



RESEARCH ARTICLE

10.1029/2019GC008679

Key Points:

- A geographic bin stacking algorithm was developed to construct “virtual stations,” where seismic waves are slowness-stacked to improve signal-to-noise ratios
- Global S waves, multibounce S waves S_n ($n = 2, 3, 4, 5, 6$), Sdiff, and core-reflected ScS waves and higher multiples ScSn ($n = 2, 3, 4, 5$) for both minor and major arc paths, where visible, were measured
- Sampling of the mantle in the southern hemisphere is improved using multibounce data

Correspondence to:

E. J. Garnero,
garnero@asu.edu

Citation:

Lai, H., & Garnero, E. J. (2020). Travel time and waveform measurements of global Multibounce seismic waves using virtual station seismogram stacks. *Geochemistry, Geophysics, Geosystems*, 21, e2019GC008679. <https://doi.org/10.1029/2019GC008679>

Received 3 SEP 2019

Accepted 20 DEC 2019

Accepted article online 22 DEC 2019

Travel Time and Waveform Measurements of Global Multibounce Seismic Waves Using Virtual Station Seismogram Stacks

Hongyu Lai¹ and Edward J. Garnero¹¹School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA

Abstract We construct geographically localized bin stacks of waveforms, called virtual stations, to enhance signal-to-noise ratios (SNRs) for travel time and waveform measurements of multibounce S and ScS phases (S up to S6 and ScS up to ScS5), as well as direct S, ScS, and Sdiff, on tangential component data. Major arc S and ScS multibounce waves were also measured. Virtual station data are referenced to empirical wavelets constructed from direct S waves for each event. The virtual station approach is useful for low SNR data, bolstering wave path coverage in the southern hemisphere. Goodness of fit measurements between the adapted empirical wavelet and virtual station waveforms are documented, as well as SNRs, allowing for objective definition of travel time measurement quality. From a data set of 360 earthquakes and 8,407 seismographic stations, nearly 4 million records were utilized to construct 248,657 virtual station stacked seismograms, which were compared to best-fitting empirical wavelets. After human inspection of virtual station results, 8,871 travel time measurements were retained from 19 different minor and major arc seismic wave types. Higher multibounce data improve sampling of the southern hemisphere. From 188,003 single seismograms, 3,331 multibounce wave measurements were also made. Comparisons of single seismogram and virtual station stack measurements show a consistent bias: Virtual stack onset times are systematically early due to a broadening effect from stacking records with arrival time differences, which we correct for. The travel time and waveform measurements are publicly available.

1. Introduction

Seismic data quantity and type are central to the imaging of Earth's heterogeneous mantle. While surface waves and normal modes are fundamentally important in seismic imaging of the planet, seismic body waves provide the most detailed information of fine scale heterogeneity within the deep interior. For this reason, tomographic imaging routinely employs a large number of body waves of many different seismic phases (e.g., Auer et al., 2014; Durand et al., 2017; French & Romanowicz, 2015; Grand, 2002; Houser et al., 2008; Koelemeijer et al., 2016; Kustowski et al., 2008; Li et al., 2008; Moulík & Ekström, 2014; Panning & Romanowicz, 2006; Ritsema et al., 2011; Simmons et al., 2010). Comparisons of different global seismic tomography models reveal very similar large-scale structures in the lower mantle (Cottaar & Lekic, 2016; Garnero et al., 2016; Lekic et al., 2012). However, the intermediate- to small-scale features vary between models (e.g., Becker & Boschi, 2002; Garnero et al., 2016), suggesting uncertainties in the short to intermediate wavelength structure from model to model.

Greater availability of global seismic network data makes possible large data set construction, which may help bolster seismic coverage for some regions. To this end, in a previous study (Lai et al., 2019; hereafter Lai19), we used an empirical wavelet construction method to document seismic wave travel time and waveform measurements for the six phases S, Sdiff, SS, SSS (S3), ScS, and ScSScS (ScS2). An empirical wavelet for each earthquake was iteratively constructed from observed S waves and used as a reference waveform for correlative comparisons with observed waveforms for that event. The method involved adapting the empirical wavelet to best-fit every phase of interest, ultimately yielding ~250 K travel time and waveform measurements for the above six phases. However, the wave path sampling of the deep interior from this data set is significantly better in the northern hemisphere than in the southern hemisphere (roughly five times more sampling in the northern hemisphere). Also, many of the data are from common source-receiver geometries (which has not significantly changed over the decades), such as key subduction zones to continents with long-standing seismic networks. In an attempt to mitigate path geometry redundancy, Lai19 prioritized

events in unique locations, by pursuing earthquakes with a maximum distance to other earthquakes in the data set. This still resulted in a sampling bias toward the northern hemisphere.

In an attempt to address the wave path coverage bias, here we investigate multibounce seismic waves: S and ScS waves that bounce multiple times off the surface (S and ScS) and core-mantle boundary (ScS). While some of these waves are used in tomographic imaging, e.g., ScS multibounce waves up to ScSScSScSScS (ScS4) (French & Romanowicz, 2015; Ritsema et al., 2011), such multibounce energy is not currently commonly employed, and higher multiples have not been utilized. Here we pursue any and all seismic body waves apparent on global profiles of transverse component S waves which were not measured in Lai19. These include the major arc S2 and S3 (S2m and S3m, respectively, where the major arc wave path departs from the earthquake in the opposite direction from the station along the great circle path), SSSS (S4), S4 major arc (S4m), SSSSS (S5), S5 major arc (S5m), SSSSSS (S6), S6 major arc (S6m), ScSScSScS (ScS3), ScS3 major arc (ScS3m), ScSScSScSScS (ScS4), ScS4 major arc (ScS4m), ScSScSScSScSScS (ScS5), and ScS5 major arc (ScS5m). These phases, especially the major arc phases, can have very long distances and often have paths that traverse the southern hemisphere and provide new and unique wave path sampling locations.

For the phases used in Lai19 (S, SS, S3, ScS, ScS2, Sdiff), an average of roughly 20% of all seismic waves investigated were retained (this varied phase-by-phase), with the high rejection rate being predominantly related to low signal-to-noise ratio (SNR) data. To utilize more of the available seismic data and successfully pursue the commonly lower amplitude multibounce phases, in this paper we adopt a stacking-based method to improve the SNR of low amplitude (and sometimes below the noise level) multibounce energy by slowness stacking seismic waves in geographically confined regions, or “bins.” The principle benefit of an array approach is a far improved SNR of waves of interest (Frost et al., 2013; Frost & Rost, 2014; Rost & Thomas, 2002), which allows measurements of data that would otherwise be rejected from single-seismogram approaches, e.g., as that used in Lai19. We call the geographical bin stacked data in this paper virtual stations and measure travel times and waveform information using a reference empirical wavelet made from stacking S waves on an event-by-event basis (as in Lai19). In addition to employing the virtual station approach on these multibounce phases, in this study we also (1) identify wave path corridors for the phases of Lai19 for which there are data but no retained measurements (thus, we pursue enhancing SNRs with virtual stations for either unusual or noisy source-receiver geometries), and (2) subsequently investigate all multibounce phases for the possibility of single seismogram measurements.

In section 2 that follows, we first introduce our data collection and preprocessing procedures. In section 3 we show the workflow of the construction of virtual stations and section 4 presents the virtual station travel time acquisition, measurement trends, single seismogram measurements, and the global coverage of our data. We explore the dependency of our results on some of our assumptions in section 5. Conclusions follow with information on access to the freely available measurements made in this study.

2. Global Data Set

2.1. Data Collection

In this paper, we investigate earthquakes analyzed in Lai19, so that results of travel time and waveform analyses here can be directly combined with waves of that study (S, SS, S3, ScS, ScS2, Sdiff) from the same earthquakes. We collected all available data from several data agencies that distribute seismic data freely, including the Incorporated Research Institutions for Seismology (IRIS, <http://www.iris.edu>), the Observatories and Research Facilities for European Seismology (ORFEUS, <http://www.orfeus-eu.org>), the Northern California Seismic Network (NECDC, <http://www.ncedc.org/ncsn>), the F-net Broadband Seismograph Network (F-net, <http://www.fnet.bosai.go.jp>), and the Canadian National Seismic Network (CNSN, <http://www.earthquakescanada.nrcan.gc.ca/stdon/CNSN-RNSC>). See the International Federation of Digital Seismograph Networks (doi:10.7914/SN/II) for more information.

Data were collected using various webservices tools supported by the different data agencies. A two-step process was employed. First, we collected deep focus events (>50 km depth) in the time period from January 1994 to October 2017, with moment magnitude greater than 6.0. Smaller moment magnitude earthquakes displayed higher noise levels and inconsistent data quality. Deeper earthquakes tend to have simpler source time functions and less complications from crustal structure (e.g., depth phases and crustal reverberations)

interfering with phases of interest. This resulted in 733 earthquakes. The location and origin time information was taken from the International Seismological Centre (ISC) bulletin (<http://www.isc.ac.uk/iscbulletin/search/catalogue>). However, deep earthquakes are restricted geographically to subduction zones, thus resulting in restricted wave path coverage. Therefore, in a second step we collected select shallow earthquakes (source depth <50 km) for the same time period, based on a prioritization scheme that (1) favored more recent time periods (to take advantage of more abundantly available data) and (2) ranked events more favorably for having the greatest distance from the deep focus events. The shallow events thus play a fundamentally important role in expanding the global coverage provided solely by the deep earthquakes to help normalize source distribution. This amounted to 113 shallow earthquakes. All events were inspected for possible contamination from other events which occurred nearby in time; if any energy was apparent, the later event was rejected from our catalog. Earthquakes were retained if adequate SNR was present such that measurements could be confidently made. This resulted in a final collection of 360 earthquakes (247 deeper than 50 km, 113 shallower than 50 km).

2.2. Basic Preprocessing

We collected a 2-hr time window length following the earthquake origin time for all available seismic stations for all events. For each station, the horizontal components of motion for the broadband data were rotated to the great-circle path to obtain the radial (R) and transverse (T) components of motion. This paper analyzed the T components to study SH waves, thus results of this work can be combined with those of Lai19. The instrument response for every station was removed through deconvolution using pole-zero files supplied by the data agency, and data were analyzed on velocity recordings. All data were band pass filtered in the period range between 16 and 100 s. This was the same period range as used in Lai19.

An example event that occurred on 24 May 2013 (latitude: 54.87, longitude: 153.28, source depth: 608.9 km, and moment magnitude: 8.3) is used to show phases we pursue in this study. While this is an exceptionally large earthquake (the largest in our data set), it is useful because it clearly shows all phases used in this study. Figure 1 shows almost 2 hr of recordings over the entire distance range. Due to the large number of recordings we have for this event (1,801 stations), we linearly stacked records in every 1.0° distance window to produce a clearer display of the seismic arrivals that are present. Travel time curves for the PREM model (Dziewonski & Anderson, 1981) are also shown. A slightly reduced upper corner to the band pass filter (20 s instead of 16 s) was used to further suppress noise for the purpose of the figure. Direct and multibounce S waves are seen out to S6 (major arc), and multiple ScS waves are seen out to ScS5. All of the depth phases are also visible, though in this study we do not pursue the depth phases because of the sometimes-complicated wave shapes. This event makes clear the possibility of measuring the higher multibounce waves. Figure 1a also shows evidence for waves that have traveled in the minor arc direction, and then continue to make an additional round trip (i.e., 360° more). These can be seen between 12° and 30° (minor arc distance) starting at around 5,500 s. The first of these arrivals is S4 + 360°. These waves are rare in our data set so we do not include them in our analyses.

To see some of the waves of interest more clearly, Figure 1b presents two zoomed in boxes (“Z1” and “Z2”) which are shown in Figure 2. As in Figure 1, these records are distance increment stacked (but with fewer records per stack). In Figure 2a, we show data from 85–90° where strong S, SS, S3, S4, and ScS2 phases are present. In Figure 2b, we show a slightly expanded distance range, but for a much later time window to highlight later arriving multibounce energy. Clear major arc multibounce waves are seen for S (blue lines: S3m, S4m, S5m, S6m) and minor and major arc multibounce waves for ScS (light-red lines: ScS4, ScS5, ScS3m, ScS4m, and ScS5m). At distances larger than 95°, there are fewer records per stack, and the seismograms are significantly noisier. This demonstrates the benefit of stacking data to see the higher multibounce waves.

Smaller events in this study differ from the large earthquake of Figure 1 in that they may contain multibounce waves over a restricted azimuth or distance range, due to favorable radiation pattern for some path geometries, and thus these waves may not be uniformly present. As mentioned, data from shallow events can be noisier than deep focus earthquakes, but they are more abundant, which affords finding events with multibounce phases of interest for all source depths. Figure 3 shows two example events that are smaller (roughly magnitude 6), with one being very shallow. These events represent typical events from our study and highlight the presence of multibounce waves (both minor and major arc).

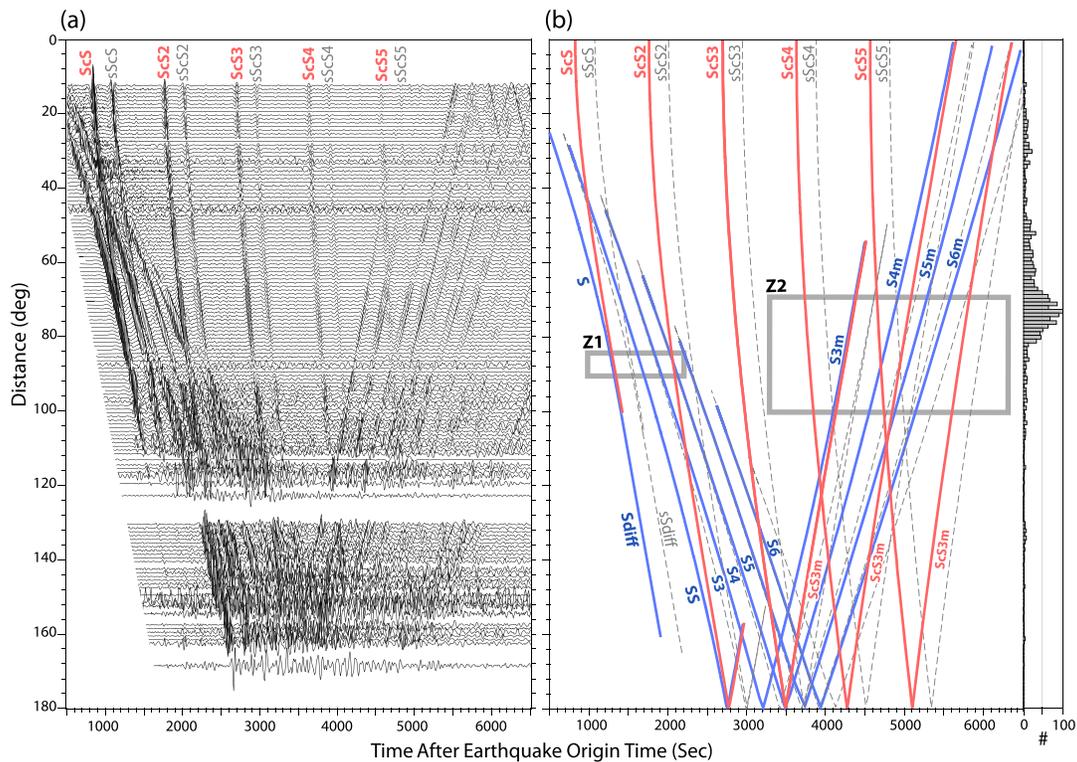


Figure 1. (a) Record section distance profile of transverse component velocity recordings for an event on 24 May 2013 (Latitude: 54.87°, Longitude: 153.28°, Depth (Z): 608.9 km, Magnitude: 8.3 Mw). Seismograms were low-pass filtered at 20 s, then linearly stacked in 1° distance bins. (b) Travel time curves for the principle phases of interest are shown, and include direct S, Sdiff, and multibounce S waves (minor and major arc) as blue lines. ScS and higher multiple bounces are also shown (light-red lines). Depth phases are shown as gray dashed lines. Any phase name ending with an “m” is a major arc wave. The two gray boxes (marked Z1 and Z2) are zoomed in time-distance regions, which are shown in Figure 2. The number of seismograms in each stack shown in panel (a) are presented on the right of panel (b) as histogram bars. The most populated stacks have up to 100 seismograms.

The ray paths of the multibounce waves pursued in this study are shown in cross-section plots in Figure 4, along with the paths of the six phases that were the focus of Lai19. The multibounce S_n and S_{nm} waves densely sample the upper part of the mantle while the ScS_n and ScS_{nm} waves sample the entire mantle and provide redundant sampling of the lowermost mantle. Figure 4f demonstrates the potential sampling density improvement when incorporating all the multibounce data together (for one earthquake and the two distances portrayed in the figure).

3. Virtual Station Seismograms

3.1. Development of an Adaptive S wave Empirical Wavelet

In this study we adopt the method used in Lai19 to construct an average shape of the S wave, on an event-by-event process. The method is briefly summarized here (and more details and examples can be found in Lai19). First, S waves are used because they are the first arrival (before core diffraction distances) and have the best SNR compared to the energy in front of the phase. For each earthquake, we stack all seismograms at distances larger than 30° to avoid complications from upper mantle triplications (e.g., Grand & Helmerger, 1984; Tao et al., 2017) and smaller than core diffraction (roughly 100 to 103°, depending on source depth) to minimize possible attenuation, scattering, or multipathing effects from long paths at the base of the mantle (e.g., Ni et al., 2002, 2003, 2005; Ritsema et al., 1998). Waves are initially stacked on the PREM-predicted time, and the resulting stack is used to iteratively realign every record to the previous stack (using cross correlation, with the cross-correlation coefficient (CCC) and SNR used to weight records in subsequent stack iterations), then restacked. When the stack converges to a shape that no longer changes, we define this as the Generalized Empirical Wavelet (GEW). However, the GEW was constructed with some records that are broader or narrower than the GEW. Therefore, the shape of the GEW onset is an average (i.e.,

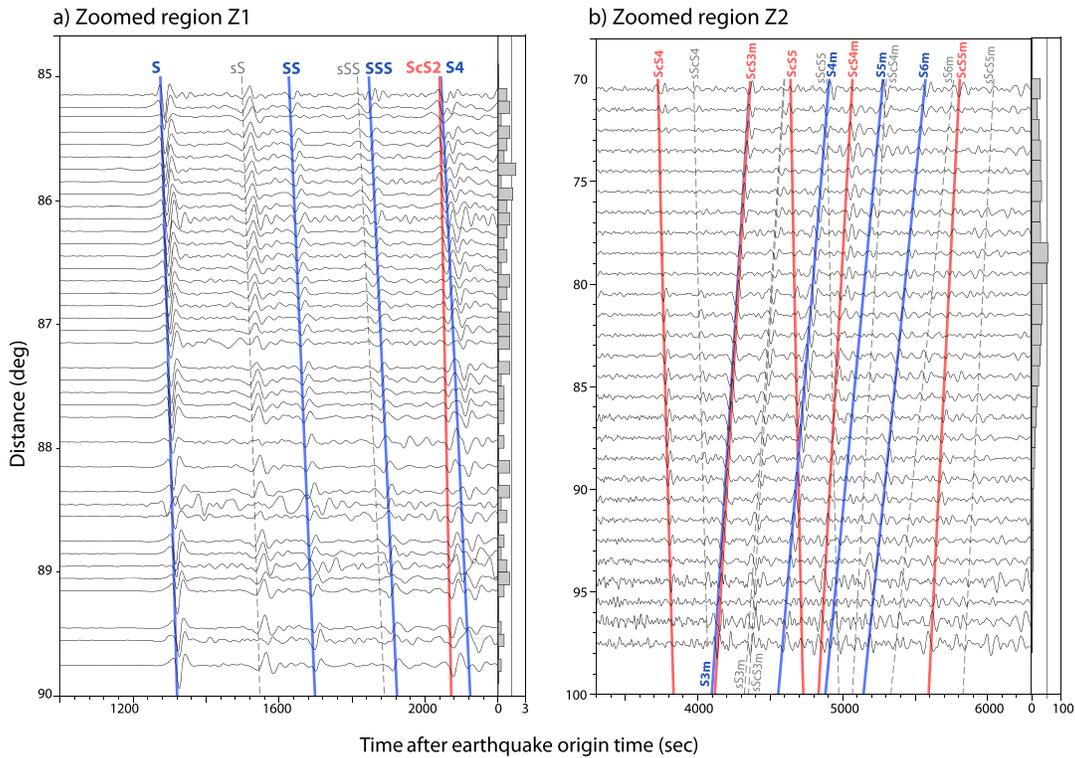


Figure 2. Zoomed in record sections for same event of Figure 1 and regions demarked in Figure 1b. As in Figure 1, records are stacks of small distance ranges. Tangential component velocity records are shown. (a) Zoomed area Z1 showing stacked records exhibiting phases S, SS, S3, ScS2, and S4. Dashed lines correspond to depth phases (colors are as in Figure 1). The distance increment used in stacking is 0.1° , and number of records in each stack is shown in the histogram to the right of the record section. (b) Zoomed area Z2 showing stacked records exhibiting phases S3m, S4m, S5m, S6m, ScS4, ScS5, ScS3m, ScS4m, ScS5m, and depth phases. A larger time window is shown than for panel (a), and the distance increment used in the stacking is 1° . All else is as in panel (a).

blurred) shape. To arrive at a more representative shape function, every record is made to fit the GEW by expanding or shrinking it, then the GEW stack is updated once more. Lai19 dubbed this the Stretched Empirical Wavelet (SEW), which has the advantage over the GEW in that it has a more representative onset shape for all records, if the SEW shape is adapted to fit individual records.

Thus, for each event, an SEW was constructed for comparison to every wave of interest. The SEW was adapted to fit to each observed wave by either broadening it from convolution with a t^* operator (Futterman, 1962) (to match records broader than the SEW) or narrowing it (to match records narrower than the SEW). An onset time is fixed to the adapted SEW through correlation with a Gaussian function in the time domain, G_t , which has the onset time hardwired to the onset of the Gaussian (as defined as the time associated with 0.01 amplitude for a Gaussian with the peak amplitude defined as 1). The Gaussian function is defined as follows:

$$G_t = e^{-\frac{t^2}{2g^2}}, \quad (1)$$

where i is the number of time points (and thus the length of the function in time points), e is Euler's number, and g is the Gaussian factor (which controls the width of the function).

This process is automated and provides a stable and objective manner for travel time determination, as well as documentation of wave shape broadness, and other factors (described in section 3.3). Figure 5 shows an example of a virtual station stack of a major arc ScS5 (ScS5m, introduced in section 3.3), the best-fitting t^* ed SEW trace, and the best-fitting Gaussian function with the onset time determination. The onset of ScS5m is well captured by this approach.

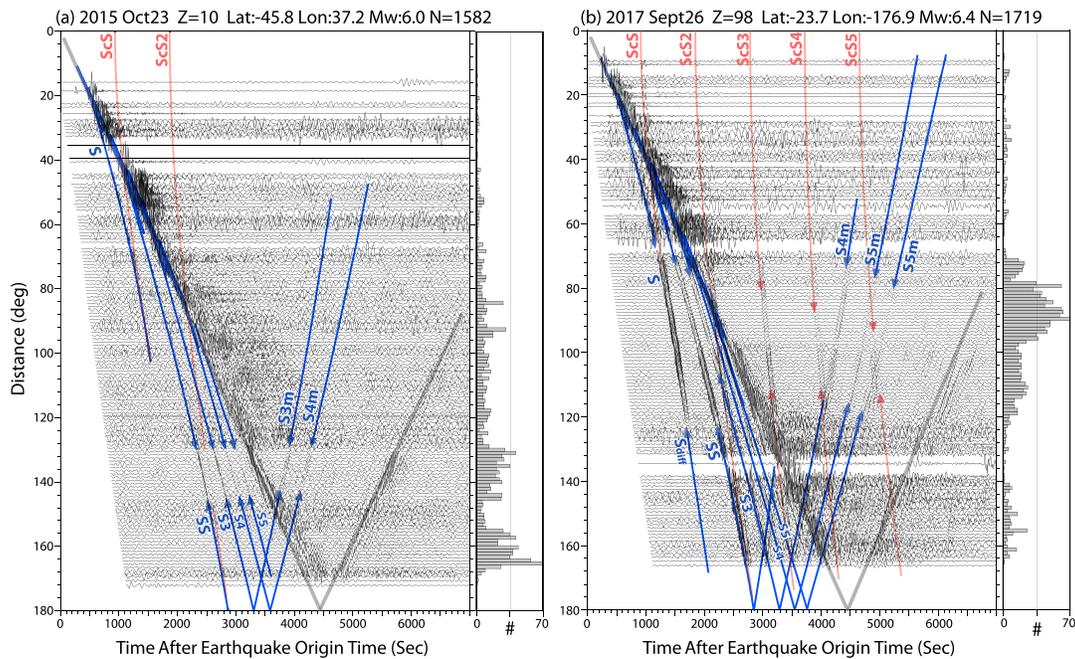


Figure 3. Transverse component distance-increment stacked record sections as in Figure 1 (both panels use 1° for distance stacking). (a) A shallow ($Z = 10$ km) and relatively small ($M_w 6.0$) southwest Indian rise event. Minor arc SS, S3, S4, and S5 are clearly apparent (between arrows on the travel-time curves), as are the major arc S3m and S4m. (b) A relatively shallow ($Z = 98$ km) and small ($M_w 6.4$) Fiji event. Both minor and major arc multibounce S waves are apparent, as are higher multiple ScS waves.

3.2. Virtual Station Grid Parameterization

To set up virtual station stacking, we first defined a grid of equal area cells on Earth's surface. A network of grid points separated by 700 km in latitude and longitude was established, each with a radius of 500 km (thus, the entire globe is covered by the grid cells). Several different grid cell separations and radii were investigated, but smaller grid cells resulted in significantly fewer viable virtual stations away from dense seismic networks. Every earthquake and station were assigned to the grid cell within which they were located. Thus, every grid cell has a list of associated events and stations that are within it (if any). Then, for every event-station grid pair, we searched the station grid cell for the existence of records for every phase of interest. If the number of records for a phase of interest was greater than a user-defined threshold, we proceeded to construct a virtual station stack for that phase of interest. In this study, if there are three or more records in the grid cell, we proceeded with making the virtual station stack. We experimented with different thresholds. The minimum of three was chosen as a compromise between having ample virtual stations globally (larger threshold values limited virtual station global coverage) and not having enough stations populating any grid cell (fewer than three commonly yielded noisy stacks). The 360 earthquakes and 8,407 unique stations (Figure 6a) in our data set resulted in 289 unique virtual station grid cells (shown as the blue circles in Figure 6b).

For all the phases of interest in this study (Figure 4) and every virtual station possibility for our data set, the grid cell center was relocated to be the geographic center of the contributing stations within the initial virtual station bin. Updated (relocated) grid center locations resulted in roughly ~ 248 K unique virtual station seismograms for our phases of interest. Every virtual station is assigned its own unique identification number, and along with this number we store the seismic phase name, the relocated virtual station latitude and longitude, earthquake information, epicentral distance, azimuth, and back azimuth. The locations of virtual stations are shown as small black triangles in Figure 6b. They differ from each other within a given virtual station cell because each was relocated to be the center of the stations available for the particular phase and earthquake being stacked

For the six phases of Lai19, the 20% retention rate (roughly 250 K measurements were retained from ~ 1.4 million seismograms investigated) was of course accompanied by an $\sim 80\%$ record rejection rate, commonly because of low SNR. Therefore, in addition to the multibounce waves not investigated in Lai19, we

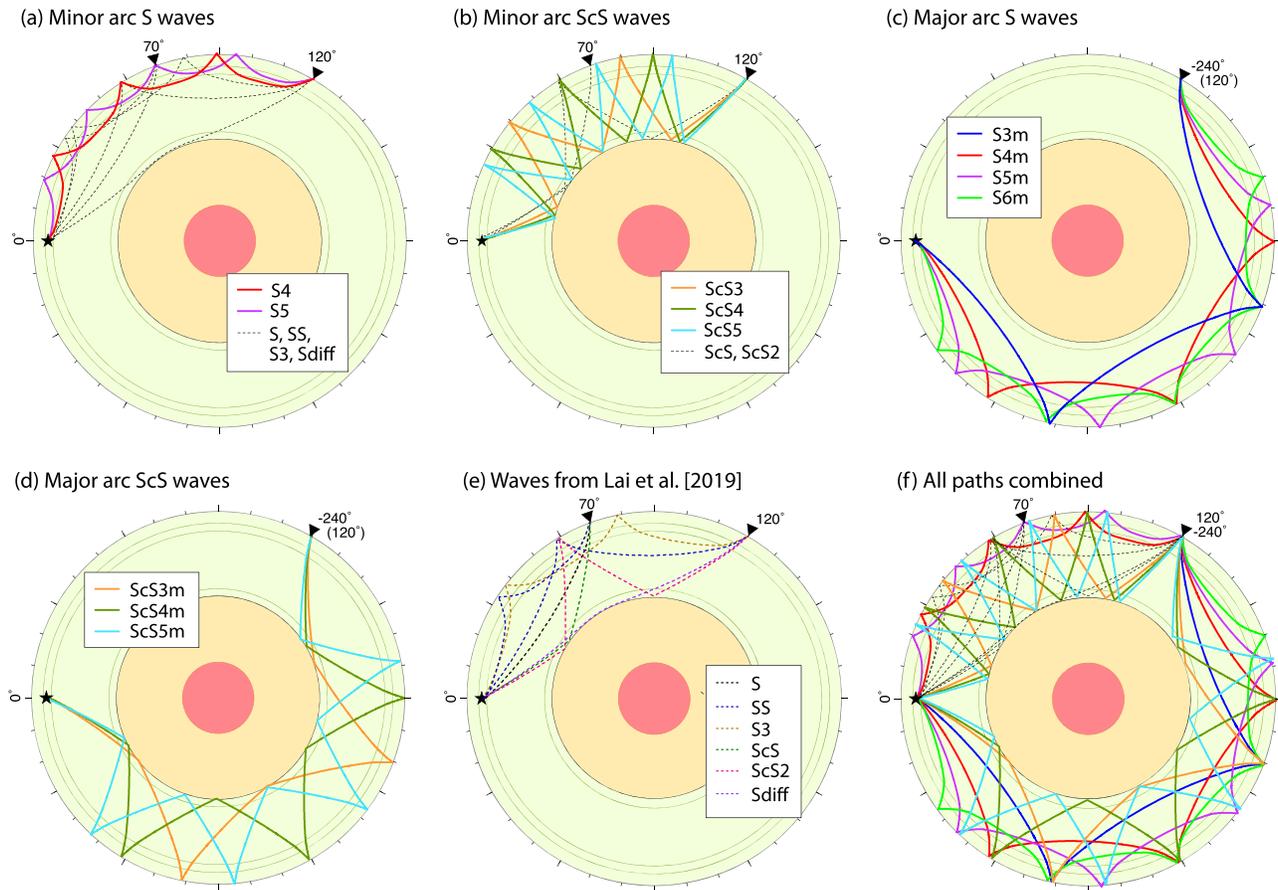


Figure 4. Cross-sections showing ray paths of a 500 km deep earthquake for (a) wave paths of minor arc multibounce S waves S4 and S5 (at 120°) and paths of S, SS, S3, and Sdiff studied in Lai19 at 70° and 120°; (b) as in (a), but minor arc multibounce ScS waves are shown (ScS3, ScS4, and ScS5 of this study, ScS and ScS2 of Lai19); (c) major arc multibounce S waves (S3m, S4m, S5m, S6m at 120° minor arc distance, 240° distance in the major arc direction); (d) as in (c), but major arc multibounce ScS waves (ScS3m, ScS4m, and ScS5m); (e) the six seismic waves studied in Lai19; (f) all of the paths of (a) through (e) combined, with all phases pursued in this study different from Lai19 given in color (the gray dashed lines are the six phases of Lai19).

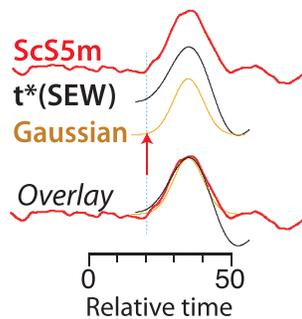


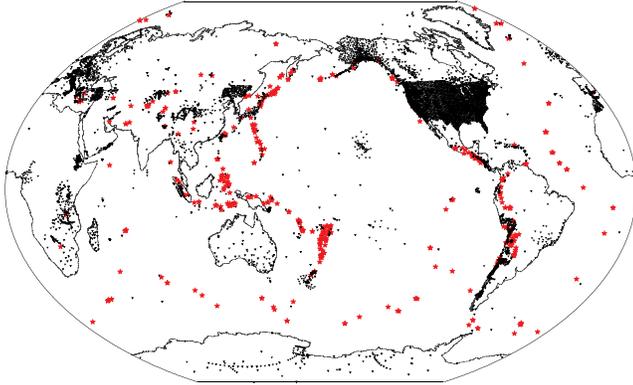
Figure 5. Example comparison of a major arc ScS5 virtual station stack (ScS5m, red traces) with an SEW that has been broadened by convolution with a t^* operator to best-fit the ScS5m stack. The Gaussian function that best-fits the broadened SEW is also shown (orange trace) along with a red arrow which indicates the automatically assigned onset time. The traces are overlaid to show the wave shape comparisons. Event information: 15 April 2017, Latitude: -23.3° , Longitude: -67.8° , Z: 155 km, Mw: 6.3, minor arc epicentral distance 66.9° , major arc distance 293.1° .

implemented the virtual station approach with the basic six phases (S, SS, S3, ScS, ScS2, and Sdiff) with the goal of extracting more measurements from the previously rejected records of Lai19, as they could potentially provide new sampling geometries. With this goal in mind, we preclude virtual station construction if we already possess three or more measurements for any virtual station geometry for the six phases of Lai19. Thus, the virtual station approach for the basic six phases is used to fill in unsampled regions and to bolster poorly sampled corridors.

3.3. Virtual Station Seismogram Stacking and Measurements

A travel time prediction is computed for the PREM model for every seismogram (and phase of interest) in the retained virtual station cells. For each virtual station cell, phases are aligned on the PREM-predicted time, and weighted according to the distance to the relocated virtual station bin center with a Gaussian factor (a weight of 1 is assigned at the bin center that reduces to 0.5 at the edge of the virtual station circle). We utilize focal mechanisms from globalcmt.org (Dziewonski et al., 1981; Ekström et al., 2012) to predict the radiation pattern amplitude of all phases (between -1.0 and 1.0) which is used to predict the polarity of all phases of

(a) Event and station coverage



(b) Virtual station cells and stations

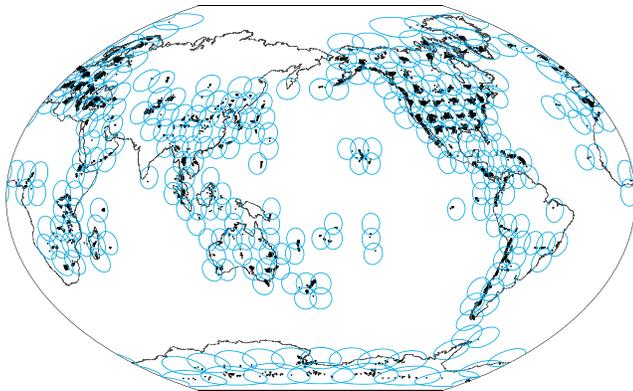


Figure 6. (a) Geographical locations of 360 earthquakes (red stars) and 8,407 seismographic stations (black triangles) used in this study to explore virtual station construction. This is the same data set as Lai19. (b) Locations of 289 initial virtual station grid cells (blue circles), and roughly 247,000 distinct relocated virtual station locations (black triangles).

between the start and stop time of the phase of interest, as previously described as one-half period) and the noise window (in a window from -100 to -20 s relative to the PREM-predicted time for the phase of interest) are divided by each other. We denote this as SNR_{VS} . We also document the average of the SNRs computed for each record used to create a virtual stack seismogram (notated as $\overline{\text{SNR}}$). Additionally, we computed SNR on the virtual stack by dividing by the maximum peak-to-trough amplitude within one period (defined from the phase of interest) and in the noise (in the same 80 s noise window as SNR_{VS}). We denote this as $\text{SNR}_{\text{peak_trough}}$. The $\text{SNR}_{\text{peak_trough}}$ measurement is more affected by a single large pulse in the noise window than SNR_{VS} . We document if any known seismic waves are predicted to arrive in the time window used to estimate the noise level. While other SNR measurement approaches are possible (e.g., Sheriff & Geldart, 1995), we use the same as in Lai19 for comparison purposes, so users of both data sets have a common SNR measurement approach.

Waveform distortions may occur for some path geometries from a variety of sources, including multipathing, where wave energy can take different paths when tangential to or in the vicinity of large-scale heterogeneities and significantly broaden pulses (Cormier, 1989; Ni et al., 2002, 2005; Ni & Helmberger, 2003a, 2003b; Vidale, 1987) or scattering from heterogeneity, either fine-scale which can diminish high frequencies or scales comparable to (or larger than) the dominant seismic wavelength which can result in additional scattered arrivals that can manifest as precursory or postcurious energy (Bréger & Romanowicz, 1998; Flanagan & Shearer, 1998; Rost et al., 2008; Rost & Earle, 2010; Rychert & Shearer, 2010; Toh et al., 2005). We document waveform differences between the observed virtual stack seismograms and the best-fitting SEW through measurement of the average of their amplitude differences when aligned at their maximum cross correlation. We define this as misfit:

interest. For any virtual stations possessing records predicted to be in different quadrants of the radiation pattern, records are flipped to the same polarity before virtual station stacking. Finally, all weighted records are stacked.

Three examples of virtual station stacking are presented in Figure 7. The virtual station construction process is shown for a minor arc ScS4 in Figure 7b, for a deep focus Fiji event recorded in northern Japan. The 21 original seismograms are stacked along the PREM-predicted ScS4 slowness, resulting in a much-improved SNR (5.0) in comparison to the average SNR of the constituent records (1.2). The SNR measurement method is defined below. A major arc example for ScS5m is presented in Figure 7c, for a deep focus South American earthquake recorded at 96 stations in a virtual station cell in the northeastern United States. Again, a vast improvement in SNR is apparent (3.6 for the stack compared to 1.2 for the average of the constituent traces). Lastly, a shallow Alaskan earthquake recorded in a virtual station bin with 80 stations is shown for S5m (Figure 7d). Again, the SNR improvement is apparent.

For all virtual station stacks (e.g., Figure 7), best-fits of the stretched SEW to the virtual station stack and the Gaussian to the stretched SEW are determined by cross correlation (as in Lai19). We also measure the CCC between the virtual station stack with the GEW (the unstretched, initial empirical wavelet). These correlation coefficients along with the measured and predicted arrival times are also stored. For the purpose of estimating wave period, we approximate a start and stop time of the arrival of interest using the time associated with a 0.1 amplitude level (with the arrival peak defined as 1) to denote the beginning and end of the best-fitting SEW. The time length between these two times defines an estimate of half wave period of the phase of interest in velocity recordings.

As noted in Lai19, there are several ways to document the SNR. For virtual stations, we adapted the average amplitude approach, whereby the average amplitude in both a signal window (approximated as the time

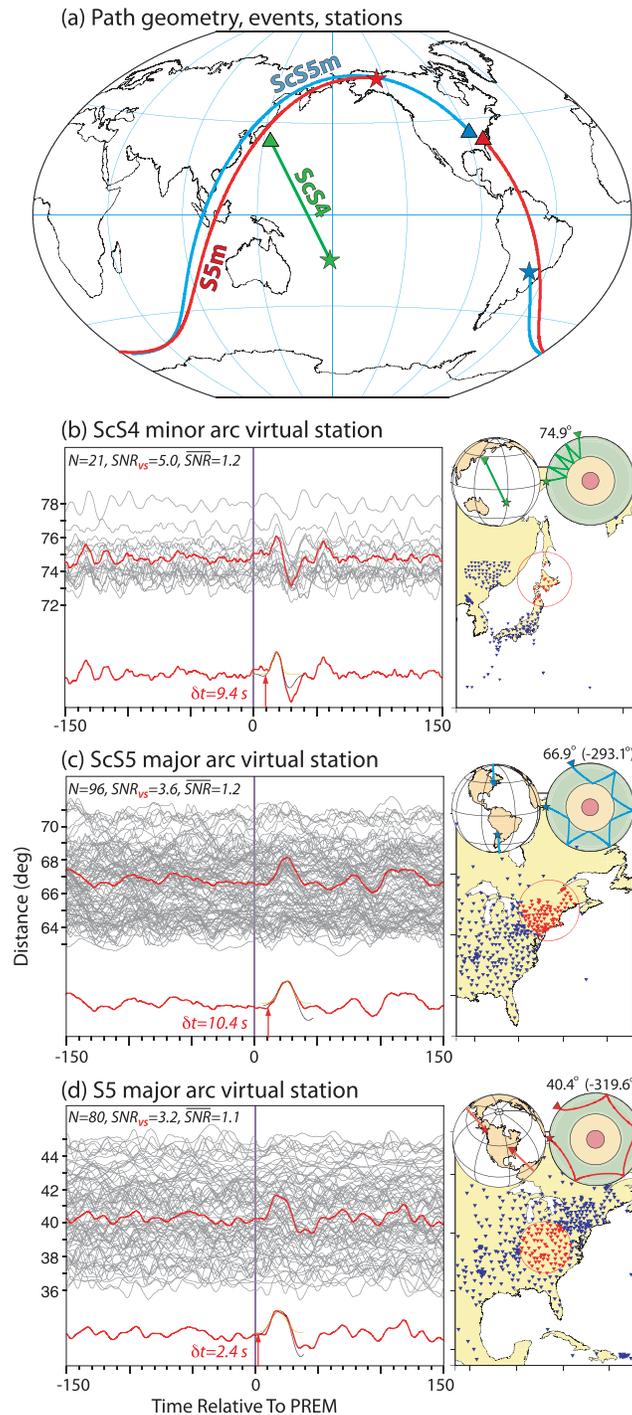


Figure 7. Virtual Stacking Examples. (a) Map showing great circle ray path geometries of examples shown in rest of the figure, with earthquake (stars) and virtual station bin center (triangles) locations. (b) Record section of 21 transverse component velocity recordings of minor arc ScS4 (left, gray traces) for a virtual station bin in northern Japan (map to the right, red triangles are those of the traces on the left, blue triangles are the available stations for that event). Event information: 24 February 2017, Latitude: -23.3° , Longitude: -178.8° , Z: 415 km, Mw: 6.9. Traces are aligned on the PREM time predicted for ScS4. Path geometry is also shown on the right (above the map). Virtual station stack is the red trace plotted at the distance of the virtual station bin center (74.9°), and also plotted below with the best-fitting SEW and Gaussian function, along with the onset time predicted by the empirical wavelet method (red arrow), here, 9.4 s after the PREM prediction. The number of stations in the stack (N), the SNR of the virtual station stack (SNR_{VS}), and the average SNR of the contributing records (SNR) are printed in the upper left. (c) As with (b), except for ScS5m (ScS5 major arc). Both minor and major arc distances are indicated above the ray path cross-section panel (upper right). Event information: 15 April 2017, Latitude: -23.3° , Longitude: -67.8° , Z: 155 km, Mw: 6.3. (d) as with (c), except for S5m (S5 major arc). Event information: 1 May 2017, Latitude: 59.8° , Longitude: -136.7° , Z: 3 km, Mw: 6.3.

$$Misfit = \frac{\sum_{i=1}^n |A_i^{VS} - A_i^{best-fit SEW}|}{n}, \quad (2)$$

where A_i^{VS} and $A_i^{best-fit SEW}$ are the amplitudes of virtual stack and best-fitting SEW (e.g., the red and black traces, respectively, in Figure 5) at the i^{th} point, measured across a one period window, and n is the number of points in this window. When measured over one period the phase of interest, we denote this as $Misfit_{SIGNAL}$. We also compute the misfit over one period in the time window immediately preceding and following the signal window, defined as $Misfit_{PRE}$ and $Misfit_{POST}$, respectively.

Lai19 introduced a comprehensive weighting scheme for the basic six phases they investigated, as a way to empirically establish comparative data quality between the different phases, for the purpose of future imaging experiments. We use the same approach here, where five individual attributes (SNR_{VS} (w_{SNR}), the CCC between the stretched SEW and the virtual station stack (w_{CCC}), and the misfit measurements of the main phase, precursor and postcursor time windows w_{Misfit_signal} , w_{Misfit_pre} , and w_{Misfit_post} , respectively) are multiplied to define a comprehensive weight $w_{comprehensive}$:

$$w_{comprehensive} = w_{SNR} \times w_{CCC} \times w_{Misfit_signal} \times w_{Misfit_pre} \times w_{Misfit_post}. \quad (3)$$

The weighting factors on the right side of equation (3) are as defined in Figure 13 of Lai19 and kept the same so the data set presented in this paper can be easily incorporated with that of Lai19 using the same comprehensive weight values. We briefly reiterate those functions here: w_{SNR} is 1 for $SNR \geq 5$, and linearly decreases to 0.5 at $SNR = 2$, and is fixed at 0.5 for $SNR < 2$; w_{CCC} is 1 for $CCC \geq 0.98$, and linearly decreases to 0.5 at $CCC = 0.92$, and is fixed at 0.5 for $CCC < 0.92$; w_{Misfit_signal} is 1 for $Misfit_{SIGNAL} \leq 0.05$, and linearly decreases to 0.5 at $Misfit_{SIGNAL} = 0.30$, and is fixed at 0.5 for $Misfit_{SIGNAL} > 0.30$; w_{Misfit_pre} is 1 for $Misfit_{PRE} \leq 0.10$, and linearly decreases to 0.5 at $Misfit_{PRE} = 0.20$, and is fixed at 0.5 for $Misfit_{PRE} > 0.20$; and w_{Misfit_post} is 1 for $Misfit_{POST} \leq 0.50$, and linearly decreases to 0.2 at $Misfit_{POST} = 1.0$, and is fixed at 0.2 for $Misfit_{POST} > 1.0$. While empirically developed, this weighting factor presents a simple approach of comparing and ranking measurements made in this study.

3.4. Quality Control

After implementing virtual station stacking for all multibounce phases and the phases of Lai19 for the 360 events, 248,657 virtual stations stacks were constructed (~181,451 for new multibounce phases introduced here, and 67,206 for the six basics phases of Lai19). The number of seismograms used in this process was 3,961,572. The multibounce phases have significantly longer travel paths in the mantle, and thus are lower amplitude and more attenuated than direct waves. We thus routinely found lower SNR for multibounce waves than for direct waves. Here we followed the approach of Lai19 and constructed Portable Document Format (PDF) files displaying all virtual station seismograms with the best-fitting SEW (and the Gaussian that best-fits the SEW) plotted on top of the phase of interest. This was done on an earthquake-by-earthquake and phase-by-phase basis. Sixteen virtual station seismograms were plotted per PDF page, which allows simple and effective user interactive reviewing of the virtual station stacks, SEWs, and onset time estimations. We empirically determined SNR, CCCs, and misfit values for which our algorithm retained or rejected virtual station stacks. Specifically, a record was retained if $SNR_{VS} \geq 1.8$, the CCC between the virtual station stack and the best-fitting SEW was greater than 0.85, and the misfit measurement $Misfit_{PRE}$ was less than 0.3. Records slated to be retained are shown in the PDF with a red X in an interactive box to the right of the waveform; rejected records are shown with the box left unchecked. Our algorithm does not select traces if predictions for interfering phases arrive less than 25 s from the PREM-predicted time. Since records are slowness stacked on the phase of interest, many interfering phases do not stack coherently, so the user can update the selection to be retained. An additional box to the left of the waveform can be used for alphanumeric input, if needed. The retained or rejected choices made by the algorithm can be updated by the reviewer of the files, then saved, and the choices subsequently extracted from the PDF files by scripts. Records selected by the algorithm are rejected if the onset times in the virtual station stack are not clearly defined. Alternatively, sometimes the algorithm did not pick a good record due to a low SNR value, owing to some energy far in front of the phase of interest. Thus, the user can reselect any good data that may have been rejected. Figure 8 shows an excerpt of part of a page from a PDF catalog for S4m.

SSSS (S4) major arc empirical wavelet fits

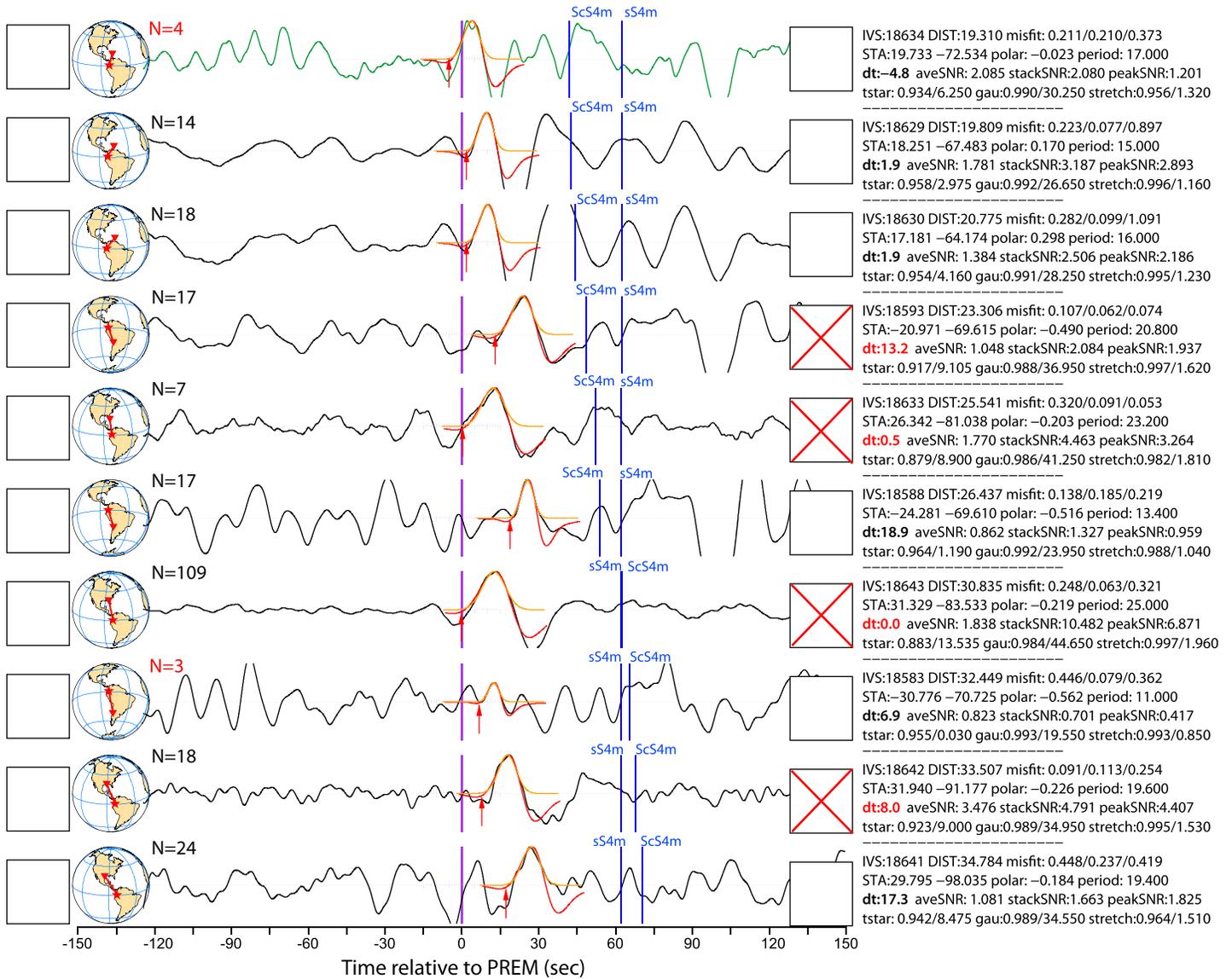


Figure 8. Partial page of an S4m virtual station PDF catalog plot for a 144 km deep event on 9 February 2013 in South America. Virtual station stacks are black traces, adapted empirical wavelets (SEWs) and Gaussian wavelets (red and orange traces, respectively) are shown with onset time predictions (red arrows) as determined by the SEW fitting algorithm. All traces are plotted relative to the PREM-predicted time (purple lines at time = 0, the predicted arrival time of S4m). Predictions for two expected arrivals (the depth phase sS4m and ScS4m) are notated with vertical blue lines and text labels. Other features, from left to right, include user text input box, reference globe (showing earthquake (red star), station (red triangle), the minor arc great circle path (red line)), the number of seismograms (N) used in the virtual station stack (which is red when $N < 5$), the retain/reject box (red X indicates record is retained), and a text block with information about the virtual station: ID#, epicentral distance (DIST), pre/signal/post misfit measurements, station (STA) latitude and longitude, predicted radiation pattern amplitude (polar), estimated period, travel time anomaly relative to PREM (dt, shown here in red for selected records), $\overline{\text{SNR}}$ (aveSNR), SNR_{VS} (stackSNR), $\text{SNR}_{\text{peak_trough}}$ (peakSNR), and information about the t^* operator, Gaussian factor, and stretching values.

Examples of empirical wavelets fitting multibounce waves are shown in Figure 9 for all phases studied here that were not measured in Lai19 (as well as SS, S3, and ScS2 that were studied in Lai19). To objectively present virtual station stack quality, 10 traces were randomly chosen from the population of each wave type, then five traces were chosen of the 10 for display in the figure. While the SNR is variable across the virtual station stacks, as is the number of contributing records to each stack, the resulting phase of interest is clear and the SEW matching is robust.

4. Results: Travel Times, Trends, and Wave Path Coverage

4.1. Virtual Station Travel Time Data Set

Here we present the travel time data set that we have constructed with virtual station seismograms. Measurements were made for 19 distinct seismic phases: S, Sdiff, SS, SSm, S3, S3m, S4, S4m, S5, S5m, S6m, ScS, ScS2, ScS3, ScS3m, ScS4, ScS4m, ScS5, and ScS5m. Roughly 250 K virtual station seismograms were constructed, then algorithmically retained or rejected, then subsequently reviewed by the authors to validate or update the retain/reject choices. After this process, 8,871 virtual station seismograms were retained (~3.6% retention rate). This number is significantly smaller than the average acceptance rate for the six basic phases of Lai19, which is around 20%. However, the Lai19 retention rate was dominated by direct S waves (slightly less than one half of their data set of ~252 K retained measurements), in which 56.8% of viewed S waves were retained. The number of retained virtual stack seismograms are listed in Table 1. As mentioned earlier, longer wave paths (e.g., higher multiple major arc phases) are lower amplitude from geometric spreading and attenuation, thus typically have lower success rates. In the case of SSm, it has fewer possibilities for analysis due to its distance range limit, which is roughly a 20° window from 160° to 180° minor arc distance (thus, 180° to 200° major arc distance); larger distances for this phase result in proximity to the diffracted wave SdiffSdiff, which was not pursued here, primarily because the ambiguity in the location of diffraction. In general, larger events have larger acceptance rates (Figure 10), though variability is apparent (and expected) due to differences in radiation patterns to favorable path geometries with abundant stations. Figure 10 presents the range of acceptance rate values using quantiles (one-fourth population ranges) to clearly show the range of the main half of the populations.

The number of records used in each virtual station stack is variable, from the minimum number of acceptable records (three records) up to a maximum of 331 stations. Figure 11 presents a histogram of the frequency distribution of the number of records used in making virtual station stacks. For locations possessing dense networks (e.g., EarthScope's USArray, see <http://earthscope.org>), the number of records used is large. For example, 626 of the virtual station stacks using EarthScope data have over 100 contributing records. Roughly 43% of the virtual station stacks have between three and 10 contributing records, and ~62% of the virtual station stacks have between three and 20 contributing records. As mentioned in discussion of Figure 7, the SNR of virtual station stacks is larger than that of the contributing records for the vast majority of our data (Figure 12). Data points in Figure 12 were computed for virtual stacks made with three or more contributing records. While significant scatter is present in the plot, the SNR of virtual stacks are to first order three times larger than that of the average SNR of all stations contributing to the virtual station stack. The points below the 1:1 line in Figure 12 may represent cases where the noise sums constructively in virtual station stacks, lowering their SNRs, or a high SNR in a single record that does not constructively stack in the virtual station (e.g., due to a negative polarity noise pulse on a poor record), thus resulting in a higher SNR average of the contributing records. Some scatter is likely due to virtual station stack and single seismogram measurements made for data from a wide range of radiation pattern amplitudes for waves of interest, as well as shallow (noisier) versus deep (cleaner) events, and higher multiples typically having lower SNR than lower multiples.

4.2. Single-Seismogram Multibounce Wave Measurements and Virtual Stack Corrections

After visual inspection of summary data record sections for all events, single seismogram multibounce waves were apparent for several earthquakes. As a follow-up step, we therefore processed raw single seismogram data in the same fashion as we did with virtual station stacks: A best-fitting SEW was fit to all visible multibounce single seismogram phases (as in section 3.1), PDF files were made displaying the best-fitting SEW and Gaussian, then the human review process was conducted for retaining or rejecting data (as in section 3.4). This allowed documentation and measurement of the infrequent, but present, multibounce waves on single seismograms. Table 2 presents the retained single seismogram measurement counts.

The 3,298 retained single seismogram picks were from records used in virtual station stacks, and thus allow us to compare single-seismogram measurements to the stacks they contributed to. For virtual station stacks having five or more single seismogram measurements, the contributing record onset travel times were averaged and compared to the onset time measured for the virtual station stack in Figure 13a (for all measured phases). A least squares best-fit line is also shown (with an R^2 value of 0.86) and indicates that virtual station

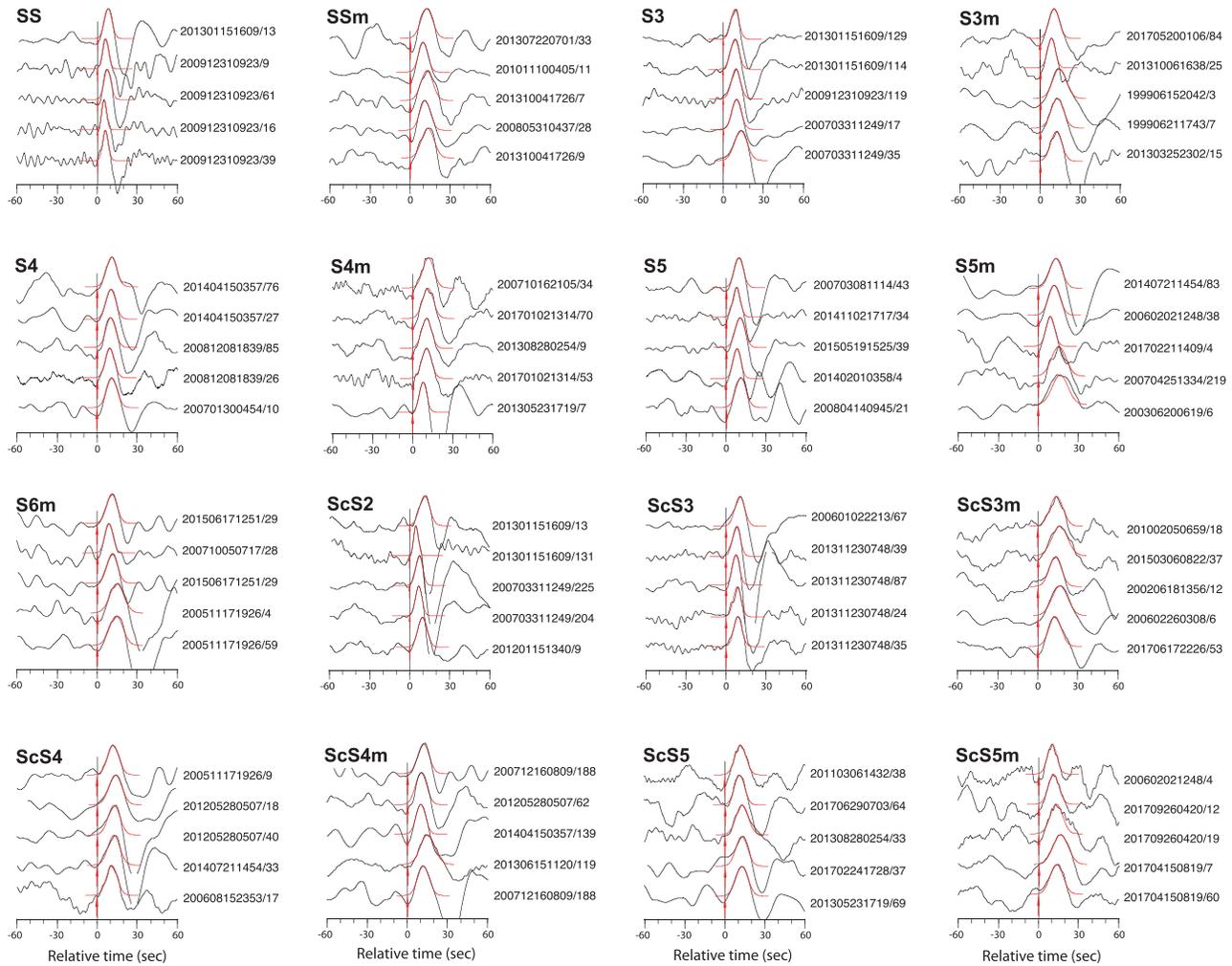


Figure 9. Five virtual station seismogram examples for the 16 of the multibounce phase types studied here (black traces) plotted with the best-fit stretched empirical wavelet (SEW, red traces) and the onset time determined from the Gaussian function that best-fits the SEW (red arrows). Event information and the number of seismograms used in the virtual station stack are to the right of each trace, as yyymmddhhmm/N (yyyy = year; mm = month; dd = day; hh = hour; mm = minute; N = # records). Records are aligned at zero time for ease of viewing, and all maximum amplitudes are normalized to unity.

stack onset time determinations are several seconds earlier than the average of the onset times of single seismograms measured for that same bin, but well correlated. This can be expected—the onset of the virtual station stack (and thus the onset time estimation) is influenced by energy from the earliest arriving phases in the stack (thus, earlier than the average onset time of the contributing records). In Figure 13b we compare the best-fit Gaussian of virtual station stacks to the average of the best-fit Gaussians of the single seismograms that contributed to each virtual station Gaussian; each was referenced to the Gaussian of the unstretched SEW for each earthquake (this latter step was taken to remove the effect of event size so data are comparable from different events). This comparison highlights that the Gaussians of virtual station stacks are broader than the average of the contributing seismogram Gaussians. This is also expected, for a given distribution of travel time perturbations in a virtual station bin, the temporally distributed contributing records will result in a broader stack. The trend lines in Figures 13a and 13b were weighted using a combination of the SNR of the contributing stations and the SNR of the virtual station stack. We did not find any dependency on specific seismic phases, thus the trendline in Figure 13b,

$$g_{S.S.-SEW}^{mean} = 0.333g_{VS-SEW} + 1.405, \quad (4)$$

while scatter is present, can be used to estimate a prediction of the average Gaussian factor for the contributing stations from a virtual station Gaussian. Therefore, an adjusted arrival time can be estimated from a

Table 1
Phase-by-Phase Listing of Number of Virtual Station Seismograms Retained for Measurements, Total Number of Viewed Virtual Station Seismograms, and the Ratio of These, i.e., the Percentage of Records That are Retained, i.e., the Retention Rate

Phase	Retained	Total viewed	Retention Rate (%)
S	561	2,045	27.43
Sdiff	438	5,514	7.94
SS	1,652	19,724	8.37
SSm	280	2,766	10.12
S3	1,767	15,011	11.77
S3m	719	20,434	3.51
S4	444	6,847	6.48
S4m	387	25,079	1.54
S5	142	1,984	7.15
S5m	83	16,412	0.50
S6m	41	16,527	0.24
ScS	464	6,414	7.23
ScS2	734	18,498	3.96
ScS3	405	15,980	2.53
ScS3m	216	11,979	1.80
ScS4	391	24,737	1.58
ScS4m	24	12,054	0.19
ScS5	69	12,371	0.55
ScS5m	54	14,281	0.37
TOTAL	8,871	248,657	3.57

scaled Gaussian (using its onset, as introduced in section 3.1), which will more appropriately depict the average onset time of records contributing to the virtual station stack. We thus develop a correction to shift virtual station stack onset time measurements to a prediction of the average onset time of the contributing seismograms. After correction, the adjusted virtual station stack times agree well with the average time anomalies of the contributing stations (Figure 13c). We apply this correction to all virtual station times. The correction (from equation (4)) will improve with more measurements, which might allow source depth or magnitude dependencies to be determined.

For all retained measurements, an ASCII file is made available with measurements and measured attributes made here, including the measured and corrected travel times relative to the PREM model. This file is freely shared via Zenodo (see Acknowledgments). Table 3 presents the tabulated information descriptions for the virtual station stack measurements file, and Table 4 presents the information descriptions for the single seismogram measurements file.

4.3. Wave Path Sampling Coverage

Virtual station construction results in a stack with an improved SNR over single seismograms, making possible measurements from multibounce phases that are otherwise difficult to investigate, resulting in new path geometries, especially with the major arc data. In Figure 14 we present wave path sampling as a function of latitude, by counting rays in $5^\circ \times 5^\circ$

cells. The coverage in the upper mantle (0–660 km, top row of panels) is compared to the deepest 300 km of the mantle (2,591–2,891 km, bottom row of panels). We first show the latitudinal sampling of the six phases of Lai19 (first column). There is a clear bias of greater sampling in the northern hemisphere, which is dominantly due to the large number of S waves in that data set, and the predominance of stations in the northern hemisphere. The coverage of all virtual station stacks measured here are shown in the second

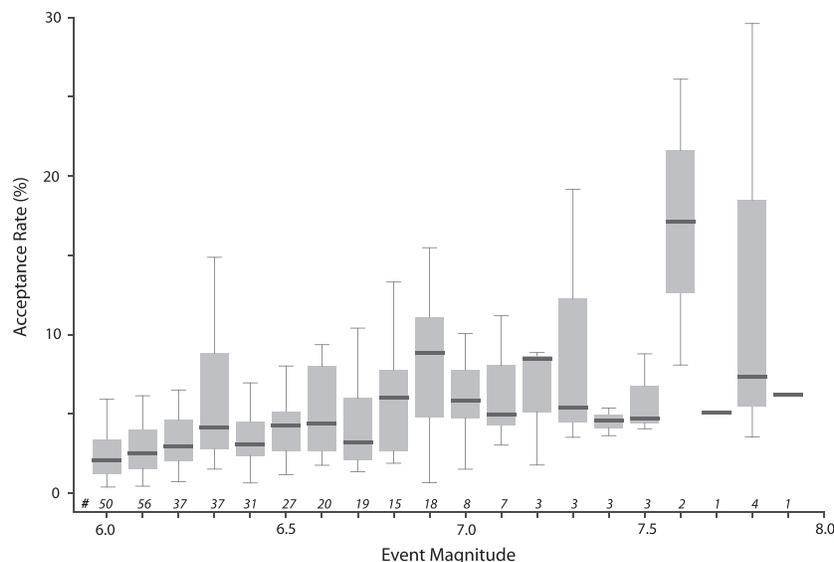


Figure 10. Virtual station measurement acceptance rate averaged in earthquake magnitude bins, for all measured phases of Table 1. Thick horizontal bar in the gray shaded box represents the median value for each magnitude bin population. The gray boxes show the range between the 25th and 75th quantile percentage ranges of the population (Q1 and Q3, respectively). The vertical bars outside of the gray boxes extend upward to the largest value in the population within the maximal ($Q3 + 1.5(Q3 - Q1)$) range and downward to the smallest value in the minimal ($Q1 - 1.5(Q3 - Q1)$) range. The italicized numbers above the magnitude axis correspond to the number of earthquakes used in the acceptance rate averages.

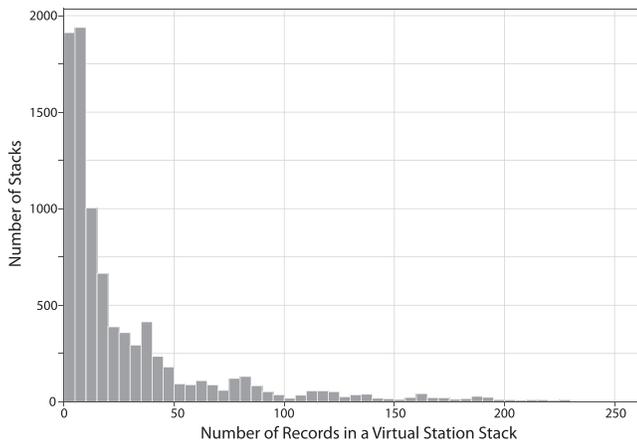


Figure 11. Frequency histogram showing the number of virtual station stacks having different numbers of contributing records. Over half of the population is in the first three bars (between three and 30 records). See text for more information.

column. The latitudinal sampling is significantly improved in the southern hemisphere, particularly in the upper mantle. The third column shows the coverage from virtual station stacks of multibounce S waves (Sn , $n = 2-6$, minor and major arc). These improve sampling in both hemispheres. The multibounce ScS wave coverage (ScSn, $n = 2-5$, minor and major arc) are presented in the fourth column, and only slightly improve southern hemisphere sampling. This is primarily due to the very low number of major arc ScSn measurements, and the fact that the higher multiple ScSn commonly sample the same corridors of ScS (i.e., a single reflection, $n = 1$). Latitudinal coverage for all minor arc virtual station stacks are shown in the fifth column and is similar (but slightly improved) to the ScSn coverage. However, if only major arc virtual station paths are considered (sixth and final column), we see the best relative southern hemisphere sampling. This highlights the potential benefit of major arc paths in whole mantle imaging.

Lateral ray path coverage maps for all virtual station paths for the upper 660 km and deepest 300 km of the mantle are presented in Figure 15a.

Both depth shells have reasonably good coverage (with the upper mantle being better sampled). Figure 15b shows a sampling coverage density map, in $5^\circ \times 5^\circ$ cells for the same depth shells. While it is clear that the upper mantle has slightly greater sampling, as suggested in Figure 14, the

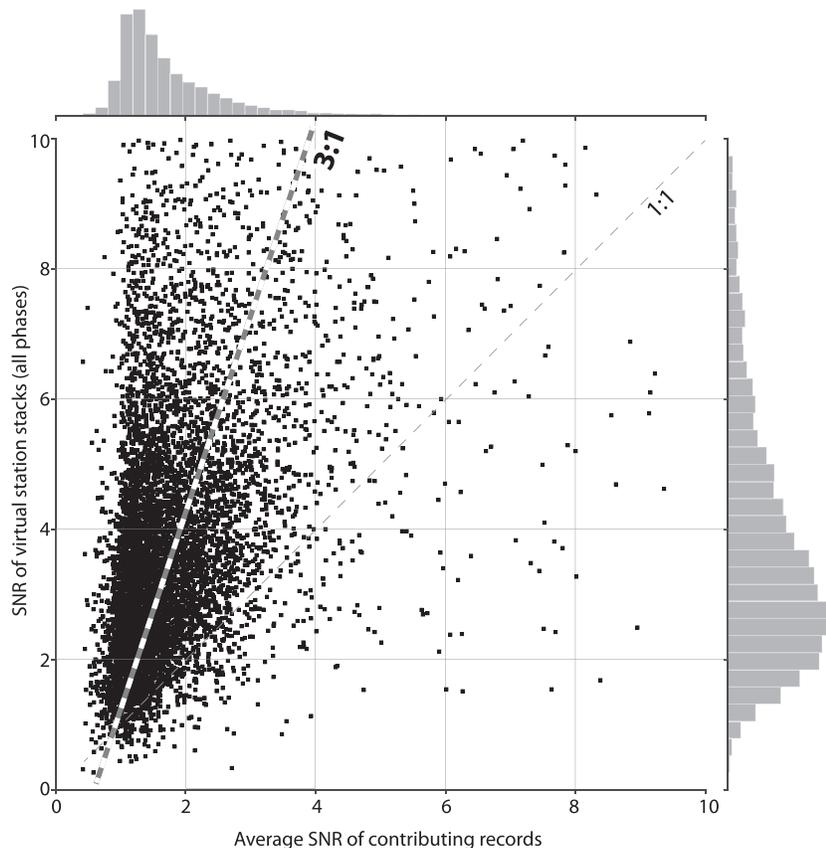


Figure 12. Comparison of virtual station stack SNRs (vertical axis) to the average SNR of the contributing records for each stack (horizontal axis). Included data virtual station stacks have a minimum of three contributing records. A line with 3-to-1 slope compares well with the data points. The histogram at the top of the figure corresponds to the number of SNR averages taken as a function of their average SNR; the histogram to the right corresponds to the number of virtual stacks as a function of SNR measurement.

Table 2
Phase-by-Phase Listing of Number of Single Seismogram Measurements Retained, Viewed, and the Retention Rate of Each Phase

Phase	Retained	Total viewed	Retention Rate (%)
S3m	874	36,298	2.42
S4m	223	35,165	0.64
S5m	20	11,688	0.19
S6m	4	2,252	0.18
ScS3	1,668	39,364	4.26
ScS4	451	36,787	1.24
ScS4m	17	4,721	0.36
ScS5	41	21,728	0.22
TOTAL	3,298	188,003	1.75

southern hemisphere is fairly well sampled. To illustrate the sampling density improvement when using higher multiple bounce waves (especially major arc data) investigated here, compared to phases used in Lai19 (S, SS, S3, Sdiff, ScS, ScS2), we present a representation of sampling density improvement in Figure 15c. Here improvement is defined as percentage increase of sampling in the cell by the addition of the virtual station data (as the number of virtual station paths divided by the number of paths in the Lai19 data set, times 100). The warmer colors mark the most improved regions, with orange and red marking more than a 100% improvement or more. The southern hemisphere coverage improves significantly from the virtual station stack data (especially the major arc data).

5. Discussion

In this paper we present a virtual station stacking method to exploit the use of multibounce data that is otherwise typically too low in amplitude to be confidently measured on single seismograms (i.e., the multibounce phase-of-interest is not out of the noise level enough to warrant measurement). The virtual station stacking increases the SNR, and combined with the empirical wavelet approach, provides an objective method to measure travel time and waveform information. Some systematic broadening was apparent in the virtual station stacks and corrected for. This method was deployed on all measurable phases in our data set. Single seismograms were also measured where possible, and all measurements were visually inspected. Figure 13 highlights the fact that virtual station stacks are broadened compared to the constituent seismograms (where measurable). A correction for this effect allows a more confident measure of virtual stack travel time. An empirical comprehensive weight (equation (3)) permits a relative weight for travel times presented here.

However, even with an improved SNR over constituent contributing seismograms, virtual stations average typically noisy seismograms within a geographical bin, and uncertainties exist. For example, we chose to stack stations within a geographic bin having a 5° radius from the grid center, with a spatial Gaussian weighting factor (G_R) relative to the grid center:

$$G_R = e^{-\frac{R^2}{2\sigma_R^2}} \tag{5}$$

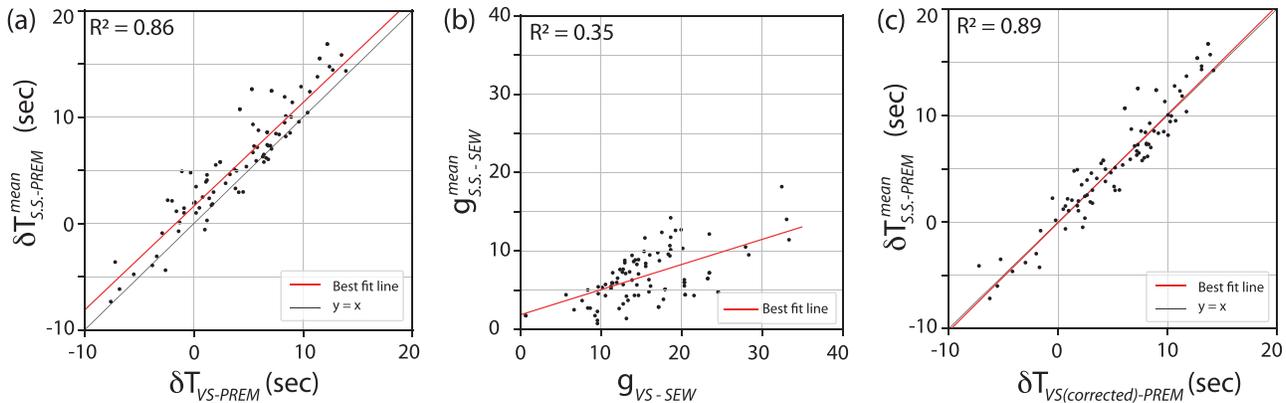


Figure 13. (a) For virtual station stacks containing single seismogram measurements, the average travel time anomaly (with respect to PREM), $\delta T_{S.S.-PREM}^{mean}$, is compared to the travel time anomaly of the virtual station stacks they contributed to, $\delta T_{VS-PREM}$ (for all phases). A best-fit line and R^2 value are also shown. (b) The average of the best-fit Gaussians of the single seismograms from (a), $g_{S.S.-SEW}^{mean}$, compared to the best-fit Gaussian of the virtual station stack, g_{VS-SEW} , where both were referenced to the Gaussian of the SEW for their corresponding events. A best-fit trend line and R^2 value are also shown. (c) After the Gaussian of the virtual station stack is corrected using the trend line in (b), an onset time is determined, and compared to the average of the onset times of the contributing single seismogram records. The updated trend line has an R value of 0.89, and agrees well with the $y = x$ line.

Table 3

List of Virtual Station Stack Measurements and Attributes Computed in This Study That are Shared in the Archived Data File Number 1 (see Acknowledgments)

#	Information	Description
1	VS number	An integer code to represent VS station name
2	Distance	Epicentral distance between event and VS grid center in degrees
3	Stack number	Number of records that is used to stack for virtual stations
4	VS latitude	VS location latitude in degrees
5	VS longitude	VS location longitude in degrees
6	Event latitude	Earthquake hypocentral location latitude in degrees ^a
7	Event longitude	Earthquake hypocentral location longitude in degrees ^a
8	Event depth	Earthquake hypocentral location depth in km ^a
9	Event magnitude	Earthquake moment magnitude ^a
10	Origin time	Earthquake origin time ^a
11	Azimuth	Azimuth from earthquake to VS grid (in degrees)
12	Back azimuth	Back azimuth measured at VS clockwise back to earthquake (in degrees)
13	Phase name	Either S, Sdiff, SS, SSm, S3, S3m, S4, S4m, S5, S5m, S6m, ScS, ScS2, ScS3, ScS3m, ScS4, ScS4m, ScS5, or ScS5m
14	Predicted time	Travel time prediction of the PREM model
15	Measured time	Travel time anomaly of phase onset relative to PREM (observed minus PREM)
16	Corrected time	Travel time anomaly of phase onset relative to PREM (observed minus PREM) corrected for over-broadened virtual station stack
17	Phase start	The start time, relative to PREM prediction, of the beginning of the time window used to define one pulse width of VS phase of interest on velocity recordings, measured at the 10% amplitude level preceding the wave peak (used to auto-define the Misfit measurement windows)
18	Phase end	The end time, relative to the PREM prediction, of the end of the time window used to define one pulse width of VS phase of interest on velocity recordings, measured at the 10% amplitude level following the wave peak (used to auto-define the Misfit measurement windows)
19	SNR _{VS}	VS SNR measurement from the average amplitude of the signal to the average amplitude of the noise
20	$\overline{\text{SNR}}$	Average SNR of all records used to create a VS stack
21	SNR _{peak_trough}	VS SNR measurement from the maximum peak-to-trough value within one period of the signal compared to noise
22	Misfit _{SIGNAL}	The average difference between the VS phase and the best-fit SEW over one period (as in equation (2))
23	Misfit _{PRE}	As above, except over one period preceding the phase of interest
24	Misfit _{POST}	As above, except over one period following the phase of interest
25	t*	The best-fit t* value that, when convolved with the SEW, gives the best-fit to records that are broader than the SEW
26	Stretch factor	A measure of the amount the SEW has to be narrowed to fit records that are narrower than the SEW
27	CCC _[rec,SEW]	Cross-correlation coefficient between observed record and the best-fitting SEW adapted to the record
28	CCC _[rec,GEW]	Cross-correlation coefficient between observed record and the GEW, which measures the record's fit to the average S wave phase shape for the event
29	$g_{\text{best-fit_SEW}}$	Gaussian factor of the best-fitting Gaussian function (g, equation (1)) to a record's best-fitting SEW
30	$g_{\text{event_SEW}}$	Gaussian factor of the best-fitting Gaussian function (g, equation (1)) to the GEW for the event
31	Misfit _g	The misfit measured between $g_{\text{best-fit_SEW}}$ and $g_{\text{event_SEW}}$ (computed as in equation (2)) which provides a different measure of record broadening
32	W _{comprehensive}	An empirical comprehensive weight value for each data (see equation (3))
33	Noise window traffic flag	Records that have "traffic" (interfering seismic waves) predicted to arrive in the noise window (of the SNR measurement) are flagged as 1, otherwise 0 ^b
34	Period	Estimated period of the phase, from the start and end of the pulse measured at 0.1 amplitude (when peak is set to 1) measured on displacement recordings
35	Polarity	The predicted amplitude between [-1,1] using the SH radiation pattern for the phase of interest

Note. The number in the first column of the table below corresponds to the column number in archived ASCII file. VS = Virtual Station.

^aAs provided by ISC (International Seismological Center). ^bSome records do not show energy corresponding to predicted times of interfering phases. Thus, in some cases, an entry of 1 may correspond to a retained measurement.

Here R is radius in degrees. We used a Gaussian factor of $g_R = 4$, which gives $G_R(R = 5) = 0.5$, thus the weighting tapers from 1 at the grid center ($R = 0$) to 0.5 at $R = 5$. The large stacking radius was chosen to include more stations in order to permit more virtual station stacks, as well as an attempt to improve the SNR. We present the effect of different g_R factors on the radius weighting in Figure 16a. Four different g_R values are shown. The effect of different g_R is shown for a multibounce ScS wave example (virtual station stack of ScS3m, made from 24 contributing records) in Figure 16b. The virtual station stacks for the same four g_R values are presented and show the coherent stacking of ScS3m for $g_R \geq 2$. Larger g_R results in improved SNR, and while $g_R = 6$ has slightly better SNR than $g_R = 4$, we choose the latter to minimize the blurring effect of the greater weight for the more distant stations contributing to the stack. This is

Table 4

List of Single Seismogram Measurements and Attributes Computed in This Study That are Shared in the Archived Data File Number 2 (see Acknowledgments)

#	Information	Description
1	Station name	The 3–5 character station name code
2	Network name	The two-digit code for the seismographic network
3	Distance	Epicentral distance between earthquake and station in degrees
4	Station latitude	Station location latitude in degrees ^a
5	Station longitude	Station location longitude in degrees ^a
6	Event latitude	Earthquake hypocentral location latitude in degrees ^b
7	Event longitude	Earthquake hypocentral location longitude in degrees ^b
8	Event depth	Earthquake hypocentral location depth in kmb
9	Event magnitude	Earthquake moment magnitude ^b
10	Origin time	Earthquake origin time ^b
11	Azimuth	Azimuth from earthquake to station (in degrees)
12	Back azimuth	Back azimuth measured at station clockwise back to earthquake (in degrees)
13	Phase name	Either S3m, S4m, S5m, S6m, ScS3, ScS4, ScS4m, or ScS5
14	Predicted time	Travel time prediction of the PREM model
15	Measured time	Travel time anomaly of phase onset relative to PREM (observed minus PREM)
17	Phase start	The start time, relative to PREM prediction, of the beginning of the time window used to define one pulse width of phase of interest on velocity recordings, measured at the 10% amplitude level preceding the wave peak (used to auto-define the Misfit measurement windows)
18	Phase end	The end time, relative to the PREM prediction, of the end of the time window used to define one pulse width of phase of interest on velocity recordings, measured at the 10% amplitude level following the wave peak (used to auto-define the Misfit measurement windows)
19	SNR _{average_amp}	The signal-to-noise measurement from the average amplitude of the signal to the average amplitude of the noise
20	SNR _{peak-trough}	The signal-to-noise measurement from the maximum peak-to-trough measurement within one period of the signal compared to noise
21	SNR _{max_peak}	The signal-to-noise measurement from the maximum peak in the signal window compared to the maximum peak in the entire noise window
22	Misfit _{SIGNAL}	The average difference between the phase and the best-fit SEW over one period
23	Misfit _{PRE}	As above, except over one period preceding the phase
24	Misfit _{POST}	As above, except over one period following the phase
25	t*	The best-fit t* value that, when convolved with the SEW, gives the best-fit to records that are broader than the SEW
26	Stretch factor	A measure of the amount the SEW has to be narrowed to fit records that are narrower than the SEW
27	CCC _[rec,SEW]	Cross-correlation coefficient between observed record and the best-fitting SEW adapted to the record
28	CCC _[rec,GEW]	Cross-correlation coefficient between observed record and the GEW, which measures the record's fit to the average S wave phase shape for the event
29	g _{best-fit_SEW}	Gaussian factor of the best-fitting Gaussian function (g, Equation (1)) to a record's best-fitting SEW
30	g _{event_SEW}	Gaussian factor of the best-fitting Gaussian function (g, Equation (1)) to the GEW for the event
31	Misfit _g	The misfit measured between g _{best-fit_SEW} and g _{event_SEW} (computed as in equation (2)) which provides a different measure of record broadening
32	W _{comprehensive}	An empirical comprehensive weight value for each data (see equation (3))
33	Noise window traffic flag	Records that have “traffic” (interfering seismic waves) predicted to arrive in the noise window (of the SNR measurement) are flagged as 1, otherwise 0
34	Period	Estimated period of the phase, from the start and end of the pulse measured at 0.1 amplitude (when peak is set to 1) measured on displacement recordings
35	Polarity	The predicted amplitude between [−1,1] using the SH radiation pattern for the phase of interest

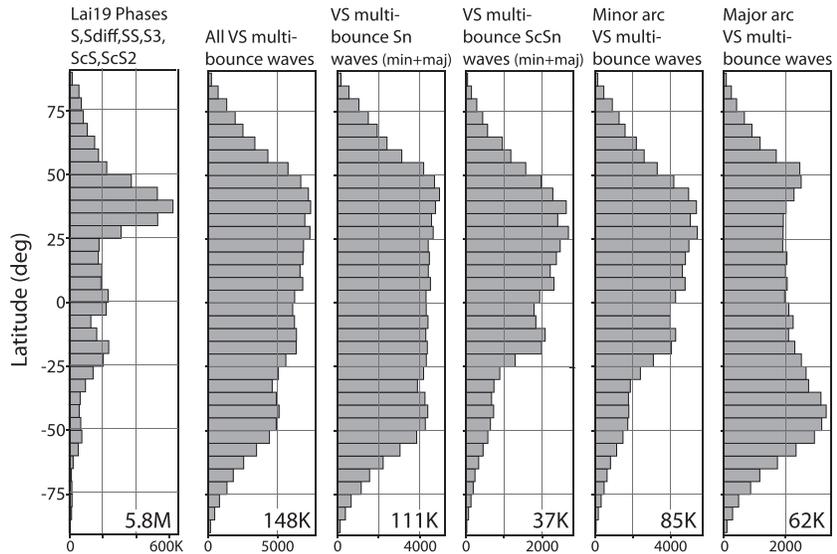
Note. The number in the first column of the table below corresponds to the column number in archived ASCII file.

^aAs provided by the data agencies listed in section 2.1. ^bAs provided by ISC (International Seismological Center).

apparent for $g_R = 6$ from the slightly larger standard deviation. We note the onset time of the different virtual station stacks do not change for the different g_R .

While the onset assignment to data using the empirical wavelet method is objective, and uncertainties in onset time estimations are likely within ± 1 s for the wavelet at hand, the broadened virtual station stacks relative to individual stations present additional uncertainties (as detailed in Figure 13). While we correct for the wave broadening to shift the onset time estimations, there is some scatter in the comparison of the averaged onset times of constituent records compared to the corrected virtual station stack onset times (Figure 13c). Several seconds of scatter is apparent, and sometimes larger. We have omitted virtual station stacks if their onset times differ from the average of onset times of contributing station averages by more than 8 s (the one-half period of the upper corner of the bandpass filter used on the data, and the average

(a) 5x5 deg cell path coverage: 0-660 km depth shell



(b) 5x5 deg cell path coverage: 2591-2891 km depth shell

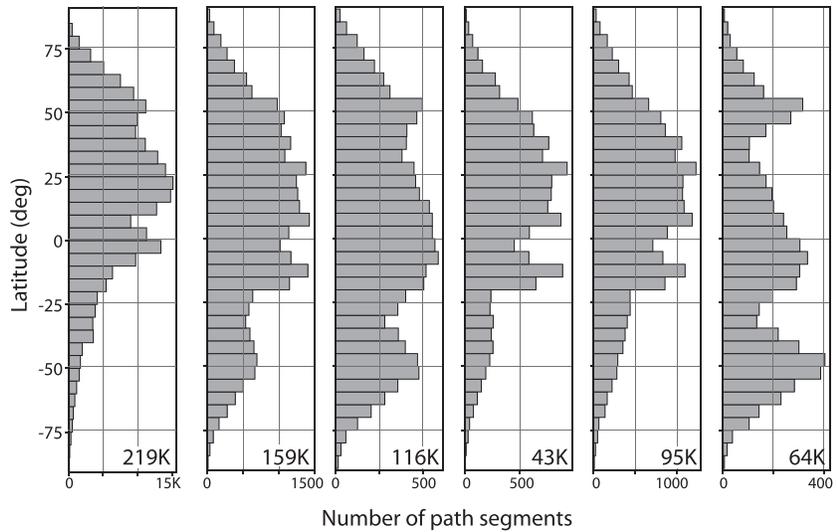


Figure 14. Ray path coverage as a function of latitude, where the number of rays were counted in all $5^\circ \times 5^\circ$ cells, for a variety of different data groupings, for (a) the entire upper mantle (between 0 and 660 km), and (b) the lowermost 300 km of the mantle. The total number of path segments is written in the lower right of each panel. The single seismogram data set of Lai19 (for S, Sdiff, SS, S3, ScS, and ScS2) is shown in the leftmost column. The second column presents coverage of all virtual station stacks measured in this paper. The third column shows the coverage of all major and minor arc multibounce S_n waves ($n = 2-6$). Column 4 is the same as column 3, except for ScS_n ($n = 2-5$). Column 5 presents coverage for all minor arc phases and column 6 presents coverage for all major arc phases. See text for more details.

period of most of the data) in the line fitting in Figure 13; this should preclude phase misidentification. We emphasize the importance of the comprehensive weight we give to all measurements, which depends upon factors that characterize SNR and goodness of fit of the SEW to the phases of interest.

We further note that our coverage discussions (e.g., Figures 14 and 15) were based upon infinite frequency ray paths. The average period of minor and major arc multibounce wave virtual station stacks measured here are 17.6 and 22.5 s, respectively. These can be considered small to intermediate period. While ray computations for coverage may be a reasonable approximation for this period, the sensitivity of these waves, especially those with long paths (e.g., all major arc phases), spans a volume which should be taken into

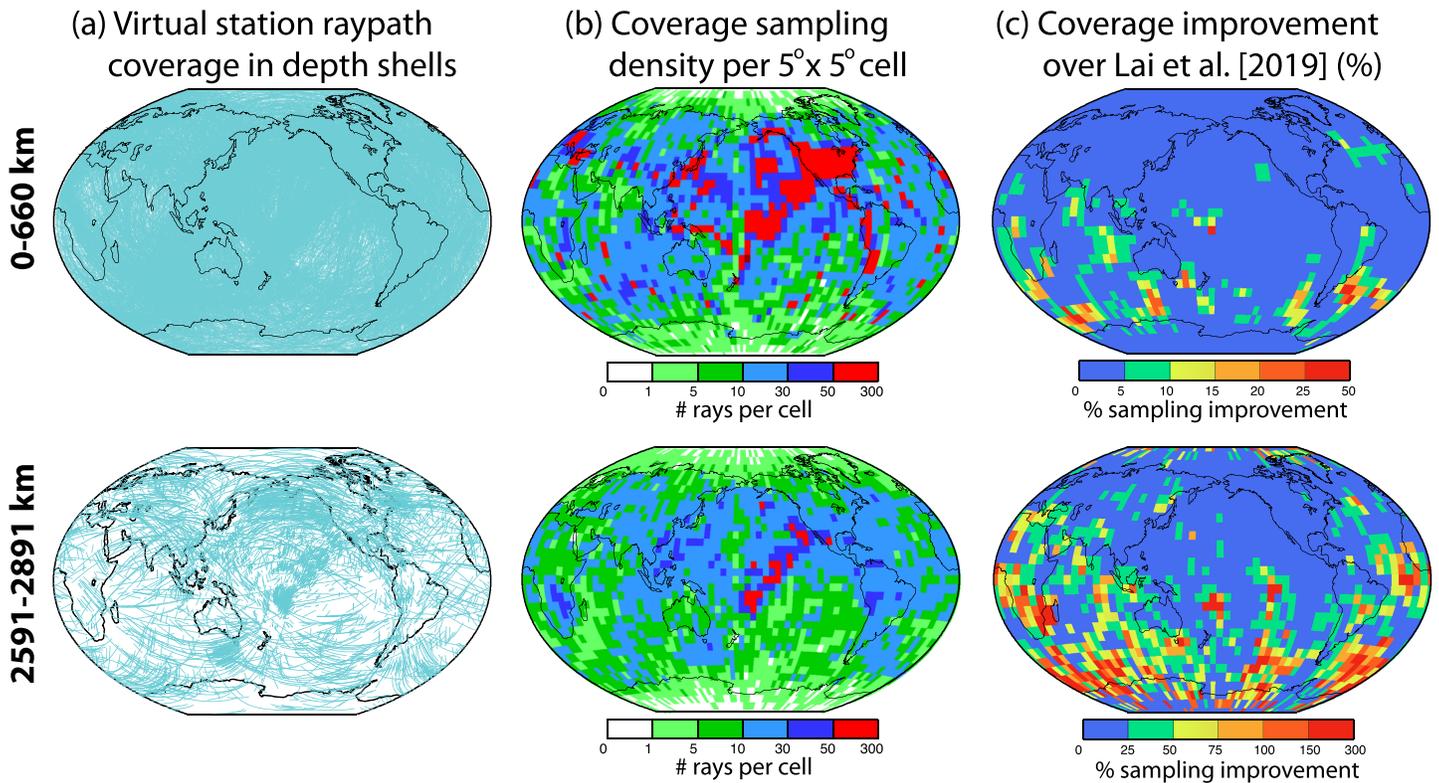


Figure 15. (a) Ray paths (blue lines) of all virtual station stacks measured in this study, for the upper mantle (top row) and deepest 300 km of the mantle (bottom row). (b) Coverage sampling density in $5^\circ \times 5^\circ$ cells, in the upper and lowermost mantle, for the ray paths in (a). Scale bars present the number of rays counted in each cell. (c) Sampling density coverage improvement over Lai19 (in %) by taking the panels in (b) and dividing by the same for the ray paths of Lai19. Thus, 100 represents the same number of rays in each data set. The warmer colors (strongest sampling improvement) are mostly present in the southern hemisphere, especially for the deepest mantle.

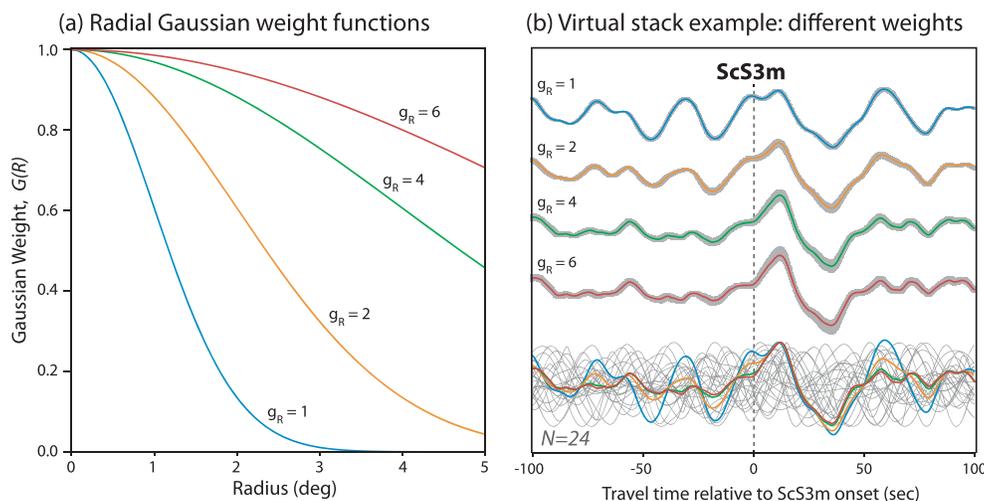


Figure 16. (a) Relationship between distance from grid center and Gaussian weight G_R for different Gaussian factors g_R . This study used $g_R = 4$. (b) Virtual station stack example for the major arc multibounce wave ScS3m. Virtual station stacks are shown for the g_R of panel (a), along with ± 1 standard deviation (with amplitudes normalized to the maximum energy in the time window). The bottom of the panel displays the 24 records used to make the stack, along with the four virtual station stacks overlaid on them (the virtual station stacks are normalized to the amplitude to the of the estimated ScS3m peak). Event information: 18 January 2011, Latitude = 28.7° , Longitude = 64.0° , Z = 80 km, Mw = 7.2. Virtual Station information: Latitude -14.1° , Longitude = -70.8° .

consideration for imaging purposes. Other corrections were not applied here and should be similarly taken into consideration, like crustal corrections (e.g., Artemieva & Mooney, 2001; Laske et al., 2013) and ellipticity corrections (e.g., Kennett & Gudmundsson, 1996).

This study focused on SH waves, so results can be combined with results of Lai19. The P-SV system is more complicated due to P-SV coupling and conversions, and is thus left for future work. However, the empirical wavelet approach of Lai19 should work with single seismogram measurements of P-SV phases, as should this virtual station stacking approach presented here.

6. Conclusions

We presented a virtual station stacking algorithm that computes a seismogram stack for geographical bins to improve SNRs of typically low amplitude seismic waves. The main focus of this study was multibounce waves not measured in Lai19, namely S4, S5 (and major arc counterparts S2m, S3m, S4m, S5m, and S6m), ScS3, ScS4, and ScS5 (and major arc counterparts ScS3m, ScS4m, ScS5m). We also constructed virtual station stacks for phases of S, SS, S3, ScS, ScS2, and Sdiff, for poorly or unsampled corridors of Lai19. For a data set of 360 global earthquakes and 8,407 seismographic stations, we defined 289 virtual station grid cells, which were used to compute 248,657 virtual station stacks of 19 different seismic phases. After visual inspection, 8,871 of these were retained. Onset travel time and waveform information was retained. We also searched for single seismogram multibounce waves which were present for some stronger earthquakes and identified and measured 3,331 phases. These measurements were compared to the virtual station stacks and used to develop a wave shape width correction to the virtual station stacks, which are broadened by summing data containing time shifts. This resulted in an onset time correction that was applied to all virtual station stack measurements. The resulting wave path coverage from the multibounce data significantly helps to bolster sampling in the southern hemisphere. All measurements and data attributes are made publicly available.

Acknowledgments

This research was financially supported by funds from NSF EAR-1648817. The detailed comments of two anonymous reviewers greatly helped to improve the manuscript. We thank several seismic data centers that provided seismic data for this study, including IRIS DMC (Incorporated Research Institutions for Seismology, Data Management Center), ORFEUS (Observatories & Research Facilities for European Seismology), NECD (Northern California Earthquake Data Center), F-net (F-net Broadband Seismograph Network), CNSN (Canadian National Seismic Network). We also thank EarthScope program for the freely available USArray data, which provided large numbers of high-quality seismic data. Moreover, we thank the principle investigators, individuals, and organizations that deployed the seismic networks used in this study. A total number of 308 network codes was used in this study, which correspond to nearly 1,600 unique network deployments from various organizations and groups. We have compiled a list of detailed information of each individual network used in this study, including the network name, network operator, network country, network website if available, network deployment country, and digital object identifier if available. Please refer to the supplementary information for details. Our final data set is available online (<https://doi.org/10.5281/zenodo.3247093>).

References

- Artemieva, I. M., & Mooney, W. D. (2001). Thermal thickness and evolution of Precambrian lithosphere: A global study. *Journal of Geophysical Research*, *106*(B8), 16,387–16,414. <https://doi.org/10.1029/2000JB900439>
- Auer, L., Boschi, L., Becker, T. W., Meyer, T. N., & Giardini, D. (2014). Savani: A variable resolution whole-mantle model of anisotropic shear velocity variations based on multiple data sets. *Journal of Geophysical Research: Solid Earth*, *119*, 3006–3034. <https://doi.org/10.1002/2013JB010773>
- Bréger, L., & Romanowicz, B. (1998). Three-dimensional structure at the base of the mantle beneath the Central Pacific. *Science*, *282*(5389), 718–720. <https://doi.org/10.1126/science.282.5389.718>
- Becker, T. W., & Boschi, L. (2002). A comparison of tomographic and geodynamic mantle models. *Geochemistry, Geophysics, Geosystems*, *3*(1), 168. <https://doi.org/10.1029/2001gc000168>
- Cormier, V. F. (1989). Slab diffraction of S waves. *Journal of Geophysical Research*, *94*(B3), 3006–3024. <https://doi.org/10.1029/JB094iB03p03006>
- Cottaar, S., & Lekic, V. (2016). Morphology of seismically slow lower-mantle structures. *Geophysical Journal International*, *207*(2), 1122–1136. <https://doi.org/10.1093/gji/ggw324>
- Durand, S., Debayle, E., Ricard, Y., Zanolli, C., & Lambotte, S. (2017). Confirmation of a change in the global shear velocity pattern at around 1000 km depth. *Geophysical Journal International*, *211*(3), 1628–1639. <https://doi.org/10.1093/gji/ggx405>
- Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, *25*(4), 297–356. [https://doi.org/10.1016/0031-9201\(81\)90046-7](https://doi.org/10.1016/0031-9201(81)90046-7)
- Dziewonski, A. M., Chou, T. A., & Woodhouse, J. H. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research*, *86*(B4), 2825–2852. <https://doi.org/10.1029/JB086iB04p02825>
- Ekström, G., Nettles, M., & Dziewonski, A. M. (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, *200–201*, 1–9. <https://doi.org/10.1016/j.pepi.2012.04.002>
- Flanagan, M. P., & Shearer, P. M. (1998). Global mapping of topography on transition zone velocity discontinuities by stacking SS precursors. *Journal of Geophysical Research*, *103*(B2), 2673–2692. <https://doi.org/10.1029/97JB03212>
- French, S. W., & Romanowicz, B. (2015). Broad plumes rooted at the base of the Earth's mantle beneath major hotspots. *Nature*, *525*(7567), 95–99. <https://doi.org/10.1038/nature14876>
- Frost, D. A., & Rost, S. (2014). The P-wave boundary of the Large-Low Shear Velocity Province beneath the Pacific. *Earth and Planetary Science Letters*, *403*, 380–392. <https://doi.org/10.1016/j.epsl.2014.06.046>
- Frost, D. A., Rost, S., Selby, N. D., & Stuart, G. W. (2013). Detection of a tall ridge at the core–mantle boundary from scattered PKP energy. *Geophysical Journal International*, *195*(1), 558–574. <https://doi.org/10.1093/gji/ggt242>
- Futterman, W. I. (1962). Dispersive body waves. *Journal of Geophysical Research*, *67*(13), 5279–5291. <https://doi.org/10.1029/JZ067i013p05279>
- Garnero, E. J., McNamara, A. K., & Shim, S.-H. (2016). Continent-sized anomalous zones with low seismic velocity at the base of Earth's mantle. *Nature Geoscience*, *9*(7), 481–489. <https://doi.org/10.1038/ngeo2733>
- Grand, S. P. (2002). Mantle shear–wave tomography and the fate of subducted slabs. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *360*(1800), 2475–2491. <https://doi.org/10.1098/rsta.2002.1077>
- Grand, S. P., & Helmberger, D. V. (1984). Upper mantle shear structure of North America. *Geophysical Journal International*, *76*(2), 399–438. <https://doi.org/10.1111/j.1365-246X.1984.tb05053.x>

- Houser, C., Masters, G., Shearer, P., & Laske, G. (2008). Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms. *Geophysical Journal International*, *174*(1), 195–212. <https://doi.org/10.1111/j.1365-246X.2008.03763.x>
- Kennett, B. L. N., & Gudmundsson, O. (1996). Ellipticity corrections for seismic phases. *Geophysical Journal International*, *127*(1), 40–48. <https://doi.org/10.1111/j.1365-246X.1996.tb01533.x>
- Koelemeijer, P., Ritsema, J., Deuss, A., & Van Heijst, H. J. (2016). SP12RTS: A degree-12 model of shear- and compressional-wave velocity for Earth's mantle. *Geophysical Journal International*, *204*(2), 1024–1039. <https://doi.org/10.1093/gji/ggv481>
- Kustowski, B., Ekström, G., & Dziewonski, A. M. (2008). Anisotropic shear-wave velocity structure of the Earth's mantle: A global model. *Journal of Geophysical Research*, *113*, B06306. <https://doi.org/10.1029/2007JB005169>
- Lai, H., Garnero, E. J., Grand, S. P., Porritt, R. W., & Becker, T. W. (2019). Global travel time data set from adaptive empirical wavelet construction. *Geochemistry, Geophysics, Geosystems*, *20*(5), 2175–2198. <https://doi.org/10.1029/2018GC007905>
- Laske, G., Masters, G., Ma, Z., & Pasyanos, M. (2013). Update on CRUST1.0—A 1-degree global model of Earth's Crust. *Geophysical Research Abstracts*, *15*, EGU2013-2658
- Lekic, V., Cottaar, S., Dziewonski, A., & Romanowicz, B. (2012). Cluster analysis of global lower mantle tomography: A new class of structure and implications for chemical heterogeneity. *Earth and Planetary Science Letters*, *357*–*358*, 68–77. <https://doi.org/10.1016/j.epsl.2012.09.014>
- Li, C., van der Hilst, R. D., Engdahl, E. R., & Burdick, S. (2008). A new global model for P wave speed variations in Earth's mantle. *Geochemistry, Geophysics, Geosystems*, *9*, Q05018. <https://doi.org/10.1029/2007GC001806>
- Moulik, P., & Ekström, G. (2014). An anisotropic shear velocity model of the Earth's mantle using normal modes, body waves, surface waves and long-period waveforms. *Geophysical Journal International*, *199*(3), 1713–1738. <https://doi.org/10.1093/gji/ggu356>
- Ni, S., Cormier, V. F., & HelMBERGER, D. V. (2003). A comparison of synthetic seismograms for 2D structures: Semianalytical versus numerical. *Bulletin of the Seismological Society of America*, *93*(6), 2752–2757. <https://doi.org/10.1785/0120030011>
- Ni, S., & HelMBERGER, D. V. (2003a). Further constraints on the African superplume structure. *Physics of the Earth and Planetary Interiors*, *140*(1–3), 243–251. <https://doi.org/10.1016/j.pepi.2003.07.011>
- Ni, S., & HelMBERGER, D. V. (2003b). Ridge-like lower mantle structure beneath South Africa. *Journal of Geophysical Research*, *108*(B2), 2094. <https://doi.org/10.1029/2001JB001545>
- Ni, S., HelMBERGER, D. V., & Tromp, J. (2005). Three-dimensional structure of the African superplume from waveform modelling. *Geophysical Journal International*, *161*(2), 283–294. <https://doi.org/10.1111/j.1365-246X.2005.02508.x>
- Ni, S., Tan, E., Gurnis, M., & HelMBERGER, D. (2002). Sharp sides to the African superplume. *Science*, *296*(5574), 1850–1852. <https://doi.org/10.1126/science.1070698>
- Panning, M., & Romanowicz, B. (2006). A three-dimensional radially anisotropic model of shear velocity in the whole mantle. *Geophysical Journal International*, *167*(1), 361–379. <https://doi.org/10.1111/j.1365-246X.2006.03100.x>
- Ritsema, J., Deuss, A., Van Heijst, H. J., & Woodhouse, J. H. (2011). S40RTS: A degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltimes and normal-mode splitting function measurements. *Geophysical Journal International*, *184*(3), 1223–1236. <https://doi.org/10.1111/j.1365-246X.2010.04884.x>
- Ritsema, J., Ni, S., HelMBERGER, D. V., & Crotwell, H. P. (1998). Evidence for strong shear velocity reductions and velocity gradients in the lower mantle beneath Africa. *Geophysical Research Letters*, *25*(23), 4245–4248. <https://doi.org/10.1029/1998GL900127>
- Rost, S., & Earle, P. S. (2010). Identifying regions of strong scattering at the core–mantle boundary from analysis of PKKP precursor energy. *Earth and Planetary Science Letters*, *297*(3–4), 616–626. <https://doi.org/10.1016/j.epsl.2010.07.014>
- Rost, S., Garnero, E. J., & Williams, Q. (2008). Seismic array detection of subducted oceanic crust in the lower mantle. *Journal of Geophysical Research*, *113*, B06303. <https://doi.org/10.1029/2007JB005263>
- Rost, S., & Thomas, C. (2002). Array seismology: Methods and applications. *Reviews of Geophysics*, *40*(3), 1008. <https://doi.org/10.1029/2000RG000100>
- Rychert, C. A., & Shearer, P. M. (2010). Resolving crustal thickness using SS waveform stacks. *Geophysical Journal International*, *180*(3), 1128–1137. <https://doi.org/10.1111/j.1365-246X.2009.04497.x>
- Sheriff, R. E., & Geldart, L. P. (1995). *Exploration seismology* (2nd ed., p. 592). Melbourne Australia: Cambridge University Press.
- Simmons, N. A., Forte, A. M., Boschi, L., & Grand, S. P. (2010). GyPSuM: A joint tomographic model of mantle density and seismic wave speeds. *Journal of Geophysical Research*, *115*, B12310. <https://doi.org/10.1029/2010JB007631>
- Tao, K., Grand, S. P., & Niu, F. (2017). Full-waveform inversion of triplicated data using a normalized-correlation-coefficient-based misfit function. *Geophysical Journal International*, *210*(3), 1517–1524. <https://doi.org/10.1093/gji/ggx249>
- Toh, A., Romanowicz, B., Capdeville, Y., & Takeuchi, N. (2005). 3D effects of sharp boundaries at the borders of the African and Pacific Superplumes: Observation and modeling. *Earth and Planetary Science Letters*, *233*(1–2), 137–153. <https://doi.org/10.1016/j.epsl.2005.01.037>
- Vidale, J. E. (1987). Waveform effects of a high-velocity, subducted slab. *Geophysical Research Letters*, *14*(5), 542–545. <https://doi.org/10.1029/GL014i005p00542>