

Joint inversion of gravity and arrival time data from Parkfield: New constraints on structure and hypocenter locations near the SAFOD drill site

S. Roecker,¹ C. Thurber,² and D. McPhee³

Received 30 December 2003; revised 26 March 2004; accepted 30 March 2004; published 18 May 2004.

[1] Taking advantage of large datasets of both gravity and elastic wave arrival time observations available for the Parkfield, California region, we generated an image consistent with both types of data. Among a variety of strategies, the best result was obtained from a simultaneous inversion with a stability requirement that encouraged the perturbed model to remain close to a starting model consisting of a best fit to the arrival time data. The preferred model looks essentially the same as the best-fit arrival time model in areas where ray coverage is dense, with differences being greatest at shallow depths and near the edges of the model where ray paths are few. Earthquake locations change by no more than about 100 m, the general effect being migration of the seismic zone to the northeast, closer to the surface trace of the San Andreas Fault. *INDEX TERMS*: 1219 Geodesy and Gravity: Local gravity anomalies and crustal structure; 7215 Seismology: Earthquake parameters; 7230 Seismology: Seismicity and seismotectonics; 8180 Tectonophysics: Tomography. **Citation**: Roecker, S., C. Thurber, and D. McPhee (2004), Joint inversion of gravity and arrival time data from Parkfield: New constraints on structure and hypocenter locations near the SAFOD drill site, *Geophys. Res. Lett.*, *31*, L12S04, doi:10.1029/2003GL019396.

1. Introduction

[2] The main objective of the San Andreas Fault Observatory at Depth (SAFOD) is to drill and instrument a borehole across the San Andreas Fault Zone near Parkfield, California (Figure 1) in order to address fundamental questions about the processes controlling faulting and earthquake generation within that fault. From a broad perspective, an accurate characterization of the elastic wave speed structure in the vicinity of SAFOD is required to permit meaningful extrapolations of what essentially will be point samples of the fault zone. More parochially, there is a need to constrain the locations of earthquakes near the drill site that will serve as targets for the drilling. The results of previous investigations in Parkfield [e.g., *Michellini and McEvilly*, 1991; *Eberhart-Phillips and Michael*, 1993; *Thurber et al.*, 2003, 2004; S. Roecker et al., Refining the image of the San Andreas Fault near Parkfield, California using a finite difference travel time computation

technique, submitted to *Tectonophysics*, 2003, hereinafter referred to as Roecker et al., submitted manuscript, 2003] demonstrate an intimate link between the presumed wave speed structure and hypocenter location in the vicinity of the drill site.

[3] Tomographic images constructed largely from micro-earthquake P and S wave arrival times corroborate those derived from magnetotelluric (MT) studies [*Unsworth et al.*, 2000; *Thurber et al.*, 2003], but MT observations generally bear only upon the interpretation of a seismically derived image, and not on the image itself. Gravity observations, on the other hand, can add constraints to seismic images, at least to the extent that P wave speeds of rocks can be estimated from their densities. Because of differences in sampling and sensitivity, the simultaneous inversion of gravity and arrival time datasets potentially can improve the accuracy and resolution of models derived independently. Additionally, while current P wave speed models for SAFOD are grossly consistent with those inferred from gravity data, an analysis by D. McPhee (personal communication, 2003) found significant discrepancies between the observed gravity and the predictions of the wave speed models (Figure 2, top). Their results raise questions about the feasibility of generating a subsurface image that satisfies both sets of observations, what that image might reveal about subsurface structure, and what the effects such a model might have on our definition of the seismic zone in general and on the location of the SAFOD target events in particular. We therefore attempted to construct an image from a joint inversion of gravity and arrival time data that would satisfy both types of observation.

2. Description of the Data

[4] Most of the seismic data used in this study are from microearthquakes recorded by the Parkfield Area Seismic Observatory (PASO) array from July, 2000 to October, 2002 (Figure 1). The details of PASO are described elsewhere (e.g., Roecker et al., submitted manuscript, 2003). At its peak deployment, PASO consisted of 59 stations hosting a variety of short period and broadband sensors. These data are augmented by several sources, the most important being UC-Berkeley's High Resolution Seismic Network (HRSN), the USGS CALNET seismic network, downhole sensors recording in the SAFOD pilot hole, and recordings of 17 active sources (shots) in the region. Also included are subsets of data from the active source experiment of *Catchings et al.* [2002] and a small active/passive deployment by *Li et al.* [1997]. The final dataset consists 44,530 P-wave observations and 47,546 S-wave observations from 686 earthquakes, 17 PASO shots, 82 shots (at up to 63

¹Department of Earth and Environmental Science Rensselaer Polytechnic Institute, Troy, New York, USA.

²Department of Geology and Geophysics, University of Wisconsin-Madison, Madison, Wisconsin, USA.

³U.S. Geological Survey, Menlo Park, California, USA.

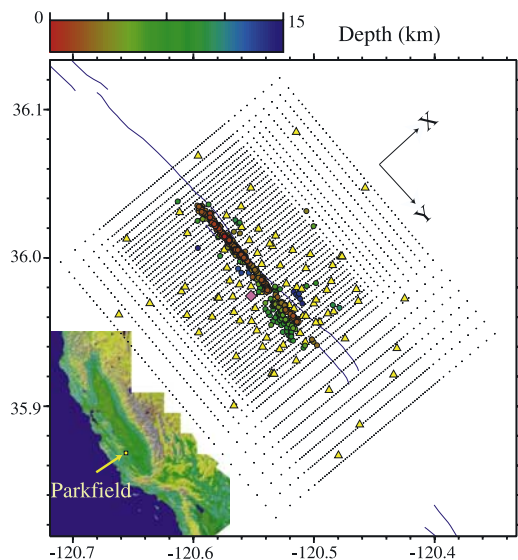


Figure 1. Location map for the Parkfield area. Yellow triangles indicate seismic stations used in this study. The nodes of the coarse scale grid are plotted as small crosses. The SAFOD drill site is indicated by a pink diamond. Blue lines map the SAF surface trace. Circles are epicenters of events located in the best fit arrival-time model; colors filing the circles correspond to depth as shown in the palette.

stations) from the *Catchings et al.* [2002] experiment, three small check shots near the drill site, and three shots and five earthquakes from the *Li et al.* [1997] deployment.

[5] The gravity data (Figure 2) consist of 903 readings compiled from three data sets. Regional data are from *Simpson et al.* [1988] and are spaced at about 1.6 km or less over most of the area. In 1999, closely spaced (~ 40 m) stations were collected along one SW-NE profile crossing the SAFOD pilot hole and colocated with the seismic line of *Catchings et al.* [2002]. In 2000, a grid of stations spaced ~ 400 m apart was collected to more precisely define the gravity field around the proposed drill site. All observed gravity values are referenced to the International Gravity Standardization net (IGSN71) [Morelli, 1974] and reduced to Free Air Anomalies, and Complete Bouguer Anomalies (reduction density 2670 kg/m^3) according to the Geodetic Reference system 1967 [International Association of Geodesy, 1971] and standard formulas [Telford et al., 1976]. The nominal uncertainty on the raw readings due to the instrument, base station, and terrain corrections, is 0.2 mgals.

[6] A final source of information used in the inversion comes from the borehole measurements of in situ V_p and V_s by *Boness and Zoback* [2004]. These measurements are included as soft constraints on the inversion.

3. Analysis

[7] We parameterize the medium by a Cartesian grid with a constant spacing of 200 m between nodes. We use the epicenter of the 1966 Parkfield event (35.96°N , 120.5047°W) as the origin, and rotate the grid 137.2° clockwise so that the positive Y axis points SE along the strike of the fault trace in the vicinity of the drill site (Figure 1). The positive X axis points NE. Z is positive

down and Z = 0 corresponds to mean sea level. A model is specified by assigning a wave speed to each node in the grid; wave speeds between nodes are determined via trilateral interpolation. Generally it will not be useful for all of these points to be independent variables, particularly along the edges of the model where data coverage is sparse, and therefore we accumulate partial derivatives for a coarse mesh (Figure 1) by summing the contributions of the enclosed elements of the fine grid.

[8] Travel times are calculated using the finite difference technique of *Vidale* [1988] as implemented by *Hole and Zelt* [1995]. Hypocenter derivatives are included explicitly even though the corresponding perturbations are not used (they are adjusted in a separate step) to prevent hypocenter perturbations from being artificially minimized (S. Roecker et al., submitted manuscript, 2003). The locations of shots are assumed known, but the origin time is permitted to vary to allow for timing errors or near-source delays. Wave speed perturbations are subject to a combination of a priori (through constraint equations) and a posteriori (moving window average) smoothing to discourage poorly constrained short wavelength features from appearing in the model. Following the experience of previous investigators [e.g., *Eberhart-Philips et al.*, 1995], we use S wave observations in combination with P to solve for V_p/V_s rather than V_s directly. Thus, while we do not presume any explicit relation between shear wave speeds and density, they are implicitly linked through this parameterization.

[9] Gravity observations are included by presuming a functional relationship between density and wave speed based on laboratory measurements. Assuming basement

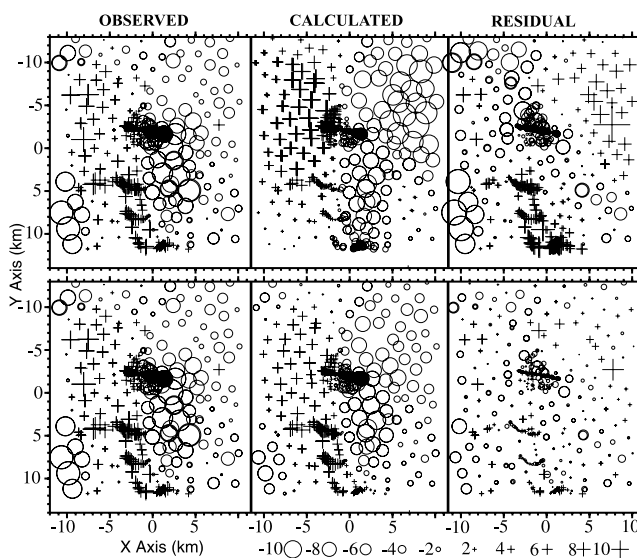


Figure 2. Gravity data used in this analysis. Columns from left to right correspond to the observed Bouguer anomalies, the anomalies calculated from a starting wave speed model, and the difference between the two. Circles are negative, pluses are positive, with size corresponding to magnitude in mgals as shown at the bottom of the figure. The upper row corresponds to the best-fit arrival time model, and the lower row to the final model produced by the simultaneous inversion of arrival time and gravity data. The orientation of the X and Y axis is shown in Figure 1.

rocks are located wherever seismic velocities exceed 6.0 km/s, P wave speeds above about 6.0 km/s follow the *Christensen and Mooney* [1995] relationship for crystalline rocks at a depth of 10 km ($\rho = 989.3 + 289.1 \cdot V_p$). P wave speeds below about 5.5 km/s use the relationship of *Gardner et al.* [1974] ($\rho = 0.23V_p^{0.25}$) that was empirically derived from sedimentary rocks. The uncertainty in the calculated densities is about 0.1–0.2 gm/cc. A polynomial is used to derive density in a smooth fashion between 5.5 and 6.0 km/s; values in this range are close to 2670 kg/m³. A density is thus computed for each node of the fine grid tomography model, and this value is presumed constant within a 200 m dimension cube with the node at the center. Because we model variations in gravity rather than absolute values, we demean the densities at each depth level in the grid. Note that demeaning is equivalent to assigning the same mean density at a given depth to all points outside the model, which also serves to reduce the influence of unmodeled densities near the edges of the structure. The anomaly generated at any observation point by the resulting $\delta\rho$ is calculated using the formula for a rectangular parallelepiped given by *Talwani* [1973].

[10] Although the construction of a joint gravity-slowness inversion is straightforward in principle, implementation requires additional thought. Because a gravity observation is affected by the density of every element in the model, the addition of gravity to the arrival time inverse problem has two potentially deleterious effects: (1) the matrix of normal equations is no longer sparse, and (2) the inherent non-uniqueness of gravity modeling requires the application of additional constraints to prevent unreasonable variations in structure from degrading the image. We therefore tested several strategies in implementing a joint inversion. We first tried alternating independent inversions of the arrival time and gravity observations, but, because this procedure is not sensitive to potential inconsistencies in the data, oscillatory perturbations ensued. We next tried to adjust a best-fitting arrival time model to fit the gravity observations alone, and then use the resulting slowness perturbations as "soft" constraints on the arrival time inversion. While the iterated perturbations generally were more stable than in the alternating approach, the gravity observations were only marginally better fit than they were by the initial model. This is most likely because the gravity observations themselves are not explicitly included in the inversion.

[11] We finally attempted a complete simultaneous inversion of arrival times and gravity observations. Initial trials showed that perturbations to the variables most sensitive to gravity data could be stabilized by, in addition to smoothing by proximal node coupling, adding constraints, that encourage the perturbed model to stay close to the starting model (a single grid is used for the joint inversion). Of all three different starting models we attempted (a one-dimensional model, a best-fit gravity-only model, and a best-fit arrival time-only model), the model derived from the best-fit arrival time-only starting model also best-fit the entire data set. The other starting models resulted either in significantly poorer fits to the arrival time data or tended to converge at an

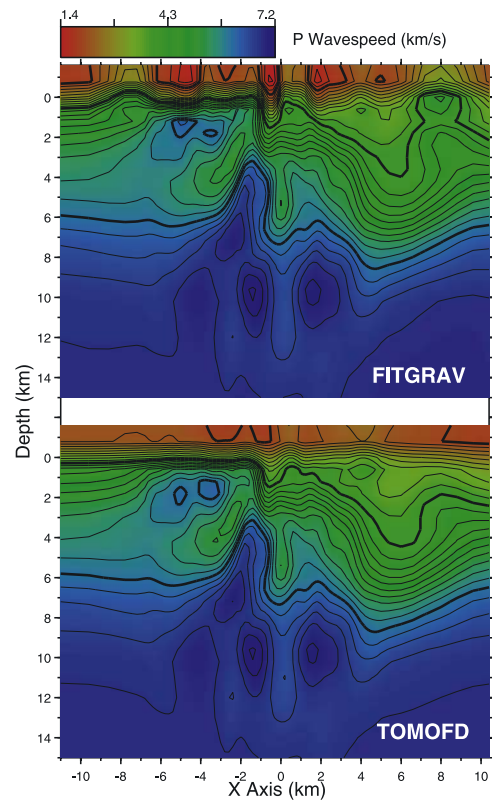


Figure 3. Vertical cross-section of P wave speeds for the best-fit arrival time model (lower) and through the combined gravity and arrival time model (upper) taken at $Y = -2.2$ km, which is close to the drill site. Contour interval is 0.2 km/s. Hypocenters of earthquakes used in the inversions located within 500 meters of the section are plotted as open circles. Locations of the trace of the San Andreas Fault (SAF) and the Pilot Hole (PH) indicated by labeled vertical lines.

excruciatingly slow rate. In general, this result is due to perturbations controlled mostly by gravity observations being much more underdetermined (and hence requiring the additional stability constraints). We note that this strategy of sequentially freeing variables based on the degree to which they are under or over determined was implemented by previous investigators under the rubric "progressive inversion" [e.g., *Roecker*, 1982; *Pavlis and Booker*, 1983].

[12] The preferred model shown here was achieved after three iterations of the simultaneous inversion; further iteration resulted in no significant reduction in the variance of the misfit. The a posteriori data variance is 0.026 s² for the arrival time data and 0.103 mgal² for the gravity observations (Figure 2), both of which are close to the anticipated noise levels (0.02 s² and 0.10 mgal²) of the respective datasets. Thus, we conclude that it is possible to generate a model consistent with both sets of observations.

4. Discussion

[13] A comparison of the model generated by inversion of arrival times only (i.e., the starting model) with that generated from the joint inversion of arrival times and

gravity observations (Figure 3) shows that the main features of the arrival time model remain intact, and hence the interpretations offered by previous investigators of these variations [e.g., Thurber *et al.*, 2003; S. Roecker *et al.*, submitted manuscript, 2003] remain the same. Specifically, the regions near the fault zone are characterized by low V_p and high V_p/V_s , trends interpreted as evidence of a fluid-rich fault zone [Thurber *et al.*, 2003]. To the SW of the fault, the higher V_p and lower V_p/V_s values typical of granites are present, consistent with the presence of Salinian basement [Page, 1981]. The principal differences between these models are in the shallow structure (<2 km depth) and in the regions near the edges of the model (wave speeds below 4 km are virtually identical). Because these areas tend to be poorly resolved by the arrival time data, it would appear that the primary reason for the original misfit was not so much that the arrival time images were wrong as that the arrival time data were incapable of discerning structure in some regions that influence gravity observations.

[14] One might expect that the overall similarity between the arrival time and gravity/arrival time models would imply that the corresponding hypocenter locations would not be significantly different. Indeed, relocated epicenters changed by less than a few 10's of meters along strike (Y axis). However, the characterization of the gradient across the fault has already been shown to be important for locations [e.g., Rittger *et al.*, 2001], and the steeper (by ~12%) gradient across the fault in the upper few km of the gravity/arrival time model does cause the epicenters in general to migrate to the NE (i.e., along the X axis and closer to the San Andreas Fault trace at $X = 0$) by about 100 meters. In particular, the SAFOD target event of 17 March, 2001 relocated 146 m to the NE, outside of the 2σ uncertainties determined by Rittger *et al.* [2001]. One might expect that the perturbations in near surface structure would have a similar effect on depth estimates, and indeed these changes are on the order of 100 m, but these changes are nevertheless still within the expected uncertainties for this parameter [e.g., Thurber *et al.*, 2004].

5. Conclusions

[15] We have demonstrated that it is possible to generate a wave speed model of the region near the SAFOD drill site that is consistent with both gravity and arrival time observations. The arrival-time based model did not need to change fundamentally in order to accommodate gravity observations, a result that increases confidence in the original arrival time tomography model. Nevertheless, gravity data do provide additional information about the nature of the wave speed gradient across the fault, which in turn helps to refine our estimates of the hypocenter locations needed to guide the SAFOD drilling.

[16] **Acknowledgments.** This work was supported by NSF Continental Dynamics Program grants EAR 98-14192 (UW) and EAR 98-14155 (RPI). Equipment was provided by the IRIS/PASSCAL program. We received significant help from Marcos Alvarez, Tim Parker, and Noel Barstow of PASSCAL, Glen Offield, Frank Vernon, and Jennifer Eakins of UCSD, Shirley Baher, Bill Unger, Haijiang Zhang and Neal Lord of UW, Mitch Gold, Jeremiah Armitage, Jacqueline Krajewski, and Mike Klusman of RPI, and Steve Hickman, Tom Burdette, Gary Fuis, John Langbein, and Robert Summers of the USGS. We thank the landowners in the Parkfield area for granting us access to their property.

References

- Boness, N. L., and M. D. Zoback (2004), Stress-induced seismic velocity anisotropy and physical properties in the SAFOD Pilot Hole in Parkfield, CA, *Geophys. Res. Lett.*, *31*, doi:10.1029/2003GL019020, in press.
- Catchings, R. D., M. J. Rymer, M. R. Goldman, J. A. Hole, R. Huggins, and C. Lippus (2002), High-resolution seismic velocities and shallow structure of the San Andreas Fault zone at Middle Mountain, Parkfield, California, *Bull. Seismol. Soc. Am.*, *92*, 2493–2503.
- Christensen, N. I., and W. D. Mooney (1995), Seismic velocity structure and composition of the continental crust; a global view, *J. Geophys. Res.*, *100*, 9761–9788.
- Eberhart-Phillips, D., and A. J. Michael (1993), Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield region, central CA, *J. Geophys. Res.*, *98*, 15,737–15,758.
- Eberhart-Phillips, D., W. Stanley, B. Rodriguez, and W. Lutter (1995), Surface seismic and electrical methods to detect fluids related to faulting, *J. Geophys. Res.*, *100*, 12,919–12,936.
- Gardner, G. H. F., L. W. Gardner, and A. R. Gregory (1974), Formation velocity and density: The diagnostic basics for stratigraphic traps, *Geophysics*, *39*, 770–780.
- Hole, J. A., and B. C. Zelt (1995), 3-D finite-difference reflection travel-times, *Geophys. J. Int.*, *121*, 427–434.
- International Association of Geodesy (1971), Geodetic reference system 1967, *Int. Assoc. Geod. Spec. Publ.* *3*, 116 pp., Paris.
- Li, Y. G., W. L. Ellsworth, C. H. Thurber, P. E. Malin, and K. Aki (1997), Fault-zone guided waves from explosions in the San Andreas Fault at Parkfield and Cienega Valley, California, *Bull. Seismol. Soc. Am.*, *87*, 210–221.
- Micheline, A., and T. V. McEvelly (1991), Seismological studies at Parkfield, I. Simultaneous inversion for velocity structure and hypocenters using cubic B-splines parameterization, *Bull. Seismol. Soc. Am.*, *81*, 524–526.
- Morelli, C., (Ed.) (1974), The international gravity standardization net 1971, *Int. Assoc. Geod. Spec. Publ.* *4*, 194 pp., Paris.
- Page, B. M. (1981), The southern Coast Ranges, in *The Geotectonic Development of California*, edited by W. G. Ernst, pp. 329–417, Prentice-Hall, Englewood Cliffs, N. J.
- Pavlis, G. L., and J. R. Booker (1983), Progressive multiple event location (PMEL), *Bull. Seismol. Soc. Am.*, *73*, 1753–1777.
- Rittger, K., S. W. Roecker, C. Thurber, M. Gold, L. Powell, K. Roberts, G. Offield, and J. Eakins (2001), Seismic imaging of the San Andreas Fault at Parkfield, California using data from PASO-DOS, *Eos Trans. AGU*, *82*(47), Fall Meet. Suppl., S41A-0575.
- Roecker, S. W. (1982), Velocity structure of the Pamir-Hindu Kush region: Possible evidence of subducted crust, *J. Geophys. Res.*, *87*, 945–959.
- Simpson, R. W., R. C. Jachens, and C. M. Wentworth (1998), Average topography, isostatic residual gravity, and aeromagnetic maps of the Parkfield region, California, *U.S. Geol. Surv. Open File Rep.*, *89-13*, 3 sheets, scale 1:250,000.
- Talwani, M. (1973), Computer usage in the computation of gravity anomalies, *Methods Comput. Phys.*, *13*, 343–389.
- Telford, W. M., L. P. Geldart, R. E. Sheriff, and D. A. Keys (1976), *Applied Geophysics*, 860 pp., Cambridge Univ. Press, New York.
- Thurber, C., S. Roecker, K. Roberts, M. Gold, L. Powell, and K. Rittger (2003), Earthquake locations and three-dimensional fault zone structure along the creeping section of the San Andreas Fault near Parkfield, CA: Preparing for SAFOD, *Geophys. Res. Lett.*, *30*(3), 1112, doi:10.1029/2002GL016004.
- Thurber, C., S. Roecker, H. Zhang, S. Baher, and W. Ellsworth (2004), Fine-scale structure of the San Andreas fault zone and location of the SAFOD target earthquakes, *Geophys. Res. Lett.*, *31*, L12S02, doi:10.1029/2003GL019398.
- Unsworth, M., P. Bedrosian, M. Eisel, G. Egbert, and W. Siripunvaraporn (2000), Along strike variations in the electrical structure of the San Andreas Fault at Parkfield, California, *Geophys. Res. Lett.*, *27*, 3021–3024.
- Vidale, J. E. (1988), Finite-difference calculation of travel times, *Bull. Seismol. Soc. Am.*, *78*, 2062–2076.
- D. McPhee, U.S. Geological Survey, MS 989,345 Middlefield Rd., Menlo Park, CA 94025, USA.
- S. Roecker, Department of Earth and Environmental Science Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590, USA. (roecks@rpi.edu)
- C. Thurber, Department of Geology and Geophysics, University of Wisconsin-Madison, 1215 West Dayton St., Madison, WI 53706, USA.