OBSERVING AND MODELING ELASTIC SCATTERING IN THE DEEP EARTH

Peter M. Shearer and Paul S. Earle

Abstract

Seismic scattering in the deep Earth below the mantle transition zone is observed from precursors and codas to a number of body-wave arrivals, including P, P_{diff}, PKP, $PKiKP$, $PKKP$, and $P'P'$. Envelope-stacking methods applied to large teleseismic databases are useful for resolving the globally averaged time and amplitude dependence of these arrivals. Stacks of P coda near 1 Hz from shallow earthquakes exhibit significant variations among different source and receiver locations, indicating lateral variations in scattering strength. At least some deep-mantle, core-mantle boundary, and inner-core scattering is indicated by the observations, but the strength and scale length of the random velocity heterogeneity required to explain the data are not yet firmly established. Monte Carlo seismic "particle" algorithms, based on numerical evaluation of radiative transfer theory with Born scattering amplitudes for random elastic heterogeneity, provide a powerful tool for computer modeling of scattering in the whole Earth because they preserve energy and can handle multiple scattering through depth-varying heterogeneity models. Efficient implementation of these algorithms can be achieved by precomputing ray tracing tables and discretized scattering probability functions.

© 2008 Elsevier Inc.

1. Introduction

Observing and modeling seismic scattering in the mantle and core is important because of the constraints these studies provide on small-scale heterogeneity. However, investigating seismic scattering in the deep Earth is challenging because strong lithospheric scattering can mask scattered arrivals from deeper in the mantle. By correctly identifying the scattering origin of PKP precursors, Cleary and Haddon (1972) found the first definitive evidence for deep-Earth scattering. PKP precursors have an unusual ray geometry that provides a unique window into scattering within the lowermost mantle and at the core-mantle boundary (CMB). Early modeling of PKP precursors focused on the (CMB) region as their likely source and used single-scattering theory applied to random media models to provide a first-order fit to precursor amplitudes (e.g., Haddon and Cleary, 1974; Doornbos, 1978; Bataille and Flatté, 1988). However, more recent work (Hedlin *et al.*, 1997; Cormier, 1999; Margerin and Nolet, 2003a,b) showed that the scattering must extend at least 600 km into the mantle above the core, and it seems likely that some amount of scattering is present throughout the mantle.

In addition to PKP precursors, there are a number of other seismic observations that suggest deep scattering, including P_{diff} coda (Bataille *et al.*, 1990; Tono and Yomogida, 1996; Bataille and Lund, 1996; Earle and Shearer, 2001). $PKKP$ precursors (Chang and Cleary, 1978, 1981; Doornbos, 1980; Earle and Shearer, 1997), and $PKiKP$ coda (Vidale and Earle, 2000; Vidale *et al.*, 2000; Poupinet and Kennett, 2004; Koper *et al.*, 2004). In principle, the coda of deep-turning P and S waves is sensitive to lower-mantle scattering,

but it is tricky to separate this scattering from the much stronger scattering that occurs in the shallow mantle and crust. Nonetheless, several recent studies have found evidence in these phases for a deep-scattering contribution (Cormier, 2000; Lee *et al*., 2003; Shearer and Earle, 2004).

All of these results are valuable because they provide estimates on the strength of heterogeneity in the deep mantle and inner core at scale lengths (e.g., ∼10 km) much smaller than those that can be imaged using tomographic methods. These velocity anomalies are almost certainly compositional in origin because small-scale thermal perturbations would quickly diffuse away, and thus they provide insight regarding the degree of mixing in mantle convection models. It is therefore important to further develop seismic observations to resolve additional details regarding the heterogeneity, including its strength, scale length, and depth dependence.

Accurate modeling of the seismic observations requires a more complete theory than methods based on the Born approximation, which do not conserve energy and ignore the effects of multiple scattering. For example, Margerin and Nolet (2003a,b) found that Born theory is accurate for whole-mantle scattering models only when the deep-velocity heterogeneity is less than 0.5%. Finite difference/element methods can handle velocity models of arbitrary complexity but are not yet numerically feasible for global scattering problems at high frequencies (1 Hz). Faster algorithms are possible through use of the parabolic and Markov approximations (e.g., Sato and Fehler, 2006), but three-dimensional global calculations remain difficult. Here, we focus on Monte Carlo methods based on radiative transfer theory that simulate the random walk of millions of seismic energy "particles." Although these methods discard phase information, they are a powerful and practical approach to modeling whole-Earth, high-frequency scattering.

We begin by describing the processing and stacking methods that are suited for global seismic observations and then present some specifics regarding how the Monte Carlo method can be efficiently implemented. Results from the Shearer and Earle (2004) analysis of P coda will be highlighted, but we describe our algorithms in more detail and present some new results concerning lateral variability in teleseismic P coda.

2. Data Stacking

Waveform stacking has several advantages over analysis of individual seismograms:

1. It generally increases the signal-to-noise ratio, making it possible to identify and characterize weak seismic arrivals that are hard to resolve on single records.
2. It reduces the volume of data to be modeled to a more manageable level. For example, the information in thousands of global seismograms can be reduced to a single time-versus-epicentral-distance image of the average wave field.
3. It can provide a spatially averaged measure of the wave field that is less biased than results from small numbers of seismograms. By processing all of the data, it reduces the selection bias problem that may affect studies that focus on the most visible or anomalous phases in individual records.

Conventional seismogram-stacking methods do not work well for imaging scattered seismic arrivals because coda waves are generally incoherent among the different recording stations, especially at the high frequencies where the scattered wave field is typically observed. In other words, the timing of the peaks and troughs in the

seismograms varies randomly among the records. Small-scale arrays are an important exception to this and have provided valuable constraints on the slowness and back-azimuth of scattered arrivals. However, our focus in this chapter concerns the use of single stations that are too far apart for standard array processing techniques to work. In this case, it is necessary to develop stacking methods that work for incoherent data.

One approach is to discard the phase information in the seismograms and consider only their energy content as defined by their envelope functions. This method has been used successfully to image *PKP* and *PKKP* precursors (Hedlin *et al*., 1997; Earle and Shearer, 1997; Shearer *et al*., 1998), P_{diff} coda (Earle and Shearer, 2001), and P coda (Shearer and Earle, 2004).

Figure 1 illustrates the envelope-function stacking technique applied to P and its coda (Shearer and Earle, 2004). Before stacking, we manually review the data and remove seismograms with dropouts, data glitches, or contaminating arrivals from aftershocks or local earthquakes. Once the data are cleaned, they are processed and stacked using the following steps shown in Fig. 1:
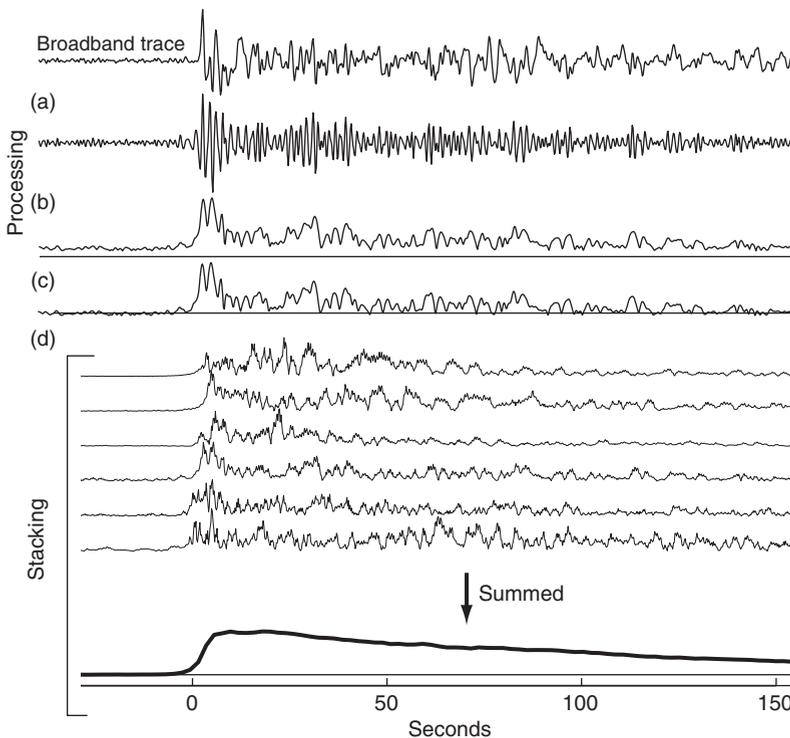

FIG. 1. Illustration of the different processing steps in the envelope-stacking technique. Starting with each original broadband trace (a) band-pass filtering is applied, (b) the envelope function is computed, (c) the power in the preevent noise is removed. The processed traces are then stacked (d). Notice the large variability seen in the individual seismograms compared to the smooth stacked trace at the bottom.

(a) The traces are band-pass filtered between 0.5 Hz and 2.5 Hz. This frequency band falls in a low noise region of the spectrum and provides the greatest sensitivity to deep-Earth structure with scale length of about 10 km.

(b) The envelope function (e.g., Kanasewich, 1981) is calculated for each seismogram.

(c) We assume that the noise and signal are uncorrelated so that, upon averaging, their energies will sum. Thus, to account for varying noise levels between traces, the envelope functions are squared and the average noise in a time window preceding the reference phase (in this example the P wave) is subtracted from the entire trace. Then, the square root of the noise corrected trace is taken. The squared envelope is used because the recorded signal is the square root of the sum of the squared noise plus the squared signal.

(d) The final stack is made by normalizing the traces to their maximum amplitudes, aligning them on the reference phase arrival time, and averaging all traces in a target distance window as a function of time.

We stack the traces in amplitude rather than power because we have found that this produces slightly more robust results than power stacks. However, in practice the differences between amplitude and power stacks are usually fairly small.

2.1. Shallow- Versus Deep-Earthquake Teleseismic P Coda

Results of this stacking method applied to teleseismic P coda are plotted in Fig. 2 (originally in Shearer and Earle, 2004), which compares stacks for over 7500 records from shallow earthquakes (depth $\leq$ 50 km) and 650 records from deep earthquakes (depth $\geq$ 400 km). Data are taken from vertical- component seismograms from $M_W = 6$ to 7 events in the IRIS FARM archive from 1990 to 1999. The stacked envelopes are binned at 5° distance intervals and 2-s time intervals. This figure shows the striking difference in teleseismic coda strength between shallow and deep earthquakes. The shallow-event coda is much stronger and longer-lasting than the deep-event coda. At 50 s following the P arrival, it is 2–5 times larger in amplitude (4–25 times larger in energy). This difference indicates that teleseismic P coda from shallow events is dominated by near-source scattering above 600-km depth. Note that both stacks should have equal coda contributions from near-receiver scattering, but the energy difference between the shallow and the deep coda is much more than a factor of two. However, as discussed in Shearer and Earle (2004), this does not necessarily imply lateral variations in scattering strength with stronger scattering in active earthquake areas. In fact, with Monte Carlo modeling, it is possible to achieve a reasonable fit to both the shallow- and deep-earthquake coda amplitudes with a single model in which scattering strength varies only with depth.

2.2. Regional Variations in Teleseismic P Coda Amplitude

The Shearer and Earle (2004) teleseismic P coda study considered only spherically averaged coda amplitudes. Here, we analyze data from this study in more detail to identify variations in coda levels among different source regions and station locations. Fig 3 illustrates such a difference between two Asian stations as seen in coda stacks of 120 quakes recorded at YAK and 163 quakes recorded by AAK (all quake depths are less than 50 km). Station AAK has consistently higher P coda amplitudes than station YAK.
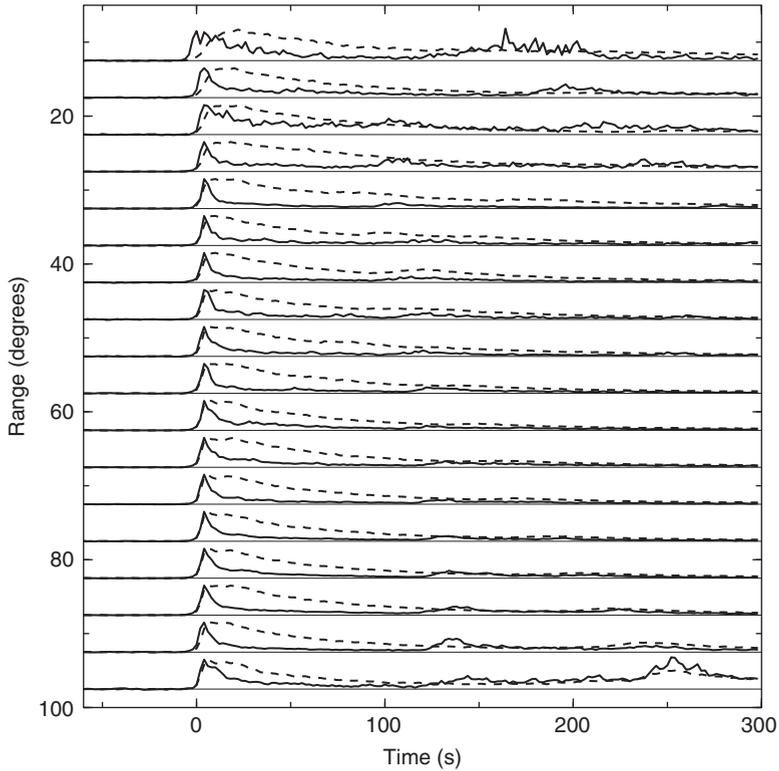
Fig. 2. A comparison between the shallow-event, envelope-function stack (dashed), and the deep-event stack (solid line). Time is relative to P and the stacks have been scaled to the same P-wave maximum amplitude. Note the much more extended coda from the shallow events. Figure taken from Shearer and Earle (2004).

To characterize these variations systematically, we measure the average ratio of the P coda amplitude to the P amplitude at source-receiver distances between 60° and 95° (the observed ratio changes little over this distance interval). The P amplitude is measured as the average absolute value in the demeaned trace in a 40-s window starting at the predicted P arrival time. The coda amplitude is measured similarly in a 60-s window starting 60 s after the P arrival time. To account for possible biasing effects related to the specific subset of events recorded by each station, we assume that the logarithm of this relative coda level, c, for the ith source and the jth receiver can be approximated as the sum of a source term, q, and a receiver term, r,

$$\log(c_{ij}) = \log(q_i) + \log(r_j). \tag{1}$$

This equation does not have a direct physical basis; it is an empirical approach to test how much of the variation in the coda amplitudes can be explained with a simple decomposition into source- and receiver-side contributions. Because we have many receivers for
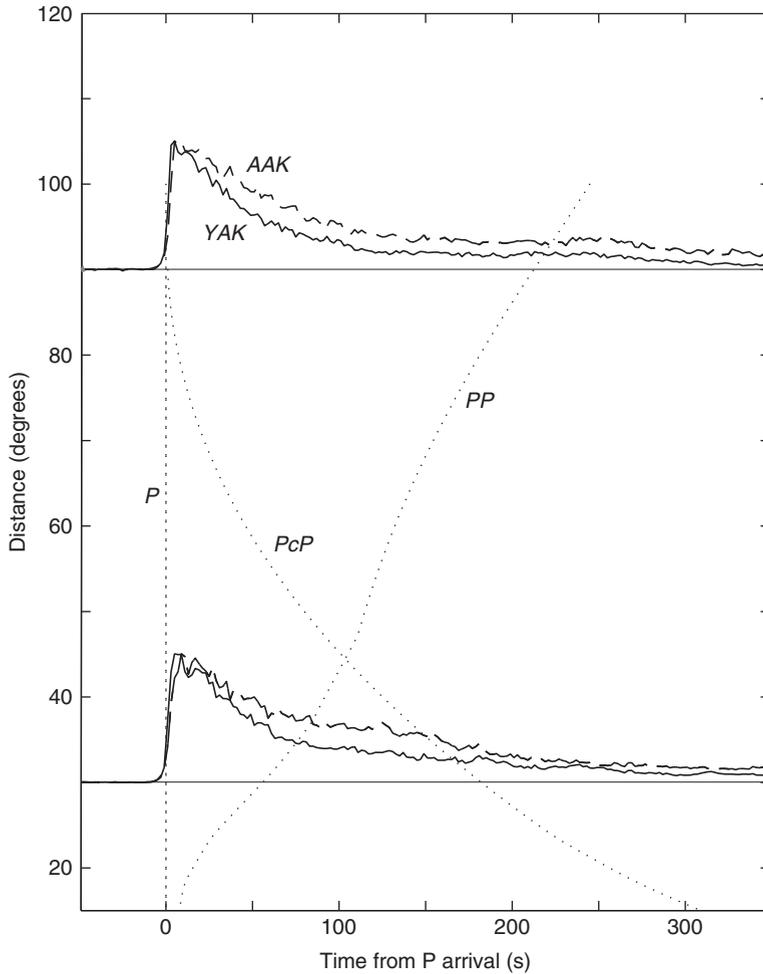
FIG. 3. A comparison of stacked P coda amplitudes for shallow earthquakes (≤50 km) between stations AAK (dashed) and YAK (solid) at 30 and 90. Note the higher coda amplitudes for AAK.

each source and many sources for each receiver, this is an over-determined problem, which we solve using a robust, iterative least squares approach. To remove the non-uniqueness in this equation (a constant could be added to the log q terms and subtracted from the log r terms), we constrain the average log r to be zero.

Fig. 4 plots the resulting individual source and receiver terms q and r. Before plotting, we scale the receiver terms to have the same median value as the source terms. Symbol size is proportional to the log deviation from the median amplitude ratio of 0.432. Plus symbols have relatively high coda levels and diamonds have relatively low coda levels. Because results can be quite variable for small numbers of traces, we only include terms constrained by at least 10 traces. Overall, the quake coda levels have about twice the
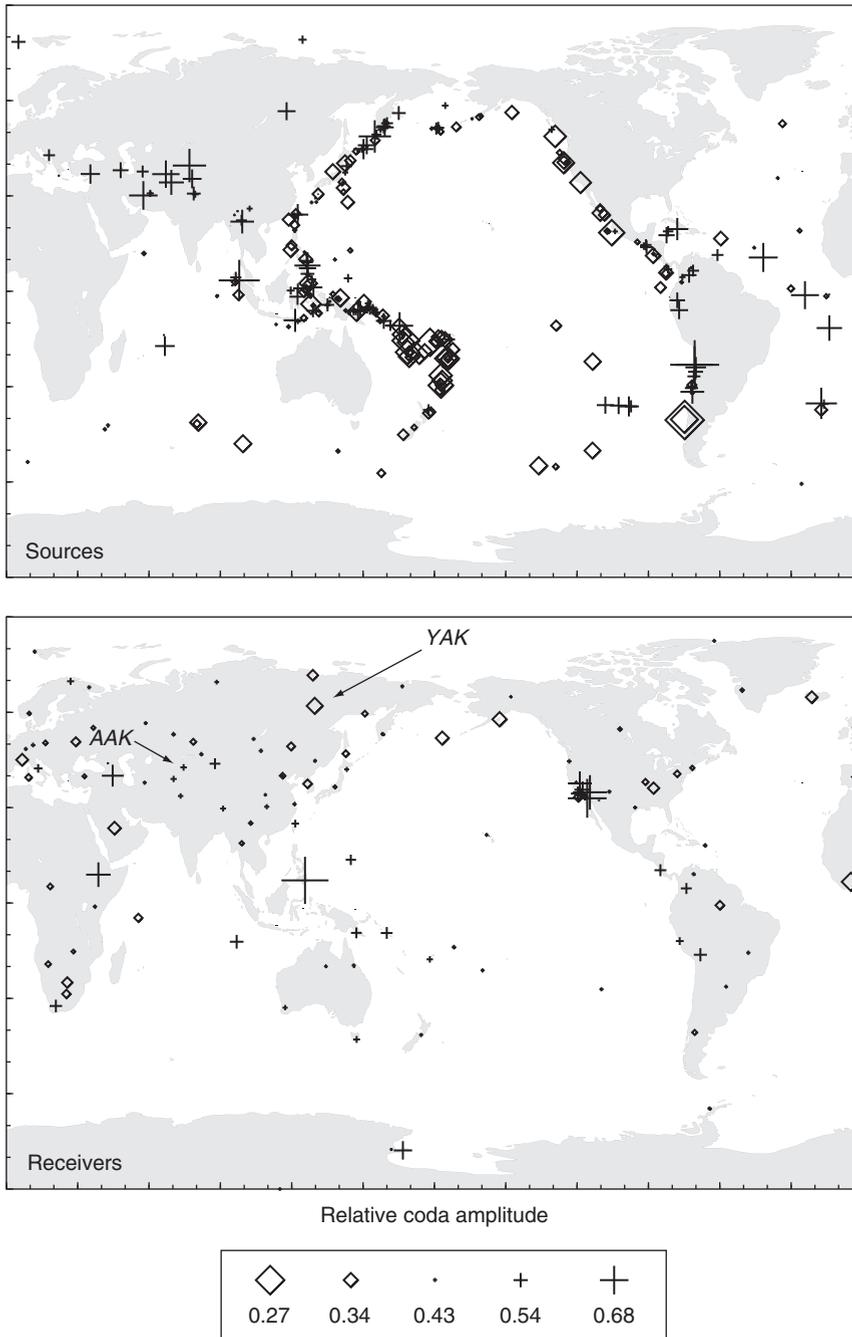
FIG. 4. Differences in teleseismic P coda amplitudes at source locations (top) and station locations (bottom). Coda levels higher than the global median of 0.43 are plotted as + symbols and weaker levels are shown as diamonds, with the size of the symbol scaling as the logarithm of the perturbation. Stations AAK and YAK (see Fig. 3) are labeled in the bottom figure.

variability of the station coda levels. The median variations in quake coda levels are about $\pm 16\%$ while the median variations in station coda levels are about $\pm 8\%$. This is consistent with near-source rather than near-receiver scattering being the dominant contributor to teleseismic P coda, as is also implied by the differences in the shallow- and deep-earthquake stacks plotted in Fig. 2. The relatively sharp changes in coda levels that can be seen in these plots over short distance intervals also indicates that the bulk of the coda originates from near-surface scattering, rather than from deeper in the mantle.

Both the quake and the station terms exhibit some degree of spatial coherence. Stronger coda is seen from events in central Asia, the Kurils and South America and weaker coda is observed in Japan, western North America, and the southwest Pacific. The station terms are less coherent but generally have weaker coda in northeast Asia, parts of Africa, and North America outside of California. There is little, if any, spatial correlation between the source and the receiver terms. This may indicate that the near-source and near-receiver scattering processes are different (i.e., more S-to-P near the source, more P-to-P near the stations), or could reflect strong variations in heterogeneity very close to earthquake source regions that do not always extend far enough to produce correlated variations in coda strength for teleseismic arrivals at nearby seismic stations.

Our station term results have only limited correlation with the scattering Q (Q_{Sc}) estimates at 1 Hz of Korn (1990, 1993) from P coda for 9 stations in Australia and around the Pacific. We observe stronger than average P coda for stations GUMO and TATO, which have lower than average Q_{Sc} at 1 Hz in Korn (1993). In contrast, we also observe relatively strong P coda for NWAO and weak P coda for AFI, opposite to Korn's results. Because of differences between our simple station term inversion method and the energy flux modeling approach of Korn (1990, 1993), it is not clear how well correlated results should be between the methods.

3. Monte Carlo Methods

Monte Carlo methods have been used in physics since the 1950s to model radiation transport by using a computer to simulate the random scattering of large numbers of individual particles [see Dupree and Fraley (2002) for a recent introduction to many of these techniques]. The Monte Carlo approach uses computer-generated random numbers to sample the different possible variables in a problem. For example, neutron scattering can be simulated by tracking the behavior of individual neutrons, radiated in random directions from a source and scattered in random directions during their propagation, thus in effect simulating the results of an actual experiment. In general, the accuracy of the solution grows with the number of particle trajectories that are computed and thus Monte Carlo methods have become increasingly useful as faster computers have become available. Typically, the algorithms converge such that the variance of the results decreases as $1/\sqrt{n}$, where n is the number of particles.

The concept of seismic "particles" may not seem useful upon initial consideration because there is no wave-particle duality for seismic waves, as exists for electromagnetic waves. However, if one is willing to consider energy transport alone and discard phase information in seismic records, then a particle-based, Monte Carlo approach can be very valuable. It is particularly suited to studying scattering at high frequencies, where the waveforms are incoherent and typical modeling efforts consider only the envelope

function. It bridges the gap between Born theory for weak scattering and computationally intensive finite difference/element calculations for complicated models.

3.1. Seismology Applications

The first use of the Monte Carlo approach in seismology was by Gusev and Abubakirov (1987) who modeled acoustic wave scattering in a uniform whole space using particles randomly radiated from an isotropic source. They assumed a constant probability of scattering per unit volume, resulting in an exponential distribution of path lengths. They did not explicitly include intrinsic attenuation but noted that it could easily be modeled in the constant Q case by multiplying the energy of each particle by $e^{-2\pi ft/Q}$. They considered both isotropic scattering and forward scattering with a Gaussian angle distribution and showed that their results agree with the diffusion model at large lapse times, but that only the forward scattering model produces realistic pulse broadening and coda envelopes at short distances. Abubakirov and Gusev (1990) presented a more detailed account of the Monte Carlo technique and used a forward scattering model to compute master curves describing the relationships between the mean free path and both pulse broadening and the intensity ratio of the direct and the scattered waves. Applying these results to S-coda observations in Kamchatka, they obtained S-wave mean free paths of 100–150 km over a 1.5–6 Hz frequency range. Gusev and Abubakirov (1996) expanded their Monte Carlo method to include scattering angles predicted by specific models of random velocity heterogeneity, including Gaussian and power-law media, and argued that a power-law exponent of 3.5–4 is in qualitative agreement with the features of observed S-wave envelopes.

Hoshiba (1991) used a Monte Carlo method to model isotropic S-wave scattering in a uniform whole space, and demonstrated that the results agree with Born theory for weak scattering and with the radiative transfer theory of Wu (1985) and the diffusion model for strong scattering. Hoshiba (1994, 1997) extended his method to include depth-dependent scattering strength, layered velocity models, and intrinsic attenuation. He simulated SH-wave reflection and transmission coefficients at layer interfaces as probabilities of reflection or transmission, which is a practical way to handle the energy partitioning at interfaces without the complexities of generating additional particles.

Margerin *et al*. (1998) modeled isotropic S-wave scattering using a Monte Carlo method for a layer over half-space model (i.e., the crust and the upper mantle) and included both surface and crust-mantle boundary (Moho) reflected/transmitted phases, using probabilities to handle reflection and transmission coefficients, but did not model S polarity, phase conversions, and intrinsic attenuation. Margerin *et al*. (2000) extended their method to fully elastic waves, including S polarity and phase conversions, and modeled scattering off randomly distributed spherical inclusions within a uniform whole space. They explored the dependence of their results of the relative size of the spheres compared to the seismic wavelengths. Margerin and Nolet (2003a,b) applied a Monte Carlo method to model *PKP* precursors with whole Earth P-to-P scattering in the mantle. They showed that their results were in good agreement with geometrical ray theory for the main *PKP* arrivals and that the scattered arrivals agreed with Born theory for weak random velocity heterogeneity.

Bal and Moscoso (2000) included S-wave polarizations in Monte Carlo simulations of randomly heterogeneous lithosphere and showed that S waves can become depolarized after multiple scattering. Yoshimoto (2000) used a finite difference ray tracing method to

implement a Monte Carlo method for a complex velocity profile for the lithosphere and found that ray bending caused by velocity gradients and the Moho can have large effects on the shape of the S coda envelope.

Wegler *et al*. (2006) and Przybilla *et al*. (2006) performed a series of tests of Monte Carlo simulations based on radiative transfer theory, for both acoustic and fully elastic P- and S scattering, and compared their results to those predicted by various analytical solutions in 3D and finite difference solutions to the full wave equation in 3D. In general, they found good agreement between the Monte Carlo approach and other methods, except in the case of extreme velocity perturbations, such as can occur in volcano seismology.

Although all of these methods work by computing trajectories for a large number of particles, they differ in some important details. A fundamental distinction can be made between two different approaches: (1) algorithms that simply count the number of particles that hit different cells in the model (e.g., Gusev and Abubakirov, 1987, 1996; Yoshimoto, 2000; Shearer and Earle, 2004) and (2) those that compute the probability of particles at a series of discrete receivers (e.g., Hoshiba, 1991, 1994, 1997; Margerin *et al*., 1998, 2000). The former provide the energy density at every point in the model and thus can be termed global methods in contrast to the local nature of the calculations in (2). Although for some models there are computational advantages to (2), the simplicity and flexibility of (1) have made it a more popular choice, given the speed and storage capabilities of modern computers.

3.2. Monte Carlo Implementation

The theoretical basis for the Monte Carlo approach is provided by radiative transfer theory (e.g., Wu, 1985; Sato, 1995; Ryshik *et al*., 1996; Bal and Moscoso, 2000; Margerin, 2005). We will not review this theory here. Instead, we will focus on the practical aspects of writing a computer program to perform seismic Monte Carlo calculations in an efficient manner for radially symmetric Earth models. As an example, we will give details of the global elastic algorithm of Shearer and Earle (2004).

The fundamental principle involved is that each seismic "particle" represents an energy packet and that our treatment of the particles (i.e., propagation, reflection/ transmission, phase conversions, and scattering) should be designed to conserve energy in a logically consistent manner. We will use geometrical ray theory to compute particle trajectories. A nice aspect of the particle approach is that geometrical spreading terms are not required because the energy reduction with distance is naturally included as the decrease in particle density as the particles spread out from the source. The energy partitioning that occurs at interfaces or scattering points is handled not by splitting the energy into different particles but by assigning appropriate probabilities to the changes in particle directions and using computer-generated random numbers to sample these probability distributions. Thus, we track only one particle at a time, making the code straightforward to parallelize and run on multiple processors, if desired.

The output of a Monte Carlo simulation will typically appear noisy at first (i.e., producing spiky and irregular envelopes) when only a small number of particles are computed, but will become increasingly smooth as more particles are included in the calculation. The number of particles required to give adequate results will vary, depending upon the details of the model, the portion of the output wave field that is of greatest interest, and how much resolution in time and space is desired.

3.3. The Monte Carlo Source

Despite the fact that earthquake radiation patterns are not uniform, most Monte Carlo simulations assume isotropic radiation from the source. This can be justified for two reasons: (1) At high frequencies, observed P and S amplitudes show considerable scatter compared to that predicted by double-couple sources (e.g., Nakamura *et al.*, 1999; Hardebeck and Shearer, 2003) presumably caused by strong crustal and near-surface focusing and scattering effects. Thus, details of the radiation pattern tend to be lost during high-frequency wave propagation. (2) Results are often averaged over many earthquakes and stations at different azimuths from the source. This will lessen the bias caused for individual ray paths by neglecting radiation pattern effects. However, it is important to recognize that some bias may still be present. For example, if most earthquakes in a region are strike-slip, the expected P-wave fraction of energy radiated in the near-vertical direction is much less than that predicted by assuming an isotropic source. This bias cannot be removed by averaging over azimuth. Isotropic average radiation will only occur for a truly random distribution of focal mechanisms, which is unlikely to be the case for real Earth observations.

If the background Earth model is radially symmetric, then the expected average energy observed on Earth's surface from an isotropic source will vary only as a function of distance. It thus makes sense to combine the energy from all of the particles hitting the surface at a particular distance (regardless of azimuth) to compute the average predicted wave field. It follows that because of the symmetry of the problem and the randomness of individual scattering events, it is only necessary to shoot the particles at a single azimuth from the source. In this case, however, the number or the energy of the particles must be weighted as $\sin \theta$, where θ is the takeoff angle from the vertical, to account for the greater number of particles expected at more horizontal takeoff angles for a spherically isotropic source.

For efficiency it is usually desirable for ray tracing and other information to be computed and stored at certain discrete values, in which case the sampling will be limited to these values. It is not necessary for the takeoff angle sampling to be uniform, provided suitable weights are assigned to the particles. This technique is termed event biasing in the Monte Carlo literature. For example, the rays could be evenly sampled in ray parameter rather than angle or proportionally more rays could be fired at steeper angles for better sampling of core phases compared to crustal phases. Both upgoing and downgoing rays from the source should be included unless the source is assumed to be exactly at the surface. In Shearer and Earle (2004), we spray rays evenly spaced in 10,000 values of ray parameter (from $p = 0$ to $p = 1/c$, where c is the P or S velocity at the source), and set the energy of the ith particle proportional to $(\theta_i - \theta_{i-1}) \sin \theta_i$, where the takeoff angle $\theta_i = \sin^{-1}(cp_i)$. Note that $c\, dp = \cos \theta\, d\theta$, $\theta_i - \theta_{i-1} \approx d\theta_i$, and thus $(\theta_i - \theta_{i-1}) \sin \theta_i$ is proportional to $\tan \theta$ when dp is constant.

For S waves, it is simplest to assume random polarizations for the particles leaving the source, again making the assumption that radiation pattern differences will tend to average out when results are combined from many different sources and receivers. In addition, multiple scattering will at some point remove the information about the original source polarization. In complete modeling of both P and S waves, an S-to-P total initial energy scaling factor, $q = E_i^S/E_i^P$, must be assumed. This can be done by radiating q times more S particles or by radiating equal numbers of P and S particles but assigning q times more energy to the S particles. The latter approach is more efficient for resolving

both P and S phases. Theoretical results for a double-couple source suggest that $q = 23.4$ for a Poisson solid (e.g., Sato, 1984) and this factor was used by Shearer and Earle (2004). However, this assumes that the S and P corner frequencies are identical and earthquake source studies indicate that the P-wave corner frequency is typically higher than the S-wave corner frequency, with observations ranging from $q = 9$ to 14 (Boatwright and Fletcher, 1984; Abercrombie, 1995; Prieto *et al*., 2004). These values are for total radiated energy integrated over the entire frequency band and are not necessarily appropriate at the single fixed frequency used in each Monte Carlo calculation. The question of the relative sizes of P and S radiation from earthquakes is an active area of research to which scattering studies may be able to contribute by better separation of intrinsic attenuation, scattering attenuation, and source effects in observed earthquake spectra.

Because the number of radiated particles will vary depending upon how long the program is kept running, it is simplest to initially consider only relative energy at the source and perform the normalization to absolute energy at a later stage. This may be accomplished by keeping track of the total initial energy of the radiated particles and multiplying the observed energy by the ratio of the desired radiated energy to the total initial particle energy. In many applications, only the relative time versus distance behavior of the wave field is important, in which case the calibration to absolute amplitude is not required.

3.4. Particle Trajectories

Most seismic applications of the Monte Carlo approach have assumed acoustic or elastic body-wave propagation and used ray theory to track the particle trajectories. For simple whole-space or homogeneous layer models, the ray paths are straight lines. However, for more realistic models containing velocity gradients the ray paths are curved. To save computer time when computing results for millions of particles, it is advantageous in this case to precompute ray tracing results (time and distance) at discrete values of ray parameter, p, and depth in the model. The Shearer and Earle (2004) algorithm computed dt and dx values within 10-km-thick layers in the model for 10,000 values of p, saving the results in separate arrays for P and S waves. Another array records whether the ray passes through, reflects off, or turns (changes direction) within each layer. Of course, a specific velocity versus depth model must be assumed. Standard models such as PREM (Dziewonski and Anderson, 1981) and IASP91 (Kennett, 1991) predict body-wave travel times that generally agree within a few seconds, but differences in the velocity gradients among the models can produce significant differences in ray theoretical amplitudes.

Modeling of S-wave polarizations is complicated by the fact that the polarization will rotate along curved ray paths. In radiative transfer theory, S-wave polarizations can be handled using the Stokes parameters and this is the approach described in Margerin *et al*. (2000) and Bal and Moscoso (2000). Shearer and Earle (2004) adopt the simpler scheme of assigning S polarity as an angle in a local SV versus SH coordinate system, an angle that will remain constant along curved ray paths in radially symmetric models, although it will of course change following reflection/transmission or scattering events. This approach is more limited than the Stokes method because it assumes that the S polarization is always linear, whereas phase shifts between the SV and SH components can occur for some reflections at interfaces. Thus, the Shearer and Earle (2004) algorithm should be considered only approximate for S polarizations.
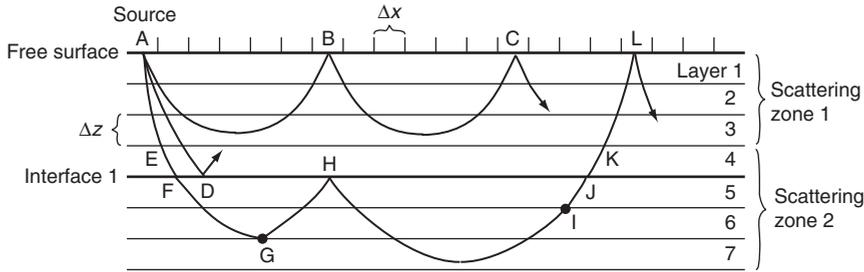
FIG. 5. Example particle trajectories for a Monte Carlo simulation method. The one-dimensional background velocity model is defined as velocity points at a depth spacing of $\triangle z$. In this example, there is a velocity jump between layers 4 and 5 and the mean free path is constant within 2 different scattering zones. See text for a description of the particle behavior.

Figure 5 shows some examples of particle trajectories. Here, the model is discretized at depth intervals of $\triangle z$. Consider the simple case of ray path ABC, which does not experience any scattering. The ray is radiated downward from the source at a specific ray parameter. The program marches downward through the layers, adding the precomputed time and distance increments for each ray segment. In layer 3, the ray experiences a turning point and its vertical direction changes. The program then goes up through the layers and continues to accumulate time and distance increments. The direction changes again when the ray reflects off the free surface at B.

The free surface is important because it is normally where we output results of Monte Carlo simulations to compare with observations. For isotropic sources and Earth models that are radially symmetric in their bulk properties (i.e., excluding random small-scale perturbations), the observed wave field is a function of time and epicentral distance only. It is therefore convenient to discretize the output into small bins in a time versus distance array (the distance increment $\triangle x$ is shown in Fig. 5), into which the accumulated energy from each arriving seismic particle is summed. At this point, separate results can be saved for P- and S energy (derived from the wave type) as well as the vertical, radial, and tangental components (derived from the local ray angle, wave type, and S polarization). Later, these energies are normalized by the surface area in each bin (e.g., to account for the greater surface area between $90°$ and $91°$ from the source compared to between $10°$ and $11°$ from the source). In fully elastic calculations, P-to-S and S-to-P conversions occur at the free surface and must be included in the modeling (see Section 3.4.1 below).

Following its reflection at point B, the ray again travels downward until it turns and reflects again at surface point C, where the ray energy is added to a different part of the time–distance array. The calculation for each particle continues until a maximum time limit is reached, at which point the algorithm starts over with a new particle from the source.

3.4.1. Interfaces

All standard Earth models contain significant velocity changes at the surface, the Moho, the CMB, and the inner-core boundary (ICB) and minor velocity jumps near 410- and 660-km depth. Thus, it is important for Monte Carlo simulations to correctly model

the reflection and the transmission behavior at interfaces. As described in Hoshiba (1997) and Margerin and Nolet (2003a), this is handled by computing energy-normalized reflection and transmission coefficients and converting them into probabilities that are used to pick (based on a computer-generated random number) the wave type (P or S), direction (up or down), and polarization (in the case of S waves) for a single particle that leaves the interface. In this way the energy partitioning at each interface is modeled stochastically as the average response of thousands of individual particles. Because these are spherical interfaces, Snell's law is obeyed and the ray parameter does not change.

In Fig. 5, there are two interfaces, one at the free surface and one between layers 4 and 5, where the velocity jumps discontinuously. For program efficiency, the reflection and the transmission coefficients are precomputed for all of the discrete values of ray parameter used to calculate the ray paths. Assuming the ray path ABC represents a P wave, there is some energy converted to SV upon each free surface reflection. Thus, there was a random chance that the particle might have changed to an S wave (with the probability determined by the energy normalized reflection coefficient), but in this example we assume that it did not. Ray path AD hits interface 2 and is reflected. However, note that other particles traveling along the exact same path may be transmitted though the interface. Whether a particle is reflected or transmitted will depend upon the value of a random number generated by the computer whenever the particle hits an interface.

The interface energy partitioning described above will result in the majority of particles going into the seismic phases with the largest reflection and transmission coefficients and thus into the highest energy parts of the wave field. This may not always be desirable if a target phase of interest is of relatively low amplitude because of a small reflection coefficient along its ray path, in which case most of the particles are "wasted" and comparatively few particles will illuminate the phase. Examples of such phases include *PcP, PKiKP*, and *PKKP* along much of their travel time curves. To improve the performance of Monte Carlo algorithms in these cases, the appropriate reflection coeffi-cients can be artificially increased, provided that the energy of the reflected particles is decreased and the energy of transmitted particles is increased, such that average energy over many particles is preserved even if energy conservation is violated for individual particles. This is an example of a particle biasing technique, which is a common approach in many Monte Carlo analyses in physics, although to our knowledge it has not yet been applied in seismology.

3.4.2. Scattering Events

Scattering strength may be described either in terms of the probability of scattering as the particle passes through a given volume or as the mean free path between scattering events. The scattering coefficient, g, is defined as the scattering power per unit volume per unit solid angle (e.g., Sato, 1977) and has units of reciprocal length. The total scattering coefficient, g_0, is defined as the average of g over all directions and can also be expressed as

$$g_0 = \ell^{-1}, \tag{2}$$

where ℓ is the mean free path. In general, these parameters will vary with depth in the Earth for physically based random heterogeneity models because they are dependent on

the seismic wave number k. Thus, one approach would be to test for a scattering event by generating random numbers at short intervals along the ray path. This is the best method in some respects because it can accurately handle depth-dependent scattering of arbitrary complexity and is straightforward to code. However, it requires generation of a large number of random numbers along the ray paths. Thus, it is more efficient to approximate the scattering probability as constant within large depth intervals, in which case the path length r to a scattering event is given by an exponentially distributed random number with mean value ℓ^P or ℓ^S for P waves and S waves, respectively. Thus, individual values of r for P and S waves are computed as

$$r_P = -\ell^P \ln x$$
$$r_S = -\ell^S \ln x, \tag{3}$$

where x is a random number between 0 and 1. This is the approach taken by Margerin and Nolet (2003a,b) and Shearer and Earle (2004) and is accurate assuming that the mean free path is much larger than the ray path segments in the model.

Note that the depths separating the intervals of different scattering probabilities need not coincide with the velocity interfaces in the model. Whenever a particle enters a layer with a different scattering probability, a random number determines the ray path length to the next scattering event. As the computer tracks the ensuing particle trajectory, if the accumulated path length within the layer exceeds this number, then a scattering event occurs. If the particle leaves the layer and enters a layer with a different scattering probability, then a new random number is generated for the new layer.

Consider ray AEFGHIJKL in Fig. 5. In this case, there is a uniform scattering probability in layers 1–3 and a different scattering probability in layers 4–7. When the ray leaves the source, a random number determines the path length to the next scattering point according to Eq. (3) and the mean free path in the top scattering zone. For this example, this length exceeds the distance AE and the particle is not scattered. At point E, a new random path length is computed from the mean free path for the lower scattering zone. This path length is exceeded by the downgoing ray somewhere in layer 6 and a scattering event occurs at point G (for coding simplicity, scattering events are forced to occur at boundaries between layers). The random orientation of the scattered ray is then computed (see Section 3.5), which in general will involve a change in ray parameter, ray vertical direction (upgoing or downgoing), ray azimuth and may also involve a change in wave type (P or S). A new path length to the next scattering event is computed for the scattered ray. The particle is reflected at H and another scattering event occurs at I. The scattered particle is then transmitted at J and leaves the lower scattering zone at K, at which point a path length is computed for the distance to the next scattering event in the upper scattering zone. The particle hits the free surface at L and the energy of the particle is added to the appropriate bin in the time versus distance array.

3.5. Scattering Angles

Once a scattered event occurs, the next step is to assign a new particle trajectory, and in the case of fully elastic simulations to assign the new wave type (P or S) and S-wave polarization. The simplest approach is to assume that the scattering is isotropic, that is,

the probability is equal in all directions. However, Gusev and Abubakirov (1987) and Abubakirov and Gusev (1990) showed that this approach does not lead to realistic early coda and pulse broadening of the direct arrival. To fit most seismic observations, some form of anisotropic scattering is required with more scattering in the forward direction. This can be achieved with empirical scattering functions, such as the Gaussian function analyzed by Gusev and Abubakirov. However, this approach does not provide a connection to the physical properties of the velocity and density perturbations that are causing the scattering.

Thus, most recent Monte Carlo simulations in seismology compute scattering probabilities based on a physical model of the scattering medium, which is generally done using Born scattering coefficients for random heterogeneity models. Although Born theory is for single scattering, it can be used to model multiple scattering when the distance between scattering events is much longer than the seismic wavelength and the scale length of the random heterogeneity, which is generally the case for elastic scattering in the Earth. The required conditions can be expressed as (e.g., Wegler *et al.*, 2006)

$$\ell \gg \lambda/2\pi \tag{4}$$

and

$$\ell \gg a, \tag{5}$$

where ℓ is the mean free path, λ is the seismic wavelength, and a is the heterogeneity correlation distance. Comparisons to finite difference calculations have confirmed the validity of Born theory to compute scattering probability in radiative transfer theory (Wegler *et al.*, 2006), except in cases of extreme scattering from very heterogeneous media (such as may occur in volcano seismology). For the P coda simulations described here, the smallest mean free path is 82 km, which is much longer than the wavelength of mantle P waves at 1 Hz and the 8-km correlation length of the random heterogeneity models.

We now summarize the Born results that are necessary to implement a fully elastic Monte Carlo method that includes P-to-S and S-to-P scattering, using the appropriate results from Sato and Fehler (1998; hereafter referred to as S&F; see also Wu, 1985, and Wu and Aki, 1985a,b, for more details on Born theory in seismology). The simplest equations are obtained when the P velocity α and S velocity β are assumed to have the same fractional velocity fluctuations (S&F, 4.47):

$$\xi(\mathbf{x}) = \frac{\delta\alpha(\mathbf{x})}{\alpha_0} = \frac{\delta\beta(\mathbf{x})}{\beta_0}, \tag{6}$$

where α_0 and β_0 are the mean P and S velocities of the medium. We further assume that the fractional density fluctuations are proportional to the velocity variations (S&F, 4.48):

$$\frac{\delta\rho(\mathbf{x})}{\rho_0} = \nu\xi(\mathbf{x}), \tag{7}$$

where ν is the density/velocity fluctuation scaling factor.

The basic scattering patterns are given by (S&F, 4.50)

$$X_r^{PP}(\psi, \zeta) = \frac{1}{\gamma_0^2} \left[v \left(-1 + \cos \psi + \frac{2}{\gamma_0^2} \sin^2 \psi \right) - 2 + \frac{4}{\gamma_0^2} \sin^2 \psi \right]$$

$$X_\psi^{PS}(\psi, \zeta) = -\sin \psi \left[v \left(1 - \frac{2}{\gamma_0} \cos \psi \right) - \frac{4}{\gamma_0} \cos \psi \right]$$

$$X_r^{SP}(\psi, \zeta) = \frac{1}{\gamma_0^2} \sin \psi \cos \zeta \left[v \left(1 - \frac{2}{\gamma_0} \cos \psi \right) - \frac{4}{\gamma_0} \cos \psi \right]$$

$$X_\psi^{SS}(\psi, \zeta) = \cos \zeta [v(\cos \psi - \cos 2\psi) - 2 \cos 2\psi]$$

$$X_\zeta^{SS}(\psi, \zeta) = \sin \zeta [v(\cos \psi - 1) + 2 \cos \psi], \tag{8}$$

where X_r^{PP} is the radial component of P-to-P scattering, X_ψ^{PS} is the ψ component of P-to-S scattering, and so on. The angles ψ and ζ are defined as in Fig. 6 and the velocity ratio $\gamma_0 = \alpha_0/\beta_0$.

Assuming a random media model, the scattered power per unit volume is given by the scattering coefficients for the various types of scattering (P to P, P to S, etc.) (S&F, 4.52):

$$g^{PP}(\psi, \zeta; \omega) = \frac{l^4}{4\pi} |X_r^{PP}(\psi, \zeta)|^2 P \left(\frac{2l}{\gamma_0} \sin \frac{\psi}{2} \right)$$

$$g^{PS}(\psi, \zeta; \omega) = \frac{1}{\gamma_0} \frac{l^4}{4\pi} |X_\psi^{PS}(\psi, \zeta)|^2 P \left(\frac{l}{\gamma_0} \sqrt{1 + \gamma_0^2 - 2\gamma_0 \cos \psi} \right)$$

$$g^{SP}(\psi, \zeta; \omega) = \gamma_0 \frac{l^4}{4\pi} |X_r^{SP}(\psi, \zeta)|^2 P \left(\frac{l}{\gamma_0} \sqrt{1 + \gamma_0^2 - 2\gamma_0 \cos \psi} \right)$$

$$g^{SS}(\psi, \zeta; \omega) = \frac{l^4}{4\pi} (|X_\psi^{SS}(\psi, \zeta)|^2 + |X_\zeta^{SS}(\psi, \zeta)|^2) P \left(2l \sin \frac{\psi}{2} \right), \tag{9}$$

where $l = \omega/\beta_0$ is the S wave number for angular frequency ω, P is the power spectral density function (PSDF) for the random media model (see S&F, pp. 14–17). A popular choice for P is the exponential autocorrelation function, in which case we have (S&F, 2.10)

$$P(m) = \frac{8\pi\varepsilon^2 a^3}{(1 + a^2 m^2)^2}, \tag{10}$$

where a is the correlation distance, ε is the root mean square (RMS) fractional fluctuation ($\varepsilon^2 = \langle \xi(\mathbf{x})^2 \rangle$), and m is the wave number [i.e., the argument of the P functions in Eq. (9) above]. For example,

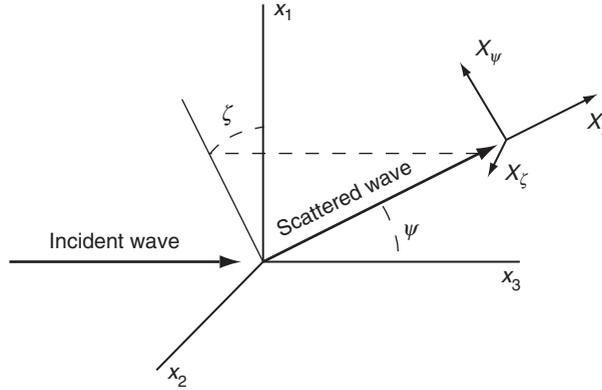$$P \left(2l \sin \frac{\psi}{2} \right) = \frac{8\pi\varepsilon^2 a^3}{\left(1 + 4a^2 l^2 \sin^2 \frac{\psi}{2} \right)^2}. \tag{11}$$

The PSDF defines the strength of the heterogeneity as a function of its scale length and controls how the amplitude of the scattering varies with seismic wavelength. Observations at a single frequency, such as the P coda results presented in this chapter, mainly constrain the heterogeneity at scale lengths near the seismic wavelength and cannot determine the PSDF very completely. Analysis of broadband data and consideration of scattered arrival amplitudes as a function of frequency will be necessary to make quantitative estimates of the PSDF.

The total scattering coefficients, g_0^{PP}, and so on, are given by the averages of these coefficients over the unit sphere. The mean free path ℓ for a ray between scattering events is given by the reciprocals of these coefficients:

$$\ell^P = \frac{1}{g_0^{PP} + g_0^{PS}}$$

$$\ell^S = \frac{1}{g_0^{SP} + g_0^{SS}} \tag{12}$$

and these values are used to assign path lengths using random numbers as described above. When a scattering event occurs, a second random number is used to decide whether the scattered wave is P or S, according to the relative sizes of g_0^{PP} and g_0^{PS} for an incident P wave or g_0^{SP} and g_0^{SS} for an incident S wave. A third random number (see below) is then used to determine the scattering angle (ψ and ζ) and the S polarization (if required). The particle then travels along its new ray direction until the next scattering event.

Despite the apparent complexity of the scattering equations, there are only three free parameters used to describe this model: the RMS fractional fluctuation ε, the correlation distance a, and the velocity density scaling factor v. Of course, a more general PSDF than the exponential model would require more parameters. Larger values of v will generally

increase the amount of backward scattering. Shearer and Earle (2004) use $v = 0.8$, an estimate for the lithosphere obtained using Birch's law (S&F p. 101). Simpler equations can be obtained for the case of velocity perturbations alone or for purely acoustic waves. The random heterogeneity described by Eq. (11) is isotropic so that the scattering properties do not depend upon the angle of the incident wave. However, anisotropic PSDFs may be important in some regions of the Earth, in which case more free parameters will be required to define the model, which may also affect the relative strength of forward versus backscattering (e.g., Hong and Wu, 2005).

An efficient way to implement these scattering kernels in a computer program is to precompute the scattered power and S-wave polarizations at a series of small intervals of solid angle. The probability of scattering at each discrete angle is then given by its relative power and a single random number can be used to assign the scattered ray path. For example, consider the scattering pattern plotted in Fig. 7 [the S-to-P coefficient g^{SP} in Eq. (9), computed for $\gamma_0 = \sqrt{3}, v = 0.8, \beta_0 = 6/\sqrt{3}$ km/s, $\omega = 2\pi, \varepsilon = 0.01$, and $a = 1$ km], which is plotted at 6° increments in ψ and ζ. Assign a unique cell number to each of the n cells in the scattering surface and save the normalized scattering probabilities in a one-dimensional array, P, of dimension n. Define a second array, S, with the cumulative probabilities in P, that is, $S(1) = P(1)$, $S(i + 1) = S(i) + P(i + 1)$, $S(n) = 1$. The scattered ray angle is then defined by the smallest value of S that is larger than a computed random number between 0 and 1.

This approach has the computational advantage that the scattering probability arrays are computed only once and then angles are obtained for individual scattering events during the Monte Carlo simulation from a single random number, without the need to recompute any of the terms in Eq. (9). The accuracy of this approach depends upon how finely the scattering angles are sampled. Shearer and Earle (2004) used an angle spacing of 1.8°. If desired, additional random numbers can be used to add a small amount of scatter to the ray angle so that the scattered ray angles are not restricted to the exact angles of the pre-computed cells. The final step is to convert from the ray-based coordinate system used to define the scattering angles (i.e., as plotted in Fig. 7) to the absolute ray parameters needed to continue propagating the particle in the model. These include the ray azimuth (degrees from north), the ray parameter (approximated as the closest ray parameter in the precomputed table), and the ray vertical direction (upgoing or downgoing).

The scattering depends on the wave number, which is a function both of the wave frequency and the local background seismic velocity. Thus, Monte Carlo calculations that include scattering must be performed for a specific frequency. In addition, Earth's changing velocity with depth results in scattering kernels that vary continuously with depth. For the most accurate whole-Earth calculations, the kernels could be computed and stored at 10-km depth intervals. But this would require random numbers to be generated every 10 km along each ray path, greatly slowing the code. Thus, for practical purposes, it is useful to approximate the scattering properties, including the mean free path, as constant within fairly coarse depth intervals. Margerin and Nolet (2003a) assume that the mean free path is constant within the entire mantle. Shearer and Earle (2004) use four mantle layers, separated at depths of 200, 600, and 1700 km.

3.6. Intrinsic Attenuation

Energy converted to heat or crystal dislocations during wave propagation is termed intrinsic attenuation (as opposed to scattering attenuation in which some energy in the main
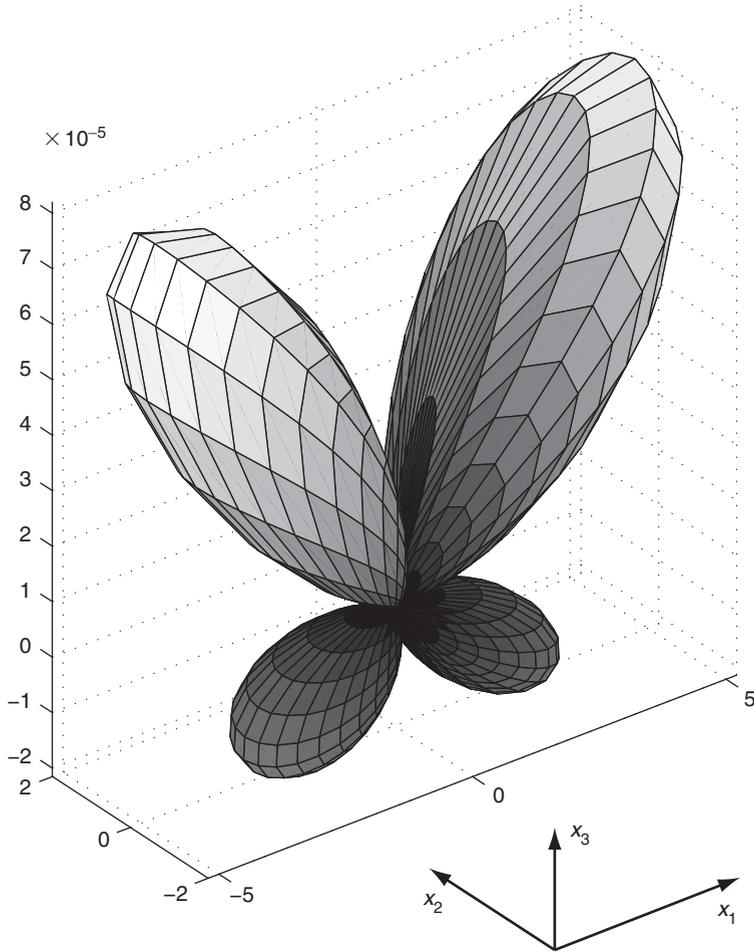
FIG. 7. An example of an S-to-P scattering pattern as a function of ray angle, computed using the Born equations for a random media model. The distance from the origin gives the value of the coefficient g^{SP} in Eq. (9). The incident S wave is traveling in the x_3 direction and is polarized in the x_1 direction.

pulse is scattered into other seismic waves). Intrinsic attenuation defined by a quality factor Q will reduce the wave energy by $e^{-\omega t/Q}$, where $\omega = 2\pi f$ is the angular frequency and t is the travel time along the ray. For global Earth models, Q varies strongly with depth so this correction is most easily performed in Monte Carlo methods by accumulating a value of $t^* = \int dt/Q$ along each particle. The energy at any desired point (such as when the particle hits the surface) is then computed using the reduction factor $e^{-\omega t^*}$.

For fully elastic calculations, both P- and S-wave attenuation must be defined. The PREM Earth model (Dziewonski and Anderson, 1981) contains depth-dependent values for both bulk and shear attenuation, $^{\kappa}Q$ and $^{\mu}Q$, from which can be computed P and S factors, $^{\alpha}Q$ and $^{\beta}Q$. However, the PREM values are accurate only at frequencies below

about 0.2 Hz because of the frequency dependence of Q at higher frequencies (e.g., Sipkin and Jordan, 1979; Lundquist and Cormier, 1980; Anderson and Given, 1982; Der *et al.*, 1986; Warren and Shearer, 2000). At 1 Hz, there is much less attenuation than predicted by the PREM model. Warren and Shearer (2000) analyzed high-frequency P and PP spectra and proposed a frequency-dependent mantle P-attenuation model that is generally consistent with prior work. At 1 Hz, $^\alpha Q = 227$ from 0- to 220-km depth and $^\alpha Q = 1383$ from 220 km to the CMB. Corresponding S attenuation can be computed using $^\beta Q = (4/9)^\alpha Q$, a commonly used approximation that assumes a Poisson solid and that all attenuation is in shear. The outer core is generally assumed to have zero attenuation, but the inner core is observed to be strongly attenuating (e.g., Bhattacharyya *et al.*, 1993; Yu and Wen, 2006), with $^\alpha Q$ values varying between about 200 and 600 and some evidence of depth dependence.

It is likely that most high-frequency estimates of Q derived from teleseismic body waves contain a mixture of both intrinsic and scattering attenuation. Thus, these published values are only useful as a starting point for intrinsic Q in whole-Earth Monte Carlo simulations; the true intrinsic Q values are likely to be higher once scattering effects are included. Shearer and Earle (2004) found this to be the case in their Monte Carlo modeling of teleseismic P amplitudes and coda at 1 Hz, for which they obtained $^\alpha Q = 450$ from 0 km to 200 km and $^\alpha Q = 2500$ between 200 km and the CMB, values significantly higher than the Warren and Shearer (2000) Q values derived from P and PP spectra.

4. FIT TO TELESEISMIC P CODA

Figure 8 shows the fit achieved to stacks of teleseismic P coda at 1 Hz by Shearer and Earle (2004) using their Monte Carlo method. The bottom plots show stacks of P coda amplitudes (obtained using the method described in Section 2) relative to the maximum P-wave amplitude for both shallow and deep earthquakes. These plots discard absolute P amplitude information, which is shown separately in the top plots. P amplitude versus distance is particularly sensitive to the intrinsic attenuation in the mantle. To model these observations, Shearer and Earle (2004) found that most scattering occurs in the lithosphere and upper mantle, but that a small amount of lower-mantle scattering was also required. Their preferred exponential autocorrelation random heterogeneity model contained 4% RMS velocity heterogeneity at 4-km scale length from the surface to 200 km depth, 3% heterogeneity at 4-km scale between 200 km and 600 km, and 0.5% heterogeneity at 8-km scale length between 600 km and the CMB. They assumed equal and correlated P and S fractional velocity perturbations and a density/velocity scaling ratio of 0.8. Intrinsic attenuation was $^\alpha Q_I = 450$ above 200 km and $^\alpha Q_I = 2500$ below 200 km, with $^\beta Q_I = (4/9)\,^\alpha Q_I$ (an approximation that assumes all the attenuation is in shear). This model produced a reasonable overall fit, for both the shallow- and deep-event observations, of the amplitude decay with epicentral distance of the peak P amplitude and the P coda amplitude and duration (see Fig. 8). These numbers imply that at 1 Hz, the total attenuation is dominated by scattering in the upper mantle and by intrinsic energy loss in the lower mantle.

To show the sensitivity of coda amplitudes to changes in the strength of the heterogeneity, Fig. 9 plots Monte Carlo predictions for a model with 30% less RMS heterogeneity (at all mantle depths) and a model with 30% more RMS heterogeneity. The resulting mean free paths in the upper 200 km are 82 km, 140 km, and 283 km for the three
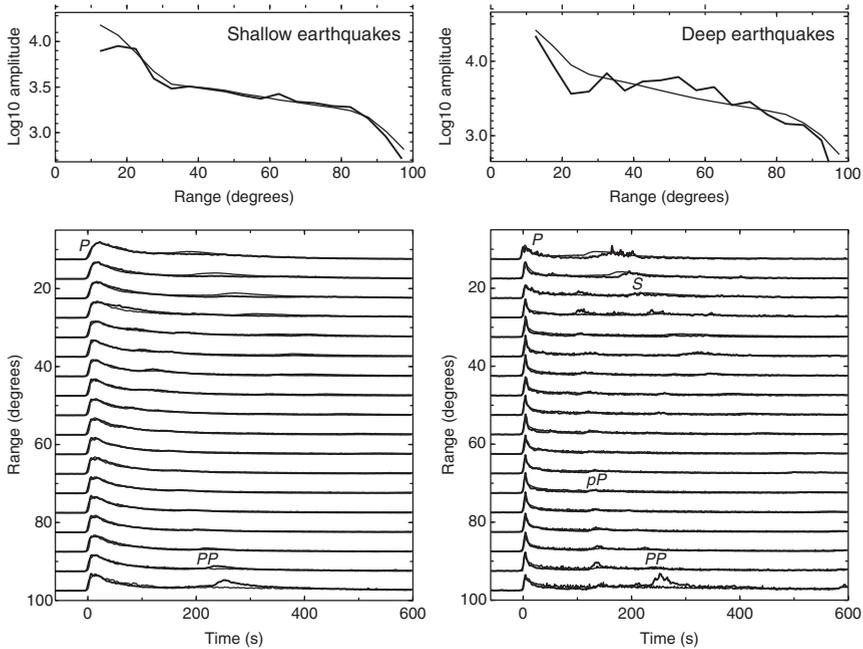
FIG. 8. Comparisons between envelope function stacks of teleseismic P-wave arrivals (solid lines) with predictions of a Monte Carlo simulation for a whole-Earth scattering model (thin lines) as obtained by Shearer and Earle (2004). The left panels show results for shallow earthquakes (≤ 50 km); the right panels are for deep earthquakes (≥ 400 km). The top panels show peak P-wave amplitude versus epicentral distance. The bottom panels show coda envelopes in $5°$ distance bins plotted as a function of time from the direct P arrivals, with amplitudes normalized to the same energy in the first 30 s. Figure adapted from Shearer and Earle (2004).

models. As expected, the coda amplitudes are very sensitive to the strength of the heterogeneity. At distances betwen 40 and 100 degrees, the differences in coda level are difficult to see at times later than about 150 s. However, at closer distances the differences persist to much longer times.

All of the P-coda observations and modeling presented here are for the vertical component. However, it is also possible to constrain mantle heterogeneity and scattering by studying the transverse component of teleseismic P coda (Nishimura *et al*., 2002; Kubanza *et al*., 2006). Our Monte Carlo code computes and saves all three components of output, but we have not yet analyzed the transverse component results.

5. CONCLUSIONS

Envelope-stacking methods and Monte Carlo modeling provide a powerful set of tools for analyzing whole-Earth scattering. Detailed applications of these approaches to a variety of seismic phases have only begun, but promise to provide reliable constraints on the average strength of small-scale heterogeneity as a function of depth in the mantle
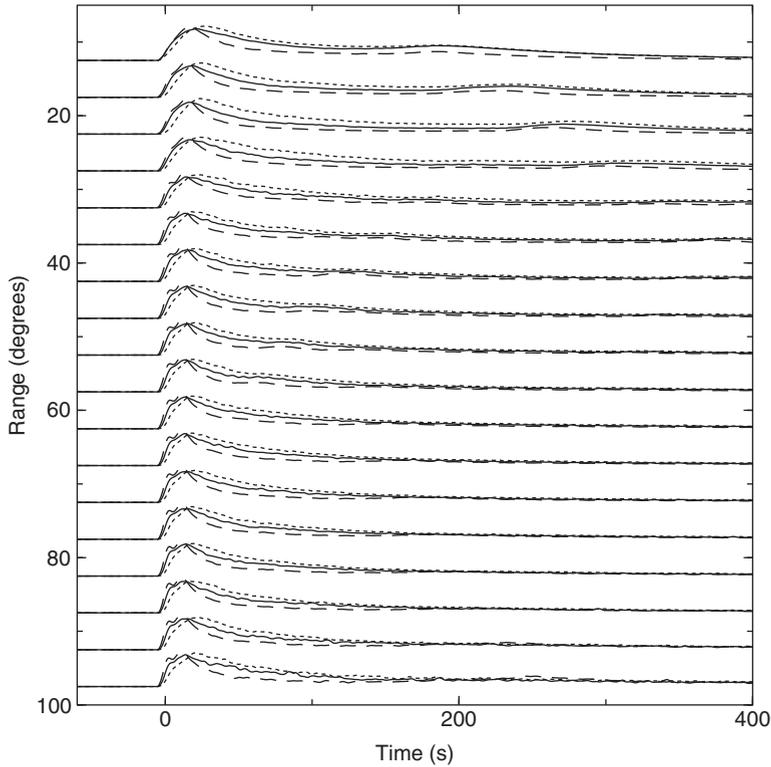
FIG. 9. A comparison of Monte Carlo-predicted P-coda amplitudes for the Shearer and Earle (2004) heterogeneity model (solid line), a model with 30% more root mean square (RMS) heterogeneity (dotted line), and a model with 30% less RMS heterogeneity (dashed line). The Shearer and Earle exponential autocorrelation random heterogeneity model contains 4% RMS velocity heterogeneity at 4-km scale length from the surface to 200-km depth, 3% heterogeneity at 4-km scale between 200 km and 600 km, and 0.5% heterogeneity at 8-km scale length between 600 km and the core-mantle boundary. It contains equal and correlated P and S fractional velocity perturbations and a density/velocity scaling ratio of 0.8. Intrinsic attenuation is $^{\alpha}Q_I = 450$ above 200 km and $^{\alpha}Q_I = 450$ above 200 km and $^{\alpha}Q_I = 2500$ below 200 km, with $^{\beta}Q_I = (4/9)\ ^{\alpha}Q_I$.

and inner core and the relative strength of intrinsic and scattering attenuation mechanisms. In addition to their value in resolving details of Earth structure, these results should help in making better estimates of earthquake source spectra, including the relative sizes of P and S corner frequencies, high-frequency spectral decay rates, and the ratio of radiated S energy to radiated P energy. Whole-Earth scattering studies and Monte Carlo simulations will become increasingly practical as global seismic data become more readily available and computer speeds continue to improve. In addition, it is clear that significant lateral variations in scattering strength exist in many regions, as can be seen in simple comparisons of teleseismic P coda levels among different stations and sources.

References

Abercrombie, R.E. (1995). Earthquake source scaling relationships from –1 to 5 M_L using seismograms recorded at 2.5-km depth. *J. Geophys. Res.* **100,** 24,015–24,036.

Abubakirov, I.R., Gusev, A.A. (1990). Estimation of scattering properties of lithosphere of Kamchatka based on Monte-Carlo simulation of record envelope of a near earthquake. *Phys. Earth Planet. Inter.* **64,** 52–67.

Anderson, D.L., Given, J.W. (1982). Absorption band Q model for the Earth. *J. Geophys. Res.* **87,** 3893–3904.

Bal, G., Moscoso, M. (2000). Polarization effects of seismic waves on the basis of radiative transport theory. *Geophys. J. Int.* **142,** 571–585.

Bataille, K., Flatté, S.M. (1988). Inhomogeneities near the core-mantle boundary inferred from short-period scattered PKP waves recorded at the global digital seismograph network. *J. Geophys. Res.* **93,** 15,057–15,064.

Bataille, K., Lund, F. (1996). Strong scattering of short-period seismic waves by the core-mantle boundary and the P-diffracted wave. *Geophys. Res. Lett.* **23,** 2413–2416.

Bataille, K., Wu, R.S., Flatté, S.M. (1990). Inhomogeneities near the core-mantle boundary evidenced from scattered waves: A review. *Pure Appl. Geophys.* **132,** 151–173.

Bhattacharyya, J., Shearer, P., Masters, G. (1993). Inner core attenuation from short-period *PKP* (*BC*) versus *PKP*(*DF*) waveforms. *Geophys. J. Int.* **114,** 1–11.

Boatwright, J., Fletcher, J.B. (1984). The partition of radiated seismic energy between *P* and *S* waves. *Bull. Seismol. Soc. Am.* **74,** 361–376.

Chang, A.C., Cleary, J.R. (1978). Precursors to PKKP. *Bull. Seismol. Soc. Am.* **68,** 1059–1079.

Chang, A.C., Cleary, J.R. (1981). Scattered PKKP: Further evidence for scattering at a rough core-mantle boundary. *Phys. Earth Planet. Inter.* **24,** 15–29.

Cleary, J.R., Haddon, R.A.W. (1972). Seismic wave scattering near the core-mantle boundary: A new interpretation of precursors to PKP. *Nature* **240,** 549–551.

Cormier, V.F. (1999). Anisotropy of heterogeneity scale lengths in the lower mantle from PKIKP precursors. *Geophys. J. Int.* **136,** 373–384.

Cormier, V.F. (2000). D'' as a transition in the heterogeneity spectrum of the lowermost mantle. *J. Geophys. Res.* **105,** 16,193–16,205.

Der, Z.A., Lees, A.C., Cormier, V.F. (1986). Frequency dependence of Q in the mantle underlying the shield areas of Eurasia, part III, the Q model. *Geophys. J.R. Astron. Soc.* **87,** 1103–1112.

Doornbos, D.J. (1978). On seismic wave scattering by a rough core-mantle boundary. *Geophys. J.R. Astron. Soc.* **53,** 643–662.

Doornbos, D.J. (1980). The effect of a rough core-mantle boundary on PKKP. *Phys. Earth Planet. Inter.* **21,** 351–358.

Dupree, S.A., Fraley, S.K. (2002). A Monte Carlo Primer, A Practical Approach to Radiation Transport. Kluwer Academic/Plenum, New York.

Dziewonski, A.M., Anderson, D.L. (1981). Preliminary reference Earth model (PREM). *Phys. Earth Planet. Inter.* **25,** 297–365.

Earle, P.S., Shearer, P.M. (1997). Observations of *PKKP* precursors used to estimate small-scale topography on the core-mantle boundary. *Science* **277,** 667–670.

Earle, P.S., Shearer, P.M. (2001). Distribution of fine-scale mantle heterogeneity from observations of Pdiff coda. *Bull. Seismol. Soc. Am.* **91,** 1875–1881.

Gusev, A.A., Abubakirov, I.R. (1987). Monte-Carlo simulation of record envelope of a near earthquake. *Phys. Earth Planet. Inter.* **49,** 30–36.

Gusev, A.A., Abubakirov, I.R. (1996). Simulated envelopes of non-isotropically scattered body waves as compared to observed ones: Another manifestation of fractal heterogeneity. *Geophys. J. Int.* **127,** 49–60.

Haddon, R.A.W., Cleary, J.R. (1974). Evidence for scattering of seismic PKP waves near the mantle-core boundary. *Phys. Earth Planet. Inter.* **8,** 211–234.

Hardebeck, J.L., Shearer, P.M. (2003). Using *S/P* amplitude rations to constrain the focal mechanisms of small earthquakes. *Bull. Seismol. Soc. Am.* **93,** 2434–2444.

Hedlin, M.A.H., Shearer, P.M., Earle, P.S. (1997). Seismic evidence for small-scale heterogeneity throughout Earth's mantle. *Nature* **387,** 145–150.

Hong, T.-K., Wu, R.-S. (2005). Scattering of elastic waves in geometrically anisotropic media and its implication to sounding of heterogeneity in the Earth's deep interior. *Geophys. J. Int.* **163,** 324–338.

Hoshiba, M. (1991). Simulation of multiple-scattered coda wave excitation based on the energy conservation law. *Phys. Earth Planet. Inter.* **67,** 123–136.

Hoshiba, M. (1994). Simulation of coda wave envelope in depth dependent scattering and absorption structure. *Geophys. Res. Lett.* **21,** 2853–2856.

Hoshiba, M. (1997). Seismic coda wave envelope in depth-dependent S wave velocity structure. *Phys. Earth Planet. Inter.* **104,** 15–22.

Kanasewich, E.R. (1981). Time Sequence Analysis in Geophysics. University of Alberta Press, Edmonton, Alberta, Canada.

Kennett, B.L.N. (1991). IASPEI 1991 Seismological Tables. Research School of Earth Sciences, Australian National University, Canberra, ACT.

Koper, K.D., Franks, J.M., Dombrovskaya, M. (2004). Evidence for small-scale heterogeneity in Earth's inner core from a global study of PKiKP coda waves. *Earth Planet. Sci. Lett.* **228,** 227–241.

Korn, M. (1990). A modified energy flux model for lithospheric scattering of teleseismic body waves. *Geophys. J. Int.* **102,** 165–175.

Korn, M. (1993). Determination of site-dependent scattering *Q* from *P*-wave coda analysis with an energy flux model. *Geophys. J. Int.* **113,** 54–72.

Kubanza, M., Nishimura, T., Sato, H. (2006). Spatial variation of lithospheric heterogeneity on the globe as revealed from transverse amplitudes of short-period teleseismic P-waves. *Earth Planets Space* **58,** e45–e48.

Lee, W.S., Sato, H., Lee, K. (2003). Estimation of S-wave scattering coefficient in the mantle from envelope characteristics before and after the ScS arrival. *Geophys. Res. Lett.* **30,** 24, 2248, doi:10.1029/2003GL018413.

Lundquist, G.M., Cormier, V.F. (1980). Constraints on the absorption band model of Q. *J. Geophys. Res.* **85,** 5244–5256.

Margerin, L. (2005). Introduction to radiative transfer of seismic waves. In: Seinsmic Data Analysis and Imaging with Global and Local Arrays, AGU Monograph Series American Geophysical Union.

Margerin, L., Nolet, G. (2003a). Multiple scattering of high-frequency seismic waves in the deep Earth: Modeling and numerical examples. *J. Geophys. Res.* **108,** B5, 2234, doi:10.1029/2002JB001974.

Margerin, L., Nolet, G. (2003b). Multiple scattering of high-frequency seismic waves in the deep Earth: PKP precursor analysis and inversion for mantle granularity. *J. Geophys. Res.* **108,** B11, 2514, doi:10.1029/2003JB002455.

Margerin, L., Campillo, M., van Tiggelen, B.A. (1998). Radiative transfer and diffusion of waves in a layered medium: New insight into coda Q. *Geophys. J. Int.* **134,** 596–612.

Margerin, L., Campillo, M., Tiggelen, B.V. (2000). Monte Carlo simulation of multiple scattering of elastic waves. *J. Geophys. Res.* **105,** 7873–7892.

Nakamura, A., Horiuchi, S., Hasegawa, A. (1999). Joint focal mechanism determination with source-region station corrections using short-period body-wave amplitude data. *Bull. Seismol. Soc. Am.* **89,** 373–383.

Nishimura, T., Yoshimoto, K., Ohtaki, T., Kanjo, K., Purwana, I. (2002). Spatial distribution of lateral heterogeneity in the upper mantle around the western Pacific region as inferred from analysis of transverse components of teleseismic P-coda. *Geophys. Res. Lett.* **29,** 23, 10.1029/2002GL015606.

Poupinet, G., Kennett, B.L.N. (2004). On the observation of high frequency PKiKP and its coda in Australia. *Phys. Earth Planet. Inter.* **146,** 497–511.

Prieto, G.A., Shearer, P.M., Vernon, F.L., Kilb, D. (2004). Earthquake source scaling and self-similarity estimation from stacking *P* and *S* waves. *J. Geophys. Res.* **109,** B08310, doi: 10.1029/2004JB003084.

Przybilla, J., Korn, M., Wegler, U. (2006). Radiative transfer of elastic waves versus finite difference simulations in two-dimensional random media. *J. Geophys. Res.* **111,** doi: 10.1029/2005JB003952.

Ryshik, L.V., Papanicolaou, G.C., Keller, J.B. (1996). Transport equations for elastic and other waves in random media. *Wave Motion* **24,** 327–370.

Sato, H. (1977). Energy propagation including scattering effects, single isotropic scattering approximation. *J. Phys. Earth* **25,** 27–41.

Sato, H. (1984). Attenuation and envelope formation of three-component seismograms of small local earthquakes in randomly inhomogeneous lithosphere. *J. Geophys. Res.* **89,** 1221–1241.

Sato, H. (1995). Formulation of the multiple non-isotropic scattering process in 3-D space on the basis of energy transport theory. *Geophys. J. Int.* **121,** 523–531.

Sato, H., Fehler, M.C. (1998). Seismic Wave Propagation and Scattering in the Heterogeneous Earth. Springer-Verlag, New York.

Sato, H., Fehler, M.C. (2006). Synthesis of seismogram envelopes in heterogeneous media. In: Advances in Geophysics, Vol. 48, Elsevier, doi: 10.1016/S0065–2687(06)48010–9.

Shearer, P.M., Earle, P.S. (2004). The global short-period wavefield modelled with a Monte Carlo seismic phonon method. *Geophys. J. Int.* **158,** 1103–1117.

Shearer, P.M., Hedlin, M.A.H., Earle, P.S. (1998). PKP and PKKP precursor observations: Implications for the small-scale structure of the deep mantle and core. In: The Carlo-Mantle Boundary Region, Geodynamics 28, American Geophysical Union, pp. 37–55.

Sipkin, S.A., Jordan, T.H. (1979). Frequency dependence of Q_{ScS}. *Bull. Seismol. Soc. Am.* **69,** 1055–1079.

Tono, Y., Yomogida, K. (1996). Complex scattering at the core-mantle boundary observed in short-period diffracted P-waves. *J. Phys. Earth* **44,** 729–744.

Vidale, J.E., Earle, P.S. (2000). Fine-scale heterogeneity in the Earths inner core. *Nature* **404,** 273–275.

Vidale, J.E., Dodge, D.A., Earle, P.S. (2000). Slow differential rotation of the Earth's inner core indicated by temporal changes in scattering. *Nature* **405,** 445–448.

Warren, L.M., Shearer, P.M. (2000). Investigating the frequency dependence of mantle Q by stacking P and PP spectra. *J. Geophys. Res.* **105,** 25,391–25,402.

Wegler, U., Korn, M., Przybilla, J. (2006). Modeling full seismograms envelopes using radiative transfer theory with Born scattering coefficients. *Pure Appl. Geophys.* **163,** 503–531, doi: 10.1007/s00024–005–0027–5.

Wu, R.-S. (1985). Multiple scattering and energy transfer of seismic waves–separation of scattering effect from intrinsic attenuation–I. Theoretical modelling. *Geophys. J.R. Astron. Soc.* **82,** 57–80.

Wu, R., Aki, K. (1985a). Scattering characteristics of elastic waves by an elastic heterogeneity. *Geophysics* **50,** 582–595.

Wu, R., Aki, K. (1985b). Elastic wave scattering by a random medium and the small-scale inhomogeneities in the lithosphere. *J. Geophys. Res.* **90,** 10,261–10,273.

Yoshimoto, K. (2000). Monte Carlo simulation of seismogram envelopes in scattering media. *J. Geophys. Res.* **105,** 6153–6161.

Yu, W.Y., Wen, L. (2006). Seismic velocity and attenuation structures in the top 400 km of the Earth's inner core along equatorial paths. *J. Geophys. Res.* **111,** B07308, doi: 10.1029/2005JB003995.