

RESEARCH ARTICLE

10.1029/2019JB018900

Key Points:

- Three-dimensional shear velocity structure of the Klyuchevskoy area was determined using coupled transdimensional Monte Carlo inversions
- Slow velocity anomalies suggest magma storage beneath active volcanoes at 2–5 km depth (below sea level) but not in the midcrust
- Sediments filling the Central Kamchatka Depression are 8 km deep, consistent with an origin of the depression as a fore-arc basin

Supporting Information:

- Supporting Information S1
- Data Set S1
- Data Set S2

Correspondence to:

Christoph Sens-Schönfelder
sens-schoenfelder@gfz-potsdam.de

Citation:

Green, R. G., Sens-Schönfelder, C., Shapiro, N., Koulakov, I., Tilmann, F., Dreiling, J., et al. (2020). Magmatic and sedimentary structure beneath the Klyuchevskoy volcanic group, Kamchatka, from ambient noise tomography. *Journal of Geophysical Research: Solid Earth*, 125, e2019JB018900. <https://doi.org/10.1029/2019JB018900>

Received 15 OCT 2019

Accepted 14 FEB 2020

Accepted article online 15 FEB 2020

Magmatic and Sedimentary Structure beneath the Klyuchevskoy Volcanic Group, Kamchatka, From Ambient Noise Tomography

Robert G. Green¹ , Christoph Sens-Schönfelder¹ , Nikolai Shapiro^{2,3} , Ivan Koulakov^{4,5} , Frederik Tilmann^{1,6} , Jennifer Dreiling¹ , Birger Luehr¹, Andrey Jakovlev^{4,5} , Ilyas Abkadyrov^{4,7}, Dima Droznin⁸, and Evgeny Gordeev⁷

¹Department of Geophysics, Helmholtzzentrum Potsdam, Deutsches GeoForschungsZentrum (GFZ), Potsdam, Germany, ²Institut de Sciences de la Terre, Université Grenoble Alpes, CNRS (UMR5275), Grenoble, France, ³Schmidt Institute of Physics of the Earth, Russian Academy of Sciences, Moscow, Russia, ⁴Trofimuk Institute of Petroleum Geology and Geophysics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russian Federation, ⁵Department of Geology and Geophysics, Novosibirsk State University, Novosibirsk, Russian Federation, ⁶Institute for Geological Sciences, Freie Universität Berlin, Berlin, Germany, ⁷Institute of Volcanology and Seismology, FEB RAS, Petropavlovsk-Kamchatsky, Russia, ⁸Kamchatka Branch of the Geophysical Service, Russian Academy of Sciences, Petropavlovsk-Kamchatsky, Russian Federation

Abstract The Klyuchevskoy Volcanic Group is a cluster of the world's most active subduction volcanoes, situated on the Kamchatka Peninsula, Russia. The volcanoes lie in an unusual off-arc position within the Central Kamchatka Depression (CKD), a large sedimentary basin whose origin is not fully understood. Many gaps also remain in the knowledge of the crustal magmatic plumbing system of these volcanoes. We conducted an ambient noise surface wave tomography, to image the 3-D shear wave velocity structure of the Klyuchevskoy Volcanic Group and CKD within the surrounding region. Vertical component cross correlations of the continuous seismic noise are used to measure interstation Rayleigh wave group and phase traveltimes. We perform a two-step surface wave tomography to model the 3-D V_{sv} velocity structure. For each inversion stage we use a transdimensional Bayesian Monte Carlo approach, with coupled uncertainty propagation. This ensures that our model provides a reliable 3-D velocity image of the upper 15 km of the crust, as well as a robust assessment of the uncertainty in the observed structure. Beneath the active volcanoes, we image small slow velocity anomalies at depths of 2–5 km but find no evidence for magma storage regions deeper than 5 km—noting the 15 km depth limit of the model. We also map two clearly defined sedimentary layers within the CKD, revealing an extensive 8 km deep sedimentary accumulation. This volume of sediments is consistent with the possibility that the CKD was formed as an Eocene-Pliocene fore-arc regime, rather than by recent (<2 Ma) back-arc extension.

Plain Language Summary The Klyuchevskoy Volcanic Group is a cluster of 13 volcanoes on the Kamchatkan corner of the Pacific ring of fire. The volcanoes regularly produce large eruptions, but good knowledge of the magma plumbing system beneath the surface is still lacking. Why the Klyuchevskoy Volcanic Group volcanoes lie in the location they do, in a large low-lying depression, is also unexplained. We undertook a seismic experiment and used the data to produce a 3-D velocity image of the subsurface beneath the volcanoes and the depression. We found that small regions of slow seismic velocity are located beneath the active volcanoes, at 2–5 km depth below sea level. This slower velocity is probably caused by magma lying within the porous fracture spaces in this rock. The seismic velocities are much faster beneath the dormant volcanoes, suggesting they have no magma beneath them. With our velocity image, we also find that the Central Kamchatka Depression is very deep, filled with over 8 km of sediments. This supports an idea that the sediments accumulated as a fore-arc basin over many millions of years, since 40 Ma, when the active line of volcanoes was found 100 km to the west.

1. Introduction

1.1. Volcanism and Neotectonics of the KVG

The Klyuchevskoy Volcanic Group (KVG) lies on the Kamchatka peninsula in Russia at the intersection of the Kuril-Kamchatka arc and the transcurrent western end of the Aleutian arc (Figure 1a). Comprising 13

© 2020. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

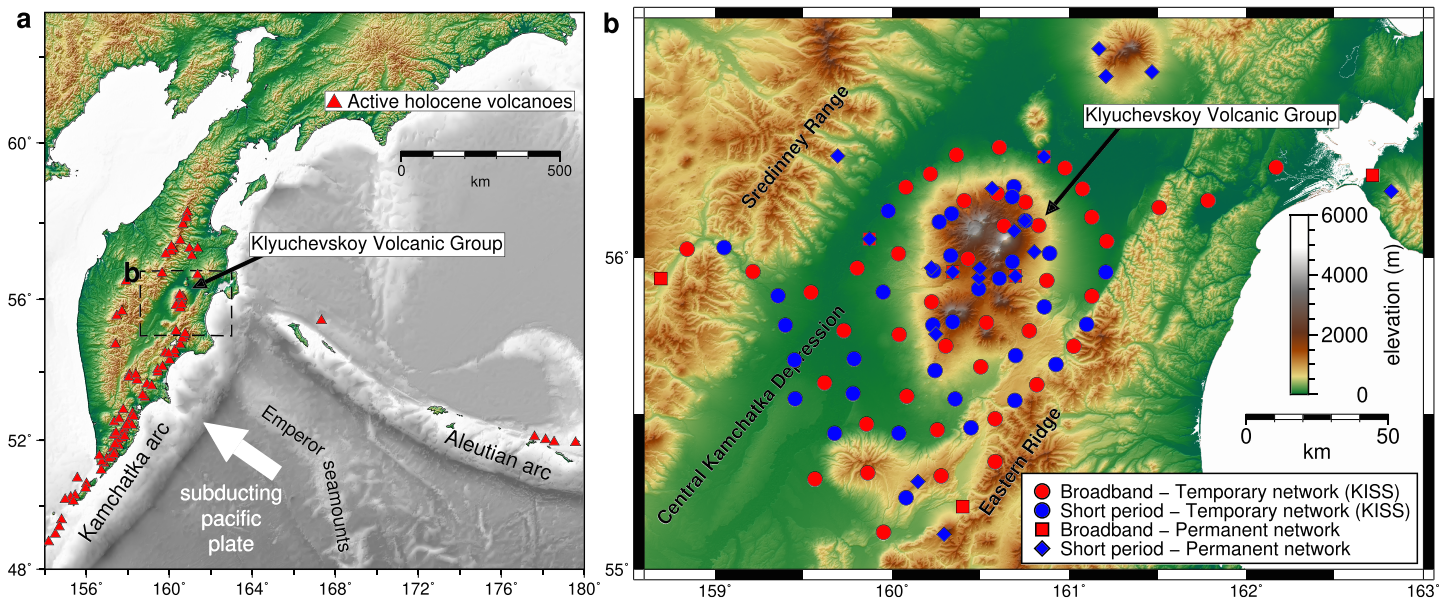


Figure 1. Regional tectonic setting a and local map b of the Klyuchevskoy volcanic group.

closely spaced dormant and active stratovolcanoes, it is one of the largest and most active clusters of subduction volcanoes in the world. Over the past 100 Kyr, the namesake of the group, Klyuchevskoy, has had an average eruption rate of 1 m^3 of rock per second (Fedotov et al., 2010). The extraordinary productivity can be related to the unique tectonic setting of the group (Figure 1a), situated not only on the northwestern cusp of the subducting Pacific plate but also just where the Hawaiian-Emperor seamount chain subducts beneath Kamchatka. Competing geodynamical models exist to explain the elevated volcanism, including mantle flow around the edge of the subducting slab, melting of the slab (Yogodzinski et al., 2001), enhanced fluid release from the thick highly hydrated seamount crust (Dorendorf et al., 2000), and slab detachment (Levin et al., 2002).

In addition to the high productivity, the group is offset westward from the current coastal volcanic arc, the Eastern Ridge (ER) (Figure 1b), possibly due to a decrease in dip of the subducting plate. This dip change is inferred from the shape of Wadati-Benioff zone seismicity (Gorbatov et al., 1997), but it is not resolved within current regional tomographic imaging (Koulakov, Dobretsov, et al., 2011). In its off-arc position, the KVG volcanic massif protrudes prominently from the topographically flat Central Kamchatka Depression (CKD), which is filled with alluvial and volcanoclastic sediments. The depression separates the active ER from the previously active volcanic arc, the Sredinny Range (SR) to the west.

Basement crust rarely outcrops across this area of Kamchatka due to the extensive surface volcanic deposits, but tectonic scale terrane accretion studies and erupted xenoliths show that the Olyutorsky arc terrane (commonly also referred to as the Achaivayam-Valaginskaya arc) makes up the upper crustal basement in this area of Kamchatka (Park et al., 2002). It is composed of cretaceous-paleocene sedimentary and volcanic rocks, emplaced by the Eocene obduction of the Olyutorsky arc (Hourigan et al., 2009; Konstantinovskaia, 2001; Shapiro & Soloviev, 2009) onto northeast Asia, which forms the majority of the Kamchatka peninsula. The tectonic genesis of the CKD and when (since the Eocene) subsidence initiated is not fully understood, but the first volcanic growth of the early KVG occurred at approximately 300 ka (Fedotov et al., 2010). Klyuchevskoy volcano is the youngest of the group at just 6 Kyr old (Braitseva et al., 1995). The 13 volcanoes of the KVG show a range of eruptive styles and compositions, from Hawaiian-type fissure eruptions (e.g., Tolbachik) to Andesitic explosive eruptions (e.g., Bezymianny). The massif is built of predominantly (four-fifth) basalt-andesite and about one-fifth andesite (Churikova et al., 2015; Fedotov et al., 2010).

Frequent volcano-tectonic seismicity is associated with magmatic activity at the KVG and is mainly focused beneath the Klyuchevskoy summit area (Fedotov & Zharinov, 2007; Khubunaya et al., 2018). During eruptions, continuous tremor signals can be located (Droznin et al., 2015), tracked (Soubestre et al., 2019), and

associated with a volcanic source based on their characteristic network-based tremor fingerprint (Soubestre et al., 2018). Long-period seismicity (Shapiro et al., 2017) occurs both at shallow depths, less than 5 km, and also at around 30 km depth, the inferred location of a gradational crust–mantle transition (Nikulin et al., 2010). The deep long period seismicity is more active during repose periods and surges in deep seismicity transfer into elevated shallow activity, which immediately precedes eruptions (Fedotov & Zharinov, 2007; Senyukov et al., 2009; Shapiro, Droznin, et al., 2017). The fast migration suggests good connectivity between deep magma reservoirs and the surface.

Past seismological imaging provides a multiscale view of the structure across the KVG and surrounding area. Regional tomographic imaging (Koulakov, Dobretsov, et al., 2011; Lees et al., 2007) has mapped the shape of the subducting Pacific plate, showing its termination just to the north of the KVG. Regional ambient noise tomography (Droznin et al., 2015) across the Kamchatka peninsula imaged shear wave velocities as low as 2 km/s in the Klyuchevskoy region.

The deep structure of the magmatic plumbing system beneath the KVG has been investigated previously with body wave tomography from local earthquakes recorded by the permanent monitoring network. Progressive studies have built on an ever-improving ray coverage through permanent network expansion and temporary deployments (Khubunaya et al., 2007; Koulakov et al., 2017; Koulakov, Gordeev, et al., 2011; Koulakov et al., 2013; Lees et al., 2007; Ivanov et al., 2016). The main features highlighted by these studies have been that of a moderate V_p , low V_s , and high V_p/V_s feature at approximately 30 km depth, coinciding with the abundant deep long period seismicity beneath Klyuchevskoy. The low-velocity anomaly also correlates well with both an elevated electrical conductivity anomaly (Moroz, 1991) and elevated apparent b value anomaly in the seismicity (Senyukov, 2013). Interpretations differ as to whether this anomaly lies in the crust (Khubunaya et al., 2018) or within the mantle [Levin et al. 2014], but there is wide consensus that the data evidence a large zone of magma accumulation (at 25–35 km depth), from which the focused volcanism of the KVG is fed. Petrological interpretations are also consistent with this as eruptive products indicate crystallization in a deep (~30 km) reservoir [Levin et al. 2014, Khubunaya et al., 2018].

From 3-D body wave tomographic imaging, some studies identify further low-velocity zones at intermediate depth (8–13 km) and shallow low-velocity zones just beneath the Klyuchevskoy edifice, which are also interpreted as potential magma reservoirs (Koulakov et al., 2013; Koulakov, Gordeev, et al., 2011). Indeed, a multilevel plumbing system seems necessary to explain the complex compositional variation within the KVG (Dobretsov et al., 2012). However, the tomographic evidence is hampered by low resolution in the upper crust that suffers from smearing of anomalies along the predominantly vertical raypaths. The petrological evidence for intermediate depth storage regions is also contrary. Ozerov et al. (1997) point toward an absence of significant accumulations above 20 km. In addition, Khubunaya et al. (2018) demonstrate strong evidence for the deeper reservoir and a shallow reservoir at ~5 km depth but no further intermediate storage.

As such, the complex nature of the magmatic system at intermediate and upper crustal depths remains unclear. This motivates our tomographic investigation imaging subsurface structures across a larger area than the immediate vicinity of Klyuchevskoy volcano, which the previously mentioned studies have been restricted to. To do this, we employ newly collected data from the first seismic survey of the entire KVG. Named KISS (*Klyuchevskoy Investigation - Seismic Structure of an Extraordinary Volcanic System*; Shapiro et al., 2017), this temporary passive seismic network was deployed at a close and regular station spacing of 10–15 km across a 150 by 150 km region. The network was deployed in collaboration with the Kamchatkan Branch of the Geophysical Survey of the Russian Academy of Sciences, such that the temporary stations were sited in complementary locations to the permanent network. In this study we present the first major results from this large data collection project using ambient noise tomography to construct a 3-D image of the shear wave velocity structure of the upper 15 km covering the entire KVG and its surrounding region.

1.2. Ambient Noise Tomography

Surface waves have long been used to constrain the subsurface shear velocity structure of the Earth, via the measurement and inversion of dispersive phase and group velocities. However, since the development of ambient noise interferometry to construct interstation signals, the ability to image regional and local structure on a scale of tens of kilometers and below has improved dramatically (Shapiro et al., 2005; Shapiro &

Campillo, 2004). Ambient noise-derived surface waves are usually energetically rich in periods of 5–20 s, therefore providing good sensitivity to at least the upper 20 km, which is ideally suited to our target of investigating the magmatic plumbing and crustal structure in the KVG.

1.3. Monte Carlo Surface Wave Tomography

Tomographic inversion is usually solved by a gradient-based optimization, where the seismic velocity model is iteratively adjusted to converge toward an “optimum” solution. The optimum solution is the velocity model with the best fit to the observed data, while adhering to regularization constraints (damping and smoothing), which condition the inversion and retain interpretability of the final model. The “tuning” or selection of these regularization parameters affects the recovered optimum model. It is therefore vital to evaluate the resolution and recoverability of key features in order to ascertain if they are stable and consistent under the choice of different “tuning” parameters. Even with sensitivity tests, a quantitative estimate of the velocity model uncertainty is difficult, as the effect of the regularization on the uncertainty is hard to quantify and because uncertainty estimates often neglect the nonlinear trade-offs between velocity anomalies and raypaths.

An approach to counter these shortcomings of regularized inversions is probabilistic methods, which search the model space with various strategies through vast numbers of forward simulations. In contrast to the “optimum model,” the solution is an ensemble of Earth models and their associated probabilities, which fully quantifies the degree of knowledge we have about the trade-offs in modeling the observed data. Computational restrictions prevent an exhaustive grid search, and as such, a popular sampling method in geophysical inversion is the Markov-chain Monte Carlo search (MCMC) with the Metropolis-Hasting acceptance rule (Gallagher et al., 2009; Malinverno, 2002). Here, many parallel, (semi-)independent processing chains generate Earth models and associated probabilities that sample the posterior probability distribution of the models, concentrating the sampling in the high probability space of the posterior distribution.

As a Bayesian probabilistic framework for surface wave tomography, Bodin and Sambridge (2009) proposed a transdimensional scheme using a reversible jump MCMC search (rjMCMC) (Green, 1995), where the number and location of irregularly shaped grid cells are free to vary through the Markov chains. A significant advantage of the transdimensional approach results from the “natural parsimony” of Bayesian inference. This is a Bayesian “Ockham’s Razor,” as the probability expressions favor simpler models over more complex models with equal data fit. Combined with the dynamic grid node parameterization, this results in a mesh that adapts such that the level of spatially variable detail is determined by the quality of retrievable information in the data. An explicit regularization to determine the level of smoothing is therefore not necessary, and the effect of grid size choices or raypath distribution is accounted for within the posterior distribution.

Data uncertainties, which are commonly unknown, can also be incorporated as nuisance parameters in the inversion (Bodin et al., 2012). The posterior distribution will contain it within reasonable variabilities and trade-off in both nuisance and model parameters that still allow an acceptable fit to the data. The philosophy of Monte Carlo tomography is therefore somewhat different to regularized inversions. One need not test the trade-off between data fit and model complexity or whether certain features can be recovered by the observations. This information is already contained within the posterior probability distribution and can be realized by (as a simplest projection) taking the mean and standard deviation of the posterior distribution as the expected value and associated uncertainty. In this framework, if anomalies lie below the magnitude of the uncertainty, then the inference is that they are not statistically relevant and should not be interpreted.

Here we follow this probabilistic approach and present the first surface wave tomographic imaging of the crustal structure of the KVG. We construct interstation empirical Green’s Functions through cross correlation of ambient seismic noise (section 2.1). Using two alternative methods, we make complementary interstation measurements of Rayleigh wave group traveltimes (section 2.2) and raypath average phase velocities (section 2.3). The 3-D shear velocity structure is obtained in a coupled two-step procedure, first determining period-wise group and phase velocity maps (section 3) and from these inverting for the 1-D V_{sv} structure with depth across many geographically local dispersion curves (section 4). An alternative approach of a one-step inversion to the 3-D model has been proposed and tested with synthetic data (Zhang et al., 2018), but the merits of application to geologically realistic structures remain under discussion. In the two-step procedure, we perform both inversion steps using a transdimensional Bayesian Monte Carlo search, and the

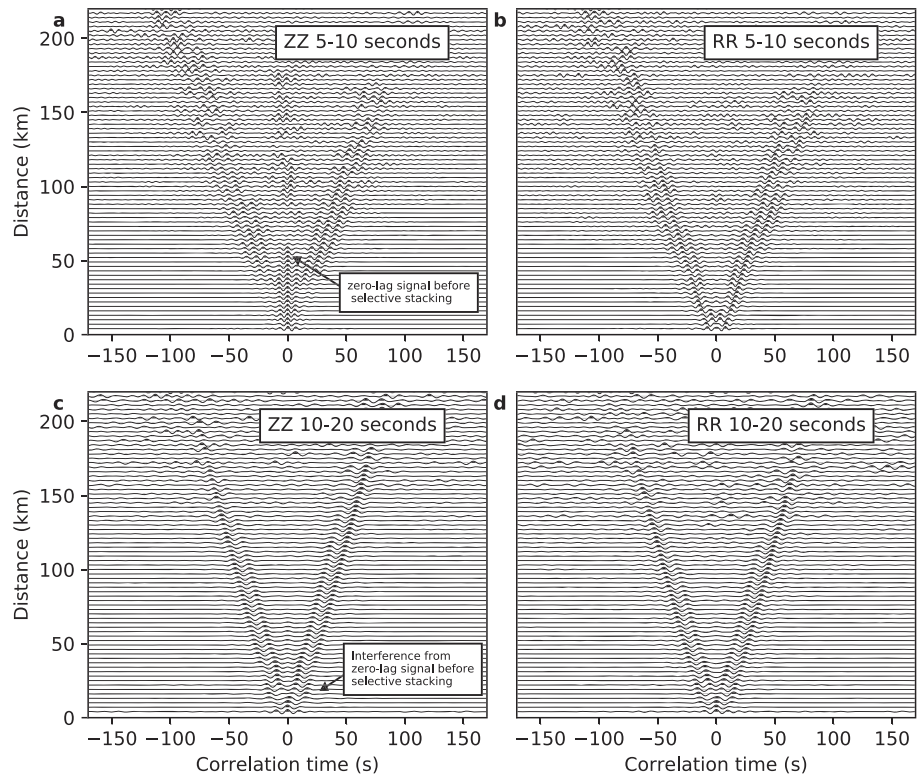


Figure 2. Cross-correlation trace record sections. Traces are stacked in 3 km intervals of interstation distance. Vertical components (a and c) show a contaminating signal from *P* wave noise sources at zero lag. It is vertically polarized and therefore absent from radial components (b and d).

uncertainty from the 2-D solutions is propagated as a relative data noise for the 1-D inversions in the second step.

2. Noise Cross Correlations and Rayleigh Wave Traveltime Measurements

2.1. Cross Correlation and Stacking

We use continuous three-component seismic records from the temporarily deployed KISS network and simultaneous records of the permanent seismic network of the Kamchatkan Branch of the Geophysical Survey of the Russian Academy of Sciences. The KISS experiment deployed 30 Trillium Compact (120 s), 14 Guralp CMG-6 T/6TD (30 s), 9 R-Sensors CME-4111, and 30 Mark L-4C-3D (1 s), for 1 year from summer 2015 to summer 2016. While the Mark instruments are 1 s passive sensors, they are calibrated before each experiment and have proven to be useful for ambient noise records up to 20 s (Jaxybulatov et al., 2014). We find this to be the case again in this study and use them as semibroadband sensors. The CME sensors are passive instruments with a reported flat response between 60 s and 50 Hz, but they do not perform well at periods longer than 1 s, and we treat them as short-period sensors.

From the permanent Kamchatkan Branch of the Geophysical Survey of the Russian Academy of Sciences network, we use data from 7 broadband stations and 18 short-period stations. Broadband stations are equipped with Guralp 3 T (120 s) and Guralp 6 T (30 s) seismometers. Short-period stations are equipped with three-component SM3KV velocimeters (0.7–20 Hz). Data from the remotely located short-period stations are transmitted via analog FM radio telemetry, and the combined instrument-transmission response is regularly calibrated. All data for the experimental period are archived at the GEOFON data center (Shapiro et al., 2015) along with experiment and data preparation reports. In total (after accounting for station losses), we use continuous three-component records of 50 broadband, 27 semibroadband, and 28 short-period stations.

The procedure to compute daily three-component ambient noise cross correlations is described in the supporting information (Text S1). In this study we concern ourselves with Rayleigh wave tomography and

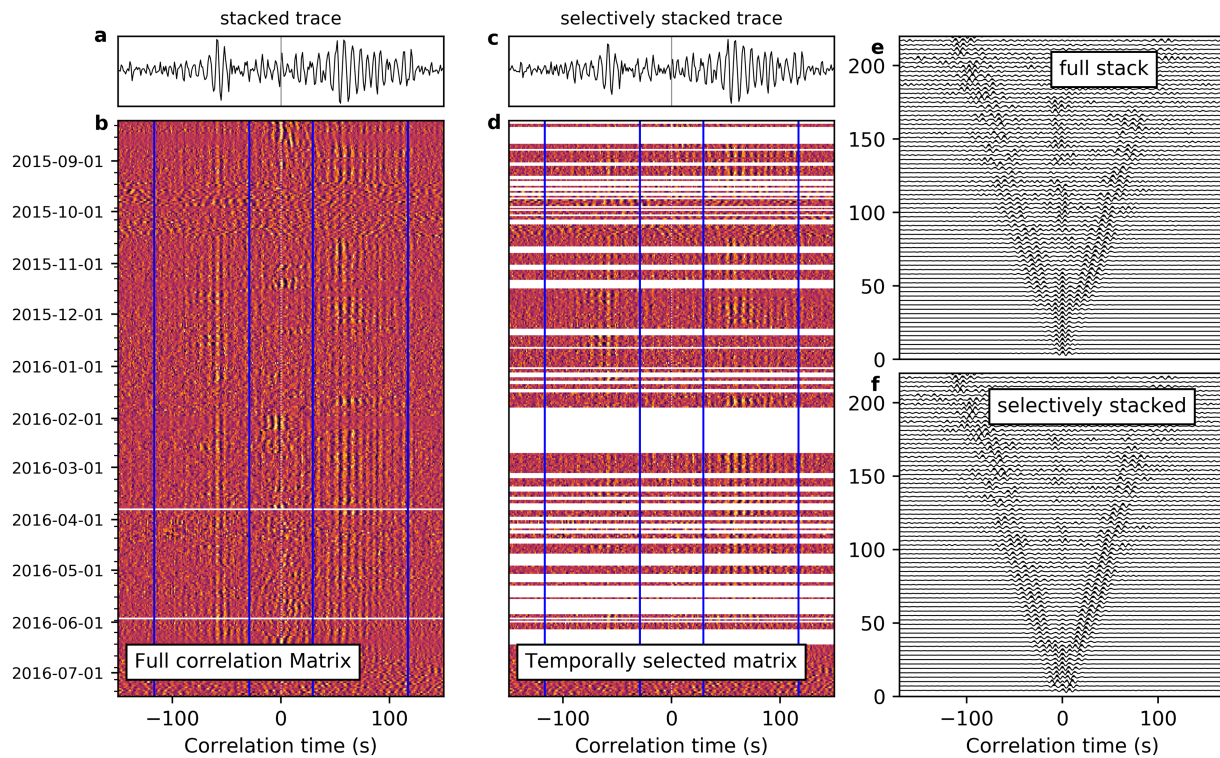


Figure 3. Temporal selective stacking (vertical components). Correlation matrix (b) displays daily correlation traces through the deployment for pair IR10-LGN (see Figure 4i). Blue lines indicate the Rayleigh wave arrival band. Time periods where the high amplitude zero-lag pulse dominates the trace are removed from matrix (d). When traces are linearly stacked (c), the energy at zero lag is suppressed relative to full stack (a). The data set-wide removal is demonstrated by record sections (e and f).

perform further analysis on vertical component cross correlations. After linear stacking of all daily cross-correlation functions, the traces show prominent Rayleigh waves dominating the interstation empirical Green's Functions (Figure 2).

Cross-correlation traces also show a strong pulse at 0 s lag time in the vertical but not horizontal components (Figures 2a and 2c). The strength of this zero-lag signal is variable but sometimes persistently strong over time periods of days and sometimes weeks (Figure 3b) and is the result of the correlation of strong *P* wave energy, originating not from earthquakes but most likely from microseism ambient noise sources at teleseismic distances. Analysis of similar zero-lag signals has confirmed their origin as strong storm-generated microseism events and located the source regions at teleseismic distances with beamforming techniques (Landès et al., 2010; Retailleau et al., 2018). However, as this signal contaminates the observation of interstation surface waves especially at short distances, we removed days that show a strong zero-lag pulse from the stack (Figure 3d), enabling reliable measurement of Rayleigh wave traveltimes at closer separation distances (Figure 3f).

We investigated the use of both linear and phase weighted stacking (PWS) (Schimmel et al., 2011; Ventosa et al., 2017) of the daily cross-correlation traces to provide a single high-quality trace for each interstation pair. Through a cross validation of the two methods (reviewed in Figure S1), we found that PWS provides reliable Rayleigh wave dispersion measurements, very similar to those obtained through linear stacking for paths with large signal-to-noise ratio (SNR), but enables the analysis of low SNR pairs. PWS is therefore used for picking Rayleigh wave group traveltimes (section 2.2), but linear stacking is used for phase velocity measurements (section 2.3) to retain unbiased observations of the phase.

2.2. Interstation Group Velocity Measurements

Group traveltimes of the interstation Rayleigh waves are measured using Frequency Time Analysis (Dziewonski et al., 1969; Levshin et al., 1989) of the vertical component traces. This well-established method involves performing a time-frequency transformation on the time domain correlation traces with a series of

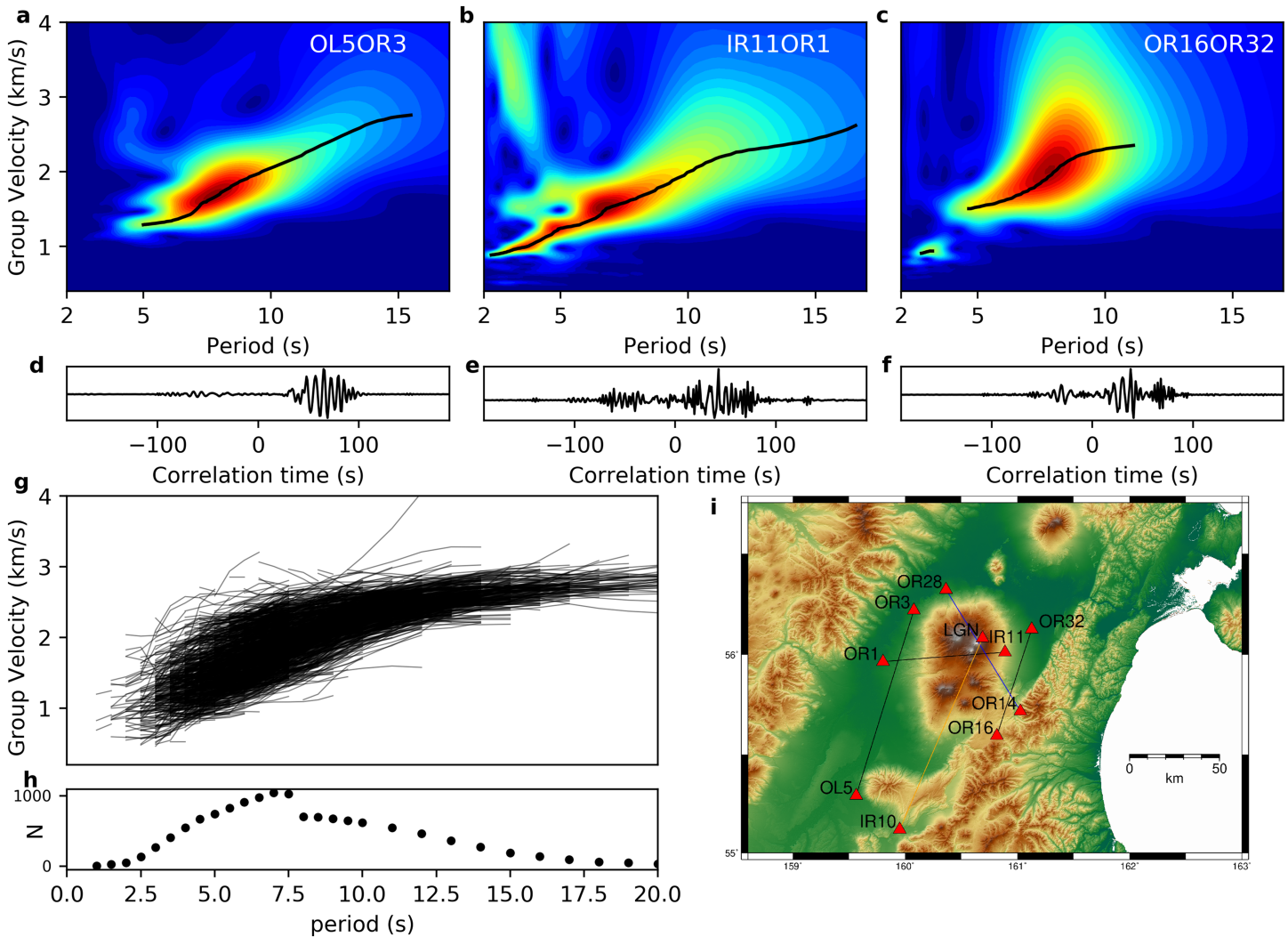


Figure 4. Interstation group velocity measurements. (a–c) FTAN amplitude diagrams for cross-correlation traces. (d–f) Black lines show dispersion curves at instantaneous frequency. (g) all manually picked dispersion curves. (h) Number of observations per period. Map (i) shows station pairs for this figure (black lines), Figure 3 (orange), and Figure 5 (blue).

narrow band filters and measuring the arrival time of the peak in the envelope. Where interstation group velocities are referred to (rather than traveltimes), these assume the distance along a straight-line raypath. Due to waveform complexities from scattering and multipathing and the goal to measure group traveltimes at periods as short as 2 s, we manually picked the group velocity dispersion curves. Automatic algorithms to track the dispersion ridge did not perform well as a result of sharp gradients and discontinuities. A highly responsive and efficient software program (XDC) (Ryberg et al., 2017) was optimized for this manual task, such that picking the large data set was feasible and efficient.

As shown later, we observed Rayleigh wave dispersion curves with large gradients of the velocity along the period axis. These are due to the high contrasts along low velocity basins and required an additional optimization of the Frequency Time Analysis filters. We modify the filter bandwidth such that rather than using a fixed percentage of the center frequency, the fractional width of the filter increases to lower frequencies. This inhibits oscillations in the dispersion curves at long periods while allowing a narrower width at short periods to follow sharp gradients and minimize discontinuities. Once the width variation was defined (from 62% at 0.025 Hz to 30% at 0.25 Hz), the same filters were applied when picking every Rayleigh wave. Where velocity gradients were so large that they could not be resolved and the dispersion ridge appear to jump (e.g., Figure 4 c), we treated each segment of the dispersion curve separately. After picking the dispersion curves, the

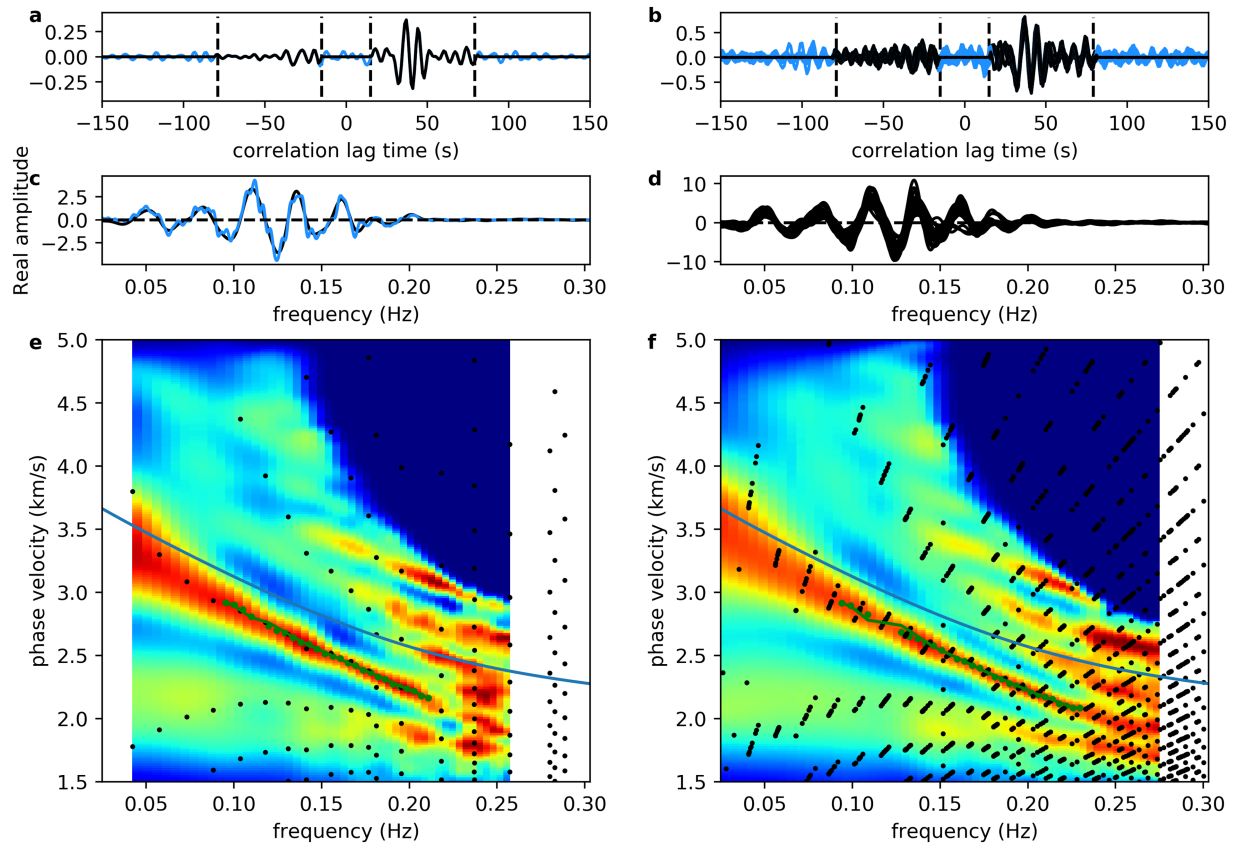


Figure 5. Phase velocity measurement technique. Time domain correlation traces of pair OR14-OR28 (pair location in Figure 4i) for complete stack (a) and monthly stacks (b). Black lines show the traces after windowing. (c,d) Frequency domain correlation spectra for traces in (a) and (b) which are smoother after windowing (black). (e) Density heat map of frequency-velocity estimates from zero crossings of (c). Zero crossings are black circles, blue line is a reference curve, and the picked dispersion curve is show in green. (f) Density heat map of all frequency-velocity estimates from all zero crossings of (d). Picked dispersion curve can be traced to higher frequencies.

central frequency of each filter is replaced with the time derivative of the phase at the picked group arrival time, known as the instantaneous frequency of the surface wave signal within that passband (Bensen et al., 2007; Levshin et al., 1989). This change is often small but notable as a small shift in the black dispersion curve on Figures 4a–4c relative to the center of the amplitude peak.

The target set of correlation traces is defined using a SNR quality criteria. We require the coda SNR (amplitude ratio between surface wave and noise window within the coda) to be greater than 8.0 and the zero-lag SNR (noise window before the picked Rayleigh arrival) to be greater than 3.0. Traces with lower quality than these thresholds showed poor emergence of Rayleigh wave signals. Two-thousand thirty-three group dispersion curves are subsequently obtained by manual picking of the target pairs (Figure 4). We exclude travel-time observations of less than 20 s (where any remaining central pulse might still interfere) and discard observations where the wavelength exceeds half the interstation distance. We also trialed excluding observations with wavelength above one third of the interstation distance, but the tomographic inversions had similar residuals. Any pairs containing a short-period station are limited to periods below 7.5 s, resulting in the step in Figure 4h. Usable observation sets are available between the period range of 3 to 15 s, and the largest traveltime observation set of 1,043 measurements occurs at 7.0 s period (Figure 4h). Straight-ray interstation group velocities vary by up to 1.5 km/s at 5 s period. From this we expect to observe extreme heterogeneity in tomographic models.

2.3. Interstation Phase Velocity Measurements

Path average interstation Rayleigh wave phase velocities are measured from the correlation spectrum (the cross-correlation trace in the frequency domain) (Aki, 1957) following the approach popularized by

Ekström et al. (2009). Correlation spectra are smoothed by initially time windowing the surface wave arrivals in the time domain (Figures 5a and 5b). The zero crossings of the observed spectra (Figure 5c) are then associated with the zeros of a Bessel function of the first kind, to provide phase velocity estimates (with 2π ambiguity) at discrete frequencies (black dots on Figure 5e). The key procedural stage of the dispersion curve measurement is then the elimination of the 2π ambiguity by identifying the discrete velocity-frequency points that define a reasonable phase velocity dispersion curve. In short, this means identifying where to ignore erroneous zero-crossing data points, caused by spurious noise in the correlation spectra.

We use an automatic picking algorithm modified from Kästle et al. (2016), which implements an irregular moving Gaussian filter over the frequency-velocity space to define a density heat map of the discrete frequency-velocity estimates from the zero crossings. From this derived density (colored background of Figure 5e), a smooth dispersion curve is then picked (green line), following the same constraints as Kästle et al. (2016) on continuity, and difference of gradient relative to a reference curve. This algorithm has been shown to be very robust with noisy data.

We extend the method to be more robust for our data set, using the repeatability of the noise correlation measurements. The strategy is to use many measurements from shorter noise correlation stacks rather than measuring a single dispersion curve on the stack of all correlation traces. This idea has been applied successfully in Sens-Schönfelder and Eulenfeld (2019) for relative velocity measurements. For many monthly stacked correlation traces (Figure 5b), we calculate the discrete velocity-frequency estimates from the zero crossings. The scanning algorithm is then applied over many more discrete zero-crossing velocity estimates (dots on Figure 5f) and is therefore much more robust to individual erroneous zero crossings when more points are present in frequency-velocity space. This helps distinguish between the false branches of the dispersion curve and in the case of Figure 5 allows to follow the dispersion curve over an extended frequency range.

Despite the use of this phase velocity algorithm, we found that the stability of the phase velocity picking did not allow to pick as many interstation paths as for group velocities. The method is sensitive to the initial windowing in the time domain when the wave packet has a long and complex shape (e.g., Figure 4b). This is the case for the many interstation pairs that sample the river basins on either side of the main massif of the KVG. Therefore, we applied strict quality control criteria to phase velocity measurements. The velocities are required to be stable to less than 2% variation between both a range of starting reference curves and a range of time-windowing filters. The maximum frequency of the dispersion curve is limited to when a single cycle skip error becomes less than 0.1 km/s (unlikely for picking algorithm to avoid), and the minimum frequency is limited by the criteria that interstation distance exceeds 1.5 times the wavelength. The accepted and rejected raypaths for these criteria are displayed in Figures 6c and 6d. The available observation set is significantly smaller than for group velocities, with a peak observation count of 177 at 6.5 s period (Figure 6).

3. Two-Dimensional Tomography

3.1. Transdimensional rjMCMC Algorithm With Mixed Noise Distribution

As motivated in section 1.2, we use a Bayesian framework to invert single-period interstation traveltime observations. The “inversions” (for they are not truly inversions but rather searches) follow the rj-MCMC implementation of Bodin et al. (Bodin, Sambridge, Rawlinson, & Arroucau, 2012) in which the 2-D model is parameterized by a variable number of constant slowness Voronoi cells. In addition to model values (grid node position and slowness), the data noise (or uncertainty) is a further parameter to be varied through the Markov chains. We follow Tilmann et al. (2019) and express the data uncertainty distribution as a mixture of a Gaussian (of width, σ , representing measurement uncertainties) and a uniform distribution once the Gaussian falls below the uniform probability (which is defined by f , the fraction of the total number of observations which are outliers). The mixed distribution allows observations with residuals lying outside the Gaussian to be classed as outliers. Changing the model will then not affect their contribution to the probability, as can occur when all noise is treated as Gaussian. σ and f are the two nuisance (nonmodel) parameters to be determined along with the model parameters in the Monte Carlo search. We set a uniform model prior distribution between 75% of the minimum interstation slowness and 125% of the maximum interstation slowness. Priors for the numbers of cells and noise parameters are set at practical limits far beyond realistic values.

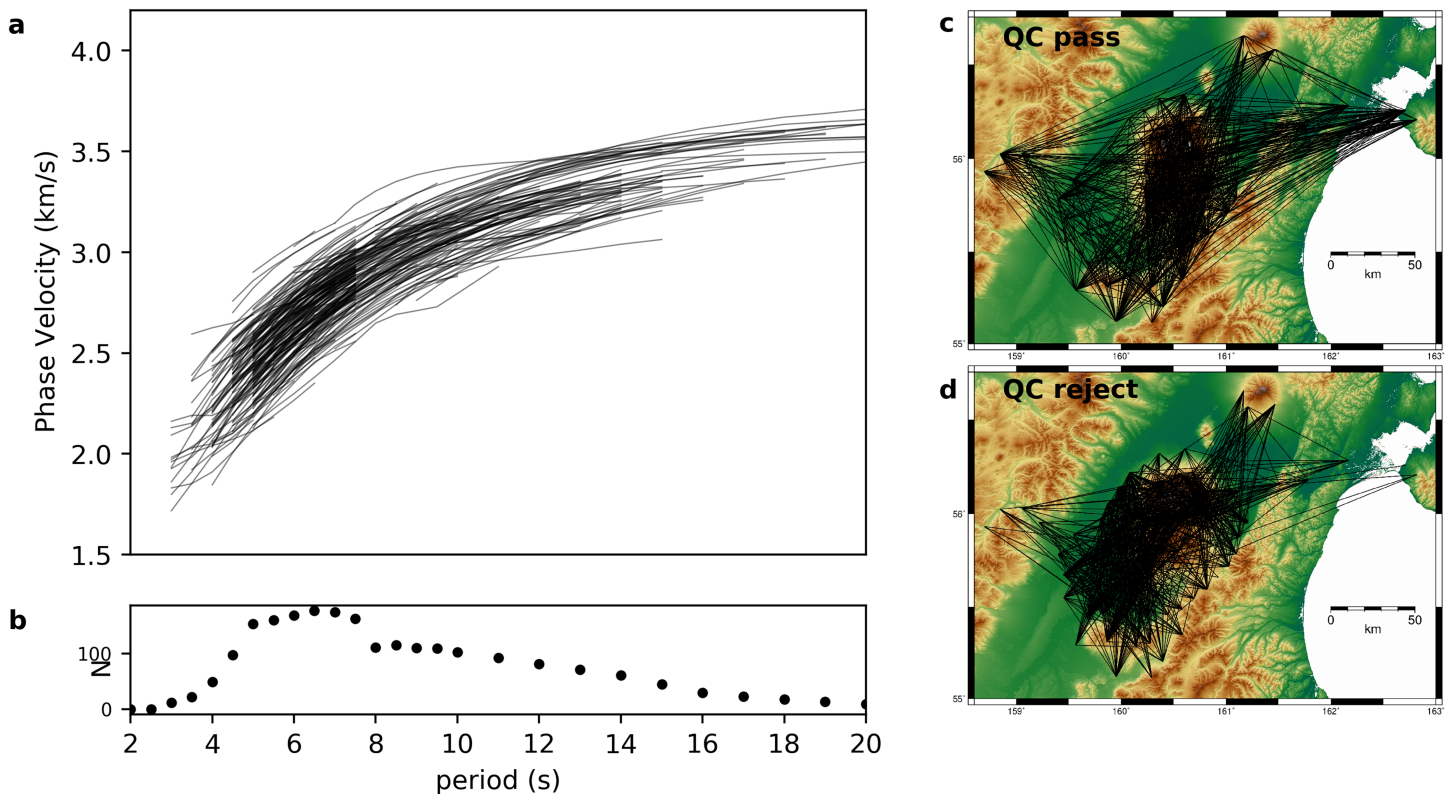


Figure 6. Interstation phase velocity measurements. (a) Dispersion curves which passed quality control. (b) Number of observations per period. Pairs that passed the quality control procedure shown on map (c) and pairs that failed on map (d).

A chain of models is generated following this algorithm:

1. Generate a starting model (\mathbf{m}) randomly from the prior distribution.
2. Generate a trial model (\mathbf{m}') by choosing randomly from one of the six possible perturbation types (move node, perturb cell slowness, add node, delete node, perturb outlier fraction, and perturb data noise).
3. Calculate the acceptance ratio based on the likelihood of the current and trial models.
4. Generate a random number from 0 to 1, and if less than the acceptance ratio, accept the trial model as the new current model ($\mathbf{m} = \mathbf{m}'$). A likelihood increase for \mathbf{m}' will always be accepted, and a moderate likelihood decrease for \mathbf{m}' will sometimes be accepted. Otherwise, reject the trial model (\mathbf{m}') and retain the current model (\mathbf{m}).
5. Return to (2) and iterate.

Following this algorithm, the Markov chains converge effectively to concentrate their sampling within the higher probability region of the posterior distribution. Models within the Markov Chains are not independent, and so unless run for an extremely long time, a single chain will be too focused within model space. Therefore, the saved models within the chains are thinned (to every 100 iterations), and 60 chains are run in parallel, such that the ensemble of models efficiently sample the posterior distribution function. Chains are run in four independent subsets of 15 chains, and within each subset, we use parallel tempering (Sambridge, 2014) to accelerate convergence and discourage local minima. This involves running four of the 15 chains in each subset (16 out of 60 in total) at “higher temperatures,” meaning that the proposal distribution for each Markov chain iteration is up to 2.8 times wider. Within each subset of 15 chains, the chains communicate every 20,000 iterations, and if a high-temperature chain has a higher-probability model than any low-temperature chain (no change to proposal distributions), then the chains swap models. Every 200,000 iterations, we recompute the raypaths by ray tracing through the mean velocity model of the current ensemble, using the Fast-Marching-Method (Rawlinson & Sambridge, 2004), an eikonal ray solver. Updating the raypaths is especially important for the data inverted here, as the strong velocity heterogeneity has a profound impact on ray geometry. We run the chains for 1 million iterations. The first half (500,000

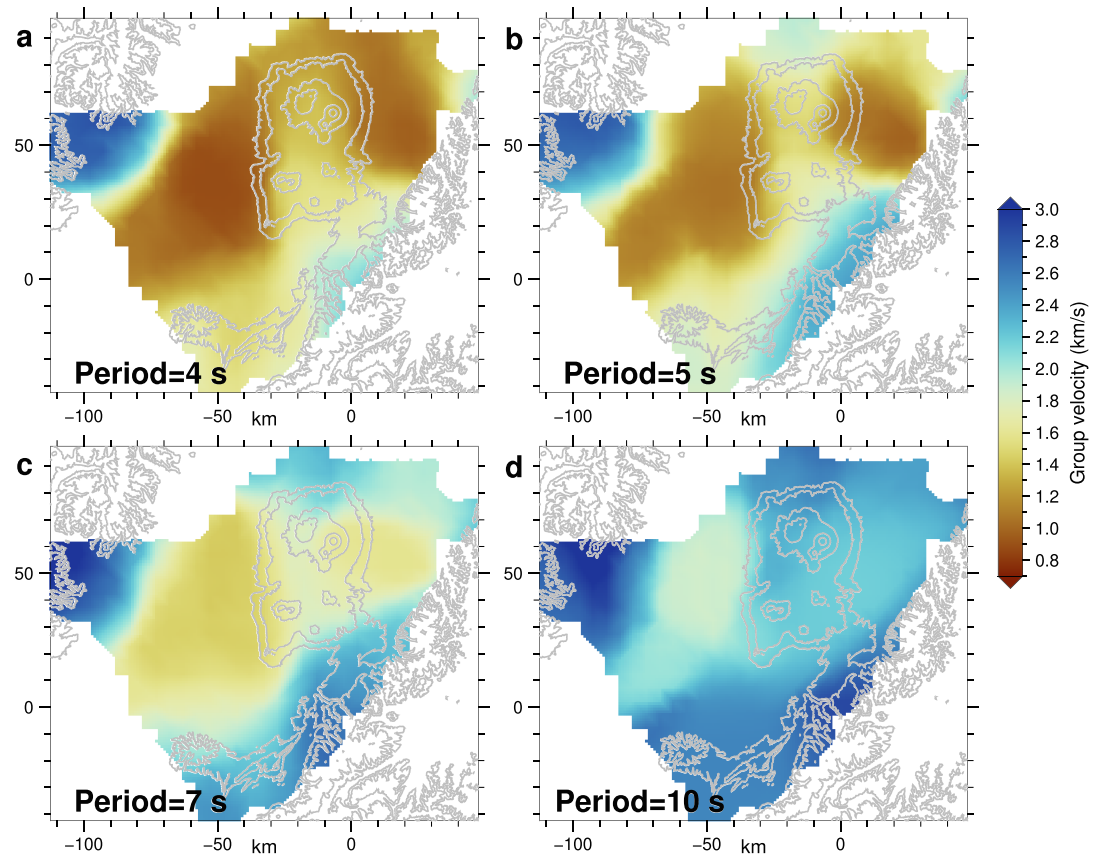


Figure 7. Group velocity maps for (a) 4, (b) 5, (c) 7 and (d) 10 seconds period. Mean of the posterior probability density function from rjMcMC tomography. Gray lines are 1,000 m contours. All period-wise results and standard deviations are in figures S3 and S4.

iterations) are discarded as the burn-in phase, which comfortably accounts for the iterations required to converge to the high-probability region of the posterior distribution. We also class the 100,000 iterations after each subsequent raypath recomputation as burn-in, as the models can take some time to adjust when the raypaths change. The subsequent posterior is formed from only low-temperature chains, constituting an ensemble of 13,200 models (after thinning).

3.2. Topographic Considerations

The high relief of the KVG poses a concern for how the topography might affect our tomographic models. On the one hand, the relief will impact on raypath distance as a purely horizontal distance will differ from a ray following the topographic surface. On the other hand, the topographic surface may also influence surface wave propagation, through scattering and interference effects, and dispersion measurements may experience a resultant bias. A discussion of these considerations is included within the supporting information and concludes that the biases' from topography are small compared to velocity variations imaged in this tomography. As such, a simpler treatment of rays that follow a horizontally flat (but nevertheless curved) path is appropriate.

3.3. Rayleigh Group and Phase Velocity Results

We perform the 2-D Monte Carlo tomography for group velocities between 3 and 15 s period (Figures S3 and S4) and for phase velocities between 6 and 20 s period (Figures S5 and S6). Once rays have been traced through the average solution of the final ensemble, we restrict the period range for further analysis due to incomplete ray coverage at the low and high ends of the period range. The average velocity images of these discarded periods are also incoherent with adjacent periods, showing either smearing or translations in the

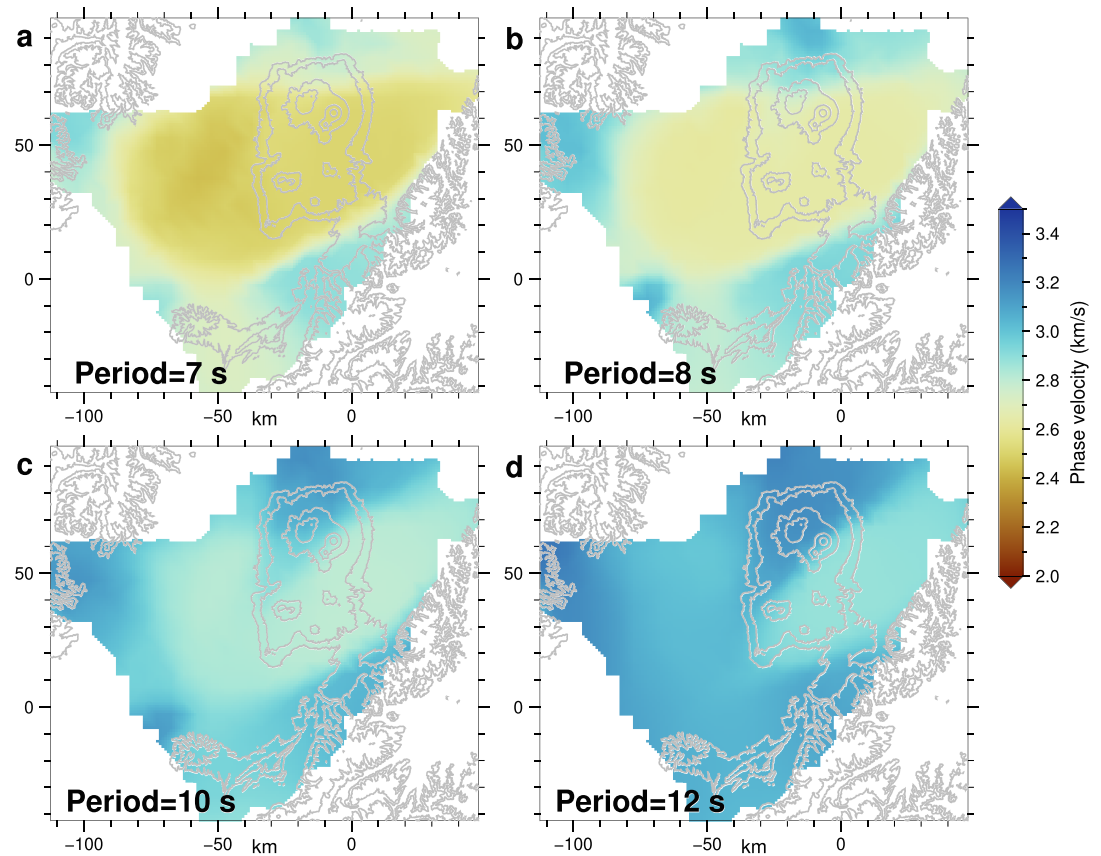


Figure 8. Phase velocity maps for (a) 7, (b) 8, (c) 10 and (d) 12 seconds period. Mean of the posterior probability density function from rjMCMC tomography. Gray lines are 1,000 m contours. All period-wise results and standard deviations are in figures S5 and S6.

position of anomaly boundaries. For group velocities, we retain periods between 4 and 14 s (Figure 7), and for phase velocities, we retain periods between 7 and 15 s (Figure 8).

4. Inversion of Local Dispersion Curves

In the second step, local dispersion curves at geographical grid nodes of 5 km spacing are inverted for a 1-D shear velocity structure with depth. The period-wise velocities and errors are selected from the mean and standard deviation of the posterior distributions of the 2-D lateral inversions, to define both phase and group velocity dispersion curves at each node (Figures 9a and 9b). Errors are propagated as relative uncertainties for each individual 1-D shear velocity inversion. The inversions follow the same transdimensional rjMCMC algorithm as in the 2-D tomography but with a 1-D Voronoi model of shear wave velocity (e.g., models in Figures 9c and 9d). Inversions are implemented using the recently developed BayHunter package (Dreiling & Tilmann, 2019). Dispersion curves are predicted using the forward modeling routine of the software package Computer Programs in Seismology (Herrmann & Ammon, 2004) and use a fixed V_p/V_s ratio of 1.75.

A prior of 0.5–5.0 km/s is used for shear wave velocity, 1–20 for the number of cells, and 1–30 km for Voronoi node depth. A joint likelihood is formulated as in Bodin et al. (Bodin et al., 2012) and Dreiling and Tilmann et al. (2019), incorporating the fit to both the phase and group velocity local dispersion curves. There is no need to define a relative weight of either data set, as this is accounted for by having a separate noise parameter for each dispersion curve. As a joint inversion, the data noise is defined with a covariance matrix (Bodin, Sambridge, Tkalčić, et al., 2012). For surface wave dispersion, we assume the noise to be uncorrelated between data points, such that the covariance is a diagonal constant matrix multiplied by the scalar noise parameter (σ). As in the 2-D inversion, the noise parameters are varied as part of the Markov

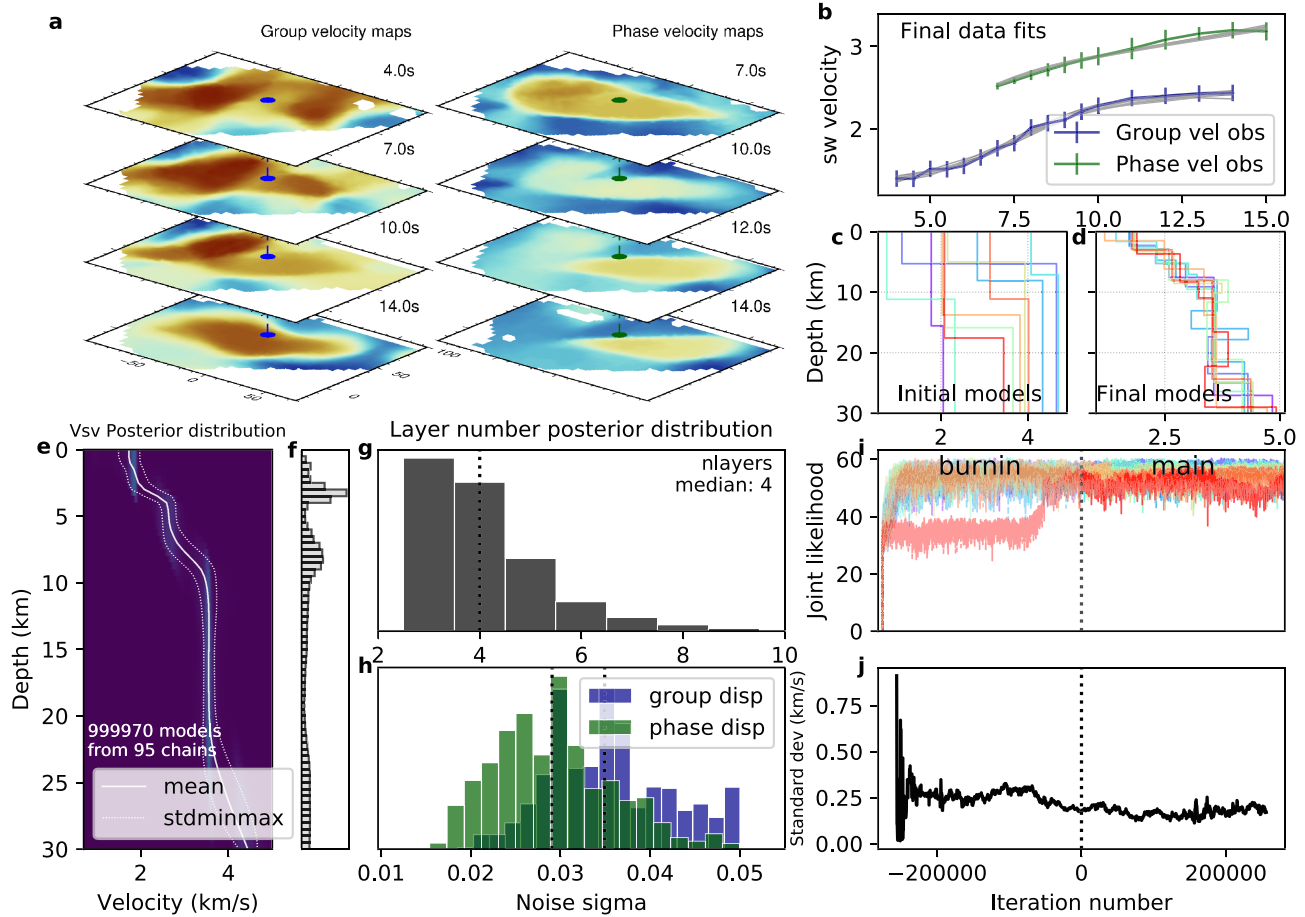


Figure 9. 1-D inversion strategy. (a) Local dispersion curves for each geographical point extracted from period-wise group and phase velocity maps. (b) Local phase and group dispersion curves for grid point $(-17.5, 52.5)$ marked in (a) with the final forward model from chains 1–10 (gray lines). Initial (c) and final (d) velocity models of chains 1–10. (e) Posterior distribution of Vsv models as 2-D colored histogram. One chain out of 100 is an outlier. (f) Posterior distribution of interface depths in Vsv models. (g) Posterior distribution of number of layers in models. (h) Posterior distribution of noise sigma for group and phase. (i) Joint likelihood evolution of chains 1–10. (j) Evolution of interchain standard deviation between current model velocity at 5 km depth.

chains, but here the diagonals of the covariance matrix are replaced by normalized weights, defined from the uncertainties of the 2-D velocity maps at the respective location.

We run the Monte Carlo search for 500,000 iterations with 100 independent chains. For the first 1% of iterations, we restrict the MCMC algorithm so that an increase in dimensions (number of layers) is prohibited. This allows the simplest model of one layer over a half-space to adjust to reasonable shear velocity values before more layers are added. The first 250,000 iterations are discarded as burn-in (Figures 9i and 9j). In the ensemble of models that constitute the posterior distribution, the layer distribution (Figure 9g) is non-Gaussian and tends to show a mode of three layers at most inversion locations.

Outlier chains that become stuck sampling a local maximum of the likelihood function rather than the global maximum are identified based on two selection methods. First, we discard individual chains with a median likelihood that differs by more than a certain threshold (40) from the median of all chains. Second, we identify outlier chains where a forward modeling numerical instability occurs, by selecting any chain where the model has a shear velocity of greater than 4.0 km/s in the top layer. These models are generally underlain by large low velocity zones and are well known to cause instabilities in the calculation of the associated dispersion curves. In fact, rejecting models with unrealistically high velocities at the surface is more effective than restricting the vertical gradient of the shear velocity. Models from accepted chains are also thinned to a maximum of 1 million models to define the posterior distribution (Figure 9e).

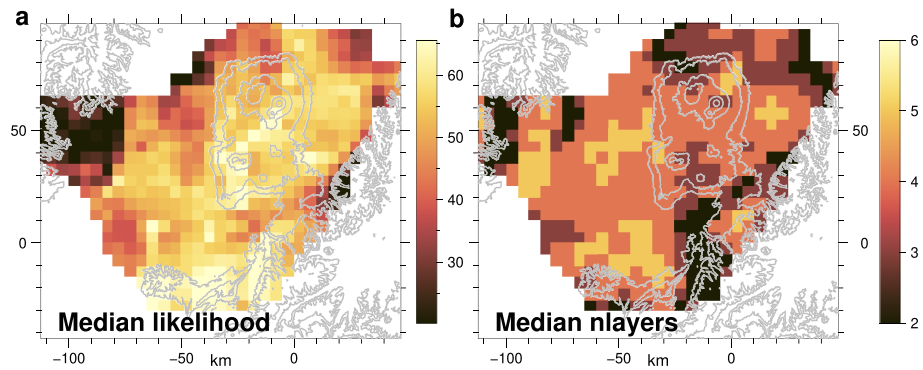


Figure 10. Condition of individual 1-D inversion results. (a) Median likelihood of the posterior distribution. (b) Median number of layers of the posterior distribution. Results displayed discretely for each inversion node.

We show the median likelihood of the set of final 1-D models in Figure 10a as a measure of the quality of data fit. The local dispersion curves are most coherent and smooth in the center of the tomographic domain. As such, the quality of fit is worst in the far western edge of the region (Figure 10a). Here the 2-D tomographic solutions have poor constraints on the very fast velocities as there are very few raypath intersections.

Individual 1-D depth inversions determine the shear velocity with reference to the free surface, and so in order to build a correct undistorted 3-D volume, the depth from surface must be corrected to a consistent depth relative to sea level. To do this, we smooth the digital elevation model using a Gaussian filter with a 50 km radius and then use this local elevation value to shift the results of each 1-D inversion. Köhler et al. (2012) found by empirical analysis that the shortest wavelength of topography, which a Rayleigh wave is sensitive to, is approximately 2.5 times the wavelength of a Rayleigh wave. Hence, the choice of the 50 km filter is a reasonable average for our period range of 4–14 s (~15–90 km wavelengths). If too short a wavelength is used for the filter, the short wavelength topography is projected into the 3-D velocity volume, and at shallow depths, the horizontal velocity slices simply resemble the surface topography.

5. Results—Three-Dimensional V_{sv} Velocity Structure

After the results from the individual 1-D shear velocity inversions have been adjusted for their elevation above sea level, we assemble all grid nodes into a 3-D shear velocity volume. In the following, depth is therefore always defined relative to sea level. A 1-D depth-dependent average is made as a reference of the entire volume, excluding grid nodes to the west of $x = -100$ km, where extreme high velocities correlate with poor fit (Figure 10), and large standard deviations at shallow depths (Figures S7–S10). The shear velocity deviation is then defined as the relative deviation from the depth-dependent reference curve in percent (Figure S11). The shear velocity deviation is presented as horizontal depth sections in Figure 11 and as vertical cross sections in Figure 12. A more complete depth range is presented in Figures S7–S10, and further vertical cross sections are shown in Figures S12 and S13. We note that as expected from a Bayesian Monte Carlo approach, the standard deviation of V_{sv} (Figures S7–S10) shows high values not just at extreme velocity values but at extreme velocity gradients on the edges of anomalies, where the uncertainty from the 2-D inversion maps was high.

6. Discussion and Interpretation

Our tomographic model provides good constraints on velocities of the upper 15 km, which allows us to resolve a number of key features beneath the KVG and the surrounding CKD.

6.1. Sedimentary Basins of the CKD

The most striking feature of the 3-D velocity model is the clearly resolved lateral structure of the CKD sedimentary basin. The slow velocities of less than 1.4 km/s at shallow depths (Figures 11a and 11b) follow the low elevation of the river basins to the west, north, and east of the high-elevation volcanic massif. This

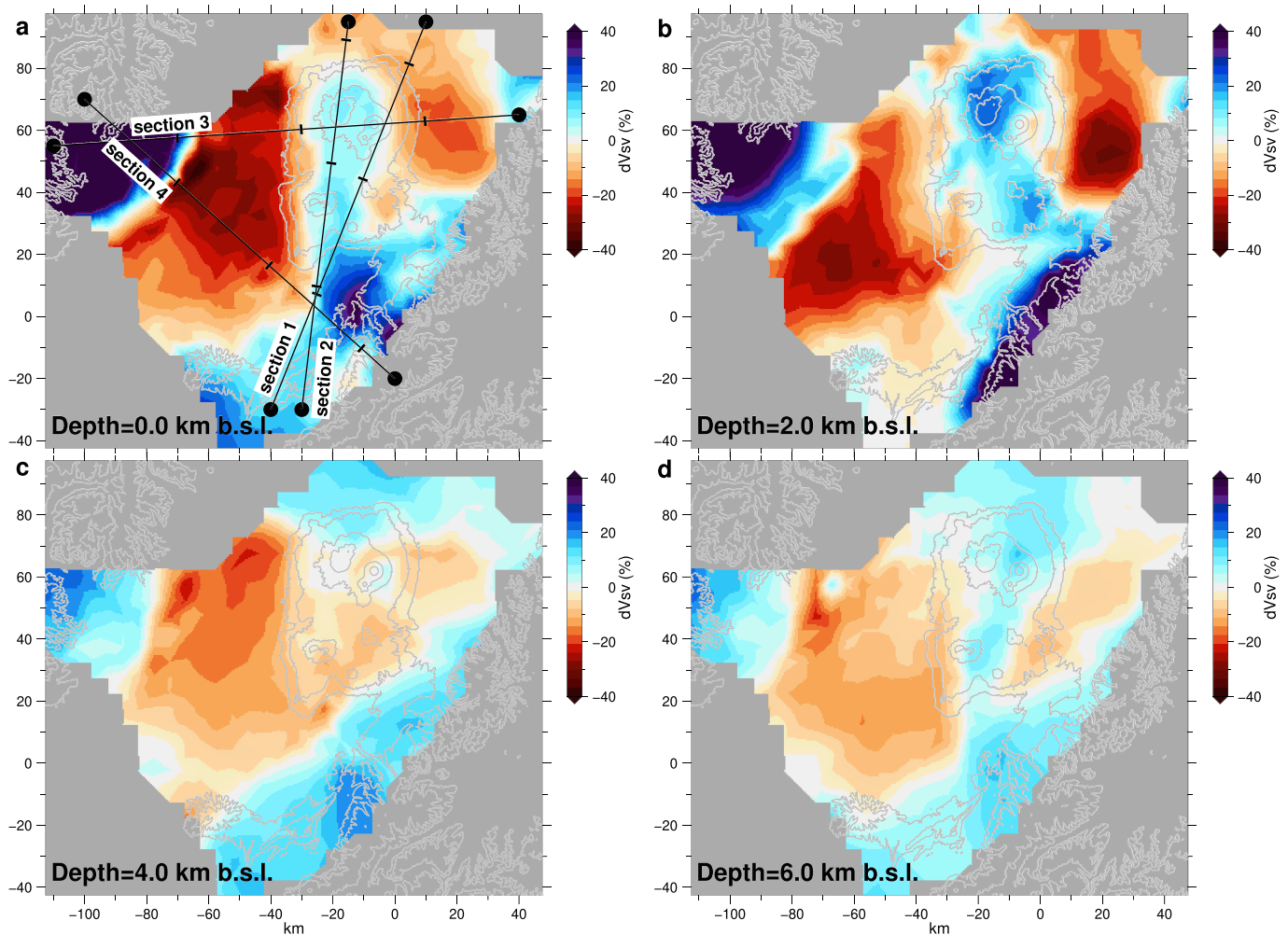


Figure 11. Horizontal slices through 3-D volume showing shear velocity (V_{sv}) deviation at (a) sea level, (b) 2 km, (c) 4 km and (d) 6 km below sea level. Profile lines on (a) indicate the location of cross sections in Figure 12. Ticks mark 40, 80, and 120 km along the profile. Gray lines are 1,000 m contours.

contrasts sharply with faster velocities where topography rises either as the KVG or the shoulders of the depression. This is not just an effect of our topographic depth corrections but is obvious also in maps of the shear velocity at the surface. The velocity contrasts primarily reflect the juxtaposition of sediments within the basin, against either the arc terrane basement rocks of the SR and ER or the basaltic-andesite of the KVG. The exact position of the western boundary of the basin (with the SR) is poorly defined in the tomography and is reflected in the uncertainties at this sharp velocity gradient (Figures S3 and S7).

The basins on both sides of the KVG massif show a shallow layer of very slow velocities (<1 km/s at their slowest) in the upper 3 km (Figures 11a, 11b, 12e, and 12g). In cross-sectional view, this shallow layer is symmetric across the volcanic massif (reds in Figures 12e and 12f). In section 4 (Figures 12g and 12h), we see that the structure through the western basin is very uniform. We interpret this upper 3 km, labeled as Layer S1 in interpretive Figure 13, to represent young unconsolidated sediments. A sharp velocity gradient is seen at the base of Layer S1 at ~ 3 km depth. Using the posterior ensemble of model results from the shear velocity inversion, we trace the peak of the interface depth distribution (as in Figure 9f) to map the depth of this shallow layer (S1) across the basin. This interface depth map (Figure 13a) shows a gentle thickening of Layer S1 toward the southwest and has a similar thickness to the north and east of KVG. Despite the relative uniformity of depth, the velocities are slowest in the basin on the west of the KVG, compared to the north and east of the KVG, which possibly reflects higher water saturation of the sediments here, where the Kamchatka River flows through the CKD.

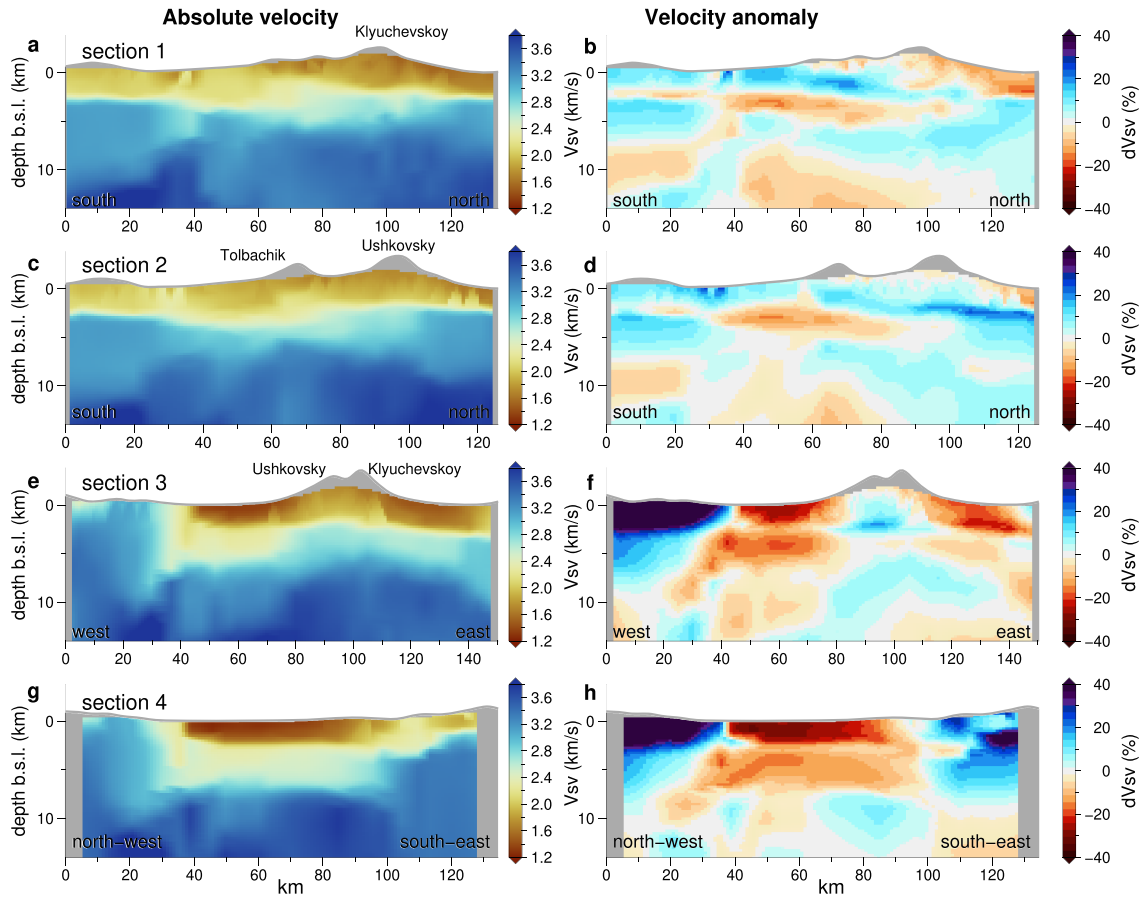


Figure 12. Cross sections through 3-D shear velocity (V_{sv}) volume. Figure 11a indicates locations of cross sections and distance markers. Left-hand column (a, c, e, and g) shows absolute V_{sv} anomalies, and right-hand column (b, d, f, h) shows V_{sv} anomalies. Surface topographic profile in gray, and the 50 km filtered topography is evident from boundaries of the velocity model.

Below the 3 km interface, slower than average velocities persist deeper (to 7–8 km) in the western basin (yellow of Figure 12e), but on the east the underlying velocities are closer to the average. We interpret the 2–2.4 km/s material below the 3 km interface in the western basin (labeled S2 on Figure 13) to still represent sediments, both because of their low absolute velocities and because a further sharp velocity gradient is seen below at 8 km depth where the shear velocity increases toward the average in the region. Thus, the interface at 8 km (base of S2) likely represents the top of the basement, and the interface at 3 km (S1-S2) is an intra-basin interface. The material between 3 and 8 km, Layer S2 in Figure 13, then represents more consolidated sediments, which have been deposited since subsidence began in the CKD. The interface at the base of Layer S2 can also be traced from the interface depth distributions and is shown in Figure 13b. In the basin to the north and east, the velocities below 3 km are not as slow (2.4–2.8 km/s) as in the western basin, and as such, there is not such a clear velocity jump at ~8 km depth. Layer S2 is therefore not well defined in these regions, which suggests a shallower transition to basement on the east compared to the west of the KVG. Our mapping of the basement provides the first spatially extensive information about the depth of the CKD, which reaches a maximum of 8 km below sea level.

Some localized data are available from active source seismic surveys, which were conducted in the CKD and across the KVG in the 1970s [Balesta et al. 1977], and offer some corroboration of our findings. Two shallow horizons in Balesta et al. (1977) can be correlated with our observations here. The sharp gradient in our model at 8 km (base of S2) correlates well with a V_p interface (V_p 5.7–6.2 km/s), which Balesta et al. (1977) interpret (like us) to be the top of the crystalline basement. As in our interpretation, the depth of the top of the basement is shallower to the east of the KVG compared to the western side. A shallower V_p

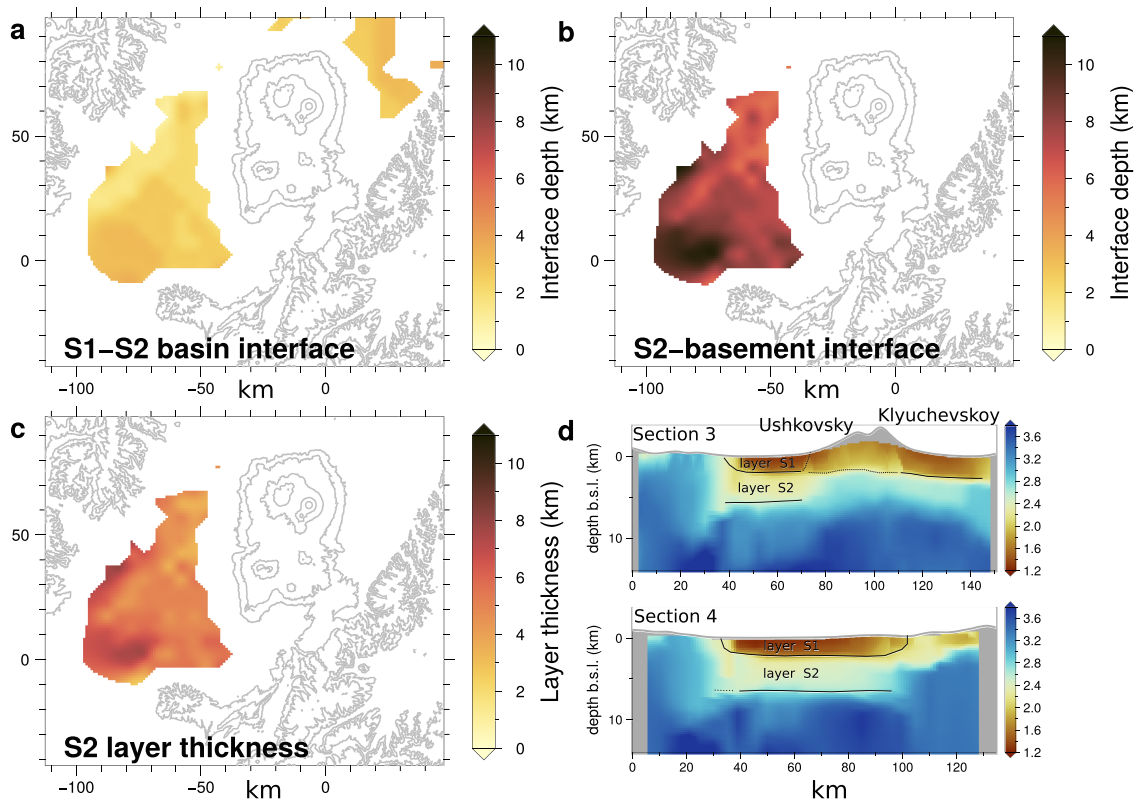


Figure 13. Layering in the Central Kamchatka depression. (a) Map of interface depth between layers S1 and S2 (d). (b) Map of interface depth at the base of layer S2. Interface depth is the peak of the interface depth distribution from all models of the posterior solution. (c) Thickness of layer S2.

interface (V_p 5.0–5.4) is similarly found at a depth of approximately 3 km within the CKD to the west of but is less continuously mapped in V_p as in our V_s results (S1-S2 interface).

Our mapping of the sediments and basement has implications for tectonic models of the genesis of the CKD. Current theories propose that the depression is either a back-arc basin (with extension behind the modern ER) (Park et al., 2002) or alternatively is a longer-lived fore-arc basin associated with the older SR arc (Avdeiko et al., 2007; Portnyagin et al., 2007).

The back-arc basin theory holds that extension initiates behind the ER, which is advancing northward following the migrating Miocene–Pliocene collisions of the Kronotsky arc terrane (Alexeiev et al., 2006; Lander & Shapiro, 2007; Pedoja et al., 2013). This would mean that the extension is rather recent in the region of KVG, as the last cape of the Kronotsky arc collision was accreted onto the eastern Kamchatka coastline around 2 Ma (Kozhurin & Zelenin, 2017). The mapped 8 km of accumulated sediments would therefore require an extremely high subsidence and sedimentation rate of approximately 4 m/Kyr. There are also additional problems with the geometry of this theory, as the northward migration of the Kronotsky arc collisions would predict the CKD to be wider toward the south, which is not observed. A further implication of back-arc spreading is that a crustal thinning signature would be expected within the depression. Assuming isostatic compensation, an isostatic balance can be constructed between the CKD and the approximately 35 km thick crust (Iwasaki et al., 2013; Levin et al., 2002; Nikulin et al., 2010) on the CKD shoulders. Using continental crust density of 2.7 g/cm³, mantle density of 3.4/cm³, sediment density of 2.2/cm³ (for the 5 km thick Layer S2), and unconsolidated sediment density of 1.6/cm³ (for the 3 km thick Layer S1), the thinned crustal thickness within the CKD would be expected to be approximately 18 km. Future receiver function and noise autocorrelation investigations using the KISS data set should provide the resolution to determine whether such a change in Moho depth from 35 to 26 km (18 km + 8 km of sediments) can be observed.

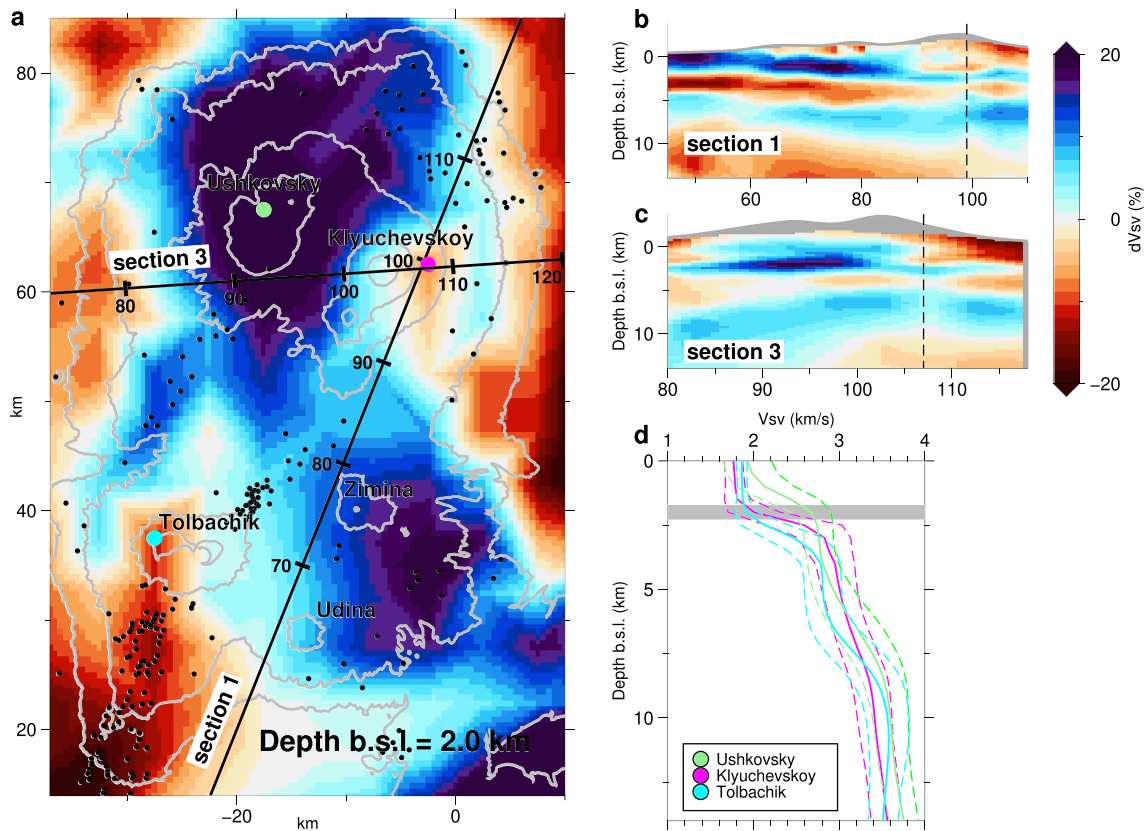


Figure 14. Zoom in on volcanic massif of KVG. (a) Horizontal section at 2 km depth with distance indicators for cross Sections 1 (b) and 3 (c). Black circles show monogenetic cones. Dotted vertical line indicates the intersection of the two profiles. Colored dots mark position of 1-D profiles in (d). (d) Comparison of 1-D shear velocity profiles under volcanoes. Mean V_s is solid line, with uncertainty range in dotted lines. Gray band marks the depth of horizontal section (a).

The alternative neotectonic theory is that the CKD is a long-lived fore-arc basin to the SR (Avdeiko et al., 2007; Portnyagin et al., 2007), formed during Eocene-Pliocene subduction prior to the development of the modern arc, the ER. This would then explain the thick sedimentary sequence with more modest accumulation rates (approximately 0.2 m/Kyr). We also speculate that the two sedimentary layers within the CKD could represent different phases of subsidence. In this scenario the deeper and thicker Layer S2 (Figure 13) formed within an Eocene-Pliocene fore-arc regime ahead of the SR volcanic arc. Then following Pliocene development of the ER, a faster back-arc extensional subsidence may have triggered an influx of sediments, forming the slower velocity and less consolidated sediments of Layer S1 (Figure 13).

6.2. Magmatic Structure beneath the KVG

The broad scale crustal root of the volcanic massif can be seen in Cross Sections 1 and 2 of Figures 12a and 12c, which transect the transition from crust on the shoulders of the CKD, through the KVG volcanic massif. Beneath the volcanic massif, we observe a lense of slow velocities relative to the average for the area. The influence of this slow velocity root beneath the KVG extends to about 6 km depth. This suggests that the accumulation of magma storage bodies, as sills or dykes, is concentrated shallower than 6 km depth. Below this, we observe reduced lateral velocity heterogeneity, and as such, there is an absence of evidence for magma storage bodies deeper than 6 km.

Within the volcanic massif of the KVG, the amplitude of anomalies is smaller than the contrasts of the sedimentary basins. However, notable structural features are present when discussed alongside an interrogation of their uncertainty. At depths of 2–3 km, slow velocities are present beneath the currently active volcanoes, while nonactive volcanoes of the KVG show fast velocity anomalies (Figure 14). In particular, a small slow velocity anomaly can be seen below Klyuchevskoy (intersection of profiles on Figure 14). While this feature is small, it is distinctly separate from the slow velocities of the river

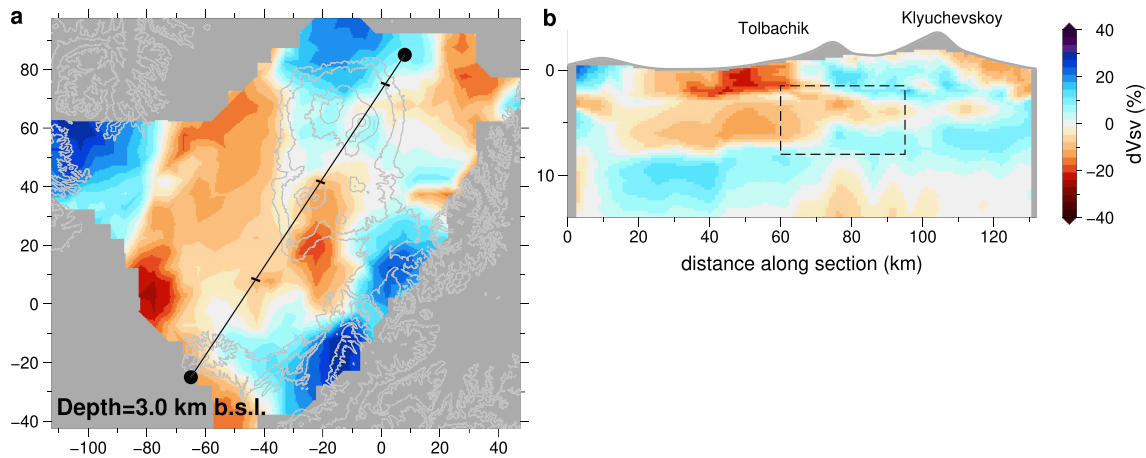


Figure 15. Horizontal section at 3 km depth (a) with distance indicators for cross section (b) which highlights with a dashed box the slow velocity anomaly which extends from sedimentary layer S2 (Figure 13) below the southern half of the KVG.

basin to the east. At this same depth in Figure 14, we see a N-S oriented slow velocity anomaly under Tolbachik and the system of fissures and monogenetic cones, which extends SSW (Tolbackinsky Dol). This location is the site of a major fissure eruptions, of which the two most recent and famous ones occurred in 1975–1976 and 2012–2013.

The correlation of the slow velocities with recent volcanic activity suggests a possible relation of these features with either magma mixing or magma storage regions. Meanwhile, in Figure 14, Zimina and Udina volcanoes in the southeast and Ushkovsky Volcano in the northwest part of the KVG show much faster velocities, representing possibly solidified cumulates from these less active volcanic centers. The fastest anomaly, Ushkovsky Volcano, is a predominantly basaltic volcanic center and has erupted only three times in the Holocene period. Crucially, the posterior ensemble of models from the Monte Carlo tomography can be used to demonstrate the significance of these velocity anomalies. In Figure 14d, the shear velocity model at Ushkovsky (fast) is compared with Klyuchevskoy and Tolbachik (slow), showing that the uncertainty bounds of the velocity between these points are distinctly separated in the shallow part of the model. Therefore, even though these anomalies are smaller in amplitude than the contrasts to the sedimentary basins, they are still significant and interpretable.

Between 3 and 5 km depth, we also observe a larger region of slow velocities beneath the southern half of the KVG massif, under both Tolbachik and Udina volcanoes (Figures 11c, S7, and S8). This anomaly could be interpreted as a significant region of melt accumulation under these volcanoes in the southern half of the KVG. However, no supporting evidence for this exists, either in other geophysical models or in the eruptive activity of the southern volcanoes in comparison to Klyuchevskoy volcano. As an alternative, we propose that these much slower velocities could be of sedimentary origin, indicating that the volcanoes of the southern half of the KVG have grown out through and on top of a significant sedimentary package. This hypothesis is supported by a cross section displayed in Figure 15, where the slow velocity anomaly in question is continuously linked with the slow velocities of Sedimentary Layer S2 in the CKD. Future radial anisotropy tomography studies might help to test this hypothesis and our magma storage interpretations (Figure 14), by distinguishing between anisotropic sill-like magmatic storage (Jaxybulatov et al., 2014; Jiang et al., 2018) and sedimentary material.

From 6 km and deeper, we see faster velocities under the volcanic massif (Figures 11d and S7–S10), suggesting that deeper than 6 km, the active magma storage regions are underlain by solidified cumulates. Distinct slow anomalies are restricted to upper crustal levels, between 2 and 5 km below sea level, indicating that magma storage regions for the eruptive activity of the KVG are shallow but not immediately below the volcanic edifices. This interpretation of magma storage at 2–5 km is supported by petrological evidence of Khubunaya et al. (2018) for a shallow storage region where aluminous basalts fractionate and evolve. The absence of magma storage bodies at intermediate depths (~8–15 km) in our model means that we cannot confirm the findings of Koulakov et al. (Kästle et al., 2016; Koulakov et al., 2013; Koulakov, Gordeev,

et al., 2011). However, the seismic expression of the intermediate depth storage bodies appears to show significant temporal variability (Koulakov et al., 2013). This may also be an effect of a loss of lateral resolution for the longer period surface waves sampling depths of ~10 km. A joint body wave and surface wave inversion is the strategy required to combine the complementary sensitivities of these two observations and should be the objective of future research.

We can summarize therefore that within seismic imaging to date, the only slow velocity magmatic features that are repeatedly reproduced in current imaging are a deep storage region at about 30 km depth and shallow storage regions at 2–5 km depth. Observations of long-period earthquakes (the generation mechanism of which is believed to be hydrothermal or volcanic in origin) show a similar pattern of shallow and deep separation. Long-period earthquakes occur in two depth ranges, above 5 km and around 30 km (Senyukov, 2013), and migration of event rates indicates a hydraulic connection between these two clusters (Shapiro, Sens-Schönfelder, et al., 2017). Interpreted together with petrological evidence for storage at 5 km and greater than 30 km depth, this seems to indicate that prolonged storage regions are absent within the midcrust.

7. Conclusions

We map two sedimentary layers in the basins surrounding the KVG, revealing a depth of the basement of 8 km below sea level in the CKD. This volume of sediments within the sedimentary basin supports a hypothesis that the CKD is a fore-arc basin, formed during Eocene-Pliocene subduction on the Proto-Kamchatka subduction zone and associated with volcanism of the SR arc. The intrabasin interface in the western part of the CKD indicates a change in the sedimentation process that might have been induced by the Miocene–Pliocene collisions of the Kronotsky arc.

The volcanic massif of the KVG is characterized by much lower amplitude anomalies in seismic velocity compared to the anomalies produced by the sediment-volcanic contrast. Shallow velocity heterogeneity is never-the-less significant as verified by the uncertainty analysis using the Monte Carlo ensemble result and reveals a correlation between slow velocities and the currently active volcanoes, whereas the less active volcanoes are aligned along fast velocities at shallow depth. Below 6 km depth, velocities are higher, and the structure is relatively homogenous. We therefore find evidence only for magma storage beneath the active volcanoes between depths of 2 and 5 km below sea level.

Acknowledgments

R. G. G. is supported by an independent Postdoctoral Fellowship from the Alexander von Humboldt Foundation (funded by the German Federal Ministry for Education and Research). This study was financially supported by the Russian Ministry of Education and Science (Grant N 14.W03.31.0033) and by the European Research Council (ERC) under the European Union Horizon 2020 Research and Innovation Programme (Grant Agreement 787399-SEISMAZE). Fast Marching code for forward modeling originally downloaded from Matlab file exchange contribution 24531. Seismological data are available from the GEOFON data center of GFZ-Potsdam (<https://geofon.gfz-potsdam.de/>). Derived products from this publication, including seismic velocity volumes, are available within the supporting information. The authors declare no financial conflicts of interest.

References

- Aki, K. (1957). Space and time spectra of stationary waves with special reference to microtremors. *Bulletin. Earthquake Research Institute, University of Tokyo*, 35, 415–456.
- Alexeev, D. V., Gaedicke, C., Tsukanov, N. V., & Freitag, R. (2006). Collision of the Kronotskiy arc at the NE Eurasia margin and structural evolution of the Kamchatka-Aleutian junction. *International Journal of Earth Sciences*, 95, 977–993.
- Avdeiko, G. P., Savelyev, D. P., Palueva, A. A., & Popruzhenko, S. V. (2007). Evolution of the Kurile-Kamchatkan volcanic arcs and dynamics of the Kamchatka-Aleutian junction. In J. Eichelberger, E. Gordeev, P. Izbekov, M. Kasahara, & J. Lees (Eds.), *Volcanism and subduction: The Kamchatka Region, Geophysical Monograph Series* (Vol. 172, pp. 37–55). USA: AGU. <https://doi.org/10.1029/172GM04>
- Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., et al. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, 169(3), 1239–1260. <https://doi.org/10.1111/j.1365-246X.2007.03374.x>
- Bodin, T., Sambridge, M., Rawlinson, N., & Arroucau, P. (2012). Transdimensional tomography with unknown data noise. *Geophysical Journal International*, 189(3), 1536–1556.
- Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., & Rawlinson, N. (2012). Transdimensional inversion of receiver functions and surface wave dispersion. *Journal of Geophysical Research*, 117, B02301. <https://doi.org/10.1029/2011JB008560>
- Braitseva, O. A., Melekestev, I. V., Ponomareva, V. V., & Sulzerzhitsky, L. D. (1995). Ages of calderas, large explosive craters and active volcanoes in the Kuril-Kamchatka region, Russia. *Bulletin of Volcanology*, 57, 383–402.
- Churikova, T. G., Gordeychik, B. N., Edwards, B. R., Ponomareva, V. V., & Zelenin, E. A. (2015). The Tolbachik volcanic massif: A review of the petrology, volcanology and eruption history prior to the 2012–2013 eruption. *Journal of Volcanology and Geothermal Research*, 307, 3–21.
- Dobretsov, N. L., Koulakov, I. Y., & Litasov, Y. D. (2012). Migration paths of magma and fluids and lava compositions in Kamchatka. *Russian Geology and Geophysics*, 53, 1253–1275.
- Dorendorf, F., Wiechert, U., & Wörner, G. (2000). Hydrated sub-arc mantle: A source for the Kluchevskoy volcano, Kamchatka/Russia. *Earth and Planetary Science Letters*, 175, 69–86.
- Dreiling, J., & Tilmann, F. (2019). BayHunter–McMC transdimensional Bayesian inversion of receiver functions and surface wave dispersion. DOI: <http://doi.org/10.5880/GFZ.2.4.2019.001>
- Droznin, D. V., Shapiro, N. M., Droznina, S. Y., Senyukov, S. L., Chebrov, V. N., & Gordeev, E. I. (2015). Detecting and locating volcanic tremors on the Klyuchevskoy group of volcanoes (Kamchatka) based on correlations of continuous seismic records. *Geophysical Journal International*, 203(2), 1001–1010.

- Dziewonski, A., Bloch, S., & Landisman, M. (1969). A technique for the analysis of transient seismic signals. *Bulletin of the Seismological Society of America*, 59(1), 427–444.
- Ekström, G., Abers, G. A., & Webb, S. C. (2009). Determination of surface wave phase velocities across USArray from noise and Aki's spectral formulation. *Geophysical Research Letters*, 36, L18301. <https://doi.org/10.1029/2009GL039131>
- Fedotov, S. A., & Zharinov, N. A. (2007). On the eruptions, deformation, and seismicity of Klyuchevskoy Volcano, Kamchatka in 1986–2005 and the mechanisms of its activity. *Journal of Volcanology and Seismology*, 1, 71–97.
- Fedotov, S. A., Zharinov, N. A., & Gontovaya, L. I. (2010). The magmatic system of the Klyuchevskaya group of volcanoes inferred from data on its eruptions, earthquakes, deformation, and deep structure. *Journal of Volcanology and Seismology*, 4, 1–33.
- Gallagher, K., Charvin, K., Nielsen, S., Sambridge, M., & Stephenson, J. (2009). Markov chain Monte Carlo (MCMC) sampling methods to determine optimal models, model resolution and model choice for earth science problems. *Marine and Petroleum Geology*, 26(4), 525–535.
- Gorbatov, A., Kostoglodov, V., Suárez, G., & Gordeev, E. (1997). Seismicity and structure of the Kamchatka Subduction zone. *Journal of Geophysical Research*, 102, 17,883–17,898. <https://doi.org/10.1029/96JB03491>
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4), 711–732.
- Herrmann, R. B., & Ammon, C. J. (2004). *Computer Programs in Seismology: Surface Waves, Receiver Functions and Crustal Structure*. Saint Lois, USA: Saint Lois University.
- Hourigan, J. K., Brandon, M. T., Soloviev, A. V., Kirmasov, A. B., Garver, J. I., Stevenson, J., & Reiners, P. W. (2009). Eocene arc-continent collision and crustal consolidation in Kamchatka. *Russian Far East. Am J Sci.*, 309, 333–396.
- Ivanov, A. I., Koulakov, I. Y., West, M., Jakovlev, A. V., Gordeev, E. I., Senyukov, S., & Chebrov, V. N. (2016). Magma source beneath the Bezmianny volcano and its interconnection with Klyuchevskoy inferred from local earthquake seismic tomography. *Journal of Volcanology and Geothermal Research*, 323, 62–71.
- Iwasaki, T., Levin, V., Nikulin, A., & Iidaka, T. (2013). Constraints on the Moho in Japan and Kamchatka. *Tectonophysics*, 609, 184–201.
- Jaxybulatov, K., Shapiro, N. M., Koulakov, I., Mordret, A., Landès, M., & Sens-Schönfelder, C. (2014). A large magmatic sill complex beneath the Toba caldera. *Science*, 346(6209), 617–619. <https://doi.org/10.1126/science.1256785>
- Jiang, C., Schmandt, B., Farrell, J., Lin, F. C., & Ward, K. M. (2018). Seismically anisotropic magma reservoirs underlying silicic calderas. *Geology*, 46(8), 727–730. <https://doi.org/10.1130/G45104.1>
- Kästle, E. D., Soomro, R., Weemstra, C., Boschi, L., & Meier, T. (2016). Two-receiver measurements of phase velocity: Cross-validation of ambient-noise and earthquake-based observations. *Geophysical Journal International*, 207(3), 1493–1512. <https://doi.org/10.1093/gji/ggw341>
- Khubunaya, S. A., Gontovaya, L. I., Sobolev, A. V., & Khubunaya, V. S. (2018). On the magma chambers beneath Klyuchevskoi volcano, Kamchatka. *Journal of Volcanology and Seismology*, 12(2), 85–99.
- Khubunaya, S. A., Gontovaya, L. I., Sobolev, A. V., & Nizkous, I. V. (2007). Magma chambers beneath the Klyuchevskoy volcanic group (Kamchatka). *J Volcanol Seismol.*, 1, 98–118.
- Köhler, A., Weidle, C., & Maupin, V. (2012). On the effect of topography on surface wave propagation in the ambient noise frequency range. *J Seismol*, 16(2), 221–231. <https://doi.org/10.1007/s10950-011-9264-5>
- Konstantinovskaia, E. (2001). Arc-continent collision and subduction reversal in the Cenozoic evolution of the Northwest Pacific: An example from Kamchatka (NE Russia). *Tectonophysics*, 333, 75–94.
- Koulakov, I., Abkadyrov, I., Arifi, N. A., Deev, E., Droznina, S., Gordeev, E. I., et al. (2017). Three Different Types of Plumbing System beneath the neighboring active volcanoes of Tolbachik, Bezmianny, and Klyuchevskoy in Kamchatka. *J. Geophys. Res. Solid Earth*, 122, 3852–3874. <https://doi.org/10.1002/2017JB014082>
- Koulakov, I., Gordeev, E. I., Dobretsov, N. L., Vernikovskiy, V. A., Senyukov, S., & Jakovlev, A. (2011). Feeding volcanoes of the Klyuchevskoy group from the results of local earthquake tomography. *Geophysical Research Letters*, 38, L09305. <https://doi.org/10.1029/2011GL046957>
- Koulakov, I., Gordeev, E. I., Dobretsov, N. L., Vernikovskiy, V. A., Senyukov, S., Jakovlev, A., & Jaxybulatov, K. (2013). Rapid changes in magma storage beneath the Klyuchevskoy group of volcanoes inferred from time-dependent seismic tomography. *Journal of Volcanology and Geothermal Research*, 263, 75–91. <https://doi.org/10.1016/j.jvolgeores.2012.10.014>
- Koulakov, I. Y., Dobretsov, N. L., Bushenkova, N. A., & Yakovlev, A. V. (2011). Slab shape in subduction zones beneath the Kurile-Kamchatka and Aleutian arcs based on regional tomography results. *Russian Geology and Geophysics*, 52, 650–667.
- Kozhurin, A., & Zelenin, E. (2017). An extending island arc: The case of Kamchatka. *Tectonophysics*, 706. <https://doi.org/10.1016/j.tecto.2017.04.001>
- Lander, A. V., & Shapiro, M. N. (2007). The origin of the modern Kamchatka Subduction zone. *Geophys Monogr Ser.*, 172, 57–64.
- Landès, M., Hubans, F., Shapiro, N. M., Paul, A., & Campillo, M. (2010). Origin of deep ocean microseisms by using teleseismic body waves. *Journal of Geophysical Research: Solid Earth*, 115, B5302. <https://doi.org/10.1029/2009JB006918>
- Lees, J. M., Symons, N., Chubarova, O., Gorelchik, V., & Ozerov, A. (2007). Tomographic images of Klyuchevskoy volcano P-wave velocity. In J. Eichelberger, E. Gordeev, P. Izbekov, M. Kasahara, & J. Lees (Eds.), *Volcanism and subduction: The Kamchatka Region, Geophysical Monograph Series* (Vol. 172, pp. 293–302). <https://doi.org/10.1029/172GM21>
- Lees, J. M., Vandecar, J., Gordeev, E., Ozerov, A., Brandon, M., Park, J., & Levin, V. (2007). Three dimensional images of the Kamchatka-Pacific plate cusp. In J. Eichelberger, E. Gordeev, P. Izbekov, M. Kasahara, & J. Lees (Eds.), *Volcanism and subduction: The Kamchatka Region, Geophysical Monograph Series* (Vol. 172, pp. 65–75). USA: AGU. <https://doi.org/10.1029/172GM06>
- Levin, V., Park, J., Brandon, M., Lees, J., Peyton, V., Gordeev, E., Ozerov, A. (2002). Crust and upper mantle of Kamchatka from teleseismic receiver functions. *Tectonophysics*, 358, 233–265.
- Levin, V., Shapiro, N., Park, J., & Ritzwoller, M. (2002). Seismic evidence for catastrophic slab loss beneath Kamchatka. *Nature*, 418(6899), 763–767. <https://doi.org/10.1038/nature00973>
- Levshin, A. L., Yanovskaya, T. B., Lander, A. V., Bukchin, B. G., Barmin, M. P., Ratnikova, L. I., & Its, E. N. (1989). Recording, identification, and measurement of surface wave parameters. In *Seismic Surface Waves in a Laterally Inhomogeneous Earth* (pp. 131–182). Pub: Kluwer Academic.
- Malinverno, A. (2002). Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophysical Journal International*, 151(3), 675–688.
- Moroz, Y. F. (1991). *Elektroprovodnost' Zemnoi Kory I Verkhnei Mantii Kamchatki (the Electric Conductivity of the Crust and Upper Mantle in Kamchatka)*. Moscow: Nauka.

- Nikulin, A., Levin, V., Shuler, A., & West, M. (2010). Anomalous seismic structure beneath the Klyuchevskoy group, Kamchatka. *Solid Earth*, 37, L14311. <https://doi.org/10.1029/2010GL043904>
- Park, J., Levin, V., Brandon, M., Lees, J., Peyton, V., Gordeev, E., & Ozerov, A. (2002). A dangling slab, amplified arc volcanism, mantle flow and seismic anisotropy in the Kamchatka plate corner. *AGU Geodynamics Series*, 30, 295–324.
- Pedoja, K., Authemayou, C., Pinegina, T., Bourgeois, J., Nexer, M., Delcaillau, B., & Regard, V. (2013). “Arc-continent collision” of the Aleutian-Komandorsky arc into Kamchatka: Insight into quaternary tectonic segmentation through Pleistocene marine terraces and morphometric analysis of fluvial drainage. *Tectonics*, 32, 827–842. <https://doi.org/10.1002/tect.20051>
- Portnyagin, M., Bindeman, I., Hoernle, K., & Hauff, F. (2007). Geochemistry of primitive lavas of the Central Kamchatka depression: Magma generation at the edge of the Pacific plate. In J. Eichelberger, E. Gordeev, P. Izbekov, M. Kasahara, & J. Lees (Eds.), *Volcanism and subduction: The Kamchatka Region, Geophysical Monograph Series* (Vol. 172, pp. 199–239). USA: AGU. <https://doi.org/10.1029/172GM16>
- Rawlinson, N., & Sambridge, M. (2004). Wave front evolution in strongly heterogeneous layered media using the fast marching method. *Geophysical Journal International*, 156(3), 631–647.
- Retailleau, L., Landès, M., Gualtieri, L., Shapiro, N. M., Campillo, M., Roux, P., & Guilbert, J. (2018). Detection and analysis of a transient energy burst with beamforming of multiple teleseismic phases. *Geophysical Journal International*, 212(1), 14–24. <https://doi.org/10.1093/gji/ggx410>
- Ryberg, T., Geissler, W. H., Jokat, W., & Pandey, S. (2017). Uppermost mantle and crustal structure at Tristan da Cunha derived from ambient seismic noise. *Earth and Planetary Science Letters*, 471, 117–124. <https://doi.org/10.1016/j.epsl.2017.04.049>
- Sambridge, M. (2014). A parallel tempering algorithm for probabilistic sampling and multimodal optimization. *Geophysical Journal International*, 196(1), 357–374. <https://doi.org/10.1093/gji/ggt342>
- Schimmel, M., Stutzmann, E., & Gallart, J. (2011). Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale. *Geophysical Journal International*, 184, 494–506.
- Sens-Schönfelder, C., & Eulenfeld, T. (2019). Probing the in situ elastic nonlinearity of rocks with earth tides and seismic noise. *Physical Review Letters*, 122(13), 138,501. <https://doi.org/10.1103/PhysRevLett.122.138501>
- Senyukov, S. L. (2013). *Forecasting of the Eruptions of Volcanoes Klyuchevskoy and Bezymianny at Kamchatka [in Russian]*. Russian Federation: Lambert Academic.
- Senyukov, S. L., Droznina, S. Y., Nuzhdina, I. N., Garbuzova, V. T., & Kozhevnikova, T. Y. (2009). Studies in the activity of Klyuchevskoi volcano by remote sensing techniques between January 1, 2001 and July 31, 2005. *Journal of Volcanology and Seismology*, 3, 191–199.
- Shapiro, M. N., & Soloviev, A. V. (2009). Formation of the Olyutorsky-Kamchatka foldbelt: A kinematic model. *Russian Geology and Geophysics*, 50, 668–681. <https://doi.org/10.1016/j.rgg.2008.10.006>
- Shapiro, N. M., & Campillo, M. (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, 31, L07614. <https://doi.org/10.1029/2004GL019491>
- Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H. (2005). High-resolution surface wave tomography from ambient seismic noise. *Science*, 307(5715), 1615–1618. <https://doi.org/10.1126/science.1108339>
- Shapiro, N. M., Droznin, D. V., Droznina, S. Y., Senyukov, S. L., Gusev, A. A., & Gordeev, E. I. (2017). Deep and shallow long-period volcanic seismicity linked by fluid-pressure transfer. *Nature Geoscience*, 10(6), 442–445.
- Shapiro, N. M., Sens-Schönfelder, C., Lühr, B. G., Weber, M., Abkadyrov, I., Gordeev, E. I., et al. (2015). Klyuchevskoy volcanic group experiment (KISS). GFZ data services. Seismic network. <https://doi.org/10.14470/K47560642124>
- Shapiro, N. M., Sens-Schönfelder, C., Lühr, B. G., Weber, M., Abkadyrov, I., Gordeev, E. I., et al. (2017). Understanding Kamchatka's extraordinary volcano cluster. *Eos*, 98. <https://doi.org/10.1029/2017EO071351>
- Soubestre, J., Seydoux, L., Shapiro, N. M., de Rosny, J., Droznin, D. V., Droznina, S. Y., et al. (2019). Depth migration of seismovolcanic tremor sources below the Klyuchevskoy volcanic group (Kamchatka) determined from a network-based analysis. *Geophysical Research Letters*, 46, 8018–8030. <https://doi.org/10.1029/2019GL083465>
- Soubestre, J., Shapiro, N. M., Seydoux, L., de Rosny, J., Droznin, D. V., Droznina, S. Y., et al. (2018). Network-based detection and classification of seismovolcanic tremors: Example from the Klyuchevskoy volcanic group in Kamchatka. *Journal of Geophysical Research: Solid Earth*, 123, 564–582. <https://doi.org/10.1002/2017JB014726>
- Tilmann, F. J., Sadeghisorkhani H., & Mauerberger, A. (2019). Another look at the treatment of data uncertainty in Markov chain Monte Carlo inversion and other probabilistic methods. *EarthArXiv*. <https://doi.org/10.31223/osf.io/bcs4j>
- Ventosa, S., Schimmel, M., & Stutzmann, E. (2017). Extracting surface waves, hum and normal modes: Time-scale phase-weighted stack and beyond. *Geophysical Journal International*, 211, 30–44.
- Yogodzinski, G. M., Lees, J. M., Churikova, T. G., Dorendorf, F., Wöerner, G., & Volynets, O. N. (2001). Geochemical evidence for the melting of subducting oceanic lithosphere at plate edges. *Nature*, 409(6819), 500–504. <https://doi.org/10.1038/35054039>
- Zhang, X., Curtis, A., Galetti, E., & de Ridder, S. (2018). 3-D Monte Carlo surface wave tomography. *Geophysical Journal International*, 215(3), 1644–1658. <https://doi.org/10.1093/gji/ggy362>