JGR Solid Earth

RESEARCH ARTICLE
10.1029/2018JB017055

Crustal Imaging With Bayesian Inversion of Teleseismic *P* Wave Coda Autocorrelation

Mehdi Tork Qashqai[1], Erdinc Saygin[1,2], and B. L. N. Kennett[3]

[1]Deep Earth Imaging, Future Science Platform, CSIRO, Perth, Western Australia, Australia, [2]Department of Physics, School of Physics, Mathematics and Computing, Faculty of Engineering and Mathematical Sciences, University of Western Australia, Perth, Western Australia, Australia, [3]Research School of Earth Sciences, Australian National University, Canberra, ACT, Australia

Key Points:
- We develop a new approach for crustal imaging by utilizing teleseismic *P* wave coda autocorrelograms
- We retrieve zero-offset reflections from teleseismic *P* wave coda to image the crustal structure of Australia through a Bayesian inversion
- The results of the probabilistic inversion are consistent with previous crustal models derived from multiple classes of information

Supporting Information:
- Supporting Information S1

Correspondence to:
M. Tork Qashqai,
mehdi.torkqashqai@csiro.au

Citation:
Tork Qashqai, M., Saygin, E., & Kennett, B. L. N. (2019). Crustal imaging with Bayesian inversion of teleseismic *P* wave coda autocorrelation. *Journal of Geophysical Research: Solid Earth*, *124*, 5888–5906. https://doi.org/10.1029/2018JB017055

Received 28 NOV 2018
Accepted 18 APR 2019
Accepted article online 23 APR 2019
Published online 29 JUN 2019

Abstract The autocorrelation of the seismic transmission response of a layered medium (autocorrelogram), in the presence of a free surface, corresponds to the reflection response. Despite many studies on the imaging of local structures through retrieval and forward modeling of stacked autocorrelograms, there is limited work on the inversion of these data. In this study, we demonstrate that the probabilistic inversion of autocorrelograms is efficient and can be used as an alternative imaging tool when other approaches are not applicable. Here, we calculate autocorrelograms of teleseismic *P* wave coda recorded on more than 1,200 permanent and temporary seismic stations across Australia and utilize a Bayesian framework to invert these data for crustal imaging. The results show patterns of structures consistent with those seen in previous crustal models constructed from receiver function, seismic reflection, and refraction methods. The new approach can therefore image large-scale crustal structures comparable to those from other seismological methods.

1. Introduction

In recent years, the autocorrelation of both the diffusive wavefield (e.g., ambient noise energy) and teleseismic coda waves has become a popular approach for extracting local structure beneath individual seismic stations, including the depth to the crust-mantle interface (Moho) and the lithosphere-asthenosphere boundary. The idea of seismic wave autocorrelation was developed by Kunetz and d'Erceville (1962) and Claerbout (1968) for plane waves at normal incidence to a horizontally stratified acoustic medium with a free surface, where one side of autocorrelation of the seismic transmission response (generated by a deep source) corresponds to the reflection response beneath the station as if there was a virtual source at the location of the receiver (zero-offset reflection). Subsequently, Frasier (1970) extended this theorem for the case of non-normal incidence propagation of elastic waves and Wapenaar (2004) extended this to 2-D and 3-D acoustic and elastic media.

Despite the success of many studies on the processing and/or forward modeling of autocorrelograms (e.g., Becker & Knapmeyer-Endrun, 2018; Clayton, 2018; Daneshvar et al., 1995; Gorbatov et al., 2013; Heath et al., 2018; Ito & Shiomi, 2012; Kennett et al., 2015; Kennett & Sippl, 2018; Nishitsuji et al., 2016; Oren & Nowack, 2017; Pham & Tkalčić, 2017, 2018; Pham et al., 2018; Romero & Schimmel, 2018; Ruigrok & Wapenaar, 2012; Saygin et al., 2017; Sun & Kennett, 2016, 2017; Sun et al., 2018; Taylor et al., 2016; Tibuleac & von Seggern, 2012), to our best knowledge, there are no published studies on the inversion of autocorrelograms for mapping major discontinuities in the crust and upper mantle. Here, we investigate the inversion of autocorrelograms for crustal imaging.

As a test bed for the inversion of autocorrelograms, we use a Bayesian framework to invert autocorrelograms for imaging the crustal-scale V_p and Moho structures across Australia. For this purpose, we retrieve stacked autocorrelograms from teleseismic waveforms containing *P* wave coda, recorded on the vertical components of the permanent and temporary seismic stations across Australia (Figure 1). By checking the consistency of crustal structures (V_p and Moho depth estimates) obtained from the inversion with those from the Australian Seismological Reference Model (AuSREM; Salmon et al., 2013), here we mainly focus on the validation of the results to show the potential of this approach for crustal imaging, rather than seek to obtain new crustal models for Australia. Our results reveal crustal features that are linked well to those seen in the AusREM.

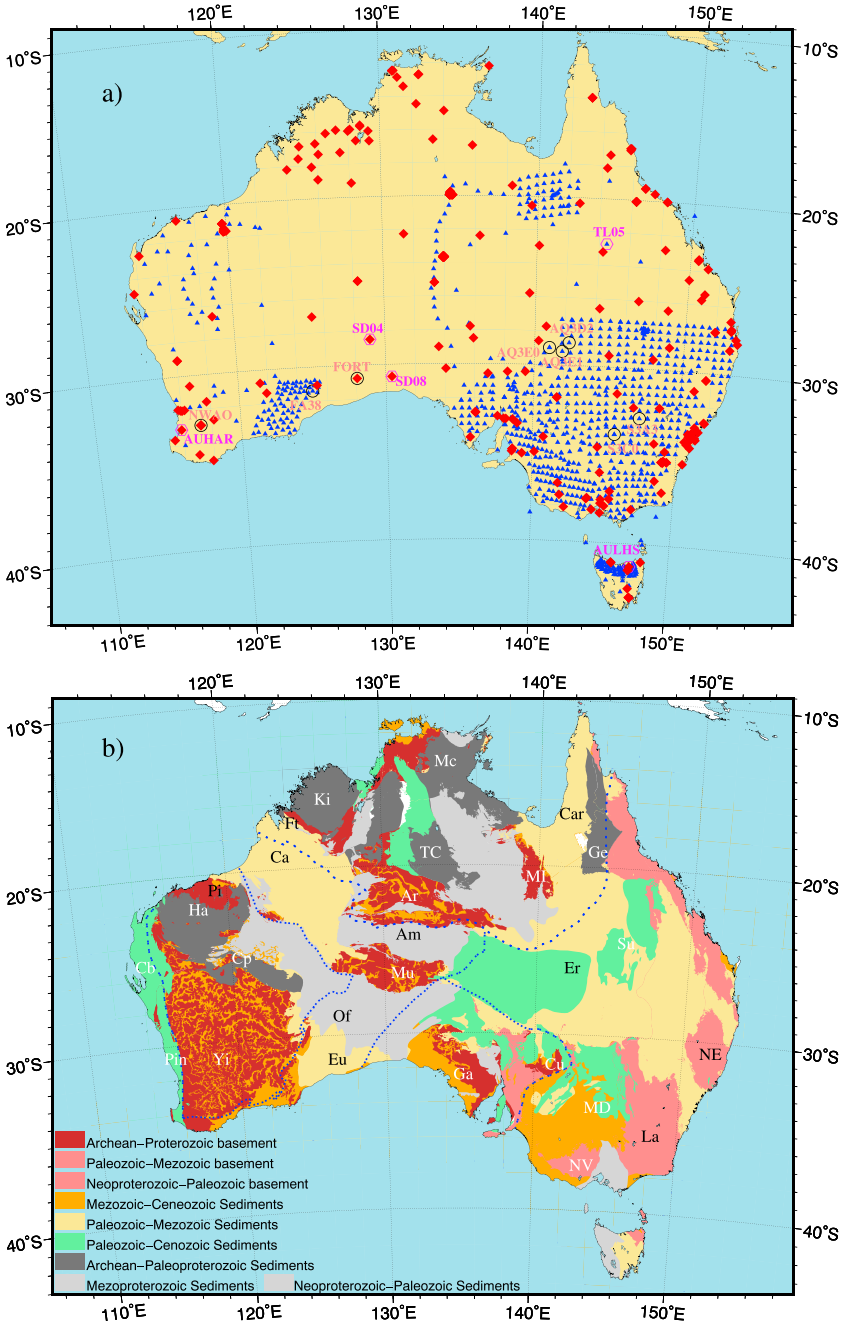

Figure 1. (a) Map of Australia with temporary (blue triangles) and permanent (red triangles) seismic stations used in this study. Locations of stations for which 1-D examples are given in section 7.1 (circles) and in the supporting information (hexagons) are also shown. (b) A simplified version of the main tectonic features of Australia (Raymond et al., 2018). Boundaries of the major cratons are marked by blue dotted lines. Key to the marked features: Am–Amadeus basin; Ca–Canning basin; CP–Capricorn orogen; Cb–Carnarvon basin; Er–Eromanga basin; Eu–Eucla basin; Ga–Gawler craton; Ha–Hamersley basin; Ki–Kimberley block; La–Lachlan orogen; Mi–Mt Isa block; Of–Officer basin; Pi–Pilbara craton; Pin–Pinjarra orogen; NE–New England orogen; Yi–Yilgarn craton; Mu–Musgrave block; Ar–Arunta block; Ft–Fitzroy trough; Mc–McArthur basin; MD–Murray-Darling basin; Su–Surat basin; Ge–Georgetown inlier; Cu–Curnamona craton; Car–Carpentaria basin; Tc–Tennant creek; NV–Newer Volcanic province.

The AuSREM model is built by integrating results of multiple seismic methods. In the following paragraphs, we briefly describe the strengths and limitations of some of the methods that were used in constructing the crustal component of AusREM such as seismic reflection, refraction, and receiver function techniques. Then, we present a short summary of the advantages of using the inversion of teleseismic *P* wave coda autocorrelogram. The limitations of our approach will be discussed later in section 7.3.

Seismic reflection and refraction profiling can be expensive to acquire, but they provide high resolution images of the Earth in 2-D and 3-D. In the deep seismic reflection profiling, the accuracy of the velocity analysis depends on the curvature of the travel time curves associated with reflections. This curvature decreases with depth, degrading the sensitivity of the velocity analysis at deeper parts of the Earth. Therefore, seismic velocities are poorly defined in the middle and lower crust, whereas seismic velocities in the upper crust are well constrained. In contrast, the deep refraction profiling needs much larger offsets than the conventional reflection method but can give better estimates of seismic velocities from the deeper parts of the Earth's crust. Hence, seismic reflection and refraction data sets complement each other and joint analyses or interpretation of them (as it is the case in the AuSREM work) can produce a robust estimation of the crustal-scale velocity structure.

Single-station receiver function imaging techniques (e.g., Ammon et al., 1990; Dueker & Sheehan, 1998; Kind et al., 2012; Langston, 1979; Phinney, 1964; Rychert & Harmon, 2016; Yuan & Kind, 2016; Wittlinger et al., 2004; Zhu & Kanamori, 2000) rely on mode conversions of body waves (*S*-to-*P* and *P*-to-*S*). They utilize teleseismic waveforms recorded on the vertical and horizontal components of a seismometer and are powerful tools in estimating the crustal and upper mantle structure beneath seismic stations. The *P* receiver functions (*P*-to-*S* converted phases) are sensitive to shear wave velocity (V_s), while the *S* receiver functions are mostly sensitive to V_p, and both are very sensitive to the seismic velocity discontinuities. *S*-to-*P* converted phases associated with the upper mantle discontinuities are not masked by crustal multiples/reverberations, which is not generally the case with *P*-to-*S* converted phases. Therefore, they are more useful for estimating the depth of major discontinuities in the upper mantle such as midlithospheric discontinuity and the lithosphere-asthenosphere boundary (Kind et al., 2012; Rychert & Harmon, 2016). However, they can also be used for crustal imaging beneath ice sheets, as conversions from crustal discontinuities beneath ice are not affected by ice multiples (Hansen et al., 2009) in *S* receiver functions. *P* receiver functions, on the other hand, are mostly used to image the Moho and crustal properties (e.g., V_p/V_s; e.g., Agostinetti & Malinverno, 2010; Langston, 1979; Phinney, 1964; Zhu & Kanamori, 2000).

In a similar way to the *S* receiver function, the teleseismic *P* wave coda autocorrelation is mostly sensitive to V_p changes at depths. However, the inversion of teleseismic *P* wave coda autocorrelation has some advantages over the inversion of *S* receiver functions for recovering the crustal V_p structure. *S*-to-*P* converted phases are much noisier than *P*-to-*S* converted phases since they arrive later in the teleseismic waveforms with other scattered arrivals. As these signals are lower frequency due to attenuation compared to *P*-to-*S* phases, they carry information with lower resolution from the major discontinuities (Hansen & Schmandt, 2017; Rychert & Harmon, 2016). In general, teleseismic *P* wave coda autocorrelograms are less noisy than *S* receiver functions and can be used over a wide range of frequencies. In fact, the inversion of autocorrelograms can potentially be used to explore shallow crustal substructures (sedimentary layer, cover thickness, and ice properties) by using a higher-frequency band of teleseismic waveforms (Daneshvar et al., 1995; Pham & Tkalčić, 2017, 2018; Romero & Schimmel, 2018; Saygin et al., 2017). In theory, the autocorrelation of seismic transmission response generated by a deep source (teleseismic earthquake) is an approximation to a zero-offset seismic reflection signal in the active source seismic reflection method (Claerbout, 1968; Kennett, 2015; Ruigrok & Wapenaar, 2012; Saygin et al., 2017) and its acquisition and processing are more cost effective, thereby it can be useful for seismic imaging of the subsurface particularly for imaging the middle and lower crustal layers where the seismic reflection imaging method has problems (Clayton, 2018).

As the standard autocorrelation technique only needs one component of the seismogram (e.g., the vertical component), it has a further advantage over the receiver function methods in areas where only the vertical component of seismograms are available, for example, earlier deployments of short-period stations in Australia. Furthermore, it has been shown by Sun et al. (2018) that *P* wave reflectivity traces obtained from global events with epicentral distances between 30° and 90° (slowness range of 0.04–0.08 s/km) are scarcely affected by conversions and multiples, which is not the case for receiver function methods.

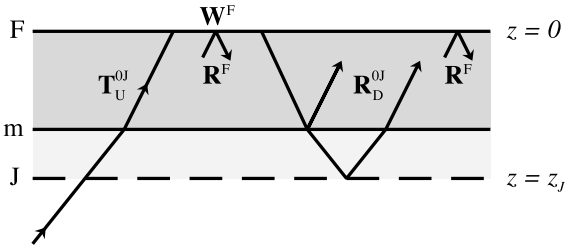

Figure 2. Representation of the transmission zone from level z_J to the surface (see section 2 and equation (2)). The lower boundary is taken below the base of the crust at m. The action of the transmission and reflection matrices is represented schematically.

The inversion of teleseismic P wave coda autocorrelograms has therefore a considerable potential for imaging subsurface structure across a broad range of scales. In this study, we limit ourselves to using the teleseismic P wave coda for imaging deep crustal structures, but diffusive seismic wavefields (ambient noise) and/or other components (horizontal components) can also be used in the inversion framework. In the following sections, we first provide an overview of the theory behind the nature of autocorrelograms obtained from teleseismic arrivals. Then, following a brief description of the data selection and processing steps in section 3, we summarize the inversion approach and model parameterization in sections 4 and 5. The resolving power of the approach is examined by applying the inversion to synthetic data, and these results are shown in section 6. In section 7, we discuss and validate results by using a comparison of our crustal V_p and Moho models with the corresponding models from AuSREM.

2. Nature of Autocorrelograms for Teleseismic Arrivals

For the onset of the wavetrain from a teleseism we can represent the response at a station in terms of incident plane waves with a single slowness component p. The frequency domain response is then (see, e.g., Kennett, Kennett, 2001, Chapter 16, and Kennett, 2002, Chapter 28)

$$\mathbf{w}_0(p,\omega) = S_I(p,\omega)\mathbf{Z}^T(p,\omega)\mathbf{C}_R^T(p,\omega)\Phi(p,\omega), \tag{1}$$

where $\Phi(p,\omega)$ represents the phase propagation effects from the source, $\mathbf{Z}(p,\omega)$ includes amplitude effects prior to arrival at the zone below the station, and $S_I(p,\omega)$ the combined effects of excitation and the instrumental response. $\mathbf{C}_R(p,\omega)$ describes the effect of structure local to the station. The superscript ^{T} denotes the transpose of the various matrices.

We take a transmission zone from a level $z = z_J$ below the base of the crust to the surface at $z = 0$, as sketched in Figure 2. The local structural term $\mathbf{C}_R(p,\omega)$ can then be written as

$$\mathbf{C}_R(p,\omega) = \mathbf{W}_F(p)\mathbf{T}_U^{FJ}(p,\omega) = \mathbf{W}_F(p)[\mathbf{I} - \mathbf{R}_D^{0J}(p,\omega)\mathbf{R}_F(p)]^{-1}\mathbf{T}_U^{0J}(p,\omega), \tag{2}$$

in terms of the transmission matrix through the zone 0J: $\mathbf{T}_U^{0J}$, the reflection matrix at the free surface $\mathbf{R}_F$, reflection back from 0J: $\mathbf{R}_D^{0J}$, and the amplification of ground motion at the free surface represented by $\mathbf{W}_F$. $\mathbf{I}$ is the identity matrix. We introduce also the modified reflection matrix

$$\mathbf{R}_D^{FJ}(p,\omega) = [\mathbf{I} - \mathbf{R}_D^{0J}(p,\omega)\mathbf{R}_F(p)]^{-1}\mathbf{R}_D^{0J}(p,\omega). \tag{3}$$

In terms of propagating waves normalized to unit energy transport in the vertical direction, there are general symmetry properties for reflection and transmission (Kennett et al., 1978)

$$\mathbf{R}_D^{0J}(p) = [\mathbf{R}_D^{0J}(p)]^T, \quad \mathbf{T}_U^{0J}(p) = [\mathbf{T}_D^{0J}(p)]^T, \quad \mathbf{R}_D^{FJ}(p) = [\mathbf{R}_D^{0J}(p)]^T, \quad \mathbf{T}_U^{FJ}(p) = [\mathbf{T}_D^{0J}(p)]^T \tag{4}$$

For an isotropic medium, the reflection matrix $\mathbf{R}_D^{0J}$ is block diagonal with a 2×2 matrix for P-SV waves and a single element for SH waves. The symmetry property (4) corresponds to symmetry in conversion P-SV and SV-P. In the case of anisotropy the matrix is full, but the symmetry properties (4) still hold for the various components.

For a perfectly elastic medium we have a further property for propagating waves:

$$[\mathbf{R}_D^{0J}(p)]^{T*}\mathbf{R}_D^{0J}(p) + [\mathbf{T}_D^{0J}(p)]^{T*}\mathbf{T}_D^{0J}(p) = \mathbf{I}, \tag{5}$$

where the star denotes a complex conjugate. The spectral relation (5) corresponds to a direct link between the autocorrelograms of reflection and transmission processes.

For a perfectly elastic medium, the modified reflection and transmission matrices have the relation

$$\mathbf{I} + \mathbf{R}_F(p)[\mathbf{R}_D^{FJ}(p,\omega)]^* + [\mathbf{R}_D^{FJ}(p,\omega)]^T[\mathbf{R}_F(p)]^T = [\mathbf{T}_U^{FJ}(p,\omega)]^*[\mathbf{T}_U^{FJ}(p,\omega)]^T, \tag{6}$$

where transmission symmetry has been employed to represent the right-hand side in terms of upward transmission. Frasier (1970) was the first to obtain a relation equivalent to equation (6) for the *P-SV* system, and a convenient derivation is provided by Ursin (1983). Recognizing that the Fourier transform of a spectrum is the corresponding autocorrelogram, we see that equation (6) implies that the autocorrelation of the transmission response yields the combination of forward and reversed time reflection response. This relation, in the time domain, was recognized for purely vertically traveling waves by Kunetz and d'Erceville (1962) and Claerbout (1968), but we see that it has a much broader validity for purely propagating waves in transmission. For a weakly attenuative medium we can envisage that equation (6) will remain a reasonable approximation, but cannot be exact. Equation (6) leads directly to a relation for $\mathbf{C}_R(p,w)$:

$$[\mathbf{C}_R(p,\omega)]^{T*}\mathbf{C}_R(p,\omega) = [\mathbf{W}_F(p)]^*[\mathbf{W}_F(p)]^T + [\mathbf{W}_F(p)]^*\left(\mathbf{R}_F(p)[\mathbf{R}_D^{FJ}(p,\omega)]^* + [\mathbf{R}_D^{FJ}(p,\omega)]^T[\mathbf{R}_F(p)]^T\right)[\mathbf{W}_F(p)]^T. \tag{7}$$

Let us now consider the autocorrelation of a single slowness component of the surface response by working in the frequency domain:

$$[\bar{\mathbf{w}}_0(p,\omega)]^{T*}\bar{\mathbf{w}}_0(p,\omega) = S_I^*(p,\omega)[\mathbf{Z}(p,\omega)]^*[\mathbf{C}_R(p,\omega)]^*[\Phi(p,\omega)]^{T*}\Phi(p,\omega)\mathbf{C}_R^T(p,\omega)\mathbf{Z}^T(p,\omega)S_I(p,\omega). \tag{8}$$

In equation (8) we can recognize the spectrum of the combined excitation and instrument response $S_I^*(\omega)S_I(\omega)$. The propagation phase term $[\Phi(p,\omega)]^{T*}\Phi(p,\omega)$ will reduce to the identity in the far field from any source, since the components are orthogonal and the phase cancels out between the two complex conjugate terms. We are then left with the far-field contribution:

$$[\mathbf{w}_0(p,\omega)]^{T*}\mathbf{w}_0(p,\omega) = [\mathbf{Z}(p,\omega)]^*[\mathbf{C}_R(p,\omega)]^*\mathbf{C}_R^T(p,\omega)\mathbf{Z}^T(p,\omega)[S_I^*(p,\omega)S_I(p,\omega)]. \tag{9}$$

We now can recognize the spectrum of the transmission term $[\mathbf{C}_R(p,\omega)]^*\mathbf{C}_R^T(p,\omega)$ modulated by external amplitude effects. Recognizing that the Fourier transform of a product produces a convolution in the time domain, we obtain

$$\mathcal{A}[\mathbf{w}_0(p,t)] = \mathcal{A}[S_I(p,t)] * \mathcal{A}[\mathbf{C}_R^T(p,t)] * \hat{\mathbf{Z}}(p,t) \tag{10}$$

where we have written $\mathcal{A}[\ \ ]$ for the autocorrelogram. The last term $\hat{\mathbf{Z}}(p,t)$ in equation (10) summarizes the amplitude effects on the wavefield before it arrives at the base of the local structure. Whatever the nature of the propagation before energy arrives at the level z_J on its way to the surface, the first two terms on the right-hand side of equation (10) will be present, whereas $\hat{\mathbf{Z}}(p,t)$ will be variable. We see that the autocorrelation of the surface displacement at a particular station contains scaled information on the reflection response of the structure beneath the station including free surface reverberations, but this is convolved with the combined effects of excitation by distant sources and the instrumental response.

When we concentrate attention on just the vertical component of ground motion, the dominant terms in $\mathcal{A}[\mathbf{C}_R(p,t)]$ will arise from *P* waves with some conversions for arrivals further away from the vertical.

When we construct the response at an individual station, we combine the estimates of $\mathcal{A}[\mathbf{w}_0]$ from many teleseismic events, with compensation for differential moveout between the different slowness arrivals. The stacking will emphasize the coherent contribution from $\mathcal{A}[S_I] * \mathcal{A}[\mathbf{C}_R^T]$. Thus, we get a modulated version of the reflection response beneath the station, including the effects of the free surface.

The process of autocorrelation emphasizes the arrivals that have a systematic pattern of delays. In the case of upward transmission through the crust, waves reflected back from the free surface such as Pp_mp will have the equivalent delay pattern to a simple reflection from the crust-mantle interface p_mp (Figure 3). We can then think of the stacking process as enhancing such internal crustal reflections which are common to a wide class of incident waves.

3. Data Selection and Processing

We construct stacked autocorrelograms for over 1,200 permanent and temporary seismic stations across Australia (Figure 1). For permanent stations, operated by Geoscience Australia, we use all seismic data

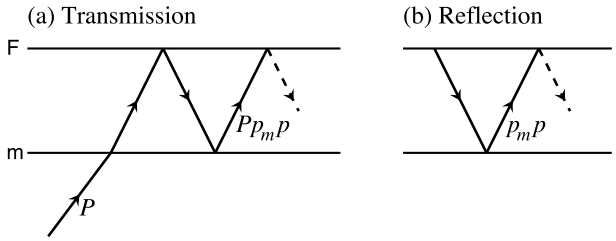

Figure 3. Equivalence of (a) transmission and (b) reflection results in the presence of a free surface (F) and reflector (m), the dashed path indicates further surface reflections.

recorded from the starting date of their operations until 31 March 2018. Temporary stations typically have an operation period of a few months up to two years. For these stations, we use all of the seismic waveforms recorded during their operation period. We choose teleseismic events with $M_w \geq 5.5$ and epicentral distances between 30° and 90° (slowness range of 0.04–0.08 s/km). At these epicentral distances, incoming *P* waves are steeply incident and mostly dominate the vertical component. The *P* wave coda is extracted from a time window of 10 to 40 s after the theoretical P arrivals predicted by the ak135 model.

In the next stage, the seismic records with quality issues (e.g., dead traces, and gaps) are automatically rejected, the mean of each seismogram is first removed and the resulting waveform is downsampled to 10 Hz. To obtain the zero-offset reflection response (autocorrelogram) beneath each station, a fourth-order Butterworth band-pass filter with corner frequencies of 0.5 and 1.5 Hz is applied before and after the autocorrelation. Traces are then normalized to unit amplitude. The signal-to-noise ratio and the response of the receiver structure are enhanced through summation (stacking) of the autocorrelation waveforms. The effects of different source time functions are also suppressed by the band-pass filtering and stacking processes. Of particular concern is contamination of the selected time window with the secondary deterministic phases such as reflections from the core-mantle boundary (PcP). By using the AK135 velocity model (Kennett et al., 1995), we calculate the theoretical arrival times for the direct *P* waves and PcP reflections. Our calculations show that PcP reflections associated with events having epicentral distances between 63° and 77° (ray parameters of 0.059–0.05 s/km) arrive in our selected time window (Figure S1 in the supporting information). However, this is not a major concern in this study as the majority of the events for all the seismic stations have average ray parameters greater than 0.06 s/km (epicentral distances ≤62° and ≥30°). Thus, the deterministic component of the *PcP* phases are attenuated by the stacking process. Finally, one standard deviation (1σ) bounds for stacked autocorrelograms are also obtained from the variance of the stacks and used as uncertainty measures for data during the inversion. An example of selected events and resulting autocorrelograms for station AQ3E7 is given in Figure 4.

Further, it is worthwhile analyzing other time windows of the teleseismic *P* wave coda for retrieving the stacked autocorrelograms. To do this, we repeat the above processing steps for other time windows (see supporting information Figure S2). Interestingly, we find that the stacked autocorrelograms resulting from different segments of *P* wave coda autocorrelograms are consistent with each other and there are no significant variations among them.

It is also of particular interest to work out a minimum number of events that are needed for robust estimation of the crustal structure below seismic stations. We investigate this for three seismic stations AQ3E0, FORT, S3C0 (see Figure 1), and the results are given in the supporting information (Figures S3–S5). We conclude that depending on the complexity of the structure below the receiver, one may need at least 30 events (for a simple structure) to 150 events (for a complex structure) to obtain a comparable velocity structure to that estimated from an inversion whose input (stacked autocorrelogram) is calculated from the full suite of available events.

4. Bayesian Inversion of Autocorrelograms

We employ a Bayesian inference approach to solve the nonlinear inverse problem $\mathbf{d} = g(\mathbf{m})$, where $\mathbf{d}$ is the data vector containing observations (here the autocorrelation of teleseismic *P* wave coda) and $g(\mathbf{m})$ is synthetic data computed using model $\mathbf{m}$. In this framework, the general solution to the inverse problem is represented by a posterior probability density function (Tarantola, 2005):

$$\sigma(\mathbf{m}) \propto \rho(\mathbf{m})L(\mathbf{m}), \tag{11}$$

where $\rho(\mathbf{m})$ is the prior probability distribution over $\mathbf{m}$ describing information about model parameters obtained independently of data $\mathbf{d}$; $L(\mathbf{m})$ is the *likelihood function*, which is a measure of misfit between data ($\mathbf{d}$) and predictions ($g(\mathbf{m})$). The posterior distribution is often high dimensional and does not possess a simple form in most applications of interest. In such circumstances, Markov chain Monte Carlo sampling approaches are used to solve the inverse problem.

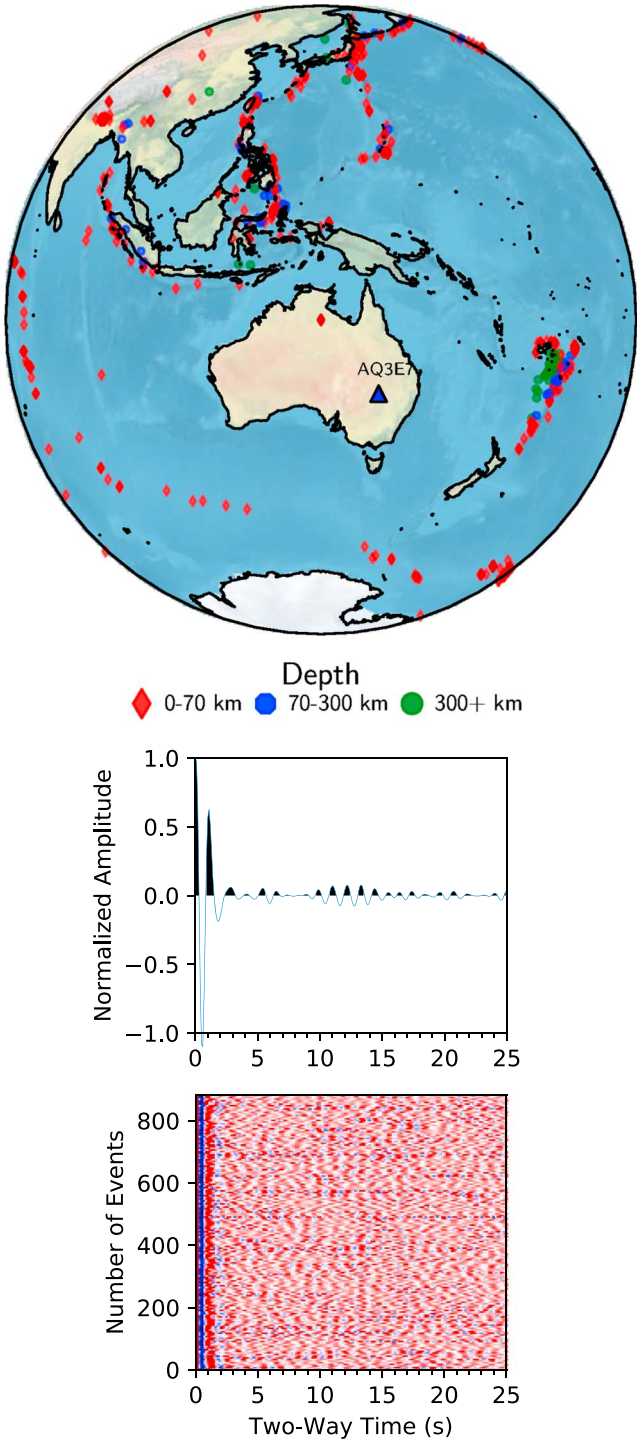

Figure 4. The top panel shows an example of earthquake distribution, and the middle panel indicates the stack of all autocorrelograms, shown in the bottom panel, for station AQ3E7. The bottom panel illustrates that the arrivals associated with the Moho are aligned and visible between 10 and 15 s.

In the typical case of independent and identically (normally) distributed observations, we can adopt the following form for the likelihood function (e.g., Tarantola, 2005):

$$L(\mathbf{m}) \propto exp\left(-\frac{1}{2}\Phi(\mathbf{m})\right), \tag{12}$$

where

$$\Phi(\mathbf{m}) = \sum_{i=1}^{N}\left[\frac{\mathbf{d_i^{obs}} - g_i(\mathbf{m})}{\sigma_i}\right]^2 \tag{13}$$

is *the misfit function*, i refers to data points, N is the total number of data points, and σ_i is 1σ uncertainty associated with data at point i.

We use the Delayed Rejection Adaptive Metropolis (DRAM) algorithm (Haario et al., 2006) to sample the posterior distribution (equation (11)). DRAM is a combination of Adaptive Metropolis (AM; Haario et al., 2001) and Delayed Rejection (DR; Mira, 2001) algorithms and has been widely used in many geophysical applications (Afonso et al., 2016; Ball et al., 2014; Tork Qashqai et al., 2016, 2018). The DRAM algorithm benefits from the advantages of both DR and AM sampling methods. The traditional Metropolis-Hastings (MH) sampling algorithm (Metropolis et al., 1953) considers only one proposed distribution at each simulation for the model and tries to accept or reject a sample drawn from the proposal. When the variance of the proposed distribution is chosen to be too high, most of the models may be rejected. This may be addressed by employing the DR algorithm that seeks to rectify problems associated with inappropriate proposals for the model by allowing for a second (or more) modified proposal(s) before rejecting a sample (e.g., searching a region in the parameter space closer to the current position). Thus, DRAM improves the efficiency of sampling by reducing the number of rejected models.

A proper choice of proposal distribution for the MH algorithm has been identified as a critical factor for the convergence of the algorithm. In the case where the variance of the proposal distribution is chosen extremely small in the MH algorithm, the algorithm proposes small jumps, and therefore, it takes a long time to converge. A possible solution to this can be provided by the AM algorithm that uses the history of the chain (accepted models so far) in order to update the proposal distribution after some non-adaptation time. The adaptation process can be iterated at regular intervals.

The number of adaptation, the optimum number of samples needed for adaptation (non-adaptation time), and the number of samples between adaptations (intervals) depend on the problem. In this study, the total number of simulations per station is 200,000. Adaptation starts after 100,000 simulations (non-adaptation time), and the proposal distribution of the model is updated 40 times (every 2,500 iterations). Therefore, the inversion benefits from the DR algorithm during the entire run time of the inversion with the additional benefit from the AM algorithm after the non-adaptation period. The inversion approach is computationally efficient. For the current parameter settings and dimensionality of our problem (4 crustal layers; 15 parameters), the total run time of the parallelized inversion including twelve 1-D inversions, each of which has 200,000 simulations, is 323 s on a Linux machine with 12 processors (one processor for each inversion). However, the run time increases as the dimensionality and the complexity of the target and prior distributions grow.

5. Model Parameterization, Priors, and Forward Problem

The Earth beneath each seismic station is parameterized with four horizontal and isotropic crustal layers over an upper mantle layer (half-space). Each crustal layer is described by three main parameters: density (ρ), thickness variation (Δh), and V_p/V_s. Density and V_p/V_s parameters in the upper mantle, as well as the slowness are also treated as unknown and directly derived from the inversion. In the inversion, the thickness of the crust is obtained by randomly perturbing a four-layer version of the CRUST 1.0 thickness model (Laske et al., 2013). We show in section 6 that the final crustal thickness model estimated from the inversion does not depend on this initial crustal thickness model.

The prior probability distributions of all model parameters are uniform, and their minimum and maximum bounds are given in Table 1. The assigned prior ranges for parameters are broad enough to consider all plausible models. During the inversion, when ρ is sampled from its prior distribution by the DRAM algorithm, we use the empirical relationship proposed by Brocher (2005) to relate it to V_p. V_s is then calculated for

Table 1
The List of All Unknown Parameters and Priors

Main parameters	Minimum	Maximum
Density (first layer-kg/m^3)	2,400	2,850
Density (second layer-kg/m^3)	2,400	2,850
Density (third layer-kg/m^3)	2,700	3,000
Density (fourth layer-kg/m^3)	2,700	3,000
V_p/V_s (in all crustal layers)	1.65	1.9
Δh (first layer-km)	−3.5	3.5
Δh (second layer-km)	−8.0	8.0
Δh (third layer-km)	−8.0	8.0
Δh (fourth layer-km)	−8.0	8.0
Density (in the upper mantle-kg/m^3)	3,200	3,500
Slowness (s/km)	0.04	0.08

each layer from the computed V_p and the unknown V_p/V_s parameter (drawn from its prior distribution by the DRAM algorithm; $V_s = \frac{V_p}{V_p/V_s}$).

Sampled parameters are used in a modified version of the reflectivity code of Randall (1989; *respknt*) to generate synthetic seismograms. This algorithm is based on the reflectivity matrix approach of Kennett (1983) for a cylindrically symmetric medium. The synthetic vertical seismograms at the free surface are then autocorrelated to estimate the reflection response (autocorrelogram) of the 1-D Earth beneath each station. In order to have the same frequency content as the observed data for synthetic data, predictions are band-pass filtered before and after autocorrelation using the same band-pass filter as for data. We do not assume any discontinuity in the upper mantle.

6. Synthetic Recovery Tests

Four sets of synthetic data associated with four 1-D Earth models are inverted to demonstrate the robustness of the technique in recovering crustal reflections and their properties (Figure 5). The synthetic data sets are constructed by solving the forward problem for four different known Earth models. Vertical components of all synthetic seismograms are computed using a modified version of the *respknt* program (Randall, 1989) using a slowness of 0.06 s/km. The same processing workflow as applied to real data is used here to extract synthetic *P* wave reflectivity traces (autocorrelograms). Low-frequency random Gaussian noise with a standard deviation of 0.015 is added to each of the synthetic data sets (see Figure S7 in the supporting information).

Model 1 (Figure 5a) consists of a four-layer crustal model with a relatively thick (5 km) sedimentary layer on the top. Model 2 (Figure 5b) has the same velocity structure as model 1 but with different layer thicknesses. Model 3 (Figure 5c) and model 4 (Figure 5d) are rather different to the previous two models. Model 3 is a three-layer Earth model with a shallower Moho depth. We create this 1-D model to see if the true number of layers with other crustal properties can be identified by the inversion using a four-layer crust parameterization. In model 4, we include a low-velocity upper crustal layer to illustrate the power of the inversion in recovering negative velocity anomalies. As in real data, crustal thicknesses are perturbed around a reference crustal thickness model. Here, we use the same prior bounds as used for the real data (Table 1). Crustal thicknesses for each of the synthetic 1-D models are randomly perturbed around wrong crustal thickness models (see Tables S1–S4), which are far from the true models.

Probability density plots for the posterior distributions of V_p and crustal discontinuities are drawn from the best 2,000 accepted models, as well as fits to observed data clearly show that the Bayesian inversion reliably recovers the "true" crustal models in the presence of noise in the autocorrelogram. Given the sensitivity of autocorrelograms to the discontinuities, and the wide ranges used for priors, the posterior distribution for some depths and V_p are not fully resolved in models 1 and 2. However, as expected, the mean solution (blue) is a smooth version of the true model and the data fit is excellent. In model 3, despite using a four-layer parameterization in the inversion, the true model (three-layered Earth) is well recovered by the posterior

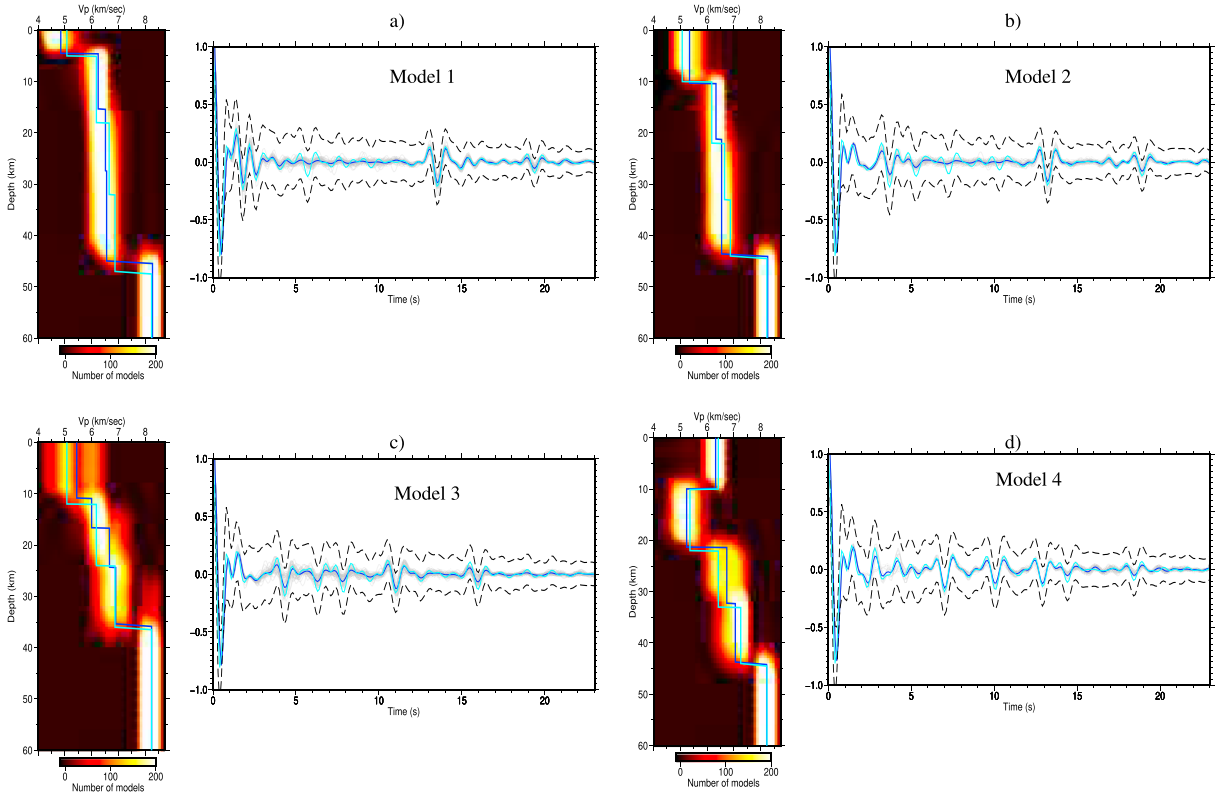

Figure 5. Synthetic tests for four different 1-D synthetic earth models (a–d) to show the resolving power of the technique in recovering crustal reflections and their properties. In (a)–(d), the left panels are the posterior distribution of the V_p structure from the best 2,000 accepted models. The right panels display the fit to stacked autocorrelograms contaminated by random Gaussian noise. Gray colors are the best 2,000 predictions. Dashed lines are 1σ observational uncertainty level obtained during the stacking of the observed autocorrelograms. In (a)–(d), the cyan color is the synthetic data (right) or the synthetic structure (left) and the blue color indicates the mean solution.

ensemble of models. In general, our tests show that a parameterization with a sedimentary layer over three crystalline layers can better fit data at some locations compared to a three-layered crust parameterization. In model 4, where the velocity jump across the crust-mantle boundary is less than the other models, the true structure is fully resolved and is within the high probability regions of the posterior distribution. Slowness values for all models are also fully recovered, and their associated posterior distributions are shown in Figure S8 (supporting information). Synthetic tests, especially for the model 1 and model 3, suggest that if the velocity jump across crustal discontinuities is smaller than 1σ posterior uncertainties, and/or the inversion is overparameterized (e.g., by one layer), we might obtain a smooth version of the true subsurface structure. The recovery of the synthetic V_s structures for the models 1 to 4 is also given in the supporting information (Figure S9).

6.1. The Exponential Weighting Function

A challenge that emerged during the inversion of real data was the effect of side lobes of the zero-lag autocorrelation peak on the inversion results. This effect becomes more pronounced for shallow reflections, where it can lead to a poor fit to the crustal reflections at some stations. To solve this problem, the autocorrelation of the teleseismic *P* wave coda is multiplied by an exponential weighting function to dampen the high amplitudes of the autocorrelogram's side lobes (e.g., ~ 0–2.0 s):

$$g(\mathbf{t}) = \mathbf{a} \times \mathbf{t}^{\mathbf{b}} \times \exp(\frac{\mathbf{t}}{\mathbf{c}}), \tag{14}$$

where $\mathbf{t}$ is the two-way travel time (0–30 s), $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{a}$ are empirical constants determined by trial and error to suppress the amplitudes of the autocorrelograms around the side lobes. After a number of tests, $\mathbf{b}$ and $\mathbf{c}$ are fixed ($\mathbf{b} = 1.3$ and $\mathbf{c} = 1{,}000$), while different $\mathbf{a}$ values are examined for decreasing the amplitudes of the autocorrelation's side lobes.

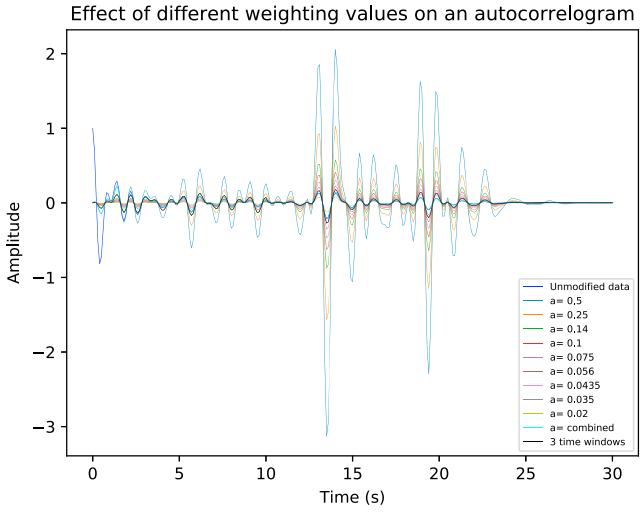

Figure 6. This figure illustrates that each of the subjective weighting values (**a** coefficients in the exponential weighting function introduced in section 6.1) can change the amplitudes of a synthetic autocorrelogram significantly if they are individually applied on the entire autocorrelation window. The autocorrelogram is modeled using the synthetic model 1 given in Figure 5. The amplitudes of the autocorrelogram after applying the exponential function with different weighting values are compared with the amplitude of the unmodified autocorrelogram (blue autocorrelogram) and the autocorrelogram obtained from applying equation (14) with the chosen weighting values a_1 to a_9 (cyan autocorrelogram) and time windows. The result for applying the exponential weighting function with the three time windows (black autocorrelogram) instead of nine time windows (see section 6.1) is also given for comparison purposes.

In Figure 6, we show the results of applying the exponential weighting function, equation (14), on a synthetic autocorrelation using different **a** values. The synthetic autocorrelogram (thick blue autocorrelogram in Figure 6) is modeled using the synthetic velocity model 1 (see section 6). As can be seen from the figure, relatively high **a** values (**a** = 0.5, **a** = 0.25, **a** = 0.14, **a** = 0.1) can be chosen to dampen the amplitudes of the side lobes. However, they also modify the rest of the autocorrelogram significantly including the Moho reflection (around 13.5 s). On the other hand, if very small **a** values are chosen, the amplitudes of the later arrivals (10–30 s) are left relatively unchanged, whereas the amplitudes of the first 10 s of the trace are notably decreased. To overcome this problem, the entire autocorrelogram is divided into nine small time windows and different time windows are multiplied by the equation (14) with different **a** values to only change the amplitudes of autocorrelation at first 2–3 s and leave the rest of trace relatively unchanged. Values of **b** and **c** are kept fixed for all windows and set to 1.3 and 1000, respectively. Values of a_1 to a_9 and the nine time windows are determined as follows: a_1=0.5 ($0 \le t \le 1.5$ s), $a_2 = 0.25$ ($1.6 \le t \le 3.4$ s), $a_3 = 0.14$ ($3.5 \le t \le 5.0$ s), $a_4 = 0.1$ ($5.1 \le t \le 6.3$ s), $a_5 = 0.075$ ($6.4 \le t \le 8.2$ s), $a_6 = 0.056$ ($8.3 \le t \le 9.7$ s), $a_7 = 0.0435$ ($9.8 \le t \le 12.3$ s), $a_8 = 0.035$ ($12.4 \le t \le 16.4$ s), $a_9 = 0.02$ (16.5 s $\le t$).

When working with real data, the amplitude pattern may not be as the same as that for the synthetic autocorrelograms. Therefore, the **a**, **b**, and **c** values are carefully selected to preserve the amplitude of later arrivals ($t > 3.5$ s) as far as possible. In Figure 6, we show that applying an exponential weighting function with three time windows and **a** coefficients (black autocorrelogram) including $a_1 = 0.25$ ($0 \le t \le 5$ s), $a_2 = 0.14$ ($5.1 \le t \le 10$ s), and $a_3 = 0.0435$ ($10.1 \le t$ s) cannot preserve the amplitudes of the original autocorrelogram at t≥5 s as far as the chosen one with nine time windows.

This weighting process improved data fit and led to obtaining a more reliable structure at locations where the inversion failed to fit data properly (especially Moho reflections). Examples of such improvements for some stations, where their locations indicated by magenta hexagons in Figure 1, are given in the supporting information. By using synthetic and real data, we show in the supporting information (Figures S10–S15) that applying the weighting function does not significantly change the results for the deep crust (e.g., Moho) but slightly increases the V_p uncertainty (broadening the posterior distribution of V_p) at shallow depths. In the supporting information (Figure S16), we also show the map of minimum misfit obtained from 1-D inversions without applying the weighting function (see sections 7.1 and 7.2 for more details about the interpolation scheme). A comparison between this map and Figure 8 (see next section) implies that applying equation (14) improved the data fitting.

7. Results and Discussion

7.1. 1-D Examples and Misfit

In this section, we focus on 1-D examples of V_p structures obtained from the inversion of real data and the misfit between the data and predictions. The posterior distribution of V_p structure and also predicted autocorrelograms obtained from the best 2,000 accepted models for eight seismic stations (black circles in Figure 1) are plotted in Figure 7 as examples. The misfit between observed and predicted autocorrelograms is satisfactory at most of the stations. However, the number of events was ≤10 for a few stations, which produced unrealistic observational uncertainties during the stacking. Therefore, the inversion was unsuccessful to give a satisfactory level of a misfit at these locations (the inversion accepted only a few models). These stations are indicated by red circles in Figure 8, which shows the 2-D map of misfit between the observed and predicted data (divided by the total number of data points). This map is obtained through an interpolation of the minimum misfit values (e.g., the misfit value associated with the model that best describes data) of 1-D

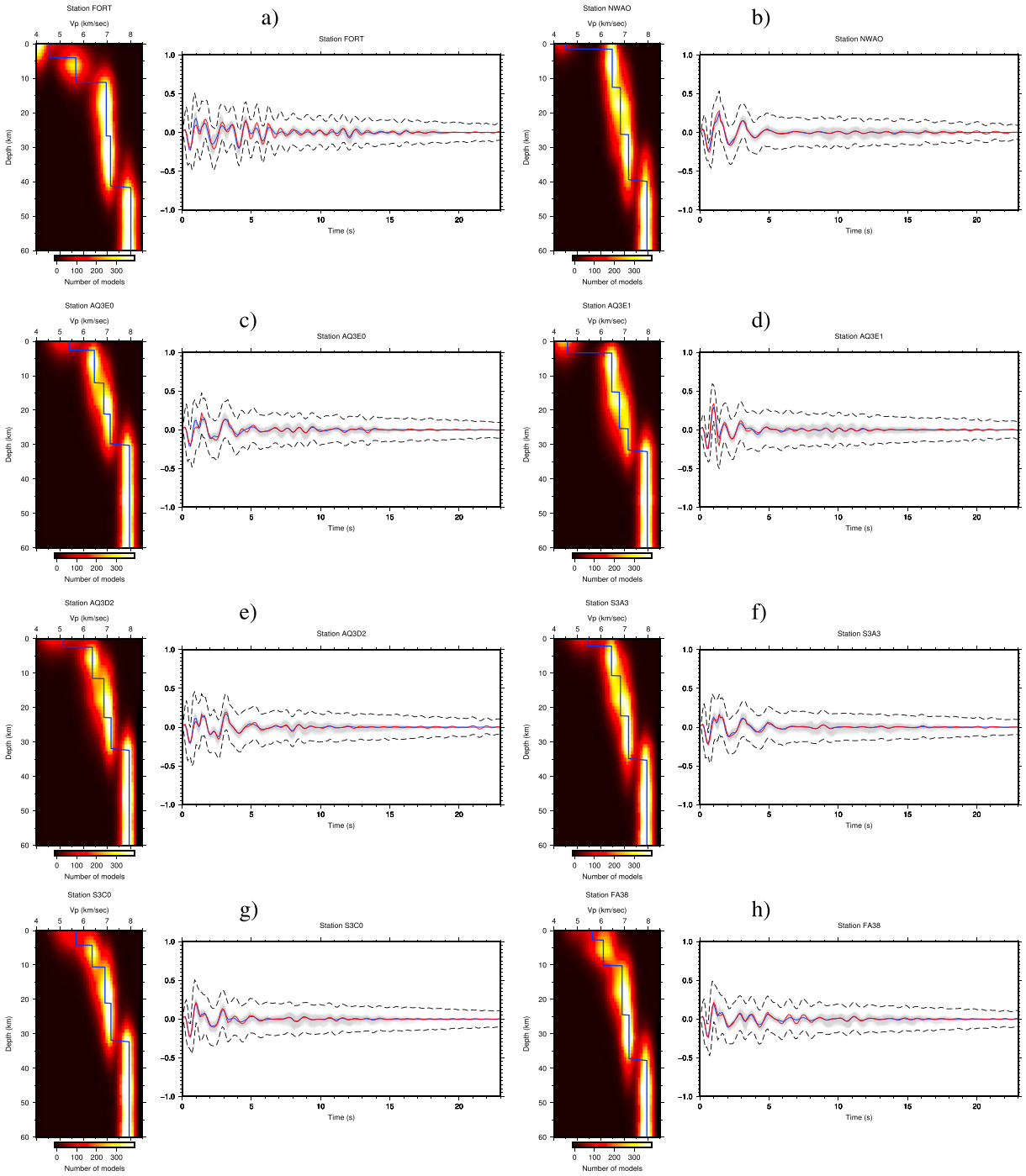

Figure 7. 1-D examples showing fits to the stacked autocorrelograms (right panels in a–h) and the probability density plots of the recovered V_p structure (left panels in a–h) for eight seismic stations shown by circles in Figure 1. In showing the V_p structures, the blue colors are the mean solutions while in the figures representing fits to data red and blue colors are data and the mean of the best 2,000 predictions (gray colors), respectively. Dashed lines are 1σ observational uncertainty level obtained during the stacking of the observed autocorrelograms. The exponential weighting function (equation (14)) has been applied to all autocorrelograms during the inversion.

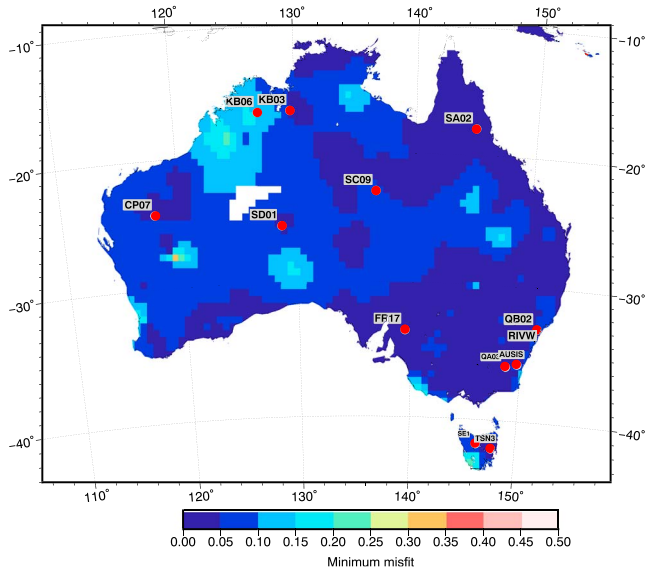

Figure 8. The 2-D map of minimum misfit between the observed and predicted data (unitless). The map obtained via the interpolation of the minimum misfit values (obtained via equation (11)) associated with 1-D inversions. These values are divided by the total number of data points (301 in the present work) before the interpolation. For some stations, the number of events was ≤10, which led to obtaining unrealistic observational uncertainties during the stacking. Therefore, the inversion was unsuccessful to fit data appropriately (e.g., low acceptance ratio). The misfit values associated with these stations are not used in producing the misfit map.

inversions. The misfit values associated with stations given in Figure 8 are excluded from the interpolation calculation.

Figures 7 and 8 demonstrate that the fit to data at most of the locations is generally good (misfit ≤0.1) and thus results from the inversion are reliable and can be considered as the first-order estimate of the deep crustal properties. We assume a 1-D Earth model with four uniform isotropic crustal layers, each of which is characterized by uniform physical properties. Additionally, we assume no discontinuities in the upper mantle. Therefore, it is obvious that at locations where the nature of the structure beneath the station violates our assumptions, the teleseismic *P* wave coda becomes far more complex and the misfit values are expected to be relatively high. In such cases, the 1-D structures and the inferred 3-D structure (see next section) around those particular locations have higher uncertainties.

7.2. Moho Structure

Results from the 1-D inversions at the full suite of stations are combined through interpolation (stations given in Figure 8 are excluded) to image Moho and pseudo 3-D V_p structures (and their related uncertainties) across the continent (section 7.3). The interpolation technique from the Generic Mapping Tools package (Smith & Wessel, 1990; Wessel et al., 2013) is used to construct a grid with 0.5° × 0.5° resolution for crustal structure across the Australian continent. The mean values of the posterior distributions of crustal parameters, for example, Moho and V_p, are used as input for interpolation calculations.

The Moho structure and its 1σ uncertainty, derived from its posterior distribution, are shown in Figures 9a and 9b, respectively. In the supporting information (Figure S17), we also provide posterior distributions (as histograms) of the depth to Moho for stations whose 1-D V_p structures are shown in Figure 7. The 1σ uncertainties are ≤10% of the maximum Moho depth (≤6.0 km). Since the autocorrelograms are more sensitive to velocity/density discontinuities (similar to receiver functions), one of the possible reasons for the high uncertainties (e.g., 5 km) is the lack of strong impedance contrast at the crust-mantle boundary at some locations, especially in east and southeast of Australia (Kennett et al., 2011, 2015). The wide prior ranges used for crustal thickness variations (Table 1) could also contribute to the higher uncertainties. In Figure 9c we show the Moho map obtained by Salmon et al. (2013; the AuSREM Moho model), which has been derived using multiple seismic methods, such as deep seismic reflection profiles, receiver functions studies, and seismic refraction experiments.

In Figure 9d, we display the difference between our Moho depth model and the AuSREM Moho model, where the negative values denote deeper Moho depths obtained in this study. In many parts of the continent, the velocity contrast across the crust-mantle boundary is gradational, and its thickness varies between 2 and 8 km (Salmon et al., 2013). Part of the reason for shallower Moho found in the present study at some locations is that our parameterization cannot capture this transition. Thus, we believe that shallower Moho depths obtained here are likely the top of this transition zone (e.g., east of Australia). Another reason for differences between our Moho map and the AuSREM Moho map is that the AuSREM model employed multiple data sources, which provide a different pattern of coverage. For example, while we have a limited number of stations in central Australia, the AuSREM Moho values in this area were estimated from several seismic refraction and reflection surveys. Therefore, we have limited resolution after the interpolation in central Australia, leading to large differences with the AuSREM Moho values in this region.

Nevertheless, the overall pattern of our Moho depth model is highly comparable to that obtained by Salmon et al. 2013; Figure 9c). The majority of the difference between our model and the AuSREM Moho model is mostly negligible (±2 km) and is between 2 and 6 km at some localities across the continent (e.g., central Australia). These differences lie within the 1σ uncertainty of the posterior distribution of the Moho depth. The consistency of our results with the AuSREM Moho model is suggesting that the Bayesian inversion of autocorrelograms can produce comparable results to those from receiver function and seismic reflection

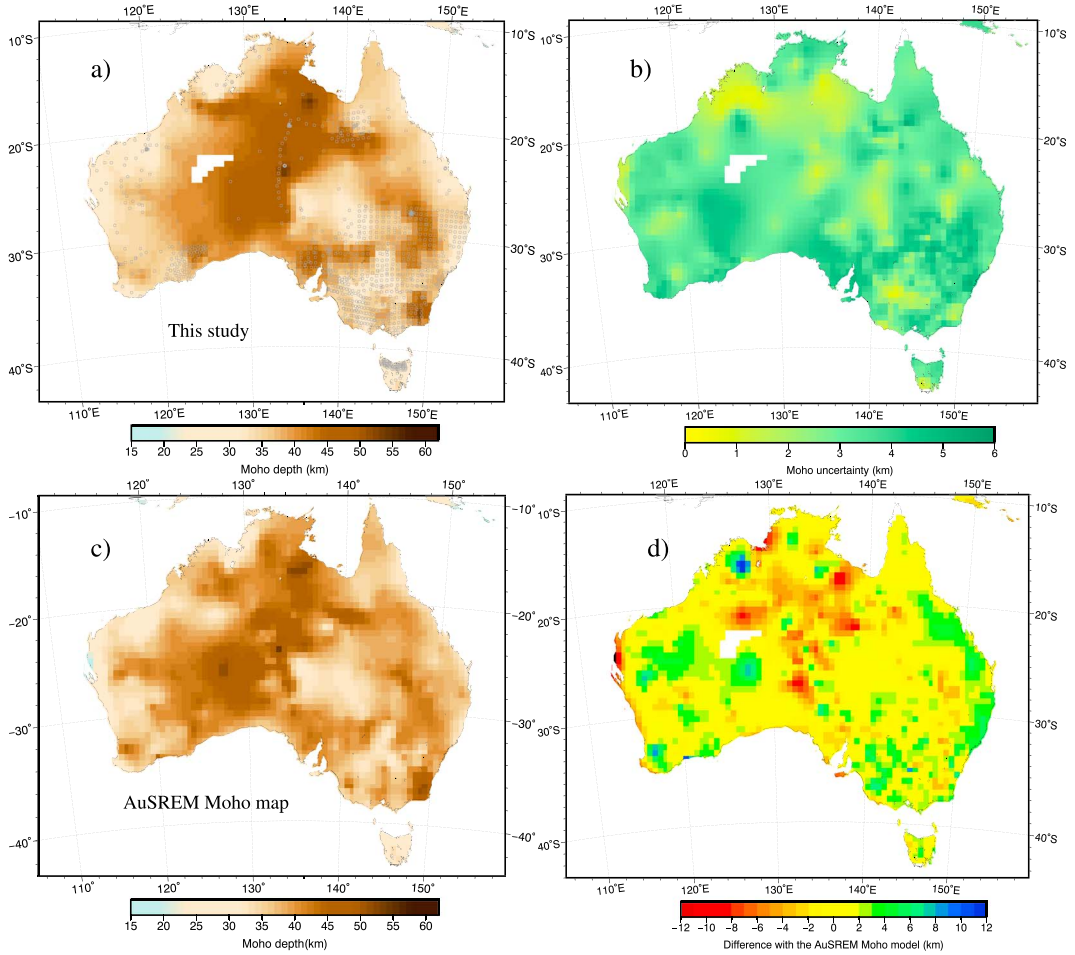

Figure 9. (a) The Moho structure map estimated from the inversion of autocorrelograms. Seismic stations used in this study are superimposed (gray circles). (b) The 1σ uncertainty map constructed from the posterior distribution of the Moho depth across the continent. (c) Map of the Australian Seismological Reference Model (AuSREM) Moho model estimated by (Salmon et al., 2013). (d) The map of difference between (a) and (c). Negative values indicate deeper Moho found in this study.

methods. It is noteworthy that we find deeper Moho in the southeast of the Eromanga Basin and underneath the Murray Basin and also in an area with a relatively thick crust near 33°S and 140°N, as seen in deep reflection profiling (Kennett, 2015) and receiver functions (Fontaine et al., 2013) images. These results are also consistent with those obtained by Kennett et al. (2015) using spatial stacking of high-frequency autocorrelograms. Shallower Moho depths imaged here around 37°S and 148°N also matches the results of Kennett et al. (2015). In summary, the dissimilarity between our Moho model and the AuSREM Moho model may be generally due to different approaches and/or data sets with disparate sensitivities and coverage being used.

7.3. V_p Structure

As the tectonic interpretation of the V_p structure is beyond of the scope of this paper, in this section we only summarize the general pattern of our V_p model and compare it with the AuSREM model (Salmon et al., 2013). For this purpose, depth slices at depths of 5, 10, 20, 30, 40, and 50 km are given in Figure 10. The same depth slices are also depicted for V_s in the supporting information (Figure S18). Some selected vertical slices from the inverted 3-D V_p structure along profiles AA$''$, BB$''$, CC$''$, and DD$''$ (shown by dashed lines in Figure 10) are plotted in Figure 11.

The primary control on the AusREM crustal *P* wave velocity model comes from the seismic refraction profiles, which are sparse and do not provide full coverage of the majority of Australia. Additional information has been estimated from receiver function studies and ambient noise tomography (Saygin & Kennett, 2012) via conversion of V_s to V_p. The latter provides a good continent-wide coverage down to 25 km depth. Below

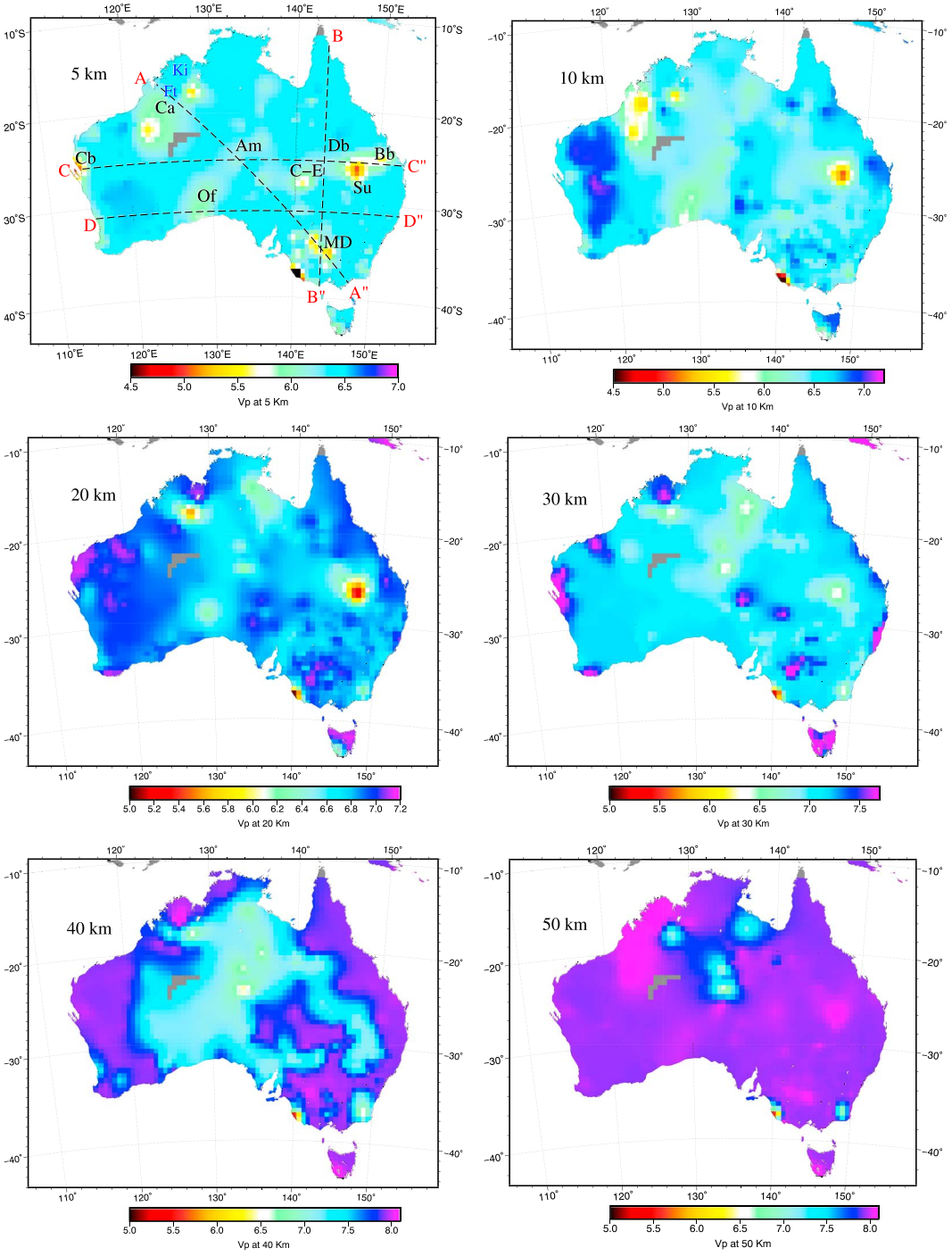

Figure 10. Depth slices of V_p structure at 5, 10, 20, 30, 40, and 50 km. Profiles AA$''$ to DD$''$ are also shown by dashed lines in the 5 km depth slice. Abbreviations: Carnarvon basin (Cb), Amadeus basin (Am), Canning basin (Ca), Cooper-Eromanga basin (C-E), Officer basin (Of), Murray-Darling basin (MD), Surat basin (Su), Kimberley Block (Ki), Drummond basin (Db), Fitzroy trough (Ft), and Bowen basin (Bb).

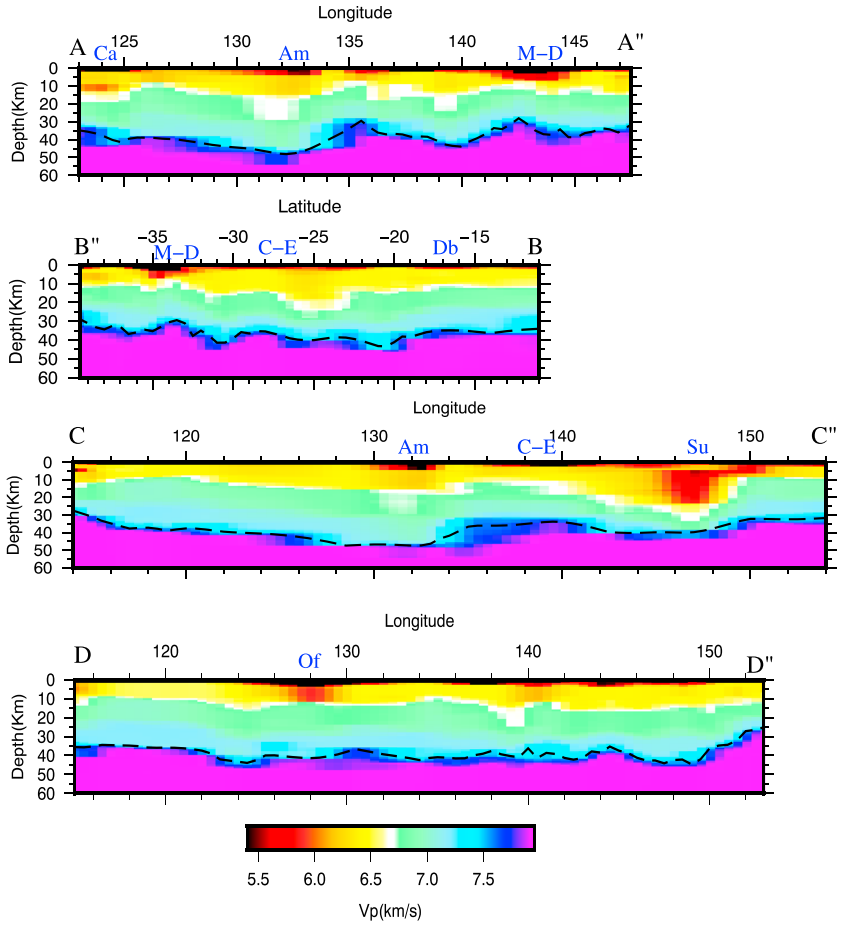

Figure 11. 2-D cross sections of V_p along profiles AA'' to DD''. Moho along each profile is shown by a black dashed line. The abbreviations are as given in Figure 10.

this depth, results of V_p from ambient noise tomography are not reliable as discussed in Salmon et al. (2013). Therefore, for depths ≥25 km, Salmon et al. (2013) used only the estimate of the V_p obtained from refraction studies with lower-weighted extra information from receiver function studies. Even though the AuSREM Moho model, obtained from seismic reflection and receiver function studies, has moderate coverage across the continent, the AuSREM V_p velocity model is less constrained by the limited coverage of seismic refraction experiments. However, we have better coverage across the entire continent by exploiting teleseismic *P* wave coda recorded on the permanent and portable seismic stations for constraining the V_p structure. An important point about the AuSREM model is that it will evolve over time when new data and methods become available. As an example, Kennett et al. (2015) updated the Moho depth of southeast Australia by spatially stacking of the crustal *P* wave reflectivity estimated from autocorrelograms. Recently, the Moho structure of Australia has been updated by Kennett et al. (2018) using the full range of available information by the end of 2016, including those extracted from many data points (seismic stations) that were not available for the model produced by Salmon et al. (2013).

Despite differences between our absolute V_p values and those from AuSREM, the general pattern of our V_p structure is consistent with the AuSREM crustal model (Figure 10). Although we have limited resolution at shallow depths, the signature of thick sediments is evident by low V_p at 5 km depth in the Canning, Officer, Amadeus, Cooper-Eromanga, and Drummond basins, as well as the Murray-Darling, Surat, and the Bowen basins. Arroucau et al. (2010) also found a low shear wave velocity anomaly near west of Murray-Darling basin at middle-upper crustal depths and related this anomaly to part of the infrabasin that is thicker and broader than previously imaged. The location of low V_p in the 5 km depth slice in the Canning basin also matches with the location of low V_s obtained from the ambient noise tomography of Saygin and Kennett (2012) at the same depth.

At 10 and 20 km depths V_p structure indicates high wave speeds in the Western Australia and a few localities across the continent. Interestingly, the influence of thick sediments is still visible at the Officer, Surat, and Canning basins. In general, the V_p values at depth slices of 10, 20, 30, and 40 km show high velocities in the range of 6.0–7.15 km/s. For the deep crust (e.g., 30 to 50 km depth), there is a good correlation between the pattern of V_p and Moho depth (Figure 9a) structures, indicating that areas with deeper Moho are characterized by lower seismic velocities and vice versa. A distinct feature of our V_p model and AuSREM crustal V_p structures is the narrow northwest/southeast trending zone of lower velocity (relative to its surroundings) on the southern borders of Kimberly block seen in the depth slice at 30–40 km. This zone approximately coincides with the location of the Fitzroy trough (Ft) that has a narrow zone of thick sediments (Figure 3 in Salmon et al., 2013).

Our velocities are higher than AuSREM velocities, as we incorporated the higher powers of V_p when linking density to V_p, which was not the case in Salmon et al. (2013). Also, we have not had to make any conversions between V_s and V_p. Another caveat is that in the case of gradational transition from the crust to the upper mantle, due to the employed parameterization, the inversion may slightly overestimate lower crustal velocities to fit the observed data by sampling the higher velocities to reduce the velocity contrast across the crust-mantle boundary. Therefore, the velocity distributions presented here likely indicate the upper bound of the actual crustal velocity distribution. Such differences are not surprising given the fact that the two V_p models have been estimated through (i) different data sets with different inherent sensitivities and (ii) different methods with dissimilar constraints or assumptions being used. Here we should emphasize that our approach does not depend on conversions between wave types and allows access to a wide range of frequencies. Also, except for very complex near-surface structures, the reflections extracted from autocorrelations are less contaminated by the free surface multiples compared to the receiver functions (Kennett et al., 2015; Kennett & Sippl, 2018; Sun & Kennett, 2016). Therefore, the inversion of these data can be more feasible than the inversion of receiver functions. However, the limitations of the inversion approach arise from the use of simplified model/parameterization as well as relatively high uncertainty due to single data set inversion, which will be addressed in future work.

8. Conclusion

We have developed a new approach for crustal imaging by utilizing teleseismic *P* wave coda autocorrelograms for a large number of permanent and temporary seismic stations. We have demonstrated that the Bayesian inversion of teleseismic *P* wave coda autocorrelograms is a useful and efficient approach to map deep crustal structure and can serve as a framework for future developments and improvements. Our synthetic tests show that we are able to recover a first-order estimate of the true crustal structure using this kind of data in the presence of seismic noise. We have applied the approach to all available permanent and portable seismic stations (over 1,200) across Australia to image Moho depth and crustal structure. The inversion results produce patterns of structures highly consistent with the previous crustal model obtained from many other seismological sources and methods. Our results confirm the utility of this approach, which removes the need for subjective picking of the Moho reflections in autocorrelation images. The framework used here can be further developed to taking into account the gradational transition from the crust to the mantle. The method is more cost effective than active seismic surveys/methods and can be used as an alternative tool for imaging shallow and deep crustal features when inversion of the receiver functions or deep seismic reflection profiling is not applicable.

Acknowledgments
Data recorded on permanent stations (operated by Geoscience Australia) were retrieved from IRIS Data Management Center. Data recorded by portable stations were archived at the Research School of Earth Sciences, Australian National University, for example, AQ3 (AusPass, 2014; see http://auspass.edu.au/research/AuSeis.html). We thank ANU RSES staff for providing seismic data recorded on portable stations. The crustal models obtained from the inversion and their associated uncertainties are available from the AusPass repository (http:/auspass.edu.au/). We thank Julia Pfeffer and the AusPass development team for their assistance in this regard. *ObsPy* (https://doi.org/10.1785/gssrl.81.3.530) were used to download and process seismic data. Figures were made using Python *Matplotlib* packages (https://matplotlib.org/) and the Generic Mapping Tools (GMT; Wessel et al., 2013). The authors wish to thank the Editor, Martha Savage, the Associate Editor, Andreas Fichtner, and the two reviewers, Anya M. Reading and Tuncay Taymaz, for their constructive comments that improved the quality of the manuscript.

References

Afonso, J. C., Rawlinson, N., Yang, Y., Schutt, D. L., Jones, A. G., Fullea, J., & Griffin, W. L. (2016). 3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle: III. Thermochemical tomography in the Western-Central U.S. *Journal of Geophysical Research: Solid Earth*, *121*, 7337–7370. https://doi.org/10.1002/2016JB013049

Agostinetti, N. P., & Malinverno, A. (2010). Receiver function inversion by trans-dimensional Monte Carlo sampling. *Geophysical Journal International*, *181*(2), 858–872. https://doi.org/10.1111/j.1365-246X.2010.04530.x

Ammon, C. J., Randall, G. E., & Zandt, G. (1990). On the nonuniqueness of receiver function inversions. *Journal of Geophysical Research*, *95*(B10), 15,303–15,318. https://doi.org/10.1029/JB095iB10p15303

Arroucau, P., Rawlinson, N., & Sambridge, M. (2010). New insight into Cainozoic sedimentary basins and Palaeozoic suture zones in southeast Australia from ambient noise surface wave tomography. *Geophysical Research Letters*, *37*, L07303. https://doi.org/10.1029/2009GL041974

AusPass (2014). AQ3 SW Queensland and NW New South Wales. The Australian National University. Dataset/Seismic Network. https://doi.org/10.7914/SN/4J_2014

Ball, J. S., Sheehan, A. F., Stachnik, J. C., Lin, F.-C., & Collins, J. A. (2014). A joint Monte Carlo analysis of seafloor compliance, Rayleigh wave dispersion and receiver functions at ocean bottom seismic stations offshore New Zealand. *Geochemistry, Geophysics, Geosystems*, *15*, 5051–5068. https://doi.org/10.1002/2014GC005412

Becker, G., & Knapmeyer-Endrun, B. (2018). Crustal thickness across the Trans-European Suture Zone from ambient noise autocorrelations. *Geophysical Journal International*, *212*(2), 1237–1254. https://doi.org/10.1093/gji/ggx485

Brocher, T. M. (2005). Empirical relations between elastic wavespeeds and density in the Earth's Crust. *Bulletin of the Seismological Society of America*, *95*(6), 2081–2092. https://doi.org/10.1785/0120050077

Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, *33*(2), 264–269. https://doi.org/10.1190/1.1439927

Clayton, R. (2018). Imaging the subsurface with ambient noise autocorrelations, SEG Technical Program Expanded Abstracts 2018, SEG Technical Program Expanded Abstracts, pp. 4852-4856, Society of Exploration Geophysicists. https://doi.org/10.1190/segam2018-2978420.1

Daneshvar, M. R., Clay, C. S., & Savage, M. K. (1995). Passive seismic imaging using microearthquakes. *Geophysics*, *60*(4), 1178–1186. https://doi.org/10.1190/1.1443846

Dueker, K. G., & Sheehan, A. F. (1998). Mantle discontinuity structure beneath the Colorado Rocky Mountains and High Plains. *Journal of Geophysical Research*, *103*(B4), 7153–7169. https://doi.org/10.1029/97JB03509

Fontaine, F. R., Tkalčić, H., & Kennett, B. L. N. (2013). Imaging crustal structure variation across southeastern Australia. *Tectonophysics*, *582*, 112–125. https://doi.org/10.1016/j.tecto.2012.09.031

Frasier, C. W. (1970). Discrete time solution of plane P-SV waves in a plane layered medium. *Geophysics*, *35*(2), 197–219. https://doi.org/10.1190/1.1440085

Gorbatov, A., Saygin, E., & Kennett, B. L. N. (2013). Crustal properties from seismic station autocorrelograms. *Geophysical Journal International*, *192*(2), 861–870. https://doi.org/10.1093/gji/ggs064

Haario, H., Laine, M., Mira, A., & Saksman, E. (2006). DRAM: Efficient adaptive MCMC. *Statistics and Computing*, *16*(4), 339–354. https://doi.org/10.1007/s11222-006-9438-0

Haario, H., Saksman, E., & Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, *7*(2), 223–242.

Hansen, S. E., Julià, J., Nyblade, A. A., Pyle, M. L., Wiens, D. A., & Anandakrishnan, S. (2009). Using *S* wave receiver functions to estimate crustal structure beneath ice sheets: An application to the Transantarctic Mountains and East Antarctic craton. *Geochemistry, Geophysics, Geosystems*, *10*, Q08014. https://doi.org/10.1029/2009GC002576

Hansen, S. M., & Schmandt, B. (2017). *P* and *S* wave receiver function imaging of subduction with scattering kernels. *Geochemistry, Geophysics, Geosystems*, *18*, 4487–4502. https://doi.org/10.1002/2017GC007120

Heath, B. A., Hooft, E. E., & Toomey, D. R. (2018). Autocorrelation of the seismic wavefield at Newberry Volcano: Reflections from the magmatic and geothermal systems. *Geophysical Research Letters*, *45*, 2311–2318. https://doi.org/10.1002/2017GL076706

Ito, Y., & Shiomi, K. (2012). Seismic scatterers within subducting slab revealed from ambient noise autocorrelation. *Geophysical Research Letters*, *39*, L19303. https://doi.org/10.1029/2012GL053321

Kennett, B. L. N. (1983). *Seismic wave propagation in stratified media*. Cambridge: Cambridge University Press.

Kennett, B. (2001). *The seismic wavefield, Volume I: Introduction and theoretical development*. Cambridge: Cambridge University Press.

Kennett, B. (2002). *The seismic wavefield, Volume II: Interpretation of seismograms on regional and global scales*. Cambridge: Cambridge University Press.

Kennett, B. L. N. (2015). Lithosphere-asthenosphere P-wave reflectivity across Australia. *Earth and Planetary Science Letters*, *431*, 225–235. https://doi.org/10.1016/j.epsl.2015.09.039

Kennett, B. L. N., Chopping, R., & Blewett, R. (2018). *The Australian continent: A geophysical synthesis*. Canberra: ANU Press and Geoscience Australia. http://doi.org/10.22459/AC.08.2018

Kennett, B. L. N., Engdahl, E. R., & Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, *122*(1), 108–124. https://doi.org/10.1111/j.1365-246X.1995.tb03540.x

Kennett, B. L. N., Kerry, N. J., & Woodhouse, J. H. (1978). Symmetries in the reflection and transmission of elastic waves. *Geophysical Journal International*, *52*(2), 215–229. https://doi.org/10.1111/j.1365-246X.1978.tb04230.x

Kennett, B. L., Salmon, M., Saygin, E., & Group, A. W. (2011). AusMoho: The variation of Moho depth in Australia. *Geophysical Journal International*, *187*(2), 946–958. https://doi.org/10.1111/j.1365-246X.2011.05194.x

Kennett, B. L. N., Saygin, E., & Salmon, M. (2015). Stacking autocorrelograms to map Moho depth with high spatial resolution in southeastern Australia. *Geophysical Research Letters*, *42*, 7490–7497. https://doi.org/10.1002/2015GL065345

Kennett, B., & Sippl, C. (2018). Lithospheric discontinuities in Central Australia. *Tectonophysics*, *744*, 10–22. https://doi.org/10.1016/j.tecto.2018.06.008

Kind, R., Yuan, X., & Kumar, P. (2012). Seismic receiver functions and the lithosphere–asthenosphere boundary. *Tectonophysics*, *536–537*, 25–43. https://doi.org/10.1016/j.tecto.2012.03.005

Kunetz, A., & d'Erceville, E. (1962). Sur certaines propriétés d'une onde acoustique plane de compression dans une milieu stratifié. *Annales de Geophysique*, *18*, 351–359.

Langston, C. A. (1979). Structure under Mount Rainier, Washington, inferred from teleseismic body waves. *Journal of Geophysical Research*, *84*(B9), 4749–4762. https://doi.org/10.1029/JB084iB09p04749

Laske, G., Masters, G., Ma, Z., & Pasyanos, M. (2013). Update on CRUST1.0—A 1-degree global model of Earth's crust. EGU General Assembly 2013.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092. https://doi.org/10.1063/1.1699114

Mira, A. (2001). On Metropolis-Hastings algorithms with delayed rejection. *Metron - International Journal of Statistics*, *0*(3-4), 231–241.

Nishitsuji, Y., Rowe, C. A., Wapenaar, K., & Draganov, D. (2016). Reflection imaging of the Moon's interior using deep-moonquake seismic interferometry. *Journal of Geophysical Research: Planets*, *121*, 695–713. https://doi.org/10.1002/2015JE004975

Oren, C., & Nowack, R. L. (2017). Seismic body-wave interferometry using noise autocorrelations for crustal structure. *Geophysical Journal International*, *208*(1), 321–332. https://doi.org/10.1093/gji/ggw394

Pham, T.-S., & Tkalčić, H. (2017). On the feasibility and use of teleseismic P wave coda autocorrelation for mapping shallow seismic discontinuities. *Journal of Geophysical Research: Solid Earth*, *122*, 3776–3791. https://doi.org/10.1002/2017JB013975

Pham, T.-S., & Tkalčić, H. (2018). Antarctic ice properties revealed from teleseismic P-wave coda autocorrelation. *Journal of Geophysical Research: Solid Earth*, *123*, 7896–7912. https://doi.org/10.1029/2018JB016115

Pham, T. -S., Tkalčić, H., Sambridge, M., & Kennett, B. L. (2018). Earth's correlation wavefield: Late coda correlation. *Geophysical Research Letters*, *45*, 3035–3042. https://doi.org/10.1002/2018GL077244

Phinney, R. A. (1964). Structure of the Earth's crust from spectral behavior of long-period body waves. *Journal of Geophysical Research*, *69*(14), 2997–3017. https://doi.org/10.1029/JZ069i014p02997

Randall, G. E. (1989). Efficient calculation of differential seismograms for lithospheric receiver functions. *Geophysical Journal International*, *99*(3), 469–481. https://doi.org/10.1111/j.1365-246X.1989.tb02033.x

Raymond, O. L., Totterdell, J. M., Stewart, A. J., & Woods, M. A. (2018). Australian Geological Provinces, 2018.01 edition [digital dataset]. Geoscience Australia, Commonwealth of Australia, Canberra. http://www.ga.gov.au

Romero, P., & Schimmel, M. (2018). Mapping the basement of the Ebro Basin in Spain with seismic ambient noise autocorrelations. *Journal of Geophysical Research: Solid Earth*, *123*, 5052–5067. https://doi.org/10.1029/2018JB015498

Ruigrok, E., & Wapenaar, K. (2012). Global-phase seismic interferometry unveils P-wave reflectivity below the Himalayas and Tibet. *Geophysical Research Letters*, *39*, L11303. https://doi.org/10.1029/2012GL051672

Rychert, C. A., & Harmon, N. (2016). Stacked P-to-S and S-to-P receiver functions determination of crustal thickness, Vp, and Vs: The H-V stacking method. *Geophysical Research Letters*, *43*, 1487–1494. https://doi.org/10.1002/2015GL067010

Salmon, M., Kennett, B. L. N., & Saygin, E. (2013). Australian Seismological Reference Model (AuSREM): crustal component. *Geophysical Journal International*, *192*(1), 190–206. https://doi.org/10.1093/gji/ggs004

Saygin, E., Cummins, P. R., & Lumley, D. (2017). Retrieval of the P wave reflectivity response from autocorrelation of seismic noise: Jakarta Basin, Indonesia. *Geophysical Research Letters*, *44*, 792–799. https://doi.org/10.1002/2016GL071363

Saygin, E., & Kennett, B. L. N. (2012). Crustal structure of Australia from ambient seismic noise tomography. *Journal of Geophysical Research*, *117*, B01304. https://doi.org/10.1029/2011JB008403

Smith, W. H. F., & Wessel, P. (1990). Gridding with continuous curvature splines in tension. *Geophysics*, *55*(3), 293. https://doi.org/10.1190/1.1442837

Sun, W., Fu, L.-Y., Saygin, E., & Zhao, L. (2018). Insights into layering in the cratonic lithosphere beneath Western Australia. *Journal of Geophysical Research: Solid Earth*, *123*, 1405–1418. https://doi.org/10.1002/2017JB014904

Sun, W., & Kennett, B. L. (2016). Receiver structure from teleseisms: Autocorrelation and cross correlation. *Geophysical Research Letters*, *43*, 6234–6242. https://doi.org/10.1002/2016GL069564

Sun, W., & Kennett, B. L. N. (2017). Mid-lithosphere discontinuities beneath the western and central North China Craton. *Geophysical Research Letters*, *44*, 1–9. https://doi.org/10.1002/2016GL071840

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

Taylor, G., Rost, S., & Houseman, G. (2016). Crustal imaging across the North Anatolian Fault Zone from the autocorrelation of ambient seismic noise. *Geophysical Research Letters*, *43*, 2502–2509. https://doi.org/10.1002/2016GL067715

Tibuleac, I. M., & von Seggern, D. (2012). Crust-mantle boundary reflectors in Nevada from ambient seismic noise autocorrelations. *Geophysical Journal International*, *189*(1), 493–500. https://doi.org/10.1111/j.1365-246X.2011.05336.x

Tork Qashqai, M., Afonso, J. C., & Yang, Y. (2016). The crustal structure of the arizona transition zone and southern colorado plateau from multiobservable probabilistic inversion. *Geochemistry, Geophysics, Geosystems*, *17*, 4308–4332. https://doi.org/10.1002/2016GC006463

Tork Qashqai, M., Afonso, J. C., & Yang, Y. (2018). Physical state and structure of the crust beneath the western-central United States from multi-observable probabilistic inversion. *Tectonics*, *37*, 3117–3147. https://doi.org/10.1029/2017TC004914

Ursin, B. (1983). Review of elastic and electromagnetic wave propagation in horizontally layered media. *Geophysics*, *48*(8), 1063. https://doi.org/10.1190/1.1441529

Wapenaar, K. (2004). Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation. *Physical Review Letters*, *93*(254), 301. https://doi.org/10.1103/PhysRevLett.93.254301

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., & Wobbe, F. (2013). Generic Mapping Tools: Improved version released. *Eos, Transactions American Geophysical Union*, *94*(45), 409–410. https://doi.org/10.1002/2013EO450001

Wittlinger, G., Farra, V., & Vergne, J. (2004). Lithospheric and upper mantle stratifications beneath Tibet: New insights from Sp conversions. *Geophysical Research Letters*, *31*, L19615. https://doi.org/10.1029/2004GL020955

Yuan, X., & Kind, R. (2016). Receiver functions with *S* waves. In M. Beer, I. A. Kougioumtzoglou, E. Patelli, & I. S.-K. Au (Eds.), *Encyclopedia of earthquake engineering* (pp. 1–16). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-36197-5_374-1

Zhu, L., & Kanamori, H. (2000). Moho depth variation in Southern California from teleseismic receiver functions. *Journal of Geophysical Research*, *105*(B2), 2969–2980. https://doi.org/10.1029/1999JB900322