Accuracy and Precision of Earthquake Location Programs: Insights from a Synthetic Controlled Experiment

Yifan Yu*[1], William L. Ellsworth[1], and Gregory C. Beroza[1]

Abstract

Earthquake location programs employ diverse approaches to address the challenges posed by incomplete knowledge and simplified representation of complex Earth structures. Assessing their reliability in location and uncertainty characterization remains challenging as benchmark datasets with known event locations are rare, and usually confined to particular sources, such as quarry blasts. This study evaluates eight earthquake location methods (GrowClust, HypoDD, Hypoinverse, HypoSVI, NonLinLoc, NonLinLoc_SSST, VELEST, and XCORLOC) through a controlled synthetic computational experiment on 1000 clustered earthquakes based on the setting of the 2019 Ridgecrest, California, earthquake sequence. We construct a travel-time dataset using the fast-marching method and a 3D velocity model extracted from the Community Velocity Model, supplemented with a von Karman perturbation to represent small-scale heterogeneity, and including elevation effects. Picking errors, phase availability, and outliers are introduced to mimic difficulties encountered in seismic network monitoring. We compare the location results from eight programs applied to the same travel-time dataset and 1D velocity structure against the ground-truth locations. For this aftershock sequence, our results reveal the superior accuracy and precision of differential time-based location methods compared to single-event location methods. The results validate the significance of compensating for deviations from assumed 1D velocity structure either by path or site correction modeling or by cancellation of unmodeled structure using differential arrival times. We also evaluate the uncertainty quantification of each program and find that most of the programs underestimate the errors.

Cite this article as Yu, Y., W. L. Ellsworth, and G. C. Beroza (2024). Accuracy and Precision of Earthquake Location Programs: Insights from a Synthetic Controlled Experiment, *Seismol. Res. Lett.* **96**, 1860–1874, doi: 10.1785/0220240354.

Introduction

Accurate earthquake locations, normally provided in the format of earthquake catalogs are essential for most seismological applications such as characterizing seismic hazards (Park *et al.*, 2020; Tan *et al.*, 2021), exploring failure processes (Ross *et al.*, 2019), and illuminating subsurface structures (Wilding *et al.*, 2023). The earthquake location problem uses observed arrival times of direct *P* and *S* phases to estimate the hypocentral parameters of earthquake location and origin time. The relationship between arrival times and hypocentral parameters is nonlinear (Thurber, 1985), and is usually solved iteratively as a linearized problem. Earthquake location results can be strongly influenced by the complexities of solving an inverse problem together with limited knowledge and simplified representation of Earth's structure (Billings *et al.*, 1994). Despite continuous improvements, our understanding of the Earth's 3D seismic velocity structure is often low resolution at large spatial scales and incomplete or absent at small spatial scales, especially in geologically complex regions, such as near active faults where earthquakes occur (Thurber *et al.*, 1997). Faults commonly juxtapose dissimilar lithologies creating complicated 3D velocity structures that can result in large location errors (Wesson, 1971). Moreover, the varying elevations of seismic stations can introduce errors related to the underlying geology (Wesson *et al.*, 1973). Even though several location programs accept 3D velocity models (e.g., GrowClust3D, Trugman *et al.*, 2022), empirical 1D velocity profiles are most commonly used in earthquake location workflows. In operational seismic monitoring, sparse station distributions and suboptimal network geometry can lead to instability in the inverse problem (Gomberg *et al.*, 1990; Kraft *et al.*, 2013). In addition, errors

1. Department of Geophysics, Stanford University, Stanford, California, U.S.A., https://orcid.org/0000-0001-5986-9702 (YY); https://orcid.org/0000-0001-8378-4979 (WLE); https://orcid.org/0000-0002-8667-1838 (GCB)

*Corresponding author: yuyifan@stanford.edu

in waveform processing compromise the reliability of arrival-time measurements (Schaff *et al.*, 2004).

Numerous earthquake location algorithms have been developed to meet the challenges of accurate hypocenter estimation. Most of these methods focus on novel approaches to solving the nonlinear inverse problem and strategies to suppress the effect of unmodeled velocity structure. The original linearized location algorithm (Geiger, 1910) was developed over a century ago. An important subsequent development was the development of joint hypocentral location algorithms in which multiple earthquakes were located simultaneously (Douglas, 1967). That enabled the subsequent development of simultaneous estimation of hypocentral parameters and seismic velocity structure (Kissling *et al.*, 1994). Improvements in algorithms and computation eventually allowed the development of nonlinear approaches (Tarantola and Valette, 1981).

More recent advances tend to highlight computational capability. Grid search and stochastic direct search have gained popularity as alternatives to linearization, such as, HypoSVI (Smith *et al.*, 2022), NonLinLoc (Lomax *et al.*, 2000, 2009), and NonLinLoc_SSST (Lomax and Savvaidis, 2022). Differential time-based location programs are designed to suppress path effects by pairing nearby events rather than treating them individually (Jordan and Sverdrup, 1981; Got *et al.*, 1994; Waldhauser and Ellsworth, 2000; Lin and Shearer, 2006) (e.g., GrowClust, Trugman and Shearer, 2017; HypoDD, Waldhauser, 2001; and XCORLOC, Lin, 2018). Additional strategies, such as adaptive weighting, static station terms, and source-specific station terms (SSST) have been introduced to accommodate systematic unmodeled geological structure, either along the path or at the station, and potential instrumental errors (Frohlich, 1979; Richards-Dinger and Shearer, 2000; Lin, 2018). Although these programs have been used to locate earthquakes to great effect, there are differences in their outputs that motivate a comprehensive evaluation of how best to use them, how accurate their hypocentral estimates are, and how well they characterize error.

To address this need, we developed a realistic benchmark dataset with known event locations that can be used to assess the performance of different algorithms for estimating location and quantifying uncertainty. Some work has been done in this area already. Lin and Shearer (2005) compare three relative earthquake location techniques on a simple synthetic experiment, and Wuestefeld *et al.* (2018) test different objective functions for earthquake location in a simplified downhole acquisition scenario. Pyle *et al.* (2023) evaluate accuracy and precision for locating nine shallow earthquakes in the Rock Valley fault zone, southern Nevada. To understand performance of location programs more generally, however, requires a test that replicates real-world seismic monitoring of an earthquake dataset with a substantial number of events over a spatially extended region. Given the prohibitive costs and inconvenience of conducting controlled explosion experiments, which are rarely conducted at sufficient depth near-fault zones and do not fully replicate earthquake processes (Thurber *et al.*, 1997, 2004), we take the approach of a computational experiment. This study assesses the effectiveness of eight widely used earthquake location programs (GrowClust, HypoDD, HYPOINVERSE, HypoSVI, NonLinLoc, NonLinLoc_SSST, VELEST, and XCORLOC) for accurate hypocenter estimation and reliable uncertainty quantification using a realistic controlled synthetic experiment based on the setting of the 2019 Ridgecrest earthquake sequence.

We construct a realistic 3D velocity model that incorporates the complex effects of elevation variations observed in the Ridgecrest region, where elevated granitic mountains have higher seismic velocities and sedimentary basins have lower seismic velocities (Fig. 1). This model is used to generate a travel-time dataset using an eikonal-function-based fast marching method travel-time calculator (White *et al.*, 2020). We also introduce realistic variations in phase availability, picking errors, and arrival-time outliers in the dataset. As described in below (Fig. 2), we first set up the forward problem to construct the travel-time dataset. We then run eight earthquake location programs on the same dataset with the same 1D velocity profile. The outputs are compared with the ground-truth (GT) locations to analyze their performance in terms of accuracy and precision.

The accuracy error quantifies the discrepancy between each event location in the program's output and the GT location. Precision error, on the other hand, measures the average difference in distances between event pairs formed by each hypocenter and its neighboring events within 2 km distance. The precision error assesses the preservation of relative structures such as fault-plane geometry. We use the Chamfer distance, a standard metric for evaluating shape dissimilarity between two point clouds, to assess each program's ability to recover structures. The Chamfer distance ($d(S_1,S_2)$) is defined as

$$d(S_1,S_2) = \sum_{x\in S_1} \min_{y\in S_2} \|x-y\|_2^2 + \sum_{y\in S_2} \min_{x\in S_1} \|x-y\|_2^2, \quad (1)$$

in which S_1, S_2 are two point clouds, and here refers to the hypocenter from GT dataset and each program's output.

Our results demonstrate that differential time-based earthquake location programs achieve higher accuracy and precision than single-event location programs. Single-event location programs with various forms of station corrections implemented, and differential time-based programs, excel in hypocenter determination. This validates the importance of compensating for 1D velocity structure approximations through either direct modeling of path or site contributions or by suppressing it through pairwise cancellation of these effects.

We compare the uncertainty quantification for each program with GT using the 95% confidence ellipse. The results suggest a need to revisit location error estimates as most of the programs underestimate the errors. In the Discussion, we argue that a 1D velocity profile approximation will introduce

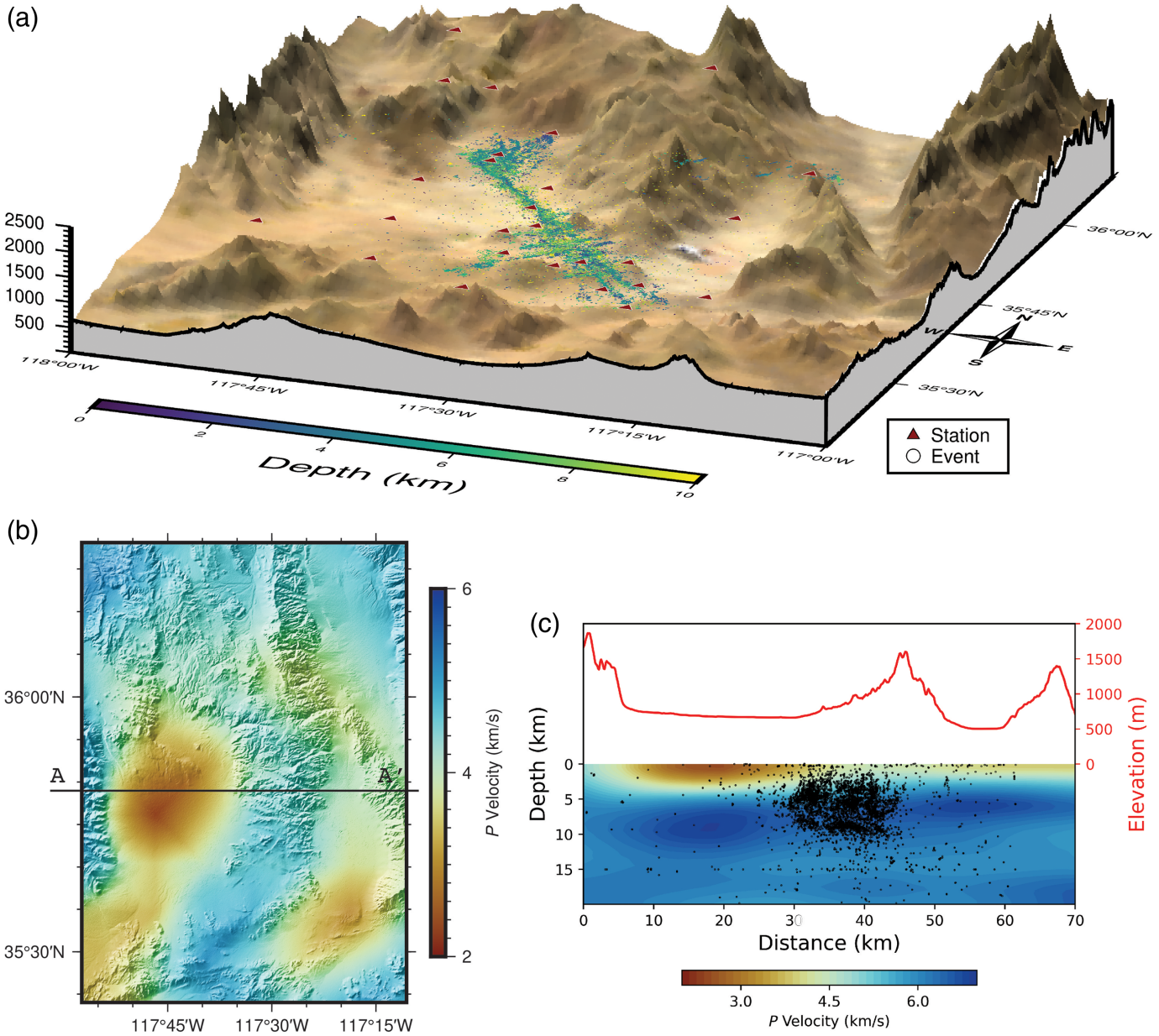


Figure 1. (a) Distribution of seismic stations and seismicity in the QTM catalog from the Southern California Earthquake Data Center (SCEDC) during the 2019 Ridgecrest earthquake sequence (Ross *et al.*, 2019; Hauksson *et al.*, 2020). (b) Seismic velocity map of the modeled area, with shaded elevation data. Both velocity and elevation data are extracted from the Statewide California Earthquake Center (SCEC) Community Velocity Model (CVM) (Small *et al.*, 2017). The correlation between low seismic *P*-wave velocity and low elevation in the sedimentary basin is highlighted. (c) Seismic velocity and elevation along the A–A′ profile, with black dots indicating seismicity from the SCEDC QTM catalog. The color version of this figure is available only in the electronic edition.

a bias that fundamentally limits the accuracy of depth determinations. We also note that our computational synthetic test experiment can be easily extended to other regions or scenarios at local-to-regional distance scales.

Overview of Eight Earthquake Location Programs

In this section, we review algorithms used to estimate hypocentral locations and uncertainty, implemented in both single-event location programs (Hypoinverse, HypoSVI, NonLinLoc, NonLinLoc_SSST, and VELEST) and differential time-based location programs (GrowClust, HypoDD, and XCORLOC). Although some single-event location programs utilize information across multiple events to model station corrections or perform simultaneous inversion, for clarity and convenience, we categorize these eight programs into these two groups.

Single-event location programs

Similar to earlier linearized earthquake location products built by U.S. Geological Survey (HYPOLAYR, Eaton, 1970; HYPO71,

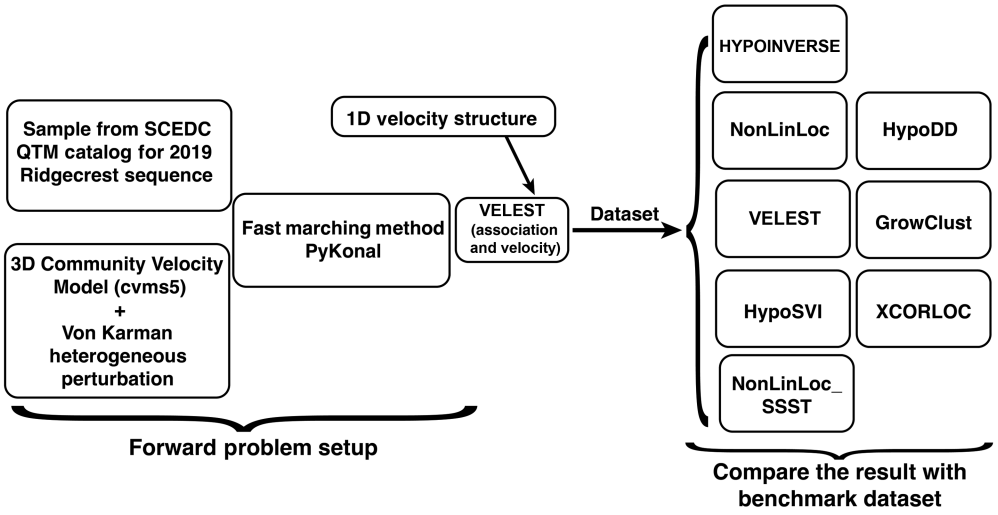


Figure 2. Workflow for evaluating eight earthquake location programs. The process begins by using PyKonal (White *et al.*, 2020) to calculate travel times on a realistic 3D velocity model (Fig. 3) for randomly sampled event locations. After introducing phase-picking errors and outliers, the dataset, along with the 1D velocity model for the 2019 Ridgecrest earthquake sequence (Shelly, 2020) is input into VELEST (station terms disabled), serving as a phase association step, to estimate the starting locations and to determine a suitable 1D velocity structure to be used for all eight programs.

Lee and Lahr, 1972; and HYPOELLIPSE, Lahr, 1979), HYPOINVERSE-2000 iteratively solves a set of linearized equations to minimize the root mean square (rms) travel-time residual, utilizing adaptive damping and cutoff criteria. Because of its simplicity and long-standing use, HYPOINVERSE plays a key role in seismic network monitoring (e.g., Southern California Earthquake Data Center [SCEDC], Hutton *et al.*, 2006). VELEST (version 3.2, Kissling *et al.*, 1994) simultaneously determines hypocenters, 1D velocity structure, and station terms by incorporating linearized updates to all unknowns into the Geiger equations. Rather than linearize the nonlinear problem, NonLinLoc uses a probabilistic, global-search method to estimate the locations. The algorithm's efficiency is enhanced by the equal-differential time (EDT) likelihood function and the directed OctTree sampling method (Lomax *et al.*, 2000). It is adopted in some regional seismic network operations (e.g., TexNet, Savvaidis *et al.*, 2019). By incorporating the corrections from the SSST method, which smoothly vary throughout the space to provide source-position-dependent corrections for each station, NLL_SSST, and NLL_SSST_Coherence, which further uses the coherence measurements from the event pairs provide better location estimates (Lomax and Savvaidis, 2022). We include both NonLinLoc and NonLinLoc_SSST in this study. Following a Bayesian inference approach for the multimodal posterior distributions in hypocentral inverse problems, HypoSVI implements Stein variational inference (Smith *et al.*, 2022). To accelerate the extensive travel-time computation, HypoSVI uses EikoNet, an efficient machine learning-based travel-time calculator (Smith *et al.*, 2021).

For error estimation, HYPOINVERSE and VELEST calculate error ellipses based on the covariance matrix from reading error-weighted travel-time residuals. Single-event mode with no iteration performed can be used to assess error statistics in VELEST (Kissling, 1995). NonLinLoc and its extended programs obtain the covariance matrix from normalized samples, extracting uncertainty estimates based on the distribution of sampled locations using a Gaussian estimator. HypoSVI obtains uncertainty estimates by analyzing the 95th percentile of sampled locations along each dimension (Smith *et al.*, 2022). We note that covariance estimates for nonlinear problems are only valid in a linear approximation.

Differential time-based location programs

The other three location methods we assess in this study leverage differential travel times between nearby event pairs to reduce path effects, thereby achieving high precision (Jordan and Sverdrup, 1981; Got *et al.*, 1994; Waldhauser and Ellsworth, 2000). HypoDD minimizes the L2 norm of differential travel-time misfits using either SVD or least-squares (LSQR) algorithms. It groups events based on interevent distances and the number of observations (Waldhauser, 2001). GrowClust was developed to overcome the instability and computational limits posed by ill-conditioned matrices from large datasets. It features the minimization of the L1 norm and the use of travel-time table searches. GrowClust clusters events based on a linkage selection criterion, such as waveform similarity (Trugman and Shearer, 2017). Recent updates to GrowClust include accepting 3D velocity models and implementing parallel computing to further enhance its capabilities (Trugman *et al.*, 2022). Likewise, XCORLOC minimizes hybrid misfits (a combination of L1 and L2 norms) through travel-time table searches. In addition, the package includes the inversion for static station corrections and SSST; however, it requires additional information on the mean epicentral distances to model the SSST (Lin, 2018).

HypoDD estimates error through the covariance matrix, with the SVD mode recommended to avoid underestimation due to the damping effects inherent in LSQR mode (Waldhauser, 2001). Given computational limitations, this is

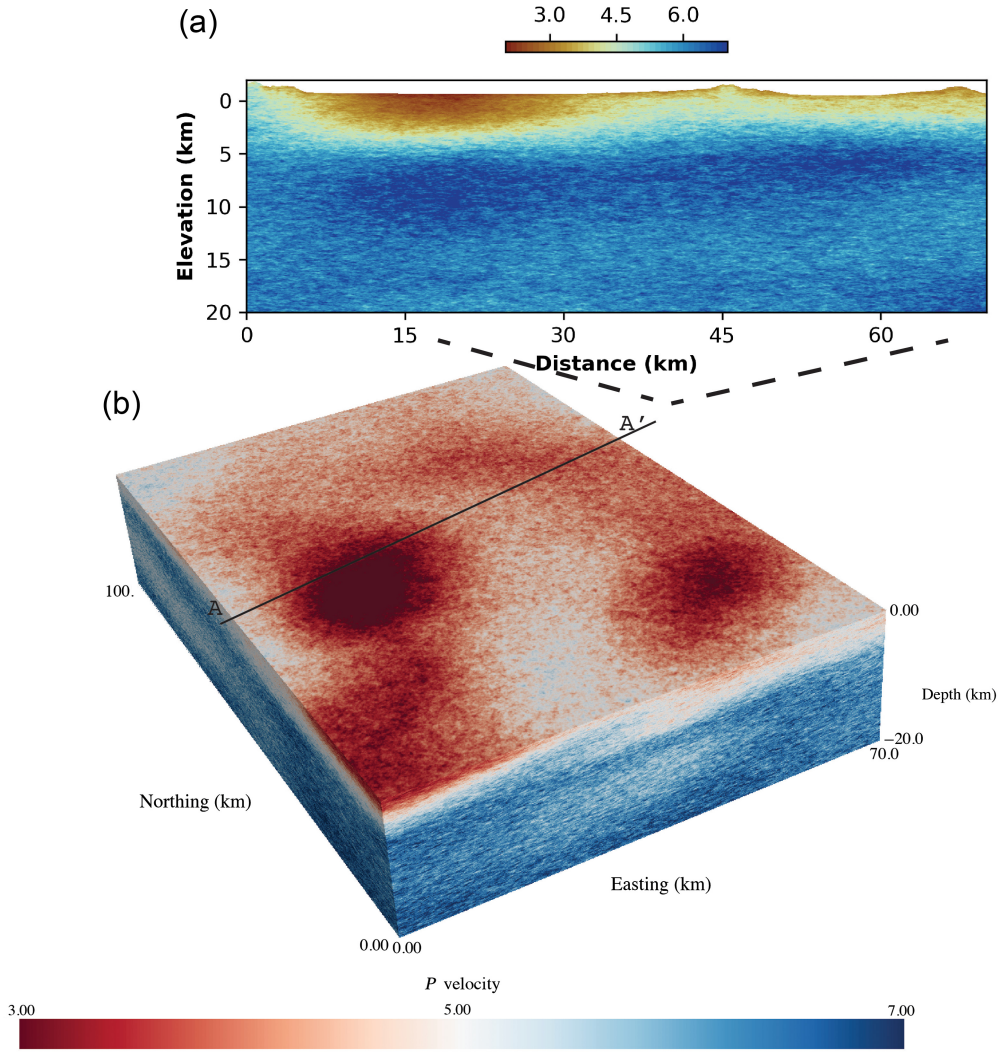


Figure 3. (a) The A–A′ profile of the 3D *P*-wave velocity model used in this study, corresponding to the profile shown in Figure 1. (b) The 3D *P*-wave seismic velocity model used in this study, with a von Karman perturbation superimposed on the SCEC CVMS5 to represent realistic small-scale heterogeneity. The model grid dimensions are 70 km × 100 km horizontally and 20 km in depth, plus realistic elevation variations (up to 2 km). The grid resolution is 100 m × 100 m horizontally and 50 m vertically. The color version of this figure is available only in the electronic edition.

typically done by running small datasets in SVD mode. Both GrowClust and XCORLOC use bootstrapping to quantify uncertainty.

Evaluating the Programs

In this section, we outline the workflow designed to assess the performance of each earthquake location program. As depicted in Figure 2, the workflow consists of three parts: (1) constructing the synthetic travel-time dataset, (2) running eight location programs, and (3) comparing the output with GT. We begin this section with a detailed description of the synthetic dataset setup, followed by an overview of each program's configuration. Although the parameters vary across programs, we attempt to adhere to the recommendations in their respective manuals and report any adjustments we make. The running scripts were also shared with the program authors for their review. Finally, we present the methods used to assess the performance of earthquake location programs by measuring accuracy, precision, and uncertainty against GT. We introduce metrics for evaluating the accuracy, precision, and reliability of uncertainty estimates, providing a comprehensive comparison of each program's output.

Forward problem setup

To simulate a realistic seismic monitoring scenario, we consider stations at the same location as the available stations covering the Ridgecrest region from those listed at the SCEDC (Hauksson *et al.*, 2020) (Fig. 1). Our GT hypocenter dataset consists of 1000 events randomly sampled from the 2019 Ridgecrest Earthquake QTM catalog (Ross *et al.*, 2019). We first select clusters containing more than 400 events in the QTM catalog, and then sample from these well-populated clusters. Because we do not calculate the waveforms, this filtering process is required for applying those methods that use additional information, such as waveform similarity, to form clusters. Leveraging the information inherited from the QTM catalog, which is constructed using GrowClust, the valid waveform similarities (cross-correlation values) are retained and used for those differential time-based location programs. We note that using the waveform similarities based on the real data in this synthetic experiment has the potential to introduce some unknown bias into the results.

To construct a geologically accurate 3D velocity model, we first obtain a base model from the Southern California Earthquake Center (SCEC) Community Velocity Model (CVMS5) (Lee *et al.*, 2014; Small *et al.*, 2017), spanning 70 km in easting, 100 km in northing, and 20 km in depth, with a grid resolution of 100 m in the east and north directions, and 50 m vertically (Fig. 3). The model contains only large-scale variations

due to the smoothing effect in the full waveform inversion used to develop the base model, so we introduce realistic small-scale heterogeneity into the model using the von Karman perturbation model. The von Karman correlation function is as follows:

$$P_{vK}(k_x,k_y,k_z) = \frac{2^d \pi^{d/2} \varepsilon^2 a_x a_y a_z \Gamma(\kappa + d/2)}{\Gamma(|\kappa|)(1 + a_x^2 k_x^2 + a_y^2 k_y^2 + a_z^2 k_z^2)^{\kappa+d/2}}, \quad (2)$$

in which d is the Euclidean dimension ($d = 3$); Γ is the Gamma function; κ is the Hurst exponent ($\kappa = 0.04$); a is the correlation length for each direction ($a_x = a_y = 100$ m, $a_z = 50$ m); and ε is the fractional magnitude of the fluctuation ($\varepsilon = 0.05$) (Nakata and Beroza, 2015).

To represent elevation effects (Fig. 1), we classify each grid as either granitic ($V_P > 4$ km/s) or sedimentary rock ($V_P \leq 4$ km/s) based on the seismic P-wave velocity at the sea level datum. The seismic velocity is then adjusted to decrease with elevation, following different slopes determined by rock type in Brocher (2008). The elevation data are obtained from the SCEC CVM.

We use the PyKonal package (White *et al.*, 2020) to compute the travel-time dataset. For each station, we set the probability of recording P phases at 67% and for S phases at 50%. This results in an average of 18 P phases and 14 S phases per event across 27 stations. We add reading errors using a Laplacian distribution with a location parameter of 0 ms and scale parameters of 20 ms for P phases and 40 ms for S phases. In addition, we include 1% P and 4% S phase outliers, both positive and negative, with deviations uniformly distributed between 0.4 and 1.0 s for P and between 0.4 and 1.4 s for S. We round off the dataset to a precision of 10 ms to represent a 100 Hz sampling rate. The travel-time dataset and the empirical 1D velocity profile in the Ridgecrest region (Shelly, 2020) are input to VELEST (station term disabled), which serves as a phase association tool to prepare the starting locations from travel times and to determine a revised 1D velocity profile that better fits our 3D velocity model.

To accommodate the other programs (GrowClust, HypoDD, and XCORLOC), we need a differential travel-time dataset. We use the *ph2dt* function in HypoDD to pair the neighboring events, allowing for a maximum of 65 neighbors within 15 km for each event. For the value of differential times, smaller phase errors in Laplacian distribution (location parameter: 0 ms, scale parameter: 11 ms for both phases), along with fewer outliers (0.5% for P phases and 2% for S phases) are added to the results from PyKonal calculation, as Schaff and Waldhauser (2005) show that the differential times are more accurate. We fit the Laplacian distribution to rms residuals of P and S phases from the QTM catalog, yielding location parameters of 0 ms and scale parameters of 22 and 20 ms, respectively. Based on other potential misfit sources, we adopted a scale parameter of 11 ms for both phases. In addition, they are rounded to 1 ms precision because the differential time is able to achieve subsample precision from spectral coherence weighted cross-correlation measurements (Poupinet *et al.*, 1984). Exploiting the information inherited from the QTM catalog, the value of waveform similarity is assigned to 1.0 if two events are in the same cluster, 0.3 otherwise. Finally, we calculate the distances between event-pair midpoints and stations to fulfill the mean epicentral distance requirements of XCORLOC.

Running the location programs

For single-event location programs, we use the same travel-time dataset and revised 1D velocity model as inputs. For the differential time-based location programs, we use the differential time dataset only, and the same revised 1D velocity model as inputs. Although HypoDD and XCORLOC can use both absolute and differential travel time as input, we only consider the differential times. We set station elevations to zero for a fair comparison as some of the programs treat all stations at zero elevation, the results of using real station elevations are shown in Table A1.

Starting with a trial depth of 8 km, HYPOINVERSE converges to a final solution employing default updating rms and distance weighting strategies per iteration, with an assumed reading and timing error of 0.04 s. VELEST runs in the simultaneous mode to estimate the location with station terms enabled. After 30 iterations, we execute VELEST in single-event mode only to compute uncertainty statistics with locations from the simultaneous mode fixed and no inversion iteration performed. After building a travel-time grid of the study area with a grid size of 0.1 km using azimuthal equidistant projection, NonLinLoc calculates final solutions by OctTree sampling to maximize the equal differential time likelihood function weighted by the variance of origin-time estimates over all pairs of readings (EDT_OT_WT). After four iterations of generating the SSST correction grid, NonLinLoc_SSST performs location in a similar configuration with NonLinLoc but exploits the SSST corrections. When running HypoSVI, we first train EikoNet on the velocity profile using a batch size of 64, with 1,000,000 training and 100,000 testing data for 100 epochs at a learning rate of 1×10^{-3}. For HypoSVI, we retain the recommended parameters in the example files and allow one iteration of static station terms and 10 iterations of SSST.

For HypoDD we cluster all events into a single cluster, requiring a minimum of eight observations per pair. After three sets of damped LSQR iterations, the relocated output of hypoDD is divided into small sets for input into SVD mode for error calculation because error estimation in LSQR mode is restricted to the diagonal variances and is dominated by the damping factor (Waldhauser, 2001). We use the "trace" mode in GrowClust to calculate the travel-time table with a 0.5 km horizontal and 0.2 km depth resolution, utilizing a 1D velocity profile. By setting the minimal coefficient of event-pair

similarity to 0.75, GrowClust keeps events in the same cluster as the QTM catalog from which we derived GT locations. The program conducts 100 bootstrapping iterations for uncertainty analysis. Employing the same grid dimension of the travel-time table, XCORLOC executes 10 iterations of SSST, 5 iterations of waveform cross-correlation location, as well as 100 bootstrapping iterations.

Postprocessing and metrics

After measuring the horizontal and vertical distances between each program's output and the GT dataset, we compute the average distances to determine the horizontal and vertical mean accuracy errors. We then calculate the interevent distance misfits, defined as the distances between each event and its neighboring event, between the output and the GT. We average these misfits to obtain the mean precision errors in both the horizontal and vertical directions. We evaluate the Chamfer distance to quantify the dissimilarity between predicted and GT locations comprehensively in the form of point clouds.

Beyond assessing the accuracy and precision of the programs, we also evaluate the uncertainty output. The uncertainty outputs vary in form in the different algorithms, so we first convert them to 95% confidence intervals. Although the earthquake location problem is 4D (i.e., solving for four unknowns of origin time and three spatial coordinates), some programs exclude origin time from the error analysis. Therefore, our focus is on error output in spatial coordinates. For programs that analyze errors based on the covariance matrix, we convert the output by multiplying by a factor derived from the assumption of a chi-square distribution with three degrees of freedom. We note that in VELEST, the default output error is the unit covariance matrix, so we first multiply it by the data variance. For bootstrapping-based programs, we assume the bootstrapping results follow a Gaussian distribution, convert the median error to the mean, and then apply a similar factor based on the chi-square distribution. For NonLinLoc and NonLinLoc_SST, the error distribution is assumed to follow the $\Delta\chi^2$ distribution with three degrees of freedom, as described in the program's manual. To evaluate the effectiveness of the calculated 95% confidence interval in characterizing the uncertainty, we define two metrics: the inclusion rate of hypocenters and the inclusion rate of event pairs. The inclusion rate of hypocenters represents the percentage of outputs where the true location falls within the estimated location plus the 95% confidence ellipse. It provides a measure of the accuracy of the location uncertainty estimates. The inclusion rate of event pairs assesses the relative precision of event pair locations. For each event, we identify neighboring events within a 2 km radius. We then count the ratio of pairs where the true interevent distance is within the estimated distance plus the combined error interval.

Results

Figure 4 compares the results of single-event location programs with the GT. In map view, all programs capture the overall trend of the orthogonal fault system; however, differences in resolving the fine details of fault structures are apparent, particularly the linear alignments at the southeastern end of the main fault. Although these trends are visible in the GT locations, the results from both Hypoinverse and NonLinLoc appear more diffuse. In contrast, NonLinLoc_SSST, HypoSVI, and Velest produce results that are more concentrated around these features. In a similar manner, the southwestern orthogonal branch is better resolved by NonLinLoc_SSST. This comparison indicates that while NonLinLoc and HYPOINVERSE can recover the basic epicentral distribution, further refinement is possible for detailed seismicity studies. This is supported by the error histograms. NonLinLoc_SSST exhibits significantly lower errors. The locations from VELEST show improved resolution of GT compared to NonLinLoc and HYPOINVERSE, which have similar horizontal errors of less than 3–4 km. Interestingly, despite HypoSVI's effectiveness in resolving seismic structures, as evidenced by the large number of events with errors less than 1 km, there is still a small number of events with larger errors. In the depth profile (Fig. 5), distinctions are more pronounced. None of the programs successfully recovered the outlier event located around 8 km distance and 10 km depth. Most programs struggle to constrain the distribution of GT seismicity around 8 km distance and 4 km depth, with HypoSVI and HYPOINVERSE producing particularly diffuse results. The histograms demonstrate that VELEST and NonLinLoc_SSST have the most consistent performance in depth accuracy. Although HypoSVI shows minimal errors for the majority of events, it produces some outliers. Even though the programs use the same dataset, the strategies employed by NonLinLoc_SSST, HypoSVI, and VELEST, such as averaging station terms across all events, gain more information from the travel time to recover better hypocenter locations.

Figure 6 compares the results of differential time-based location programs with GT. The structures manifest strong coherence with the true locations, which aligns with the empirical findings that additional relocation using waveform cross-correlation can better delineate seismicity and highlight geological structures. The results from the three programs closely match the GT, with histograms indicating horizontal errors of less than 1 km; however, a few events from GrowClust exhibit larger errors, which is particularly evident in the depth profile. HypoDD and XCORLOC successfully recover the outlier event at ~8 km distance and 10 km depth, while GrowClust fails to do so. In addition, GrowClust tends to locate some events slightly deeper, as reflected in the error histograms. This issue is likely due to the small number of events in the clusters these events and the outliers belong to that lack sufficient constraints to recover locations. HypoDD and XCORLOC handle all events within a single cluster. Because of the design of this experiment, GrowClust inherits the cluster information from

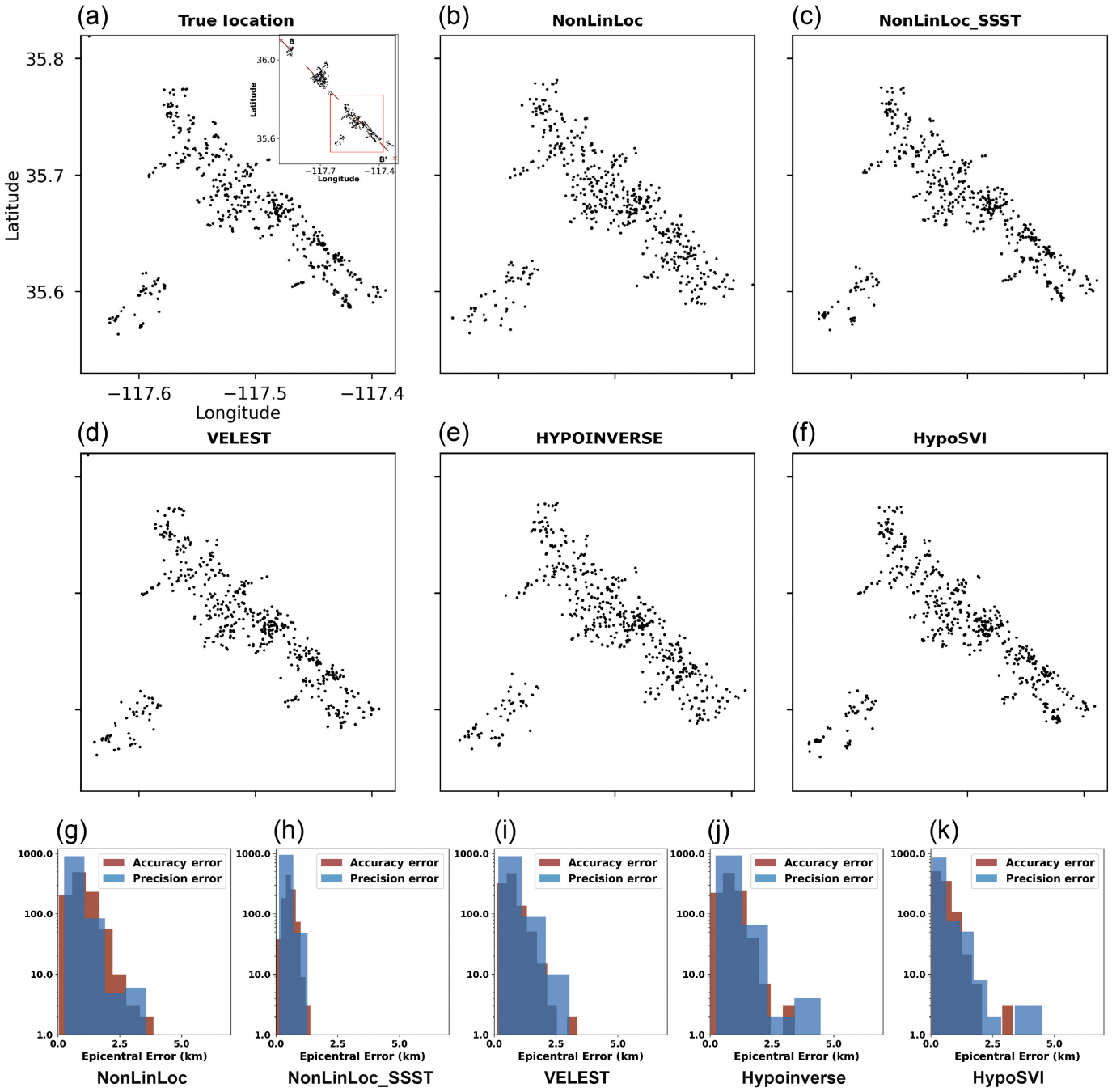


Figure 4. (a–f) Map view of the epicentral locations of ground truth (GT) and the outputs of each single-event location programThe inset map in (a) shows the focused region and the B–B′ profile used for the depth analysis. (g–k) Histograms of horizontal error distributions for each program. In each histogram, the *x*-axis represents error magnitude in kilometers, and the *y*-axis shows the number of events on a logarithmic scale. Red bars indicate accuracy errors, while blue bars represent precision errors. The color version of this figure is available only in the electronic edition.

the QTM catalog, so there are some clusters with few events. Our results suggest that ensuring a sufficient number of events per cluster is important for GrowClust's performance. Overall, histograms of accuracy and precision errors indicate that relative structure exhibits lower error than absolute location and horizontal errors are less than depth errors. The accurate and precise results from differential time-based location programs suggest that by suppressing path effects, the biases from the 1D velocity structure are successfully corrected.

Table 1 summarizes the performance of eight programs based on the metrics introduced earlier. The table reports mean accuracy and precision errors for both horizontal and depth directions. Results considering correct elevation are presented in Table A1. Generally, the location errors are on the order of hundreds of meters, with depth errors larger than horizontal errors. HypoDD achieves the highest accuracy in the horizontal

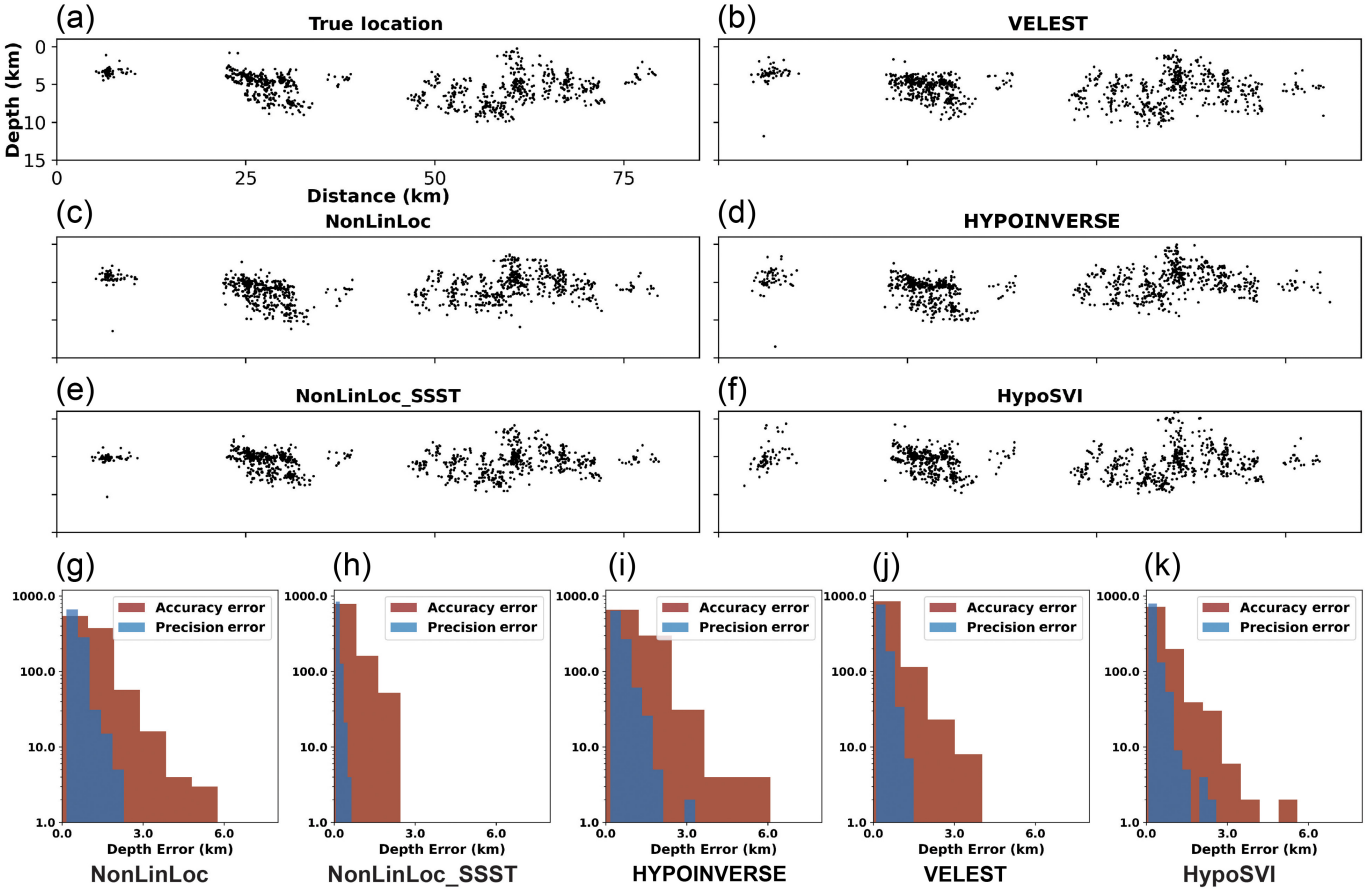


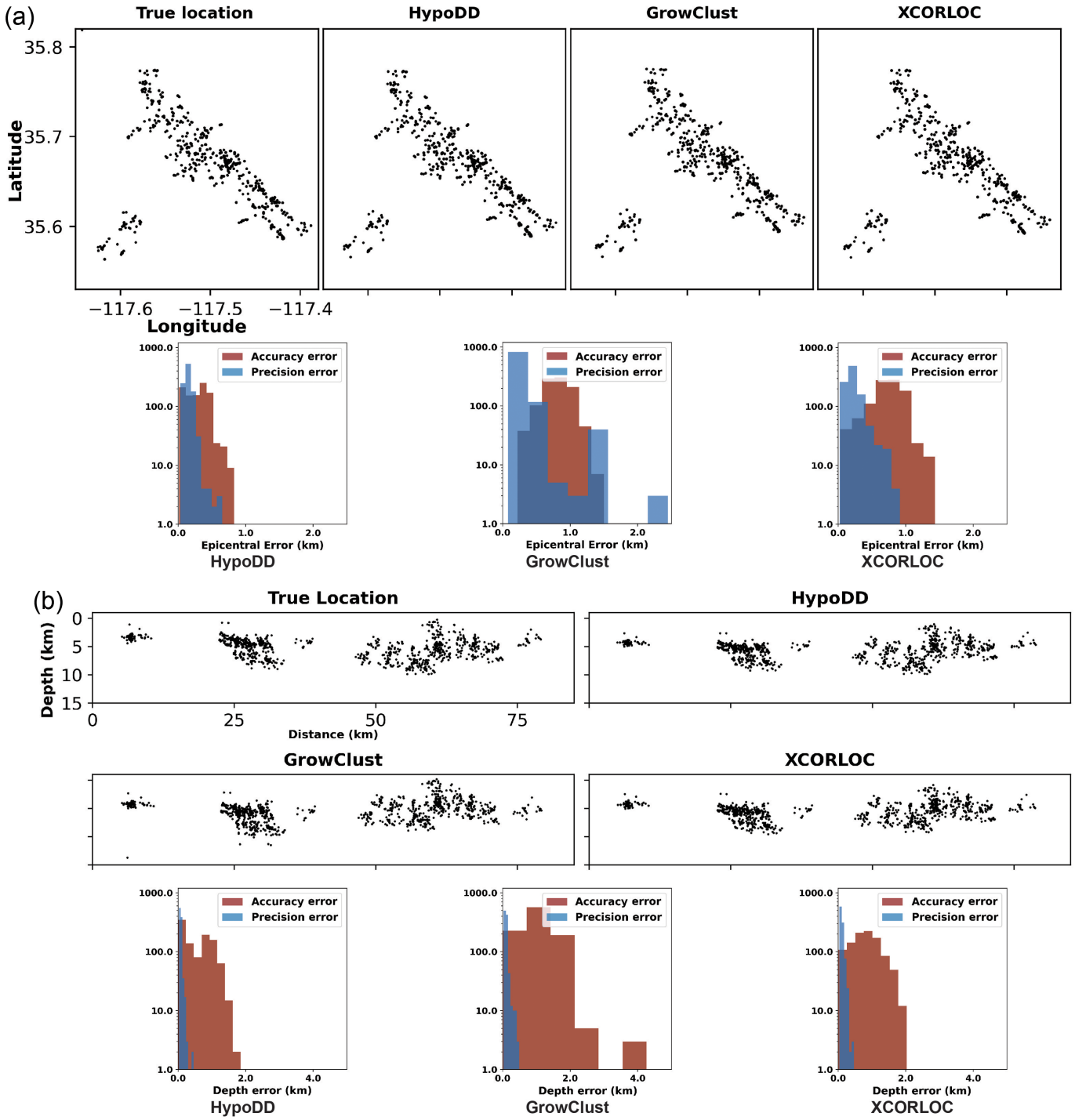
Figure 5. (a–f) Depth profiles along the B–B′ line for the GT and each single-event location program's output. (g–k) Histograms of depth error distributions with red bars for accuracy errors and blue bar for precision errors. The *x*-axis represents error magnitude, and the *y*-axis indicates event occurrence. The color version of this figure is available only in the electronic edition.

direction, whereas HypoDD, VELEST, NonLinLoc_SSST, and HypoSVI perform similarly in-depth estimation. As expected, programs using differential time data with additional information embedded, HypoDD, GrowClust, and XCORLOC demonstrate superior mean precision. The large Chamfer distance observed in GrowClust is likely due to the presence of larger errors for events that fall within small clusters.

Table 2 summarizes the performance of eight programs in uncertainty quantification. Inclusion rates of hypocenters, which measure the fraction of events that the true location falls within the estimated location plus the 95% confidence ellipse are not reported for differential time-based programs, as those programs are primarily designed to estimate interevent distance uncertainty. This difference is evident in the scale of their uncertainty outputs: differential time-based programs yield uncertainties on the order of hundreds of meters, whereas single-event programs yield uncertainties on the order of kilometers. Among the programs, NonLinLoc, NonLinLoc_SSST, and VELEST exhibit robust error estimation capabilities for hypocenters. In contrast, HYPOINVERSE shows lower performance, highlighting the inherent complexity and challenges of error analysis in earthquake locations. Despite considering the 95% confidence interval, all programs report an inclusion rate of hypocenters significantly below 95%, indicating that these algorithms tend to underestimate location uncertainties. This should motivate improvements in uncertainty analysis. For event pairs, the high inclusion rates of single-event programs are due to their generally larger uncertainty outputs. The comparison between event pairs and hypocenters also suggests the effect of regional systematic shift increasing location uncertainty while interevent error is less affected. HypoDD's covariance matrix-based uncertainty functionality proves effective, whereas the bootstrapping-based approaches of GrowClust and XCORLOC show limitations in error quantification.

Discussion

Systematic shifts in hypocentral locations have been widely documented from regional to teleseismic monitoring scales (Engdahl *et al.*, 1998; Syracuse and Abers, 2009). These shifts can stem from factors such as station geometry, the presence of high-velocity layers, or the simplification inherent in assuming a 1D velocity structure. In our analysis, we observed such inconsistent systematic depth shifts across the study region in the results (Fig. 7a). The controlled synthetic experiment



framework allows us to investigate directly the relationship between the 3D velocity model and the 1D velocity profile used by the location programs. By comparing the 1D velocity profile with randomly sampled velocity cross-sections from the 3D model, we observed that velocity cross sections from areas of accurate locations align more closely with the 1D velocity profile (Fig. 7b,c). This unsurprising result highlights the limited capability of 1D velocity profiles to recover precise earthquake hypocenters. To enhance seismic monitoring accuracy, the biases introduced by neglecting 3D velocity variations must be reduced. The advances in computational capabilities have

Figure 6. (a) Map view of the epicentral locations of GT and the output of each differential time-based location program, with histograms detailing the distribution of errors. In the histograms, the *x*-axis represents error magnitude in kilometers, and the *y*-axis shows the number of events on a logarithmic scale. The red bars indicate accuracy errors, and the blue bars represent precision errors. (b) Depth profiles along the B–B′ line for the GT and each program's output. The accompanying histograms follow the same format, with the *x*-axis showing error magnitude and the *y*-axis indicating event occurrence, with red for accuracy and blue for precision. The *x*-axis range in these histograms is smaller than in Figures 4 and 5. The color version of this figure is available only in the electronic edition.

TABLE 1
Summary of Performance Metrics for Eight Earthquake Location Programs Using Zero Station Elevation

	Mean Accuracy Error (km)		Median Accuracy Error (km)		Chamfer	Mean Precision Error (km)	
Program	**Horizontal**	**Depth**	**Horizontal**	**Depth**	**Distance**	**Horizontal**	**Depth**
HYPOINVERSE	0.824	1.118	0.768	1.034	1.617	0.571	0.684
VELEST	0.694	0.560	0.576	0.380	1.172	0.379	0.663
NonLinLoc	0.953	0.969	0.788	1.075	1.626	0.581	0.694
NonLinLoc_SSST	0.542	0.583	0.533	0.475	1.219	0.163	0.298
HypoSVI	0.498	0.580	0.390	0.332	1.092	0.316	0.452
HypoDD	0.315	0.571	0.356	0.526	0.974	0.080	0.141
XCORLOC	0.746	0.818	0.756	0.808	1.507	0.101	0.256
GrowClust	0.846	1.069	0.848	1.097	1.694	0.148	0.310

TABLE 2
Inclusion Rates of Hypocenters and Event Pairs for Each Earthquake Location Program

Program	Hypocenters	Event Pairs
HYPOINVERSE	13.1%	85.7%
VELEST	80.3%	97.0%
NonLinLoc	85.0%	99.7%
NonLinLoc_SSST	82.3%	99.9%
HypoSVI	75.0%	98.2%
HypoDD	n/a	99.5%
XCORLOC	n/a	40.8%
GrowClust	n/a	84.4%

enabled the integration of 3D velocity structures into location algorithms offering a promising approach to reduce location error. Trugman *et al.* (2022) demonstrated that using a 3D velocity model can resolve seismicity patterns that remain obscured in a 1D model.

To understand the relative contributions of different error sources, we compare the residuals returned by the location programs with assigned picking errors and 1D approximation deviations. The median residual from HYPOINVERSE's output is 130 ms, reflecting the combined effects of multiple potential factors. In our synthetic tests, the assigned picking errors follow a Laplacian distribution (location parameter = 0 ms, scale parameter = 20 ms) for *P* arrivals. For a source located at the center in our model and 8 km depth, calculating travel times at zero elevation using the 1D velocity model yields a median error of 79 ms compared to the 3D model. This analysis reveals that the simplified velocity model contributes significantly more to the total location misfits than the assigned picking errors. The large observed residuals (130 ms) suggest additional complexity in the location problem that introduces more uncertainties other than these two factors.

Standard network operations typically use single-event location methods, relying on automated or manual measurements; however, outliers in phase arrival times due to path effects or low signal-to-noise ratios can affect the accuracy and precision of the results (Rowe *et al.*, 2002). Our results suggest that single-event location programs may fail to correct outlier events, and that differential time-based location programs hold promise for recovering them. In doing so, it is essential to have an appropriate cluster size to ensure the reliability of location results in differential time-based location programs. Although differential time measurements are often associated with cross-correlation calculations, they can be effectively derived from arrival-time data that are routinely available for all seismic events. This broader applicability has been demonstrated through the successful implementation of differential location catalogs in near-real-time monitoring, particularly with the double-difference real-time (DDRT) catalog system operated by the Northern California Seismic Network. The DDRT system dynamically locates new events relative to previously determined locations in the catalog by utilizing both arrival-time differences and cross-correlation measurements (Waldhauser, 2009). The system's successful deployment in various monitoring contexts, from regional networks to submarine volcano monitoring (Waldhauser *et al.*, 2020) illustrates that differential location methods can be effectively integrated into routine network operations. Beyond differential methods, our work also demonstrates how single-event location procedures in network monitoring can be enhanced through the integration of station corrections. By inputting station corrections derived from VELEST into HYPOINVERSE, the results show promising improvements in location accuracy: the mean horizontal error decreased from

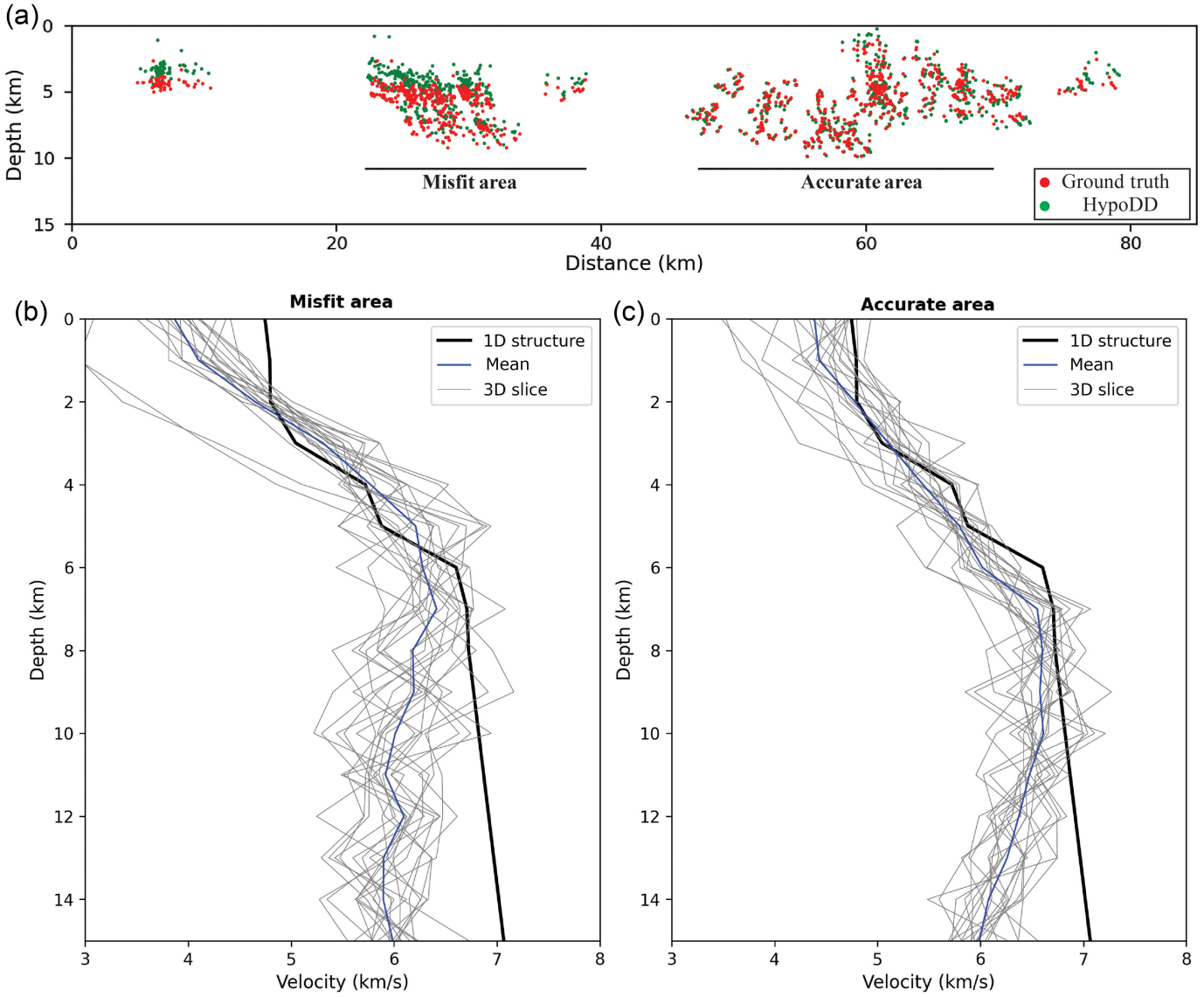


Figure 7. (a) Depth profile comparing the results of HypoDD with the GT. There is a minimal systematic shift in the denoted accurate area, whereas the misfit area shows a significant upward shift. (b) The plot of the 1D velocity structure used by the program, including randomly selected velocity cross sections from the 3D velocity model in the misfit area, along with the mean of these cross sections. (c) The plot of the 1D velocity structure used by the program, including randomly selected velocity cross sections from the 3D velocity model in the accurate area, along with the mean of these cross sections. The 1D velocity profile fits the accurate area better, which leads to negligible shifts. The discrepancy of velocity in the misfit area leads to systematic shifts. The color version of this figure is available only in the electronic edition.

0.824 to 0.534 km, and the depth error reduced from 1.118 to 0.650 km. Incorporating advanced statistical optimization methods to identify and mitigate outliers during the inversion process could also improve results, for example, events are deleted or downweighted during iterations if they lose connection to other events or exceed the maximum expected delay time (Waldhauser, 2001).

Quantifying uncertainties in location results is challenging due to a combination of measurement inaccuracy, computational complexity, and nonlinearity (Hardebeck and Husen, 2010). Our findings indicate that most programs tend to underestimate uncertainties, highlighting the need for improved uncertainty analysis. Although uncertainties are often assessed based on phase-picking quality and travel-time misfits, it is essential to integrate epistemic factors such as velocity biases, as well as station coverage and geometry into the analysis (Garcia-Aristizabal *et al.*, 2020). Practically, users should rely on the output of the error from the programs when interpreting results and use other aspects like azimuthal gap and the ratio of hypocentral depth to station distance together for comprehensive error analysis (Hardebeck and Husen, 2010). In addition, probabilistic analyses, particularly incorporating prior knowledge of seismic sources could provide valuable insight into reducing location uncertainty (Lomax and Savvaidis, 2019).

Although we endeavored to develop a geologically plausible model, our experiment does not fully capture the complexity of Earth's structure. Several commonly observed geological features that complicate seismic wave propagation will hinder the determination of earthquake locations, including vertical and horizontal velocity discontinuities, low-velocity layers, and fault damage zones, all of which are omitted from our synthetic test. Moreover, although our model introduces short and long-wavelength heterogeneities from the von Karman and CVMs, it may neglect intermediate wavelength heterogeneity. Anisotropy, as well as nonstationary velocity structure heterogeneity, could both be included and might be more realistically assessed through waveform modeling.

Conclusions

This study systematically evaluates the performance of eight earthquake location programs in estimating hypocenters and quantifying uncertainty. Through a realistic controlled synthetic experiment based on the 2019 Ridgecrest earthquake sequence, we show the superiority of differential time-based location methods over single-event location methods in terms of accuracy and precision. Single-event programs, NonLinLoc_SSST, VELEST, and HypoSVI, which account for station corrections by exploiting the information across all events stand out. This demonstrates the effectiveness of programs that account for the influence of velocity heterogeneity, either by direct modeling or through cancellation. Our analysis reveals the tendency for each program to underestimate uncertainty, which motivates a re-evaluation of earthquake location uncertainty characterization. We also find systematic biases in-depth estimation when using a locally unrepresentative 1D velocity profile.

By elucidating the performance of different earthquake location methodologies, we aim for this study to help guide researchers in the use of earthquake location programs, and to help developers of those programs make improvements. Toward that end, we provide usage instructions and test dataset for these eight programs in Data and Resources, where we documented, for example, our usage of HypoDD: we first include all events in one cluster a run in LSQR mode, then separately run subsets in SVD mode for uncertainty. Although our study focuses on a selected set of programs, we acknowledge that there are other earthquake location tools that we did not evaluate. We encourage others to use the provided test data to assess additional programs of their choice and contribute further to the advancement of accuracy and precision of earthquake locations.

Although this work focuses on a well-monitored earthquake sequence, diverse monitoring scenarios such as offshore events and sparse network distributions further complicate the location problems. The testing framework developed in this study provides a systematic approach to evaluate location methods across these challenging scenarios. We encourage future studies to apply this framework to quantitatively assess the performance of different location algorithms under various monitoring conditions. The synthetic testing framework established in this study offers a platform to test other earthquake location programs and builds a potential foundation for other research efforts. Most obviously, it can be readily applied to other regions and extended to broader scenarios, including regional and teleseismic scales, offering a solid basis for the exploration and refinement of earthquake location techniques.

Data and Resources

All eight earthquake location programs came from published sources listed in the references and can be accessed publicly at their host websites. The Southern California Earthquake Center (SCEC) Community velocity model (SVM) can be obtained at https://github.com/SCECcode/ucvm, the QTM catalog and stations are hosted in the Southern California Earthquake Data Center (SCEDC) at https://scedc.caltech.edu. The codes and data used in this study are hosted at https://github.com/YuYifan2000/comparison_hypoDD_GrowClust/. All websites were last accessed in November 2024.

Declaration of Competing Interests

The authors acknowledge that there are no conflicts of interest recorded.

Acknowledgments

The authors thank Peter Shearer and the anonymous reviewer for their comments and suggestions. The authors thank Guoqing Lin, Anthony Lomax, Daniel Trugman, Zachary Ross, and Felix Waldhauser for their help and guidance in running the location programs. The figures are plotted using Matplotlib (Hunter, 2007) and Generic Mapping Tools (Wessel *et al.*, 2019). This research was supported by the Statewide California Earthquake Center (Award Number 24200, Contribution Number 14069). Southern California Earthquake Center (SCEC) is funded by National Science Foundation (NSF) Cooperative Agreement EAR-2225216 and U.S. Geological Survey (USGS) Cooperative Agreement G24AC00072-00.

References

Billings, S. D., M. S. Sambridge, and B. L. N. Kennett (1994). Errors in hypocenter location: Picking, model, and magnitude dependence, *Bull. Seismol. Soc. Am.* **84,** no. 6, 1978–1990.

Brocher, T. M. (2008). Compressional and shear-wave velocity versus depth relations for common rock types in Northern California, *Bull. Seismol. Soc. Am.* **98,** no. 2, 950–968.

Douglas, A. (1967). Joint epicentre determination, *Nature* **215,** no. 5096, 47–48.

Eaton, J. P. (1970). HYPOLAYR, a computer program for determining hypocenters of local earthquakes in an earth consisting of uniform flat layers over a half space, *U.S. Geol. Surv. Tech. Rept. 69-85.*

Engdahl, E. R., R. van der Hilst, and R. Buland (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am.* **88,** no. 3, 722–743.

Frohlich, C. (1979). An efficient method for joint hypocenter determination for large groups of earthquakes, *Comput. Geosci.* **5,** no. 3, 387–389.

Garcia-Aristizabal, A., S. Danesi, T. Braun, M. Anselmi, L. Zaccarelli, D. Famiani, and A. Morelli (2020). Epistemic uncertainties in local earthquake locations and implications for managing induced seismicity, *Bull. Seismol. Soc. Am.* **110,** no. 5, 2423–2440.

Geiger, L. (1910). Herdbestimmung bei Erdbeben aus den Ankunftszeiten, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1910,** 331–349 (in German).

Gomberg, J. S., K. M. Shedlock, and S. W. Roecker (1990). The effect of S-wave arrival times on the accuracy of hypocenter estimation, *Bull. Seismol. Soc. Am.* **80,** no. 6A, 1605–1628.

Got, J.-L., J. Fréchet, and F. W. Klein (1994). Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea, *J. Geophys. Res.* **99,** no. B8**,** 15,375–15,386.

Hardebeck, J., and S. Husen (2010). Earthquake location accuracy, *Community Online Resource for Statistical Seismicity Analysis*, doi: 10.5078/corssa-55815573.

Hauksson, E., C. Yoon, E. Yu, J. R. Andrews, M. Alvarez, R. Bhadha, and V. Thomas (2020). Caltech/USGS Southern California Seismic Network (SCSN) and Southern California Earthquake Data Center (SCEDC): Data availability for the 2019 Ridgecrest sequence, *Seismol. Res. Lett.* **91,** no. 4, 1961–1970.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.* **9,** no. 3, 90–95.

Hutton, K., E. Hauksson, J. Clinton, J. Franck, A. Guarino, N. Scheckel, D. Given, and A. Yong (2006). Southern California seismic network update, *Seismol. Res. Lett.* **77,** no. 3, 389–395.

Jordan, T. H., and K. A. Sverdrup (1981). Teleseismic location techniques and their application to earthquake clusters in the South-Central Pacific, *Bull. Seismol. Soc. Am.* **71,** no. 4, 1105–1130.

Kissling, E. (1995). *Velest User's Guide*, Institute of Geophysics, ETH Zuerich.

Kissling, E., W. L. Ellsworth, D. Eberhart-Phillips, and U. Kradolfer (1994). Initial reference models in local earthquake tomography, *J. Geophys. Res.* **99,** no. B10, 19,635–19,646.

Kraft, T., A. Mignan, and D. Giardini (2013). Optimization of a large-scale microseismic monitoring network in northern Switzerland, *Geophys. J. Int.* **195,** no. 1, 474–490.

Lahr, J. C. (1979). HYPOELLIPSE: A computer program for determining local earthquake hypocentral parameters, magnitude, and first motion pattern, *U.S. Geol. Surv. Tech. Rept. 79-431*.

Lee, E.-J., P. Chen, T. H. Jordan, P. B. Maechling, M. A. M. Denolle, and G. C. Beroza (2014). Full-3-D tomography for crustal structure in Southern California based on the scattering-integral and the adjoint-wavefield methods, *J. Geophys. Res.* **119,** no. 8, 6421–6451.

Lee, W. H. K., and J. C. Lahr (1972). *HYPO71: A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes*, *U.S. Geol. Surv. Tech. Rept. 72-224*.

Lin, G. (2018). The Source-specific station term and waveform cross-correlation earthquake location package and its applications to California and New Zealand, *Seismol. Res. Lett.* **89,** no. 5, 1877–1885.

Lin, G., and P. Shearer (2005). Tests of relative earthquake location techniques using synthetic data, *J. Geophys. Res.* **110,** no. B4, doi: 10.1029/2004JB003380.

Lin, G., and P. Shearer (2006). The COMPLOC earthquake location package, *Seismol. Res. Lett.* **77,** no. 4, 440–444.

Lomax, A., and A. Savvaidis (2019). Improving absolute earthquake location in west Texas using probabilistic, proxy ground-truth station corrections, *J. Geophys. Res.* **124,** no. 11, 11,447–11,465.

Lomax, A., and A. Savvaidis (2022). High-precision earthquake location using source-specific station terms and inter-event waveform similarity, *J. Geophys. Res.* **127,** no. 1, e2021JB023190, doi: 10.1029/2021JB023190.

Lomax, A., A. Michelini, and A. Curtis (2009). Earthquake location, direct, global-search methods, in *Encyclopedia of Complexity and Systems Science*, R. A. Meyers (Editor), 1–33, Springer, New York, New York.

Lomax, A., J. Virieux, P. Volant, and C. Berge-Thierry (2000). Probabilistic earthquake location in 3D and layered models, in *Advances in Seismic Event Location*, C. H. Thurber and N. Rabinowitz (Editors), Modern Approaches in Geophysics, 101–134, Springer, Dordrecht, Netherlands.

Nakata, N., and G. C. Beroza (2015). Stochastic characterization of mesoscale seismic velocity heterogeneity in Long Beach, California, *Geophys. J. Int.* **203,** no. 3, 2049–2054.

Park, Y., S. M. Mousavi, W. Zhu, W. L. Ellsworth, and G. C. Beroza (2020). Machine-learning-based analysis of the Guy-Greenbrier, Arkansas earthquakes: A tale of two sequences, *Geophys. Res. Lett.* **47,** no. 6, e2020GL087032, doi: 10.1029/2020GL087032.

Poupinet, G., W. L. Ellsworth, and J. Frechet (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras Fault, California, *J. Geophys. Res.* **89,** no. B7, 5719–5731.

Pyle, M. L., T. Chen, L. Preston, M. Scalise, C. Zeiler, and K. D. Smith (2023). How good is your location? Comparing and understanding the uncertainties in location for the 1993 Rock Valley sequence, *Seism. Rec.* **3,** no. 4, 259–268.

Richards-Dinger, K. B., and P. M. Shearer (2000). Earthquake locations in southern California obtained using source-specific station terms, *J. Geophys. Res.* **105,** no. B5, 10,939–10,960.

Ross, Z. E., B. Idini, Z. Jia, O. L. Stephenson, M. Zhong, X. Wang, Z. Zhan, M. Simons, E. J. Fielding, S.-H. Yun, *et al.* (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence, *Science* **366,** no. 6463, 346–351.

Rowe, C. A., R. C. Aster, B. Borchers, and C. J. Young (2002). An automatic, adaptive algorithm for refining phase picks in large seismic data sets, *Bull. Seismol. Soc. Am.* **92,** no. 5, 1660–1674.

Savvaidis, A., B. Young, G. D. Huang, and A. Lomax (2019). TexNet: A statewide seismological network in Texas, *Seismol. Res. Lett.* **90,** no. 4, 1702–1715.

Schaff, D. P., and F. Waldhauser (2005). Waveform cross-correlation-based differential travel-time measurements at the Northern California Seismic Network, *Bull. Seismol. Soc. Am.* **95,** no. 6, 2446–2461.

Schaff, D. P., G. H. R. Bokelmann, W. L. Ellsworth, E. Zanzerkia, F. Waldhauser, and G. C. Beroza (2004). Optimizing correlation techniques for improved earthquake location, *Bull. Seismol. Soc. Am.* **94,** no. 2, 705–721.

Shelly, D. R. (2020). A high-resolution seismic catalog for the initial 2019 Ridgecrest earthquake sequence: Foreshocks, aftershocks, and faulting complexity, *Seismol. Res. Lett.* **91,** no. 4, 1971–1978.

Small, P., D. Gill, P. J. Maechling, R. Taborda, S. Callaghan, T. H. Jordan, K. B. Olsen, G. P. Ely, and C. Goulet (2017). The SCEC unified community velocity model software framework, *Seismol. Res. Lett.* **88,** no. 6, 1539–1552.

Smith, J. D., K. Azizzadenesheli, and Z. E. Ross (2021). EikoNet: Solving the Eikonal equation with deep neural networks, *IEEE Trans. Geosci. Remote Sens.* **59,** no. 12, 10,685–10,696.

Smith, J. D., Z. E. Ross, K. Azizzadenesheli, and J. B. Muir (2022). HypoSVI: Hypocentre inversion with Stein variational inference and physics informed neural networks, *Geophys. J. Int.* **228,** no. 1, 698–710.

Syracuse, E. M., and G. A. Abers (2009). Systematic biases in subduction zone hypocenters, *Geophys. Res. Lett.* **36,** no. 10, doi: 10.1029/2009GL037487.

Tan, Y. J., F. Waldhauser, W. L. Ellsworth, M. Zhang, W. Zhu, M. Michele, L. Chiaraluce, G. C. Beroza, and M. Segou (2021). Machine-learning-based high-resolution earthquake catalog reveals how complex fault structures were activated during the 2016–2017 Central Italy sequence, *Seism. Rec.* **1,** no. 1, 11–19.

Tarantola, A., and B. Valette (1981). Inverse problems = Quest for information, *J. Geophys. Res.* **50,** no. 1, 159–170.

Thurber, C. H. (1985). Nonlinear earthquake location: Theory and examples, *Bull. Seismol. Soc. Am.* **75,** no. 3, 779–790.

Thurber, C., S. Roecker, W. Ellsworth, Y. Chen, W. Lutter, and R. Sessions (1997). Two-dimensional seismic image of the San Andreas Fault in the Northern Gabilan Range, central California: Evidence for fluids in the fault zone, *Geophys. Res. Lett.* **24,** no. 13, 1591–1594.

Thurber, C., S. Roecker, H. Zhang, S. Baher, and W. Ellsworth (2004). Fine-scale structure of the San Andreas fault zone and location of the SAFOD target earthquakes, *Geophys. Res. Lett.* **31,** no. 12, doi: 10.1029/2003GL019398.

Trugman, D. T., and P. M. Shearer (2017). GrowClust: A hierarchical clustering algorithm for relative earthquake relocation, with application to the Spanish Springs and Sheldon, Nevada, earthquake sequences, *Seismol. Res. Lett.* **88,** no. 2A, 379–391.

Trugman, D. T., C. J. Chamberlain, A. Savvaidis, and A. Lomax (2022). GrowClust3D.jl: A Julia package for the relative relocation of earthquake hypocenters using 3D velocity models, *Seismol. Res. Lett.* **94,** no. 1, 443–456.

Waldhauser, F. (2001). hypoDD – A program to compute double-difference hypocenter locations, *U.S. Geol. Surv. Open-File Rept. 2001-113*, doi: 10.3133/ofr01113.

Waldhauser, F. (2009). Near-real-time double-difference event location using long-term seismic archives, with application to Northern California, *Bull. Seismol. Soc. Am.* **99,** no. 5, 2736–2748.

Waldhauser, F., and W. L. Ellsworth (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California, *Bull. Seismol. Soc. Am.* **90,** no. 6, 1353–1368.

Waldhauser, F., W. S. D. Wilcock, M. Tolstoy, C. Baillard, Y. J. Tan, and D. P. Schaff (2020). Precision seismic monitoring and analysis at axial seamount using a real-time double-difference system, *J. Geophys. Res.* **125,** no. 5, e2019JB018796, doi: 10.1029/2019JB018796.

Wessel, P., J. F. Luis, L. Uieda, R. Scharroo, F. Wobbe, W. H. F. Smith, and D. Tian (2019). The Generic Mapping Tools Version 6, *Geochem. Geophys. Geosys.* **20,** no. 11, 5556–5564.

Wesson, R. L. (1971). Travel-time inversion for laterally inhomogeneous crustal velocity models, *Bull. Seismol. Soc. Am.* **61,** no. 3, 729–746.

Wesson, R. L., J. C. Roller, and W. H. K. Lee (1973). Time-term analysis and geological interpretation of seismic travel-time data from the Coast Ranges of central California, *Bull. Seismol. Soc. Am.* **63,** no. 4, 1447–1471.

White, M. C. A., H. Fang, N. Nakata, and Y. Ben-Zion (2020). PyKonal: A Python package for solving the Eikonal equation in spherical and cartesian coordinates using the fast marching method, *Seismol. Res. Lett.* **91,** no. 4, 2378–2389.

Wilding, J. D., W. Zhu, Z. E. Ross, and J. M. Jackson (2023). The magmatic web beneath Hawai'i, *Science* **379,** no. 6631, 462–468.

Wuestefeld, A., S. M. Greve, S. P. Näsholm, and V. Oye (2018). Benchmarking earthquake location algorithms: A synthetic comparison, *Geophysics* **83,** no. 4, KS35–KS47.

Appendix

Appendix provides additional materials supporting this study. Table A1 presents the additional results from running location programs while incorporating the correct station elevations.

TABLE A1
Summary of Performance Metrics for Eight Earthquake Location Programs Under Correct Station Elevation

	Mean Accuracy Error (km)		Chamfer	Mean Precision Error (km)	
Method	**Horizontal**	**Depth**	**Distance**	**Horizontal**	**Depth**
HYPOINVERSE	0.824	1.118	1.617	0.571	0.684
VELEST	0.696	0.559	1.170	0.380	0.656
NonLinLoc	0.953	0.969	1.626	0.580	0.694
NonLinLoc_SSST	0.440	0.538	1.071	0.158	0.307
HypoSVI	0.429	0.338	0.828	0.317	0.383
HypoDD	0.315	0.571	0.974	0.080	0.141
XCORLOC	0.746	0.818	1.507	0.101	0.256
GrowClust	0.846	1.066	1.701	0.148	0.304

Manuscript received 6 September 2024
Published online 3 December 2024