High-precision, absolute earthquake location using source-specific station terms and inter-event waveform similarity

Anthony Lomax[1,1] and Alexandros Savvaidis[2,2]

[1]ALomax Scientific
[2]University of Texas at Austin

November 30, 2022

Abstract

Earthquake monitoring and many seismological studies depend on absolute earthquake locations from phase arrival-times. We present an absolute earthquake location procedure (NLL-SSST-coherence) which approaches the precision of waveform-based, relative location and is applicable with few seismic stations. NLL-SSST-coherence is based on the probabilistic, global-search NonLinLoc (NLL) location algorithm which defines a probability density function (PDF) in 3D space for absolute hypocenter location and is highly robust to outlier data. NLL-SSST-coherence location first reduces velocity model error through iteratively generated, smooth, source-specific, station travel-time corrections (SSST). Next, arrival-time error is reduced by consolidating location information across events based on inter-event waveform coherency. If the waveforms at a station for multiple events are very similar (have high coherency) up to a given frequency, then the distance separating these "multiplet" events is small relative to the seismic wavelength at that frequency. NLL-coherence relocation for a target event is a stack over 3D space of the NLL-SSST location PDF for the event and the PDF's for other multiplet events, each weighted by its waveform coherency with the target. NLL-coherence relocation requires waveforms from only one or a few seismic stations, enabling precise, absolute relocation with sparse networks, for foreshocks and early aftershocks of significant events before installation of temporary stations, and for older data sets with few waveform data. We show the behavior and performance of NLL-SSST-coherence with synthetic and ground-truth tests, and through application and comparison to relative locations for California earthquake sequences with dense and sparse station coverage.

Journal of Geophysical Research Solid Earth

Supporting Information for

High-precision, earthquake location using source-specific station terms and inter-event waveform similarity

Anthony Lomax[1] and Alexandros Savvaidis[2]

[1]ALomax Scientific, Mouans-Sartoux, France.

[2]Bureau of Economic Geology, The University of Texas at Austin, Austin, TX.

Contents of this file

Additional Supporting Information (Files uploaded separately)

Captions for Movie S1

Introduction

The supporting information for this article includes a 3D, fly-around animation of the NLL-SSST-coherence relocations of Lone Pine seismicity.

Movie S1. (ms01_OwensLake2020_Alomax_NLL-SSST-coherence_movie_20210511.mp4) 2020 Mw 5.8 Lone Pine, California earthquake sequence relocations. Fly-around animation of 2020-01-12 to 2021-02-15 hypocenters from the NLL-SSST-coherence relocations. Hypocenter color shows origin time, symbol size is proportional to magnitude. Hypocenter color shows origin time, symbol size is proportional to magnitude. NLL-SSST-coherence hypocenters are randomly shifted 0.05 km to avoid overlapping symbols. White triangle to the west shows the only nearby seismic station (CI.CWC) available for relocation. Green lines show faults from the USGS Quaternary fault and fold database for the United States. Background topography image from OpenTopgraphy.org.

High-precision earthquake location using source-specific station terms and inter-event waveform similarity

Anthony Lomax[1] and Alexandros Savvaidis[2]

[1]ALomax Scientific, Mouans-Sartoux, France.

[2]Bureau of Economic Geology, The University of Texas at Austin, Austin, TX.

Corresponding author: Anthony Lomax (anthony@alomax.net),

Key Points:

- We use source-specific station terms and waveform similarity to achieve high-precision earthquake location (NLL-SSST-coherence).
- NLL-SSST-coherence approaches the precision of relative location methods and can give better depth constraint when station coverage is poor.
- NLL-SSST-coherence requires waveforms from only one or a few stations and thus is applicable with sparse networks and older sequences.

Abstract

Earthquake monitoring and many seismological studies depend on earthquake locations from phase arrival-times. We present an extended, arrival-time earthquake location procedure (NLL-SSST-coherence) which approaches the precision of differential-timing based, relative location methods and is applicable with few seismic stations. NLL-SSST-coherence is based on the probabilistic, global-search NonLinLoc (NLL) location algorithm which defines a probability density function (PDF) in 3D space for hypocenter location and is highly robust to outlier data. NLL-SSST-coherence location first reduces velocity model error through iteratively generated, smooth, source-specific, station travel-time corrections (SSST). Next, arrival-time error is reduced by consolidating location information across events based on inter-event waveform coherency. If the waveforms at a station for multiple events are very similar (have high coherency) up to a given frequency, then the distance separating these “multiplet” events is small relative to the seismic wavelength at that frequency. NLL-coherence relocation for a target event is a stack over 3D space of the NLL-SSST location PDF for the event and the PDF’s for other multiplet events, each weighted by its waveform coherency with the target. NLL-coherence relocation requires waveforms from only one or a few

seismic stations, enabling precise relocation with sparse networks, for foreshocks and early aftershocks of significant events before installation of temporary stations, and for older data sets with few waveform data. We show the behavior and performance of NLL-SSST-coherence with synthetic and ground-truth tests, and through application and comparison to relative locations for California earthquake sequences with dense and sparse station coverage.

Plan language summary

Earthquake monitoring, early-warning, public information and understanding depend on standard locations of earthquakes in geographical space. Specialized, relative location methods extend standard locations to determine more precisely the positions of nearby earthquakes with respect to each other. We present a standard earthquake location procedure (NLL-SSST-coherence) which approaches the precision of relative location methods while being more generally applicable and efficient. NLL-SSST-coherence uses the NonLinLoc (NLL) location algorithms which determine an earthquake location as a probability cloud in 3D space and work well with poor quality seismogram recordings. NLL-SSST-coherence location first reduces effects of limited knowledge of seismic wavespeeds in the Earth through spatial averaging of wavespeed errors (SSST). Next, it reduces effects of error in measuring the timing of earthquake energy arrival at seismic stations by consolidating location information between nearby earthquakes. Nearby events are identified by their seismogram waveforms which are very similar, wiggle for wiggle – they have high coherence. NLL-SSST-coherence relocation enables precise earthquake relocation with sparse networks, for foreshocks and early aftershocks of significant earthquakes, and for older earthquake sequences. We show the performance of NLL-SSST-coherence with simulated and real data tests, and through application to California earthquake sequences with dense and sparse station coverage.

1 Introduction

Earthquake locations are fundamental to earthquake, volcano, glacier and nuclear test monitoring, and much seismological research and understanding. These locations are obtained from arrival times of seismic phase energy and show where a seismic event occurred relative to tectonic, geographic and urban features, along with the time of the event (Li et al., 2020; Lomax et al., 2014; C. Thurber & Rabinowitz, 2000). Sets of earthquake locations form seismicity which defines faulting structures and areas of earthquake and volcanic hazard. Space-time patterns in seismicity determine the geometry and activity on individual faults, the stages of earthquake initiation and the causes of human induced seismicity.

Relative to the needs of modern seismological study, standard earthquake location (event by event location using absolute arrival-times) often has low accuracy and precision, where accuracy is closeness to a usually unknown ground-truth, and precision is relative location accuracy – the correctness of the relative positions of nearby hypocenters. For example, association of possibly induced seismicity with human activities require high accuracy in absolute epicenter and depth of seismicity (Lomax & Savvaidis, 2019), while earthquake and tsunami early-warning and rapid estimation of rupture and ground shaking hazard require accurate hypocentral depth determinations (Bernardi et al., 2015). Study of the complexity and fine scale structure of fault systems, and relating these to geologic structures, fracture systems, stress patterns and geo-fluids requires both high accuracy and high precision. Accurate determination of hypocentral depth is particularly difficult as it requires that seismic stations are well distributed above and around the seismicity, and even then the obtained depths are strongly dependent on the accuracy of the used seismic velocity model (Gomberg et al., 1990).

Means for improving the accuracy and precision of standard earthquake locations include having stations close to and above the source zone (Billings et al., 1994; Buehler & Shearer, 2016; Gomberg et al., 1990; Hardebeck & Husen, 2010; Pavlis, 1986), use of 3D and geology-based, seismic velocity models (Darold et al., 2014; Latorre et al., 2016; e.g. Ryaboy et al., 2001; Wagner et al., 2013), station travel-time corrections (Lin & Shearer, 2005; Lomax, 2008, 2020a; e.g. Myers, 2000; Nicholson et al., 2008; Nooshiri et al., 2017; Pavlis & Hokanson, 1985a; Richards-Dinger & Shearer, 2000), ground-truth calibration (Bondár & McLaughlin, 2009; Lomax & Savvaidis, 2019; Ritzwoller et al., 2003) and use of location algorithms robust to error in the velocity models or earthquake arrival-time data (Stauder & Ryall, 1967; Ishida & Kanamori, 1978; Shearer, 1997; Lomax, 2008; Lomax et al., 2014).

High-precision, multi-event, relative location methods (Fehler et al., 2000; Frémont & Malone, 1987; Got et al., 1994; Lin et al., 2007; Nakamura, 1978; Poupinet et al., 1982; Rowe et al., 2002; Shearer, 1997, 2005; Trugman & Shearer, 2017; Waldhauser & Ellsworth, 2000) require and build upon standard locations. Relative location methods use waveform similarity and precise, cross-correlation, differential timing between events at individual stations to determine fine-scale, inter-event spatial relations. These methods can image seismicity in remarkable detail, showing narrow streaks, highly localized fault planes and sets of faulting structures (Got et al., 1994; Michele et al., 2020; Rubin et al., 1999; Waldhauser et al., 2004). However, these procedures depend on good station and ray coverage, and a model with accurate velocities and gradients of velocity (Gibbons et al., 2017; Matoza et al., 2013; Michelini & Lomax, 2004; Richards et al.,

2006) and may fail to resolve meaningful differences between events in epicenter and especially depth (Hauksson et al., 2020; Schoenball & Ellsworth, 2017), perhaps because of poor station distribution and consequent poor ray coverage around the sources, or because of low accuracy and precision in the underlying standard locations.

Here we introduce a standard, arrival-time location procedure, NLL-SSST-coherence, modified to improve relative location accuracy through use of spatially varying, source-specific station travel-time corrections (SSST) and a new, waveform coherence based, multi-event location procedure. In a first relocation stage, an event catalog is iteratively relocated while generating smoothly varying, SSST corrections throughout a 3D volume, providing a source-position dependent correction for each station and phase type. The iteration uses Gaussian smoothing kernels of decreasing size to produce final, NLL-SSST locations. Residuals from P and S arrivals and relocated events meeting minimum quality criteria are used for update at each iteration.

In a second relocation stage the relative location accuracy of the NLL-SSST locations is further increased by consolidating location information across events based on waveform coherence between events. This coherence relocation is based on the concept that if the waveforms at a station for two or more events are very similar (have high coherence) up to a highest frequency, then the distance separating these "multiplet" events is small relative to the seismic wavelength at that frequency, the events may even correspond to stress release on the same, small fault patch (Geller & Mueller, 1980; Nadeau et al., 1994; Poupinet et al., 1982, 1984).

We present a synthetic test which shows that the NLL-coherence relocation procedure correctly and significantly reduces hypocenter scatter by grouping together multiplet events as well as shifting outlier hypocenters towards their multiplet centroid. However, the procedure may over-cluster events, since well-located events strongly "attract" high-coherence multiplet events that are poorly constrained by insufficient or noisy arrival time data. We apply the NLL-SSST-coherence location procedure to a ground-truth, explosion dataset using only P arrival times and waveforms from a single station to show that the procedure gives nearly the same relative location accuracy as obtained with high-precision, correlation-based time-delay measurements and double-difference, relative relocation.

We next apply the NLL-SSST-coherence location procedure to the 2004 Mw 6.0 Parkfield, and 2020, Mw 5.8 Lone Pine California earthquake sequences and compare the results with other standard and relative location catalogs for these sequences. The NLL-SSST-coherence relocations generally show increased organization, clustering and depth resolution of seismicity over other standard location catalogs. Compared to relative location catalogs, the NLL-SSST-coherence relocations recover well smaller scale patterns and features in the seismicity, with evidence of improved, larger scale relative location accuracy when there are few stations over or near the seismicity. Application of NLL-SSST-coherence locations is also presented in Lomax (2020b) for the 2020 Mw 6.5 Monte Cristo, Nevada sequence.

These results show that the NLL-SSST-coherence location procedure approaches the precision of cross-correlation based, relative location methods, while requiring less computing time and being applicable to sparser station distributions and studies with limited waveform data.

2 The NLL-SSST-coherence procedure for high-precision earthquake location

We obtain high-precision earthquake relocations through the combined use of source-specific, station travel-time corrections (SSST) and stacking of probabilistic event locations based on inter-event waveform coherence. We use the NonLinLoc location algorithm (Lomax et al., 2000, 2014); NLL hereafter), which performs efficient, global sampling to obtain an estimate of the posterior probability density function (PDF) in 3D space for hypocenter location. This PDF provides a complete description of likely hypocentral locations with comprehensive uncertainty information, and allows robust application of waveform coherence relocation. Within NLL, we use the equal differential-time (EDT) likelihood function (Font et al., 2004; Lomax, 2005, 2008; Lomax et al., 2014; Zhou, 1994), which is highly robust in the presence of outlier data caused by large error in phase identification, measured arrival-times or predicted travel-times. We use a finite-differences, eikonal-equation algorithm (Podvin & Lecomte, 1991) to calculate gridded P and S travel-times for initial NLL locations.

2.1 Source-specific station term corrections

In a first relocation stage, NLL-SSST-coherence iteratively develops SSST corrections, which can greatly improve relative location accuracy and clustering of events (Pavlis & Hokanson, 1985b; Richards-Dinger & Shearer, 2000; Lin & Shearer, 2005; Nooshiri et al., 2017). In contrast to station static corrections (Ellsworth, 1975; Frohlich, 1979; Lomax, 2005, 2008; Tucker et al., 1968) which give a unique time correction for each station and phase type, SSST corrections vary smoothly throughout a 3D volume to specify a source-position dependent correction for each station and phase type. Spatial-varying, SSST corrections are most important when the ray paths between stations and events differ greatly across the studied seismicity, including when stations are inside the seismicity distribution, the extent of seismicity is large relative to the distance to the stations, or the depth range of events is large. SSST corrections increase in importance as error in the velocity model increases, such as when a 1D, laterally homogeneous model or a large-wavelength, smooth model is used in an area with sharp, lateral velocity contrasts or small scale, 3D heterogeneity.

Within the NonLinLoc package (Lomax et al., 2000, 2014), SSST corrections are developed iteratively with spatial smoothing of decreasing size using a Gaussian kernel (Fig. 1), this approach is similar to the shrinking box SSST approach of (Lin & Shearer, 2005). Given an initial set of gridded travel-times and event locations, 3D grids of SSST corrected travel-times for each station-phase are created iteratively by:

- At each node in the corrected travel-time grid and for each station-phase:
    - Accumulate the weighted mean of residuals, $\bar{R}$, for the station-phase for each event location exceeding specified quality criteria. The weight, *w*, is given by a modified Gaussian kernel,

      $$w = \exp(-d^2/D^2) + \epsilon, \qquad (1)$$

      where *d* is the distance between the grid node and the event hypocenter, *D* controls the smoothing width, and ϵ is a small value to give finite weight for all events and thus non-zero corrections even if all event hypocenter are far from the grid node.
    - Add $\bar{R}$ as the current SSST correction to the previous travel-time for the station-phase at the node and store at the node in the updated SSST corrected travel-time grid.
- Relocate all events using the updated SSST corrected travel-times.
- Reduce *D* and return to step 1 if $D \geq D_{min}$, the smallest required smoothing distance.

For the case of a grid node far from all event hypocenters, all weights, *w,* will be approximately ϵ, and $\bar{R}$ will be close to the station static correction for the set of locations. Similarly, if the starting value of *D* is large relative to the extent of stations and hypocenters, then $\bar{R}$ for all station-phases will be close to the station static correction for the first SSST iteration. D_{min} might be set so the corrections vary slowly on the scale of the smallest target features in the seismicity and are derived from numerous events (e.g. more than 10-100 within Dmin) in denser areas of seismicity. Additionally, a check for improvement in the suite of SSST relocation results with decreasing *D* may suggest that results at a larger *D* than D_{min} should be used for further analysis.

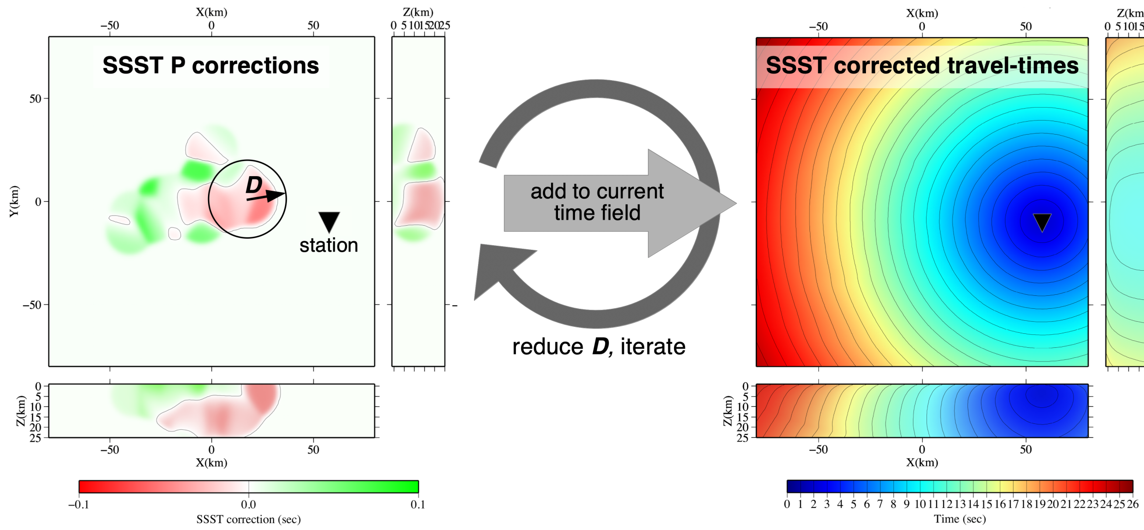


Fig. 1. Schematic of iterative development of SSST corrections. Given relocations with current travel-time fields, the weighted mean of residuals (left) is obtained with smoothing width *D* for P phases at a station (black triangle). These SSST corrections are added to the current, P travel-time field for the station to produce updated, SSST corrected travel-times (right). The smoothing width *D* is reduced and the process is iterated.

2.2 Waveform coherency relocation method

In a second relocation stage, NLL-SSST-coherence invokes a new procedure which greatly reduces aleatoric location error by consolidating information across event locations based on waveform coherency between the events. This coherency relocation, NLL-coherence, is based on the concept that if the waveforms at a station for two events are very similar (e.g. have high coherency) up to a given dominant frequency, then the distance separating these events is small relative to the seismic wavelength at that frequency (e.g., Geller & Mueller, 1980; Poupinet et al., 1984), perhaps less than about ¼ of this wavelength (Geller & Mueller, 1980; Thorbjarnardottir & Pechmann, 1987). A pair of similar events is a doublet and a set of similar events may be called a cluster, multiplet or family, these events all likely occur on a small patch of a fault with similar magnitude and source mechanism (Cattaneo et al., 1997; Ferretti, 2005; Geller & Mueller, 1980; Ishida & Kanamori, 1978; Nadeau et al., 1994; Poupinet et al., 1982). In a high-precision, microseismic study (Goertz−Allmann et al., 2017) show for waveform windows spanning both P and S waves that correlation coefficients greater than about 0.7 indicate event multiplets locate within about 0.1 km, which is about ¼ wavelength for the typical dominant waveform frequency ~20 Hz and wave velocity of ~2.5 km/s shown in their study. The results of (Goertz−Allmann et al., 2017; their figs. 4 and 6) also show lack of clustering and separation of event pairs throughout the region of studied seismicity for correlation coefficients less than about 0.5.

For detailed seismicity analysis, the precise hypocenter locations of events in multiplets can be assigned to a unique centroid point or coalesced in space through some statistical combination of the initial hypocenter locations (Jones & Stewart, 1997; Kamer et al., 2015). Alternatively, precise, differential times between like-phases (e.g. P and S) for doublet events can be measured using time- or frequency-domain, waveform correlation methods. Differential times from a sufficient number of stations for pairs of doublet events allows high-precision, relative location between the events, usually maintaining the initial centroid of the event positions (Got et al., 1994; Ito, 1985; Matoza et al., 2013; Nadeau et al., 1994; Nakamura, 1978; Poupinet et al., 1982, 1984; Waldhauser & Ellsworth, 2000).

Here we use waveform similarity directly to improve relative location accuracy without the need for differential time measurements or many stations with waveform data. We assume that high coherency between waveforms for two events implies the events are nearly co-located, and also that all of the information in the event locations, when corrected for true origin-time shifts, should be nearly identical in the absence of noise. Then, stacking procedures can be used to reduce the noise in this information and improve the location precision for individual, target events. We use the coherency between waveforms for pairs of events (i.e. the target event and all other events) at one or more stations to combine through stacking an initial set of location probability density functions (PDF's). This stack directly improves the hypocenter location for each target event by effectively combining and completing arrival time data over events and reducing noise (aleatoric error) in this data.

We measure waveform coherence as the maximum cross-correlation between two waveforms (e.g., Aster & Scott, 1993), calculated using the xcorr function in the ObsPy Python package (Beyreuther et al., 2010; Krischer et al., 2015), which performs a normalized cross-correlation in the time-domain. To form weights for stacking location PDF's (Fig. 2), positive coherences, C, above a minimum cutoff value, C_{min}, (e.g. 0.5) up to a plateau cutoff value, C_{plat}, (e.g. 0.9) are mapped linearly to the range 0.0 to 1.0, w_{lin},

$$w_{lin} = (C - C_{min}) / (C_{plat} - C_{min}). \tag{2a}$$

and then mapped through a smooth, cosine taper to form 0.0 to 1.0 stacking weights, W,

$$W = 0.5 \cos(\pi w_{lin}) + 0.5. \tag{2b}$$

Stacking weights for coherences C less than C_{min} or greater than C_{plat} are set to $W=0$ or $W=1$, respectively.

For cross-correlation, we use a waveform window that includes P and S waves so that we maintain the S-P time interval, the P coda and part of the S coda, all of which better constrain waveform similarity for the purpose of quantifying the proximity of events (Fig. 2). When the waveforms for multiple stations are available for a pair of events, we use the maximum of inter-event coherency over stations as the coherency for stacking. This choice is justified since the coherency for real, noisy waveforms is much less likely to be over-estimated than under-estimated. The number of event pairs for which coherence is calculated can be reduced by only considering pairs with initial inter-hypocenter separation within a maximum cutoff distance (e.g., Aster & Scott, 1993).

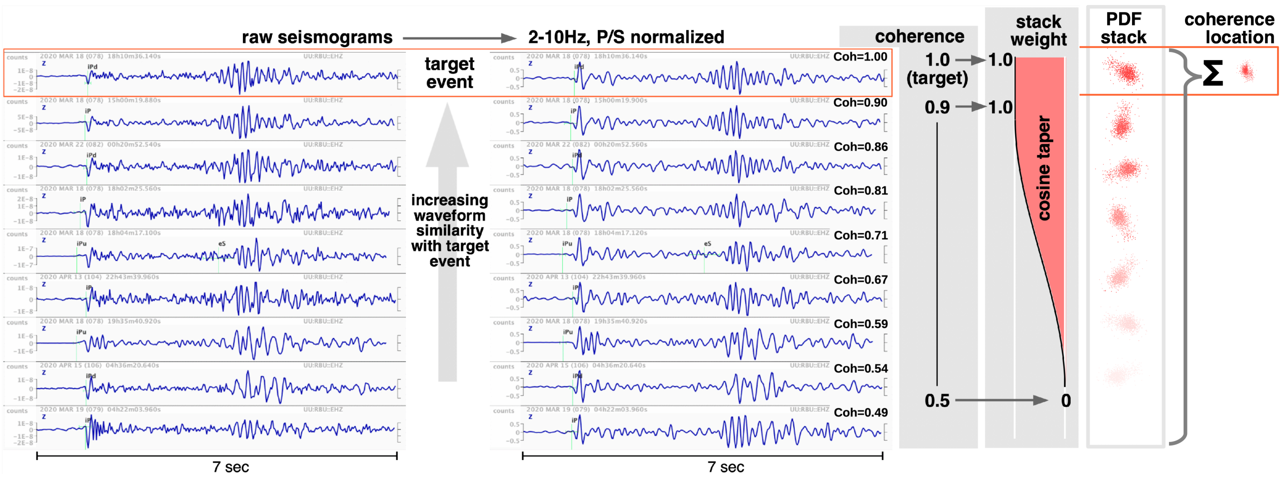


Fig. 2. Example traces showing waveform coherency, stack weights and PDF stack. (left) Raw seismograms for a set of aftershocks events. (middle) Corresponding, 2-pole, 2Hz – 10Hz bandpass filtered waveforms used for coherence calculations. Top trace is the target waveform; coherence (Coh) with the target is indicated to the upper right of each waveform. (right) Schematic representation of coherences between the target and each waveform, corresponding stacking weight after cosine taper mapping of coherence, location PDF's forming the stack (color intensity indicates stack weight), and final NLL-coherence location PDF for the target event.

The NLL-coherence procedure requires a set of initial locations and corresponding PDF's for the spatial hypocenter locations. The NLL location PDF is a probability density function over possible, 3D spatial hypocenter locations which combines information in the observed data, the prior (location search volume), and the ability of the forward problem to predict the observed data (Lomax et al., 2014; Tarantola & Valette, 1982). For each target event, the procedure forms a weighted stack of normalized PDF's over 3D space consisting of: the initial location PDF for the target event with a weight of 1, and the location PDF weighted by W for each of the other events that have inter-event coherency with the target event greater than C_{min}. The PDF stack is raised to the power of the total sum of weights, which concentrates the PDF to show uncertainty in the relative location of multiplet events; without raising to this power, the PDF's would show the generally much larger uncertainty of the original, standard locations.

PDF stacking with weighting is established in probability theory as Bayesian model averaging, a procedure for combining forecasts from multiple possible solutions (Fragoso & Neto, 2018; Hoeting et al., 1999). Given $n = 0, N\text{-}1$ similar events and invoking Bayesian model averaging, the NLL-coherence location likelihood (unnormalized PDF), L_{coh}, for a target event in the set of similar events is,

$$L_{coh}(\boldsymbol{x}|\boldsymbol{d}) = \sum_{n=0}^{N-1} L_n(\boldsymbol{x}|W_n, \boldsymbol{d})\, W_n \quad , \tag{3}$$

where $\boldsymbol{x}$ is spatial position, $\boldsymbol{d}$ is the set of waveform and pick data for all considered events, and $L_n()$ and W_n are the NLL-SSST location PDF and stacking weight, respectively, for event n.

This combined forecast is simply a weighted average of the NLL initial location PDF's for each of the similar events. Formally, in Bayesian model averaging, the weights W_n are the posterior probabilities of the solutions L_n for each event n given the data $\boldsymbol{d}$ – a measure of how plausible the solution is given the data. For NLL-coherence, under the assumption that events with similar waveforms have similar hypocenter positions, we use each solution L_n as a proxy solution for the target event and use waveform coherence to define the plausibility W_n of each L_n for constraining the target event position.

The NLL-coherence PDF stack is raised to the power of the total sum of weights to make the shape and spread (and thus measures such as variance) of the final solution comparable to those of a product of PDF's. This effect is illustrated by a sum of N normal distributions with zero mean raised to the power N,

$$\left[\sum_{0}^{N-1} \mathrm{e}^{-x^2/\sigma^2}\right]^N = \left[N\,\mathrm{e}^{-x^2/\sigma^2}\right]^N \quad , \tag{4}$$

which is proportional to the product of the same normal distributions,

$$\prod_{0}^{N-1} \mathrm{e}^{-x^2/\sigma^2} = [\mathrm{e}^{-x^2/\sigma^2}]^N \quad . \tag{5}$$

The NLL-coherence stack PDF forms the probabilistic, coherence relocation for the target event and defines all location information, such as origin time, location uncertainties, and arrival-time residuals. The NLL-coherence procedure can be implemented as a workflow using modules of the NonLinLoc package.

The stacking weights W_n can also be used to combine and, in effect, stack first-motion readings between multiplet events. A combined set of first-motion readings is formed from the target events readings with a weight of 1.0 and readings from each similar event n with weight W_n. This augmented set of readings produces a greater number of composite, better constrained focal-mechanisms than do sets of single event readings, though these composite mechanism are locally correlated across multiplet events.

For an event that has coherency with all other events less than C_{min}, the PDF stack and all location information will be identical to those for the initial location for the event. For an initial event that is poorly constrained with an extensive PDF, but which has high coherency with other, well constrained events, the stacked PDF location will closely match the locations of the well constrained events. Unlike differential-time based, relative location methods, NLL-coherence relocation can be performed with waveforms from few or even a single station. Consequently, NLL-coherence relocation is computationally efficient, allows precise relocation of seismicity when the closest station is far from the seismicity and for sparse networks, enables precise relocation of foreshocks and early aftershocks in a mainshock sequence or swarm before nearby temporary stations are installed, and can be applied to historical sequences with little available waveform data.

Traces showing different values of waveform coherency for an example, target event, and a schematic of mapping from coherence to stack weight and the stacking of PDF’s to form the NLL-coherence location are shown in Fig. 2. Fig. 3 shows example event PDF's before and after NLL-coherence location and the event PDF's that are weighted by coherence and summed to form the NLL-coherence location for the example event. Fig. 3a shows the NLL-SSST hypocenter and PDF for a target event, Fig. 3b shows the hypocenters and PDF’s for the target event and all similar events, Fig. 3c shows the NLL-coherence location for the target event after coherence weighted stacking over all the event PDF’s in Fig. 3b. Fig. 3d shows the NLL-coherence locations for all similar events from Fig. 3b after coherence weighted stacking for each event; these epicenters and PDF's show how NLL-coherence location produces clustering and organization of similar event hypocenters along with greatly reduced PDF extent, which shows formal location uncertainty.

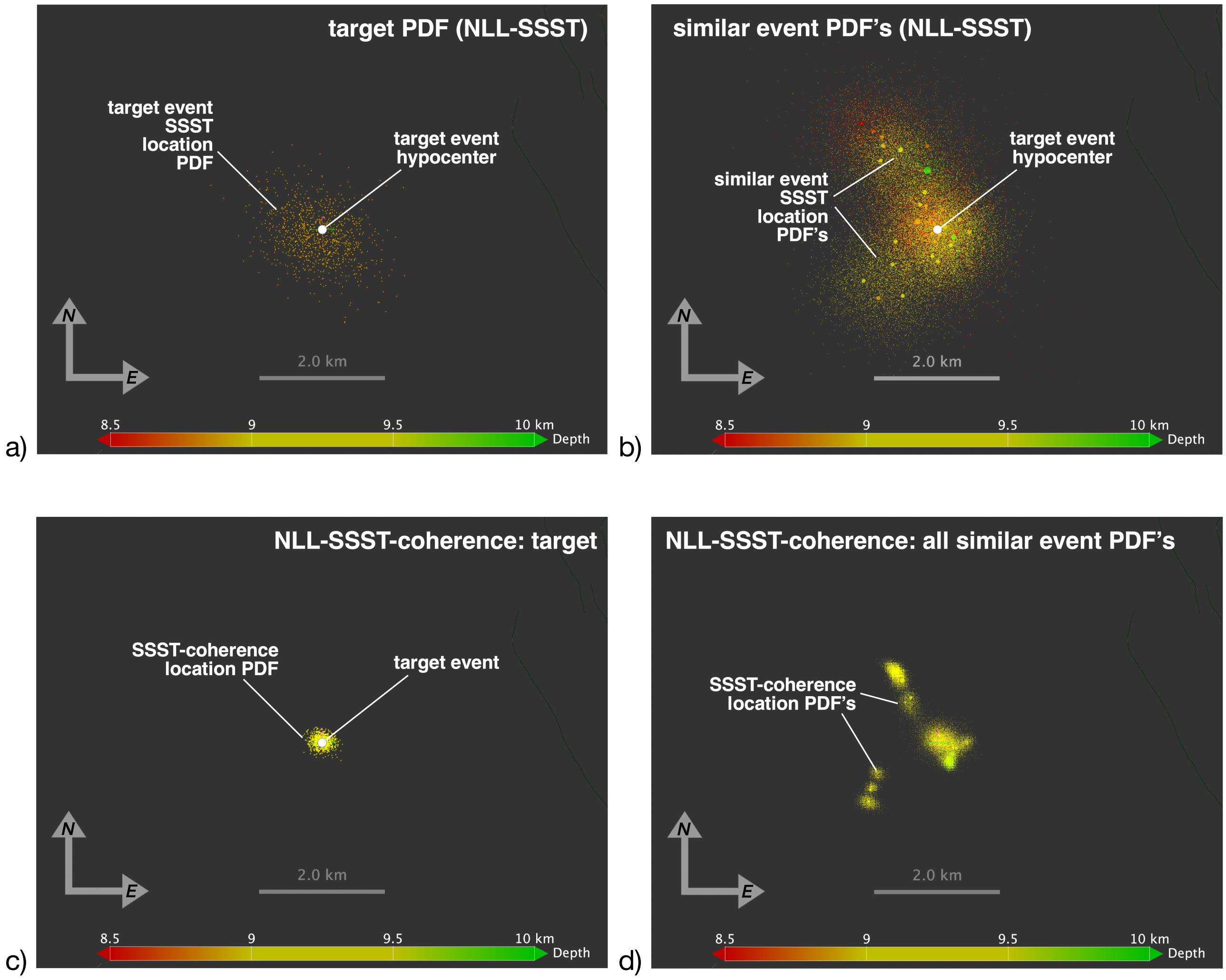


Fig. 3. Example event PDF's before and after NLL-coherence location. Epicenters (large dots) and location PDF's (yellow clouds) for NLL-SSST relocations for (a) a target event and (b) similar events with waveforms coherency C≥0.5 with respect to the target. Epicenters and stacked location PDF's after NLL-coherence location for (c) the target event alone, and (d) all events similar to the target. The PDF's in (b) are summed with stacking weights, *W*, to give the target NLL-coherence location shown in (c).

3 Synthetic test of NLL-coherency relocation

We test and illustrate the performance of the NLL-coherence relocation procedure using synthetic, noisy data for a set of irregularly clustered, multiplet, benchmark events within a realistic network station geometry. It is important to synthesize realistic epistemic and aleatoric arrival time and travel-time error, since NLL-coherence relocation is based on reducing the error in hypocenters obtained from such data.

To represent a realistic, clustered hypocenter distribution we form a set of benchmark events irregularly spaced on a circle of radius 5 km at a depth of 3 km (Fig. 4). The benchmark hypocenters are randomly placed on the circle so as to produce dense clusters of events, more sparsely spaced events, and larger gaps between event groups. This distribution represents multiplet events on small asperities, surrounding isolated events, and aseismic zones in between. The station distribution, derived from the TexNet network around Pecos, Texas (Savvaidis et al., 2019), consists of 30 stations to about 50 km distance and well distributed in azimuth around the benchmark events.

Epistemic and aleatoric errors arises from velocity model and travel-time error, noisy waveform phase onsets, phase mis-identification, and so on. To model realistically these errors, we create noisy, synthetic arrival times for the benchmark events. We first calculate exact travel-times for the benchmark hypocenters using a smooth, laterally homogeneous crustal model. Then, to introduce pseudo-realistic error and uncertainties for P and S arrival times for each event, we: 1) create a P or an S arrival time datum for each station with a probability of 0.3, 2) add a random, Gaussian timing error with standard deviation 0.04 sec for P or 0.1 sec for S to the exact arrival time, 3) randomly double this error with probability 0.5, and 4) for location, set a nominal Gaussian picking uncertainty with standard deviation of 0.02 sec for P and 0.05 sec for S. The use in step 4) of a smaller nominal uncertainty than the timing error added in steps 2) and 3) effectively introduces outlier arrival data and mimics mild to moderate epistemic error in the velocity model and arrival times. The above settings are chosen to mimic un-modeled error and resulting hypocenter scatter in the last iteration of NLL-SSST locations, not the typically much larger error and hypocenter scatter present in initial catalog locations.

Using a velocity model for relocation of this synthetic, noisy data that differs from the one used to calculate the exact travel-times for the benchmark hypocenters would reproduce large epistemic error due to velocity model error. However, we do not use a differing model here, because epistemic model error typically introduces an overall bias in hypocenter locations which would obscure the more interesting and important reduction of scatter and clustering of hypocenters that the coherence location produces. A bias or strong distortion of hypocenter locations due to large scale, epistemic velocity model error will remain in NLL-coherence relocation, as it does in all other location procedures and algorithms.

As there are no waveforms for the benchmark events, but we know their “true” hypocenter positions, we create synthetic coherences between all pairs of events based on the benchmark distance between their hypocenters. This coherence C is defined by,

$$C = \exp(-x^2/X^2), \quad (6)$$

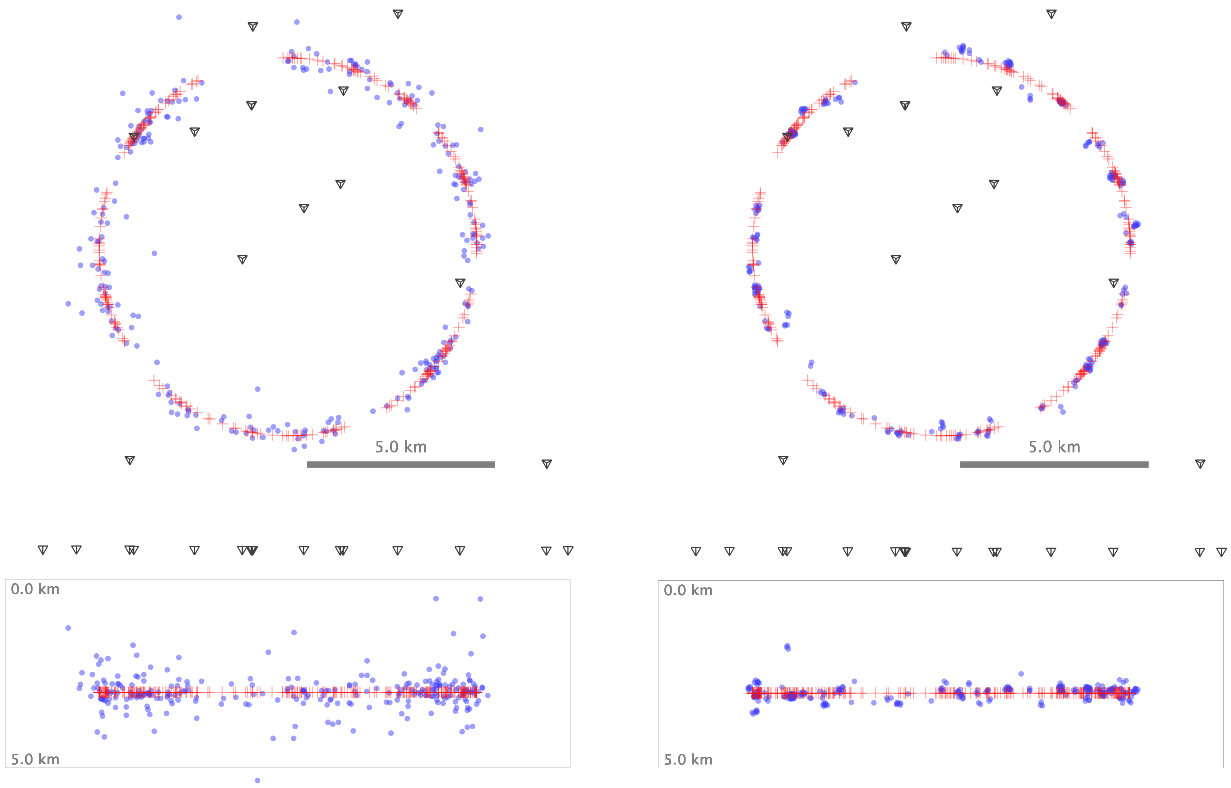
where x is the true distance between two benchmark hypocenters and X is a characteristic distance, here set to 1 km. Coherences, C, above a minimum cutoff value C_{min}=0.5 are mapped to 0.0 to 1.0 weights, W, for stacking similar event location PDF's, as described in Section 2.2 and Equation 2. Thus two events with true separation of 0.6 km will have a synthetic coherence of about 0.7 and a PDF stack weight of about 0.5.

The standard, NLL event locations using the noisy, synthetic arrival times gives the set of noisy, synthetic hypocenters shown in Fig. 4. Note the degree of scattering of these hypocenters around the circle of benchmark hypocenters, including the much larger scatter of hypocenters in depth than in epicenter, the general clustering of events around denser sets of benchmark hypocenters, and the presence of clear outlier hypocenters far from the true benchmark events.

These standard NLL locations are used as starting locations for NLL-coherence relocation, shown in Fig. 4. The coherence relocations are performed with the synthetic coherence, PDF stack weights, other settings are similar to those used for relocation of California sequences in Section 4. Relative to the standard locations (Fig. 4), note the greatly reduced scatter of coherence hypocenters around the circle of benchmark hypocenters, particularly in depth, the tight clustering of most events typically near denser sets of benchmark hypocenters, and the presence of few outlier hypocenters.

These results show several important aspects of NLL-coherence event relocation. Firstly, coherence relocation correctly and significantly reduces hypocenter scatter towards the true benchmark hypocenter locations. True epicenter error is reduced from a mean of 0.3 km for the noisy locations to a mean of 0.2 km with coherence location; true depth error standard-deviation is reduced from 0.5 km for the noisy locations to 0.2 km with coherence location. Secondly, NLL-coherence relocation correctly and significantly shifts outlier hypocenters towards the benchmark hypocenter locations. True epicenter outliers up to 1.3 km for the noisy locations are reduced to no outliers > 0.6 km with coherence location; true depth outliers up to 2.5km for the noisy locations are reduced to only one outlier > 0.5 km with coherence location.

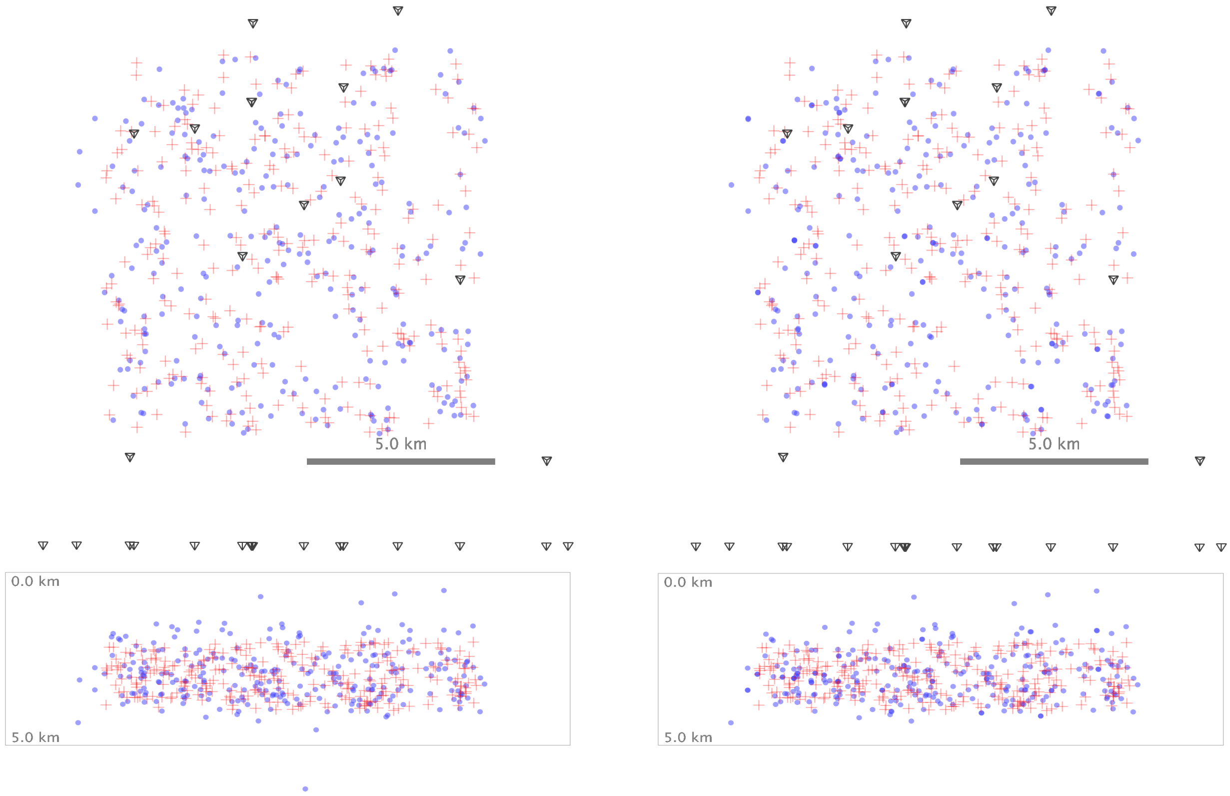
Thirdly, coherence relocation tends to cluster sets of events near denser sets of benchmark hypocenters – this implies that coherence relocation will correctly group together multiplet events as defined by coherence. However, coherence location tends to over-tightly cluster events, so that they often fill a smaller volume than the true spread of benchmark events. This is likely caused by the presence of sub-sets of well constrained, well located events with co-located PDF's of small extent which “attract” high-coherency, but poorly constrained, multiplet events with extensive PDF’s. Effectively, large arrival time error, or outlier and missing arrival time data for a poorly located event represent a loss of information which cannot be recovered, so improvement in location is obtained by substitution of the weighted stack of better constrained, multiplet event PDF's as a proxy location.



noisy synthetic events NLL-coherence relocations

Fig. 4. Synthetic circle test of NLL-coherence relocation. Map view (top row) and section view from south (bottom row). Red crosses show true benchmark hypocenters used for generating P and S arrivals. Left column shows noisy synthetic events (blue dots) obtained through standard location with realistic noise added to P and S arrivals. Right column shows NLL-coherence relocations (blue dots) of noisy synthetic events; events are shifted randomly 0.1 km to avoid overlapping symbols due to the discrete gridding of stored SSST PDF’s used for NLL-coherence relocations. Nearby station shown as inverted pyramids.

We additionally performed an identical synthetic test except using benchmark events drawn from a uniform random distribution instead of irregularly spaced on a circle. This test shows the NLL-coherence response to statistically unclustered seismicity and checks if the procedure produces artifacts such as those found by (Nicholson et al., 2000) for hypocenter coalescence procedures based solely on original locations and their uncertainties. These artifacts include an overall shrinking of the cloud of seismicity towards its barycenter, and low density holes with surrounding high density webs and clusters of seismicity not present in the benchmark distribution. The results in Fig. 5 show that the NLL-coherence seismicity has a distribution similar to that of the benchmark events and the noisy synthetic events. The NLL-coherence relocations do not show increased or decreased clustering relative to the noisy synthetic locations, and do not show any of the artifacts found by (Nicholson et al., 2000). The mean epicenter error relative to the benchmarks is 0.3 km for both the noisy and coherence locations. True depth error standard-deviation is reduced slightly from 0.5 km for the noisy locations to 0.4 km with coherence location, primarily due to the coherence relocation correctly and significantly shifting several outlier hypocenters towards the benchmark hypocenter locations. Thus the NLL-coherence procedure does not introduce artifact clustering or de-clustering for unclustered, uniform random benchmark seismicity, but does detect and remove some outlier events in the corresponding noisy synthetic events.



noisy synthetic events NLL-coherence relocations

Fig. 5. Synthetic uniform random test of NLL-coherence relocation. Map view (top row) and section view from south (bottom row). Red crosses show true benchmark hypocenters used for generating P and S arrivals. Left column shows noisy synthetic events (blue dots) obtained through location with realistic noise added to P and S arrivals. Right column shows NLL-coherence relocations (blue dots) of noisy synthetic events. Nearby station shown as inverted pyramids.

4 Ground-truth test of NLL-SSST-coherence relocation

We further test and illustrate the performance of the NLL-SSST-coherence relocation procedure through application to regional recordings of a ground truth (GT) dataset of surface explosions in Finland in 2007 as reported by (Gibbons et al., 2020). The explosions have known coordinates and are located in two tight and one extended clusters, all within an area of about 0.3 x 0.3 km (Fig. 6). (Gibbons et al., 2020) analyze and provide waveforms for 6 stations at about 59 to 208 km from the explosion sources and with good azimuthal distribution and coverage to represent a realistic, sparse regional network. We automatically pick P arrivals on the waveforms using FilterPicker (Lomax et al., 2012) with default settings, and then manually check the picks, modifying or adding a small proportion of the picks; this picking procedure mimics the automatic picking with manual revision of many regional networks. We do not pick S arrivals as they do not have clear onsets at most stations for most events. In (Gibbons et al., 2020), all locations are performs with depth fixed, since the stations at regional distance provide no depth constraint, but here we allow depth to vary between -2 and 8 km depth to better model the poor depth control of many relocation studies. Following (Gibbons et al., 2020), the ak135 velocity model (Kennett et al., 1995) is used for initial NLL locationsFig. 6. The initial locations (Fig. 6a) form clusters and scattered events over an area of about 10 x 10 km and from -2.0 km to about 6.3 km in depth.

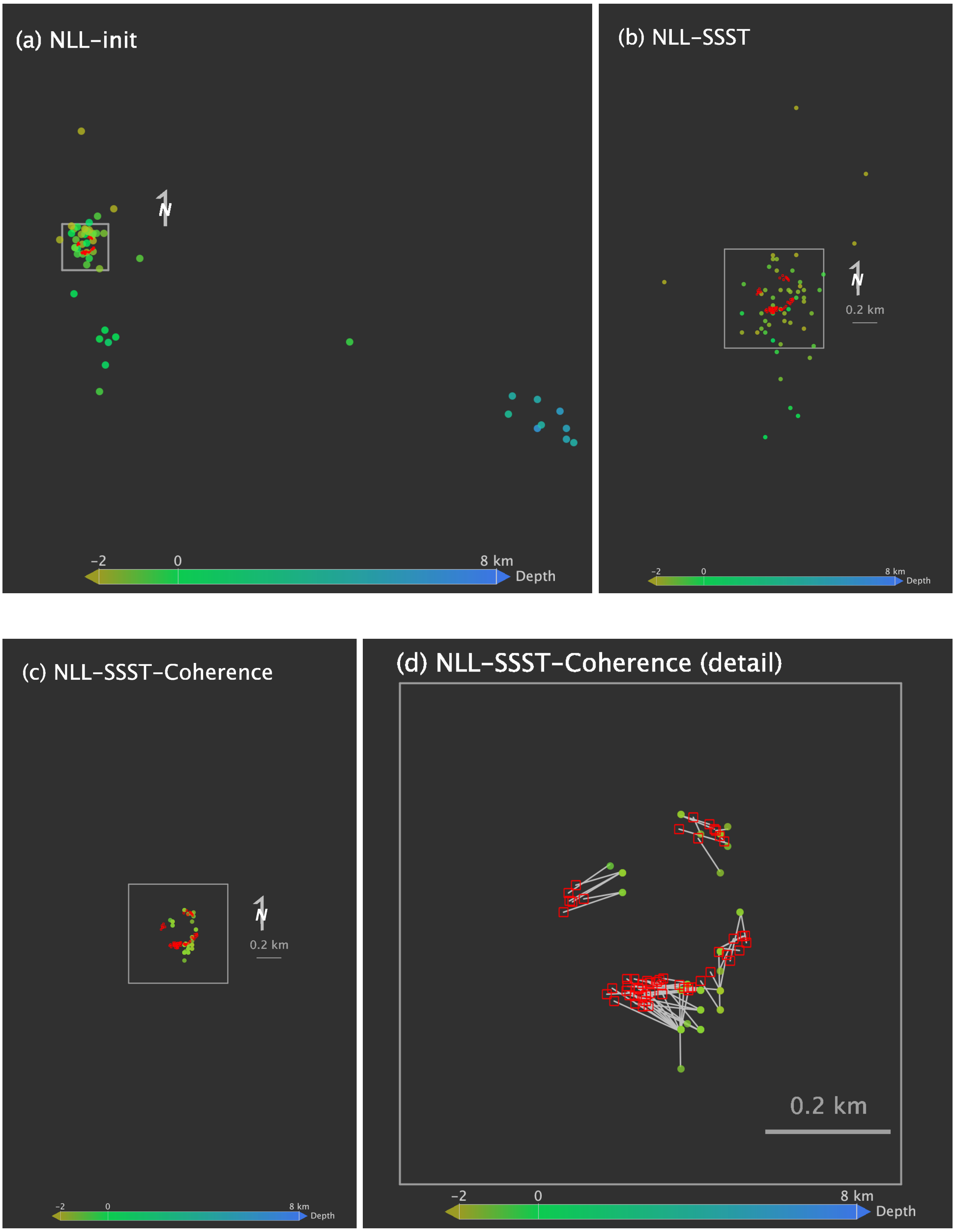

(a) NLL-init
-2
0
8 km
Depth
(b) NLL-SSST
0.2 km
-2
0
8 km
Depth
(c) NLL-SSST-Coherence
0.2 km
-2
0
8 km
Depth
(d) NLL-SSST-Coherence (detail)
0.2 km
-2
0
8 km
Depth

Fig. 6. Ground truth test of NLL-coherence relocation. Red squares show GT explosion epicenters, dots show relocated epicenters with color indicating hypocenter depth. (a) Initial NLL locations using auto and manual picked P arrivals; one event falls outside the plot about 8 km north-northeast of the GT events. (b) Relocations after 3 iterations of NLL-SSST using a very large smoothing width *D*; one event falls outside the plot about 5 km east-southeast of the GT events. (c) NLL-SSST-coherence relocations. (d) Detailed view of NLL-SSST-coherence relocations; gray lines connect corresponding GT and relocated epicenters. In all panels the gray box is 0.8 km square and the relocated epicenters are shifted so the northeast cluster of 8 events in panel (d) aligns with the corresponding GT cluster; the absolute positions of the relocated events are approximately 0.8 km northwest of the GT locations.

For NLL-SSST relocations, since the source area is small relative to the station distances, we iteratively generate SSST corrections with a fixed, large smoothing distance, *D,* of 999 km, thus effectively calculating each weighted mean of residuals, $\bar{R}$, as a station static correction (mean of all residuals). The quality criteria for an event location and station-phase to be used for calculating $\bar{R}$ are: 68% error-ellipsoid principle-axis half-width ≤8.0 km, root mean square of residuals (rms) ≤0.075 sec, P residual ≤ 0.5 sec, S residual ≤ 0.5 sec. The final NLL-SSST relocations (Fig. 6b) merge the two principle clusters and most scattered events in the initial NLL locations (Fig. 6a) into one cluster within an area of about 1 km square and a depth range of -2.0 to 0.6 km, one event remains about 5 km east-southeast of the GT events at a depth of about 6 km.

For NLL-coherence relocations, we measure coherency using vertical component waveforms from only the closest station, LP53, at about 59 km, channel XK.LP53.00.HHZ with 50 Hz sampling. The waveforms are filtered from 2-25Hz in a window from 4 sec before the predicted P arrival to 4 sec after the predicted S arrival. Cross-correlation is applied between waveforms windows sliding from -2.0 to 2.0 sec, and the 0-1 stacking weight is set following Eq. 2 over coherency values from C_{min} = 0.45 to 1.0. This procedure is applied to the final NLL-SSST relocations (Fig. 6b) for all event pairs. The NLL-SSST-coherence relocations are shown in Fig. 6c and in detail in Fig. 6d.

The NLL-SSST-coherence relocations for the explosion dataset (Fig. 6d) cover about the same extent as the GT epicenters and recover the two tight and one extended GT clusters with correct identification of member events for each clusterFig. 6. The NLL-SSST-coherence epicenters show some distortion in the relative position of the clusters (~0.1 km) and in the relative positions of events within each cluster (< ~0.05 km). But, remarkably, these distortions are not much greater than those obtained with high-precision, correlation-based time-delay measurements and double-difference relocation by (Gibbons et al., 2020; their fig. 5). Distortion of relative cluster positions is attributed by (Gibbons et al., 2020) to error in variations in velocity model slowness across the GT source region.

For the NLL-SSST-coherence relocations, ¼ of the seismic wavelength at the highest frequencies with signal energy, ~20 Hz, is about 0.075 km. This distance likely represents the lower limit of inter-event separation resolvable by the NLL-coherence procedure, which agrees well with the good NLL-coherence resolution of the relative, horizontal cluster positions, which are separated by 2-3 times the ¼ wavelength distance, and poor resolution of intra-cluster, relative event positions, separated by much less than this distance. The intra-cluster, relative event positions are most likely noisy and not robust, affected by error in the arrival picks and NLL-SSST locations, and by the tendency of NLL-coherence to over-cluster on the smallest scales.

The depth range of the NLL-SSST-coherence relocations is about -1.7 to -1.2 km, much less than the range for the initial and NLL-SSST locations, though the error in depth range relative to GT (all GT sources at about -0.5 km depth) is larger than relative epicentral error due to very poor depth constraint from the used data. However, the striking improvements of relative depth and epicenters NLL-SSST-coherence over the initial locations provides further evidence that the combined SSST and coherence procedures can correctly shift noisy and strong, outlier locations towards similar event hypocenters.

5 NLL-SSST-coherence relocation for California earthquake sequences

We next show how NLL-SSST-coherence relocation performs relative to established standard and relative location procedures using two recent earthquake sequences in California (Fig. 7). First, we examine the 2004 Mw 6.0 Parkfield sequence, which was well recorded by numerous seismic stations around and above the seismicity, to show how NLL-SSST-coherence relocation improves on standard locations and approaches the precision of waveform, cross-correlation based, relative location methods. Then we examine the 2020 Mw 5.8 Lone Pine sequence to show how NLL-SSST-coherence relocation can produce higher precision locations and better depth control than waveform, cross-correlation based, relative location when there are no seismic stations above a sequence and few nearby stations. Lomax (2020b) also presents analysis of NLL-SSST-coherence locations and comparison with routine catalog locations for the 2020 Mw 6.5 Monte Cristo, Nevada sequence.

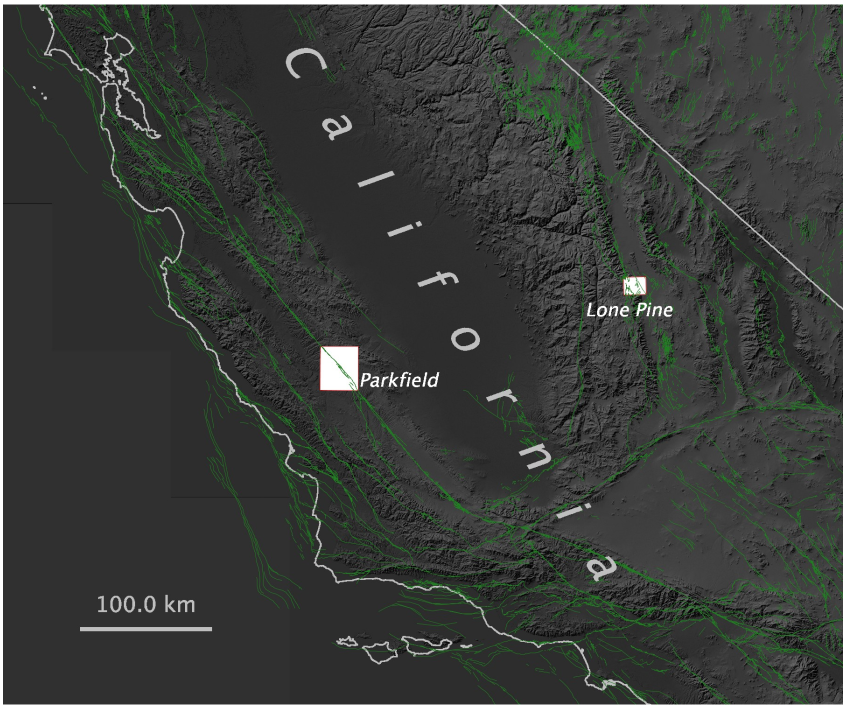


Fig. 7. Location in California of the Parkfield and Lone Pine study areas. Green lines show faults from the USGS Quaternary fault and fold database for the United States. Background topography image from OpenTopography.org.

5.1 2004 Mw 6.0 Parkfield, California

The 2004 Mw 6.0 Parkfield earthquake sequence occurred along a 40 km stretch of the San Andreas Fault Zone (SAFZ) in central California (Fig. 7) between a 150 km long, creeping section of the fault to the northwest and a locked section to the southeast that last ruptured in the 1857 Mw 7.9 Fort Tejon earthquake (Bakun et al., 2005). The 2004 sequence was recorded by a large number of well-distributed seismic stations, including borehole and high sample-rate instruments. This seismic data has been used in numerous studies to examine the velocity structure in the area and for high-precision, waveform cross-correlation based, relative location methods (Michelini & McEvilly, 1991; Nadeau et al., 1994; C. Thurber et al., 2006; Waldhauser et al., 2004; Zhang et al., 2009). The Parkfield sequence is therefore an excellent reference case for examining the performance of the NLL-SSST-coherence relocation procedure for obtaining high-precision hypocenter locations.

We obtain a catalog (USGS-NCSN Catalog) of 2828 events for the Parkfield area (latitude 35.75º to 36.05 º, longitude -120.62º to -120.30º) with M ≥ 1.5 from 1984-01-01 to 2020-21-31 from the Northern California Earthquake Data Center (NCEDC). The USGS-NCSN Catalog standard locations (NCSN-ABS), obtained using localized velocity models and station travel-time corrections, and corresponding NCSN Double-Difference Catalog locations (NCSN-DD) based on the HypoDD, double-difference, relative location method (Waldhauser, 2009; Waldhauser & Ellsworth, 2000; Waldhauser & Schaff, 2008) are shown in Fig. 8. We also obtain from NCEDC P and S arrival times, time uncertainties, first-motions and waveforms for the catalog events to use for NLL-SSST-coherence relocation.

For initial NLL location (Fig. 8b), we use the 1D, P Parkfield – Middle Mountain (PMM) velocity profile with Vp/Vs=1.78 (Oppenheimer et al., 1993), linearly interpolated between depth nodes to form a smooth model. Following (Eberhart−Phillips & Michael, 1993; Zhang et al., 2009) we modify the model with a 5% increase [decrease] in velocity for station to the southwest [northeast] of the San Andreas Fault to account for a well defined, average velocity contrast across the fault.

For the Parkfield NLL-SSST relocations (Fig. 8c), we iteratively generating SSST corrections using the NCEDC catalog events and arrival data with smoothing distances, *D,* of 32, 16, 8, and 4km, spanning from the sequence size to larger than typical SSST cluster sizes and the target, sub-kilometer location precision. The quality criteria for an event location and station-phase to be used for calculating $\bar{R}$ are: 68% error-ellipsoid principle-axis half-width ≤5.0 km, root mean square of residuals (rms) ≤0.35 sec, number of readings ≥12, azimuth gap ≤ 135°, P residual ≤ 1.0 sec, S residual ≤ 2.0 sec. Note the dramatic improvement in clustering and organization of the NLL-SSST relocations (Fig. 8c) relative to the initial NLL locations (Fig. 8b).

For the Parkfield NLL-coherence relocations, we measure coherency using waveforms from vertical component channels from four nearby stations over and around the main seismicity: NC.PHA.--.EHZ, BK.PKD.--.HHZ, BP.RMNB.--.DP1, NC.PWK.--.EHZ. The waveforms are filtered from 2-10Hz in a window from 4 sec before the predicted P arrival to 4 sec after the predicted S arrival. Cross-correlation is applied between waveforms windows sliding from -2.0 to 2.0 sec, and the 0-1 stacking weight is set following Eq. 2 over coherency values from C_{min} = 0.5 to 1.0. This procedure is applied to the D = 4 km NLL-SSST relocations (Fig. 8c) for all event pairs with a maximum hypocenter separation of 5.0 km. The final NLL-SSST-coherence relocations are shown in Fig. 8d and are available as a CSV format table in DataSet S1.

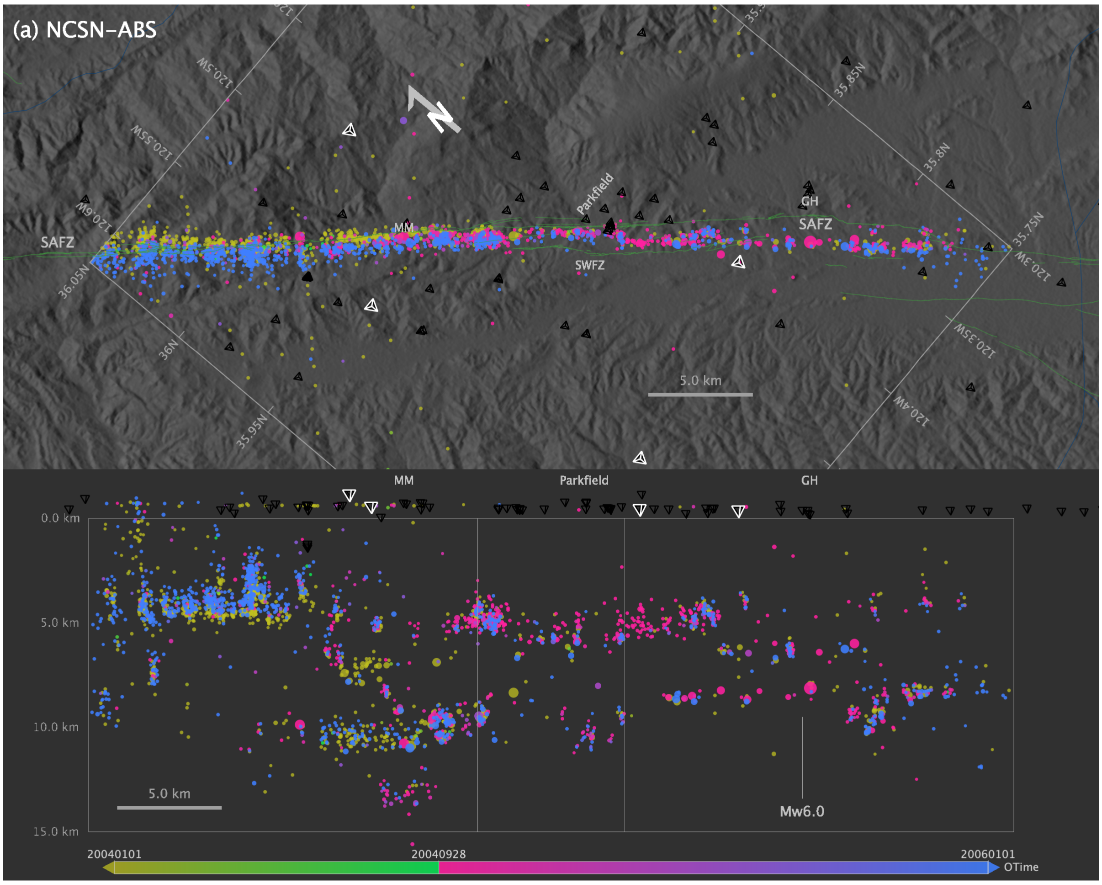
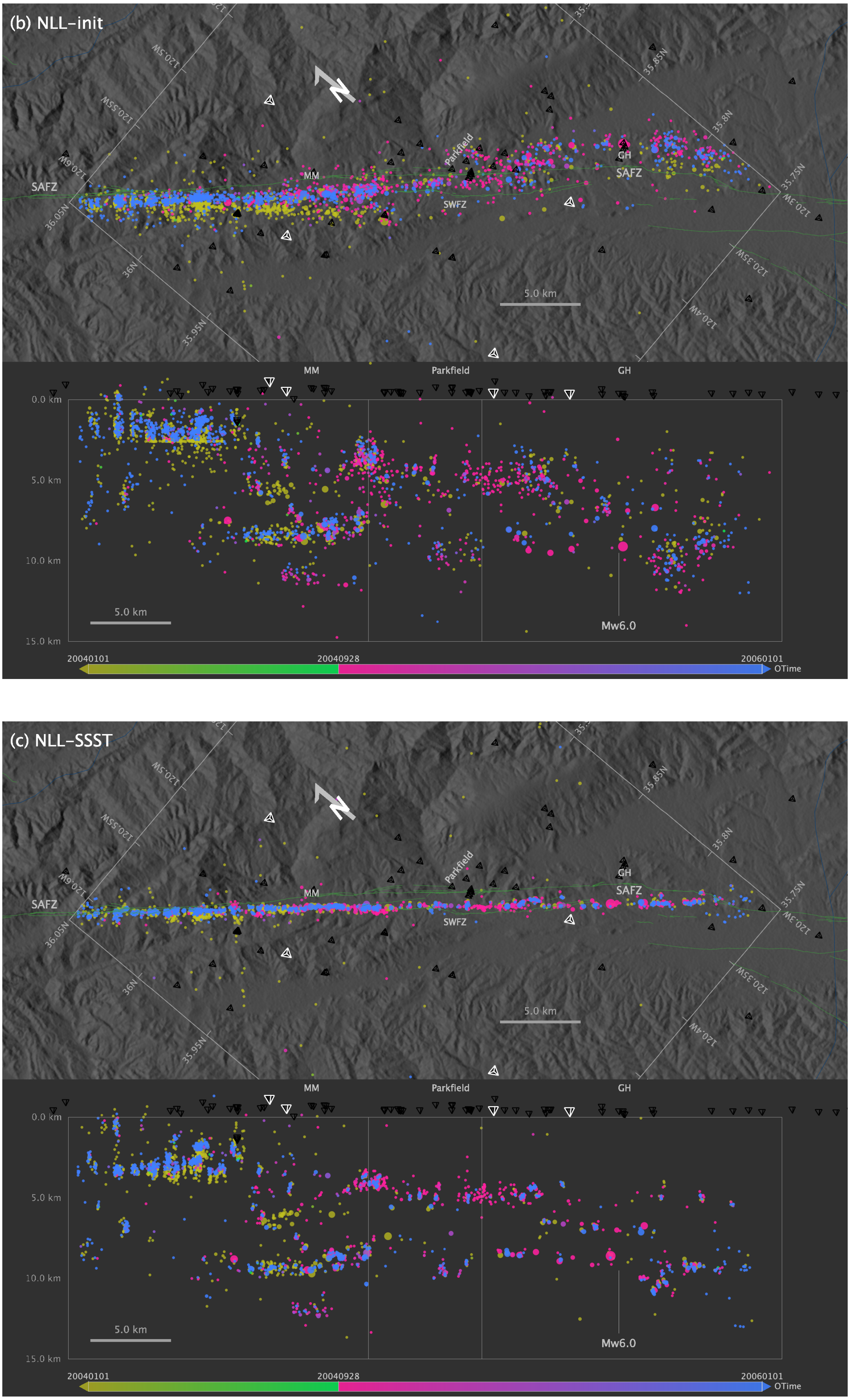
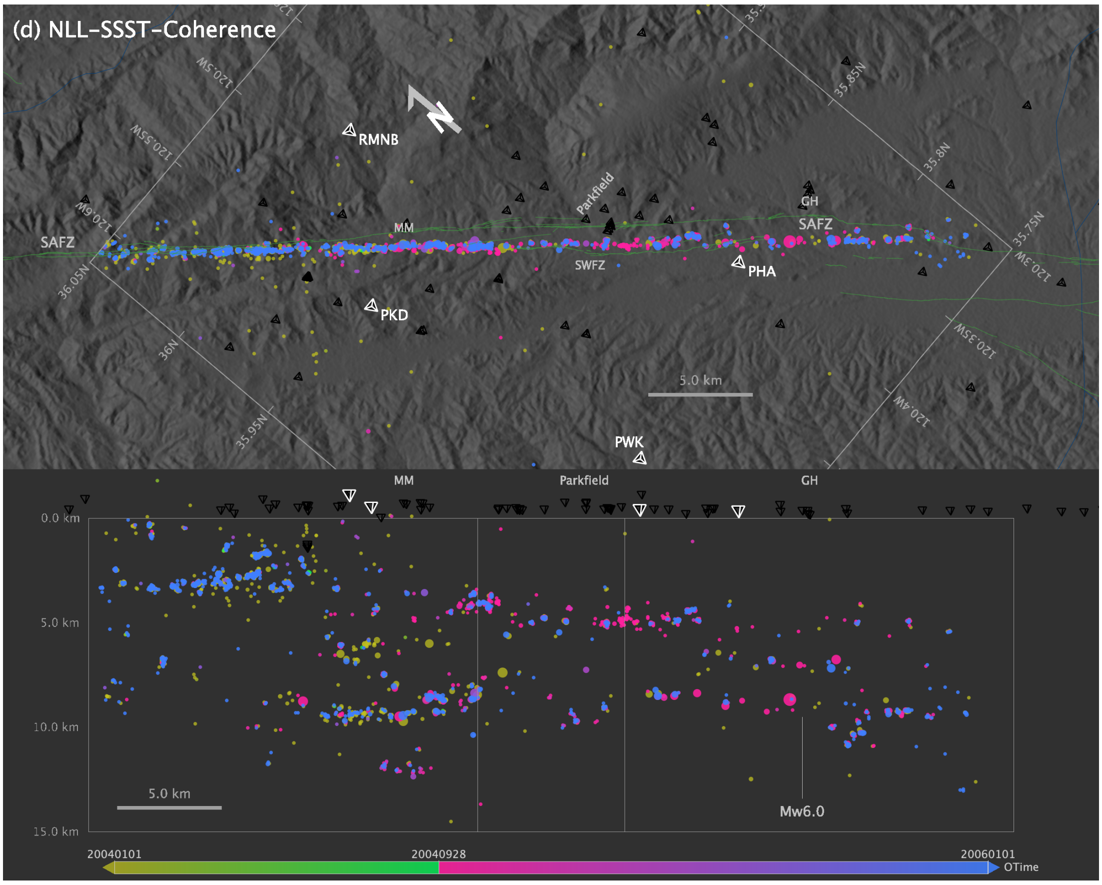
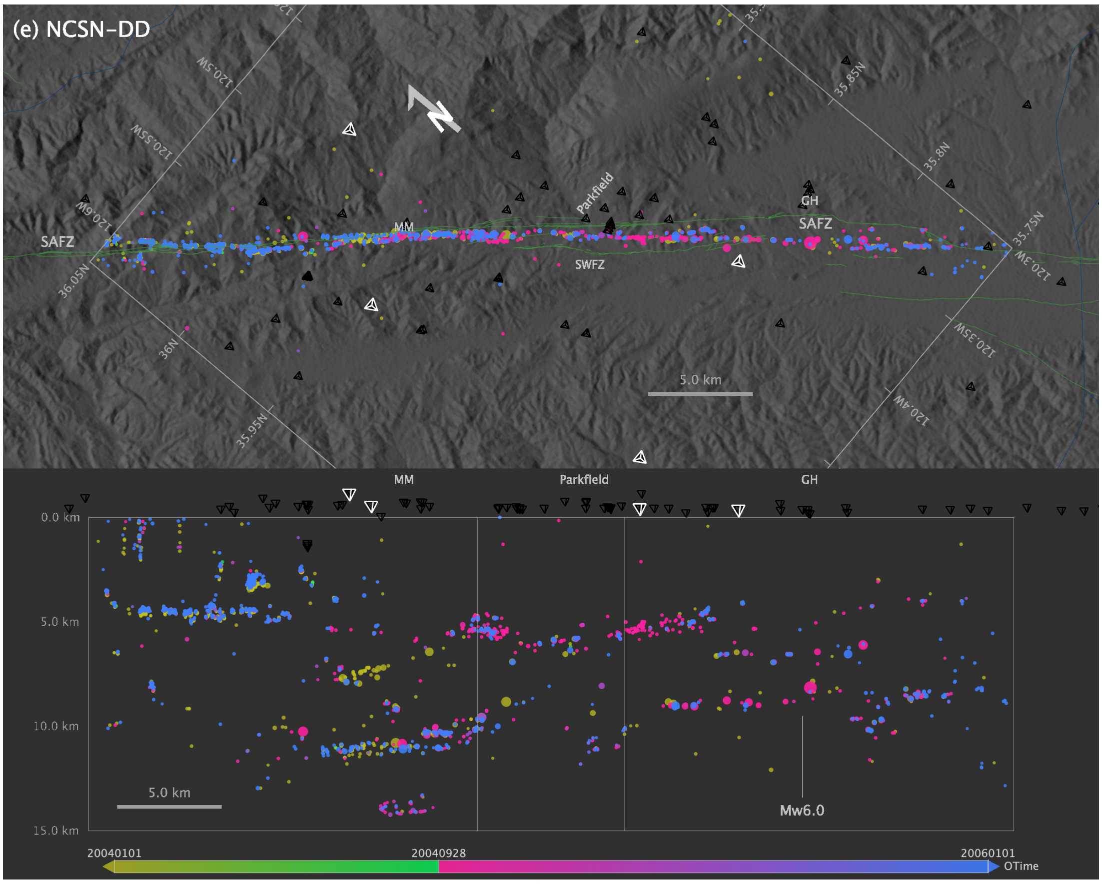


Fig. 8. 2004 M6.0 Parkfield, California earthquake sequence relocations. Map view and view from the southwest (N130W) of M ≥ 1.5, 1984-01-01 to 2020-21-31 hypocenters for the (a) NCSN-ABS, (b) initial NLL (NLL-init), (c) NLL-SSST D = 2 km relocations, (d) NLL-SSST-coherence, (e) NCSN-DD relocations. Hypocenter color shows origin time, symbol size is proportional to magnitude. NLL-SSST-coherence hypocenters in (d) are shifted randomly 0.1 km to avoid overlapping symbols. Inverted pyramids shows nearby seismic stations used for relocation; stations used for NLL-coherence waveform correlation emphasized in white and labelled with station codes in panel (d). Green lines show faults from the USGS Quaternary fault and fold database for the United States, with SAFZ denoting the San Andreas Fault Zone, SWFZ – Southwest Fracture Zone, MM – Middle Mountain, GH – Gold Hill. Background topography image from OpenTopgraphy.org.



(b) NLL-init
SAFZ
MM
Parkfield
SWFZ
GH
SAFZ
5.0 km
MM
Parkfield
GH
0.0 km
5.0 km
10.0 km
15.0 km
5.0 km
Mw6.0
20040101
20040928
20060101
OTime
(c) NLL-SSST
SAFZ
MM
Parkfield
SWFZ
GH
SAFZ
5.0 km
MM
Parkfield
GH
0.0 km
5.0 km
10.0 km
15.0 km
5.0 km
Mw6.0
20040101
20040928
20060101
OTime


(d) NLL-SSST-Coherence
RMNB
PKD
PHA
PWK
MM
Parkfield
GH
SAFZ
SWFZ
5.0 km
Mw6.0
0.0 km
5.0 km
10.0 km
15.0 km
20040101
20040928
20060101
OTime


(e) NCSN-DD
MM
Parkfield
GH
SAFZ
SWFZ
5.0 km
Mw6.0
0.0 km
5.0 km
10.0 km
15.0 km
20040101
20040928
20060101
OTime

The Parkfield NCSN-ABS (Fig. 8a) and NLL-SSST (Fig. 8c) relocations are similar, both showing a concentration of seismicity around a near vertical plane under the SAFZ, large scale, horizontal banding at depth, and vertical scatter in epicenters likely due to location error. There are clear differences and distortions in these sets of standard locations due to the use of different velocity models, station corrections and location procedures in the two catalogs—the NLL-SSST hypocenters are roughly 1km shallower [deeper] than the NCSN-ABS hypocenters in the northwestern 2/3 [southeastern 1/3] of the study zone and there is a notable shift in epicenter and depth of the M6.0 2004 mainshock hypocenter. The NLL-SSST locations also image a single, almost planer SAFZ across the study area, while the NCSN-ABS epicenters suggest an SAFZ composed of several near-planar segments with slight differences in strike and dip. All of these differences in NCSN-ABS and NLL-SSST standard locations pass to and are preserved in the NCSN-DD and NLL-SSST-coherence locations, respectively.

The Parkfield NLL-SSST-coherence (Fig. 8d) and NCSN-DD relocations (Fig. 8e) show similar large scale organization and smaller scale clustering of seismicity, and similar improvement relative to the NLL-SSST and NCSN-ABS locations. But in most areas the NCSN-DD locations define clearer concentration and lineation of hypocenters on an intermediate scale (~1-3km) and fewer isolated hypocenters than NLL-SSST-coherence. These differences are likely due to the explicit mapping in DD locations of high-precision, cross-correlation, differential times to relative hypocenter positions, while the NLL-SSST-coherence procedure performs a more rudimentary coalescence of NLL-SSST hypocenters for similar events.

The larger scale organization and smaller scale clustering of NLL-SSST-coherence in depth section (Fig. 8d) resembles closely the results of (C. Thurber et al., 2006) obtained with double-difference relocations in a 3D, tomographic velocity model. However, in contrast to (C. Thurber et al., 2006) and most other previous studies, and the NCSN-DD locations (Fig. 8e), the NLL-SSST-coherence seismicity falls on a single, near-vertical and almost planer surface across the study area (Fig. 8d). These seismicity patterns and results show NLL-SSST-coherence captures well features of the seismicity on all scales, and suggests real improvement in larger-scale location precision over the initial NLL locations and other studies, primarily due to corrections and resulting location shifts in the NLL-SSST procedure. Between Middle Mountain and Gold Hill, the near-vertical fault surface imaged by NLL-SSST-coherence underlies the surface trace of the Southwest Fracture Zone (SWFZ) and not the main SAF trace to the northeast, in agreement with the (C. Thurber et al., 2006) relocations and with observations of co-seismic slip on the SWFZ (Rymer et al., 2006). This largest scale position of epicenters, however, is mainly controlled by our imposed, 10% contrast across the

SAFZ in the model used for initial NLL location, and not by an overall shift in epicenters due to the NLL-SSST-coherence procedures.

5.2 2020 Mw 5.8 Lone Pine, California

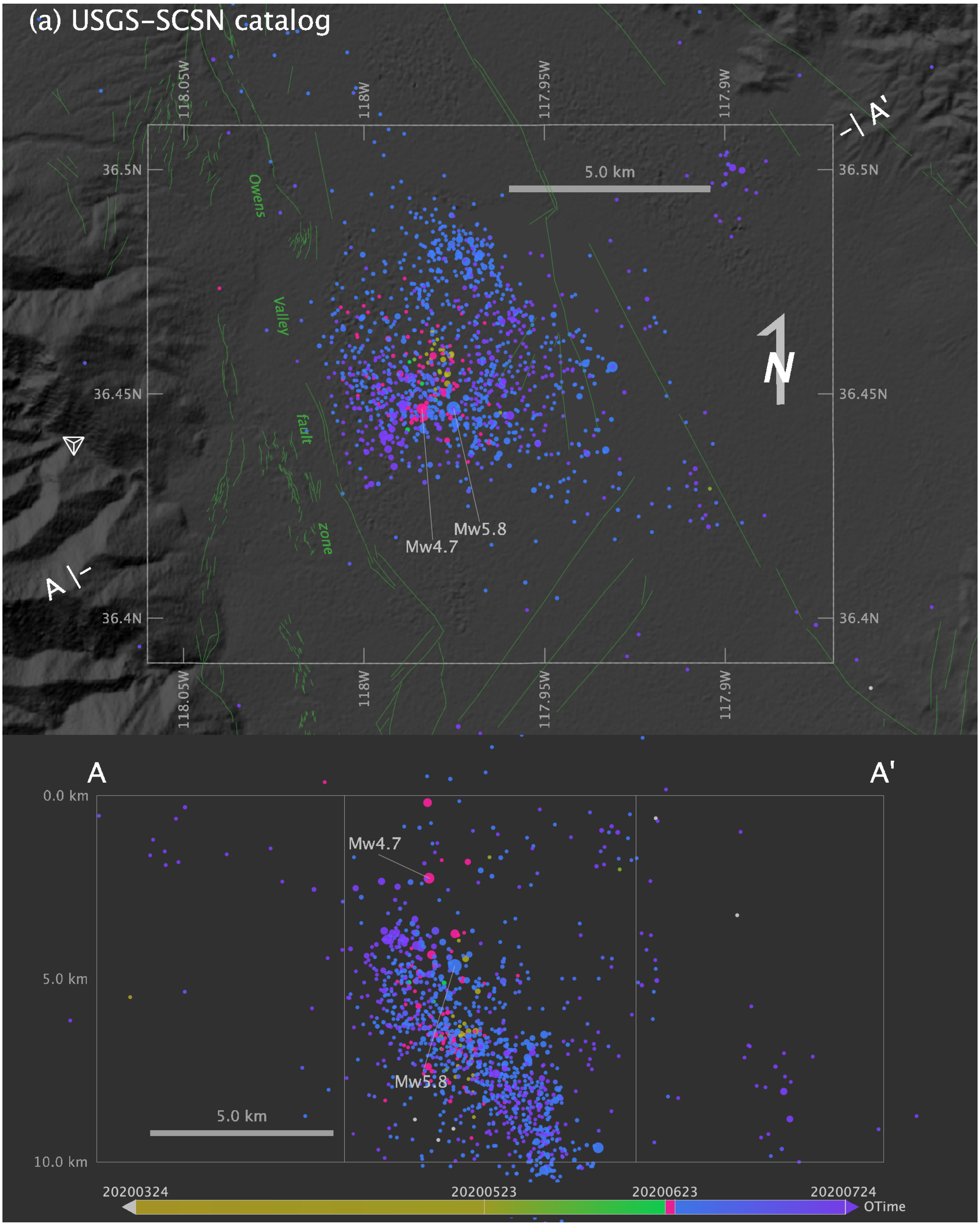
The 2020 Mw 5.8 Lone Pine, California earthquake sequence occurred along the Owens Valley fault zone (OVFZ) in eastern California (Fig. 7 and 8) near the southern end of the 1872 Mw ~7.5 Owens Valley earthquake rupture (Hauksson et al., 2020). The sequence includes mainly normal faulting events on an ~5 x 5 km, east-northeast dipping zone and has a distinct, multi-stage series of foreshocks including an Mw 4.6 event with aftershocks (Hauksson et al., 2020). The 2020 Lone Pine sequence was recorded by only a few seismic stations within ~20 km and one station at ~10 km from the mainshock, but no stations above the sequence, and thus demonstrates the performance of the NLL-SSST-coherence relocation procedure for the case of poor seismic network coverage.

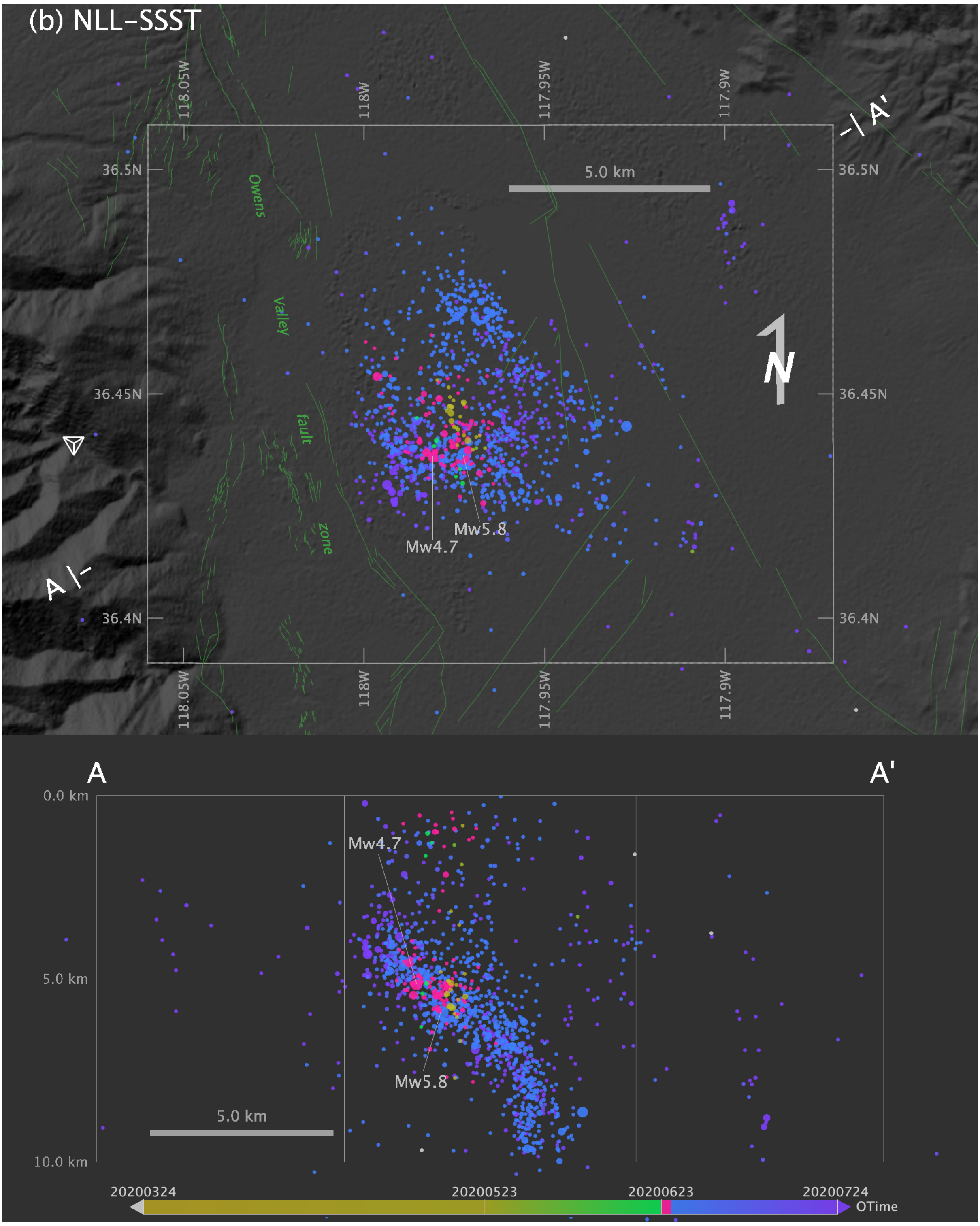
We obtain a catalog (USGS-SCSN catalog) of 1326 events (M 0.1-5.8) from 2020-01-01 to 2021-02-15 for the area of the Lone Pine sequence (within 20km of latitude 36.45º, longitude -118.00º) from USGS-earthquake hazards (Benz, 2017), with corresponding Southern California Seismic Network (SCSN) arrival phase types, times, time uncertainties, and first motions accessed from the Southern California Earthquake Data Center (SCEDC, 2013) and USGS-earthquake hazards. Waveforms for NLL-SSST-coherence relocation were obtained from SCEDC. To stabilize hypocenter depths for the three largest events, only the earliest two S arrival times (for the Mw 4.7, 2020-06-23 00:25 and Mw 5.8, 2020-06-24 17:40 events) or three S times (for the Mw 4.6, 2020-06-24 17:59 event) are used for location.

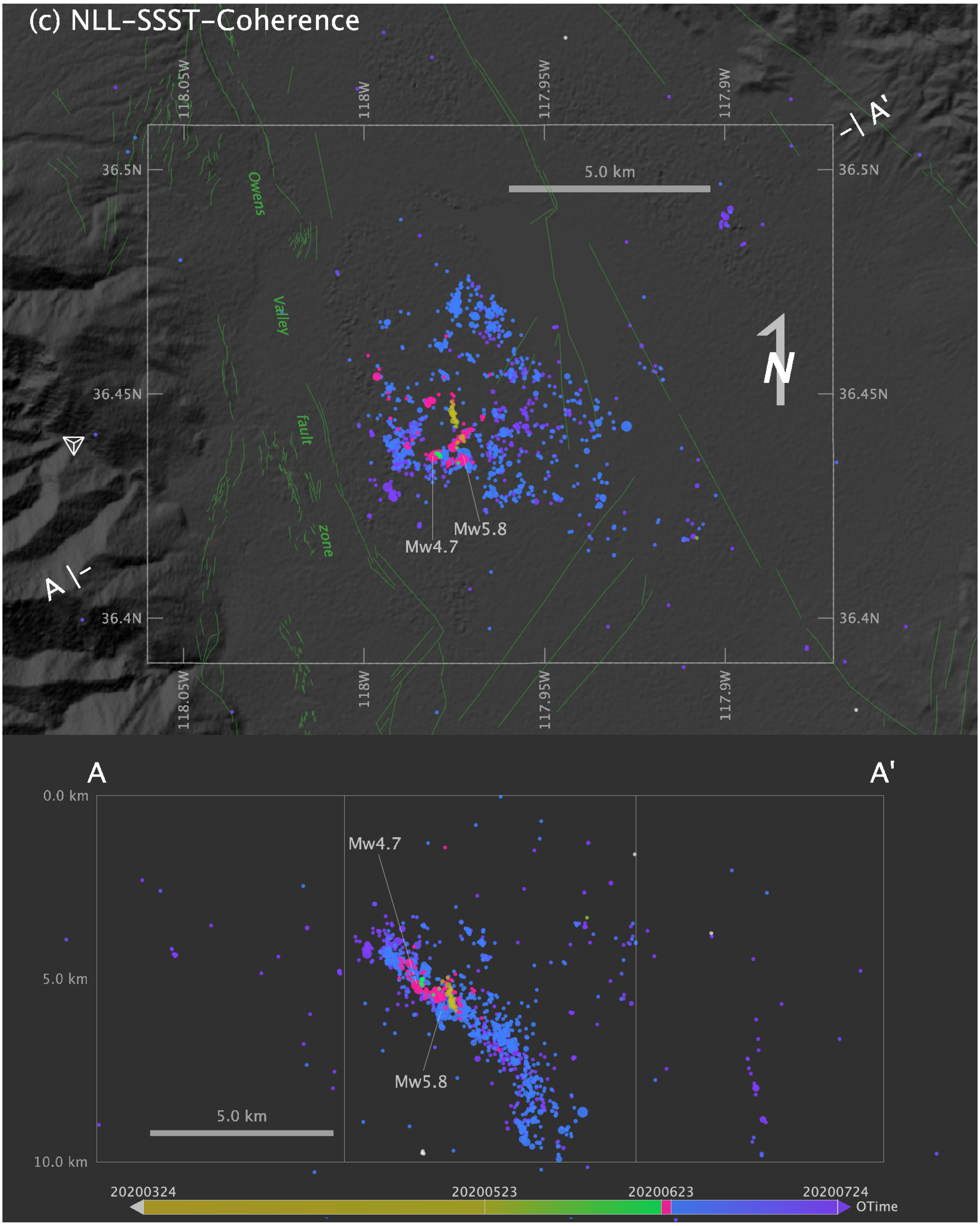
For initial NLL relocations we calculate travel-times in a smoothed version (KS-smooth) (Lomax, 2020b) of the KS seismic P-wave velocity model used for 2008 Mogul, northwest Nevada sequence relocations (von Seggern et al., 2015), with constant Vp/Vs=1.73 to obtain S travel-times.

For the Lone Pine NLL-SSST relocations, we iteratively generating SSST corrections using the USGS-SCSN catalog events and arrival data with smoothing distances, *D*, of 16, 8, 4, and 2km, spanning from larger than the sequence size to larger than typical SSST cluster sizes and the target, sub-kilometer location precision (Fig. 9b). The quality criteria for an event location and station-phase to be used for calculating $\bar{R}$ are: 68% error-ellipsoid principle-axis half-width ≤5.0 km, root mean square of residuals (rms) ≤0.35 sec, number of readings ≥12, azimuth gap ≤ 135°, P residual ≤ 1.0 sec, S residual ≤ 2.0 sec.

For the Lone Pine coherence relocations, we measure coherency using waveforms from vertical component channels from 7 nearby stations over a wide azimuth range: CI.CWC.--.HHZ, CI.DAW.--.HHZ, CI.CGO.--.HHZ, CI.WMF.--.HHZ, CE.44015.10.HNZ. Waveforms are filtered from 2-10Hz in a window from 4 sec before the predicted P arrival to 4 sec after the predicted S arrival. Cross-correlation is applied between waveforms windows sliding from -2.0 to 2.0 sec, and the 0-1 stacking weight is set following Eq. 2 over coherency values from C_{min} = 0.5 to 1.0. This procedure is applied to the D = 4 km NLL-SSST relocations (Fig. 9b), which exhibit more organization than the D = 2 km locations, for all event pairs with a maximum hypocenter separation of 5.0 km. The final NLL-SSST-coherence relocations are shown in Fig. 9c and Movie S1 and available as a CSV format table in DataSet S2.

662 
(a) USGS-SCSN catalog
118.05W
118W
117.95W
117.9W
36.5N
36.45N
36.4N
5.0 km
A
A'
Owens
Valley
fault
zone
N
Mw5.8
Mw4.7
0.0 km
5.0 km
10.0 km
20200324
20200523
20200623
20200724
OTime


(b) NLL-SSST
118.05W
118W
117.95W
117.9W
36.5N
36.45N
36.4N
5.0 km
Owens
Valley
fault
zone
Mw5.8
Mw4.7
A
A'
0.0 km
5.0 km
10.0 km
20200324
20200523
20200623
20200724
OTime

664 

(c) NLL-SSST-Coherence
118.05W
118W
117.95W
117.9W
36.5N
36.45N
36.4N
A'
A
5.0 km
Owens
Valley
fault
zone
N
Mw5.8
Mw4.7
0.0 km
5.0 km
10.0 km
20200324
20200523
20200623
20200724
OTime

Fig. 9. 2020 Mw 5.8 Lone Pine, California earthquake sequence relocations. Map and cross section (A–A′) of 2020-01-12 to 2021-02-15 hypocenters from the (a) USGS-SCSN catalog, (b) NLL-SSST *D* = 4 km relocations, (c) NLL-SSST-coherence relocations. Hypocenter color shows origin time, symbol size is proportional to magnitude. NLL-SSST-coherence hypocenters in (c) are shifted randomly 0.05 km to avoid overlapping symbols. White triangle to the west show the only nearby seismic station (CI.CWC) available for relocation. Green lines show faults from the USGS Quaternary fault and fold database for the United States. The hypocenter colors and the orientation of the cross section (A–A′) correspond to Figs. 2 and 3b, respectively in (Hauksson et al., 2020). Background topography image from OpenTopgraphy.org.

We compare the NLL-SSST-coherence relocations for the Lone Pine sequence to the two sets of location results presented in (Hauksson et al., 2020): one set from a waveform relocation procedure (Hauksson et al., 2012) which clusters events from the USGS-SCSN catalog and then uses differential travel-times for relative relocation within each cluster (HS catalog, 1052 events; (Hauksson et al., 2020; their fig. 3), and a second set from application of template-matching (Ross et al., 2018) to augment the USGS-SCSN catalog with numerous, newly detected events followed by relative relocation with cross-correlation, differential times using GrowClust (Trugman & Shearer, 2017) (QTM catalog, ~24;000 events; (Hauksson et al., 2020; their fig. 2).

The NLL-SSST-coherence hypocenters for the Lone Pine sequence (Fig. 9c) show a similar overall extent and shape, and similar areas of main clustering of seismicity and location of main events as the hypocenters from the USGS-SCSN catalog (Fig. 9a) and the HS and QTM epicenters of (Hauksson et al., 2020). On a smaller scale (< 1 km), the NLL-SSST-coherence epicenters show cluster shapes and lineations that roughly match most denser clouds of seismicity in the QTM and HS catalog, though the NLL-SSST-coherence epicenters are typically sparser with more concentrated clusters than those in the QTM and HS catalogs.

In (Hauksson et al., 2020) the depth distribution of events is only presented for the HS catalog, possibly because of a lack of constraint on depth in the QTM procedure due to lack of stations near or over the sequence. The HS catalog depth distribution in section view (Hauksson et al., 2020; their fig. 3) shows a broad zone of southeast dipping seismicity possibly composed of several more steeply southeast dipping segments. This distribution, along with fault-plane dips from moment tensor inversion, is interpreted by (Hauksson et al., 2020) to show volumetric deformation during the sequence. The NLL-SSST-coherence hypocenters in the same section view (Fig. 9c) show a narrower, northeast dipping, main zone of seismicity with, at its base, an apparently connected, near-vertically dipping zone. Numerous, shallow NLL-SSST hypocenters above the northeast dipping zone of seismicity and other scattered seismicity (Fig. 9b) are shifted as much as 5 km into the main dipping zone by the NLL-coherence procedure (Fig. 9c).

An oblique view from the northwest, nearly along the slip direction on the preferred, east-dipping fault-plane of the SCSN, mainshock moment-tensor (Fig. 10; Movie S1, last frame) gives clearer alignments of NLL-SSST-coherence hypocenters. This view suggests a complex “S”-shaped faulting structure composed at its top and base of sub-parallel sets of steeply southeast dipping sub-faults. These sets of sub-faults bracket a single faulting surface or narrower set of steeply dipping sub-faults that may have hosted much of the mainshock rupture. There are also several “satellite” structures parallel to the sub-faults and up to 5 km from the main faulting structure. This geometry agrees with the interpretation (Hauksson et al., 2020) of heterogeneous volumetric deformation, and furthermore suggests that aftershock and perhaps mainshock faulting occurs on sets of steeply northeast dipping, sub-parallel faults with oblique, normal and right-lateral slip. These apparent fault sets and the preferred, mainshock fault plane align with the Sierra Nevada frontal fault to the northwest of the sequence and west of Lone Pine.

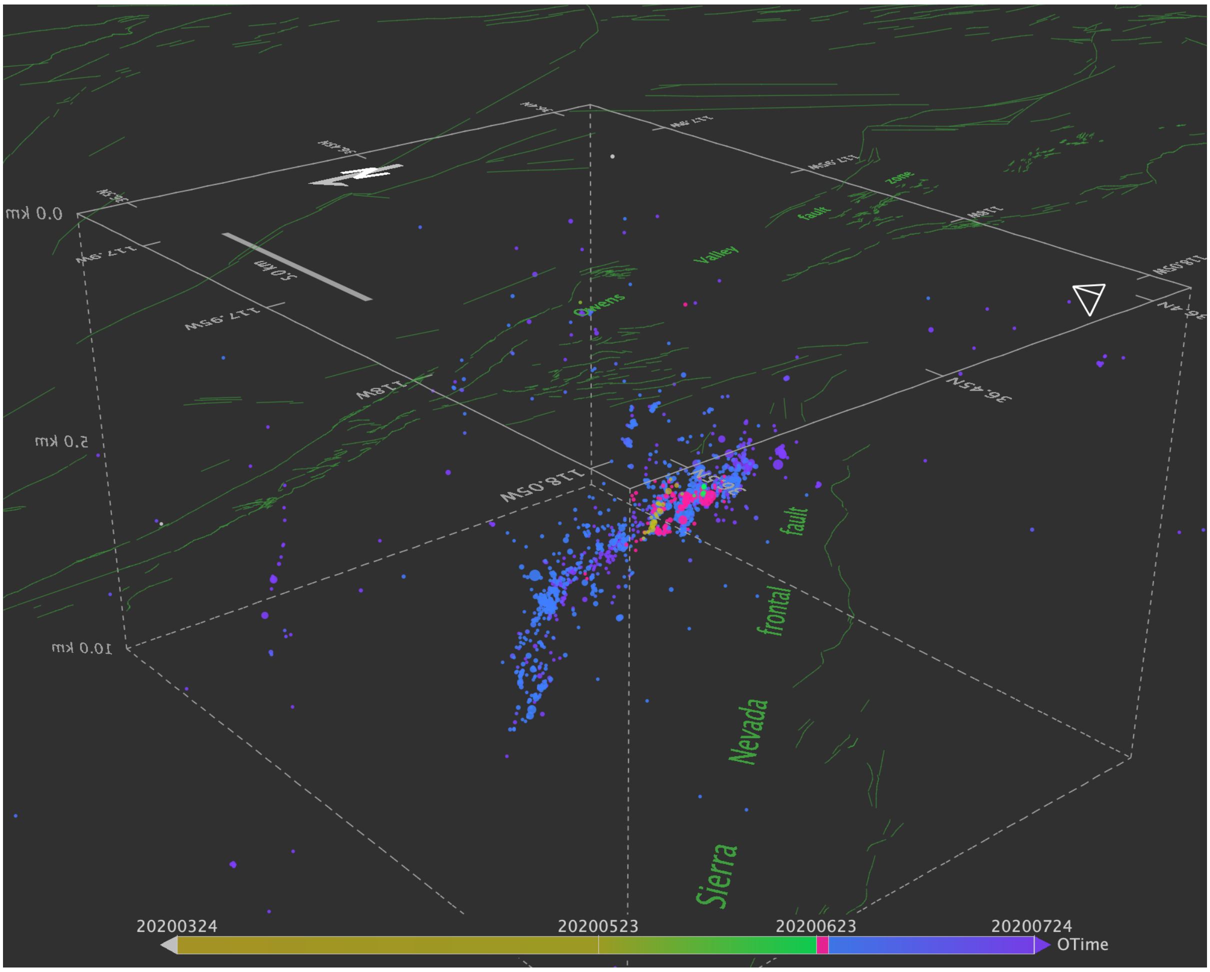


Fig. 10. 2020 Mw 5.8 Lone Pine, California earthquake sequence relocations. Oblique view from N50°E and plunging 25° of 2020-01-12 to 2021-02-15 hypocenters from the NLL-SSST-coherence relocations. Hypocenter color shows origin time, symbol size is proportional to magnitude. Other map elements as in Fig. 9. See also Movie S1, last frame.

The NLL-SSST-coherence results for the Lone Pine sequence (Fig. 9c) also show clearly a three-stage foreshock sequence starting in March 2020 with a first stage of seismicity along an ~1km long, north-south trend north of the eventual Mw 5.8 hypocenter (dark yellow events). A second stage begins on 22 June when seismic activity shifts to a small cluster (green events) ~1km west of the future Mw 5.8 hypocenter. A third stage begins at this cluster with the Mw 4.6 foreshock on 23 June followed by aftershock over an area of ~2 x 2 km (magenta events). 41 hours after the Mw 4.6 event the Mw 5.8 mainshock occurs on 24 June with aftershocks (blue events) covering an area of about 5 x 5 km. These results suggest a more concentrated and organized foreshock distribution that shown in the high-precision QTM catalog of (Hauksson et al., 2020).

6 Discussion

SSST and NLL-coherence together greatly increase relative location accuracy within a standard, arrival-time location framework. SSST does this by removing common-mode travel-time residuals at available stations as a function of hypocentral position, which reduces location bias between nearby events located with differing sets of stations or phase typess. NLL-coherence location achieves high precision by stacking probabilistic location PDF's of nearly co-located, multiplet events, as measured by waveform similarity. This stacking of PDF's effectively reduces aleatoric error and suppresses outliers in the underlying arrival times, while filling in missing arrival time data across multiplet events, resulting in a spatial coalescence of location for events with similar waveforms. The similarity of the NLL-SSST-coherence and double-difference, cross-correlation based, relative hypocenter positions for Parkfield at all but the smallest scales suggest that large and intermediate scale improvements in precision for relative location is possible solely through corrections such as SSST and coalescence of event multiplets guided by waveform similarity. However, our synthetic study, our comparison with double-difference relative location results for the 2004 Parkfield sequence, and results for the 2020 Monte Cristo sequence (Lomax, 2020b) show that this coalescence may tend to over-tightly cluster events at smallest scales, while potentially not resolving lineations and other extended features of the seismicity at this scale. The Parkfield results also suggest possible improvement in larger scale, relative location accuracy, primarily due to the NLL-SSST procedure.

In contrast to the coherence-weighted stacking of PDF's in NLL-coherence, cross-correlation based, relative location methods such as HypoDD or GrowClust achieve high to very high precision through explicit, inter-event, differential location involving inversion of precise arrival-time differences mapped into differences in distance along available rays. For relocation studies with good station coverage, and thus good ray coverage around the hypocenters, these relative location methods should achieve higher precision than NLL-SSST-coherence. However, for cases of poor station and ray coverage, NLL-SSST-coherence may produce higher relative location accuracy and better depth control than do cross-correlation based, relative location methods, as indicated by our results for the 2020 Lone Pine sequence and supported by the striking improvements of relative depths obtained for the Finland GT test.

NLL-coherence location requires waveform cross-correlation on only one or a few channels, while cross-correlation based, relative location procedures often use cross-correlation on P and S arrival windows for vertical and horizontal channels at all or most available stations. For example, for HypoDD relocation of 20 years of Northern California seismicity with around 500 stations, (Richards et al., 2006; Waldhauser & Schaff, 2008) perform about 26 billion P and S wave cross-correlations on 100Hz, vertical-component channels between all event pairs within 5 km out of 225,000 total events, giving a mean of about 230,000, 1-2 sec window cross-correlations per event. A similar procedure limited to 50 stations per event might still require around 23,000 cross-correlations per event. For application of NLL-coherence with 50 Hz waveforms from 4 stations and about 1000 events within 5 km, as in our Lone Pine example, about 2000, ~10 sec window cross-correlations per event are performed. NLL-coherence thus typically requires less computing time and resources than cross-correlation based, relative location methods. Excluding waveform download, the NLL-SSST-coherence processing pipeline in this study requires about 1.5 hours for Lone Pine (1326 events) and 4 hours for Parkfield (2793 events) on an 8 core, 3.6 GHz Intel® Core i9 workstation with shell or software parallelization of NLL location, NLL-SSST calculation and cross-correlations, but not using a GPU. This efficiency and the need for few waveform channels means NLL-SSST-coherence can provide rapidly high-precision, near-realtime relocation of new seismicity if the SSST corrections have pre-calculated from previous events in the area.

Additionally, since NLL-coherence requires waveforms on a single (vertical) channel from only one or a few stations, it can be applied with foreshocks and early events in a sequence before temporary stations are deployed, to older sequences where limited, digital, waveform data is available, or even to historical sequences if good quality analog records can be digitized. NLL-SSST-coherence relocation for over 12,000 events of 2020 Monte Cristo sequence (Lomax, 2020b) was successfully performed with only 2 waveform channels, one from a permanent station outside the sequence but available before and throughout the sequence, and another from a temporary station near the sequence and available from a few days after the mainshock.

The apparent tendency of NLL-coherence to over-tightly cluster events at smallest scales is an important issue, as it may limit the smallest scale at which NLL-coherence results should be interpreted. This scale may related to a fraction of the wavelength of the highest or dominant frequencies in the waveforms, e.g. 0.1 to 0.2 km for the California sequences presented here, but may also vary with the quality of the NLL-SSST locations and PDF's. We have also noticed that very nearby stations with simple waveforms (short S-P interval, little wave scattering) may have a tendency to produce high coherence values for events that are not nearby relative to target scales in a study. Strictly, this does not necessarily violate the ¼ wavelength rule, as the simple waveforms often have a relatively low dominant period. But this phenomena can lead to some false shifting of poorly constrained events into nearby event clusters. Further understanding of both these issues to improve the NLL-coherence procedure requires analysis and better understanding of the variation of waveform coherence with different inter-event and stations distances and azimuths, and with differing event sizes and source properties.

Cross-correlation based, relative location procedures require standard location results to form starting locations, to identify nearby, potential multiplet events, and to constrain the centroid of relative location hypocenters. NLL-SSST or NLL-SSST-coherence can be used to get an optimal set of standard, starting locations for applications of such relative location procedures. These optimal starting locations may be of particular importance for seismicity studies with poor station coverage or depth control. All standard and relative location methods remain subject to absolute location error and loss of accuracy due to error in the reference velocity model and insufficient station coverage. This absolute location error is carried into relative location results from the underlying, starting, standard location results.

5 Conclusions

We have introduced a new procedure (NLL-SSST-coherence) for high-precision, probabilistic, standard earthquake location which uses source-specific station corrections (NLL-SSST) and inter-event waveform similarity measured by cross-correlation coherence (NLL-coherence). NLL-SSST and NLL-coherence together greatly increase location precision over initial seismicity catalogs. We illustrated the behavior and performance of the NLL-SSST-coherence procedure through a synthetic example, ground-truth relocations, and relocation of two California earthquake sequences.

These results show that NLL-SSST-coherence location approaches the precision of cross-correlation based, relative location methods. Moreover, the results suggest that for sequences with few or no nearby stations NLL-SSST-coherence location may produce more stable and meaningful hypocenter locations, especially in depth, than cross-correlation based, relative location methods. NLL-SSST-coherence can also be used to get an optimal set of starting locations before application of relative location procedures.

NLL-SSST-coherence requires less computing time and resources than cross-correlation based, relative location methods, and can be applied with foreshocks and early events in a sequence before temporary station deployments and to older sequences with few waveform data.

Data and resources

The supporting information for this article includes a 3D, fly-around animation of the NLL-SSST-coherence relocations of Lone Pine seismicity (Movie S1). CSV tables of the final, NLL-

SSST-coherence earthquake relocation catalogs for Parkfield and Lone Pine are available at the Zenodo dataset repository (Lomax & Savvaidis, 2021b). An archive containing a directory structure, files and instructions for installing, configuring and running NLL-SSST-coherence for a subset of Parkfield events is available at the Zenodo dataset repository (Lomax & Savvaidis, 2021a).

The earthquake catalogs and corresponding phase arrival times, waveforms and metadata were accessed: for the Finland GT study from (Gibbons et al., 2020) and through Résif-Epos at https://www.resif.fr (last accessed April 2021) and http://doi.org/10.17616/R37Q06 (last accessed April 2021); for the 2004 Parkfield relocations through the Northern California Earthquake Data Center (NCEDC), http://doi.org/10.7932/NCEDC (last accessed April 2021); for the 2020 Lone Pine relocations through USGS-earthquake hazards available at https://www.usgs.gov (last accessed April 2021) and https://earthquake.usgs.gov/earthquakes/search (last accessed April 2021) and corresponding phase arrival times from SCEDC (2013) accessed from http://service.scedc.caltech.edu/fdsnws/event/1/ (last accessed April 2021) and https://earthquake.usgs.gov/earthquakes/search (last accessed April 2021). The USGS Quaternary fault and fold database for the United States is available at: https://www.usgs.gov/natural-hazards/earthquake-hazards/faults (last accessed April 2021).

All earthquake relocations were performed with NonLinLoc (Lomax et al., 2001, 2014); http://www.alomax.net/nlloc; https://github.com/alomax/NonLinLoc; last accessed April 2021). SeismicityViewer (http://www.alomax.net/software, last accessed April 2021) was used for 3D seismicity analysis and plotting, SeisGram2K (http://www.alomax.net/software, last accessed April 2021) was used for seismogram analysis and plotting, ObsPy (Beyreuther et al., 2010; Krischer et al., 2015), (http://obspy.org, last accessed April 2021) for reading seismicity catalogs and for coherence calculations, and LibreOffice (https://www.libreoffice.org, last accessed April 2021) for word processing, spreadsheet calculations and drawings.

Acknowledgments

We gratefully thank Pierre Henry, Alberto Michelini, Andrew Curtis, two reviewers and the editors for valuable discussions, comments and corrections. The development of NLL-SSST-coherence and the work reported here was supported by State of Texas through TexNet and by ALomax Scientific personal funds.

References

Aster, R. C., & Scott, J. (1993). Comprehensive characterization of waveform similarity in microearthquake data sets. *Bulletin of the Seismological Society of America*, *83*(4), 1307–1314. Retrieved from https://pubs.geoscienceworld.org/ssa/bssa/article/83/4/1307/119769/Comprehensive-characterization-of-waveform

Bakun, W. H., Aagaard, B., Dost, B., Ellsworth, W. L., Hardebeck, J. L., Harris, R. A., et al. (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature*, *437*(7061), 969–974. https://doi.org/10.1038/nature04067

Benz, H. (2017). Building a National Seismic Monitoring Center: NEIC from 2000 to the Present. *Seismological Research Letters*, *88*(2A), 265–269. https://doi.org/10.1785/0220170023

Bernardi, F., Lomax, A., Michelini, A., Lauciani, V., Piatanesi, A., & Lorito, S. (2015). Appraising the Early-est earthquake monitoring system for tsunami alerting at the Italian Candidate Tsunami Service Provider. *Natural Hazards and Earth System Sciences*, *15*(9), 2019–2036. https://doi.org/10.5194/nhess-15-2019-2015

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, *81*(3), 530–533. https://doi.org/10.1785/gssrl.81.3.530

Billings, S. D., Sambridge, M. S., & Kennett, B. L. N. (1994). Errors in hypocenter location: Picking, model, and magnitude dependence. *Bulletin of the Seismological Society of America*, *84*(6), 1978–1990.

Bondár, I., & McLaughlin, K. L. (2009). A New Ground Truth Data Set For Seismic Studies. *Seismological Research Letters*, *80*(3), 465–472. https://doi.org/10.1785/gssrl.80.3.465

Buehler, J. S., & Shearer, P. M. (2016). Characterizing Earthquake Location Uncertainty in North America Using Source–Receiver Reciprocity and USArrayShort. *Bulletin of the Seismological Society of America*, *106*(5), 2395–2401. https://doi.org/10.1785/0120150173

Cattaneo, M., Augliera, P., Spallarossa, D., & Eva, C. (1997). Reconstruction of seismogenetic structures by multiplet analysis: An example of Western Liguria, Italy. *Bulletin of the Seismological Society of America*, *87*(4), 971–986. Retrieved from https://pubs.geoscienceworld.org/ssa/bssa/article-abstract/87/4/971/120243/Reconstruction-of-seismogenetic-structures-by

Chamberlain, C. J., Hopp, C. J., Boese, C. M., Warren‐Smith, E., Chambers, D., Chu, S. X., et al. (2018). EQcorrscan: Repeating and Near‐Repeating Earthquake Detection and Analysis in Python. *Seismological Research Letters*, *89*(1), 173–181. https://doi.org/10.1785/0220170151

Darold, A., Holland, A., Chen, C., & Youngblood, A. (2014). *Preliminary Analysis of Seismicity Near Eagleton 1-29, Carter County, July 2014* (Open-File Report No. OF2-2014). Oklahoma Geolical Survey. Retrieved from http://ogs.ou.edu/docs/openfile/OF2-2014.pdf

Eberhart‐Phillips, D., & Michael, A. J. (1993). Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield Region, central California. *Journal of Geophysical Research: Solid Earth*, *98*(B9), 15737–15758. https://doi.org/10.1029/93JB01029

Ellsworth, W. L. (1975). Bear Valley, California, earthquake sequence of February-March 1972. *Bulletin of the Seismological Society of America*, *65*(2), 483–506. Retrieved from https://pubs.geoscienceworld.org/ssa/bssa/article-abstract/65/2/483/117416/Bear-Valley-California-earthquake-sequence-of

Fehler, M., Phillips, W. S., House, L., Jones, R. H., Aster, R., & Rowe, C. (2000). Improved Relative Locations of Clustered Earthquakes Using Constrained Multiple Event Location. *Bulletin of the Seismological Society of America*, *90*(3), 775–780. https://doi.org/10.1785/0119990095

Ferretti, G. (2005). An Improved Method for the Recognition of Seismic Families: Application to the Garfagnana-Lunigiana Area, Italy. *Bulletin of the Seismological Society of America*, *95*(5), 1903–1915. https://doi.org/10.1785/0120040078

Font, Y., Kao, H., Lallemand, S., Liu, C.-S., & Chiao, L.-Y. (2004). Hypocentre determination offshore of eastern Taiwan using the Maximum Intersection method. *Geophysical Journal International*, *158*(2), 655–675. https://doi.org/10.1111/j.1365-246X.2004.02317.x

Fragoso, T. M., & Neto, F. L. (2018). Bayesian model averaging: A systematic review and conceptual classification. *International Statistical Review*, *86*(1), 1–28. https://doi.org/10.1111/insr.12243

Frémont, M.-J., & Malone, S. D. (1987). High precision relative locations of earthquakes at Mount St. Helens, Washington. *Journal of Geophysical Research: Solid Earth*, *92*(B10), 10223–10236. https://doi.org/10.1029/JB092iB10p10223

Frohlich, C. (1979). An efficient method for joint hypocenter determination for large groups of earthquakes. *Computers & Geosciences*, *5*(3), 387–389. https://doi.org/10.1016/0098-3004(79)90034-7

Geller, R. J., & Mueller, C. S. (1980). Four similar earthquakes in central California. *Geophysical Research Letters*, *7*(10), 821–824. https://doi.org/10.1029/GL007i010p00821

Gibbons, S. J., Pabian, F., Näsholm, S. P., Kværna, T., & Mykkeltveit, S. (2017). Accurate relative location estimates for the North Korean nuclear tests using empirical slowness corrections. *Geophysical Journal International*, *208*(1), 101–117. https://doi.org/10.1093/gji/ggw379

Gibbons, S. J., Kværna, T., Tiira, T., & Kozlovskaya, E. (2020). A benchmark case study for seismic event relative location. *Geophysical Journal International*, *223*(2), 1313–1326. https://doi.org/10.1093/gji/ggaa362

Goertz−Allmann, B. P., Gibbons, S. J., Oye, V., Bauer, R., & Will, R. (2017). Characterization of induced seismicity patterns derived from internal structure in event clusters. *Journal of Geophysical Research: Solid Earth*, *122*(5), 3875–3894. https://doi.org/10.1002/2016JB013731

Gomberg, J. S., Shedlock, K. M., & Roecker, S. W. (1990). The effect of S-wave arrival times on the accuracy of hypocenter estimation. *Bulletin of the Seismological Society of America*, *80*(6A), 1605–1628.

Got, J.-L., Fréchet, J., & Klein, F. W. (1994). Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea. *Journal of Geophysical Research*, *99*(B8), 15375. https://doi.org/10.1029/94JB00577

Hardebeck, J., & Husen, S. (2010). Earthquake location accuracy. https://doi.org/10.5078/CORSSA-55815573

Hauksson, E., Yang, W., & Shearer, P. M. (2012). Waveform Relocated Earthquake Catalog for Southern California (1981 to June 2011). *Bulletin of the Seismological Society of America*, *102*(5), 2239–2244. https://doi.org/10.1785/0120120010

Hauksson, E., Olson, B., Grant, A., Andrews, J. R., Chung, A. I., Hough, S. E., et al. (2020). The Normal−Faulting 2020 Mw 5.8 Lone Pine, Eastern California, Earthquake Sequence. *Seismological Research Letters*. https://doi.org/10.1785/0220200324

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: a tutorial (with comments by M. Clyde, David Draper and E. I. George, and a rejoinder by the authors. *Statistical Science*, *14*(4), 382–417. https://doi.org/10.1214/ss/1009212519

Ishida, M., & Kanamori, H. (1978). The foreshock activity of the 1971 San Fernando earthquake, California. *Bulletin of the Seismological Society of America*, *68*(5), 1265–1279. Retrieved from https://pubs.geoscienceworld.org/ssa/bssa/article/68/5/1265/117791/The-foreshock-activity-of-the-1971-San-Fernando

Ito, A. (1985). High Resolution Relative Hypocenters of Similar Earthquakes by Cross-Spectral Analysis Method. *Journal of Physics of the Earth*, *33*(4), 279–294. https://doi.org/10.4294/jpe1952.33.279

Jones, R. H., & Stewart, R. C. (1997). A method for determining significant structures in a cloud of earthquakes. *Journal of Geophysical Research: Solid Earth*, *102*(B4), 8245–8254. https://doi.org/10.1029/96JB03739

Kamer, Y., Ouillon, G., Sornette, D., & Wössner, J. (2015). Condensation of earthquake location distributions: Optimal spatial information encoding and application to multifractal analysis of south Californian seismicity. *Physical Review E*, *92*(2), 022808. https://doi.org/10.1103/PhysRevE.92.022808

Kennett, B. L. N., Engdahl, E. R., & Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, *122*(1), 108–124. https://doi.org/10.1111/j.1365-246X.1995.tb03540.x

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., & Wassermann, J. (2015). ObsPy: a bridge for seismology into the scientific Python ecosystem. *Computational Science & Discovery*, *8*(1), 014003. https://doi.org/10.1088/1749-4699/8/1/014003

Latorre, D., Mirabella, F., Chiaraluce, L., Trippetta, F., & Lomax, A. (2016). Assessment of earthquake locations in 3−D deterministic velocity models: A case study from the

Altotiberina Near Fault Observatory (Italy). *Journal of Geophysical Research: Solid Earth*, *121*(11), 8113–8135. https://doi.org/10.1002/2016JB013170

Li, L., Tan, J., Schwarz, B., Staněk, F., Poiata, N., Shi, P., et al. (2020). Recent Advances and Challenges of Waveform−Based Seismic Location Methods at Multiple Scales. *Reviews of Geophysics*, *58*(1). https://doi.org/10.1029/2019RG000667

Lin, G., & Shearer, P. (2005). Tests of relative earthquake location techniques using synthetic data. *Journal of Geophysical Research: Solid Earth*, *110*(B4). https://doi.org/10.1029/2004JB003380

Lin, G., Shearer, P. M., & Hauksson, E. (2007). Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005. *Journal of Geophysical Research*, *112*(B12), B12309. https://doi.org/10.1029/2007JB004986

Lomax, A. (2005). A Reanalysis of the Hypocentral Location and Related Observations for the Great 1906 California Earthquake. *Bulletin of the Seismological Society of America*, *95*(3), 861–877. https://doi.org/10.1785/0120040141

Lomax, A. (2008). Location of the Focus and Tectonics of the Focal Region of the California Earthquake of 18 April 1906. *Bulletin of the Seismological Society of America*, *98*(2), 846–860. https://doi.org/10.1785/0120060405

Lomax, A. (2020a). Absolute Location of 2019 Ridgecrest Seismicity Reveals a Shallow Mw 7.1 Hypocenter, Migrating and Pulsing Mw 7.1 Foreshocks, and Duplex Mw 6.4 Ruptures. *Bulletin of the Seismological Society of America*, *110*(4), 1845–1858. https://doi.org/10.1785/0120200006

Lomax, A. (2020b). The 2020 Mw 6.5 Monte Cristo Range, Nevada earthquake: relocated seismicity shows rupture of a complete shear-crack system. https://doi.org/10.31223/X5X015

Lomax, A., & Savvaidis, A. (2019). Improving Absolute Earthquake Location in West Texas Using Probabilistic, Proxy Ground-Truth Station Corrections. *Journal of Geophysical Research: Solid Earth*, *124*(11), 11447–11465. https://doi.org/10.1029/2019JB017727

Lomax, A., & Savvaidis, A. (2021a). *Files and instructions for running NLL-SSST-coherence for a subset of Parkfield events*. Zenodo. https://doi.org/10.5281/zenodo.4756709

Lomax, A., & Savvaidis, A. (2021b). NLL-SSST-coherence earthquake relocation catalogs for the Parkfield and Lone Pine, California earthquake sequence. https://doi.org/10.5281/zenodo.4756181

Lomax, A., Virieux, J., Volant, P., & Berge-Thierry, C. (2000). Probabilistic Earthquake Location in 3D and Layered Models. In C. H. Thurber & N. Rabinowitz (Eds.), *Advances in Seismic Event Location* (Vol. 18, pp. 101–134). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-015-9536-0_5

Lomax, A., Zollo, A., Capuano, P., & Virieux, J. (2001). Precise, absolute earthquake location under Somma-Vesuvius volcano using a new three-dimensional velocity model. *Geophysical Journal International*, *146*(2), 313–331. https://doi.org/10.1046/j.0956-540x.2001.01444.x

Lomax, A., Satriano, C., & Vassallo, M. (2012). Automatic Picker Developments and Optimization: FilterPicker--a Robust, Broadband Picker for Real-Time Seismic Monitoring and Earthquake Early Warning. *Seismological Research Letters*, *83*(3), 531–540. https://doi.org/10.1785/gssrl.83.3.531

Lomax, A., Michelini, A., & Curtis, A. (2014). Earthquake Location, Direct, Global-Search Methods. In R. A. Meyers (Ed.), *Encyclopedia of Complexity and Systems Science* (pp. 1–33). New York, NY: Springer New York. https://doi.org/10.1007/978-3-642-27737-5_150-2

Matoza, R. S., Shearer, P. M., Lin, G., Wolfe, C. J., & Okubo, P. G. (2013). Systematic relocation of seismicity on Hawaii Island from 1992 to 2009 using waveform cross correlation and cluster analysis. *Journal of Geophysical Research: Solid Earth*, *118*(5), 2275–2288. https://doi.org/10.1002/jgrb.50189

Michele, M., Chiaraluce, L., Stefano, R. D., & Waldhauser, F. (2020). Fine-Scale Structure of the 2016–2017 Central Italy Seismic Sequence From Data Recorded at the Italian National Network. *Journal of Geophysical Research: Solid Earth*, *125*(4), e2019JB018440. https://doi.org/10.1029/2019JB018440

Michelini, A., & Lomax, A. (2004). The effect of velocity structure errors on double-difference earthquake location. *Geophysical Research Letters*, *31*(9), n/a-n/a. https://doi.org/10.1029/2004GL019682

Michelini, A., & McEvilly, T. V. (1991). Seismological studies at Parkfield. I. Simultaneous inversion for velocity structure and hypocenters using cubic B-splines parameterization. *Bulletin of the Seismological Society of America*, *81*(2), 524–552. Retrieved from https://pubs.geoscienceworld.org/ssa/bssa/article-abstract/81/2/524/119448/Seismological-studies-at-Parkfield-I-Simultaneous

Myers, S. C. (2000). Improving Sparse Network Seismic Location with Bayesian Kriging and Teleseismically Constrained Calibration Events. *Bulletin of the Seismological Society of America*, *90*(1), 199–211. https://doi.org/10.1785/0119980171

Nadeau, R., Antolik, M., Johnson, P. A., Foxall, W., & McEvilly, T. V. (1994). Seismological studies at Parkfield III: microearthquake clusters in the study of fault-zone dynamics. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, *31*(6), 271. https://doi.org/10.1016/0148-9062(94)90077-9

Nakamura, Y. (1978). A1 moonquakes: source distribution and mechanism. *Proc. Lunar and Planet. Sci. Conf. 9th*, 3589–3607. Retrieved from https://ci.nii.ac.jp/naid/10006236523/en/

Nicholson, T., Clarke, D., & Townend, J. (2008). Regional earthquake location using empirical traveltimes in a region of strong lateral velocity heterogeneity. *Geophysical Journal International*, *175*(2), 560–570. https://doi.org/10.1111/j.1365-246X.2008.03858.x

Nooshiri, N., Saul, J., Heimann, S., Tilmann, F., & Dahm, T. (2017). Revision of earthquake hypocentre locations in global bulletin data sets using source-specific station terms. *Geophysical Journal International*, *208*(2), 589–602. https://doi.org/10.1093/gji/ggw405

Oppenheimer, D. H., Klein, F. W., Eaton, J. P., & Lester, F. W. (1993). *The Northern California Seismic Network bulletin, January-December 1992* (No. 93–578). *Open-File Report*. U.S. Geological Survey,. https://doi.org/10.3133/ofr93578

Pavlis, G. L. (1986). Appraising earthquake hypocenter location errors: A complete, practical approach for single-event locations. *Bulletin of the Seismological Society of America*, *76*(6), 1699–1717.

Pavlis, G. L., & Hokanson, N. B. (1985a). Separated earthquake location. *Journal of Geophysical Research: Solid Earth*, *90*(B14), 12777–12789. https://doi.org/10.1029/JB090iB14p12777

Pavlis, G. L., & Hokanson, N. B. (1985b). Separated earthquake location. *Journal of Geophysical Research: Solid Earth*, *90*(B14), 12777–12789. https://doi.org/10.1029/JB090iB14p12777

Podvin, P., & Lecomte, I. (1991). Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools. *Geophysical Journal International*, *105*(1), 271–284. https://doi.org/10.1111/j.1365-246X.1991.tb03461.x

Poupinet, G., Glangeaud, F., & Cote, P. (1982). P-Time delay measurement of a doublet of microearthquakes. In *ICASSP '82. IEEE International Conference on Acoustics, Speech, and Signal Processing* (Vol. 7, pp. 1516–1519). Paris, France: Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICASSP.1982.1171796

Poupinet, G., Ellsworth, W. L., & Frechet, J. (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras Fault, California. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 5719–5731. https://doi.org/10.1029/JB089iB07p05719

Richards, P. G., Waldhauser, F., Schaff, D., & Kim, W.-Y. (2006). The Applicability of Modern Methods of Earthquake Location. *Pure and Applied Geophysics*, *163*(2–3), 351–372. https://doi.org/10.1007/s00024-005-0019-5

Richards-Dinger, K. B., & Shearer, P. M. (2000). Earthquake locations in southern California obtained using source-specific station terms. *Journal of Geophysical Research: Solid Earth*, *105*(B5), 10939–10960. https://doi.org/10.1029/2000JB900014

Ritzwoller, M. H., Shapiro, N. M., Levshin, A. L., Bergman, E. A., & Engdahl, E. R. (2003). Ability of a global three-dimensional model to locate regional events: GLOBAL MODEL AND REGIONAL EVENTS. *Journal of Geophysical Research: Solid Earth*, *108*(B7). https://doi.org/10.1029/2002JB002167

Ross, Z. E., Meier, M.-A., & Hauksson, E. (2018). P Wave Arrival Picking and First-Motion Polarity Determination With Deep Learning. *Journal of Geophysical Research: Solid Earth*, *123*(6), 5120–5129. https://doi.org/10.1029/2017JB015251

Rowe, C. A., Aster, R. C., Phillips, W. S., Jones, R. H., Borchers, B., & Fehler, M. C. (2002). Using Automated, High-precision Repicking to Improve Delineation of Microseismic Structures at the Soultz Geothermal Reservoir. *Pure Appl. Geophys.*, *159*, 34.

Rubin, A. M., Gillard, D., & Got, J.-L. (1999). Streaks of microearthquakes along creeping faults. *Nature*, *400*(6745), 635–641. https://doi.org/10.1038/23196

Ryaboy, V., Baumgardt, D. R., Firbas, P., & Dainty, A. M. (2001). Application of 3-D Crustal and Upper Mantle Velocity Model of North America for Location of Regional Seismic Events. *Pure and Applied Geophysics*, *158*(1), 79–103. https://doi.org/10.1007/PL00001169

Rymer, M. J., Tinsley, J. C., Treiman, J. A., Arrowsmith, J. R., Clahan, K. B., Rosinski, A. M., et al. (2006). Surface Fault Slip Associated with the 2004 Parkfield, California, Earthquake. *Bulletin of the Seismological Society of America*, *96*(4B), S11–S27. https://doi.org/10.1785/0120050830

Savvaidis, A., Young, B., Huang, G. D., & Lomax, A. (2019). TexNet: A Statewide Seismological Network in Texas. *Seismological Research Letters*, *90*(4), 1702–1715. https://doi.org/10.1785/0220180350

SCEDC. (2013). Southern California Earthquake Data Center. Caltech. Dataset. https://doi.org/10.7914/SN/CI

Schoenball, M., & Ellsworth, W. L. (2017). Waveform−Relocated Earthquake Catalog for Oklahoma and Southern Kansas Illuminates the Regional Fault Network. *Seismological Research Letters*, *88*(5), 1252–1258. https://doi.org/10.1785/0220170083

von Seggern, D. H., Anderson, J. G., Tibuleac, I. M., & Biasi, G. P. (2015). Double-Difference Location and Ground-Truth Classification of the 2008 Mogul, Nevada, Very Shallow Earthquake Sequence. *Seismological Research Letters*, *86*(1), 146–157. https://doi.org/10.1785/0220140121

Shearer, P. M. (1997). Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence. *Journal of Geophysical Research: Solid Earth*, *102*(B4), 8269–8283. https://doi.org/10.1029/96JB03228

Shearer, P. M. (2005). Southern California Hypocenter Relocation with Waveform Cross-Correlation, Part 2: Results Using Source-Specific Station Terms and Cluster Analysis. *Bulletin of the Seismological Society of America*, *95*(3), 904–915. https://doi.org/10.1785/0120040168

Stauder, W., & Ryall, A. (1967). Spatial distribution and source mechanism of microearthquakes in Central Nevada. *Bulletin of the Seismological Society of America*, *57*(6), 1317–1345.

Tarantola, A., & Valette, B. (1982). Inverse Problems = Quest for Information. *Journal of Geophysics*, 16.

Thorbjarnardottir, B. S., & Pechmann, J. C. (1987). Constraints on relative earthquake locations from cross-correlation of waveforms. *Bulletin of the Seismological Society of America*, *77*(5), 1626–1634.

Thurber, C., & Rabinowitz, N. (Eds.). (2000). *Advances in Seismic Event Location*. Springer Netherlands. https://doi.org/10.1007/978-94-015-9536-0

Thurber, C., Zhang, H., Waldhauser, F., Hardebeck, J., Michael, A., & Eberhart-Phillips, D. (2006). Three-Dimensional Compressional Wavespeed Model, Earthquake Relocations, and Focal Mechanisms for the Parkfield, California, Region. *Bulletin of the Seismological Society of America*, *96*(4B), S38–S49. https://doi.org/10.1785/0120050825

Trugman, D. T., & Shearer, P. M. (2017). GrowClust: A Hierarchical Clustering Algorithm for Relative Earthquake Relocation, with Application to the Spanish Springs and Sheldon, Nevada, Earthquake Sequences. *Seismological Research Letters*, *88*(2A), 379–391. https://doi.org/10.1785/0220160188

Tucker, W., Herrin, E., & Freedman, H. W. (1968). Some statistical aspects of the estimation of seismic travel times. *Bulletin of the Seismological Society of America*, *58*(4), 1243–1260. Retrieved from https://pubs.geoscienceworld.org/ssa/bssa/article/58/4/1243/116593/Some-statistical-aspects-of-the-estimation-of

Wagner, M., Husen, S., Lomax, A., Kissling, E., & Giardini, D. (2013). High-precision earthquake locations in Switzerland using regional secondary arrivals in a 3-D velocity model. *Geophysical Journal International*, *193*(3), 1589–1607. https://doi.org/10.1093/gji/ggt052

Waldhauser, F. (2009). Near-Real-Time Double-Difference Event Location Using Long-Term Seismic Archives, with Application to Northern California. *Bulletin of the Seismological Society of America*, *99*(5), 2736–2748. https://doi.org/10.1785/0120080294

Waldhauser, F., & Ellsworth, W. L. (2000). A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California. *Bulletin of the Seismological Society of America*, *90*(6), 1353–1368. https://doi.org/10.1785/0120000006

Waldhauser, F., & Schaff, D. P. (2008). Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods. *Journal of Geophysical Research: Solid Earth*, *113*(B8). https://doi.org/10.1029/2007JB005479

Waldhauser, F., Ellsworth, W. L., Schaff, D. P., & Cole, A. (2004). Streaks, multiplets, and holes: High-resolution spatio-temporal behavior of Parkfield seismicity. *Geophysical Research Letters*, *31*(18). https://doi.org/10.1029/2004GL020649

Zhang, H., Thurber, C., & Bedrosian, P. (2009). Joint inversion for Vp, Vs, and Vp/Vs at SAFOD, Parkfield, California. *Geochemistry, Geophysics, Geosystems*, *10*(11). https://doi.org/10.1029/2009GC002709

Zhou, H. (1994). Rapid three-dimensional hypocentral determination using a master station method. *Journal of Geophysical Research*, *99*(B8), 15439. https://doi.org/10.1029/94JB00934