

Appendix C

Seismicity Reviews

C-1: Report: Central California Offshore Earthquake Assessment
Phases 1-3, Dr. Clifford Thurber

C-2: Central California Offshore Earthquake Location Assessment, Dr. Felix
Waldhauser

Appendix C-1

Report: Central California Offshore Earthquake Assessment
Phases 1-3
Dr. Clifford Thurber

Phases 1 and 2 Report: Central California Offshore Earthquake Assessment

PG&E Contract No. 2500174779

Clifford Thurber

March 18, 2009

I report on my analysis of the locations of events offshore California, 35.1° to 35.3° N, with focus on an apparent NW-SE lineation identified from the tomographic inversion and relocation results of Dr. Jeanne Hardebeck, USGS (the "target events"). The tasks for Phases 1 and 2 are as follows:

Phase 1 - assessment of Jeanne Hardebeck's location/tomography results

a) Evaluate the parameters and other aspects used in the completed tomographic inversion work

Set up and run tomoDD inversion using original parameters

Assess sensitivity of results to parameter changes: relative weighting of data types, inversion sequence, event linkage

b) Assess suitability of tomographic velocity model grid

c) Evaluate potential for S minus P time moveout analysis to test reality of offshore lineations

d) Provide report including my preliminary professional assessment of the off-shore lineation

Phase 2 - assessment of offshore earthquake location uncertainties

a) Carry out relocation analysis using two different, existing regional 3D velocity models

b) Repeat tomographic inversions using these two different models as starting models

c) Delete-one tests of sensitivity of locations to observing stations

d) Evaluation of realistic location uncertainty estimates

e) Provide report including my detailed professional assessment of the robustness of the location results for the off-shore events

My first step in Phase 1 was an evaluation of the parameters used in and other aspects of Hardebeck's tomographic inversion work using tomoDD (Zhang and Thurber, 2003). In my opinion, the inversion grid adopted by Hardebeck is too finely spaced, with 2 km grid intervals in

all three Cartesian directions throughout most of the model volume. The available data cannot adequately constrain such a spatially fine model. Hardebeck appears to have compensated adequately for the excessively fine grid by using relatively high model perturbation smoothing weights. I was unable to replicate her tomographic inversion due to her dense grid parameterization (nearly 25,000 nodes) combined with undetermined compiler differences. The results I obtained using a decimated version of Hardebeck's model, however, show moderately small location differences for the target events, but unstable locations for a number of other events (Figure 1a, b, and c). The location results are insensitive to the details of the tomoDD parameter choices, as tests on a broad range of key parameter values for the decimated model do not result in significant location differences for the target events. The parameters tested are the model perturbation smoothing weight (20, 200, 2000) and linking distance (60 to 1 for cross-correlation data, 60 to 2 for catalog data). A smoothing weight of 2000 yields locations that are indistinguishable from those for 200 (the reference value, used by Hardebeck). For a smoothing weight of 20, some target events shift on the order of a kilometer, but most move less than 500 m. Changes in linking distance make at most minor changes to the locations, except for a small number of unstable events.

My summary evaluation of Hardebeck's tomography work and the associated location results is that a more conservative (i.e. coarser) velocity model grid would have been more appropriate, but with the heavy smoothing applied, stable tomography results were obtained. My tests using a decimated version of Hardebeck's starting model show very little sensitivity to a broad range of parameter values. The use of different models and different data subsets results in substantially larger hypocenter changes for the target events, however, as discussed below.

To complete the phase 1 tasks, I made a preliminary evaluation of the potential for using an S minus P time moveout analysis to test the reality of the offshore lineation. Unfortunately, the number of earthquakes with a significant number of S picks is quite small. For the earthquakes associated with the lineation, there is only one event with 7 S picks, and there are 7 events with 5 and 19 events with 4 S picks. A simple moveout analysis also requires events at comparable depths. I have not been able to identify a suitable subset or even a pair of events within the lineation that could be used for a meaningful moveout analysis. There is a potential for determining regional (average) Vp/Vs values from the use of Wadati diagrams on the available S picks. I determined single-event Vp/Vs values ranging between 1.72 and 1.79.

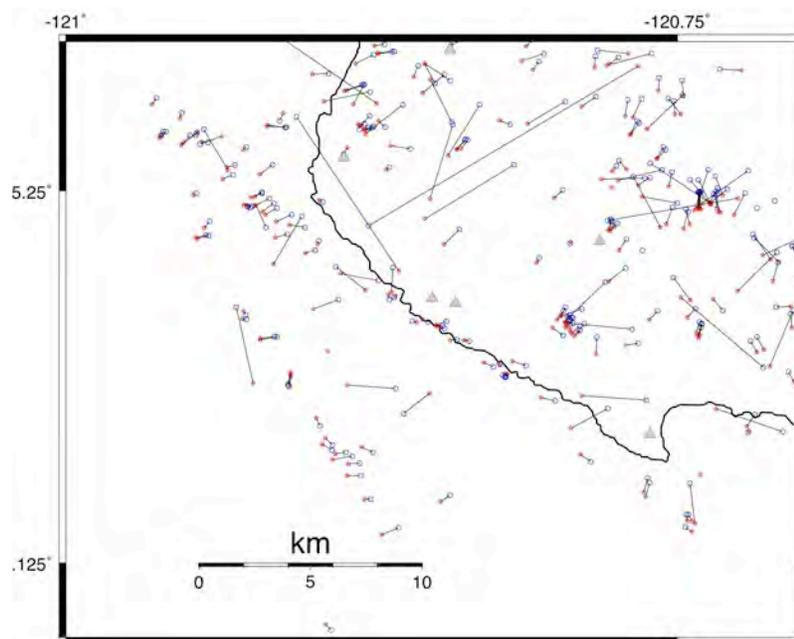


Figure 1a

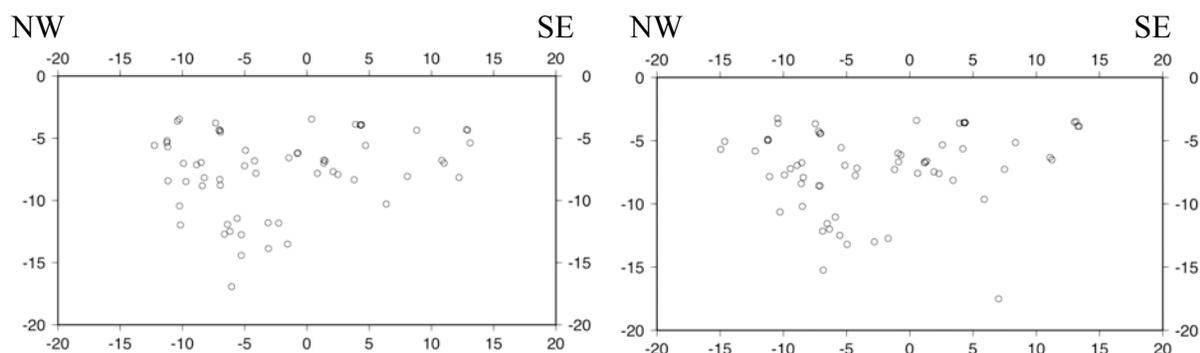


Figure 1b

Figure 1c

Figure 1a compares the epicenters from Hardebeck's tomographic inversion (red stars) to epicenters derived from a tomographic inversion using a decimated version of Hardebeck's model as a starting model with all of the available data (blue circles). Note that there are numerous events with unstable locations, but the vast majority of the events move very little, especially those in the offshore lineation (the target events). The relatively linear character of the offshore events is preserved quite well compared to Hardebeck's results. In subsequent relocation maps, Hardebeck's locations are also used as the reference events (red stars). Figure 1b shows a depth section for the target events. Events within 2 km of a great circle through the target events (center point at 35.21° , -120.86° , azimuth 160° CCW from North) are shown. The locations in section view are comparable to those provided to us by Hardebeck (Figure 1c), although the events that fall within the 2-km-wide zone used for the projection differ slightly between the two plots.

Next I present the results from Phase 2, evaluating the sensitivity of the offshore epicenters to the velocity model and data subsets used to locate them, and delete-one-station tests to derive estimates of relative location uncertainty. In Figures 2 through 5, I show a series of epicenter maps and cross-sections illustrating the relative stability (or instability) of the hypocenters in the region relative to the use of different velocity models and combinations of data.

Figures 2 and 3 show the effect on the locations of using two other regional 3D models as starting models for the tomographic inversion. For Figure 2, the initial model is extracted from the statewide 3D V_p model of Lin et al. (2009), with a constant V_p/V_s value of 1.78 used to generate the starting V_s model. More events show unstable location perturbations compared to the decimated Hardebeck model. The model of Thurber et al. (2006) for the greater Parkfield region is used as the starting model for the results in Figure 3. Significant location shifts are obtained that are largely systematic in character, with most events moving to the northeast by 2 to 4 km. These results must be viewed with caution, however, because the region of interest is near the corner of the Thurber et al. (2006) study region, so the model gridding may not be adequate to represent the 3D structure faithfully.

The above tests largely reflect uncertainties in absolute locations for the earthquakes. An extreme example is Figure 3a, in which a tomographic solution is obtained that results in a very systematic and relatively large location shift. Although we can identify a likely cause for the presumably poor performance of this solution, in general it would be difficult to state with authority whether one set of tomographic results is superior to another.

Previous experience shows that including active-source data, especially for shots recorded at network stations, helps remove some of the nonuniqueness that is inherent in passive-source (earthquake) tomographic inversions, and can help lead to reasonably accurate relative locations (Thurber et al., 2003, 2004). There have been previous active-source experiments in the study region, including off-shore/on-shore work. Such data should be incorporated in future tomographic modeling work. If network picks can be obtained for these shots from archived waveforms or catalog records, or alternatively for known quarries in the region, such data could be quite valuable in constraining absolute locations better than the present earthquake-only dataset can.

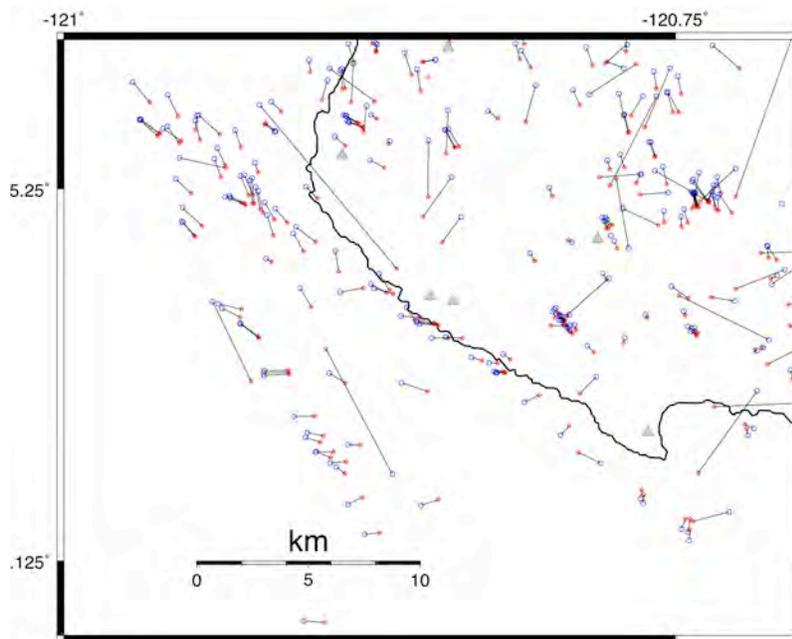


Figure 2a

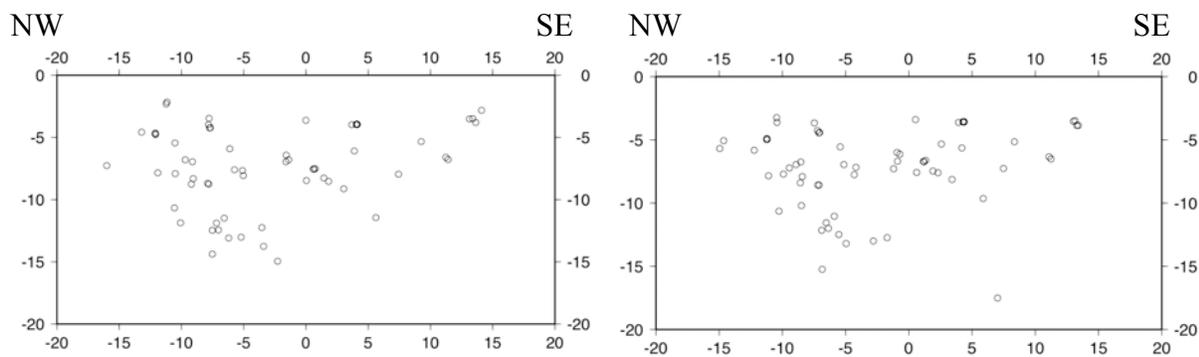


Figure 2b

Figure 2c

Figure 2a compares the epicenters from Hardebeck's tomographic inversion (red stars) to epicenters derived from a tomographic inversion using an initial model extracted from the statewide 3D V_p model of Lin et al. (2009) (blue circles), with a constant V_p/V_s value of 1.78 used to generate the starting V_s model, with all of the available data. There are a surprisingly large number of events with unstable locations, but the majority move less than a kilometer, especially those in the offshore lineation (the target events). The linear character of the offshore events is again preserved reasonably well. Figure 2b shows a depth section for the target events from their Figure 2a locations. As before, events within 2 km of a great circle through the target events (center point at 35.21° , -120.86° , azimuth 160° CCW from North) are shown. The locations in section view are similar to those provided by Hardebeck (Figure 2c), although the events that fall within the 2-km-wide zone used for the projection differ slightly between the two plots.

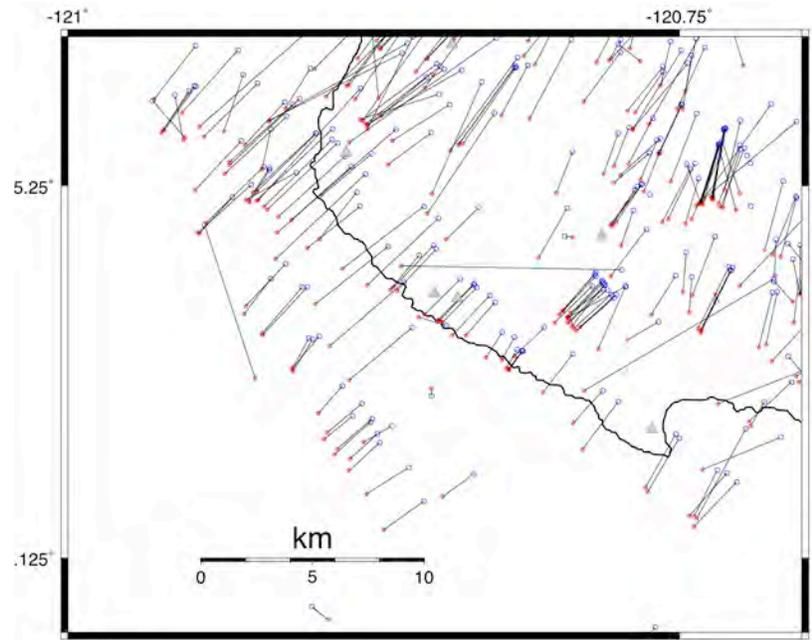


Figure 3a

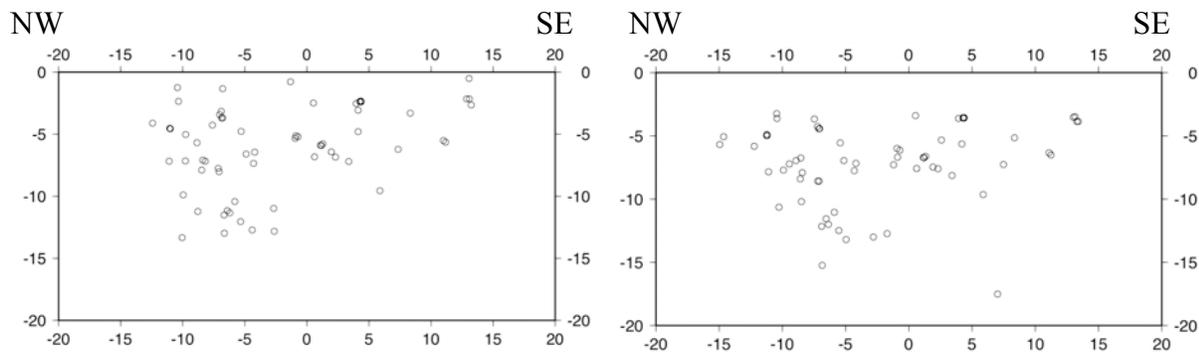


Figure 3b

Figure 3c

Figure 3a compares the epicenters from Hardebeck's tomographic inversion (red stars) to epicenters derived from a tomographic inversion using an initial 3D Vp model from Thurber et al. (2006) for the greater Parkfield region (blue circles), with a constant Vp/Vs value of 1.78 used to generate the starting Vs model, and with all of the available data. There is a systematic northeastward shift of 2 to 4 km for most events. Figure 3b shows a depth section for the target events from their Figure 3a locations. As before, events within 2 km of a great circle through the target events (center point at 35.21°, -120.86°, azimuth 160° CCW from North) are shown. The locations in section view are quite different those provided by Hardebeck (Figure 3c), with the events being systematically shallower in Figure 3b.

Withholding of either the cross-correlation data or the S data also leads to some location scatter, although much less than for the results shown in Figures 2 and 3. In Figure 4, we repeat the tomographic inversion using the decimated version of Hardebeck's model as a starting model, but remove the cross-correlation data from the inversion. Similarly, in Figure 5, we repeat the tomographic inversion using the decimated version of Hardebeck's model, but remove the S-wave data from the inversion. In both cases, the locations are much more stable than I expected. The location shifts for the target events nearest the coast in particular are quite small, less than 500 m in epicenter in many cases. In depth, the location differences are relatively subtle.

It is quite important to assess the level of relative location uncertainty in addition to characterizing the absolute location uncertainty. One way to approach this is to selectively and repeatedly remove data from the inversion and use the resulting variability in the locations to provide an empirical estimate of the relative location uncertainty. I used the strategy of deleting individual stations from the inversion, one at a time, and evaluating the standard deviations of the location coordinates (in the Cartesian reference frame of the cross-sections shown above). I sequentially deleted one of the five coastal stations ARHV, DCHV, LMHV, PBHV, and PSHV, chosen because of their locations and their provision of numerous S-wave picks. Thus, removal of these stations should have the largest possible impact on the stability of the relative locations. This jackknife-like estimate of relative location uncertainty yields values of 140 m in the direction parallel to the lineation, 190 m perpendicular to that, and 280 m in depth. This can be compared to the hypoDD-style (Waldhauser, 2001) location uncertainty estimates provided by tomoDD that are typically 3 to 10 m.

On the basis of the above tests, my firm conclusion is that the offshore seismic lineation feature discovered by Hardebeck is real. Joint tomographic inversions with different starting models yield locations that show systematic shifts but tend to preserve the lineation relatively clearly. The locations are robust to the removal of cross-correlation data or S-wave data from the inversion. An empirical estimate of relative location uncertainty yields values that are much higher than the typical optimistic values produced from the double-difference approach, but they are still sufficiently small to give sufficient confidence to Hardebeck's results. Further geophysical investigation of this feature is warranted, in my opinion.

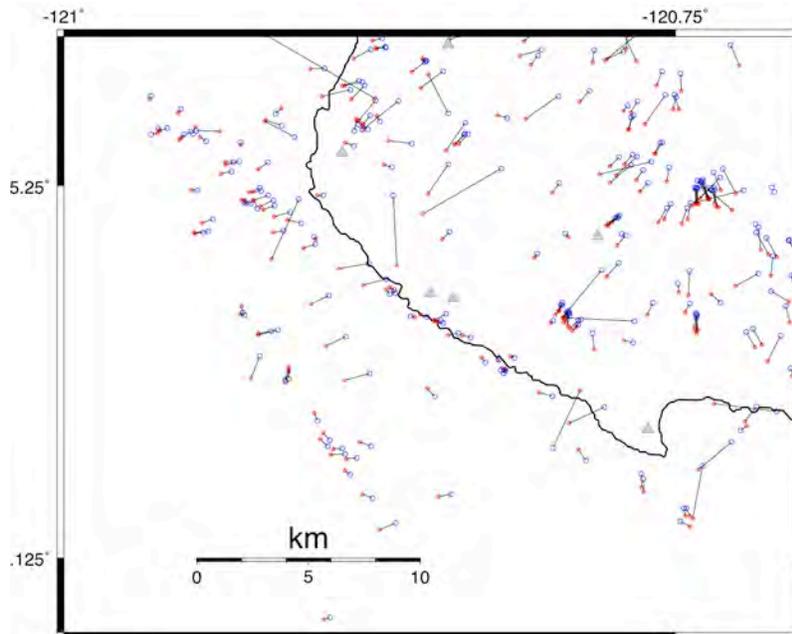


Figure 4a

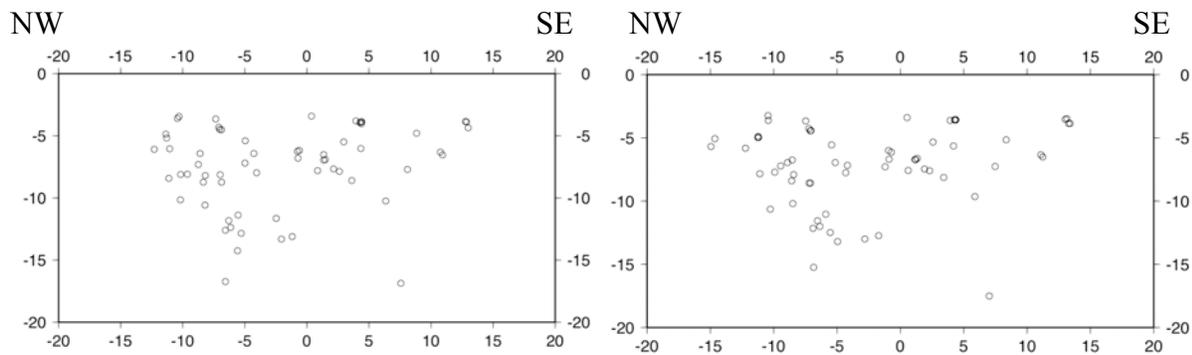


Figure 4b

Figure 4c

Figure 4a compares the epicenters from Hardebeck's tomographic inversion (red stars) to epicenters derived from a tomographic inversion using the same input model but excluding the waveform cross-correlation data (blue circles). The location shifts are generally small, especially for the target earthquakes. This indicates that the catalog absolute and differential times by themselves are adequate to yield well-constrained locations. Figure 4b shows a depth section for the target events from their Figure 4a locations. Again, events within 2 km of a great circle through the target events (center point at 35.21° , -120.86° , azimuth 160° CCW from North) are shown. The locations in section view are quite similar to those provided by Hardebeck.

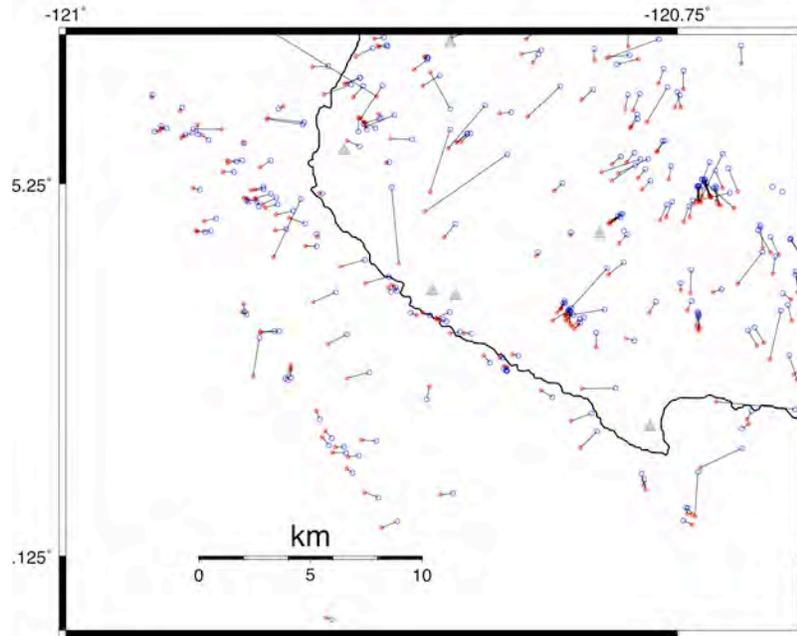


Figure 5a

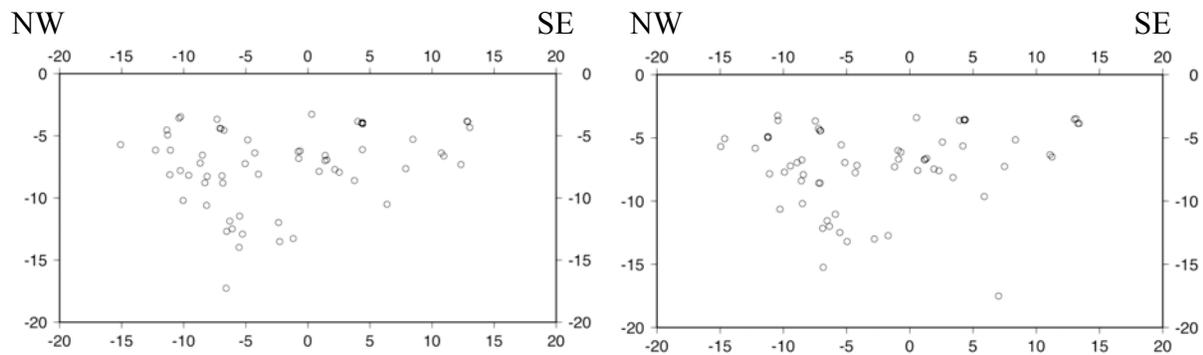


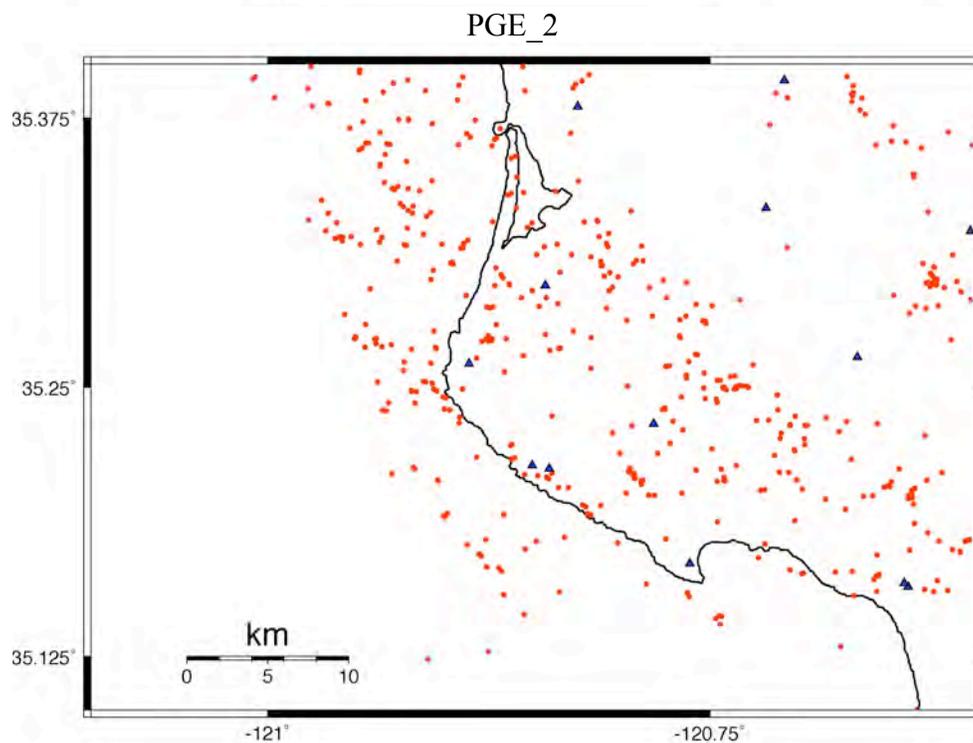
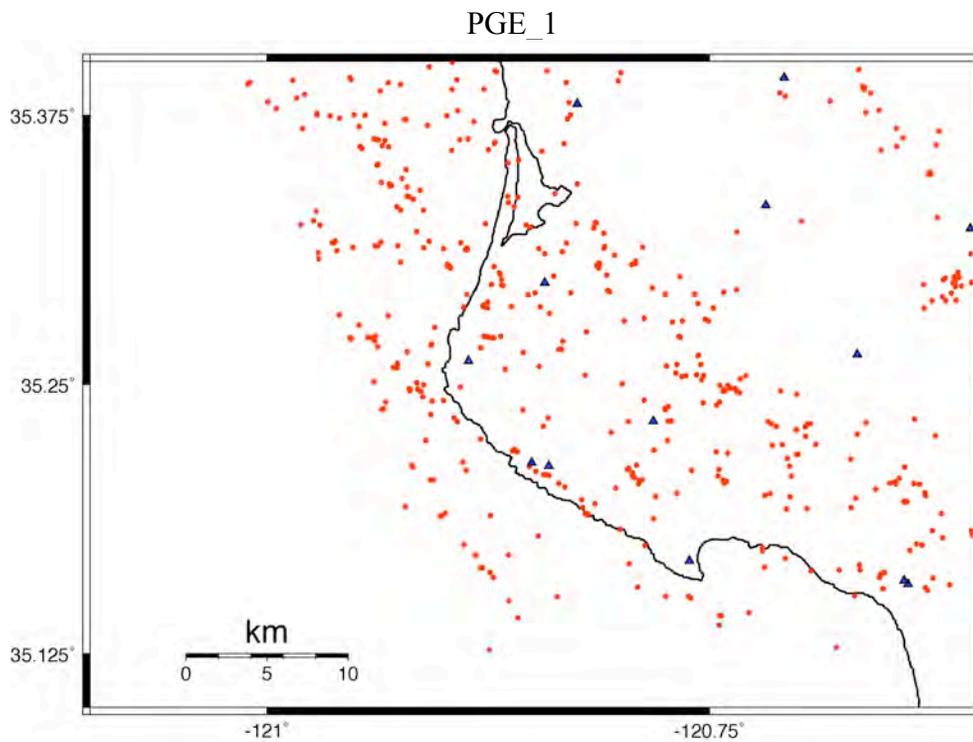
Figure 5b

Figure 5c

Figure 5a compares the epicenters from Hardebeck's tomographic inversion (red stars) to epicenters derived from a tomographic inversion using the same input model but excluding the S-wave data (blue circles). The location shifts are generally small, especially for the target earthquakes. This indicates that the catalog absolute and differential times by themselves are adequate to yield well-constrained locations. Figure 4b shows a depth section for the target events from their Figure 4a locations. Again, events within 2 km of a great circle through the target events (center point at 35.21° , -120.86° , azimuth 160° CCW from North) are shown. The locations in section view are extremely similar to those provided by Hardebeck.

Appendix

For completeness, I include here two early relocation plots that have been superseded by the results in the main body of the report.



References

- Lin, G., C. H. Thurber, H. Zhang, E. Hauksson, P. M. Shearer, F. Waldhauser, T. M. Brocher, and J. Hardebeck, A California statewide three-dimensional seismic velocity model from both absolute and differential times, *Bull. Seism. Soc. Am.*, submitted, 2009.
- Thurber, C., S. Roecker, K. Roberts, M. Gold, L. Powell, and K. Rittger, 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, 10.1029/2002GL016004, 2003.
- Thurber, C., S. Roecker, H. Zhang, S. Baher, and W. Ellsworth, Fine-scale structure of the San Andreas fault and location of the SAFOD target earthquakes, *Geophys. Res. Lett.*, 31, L12S02, doi 10.1029/2003GL019398, 2004.
- Thurber, C., H. Zhang, F. Waldhauser, J. Hardebeck, A. Michael, and D. Eberhart-Phillips, Three-dimensional compressional wavespeed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region, *Bull. Seism. Soc. Am.*, 96, S38-S49, 2006.
- Waldhauser, F., hypoDD: A computer program to compute double-difference hypocenter locations, *U.S. Geol. Surv. Open File Rep.*, 01-113, 25 pp., 2001.
- Zhang, H., and C. H. Thurber, Double-difference tomography: the method and its application to the Hayward fault, California, *Bull. Seism. Soc. Am.*, 93 1875-1889, 2003.

Phase 3 Report: Central California Offshore Earthquake Assessment

PG&E Contract No. 2500174779

Clifford H. Thurber

Department of Geoscience

University of Wisconsin-Madison

1215 W. Dayton St.

Madison, WI 53706

August 18, 2009

Phase 3 Report: Central California Offshore Earthquake Assessment

PG&E Contract No. 2500174779

Clifford Thurber

August 18, 2009

I report on my extended analysis of the locations of events offshore California, 35.1° to 35.3° N, focusing on an apparent NW-SE-trending seismic lineation identified from the tomographic inversion and relocation results of Dr. Jeanne Hardebeck, USGS (the "target" earthquakes). The tasks for Phase 3 are as follows:

Task 1 - Waveform cross-correlation analysis.

- a) Redo the waveform cross-correlation analysis with the same method as Jeanne Hardebeck used.
- b) Redo the waveform cross-correlation analysis with a more sophisticated method.

Task 2 - Replicate Jeanne Hardebeck's exact tomoDD results.

Task 3 - Redo Jeanne Hardebeck's tomoDD analysis with improved cross-correlation data and adjusted velocity model and parameters.

Task 4 - Write a fully documented report providing all the important input and output files.

Task 1 - Waveform cross-correlation analysis

a) Redo the waveform cross-correlation analysis with the same method as Jeanne Hardebeck used.

The first step in Task 1a was the reconstruction of the waveform database used by Hardebeck, who provided me with the raw waveforms in multiple compressed tar files. Once the necessary directory structure was established, and the byte order of the waveform data was determined, I started to work with Hardebeck's cross-correlation code (xcor) and found that it did not function properly on our computer system. Debugging required changing to a different

computer platform and swapping the data byte order. Eventually, debugging succeeded and I was able to replicate the cross-correlation results for 3 stations chosen at random by Hardebeck and provided to me, up to an occasional round-off difference in the sixth decimal place (e.g. 0.670000 instead of 0.669999). Thus, I was able to confirm the validity of the original cross-correlation data.

b) Redo the waveform cross-correlation analysis with a more sophisticated method.

The first step in Task 1b was reorganization of the waveform database for compatibility with an independent cross-correlation code, `corr_sac` (Rowe et al., 2002), along with setting up the necessary auxiliary files. One of `corr_sac`'s advantages over standard cross-correlation computations (such as `xcor`) is the use of dynamic coherency-weighted spectral filtering. The goal is to down-weight incoherent frequency bands while avoiding removal of useful signal, which occurs in standard cross-correlation analysis when a fixed frequency band is used for data filtering. Uncertainties for the differential times are determined using multiple narrow-band estimates. For reliably high correlation values (typically 0.9 and above), the cross-spectral phase method is used to determine a subsample differential time (Poupinet et al., 1984), using multi-taper spectral estimation (Thomson, 1982). Based on the quality of the relocation results shown below, I believe I succeeded in obtaining a very high quality set of cross-correlation data from my analysis with `corr_sac`. I also obtained substantially more high-correlation differential times: Hardebeck obtained 17,271 differential times above her specified acceptance threshold of 0.8 for the PG&E stations (listed in Table 1) using `xcor`, whereas I was able to obtain 41,614 differential times above that same threshold for the same set of stations and events with `corr_sac`. Spot checks of individual differential times values show that differences, where present, are typically on the order of one to a few samples.

Task 2 - Replicate Jeanne Hardebeck's exact tomoDD results

In Phases 1 and 2 of my work, I attempted to replicate the original double-difference (DD) tomography results of Hardebeck using tomoDD (Zhang and Thurber, 2003). Initial problems involved compilation issues, but even when those issues were overcome previously, I was unable to replicate the original results precisely. This problem continued in the current Phase 3 work, until I determined from a detailed comparison of Hardebeck's tomoDD log files

versus mine that Hardebeck's inversion used a different number of stations than my mine did despite my use of the exact files provided to me by her. With the station list identified as the likely source of the discrepancies between the results, I discovered that the station file provided to me post-dated the actual inversion run I was trying to match. Using the matching station file (which contained an error that led to Hardebeck modifying it before sending it to me) and the tomoDD executable compiled on Hardebeck's computer system, I was finally able to exactly replicate her inversion and relocation results. A map view and cross-section of my relocations that match the original are shown in Figure 1. Specifically, my locations in Figure 1 exactly match those of Hardebeck's February 26, 2009, tomoDD results. Thus, I was ultimately able to confirm the validity of Hardebeck's original tomoDD results.

Task 3 - Redo Jeanne Hardebeck's tomoDD analysis with improved cross-correlation data and adjusted velocity model and parameters

My main concern about Hardebeck's original tomoDD results was the extremely fine model grid she employed in her inversions, with 2 km node spacing in all directions throughout most of the inversion model. For a large study region (100 km by 100 km) with a coarse station spacing (generally 10 km and greater), it is my opinion that use of a very fine grid (in this case, 2 km) is potentially pushing the limits for achieving a stable inversion result. Hardebeck used a high smoothing weight in her inversions, which appears to have compensated adequately for the fine gridding. In my Phase 1 and 2 work, I tested the use of two different initial models, one based on Subregion 3 of the statewide tomography study of Lin et al. (2009) with 10 km horizontal gridding, and another from the Parkfield regional model of Thurber et al. (2006) with variable gridding (3 to 20 km). The results I obtained with the former initial model were more stable and the velocity model grid covered the present study area more uniformly, so I ultimately adopted a modified version of the Lin et al. (2009) model grid, with 10 km horizontal gridding covering a smaller region, to use for most of my Phase 3 tomoDD inversions - see Figure 2 a and b. For additional comparisons, I also decimated Hardebeck's model to 4 km and 10 km horizontal gridding (and with coarser vertical gridding as well) and carried out a pair of inversions with each decimated model.

I carried out 24 inversions using tomoDD with varying cross-correlation datasets, smoothing weights, data weighting, and solution step sequences. Key aspects of the inversions

are listed in Table 2. The locations of the "target" earthquakes vary in detail from one inversion to another, but significantly less so than observed in the inversions for my Phase 1 and 2 report. In the present work, however, the use of what I consider to be more appropriate initial models (two with a 10-km grid spacing that is more compatible with the station spacing, and one with a grid interval double that of Hardebeck's) combined with an increased cross-correlation dataset for the PG&E stations resulted in smaller location variations among inversions compared to my previous results.

My preferred results are shown in Figure 3, and the control parameters for that inversion are provided in the Appendix. My preference for this result is based on (1) my opinion that my cross-correlation dataset is of higher quality than that of Hardebeck's and (2) the slightly better data fit for the results shown in Figure 3 compared to that for the results shown in Figures 4 through 7 (the residuals are a few percent smaller on average). The relocations from several other runs with different data and/or initial models are shown in Figures 4 through 7. For Figure 3, absolute and catalog differential P and S waves were used, and the only cross-correlation data included were from my reanalysis of the PG&E station data. For comparison, the location results from a P-wave-only inversion are shown in Figure 4. The differences are subtle. The relocations in Figure 5 were obtained by adding Hardebeck's NCSN and SCSN cross-correlation data, with P and S absolute and catalog differential times included. The differences from Figures 3 and 4 are quite minor. For the relocations in Figures 6 and 7, decimated versions of Hardebeck's initial model with 4 and 10 km gridding were used, and the inversions were otherwise identical to the one producing the results in Figure 4. In the case of Figure 6, the target lineation lies slightly closer to the coastline than the others. Overall, the results are remarkably stable to a variety of significant changes to the inversion.

Comparing my results to those from my replication of Hardebeck's inversion (Figure 1), the "target" earthquake lineation appears to be slightly more sharply defined in Figures 3 through 7 compared to the replicated Hardebeck results, except for the seismic activity off the tip of San Luis Point. Similarly, the seismicity associated with the Hosgri Fault farther offshore also appears somewhat better aligned in the new results compared to Hardebeck's. In depth, the "target" earthquakes in my inversion results appear slightly more clustered into limited depth ranges (2-3 km, 5-7 km, 10-12 km) than in Hardebeck's results, something that is a characteristic of numerous California strike-slip faults (Rubin et al., 1998; Waldhauser et al., 1999; Schaff et

a., 2002; Waldhauser et al., 2004). There is a general theme in these papers that the repeating events indicate relatively steady state loading due to creep surrounding these stuck patches. Nadeau and coworkers have papers that estimate creep rate from the moments and frequency of repeaters, for example Schmidt et al. (2005). I attribute the increased sharpening and clustering to the quality of the new cross-correlation data. As discussed below, it is not possible to state with certainty whether the original locations in Figure 1 or the new locations in Figure 3 are superior to the other.

In my previous report, I discussed delete-one-station estimates of location uncertainty, which provides robust estimates of relative locations, not absolute locations, because the same velocity model is used for all realizations. Those estimates were on the order of 150-300 meters. The theory for estimating absolute location uncertainties where the uncertainty in the velocity model is considered explicitly (Pavlis, 1986) is complex, and such an analysis is beyond the scope of this project. A rough estimate of absolute uncertainty can be obtained by considering the variations in absolute locations resulting from the use of different starting velocity models and different control parameter settings. Figures 8 a and b show examples of location differences between Hardebeck's results and two of my solutions (Run 11s, from Figure 3, and Run 21s, from Figure 7). Shifts of up to about 500 meters are observed for the target earthquakes, with larger shifts for the events along the Hosgri Fault system farther offshore. I consider 500 meters to be a reasonable estimate of the absolute location uncertainty for the target earthquake lineation. The fact that the locations hardly change whether or not S waves are included in an inversion (c.f. Figures 3 and 4) supports this relatively small absolute location uncertainty estimate.

Task 4 - Write a fully documented report providing all the important input and output files

This report contains the description of the key results that have emerged from my work. In addition, I am providing a compressed tar file of the input parameter files and the final hypocenter location files for the inversion results shown in the figures, as well as the cross-correlation results I obtained for the PG&E stations. Upon request, I can provide any other files that are deemed useful.

Conclusions and Recommendations

On the basis of my research in this third phase of work, my previous conclusion is confirmed: that the coastal seismic lineation feature discovered by Hardebeck is real. Joint tomographic inversions with different datasets and control parameters yield locations that show systematic shifts but tend to preserve the lineation quite clearly, and in some cases sharpen its appearance. A similar sharpening of the seismicity trend associated with the Hosgri Fault situated even farther from the coast, adds confidence to the reality of the coastal lineation. Further geophysical investigation of this feature is clearly warranted.

There are a number of reasons why seismic tomography by itself cannot provide definitive results on absolute locations of earthquakes. An obvious reason is data noise, which generally will be worse for S waves. Data noise can lead to both relative and absolute location errors - random pick errors produce location scatter and systematic pick errors cause systematic location shifts. It is also not possible to represent the Earth's structure exactly. The seismic velocity structure of the Earth varies over a wide range of scales, and seismic tomography can only capture part of that range. First-arrival tomography cannot be used to detect and model sharp discontinuities, and it inevitably produces a blurred model. A more subtle problem is that because the seismic velocity model is not perfectly known, first-arrival ray paths will be incorrect in general and sometimes can be seriously in error (e.g., an up-going wave instead of a down-going wave). Finally, the lack of strong constraint on shallow structure due to the absence of very shallow earthquake sources and the sparse distribution of stations makes the well-known depth-origin time trade-off a serious problem for obtaining absolute depths, especially when the earthquakes lie outside the network.

My previous experience with local earthquake tomography shows that including active-source data, especially for shots recorded at network stations, helps remove some of the non-uniqueness that is inherent in passive-source (earthquake) tomographic inversions, and can help lead to reasonably accurate absolute locations (Thurber et al., 2003, 2004). There have been previous active-source experiments in this study region, including offshore/onshore work. The active-source data available through the California earthquake data centers and USGS colleagues was incorporated in my tomographic modeling work (Figure 2a), which included picks from the PG&E (Lines 1 and 3) and Morro Bay refraction lines, but the data I was provided did not include PG&E network picks for these refraction lines. If network picks can be made available

for these shots from archived waveforms or catalog records, they could be quite valuable in constraining absolute locations better than the present dataset can. Alternatively, if it is possible to identify events that can be unambiguously associated with known quarries in the region, as was done by Lin et al. (2007) for southern California, that could provide constraints that are nearly as strong as active-source data. In addition, there may be other active-source data not available through the California earthquake data centers that could be instrumental in improving absolute location constraints, or perhaps a new offshore-onshore dataset could be acquired.

I also strongly recommend an effort to determine whether the location of the lineation derived from my work or Hardebeck's work can be associated with features suggestive of faulting in other sources of data, although the likelihood that this feature represents strike-slip faulting with possibly limited total offset may make such work difficult. Marine seismic reflection data from offshore of Buchon Point might reveal features that could constrain the absolute location of the causative fault. Onshore high-resolution seismic profiling, trenching, or other paleoseismic work might reveal if the fault comes onshore, and if so where. High-resolution potential field data may provide key clues to the absolute position of the fault. There also may be results already published in the gray literature that could yield critical information.

Table 1. List of PG&E stations for which xcor cross-correlation processing was done by Hardebeck and corr_sac cross-correlation processing was done by Thurber.

<u>STATION NAME</u>	<u>LATITUDE</u>	<u>LONGITUDE</u>
AR	35.4632	-120.9783
BL	35.5338	-120.9067
BP	35.7537	-121.1432
CS	35.6005	-121.0940
DC	35.2122	-120.8408
DC2	35.2138	-120.8503
DP	35.2330	-120.7817
EC	35.3333	-120.7182
LM	35.3803	-120.8247
LQ	35.6688	-120.9907
LS	35.2973	-120.8430
ML	35.3225	-120.6025
NP	35.0863	-120.4725
OC	34.8500	-120.4772
OF	35.0160	-120.6055
PB	35.7090	-121.2715
PS	34.9022	-120.6217
RP	35.7772	-121.3058
SH	35.1682	-120.7613
WR	35.4588	-120.8817

Table 2. Information on selected key inversion control parameters and input data for the tomoDD inversions carried out in this project. Modified L is model modified from the Lin et al. Subregion 3 model; Dec. H is decimated version of Hardebeck's model at 4 or 10 km grid spacing. Cross-correlation datasets are as follows: H, Hardebeck's original xcor; T0.7, Thurber's corr_sac with 0.7 threshold; T0.8, Thurber's corr_sac with 0.8 threshold, P+S; T0.8P, Thurber's corr_sac with 0.8 threshold, P only. Inversion sequences included 3, 4, or 5 steps: 3 = hypocenter only, joint inversion, and hypocenter only; 4 = joint, hypocenter, joint, hypocenter; 5 = hypocenter, joint, hypocenter, joint, hypocenter. Data weighting combinations: h = high, e = equal; A = absolute, CT=catalog, CC=cross-correlation (for example, hA means high weight on absolute data, eACT means equal weight on absolute and catalog data, etc.).

Run	Initial V	P or P+S?	CC data	Smooth Wt	Inversion sequence
0	Modified L	P+S	H	20	3hA, 3eACTCC
1	Modified L	P+S	H	20	3eACT, 3hCT, 3hCC
2	Modified L	P+S	H+T0.7	20	3hA, 3hCT, 3hCC
3	Modified L	P+S	H+T0.7	200	3hA, 3eACT, 3hCC
4a	Modified L	P+S	H+T0.7	100	4eACT, 4hCC
4b	Modified L	P+S	H+T0.8	100	4eACT, 4hCC
5a	Modified L	P+S	H	20	4eACT, 4hCC
5b	Modified L	P+S	T0.8P	20	4eACT, 4hCC
6	Modified L	P+S	H+T0.7	200	5eACT, 5hCC
7a	Modified L	P+S	T0.8	200	5eACT, 5hCC
7b	Modified L	P+S	T0.8P	200	5eACT, 5hCC
8a	Modified L	P+S	H+T0.7	200	5hA, 3eACT, 5hCC
8b	Modified L	P+S	T0.8P	200	5hA, 3eACT, 5hCC
9	Modified L	P+S	H	200	5hA, 3eACT, 5hCC
10a	Modified L	P+S	T0.8	100	5hA, 3eACT, 5eCTCC
10b	Modified L	P+S	T0.8P	100	5hA, 3eACT, 5eCTCC
11p	Modified L	P	T0.8P	100	5eACT, 3hCT, 5hCC
11s	Modified L	P+S	T0.8P	100	5eACT, 3hCT, 5hCC
12p	Modified L	P	H+T0.8P	100	5eACT, 3hCT, 5hCC
12s	Modified L	P+S	H+T0.8P	100	5eACT, 3hCT, 5hCC
20p	Dec. H 4	P	T0.8P	100	5eACT, 3hCT, 5hCC
20s	Dec. H 4	P+S	T0.8P	100	5eACT, 3hCT, 5hCC
21p	Dec. H 10	P	T0.8P	100	5eACT, 3hCT, 5hCC
21s	Dec. H 10	P+S	T0.8P	100	5eACT, 3hCT, 5hCC

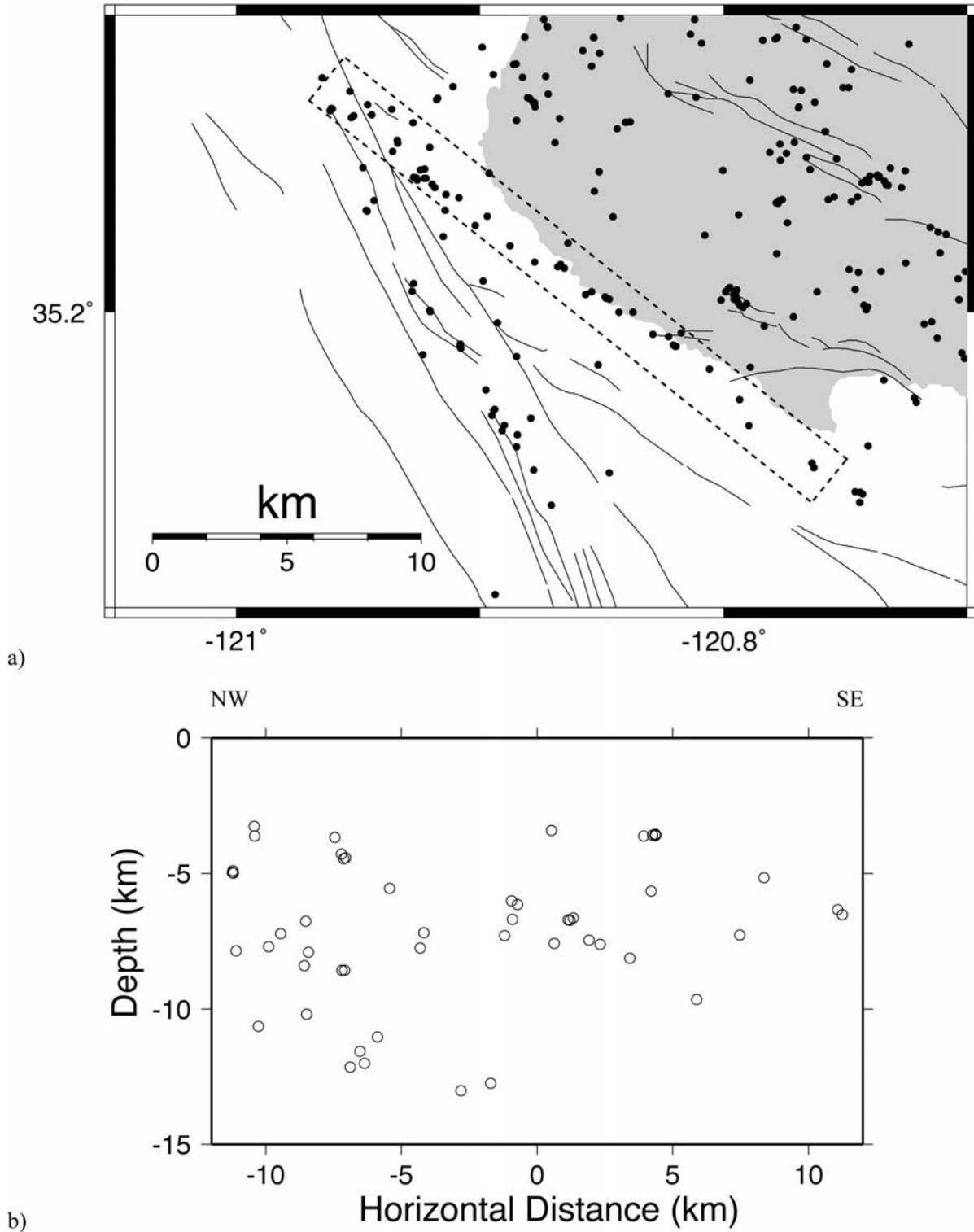


Figure 1. (a) Map view and (b) cross-section (center point at 35.21°, -120.86°, azimuth 129° CW from North, half-width 1 km) of replicated locations from Hardebeck's tomographic inversion. The box in (a) indicates the earthquakes plotted in (b).

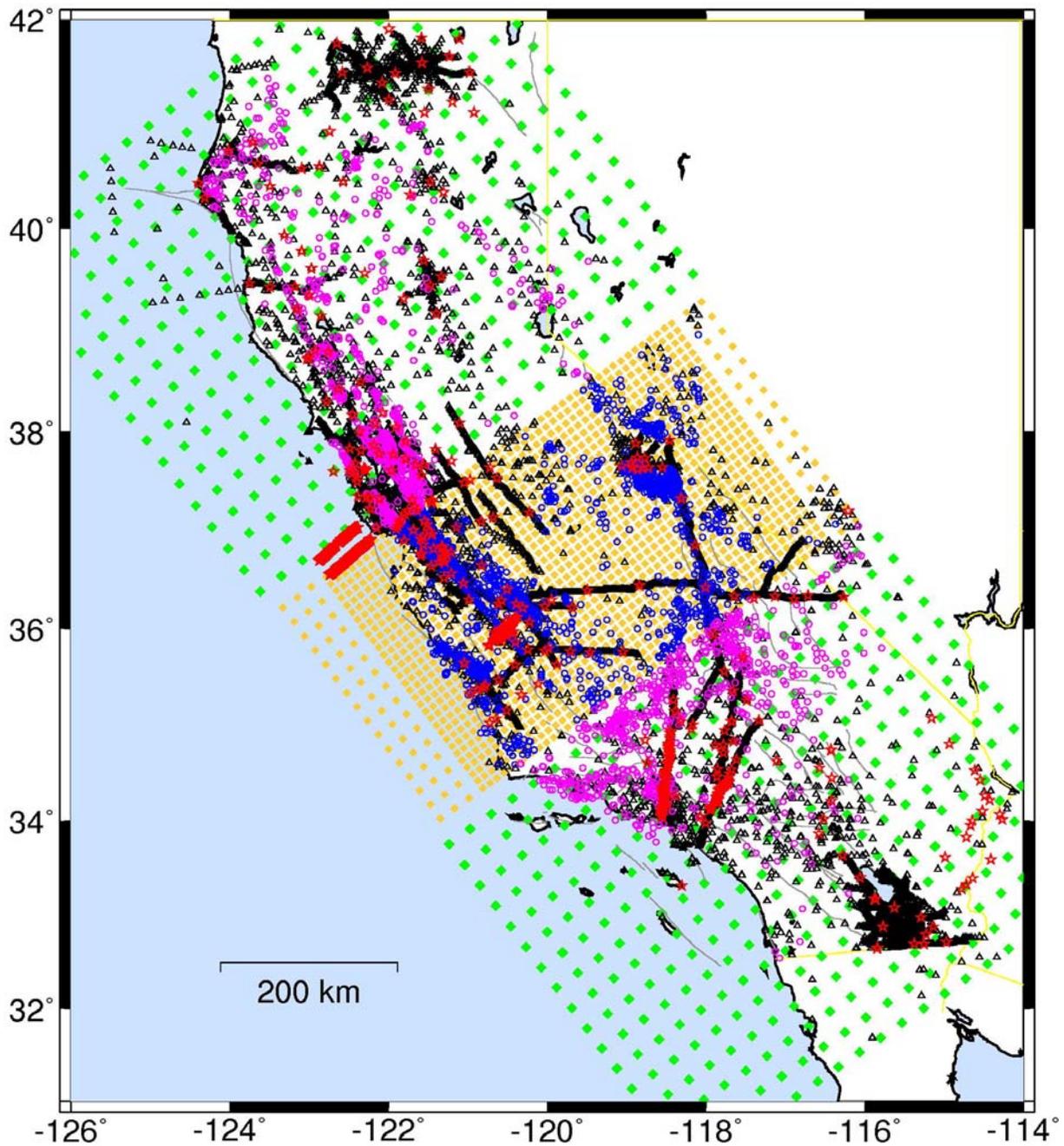


Figure 2a. Map showing the distribution of all the earthquakes (colored circles), explosions (red stars), seismic stations (black triangles) and model grid nodes (colored diamonds) included for the DD tomography inversions for Subregion 3 of Lin et al. (2009), from which the initial velocity model used here was extracted.

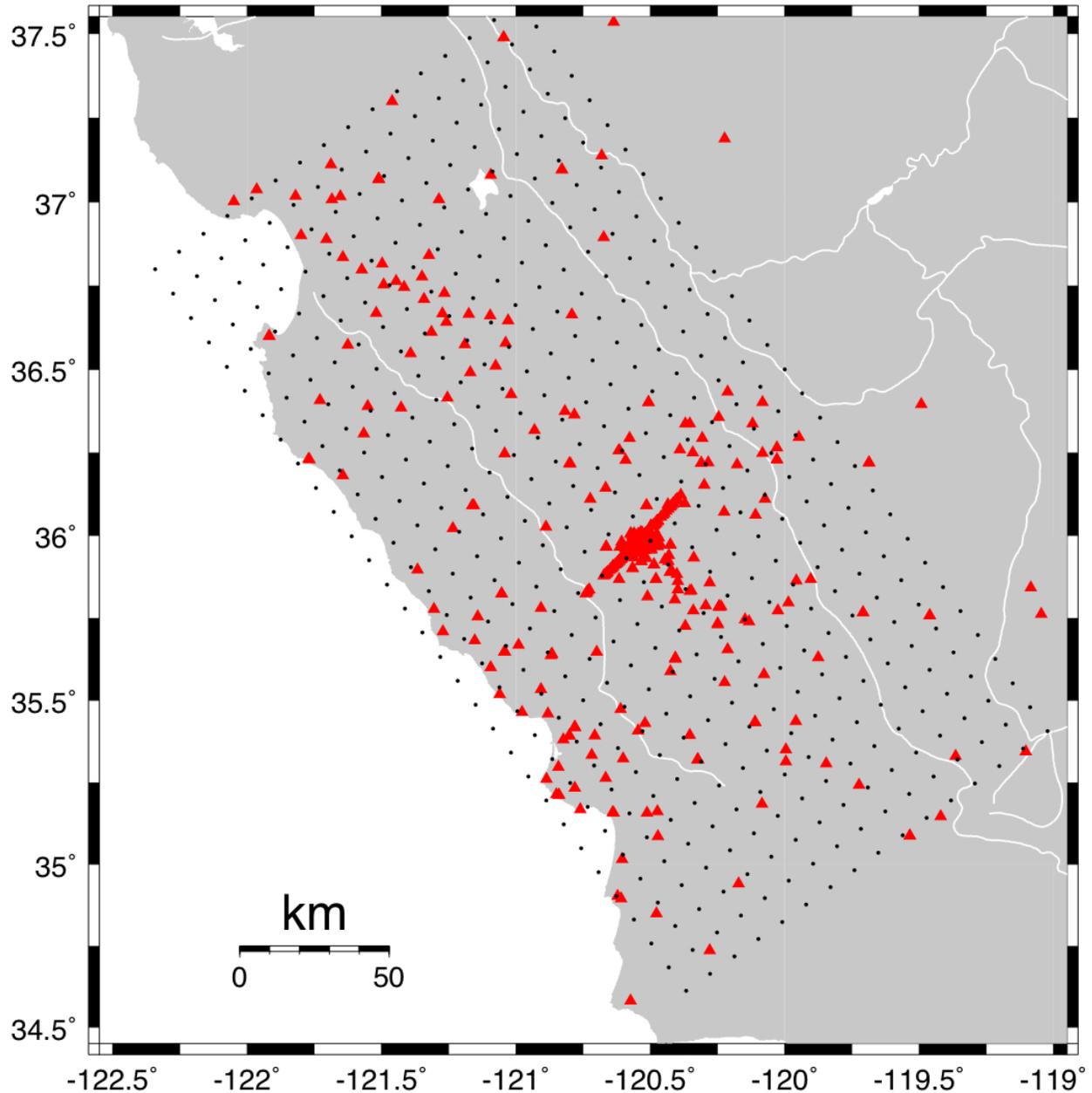


Figure 2b. Map showing the distribution of seismic stations (red triangles) and model grid nodes (black dots) for my tomographic inversions, modified from the work of Lin et al. (2009). The grid has a uniform spacing of 10 km in the horizontal directions, and has nodes at -1, 1, 4, 8, 14, 20, 27, 35, and 45 km depth.

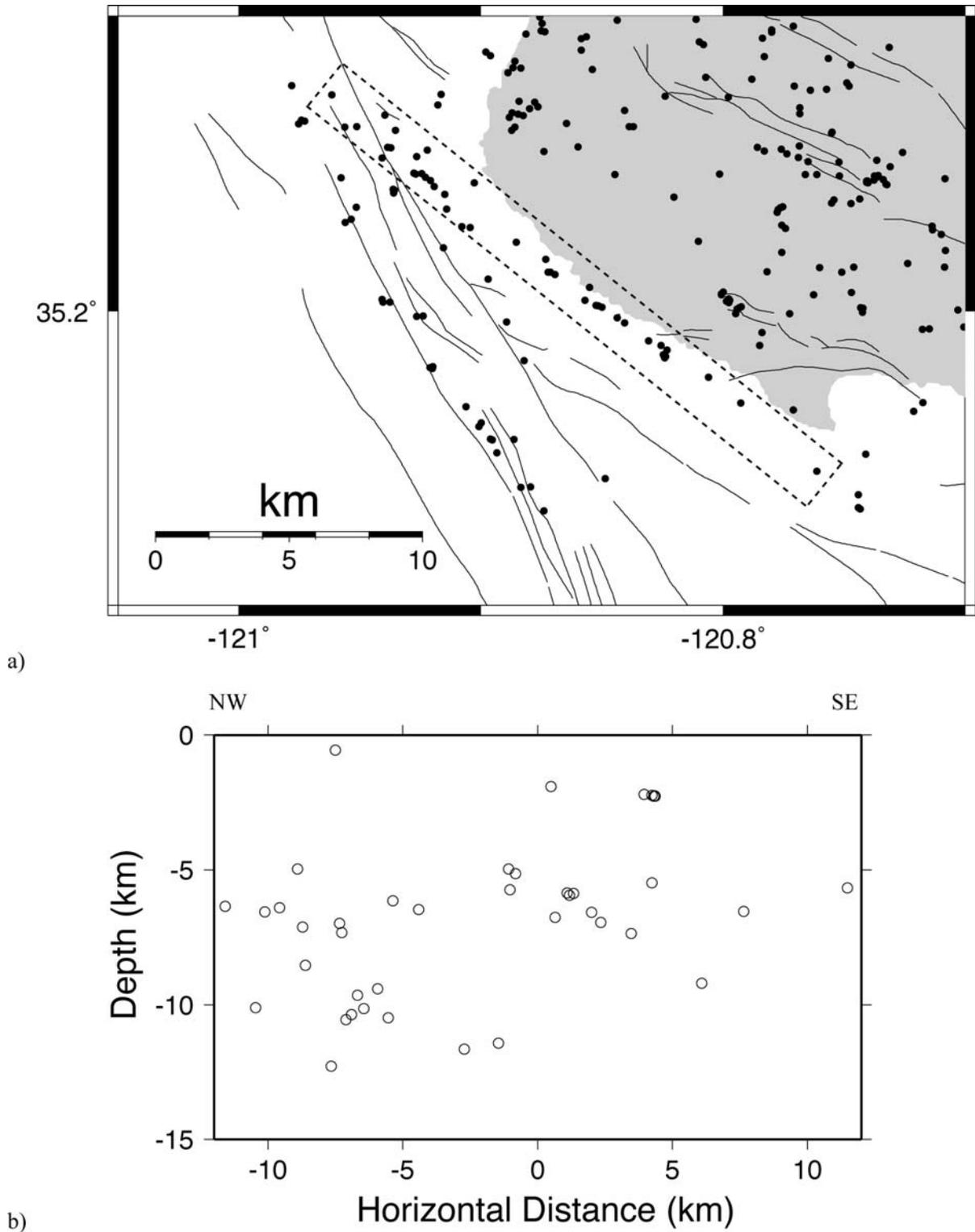
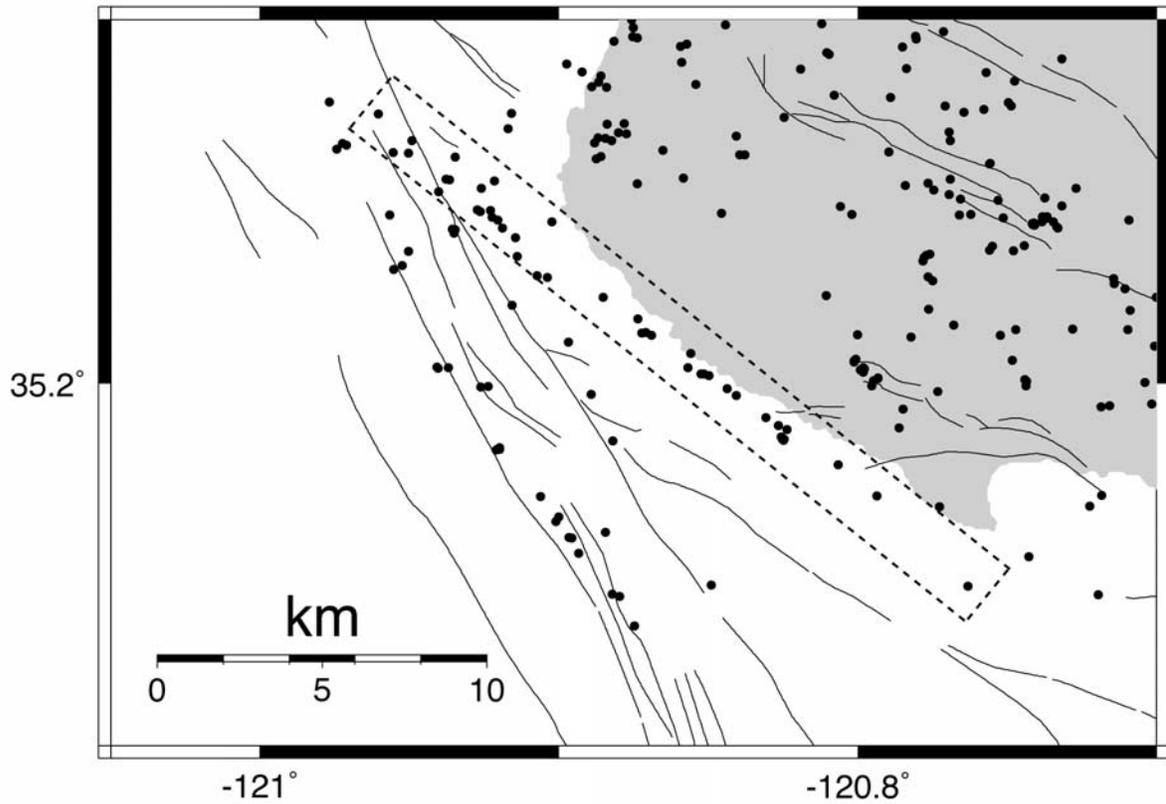
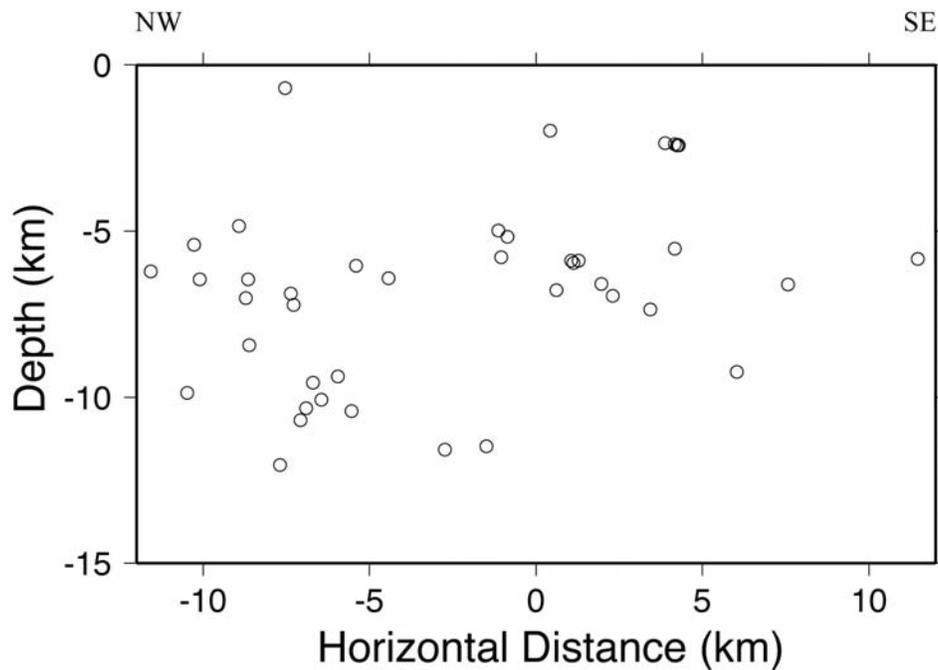


Figure 3. (a) Map view and (b) cross-section (center point at 35.21°, -120.86°, azimuth 129° CW from North, half-width 1 km) of locations from a tomoDD inversion using P and S waves, the modified initial model from Lin et al. (2009), and cross-correlation data from my reanalysis of PG&E data. The box in (a) indicates the earthquakes plotted in (b). Run 11s in Table 2.



a)



b)

Figure 4. (a) Map view and (b) cross-section (center point at 35.21° , -120.86° , azimuth 129° CW from North, half-width 1 km) of locations from a tomoDD inversion using the modified initial model from Lin et al. (2009), cross-correlation data from my reanalysis of PG&E data, and P waves only. The box in (a) indicates the earthquakes plotted in (b). Run 11p in Table 2.

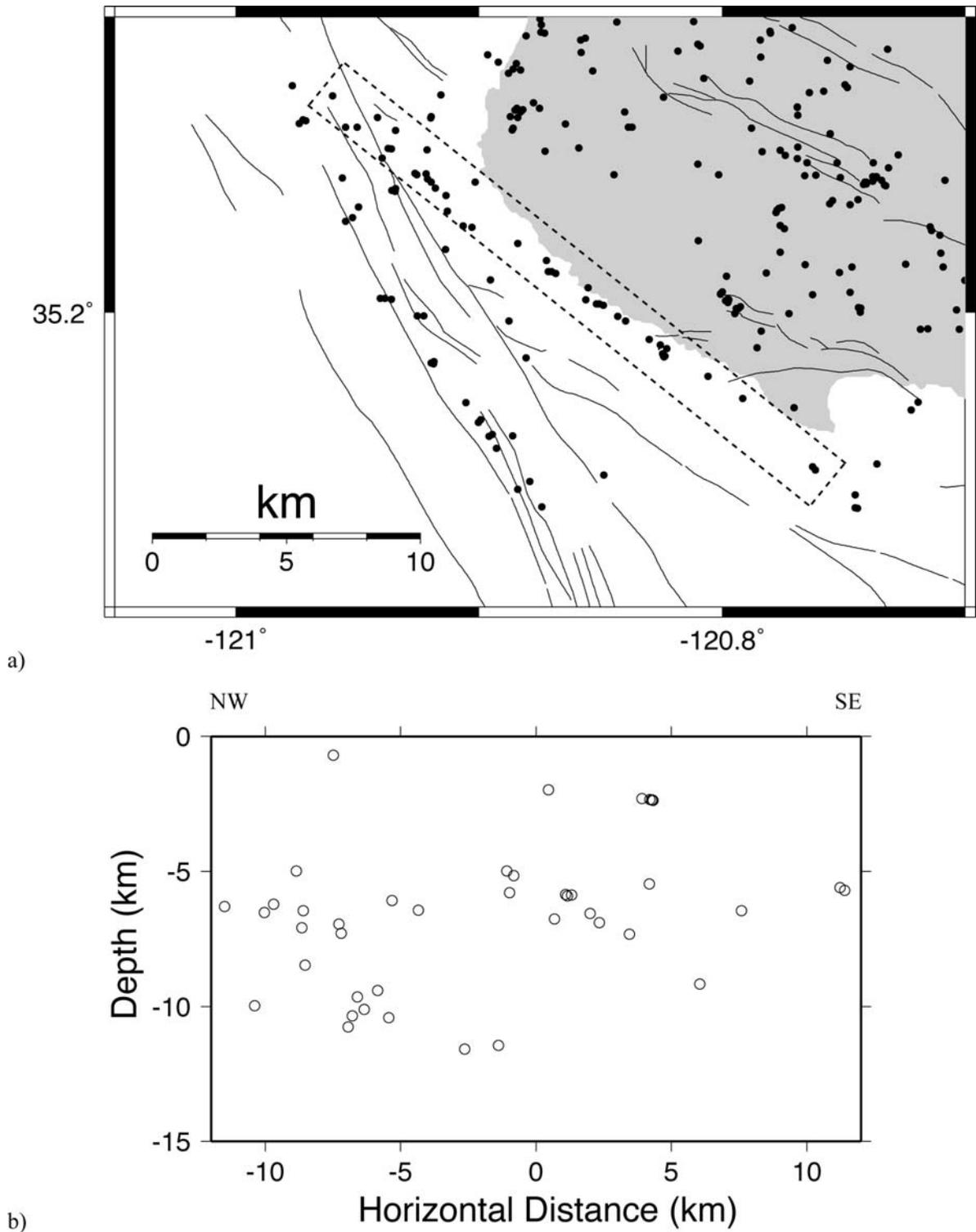
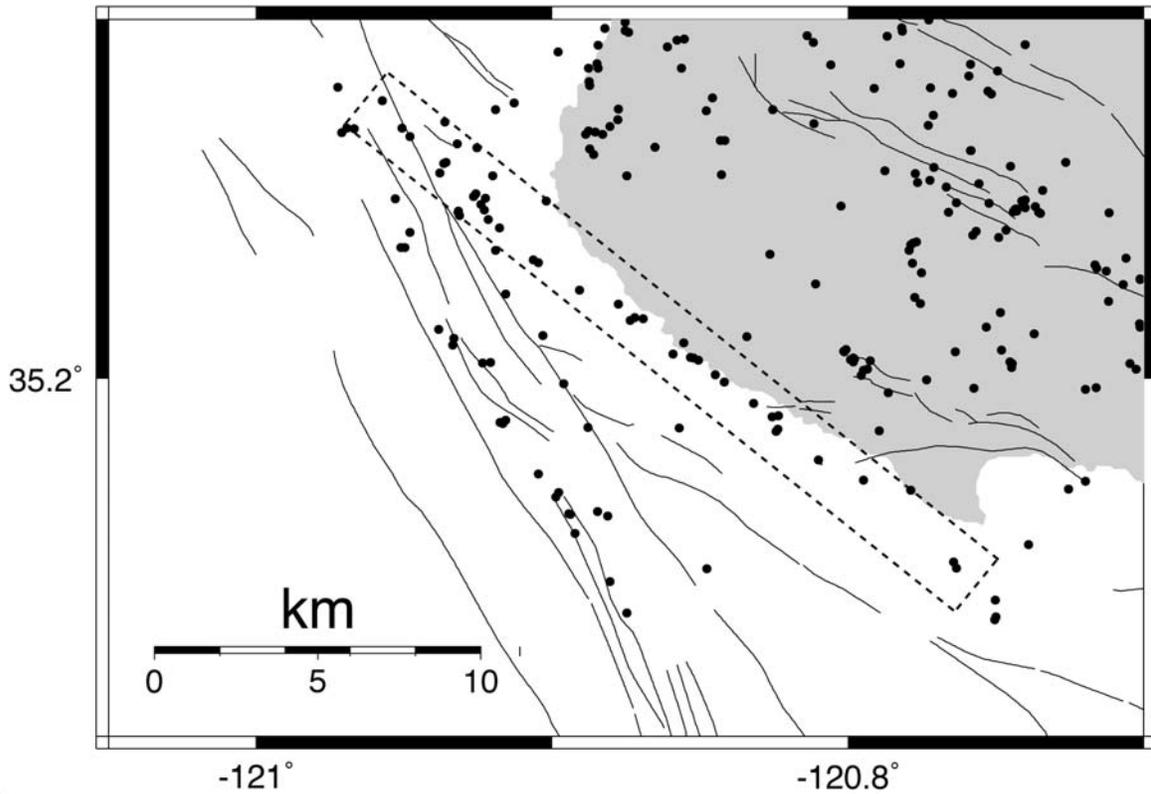
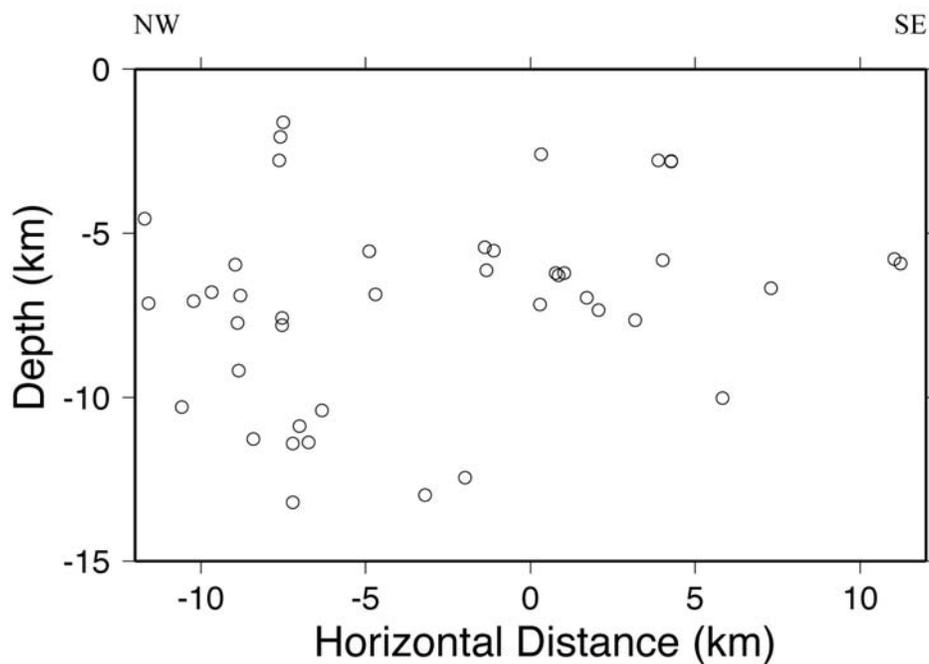


Figure 5. (a) Map view and (b) cross-section (center point at 35.21° , -120.86° , azimuth 129° CW from North, half-width 1 km) of locations from a tomoDD inversion using P and S waves, the modified initial model from Lin et al. (2009), and cross-correlation data from my reanalysis of PG&E data plus Hardebeck's cross-correlation data for NCSN and SCSN stations. The box in (a) indicates the earthquakes plotted in (b). Run 12s in Table 2.

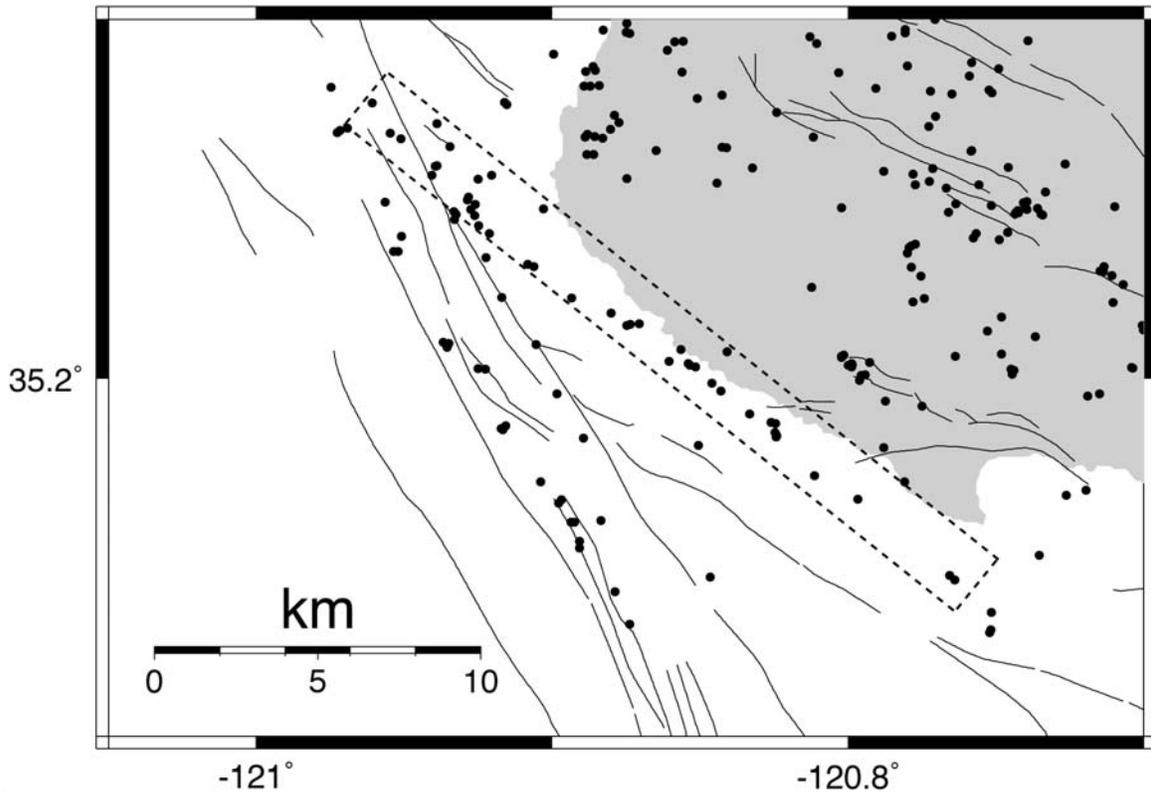


a)

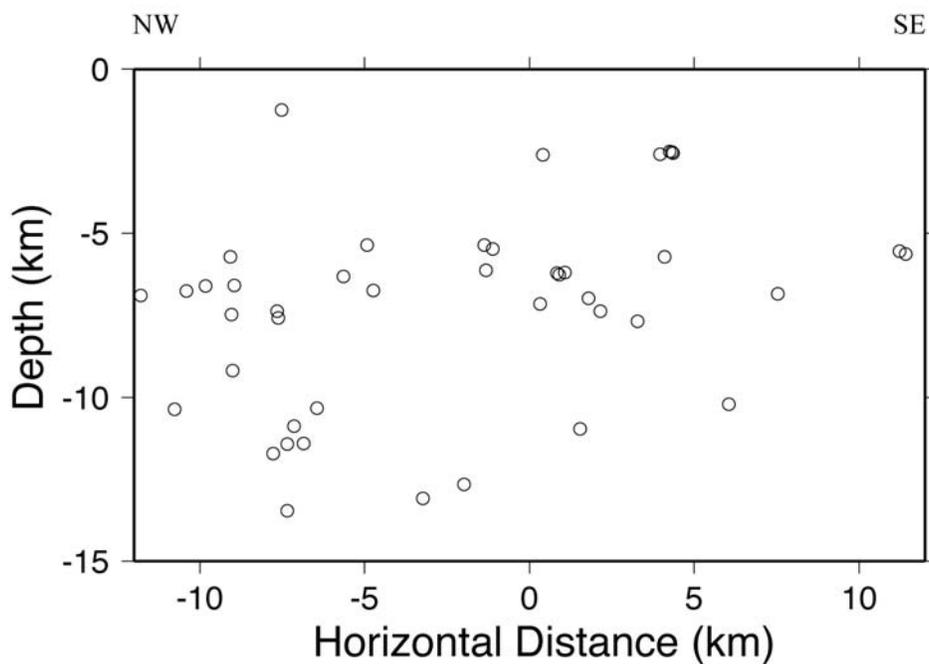


b)

Figure 6. (a) Map view and (b) cross-section (center point at 35.21°, -120.86°, azimuth 129° CW from North, half-width 1 km) of locations from a tomoDD inversion using P and S waves, the 4-km grid decimated Hardebeck model, and cross-correlation data from my reanalysis of PG&E data. The box in (a) indicates the earthquakes plotted in (b). Run 20s in Table 2.



a)



b)

Figure 7. (a) Map view and (b) cross-section (center point at 35.21°, -120.86°, azimuth 129° CW from North, half-width 1 km) of locations from a tomoDD inversion using P and S waves, the 10-km grid decimated Hardebeck model, and cross-correlation data from my reanalysis of PG&E data. The box in (a) indicates the earthquakes plotted in (b). Run 21s in Table 2.

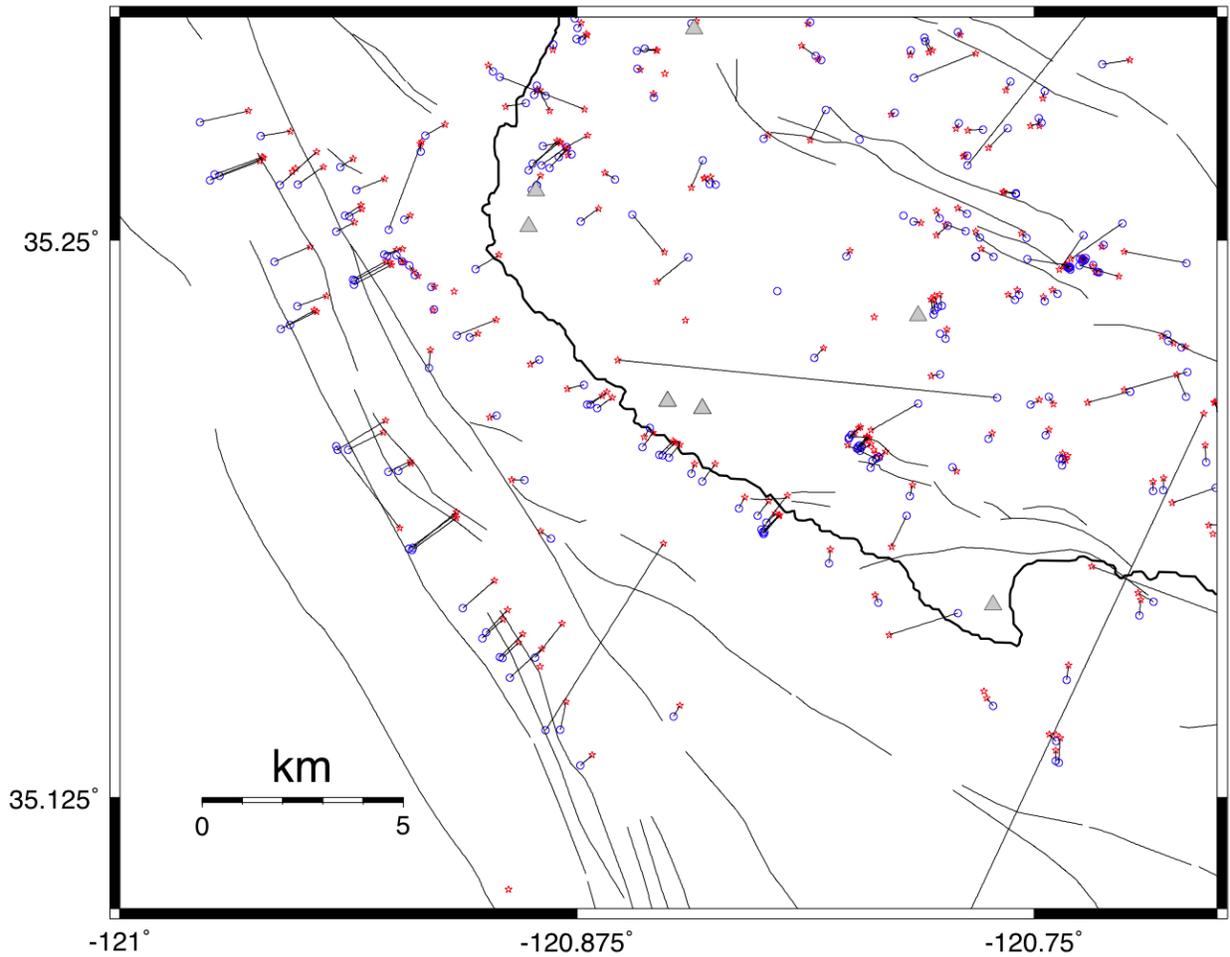


Figure 8. (a) Comparison of the epicenters from Hardebeck's tomographic inversion (red stars) to the epicenters shown in Figure 3 (blue circles; Run 11s in Table 2). The location shifts are relatively small for the target earthquakes, generally a few hundred meters seaward. The location shifts for events along the Hosgri Fault system are substantially larger, ending up on the order of a kilometer farther away from the coast.

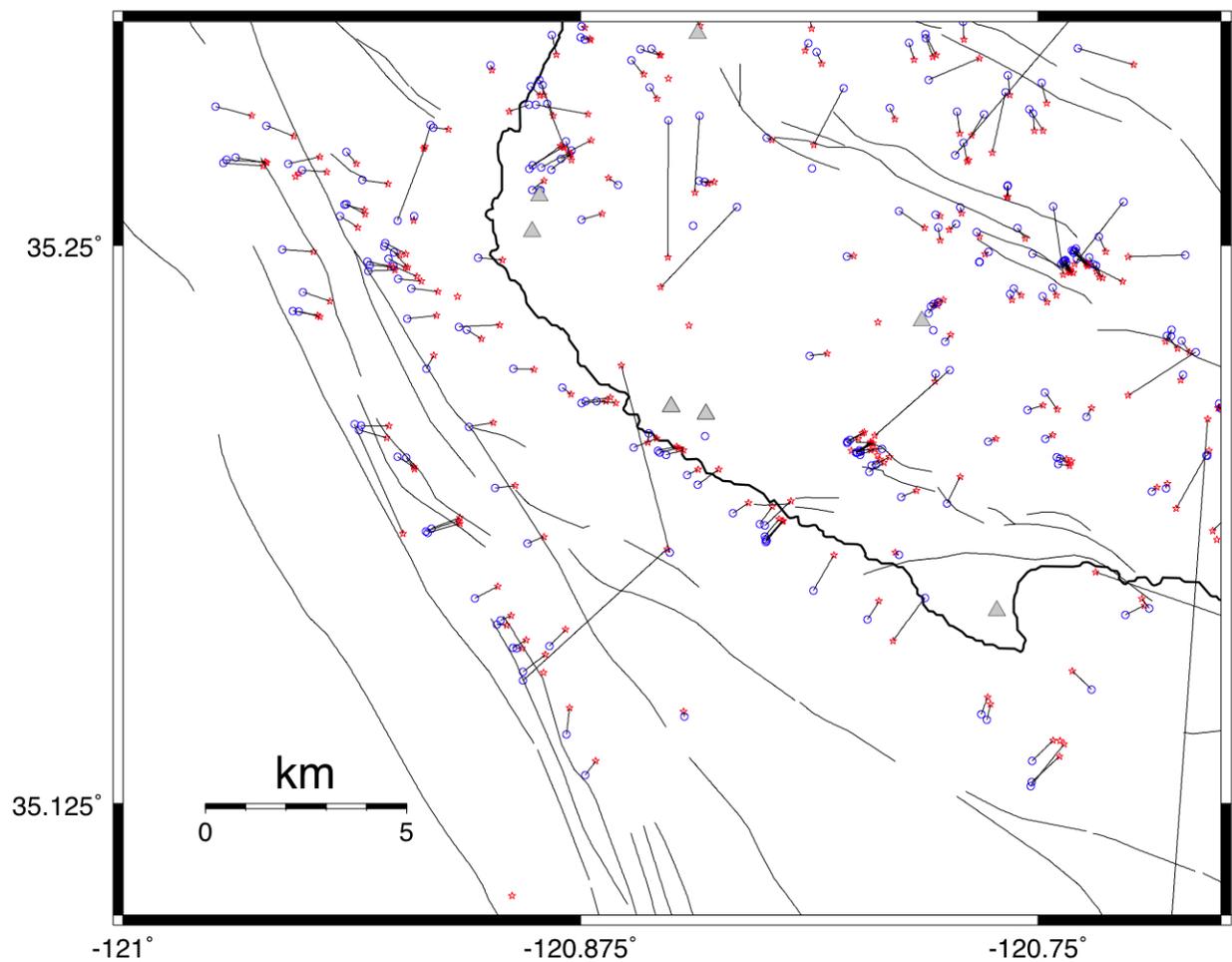


Figure 8. (b) Comparison of the epicenters from Hardebeck's tomographic inversion (red stars) to the epicenters shown in Figure 7 (blue circles; Run 21s in Table 2). The location shifts are slightly larger for the target earthquakes, with some moving up to about a half kilometer seaward. The location shifts for events along the Hosgri Fault system are not as large as in Figure 8a, ending up on the order of half a kilometer farther away from the coast.

Appendix

List of parameter settings used for the tomoDD inversion shown in Figure 3.

```
*--- data type selection:
* IDAT:  0 = synthetics; 1= cross corr; 2= catalog; 3= cross & cat
* IPHA:  1= P; 2= S; 3= P&S
* DIST: max dist [km] between cluster centroid and station
* IDAT  IPHA  DIST
*     3    3    200
*
*--- event clustering:
* OBSCC:  min # of obs/pair for crosstime data (0= no clustering)
* OBSCT:  min # of obs/pair for network data (0= no clustering)
* OBSCC  OBSCT  CC_format
*     0     0     2
*
*--- solution control:
* ISTART:  1 = from single source; 2 = from network sources
* ISOLV:   1 = SVD, 2=lsqr
* NSET:    number of sets of iteration with specifications following
* ISTART  ISOLV  NSET  weight1  weight2  weight3  air_depth
*     2     2    13    100    100    100    -1.7
* i3D  delt1  ndip  iskip  scale1  scale2  iuses
*     2     0     9     1     0.5    1.00    2
* xfac  tlim   nitpb(1) nitpb(2) stepl
*     1.3  0.0005  50     50     0.5
* lat_Orig lon_Orig Z_Orig iorig rota
*    36.5000 -120.0000  0     1    -36.0
*
*--- data weighting and re-weighting:
* NITER:    last iteration to used the following weights
* WTCCP, WTCCS:  weight cross P, S
* WTCTP, WTCTS:  weight catalog P, S
* WRCC, WRCT:    residual threshold in sec for cross, catalog data
* WDCC, WDCT:    max dist [km] between cross, catalog linked pairs
* DAMP:         damping (for lsqr only)
*
* --- CROSS DATA ----- ----CATALOG DATA ----
* NITER WTCCP WTCCS WRCC WDCC WTCTP WTCTS WRCT WDCT WTCD DAMP JOINT THRES
* 2      0.01 0.01 -9 -9 1.0 0.5 -9 -9 1 200 0 2
* 2      0.01 0.01 -9 -9 1.0 0.5 -9 -9 1 200 1 2
* 2      0.01 0.01 -9 -9 1.0 0.5 -9 -9 1 200 0 2
* 2      0.01 0.01 -9 -9 1.0 0.5 -9 10 1 200 1 2
* 2      0.01 0.01 -9 -9 1.0 0.5 -9 10 1 200 0 2
* 2      0.1 0.05 -9 -9 1.0 0.5 -9 10 .1 200 0 2
* 2      0.1 0.05 -9 -9 1.0 0.5 -9 10 .1 200 1 2
* 2      0.1 0.05 -9 -9 1.0 0.5 -9 10 .1 200 0 2
* 2      1 0.5 8 -9 1.0 0.50 6 10 .01 200 0 2
* 2      1 0.5 8 -9 1.0 0.50 6 10 .01 200 1 2
* 2      1 0.5 8 5 1.0 0.50 6 10 .01 200 0 2
* 2      1 0.5 8 5 1.0 0.50 6 10 .01 200 1 2
* 2      1 0.5 8 5 1.0 0.50 6 10 .01 200 0 2
```

References

- Du, W., C. H. Thurber, and D. Eberhart-Phillips, Earthquake relocation using cross-correlation time delay estimates verified with the bispectrum method, *Bull. Seism. Soc. Am.*, 94, 856-866, 2004.
- Lin, G., P. Shearer, E. Hauksson, and C. Thurber, A three-dimensional crustal seismic velocity model for Southern California from a composite event method, *J. Geophys. Res.*, 112, B11306, doi:10.1029/2007JB004977, 2007.
- Lin, G., C. H. Thurber, H. Zhang, E. Hauksson, P. M. Shearer, F. Waldhauser, T. M. Brocher, and J. Hardebeck, A California statewide three-dimensional seismic velocity model from both absolute and differential times, *Bull. Seism. Soc. Am.*, revised and resubmitted, 2009.
- Pavlis, G. L., Appraising earthquake hypocenter location errors: a complete, practical approach for single-event locations, *Bull. Seismol. Soc. Am.* 76, 1699-1717, 1986.
- Poupinet, G., Ellsworth, W.L., Frechet, J., Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras fault, California, *J. Geophys. Res.*, 89, 5719-5731, 1984.
- Rowe, C. A., R. C. Aster, B. Borchers and C. J. Young, An automatic, adaptive algorithm for refining phase picks in large seismic data sets, *Bull. Seism. Soc. Am.*, 92, 1660-1674, 2002.
- Rubin, A. M., D. Gillard, and J.-L. Got, Streaks of microearthquakes along creeping faults, *Nature*, 400, 635-641, 1999.
- Schaff, D. P., G. H. R. Bokelmann, G. C. Beroza, F. Waldhauser, and W. L. Ellsworth, High resolution image of Calaveras Fault seismicity, *J. Geophys. Res.* 107, no. B9, 2186, doi 10.1029/2001JB000633, 2002.
- Schmidt, D. A., R. Bürgmann, R. M. Nadeau, and M. D'Alessio, Distribution of aseismic slip rate on the Hayward fault inferred from seismic and geodetic data, *J. Geophys. Res.* 110, B08406, doi 10.1029/2004JB003397, 2005.
- Thomson, D. J., Spectrum estimation and harmonic analysis, *Proc. IEEE*, 70, 1055-1096, 1982.
- Thurber, C., S. Roecker, K. Roberts, M. Gold, L. Powell, and K. Rittger, 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, 10.1029/2002GL016004, 2003.

- Thurber, C., S. Roecker, H. Zhang, S. Baher, and W. Ellsworth, Fine-scale structure of the San Andreas fault and location of the SAFOD target earthquakes, *Geophys. Res. Lett.*, 31, L12S02, doi 10.1029/2003GL019398, 2004.
- Thurber, C., H. Zhang, F. Waldhauser, J. Hardebeck, A. Michael, and D. Eberhart-Phillips, Three-dimensional compressional wavespeed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region, *Bull. Seism. Soc. Am.*, 96, S38-S49, 2006.
- Waldhauser, F., W. L. Ellsworth, and A. Cole, Slip-parallel seismic lineations along the northern Hayward fault, California, *Geophys. Res. Lett.* 26, 3525-3528, 1999.
- Waldhauser, F., W. L. Ellsworth, D. P. Schaff, and A. Cole, Streaks, multiplets, and holes: High-resolution spatio-temporal behavior of Parkfield seismicity, *Geophys. Res. Lett.*, 31, L18608, doi:10.1029/2004GL020649, 2004.
- Zhang, H., and C. H. Thurber, Double-difference tomography: the method and its application to the Hayward fault, California, *Bull. Seism. Soc. Am.*, 93, 1875-1889, 2003.

Appendix C-2

Central California Offshore Earthquake Location Assessment

Dr. Felix Waldhauser

Project Report

Project Name: Central California Offshore Earthquake Location Assessment

Project Number: 2500179119

Prepared by:

Dr. Felix Waldhauser
423 W 120th Street, Apt 88
New York, NY 10027
Tel: 212 678 4804
Email: felixw@ldeo.columbia.edu

Prepared for:

Dr. Marcia McLaren, Geosciences, PG&E

Summary

A series of double-difference relocations is carried out using the *hypoDD* algorithm of Waldhauser and Ellsworth (2000), together with different velocity models and *hypoDD* parameters, to evaluate the robustness of earthquake locations along the Shoreline Fault Zone (SFZ). The SFZ is a linear seismicity feature off-shore and southeast of Point Buchon, Central California, between latitudes 35.1°N and 35.3° N. Existing seismic arrival time picks and cross-correlation delay times, as well as new synthetic data, were used to relocate the events and estimate their location uncertainty and robustness. This study finds that the resolution in relative hypocenter locations, given the available data and network configuration, is on average better than 0.2 km in horizontal and 0.7 km in vertical directions. This study confirms the spatial distribution of the seismic activity along the Shoreline Fault Zone as imaged by J. Hardebeck, USGS.

1. Objectives

The focus of this study is to test the robustness of earthquake locations estimated by Jeanne Hardebeck (USGS) with the tomoDD program (Zhang and Thurber, 2003) within the linear seismicity feature offshore and southeast of Point Buchon, Central California, between latitudes 35.1°N and 35.3° N (Figure 1), referred to in the following as the Shoreline Fault Zone (SFZ). This is done by re-analyzing existing pick and cross-correlation based differential times (Jeanne Hardebeck, pers. Comm.), as well as synthetic travel times. The goal is to estimate hypocenter locations and associated location uncertainties using the double-difference algorithm *hypoDD* (Waldhauser and Ellsworth, 2000; Waldhauser, 2001) together with 1D and 3D seismic velocity models. The various location estimates that result from this study are then compared with the locations obtained by J. Hardebeck.

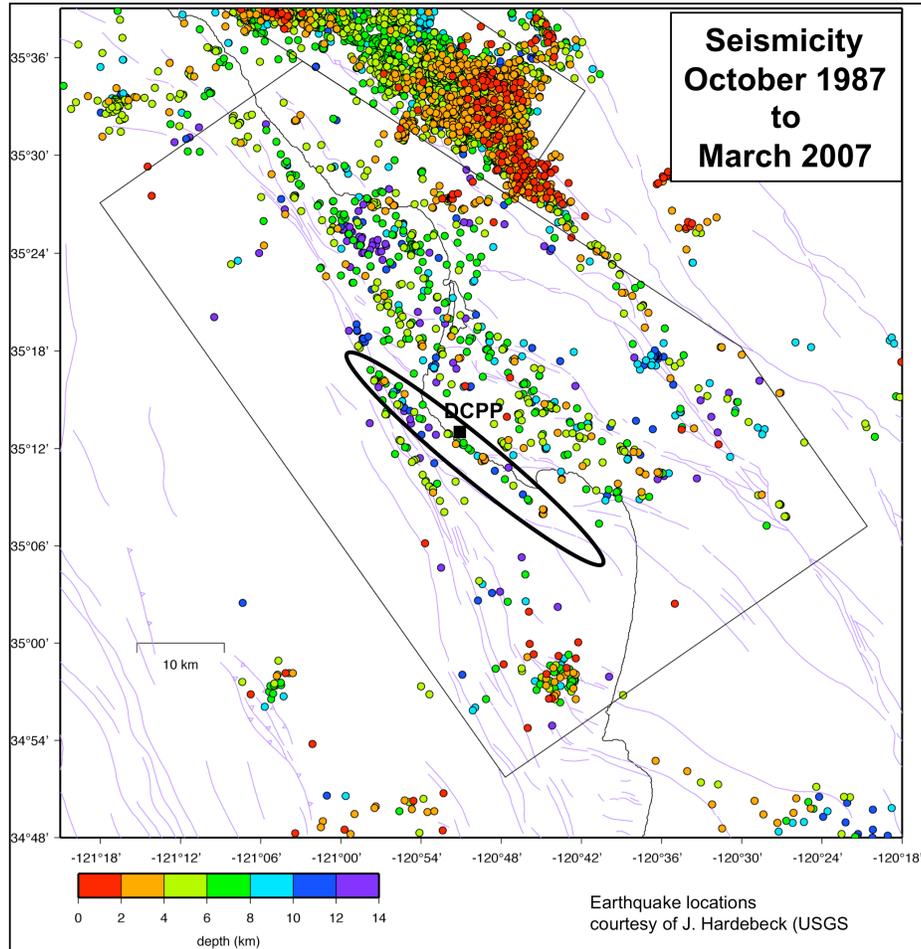


Figure 1 Study area in Central California. Epicenters color-coded by depth. Coastline and mapped surface traces of faults are shown. Polygon includes events analyzed in this study. Elongated ellipse includes events along the Shoreline Fault Zone (SFZ). DCPP: Diablo Canyon Power Plan. Figure courtesy of J. Hardebeck, USGS.

2. Double-Difference Relocation Analysis

The *hypoDD* algorithm minimizes, in a weighted least squares sense, residuals between observed and predicted travel time differences for pairs of earthquakes to solve for adjustments in the vectors connecting the hypocenters through the partial derivatives of the travel times for each event with respect to the unknown. The method is described in detail in Waldhauser and Ellsworth (2000) and Waldhauser (2001). We apply the 1-D model of McLaren and Savage (2001) and the 3-D tomoDD model provided by Jeanne Hardebeck to the original data set of 1418 earthquakes included in the polygon shown in Figure 1, and two smaller subsets of the data that focus on the Shoreline Fault Zone. A

version of the *hypoDD* program that uses 3-D velocity models has previously been used in Rietbrock and Waldhauser (2004), and has also been applied in Thurber et al. (2005), and Waldhauser and Schaff (2008).

In all *hypoDD* runs for which results are shown in the following, we require each event pair to have at least four delay time links. Only stations within 150 km distance from each event pair are used. The SimulPS locations of Jeanne Hardebeck are used as starting locations in all cases. The system of double-difference equations is solved using the conjugate gradient solver *LSQR* (Paige and Saunders, 1982). In the following figures the various *hypoDD* location results are always compared to the final *tomoDD* hypocenter locations provided by Jeanne Hardebeck (in the following also referred to as JH locations).

2.1 Delay time data

The analysis includes 1418 earthquakes in the central coastal California region within the polygon provided by Jeanne Hardebeck (USGS), located approximately between latitudes 34.8°N and 35.5°N. In addition to the nearly 21,000 cross-correlation differential times measured by Jeanne Hardebeck, over one million delay times are formed from the phase pick data using the program *ph2dt* (Waldhauser, 2001). Parameters used to run *ph2dt* are listed in Table 1.

The accuracy and consistency of the phase arrival time picks and the cross-correlation delay times provided by Jeanne Hardebeck is evaluated by comparing the two data sets for all pair of events for which both data types are available. The distributions of differences in P-wave delay times are shown in Figure 2a for pairs of events with correlation coefficients $C_f \geq 0.7$ (2588 pairs) and $C_f \geq 0.9$ (471 pairs). The distributions have standard deviations of 0.150 s for delay times with $C_f \geq 0.7$ and 0.117 s for those with $C_f \geq 0.9$. They exhibit long but thin tails. The median absolute value is 0.015 s for both C_f thresholds, indicating both the consistency of the phase picks as well as the precision of the correlation data down to the cutoff threshold of $C_f = 0.7$. The distribution characteristics are similar for the 22 correlated pairs of earthquakes along the SFZ for which both pick and cross-correlation delay times were available (Figure 2b). The relatively small number of correlation measurements is likely due to the sparseness of the seismicity along the SFZ (i.e., the low number of event pairs with short inter-event distances), and faulting complexities that cause seismograms at common stations to be dissimilar even for nearby events.

Table 1 Parameters used to generate network of delay time links from phase pick data (program *ph2dt*, see Waldhauser 2001 for parameter description):

MINWGHT	MAXDIST	MAXSEP	MAXNGH	MINLNK	MINOBS	MAXOBS
0.0	150	5	100	6	1	60

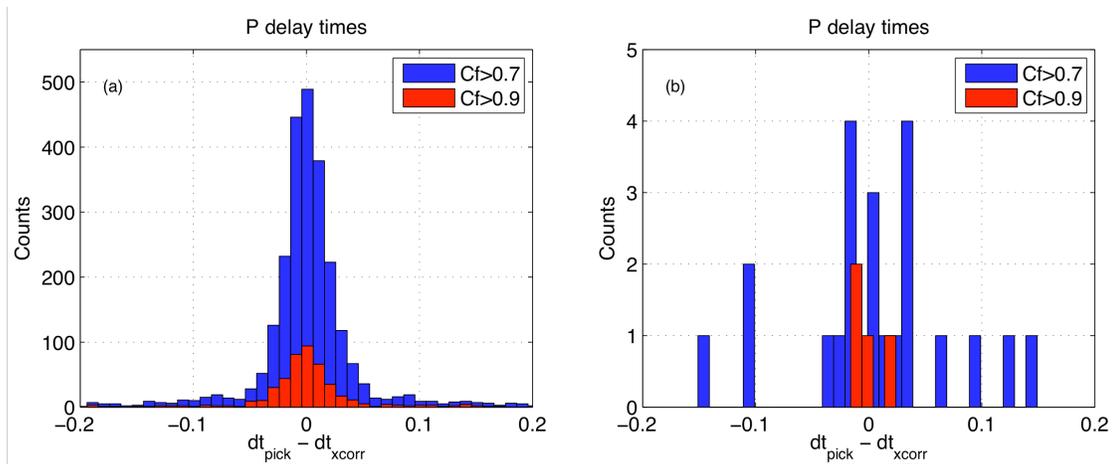


Figure 2 Histograms of differences between P-wave cross-correlation delay-times obtained by Jeanne Hardebeck and corresponding delay-times formed from the phase picks. a) Distribution shown for 2588 pairs of earthquakes within the polygon in Figure 1 with cross-correlation coefficients $Cf \geq 0.7$ and 471 pairs with $Cf \geq 0.9$. Standard deviations are 0.150 s for $Cf \geq 0.7$ and 0.117s for $Cf \geq 0.9$. Median absolute values are 0.015 s for both Cf thresholds. b) Distribution shown for 22 pairs of earthquakes along the Shoreline Fault Zone with $Cf \geq 0.7$ and 4 pairs with $Cf \geq 0.9$. Standard deviations are 0.070 s for $Cf \geq 0.7$ and 0.022 s for $Cf \geq 0.9$. Median absolute values are 0.033 s ($Cf \geq 0.7$) and 0.014 ($Cf \geq 0.9$).

2.2 Relocation with 1D model (DD-1D)

The combined pick and correlation data have been inverted using the layered 1-D model of McLaren and Savage (2001) (MS2001, blue line in Figure 3). The velocity depth function of the MS2001 model has been re-sampled for its use with *hypoDD* to avoid strong velocity contrasts (red line in Figure 3). S-wave velocities are obtained by scaling the P-velocity model by a factor of 1.73. For comparison, we also show the model used for relocating the NCSN earthquake catalog in the Central Coast region (Waldhauser and Schaff, 2008) (green line in Figure 3). The model is based on the 1-D model used by the NCSN for routine earthquake location (their model CST) (Oppenheimer et al., 1993). Notable differences between the two models are the slower velocities at shallow depths in the NCSN model.

The double-difference results for 1380 events, included in the polygon (DD1Da) are shown in map view and cross sections in Figure 4, and compared to the JH tomDD locations (gray circles). The map view is centered at $35^{\circ} 12.9769' N$ and $120^{\circ} 52.3128' E$. The root mean square (RMS) of the weighted pick delay time residuals of the relocated events is 0.071 s (down from 0.18 s before relocation), and that of the correlation time residuals is 0.005 s. The *hypoDD* parameters that were used to control the iterations and data weighting are listed in Table 2.

Because in the DD approach all events are connected through a network of weighted differential-time links, each event may influence the locations of its neighboring events. Bias can be introduced when the distribution and quality of the differential times is heterogeneous across the network, as is the case here. While the onshore events are generally well recorded by nearby well-distributed stations, most offshore events have

large azimuthal station gaps. Furthermore, events along the SFZ, for example, connect to onshore events over comparably large distances. In order to evaluate the effect of the (well-constrained) on-shore seismicity on the location of the off-shore locations, two additional DD inversions using the 1-D model have been carried out: the first one uses a smaller (zoomed) box including 349 events and centered on the SFZ (DD1Dz) (Figure 5) and one using only the 65 events along the SFZ (DD1Df) (Figure 6). The RMS values for the former DD relocations are 0.064 s for pick and 0.003 s for correlation data, and for the latter locations these values are 0.048 s and 0.003 s respectively. Except for the damping value, the weighting parameters for these additional runs were kept the same as listed in Table 2.

Table 2 *hypoDD* iteration and weighting parameters used in this study (see Waldhauser, 2001 for parameter description):

NITER	WTCCP	WTCCS	WRCC	WDCC	WTCTP	WTCTS	WRCT	WDCT
6	0.01	0.01	-999	-999	1.0	0.5	14	5
6	0.01	0.01	20	-999	1.0	0.5	7	5
6	0.01	0.01	20	-999	1.0	0.5	7	4
6	1.0	0.5	20	5	0.01	0.005	7	4
6	1.0	0.5	10	3	0.001	0.0005	7	4

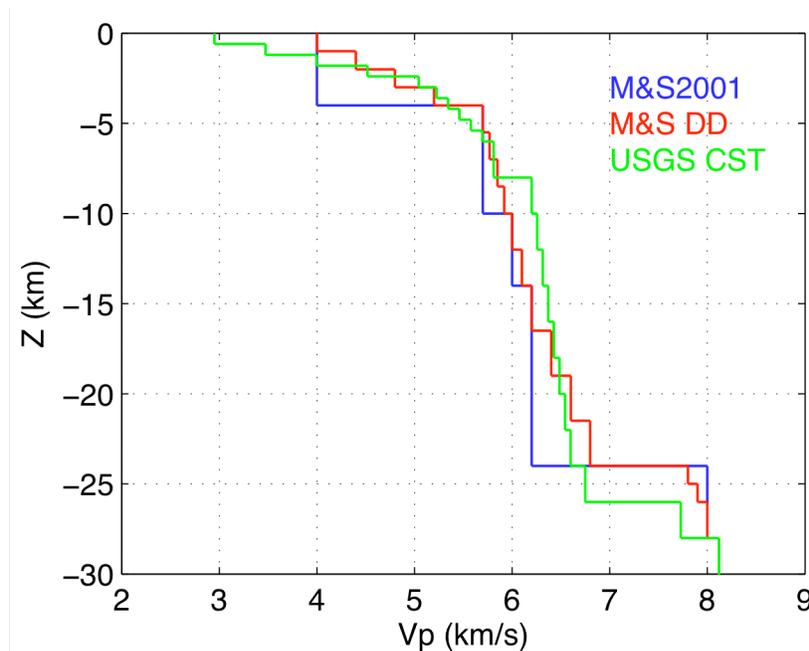


Figure 3 Velocity-depth functions for the study region: M&S2001: McLaren and Savage (2001); M&S DD: re-sampled McLaren and Savage (2001) model; USGS CST: model used by the NCSN for routine location purposes for the Central Coast (CST) region (Oppenheimer et al., 1993).

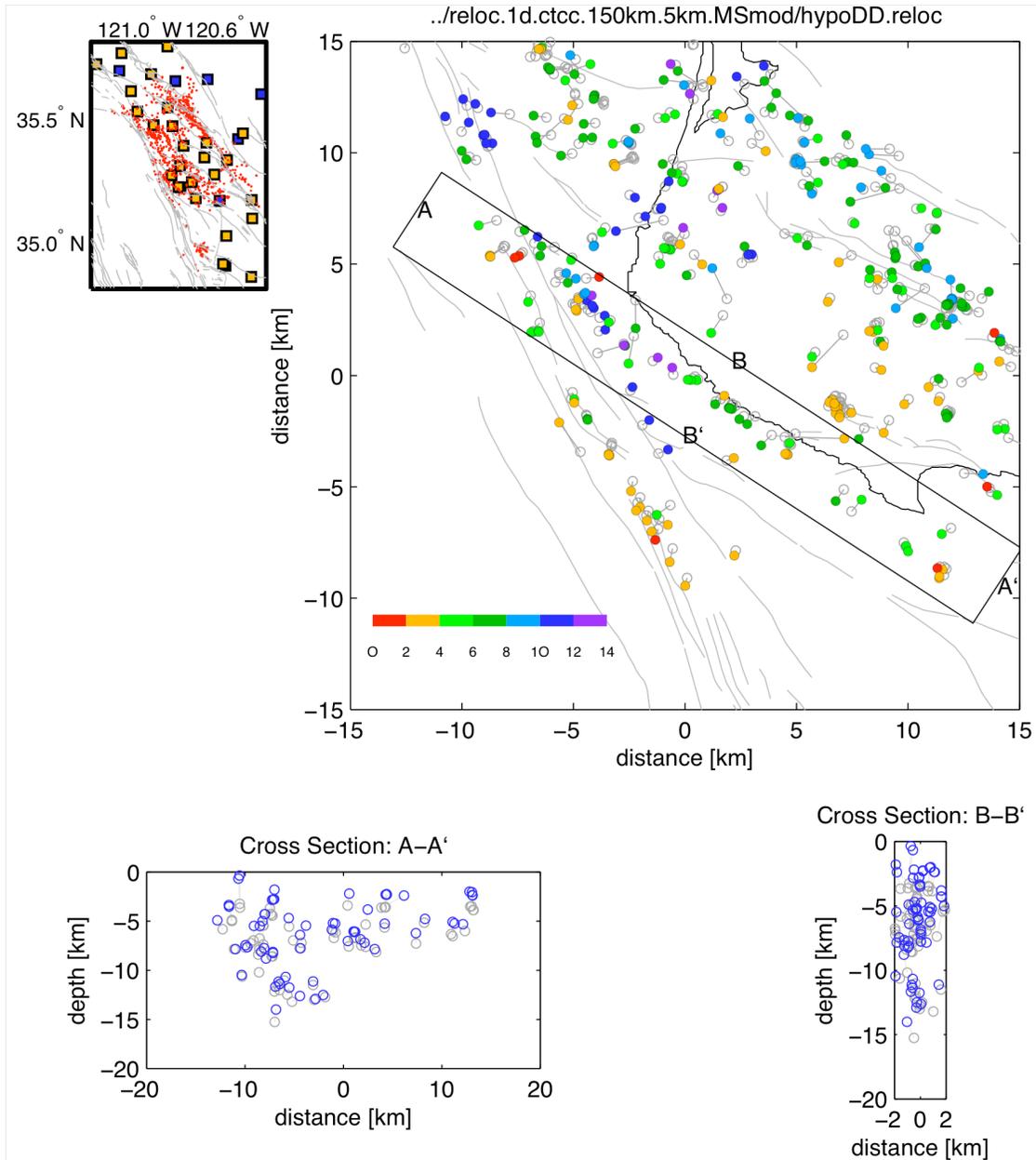


Figure 4 Map view (top panels) and cross-sections (bottom panels) of double-difference solutions (DD1Da) for 1380 earthquakes (out of 1418 selected) within the polygon (see Figure 1) computed using the 1D model shown in Figure 3 (red) and the combined data set of 938,000 P- and 112,000 S-wave picks and 8,440 P- and 4,757 S-correlation delay times. In the top left panel, blue squares denote stations that have pick data only and orange squares those that have additional cross-correlation data. Red dots are epicenter locations, gray lines fault traces, and black line coastal line. In the top right panel, color scale indicates hypocenter depths (in km). Box outlines events shown in cross-sections. Gray circles are the tomoDD locations by J. Hardebeck, with gray lines connecting to the DD locations. Note: In the top right panel, additional events outside the shown area are also included in the relocation. The top left panel shows all events relocated (in red). The map view (top right) is centered on the SFZ at 35° 12.9769' N and 120° 52.3128' E.

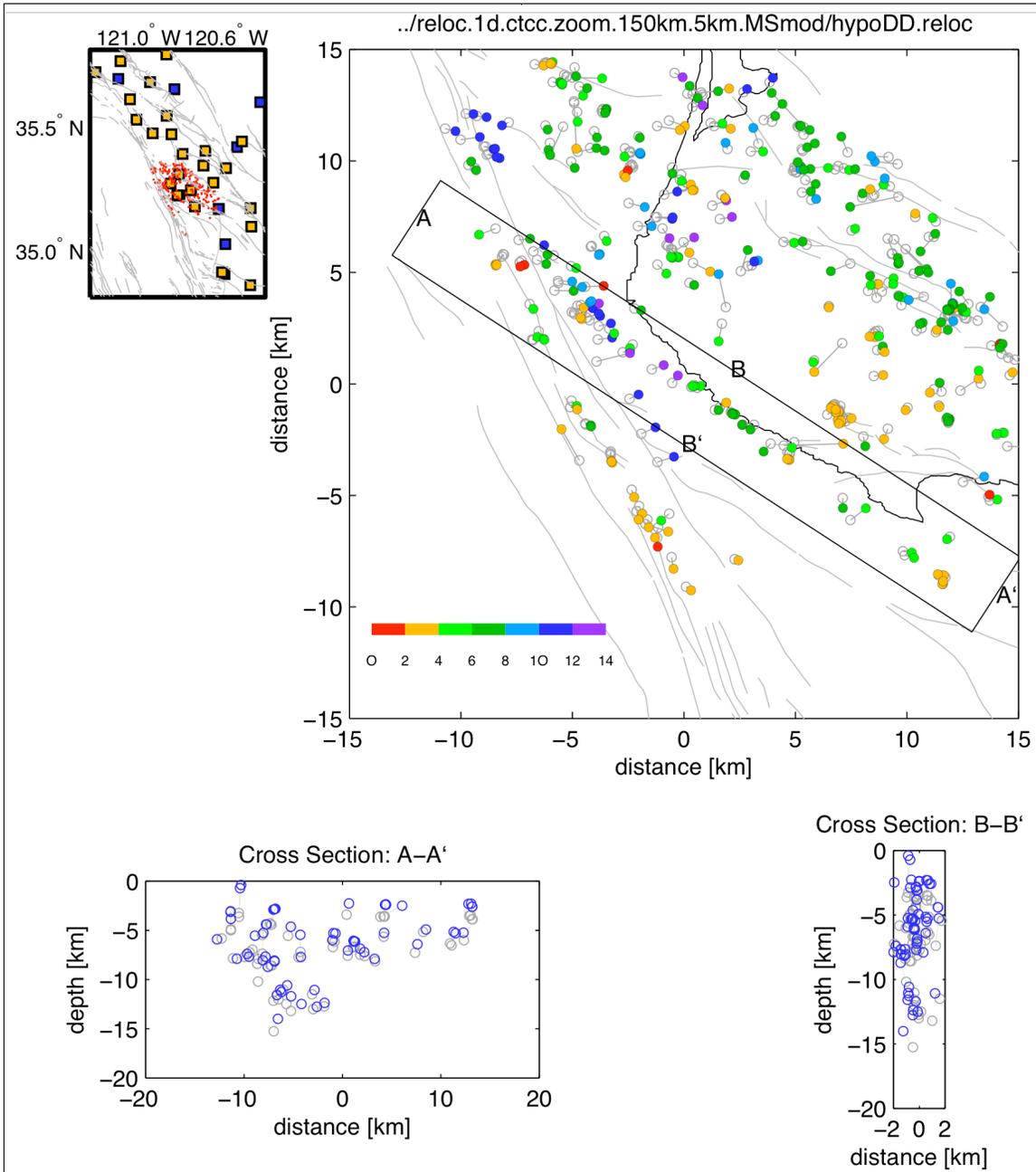


Figure 5 Same as Figure 4, but showing double-difference solutions (DD1Dz) computed for 349 earthquakes (out of 363 selected) in a smaller area centered on the Shoreline Fault Zone, using the 1D model and a combined data set of 203'000 P- and 24'000 S-wave picks and 847 P- and 653 S-wave correlation delay times.

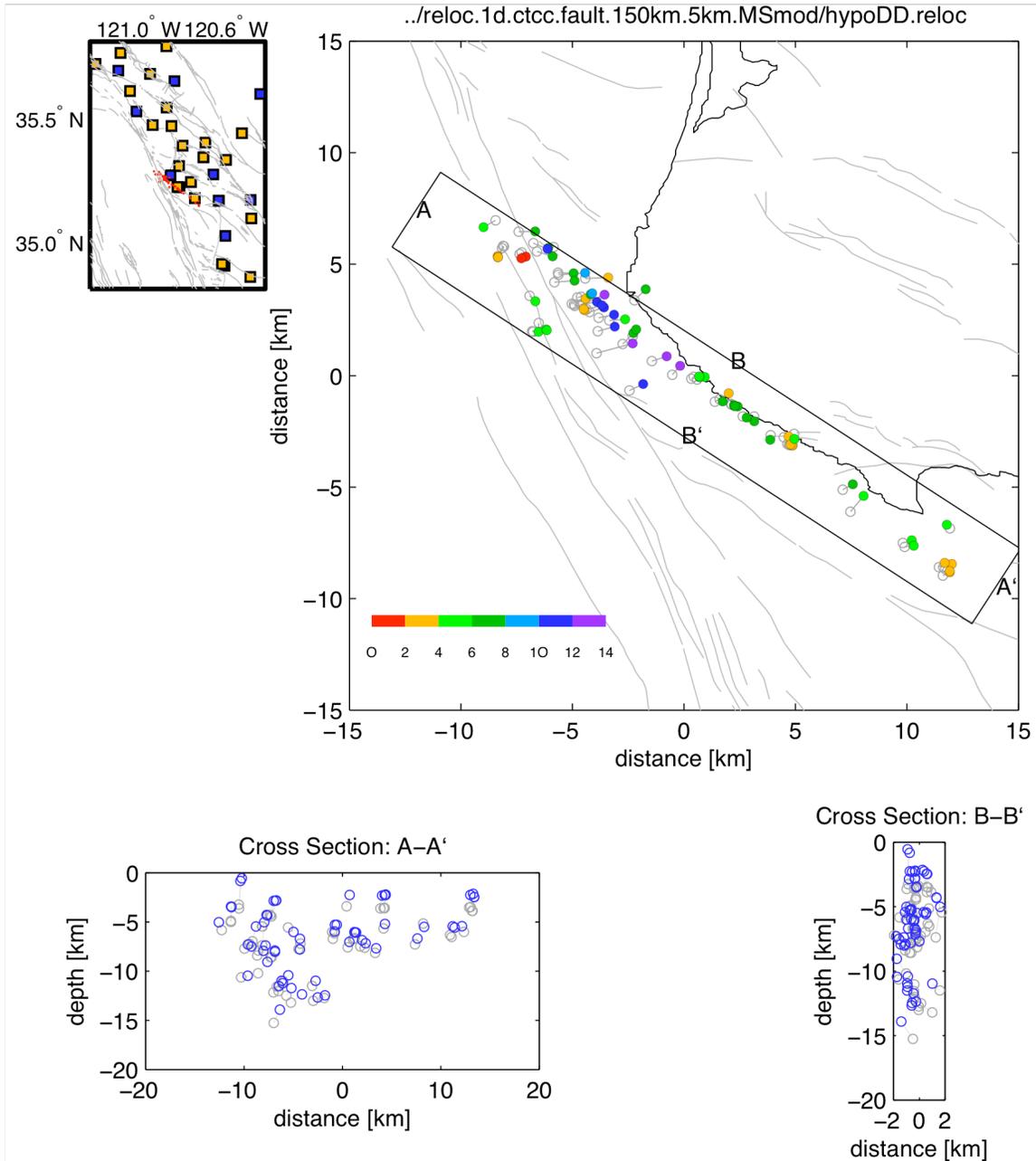


Figure 6 Same as Figure 4, but showing double-difference solutions (DD1Df) computed for 65 earthquakes along the Shoreline Fault Zone, using the 1D model and a combined data set of 14'800 P- and 2'000 S-wave picks and 137 P- and 115 S-wave correlation delay times.

2.3 Relocation with 3D model (DD-3D)

In a second step, the combined pick and correlation data have been relocated using *hypoDD* and the 3D P- and S-wave tomography model provided by Jeanne Hardebeck. The double-difference results for 1375 events (out of 1418 total) (DD3Da) within the region outlined by the polygon in Figure 1 are shown in map view and cross sections in Figure 7, together with the JH locations (gray circles). The RMS of the final pick delay time residuals for the relocated events is 0.074 s (down from 0.18 s in the initial data), and that of the correlation time residuals is 0.005 s. These values are similar to the DD1Da results. The weighting parameters are the same as used in the DD1D inversions (see Table 2).

Similar to the 1D cases, two additional 3D DD inversions were performed to evaluate the dependency of the offshore event locations on the onshore seismicity during the DD relocation procedure. Relocation results from events in the zoomed area (DD3Dz) and events along the SFZ (DD3Df) are shown in Figure 8 and Figure 9. RMS values for the DD3Dz relocations are 0.067 s for pick and 0.004 s for correlation data. For the DD3Df relocations these values are 0.059 s and 0.004 s.

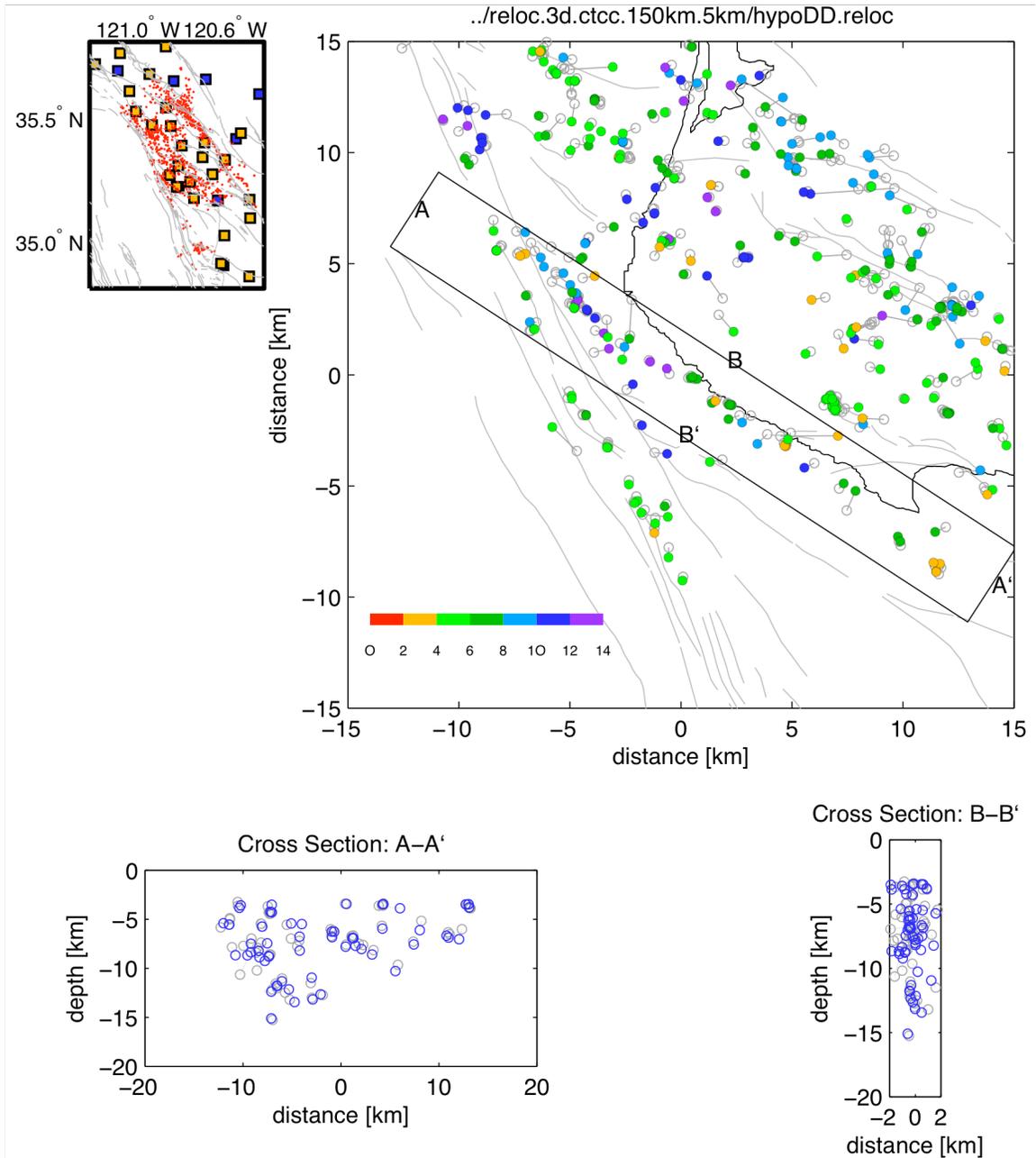


Figure 7 Same as Figure 4, but showing double-difference solutions (DD3Da) for 1375 earthquakes (out of 1418 selected) using the 3D model from Jeanne Hardebeck. The top right panel shows a subset of the events that went into the DD inversion.

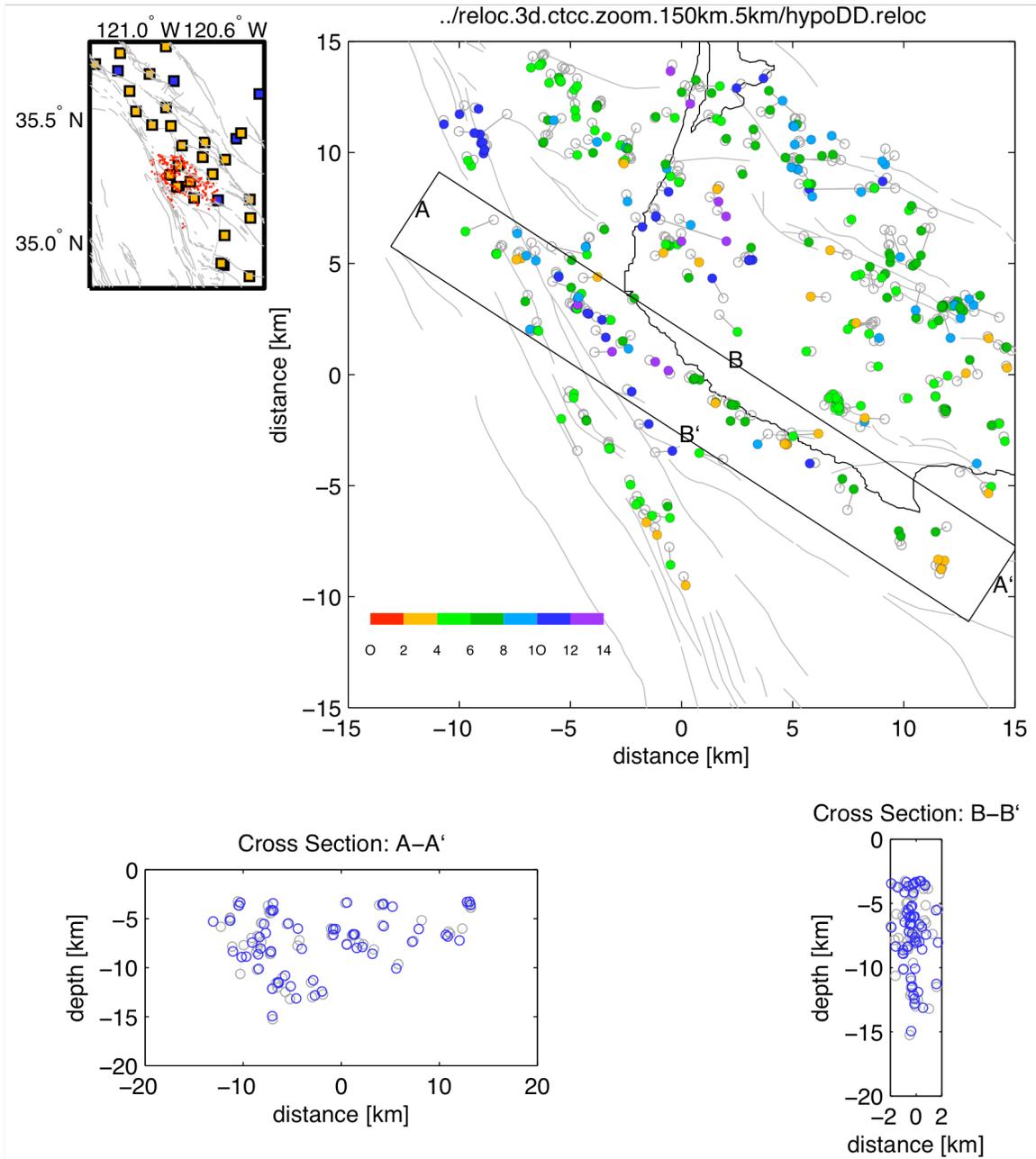


Figure 8 Same as Figure 5, but showing double-difference solutions (DD3Dz) computed for 351 earthquakes (out of 363 selected) using the 3D model from Jeanne Hardebeck. All events used in this inversion are shown in the map view.

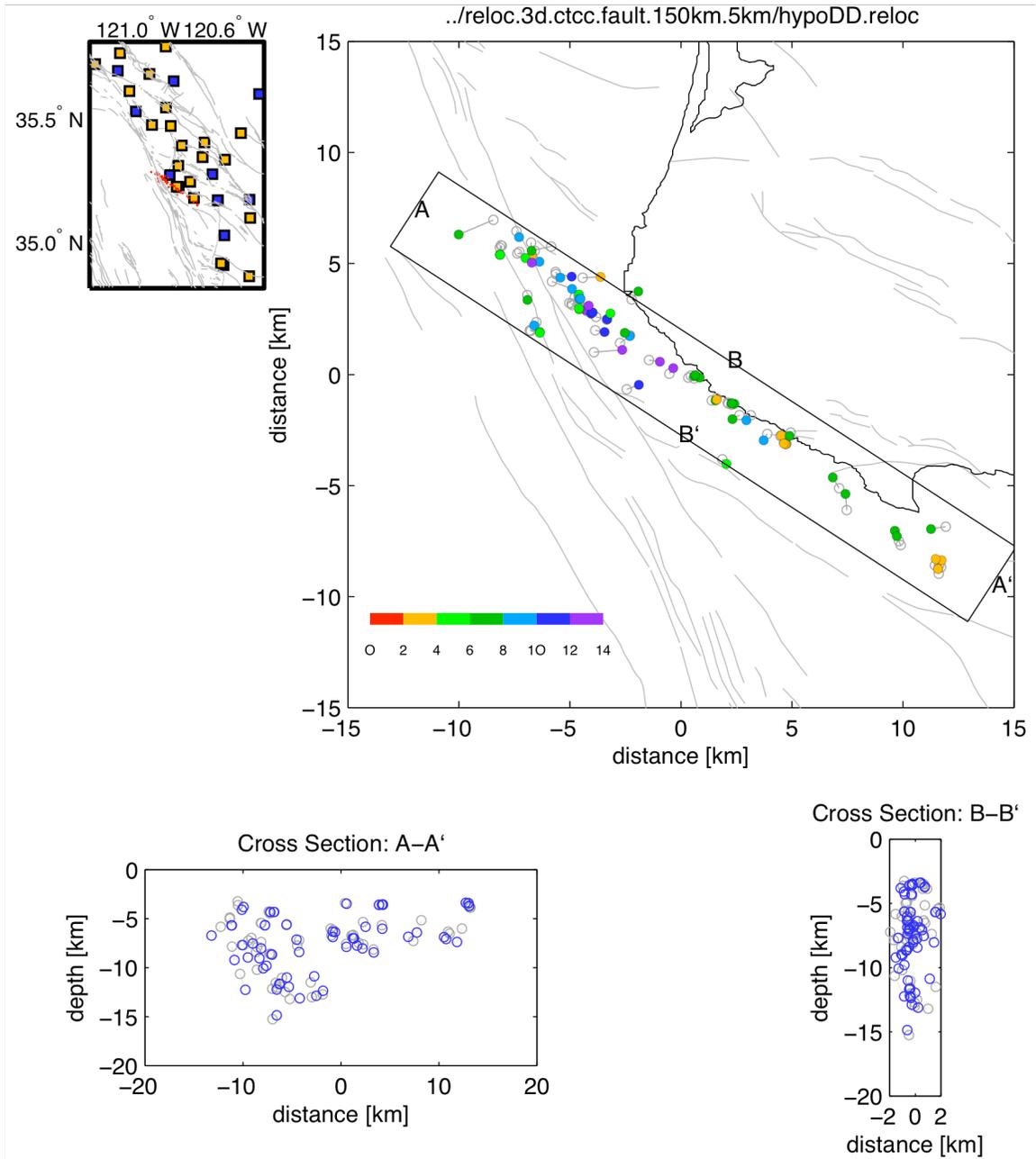


Figure 9 Same as Figure 6, but showing double-difference solutions (DD3Df) computed for 65 earthquakes using the 3D model from Jeanne Hardebeck.

2.4 Estimates of relative location uncertainty

Relative location errors are estimated for each of the 1375 3-D relocated events by bootstrapping, with replacement, the final un-weighted double-difference residual vector (Waldhauser and Ellsworth, 2000). Error ellipsoids are obtained at the 90% confidence level for 200 bootstrap samples. The distributions of the major and minor axes of the horizontal and vertical projections of these ellipsoids are shown in Figure 10. Note i) the strongly elongated lateral ellipses due to the one-side station distribution, and ii) the long tails of the error distributions associated with weakly linked events.

Figure 11 shows locations and error ellipses for the 65 events along the Shoreline Fault Zone computed with the 3D model (DD3Df, see Figure 9). The large lateral errors for two events indicate little control of their location through the delay times that remain after 30 iterations (see Table 2). However, control of these locations may have been better during the first iterations. Formal least squares errors derived from the full covariance matrix for selected number of events show error estimates that are similar to the bootstrap results.

The bootstrap error analysis is also carried out for the DD3Dz case, which includes events surrounding the SFZ (Figure 12). We observe large elongated ellipses for some (mostly onshore) events that indicate loss of control due to weak links to neighboring events. The error ellipses of the events along the SFZ, now derived by including the onshore events, are comparable to the ellipses derived from inverting the SFZ events alone (Figure 11). However, they are overall slightly smaller because of the additional data linking the events along the SFZ to events away from the fault. Events in the center of a cluster are typically best constrained, events at the edge of a cluster worst.

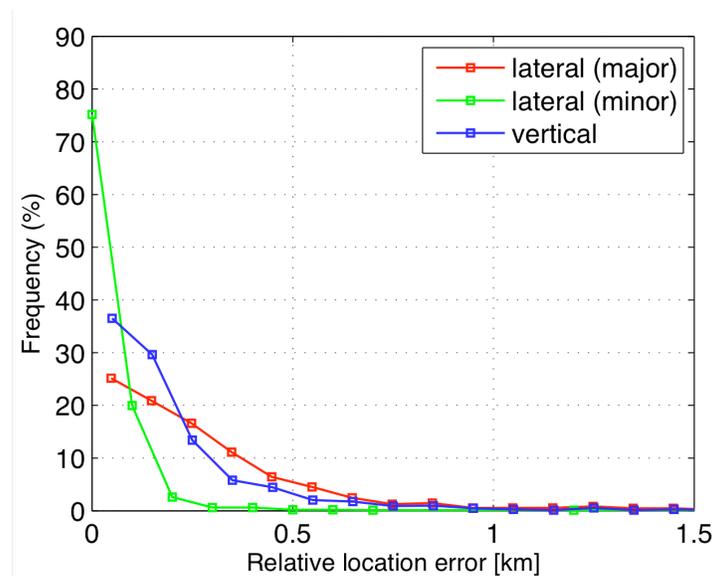


Figure 10 Histograms of lateral and vertical relative location errors of 3D double-difference solutions for 1375 events using the combined pick and correlation data (see Figure 7 for event locations). Errors are computed from the mayor axes of the horizontal and vertical projection of the 90% confidence ellipsoids obtained from a bootstrap analysis of the final double-difference vector based on 200 samples with replacement. Percentage values are computed within bins of 0.1 km.

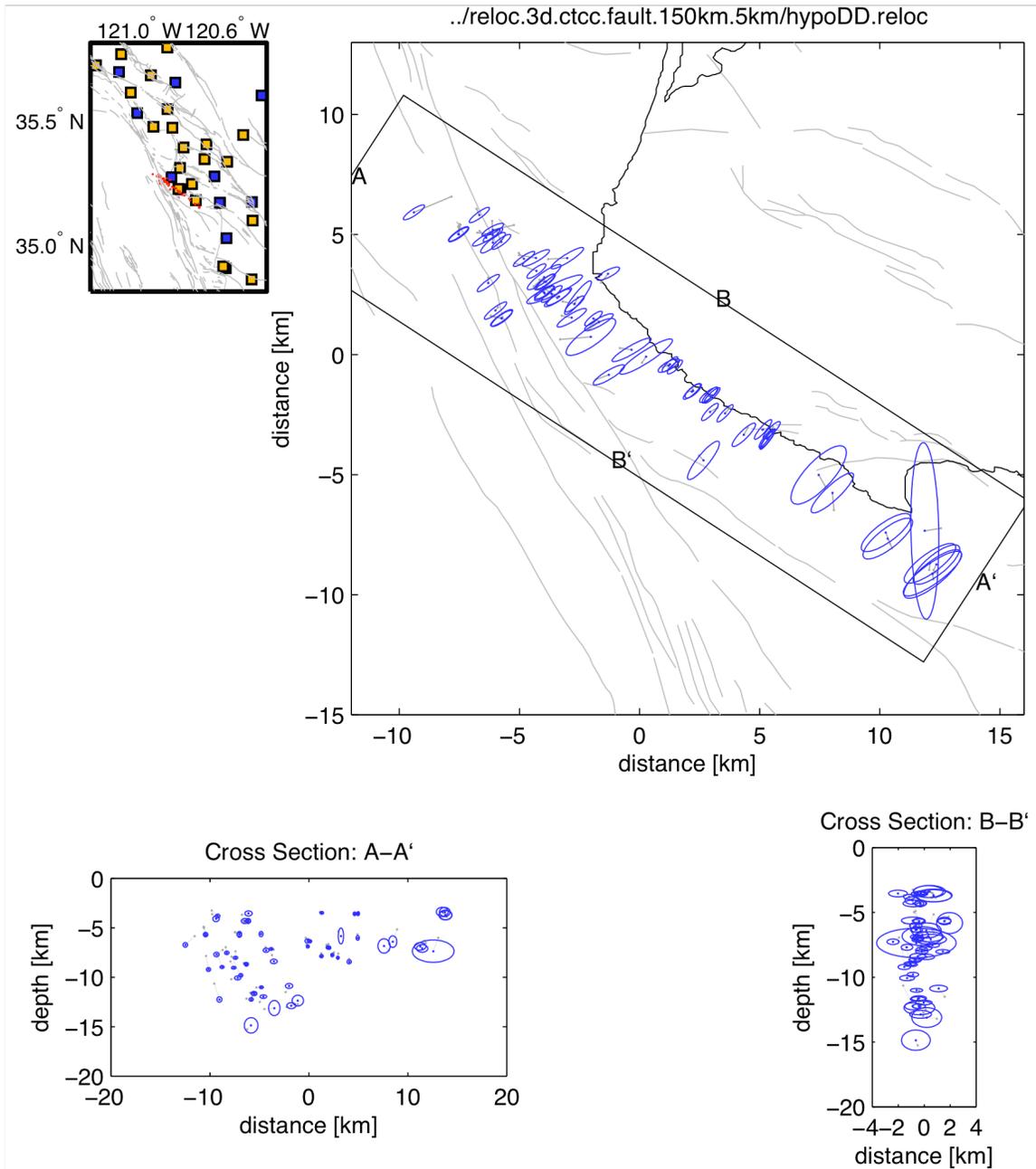


Figure 11 Relative location errors for the 65 earthquakes along the Shoreline Fault Zone using the 3D model (DD3Df, see also Figure 9). Relative horizontal and vertical error ellipses are shown at the 90% confidence level. Ellipses are computed from the mayor axes of the horizontal and vertical projection of the 90% confidence ellipsoids obtained from a bootstrap analysis of the final double-difference vector based on 200 samples with replacement. Gray lines connect to corresponding JH locations.

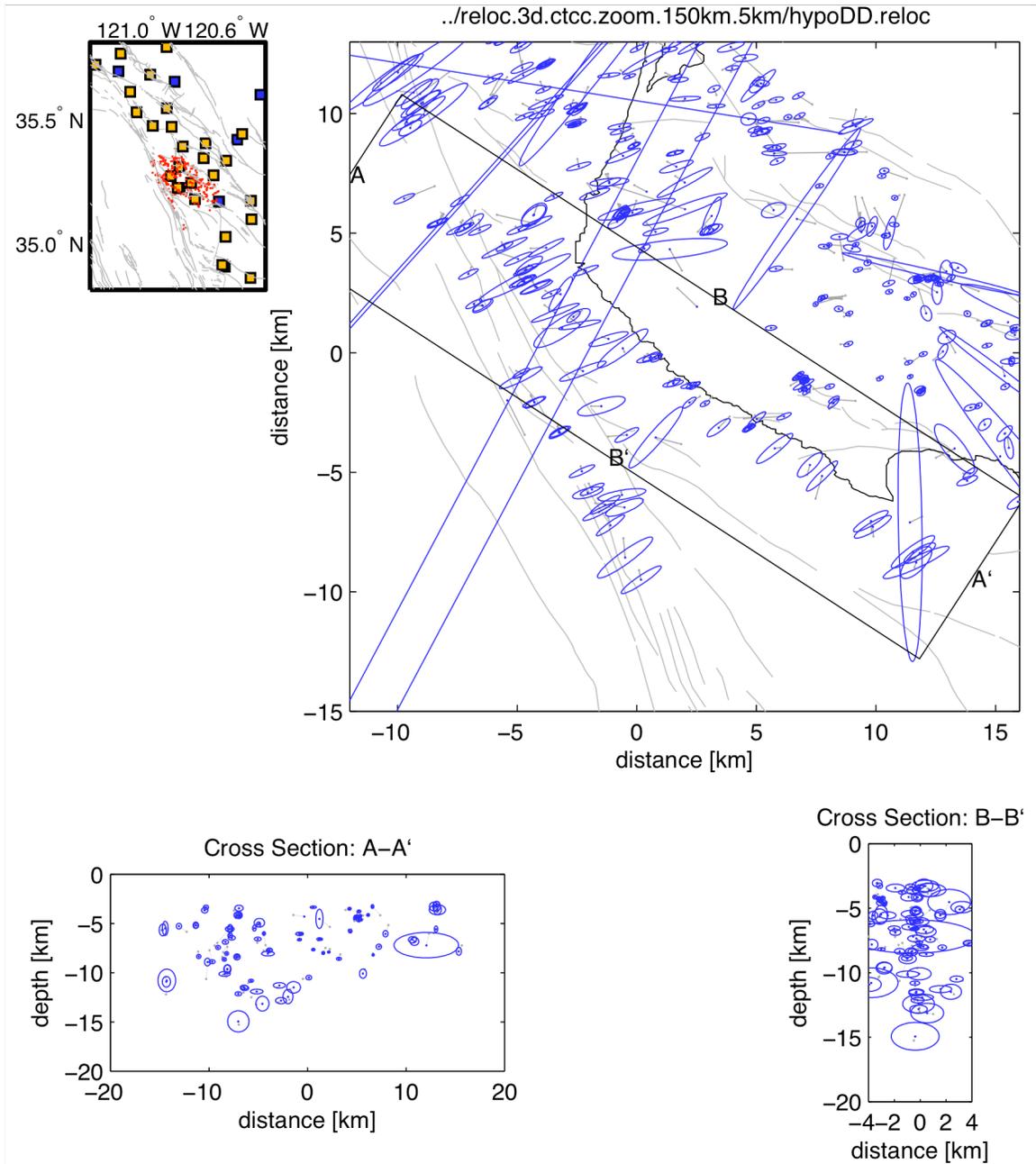


Figure 12 Same as Figure 11, but showing the double-difference locations with their relative horizontal and vertical bootstrap error ellipses at the 90% confidence level for the 3D DD solutions shown in Figure 8.

3. Synthetic tests

Tests with synthetic travel times are carried out to evaluate the resolution capabilities of the available seismic stations. Travel times are computed from 140 sources through the 3D model to 33 stations that recorded the real earthquakes within 100 km. The sources are evenly distributed (spacing = 2 km) on a vertical plane between 1 and 13 km depth (gray circles in Figure 13), mimicking the fault defined by the earthquakes along the Shoreline Fault Zone.

The sources are then randomly mislocated between -4 and 4 km in all three directions (color coded in the top right panel and blue circles in the bottom panels of Figure 13) (Table 3), and Gaussian noise with a standard deviation of 0.1 s is added to the perfect travel times. These values are conservative estimates of the uncertainties in location and pick accuracy of the initial (SimulPS) solutions. From these synthetic travel times we form delay times between nearby events and relocate the events using hypoDD with parameters similar to the ones chosen for relocating the real data. The relocation results for these synthetic data are shown in Figure 14. Average mislocations in each of the three directions are less than 200 m, and the maximum mislocation is less than 500 m (Table 3).

The mislocations are somewhat smaller when Gaussian noise with a standard deviation of 0.01 s, similar to the precision of the cross-correlation data, is added to the perfect data before relocation (Table 3). Note, however, that events along the SFZ are predominately constrained by delay times from phase picks (see Figure 2b).

Table 3 Horizontal (DX, DY) and vertical (DZ) deviations of shifted starting and relocated synthetic sources from true locations, with Gaussian noise (standard deviation = 0.1 s and 0.01 s) added to the perfect data before relocation (in km):

	DX		DY		DZ	
	Mean	Max	Mean	Max	Mean	Max
Initial (shifted)	1.9	4.0	1.9	3.9	2.0	4.0
Relocated noise=0.1 s	0.11	0.32	0.19	0.42	0.13	0.42
Relocated noise=0.01s	0.09	0.18	0.18	0.32	0.15	0.32

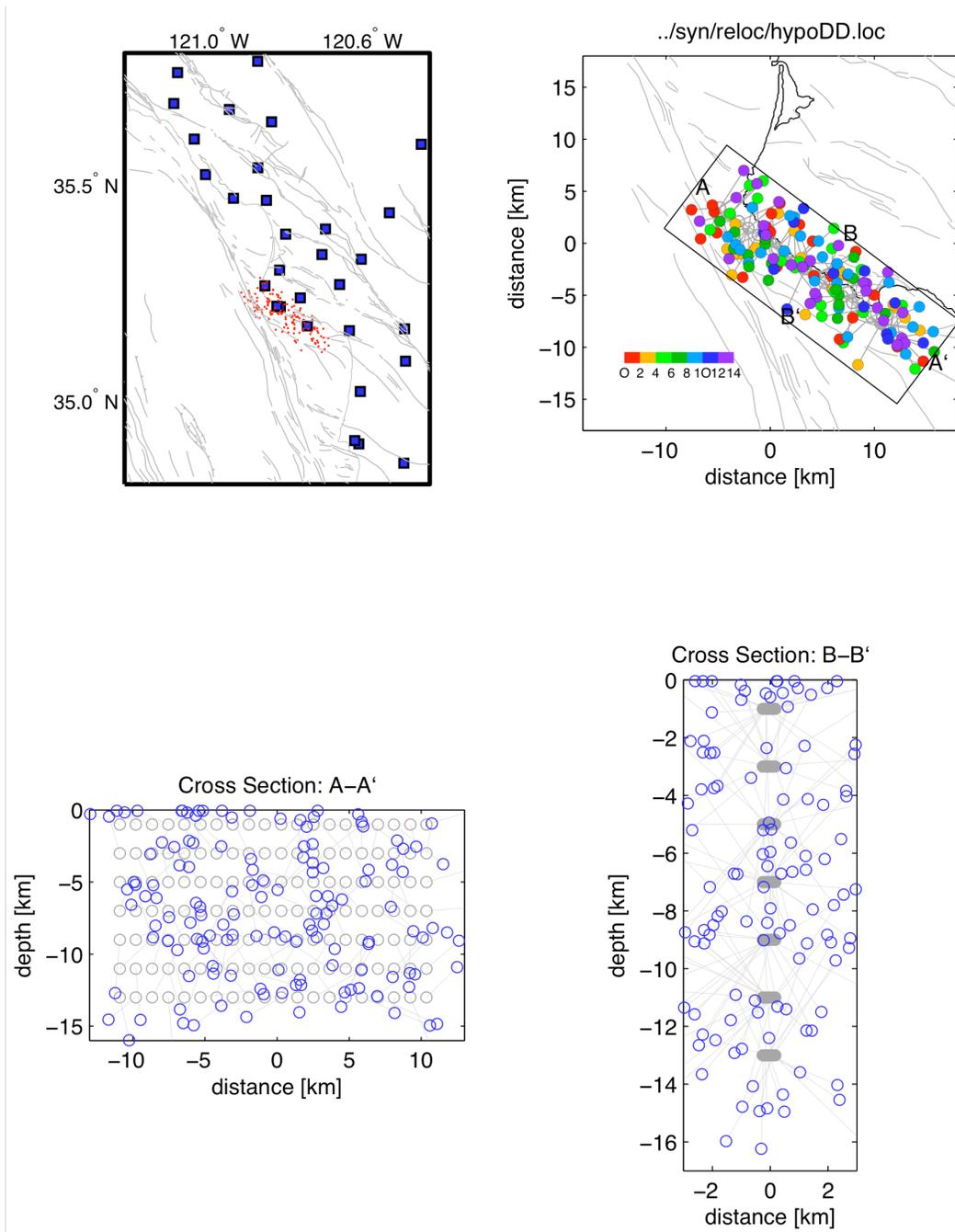


Figure 13 Test with synthetic data. Map view (top panels) and cross sections (lower panels) of synthetic sources (gray circles) for which perfect travel times are calculated through the 3D model to the stations within 100 km (blue squares in the top left panel). Depth (in km) color-coded (top right panel) and blue (bottom panel) circles show starting locations used in the DD inversion. Starting locations are obtained by randomly shifting the true locations within 4 km in each direction.

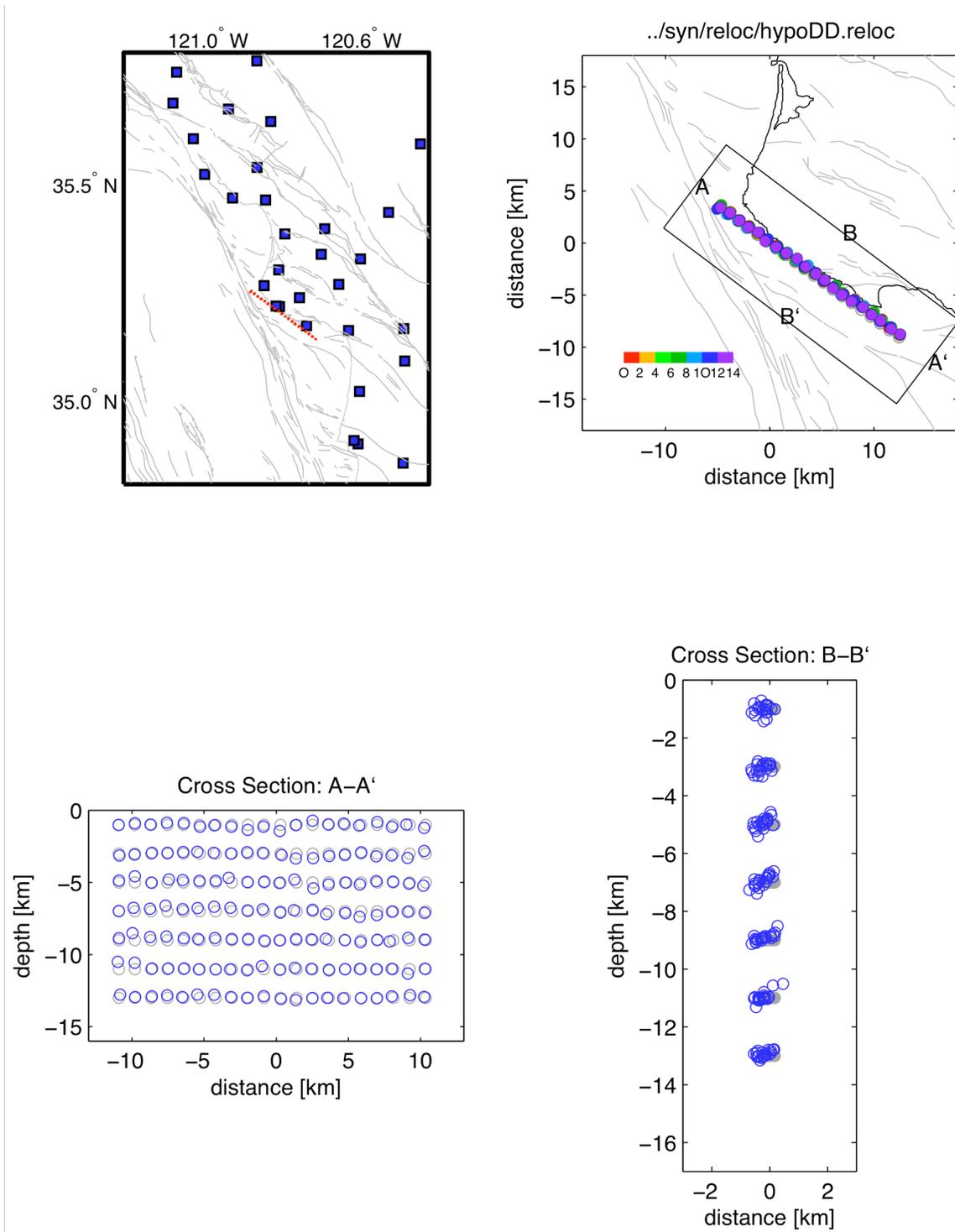


Figure 14 Same as Figure 13, but with the blue circles now representing the DD relocated sources obtained from the starting locations shown in Figure 13. Gray circles are true locations. Gaussian noise with a standard deviation of 0.1 s was added to the perfect data before relocation.

4. Conclusions

A comparison between the six different sets of earthquake locations (1D vs. 3D, for each of the large, local, and fault scale event sets) is given in Figure 15 and Figure 16. Figure 15 shows epicenter locations from all six data sets superimposed and color coded for comparison. Figure 16 shows, for each event along the SFZ, the deviation in hypocenter locations in horizontal and vertical direction from the mean of all six solutions. Two main trends can be identified. Absolute epicenter locations shift west with increasing area of relocation, and events locate deeper in 3D solutions (DD3D) compared to 1D solutions (DD1D). The west shift from fault scale DD relocations (DDf) to full area relocations (DDa) is about 500 m (Figure 16, left panel). This is presumably due to the larger distances over which the off-shore events are linked to the on-shore events in large-scale relocations, and thus the off-shore events become more sensitive to the apparent faster velocities of the 1D model compared to the true structure. That shift is smaller for 3D solutions than it is for 1D solutions, indicating that the 3D model is a better representation of the true structure. 3D solutions are on average about 1 km deeper than 1D solutions (Figure 16, right panel). Again, this points to the differences between the two velocity models used, in particular at upper crustal depths. Depths are similar for different scales of relocation areas within the 1D and 3D solutions, indicating that relative depths are less sensitive to the larger distances over which offshore events are linked to onshore events.

Table 4 summarizes the differences between the various DD locations obtained in this study and the locations obtained by Jeanne Hardebeck (JH locations). Only the events along the SFZ are included in these statistics. The median absolute deviations are typically less than a few hundred meters from the JH locations. They are highest (~700 m) in vertical directions for hypoDD solutions obtained in the 1D model. For comparison, median vertical deviations of the 3D solutions from the JH locations are ~200 m. Maximum deviations range from several hundred of meters (north-south-direction) to less than 3 km in vertical direction. They are smallest for the 3D DD solutions. The SimulPS locations that were used as starting locations in all DD inversions show the largest deviations from the JH locations (~300 m). Relative location errors for events along the SFZ, estimated from a bootstrap analysis of the final double-difference vector, include, in most cases, the locations determined by J. Hardebeck (Figure 11).

The synthetic experiment demonstrates that with the available seismic stations and data accuracy the relative locations of the SFZ events can be constrained within less than 200 m on average. The synthetic results indicate no systematic bias in the source locations and no distortion of the general geometry of the simulated vertical fault plane.

In summary, this study confirms the spatial distribution of the seismic activity along the Shoreline Fault Zone as imaged by J. Hardebeck. It is pointed out again here that the underlying seismic data used in this study is the same as used by J. Hardebeck to generate her solutions, as are the fundamental equations to solve for relative event locations. This study focused on testing the robustness of Hardebeck's solutions using different velocity models and various parameters that control the network of delay time links and the behavior of these data during their inversion for relative hypocenter locations.

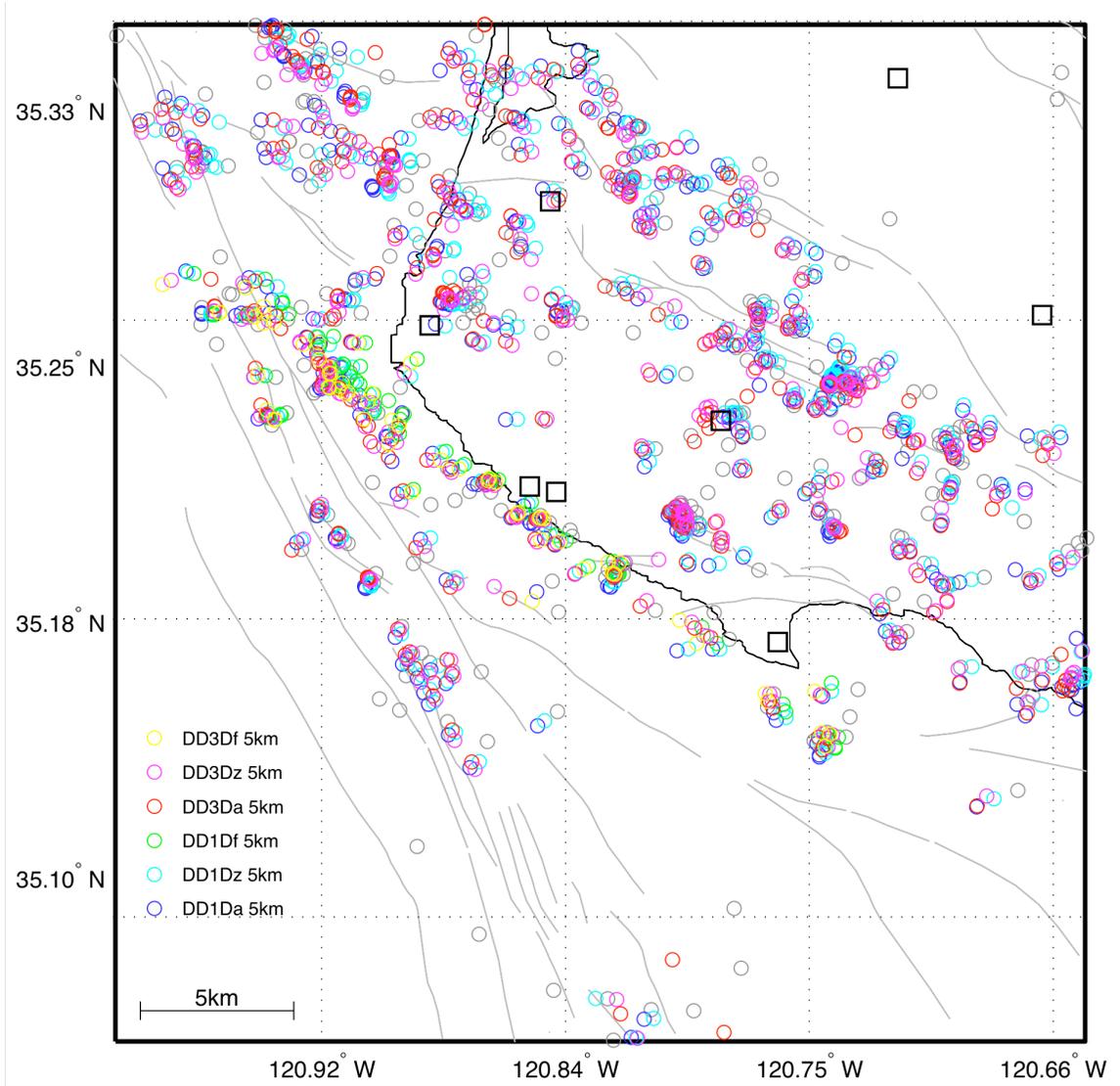


Figure 15 Superposition of epicenters from the six different inversions carried out in this study and shown in Figure 4 - Figure 9. Gray circles are JH locations. DD1Da, DD3Da: 1D and 3D solutions for all events in the polygon shown in Figure 1; DD1Dz, DD3Dz: 1D and 3D solutions for all events in rectangular area centered on the Shoreline Fault Zone; DD1Df, DD3Df: 1D and 3D solutions for events along the Shoreline Fault Zone.

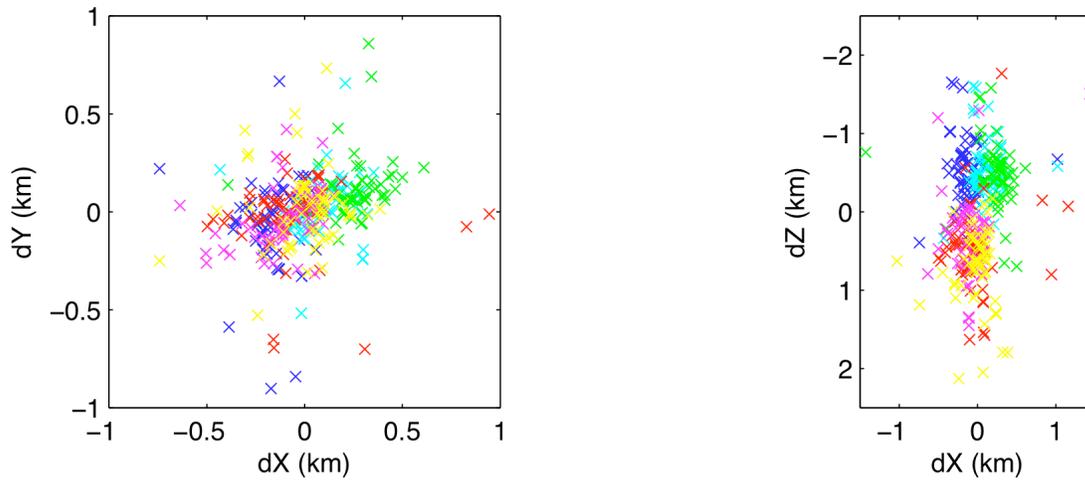


Figure 16 Relative hypocenter locations (in km) in horizontal (left panel) and vertical (right panel) direction after subtracting the mean of all six solutions for each event along the SFZ. Color-coded as in Figure 15.

Table 4 Summary of absolute differences between Jeanne Hardebeck’s tomoDD locations (JH) and the various relocation runs carried out in this study, for the 65 events along the Shoreline Fault Zone (in km):

JH vs.	DX		DY		DZ	
	Median	Max	Median	Max	Median	Max
SimulPS	0.39	1.39	0.23	1.67	0.32	2.84
DD1Da	0.18	1.47	0.21	0.58	0.73	2.97
DD1Dz	0.29	1.50	0.18	0.63	0.71.	2.91
DD1Df	0.41	1.64	0.15	0.69	0.72	2.82
DD3Da	0.12	1.25	0.12	0.95	0.20	1.73
DD3Dz	0.18	1.03	0.18	0.89	0.22	1.81
DD3Df	0.19	1.17	0.20	1.12	0.21	2.09

5. Electronic Files

- ASCII files with double-difference hypocenter locations for all case studies. The files can be downloaded from www.ideo.columbia.edu/~felixw/PGE/

6. References

- McLaren, M. K. and W. U. Savage (2001), Seismicity of South-Central Coastal California: October 1987 through January 1997, *Bull. Seism. Soc. Am.* 91, 1629-1658.
- Oppenheimer, D. H., F. W. Klein, J. P. Eaton, F. W. Lester (1993), The Northern California Seismic Network bulletin, January-December 1992, *U.S. Geol. Surv. Open-file Report*, 93-578, Menlo Park, California.
- Paige, C. C. and M. A. Saunders (1982). LSQR: Sparse linear equations and least squares problems, *ACM Transactions on Mathematical Software* 8/2, 195–209.
- Rietbrock, A. and F. Waldhauser (20014), A narrowly spaced double-seismic zone in the subducting Nazca plate, *Geophys. Res. Lett.*, 31, L10608, doi:10.1029/2004GL019610.
- Thurber, C., H. Zhang, F. Waldhauser, J. Hardebeck, A. Michael, and D. Eberhart-Phillips (2006), Three-Dimensional Compressional Wavespeed Model, Earthquake Relocations, and Focal Mechanisms for the Parkfield, California, Region, *Bull. Seism. Soc. Am.* 96, 38-49.
- Waldhauser, F. and W.L. Ellsworth (2000), A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, *Bull. Seism. Soc. Am.* 90, 1353-1368.
- Waldhauser, F. (2001), HypoDD: A computer program to compute double-difference earthquake locations, *U.S. Geol. Surv. Open-file Report*, 01-113, Menlo Park, California.
- Waldhauser, F. and D.P. Schaff (2008), Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods, *J. Geophys. Res.* 113, B08311, doi:10.1029/2007JB005479.
- Zhang, H. and C.H. Thurber (2003), Double-Difference Tomography: The Method and Its Application to the Hayward Fault, California, *Bull. Seism. Soc. Am.*, 93, 1875-1889, DOI: 10.1785/0120020190.