Globally Scattered 2011 Tohoku Tsunami Waves From a Seafloor Sensor Array in the Northeast Pacific Ocean

Monica Kohler, Daniel Bowden, Jean-paul Ampuero, Jian Shi

JGR Solid Earth

RESEARCH ARTICLE
10.1029/2020JB020221

Globally Scattered 2011 Tohoku Tsunami Waves From a Seafloor Sensor Array in the Northeast Pacific Ocean

Monica D. Kohler[1], Daniel C. Bowden[2,3], Jean-Paul Ampuero[2,4], and Jian Shi[1]

[1]Department of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, CA, USA, [2]Seismological Laboratory, California Institute of Technology, Pasadena, CA, USA, [3]Now at Institute of Geophysics, ETH Zürich, Zürich, Switzerland, [4]Now at Université Côte d''Azur, IRD, CNRS, Observatoire de la Côte d'Azur, Géoazur, France

Abstract Modeling of tsunami wave propagation for forecasting focuses on the arrival time and amplitude of the earliest tsunami waves reaching coastlines. The complex later tsunami wavefield, in which scattering is predominant, poses additional hazards due to possible constructive interference of coherent packets of wave energy. However, almost no data sets exist to characterize the geographical sources and temporal evolution of the scattered waves. Here we show how recordings of the 2011 Tohoku tsunami by an array of pressure gauge sensors in the northeastern Pacific Ocean reveal coherent waves that are produced by scattering from distant coastlines including South America and Antarctica, as well as multiple seafloor fracture zones, ridges, and island chains. Multiple signal classification analysis and backward propagation ray tracing provide tight constraints on the origin of each scattered phase and resolve simultaneous wave arrivals from different scatterers. Incoming waves from constant back azimuths occur over time durations of several hours, revealing the time persistence of specific geographical scatterers. The results can advance numerical predictions of tsunami wave impact because they provide direct evidence for the necessity of incorporating both local and distant bathymetry over a range of length scales and for long time durations, to account for the azimuthal dependence of scatterer strength.

Plain Language Summary An array of sensors on the seafloor off the coast of California recorded the global tsunami produced by the 2011 magnitude 9.0 Tohoku, Japan, earthquake. The tsunami data recorded across the array show multiple, coherent waves arriving between 1 hr and more than 40 hr after the initial tsunami wave. Such long-duration tsunami waves are of concern because they arrive at distant coastlines long after the first, direct wave arrival. They can become amplified as they are influenced by the geometries of ports and harbors, sometimes causing damaging runup and inundation. This study involved a technique that takes advantage of similarities in data among the sensors to investigate which seafloor structures in the Pacific Ocean caused the long-lasting, later-arriving tsunami waves. By projecting the waves backward in space and time, specific structures were identified by pinpointing the geographical directions from which each wave arrived. The structures include distant coastlines, including those along South America and Antarctica, as well as many seafloor fracture zones, island and seamount chains, and mid-ocean ridges. Unless later-arriving wave amplitudes are adequately modeled over the deep, open ocean regions through which a tsunami wave propagates, their effects on distant coastlines may be severely underestimated.

Key Points:
- Coherent, long-duration 2011 Tohoku tsunami coda waves were recorded by a dense pressure gauge array in open ocean offshore California
- The seafloor sources of tsunami wave scattering were identified by application of a high-resolution beamforming and backprojection method
- Sources of scattering include South America and Antarctica coastlines, multiple seafloor fracture zones, ridges, and island chains

Supporting Information:
- Supporting Information S1
- Supporting Information S2

Correspondence to:
M. D. Kohler,
kohler@caltech.edu

Citation:
Kohler, M. D., Bowden, D. C., Ampuero, J.-P., & Shi, J. (2020). Globally scattered 2011 Tohoku tsunami waves from a seafloor sensor array in the northeast Pacific Ocean. *Journal of Geophysical Research: Solid Earth*, *125*, e2020JB020221. https://doi.org/10.1029/2020JB020221

Received 19 MAY 2020
Accepted 25 OCT 2020
Accepted article online 30 OCT 2020

1. Introduction

Accurate identification of large-amplitude scattered tsunami wave arrivals has implications for forecasting because it enables effective, time-evolving tsunami warnings for specific stretches of distant coastline. Tsunami modeling in operational warning systems is necessarily based on computationally efficient but simplified procedures that focus on the main first arrival. However, later tsunami arrivals have been damaging in a number of cases, especially if the population is unaware of the persistence time of the hazard (Barberopoulou et al., 2014; Contreras et al., 2016; Koshimura et al., 2008; Kowalik et al., 2008; Pattiaratchi & Wijeratne, 2009). Similar simplifying assumptions are pervasive in the way tsunami data are currently used to infer earthquake source properties: Usually, the first tsunami arrivals are typically considered in earthquake source inversions, though recent studies focus on how later tsunami arrivals can also be used to investigate the earthquake source process (Gusman et al., 2017; Kubota et al., 2018; Suppasri et al., 2017). While the governing equations of trans-oceanic tsunami propagation are well established,

some implications of the theory have received attention only after limitations of the available data were overcome. For instance, it is thanks to recordings of huge tsunamis in the past decade that research on long-period tsunami dispersion has gained new impetus (Tsai et al., 2013). Munk (1963) was among the first to identify tsunami scattering properties such as reverberations, decay times, and basin-wide energy diffusion using long-duration data from the 1960 *M*9.5 Chile earthquake (Miller et al., 1962). Scattering by mid-ocean bathymetry features is theoretically expected to affect tsunami waves, modifying their amplitudes, energy packet travel times, duration, and propagation direction (Mofjeld et al., 2000, 2001, 2004; Saito & Furumura, 2009). However, it is difficult to validate the theory of tsunami scattering, primarily because only a handful of large or regional tsunamis have been recorded on pressure gauge arrays with interstation spacing small enough, or for long-enough continuous time durations (Fukao et al., 2008; Gusman et al., 2016; Kubota et al., 2020; Matsumoto et al., 2017; Mizutani & Yomogida, 2019; Sandanbata et al., 2018; Sheehan et al., 2019; Thomson et al., 2011; Wang et al., 2019), to resolve the source of scattered waves.

Shorter-period waves (~10–30 min) in particular are influenced by the geometries of ports, harbors, and maritime facilities; this dependence, in turn, results in later-arriving, hazardous, tsunami-induced effects such as strong, persistent currents, jets, and large eddies (Borrero et al., 2015). In addition, this period band encompasses the range of periods seen in waves causing damaging runup and inundation in different locations around the world (Okal, Fritz, Raad et al., 2006; Okal, Fritz, Raveloson et al., 2006). This band is complementary to long-period tsunami observations (Heidarzadeh & Satake, 2014), but it is much more rarely observed in arrays of multiple sensors because of limited sample rates that are almost always too low and limited recording time durations where records either stop after a few hours or are further severely downsampled at the longer times.

Here we present a high-resolution analysis of a unique data set with the aim of advancing our understanding of tsunami wave scattering. The 11 March 2011 M_w 9.0 Tohoku, Japan, earthquake resulted in a major tsunami that propagated across the Pacific Ocean. A temporary array of ocean bottom seismometers (OBS) and differential pressure gauges (DPG) was deployed far off the coast of Southern California (Figure 1a), where it recorded the subsequent tsunami with unprecedented spatial and temporal resolution (Figure 1b) (Kohler & Science Team, 2010, 2011; Lin et al., 2015). This array comprised the ALBACORE (Asthenospheric and Lithospheric Broadband Architecture from the California Offshore Region Experiment) deployment and spanned a region that was 150 km north-south by 400 km east-west. It extended into the deep open Pacific Ocean, west of the continental shelf edge. In this array, 22 stations with an average spacing of 75 km were equipped with DPGs that recorded waveform time series data continuously at 50 samples per second (sps). The spatial and temporal resolution of the tsunami data enabled by the dense ALBACORE array configuration makes it possible to investigate features in the tsunami waveforms, especially the later but still prominent and coherent tsunami waves, which have not previously been observed with this level of clarity on these spatial and temporal scales.

The ability to identify potential sources of seafloor and coastline topographic scattering of tsunami waves is necessary to characterize the physical mechanisms of the long-duration scattering process. Classical beamforming has been previously applied to tsunami waveforms recorded by arrays of sensors to determine the direction-of-arrival of the incoming wave energy packets (e.g., Hanson & Bowman, 2005; Shi et al., 2017); however, it has limited ability to discriminate between true signals and noise unrelated to the tsunami energy, even after narrowband filtering. Nor does it distinguish between contemporaneous arrivals of energy from different scattering sources. In this study, we show how the application of a high-resolution array-processing technique to the dense ALBACORE tsunami data set leads to the identification of scattered tsunami wave arrivals and the accurate estimation of their direction-of-arrival (back azimuth). We apply the Multiple Signal Classification (MUSIC) beamforming method (Schmidt, 1986) to the Tohoku tsunami data, because of its increased resolution in determining the direction-of-arrivals over conventional delay-and-sum and coherence beamforming methods.

The specific seafloor bathymetric structures that caused scattering are then identified through backprojection (reverse wave tracking) based on the direction-of-arrival estimated from narrow band tsunami wave signals. Backprojection techniques have been applied to many fields including seismological and oceanographic phenomena to define spatiotemporal features of source processes (Hayashi et al., 2011; Heidarzadeh & Satake, 2014; Meng et al., 2011; Zhang et al., 2009). The spatial and temporal evolution of

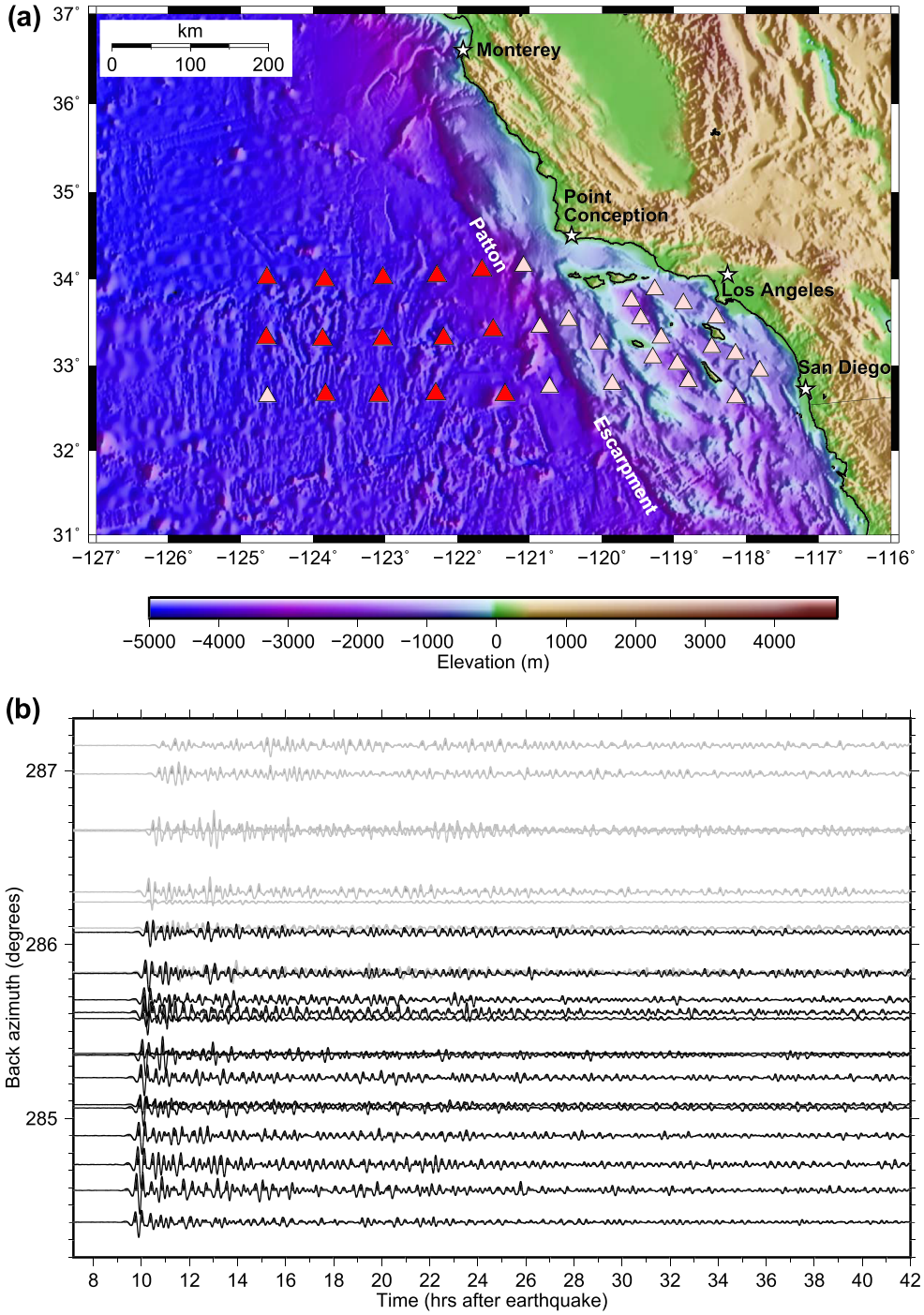

Figure 1. Offshore California ocean bottom pressure gauge locations and 2011 Tohoku tsunami waveforms. (a) Subset of differential pressure gauges that recorded the Tohoku tsunami used in this study (red triangles). (b) Tsunami waveforms, band-pass filtered 10–30 min, each normalized by the same constant to show relative amplitudes from sensor to sensor. Black waveforms correspond to subset of sensors used. Positive waveform polarities are filled in. Back azimuth = tsunami direction-of-arrival angle with respect to north between sensor and earthquake location.

earthquake ruptures has been imaged by backprojection of seismic waves recorded by dense arrays to track the source regions that generate strong and coherent high-frequency radiation (Fletcher et al., 2006; Ishii et al., 2005). Tsunami recordings by ocean bottom arrays have been used to locate a tsunami source (Fukao et al., 2008). Tsunami data assimilation has been proposed to predict later arrivals at coastlines in

real time to reconstruct the wavefield (Gusman et al., 2016; Maeda et al., 2015; Sheehan et al., 2019; Wang et al., 2019) which, although promising, requires numerical simulations and further assumptions about the wavefield.

The present work is a first application of a data-driven, high-resolution, beamforming technique to resolve tsunami scattering temporally and spatially. In this paper, we use the comprehensive definition of "scattered waves" that include reflected, refracted, and diffracted waves, as well as random waves. We use the term "coda" to refer to all observed wave energy that arrives after the first, direct, tsunami wave arrival, as recorded in the ALBACORE data. We deliberately focus on the shorter-period range because of the hazard implications of wave amplification at harbors and ports, since these are the final destinations of the scattered waves.

2. Beamforming of Tsunami Waves

The first step in our analysis is to identify coherent waves present on sliding time windows of the array data and estimate their direction-of-arrival at the array. Our analysis uses 14 DPG sensors located in the open ocean from the ALBACORE array near Southern California. It excludes sensors that were on the continental shelf (Figure 1, east of the Patton Escarpment that defines the edge of the continental shelf) to avoid complexity and bias due to local scattering sources such as nearby islands, irregularities in the Southern California coastline, and the abrupt bathymetry gradient introduced at the Patton Escarpment. It also excludes stations that did not return data due to faulty operation. To assess how the scattered wavefield depends on frequency, which controls the ratio of wavelength-to-size of scattering features, we conduct the analysis at various narrow frequency bands separately. The period bands examined here are as follows: 9–11, 11–13, 13–15, 15–17, 17–19, 19–21, 21–23, 23–25, 25–27, 27–29, and 29–31 min, corresponding to wavelengths between ~90 and ~315 km in 3,000 m water depth. A sliding window of data is used that is 3 times the average period of each band, with beamforming applied at each 100 s interval so that there is a considerable amount of overlap from one window to the next (Meng et al., 2011).

The MUSIC method (Schmidt, 1986) is a subspace algorithm especially designed to determine the direction-of-arrival of multiple waves recorded simultaneously by a dense array. For a detailed description we refer to Meng et al. (2011, 2012), whose implementation we adapted to tsunami waves. Traditional beamforming compares signals across an array, testing a range of possible time shifts according to different directions and speeds of wave propagation, until the waveforms are most similar. MUSIC beamforming is a modification that includes an eigenvalue decomposition to distinguish true signals from noise and it consequently sharpens or improves the beamform image (Meng et al., 2011, 2012; Schmidt, 1986).

To describe the MUSIC algorithm, we first start with a brief description of standard coherence beamforming. This is performed in the following way: (1) At a given time window and frequency, a cross-coherence matrix, R_{xx}, is constructed between all station pairs in the array, which measures the similarity of waveforms assuming no time shifts are applied. That matrix is

$$R_{xx} = \begin{bmatrix} x_1(\omega)x_1(\omega) & x_1(\omega)x_2(\omega) & \dots \\ x_2(\omega)x_1(\omega) & x_2(\omega)x_2(\omega) & \dots \\ \dots & \cdots & \dots \end{bmatrix} \quad (1)$$

where $x_n(\omega)$ is the signal at the nth station for frequency, ω. Because the time window contains only three wave periods on average, a multitaper method (Thomson, 1982) is applied to improve the robustness of the cross-spectrum estimate that is used in the cross-coherence matrix computations. (2) A grid search over all possible azimuths and propagation speeds is next performed. For a given azimuth and wave speed, an incident plane wave can be expressed as a phase delay a_n between two stations, where $a_n = e^{-i2\pi f\, r_n \cdot S}$ for the nth station, f is frequency, S is the slowness vector, and r_n is the position vector of the station relative to an arbitrary reference point of the array; in our case, this is a point in the middle of the pressure gauge array. (3) In traditional beamforming, the appropriate delay is then applied to each station in the matrix R_{xx}, and the total power of time-shifted traces is given as a scalar:

$$P = \frac{a^T R\, a}{a^T\, a} \tag{2}$$

where a^T denotes the transpose of a. The cumulative power, P, represents the success, or lack thereof, of the plane wave having originated from that particular direction with that particular speed (again, prescribed by the phase delays in each a_n).

Unique to the MUSIC approach, an eigenvalue decomposition of the matrix R_{xx} is first used to separate the more significant signals from the less significant "noise." The larger-amplitude eigenvalues are assumed to be true signal, and the smaller eigenvalues are the noise. Rather than use the more significant eigenvectors, it is actually more stable to collect those signals orthogonal to the noise eigenvectors. The final MUSIC pseudo-spectrum (so named because it now only uses a subset of eigenvectors) is measured as follows:

$$P_{MUSIC} = \frac{a^T a}{a^T\, R_{noise}\, a}. \tag{3}$$

An advantage of the MUSIC approach, given the tighter constraints between signal space and noise space, is that it should be able to resolve multiple signals incident upon the array simultaneously, if they exist. As such, secondary maxima of the MUSIC pseudo-spectra are also considered and deemed significant based on their amplitudes relative to that of the primary maximum. We note that the MUSIC method assumes signals in the presence of Gaussian white noise. In our application of tsunami wavefronts, this noise may include waves due to very localized scattering (i.e., within the scale of the array), localized wind effects on the water waves, and instrument errors. Because of the spatially localized nature of these noise-generating mechanisms, they are not coherent across the entire array.

In general beamforming applications, the slowness of the plane waves moving across the array is an additional unknown. However, the slowness of tsunami waves is sufficiently well known, based on bathymetry and previous analysis of the ALBACORE data (Lin et al., 2015), so here we limit the search to a narrow range of slownesses. For the beamforming, a narrow annulus of slowness is used to constrain possible large-amplitude values in the pseudo-spectrum of interest (hereafter referred to as a "hot spot" in the "heatmap") to preclude aliasing effects. The cumulative energy in this annulus is plotted for each time window in the coda waves (after the first arrival), as tsunami energy continues to arrive over the next 40 hr after the initial direct arrival. Given MUSIC's sharpness in identifying the direction-of-arrival at each time window, we use these heatmaps to manually identify significant sources of scattering ("scatterers") to further investigate, with at least 30 hot spots for each period band considered as candidates for backward-propagated ray tracing.

A disadvantage of the MUSIC approach is that amplitude information is effectively lost; traces are normalized when measuring coherence, and the eigenvalue decomposition further distorts any measure of raw signal amplitude. To obtain a more robust measure of relative wave amplitudes, we turn to the more direct delay-and-sum beamforming. Once an azimuth of interest is identified by MUSIC, the pressure gauge waveforms are simply time shifted and summed according to the direction-of-arrival indicated by MUSIC. This sum can be related to the total power observed passing over our array (Rost & Thomas, 2002). We can then compare the relative power arriving at a receiver at any time in the hours after the main arrival, reported in decibels (dB), and indicated by open circles in Figure 6 and supporting information Figures S1–S11 (top panels).

We also estimate the uncertainty of our direction-of-arrival estimates. We follow the approach of Fletcher et al. (2006) to repeat the MUSIC identification after adding 1,000 realizations of randomly phase-shifted noise to the waveforms. This lets us assess the method's robustness under background noise variations. In each time window, a single maximum of slowness direction and magnitude is selected. For this selection, rather than restricting our maximum solely to the annulus of slowness described here previously, we apply a reduced weighting to amplitudes outside the annulus; energy is reduced linearly away from the expected slowness such that aliased hot spots are unlikely to be selected. However, the pseudo-spectrum is otherwise unaffected. Given the identified maxima for the time window, data are aligned and traces are averaged to define a "master" waveform which is subtracted from all waveforms to create residual "noise."

These noise residuals are given random phase shifts for each frequency in the Fourier domain then added back to the original master waveform, and the MUSIC process is repeated. Because these waveforms are based on the same master waveform, ideally the new MUSIC pseudo-spectrum hot spot should be centered at a slowness of 0,0 in the east-west and north-south directions. The extent to which this hot spot moves around over the course of 1,000 realizations of the randomly phase-shifted noise indicates the strength and impact of the noise and thus our uncertainty. Given the variable hot spot peaks, a contour is drawn for the 95% confidence level. This confidence contour is shifted back to the original MUSIC pseudo-spectrum, and the minimum and maximum azimuths spanned are measured. Within a 95% confidence level, the back azimuth angles are within ±2.6°, and thus, we use this as the estimate of beamforming uncertainty in the following sections.

3. Backward Propagation Ray Tracing

We use the information from the coherent pseudo-spectrum amplitudes obtained in the previous step to identify the actual geographic locations of the scatterers that may have produced them. To accomplish this, we use a backward propagation ray tracing technique to trace the tsunami wave from the ALBACORE array to the scattering source. First, we compute the travel times of a virtual tsunami that originates from the center point of the 14 ALBACORE stations using the TTT travel time solver that assumes long waves (shallow water) and Huygens principle for the propagating wavefront with a user-specified bathymetry model (Amante & Eakins, 2009). However, since this approach assumes a ray approximation, we must first account for the fact that the width of our filtered tsunami rays is not infinitely small. The filtered tsunami wave travel times will be affected by bathymetry features as large as the wavelengths corresponding to the period filter bands, that is, finite-frequency effects. We thus smooth ETOPO1 bathymetry (Amante & Eakins, 2009) to the appropriate length scales corresponding to the period bands. In fact, we convert water depth to wave slowness and smooth the slowness values to the relevant length scale. We then use the smoothed bathymetry models as input to the solver to obtain the tsunami's 2-D travel time contour field for the Pacific Ocean.

We next compute the gradient of the travel time to produce the velocity field. We start the ray tracing from the virtual tsunami origin point (center of the ALBACORE array), making a small step in the direction of the back azimuth of a selected hot spot identified in the previous beamforming section. We evaluate the new ray direction at the end of the step by interpolating the computed velocity field and then make a new step. The procedure is iterated until reaching the end result, which is a backward-traced ray corresponding to a specified back azimuth using the ALBACORE array as the starting point. We repeat the ray tracing for a range of back azimuths spanning the uncertainty of our array-processing estimates. For each choice of hot spot starting point, we use five takeoff back azimuth angles, in increments of ±1.3° with respect to the center (average) of the range, to reach our beamform uncertainty of ±2.6°. The variation provides an indication of the stability or certainty (uncertainty) in our results: In some cases, a slight variation in the starting ray's takeoff angle results in a significantly different final location, and in other cases the rays robustly converge to the same place.

The backward projection is done for each of the hot spots that have been computed for times T up to 40 hr after the initial direct tsunami arrival. For given T, the backward rays end at the (x,y) location in the ocean where the observed travel time matches the total travel time of a tsunami wave scattered at (x,y). This end point satisfies the equation $T_{\text{direct}} + T = T_{\text{forward}}(x,y) + T_{\text{backward}}(x,y)$, where T_{direct} is the tsunami's shortest-path travel time from the earthquake epicenter directly to the ALBACORE array (~9.8 hr), and $T_{\text{forward}}(x,y)$ and $T_{\text{backward}}(x,y)$ are two travel time fields computed forward from the Tohoku epicenter to (x,y) and backward from ALBACORE to (x,y), respectively.

We cross reference the end points of the backward-traced rays with the corresponding bathymetry features (Tarr et al., 2010) and manually identify the regions that contributed to a particular coherent arrival. Figures 2–5 show the backward-propagated ray tracing and coherence results. Using the total energy computed by standard delay-and-sum beamforming at the back azimuth and time identified by the MUSIC analysis, we assign an amplitude to each scatterer as shown by the open circles in Figure 6 (top).

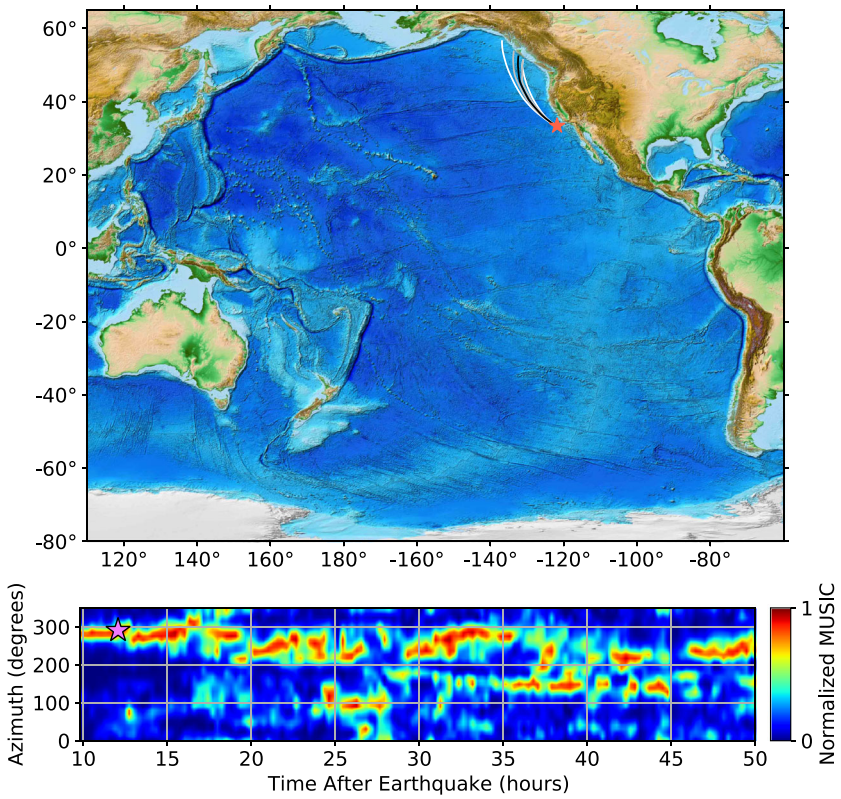

Figure 2. Backpropagation ray tracing results (top) for 17–19 min band using MUSIC pseudo-spectrum method (bottom). Arrival time shown here is 2.3 hr into the tsunami coda (first tsunami wave arrival at array is ~9.8 hr after earthquake). Red star = center of ALBACORE array. Magenta star = time point of coherence value used in this backpropagation calculation. In the bottom panel, each sliding time window's MUSIC pseudo-spectrum is normalized to one as shown in the normalized MUSIC color bar; thus, the colors correspond to time-varying maxima. In the top panel, the black ray uses the center (average) of the back azimuth takeoff point (0°), gray rays indicate ±1.3°, and white rays indicate ±2.6° spanning the uncertainty range. Rays stop at the Alaska coastline.

4. Interpretation of Scattering Sources

A feature that clearly stands out using the MUSIC pseudo-spectrum method is the long time duration of coherent waves arriving from approximately constant azimuths (Figures 2–5, bottom). The high resolution leads to clear identification of the long-term trends; the pseudo-spectrum versus azimuth figures show incoming wave arrivals from a particular back azimuth range over time durations of a few to many hours, as well as persistent arrivals that begin many hours after the initial arrival. This suggests the time persistence of specific geographical scatterers.

A comprehensive compilation of scatterers that we identified from back-propagated points in all bands between 9 and 31 min is listed in supporting information Table S1, with highlights described here. This compilation is not necessarily complete but represents a reasonable selection of unique scatterers within the entire back azimuth range. Scatterers span many major fracture zones (Mendocino, Murray, Molokai, Clipperton, Galapagos, Clarion, Marquesas, Easter, Challenger, Udintsev, and Valdivia), ridges and rises (East Pacific, Pacific-Antarctic, Louisville, Chile, Sala y Gomez, Nazca, Juan Fernandez, and Boudeuse), plateaus (Chatham and Kapingamarangi), and island chains and archipelagos (Hawaii, Tahiti, Fiji, Samoa, Solomon, Shetland, Palmer, Tuamotu, Line, Society, and Austral). The Emperor and Mellish Seamounts do not show up in our results as significant last-leg scatterers even though they have previously been identified as sources of scattered energy (Tang et al., 2012). The reason for this is that they are on a nearly direct wave path between Tohoku and Southern California. They might act as scatterers

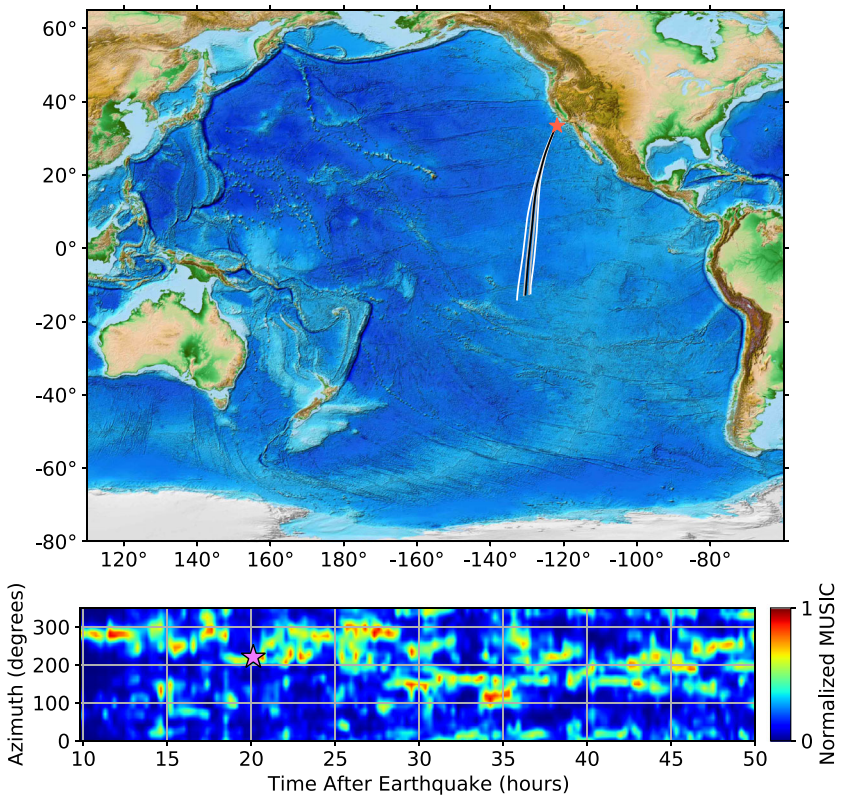

Figure 3. Backpropagation ray tracing results (top) for 13–15 min band using MUSIC pseudo-spectrum method results (bottom). Arrival time shown here is 10.3 hr into tsunami coda (first tsunami wave arrival at array is ~9.8 hr after earthquake). Red star = center of ALBACORE array. Magenta star = time point of coherence value used in this backpropagation calculation. In the bottom panel, each sliding time window's MUSIC pseudo-spectrum is normalized to one as shown in the normalized MUSIC color bar; thus, the colors correspond to time-varying maxima. In the top panel, the black ray uses the center (average) of the back azimuth takeoff point (0°), gray rays indicate ±1.3°, and white rays indicate ±2.6° spanning the uncertainty range. Rays stop at the Marquesas fracture zone in French Polynesia.

but will not cause a delay in arrival times in Southern California data and therefore will not be seen later in the coda.

Several scattering regions are common to the wave paths in all of the narrow period bands. These are the Gulf of Alaska coastline from the easternmost Aleutian Islands to northern British Columbia (e.g., Figure 2), central and Southern California, Baja California, southeast Hawaiian Islands region including the westernmost Molokai Fracture Zone, structures within Fiji and Samoa Islands, and French Polynesia (Figure 3). Clear, persistent arrivals originating from the South America and Antarctic coastlines (Figures 4 and 5) are also present in all period bands at longer times. Reflections off the western South American coastline include scattering locations between Ecuador and southern Chile, as well as several ridge and fracture zones offshore Chile (Figure 4; additional examples are in Figures S2I, S7G, and S9H). The longest-span scattered waves are the result of tsunami waves that have traveled south past the South Pacific to Antarctica and reflected off the coastline at locations in the Ross Sea, Amundsen Sea, and Bellingshausen Sea, depending on takeoff angle and intermediate scattering between those points (Figure 5; additional examples are in Figures S3G, S4H, and S5I).

All rays for the 15–17 min band for the entire 10–50 hr (after earthquake) time window are shown cumulatively in Figure 6, to illustrate the time evolution of the scatterer locations. As expected, the more distant scatterers such as Fiji, South America, and Antarctica coastlines take longer to arrive at the ALBACORE

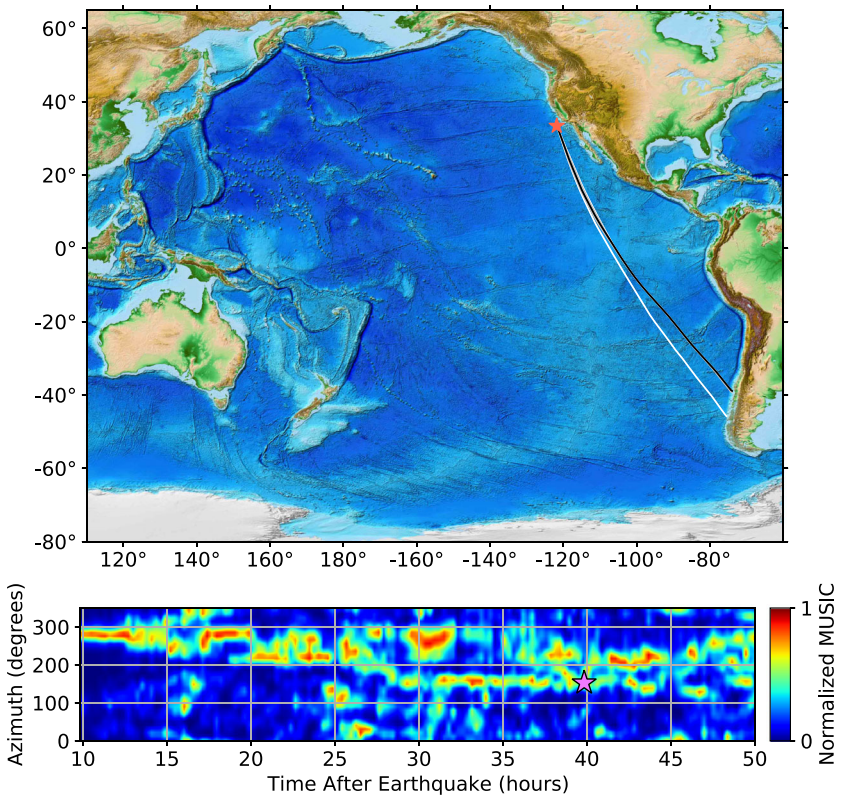

Figure 4. Backpropagation ray tracing results (top) for 15–17 min band using MUSIC pseudo-spectrum method results (bottom). Arrival time shown here is 30.0 hr into tsunami coda (first tsunami wave arrival at array is ~9.8 hr after earthquake). Red star = center of ALBACORE array. Magenta star = time point of coherence value used in this backpropagation calculation. In the bottom panel, each sliding time window's MUSIC pseudo-spectrum is normalized to one as shown in the normalized MUSIC color bar; thus, the colors correspond to time-varying maxima. In the top panel, the black ray uses the center (average) of the back azimuth takeoff point (0°), gray rays indicate ±1.3°, and white rays indicate ±2.6° spanning the uncertainty range. Gray rays converge with black ray. Rays stop at the southwest coastline of Chile.

array. However, several locations are prominent scatterers for both shorter (<5 hr into coda) and longer times (>5 hr into coda). For example, the Gulf of Alaska produces scattered arrivals within a wide range of times most likely because it is proximal to the path of waves traveling directly from the Tohoku source, and it is also the source of multiply reflected, refracted, or diffracted waves traveling within the Aleutian Island and Bering Basin region. The power of the coherent arrivals, reported in decibels relative to the strongest, early arrivals, is shown by the circle sizes in Figure 6. These decrease with increasing distance from the scatterer, where the Hawaiian Islands is associated with the largest amplitude because it is one of the earliest scatterers and has not suffered from as much geometrical spreading. While Figure 6 features some of the same bathymetry structures that are used to construct the tsunami scattering index (Mofjeld et al., 2000, 2001, 2004), the inconsistencies between our amplitudes and the scattering index values illustrate the azimuth dependence of the strength of those scatterers. We emphasize that these powers only represent the wavefield as observed at the ALBACORE array, and the actual scattering mechanism by a given bathymetric feature cannot be quantified by this single observation.

The application of the MUSIC method enables us to distinguish between two simultaneous arrivals from significantly different back azimuths because it dramatically increases the temporal and azimuthal resolution; this leads to smaller, more precise starting back azimuth angles for ray tracing. If the beamforming analysis were conducted in real time, we could easily distinguish whether the observed tsunamis are approaching

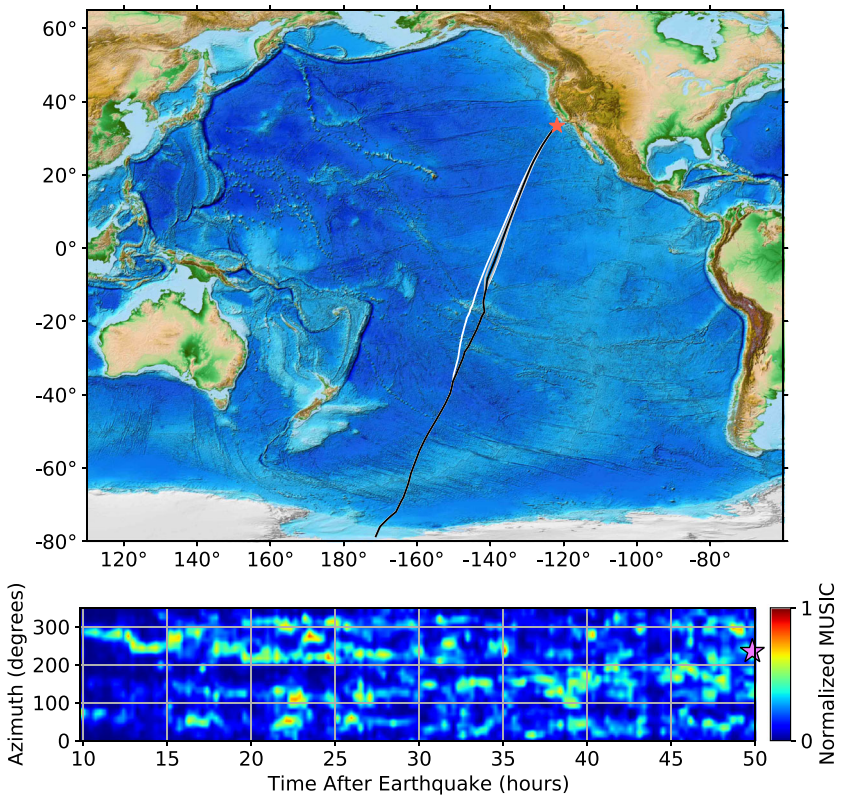

Figure 5. Backpropagation ray tracing results (top) for 9–11 min band using MUSIC pseudo-spectrum method results (bottom). Arrival time shown here is 40.0 hr into tsunami coda (first tsunami wave arrival at array is ~9.8 hr after earthquake). Red star = center of ALBACORE array. Magenta star = time point of coherence value used in this backpropagation calculation. In the bottom panel, each sliding time window's MUSIC pseudo-spectrum is normalized to one as shown in the normalized MUSIC color bar; thus, the colors correspond to time-varying maxima. In the top panel, the black ray uses the center (average) of the back azimuth takeoff point (0°), gray rays indicate ±1.3°, and white rays indicate ±2.6° spanning the uncertainty range. The three raypaths (black, gray, and white) coincide spatially in the southern end of the raypath; thus, only the black ray shows up since the others lie underneath it. Rays stop at Antarctica coastline in Ross Sea.

land or are moving out into the open ocean. In this study, MUSIC helps us determine whether there are two possible sources of scattering for similar wave travel time windows. This is apparent at various times in the period bands analyzed here (see entries in supporting information Table S1 where simultaneous arrivals from different locations are described as "split" arrivals and Figures S7D and S8A for examples). In the 17–19 min band, two possible scatterer solutions occur at 4.8 and 5.3 hr into the coda with some waves coming from the southern end of the Mid-Pacific Mountains and others from southwest Alaska (Table S1). An example from the 15–17 min band includes waves from both the Murray Fracture Zone northeast of the Hawaiian Islands and southwest Alaska coastline (at 1.2 hr into the coda). These and additional solution pairs are documented in supporting information Table S1 and Figures S1–S11.

Our results indicate that for the entire 9–31 min period range, the types and geographical locations of most of the scatterers are similar across all bands and thus similar across the corresponding tsunami wavelengths. For example, fracture zones (e.g., Clarion, Clipperton, Marquesas, and Molokai) appear to be the source of coherent waves for both shorter- and longer-period tsunami waves. We hypothesize that this is because most of these scatterers have both long- and short-wavelength heterogeneity in their geometries. Local sections of coastline are the source of multiple arrivals observed both early and late in the coda, for example, the Southern and central California coastlines and Baja California Peninsula in northern Mexico. In the case of

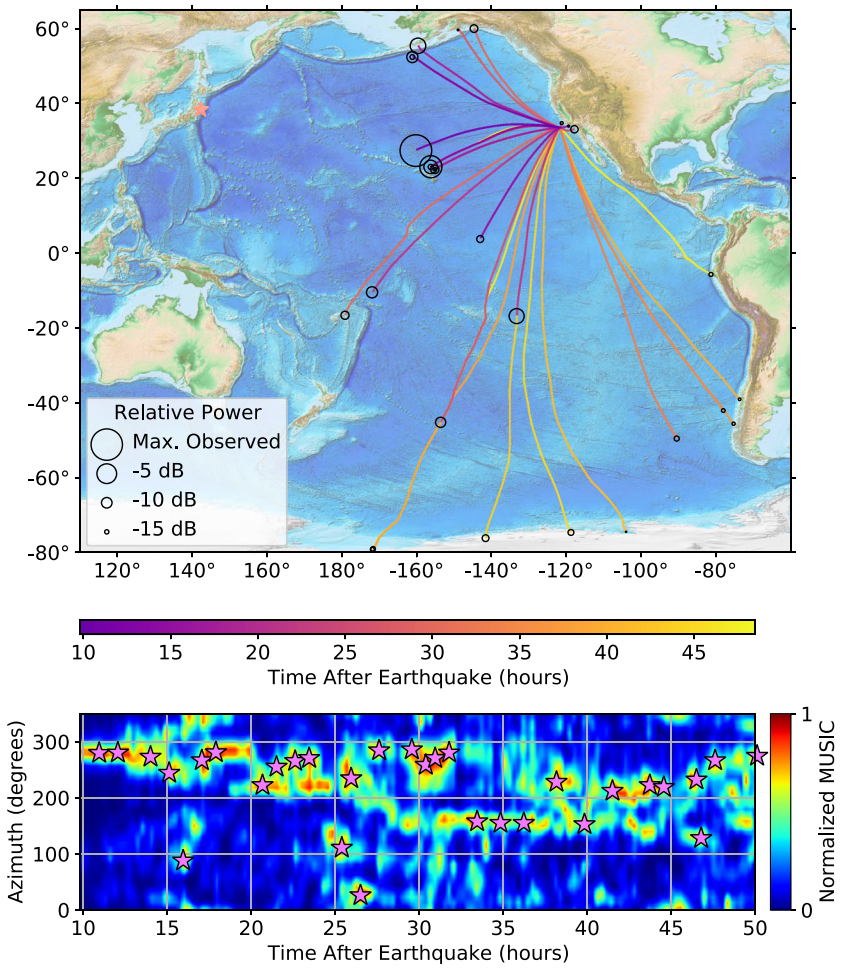

Figure 6. Relative wave amplitudes (top) as a function of time for selected 15–17 min waves and their MUSIC pseudo-spectrum values (bottom). Wave amplitudes in the top panel are indicated by circle sizes which show relative power in logarithmic units (decibels). Circle is plotted at scatterer location for selected arrivals, indicated by magenta stars, between 10 and 50 hr after earthquake. Ray colors, indicated in color bar, correspond to time after earthquake. In the bottom panel, each sliding time window's MUSIC pseudo-spectrum is normalized to one as shown in the normalized MUSIC color bar; thus, the colors correspond to time-varying maxima.

Southern California, the tsunami waves are likely being diffracted around the Point Conception headland on the Southern-central California coastline as they are arriving from the Gulf of Alaska, though we cannot rule out the possibility that they are multiple in-place reflections or the result of wave trapping between the local island plateaus, coastline, and/or continental shelf (e.g., González et al., 1995; Melgar & Ruiz-Angulo, 2018). In the 9–11 min band, arrivals from near, but not directly at, the edge of the Antarctic coastline is interpreted as possible reflections from the continental shelf offshore in the Ross Sea.

5. Conclusions

This study presents an experimental investigation into the complex, time-evolving tsunami coda wavefield—the long-duration component of the tsunami wave that arrives at distant coastlines after the first, direct wave arrival. This part of the wavefield poses significant additional hazard due to amplification of coherent packets of wave energy that can arrive at a coastline several hours after the first tsunami wave. We apply high-resolution multiple signal analysis beamforming and backpropagation ray tracing to a dense array of seafloor pressure gauge recordings made far offshore Southern California from the 2011 Tohoku tsunami to characterize up to 2 days of coda wavefield data. While previous analyses have focused on numerical

simulations, this data-based analysis reveals the variety of geographic locations in the Pacific Ocean that produced scattered phases in the coda. The combination of unique data set with high-resolution analysis method makes it possible to resolve the following: the long time persistence of geographical scatterers, multiple scatterers that produced simultaneous wave arrivals, and the azimuthal dependence of the strength of the scatterers. The scatterers for the 9–31 min period range include many fracture zones in the Pacific Ocean, island and seamount chains, mid-ocean ridges and rises, and local and distant coastlines including those in South America and Antarctica.

This study provides new targets for tsunami propagation modeling approaches. One of the challenging aspects of tsunami modeling is accurately characterizing the nonlinear behavior at shallow depths near coasts; this requires small-scale fine-resolution bathymetry models in the near-coastal regions. Our results illustrate that appropriately scaled bathymetry should be used over the entire wave path, including when the path is not direct and includes one or more significant scatterers. Such scatterers can cause distant reflected, refracted, or diffracted wave effects with amplification factors that depend on source-to-receiver azimuth. Unless the scattering amplitudes are adequately modeled over the deep, open ocean regions through which a tsunami wave propagates, their effects on distant coastlines may be severely underestimated.

Data Availability Statement

The OBS waveform data from the ALBACORE array are publicly available from the IRIS Data Management Center (https://ds.iris.edu/ds/nodes/dmc/). Specifically, the SeismiQuery web-based tool (https://ds.iris.edu/SeismiQuery/) shows ALBACORE data availability and how to access the continuous, time series data by choosing network name "2D" for years 2010 and 2011. Requests for the waveforms can be made using the BREQ_FAST software tool either by email or by web request (http://ds.iris.edu/ds/nodes/dmc/forms/breq-fast-request/); this link is also provided by SeismiQuery. The DPG waveforms are indicated by channels "LDH" and "LXH." Alternatively, the data can be accessed from IRIS using ObsPy software tools for seismology (https://docs.obspy.org/). The topography and bathymetry model used in all figures and computations is ETOPO1 provided by the NOAA National Centers for Environmental Information (https://www.ngdc.noaa.gov/mgg/global/global.html).

Acknowledgments

This manuscript greatly benefitted from thoughtful reviews provided by Frank González and an anonymous reviewer. The OBS deployment was made possible with instruments and logistical support of the Ocean Bottom Seismic Instrumentation Pool at Scripps Institution of Oceanography. The ALBACORE OBS deployment and recovery cruises were made possible with the equipment and logistical support of the University-National Oceanographic Laboratory System (UNOLS) vessel fleet and staff support at Scripps. Support for this work was provided by the NSF Hazard-SEES program Grant 1331600. J.P.A. was supported by the French government, through the UCAJEDI Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-15-IDEX-01. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience and EarthScope (SAGE) by the National Science Foundation under Cooperative Agreement EAR-1261681.

References

Amante, C., & Eakins, B. W. (2009). ETOPO1 1 arc-minute global relief model: Procedures, data sources and analysis, NOAA Technical Memorandum NESDIS NGDC-24, 19 pp.

Barberopoulou, A., Legg, M. R., Gica, E., & Legg, G. (2014). Multiple wave arrivals contribute to damage and tsunami duration on the US west coast. In Y. Kontar, V. Santiago-Fandiño, & T. Takahashi (Eds.), *Tsunami events and lessons learned. Advances in natural and technological hazards research* (Vol. 35, pp. 359–376). Dordrecht: Springer. https://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-94-007-7269-4_20

Borrero, J. C., Lynett, P. J., & Kalligeris, N. (2015). Tsunami currents in ports. *Philosophical Transactions of the Royal Society A*, *373*, 20,140,372. https://doi.org/10.1098/rsta.2014.0372

Contreras, M., Winckler, P., Sepulveda, I., Andaur-Álvarez, A., Cortes, F., Guerrero, C., et al. (2016). Field survey of the 2015 Chile tsunami with emphasis on coastal wetland and conservation areas. *Pure and Applied Geophysics*, *173*. https://doi.org/10.1007/s00024-015-1235-2

Fletcher, J. B., Spudich, P., & Baker, L. M. (2006). Rupture propagation of the 2004 Parkfield, California, earthquake from observations at the UPSAR. *Bulletin of the Seismological Society of America*, *96*, S129–S142. https://doi.org/10.1785/0120050812

Fukao, Y., Sandanbata, O., Sugioka, H., Ito, A., Shiobara, H., Watada, S., & Satake, K. (2008). Mechanism of the 2015 volcanic tsunami earthquake near Torishima, Japan. *Science Advances*, *4*, eaao0219. https://doi.org/10.1126/sciadv.aao0219

González, F. I., Satake, K., Boss, E. F., & Mofjeld, H. O. (1995). Edge wave and non-trapped modes of the 25 April 1992 Cape Mendocino tsunami. In F. Imamura & K. Satake (Eds.), *Tsunamis: 1992–1994. Pageoph topical volumes* (pp. 409–426). Basel: Birkhäuser. https://doi.org/10.1007/978-3-0348-7279-9_3

Gusman, A. R., Satake, K., Shinohara, M., Sakai, S., & Tanioka, Y. (2017). Fault slip distribution of the 2016 Fukushima earthquake estimated from tsunami waveforms. *Pure and Applied Geophysics*, *174*(8), 2925–2943. https://doi.org/10.1007/s00024-017-1590-2

Gusman, A. R., Sheehan, A. F., Satake, K., Heidarzadeh, M., Mulia, I. E., & Maeda, T. (2016). Tsunami data assimilation of Cascadia seafloor pressure gauge records from the 2012 Haida Gwaii earthquake. *Geophysical Research Letters*, *43*, 4189–4196. https://doi.org/10.1002/2016GL068368

Hanson, J. A., & Bowman, J. R. (2005). Dispersive and reflected tsunami signals from the 2004 Indian Ocean tsunami observed on hydrophones and seismic stations. *Geophysical Research Letters*, *32*, L17606. https://doi.org/10.1029/2005GL023783

Hayashi, Y., Tsushima, H., Hirata, K., Kimura, K., & Maeda, K. (2011). Tsunami source area of the 2011 off the Pacific coast of Tohoku earthquake determined from tsunami arrival times at offshore observation stations. *Earth, Planets and Space*, *63*(7), 809–813. https://doi.org/10.5047/eps.2011.06.042

Heidarzadeh, M., & Satake, K. (2014). Possible sources of the tsunami observed in the northwestern Indian Ocean following the 2013 September 24 Mw 7.7 Pakistan inland earthquake. *Geophysical Journal International*, *199*, 752–766. https://doi.org/10.1093/gji/ggu297

Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E. (2005). Extent, duration and speed of the 2004 Sumatra-Andaman earthquake imaged by the Hi-net array. *Nature*, *435*(7044), 933–936. https://doi.org/10.1038/nature03675

Kohler, M. D. & Science Team (2010). ALBACORE OBS deployment cruise report. *R/V Melville Cruise* MV1010, https://authors.library.caltech.edu/64443/

Kohler, M. D. & Science Team (2011). ALBACORE OBS recovery cruise report. *R/V New Horizon Cruise* NH1111, https://authors.library.caltech.edu/64442/

Koshimura, S., Hayashi, Y., Munemoto, K., & Imamura, F. (2008). Effect of the Emperor seamounts on trans-oceanic propagation of the 2006 Kuril Island earthquake tsunami. *Geophysical Research Letters*, *35*, L02611. https://doi.org/10.1029/2007GL032129

Kowalik, Z., Horrillo, J., Knight, W., & Logan, T. (2008). Kuril Islands tsunami of November 2006: 1. Impact at Crescent City by distant scattering. *Journal of Geophysical Research*, *113*, C01020. https://doi.org/10.1029/2007JC004402

Kubota, T., Saito, T., & Suzuki, W. (2020). Millimeter-scale tsunami detected by a wide and dense observation array in the deep ocean: Fault modeling of an Mw 6.0 interplate earthquake off Sanriku, NE Japan. *Geophysical Research Letters*, *47*, e2019GL085842. https://doi.org/10.1029/2019GL085842

Kubota, T., Saito, T., Yoshihiro, I., Yoshihiro, K., Wallace, L. M., Suzuki, S., et al. (2018). Using tsunami waves reflected at the coast to improve offshore earthquake source parameters: Application to the 2016 Mw 7.1 Te Araroa earthquake, New Zealand. *Journal of Geophysical Research: Solid Earth*, *123*, 8767–8779. https://doi.org/10.1029/2018JB015832

Lin, F.-C., Kohler, M. D., Lynett, P., Ayca, A., & Weeraratne, D. (2015). The March 11, 2011 Tohoku tsunami wavefront mapping across offshore southern California. *Journal of Geophysical Research: Solid Earth*, *120*, 3350–3362. https://doi.org/10.1002/2014JB011524

Maeda, T., Obara, K., Shinohara, M., Kanazawa, T., & Uehira, K. (2015). Successive estimation of a tsunami wavefield without earthquake source data: A data assimilation approach toward real-time tsunami forecasting. *Geophysical Research Letters*, *42*, 7923–7932. https://doi.org/10.1002/2015GL065588

Matsumoto, H., Nosov, M. A., Kolesov, S. V., & Kaneda, Y. (2017). Analysis of pressure and acceleration signals from the 2011 Tohoku earthquake observed by the DONET seafloor network. *Journal of Disaster Research*, *12*, 163–175. https://doi.org/10.20965/jdr.2017.p0163

Melgar, D., & Ruiz-Angulo, A. (2018). Long-lived tsunami edge waves and shelf resonance from the M8. 2 Tehuantepec earthquake. *Geophysical Research Letters*, *45*, 12,414–12,421. https://doi.org/10.1029/2018GL080823

Meng, L., Ampuero, J. P., Sladen, A., & Rendon, H. (2012). High-resolution back-projection at regional distance: Application to the Haiti M7.0 earthquake and comparisons with finite source studies. *Journal of Geophysical Research*, *117*, B04313. https://doi.org/10.1029/2011JB008702

Meng, L., Inbal, A., & Ampuero, J. P. (2011). A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-Oki earthquake. *Geophysical Research Letters*, *38*, L00G07. https://doi.org/10.1029/2011GL048118

Miller, G. R., Munk, W. H., & Snodgrass, F. E. (1962). Long-period waves over California's continental borderland. Part II: Tsunamis. *Journal of Marine Research*, *20*, 31–41.

Mizutani, A., & Yomogida, K. (2019). Near-fault ocean-bottom observation of an offshore earthquake by DONET: Array-based measurement of tsunami phase speed and early estimation of tsunami height, In: AGU Fall Meeting 2019 S33D-S0604D, at https://agu.confex.com/agu/fm19/meetingapp.cgi/Paper/501051

Mofjeld, H. O., Symons, C. M., Lonsdale, P., González, F. I., & Titov, V. V. (2004). Tsunami scattering and earthquake faults in the deep Pacific Ocean. *Oceanography*, *17*, 38–46.

Mofjeld, H. O., Titov, V. V., González, F. I., & Newman, J. C. (2000). Analytic theory of tsunami wave scattering in the open ocean with application to the North Pacific, *NOAA Tech. Memo*, *OAR PMEL-116*, 38 pp., PMEL, Seattle, Washington.

Mofjeld, H. O., Titov, V. V., González, F. I., & Newman, J. C. (2001). Tsunami scattering provinces in the Pacific Ocean. *Geophysical Research Letters*, *28*, 335–338.

Munk, W. H. (1963). Some comments regarding diffusion and absorption of tsunamis. In *Proceedings of tsunami meetings, tenth Pacific science congress, Honolulu, Hawaii, 1961*, *International Union of Geodesy and Geophysics (IUGG) Monograph* (Vol. 24, pp. 53–72). International Union of Geodesy and Geophysics.

Okal, E. A., Fritz, H. M., Raad, P. E., Synolakis, C. E., Al-Shijbi, Y., & Al-Saifi, M. (2006). Oman field survey after the December 2004 Indian Ocean tsunami. *Earthquake Spectra*, *22*, 203–218. https://doi.org/10.1193/1.2202647

Okal, E. A., Fritz, H. M., Raveloson, R., Joelson, G., Pancockova, P., & Rambolamanana, G. (2006). Madagascar field survey after the December 2004 Indian Ocean tsunami. *Earthquake Spectra*, *22*, 263–283. https://doi.org/10.1193/1.2202646

Pattiaratchi, C. B., & Wijeratne, E. M. S. (2009). Tide gauge observations of 2004–2007 Indian Ocean tsunamis from Sri Lanka and Western Australia. *Pure and Applied Geophysics*, *166*(1–2), 233–258. https://doi.org/10.1007/s00024-008-0434-5

Rost, S., & Thomas, C. (2002). Array seismology: Methods and applications. *Reviews of Geophysics*, *40*(3), 1008. https://doi.org/10.1029/2000RG000100

Saito, T., & Furumura, T. (2009). Scattering of linear long-wave tsunamis due to randomly fluctuating sea-bottom topography: Coda excitation and scattering attenuation. *Geophysical Journal International*, *177*, 958–965. https://doi.org/10.1111/j.1365-246X.2009.03988.x

Sandanbata, O., Watada, S., Satake, K., Fukao, Y., Sugioka, H., Ito, A., & Shiobara, H. (2018). Ray tracing for dispersive tsunamis and source amplitude estimation based on Green's law: Application to the 2015 volcanic tsunami earthquake near Torishima, south of Japan. *Pure and Applied Geophysics*, *175*(4), 1371–1385. https://doi.org/10.1007/s00024-017-1746-0

Schmidt, R. (1986). Multiple emitter location and signal parameter estimation. In *IEEE transactions on antennas and propagation* (Vol. 34, pp. 276–280). IEEE. https://doi.org/10.1109/TAP.1986.1143830

Sheehan, A. F., Gusman, A. R., & Satake, K. (2019). Improving forecast accuracy with tsunami data assimilation: The 2009 Dusky Sound, New Zealand, tsunami. *Journal of Geophysical Research: Solid Earth*, *124*, 566–577. https://doi.org/10.1029/2018JB016575

Shi, J., Kohler, M. D., Sutton, J. N., & Ampuero, J. P. (2017). Mapping coherent, time-varying wavefronts from the Tohoku tsunami into enhanced, time-dependent warning messages. in: *Proceedings of 16th* World Conference on Earthquake Engineering, 4587, at https://www.wcee.nicee.org/wcee/article/16WCEE/WCEE2017-4587.pdf

Suppasri, A., Leelawat, N., Latcharote, P., Roeber, V., Yamashita, K., Hayashi, A., et al. (2017). The 2016 Fukushima earthquake and tsunami: Local tsunami behavior and recommendations for tsunami disaster risk reduction. *International Journal of Disaster Risk Reduction*, *21*, 323–330. https://doi.org/10.1016/j.ijdrr.2016.12.016

Tang, L., Titov, V. V., Bernard, E. N., Wei, Y., Chamberlin, C. D., Newman, J. C., et al. (2012). Direct energy estimation of the 2011 Japan tsunami using deep-ocean pressure measurements. *Journal of Geophysical Research*, *117*, C08008. https://doi.org/10.1029/2011JC007635

Tarr, A. C., Villaseñor, A., Furlong, K. P., Rhea, S. & Benz, H. M. (2010). Seismicity of the Earth 1900–2007: *U.S. Geological Survey Scientific Investigations Map 3064*, 1 sheet, scale 1:25,000,000.

Thomson, D. J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, *70*, 1055–1096.

Thomson, R., Fine, I., Rabinovich, A., Mihály, S., Davis, E., Heesemann, M., & Krassovski, M. (2011). Observation of the 2009 Samoa tsunami by the NEPTUNE-Canada cabled observatory: Test data for an operational regional tsunami forecast model. *Geophysical Research Letters*, *38*, L11701. https://doi.org/10.1029/2011GL046728

Tsai, V. C., Ampuero, J. P., Kanamori, H., & Stevenson, D. J. (2013). Estimating the effect of Earth elasticity and variable water density on tsunami speeds. *Geophysical Research Letters*, *40*, 492–496. https://doi.org/10.1002/grl.50147

Wang, Y., Satake, K., Sandanbata, O., Maeda, T., & Su, H. (2019). Tsunami data assimilation of cabled ocean bottom pressure records for the 2015 Torishima volcanic tsunami earthquake. *Journal of Geophysical Research: Solid Earth*, *124*, 10,413–10,422. https://doi.org/10.1029/2019JB018056

Zhang, J., Gerstoft, P., & Shearer, P. M. (2009). High-frequency P-wave seismic noise driven by ocean winds. *Geophysical Research Letters*, *36*, L09302. https://doi.org/10.1029/2009GL037761