Benchmarking Automated Rayleigh-Wave Arrival Angle Measurements for USArray Seismograms

William D. Frazer[1], Adrian K. Doran[2,3], and Gabi Laske*[2]

Abstract

Surface-wave arrival angles are an important secondary set of observables to constrain Earth's 3D structure. These data have also been used to refine information on the alignments of horizontal seismometer components with the geographic coordinate system. In the past, particle motion has been inspected and analyzed on single three-component seismograms, one at a time. But the advent of large, dense seismic networks has made this approach tedious and impractical. Automated toolboxes are now routinely used for datasets in which station operators cannot determine the orientation of a seismometer upon deployment, such as conventional free-fall ocean bottom seismometers. In a previous paper, we demonstrated that our automated Python-based toolbox Doran–Laske-Orientation-Python compares favorably with traditional approaches to determine instrument orientations. But an open question has been whether the technique also provides individual high-quality measurements for an internally consistent dataset to be used for structural imaging. For this feasibility study, we compared long-period Rayleigh-wave arrival angles at frequencies between 10 and 25 mHz for 10 earthquakes during the first half of 2009 that were recorded at the USArray Transportable Array—a component of the EarthScope program. After vigorous data vetting, we obtained a high-quality dataset that compares favorably with an arrival angle database compiled using our traditional interactive screen approach, particularly at frequencies 20 mHz and above. On the other hand, the presence of strong Love waves may hamper the automated measurement process as currently implemented.

Cite this article as Frazer, W. D., A. K. Doran, and G. Laske (2021). Benchmarking Automated Rayleigh-Wave Arrival Angle Measurements for USArray Seismograms, *Seismol. Res. Lett.* **XX**, 1–14, doi: 10.1785/0220210189.

Supplemental Material

Introduction

On a homogenous, nonrotating Earth, Rayleigh waves are recorded only on the vertical and radial components of a three-component seismogram, and the Love wave appears only on the transverse component. Lateral heterogeneity refracts seismic surface waves away from their respective source–receiver great circle defined by the earthquake and station location (e.g., Woodhouse and Wong, 1986). This allows for some Rayleigh-wave energy to appear on the transverse component and Love-wave energy on the radial component (Fig. 1). At an observing station, the deviation of the incoming wave from the projected source back azimuth is then expressed as the arrival angle. The dispersive nature of surface waves gives these angles a frequency dependence (e.g., Larson and Ekström, 2002; Tanimoto and Prindle, 2007; Koper and Hawley, 2010). In the parametric approach to long-period surface-wave tomography between 40 and 200 s (25–5 mHz), arrival angles therefore comprise a valuable secondary set of observables to investigate the 3D structure of Earth's lithosphere and upper asthenosphere (e.g., Laske and Masters, 1996; Foster *et al.*, 2014).

For their 1996 study, Laske used an interactive screen tool to inspect the three-component seismogram and choose a proper window around the desired wave packet before three polarization vectors are constructed from the complex spectra of the three components (Laske *et al.*, 1994) (see Manual Method section for more details). The data analyst also makes decisions for which frequencies the results are acceptable or if a more optimal time window may lead to a "cleaner" result. Such an interactive tool was, and still is, extremely useful to investigate the nature of individual three-component seismograms. But it is rather cumbersome and time-prohibitive to analyze many events for a large network such as the USArray Transportable Array (TA)—a component of the Earthscope program funded by the U.S.

---

1. Department of Earth and Planetary Sciences, Yale University, New Haven, Connecticut, U.S.A.; 2. Cecil H. and Ida M. Green Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, U.S.A., https://orcid.org/0000-0003-3564-3076 (AKD); 3. Naval Undersea Warfare Center, Newport, Rhode Island, U.S.A.

*Corresponding author: glaske@ucsd.edu

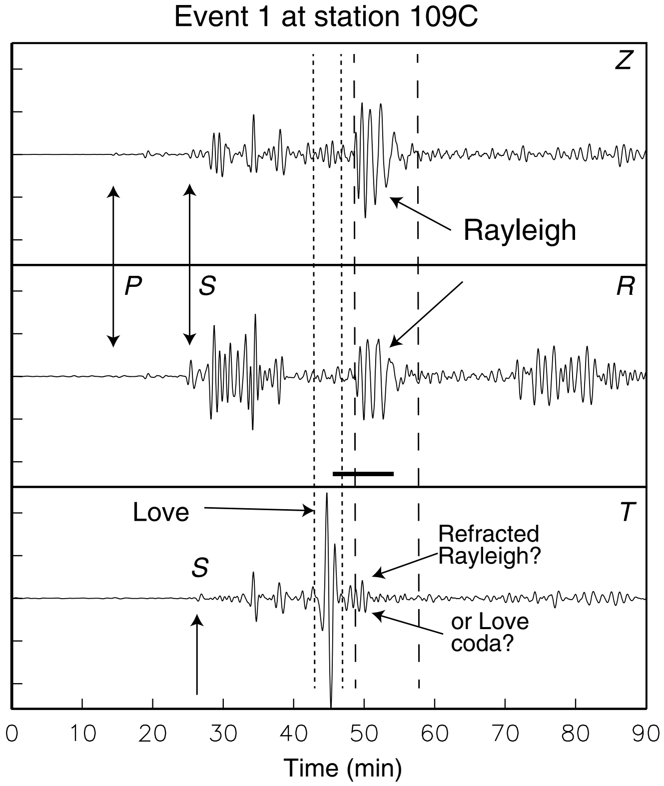

Figure 1. Rotated three-component seismogram for event 1 at station 109C to vertical (*Z*), radial (*R*), and transverse (*T*) components. The traces were filtered with a −20 dB convolution filter to pass signal between 4 and 26 mHz. Vertical-dashed lines mark the window boundaries suggested for the Love- and Rayleigh-wave train by the interactive screen tool. For Rayleigh waves, the corresponding reference group velocities are 3.96 and 3.60 km/s. The data analyst manually adjusts the windows to optimize mutual exclusion of the other wave type from analysis. For event location, see Figure 2. The fat horizontal line at the bottom of the (*R*) component marks the window of automated tool Doran–Laske-Orientation-Python (DLOPy).

National Science Foundation. The TA started operating stations in 2004 in the western United States and "rolled" across the Lower 48 States over the following decade.

Recently, Doran and Laske (2017) devised an automated tool to scan an earthquake catalog for suitable events, preprocess the data, and then determine the frequency-dependent arrival angles for Rayleigh waves. The Doran–Laske-Orientation-Python (DLOPy) tool makes use of the correlation between the vertical and the radial components for optimally tuning the two horizontal components into radial and transverse directions. It may take a few hours to process one year's worth of data, but this is vastly faster than for one person going through the data by hand, for several weeks. DLOPy has been used successfully to obtain the instrument orientation of free-fall ocean bottom seismometers (OBSs) for which the orientation of the horizontal components typically is unknown. Doran and Laske (2017) published instrument orientations for the Cascadia Initiative after benchmarking the method against Laske's original instrument orientations for the Hawaiian PLUME OBS deployment using her traditional manual tool (Rychert *et al.*, 2013). It is important to remark at this point that the orientations in Doran and Laske (2017) are simple averages over all measurements. This averaging may yield biased estimates for instrument orientation when arrival angle measurements are available from only a few back azimuths. Instrument orientations determined this way may only serve as initial estimates in which more accurate assessment is lacking. A more proper approach would be a nonlinear joint inversion of arrival angles for instrument orientation and global lateral heterogeneity following the strategy of Laske (1995).

After the successful implementation of the automated tool to determine instrument orientations, a yet unresolved question is whether individual measurements obtained by DLOPy are internally consistent and precise enough for structural imaging. The covered frequency range of both the methods differs in that the manual tool provides arrival angles at lower frequencies (5–25 mHz), whereas the automated tool covers higher frequencies (10–40 mHz). This study compares Rayleigh-wave arrival angles in the overlapping range between 10 and 25 mHz for 10 teleseismic earthquakes recorded on TA stations (Table S1; Fig. 2) and benchmarks individual measurements obtained using DLOPy against those using the traditional, manual tool. Leading into this comparison, we first tested the consistency of measurements by three different users utilizing the manual tool. We conclude that DLOPy can yield reliable results after vigorous data vetting. However, differences in the results between automated and traditional technique are typically greater than differences between users using the latter. We not only document some cases in which DLOPy provides more consistent results at high frequencies but also cases in which DLOPy yields few and inconsistent data at low frequencies. We also document some decreases of DLOPy measurements when strong Love waves were present.

Manual and Automated Method

In the following, we describe the two methods used in this study.

Manual method

The manual method is an interactive screen tool called from an xterm window. It is based on Fortran and C code in which the latter is used for disk and screen input and output, whereas the former does all of the numerical tasks. The tool provides suggested windows (dashed lines in Fig. 1) that are derived from the extrema of group velocities in the chosen frequency band, here between 5 and 25 mHz, in which the group velocities are computed using a 1D global reference model with an oceanic crust, such as a modification of Preliminary Reference Earth Model (PREM) (Dziewonski and Answerson, 1981). For most global source–receiver paths, the chosen frequency band

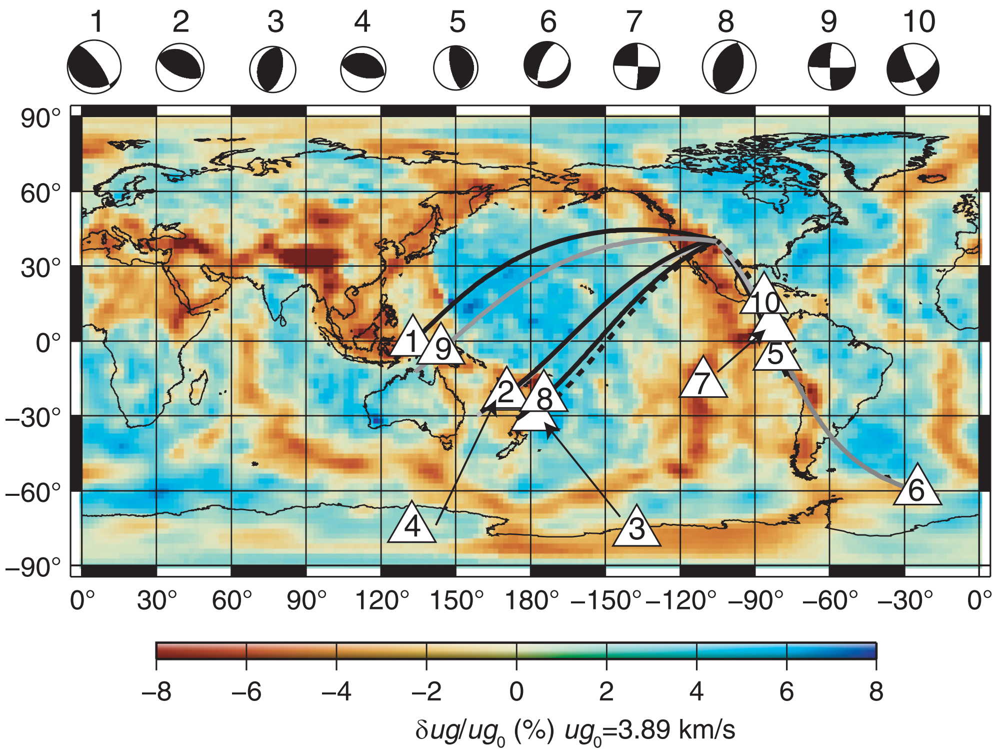

Figure 2. Location map of the 10 earthquakes used in this study (see Table S1 for details). Also shown are R_1 great-circle arcs to the center of the Transportable Array (TA) stations, as used in this study. The color-coded map is the group velocity map for 15 mHz from Ma *et al.* (2014). Beachballs along the top represent Global Centroid Moment Tensors (Ekström *et al.*, 2012).

brackets an Airy phase at about 17 mHz in which group velocity is highest (see, e.g., Laske and Widmer-Schnidrig, 2015). In the shown example, the suggested window may be acceptable, though experienced users would start slightly earlier and end some 5 s later to ensure that enough trailing signal at low frequencies enter the analysis.

The measurement method is based on the singular value decomposition of a three-component seismogram and is implemented in the frequency domain (Park, Lindberg, and Vernon, 1987). Each of the three resulting complex eigenvectors define a uniquely polarized signal whose particle motion is arbitrarily oriented in 3D space but confined to a plane (e.g., a Rayleigh wave whose particle motion is slightly tilted away from the vertical axis; Laske *et al.*, 1994). As far as tapering the waveform is concerned, Park, Vernon, and Lindberg (1987) implemented a multitaper approach. We used the first four 2.5 π prolate spheroidal-wave function eigentapers (Slepian, 1978). Because the eigentapers are orthogonal, the four spectral estimates at a given frequency are independent and allow assignment of formal measurement errors. Compared to a single boxcar window, multi-tapers cause some spectral leakage but reduce sensitivity to contamination from incoherent noise. The chosen window length depends on the epicentral distance and relative strength of the Love-wave signal. With a typical window length of 15 min, the 2.5 π tapers yield 3–4 numerically independent arrival angle estimates in the frequency range 5–25 mHz.

The singular value for a specific eigenvector (or polarization vector) indicates how much the corresponding signal contributes to the 3D waveform. For our purposes, only one signal should be present, which is the Rayleigh wave. The singular value, d_1, is then 1, and the other two, d_2 and d_3, are 0. A drop in d_1 and increases in d_2 and d_3 indicate the presence of secondary signals (e.g., a Love wave, overtones, or a Rayleigh wave suffering from multipathing) that contaminate the signal we want to analyze. The threshold below which a signal is considered contaminated is somewhat subjective and depends on the relative strength between the principal and other signals and their frequency content. Tests for the study of Laske *et al.* (1994) with synthetic waveforms for a 1D Earth revealed that $d_1 = 0.7$ is a reasonable and conservative threshold. Figure 3 shows the singular values for the seismogram in Figure 1. Here, the chosen threshold criterion holds for most of the frequencies, though some contamination may occur at frequencies around 21 mHz.

We also use the ellipticity of the signal as diagnostic for the wave type. We define ellipticity, $\epsilon = 1 - b/a$, in which b and a are the lengths of the minor and major axes of particle motion. A linearly polarized signal has high ellipticity with values up to 1, whereas nearly circular particle motion has a value close to 0. We find that Rayleigh waves typically have $\epsilon < 0.3$. Signals that yield $d_1 < 0.6$ as well as $\epsilon > 0.5$ in a wide frequency band are discarded without further consideration, and we proceed to the next seismogram. For the shown example, ϵ increases dramatically near 21 mHz, so the signal on the transverse component we describe in Figure 1 is likely Love-wave coda and not a single refracted Rayleigh wave. In such cases, we go back and choose a different window that may yield higher values of d_1, while still keeping the entire Rayleigh-wave packet in the window. In the shown example, we do not achieve much improvement. In secondary, automated data vetting, the measurements for 21 mHz are then discarded, but the results at other frequencies are used for further analysis.

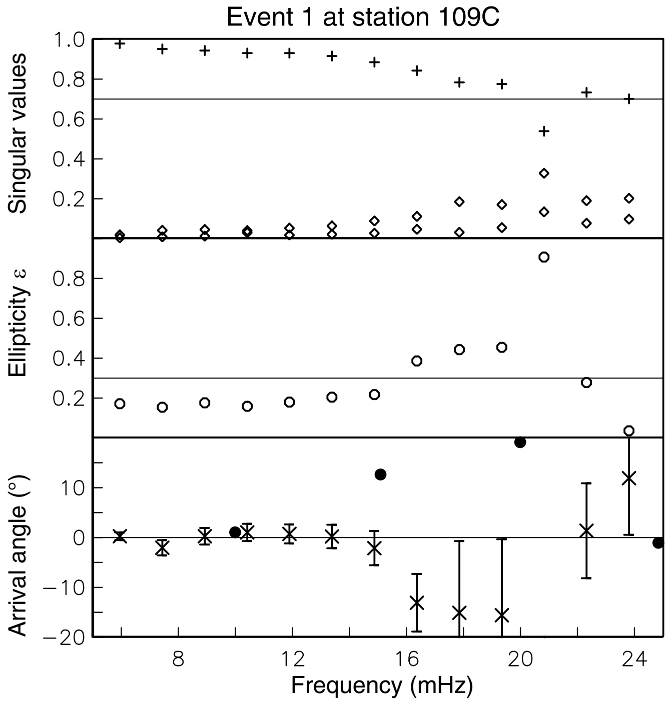

Figure 3. Select polarization parameters determined for the Rayleigh wave shown in Figure 1. Top: the three sets of singular values. A well polarized signal has the largest singular value, d_1 (crosses) close to 1, whereas the smaller singular values, d_2 and d_3 (diamonds), are close to 0. We choose 0.7 as minimum threshold for d_1 (horizontal line). Middle: ellipticity ϵ. We choose an upper threshold of 0.3 for Rayleigh waves. Bottom: arrival angles. Positive value denotes clockwise azimuth from station to source. Solid dots mark measurements using automated tool DLOPy.

DLOPy

The package DLOPy is a Python toolbox for the automated determination of arrival angles of Rayleigh waves (Doran and Laske, 2017) and is based on the method of Stachnik *et al.* (2012). Its original and primary purpose was a rapid determination of the orientation of the horizontal seismometer components of free-fall OBSs from the analysis of a relatively small set of earthquakes. DLOPy uses ObsPy (Krischer *et al.*, 2015) to download data from a data center and for basic data preprocessing. The three-component seismogram with components Z, H_1, and H_2 (vertical, horizontal 1, and horizontal 2) then is analyzed for an isolated minor or major arc Rayleigh-wave train. In a grid search for the optimal rotation angle, we seek to maximize the cross correlation between the Hilbert-transformed vertical component and the rotated H_1 component, C_{zr}. The deviation of that angle from the source–receiver great circle carries information about the instrument orientation as well as off-great circle Rayleigh-wave propagation between the particular earthquake and the station. Additional analysis of several other earthquakes as well as analysis at several frequencies and rigorous quality control ultimately yields an average that Doran and Laske (2017) declared the final instrument orientation.

The method uses a 10% Tukey window as data taper (Harris, 1978). There are two principal differences between DLOPy and the Stachnik method (Stachnik *et al.*, 2012). First, DLOPy uses multiple frequency bands to obtain multiple, frequency-dependent arrival angles, instead of using one broader frequency band that yields one frequency-averaged estimate. We process band-pass filtered records between 10 and 40 mHz, with 5 mHz increments. Second, DLOPy uses a different approach to position the data taper. Instead of a 600 s long, fixed window that starts 20 s before a predicted 4.0 km/s group arrival, we determine the frequency-dependent path-averaged group arrival time using the dispersion maps of Ma *et al.* (2014) (e.g., Fig. 2). This time determines the center of our window in which the length varies between 500 s at 40 mHz and 700 s at 10 mHz.

The DLOPy process is entirely automated in which the data analyst never inspects individual waveforms nor arrival angle results. A rigorous post-processing data vetting is, therefore, needed to pick only high-quality arrival angles for further analysis. Doran and Laske (2017) devised a three-step vetting process to determine the instrument orientation: (1) the cross-correlation coefficient, C_{zr}, has to be above a certain threshold in which Doran and Laske (2017) used $C_{zr} \geq 0.8$; (2) outliers, defined as being more than five times the median absolute deviation from the median, are removed; and (3) the bootstrap method of random sampling with replacement (Efron, 1979) is repeated 5000 times to determine the arithmetic mean and 95% confidence level as error for the instrument orientation. We follow this protocol, except that we omit step 3.

Data and Data Analysis

For this study, we considered 10 teleseismic earthquakes during the first half of 2009 (Table S1) that were recorded at the USArray TA. The TA has a station spacing of about 70 km. The chosen earthquakes have a surface-wave magnitude $M_s \leq 6.1$ and a source depth $h_0 \leq 34$ km. We chose the TA because of its high station density and having a large footprint at the same time. For such a network, lateral refraction between a source and the recording network causes significant changes in arrival angles across the network, but the station density allows assessment of systematic variations (e.g., Foster *et al.*, 2014), and such a database provides a good means for benchmarking our measurement techniques.

Analysis using the manual tool

Using the manual tool, we measured arrival angles for all 10 events. Despite its comparatively large epicentral distance, event 1 was relatively easy to analyze, as it had a well-isolated, large-amplitude Rayleigh wave for the TA that does not appear to suffer from multipathing (see Table S1 for details on each earthquake). Figure 4 documents that arrival angles measured

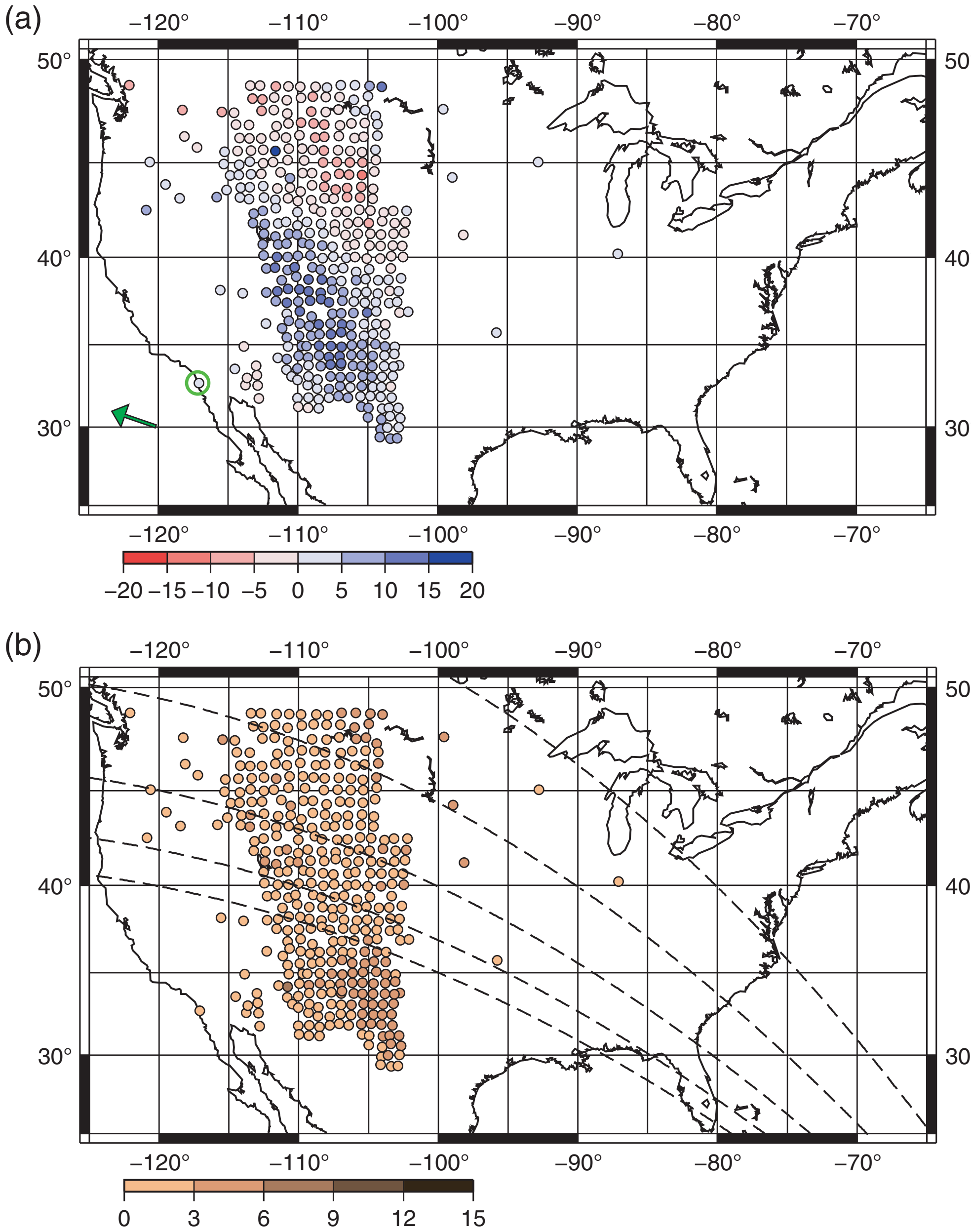

Figure 4. (a) Arrival angles at 7 mHz for event 1 determined with the manual tool by user 1. Angles are usually between −20° and +20°. Positive angles indicate arrival from the right, as viewed from the receiver to the source. The green circle marks station 109C, and the green arrow is the corresponding back azimuth to the source (displaced for unobstructed view). (b) Associated measurement errors. Dashed lines mark select source–receiver great circles. Their convergence points to the source antipode of event 1.

at 7 mHz vary significantly across the TA, but arrival angles are rather coherent, with approaches from the right of the respective receiver-source great circle in the southwestern half of the TA and approaches from the left in the northeastern part. Error bars (mostly 5° or less) indicate that the arrival angles are statistically significant. As Figure S1 reveals, the arrival angle pattern for this event is very different at 15 mHz. Setting aside potential minor systematic issues with the measurement process and/or the relative excitation of Rayleigh waves, other surface waves, and related interference, the pattern depends primarily on lateral variations in global frequency-dependent phase velocity (see, e.g., Foster *et al.*, 2014). We leave detailed testing of the arrival angle patterns against predictions for lateral heterogeneity for a future paper and concentrate in this article on the measurement fidelity itself.

Subsequent in-depth data vetting led to our decision to discard manual measurements with singular values $d_1 < 0.7$. The histograms in Figure 5 reveal that measurements at frequencies up to 15 mHz stay largely within ±20°, with standard deviations (STDs) of 4°–6° at frequencies below 15 mHz to 6° at 15 mHz. But starting at about 15 mHz, the variance in the measurements increases significantly, with STDs reaching 13° and 15° at 20 and 25 mHz. This is largely the result of the increasing influence of strong lateral variations in shallow mantle structure. Histograms also may become increasingly skewed to one side, no longer having the shape of a Gaussian distribution. This shape is controlled mainly by propagation effects of the Rayleigh wave between the source and the TA network (see also Fig. S4). At the same time, the decreasing number of data indicates the increasing difficulty to obtain precise manual measurements of a single arriving wave packet not only as a result of multipathing effects but also contamination by the Love-wave coda.

Touching briefly on the other nine earthquakes, of the eight events shown in Figure S1, all but events 3 and 4 yield a large number of data with mostly coherent patterns across the TA. Events 3 and 4 yield somewhat consistent patterns, but many

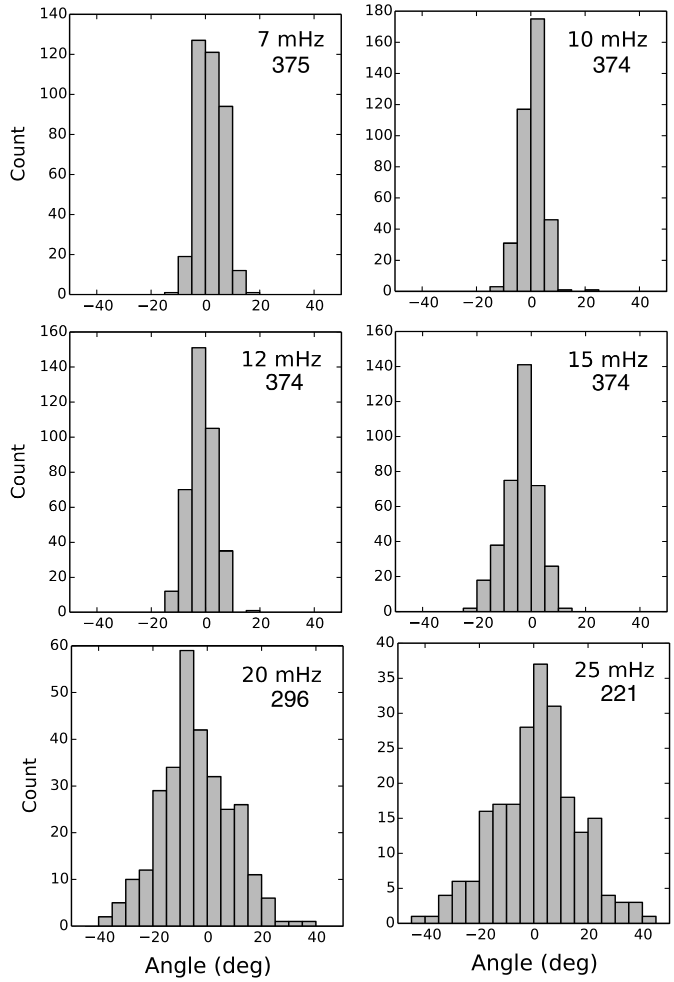

Figure 5. Arrival angle histograms for event 1 (and user 1), at given frequencies (upper-right corner), with 5° bins. Data with singular values $d_1 < 0.7$ were discarded. The number of measurements at each frequency is given below the frequency. The increasing width of the histograms with increasing frequency is a manifest of increasing lateral refraction due to 3D seismic structure.

stations had high noise levels on the horizontal components for event 4 (Table S1). Both events also produced a strong Love-wave signal at the TA, in which event 3 produced particularly large Love-wave amplitudes at many TA stations relative to the Rayleigh wave. Events 2 and 4 were in the same source region, though event 4 had a much smaller magnitude (6.1) than event 2 (7.5), which explains the low number of arrival angle data for event 4. The yield for event 5 was the lowest and is not shown. Despite the relatively small epicentral distance of some 5500 km, the horizontal components were very noisy, thereby hampering the determination of particle motion parameters with high singular values. This event also produced a long Love-wave coda that reached into the Rayleigh-wave analysis window. The waveform quality for event 10 was extremely high, but the short epicentral distance of only some 3100 km made it difficult to separate the Rayleigh wave from the Love wave.

Comparison between users

Using the manual tool, we wondered how much the resulting angles depend on the choices made by the data analyst. If the difference between DLOPy results and manual results were smaller than the variance in results by different users applying the manual tool, then this would justify the exclusive use of DLOPy in the future. We, therefore, first benchmark tested manual results obtained by different users: Will, Zemus, and Gabi (users 1, 2, and 3 in the following). Will (user 1) was a novice but a quick learner and measured angles for a few weeks. Getting used to the technique, he measured event 1 twice, in which we used the picks from the second pass only. Zemus (user 2) was a very careful undergraduate lab assistant who measured arrival angles for about a year. His angles may have the highest quality. Gabi (user 3) is an experienced user who used this tool for many years and now picks measurements rather quickly, perhaps at the expense of minimal error bars.

Map presentations of the measurements between users for event 1 are virtually identical to those in Figure 4 and also very similar for the other seven events shown in Figure S1. Difference plots confirm that disagreements in the measurements remain within a few degrees (Fig. S2). We can identify an area with systematic differences where angles of user 1 in a patch in the north are always larger that the angles measured by the other two users. This could be related to the choice of a slightly different window in which signal is allowed to contribute in one case but not the other. For this event, all the three users had nearly the same number of measurements (within a few) so that a small fraction of 1% had to be discarded in the comparison.

Overall, the correlation between datasets for pairs of users remains excellent for frequencies up to about 15 mHz (Fig. 6). Discrepancies increase with increasing frequency. Some disagreement, sometimes statistically significant, occurs at 25 mHz. Of the three pairings, measurements between user 1 and user 2 are usually slightly less consistent than for the other two pairings, which is also documented in Figure 7a. Typically, correlation between users peaks at frequencies between 8 and 12 mHz, and is lower for higher as well as lower frequencies. The drop for the latter occurs when windows are chosen too short to allow the trailing long-period portion of an inversely dispersed wave packet (group velocity declines with increasing period) be included (see Fig. 1). The reasons for the wrong choice here are the relatively strong dispersion, which leads the relatively low amplitude contributions to the seismogram, as well as the fact that spectral amplitudes decline as period increases. Because user 3 and user 2 have had more experience in using the manual tool, they tend to choose longer windows even when the wave packet visually appears to have ended. For the example shown in Figure 1, the end of the chosen window would not be at the suggested dashed line

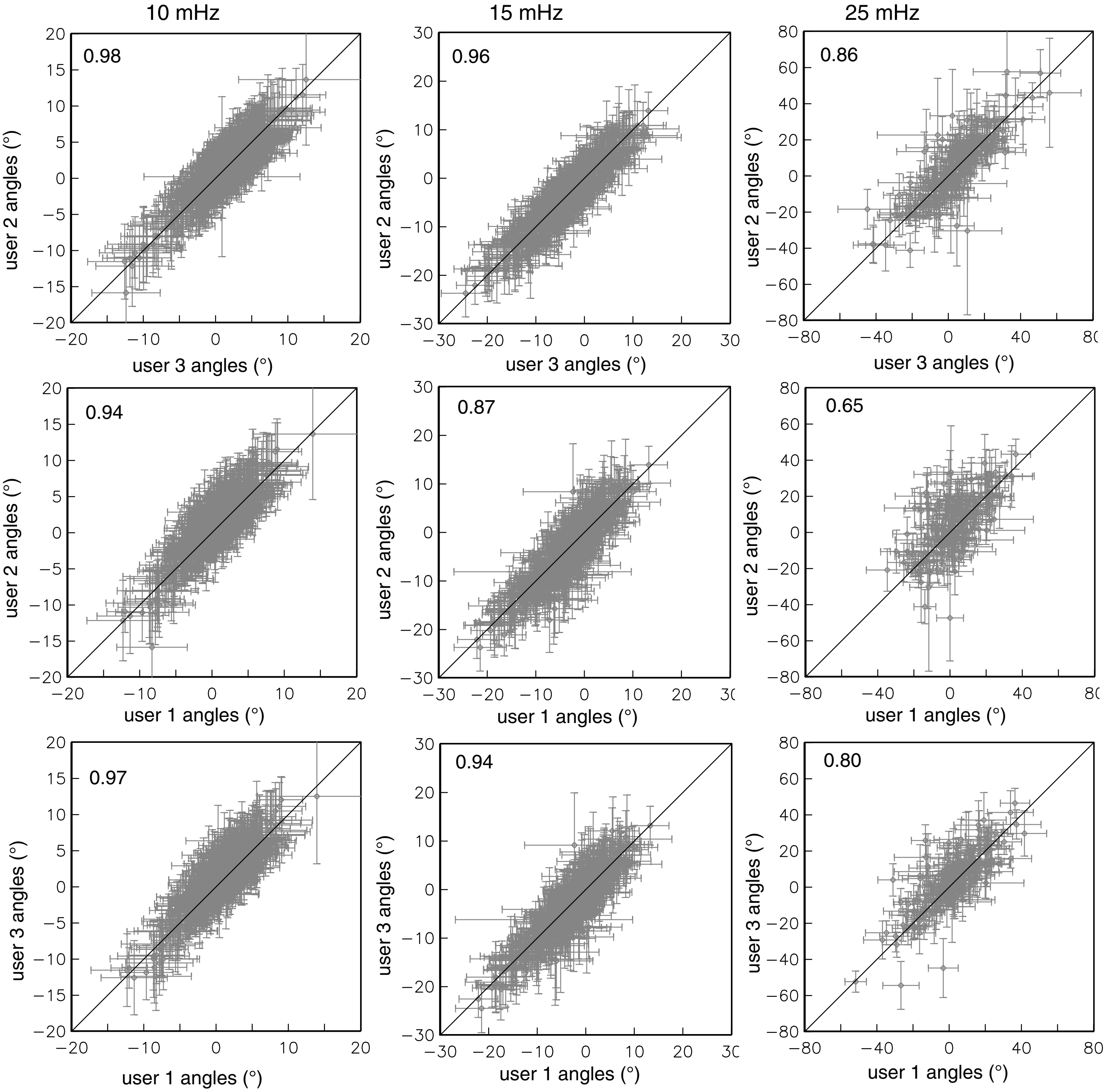

Figure 6. Measurements for event 1 taken by one user plotted against those by another user, for three frequencies. The number in the upper-left corner is the correlation coefficient. Note the different axis scales for different frequencies.

around 58 min but rather around the 65 min mark, especially when overall noise levels appear low.

We proceeded with the other nine events, including events 3, 4, 5, and 10 for which our data yield was poor. Even for these four events, measurements between users are fairly consistent (Fig. 7b), at least for frequencies 15 mHz and below. For event 5, data harvest was particularly low (only 32 measurements to compare), but the agreement between different users is extremely good. Data for event 10 scatter widely, and measurement errors are relatively large. Events 2 and 8 yielded extremely high correlation coefficients between all users, and events 3, 4, and 9 yielded high coefficients. Event 7 yielded the lowest correlation coefficients, even though waveform amplitudes appeared to have high signal levels.

We conclude that different users may choose different windows to extract the Rayleigh-wave packet from the three-component seismogram for arrival angle analysis. This may lead to divergent arrival angle measurements. However, much

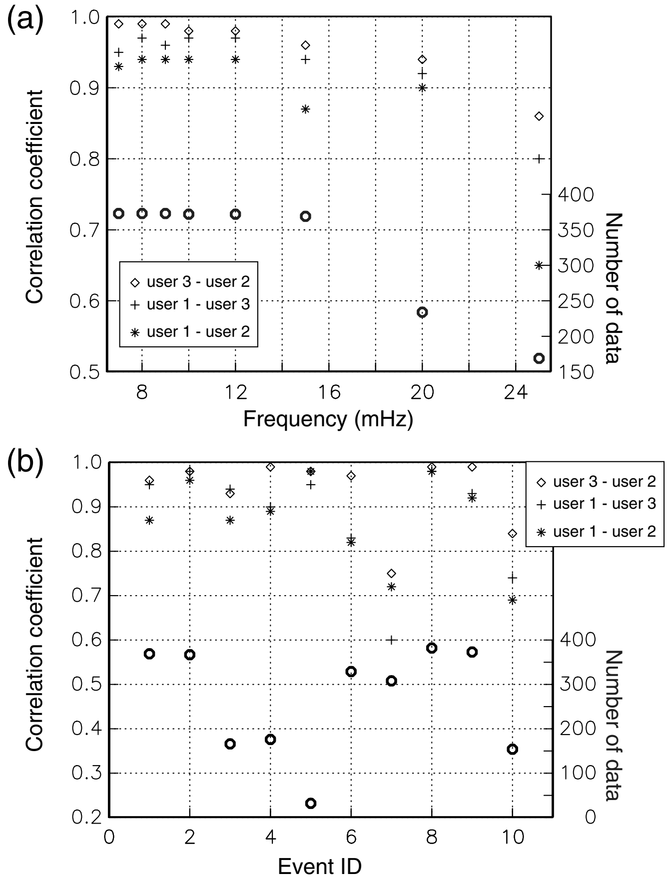

Figure 7. Summary plot of correlation coefficients: (a) measurements for event 1 between pairs of users, as function of frequency. The number of data common between user 1 and user 2 is also shown (open circles). Correlation degrades with increasing frequency, whereas the number of data also decreases.
(b) Measurements at 15 mHz for all events. The number of data common between user 1 and user 2 is also shown (open circles).

of this discrepancy is usually insignificant, and differences lie within measurement error bars.

Comparison between traditional method and DLOPy

Much like with the manual method, measurements made by DLOPy have varying quality with both frequency and event. Even though DLOPy provides measurements for frequencies up to 40 mHz, we considered the following four frequencies for the comparison with the traditional method: 10, 15, 20, and 25 mHz. The minimum correlation coefficient, C_{zr}, was initially set to 0.9. Although 0.8 may be acceptable for the determination of the instrument orientation, our rationale behind the higher threshold was a better exclusion of individual low-quality measurements. This yielded very consistent data for some events such as event 1. However, this restrictive choice left us with nearly no data for events 5 and 10, and, somewhat surprisingly, event 8, despite the fact that manual measurements for this event are of excellent quality and internally consistent (Fig. S1). Events 3 and 4 also yielded few measurements. We, therefore, decreased the threshold to $C_{zr} \geq 0.8$ as was originally used by Doran and Laske (2017).

For event 1, this yields an only marginal increase in picked arrival angles. At frequencies 10 and 15 mHz, both methods yield very similar maps (Fig. 8), in which the difference maps reveal spatially coherent differences. For example, at 10 mHz, the manual technique yields more extreme clockwise arrivals in the southwestern corner of the network, whereas it gives larger counterclockwise angles in the southeast. At 15 mHz, the most extreme discrepancies are again consistent, in which the manual technique yields larger counterclockwise angles in the northeast. The patterns in arrival angles change dramatically between frequencies, and the overall change from 10 to 15 mHz is more dramatic than the differences obtained using the different techniques. We verified that, to some extent, realistic 3D variations in Earth structure can explain such a change (Frazer *et al.*, 2018) (see also Figs. S3–S5). For long-wavelength global models, predicted arrival angles typically remain within ±10° (Laske and Masters, 1996) and so cannot fully explain an observed 20° angle (Figs. S3–S5). Angles measured manually on SPECFEM synthetic seismograms (Komatitsch and Tromp, 2002) yield angles within ±10°, though the underlying global model to construct the synthetics was rather smooth (Tromp *et al.*, 2010). Using perturbation theory and the path integral approximation of Woodhouse and Wong (1986), small-scale structure increasingly adds to off-great-circle propagation and consequently leads to larger arrival angles (Laske and Widmer-Schnidrig, 2015). But even the modern phase velocity maps of Ma *et al.* (2014) yield predicted angles that typically remain within ±10° (Figs. S3–S5).

Getting back to our measurements at 20 mHz, both methods still yield similar arrival angle maps, though we start to obtain spotty coverage with the manual technique. The disagreement becomes more extreme at 25 mHz. We refer to Figure 3 in which singular values in our manual approach decrease with increasing frequency, and some high-frequency measurements fail our quality criteria.

As with the manual measurements in Figure 5, the histograms for the DLOPy measurements in Figure 9 document increasing variance in the data as frequency increases. However, the variance for the manual measurements increases more rapidly (compare with Fig. 5). This is also evident in the data correlation plots (middle column) in which the data clouds appear increasingly squeezed in the y-axis direction. The correlation coefficient decreases (upper-left corner in each panel), and the histograms of the difference data gain in width (Fig. 9c). Although a marked variance reduction is observed at the lower three frequencies, the difference histogram for 25 mHz is much wider, indicating greater disagreement.

The correlation of the manual measurements between different users is usually greater than that between manual

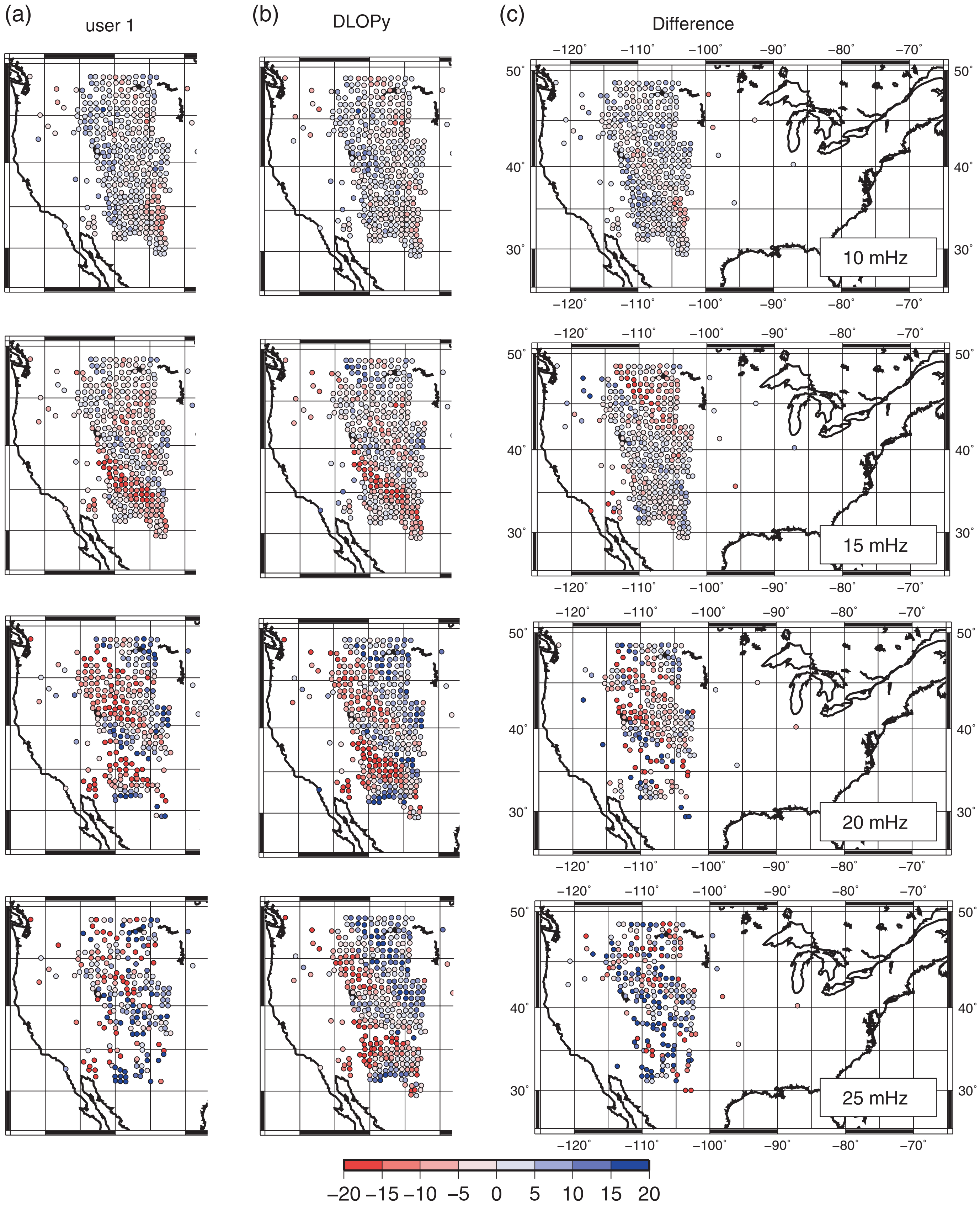

Figure 8. Arrival angle measurements for event 1 as a function of frequency. (a) Using the traditional, manual tool. (b) Using the automated DLOPy tool. (c) Difference plots.

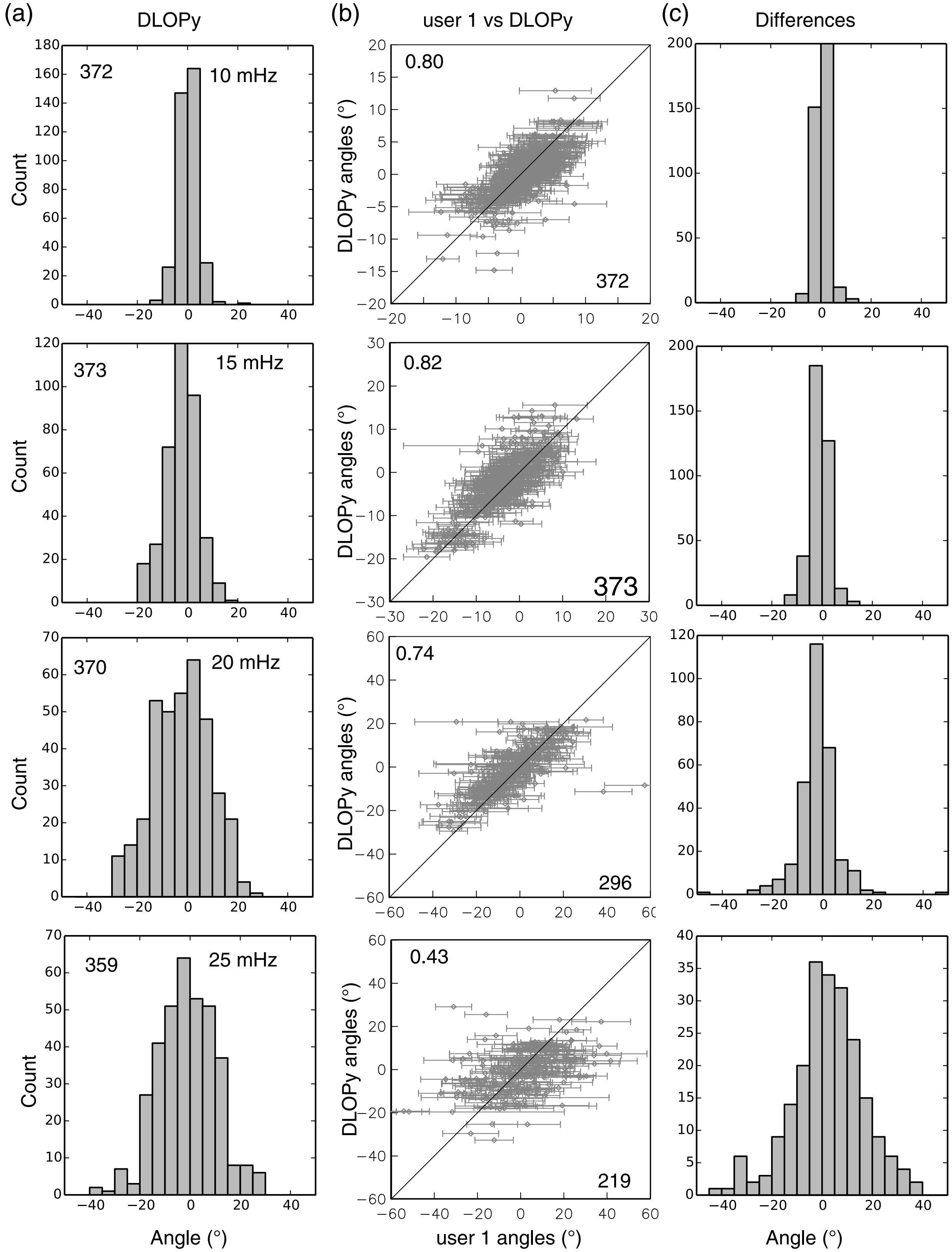

Figure 9. (a) DLOPy arrival angle histograms for event 1, at given frequencies (upper-right corner), for a quality threshold of $C_{zr} = 0.8$. The number of measurements is given in the upper-left corner. (b) DLOPy arrival angles plotted against manual measurements from user 1. The correlation coefficient is in the upper-left corner, and the number of common data is in the lower right. Note the different axis scales for different frequencies. (c) Histograms of the angle differences (user 1—DLOPy). See Figure 5 for histograms of manually measured angles from user 1.

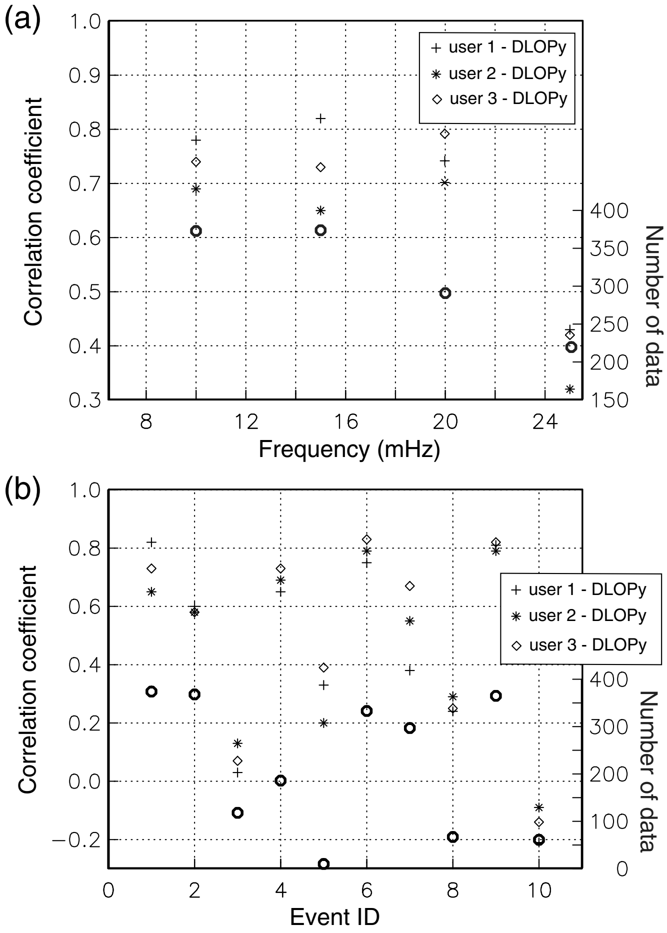

Figure 10. Summary plot of correlation coefficients. (a) For measurements for event 1 between each of the three users and DLOPy, as function of frequency. The number of common data between user 1 and DLOPy is also plotted (open circles). At all frequencies, correlation is well above the 99% confidence level (<0.3 for 100 data points) but is less than in Figure 7a. (b) Between DLOPy and manual measurements at 15 mHz for all events, for all users. The number of common data between user 1 and DLOPy is also plotted (open circles).

and automated measurements (compare Fig. 6 and middle column in Fig. 9). This is further quantified in Figure 10a. Although correlation coefficients stay well above the 99% confidence level of less than 0.2, given the number of data even at 25 mHz, the correlation coefficients stay largely below 0.8, in which correlation of the manual picks between users is 0.85 or above (Fig. 7a). We note that agreement between the DLOPy measurements and manual data of user 2 is consistently the lowest among all the three users. A possible reason for this is that user 2 tried the hardest to assure high measurement quality at frequencies below 10 mHz. Choosing a corresponding data window comes at the expense of quality for higher-frequency data, while allowing contaminating high-frequency coda signal to enter the longer analysis window. We revisit this aspect in the Discussion section.

Discussing now the other events, Figure 10b confirms that measurements at 15 mHz made with DLOPy generally agree less well with those made with the manual tool, by any user (compare with Fig. 7b). However, acceptable agreement for the most events is documented by a relatively high correlation coefficient of 0.6 or above. We have already discussed the challenges of separating the Rayleigh from the Love wave for events with short epicentral distances (events 5 and 10). The manual tool allows a user to adjust the start and length of a window to collect high-quality data (correlation coefficient 0.7 or above), whereas DLOPy produces few data, and correlation with manual data is negative.

Event 3 also did not yield good agreement between DLOPy and manual data, though the latter are of high quality and in very good agreement between the users. We have already pointed out that this event produced strong Love-wave signals. As far as spatial patterns are concerned, the manual data change rather slowly between 15 and 25 mHz (not shown). DLOPy yields nearly no data at 10 mHz, and only few data are at 15 mHz that are all in a narrow corridor perpendicular to the source–receiver great circle. DLOPy data yield are great and consistent at higher frequencies, and patterns in the maps resemble those in the manually picked maps.

This general observation is shared with event 8 as well that also produced strong Love waves at TA stations, and recall that manual data had excellent agreement between the users. Both events occurred in the Tonga–Kermadec subduction region. Also recall that event 5 exhibited a lengthy Love-wave coda. For three-component records that exhibit strong Love-wave signals, the manual tool allows the user to pick a later start time for the analysis window to optimize the isolation of the desired Rayleigh wave. As Figure 11 attests, the early start time is needed for this event to fully capture the Rayleigh wave, whereas the window computed from the modified PREM starts too late. The DLOPy window starts even earlier, thereby allowing too much of the Love wave to enter the analysis window. A drawback of automated tools, as implemented in the Stachnik *et al.* (2012) tool and in DLOPy, is that there currently is no good way to account for this, and valuable Rayleigh-wave data are discarded.

Discussion and Conclusions

This study explored the quality of individual arrival angle measurements obtained through our automated DLOPy method. This method was previously used successfully to determine the orientation of horizontal seismometer components (Doran and Laske, 2017) in which differences in the average of station-specific datasets of arrival angles were insignificant when compared with corresponding datasets obtained with our traditional manual tool (Laske *et al.*, 1994). After benchmark-testing results obtained by different users, we demonstrated here that DLOPy measurements for events with large and dominant Rayleigh waves compare favorably with manually picked data, though

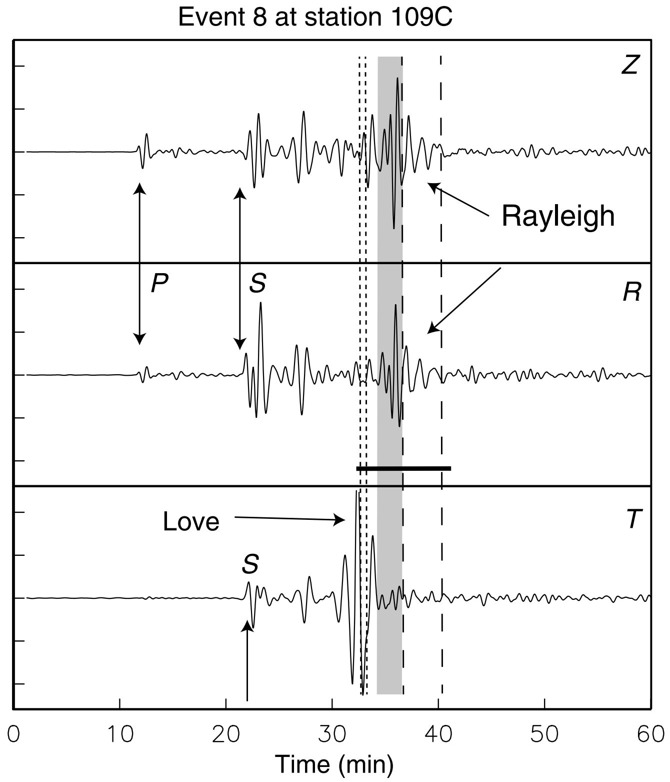

Figure 11. Rotated three-component seismogram for event 8 at station 109C to vertical (*Z*), radial (*R*), and transverse (*T*) components. For details, see Figure 1. The gray area marks the part of the Rayleigh waveform missed by the initially suggested window in the manual tool that is bounded by group velocities 3.96 and 3.60 km/s. The fat horizontal line at the bottom of the (*R*) component marks the window of automated tool DLOPy.

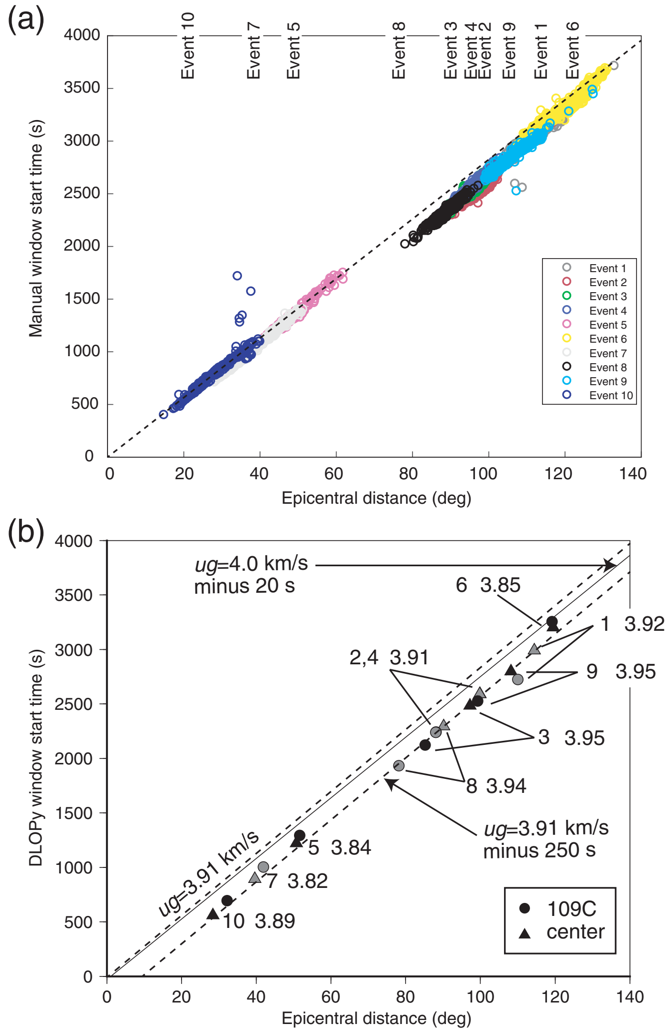

Figure 12. Start time (after event time) of the analysis window for (a) the manual technique for each event and stations in the database of user 1. The dashed line indicates an apparent group velocity that corresponds to the window start time. (b) The DLOPy tool. Symbols mark start times for two locations—station 109C in the southwest and the center of the 2009 TA (as identified in Fig. 2). At epicentral distances greater than 80°, two lines identify the pair for each earthquake. Black and gray symbols delineate different event numbers for better display. Numbers for each pair identify the event number and the average source–receiver group velocity to the center of the 2009 TA. See Figure S6 for a complete set of start times. Dashed lines mark a group velocity of 3.91 km/s, as shown in panel (a) and shifted by 250 s corresponding to the implied start time of a 500 s DLOPy window. Symbols "above" this line have lower average group velocities, whereas symbols below this line have higher group velocities. See Table S2 for detailed group velocity information. The thin-solid line corresponds to window start times used by Stachnik *et al.* (2012).

the disagreement for individual data may be larger than the variance in manual picks between the users. Both datasets have to undergo rigorous secondary screening to remove outlier data.

The manual tool was designed for high-quality arrival angle measurements between 5 and 25 mHz. Unpublished experiments for the Laske *et al.* (1994) study showed that using a fixed time window length for a strongly dispersed signal leads to a loss of fidelity at frequencies much beyond 20 mHz. For most of the waveforms analyzed here, higher frequencies arrive earlier in the time window so that a long time window, as needed to capture the lower frequencies, includes part of the high-frequency coda. Because DLOPy centers the analysis window on expected frequency-dependent group arrivals, we place more confidence in DLOPy results for frequencies higher than 20 mHz. For events with strong Love-wave signals, such as event 3 and especially event 8, the manual tool provides significantly more high-quality arrival angle data than DLOPy does. This is largely a result of the ability to choose an optimal window with the manual tool to isolate the Rayleigh wave from the Love wave.

An open question has been whether the subjective choices in the manual tool significantly deviates from automated calculations for the windows as implemented in DLOPy. Inspection of handpicked data by user 1 reveals that the start time of the window is a fairly consistent function of epicentral

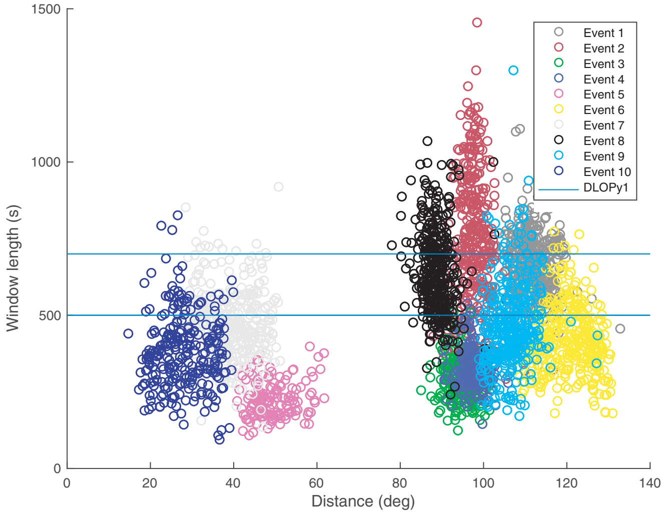

Figure 13. Length of the analysis window for the manual technique, for each event and stations in the database of user 1.

distance (Fig. 12) in which the slope suggests a group velocity of 3.91 km/s. So automation seems possible. For events associated with epicentral distances between 80° and 100°, an offset by about 180 s implies the need for an earlier window start (see also Fig. 11 and the Comparison between Traditional Method and DLOPy section). Events 3 and 8 are located in the Tonga–Kermadec region (Fig. 2), and events 2 and 4 in the Loyalty Islands region. The Rayleigh waves of all these events traverse oceanic areas associated with higher-than-average group velocities.

The automated tool by Stachnik *et al.* (2012) uses a fixed group velocity of 4 km/s, and the window starts 20 s before the computed arrival time. This does not allow for anomalously low or high source–receiver group travel times. DLOPy accounts for this by using modern dispersion maps (Ma *et al.*, 2014) to compute group travel times for each path. On the other hand, changes in this source–receiver group travel time is typically less than the 250 s offset that DLPOy computes for the start time of the window (See also Fig. S6,) or an even larger offset of 350 s for longer-period measurements; recall that DLOPy uses the group travel time for the center of the window, not its start. The proper automated start time for events 8, 3, 4, and 2 may be slightly later, perhaps 150 s before the computed group travel time, to ensure proper isolation of the Rayleigh wave. But for the events at shorter distances (events 10, 7, and 5), the DLOPy start times seem too early. Our assessment is based on 10 events, of which all but one occurred in the "seismically fast" Pacific Ocean. We feel that more testing is needed for other areas on the globe before we make formal changes to DLOPy and release a new version.

Although we can identify a clear dependence of the start time of the analysis window on epicentral distance, the situation is not as clear for the length of the window (Fig. 13). Regardless of epicentral distance, manual windows can be as short as 200 s, though such a window would provide little frequency resolution. However, they may be necessary to separate the Rayleigh wave from other signals, even at larger epicentral distances that typically allow for better separation of signals with different group velocities. We note that event 3 has particularly short windows and recall that DLOPy failed to produce data that were consistent with manual measurements.

Overall, the majority of windows are well above 450 s in length, and more distant events can have windows more than 1000 s long or even longer. Longer windows are not only desirable to increase frequency resolution on one hand but also to make sure enough long-period energy enters the analysis. As mentioned previously, long-period portions may often be difficult to identify in a strongly dispersed waveform. In the example of the seismogram of Figure 1, the window could end near the 70 min mark, so that the analysis window could cover more than 20 min. DLOPy windows are set to either 500 or 700 s, depending on frequency. For source–receiver paths with strong dispersion, even the longer DLOPy window may be too short to include significant portions of signal at longer periods. This may explain an increasing degradation of agreement with manual measurements at frequencies less than 15 mHz. At frequencies between 20 and 40 mHz, DLPOy measurements are fairly consistent, whereas manual measurements dramatically drop in quality at frequencies above 20 mHz.

We arrive at four main conclusions:

- For events that produce well-isolated Rayleigh waves at recording stations, our automated tool DLOPy yields internally consistent arrival angles that compare well with data obtained with our traditional, manual tool.
- For frequencies below 15 mHz, the manual tool typically yields more consistent data than DLOPy.
- For frequencies 20 mHz and beyond, DLOPy provides much higher-quality data than our manual tool.
- for events with strong Love waves, DLOPy may fail to yield high-quality Rayleigh-wave arrival angles.

The strict cross-correlation threshold of $C_{zr} > 0.8$ ensures that DLOPy arrival angles that are potentially contaminated by Love-wave interference are not included in a high-quality dataset. But the automated determination of the data window, as implemented in DLOPy, may lead to the exclusion of otherwise valuable Rayleigh-wave arrival angles that we could pick with our manual tool. A target of future work to improve DLPOy may include an iterative optimization process that allows the assessment of the presence of Love waves and corresponding adjustments to the analysis window.

Data and Resources

The waveforms used in this study are publicly available and were accessed through the Incorporated Research Institutions for Seismology Data Management Center (IRIS DMC; http://ds.iris

.edu/ds/nodes/dmc/). The data were accessed between January 2018 and August 2018. Doran–Laske-Orientation-Python (DLOPy) used the Python toolbox ObsPy (Krischer *et al.*, 2015) to download and process the data. DLOPy is available for download at ~https://igppweb.ucsd.edu/~adoran/DLOPy.html. SPECFEM synthetic seismograms (Tromp *et al.*, 2010) were downloaded from the IRIS DMC at https://ds.iris.edu/spud/synthetic. All websites were last accessed in September 2021.

The supplemental material to this article includes two tables on earthquake and group velocity information for the events used in this study as well as six supporting figures.

Declaration of Competing Interests

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

Acknowledgments

The research by Frazer was supported through an IRIS 2018 summer internship. The authors would like to thank Michael Hubenthal for spearheading the Incorporated Research Institutions for Seismology (IRIS) internship program and for orchestrating the student–mentor pairing for this project. Undergraduate lab assistant Zemus Diaz helped measuring arrival angles using the manual measurement tool. This project was funded by NSF Grant EAR-1722579. Some figures were created using Generic Mapping Tools (Wessel *et al.*, 2013). Comments by reviewers Göran Ekström and Joshua Russell greatly improved this article, and are much appreciated.

References

Doran, A. K., and G. Laske (2017). Ocean-bottom seismometer instrument orientations via automated Rayleigh-wave arrival angle measurements, *Bull. seismol. Soc. Am.* **107,** no. 2, 691–708, doi: 10.1785/012016165.

Dziewonski, A. M., and D. L. Answerson (1981). Preliminary reference Earth model, *Phys. Earth Planet. In.* **25,** 297–356, doi: 10.1016/0031-9201(81)90046-7.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *Ann. Stat.* **7,** 1–26.

Ekström, G., M. Nettles, and A. M. Dziewonski (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth Planet. In.* **200–201,** 1–9, doi: 10.1016/j.pepi.2012.04.002.

Foster, A., G. Ekström, and V. Hjörleifsdottir (2014). Arrival-angle anomalies across the USArray Transportable Array, *Earth Planet. Sci. Lett.* **402,** 58–68, doi: 10.1016/j.epsl.2013.12.046.

Frazer, W., G. Laske, and A. Doran (2018). Surface-wave arrival angles and wave-propagation effects at the USArray Transportable Array, *AGU Fall Meeting abstract*, Washington, D.C., S31S-0521.

Harris, F. (1978). On the use of windows for harmonic analysis with the discrete Fourier transform, *Proc. IEEE*, **66,** no. 1, 51–83, doi: 10.1109/PROC.1978.10837.

Komatitsch, D., and J. Tromp (2002). Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.* **149,** 390–412, doi: 10.1046/j.1365-246X.2002.01653.x.

Koper, K. D., and V. L. Hawley (2010). Frequency dependent polarization analysis of ambient seismic noise recorded at a broadband seismometer in the Central United States, *Earthq. Sci.* **23,** 439–447, doi: 10.1007/s11589-010-0743-5.

Krischer, L., T. Megies, R. Barsch, M. Beyreuther, T. Lecocq, C. Caudron, and J. Wassermann (2015). ObsPy: A bridge for seismology into the scientific Python ecosystem, *Comput. Sci. Discov.* **8,** 014003, doi: 10.1088/1749-4699/8/1/014003.

Larson, E. W. F., and G. Ekström (2002). Determining surface wave arrival angle anomalies, *J. Geophys. Res.* **107,** no. B6, 2127, doi: 10.1029/2000JB000048.

Laske, G. (1995). Global observation of off-great-circle propagation of long-period surface waves, *Geophys. J. Int.* **123,** 245–259, doi: 10.1111/j.1365–246X.1995.tb06673.x.

Laske, G., and G. Masters (1996). Constraints on global phase velocity maps from long-period polarization data, *J. Geophys. Res.* **101,** no. B7, 16,059–16,075, doi: 10.1029/96JB00526.

Laske, G., and R. Widmer-Schnidrig (2015). Theory and observations: Normal mode and surface wave observations, in *Treatise on Geophysics* (second ed.), G. Schubert (Editor), Vol. 1, 117–167, doi: 10.1016/B978-0-444-53802-4.00003-8.

Laske, G., G. Masters, and W. Zürn (1994). Frequency-dependent polarization measurements of long-period surface waves and their implications for global phase-velocity maps, *Phys. Earth Planet. In.*, **84,** 111–137.

Ma, Z., G. Masters, G. Laske, and M. Pasyanos (2014). A comprehensive dispersion model of surface wave phase and group velocity for the globe, *Geophys. J. Int.* **199,** 113–135, doi: 10.1093/gji/ggu246.

Park, J., C. R. Lindberg, and F. L. Vernon (1987). Multitaper spectral analysis of high-frequency seismograms, *J. Geophys. Res.* **92,** 12,675–12,684.

Park, J., F. Vernon, and C. R. Lindberg (1987). Frequency dependent polarization analysis of high-frequency seismograms, *J. Geophys. Res.* **92,** 12,664–12,674.

Rychert, C. A., G. Laske, N. Harmon, and P. M. Shearer (2013). Seismic imaging of melt in a displaced Hawaiian plume, *Nat. Geosci.* **6,** 657–660, doi: 10.1038/NGEO1878.

Slepian, D. (1978). Prolate spheroidal wave functions, Fourier analysis, and uncertainties—V: The discrete case, *Bell Syst. Tech. J.* **57,** 1371–1430.

Stachnik, J. C., A. F. Sheehan, D. W. Zietlow, Z. Yang, J. Collins, and A. Ferris (2012). Determination of New Zealand ocean bottom seismometer orientation via Rayleigh-wave polarization, *Seismol. Res. Lett.* **83,** no. 4, 704–712, doi: 10.1785/0220110128.

Tanimoto, T., and K. Prindle (2007). Surface wave analysis with beamforming, *Earth Planets Space* **59,** 453–458, doi: 10.1186/BF03352706.

Tromp, J., D. Komatitsch, V. Hjorleifsdottir, Q. Liu, H. Zhu, D. Peter, E. Boxdag, D. McRitchie, P. Friberg, C. Trabant, *et al.* (2010). Near real-time simulations of global CMT earthquakes, *Geophys. J. Int.* **183,** 381–389, doi: 10.1111/j.1365-246X.2010.04734.x.

Wessel, P., W. H. F. Smith, R. Scharroo, J. Luis, and F. Wobbe (2013). Generic mapping tools: Improved version released, *Eos Trans. AGU* **94,** 409–410, doi: 10.1002/2013EO450001.

Woodhouse, J. H., and Y. K. Wong (1986). Amplitude, phase and path anomalies of mantle waves, *Geophys. J. R. Astr. Soc.* **87,** 753–773.

Manuscript received 12 July 2021

Supplement: Benchmarking Automated Rayleigh-Wave Arrival Angle Measurements for USArray Seismograms

William D. Frazer, Adrian K. Doran and Gabi Laske

21 October 2021

Table 1. Earthquakes in 2009 used in this study

Event #	Region	MM/DD Julian Day	time UTC	Depth (km)	M_S	Distance to 109C (°)	Back Azimuth (°)	Rayleigh wave quality
1	Irian Jaya	01/03 003	19:43	17	7.6	107.0	-79.2	high, compact
2	SE of Loyalty Islands	01/15 015	7:27	27	7.5	88.1	-118.1	high, compact
3	Kermadec Islands	01/18 018	14:11	33	6.3	85.2	-130.7	low amplidude large Love
4	SE Loyalty Islands	01/21 021	17:08	24	6.1	88.0	-118.6	noisy horizontals
5	off-shore N. Peru	02/15 046	10:04	21	6.1	51.6	131.4	noisy horiz. long L coda
6	S. Sandwich Islands	02/28 059	14:33	15	6.3	119.2	145.6	mostly ok
7	S. of Panama	03/12 071	23:23	9	6.3	41.9	122.8	large amps.
8	Tonga Islands	03/19 078	18:17	34	7.6	78.1	-127.4	small very large L
9	N coast of New Guinea	04/01 091	3:54	10	6.5	99.3	-88.2	high, compact very small L
10	N of Honduras	05/28 148	8:24	10	6.3	32.2	112.5	very compact large L overlaps

Table 2. Average Group Velocity at 15 mHz for Ma et al. (2014), for two paths in this study, for each event

Event	Station 109C		center	
#	Distance	ug	distance	ug
	(deg)	(km/s)	(deg)	(km/s)
1	107.0	4.00	114.4	3.92
2	88.1	3.93	99.9	3.91
3	85.2	3.99	97.2	3.95
4	88.0	3.93	99.8	3.91
5	51.6	3.72	50.7	3.84
6	119.2	3.78	119.5	3.85
7	41.9	3.71	39.6	3.85
8	78.1	3.98	90.2	3.94
9	99.3	3.97	108.2	3.95
10	32.2	3.78	28.4	3.89

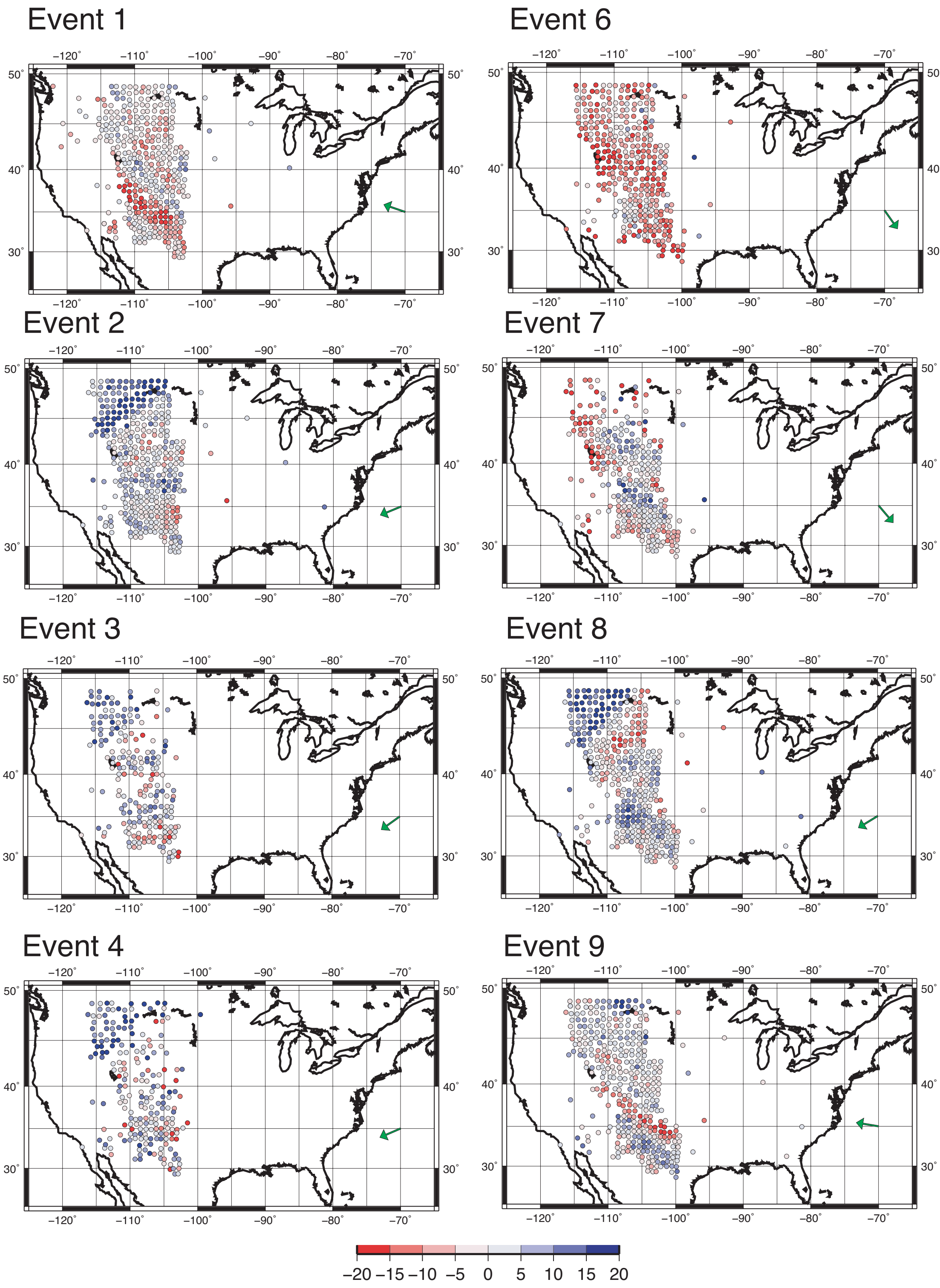

Figure S1. Arrival angles at 15 mHz for 8 events determined with the manual method by user 1. Green arrows mark the back azimuth from station 109C to the source (displaced for unobstructed view). The overall patterns are different between different events, as to be expected. But internally coherent patterns appear across the network.

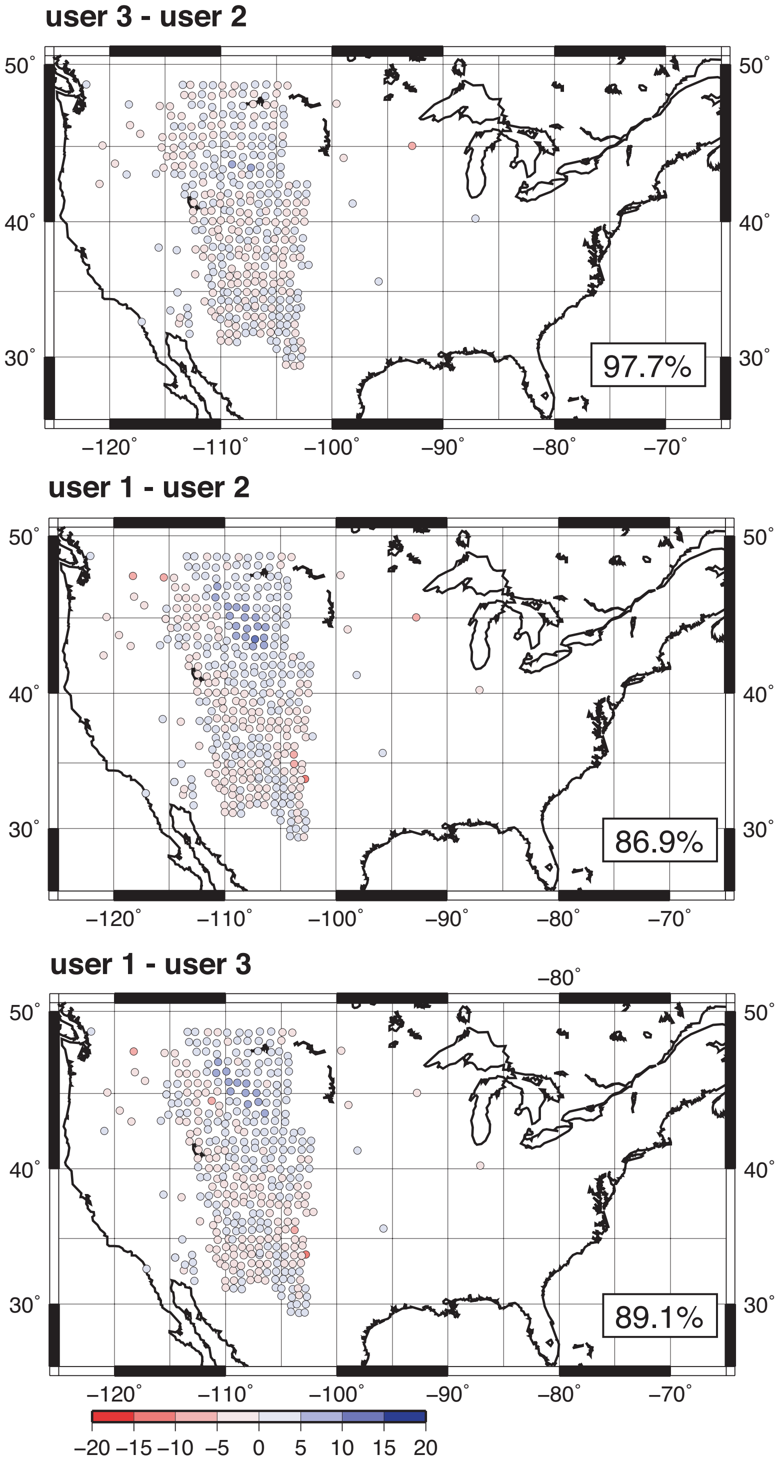

Figure S2. Differences in measured arrival angles for event 1 at 15 mHz for pairs of users (see title of each panel). Number in the box marks variance reduction for difference data with respect to reference data (second name in title).

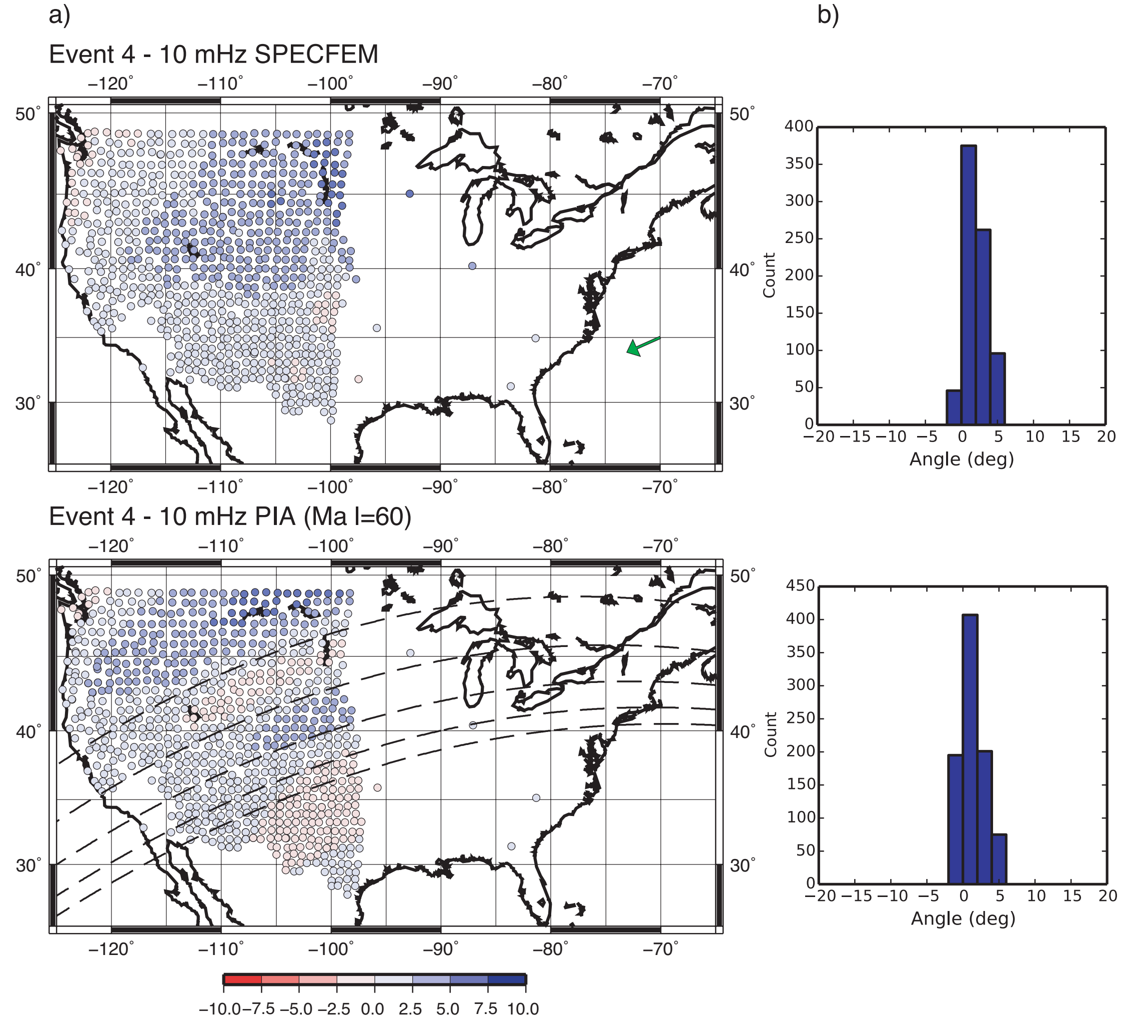

Figure S3. a) Arrival angles at 10 mHz for event #4. Top: for SPECFEM synthetics (Tromp et al., 2010). Angles were measured using our manual tool with the same parameters that we used to measure angle using the real data. The synthetic seismograms were downloaded at https://ds.iris.edu/spud/synthetic. The green arrow marks the back azimuth from station 109C to the source (displaced for unobstructed view).
Bottom: for predictions using the path integral approximation (e.g., Laske and Masters, 1996) and the phase velocity map of Ma et al. (2014). The 4-degree equal area map was expanded in surface spherical harmonics up to degree $l = 60$ for the computation of the predictions. Dashed lines mark source-receiver great circle arcs to some of the stations.
b) histograms for the data shown in a). Predictions, as determined here, span a much smaller range of angles as the real observations.

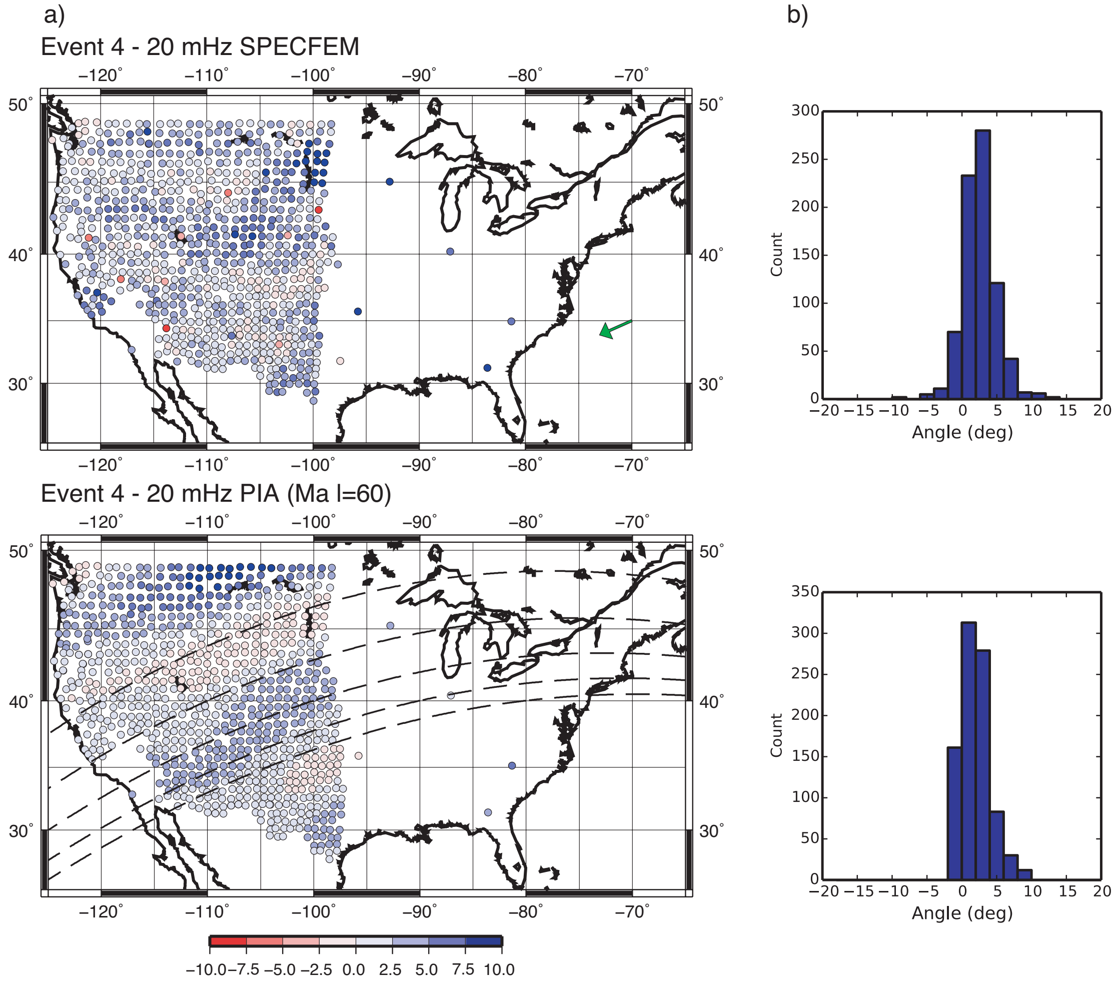

Figure S4. Arrival angles at 20 mHz for events #4. For details see Figure S3.

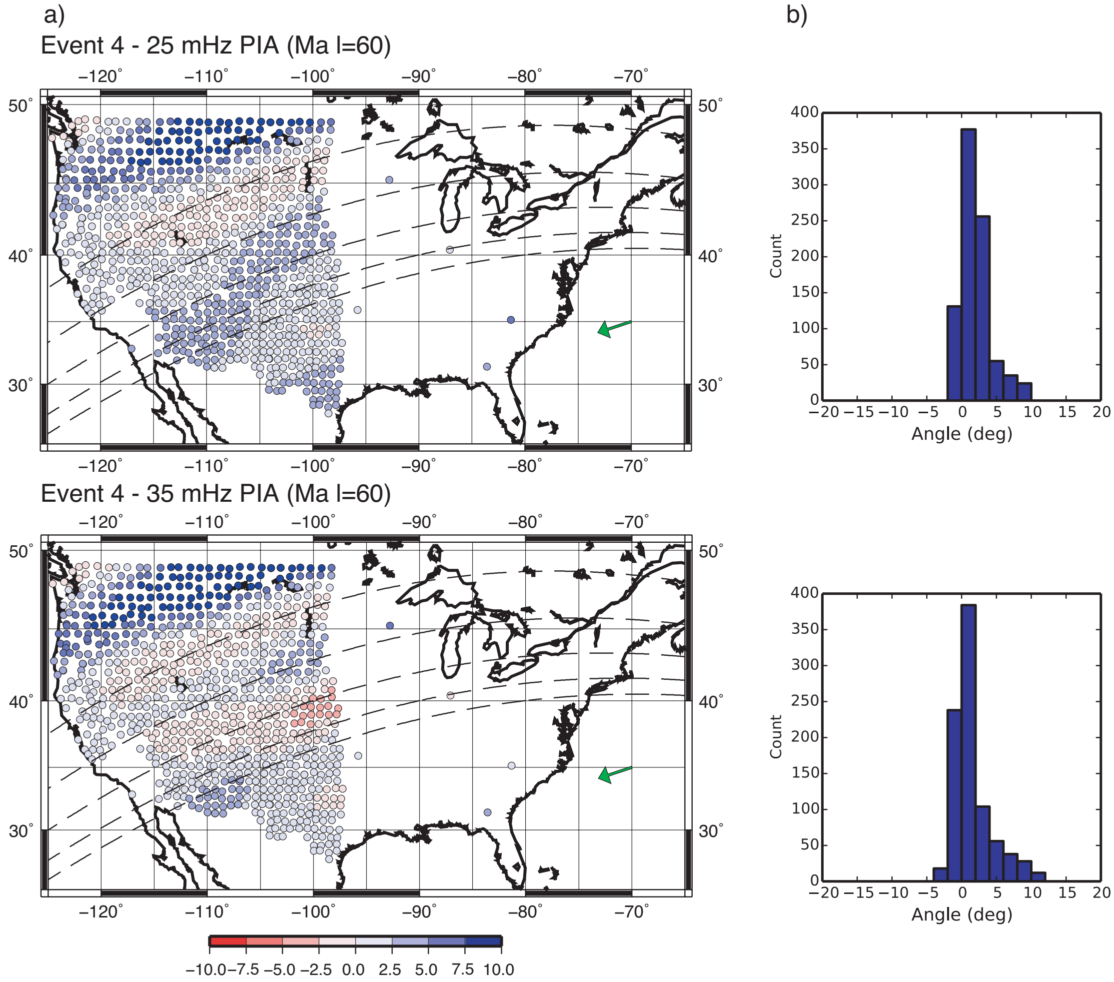

Figure S5. PIA arrival angles at 25 and 35 mHz for events #4. For details see Figure S3.

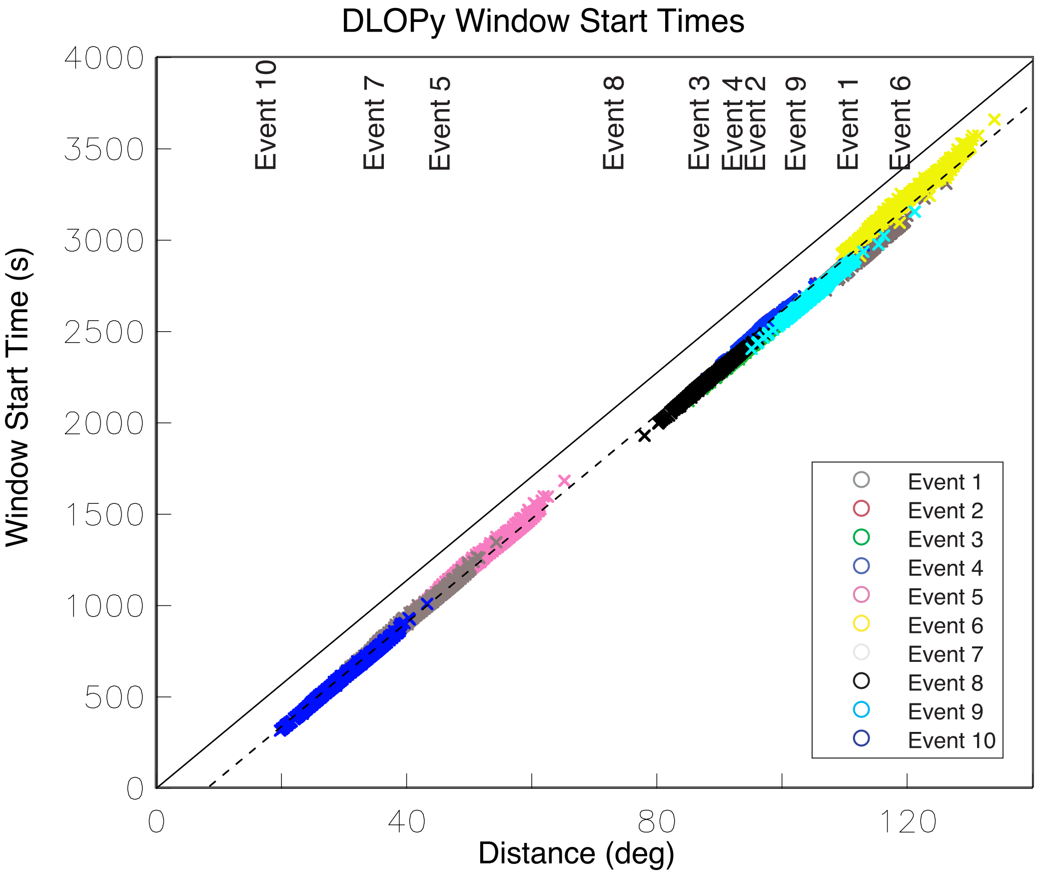

Figure S6. Calculated start time (after event time) of the analysis window for the DLOPy tool corresponding to the manual dataset of user 3 at 15 mHz. The solid line marks a group velocity of 3.91 km/s as shown in Figure 12. The dashed line represents a 250-s shift corresponding to the implied start time of a 500-s DLOPy window.