Bulletin of the Seismological Society of America, Vol. 105, No. 4, pp. –, August 2015, doi: 10.1785/0120140298

Real-Time Back Azimuth for Earthquake Early Warning

by Andreas S. Eisermann, Alon Ziv, and Gilles H. Wust-Bloch

Abstract The incentive to speed up real-time location has motivated previous researchers to go beyond standard location procedures and use not only *P*-wave arrival at some network stations but also its nonarrival at others. In addition to being sensitive to velocity model and picking uncertainties, this approach is also highly dependent on time delays due to unknowns network latencies, processing, and packet size. Thus, seeking ways to add independent real-time constraints on earthquake location are important for earthquake early warning applications. In this study, we assess the robustness of three independent real-time back-azimuth (BAZ) determination schemes, using offline records of southern California earthquakes. We find that BAZ values computed by the three methods provide equivalent levels of accuracy. By sending the three BAZ estimates to a screening module that checks for coherency and signal-to-noise ratio criteria, we show that accurate BAZ estimates are obtainable in real time, with a standard deviation of 13°. Through examination of two earthquake scenarios that use offline data, we show that the inclusion of BAZ estimates into real-time location schemes improves the performance of real-time hypocenter determination, by cutting the time it takes to obtain well-constrained hypocenters.

Introduction

The accuracy of real-time magnitude assessment and shaking prediction is limited by that of the source location. Thus, seeking ways to improve real-time earthquake location are crucial for earthquake early warning (EEW) applications. Although hypocenter determination using four or more triggers is standard practice, constraining hypocenter location in real time with fewer detections remains a challenging task (e.g., Nakamura, 1988; Ruud *et al.*, 1988; Alessandrini *et al.*, 1994; Rydelek and Pujol, 2004; Horiuchi *et al.*, 2005; Lockman and Allen, 2005; Cua and Heaton, 2007; Satriano *et al.*, 2008; Noda *et al.*, 2012). In this article, we use offline earthquake data from southern California to examine how real-time back azimuth (BAZ) may improve the performance of real-time location algorithms.

Previous studies show that real-time BAZ obtained through 3D particle motion analysis of the *P* phases at individual stations are subject to uncertainties that are unacceptably large for EEW applications. Lockman and Allen (2005) tested the accuracy of real-time BAZ in southern California, calculated using the first 0.5 s of the *P* wave; they found that the accuracy of the BAZ calculation is highly station dependent and that the level of accuracy at 75% of the network stations is insufficient for EEW algorithms utilizing 1–3 triggers. Using strong-motion data recorded by the K-NET network in Japan, Noda *et al.* (2012) showed that BAZ estimates may be improved by setting the length of the time windows to be equal to the interval between the first *P*-wave arrival and the first zero crossing of the vertical displacement seismogram. Despite the notable improvement (and time gain) with respect to the fixed interval approach, 40% of the estimates have errors that are $>30°$ (Noda *et al.*, 2012).

After describing the data used in this study, we outline a scheme for real-time BAZ determination using 3D ground motions recorded at individual stations. Then, through examples from southern California, we assess how the integration of BAZ estimates stabilizes epicenter determination. The results of this study indicate that accurate BAZ estimates are obtainable in real time, and that inserting these real-time BAZ estimates into existing hypocenter determination schemes (e.g., Horiuchi *et al.*, 2005; Satriano *et al.*, 2008) enhances their performances.

Data

The data used in this study consist of 114 sets of three-component waveforms of 17 southern California earthquakes from the California Integrated Seismic Network (CISN) database (Table 1), with magnitudes between 3.2 and 5.7 (Fig. 1). These were recorded between 1992 and 2010 by different types of sensors (4 short-period seismometers, 82 broadband seismometers, and 20 accelerometers) within 90 km of the epicenter. *P*-wave arrivals were picked automatically using the FilterPicker algorithm of Lomax *et al.* (2012). Acceleration time series were integrated once. All velocity time series were band-pass filtered between 1 and 2 Hz. In this study, BAZ were calculated using velocity time series.

Table 1
Event List

Date (yyyy/mm/dd)	Time (hh:mm:ss.sss)	Latitude (°)	Longitude (°)	Magnitude
1992/06/28	12:01:16.191	−116.32	34.12	5.7
1992/08/15	08:24:14.657	−116.40	34.09	4.7
1994/01/17	12:39:39.795	−118.54	34.27	4.9
1997/03/18	15:24:47.724	−116.82	34.97	5.3
2002/06/14	12:40:45.360	−116.29	36.72	4.6
2005/03/13	23:43:15.300	−116.85	33.30	3.5
2005/12/05	22:01:20.240	−116.32	34.85	3.2
2005/06/16	20:53:26.020	−117.01	34.06	4.9
2005/07/16	04:58:31.740	−116.32	34.84	3.8
2005/10/16	21:11:35.010	−118.16	32.45	5.0
2008/07/29	18:42:15.710	−117.76	33.95	5.4
2009/05/18	03:39:36.340	−118.34	33.94	4.7
2009/03/10	11:37:11.380	−117.87	36.39	4.4
2009/11/11	03:24:23.390	−116.05	33.25	3.6
2010/09/04	23:05:42.000	−115.25	32.18	4.7
2010/08/05	13:47:03.380	−116.05	33.27	3.4
2010/07/07	23:53:33.530	−116.49	33.42	5.4

BAZ, however, may also be calculated using acceleration or displacement time series.

Real-Time BAZ Determination

Here, we describe a new real-time BAZ determination scheme, for which accuracy level is acceptable for EEW applications. This scheme runs simultaneously three independent BAZ determination modules and then screens averaged results using signal-to-noise ratio (SNR) and coherency criteria (Fig. 2a–d). Following Noda *et al.* (2012), the data window that we use starts at the time of the first *P*-wave detection and ends at the first meaningful extremum in the vertical component (Fig. 2a). The latter is picked automatically by requiring that it exceeds 10 times the SNR level. Hereafter, we refer to the first *P* arrival and first meaningful extremum in the vertical component as $n1$ and $n2$, respectively. The $n2 - n1$ interval is typically between 0.2 and 0.3 s. In the following, we detail each of the three BAZ determination modules, analyze their respective results, and present the postscreening end product.

Single-Value Module

According to the single-value (SV) approach (Fig. 2b), BAZ is related to the 3D ground-motion amplitudes as

$$\mathrm{BAZ} = f[\phi(A_E, A_N), \mathrm{sign}(A_N), \mathrm{sign}(A_Z)], \tag{1a}$$

with

$$\phi = \arctan\left(\frac{A_E}{A_N}\right) + \pi, \tag{1b}$$

in which A_Z, A_N, and A_E are the amplitudes of the velocity seismogram along the vertical, north, and east directions, respectively, and the function f accounts for the 180° ambiguity in ϕ according to

$$f(\phi) = \begin{cases} \phi + \pi, & \text{if } A_Z \times A_N < 0, \\ \phi, & \text{otherwise.} \end{cases} \tag{1c}$$

The derivation of the above flipping condition is detailed in Appendix A. Because the SV method employs instantaneous 3D ground-motion amplitude, its result is highly sensitive to the exact manner by which that time sample is selected, and special care must be paid when choosing a good data sample. We, therefore, choose the data point corresponding to the highest horizontal amplitude within the $n1$–$n2$ data interval. The index of this data point is labeled as $n\mathrm{win}$ and is obtained using

$$n\mathrm{win} = \begin{cases} nN, & |A_N(nN)| > |A_E(nE)|, \\ nE, & \text{otherwise} \end{cases} \tag{2a}$$

with

$$nN = \underset{n1 \le n \le n2}{\mathrm{argmax}} |A_N(n)| \quad \text{and} \quad nE = \underset{n1 \le n \le n2}{\mathrm{argmax}} |A_E(n)|, \tag{2b}$$

in which nN and nE are the sample indexes at which the ground velocity amplitudes A_N and A_E are largest, respectively. Given the properties of the arctan2 function, small data errors in low SNR conditions may result in disproportionally large BAZ errors. Thus, the accuracy of the BAZ estimates is expected to be highly SNR sensitive. To identify a threshold SNR, above which large BAZ errors are infrequent, it is useful to examine the distribution of the BAZ errors as a function of the SNR. Throughout this study, the SNR are quantified as

$$\mathrm{SNR} = \min[\mathrm{SNR}_{\mathrm{ver}}, \mathrm{SNR}_{\mathrm{hor}}] \tag{3a}$$

with

$$\mathrm{SNR}_{\mathrm{ver}} = \frac{\sqrt{\frac{1}{n\mathrm{win}-n1}\sum_{n=n1}^{n\mathrm{win}} A_Z^2(n)}}{\sqrt{\frac{1}{n1}\sum_{n=0}^{n1} A_Z^2(n)}}$$

$$\text{and} \quad \mathrm{SNR}_{\mathrm{hor}} = \frac{\sqrt{\frac{1}{n\mathrm{win}-n1}\sum_{n=n1}^{n\mathrm{win}} [A_E^2(n) + A_N^2(n)]}}{\sqrt{\frac{1}{n1}\sum_{n=0}^{n1} [A_E^2(n) + A_N^2(n)]}}. \tag{3b}$$

The results of the SV approach are summarized in Figure 3a,d. A generally good agreement is found between the real-time and the catalog BAZ, with 39% of the estimates falling within ±10° of the catalog BAZ and 67% within ±20° of the catalog BAZ. In addition, the fraction of outliers decreases with increasing SNR.

Moving Average Module

The moving average (MA) approach introduced by Nakamura (1988) is equivalent to a continuously applied form of SV, with BAZ computed stepwise at every sample (Fig. 2c) according to

$$\mathrm{BAZ}_n = g[\theta^n(R_{ZE}^n, R_{ZN}^n), \mathrm{sign}(R_{ZN}^n)], \tag{4a}$$

with

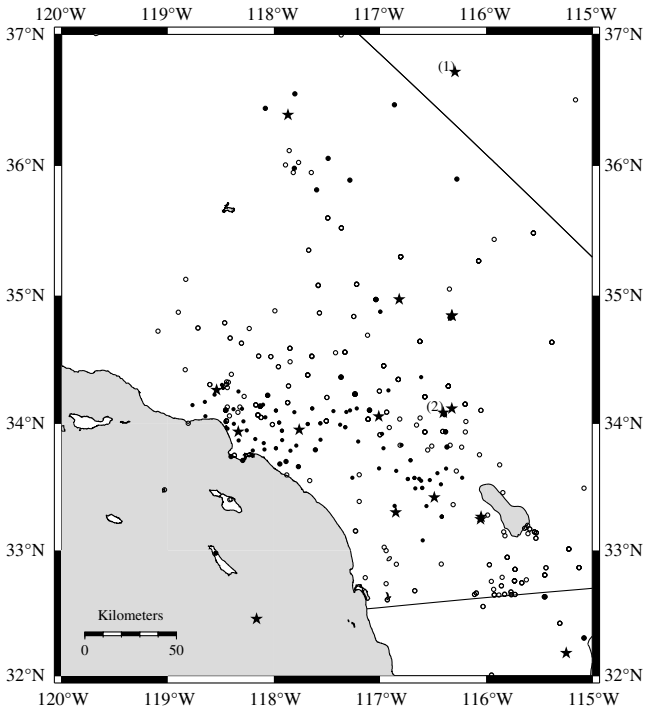

Figure 1. Location map. California Integrated Seismic Network (CISN) stations and earthquake locations are shown as circles and stars, respectively. The white circles indicate all triggered stations and the solid circles indicate stations that provided three-component data for the back-azimuth (BAZ) analysis. The two example earthquakes are labeled as (1) and (2).

$$\theta^n = \arctan\left(\frac{R^n_{ZE}}{R^n_{ZN}}\right) + \pi \tag{4b}$$

and

$$R^n_{ZE} = \alpha R^{n-1}_{ZE} + A^n_Z A^n_E \quad \text{and} \quad R^n_{ZN} = \alpha R^{n-1}_{ZN} + A^n_Z A^n_N, \tag{4c}$$

in which the function g, accounts for the 180° ambiguity in ϕ according to

$$g(\phi) = \begin{cases} \phi + \pi, & \text{if } R_{ZN} < 0, \\ \phi, & \text{otherwise} \end{cases} \tag{4d}$$

and α is a smoothing coefficient with a value close to unity. Failing to resolve the sign of R_{ZN} and R_{ZE} may introduce large errors, as it can flip the BAZ by up to 180°. Thus, here, we slightly modify Nakamura's original approach, by using $n1$–$n2$ data windows and by taking the estimate at $n2$, rather than the average value between $n1$ and $n2$. Further, we exclude samples that satisfy the following muting condition:

$$\min[\mathrm{SNR}_{\mathrm{ver}}, \mathrm{SNR}_{\mathrm{hor}}] < 1. \tag{5}$$

We find that the exclusion of these low-SNR data segments yields better estimates for the smallest magnitudes ($3 < \mathbf{M} < 4$). The relationship between the real-time and catalog BAZ for MA yields a median and a mean error of 12.3° and 31.1°, respectively (Fig. 3b). Here again, the fraction of outliers decreases with increasing SNR (Fig. 3e).

Principal Component Analysis

Two-dimensional principal component analysis (PCA) may be used to estimate the horizontal projection of the particle motion first principal component (Noda *et al.*, 2012; Fig. 2d). We construct a data matrix as follows:

$$\mathbf{D}_n = \begin{bmatrix} A^{n1}_E & \dots & A^{n2}_E \\ A^{n1}_N & \dots & A^{n2}_N \end{bmatrix}. \tag{6}$$

As before, we exclude samples that satisfy the muting condition in equation (5). The adjusted data matrix is then given by

$$A_n = \begin{bmatrix} A^{n1}_E - \langle A^n_E \rangle & \dots & A^{n2}_E - \langle A^n_E \rangle \\ A^{n1}_N - \langle A^n_N \rangle & \dots & A^{n2}_N - \langle A^n_N \rangle \end{bmatrix}, \tag{7}$$

in which the angular brackets signify the average of all samples between $n1$ and $n2$. Following singular value decomposition of $\mathbf{A}\mathbf{A}^T$, we get

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\lambda^2\mathbf{U}^T, \tag{8}$$

in which λ is a diagonal matrix containing the eigenvalues arranged in descending order, and $\mathbf{U}$ is a square matrix containing the eigenvectors corresponding to the eigenvalues in λ. The first principal component is the first row in $\mathbf{U}$. The approach for resolving the azimuthal ambiguity is detailed in Appendix B. We find that the PCA method yields an accuracy level that is similar to those of the SV and MA approaches (Fig. 3c), and that the discrepancy between real-time and catalog BAZ estimates decreases with increasing SNR (Fig. 3f).

Screening Module

Recall that SV calculates BAZ on the sample of highest horizontal amplitude within the data interval, MA averages over the entire data interval giving increasingly more weight to the latest samples and PCA gives equal weight to all data points within the data interval (Fig. 2a–d). Thus, when the three waveform components oscillate coherently (i.e., the fluctuations at all three traces are in phase), the three methods yield very similar results. When the three components oscillate initially coherently and then lose coherency, the SV module is expected to give better results than the MA. When the oscillation is initially incoherent and becomes coherent later on, the MA is expected to yield better result than the SV. Thus, a very good agreement among the three BAZ modules strongly implies that the BAZ estimation is reliable.

The results show that BAZ values computed by the three independent (SV, MA, and PCA) modules provide equivalent

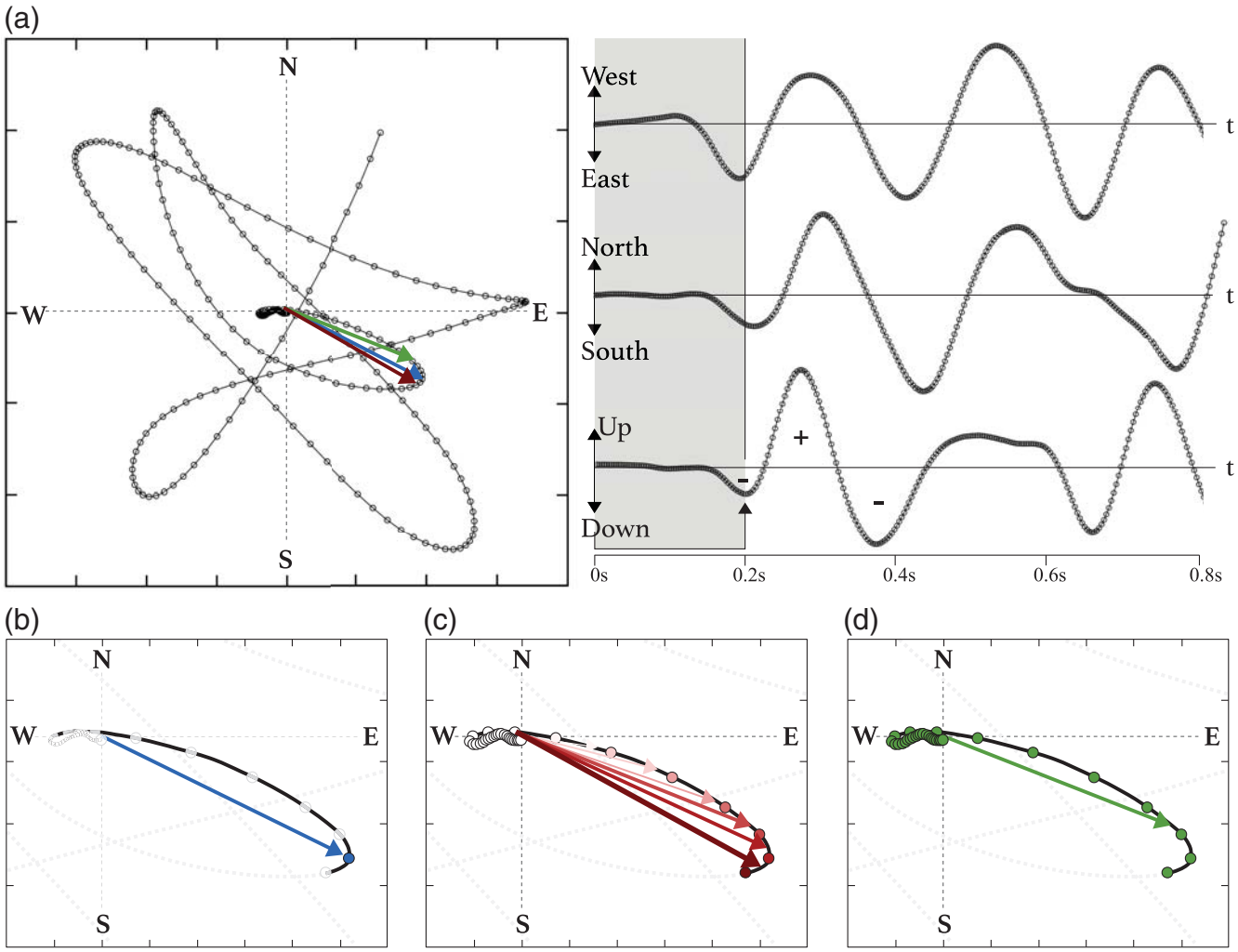

Figure 2. Diagrams illustrating the three BAZ determination schemes used in this study. (a) Particle motion of an example velocity seismogram projected onto the north–south-east–west plane and the three corresponding traces. Short data intervals are used (indicated by the gray area), because the direction of the particle motion varies rapidly with time. The data interval begins with the first *P*-wave arrival and ends at the first significant extremum on the vertical trace. (b) According to the single-value (SV) scheme, BAZ (blue arrow) is evaluated at a single data sample corresponding to the highest horizontal amplitude within the data interval (indicated by the blue circle). (c) According to the moving average (MA) scheme, BAZ (red arrow) is determined stepwise at every sample according to equation (4a)–(4d), with a weighting scheme (equation 4c) that gives increasingly more weight (indicated by increasingly thicker red arrows) to later data points. (d) In the principal component analysis (PCA), all data points are given equal weight, and the resultant vector (green arrow) corresponds to the total average of all samples. Because of the intrinsic differences between the three BAZ determination schemes, different schemes yield different estimates, and the coherency level of the three solutions is used as quality measure of the BAZ estimate.

levels of accuracy (Fig. 4). To further increase BAZ reliability and eliminate outliers, we screen the three BAZ estimates through some predetermined SNR and coherency criteria. The real-time and catalog BAZ relationship shows significant improvement when the following two screening conditions are met: SNR > 5 and all three BAZ estimations are within 3° (Figs. 4d and 5). The application of these two criteria disqualifies extreme outliers ($> 30°$) and contributes to reducing the fraction of moderate (10°–30°) BAZ errors (Fig. 4d). It results in a nearly 50% data rejection, two-thirds of which have SNR that is < 5 and are therefore irrelevant for EEW applications (Fig. 5). We conclude that computing three independent BAZ and screening them with coherency and SNR criteria provides robust real-time BAZ estimates, with a standard deviation of 13°.

Integrating BAZ into Real-Time Location Schemes

After having demonstrated that accurate BAZ estimates are obtainable in real time, we now show the inclusion of BAZ estimates into real-time location schemes enhances their performance.

The BAZ-Included Location Algorithm

The incentive to speed up location procedures requires going beyond standard location procedures and uses not only

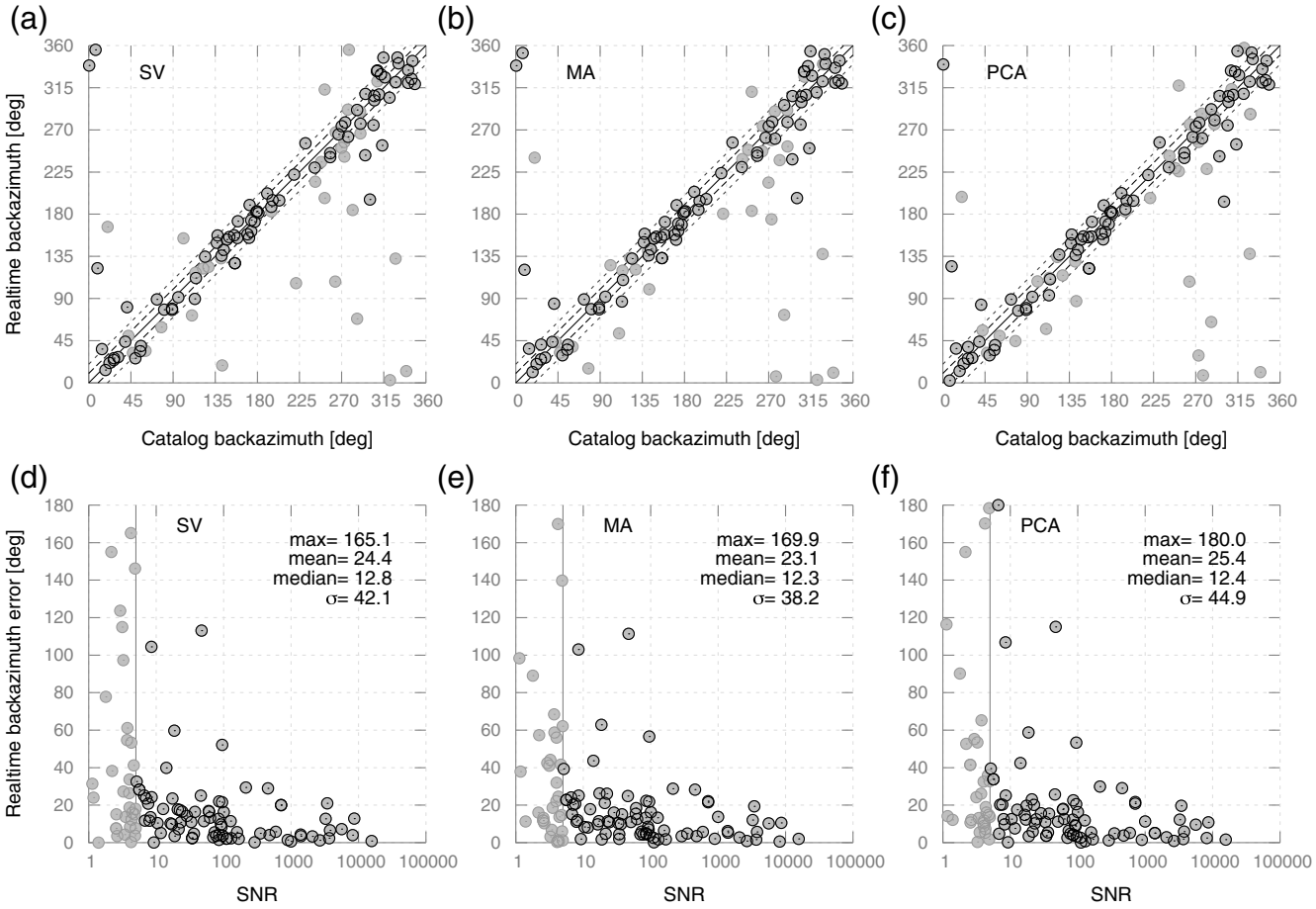

Figure 3. Diagrams summarizing the results of the three BAZ determination modules—SV, MA, and PCA. (a–c) Real-time BAZ versus catalog BAZ and (d–f) BAZ error versus signal-to-noise ratio (SNR). The gray circles indicate data with SNR < 5. Mean errors and standard deviations are reported for each BAZ method.

P-wave arrival at some network stations but also its nonarrival at others (Horiuchi *et al.*, 2005). Given a set of nonarrivals, the following set of equations may be solved (modified after Horiuchi *et al.*, 2005):

$$T_{i1}^{\text{arrival}} - T^{\text{now}} \geq T_{i1}^{\text{travel}}(\xi_{i1}, \eta) - T_k^{\text{travel}}(\xi_k, \eta) - T_k^{\text{delay}} | k \neq i1 \tag{9}$$

with $i1$ being the index of the station to which the *P* wave has arrived first and k being the index of stations to which the *P* wave has not yet arrived. The *P*-wave arrival and travel times are T^{arrival} and T^{travel}, respectively; and ξ and η are the station and earthquake coordinates, respectively; T^{now} is the clock time; and T^{delay} is the unknown time delay due to the data packet being of finite length, network latency, and processing. Use of the not-yet-arrived data (NYAD) yields a volume surrounding the first-triggered station, where dimensions decrease with elapsed time. If T^{delay} were equal to zero, NYAD volumes would start out as Voronoi polygons, reducing their size over time. In practice, however, T^{delay} is unknown at the time of T^{now}, and the initial NYAD volume may be significantly larger and geometrically complex than a standard Voronoi polygon (Fig. 6). Another consequence of nonzero T^{delay} is that the first-triggered station may not correspond to the station at which the *P* wave has first arrived (i.e., $i1$ in equation 9). To safeguard from misidentifying the station at which the *P* wave has first arrived, it is important to constantly monitor the network state-of-health and T^{delay}. Such T^{delay} estimates may be used to decide whether it is safe to identify $i1$ as the first-triggered station. In practice, the difference $T^{\text{delay}} - T^{\text{now}}$ in equation (9) is replaced by the last received time stamp of station k, thus avoiding the use of predicted latency and theoretical delay times. It is expected that data transfer delays will be reduced in the future, and NYAD estimates will become more robust. In the meantime, however, they should be used with great caution.

Once a second *P* wave is detected, the hypocenter may be further constrained to a zone in which the NYAD volume and the equal differential time (EDT) surface intersect. Given $\mathbf{M} \geq 2$ *P*-wave arrival times, a set of EDT equations may be written as

$$T_i^{\text{arrival}} - T_j^{\text{arrival}} = T_i^{\text{travel}}(\xi_i, \eta) - T_j^{\text{travel}}(\xi_j, \eta) | i \neq j, \tag{10}$$

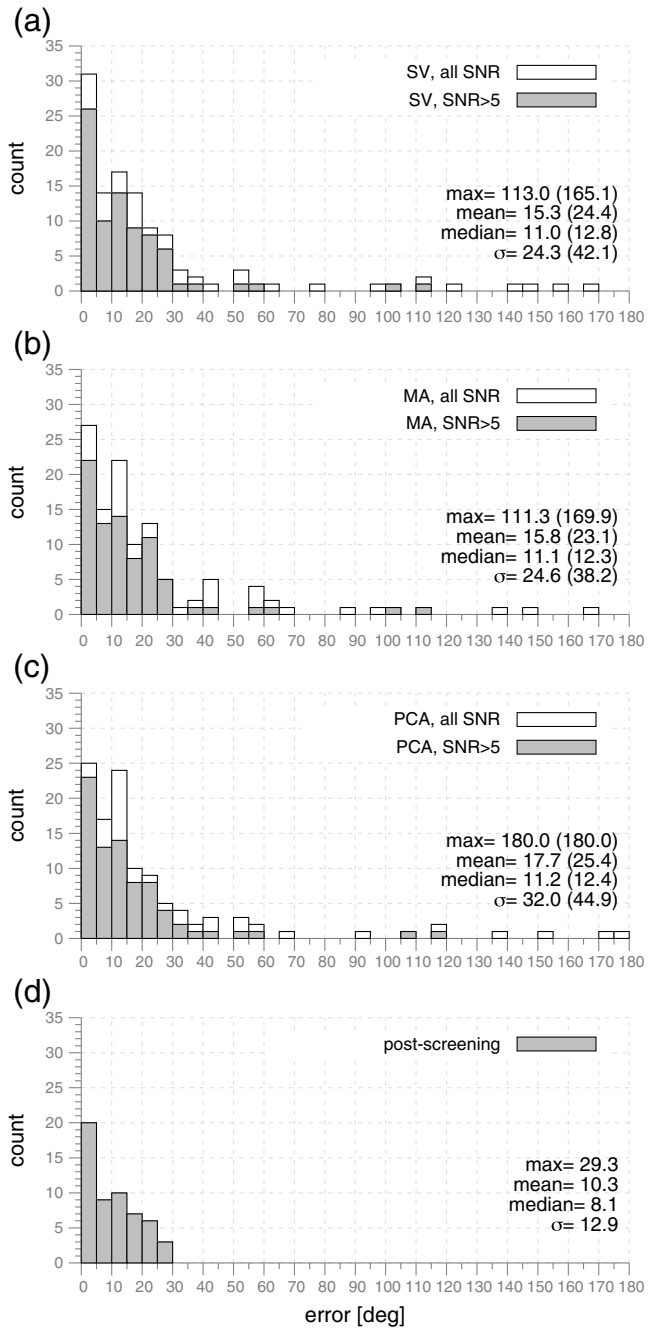

Figure 4. Histograms showing the distribution of BAZ errors for SV, MA, PCA modules, and postscreening. The white and gray columns show the entire dataset and data with SNR > 5, respectively. Mean error and error standard deviation for SNR > 5 and for the entire dataset (in parentheses) are reported on each panel.

in which i and j are the station indexes. Solution of the above equation yields a set of $\mathbf{M}(\mathbf{M}-1)/2$ EDT surfaces. With $\mathbf{M}$ 2, the hypocenter may be constrained to lie on a single-curved surface, for which thickness scales with the uncertainties in the picking and the velocity model. With $\mathbf{M}$ 3, the hypocenter may be constrained to lie within the intersection zone of three such EDT surfaces.

To obtain reliable volume constraints, both velocity model and picking uncertainties are introduced into the travel-time dependent equations. Based on equations (9) and (10), the intervals of minimum and maximum possible travel time at a point in space are compared with the uncertainty windows of picking time. This results in conservative constraint volumes, which may be notably larger than those defined by a standard approach and ensure that the location is always enclosed by the bounding constraint volume. The intersection of all given constraints at a point in time defines a solution volume, which provides accurate estimates of minimum and maximum hypocentral distances with respect to the first-triggered station. These distances, together with an appropriate attenuation law, may be used to progressively evaluate a minimal and a maximal possible event magnitude.

BAZ constraints are nearly independent of the velocity model and are very likely to become available at the same time as the NYAD information (at the time of the first P-wave detection). In light of the previously mentioned limitations of the NYAD approach, and the fact that both the NYAD and the EDT constraints are model dependent, it is expected that the integration of BAZ information into location schemes implementing these approaches (e.g., Horiuchi *et al.*, 2005; Satriano *et al.*, 2008) will enhance their performance and robustness (Joswig, 2008). The new strategy proposed here is to intersect the NYAD–EDT volumes (Horiuchi *et al.*, 2005; Satriano *et al.*, 2008) with real-time BAZ beams, for which aperture is set to ±15°. In case the screening module rejects the BAZ estimation, the real-time location scheme proceeds without it, using standard NYAD + EDT constraints. Next, we present the results of two example earthquakes, each is highlighting different aspects of the BAZ-included location scheme, and examine the extent to which the addition of the BAZ information improves the location assessment.

Example Earthquakes

Off-Network Event

The 14 June 2002 M_w 4.6 earthquake (star number 1 in Fig. 1) occurred in Nevada, outside the CISN. The actual epicenter is 59 km away from the first-triggered station. The scenario presented here is calculated using $T^{\text{delay}} = 4$ s (Fig. 7). It starts with an initially large NYAD-constrained volume that extends deep into Nevada. A consequence of the poorly constrained NYAD volume is that the maximum-estimated distance reaches 302 km, overestimating the actual distance by more than 240 km (Fig. 7a). The first BAZ is available 0.35 s later, instantaneously reducing the maximal epicentral distance from 269 to 172 km (Fig. 7b). The NYAD volume shrinks progressively, further lowering the maximal epicentral distance down to 111 km, just before the second P-wave arrival (Fig. 7c). Up to this point, the lower epicentral distance remains unconstrained. At 5.65 s, the EDT surface, defined by the second P-wave arrival, bounds the upper and lower epicentral distance to 107 and 44, respectively

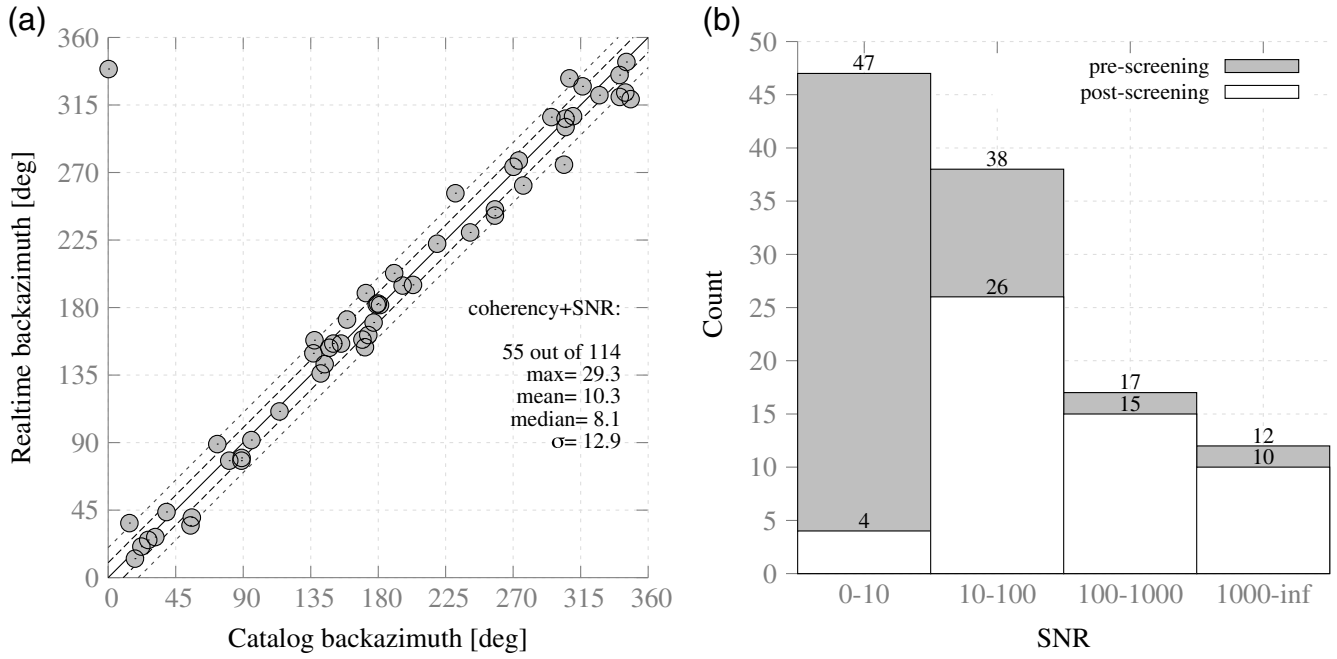

Figure 5. Summary of postscreening results. (a) Real-time BAZ versus catalog BAZ. The black-outlined circles indicate SNR > 5. (b) Histograms showing SNR distribution of postscreening BAZ. About 75% of SNR > 5 data meets the coherency criterion.

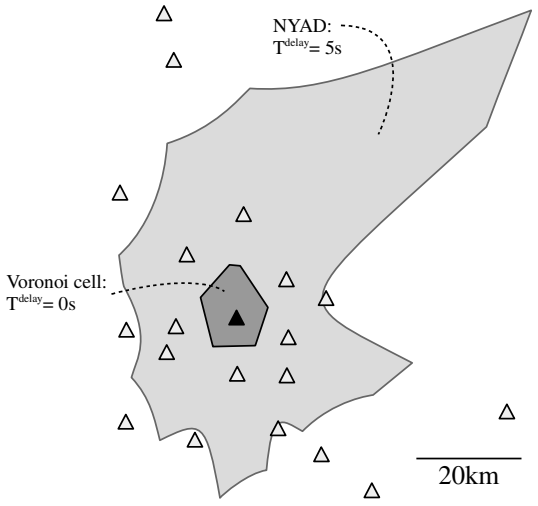

Figure 6. Diagram showing the effect of T^{delay} on a not-yet-arrived data (NYAD)-constrained volume. In a delay-free scenario, the NYAD cell starts out as Voronoi cell (dark gray) around first detection (black triangle). As T^{delay} increases, the initial cell evolves into a larger area, constructed by concave equal differential time (EDT) surfaces (light gray).

(Fig. 7d). The incorporation of the BAZ constraint raises the lower distance from 21 to 44 km, just 15 km away from the actual epicenter. At 6.05 s, the second BAZ intersects the first BAZ and the EDT surface, further confirming the solution (Fig. 7e). The third *P*-wave arrival, at 7.15 s, further constraints the event between 63 and 46 km (Fig. 7f). In this specific example, in which the first three triggered stations are almost aligned, the intersection of the three EDT surfaces defines two possible solutions, for which ambiguity can be resolved by two independent and robust BAZ estimates.

Intranetwork Event

The 15 August 1992 M_L 4.73 earthquake (star number 2 in Fig. 1) took place in the Mojave Desert, within an area of high CISN station density. Here too, the location scenario is calculated using $T^{\text{delay}} = 4$ s. The first detection occurred at a station that is located only 2 km away from the epicenter (Fig. 8a). Despite dense station coverage, the 4 s delay produces an oversized initial NYAD-constrained volume that already includes the eight nearest stations (light gray), thus resulting in an overestimation of the maximum distance by more than 85 km. The first BAZ is available 0.08 s later, instantaneously halving the maximum distance down to 45.7 instead of 87.8 km in an NYAD-only scenario (Fig. 8b). The NYAD-constrained volume progressively shrinks, lowering the maximum distance to 23.2 km (Fig. 8c), until the second detection at 2.3 s. Because of the particular event-stations layout and uncertainties in the velocity model, the width of the highly bent EDT surface is larger than the size of the NYAD-constrained cell in the direction of the maximum distance; no additional constraints are temporary gained (Fig. 8d). The second BAZ is available at 2.4 s, adding stability but no further constraints. At this stage, the minimum distance is not yet constrained. The third *P*-wave arrival at 2.95 s adds two additional EDT surfaces. Based on EDT surfaces only, the epicenter would be poorly constrained despite three available EDTs, due to the first EDT's curvature and wide extent. The added BAZ constraint drastically reduces the solution volume, determines the first minimal dis-

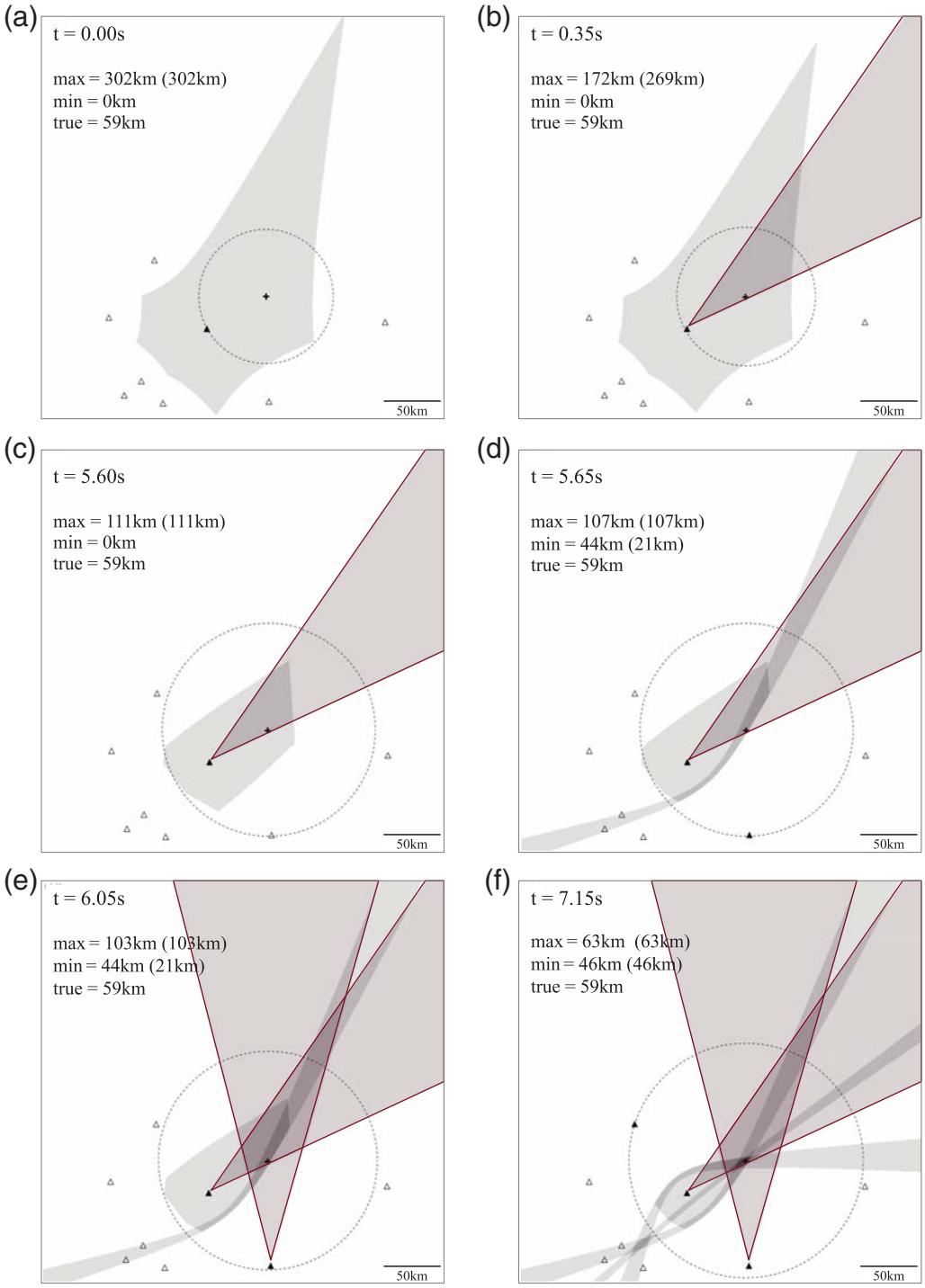

Figure 7. Snapshots of location steps for the off-network scenario. (a) At first *P*-wave arrival, the location is constrained by an initial NYAD; (b) first BAZ intersecting with shrinking NYAD; (c) first BAZ with NYAD just before *P*-wave arrival to a second station; (d) at second *P*-wave arrival, location is constrained by NYAD, one BAZ, and one EDT; (e) second BAZ intersecting NYAD, first BAZ, and first EDT; and (f) at third *P*-wave arrival, location is constrained by NYAD, two BAZ, and three EDT. The cross indicates the catalog location and the expanding dashed circle shows the *P* wavefront. The black and white triangles are triggered and nontriggered stations, respectively. Minimum and maximum distances are reported on each panel for BAZ-included and BAZ-free (in parentheses) location schemes.

tance to 1.7 km, and reduces the maximal distance down to 7 km.

The T^{delay} of 4 s used in this study is much shorter than reported values for California (Brown *et al.*, 2011), and in reality the NYAD-constrained volumes may be much larger than what is shown in Figures 7 and 8. In both scenarios, the most dramatic improvement gained by the integration of BAZ information occurs before the first available EDT. Because the in-

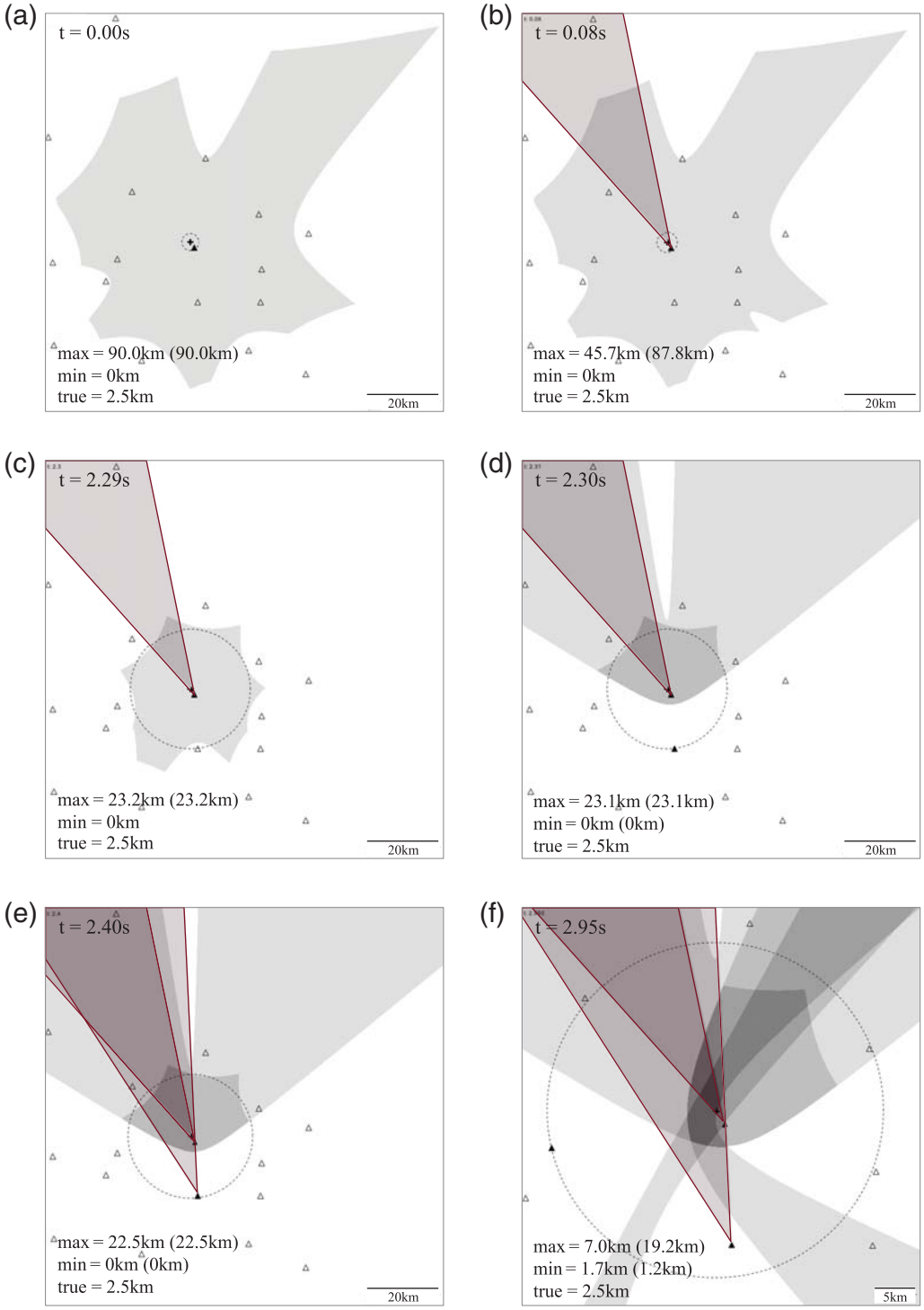

Figure 8. Snapshots of location steps for the intranetwork scenario. (a) At first *P*-wave arrival, the location is constrained by an initial NYAD; (b) first BAZ intersecting with shrinking NYAD; (c) first BAZ with NYAD just before *P*-wave arrival to a second station; (d) at second *P*-wave arrival, location is constrained by NYAD, one BAZ, and one EDT; (e) second BAZ intersecting NYAD, first BAZ, and first EDT; and (f) at third *P*-wave arrival, location is constrained by NYAD, two BAZ, and three EDT. The cross indicates the catalog location and the expanding dashed circle shows the *P* wavefront. The black and white triangles are triggered and nontriggered stations, respectively. Minimum and maximum distances are reported on each panel for BAZ-included and BAZ-free (in parentheses) location schemes.

terval between the first BAZ and the first EDT may be much longer than that between the first and the third EDT, the incorporation of BAZ results in substantial time gain. By the time of the third detection, the hypocenter is already quite well constrained by NYAD and EDT. At that stage, the addition of BAZ information mainly stabilizes the solution by adding indepen-

dent constraints and can be used to discriminate among non-unique solutions (such as the one seen in Fig. 7f).

The time gained by using BAZ estimates scales with the distance between hypocenter and the closest corner of the NYAD cell. For off-network events in which NYAD cells are often open (e.g. offshore) and regions in which station distribution is sparse, BAZ constraints hold the potential to drastically cut the location time.

Conclusions

We assess the robustness of three independent real-time BAZ determination schemes, using offline CISN seismic records from southern California. We find that BAZ values computed by the three methods provide equivalent levels of accuracy. By sending the three BAZ estimates to a screening module that checks for coherency and SNR criteria, we show that accurate BAZ estimates are obtainable in real time.

Because previous attempts at single-station BAZ calculations were subject to uncertainties that were unacceptably large for EEW applications (Lockman and Allen, 2005), currently implemented real-time hypocenter determination schemes do not make use of BAZ estimates (Horiuchi *et al.*, 2005; Satriano *et al.*, 2008). Because we demonstrate that it is possible to recover real-time BAZ, we examine the extent to which their inclusion into real-time location algorithms enhances their performance. Through examination of two earthquake scenarios that use offline data, we show the inclusion of BAZ estimates into real-time location schemes improves the performance of real-time hypocenter determination by cutting the time it takes to obtain well-constrained hypocenters. As seismic networks contain a certain fraction of misaligned sensors (Ekström and Busby, 2008; Niu and Li, 2011), and complex velocity structures may affect ray paths, the accuracy of the BAZ estimates may be receiver and/or path sensitive.

Data and Resources

The data used in this study were obtained from the Southern California Earthquake Data Center (http://www.data.scec.org; last accessed June 2014).

Acknowledgments

We thank Associate Editor Matthew Gerstenberger and two anonymous reviewers for their very constructive remarks. The financial support of the Ministry of National Infrastructures, Energy and Water Resources (Grant 0603415341) is greatly acknowledged.

References

Alessandrini, B., M. Cattaneo, M. Demartin, M. Gasperini, and V. Lanza (1994). A simple *P*-wave polarization application to earthquake location, *Ann. Geofisc.* **37,** 883–897.

Brown, H. M., R. M. Allen, M. Hellweg, O. Khainovski, D. Neuhauser, and A. Souf (2011). Development of the ElarmS methodology for earthquake early warning: Realtime application in California and offline testing in Japan, *Soil Dynam. Earthq. Eng.* **31,** 188–200.

Cua, G., and T. Heaton (2007). The virtual seismologist (VS) method: A Bayesian approach to earthquake early warning, in *Earthquake Early Warning Systems*, P. Gasparini, G. Manfredi, and J. Zschau (Editors), Springer-Verlag, Berlin, Germany, 85–132.

Ekström, G., and R. W. Busby (2008). Measurements of seismometer orientation at USArray transportable array and backbone stations, *Seismol. Res. Lett.* **79,** 554–561, doi: 10.1785/gssrl.79.4.554.

Horiuchi, S., H. Negishi, K. Abe, A. Kamimura, and Y. Fujinawa (2005). An automatic processing system for broadcasting earthquake alarms, *Bull. Seismol. Soc. Am.* **95,** 708–718.

Joswig, M. (2008). Nanoseismic monitoring fills the gap between microseismic networks and passive seismic, *First Break* **26,** no. 6, 121–128.

Lockman, A. B., and R. M. Allen (2005). Single-station earthquake characterization for early warning, *Bull. Seismol. Soc. Am.* **95,** 2029–2039.

Lomax, A., C. Satriano, and M. Vassallo (2012). Automatic picker developments and optimization: FilterPicker—a Robust, broadband picker for real-time seismic monitoring and earthquake early warning, *Seismol. Res. Lett.* **83,** no. 3, 531–540, doi: 10.1785/gssrl.83.3.531.

Nakamura, Y. (1988). On the urgent earthquake detection and alarm system (UrEDAS), in *Proc. of the 9th World Conference on Earthquake Engineering*, Tokyo–Kyoto, 2–9 August 1988, 673–678.

Niu, F., and J. Li (2011). Component azimuths of the CEArray stations estimated from *P*-wave particle motion, *Earthq. Sci.* **24,** 3–13.

Noda, S., S. Yamamoto, S. Sato, N. Iwata, M. Korenaga, and K. Ashiya (2012). Improvement of back-azimuth estimation in real-time by using single station record, *Earth Planets Space* **64,** 305–308.

Ruud, B. O., E. S. Husebye, S. D. Ingate, and A. Christofferson (1988). Event location at any distance using seismic data from single three-component station, *Bull. Seismol. Soc. Am.* **78,** 308–325.

Rydelek, P., and J. Pujol (2004). Real-time seismic warning with a 2-station subarray, *Bull. Seismol. Soc. Am.* **94,** 1546–1550.

Satriano, C., A. Lomax, and A. Zollo (2008). Real-time evolutionary earthquake location of seismic early warning, *Bull. Seismol. Soc. Am.* **98,** 1482–1494.

Appendix A

Single-Value Flipping Condition

The computed back azimuth (BAZ) is bound to a 180° (forward or backward) ambiguity, which is resolved using the fact that the BAZ vector cannot point outside the lower half-space. Considering the behavior of the arctan function, the ambiguity bound BAZ ϕ yields the correct orientation when the following conditions are met (see Fig. A1):

$$(A_Z > 0 \wedge A_N > 0 \wedge A_E > 0) \vee (A_Z < 0 \wedge A_N < 0 \wedge A_E > 0) \vee (A_Z < 0 \wedge A_N < 0 \wedge A_E < 0) \vee (A_Z > 0 \wedge A_N > 0 \wedge A_E < 0), \tag{A1a}$$

which is equivalent to

$$(A_Z > 0 \wedge A_N > 0) \vee (A_Z < 0 \wedge A_N < 0) \tag{A1b}$$

and may be simplified to

$$(A_Z \times A_N > 0). \tag{A1c}$$

If the above condition is not satisfied, the BAZ estimate has to be flipped (rotated by 180°). This yields the final BAZ estimate $f(\phi)$, according to the flipping condition (equation A1c). This flipping condition is formulated for *P* waves and is independent of the source mechanism (Fig. A1).

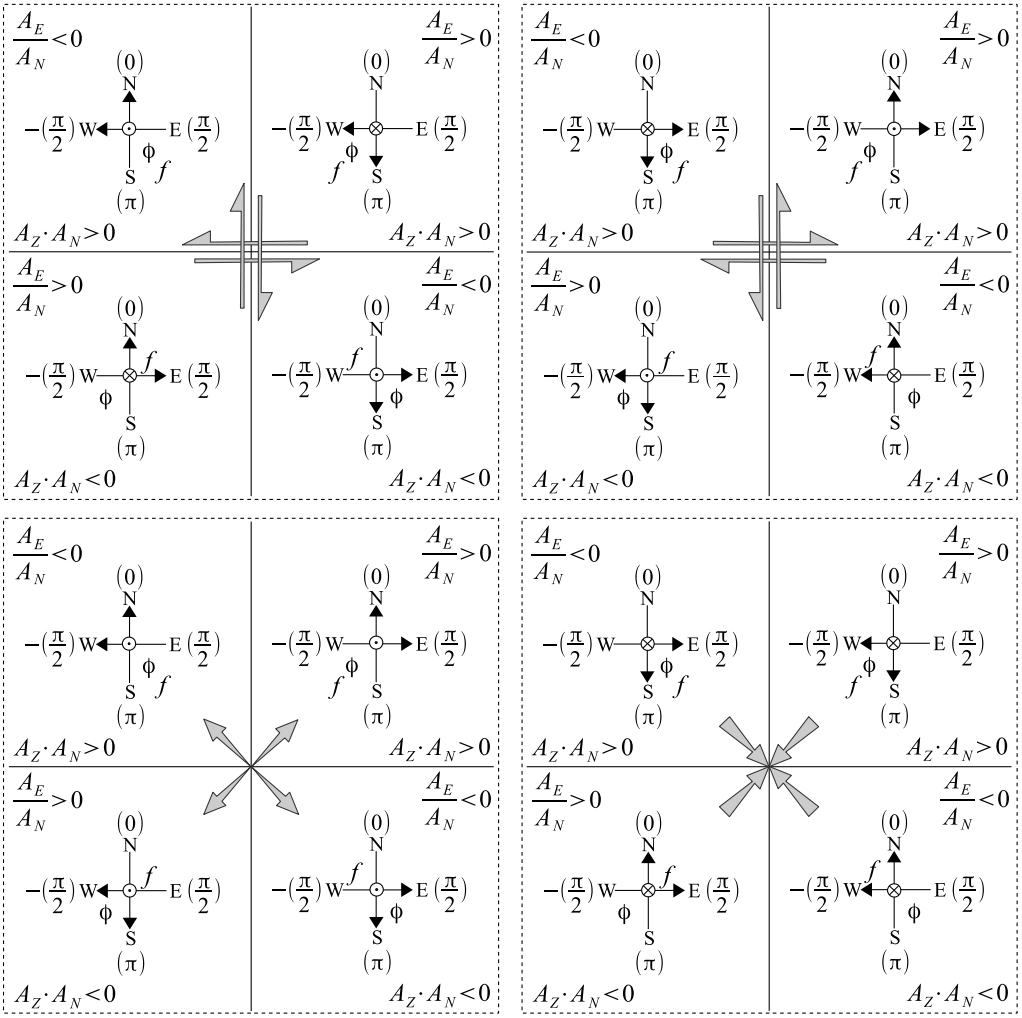

Figure A1. Schematic diagram illustrating the validity of the SV flipping condition according to equation (1c) for several source mechanisms and *P*-wave radiation patterns. (Top) Double-couple strike slip, (bottom left) explosion, (bottom right) implosion. The polarization of the particle motion and the sign of A_E/A_N are indicated for each of the four quadrants around the source. When $A_Z \times A_N$ is negative, the BAZ is flipped by 180°. For all radiation patterns, the final BAZ $f(\phi)$ points toward the source, showing the flipping condition is independent of the source mechanism.

Appendix B

Principal Component Analysis Flipping Condition

After the back azimuth (BAZ) axis is determined using 2D principal component analysis (PCA) (equations 6–8), the unique direction is determined by examining the elementary flipping condition by segment (Fig. B1). Because the polarity of the vertical trace may alternate over the analyzed data window, a weighted average direction is calculated as follows:

$$H = \sum_{n=0}^{N-1} \underbrace{(\hat{a}_n \times \hat{U})}_{a} \times \underbrace{h(A_Z[n])}_{b} \times \underbrace{|\vec{a}_n|}_{c} / N \sum_{n=0}^{N-1} |\vec{a}_n|, \quad \text{(B1)}$$

with

$$\vec{a}_n = \begin{Bmatrix} A_E[n+1] - A_E[n] \\ A_N[n+1] - A_N[n] \end{Bmatrix} \quad \text{(B2)}$$

and

$$h(z) = \begin{cases} 1, & z > 0, \\ -1, & \text{else}, \end{cases} \quad \text{(B3)}$$

in which $\hat{U}$ is the first eigenvector obtained from the PCA (equation 8) and $\hat{a}_n$ is the unit vector of $\vec{a}_n$. The elementary flipping condition $h(z)$ flips direction of each individual segment, if z is negative. Term a above assigns maximum weight to segments that are parallel to $\hat{U}$. Term b determines whether this segment has to be flipped. Term c weighs each segment proportionally to its length. Finally, H ranges from -1 to 1 and the unique BAZ vector is given by

$$\text{BAZ} = \begin{cases} \vec{U}, & H > 0, \\ -\vec{U}, & \text{else}. \end{cases} \quad \text{(B4)}$$

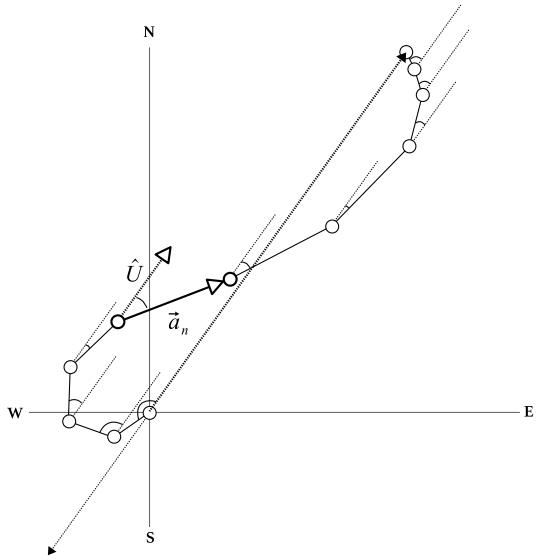

Figure B1. Schematic diagram showing the parameters of equation (B1).

Institut für Geophysik
Universität Stuttgart
Azenbergstr. 16
D-70174 Stuttgart, Germany
eisermann@tau.ac.il
(A.S.E.)

The Department of Geosciences
Tel Aviv University
Tel Aviv 69978, Israel
(A.Z., G.H.W.-B.)

Manuscript received 3 October 2014;
Published Online 2 June 2015