

## GrowClust: A Hierarchical Clustering Algorithm for Relative Earthquake Relocation, with Application to the Spanish Springs and Sheldon, Nevada, Earthquake Sequences

by Daniel T. Trugman and Peter M. Shearer

### ABSTRACT

Accurate earthquake locations are essential for providing reliable hazard assessments, understanding the physical mechanisms driving extended earthquake sequences, and interpreting fault structure. Techniques based on waveform cross correlation can significantly improve the precision of the relative locations of event pairs observed at a set of common stations. Here we describe GrowClust, an open-source, relative relocation algorithm that can provide robust relocation results for earthquake sequences over a wide range of spatial and temporal scales. The method uses input differential travel times, cross-correlation values, and reference starting locations, and applies a hybrid, hierarchical clustering algorithm to simultaneously group and relocate events within similar event clusters. The method is computationally efficient and numerically stable in its capacity to process large data sets and naturally applies greater weight to more similar event pairs. Additionally, it outputs location error estimates that can be used to help interpret the reliability and resolution of relocation results. As an example, we apply the GrowClust method to the recent Spanish Springs and Sheldon, Nevada, earthquake swarms. These sequences highlight the future potential for applying the GrowClust relocation method on a much larger scale within the region, where existing relocation results are sparse but vital for understanding the seismotectonics and seismic hazard of Nevada and eastern California.

*Electronic Supplement:* Maps of Spanish Springs and Sheldon sequences with relocation uncertainty, and tables of velocity models and relocated catalogs for the Spanish Springs and Sheldon Sequences.

### INTRODUCTION

The location of earthquake hypocenters using observations of seismic phase arrival times is a classic inverse problem in geo-

physics, with a rich history of conceptual and methodological advancements dating back more than a century. During this time, catalogs of earthquake locations became one of the most important and widely used forms of seismological data. They provide fundamental constraints on a number of important seismotectonic problems, from the resolution and imaging of fault-zone structure, to the understanding of the physical mechanisms underlying earthquake triggering and interaction, to the improvement of seismic-hazard assessments. In turn, the degree to which these and other outstanding problems can be resolved depends intrinsically on the quality and reliability of earthquake location methodology.

The travel time of a seismic phase observed at a given station depends nonlinearly on both the earthquake hypocentral coordinates and the subsurface velocity structure (Geiger, 1910; Buland, 1976; Thurber, 1985). Because of this, improvements in absolute earthquake location accuracy will always be limited by imperfect knowledge of 3D variations in the Earth structure (Thurber, 1983, 1992; Thurber and Eberhart-Phillips, 1999). However, in recent years, numerous methods have been developed which yield significant improvements to relative earthquake location accuracy through the joint relocation of pairs or clusters of linked events (Douglas, 1967; Frohlich, 1979; Jordan and Sverdrup, 1981; Got *et al.*, 1994; Shearer, 1998; Richards-Dinger and Shearer, 2000; Waldhauser, 2000; Lin *et al.*, 2007). Although initial catalog locations are routinely determined from the noisy, often emergent phase picks and associated travel times of each event in isolation, relative relocation techniques are based primarily on differential travel times of pairs of events observed at common stations. This formulation helps mitigate common-mode errors introduced by the biasing effects of unmodeled velocity structure. Furthermore, waveform cross-correlation techniques (e.g., Poupinet *et al.*, 1984; Ito, 1985; Fremont and Malone, 1987; Nadeau *et al.*, 1995; Phillips *et al.*, 1997; Rowe *et al.*, 2002) can be

used to extract differential travel times with considerably greater precision than is possible from the absolute travel times alone. Application of these relative relocation techniques can therefore result in a dramatic sharpening in seismicity and resolution of fine-scale fault structure (e.g., Rubin *et al.*, 1999; Waldhauser *et al.*, 1999; Astiz and Shearer, 2000; Shearer, 2002) compared to standard catalog locations.

Implementation of these methods is straightforward in the case of single, compact clusters of events, for which the common mode errors can be approximated as identical, but becomes more complicated for distributed seismicity for which the errors associated with unmodeled 3D velocity structure vary considerably with event location. The most widely used approach to relocate distributed seismicity is the double-difference (DD) method (Waldhauser, 2000), which sets up a linear system of equations for the differences in event locations as a function of the input 1D velocity model and differences in differential times (as well as absolute times, if desired) and then applies matrix inversion. The DD technique has been applied to problems across a wide range of scales, including relocating over 500,000 events in the northern California earthquake catalog (Waldhauser and Schaff, 2008).

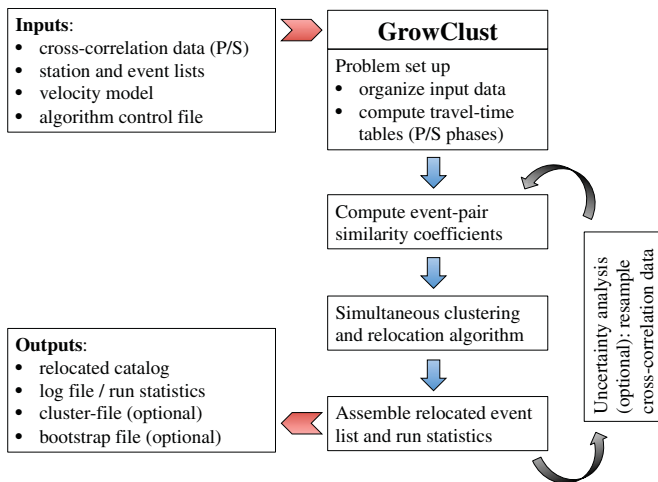
Here we describe an alternative to DD for relocating seismicity using differential times, which we term the GrowClust algorithm and which we are making available to the community as an integrated software package (see [Data and Resources](#)). Given the success of the DD algorithm and its growing user base, why did we develop a different method? One motivation was to permit applying a more robust misfit criteria than the L2 norm (least squares) that is used in standard matrix inversion. GrowClust uses the L1 norm, which is less sensitive to outliers (bad data points) in the input times and has been shown in some cases (e.g., Shearer, 1997) to yield improved results compared to L2-norm solutions. Another concern is that for many problems the matrix representing the complete solution in DD is so ill-conditioned that inversions become unstable. For example, this can occur when event clusters are linked by differential times from a single cross-correlation pair, or a chain of event pairs, and hence the stability issue becomes more pronounced for large-scale relocation problems containing multiple discrete clusters of seismicity. The DD algorithm does provide various algorithm control parameters to help stabilize the inversion by, for example, downweighting data from more distant or more poorly correlated event pairs. However, the optimal parameter choices and iteration-by-iteration sequential weights required to obtain the best locations are not always obvious to the user, and even these optimal parameter choices may not guarantee stability for certain problems. Our approach is to apply cluster analysis ideas to decide which groups of events are linked and relocated and which cluster pairs retain independent locations. Although our method has its own set of parameter choices, we designed them to be as few and as straightforward as possible. Finally, we wanted to implement a method that runs fast enough that bootstrap resampling to estimate location errors is practical on small computer systems.

The GrowClust algorithm unifies a series of programs that were developed to relocate seismicity in southern California using waveform cross correlation (Shearer, 2005; Lin *et al.*, 2007; Hauksson *et al.*, 2012). GrowClust combines cluster analysis and relocation of events within each cluster, which were previously performed as two separate steps. This provides a significant boost in computational efficiency and convergence stability for large data sets, which typically contain multiple large clusters. GrowClust uses differential travel times and cross-correlation results in a hybrid, hierarchical clustering algorithm that both groups events into clusters based on waveform similarity and relocates each event with respect to the linked events within its unique cluster. The method is computationally efficient and multiscale in its applicability to both small and large sequences. Unlike many other related relocation techniques, the GrowClust algorithm intrinsically gives greater weight to more highly correlated event pairs and does not require explicit matrix inversion.

A preliminary version of the GrowClust algorithm was first applied to perform high-precision relocation of a large data set of more than 100,000 events recorded on the Island of Hawaii from 1992 to 2009 (Matoza *et al.*, 2013). We present here a more complete description of the improved, newly open-source algorithm, which now implements a non-parametric statistical resampling technique to estimate errors in the relocated event positions, which proves useful in assessing and interpreting the relocation results. As a demonstration of the new method, we use GrowClust to relocate two prominent, recent earthquake sequences in western Nevada: (1) the 2012–2015 Spanish Springs swarm (mainshock  $M_w$  4.2), which caused nonstructural damage and significant shaking in Reno, and (2) the ongoing, 2014–present Sheldon swarm, an extensive sequence (dozens of  $M$  4 and hundreds of  $M$  3 events to date) occurring in the northwest corner of the state, where station coverage is sparse and catalog locations are particularly scattered. Large-scale, systematic event relocations in northern (Waldhauser and Schaff, 2008) and southern (Hauksson *et al.*, 2012) California have significantly advanced our understanding of seismicity, seismic hazard, and fault structure within that state. Our initial results suggest that, with methodological improvements and the recent modernization of the Nevada Seismic Network (NSN; Kent *et al.*, 2015), the same potential for scientific advancement exists within Nevada and eastern California.

## METHODS AND ALGORITHM DESCRIPTION

In this section, we detail the methodology and computational details underlying the current implementation of GrowClust. As we intend this code to be open-source (see [Data and Resources](#)), future releases of the code may include minor modifications to the algorithm presented here. We begin this section with a brief outline of the data-preprocessing steps required to obtain the necessary inputs for GrowClust. Next, we describe the basic algorithm used by GrowClust, which simultaneously groups individual earthquakes into similar event



▲ **Figure 1.** The basic GrowClust workflow chart. Users input cross-correlation data, station and event lists, a velocity model, and an algorithm control file into the GrowClust source code. GrowClust then organizes the input data and creates travel-time tables for  $P$  and  $S$  body-wave phases. After computing event-pair similarity coefficients, GrowClust initiates its simultaneous clustering and relocation algorithm. Finally, after obtaining the final, relocated event positions, the program (optionally) resamples the input data for bootstrap uncertainty analysis, or proceeds directly to output of catalog, cluster, and log files.

clusters based on the input cross-correlation data and relocates each earthquake with respect to its cluster neighbors. Finally, we describe the resampling approach that GrowClust (optionally) performs to assess uncertainties of the relocated event positions. The basic GrowClust workflow is summarized in Figure 1.

### Data Preprocessing

The fundamental input data for the GrowClust algorithm are differential travel times and cross-correlation values, obtained through waveform cross correlation of sets of earthquake event pairs observed at common stations. Both are needed, because the differential times are used to relocate the events, whereas the cross-correlation values are used to weight the data by quality and group events by waveform similarity. However, the algorithm is sensitive only to the relative values of cross-correlation coefficients, because the absolute values depend on both the data quality and auxiliary factors, such as the length of the cross-correlation window and filter type, if any is used.

GrowClust is flexible in its accommodation of cross-correlation results obtained using any method, whether time domain or frequency domain, and can use any combination of  $P$ - and  $S$ -phase cross-correlation results. For the results presented in this article, we use a time-domain cross-correlation approach that uses spline interpolation to achieve millisecond precision in differential times. Frequency-domain techniques that achieve similar precision (e.g., Poupinet *et al.*, 1984) should work equally well. For our approach, we find that applying a zero-phase band-pass filter from 1 to 10 Hz to the raw waveform data is useful in

isolating the relevant phase arrivals and in mitigating noise that may cause spurious cross-correlation results.

In addition to cross-correlation times, GrowClust requires a velocity model to compute predicted differential travel times. Though not requisite for the GrowClust method, we use travel-time tables derived from the velocity model to select cross-correlation windows around predicted  $P$ - and  $S$ -phase arrival times, a practice that can vastly expand the cross-correlation data set to include waveform pairs devoid of operator  $P$ - and  $S$ -phase picks. Finally, GrowClust requires input event and station lists that uniquely identify each earthquake (e.g., with an event identification number) and each station (e.g., with a station name). The various inputs to GrowClust, along with algorithm control parameters described in the following section, are combined in an input file read by GrowClust upon initial computation (Fig. 1).

### The GrowClust Algorithm

In this section, we provide a conceptual outline of the hybrid clustering algorithm used by GrowClust to perform relative event relocation. Upon program initiation, GrowClust reads the cross-correlation data, parameters from the algorithm control file, and input station and event lists into memory (Fig. 1). The latter contains the initial (catalog or other reference) hypocentral positions from which events are relocated. The program also constructs travel-time tables for both  $P$  and  $S$  body-wave phases based on the input velocity model, which are later used to compute the predicted differential times necessary for event relocation.

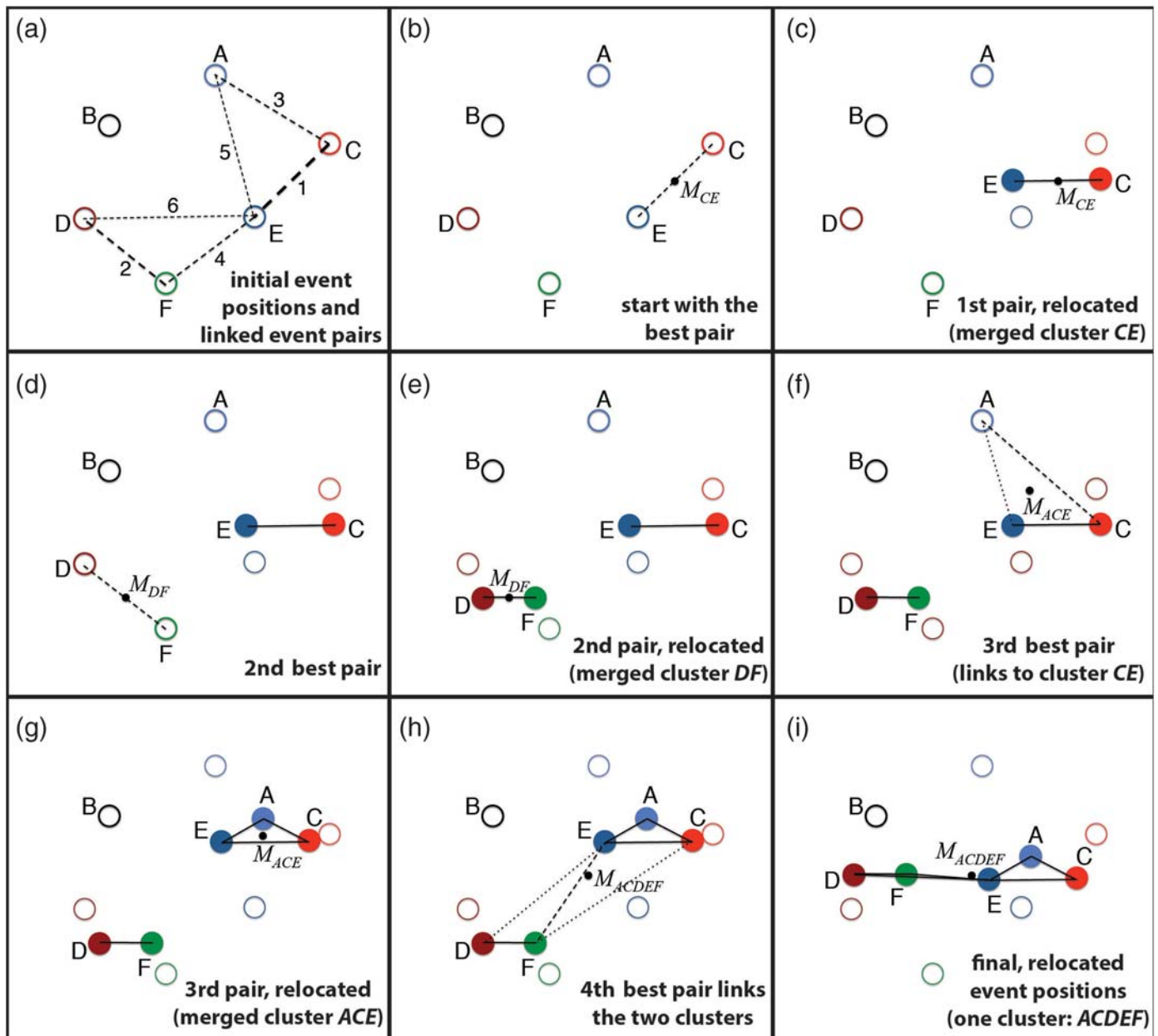
Following this initial input and data organization stage, GrowClust begins its hybrid clustering and relocation algorithm, which works as follows (see Fig. 2 for a simplified conceptual example):

1. Assign each of the  $N$  events to a distinct starting cluster number. Though many of these clusters will later be merged as part of the relocation process, the  $N$  initial clusters each begin as a single event.
2. For each event pair  $(i, j)$  compute a similarity coefficient  $Z_{ij}$  that serves as a metric to measure the data quality and waveform similarity of each distinct event pair. Here, we take  $Z_{ij}$  for an event pair to be a sum over the cross-correlation values  $r_{ij;k}$  observed at the  $k$  common stations within a maximum station distance  $\Delta_{\max}$  and that exceed a minimum value  $r_{\min}$ :

$$Z_{ij} = \sum_k r_{ij;k} \quad \forall \quad r_{ij;k} \geq r_{\min} \quad \text{and} \quad \Delta_k \leq \Delta_{\max}, \quad (1)$$

The control parameters  $\Delta_{\max}$  and  $r_{\min}$  are chosen by the user in the input control file based on the data set at hand, and cross-correlation data for  $P$  and  $S$  phases are treated equally, unless otherwise specified.

3. Sort the event pairs  $(i, j)$  by similarity coefficient  $Z_{ij}$  and process each event pair in turn, starting with the most similar. As the algorithm proceeds, there are three situations it may encounter when considering a new event pair (Fig. 2):



▲ **Figure 2.** Simplified example of GrowClust's hybrid clustering and relocation algorithm. In this example, there are six events (labeled with capital letters) and six linked event pairs with quality cross-correlation data (labeled with numbers, in sequential order of waveform similarity). Event B has no links, and hence its location is held fixed throughout the algorithm. (a) Initial catalog location of events A through F, with event pairs linked and numbered in order of decreasing waveform similarity. Note that all events begin as single-event clusters. (b) The most-similar event pair CE and its centroid  $M_{CE}$ . (c) The relocation algorithm begins by relocating the most-similar event pair about its centroid. The relocated pair now belongs to a single cluster CE. (d) The second event pair DF and its centroid  $M_{DF}$ . (e) Relocation of DF about its centroid. The relocated events are merged into a single cluster DF. (f) The third link connects event A to cluster CE. (g) Event A and cluster CE are now relocated about the total centroid  $M_{ACE}$  and merged into a single cluster ACE. Note that the relative positions of events C and E are held fixed. (h) The fourth link connects clusters DF and ACE. (i) Clusters DF and ACE are relocated about their total centroid  $M_{ACDEF}$  and merged into a single cluster. All events are now in their final, relocated positions. Note that links 5 and 6 are used in support of the final proposed cluster merger.

(a) Both event  $i$  and  $j$  are members of single-event clusters. In this situation, the events are merged into a new cluster (now with two events), and both are relocated with respect to the new cluster centroid. The relative relo-

cation algorithm uses a grid-search approach to find the relative locations that minimize the L1 norm of the residual between observed and predicted differential travel times:

$$\text{minimize } \|dR\|_{L1} = \sum_k |dtt_{ij,k} - \hat{dtt}_{ij,k}| \quad (2)$$

(Shearer, 2005). Here  $dtt_{ij,k}$  and  $\hat{dtt}_{ij,k}$  denote the observed and predicted differential travel times for the event pair  $(i, j)$  at station  $k$ , where the predicted times depend on the relative event locations and the velocity model. The grid search is performed over the relative event locations, and the L1 norm of the residuals is computed over the set of common stations with waveform observations of the event pair.

- (b) Either of events  $i$  or  $j$  are members of multievent clusters. In this situation, the algorithm performs a series of tests to decide whether to merge the clusters and relocate all events within both clusters with respect to one another. First, the algorithm searches for all other event pairs that link the two clusters. If the ratio of observed links to total possible links fails to exceed a specified threshold (e.g., 0.01), the cluster merger is rejected, and the event pair is skipped. Otherwise, the algorithm performs a test relocation of the two clusters with respect to one another using only the 10 strongest links (i.e., linking event pairs with the 10 highest  $Z_{ij}$  values) for robustness. The cluster relocation uses a multi-event generalization of the L1-norm approach applied to single-event clusters in situation (a). The relative positions of the events within each of the two distinct clusters are held fixed, whereas the two cluster centroids are adjusted about the centroid of the combined cluster. Once this trial relocation is performed, the algorithm checks to see whether the differential travel-time residuals of the newly merged cluster or the centroid shifts of the initial clusters exceed tolerance values specified by the user. These tolerance criteria help ensure the stability of the algorithm by preventing mergers that are not required by the data or that involve unreasonably large location adjustments that would violate the assumption of small location shifts implicit in relative relocation methods (Geiger, 1912). If neither criteria are violated, the two clusters are merged into a single combined cluster that contains all of events at their new relocated positions.
  - (c) Both events  $i$  and  $j$  belong to the same cluster. In this case, the algorithm simply skips the relocation of event pair, because both events have already been relocated further up the algorithm.
4. The algorithm continues processing clusters in this way until no more of them can be merged and relocated, given the algorithm control parameters. At this stage, GrowClust then computes the final run statistics for user assessment and saves these along with the relocation and clustering results for later output (Fig. 1). It is important to note that GrowClust is a purely relative relocation algorithm: events within each cluster are relocated with respect to one another, with the cluster centroids held fixed at their initial reference positions to ensure stability. As such, the events

that comprise single-event clusters due to lack of waveform similarity are not relocated by the algorithm and hence remain at their initial reference positions.

The overall strategy of GrowClust is to begin by relocating the highest quality event pairs and most similar event clusters and then hold the relative locations of these events fixed when computing additional relative locations. An advantage of this approach is that it permits a grid-search relocation method to be applied to cluster pairs in which there are only four free parameters at each relocation step: the  $dx$ ,  $dy$ , and  $dz$  offsets of one of the clusters from a reference position, plus any origin time shift ( $dt$ ). The algorithm is computationally efficient because clusters are treated as coherent event sets, which vastly reduces the effective degrees of freedom during each relocation step, because linking event pairs are not relocated independently. A further advantage of this approach is that it can be easily modified to optimize against any desired misfit norm, allowing for improved robustness through the use of the L1 norm rather than the conventional L2 norm.

It is worth noting that the internal architecture of GrowClust is well suited to some additional applications. Because the method naturally builds its results by performing new locations relative to a set of previously determined locations, it should be possible to adapt the method to compute near-real-time locations, in which new events are relocated with respect to previous events, rather than having to relocate the entire data set when new data become available. In addition, the grid-search location algorithm reads from a set of travel-time tables, which are currently computed for a 1D velocity model, but which, with only a slight increase in complexity, could be altered to work with travel-time tables based on a 3D velocity model. Finally, the procedure described below that is used to assess location uncertainties is fully parallelizable, and future implementations could be adapted to take advantage of this fact and further improve computational efficiency for large-scale problems.

### Relative Location Uncertainties

The complexity of the GrowClust algorithm precludes simple, parametric techniques for assessing formal uncertainties in the relocated hypocentral positions. However, because a measure of location uncertainty is often fundamental in the interpretation of relocation results, we developed and incorporated a non-parametric procedure within the GrowClust program that can be used to estimate location uncertainties for all relocated events. The method implements a modified bootstrap approach based on statistical resampling theory (Efron and Tibshirani, 1994).

Upon initial input to GrowClust, cross-correlation data (differential times, cross-correlation values, and associated station metadata) are organized into arrays of length  $N_{pb}$ , in which  $N_{pb}$  denotes the total number of combined  $P$ - and  $S$ -phase observations. For each bootstrap iteration, the algorithm procedure generates resampling vectors (Efron and Tibshirani, 1994) to perform efficient, random resampling of these input

arrays. The bootstrap-resampled cross-correlation data are then input to the main GrowClust algorithm, resulting in a perturbed set of event locations specific to that bootstrap iteration. Repeating the resampling procedure  $B$  times, a bootstrap distribution of hypocentral positions (longitude, latitude, depth, and origin time) is constructed for each relocated event.

Though a complete analysis of the full sampling distribution using this method may require the number of bootstrap resamples  $B$  to be of order 1000 or greater, estimates of standard errors in hypocentral parameters typically stabilize much faster ( $B \sim 50\text{--}100$ ). The nonparametric error estimates output by GrowClust are obtained from the median absolute deviations

$$\text{MAD}(X) = \text{median}(|X_i - \text{median}(X)|) \quad (3)$$

(Leys *et al.*, 2013), of the bootstrap distribution of hypocentral coordinates. This provides a more robust characterization of location uncertainty than the raw bootstrap standard errors, which may be biased if the underlying bootstrap distribution is skewed (Hesterberg *et al.*, 2003; Mammen, 2012).

## RESULTS: RELOCATION OF NEVADA SEISMICITY

Moderate-to-large, damaging earthquakes ( $M \geq 5$ ) occur more frequently in Nevada than in any state within the continental United States except California. Three  $M \geq 7$  events shook the western half of state during the twentieth century, and from the historical record  $M \geq 6$  events occurred on average every six years (VanWormer and Ryall, 1980; Smith *et al.*, 2008) in the Nevada–California border region. The strike-slip faults of the Walker Lane belt in western Nevada and eastern California are capable of hosting  $M \geq 7$  earthquakes (e.g., the 1932  $M \geq 7.1$  Cedar Mountain earthquake), and the range-bounding normal faults of the eastern Sierra frontal fault system have a history of large, mid- $M \geq 7$  events (Ramelli *et al.*, 1999; Dingler *et al.*, 2009; Wesnousky *et al.*, 2012). Though these fault systems pose a significant hazard to the population centers of Reno and Carson City, as well as rural communities throughout the region, a comprehensive study of earthquake occurrence within the Walker Lane and the state as a whole has yet to be undertaken.

To demonstrate the use and efficacy of the GrowClust method, we apply it to two recent earthquake sequences in western Nevada: the 2012–2015 Spanish Springs and 2014–present Sheldon swarms (for a regional map and station locations, see © Fig. S1, available in the electronic supplement to this article). Both sequences are prominent within Nevada’s contemporary seismic record and are worthy of scientific investigation in their own right. The Spanish Springs sequence is spatially compact, with more than 1600 events occurring over a length scale of several kilometers. The sequence was well recorded by near-source stations operated by the Nevada Seismological Laboratory (NSL), and its  $M_w$  4.2 mainshock was widely felt in the Reno area. In contrast, the Sheldon sequence occurred within the remote northwestern corner of the state,

where station coverage is sparse, and hence the initial catalog locations are poorly constrained. The Sheldon sequence was larger in scale than Spanish Springs and is of particular scientific interest due to its persistent, swarm-like seismicity, replete with several discrete clusters containing multiple  $M \geq 4$  events. In the following section, we focus primarily on the interpretation of the GrowClust-relocated event positions, while deferring a more complete assessment of the tectonics and source mechanisms of the two sequences to future studies.

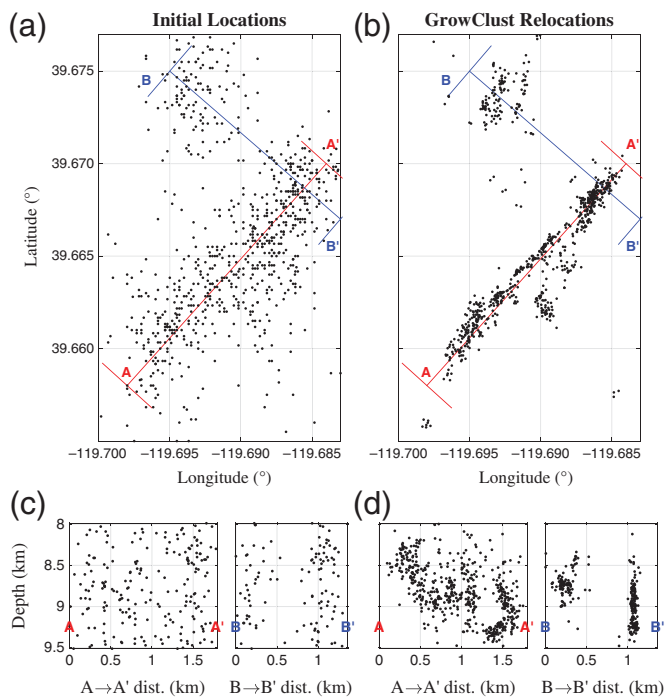
## Waveform Data and Cross Correlation

For our analysis of the Spanish Springs and Sheldon sequences, we use waveform data archived by the NSL. The NSL database is organized using an Antelope/DataScope software system (see Data and Resources), and earthquakes are routinely located by the NSL using the Antelope dblocsat2 and U.S. Geological Survey (USGS) HYPOINVERSE algorithms (Klein, 2002), assuming a reference velocity model listed in © Table S1. These event locations are then forwarded to the USGS/Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat; see Data and Resources), and we take these locations to be the initial event positions prior to performing the GrowClust relocations. We use the Antelope relational database to extract the waveforms from the events associated with each sequence.

Prior to performing waveform cross correlation, we proceed as described in the Methods and Algorithm Description section, filtering all traces from 1 to 10 Hz using a band-pass filter with a gentle roll-off that retains some energy up to 15 Hz. For event pairs in the two sequences, we compute cross-correlation functions separately for  $P$  and  $S$  phases on all available channels of all common stations. We use cross-correlation windows of  $-1.0$  to  $1.5$  s and  $-1.0$  to  $2.5$  s for  $P$  and  $S$  phases, relative to the predicted arrival times of the respective phases. The use of predicted arrival times in lieu of operator picks can significantly expand the cross-correlation data set. Differential travel times and cross-correlation coefficients are derived from the raw cross-correlation functions using a spline interpolation technique that provides millisecond sampling precision. We use the vertical-component channel for  $P$  phases (if available), and we select the horizontal channel with the highest cross-correlation coefficient for  $S$  phases (and  $P$  phases with no available vertical channel). We ensure the data quality of our cross-correlation database by further considering only those event pairs with an average cross-correlation coefficient of 0.45 across all phases and a minimum of eight phases with cross-correlation coefficients greater than 0.6.

## The 2012–2015 Spanish Springs Sequence

The 2012–2015 Spanish Springs sequence was quite active but spatially compact, with thousands of events occurring over a length scale of several kilometers. The sequence is named for its close spatial proximity to the north Reno suburb of Spanish Springs, and its largest event ( $M_w$  4.2) was widely felt throughout the city and resulted in nonstructural damage to select local buildings. The sequence was well recorded by the dense



▲ **Figure 3.** Comparison of initial Advanced National Seismic System (ANSS) catalog locations and GrowClust relocations for the 2012–2015 Spanish Springs sequence. (a, catalog) and (b, GrowClust) provide a map view comparison of the initial and relocated event positions. (c, catalog) and (d, GrowClust) provide a comparison of fault-parallel ( $A \rightarrow A'$ , red) and fault-perpendicular cross sections ( $B \rightarrow B'$ , blue).

distribution of NSN stations in the Reno area, where the local magnitude of completion is  $\sim 0.0$  (Kent *et al.*, 2015), with many recorded events of even smaller magnitude.

We use the NSL/ANSS event locations as the initial catalog positions for our GrowClust relocation analysis of this sequence. We take advantage of the dense station coverage and only consider stations within 80 km, resulting in 51 unique short-period, broadband, and strong-motion stations. We do not implement a magnitude cutoff for this sequence and hence do not expect to provide relocated positions for all events (the majority of which are  $M$  0.5 or less and were initially located by the NSL using as few as three stations). Even so, we are able to extract adequate waveform similarity to relocate 793 of the 1616 recorded events in the sequence, including all events  $M$  1.0 and greater.

GrowClust relocation results for these 793 events are presented in Figure 3, with the initial catalog positions shown for reference. Despite the dense station coverage, the initial catalog positions are highly scattered both in map view and in cross section (Fig. 3a,c), with little hint of local faulting structure. In contrast, the relocated seismicity is noticeably sharper, and the relocated event positions clearly outline several distinct faulting structures (Fig. 3b,d). The primary fault structure strikes to the northeast and is nearly vertical (cross section  $A \rightarrow A'$ ), whereas a secondary strand with a more northerly strike

branches off of the primary structure's northeast end. Fault-perpendicular cross sections ( $B \rightarrow B'$ ) provide evidence that seismicity within a secondary cluster to the northwest of the primary structure occurs on a distinct, subparallel fault strand.

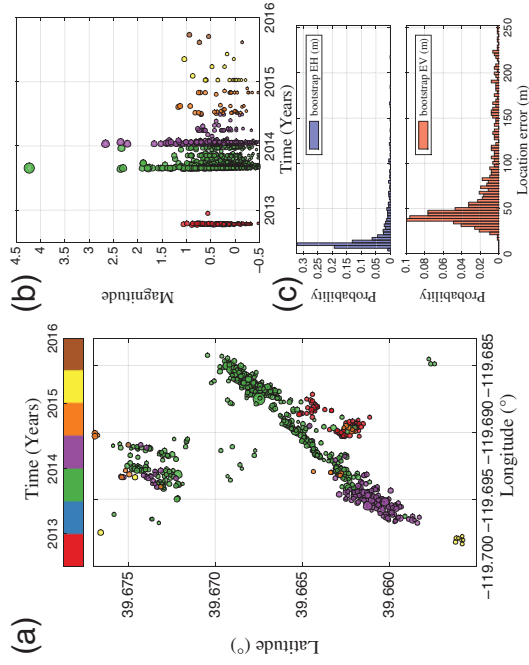
The GrowClust relocations likewise aid in the interpretation of the spatiotemporal migration of seismicity during the Spanish Springs sequence (Fig. 4). The sequence began as a short burst of  $M$  1.0 and lower seismicity in late 2012, with all events occurring on the southeasterly branch of the main fault structure. The  $M_w$  4.2 mainshock occurred on 27 August 2013 and was preceded by a vigorous foreshock sequence that migrated linearly from the southwest toward the mainshock hypocenter along the main fault strand. The subsequent aftershock sequence was extended in duration, with several distinct swarms occurring on spatially isolated sections of the different structures. Intriguingly, the relocation results reveal a hole (i.e., area free of seismicity) on the mainshock fault plane, adjacent to its hypocenter. This hole may outline the region of largest slip or near-complete stress release within the mainshock rupture zone, with aftershocks localized at the stress concentrations around its perimeter.

We further use GrowClust's nonparametric error estimation procedure to examine the lateral and vertical location uncertainties of relocated events within the sequence. Overall, the structural features of the sequence are quite well resolved, with median lateral and vertical location errors of 11 and 62 m, respectively (Fig. 4c and Fig. S2). These low nominal uncertainties, a consequence of the dense seismicity, completeness of the catalog, and good azimuthal coverage of near-source stations lend confidence to the interpretation of the salient structural and spatiotemporal features of the Spanish Springs sequence.

### The 2014–Present Sheldon Sequence

The Sheldon earthquake sequence began in July 2014, with swarms of seismicity occurring within and near the Sheldon Wildlife Refuge in the northwestern corner of Nevada. In contrast to the well-recorded Spanish Springs sequence, station coverage in this region is quite sparse, particularly before November 2014, when the near-field station COLR was installed at  $\sim 15$  km distance. The sequence is of considerable scientific interest due to the large moment release (28  $M \geq 4.0$  and 263  $M \geq 3.0$  events to date, with the largest being  $M_w$  4.8) and its persistent, swarm-like seismicity that defies conventional earthquake-triggering models. Though distant from major population centers, the larger events have been felt strongly by local residents and farming communities near Vya, Nevada.

As with Spanish Springs, we use the NSL/ANSS event locations as the initial catalog positions for our GrowClust relocation analysis. Because of the inherent limitation in near-field station coverage, we include in our analysis all nine recording stations within 250 km. To ensure adequate waveform signal-to-noise ratio, we implement the previously described quality-control criteria for our cross-correlation data (see the [Methods and Algorithm Description](#) section), and restrict our analysis to events of local magnitude 1.8 or greater occurring on or after 18 November 2014, when the COLR station was

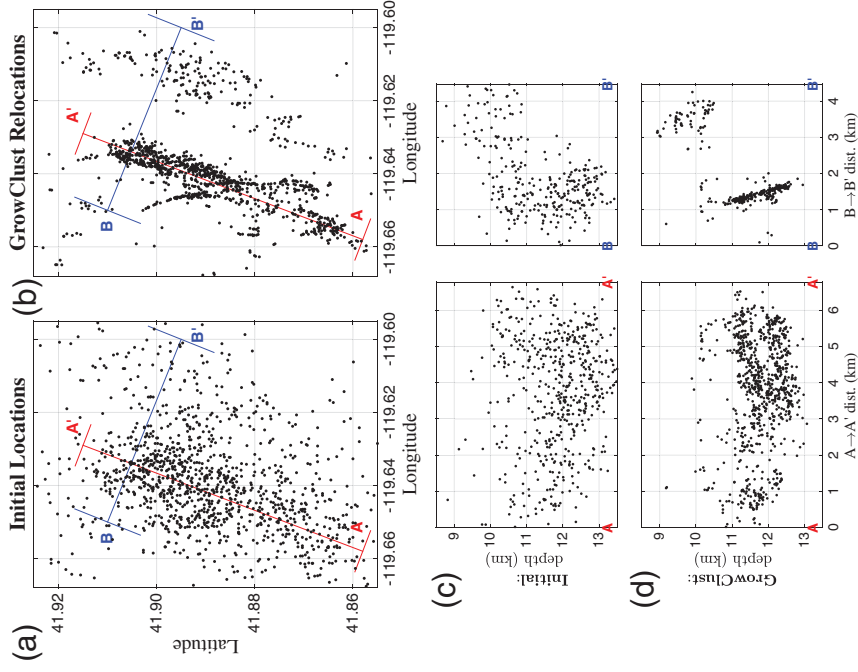


▲ **Figure 4.** Space–time evolution of the 2012–2015 Spanish Springs sequence. (a) Map view of the Spanish Springs sequence, with events color-coded by occurrence time (colors are discretized in time bins of 0.5 years from 2012.5 through 2016). (b) Magnitude as a function of time for these events (colors are consistent with panel a). (c) Histograms of horizontal (top, blue) and vertical (bottom, light red) location uncertainties for these events (obtained via 100 bootstrap resamplings of the input cross-correlation data, see the [Relative Location Uncertainties](#) section).

installed. This cutoff is slightly more conservative than the local magnitude of completion, which is  $M \sim 1.5$ . Using these criteria, our input event list consists of 1369 total events (through 28 August 2016), 1232 of which we are able to relocate.

The sparseness of station coverage in the Sheldon region is reflected in the highly scattered initial catalog positions (Fig. 5). Despite these limitations, the GrowClust relocations provide a noticeable improvement, resolving distinct fault structures from the cloud of initial positions. Seismicity is concentrated within two subparallel clusters that strike north-northeast. The western cluster is more active, containing the dominant portion of events within the sequence, and is composed of two distinct fault strands (striking north-northeast and north-northwest). Though the location uncertainties are larger for the Sheldon sequence (median horizontal and vertical errors of 117 and 91 m; Fig. 6c), they are not so large as to inhibit the resolution of *in situ* structure. This is particularly true for the western cluster, which, with its denser seismicity and hence more available similar-event pairs, is better resolved than the eastern cluster (⊕ Fig. S3). Analyses of fault-parallel and fault-perpendicular cross sections reveal that both clusters dip steeply to the east, with deeper seismicity in the western cluster (10–13 km) than in the eastern cluster (9–11 km). This structural interpretation is consistent with moment tensor analyses of the larger events, which predominantly show normal faulting on steeply dipping normal faults striking north-northeast (Ruhl *et al.*, 2016).

The improved resolution provided by the GrowClust relocations allows us to examine the spatiotemporal evolution of the



▲ **Figure 5.** Comparison of initial ANSS catalog locations and GrowClust relocations for the 2014–present Sheldon sequence. (a, catalog) and (b, GrowClust) A map view comparison of the initial and relocated event positions. (c, catalog) and (d, GrowClust) A comparison of fault-parallel ( $A \rightarrow A'$ , red) and fault-perpendicular cross sections ( $B \rightarrow B'$ , blue).

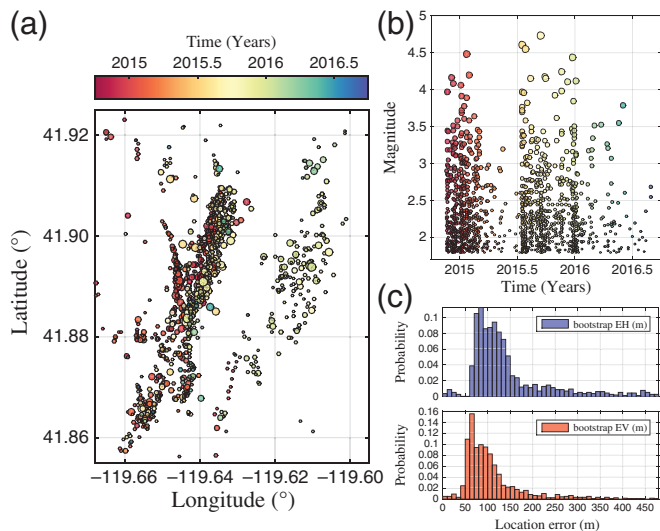
Sheldon sequence in more detail. The sequence is overall characterized by relatively continuous seismicity, but exhibits distinctive patterns of spatial migration. Early in 2015, seismicity migrates from north to south within the northwesterly branch of the major (western) cluster (Fig. 6). After a brief quiescent period, the sequence is reactivated during the second half of the year, marked by the occurrence of three events in excess of magnitude 4.5 and dozens more in excess of magnitude 3.5. The seismicity is swarm-like, with no clear mainshock, and exhibits a tendency to migrate from the southwest to the northeast, again predominantly within the western cluster. It is not until late 2015 and early 2016 that the minor (eastern) cluster becomes active, though seismicity continues unabated within the northern portion of the major (western) cluster during this time.

## DISCUSSION

### Comparison of GrowClust and HypoDD Relocation Results

It is instructive to compare the relocation results obtained using GrowClust for the Spanish Springs and Sheldon sequences to those obtained using the HypoDD double-difference tech-





▲ **Figure 6.** Space–time evolution of the 2014–present Sheldon sequence. (a) Map view of the Sheldon sequence, with events color-coded by occurrence time (colors scale varies continuously from 2014.75 to 2016.75). (b) Magnitude as a function of time for these events (colors consistent with panel a). (c) Histograms of horizontal (top, blue) and vertical (bottom, light red) location uncertainties for these events (obtained via 100 bootstrap resamplings of the input cross-correlation data, see the [Relative Location Uncertainties](#) section).

nique (Waldhauser 2000, see [Data and Resources](#)). Doing so not only provides a benchmark of sorts for GrowClust (as HypoDD is well established and justifiably popular) but also helps elucidate the more relevant differences in methodology. For the comparison, we used identical data inputs (initial hypocentral locations, velocity model, and cross-correlation data) as were used for the GrowClust results presented above. We performed sensitivity analysis on the iterative weighting scheme, algorithm damping, and other control parameters found in the HypoDD input file to obtain the best relocation results possible, given our limited experience in the technique’s use. We note that both sequences are too large in scale and too poorly conditioned to obtain singular value decomposition solutions to the DD equations, so the HypoDD results presented here are those obtained using the damped least-squares conjugate gradient (LSQR) solution (Waldhauser, 2000), and we have taken care to ensure reasonable condition numbers for each algorithm iteration. As discussed in Waldhauser (2000), the LSQR solution does not produce accurate location uncertainty estimates, precluding a formal comparison with those obtained using GrowClust. Both the GrowClust and HypoDD methods relocate a comparable fraction of the total number of events in each sequence (793 vs. 764 of the 1616 total Spanish Springs events, 1232 vs. 1282 of the 1369 total Sheldon events, for GrowClust and HypoDD, respectively). The runtime on a standard laptop computer is slightly faster for GrowClust (14.9 s and 34.2 s for the two sequences) than for HypoDD (65.0 s and 95.3 s).

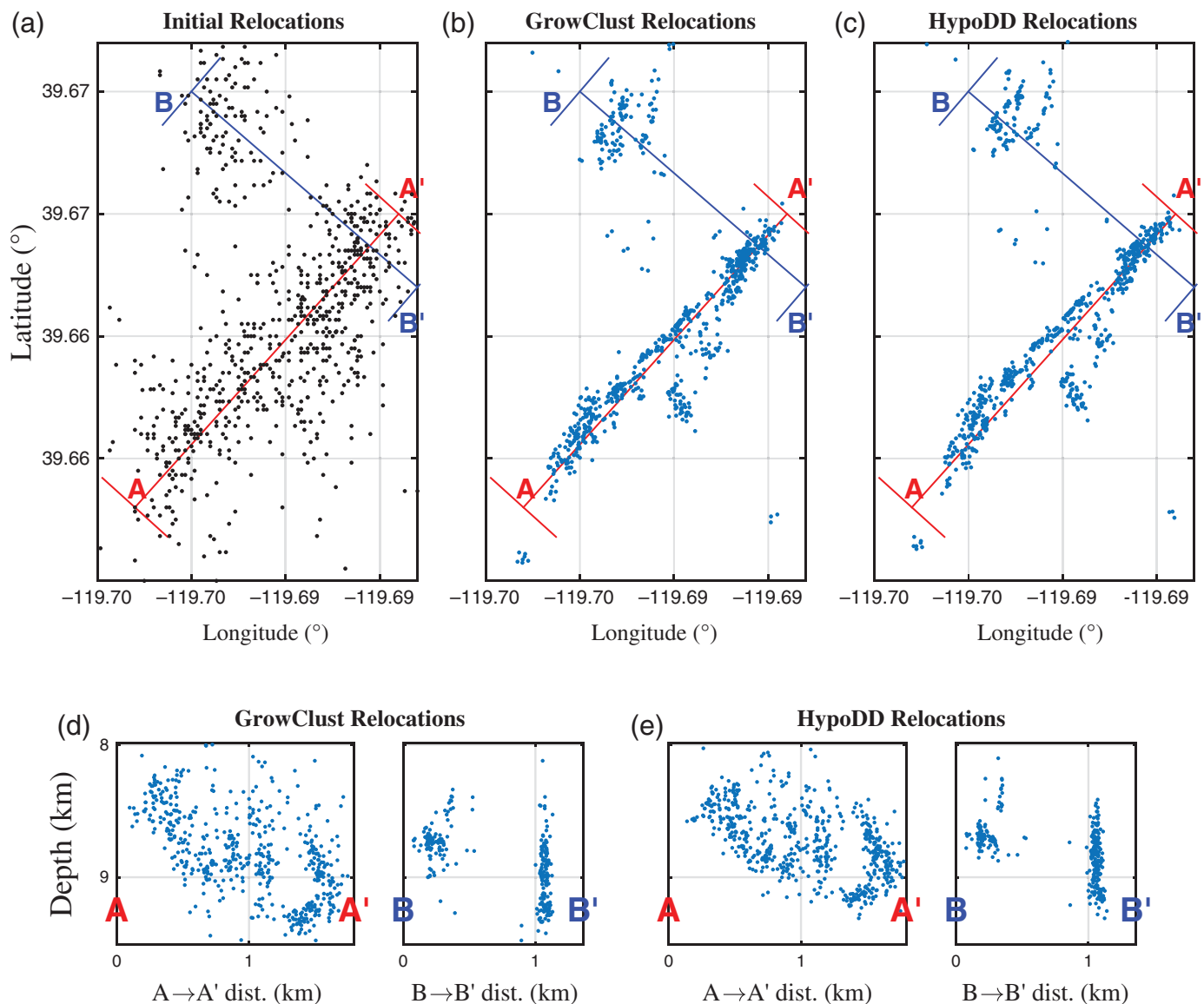
Overall, the GrowClust and HypoDD results are in good visual agreement. This is particularly true for the well-recorded

Spanish Springs sequence (Fig. 7), with both methods providing good resolution of vertically dipping fault structures and partitioning events into a single dominant cluster. The similarity in results for Spanish Springs is not surprising, given that the dense set of differential times ensures strong linkages between events (and hence, a well-conditioned system for the DD algorithm). For this reason, it is also unsurprising that there is greater disparity in the results for the Sheldon sequence (Fig. 8) in which event pairs are not as robustly linked due to the sparse station coverage. In this case, the differences in the underlying computational framework—hierarchical clustering versus matrix inversion—become more apparent. Of particular interest is the location of the smaller, eastern cluster of seismicity, which differs by more than a kilometer from method to method. For the GrowClust relocations, the centroid of this smaller cluster is fixed to its initial reference position, whereas no such constraint is imposed in HypoDD, permitting a centroid shift of greater than 1 km. The centroid of the larger, western cluster also undergoes a shift of 133 m to the northeast in the HypoDD relocations, while remaining fixed in the GrowClust relocations. GrowClust’s clustering algorithm considers the western and eastern clusters as independent units (i.e., it does not find a sufficient number of linking event pairs to join the clusters), whereas HypoDD relocates both clusters as one coherent unit, finding one or more linking chains of event pairs between the two clusters.

### Implications of Relocation Results for the Understanding of Seismotectonics in the Nevada–California Border Region

Both conventional mainshock–aftershock sequences and extended swarm-like sequences are common within the state of Nevada (Ichinose *et al.*, 1998; Smith *et al.*, 2008; Ruhl, 2015), and relocation results such as those presented here help to study and differentiate the active physical mechanisms driving these sequences. The Spanish Springs and Sheldon sequences provide important test cases in this regard and demonstrate the subtlety of the task at hand. Spanish Springs may well fall into the former classification, because its moment release is dominated by the  $M_w$  4.2 mainshock. However, complexity is pervasive even within such a compact sequence, and standard triggering models cannot adequately account for the extended duration or the spatial progression of the foreshock and aftershock sequences. The foreshock sequence is of particular interest, because its systematic linear migration toward the mainshock hypocenter lends insight into the nucleation process. Likewise, Sheldon’s distinctive spatial migration pattern, combined with its unusually persistent and swarm-like seismicity, suggest that aseismic or fluid-driven processes may be the dominant physical mechanisms driving the sequence, rather than the coseismic stress changes from isolated, individual mainshocks (Hainzl, 2002; Lohman and McGuire, 2007; Shearer, 2012; Shelly *et al.*, 2016).

The GrowClust results presented here for Spanish Springs and Sheldon establish the potential for large-scale relocation efforts, analogous to those undertaken in recent years in



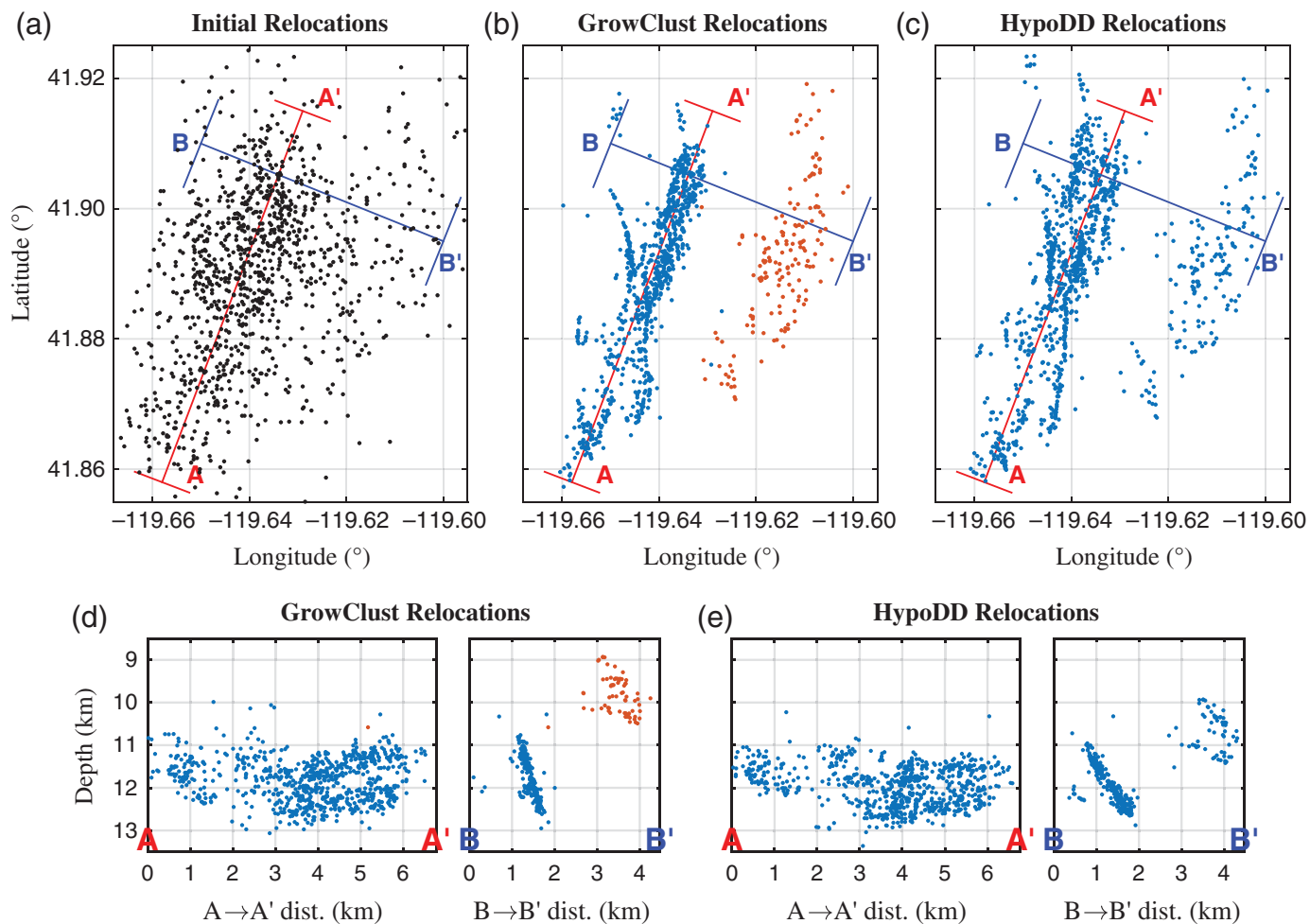
▲ **Figure 7.** Comparison of relocation results for the Spanish Springs sequence using the GrowClust and HypoDD methods. (b) and (c) Map view improvement of GrowClust and HypoDD relocations over (a) the initial catalog locations, whereas (d) and (e) compare GrowClust and HypoDD relocations in cross section. Only those events that are successfully relocated by both methods are shown, and events are color-coded by the cluster ID number of each method. For the Spanish Springs sequence, both GrowClust and HypoDD partition events into one dominant cluster.

California (Waldhauser and Schaff, 2008; Hauksson *et al.*, 2012), to elucidate subtle features of Nevada seismicity. Such efforts may prove to be particularly valuable in Nevada, where the western half of the state comprises a transition zone that accommodates  $\sim 20\%$  of the Pacific–North America plate boundary deformation and is characterized by a spatially complex distribution of intersecting normal, sinistral, and dextral faults (Ichinose *et al.*, 2003; Faults *et al.*, 2005; Wesnousky *et al.*, 2012). The relocated seismicity provides a means to image these structures in high resolution, supplying important observational constraints on the seismotectonic evolution of the Walker Lane and central Nevada seismic belts. The extensive historical record of large, damaging earthquakes in Nevada

(VanWormer and Ryall, 1980) makes apparent the importance of an improved and more rigorously quantitative understanding of seismic hazard. We plan in future studies to expand our use of GrowClust beyond individual sequences, and apply it on a much wider scale within the state and the Nevada–California border region.

## SUMMARY

GrowClust is a new and open-source earthquake relocation algorithm. It uses waveform cross-correlation input data—differential times and cross-correlation values—in a hybrid, hierarchical clustering algorithm that simultaneously groups





▲ **Figure 8.** Comparison of relocation results for the Sheldon sequence using the GrowClust and HypoDD methods. (b) and (c) The map view improvement of GrowClust and HypoDD relocations over (a) the initial catalog locations, whereas (d) and (e) compare GrowClust and HypoDD relocations in cross section. Only those events that are successfully relocated by both methods are shown, and events are color-coded by the cluster ID number of each method. For the Sheldon sequence, GrowClust partitions events into two major clusters, whereas HypoDD uses only one.

and relocates earthquakes in similar event clusters. The method is fast, flexible, and robust with respect data outliers. GrowClust also includes a built-in mechanism for performing uncertainty analysis that gives users a more complete assessment of the resolving power of the relocation results. We apply the GrowClust method to two prominent Nevada earthquake sequences: the 2012–2015 Spanish Springs and the 2014–present Sheldon swarms. The encouraging results for these examples demonstrate the scientific potential for large-scale relocation efforts within the region using the GrowClust algorithm.

## DATA AND RESOURCES

The waveform data used in this study are archived locally by the Nevada Seismological Laboratory (NSL) and are publicly available from the Incorporated Research Institutions for Seismology Data Management Center (IRIS-DMC; <http://ds.iris.edu/ds/nodes/dmc/>, last accessed September 2016). Initial catalog positions are consistent with the Advanced National

Seismic System (ANSS) Comprehensive Earthquake Catalog (<http://earthquake.usgs.gov/data/comcat/>, last accessed October 2016). We use Antelope/Datascope software (<http://www.brrt.com/>, last accessed September 2016) to extract waveforms for each event. We provide catalogs that list both initial and relocated positions for the Spanish Springs and Sheldon sequences analyzed in this article in the  electronic supplement to the article.

The GrowClust relocation codes described in this article comprise an open-source software package under the GNU General Public License v.3. The GrowClust source distribution, which includes source codes, a user guide, and an example data set, is publicly available for download at <http://igppweb.ucsd.edu/~dtrugman/> (last accessed January 2017). We use in this article the HypoDD implementation of the double-difference (DD) technique (<http://www.ldeo.columbia.edu/~felixw/hypoDD.html>, last accessed October 2016) for comparison to GrowClust relocations. 

## ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program (NSFGRFP) under Grant Number DGE-1144086. Additional support was provided by the U.S. Geological Survey–National Earthquake Hazards Reduction Program (USGS-NEHRP) Grant Number G15AS00037-2016-0072. We thank K. Smith, R. Hatch, and C. Ruhl for their guidance in the application of GrowClust to Nevada seismicity, and R. Matoza, R. Abercrombie, and A. Borsa for stimulating discussions that helped improve the algorithm design. We also thank Editor Z. Peng and two anonymous reviewers for their useful suggestions to improve this article.

## REFERENCES

- Astiz, L., and P. M. Shearer (2000). Earthquake locations in the inner continental borderland, offshore southern California, *Bull. Seismol. Soc. Am.* **90**, no. 2, 425–449, doi: [10.1785/0119990022](https://doi.org/10.1785/0119990022).
- Buland, R. (1976). The mechanics of locating earthquakes, *Bull. Seismol. Soc. Am.* **66**, no. 1, 173–187.
- Dingler, J., G. Kent, N. Driscoll, J. Babcock, A. Harding, G. Seitz, B. Karlin, and C. Goldman (2009). A high-resolution seismic CHIRP investigation of active normal faulting across Lake Tahoe basin, California–Nevada, *Geol. Soc. Am. Bull.* **121**, nos. 7/8, 1089–1107, doi: [10.1130/B26244.1](https://doi.org/10.1130/B26244.1).
- Douglas, A. (1967). Joint epicentre determination, *Nature* **215**, no. 5096, 47–48, doi: [10.1038/215047a0](https://doi.org/10.1038/215047a0).
- Efron, B., and R. J. Tibshirani (1994). *An Introduction to the Bootstrap*, CRC Press, Boca Raton, Florida.
- Faulds, J. E., C. D. Henry, and N. H. Hinz (2005). Kinematics of the northern Walker Lane: An incipient transform fault along the Pacific–North American plate boundary, *Geology* **33**, no. 6, 505, doi: [10.1130/G21274.1](https://doi.org/10.1130/G21274.1).
- Fremont, M.-J., and S. D. Malone (1987). High precision relative locations of earthquakes at Mount St. Helens, Washington, *J. Geophys. Res.* **92**, no. B10, 10,223–10,236, doi: [10.1029/JB092iB10p10223](https://doi.org/10.1029/JB092iB10p10223).
- Frohlich, C. (1979). An efficient method for joint hypocenter determination for large groups of earthquakes, *Comput. Geosci.* **5**, nos. 3/4, 387–389, doi: [10.1016/0098-3004\(79\)90034-7](https://doi.org/10.1016/0098-3004(79)90034-7).
- Geiger, L. (1910). Herdbestimmung bei Erdbeben aus den Ankunftszeiten, *K. Gessel. Wiss. Goett.* **4**, 331–349 (in German).
- Geiger, L. (1912). Probability method for the determination of earthquake epicenters from the arrival time only, *Bull. St. Louis Univ.* **8**, 60–71.
- Got, J.-L., J. Fréchet, and F. W. Klein (1994). Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea, *J. Geophys. Res.* **99**, no. B8, 15,375–15,386, doi: [10.1029/94JB00577](https://doi.org/10.1029/94JB00577).
- Hainzl, S. (2002). Indications for a successively triggered rupture growth underlying the 2000 earthquake swarm in Vogtland/NW Bohemia, *J. Geophys. Res.* **107**, no. B12, 2338, doi: [10.1029/2002JB001865](https://doi.org/10.1029/2002JB001865).
- Hauksson, E., W. Yang, and P. M. Shearer (2012). Waveform relocated earthquake catalog for southern California (1981 to June 2011), *Bull. Seismol. Soc. Am.* **102**, no. 5, 2239–2244, doi: [10.1785/0120120010](https://doi.org/10.1785/0120120010).
- Hesterberg, T., S. Monaghan, D. S. Moore, A. Clipson, and R. Epstein (2003). Bootstrap methods and permutation tests, in *The Practice of Business Statistics*, Moore, D. S., G. P. McCabe, S. L. Sclove, and W. M. Duckworth (Editors), W.H. Freeman, New York, New York, 169–183.
- Ichinose, G. A., J. G. Anderson, K. D. Smith, and Y. Zeng (2003). Source parameters of eastern California and western Nevada earthquakes from regional moment tensor inversion, *Bull. Seismol. Soc. Am.* **93**, no. 1, 61–84, doi: [10.1785/0120020063](https://doi.org/10.1785/0120020063).
- Ichinose, G. A., K. D. Smith, and J. G. Anderson (1998). Moment tensor solutions of the 1994 to 1996 Double Spring Flat, Nevada, earthquake sequence and implications for local tectonic models, *Bull. Seismol. Soc. Am.* **88**, no. 6, 1363–1378.
- Ito, A. (1985). High resolution relative hypocenters of similar earthquakes by cross-spectral analysis method, *J. Phys. Earth* **33**, no. 4, 279–294, doi: [10.4294/jpe1952.33.279](https://doi.org/10.4294/jpe1952.33.279).
- Jordan, T. H., and K. A. Sverdrup (1981). Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific, *Bull. Seismol. Soc. Am.* **71**, no. 4, 1105–1130.
- Kent, G., K. D. Smith, and G. Biasi (2015). Final report: Nevada Seismic Network, *Tech. Rept. G10AC00090*, University of Nevada, Reno, Nevada.
- Klein, F. W. (2002). User's guide to HYPOINVERSE-2000, a Fortran program to solve for earthquake locations and magnitudes, *Report 2002-171*, US Dept. of the Interior, Geological Survey, Reston, Virginia.
- Leys, C., C. Ley, O. Klein, P. Bernard, and L. Licata (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, *J. Exp. Soc. Psychol.* **49**, no. 4, 764–766, doi: [10.1016/j.jesp.2013.03.013](https://doi.org/10.1016/j.jesp.2013.03.013).
- Lin, G., P. M. Shearer, and E. Hauksson (2007). Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, *J. Geophys. Res.* **112**, no. B12, doi: [10.1029/2007JB004986](https://doi.org/10.1029/2007JB004986).
- Lohman, R. B., and J. J. McGuire (2007). Earthquake swarms driven by aseismic creep in the Salton trough, California, *J. Geophys. Res.* **112**, no. B4, doi: [10.1029/2006JB004596](https://doi.org/10.1029/2006JB004596).
- Mammen, E. (2012). *When Does Bootstrap Work?: Asymptotic Results and Simulations*, Springer Science & Business Media, New York, New York.
- Matoza, R. S., P. M. Shearer, G. Lin, C. J. Wolfe, and P. G. Okubo (2013). Systematic relocation of seismicity on Hawaii Island from 1992 to 2009 using waveform cross correlation and cluster analysis, *J. Geophys. Res.* **118**, no. 5, 2275–2288, doi: [10.1002/jgrb.50189](https://doi.org/10.1002/jgrb.50189).
- Nadeau, R. M., W. Foxall, and T. V. McEvilly (1995). Clustering and periodic recurrence of microearthquakes on the San Andreas fault at Parkfield, California, *Science* **267**, no. 5197, 503–507, doi: [10.1126/science.267.5197.503](https://doi.org/10.1126/science.267.5197.503).
- Phillips, W. S., L. S. House, and M. C. Fehler (1997). Detailed joint structure in a geothermal reservoir from studies of induced microearthquake clusters, *J. Geophys. Res.* **102**, no. B6, 11,745–11,763, doi: [10.1029/97JB00762](https://doi.org/10.1029/97JB00762).
- Poupinet, G., W. L. Ellsworth, and J. Frechet (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras fault, California, *J. Geophys. Res.* **89**, no. B7, 5719–5731, doi: [10.1029/JB089iB07p05719](https://doi.org/10.1029/JB089iB07p05719).
- Ramelli, A. R., J. W. Bell, C. M. dePolo, and J. C. Yount (1999). Large-magnitude, late Holocene earthquakes on the Genoa fault, west-central Nevada and eastern California, *Bull. Seismol. Soc. Am.* **89**, no. 6, 1458–1472.
- Richards-Dinger, K. B., and P. M. Shearer (2000). Earthquake locations in southern California obtained using source-specific station terms, *J. Geophys. Res.* **105**, no. B5, 10,939–10,960, doi: [10.1029/2000JB900014](https://doi.org/10.1029/2000JB900014).
- Rowe, C. A., R. C. Aster, W. S. Phillips, R. H. Jones, B. Borchers, and M. C. Fehler (2002). Using automated, high-precision repicking to improve delineation of microseismic structures at the Soultz Geothermal Reservoir, *Pure Appl. Geophys.* **159**, nos. 1/3, 563–596, doi: [10.1007/PL00001265](https://doi.org/10.1007/PL00001265).
- Rubin, A. M., D. Gillard, and J.-L. Got (1999). Streaks of microearthquakes along creeping faults, *Nature* **400**, no. 6745, 635–641, doi: [10.1038/23196](https://doi.org/10.1038/23196).
- Ruhl, C. (2015). Inside an earthquake swarm: Objective identification and analysis of spatiotemporal subclusters of the Mogul 2008 earthquake swarm in Reno, NV, *AGU Fall Meeting Abstracts*, San Francisco, California.

- Ruhl, C. J., K. D. Smith, and R. E. Abercrombie (2016). Moderate-to-small magnitude seismicity clusters persist in northwest Nevada for over 19 months, *SSA Annual Meeting Abstracts*, Reno, Nevada.
- Shearer, P. (2005). Southern California hypocenter relocation with waveform cross-correlation, part 2: Results using source-specific station terms and cluster analysis, *Bull. Seismol. Soc. Am.* **95**, no. 3, 904–915, doi: [10.1785/0120040168](https://doi.org/10.1785/0120040168).
- Shearer, P. M. (1997). Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence, *J. Geophys. Res.* **102**, no. B4, 8269–8283, doi: [10.1029/96JB03228](https://doi.org/10.1029/96JB03228).
- Shearer, P. M. (1998). Evidence from a cluster of small earthquakes for a fault at 18 km depth beneath Oak Ridge, southern California, *Bull. Seismol. Soc. Am.* **88**, no. 6, 1327–1336.
- Shearer, P. M. (2002). Parallel fault strands at 9-km depth resolved on the Imperial fault, southern California, *Geophys. Res. Lett.* **29**, no. 14, 19-1–19-4, doi: [10.1029/2002GL015302](https://doi.org/10.1029/2002GL015302).
- Shearer, P. M. (2012). Space-time clustering of seismicity in California and the distance dependence of earthquake triggering, *J. Geophys. Res.* **117**, no. B10, doi: [10.1029/2012JB009471](https://doi.org/10.1029/2012JB009471).
- Shelly, D. R., W. L. Ellsworth, and D. P. Hill (2016). Fluid-faulting evolution in high definition: Connecting fault structure and frequency–magnitude variations during the 2014 Long Valley Caldera, California, earthquake swarm: Fluid-faulting evolution in Long Valley, *J. Geophys. Res.* **121**, no. 3, 1776–1795, doi: [10.1002/2015JB012719](https://doi.org/10.1002/2015JB012719).
- Smith, K., D. H. von Seggern, D. Depolo, J. G. Anderson, G. P. Biasi, and R. Anooshehpour (2008). Seismicity of the 2008 Mogul-Somerset West Reno, Nevada earthquake sequence, *AGU Fall Meeting Abstracts (2008)*, vol. 53.
- Thurber, C., and D. Eberhart-Phillips (1999). Local earthquake tomography with flexible gridding, *Comput. Geosci.* **25**, no. 7, 809–818, doi: [10.1016/S0098-3004\(99\)00007-2](https://doi.org/10.1016/S0098-3004(99)00007-2).
- Thurber, C. H. (1983). Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, *J. Geophys. Res.* **88**, no. B10, 8226–8236, doi: [10.1029/JB088iB10p08226](https://doi.org/10.1029/JB088iB10p08226).
- Thurber, C. H. (1985). Nonlinear earthquake location: Theory and examples, *Bull. Seismol. Soc. Am.* **75**, no. 3, 779–790.
- Thurber, C. H. (1992). Hypocenter-velocity structure coupling in local earthquake tomography, *Phys. Earth Planet. In.* **75**, no. 1, 55–62, doi: [10.1016/0031-9201\(92\)90117-E](https://doi.org/10.1016/0031-9201(92)90117-E).
- VanWormer, J. D., and A. S. Ryall (1980). Sierra Nevada-Great Basin boundary zone: Earthquake hazard related to structure, active tectonic processes, and anomalous patterns of earthquake occurrence, *Bull. Seismol. Soc. Am.* **70**, no. 5, 1557–1572.
- Waldhauser, F. (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California, *Bull. Seismol. Soc. Am.* **90**, no. 6, 1353–1368, doi: [10.1785/0120000006](https://doi.org/10.1785/0120000006).
- 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**, no. B8, doi: [10.1029/2007JB005479](https://doi.org/10.1029/2007JB005479).
- Waldhauser, F., W. L. Ellsworth, and A. Cole (1999). Slip-parallel seismic lineations on the northern Hayward fault, California, *Geophys. Res. Lett.* **26**, no. 23, 3525–3528, doi: [10.1029/1999GL010462](https://doi.org/10.1029/1999GL010462).
- Wesnousky, S. G., J. M. Bormann, C. Kreemer, W. C. Hammond, and J. N. Brune (2012). Neotectonics, geodesy, and seismic hazard in the northern Walker Lane of western North America: Thirty kilometers of crustal shear and no strike-slip? *Earth Planet. Sci. Lett.* **329/330**, 133–140, doi: [10.1016/j.epsl.2012.02.018](https://doi.org/10.1016/j.epsl.2012.02.018).

Daniel T. Trugman

Peter M. Shearer

Cecil H. and Ida M. Green Institute of Geophysics and

Planetary Physics

Scripps Institution of Oceanography

University of California, San Diego

9500 Gilman Drive

La Jolla, California 92093-0225 U.S.A.

dtrugman@ucsd.edu

Published Online 1 February 2017