Final Report USGS/NEHRP Award No. 07HQGR0023

Rapid Determination of Near-Fault Earthquake Deformation Using LIDAR

Jean-Bernard Minster & Adrian Borsa
Institute of Geophysics and Planetary Physics
Scripps Institution of Oceanography
University of California at San Diego
La Jolla, CA 92093-0225
jbminster@ucsd.edu, aborsa@ucsd.edu

Abstract

The recently-completed airborne LIght Detection and Ranging (LIDAR) survey of the southern San Andreas, San Jacinto and Banning faults (the "B4 Survey") has delivered a high-resolution digital elevation model (DEM) of 1100 km of the most seismically active fault systems in southern California for the express purpose of providing a baseline for post-earthquake slip determination. We used the B4 Survey as a testbed to develop a processing algorithm that rapidly estimates near-fault ground deformation using simultaneous cross correlation of both topography and backscatter intensity from pre/post-earthquake LIDAR datasets. We show robust recovery of the direction and magnitude of an applied synthetic slip of 5 m in the horizontal and 0.5 m in the vertical, with excellent discrimination between areas with and without applied slip. Our results indicate that we should be able to accurately recover horizontal slip ≥0.5 m and vertical slip ≥3 cm. We also investigated misfit between overlapping data swaths to determine whether we could mitigate GPS trajectory error in order to improve comparisons between independent surveys of a fault zone. Significant calibration errors in geolocation resulted in across-swath errors throughout the B4 Survey of equal or greater magnitude than the expected vertical GPS error, postponing this aspect of the work until the data can be recalibrated and reprocessed.

Introduction

Despite the burgeoning deployment of new geodetic technologies capable of sensing ground movement on a variety of spatial and temporal scales, direct measurement of surface deformation and slip in the near-field of an earthquake rupture remains problematic. The Global Positioning System (GPS) is capable of sub-centimeter measurements of surface deformation in near real time, but even the dense Plate Boundary Observatory network has a spatial resolution on the order of several kilometers at best. Interferometric Synthetic Aperture Radar (InSAR) offers sub-centimeter accuracy, broad areal coverage and spatial resolution on the order of tens of meters, but typically suffers from decorrelation in the near-field of a rupture and in areas of dense vegetation cover. Attempts to address InSAR's near-fault decorrelation problem have led to newer techniques that estimate earthquake displacements by cross-correlating SAR amplitude or aerial/satellite pan-chromatic (e.g. SPOT) images. Although these methods yield measurements near the fault, they are limited to horizontal ground displacement and suffer from poor performance in homogeneous regions of the image.

Airborne LIDAR has the capacity to address some of the shortcomings of other measurement methods in the near-field of an earthquake rupture. Also known as airborne laser swath mapping (ALSM), the technique employs an aircraft-mounted laser to measure sub-decimeter surface topography at a spatial resolution of better than 50 cm. Spatial coverage is determined by the width of the survey corridor, which typically extends no more than a few kilometers from either side of the fault. For the purpose of earthquake slip determination, LIDAR data from a survey prior to an earthquake can be compared with data from a post-earthquake survey, the method of comparison being a question we address in this work. Compared with other methods for measuring earthquake slip, LIDAR makes up for its modest spatial coverage with a combination of superior spatial resolution [Figure 1] and unambiguous determinations of both horizontal and vertical deformation.

The problem of determining coseismic deformation in the near-field of an earthquake is relevant to both science/hazards research and hazard mitigation. One science question that would be addressed by mapping slip across the entire near field of a ruptured fault is whether field measurements of slip heterogeneity reflect actual slip versus imperfect capture of distributed slip occurring off the mapped rupture. Separately, the ability to map both the surface rupture and spatially distributed ground displacement will reveal much more about the fault geometry at depth than is revealed by the surface rupture alone. This has implications for how stresses behave on the fault, which in turn impacts earthquake rupture dynamics. Finally, knowing how the ground deforms in the entire near-field of a large earthquake adds an important observational element to seismic hazard models close to fault zones.

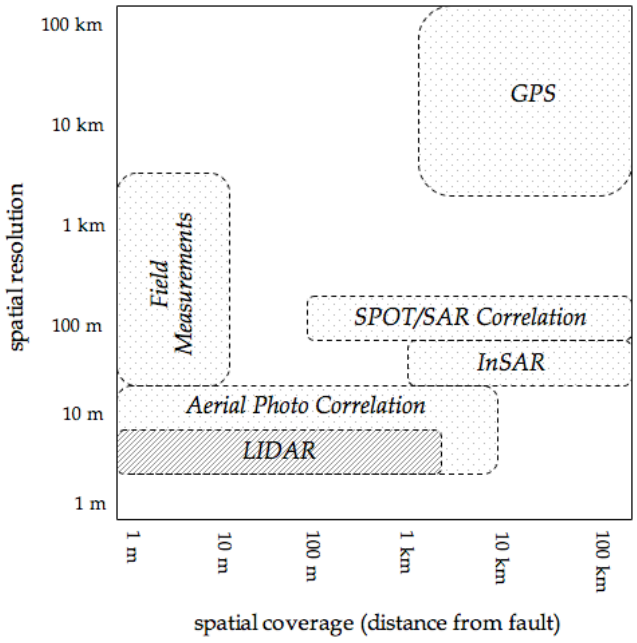

Figure 1. Estimated spatial resolution and spatial coverage of LIDAR compared to other methods of determining earthquake slip. Axes are in log scale.

Quickly identifying the main rupture trace and associated ground deformation (slumping, landslides and distributed strain) of a damaging earthquake is another primary target of our work under this grant. If LIDAR resources were quickly mobilized after an earthquake on a mapped fault, a fast, semi-autonomous processing mechanism could generate a comprehensive map of near-fault ground deformation only a few hours after the post-earthquake survey was processed. Such a map in the hands of earthquake responders would allow them to make quantitative assessments of potential damage to fault-crossing lifelines and adjacent structures, particularly in areas where distributed deformation off the fault trace would not be visually apparent.

The processing mechanism that we have developed for this project has the potential to reduce earthquake losses in the U.S. on two levels, both necessarily contingent on the

availability of a pre-earthquake DEM and post-earthquake LIDAR data acquisition. Its immediate impact in the aftermath of an earthquake would be to reduce potential economic and societal losses by making it possible to provide earthquake response teams with a deformation map to use for quickly identifying at-risk infrastructure in the vicinity of the rupture. Over the longer term, important insight would be obtained about ground motion in the near-field of an earthquake, helping to refine hazard maps in the vicinity of faults and engineering standards for near-fault structures.

Technique

This grant funded the development of an algorithm that can cross-correlate airborne LIDAR point-cloud data from separate surveys to determine the horizontal and vertical displacement field that maps points in one dataset to related points in the second dataset. Since this algorithm works on a pixel-by-pixel basis (where a "pixel" consists of data points within a bounding box of specified dimensions), it can be used to map any spatial deformation pattern for which deformation at the pixel scale is minimal.

Since this study is particularly focused on recovering the slip resulting from a moderate-to-large earthquake using airborne LIDAR data, and since there are currently no existing

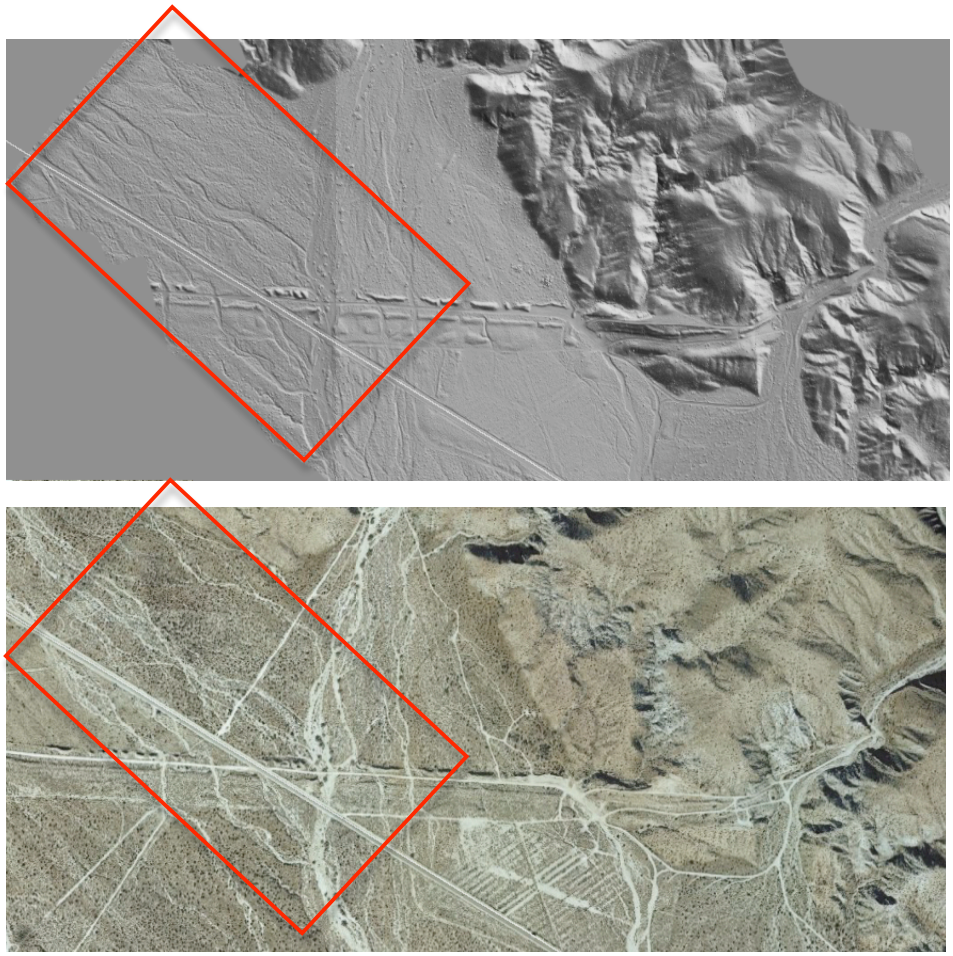

Figure 2 [Top] Topography from B4 Survey LIDAR on southern SAF near Desert Hot Springs. [Bottom] Aerial photo of above. Red outlines in both images show test area used in this study.

LIDAR datasets that span an actual earthquake, we developed our algorithm using data from the B4 survey of the southern San Andreas Fault (SAF) [Figure 2], synthetically slipped to simulate the offset that might result locally from a large (M7+) event [Figure 3]. In the specific case examined here, we assumed 5 m right-lateral horizontal slip and 0.5 m positive vertical slip on a 90° dipping fault plane roughly aligned with the trace of the SAF. All motion is confined to the northeast side of the fault (i.e. on the North American Plate), which allows us to use the unslipped southwest portion of the data as an experimental control.

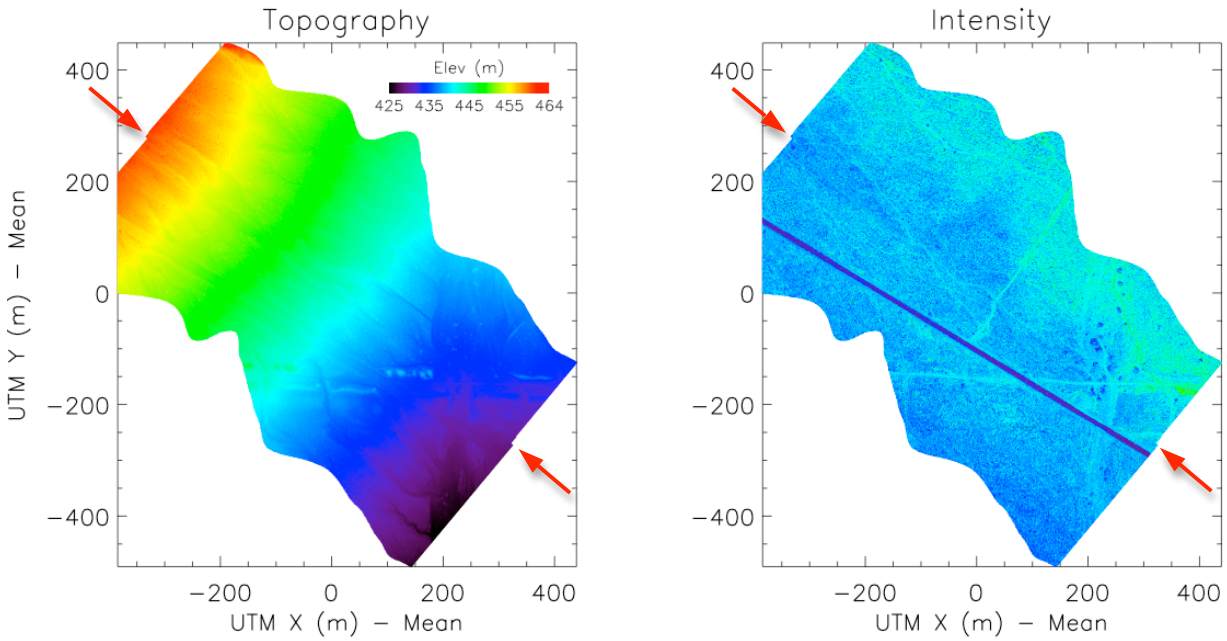

Figure 3 [Left] Raw point cloud for a small section of the B4 survey (cf. Figure 2), colored for elevation. [Right] Same as left, but with points colored for LIDAR backscatter intensity in unspecified units. Red arrows in both images show location of synthetic slip plane.

The problem of comparing point cloud datasets is conceptually similar to that of using cross-correlation analysis on 2D images. However, one cannot directly apply any of these methods to a point cloud because of the difficulty of correlating individual points between datasets. Converting a point cloud to a digital elevation model (DEM) that ostensibly describes the "true" topographic surface is a way of supplying the additional information needed to compare point clouds, via the particular autocorrelation relationship assumed when filtering and resampling the raw data to a regular grid.

Because we wish to evaluate data gradients analytically in a later step, we do not explicitly create a DEM from our data, but instead model the point cloud by fitting it with the low-order 2-D harmonic (Fourier) basis set

$$H(x,y)=\sum_{k=0}^{m}\sum_{l=0}^{n}a_{k,l}\left[\sin\left(\frac{2\pi x}{L_x}k\right)+i\cos\left(\frac{2\pi x}{L_y}l\right)\right]\left[\sin\left(\frac{2\pi y}{L_x}k\right)+i\cos\left(\frac{2\pi y}{L_y}l\right)\right] \qquad (1)$$

where x and y are the UTM coordinates of the data, L is the size of the fitting region in the coordinate directions and the complex-valued coefficients $a_{k,l}$ are found via least squares fitting to point cloud elevations. To ensure that the fit resolution is identical in both

coordinate directions, we set $L_x = L_y$ and $m = n$, although the optimal values of these parameters is an issue we discuss in the results section below. We always take L to be twice the span of the data in order to handle edge effects related to the periodic extension of the fitting window in the plane. Since the autocorrelation assumed by harmonic fitting is sinusoidal, this choice of L also means that the autocorrelation function for the lowest-order harmonic is that of a cosine window over the extent of the data, which adequately captures the long-wavelength content of the point-cloud.

Point cloud data also contains information about the intensity of ground return energy from the laser footprint [Figure 3, right]. Although intensity measurements are somewhat noisy, they are spatially coherent and can complement elevations in the determination of slip, especially in areas where the topography is flat. We model point cloud intensities exactly as we do elevations in Equation (1), yielding the fitting function $I(x, y)$. Both $I(x, y)$ and $H(x, y)$ can be calculated for the "before" dataset in advance of collection of the "after" dataset, reducing the time required for the topographic comparison when a second point cloud finally becomes available.

It is not necessary to fit the "after" point cloud to make a topographic comparison between datasets. Instead, the N points (x_i, y_i, z_i, i_i) within a small comparison pixel of the "after" dataset (i denotes the intensity measurement) are evaluated relative to $I(x, y)$ and $H(x, y)$ to estimate the offset vector $\boldsymbol{x} = <\Delta x, \Delta y, \Delta z>$ that simultaneously minimizes the elevation and intensity misfits between the two datasets. Written as a generalized least squares problem, we are trying to find the $\boldsymbol{x}$ that minimizes

$$\sum_{i=1}^{N}\left[\frac{z_i - \left[\Delta z + H\left(x_i + \Delta x, y_i + \Delta y\right)\right]}{\sigma_{z_i}}\right]^2 \text{(topography)} \; + \; \lambda\sum_{i=1}^{N}\left[\frac{i_i - \left[I\left(x_i + \Delta x, y_i + \Delta y\right)\right]}{\sigma_{i_i}}\right]^2 \text{(intensity)} \quad (2)$$

where λ is a factor that determines the relative weights of topography and intensity in the inversion. $\boldsymbol{x}$ is determined independently for each pixel, which means the algorithm is not guided by an *a priori* slip model. While this increases variability in the recovered slip, it also makes the algorithm much more useful in situations where slip or deformation is unknown. On a practical note, because I and H are functions of $\boldsymbol{x}$, Equation (2) is non-linear and must be solved iteratively.

One caveat to this study is that we do not attempt to model the effect of data error on slip recovery. Point-cloud data are the output of a multi-step geolocation process, subject to calibration error that can obscure or mimic the topographic change signal. Our results should be considered to be best-case, with further work needed to determine the potential impact of data error.

Results and Discussion: Part 1

Figure 4 shows results from the application of our algorithm to the synthetically slipped data in Figure 2. The algorithm is successful in recovering both horizontal and vertical slip, as the plots on the left show. There is a clear delineation between the unslipped control region below the slip plane and the slipped region above, recovered vectors in the slip region exhibit strong correlation in both magnitude and direction, and anomalous vectors are relatively few in number.

Slip estimation is less accurate at the boundaries of the slipped data where discontinuities are expected to impact the harmonic fit to topography and intensity. The algorithm also has some difficulty with the road running through the control region (the dark blue line in the Figure 3 intensity plot), although in the context of the rest of the control region, the vectors along the road would not be interpreted as having a geophysical origin.

The histograms on the right of Figure 4 provide a quantitative interpretation of the recovered slip and show that misfits in both the slipped and unslipped regions are nearly zero-mean and have well-behaved distributions. The interquartile range (IQR) of horizontal magnitude misfit is nearly identical in both the slipped and unslipped regions (0.51 m and 0.43 m, respectively), as is the vertical magnitude misfit (2.6 cm and 2.9 cm). This indicates that slip recovery accuracy is independent of imposed slip and suggests that the lower limit of slip detection using this algorithm is ~0.5 m in the horizontal and ~3 cm in the vertical.

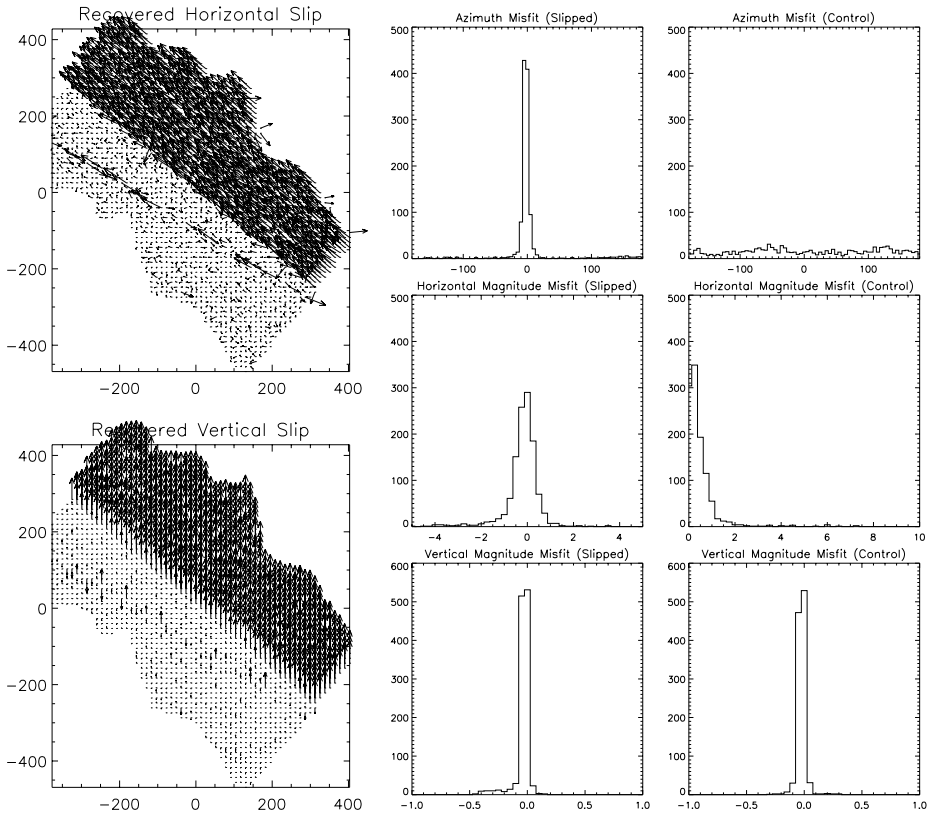

Figure 4. Results of applying the slip recovery algorithm to the data shown in Figure 2 using a pixel size of 13 m, fit resolution to "before" data of 9 m and equal weight between topography and intensity. Vector magnitudes on the recovered slip plots are normalized to the maximum value plotted, so the horizontal and vertical slip maps cannot be compared directly. Note that the recovered slip is left-lateral, due to incorrectly applied synthetic slip.

Although the solution to the least squares problem in (2) is a single optimal slip vector for each data pixel in the "after" dataset, there are three *a priori* parameters that control the particular values of $\boldsymbol{x}$ that are obtained. These are the pixel dimensions controlling the number and spatial coverage of "after" data used in each calculation; the resolution of the fit to the "before" data implied in the choices of L, m and n in $I(x, y)$ and $H(x, y)$; and the relative weight of topography versus intensity in the least squares solution.

We examined the effect of varying all these parameters over a range of values by running the analysis for Figure 4 using each parameter combination. Pixel size has the smallest effect on overall misfit for the range of values examined (10 m, 13m, 16m), although larger pixels did reduce misfits for nearly every combination of the other parameters. Smaller pixels have the advantage of increasing the resolution of the displacement map, but at a computational cost that scales as the inverse square of the reduction in pixel edge length.

Increased fit resolution to the "before" data improved misfit to the slip data, but only up to a point. We examined resolutions in the range of 4~26 m and found that resolutions from 6~10 m yielded the lowest misfits. Figure 5 shows the effect of decreasing resolution from 9 m to 13 m in the analysis used for Figure 4. While the overall pattern of the synthetic slip is adequately recovered, higher variability in the displacement field shows up as a broader distribution of misfits in both the slipped and control regions. Interestingly, increasing resolution to 4 m increased misfit in the slipped region, but reduced it in the control region. Although we did not have an opportunity to examine the misfit function, the most obvious explanation for this behavior is that the higher resolution increases the depth and/or number of alternative minima on the misfit manifold. We may want to explore alternative methods of solving for slip (e.g. simulated annealing) to address this issue.

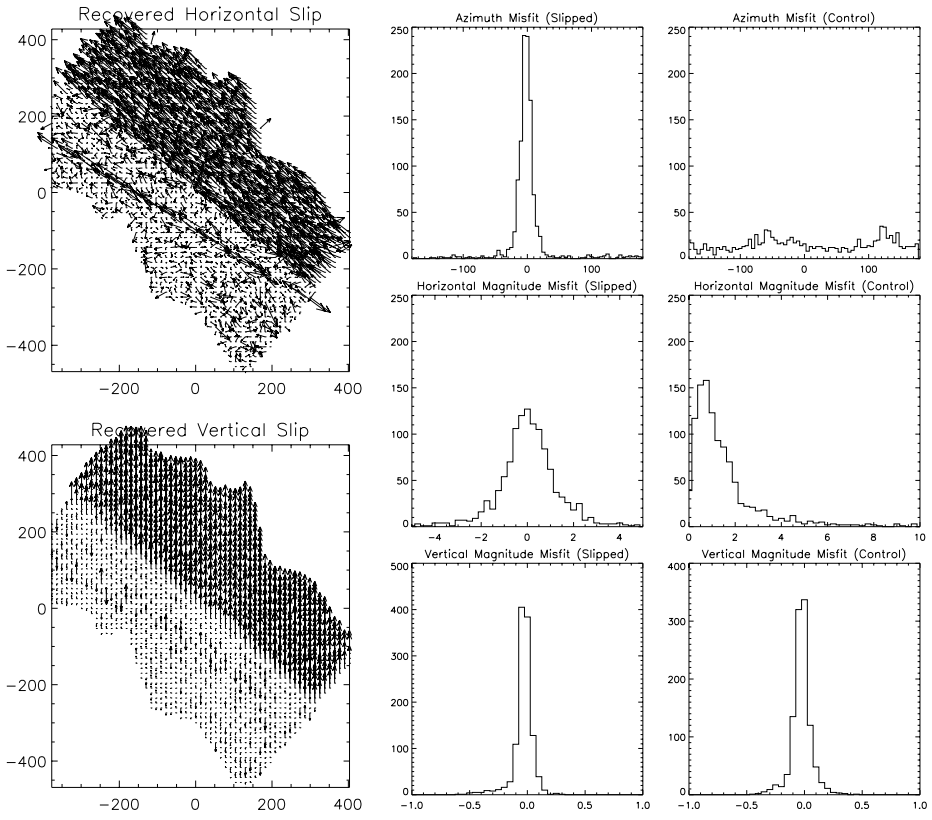

Figure 5. Same as Figure 4, but with 13 m fit resolution to "before" data instead of 9 m.

Finally, we also ran through a range of λ values to alter the relative weights of topography and intensity in the least squares solution. Misfit improves by introducing intensity data, with about equal weighting of topography and intensity providing the most benefit. As

intensity is weighted higher, it begins to negatively impact the misfit, although the λ value at which this occurs depends on the other parameters discussed above.

Results and Discussion: Part 2

Although we proposed to adapt to airborne LIDAR the GPS noise modeling algorithm we use to remove noise in vehicle-based kinematic surveys, we discovered early in our work that B4 survey data contains large and unexpected across-swath errors due to miscalibration of the LIDAR instrument and/or the aircraft inertial navigation system. Unfortunately, these errors make it impossible to independently estimate vertical GPS error from swath overlaps as is required by our noise modeling algorithm, postponing this aspect of the work until the data can be recalibrated and reprocessed.

We did considerable analysis of the calibration problem and presented it at the 2007 AGU Fall Meeting (Borsa, A.A., Bevis, M. & Hudnut, K.W. (2007), Survey-scale airborne lidar error analysis from parallel swath comparison, Eos Trans. AGU, 88(52), Fall Meet. Suppl., Abstract G51B-0433). See Appendix A for a small-format version of the poster.

Conclusions

While the work presented here has yielded promising results, we will have to better understand the shape of the cost function to optimize our approach to the displacement estimation problem. The presence of non-zero displacements in our zero-slip control region is a clear sign that our algorithm occasionally (yet systematically) picks the wrong minimum from multiple minima. The cost function for these anomalous estimates will tell us whether we should work to adjust our iterative least squares solution or investigate another optimization path entirely. Genetic algorithms are particularly suited for addressing problems with multiple minima, and we may be pressed to go this route if our current technique starts to break down when dealing with larger synthetic slip.

Nevertheless, the algorithm in its current state of development adequately addresses the hazard mitigation applications discussed in the introduction of this report and in our original proposal. To earthquake responders, consistent decimeter-level accuracy is less important than knowing the approximate size and location of ground deformation so that damage to nearby structures can be evaluated in the shortest time possible. Our algorithm accomplishes this goal, and we will make the source code available as a supplement to a journal publication that will follow this report. The real test will be when LIDAR mapping after a future earthquake generates the first pair of surveys to serve as real input into our algorithm.

Appendix A

(2007 Fall AGU Meeting, Abstract G51B-0433)