>SongMod</td>
<td>San Juan Bautista</td>
</tr>
<tr>
<td>RandomHypo06</td>
<td>Random</td>
<td>1906</td>
</tr>
<tr>
<td>RandomHypoN</td>
<td>Random</td>
<td>Rockport</td>
</tr>
<tr>
<td>RandomHypoC</td>
<td>Random</td>
<td>Bodega Bay</td>
</tr>
<tr>
<td>RandomHypoS</td>
<td>Random</td>
<td>San Juan Bautista</td>
</tr>
</tbody>
</table>

The Song distribution produces a moment magnitude 7.9 event (average slip is 4.3 m). The SongMod and Random slip distributions produce moment magnitude 7.8 events (average slip is 3.0 m).

Table 3. Coordinates of the earthquake scenario hypocenters, all of which are at an elevation of -10 km.

<table>
<thead>
<tr>
<th>Hypocenter</th>
<th>Longitude</th>
<th>Latitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>1906</td>
<td>-122.55</td>
<td>37.75</td>
</tr>
<tr>
<td>Rockport</td>
<td>-124.000</td>
<td>39.800</td>
</tr>
<tr>
<td>Bodega Bay</td>
<td>-123.016</td>
<td>38.300</td>
</tr>
<tr>
<td>San Juan Bautista</td>
<td>-121.615</td>
<td>36.872</td>
</tr>
</tbody>
</table>

Coordinates are WGS84 longitude and latitude.

Organization of the Data Files

The simulation data files are stored under two top-level directories, sf1906 and lomaprieta, which contain simulation results for the 1906 San Francisco and 1989 Loma Prieta earthquakes, respectively. The names of the files and directories use the following conventions. The names of the modeling groups are abbreviated with aagaard for Aagaard, dolenc for Dolenc and others, graves for Graves, harmsen for Harmsen and others, larsen for Larsen and others, and petersson for Petersson. In some cases we append the bandwidth to the name of the modeler with bb for T > 0.1 s, t1 for T > 1.0 s, and t2 for T > 2.0 s. We also include the type of data in the filename for geodetic displacements (geodeticdisp) and peak ground velocity (pgvgrid). In the case of the grid files, the data spans either the entire rupture length (indicated by cenca in the filename) or the San Francisco Bay area (indicated by bayarea in the filename). The directory trees are shown in figures (figures 3 through 12 follow the references).

The sf1906 directory contains three subdirectories and two files:
- **geodetic** — Calculated geodetic displacements. For comparison purposes, this subdirectory also contains displacements calculated from geodetic models and triangulation data from Song and others (2008). These files are listed in figure 4, and the file format is described in figure 7.
- **pgvpgagrid** — Calculated peak ground velocity (PGV). Broadband simulations in this subdirectory (gravesbb) also include peak ground acceleration (PGA). This data can be used to construct maps of PGV or Modified Mercalli Intensity (MMI). These files are listed in figure 5, and the file format is described in figure 8.

The lomaprieta directory contains two subdirectories and one file:
- **pgvpgagrid** — Calculated PGV. The data in this subdirectory can be used to construct maps of PGV or MMI. These files are listed in figure 6, and the file format is described in figure 8.
- **stations.txt** — List of stations used for the 1906 San Francisco earthquake simulations. The file format is described in figure 9.
- **waveforms** — Calculated velocity time histories. Within this subdirectory are five subdirectories corresponding to the five participating modeling groups. Each of these subdirectories contains a zip archive file with the waveforms for each scenario. The zip archive contains a file for each station, in which the name of the file matches the site; for example, the waveforms for station SF003 are in a file called sf003.txt. The waveform file format is described in figure 11. Most modern operating systems include utilities for extracting all or a selected subset of files from zip archives.
### Table 4. Wave propagation codes for the 1989 Loma Prieta earthquake simulations.

<table>
<thead>
<tr>
<th>Model domain</th>
<th>Aagaard</th>
<th>Dolenc and others</th>
<th>Graves</th>
<th>Harmsen and others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length, in km</td>
<td>250</td>
<td>220</td>
<td>180</td>
<td>128</td>
</tr>
<tr>
<td>Width, in km</td>
<td>110</td>
<td>135</td>
<td>84</td>
<td>52</td>
</tr>
<tr>
<td>Max. depth, in km</td>
<td>40</td>
<td>50</td>
<td>45</td>
<td>31</td>
</tr>
<tr>
<td>NW corner(1)</td>
<td>-123.7083, 38.2832</td>
<td>-123.3610, 37.9610</td>
<td>-122.9500, 37.7000</td>
<td>-122.7313, 37.6313</td>
</tr>
<tr>
<td>NE corner</td>
<td>-122.7002, 38.8944</td>
<td>-122.1150, 38.6800</td>
<td>-122.1748, 38.1326</td>
<td>-122.3116, 37.9610</td>
</tr>
<tr>
<td>SE corner</td>
<td>-121.0208, 37.0716</td>
<td>-120.6670, 37.0640</td>
<td>-121.0106, 36.8075</td>
<td>-121.2978, 37.1331</td>
</tr>
<tr>
<td>SW corner</td>
<td>-122.0180, 36.4894</td>
<td>-121.9000, 36.3600</td>
<td>-121.7858, 36.3749</td>
<td>-121.7173, 36.8069</td>
</tr>
<tr>
<td>Projection</td>
<td>none (3-D Earth)</td>
<td>ellipsoidal trans. Mercator</td>
<td>spheroidal</td>
<td>ellipsoidal trans. Mercator</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discretization</th>
<th>Method</th>
<th>unstructured finite element</th>
<th>staggered-grid finite difference</th>
<th>staggered-grid finite difference</th>
<th>staggered-grid finite difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space accuracy</td>
<td>2nd order</td>
<td>4th order</td>
<td>4th order</td>
<td>4th order</td>
<td>4th order</td>
</tr>
<tr>
<td>Time accuracy</td>
<td>2nd order</td>
<td>2nd order</td>
<td>2nd order</td>
<td>2nd order</td>
<td>2nd order</td>
</tr>
</tbody>
</table>

| Resolution | Element size, in m | variable | 125 | 150 | 50-150 |
| Bandwidth   | T > 2.0 s          | T > 1.2 s | T > 1.0 s, T > 0.1 s (2) | T > 1.0 s |
| Min. Vs, in m/s | 700     | 500       | 760       | 330       |

| Features | Topography(3) | yes | bulldozed | bulldozed | squashed |
| Water | air filled | included | sediment filled | excluded |
| Mat. properties | USGS 05.1.0 | USGS 05.1.0 | USGS 05.1.0 | USGS 05.1.0 |
| Attenuation(4) | none | none | Graves | USGS 05.1.0 | (Beroza) Graves (Wald) |

<table>
<thead>
<tr>
<th>Earthquake source</th>
<th>Method</th>
<th>offset in mesh</th>
<th>point sources</th>
<th>point sources</th>
<th>point sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of point sources(5)</td>
<td>N/A</td>
<td>34,020 (Beroza)</td>
<td>1,215 (Beroza)</td>
<td>1,099 (Beroza)</td>
<td></td>
</tr>
<tr>
<td>Fault surface</td>
<td>3D geol. model</td>
<td>plane</td>
<td>Plane</td>
<td>Plane</td>
<td></td>
</tr>
</tbody>
</table>

1Domain corners are given as WGS84 longitude and latitude.
2Graves employs a hybrid method which combines a deterministic model for frequencies T > 1.0 s with a stochastic model for frequencies 1.0 s > T > 0.1 s.
3For topography, “bulldozed” means that all material above some elevation is removed, and voids below that elevation are filled with generic material. “Squashed” means that the geologic model is vertically distorted to produce a planar top surface. (Squashing also eliminates bodies of water.)
4For material properties and attenuation, “USGS 05.1.0” refers to the U.S. Geological Survey Bay Area Velocity Model 05.1.0, and “Graves” refers to a simple attenuation model developed by Graves. See the Appendix for more information.
5“Beroza” and “Wald” refers to the source model.
File Formats

All files are ASCII text files in standard UNIX format (each lines ends with a single line-feed character, hexadecimal value 0A). Each file contains a header, followed by a data table. In the header each line begins with a # character. In the data each lines contains a series of numbers separated by spaces and/or tabs. Figures 7 through 12 describe the file formats. The files ending in .zip are zip archives. Most modern operating systems include utilities for extracting one or more files from a zip archive.

References Cited


Figures 3 through 12
10  Data Files for Ground-Motion Simulations of the 1906 San Francisco Earthquake and Scenario Earthquakes

**Figure 3.** Directory tree for 1906 San Francisco earthquake simulation results and 1989 Loma Prieta earthquake simulation results.

**Figure 4.** Geodetic displacement files for 1906 San Francisco earthquake simulation. The first part of the filename identifies the modeling group and bandwidth: aagaardt2 (Aagaard, 0.5 Hz), gravesbb (Graves, 10 Hz), gravenst (Graves, 1 Hz), harmsent1 (Harmsen and others, 1 Hz), larsent1 (Larsen and others, 1 Hz), and peterssont2 (Petersson and others, 0.5 Hz). The file beginning with song contains displacements calculated by Song and others based on geodetic models and triangulation data. The second part of the filename identifies the earthquake scenario (RandomHypo06, RandomHypoC, etc.). See figure 7 for a description of the file format.
Figure 5. Peak ground velocity files for 1906 San Francisco earthquake simulation. The first part of the filename identifies the modeling group and bandwidth: aagaardt2 (Aagaard, 0.5 Hz), gravesbb (Graves, 10 Hz), gravesti (Graves, 1 Hz), harmsent1 (Harmsen and others, 1 Hz), larsent1 (Larsen and others, 1 Hz), and peterssont2 (Petersson and others, 0.5 Hz). The second part of the filename identifies the earthquake scenario (for example, RandomHypo06, RandomHypoC). The third part of the filename identifies the region covered: bayarea (San Francisco Bay Area) or cenca (Central California). The fourth part of the filename identifies the file contents: pgvgrid (peak ground velocity) or pgvpagagrid (peak ground velocity and peak ground acceleration). See figure 8 for a description of the file format.

Figure 6. Peak ground velocity files for 1989 Loma Prieta earthquake simulations. The first part of the filename identifies the modeling group and bandwidth: aagaardt2 (Aagaard, T>2.0s), dolenct1 (Dolenc and others, T>1.0s), gravesbb (Graves, T>0.1s), gravesti (Graves, T>1.0s), and harmsent1 (Harmsen and others, T>1.0s). The second part of the filename identifies the earthquake source model (Beroza or Wald). The last part of the filename identifies the file contents: pgvgrid (peak ground velocity) or pgvpagagrid (peak ground velocity, peak ground acceleration, and modified Mercalli intensity). See figure 8 for a description of the file format.

Figure 7. Geodetic displacement file format. A geodetic displacement file contains the final horizontal geodetic displacement at each point in a set of points. The file header contains the following information: name of author; earthquake scenario or source model; date simulation was performed; bandwidth (highest seismic wave frequency included in the simulation); and description of data fields. Following the header is a series of lines, each corresponding to one point in the set of points. On each line are four numbers: the longitude, in degrees (WGS84 coordinates); the latitude, in degrees (WGS84 coordinates); the east component of displacement, in meters; and the north component of displacement, in meters.
Figure 8. Peak ground velocity file format. A peak ground velocity file contains the peak ground velocity at each point on a dense grid of points. This file format is useful for constructing maps of ground motion intensity. The file header contains the following information: name of author; earthquake scenario or source model; date simulation was performed; bandwidth (highest seismic wave frequency included in the simulation); and description of data fields. Following the header is a series of lines, each corresponding to one point in the grid of points. On each line are three numbers: the longitude, in degrees (WGS84 coordinates); the latitude, in degrees (WGS84 coordinates); and the peak ground velocity at the given point, in meters per second. Some files also include the peak ground acceleration. If so, the peak ground acceleration appears as a fourth number on each line. Some files also include both peak ground acceleration and modified Mercalli intensity. If so, the peak ground acceleration appears as a fourth number on each line, and the modified Mercalli intensity appears as a fifth number on each line.

Figure 9. Rupture trace file format. This file lists the coordinates of the rupture trace used in the 1906 San Francisco earthquake simulations. Following the header is a series of lines with the longitude and latitude (both in degrees and WGS84 coordinates). This file is named sf1906/rupturetrace.txt.

Figure 10. Waveform site list file format. This file lists the stations used in the 1906 San Francisco earthquake simulations. Following the header is a series of lines, each corresponding to one station. On each line are three numbers: the longitude, in degrees (WGS84 coordinates); the latitude, in degrees (WGS84 coordinates); and the name of the station. This file is named sf1906/stations.txt. The 11 digit number following CT is the census track number.
Figure 11. Waveform file format. A waveform file contains a synthetic seismogram. It records the ground velocity at a given station for each time step in the simulation. The waveform file header contains the following information: name of author; earthquake scenario or source model; date simulation was performed; bandwidth (highest seismic wave frequency included in the simulation); station name; longitude and latitude of the station (WGS84 coordinates); longitude and latitude of the point where the waveform data was actually computed; distance from the station to the point where the waveform data was actually computed; and description of data fields. Following the header is a series of lines, each corresponding to one time step in the simulation. On each line are four numbers: the time, in seconds relative to the origin time; the east component of ground velocity, in m/s; the north component of ground velocity, in meters per second; the up component of ground velocity, in meters per second. A waveform file is always given a name that matches the station name.

| Nominal coordinates of stations used in this study, |
| compiled from a variety of sources. |
| Columns are: |
| (1) Longitude (WGS84) |
| (2) Latitude (WGS84) |
| (3) Elevation wrt MSL (m) |
| (4) Station name |

| -121.9522 | 36.9740 | 6.6 | CAP |
| -121.5690 | 37.0090 | 60.9 | GILB |
| -121.7140 | 37.3380 | 464.7 | HALL |
| -121.3970 | 36.8480 | 90.2 | HOL |
| -121.9903 | 37.2008 | 197.5 | LEX |
| -121.8970 | 37.4300 | 7.1 | MILP |
| -121.8510 | 37.3400 | 30.8 | SJIN |
| -121.8030 | 37.2100 | 225.3 | SNJ |

Figure 12. Loma Prieta station file format. This file lists the stations used in the 1989 Loma Prieta earthquake simulations. Following the header is a series of lines, each corresponding to one station. On each line are four numbers: the longitude, in degrees (WGS84 coordinates); the latitude, in degrees (WGS84 coordinates); the elevation of the station, in meters; and the name of the station. This file is named lomaprieta/stations.txt.
Appendix
Appendix—Methodology Details

In this appendix we explain the information in tables 1 and 4. See Aagaard and others (2008a,b) for a thorough discussion of the simulations.

Problem Statement

All the ground-motion modeling codes solve the elastic wave equation in heterogeneous media, given kinematic boundary conditions on the fault surface. The kinematic boundary condition on the fault prescribes fault slip as a function of time for each point on the fault surface. The modeler is also given the 3D fault geometry; the surface topography; a 3D velocity model, which describes the 3D variation in rock properties; and a set of stations on the earth’s surface. The modeler’s job is to solve the elastic wave equation throughout the modeling domain and to report the calculated motion of each station as a function of time. These outputs are synthetic velocity time-histories.

Validation

Every ground-motion simulation has its limitations and approximations. The size of the modeling domain, spatial resolution, bandwidth, and features that can be included in a model are all limited by the available computing resources and numerical techniques. There are also considerable uncertainties in the slip-time function, fault geometry, and rock properties.

Two techniques have been used to validate the modeling results. First, the results produced by the different codes were compared to each other. Since each code employs a unique set of tradeoffs and techniques, agreement between the codes provides some assurance of the robustness of the results.

Second, the results were compared to observed data. In the case of the 1989 Loma Prieta earthquake, modern instrumental records are available to serve as a basis of comparison. For the 1906 San Francisco earthquake, no such instrumental records exist, but it is still possible to perform some comparisons based on accounts of ground-shaking intensity and damage.

Wave-Propagation Codes

Different wave-propagation codes were used by each of the ground-motion modeling groups that participated in modeling the 1906 San Francisco earthquake. Four of these codes were also used in the modeling the 1989 Loma Prieta earthquake. The characteristics of the codes are summarized in tables 1 and 4, respectively. The members of the modeling groups (listed with the primary researcher first) are:

- Aagaard — Brad Aagaard
- Dolenc — David Dolenc, Douglas Dreger, and Shawn Larsen
- Graves — Robert Graves
- Harmsen — Stephen Harmsen and Stephen Hartzell
- Larsen — Shawn Larsen, Douglas Dreger, and David Dolenc
- Petersson — Anders Petersson, Arthur Rodgers, Kathleen McCandless, Stefan Nilsson, and Bjorn Sjogreen

Aagaard used an unstructured, finite-element code that is second-order accurate in space and time. The finite-element mesh consists of 4-node tetrahedra. The element size is variable and equal to approximately 1/10th of the local wavelength of a shear wave with period 2.0 seconds. The fault surface is an internal boundary of the finite-element mesh, and fault slip is represented as a displacement discontinuity on this surface. The fault surface is smoothed by applying a Laplacian smoothing to the vertex coordinates.

The Dolenc, Graves, Harmsen, and Larsen groups used staggered-grid, finite-difference codes, which are fourth-order accurate in space and second-order accurate in time. (Dolenc and Larsen are the same code.) Petersson used a node-centered, finite-difference method that is second-order accurate in space and time. All the finite-difference codes represent fault slip as a set of evenly distributed point sources. Dolenc, Graves, Larsen, and Petersson used a uniform discretization size. Harmsen used a discretization size of 50 m in the top 700 m with a horizontal grid size of 150 m at depths greater than 700 m. The vertical grid size is 50 m down to a depth of 950 m; it then increases linearly to 150 m at a depth of 1,500 m, and to 250 m at a depth of 30 km.

Graves employed a hybrid method, that combines a deterministic finite-difference calculation for frequencies \( f < 1.0 \text{ Hz} \) with a stochastic calculation for frequencies \( 1.0 \text{ Hz} < f < 10.0 \text{ Hz} \). The short-period waves are generated by dividing the fault surface into subfaults and assuming each subfault emits randomly phased seismic waves with an omega-squared spectrum. These short-period waves are propagated through the model with a simple 1D Green’s function. This modeling technique also applies site corrections across all periods to account for near-surface, period-dependent, nonlinear site effects. We refer to Graves’s deterministic finite-difference calculations as long-period simulations \((T > 1.0 \text{ s})\) and to Grave’s hybrid simulations as broadband simulations \((T > 0.1 \text{ s})\).

Material Properties and Attenuation

All the models used the U.S. Geological Survey Bay Area Velocity Model version 05.1.0 for material properties (Brocher and others, 2006). The Aagaard, Dolenc, and Petersson simulations do not include attenuation. Larsen (and Harmsen for the Beroza source model of the 1989 Loma Prieta earthquake) used the attenuation contained in the U.S. Geological Survey Bay Area Velocity Model. Graves and Harmsen (except for Harmsen’s run with the Beroza source model of the 1989 Loma Prieta earthquake) used a simplified attenuation model developed by Graves. In this simplified
Topography and Water

For the Aagaard simulations topography and bathymetry are included by making the top surface of the finite element mesh conform to the surface topography, including the 3D curvature of the earth. The topography is smoothed using the same Laplacian filtering applied to the fault surface.

The Dolenc, Graves, Harmsen, and Petersson finite-difference codes require a planar top surface. There are two techniques for obtaining a planar top surface. Dolenc, Graves, and Petersson “bulldoze” the surface by removing all material above a certain elevation and filling in all voids below the elevation with a generic material. Bulldozing tends to remove low-velocity layers near the earth’s surface. Harmsen “squashes” the surface by vertically distorting the upper part of the velocity model so as to flatten the topography. Squashing preserves near-surface low-velocity layers at the expense of distorting the geologic units in the upper part of the velocity model (for example, in the top 1 km). The Larsen finite-difference code includes topography.

The Aagaard code does not include water (bodies of water are filled with air). The Dolenc, Larsen, and Petersson codes include water (Vs = 0) in the simulations. The Graves code fills bodies of water with generic sediment (V = 2200 m/s, V = 760 m/sec, density = 2100 kg/m³). In the Harmsen code, bodies of water are eliminated by squashing.

Slip-Time Function

The slip-time function gives fault slip as a function of time, for each point on the fault surface. All the models use the integral of Brune’s far-field time function (Brune, 1970),

\[
D(t) = D_{\text{final}} (1 - e^{-t/t_0}) (1 + \frac{t}{t_0})
\]

\[
t_0 = \frac{D_{\text{final}}}{eV_{\text{max}}}
\]

\[
V_{\text{max}} [\text{m/s}] = C_{\tau r} \sqrt{D_{\text{final}} [\text{m}]}
\]

for the slip time history where \(D(t)\) is slip as a function of time, \(t\) is time since the initiation of slip at a given point on the fault, \(t_0\) is a time constant, \(D_{\text{final}}\) is the final slip at a point, \(V_{\text{max}}\) is the peak slip rate, and \(C_{\tau r}\) is a constant that controls the rise time. For the 1989 Loma Prieta earthquake, \(C_{\tau r} = 1.5\). For the 1906 San Francisco earthquake,

\[
C_{\tau r} = \begin{cases} 
0.6 & z \geq 0 \text{ km} \\
-0.12z + 0.6 & -5.0 \text{ km} < z < 0 \text{ km} \\
1.2 & z \leq -5.0 \text{ km} 
\end{cases}
\]

where \(z\) is elevation with respect to mean sea level.

The rise time \(t_{95}\) is defined to be the time it takes for 95 percent of the final slip to occur, that is, \(D(t_{95}) = 0.95D_{\text{final}}\). For this slip-time function, the rise time is given by

\[
t_{95} = 1.745 \frac{D_{\text{final}}}{V_{\text{max}}}
\]

1906 San Francisco Source Models

The source models used for the 1906 San Francisco earthquake simulations are summarized in figures 13 through 15. The source model specifies the final slip and time of slip initiation for each point on the fault surface. There are nine source models. Two of the source models, named Song and SongMod, are models of the 1906 earthquake. The other seven source models are hypothetical scenarios, chosen to characterize the potential variability in ground motions for future events of this size on the northern San Andreas Fault.

Our Song source model (fig. 14) is based on Song and others (2008). The original Song and others source model only contains slip variations at long length-scales and a gross estimate of the rupture speed. It specifies slip on 10 km along-strike and 12 km down-dip patches, with the average rupture speed determined over portions of the fault ranging in length from 40 km to 120 km. The first 120 km of rupture north of the hypocenter propagates faster than the shear-wave speed before dropping below the shear wave speed for the remaining northward propagation. Our Song source model is the same as the original Song and others source model except that the rupture speed in the upper 5 km of the fault tapers down to 60 percent of the original Song and others value.

The SongMod source model (fig. 14) is a modified version of the Song model, which is better suited for modeling ground motions in the period range of interest (periods of 1 to 2 seconds and longer). Three modifications are applied to the Song model. First, shorter length scale variations in final slip are added to the model. The additional variations in final slip are randomly phased, and passed through a low-pass spatial filter whose amplitude falls off as the inverse square of the wavenumber for large wavenumbers. The filter’s correlation length (equal to the reciprocal of the filter’s corner wavenumber) is \(10^{0.5M_{\text{w}}-2}\) kilometers, which evaluates to 79 km or 89 km for magnitude 7.8 or 7.9 events, respectively.
Second, local variations in rupture speed are added to the model, so the rupture propagates more quickly through regions with larger final slip. This is done by adjusting the slip initiation times according to

\[ t_r = t_{\text{orig}} - t_{\text{shift}}(\text{slip}) \]

where \( t_r \) is the slip-initiation time, \( t_{\text{orig}} \) is the original slip-initiation time from the Song model, and \( t_{\text{shift}}(\text{slip}) \) is a timing perturbation that scales linearly with final slip such that \( t_{\text{shift}}=1.0 \text{ s} \) at the location of maximum final slip and \( t_{\text{shift}}= 0 \) at locations where the final slip is equal to its average value.

Third, the average slip was reduced from 4.3 m to 3.0 m, and a few small perturbations to the large length scale distributions of slip and rupture speed were made. The smaller average slip results in a moment magnitude 7.8 event compared to a moment magnitude 7.9 event. This yields a better fit to the Boatwright and Bundock intensities.

The three hypothetical scenarios SongModHypoN, SongModHypoC, and SongModHypoS (fig. 14) use the same distributions of slip and rupture speed as the SongMod source model, but they have alternative hypocenter locations: one near the northern end of the rupture (Rockport), one near the center of the rupture (Bodega Bay), and one near the southern end of the rupture (San Juan Bautista), respectively.

The four hypothetical scenarios RandomHypo06, RandomHypoN, RandomHypoC, and RandomHypoS (fig. 14) use the same hypocenters as the corresponding SongMod scenarios, but with a different distribution of final slip. The slip distribution was produced by generating a slip distribution with random phase at all wavelengths and passing it though the wavenumber squared low-pass spatial filter mentioned above, while constraining the average slip to be the same as in the SongMod model. Several such slip distributions were generated. The one selected for use has a spatial distribution of final slip that differs substantially from the SongMod source model.

For our slip time function, the rise time \( t_{95} \) is in the range of 3 to 4 seconds for slip in the range of 4 to 8 m when \( C_r = 1.2 \). This is consistent with the rise times in kinematic source inversions extrapolated to events of this size.

Figure 13. Hypocenter locations for the 1906 San Francisco earthquake simulations. Scenario epicenters are shown on maps of the entire rupture length (top left) and San Francisco Bay vicinity (top right). The thick red line shows the extent of the ruptures in the simulations, and the red stars denote the scenario epicenters. The yellow highlighted region shows the urbanized areas within the nine-county San Francisco Bay area.
Figure 14. Source models for the 1906 San Francisco earthquake simulations. The top two panels are the slip models for the 1906 earthquake, and the bottom seven panels are the slip models for hypothetical scenarios. Colors depict the magnitude of right-lateral slip, and the contours show the slip-initiation time (contour interval is 2.0 s). The nonplanar fault geometry has been mapped onto a rectangular grid for display purposes.
Figure 15. Rise times for the 1906 San Francisco earthquake simulations. Colors depict the magnitude of rise time, and the contours show the slip-initiation time (contour interval is 2.0 s). The nonplanar fault geometry has been mapped onto a rectangular grid for display purposes. For locations with zero slip, the rise time is shown as zero. Rise time is defined to be the amount of time required for 95 percent of the final slip to occur. The SongModHypoN, SongModHypoC, and SongModHypoS source models have the same rise time as the SongMod model. The RandomHypoN, RandomHypoC, and RandomHypoS source models have the same rise time as the RandomHypo06 model.

Figure 16. Source models for the 1989 Loma Prieta earthquake simulations. The Beroza model is shown on the left, and the modified Wald model is shown on the right. Colors depict the magnitude of slip, and the contours show the slip-initiation time (contour interval is 1.0 s). The down-dip coordinates of the Beroza source model have been aligned with those of the Wald source model.
1989 Loma Prieta Source Models

The source models used for the 1989 Loma Prieta earthquake simulations are summarized in figure 16. The 1989 Loma Prieta earthquake is a moment magnitude 6.9 event. There are two slip models named Beroza and Wald. The Beroza model is based on Beroza (1991), and the Wald model is based on Wald and others (1991). These two models employ significantly different representations of the source and are based on slightly different datasets.

The Beroza model represents the seismic source by using a spatially continuous slip field, and it is constrained by local stations with relatively few stations northwest of the epicenter. The rupture surface is approximated using a single plane, dipping at 70 degrees to the southwest, which strikes 130 degrees east of north. The rupture speed is 80 percent of the local shear-wave speed.

The Wald model represents the seismic source by using a piecewise uniform slip field; it is constrained by both teleseismic and local stations, with slightly more uniform azimuthal coverage than the Beroza model. The rupture surface is approximated using a single plane dipping at 70 degrees to the southwest, which strikes 128 degrees east of north. The rupture speed is 2.7 km/s, which is about 75 percent of the shear-wave speed in the primary source regions.

Two modifications were made to the Wald source model. First, the slip on the northwestern asperity was reduced by 1 m and the slip on the southeastern asperity was increased by 1 m, making the two regions of high slip more equal in moment release. Second, the slip-time function was replaced with the slip-time function described above, where peak slip rate is proportional to the square root of final slip. Figure 16 shows the modified Wald source model.

For our slip-time function, the rise time $t_{95}$ is in the range of 1 to 2 seconds for final slip in the range 1 to 3 meters, which is consistent with the rise times in the original Beroza (1991) and Wald and others (1991) models.

Modified Mercalli Intensity

Maps of Modified Mercalli Intensity (MMI) show the severity of ground shaking. MMI is usually denoted by a Roman numeral ranging from I to XII. A value of II is barely strong enough to be felt, while a value of X denotes very violent shaking with extreme damage. The MMI is defined by how shaking affects people, structures, and landforms. For our simulations we translate the peak ground velocity (PGV) and peak ground acceleration (PGA) into MMI by using the following empirical relationship developed by Wald and others (2005)

$$MMI = \begin{cases} 
\frac{1}{2} ((VII - MMI_{PGA})MMI_{PGA} + (MMI_{PGA} - V)MMI_{PGV}) & \text{if } MMI < V \\
MMI_{PGV} & \text{if } V \leq MMI < VII \\
MMI_{PGA} & \text{if } MMI \geq VII
\end{cases}$$ (7)

$$MMI_{PGA} = \begin{cases} 
2.20 \log(PGV) + 1.00 & \text{if } MMI < V \\
3.66 \log(PGV) + 1.66 & \text{if } MMI \geq V
\end{cases}$$ and

$$MMI_{PGV} = \begin{cases} 
2.10 \log(PGV) + 3.40 & \text{if } MMI < V \\
3.47 \log(PGV) + 2.35 & \text{if } MMI \geq V
\end{cases}$$ (9)
PGA is in the fraction of gravitational acceleration and PGV is in centimeters per second. To construct a map from instrumental data, ShakeMap adds false stations in areas with sparse station coverage and uses regional attenuation relationships to compute PGV and PGA values for the false stations. Then, the PGV and PGA values from both real and false stations are interpolated onto a regular grid, and the above formulas are used to compute MMI at each grid point.

For broadband simulations, which include periods $T > 0.1s$, the above formulas for MMI can be used. However, for long-period simulations, which include only periods $T > 1.0 s$, the above formulas for MMI produce values that are too low because the long-period simulations have unrealistically low PGA values. To avoid this difficulty, only the formulas that relate MMI to PGV (that is, $MMI_{PGV}$) are used for long-period simulations. For the 1906 San Francisco earthquake, the MMI value is above V in most locations, so the value of MMI is relatively insensitive to the level of PGA.
Aagaard and others—Data Files for Ground-Motion Simulations of the 1906 San Francisco Earthquake and Scenario—Data Series 413