Reservoir Imaging Using Ambient Noise Correlation From a Dense Seismic Network

M. Lehujeur1, J. Vergne1, J. Schmittbuhl1, D. Zigone1, A. Le Chenadec1, and EstOF Team1,2,3

1Université de Strasbourg, CNRS, IPGS UMR7516, et EOST UMBS830, Strasbourg, France, 2ES-Géothermie, Haguenau, France, 3GEIE-EMC, Kutzenhausen, France

Abstract In September 2014, a dense temporary seismic network (EstOF) including 288 single-component geophones was deployed during 1 month in the Outre-Forêt region of the Upper Rhine Graben (France), where two deep geothermal projects (Soulz-sous-Forêts and Rittershoffen) are currently in operation. We apply ambient seismic noise correlation to estimate the empirical Green’s function of the medium between the ~41,200 station pairs in the network. The noise correlation functions obtained are comparable to those from previous studies based on the sparse long-term networks settled in the area mostly to monitor the induced seismic activity. However, the dense spatial coverage of the EstOF network improves our ability to identify the main phases of the Green’s function. Both the fundamental mode and the first overtone of the Rayleigh waves are identified between most station pairs. P waves are also evidenced. We analyze the statistical distribution of the Rayleigh wave group velocity between station pairs as a function of the period (between 0.8 and 5 s), the station pair orientation, the distance over wavelength ratio and the signal-to-noise ratio. From these observations, we build a high-resolution three-dimensional 5 wave velocity model of the upper crust (down to 3 km deep) around the regional deep geothermal reservoirs. This model is consistent with local geological structures but also evidences nonlithological variations, particularly at depth in the basement. These variations are interpreted as large-scale temperature anomalies related to deep hydrothermal circulation.

1. Introduction

The Outre-Forêt region (northern Alsace, France) has long been known for its important geological resources. It has been extensively studied over the last century for the exploitation of oil (Haas & Hoffmann, 1929; Schnaebele, 1948) and more recently through the development of deep geothermal energy (Baillieux et al., 2013; Munck et al., 1980; Pribnow & Schellschmidt, 2000), in which the pilot geothermal project of Soulz-sous-Forêts has played a major role in the development of enhanced geothermal systems (EGS; Bressee, 1992; Genter et al., 2003, 2010; Gérard & Kappelmeyer, 1987; Huenges & Ledru, 2011; Olasolo et al., 2016). In this context, a large collection of data has contributed to the assessment of the oil and heat reservoirs in the region. The extended knowledge includes the surface geology, borehole lithology, thermal field, fault network, and numerous geophysical and geochemical surveys (e.g., magnetotellurics, gravimetry, fluid monitoring, and microseismicity), making the region attractive for testing new reservoir imaging methods.

Part of the data has been concatenated to produce a first 3-D geological model at the scale of the Upper Rhine Graben (GeORG Team, 2013). This knowledge has been completed with the recent development of the deep geothermal power plant near Rittershoffen, where several geophysical surveys have been conducted to plan the drilling of two boreholes reaching the granitic basement (Baujard et al., 2015, 2017). The microseismic activity associated with the development phase of the reservoir has also provided useful information about the orientation and the mechanical behavior of the main faults (Lenglinié et al., 2017; Maurer et al., 2015). The development of deep geothermal energy strongly relies on an appropriate knowledge of the geological structures at depth. This is particularly true for EGS projects that are targeting areas where positive temperature anomalies are related to the hydrothermal circulation of natural hot brine in fractures and fault zones located deep at the transition between the sedimentary cover and the crystalline basement (Guillou-Frottier et al., 2013; Magnenet et al., 2014).

Classical reservoir imaging, in particular, for oil and gas exploration, is based on active seismic surveys that rely on artificial impulsive sources (explosions and vibrator trucks). However, such operations have a high...
cost compared to the overall budget of a deep geothermal project and might affect the social acceptance of the project, especially in urban areas. An emerging alternative is ambient seismic noise imaging, which is a passive and low-cost approach. The feasibility of using ambient seismic noise as a source for the exploration of a geothermal field was identified several decades ago (Liaw & McEvilly, 1977). More recently, the cross correlation of ambient seismic noise records to obtain empirical Green’s functions between pairs of receivers (Lobkis & Weaver, 2001; Shapiro & Campillo, 2004) has become a standard technique for passive imaging (e.g., Lin et al., 2009; Shapiro et al., 2005; Stehly et al., 2009; Zigone et al., 2015) or monitoring (Brenguier et al., 2008; Sens-Schönfelder & Wegler, 2006) at various scales. Such techniques have been applied to geothermal fields both for exploration (Tibuleac et al., 2009; Tibuleac & Eneva, 2011) and monitoring purposes (Hillers et al., 2015; Obermann et al., 2015), relying on permanent or semipermanent seismic networks usually installed to monitor the induced seismic activity. At Soultz-sous-Forêts, the potential of such techniques was first demonstrated by Calò et al. (2013) based on a 22-station network. A comprehensive study of the properties of the ambient seismic noise recorded in the area has revealed a strong directivity of the ambient seismic noise at periods between 1 and 5 s that mainly originates from the Northern Atlantic Ocean and the Mediterranean Sea (Lehujeur et al., 2015). Such directivity patterns in the ambient seismic noise may affect the tomographic models obtained with sparse seismological networks (Lehujeur et al., 2016).

Here we focus on the EstOF temporary seismic network. We investigate the potential of dense networks operating over short time periods (i.e., 1 month) to image the deep geothermal reservoir at a regional scale (i.e., several tens of kilometers). Such dense networks with hundreds to thousands of sensors are increasingly used to build high-resolution velocity models from ambient seismic noise in various environments, such as the sea bottom (Mordret et al., 2013), urban areas (Lin et al., 2013; Nakata et al., 2015), fault zones (Ben-Zion et al., 2015; Roux et al., 2016), or volcanic edifices (Brenguier et al., 2016; Nakata et al., 2016; Wang et al., 2017). The recent development of “node”-like seismometers including all necessary components (digitizer, sensor, battery, etc.) into one single portable and wireless box is greatly contributing to the emergence of such large N-arrays for passive imaging purposes (Hand, 2014). In the following, we present the acquisition and the main characteristics of the ambient seismic noise records we obtained from the EstOF dense network in northern Alsace, France. We report on the main seismic phases that are recovered from the noise correlation processing. We determine the surface wave dispersion curves from the noise correlation functions, and we invert these curves to build a 3-D shear wave velocity model of the studied area. The model is finally discussed in the light of existing geophysical data in the region.

2. Data and Methods

2.1. Data

The EstOF network was deployed in 2014 from 25 August to 30 September (37 days). The main part of the network was composed of 259 stations settled on a 19 by 19 grid with an average interstation distance of 1.4 km. It covered a surface area of approximately 490 km² including the two deep geothermal sites Soultz-sous-Forêts and Rittershoffen, in the Outre-Forêt region, France (Figure 1), between the city of Haguenau, 30 km to the north of Strasbourg, and the French-German border. A subnet of 29 stations was deployed on a denser grid around the geothermal site Rittershoffen with an interstation distance of 500 m (Figure 1). From 8 to 18 September, the stations from the subnet were removed and used to swap the stations of the main network to recharge the batteries and download the data. The stations were equipped with Fairfield Nodal Zland nodes, which included a 10-Hz 1C vertical geophone, a 24-bit digitizer with a 2-Gb recording capacity, a GPS antenna and a lithium-ion battery with an autonomy of approximately 20 days. The sampling frequency was set to 250 samples per second. The data recovery over the whole period was

Figure 1. Map of the EstOF network (black dots). The two geothermal sites Soultz-sous-Forêts and Rittershoffen are shown as red stars. Yellow squares indicate the long-term monitoring network used by Lehujeur et al. (2016). Nodes 220, 172, and 133 (blue circles), equipped with © Fairfield Nodal Zland sensors, were installed close to the long-term stations SCHL (equipped with a Trillium compact 120-s velocimeter), LAMP and GUNS (equipped with L4C-1 Hz velocimeters). Sltz = Soultz-sous-Forêts, Rtt = Rittershoffen, Htt = Hatten, B = Betschdorf, Ha = Haguenau, See = Seebach, Sff = Soufflenheim. The solid white line corresponds to the French-German border.
approximately 97%. Few perturbing events occurred during the experiment since no stimulation was conducted at the geothermal power plants and only a few local natural earthquakes were identified during the recording period (12 events with local magnitudes between 0.9 and 2.4 in a radius of 100 km around the studied area, from the RéNaSS catalog, RéNASS, 2017). To minimize the impact for landowners, the nodes were deployed along pathways in agreement with local communities. The nodes were buried at 30 cm and coupled to the ground with 15-cm metallic spikes. The whole network was deployed in 2.5 days by seven teams of two or three people.

2.2. Methods

2.2.1. Noise Processing and Correlation

For each station, we first split the continuous noise records into 1-hr-long windows without overlaps. We compute the power spectral density (PSD, McNamara & Buland, 2004) of the raw waveforms and correct the amplitude of the PSD using the modulus of the theoretical instrumental response in acceleration. We use the PSD as a data quality estimator to reject the windows with anomalously low noise levels (below $\frac{1}{2}$ of the data).

The average shape of the probability density function of the PSD (Figure 2a) is characterized by high amplitudes for periods below 1 s that correspond to anthropogenic sources with marked daily and weekly periodicities (Figures 2b and 2c). The secondary microseismic peak that dominates the ambient seismic noise spectrum everywhere on Earth between periods 2 and 10 s (Peterson, 1993; Figure 2a, black dashed lines) generally exceeds the instrumental noise level of the nodes at periods above 1 s, despite their low cutoff period (0.1 s) and the related increase of the instrumental noise level at long periods (Figure 2a, red curve). A comparison of the noise spectrograms between one of the EstOF nodes (node 220) and station SCHL equipped with the broadband velocimeter and installed a few meters away is shown in Figures 2b and 2c. First, the figure indicates that individual EstOF nodes capture anthropogenic noise at short periods with good quality (e.g., daily variations below 1 s). Second, the figure confirms that the temporal variations of the secondary microseismic peak are detected by the EstOF sensors at periods up to 5 s and occasionally 7 s for most energetic events (see the black arrows at day 270 in Figures 2b and 2c) even if the instrumental noise levels at long periods blur the records.

Prior to cross correlation, we process noise windows as proposed by Bensen et al. (2007). Each noise window is detrended, tapered in the time domain using a cosine taper and downsampled at 25 samples per second. We uniformize the spectrum modulus between 0.1 and 10 s (spectral whitening) and clip the amplitudes in the time domain within the range ±3 times the standard deviation of the trace. No deconvolution of the instrumental response is applied to the waveforms because all the sensors of the network are the same. In such a case, the deconvolution has no effect on the noise correlation functions (NCFs) because (1) the

![Figure 2](image-url)
phase of the NCF depends only on the phase difference between the sensors, which remains the same whether or not the deconvolution is performed, and (2) the effect of the deconvolution on the noise spectrum is canceled by the spectral whitening filter. The noise correlation functions are computed for all 41,284 station pairs, averaged over time and derived in time (Roux, Sabra, Gerstoft, et al., 2005; Snieder, 2004). We arbitrarily orient the station pairs so that the signal observed on the positive side (causal side) of the NCF corresponds to sources located on the west side of the network.

2.2.2. Noise Correlation Functions, Comparison to Previous Data Set, and Dominant Wave Types

We compare the EstOF NCFs with the results of a previous study by Lehujeur et al. (2016), who used the ambient noise cross-correlation technique in the same region on a seismic network of 34 stations devoted mainly to the monitoring of seismicity (Figure 1, yellow squares). This sparse network is composed of both permanent and long-term temporary stations. The NCFs from this network were computed with several months to several years of noise. As the instruments were not similar for all the stations, the preprocessing of the noise waveforms included a deconvolution of the instrumental response (see Lehujeur et al., 2016). Some nodes of the EstOF network are located next to long-term stations (e.g., SCHL, LAMP, GUNS, Figure 1), which allows us to compare the noise correlation waveforms between the two data sets (Figure 3).

For periods between 1 and 5 s, the NCFs of the EstOF network (computed with approximately 30 days of data, Figure 3, red) have some similarities with the NCFs of the long-term network (several months to several years of data, Figure 3, black). The phase and amplitude are consistent in both networks. The asymmetry between causal and acausal sides is also consistent in this period range. Higher amplitudes are observed on the causal side due to prominent microseismic sources located in the northern Atlantic Ocean for back azimuths approximately 290–310°N (Lehujeur et al., 2015, 2016). For waveforms of both networks, we compare the ratio between the peak amplitude of direct arrivals (green time window, Figure 3) and the root-mean-square of the NCF in the coda part (blue time window, Figure 3). This ratio (Figure 3, black and red numbers, noted Rc for causal and Ra for acausal sides of the NCF) is used as an indicator of the NCF convergence (Bensen et al., 2007). The ratio is generally lower in the EstOF case, due to the shorter recording duration. However, the difference remains small compared to the difference between the recording durations used to compute the NCFs, suggesting that 30 days of noise is sufficient to reach an acceptable signal-to-noise ratio and to perform tomography. The evolution of this ratio as a function of the cumulated time (Figure 4) confirms that the NCF waveforms stabilize after 5 to 20 days, depending on the station pair and the side of the NCF. For some station pairs, we observe a decrease in the signal-to-noise ratio on the causal part of the NCF between 5 and 10 cumulative days (pairs 122–172, 122–133, Figure 4, thick curves), which coincides with a decrease in the amplitude of the secondary microseismic peak between days 245 and 250 (Figure 2b).

To further compare the wavefield reconstructed with the two data sets, we apply a band-pass filter between periods 0.28 and 5 s (0.2 to 3.5 Hz) to the NCFs and stack them in distance bins of 150 m (Figure 5). The resulting wavefield is converted from the time-distance to the frequency-wave number domain (F-K) using a 2-D Fourier transform (Figures 5b and 5d). The wavefields obtained with the two data sets show similar arrivals on both sides of the NCFs. The higher spatial resolution of the EstOF

![Figure 3. Comparison of noise correlation functions (NCFs) between three pairs of nodes in the EstOF network (220–172, 122–172, and 122–133, red) and three pairs of stations in the long-term network (SCHL-LAMP, KUHL-LAMP, KUHL-GUNS, black, NCFs after Lehujeur et al., 2016). The NCFs are band-pass filtered between 1 and 5 s. Amplitudes are expressed in arbitrary units. The numbers Rc and Ra (for causal and acausal sides, respectively) indicate the ratios between the peak amplitude in the direct arrivals window (group velocity between 0.3 and 3.5 km/s, green) and the coda window (group velocity between 0.15 and 0.3 km/s, blue). Rc and Ra are computed for both waveforms in each subplot (black and red numbers, respectively, for the long-term and EstOF networks).](image)

![Figure 4. Convergence of the noise correlation function for three pairs of nodes in the EstOF network (220–172, 122–172, and 122–133). The ratio (R) between the peak amplitude in the direct arrival window (group velocity between 0.3 and 3.5 km/s) and the coda window (group velocity between 0.15 and 0.3 km/s) is measured for both causal (thick curves) and acausal sides (thin curves).](image)
Figure 5. Comparison between the noise correlation functions obtained with the EstOF short-term dense seismic network (a, b) and the long-term sparse network for seismic monitoring of the area (c, d, correlation functions after Lehujeur et al., 2016). The filtered correlation functions are averaged in distance bins of 150 m (a, c) and converted to the frequency-wave number domain (b, d). Black arrows indicate repetitive artifacts from the EstOF instruments in the time-distance (a) and frequency-wave number (b) domains.

Figure 6. Average noise correlation wavefield obtained with the EstOF network. (a) Noise correlation functions averaged over distance bins of 150 m and cleaned up in the frequency-wave number domain to retain phase velocities between 0.7 and 7 km/s, periods between 0.28 and 5 s (0.2–3.5 Hz) and wavelengths below 20 km. (b, c) Modulus of the 2-D Fourier transform of the noise correlation functions, displayed in the period-phase velocity domain for causal and acausal sides, respectively. Amplitudes are color coded and expressed in decibels. White and red squares indicate the theoretical Rayleigh wave group velocity dispersion curves for the fundamental mode (R0) and the first overtone (R1), respectively. The label P indicates P waves.
network improves our ability to identify the main phases of the signal in both the time-distance (Figures 5a and 5c) and frequency-wave number domains (Figures 5b and 5d). This improvement is particularly significant for the acausal side of the correlation functions. The EstOf NCFs suffer from repetitive spikes that occur with a periodicity of 2 s in the time domain (see black arrows in Figure 5a) and an apparent zero wave number (i.e., infinite phase velocity, see black arrows in Figure 5b). Such artifacts are not visible on the long-term networks equipped with higher-quality seismometers (Figures 5c and 5d). They are likely caused by the GPS time synchronization within the EstOf sensors, as reported in other studies that used similar equipment (Wang et al., 2017).

To analyze the dominant wave types that emerge from the signal, we filter the average NCFs of the EstOf network (Figure 5a) in the F-K domain (or equivalently in the period-phase velocity domain). The F-K filter isolates phase velocities between 0.7 and 7 km/s, periods between 0.28 and 5 s (0.2–3.5 Hz) and wavelengths below 20 km (Figure 6). The fundamental mode and first overtone of the Rayleigh waves dominate at periods between 0.8 and 5 s (Figures 6b and 6c). These two arrivals are in good agreement with the theoretical dispersion curves predicted using the one-dimensional depth velocity profile for Soultz-sous-Forêts (Beauce et al., 1991; Charléty et al., 2006, Figures 6b and 6c, red and white squares). We also identify arrivals with higher velocity, approximately 5 km/s, which we interpret as P waves (see label P on Figure 6c). This wave type has already been observed in some previous studies based on ambient noise correlation (e.g., Nakata et al., 2016; Roux, Sabra, Kuperman, Roux, 2005). To verify this statement, we compute the theoretical P wave traveltimes in the 1-D velocity model for Soultz-sous-Forêts (Figure 7a). To the field, which has a higher signal-to-noise ratio, making tomography easier.

2.3. Group Velocity Dispersion Measurements

Using both causal and acausal NCFs, we obtain more than 80,000 estimates of the medium Green’s functions between EstOf node pairs. We decompose them in the time-frequency domain using the multiple Gaussian filter approach (e.g., Dziewonski et al., 1969; Levshin et al., 1992). We build dispersion diagrams using the envelope of the analytical signal filtered around several periods (Figure 8). For each period, we normalize the envelope amplitude by the standard deviation of the filtered trace (Figure 8, color scale). We automatically pick all the local maxima of the diagrams using the discrete derivatives of the envelope. For each local maximum (Figure 8, black circles), we collect (1) the instantaneous period  T, (2) the group velocity V, (3) the amplitude of the dispersion diagram at the pick location SNR, (4) the period-dependent ratio between the interstation distance and the wavelength d/λ (computed using the theoretical phase dispersion curve of the fundamental mode of the Rayleigh waves in the Soulz model, Figure 6c, white curve), and (5) the back azimuth measured in degrees, clockwise from north (β ranging from 180 to 360°N for the causal side and from 0 to 180° for the acausal side of the NCFs, according to the chosen convention for the orientation of the station pairs).

Such an approach avoids analyzing all the individual dispersion diagrams and can detect multiple modes as observed on many station pairs for periods between 1 and 2 s (Figure 8 and Text S2 in the supporting information). The large number of automatic picks obtained for periods below 1 s (Figure 8) illustrates the difficulty of recovering Green’s function in the anthropogenic period band in this region (Lehujeur et al., 2018).

Figure 7. Comparison between the theoretical traveltimes predicted using ray tracing in the 1-D P wave velocity model for Soultz-sous-Forêts (a) and the noise correlation functions of the EstOf network (b). The causal and acausal sides from Figure 6a are band-pass filtered between 0.28 and 1.67 s (0.6–3.5 Hz) and averaged to highlight the P wave phase.

Figure 8. Example of a group velocity dispersion diagram obtained for the causal side of the EstOf station pair 036–210 (Figure 1). Colors indicate the ratio between the envelope of the analytical signal and the standard deviation of its real part for each instantaneous period. Dots indicate the local maxima retained by the automatic picking algorithm. White curves indicate the theoretical group velocity dispersion curves for the fundamental mode and the first two overtones using the Soultz-sous-Forêts velocity model.
Although unlikely, dispersion picks with high group velocities are preserved to analyze their repeatability over the whole data set.

To get an overview of the automated pick distribution for all station pairs, we isolate slices across the five-dimensional (5-D) pick domain (T, V, β, SNR, d/λ), (see Figure 9). The distribution of periods (T) versus group velocities (V) for fixed ranges of back azimuth (β), distance over wavelength (d/λ), and SNR (Figure 9a) is a plane section through this 5-D domain. It highlights a bimodal average group velocity dispersion curve in good agreement with the theoretical curves for the fundamental mode and first overtone (white and red squares, respectively).

For periods above 2 s, we notice that the group velocity (V) stabilizes for back azimuths (β) near the two dominant noise directions observed in this region (Figure 9b): the Atlantic Ocean (back azimuth ~150°) and the Mediterranean Sea (back azimuth ~280°, Lehujeur et al., 2016). These directions are characterized by high pick densities in all the plane sections of the 5-D distribution (Figures 9f, 9i, and 9j), and they slightly change with the period (T, Figure 9j). For some directions, the picked group velocity diverges toward unrealistically high values (see the white dashed curve in Figure 9b). This divergence of the apparent group velocity as a function of the back azimuth results from directive and not fully diffuse seismic noise (Pedersen & Krüger, 2007). The azimuthal biases increase with increasing period (and wavelength, e.g., Yao & Van Der Hilst,
2009) and become very strong for periods above 4 s (Figure 10b). These directivity patterns in the noise also affect the phase of the correlation functions (Text S1).

To improve the Rayleigh wave tomography and avoid spurious effects, we select clusters of picks in this 5-D domain for the inversion. To define the boundaries of each cluster, we first adjust the period range to a narrow band centered on a specific period and adjust the velocity range accordingly depending on the targeted mode number (e.g., Figures 9a, 10a, 10d). We then adjust the back azimuth range to avoid directions where the velocity measurements are presumed to be strongly affected by the azimuthal biases (Figures 9b, 10b, and 10e). The distance over wavelength ratio is taken above 2 when possible (Figure 9c). For periods above 3 s, we are forced to reduce this threshold to 1 (Figures 10a–10c) to include enough interstation paths for tomography. This choice also contributes to strong azimuthal biases visible in the back azimuth group velocity domain, with significant pick densities for group velocities up to 5 km/s and above (Figure 10b). We finally adjust the lower SNR boundary in the SNR-velocity domain, since low-SNR picks are often associated with unrealistic velocity values (Figure 9d). For periods below 2 s, we isolate picks with SNR above 4, which is observed to improve the distinction between the fundamental mode and the first overtone (Figure 10f). We obtain 15 clusters of picks from periods near 0.6 to 4.2 s, and we attribute a mode number to each of them (0 for the fundamental mode or 1 for the first overtone) depending on its position in the period group-velocity domain (Figure 11). The dispersion curves obtained are relatively smooth and consistent between similar interstation paths. These clusters of picks are then used as input for Rayleigh wave tomography.

3. Results

3.1. Group Velocity Maps

For each cluster of picks (i.e., each period and mode number), we compute a group velocity map using “Gaussian surface wave tomography”, which assumes straight interstation rays and approximates the surface wave sensitivity kernels by a narrow region surrounding the interstation path.
Taking into account the true diffraction kernels and the ray bending effects could increase the accuracy of the resulting maps (Ritzwoller et al., 2002). However, it would render the tomographic inversion nonlinear and is beyond the scope of this study. The period range extends from 0.6 to 4.5 s for the fundamental mode and from 0.7 to 1.9 s for the first overtone (Figure 11). The inversion is regularized with a damping parameter, which controls the relative weight of the quadratic data misfit and the norm of the model, and a smoothing parameter to control the nondiagonal terms of the covariance matrix of the prior model (we assume that the velocities in two cells \(i, j\) of the dispersion map have a linear correlation coefficient with the form \(\exp(-d_{ij}/s)\), where \(d_{ij}\) is the distance between the two cells and \(s\) is the smoothing coefficient). These regularization parameters are adjusted individually for each map to minimize both the misfit between the observed and predicted traveltimes and the norm of the velocity anomalies. For each map, we compute the resolution matrix following Tarantola (2005). Pixels outside the resolved zone (i.e., cells for which the diagonal term of the resolution matrix is too low) are masked in the final maps. The resulting maps are shown in Figure 12 at different periods both for the fundamental mode (Figures 12a–12d) and for the first overtone (Figures 12e and 12f). The average velocity obtained increases with the period (Figure 12, see the reference velocity in the subplot titles). The relative velocity variations of the group velocity exhibit similar patterns for all periods and mode numbers. In particular, the prominent high-velocity zone observed to the northwest is in good agreement with the northern Vosges massif. Further details are provided in section 4.

### 3.2. Depth Inversion

To build a three-dimensional \(S\) wave velocity model of the area, we invert the bimodal (fundamental mode and first overtone) Rayleigh wave dispersion curves obtained in each pixel of the group velocity maps (Figure 12). Such inversion can be performed using a linearized approach (Aki & Richards, 2002; Dorman & Ewing, 1962; Herrmann, 2013; Xia et al., 1999), which is fast but highly sensitive to the depth model chosen to initiate the inversion. Here we use a Monte Carlo approach (e.g., Maraschini & Foti, 2010; Shapiro & Ritzwoller, 2002; Socco & Boiero, 2008), which can explore a larger parameter space and identify multiple solutions. The one-dimensional depth model is parameterized with nine layers including a half-space at the bottom (Table 1). We invert for the \(S\) wave velocity and the depths of the interfaces. Since the computation of Rayleigh wave group velocity dispersion curves further requires a \(V_p\) and a density model, we also invert for the \(V_p\) over \(V_s\) ratio and the density in each layer, but we impose a stronger prior constraint on these variables based on borehole observations as detailed below. We estimate the posterior probability density
distribution of the model space by combining prior probability density functions of the model and data spaces (Tarantola, 2005). We estimate the quality of a depth model using

$$\log lik (m) = \log p_M (m) + \log p_D (g (m))$$

(1)

where $\log lik$ is the logarithm of the likelihood, $m$ is an array of parameters corresponding to a depth model, $p_M$ is the prior probability density function of the model space, $p_D$ is the probability density function of the data space, and $g$ is the theory function used to compute dispersion curves. Taking the logarithm of the likelihood increases the stability of the metropolis algorithm.

The prior probability density function of the model space ($p_M$ in equation (1)) is defined as the product of uniform probability density functions on each parameter of the model (Table 1). The prior constraints on $V_s$ and $V_p/V_s$ in each layer are adjusted after borehole observations at Rittershoffen (Maurer et al., 2016) and Soultz-sous-Forêts (Beauce et al., 1991; Charléty et al., 2006; Cuenot et al., 2008; Dorbath et al., 2009). These measurements have revealed $V_p$ over $V_s$ ratios of approximately 1.73–1.75 in the granitic basement and very high values up to 2.15 in the upper part of the sedimentary cover. The $S$ wave velocity of the half space is constrained to the narrow range 3.2–3.6 km/s according to velocities observed in the granitic basement at both Rittershoffen and Soultz-sous-Forêts (Beauce et al., 1991; Maurer et al., 2016). The prior constraints on density parameters are set after gravimetric models of the Upper Rhine Graben (Rotstein et al., 2006). To increase the smoothness of the depth models, we impose additional prior constraints to the offsets between neighboring layers using uniform probability laws. The $V_p$ and $V_s$ offsets are restricted to the range $-0.5$ to $+1.5$ km/s. We impose a decreasing $V_p$ over $V_s$ ratio with depth (offsets bounded to the range $-1.0$, $0.0$) and increasing density (offsets between $0.0$ and $+1.0$ g/cm$^3$).

The prior probability density function of the data space ($p_D$ in equation (1)) is defined using lognormal probability laws for each point in the dispersion curve, and we assume that the group velocity measurements are independent (i.e., we use a diagonal data covariance matrix, Tarantola, 2005).

For each location, we invert the observed dispersion curves with 12 independent Markov chains running in parallel. The moves between subsequent models explored by the random walk process are governed by a Gaussian proposal probability density function, whose covariance matrix is taken as diagonal. The diagonal terms are adjusted along the inversion to stabilize the acceptance ratio of the chains to approximately 25% (i.e., on average, one model is kept for four tests). The forward computation of the dispersion curves for each depth model is done with the programs by Herrmann (Computer Program in Seismology, Herrmann, 2013). Each chain runs until 1,000 models have been retained. We end up with 12,000 models kept from the $\sim$48,000 tests for each inversion. The first models generated by the chains have very low likelihood values due to unsatisfied prior conditions. We also observe Markov chains that are trapped in local maxima of the posterior probability density function (for instance, models that fit the data reasonably well but do not fulfill the prior conditions). To exclude such models, we select only the 2,000 best models found by the 12 chains. This number is set empirically based on the analysis of the evolution of the model likelihood of the Markov chains. The median of these 2,000 best models is retained as the solution to the inversion. We use the median rather than the mean since it is less sensitive to outliers. We run 498 inversions independently for each pixel in the resolved zone of the dispersion maps (Figure 12). For all inversions, we obtain a relatively uniform distribution of $V_p/V_s$ and density between the imposed boundaries (not shown) due to the low sensitivity of Rayleigh wave velocities to these parameters (e.g., Xia et al., 1999, see Text S3). We combine the 1-D $S$ wave velocity models obtained in every pixel to form a 3-D model (Figure 13).

To quantify the data fitting, we compute the dispersion curves corresponding to the solution of the depth inversions (Figures 14b–14d, thick black depth models) and compare them to the dispersion curves from the dispersion maps (Figures 14b–14d, red dispersion curves). We observe lower likelihood values in the northwestern and southeastern part of the model (Figure 14a), which probably results from an increased complexity of the model in these regions. The solution obtained near the geothermal site of Rittershoffen...
Figure 13. Three-dimensional S wave velocity model. (a–f) Horizontal slices at different depths. Colors indicate velocity variations relative to a reference velocity indicated in the subplot title. The black dashed line indicates the position of the section of subplot (g). (g) Northwest-southeast section across the model. Amplitudes correspond to velocity variations relative to the median velocity at each depth (upper section) or to absolute velocity (lower section). The S and R labels indicate the positions of the Soultz-sous-Forêts and Rittershoffen sites, respectively. The horizontal dashed lines indicate the depth of the slices (a–f). Black and white solid lines indicate the main faults and the top of the granitic basement, respectively (GeORG Team, 2013).

is in good agreement with sonic measurements performed in borehole GRT1 (Figure 14d, purple curve, from Maurer et al., 2016).

4. Discussion

The primary pattern that emerges from both the group velocity dispersion maps and the inverted S wave velocity model is a positive velocity anomaly to the northwest, which corresponds to the northern Vosges massif (Figures 12 and 13, see the high-velocity zone to the northwest) separated from the sedimentary plain of the Rhine by the Rhenish fault, a major geological discontinuity characterized by marked changes in the topography (Hochwald crest, Figure 1) and the surface geology (Ménillet et al., 1989) and dominating the Bouguer anomaly map of the Upper Rhine Graben (Corpel & Debeglian, 1994; Edel et al., 2007; Rotstein et al., 2006; Schill et al., 2010).

The S wave velocity obtained at shallow depth is correlated with the surface geology. Three patches with low-velocity anomalies are observed on the eastern side of the map (Figure 13a, between Haguenau and Soufflenheim, southeast of Rittershoffen and near Seebach). These three zones are in good agreement with low-density zones revealed by gravity data (Corpel & Debeglian, 1994) and may correspond to areas covered with shallow and poorly consolidated layers formed by fluvial, lacustrine, and marshy deposits during the Pliocene and loess deposits during the Quaternary (Aichholzer et al., 2016; GeORG Team, 2013). Such low-
velocity anomalies do not exist in the southwestern quadrant of the map (from Haguenau to Betschdorf) and north of Soultz-sous-Forêts, where most of the Pliocene-Quaternary formations have been eroded, together with the upper part of the Tertiary layers (Froidefontaine and Pechelbronn formations, Aichholzer et al., 2016).

At greater depths, the similarity between the observed velocity and the known lithology is not straightforward. The northeastern zone of the map (Seebach area) exhibits low velocities at depths greater than 1.4 km, which might correspond to a zone with thicker and more fractured sedimentary cover (GeORG Team, 2013). We do not identify a clear correlation between the variations in the S wave velocity and the known depth of the granitic basement, observed at 1.4-km depth at Soultz-sous-Forêts and 2.2 km at Rittershoffen (Aichholzer et al., 2016; Baujard et al., 2017). Furthermore, a velocity contrast between the two sides of the Rhenish fault is visible at depths below ~1.5 km (Figures 13d–13g), where the fault is presumed to separate two similar geological units, that is, the granitic basement (Debelmas, 1974; Kappelmeyer et al., 1992; Schnaebele, 1948). This effect could result from the poor sensitivity of surface waves to deep interfaces (Text S3), leading to a low vertical resolution for the velocity model. We cannot rule out the

Figure 14. Fit to the dispersion curves for each location in the model. (a) Log-likelihood of the solution of the depth inversion (median of the 2,000 best models) for each location in the model (color bar). A value of 0 corresponds to a perfect fit to the observed dispersion curves. (b, c) Results of the depth inversion at three example locations (black dots on map a). The dispersion curves to invert are displayed in red. RU0 and RU1 correspond to the group velocities of the fundamental mode and first overtone of the Rayleigh waves, respectively. The colored curves correspond to the 2,000 best models obtained, and in the corresponding dispersion curves, colors correspond to log-likelihood values. Dashed black models indicate the prior boundaries imposed on the inversion. Solid black models indicate the 1st, 16th, 84th, 99th (thin), and 50th (thick) percentiles of the posterior distribution. Purple curve on subplot (d) corresponds to S wave sonic measurements performed in borehole GRT1 at Rittershoffen (Maurer et al., 2016).
possibility that the observed deep velocity contrasts are caused by underestimated velocity anomalies in the shallow part of the model, leading the depth inversion process to deteriorate the deepest part of the model in order to fit the observed dispersion maps. This question could be addressed either by adding more prior constraints on the shallow part of the model based on the numerous borehole measurements conducted in the area over the last decades to study the first kilometer of sediments (e.g., Aichholzer et al., 2016; Dezayes et al., 2005; Hooijkaas et al., 2006) or by analyzing Rayleigh waves at periods below 0.8–1 s (i.e., within the anthropogenic period band), since these periods can be helpful to constrain the shallow structures and therefore improve the inversion at depth (Lehujeur et al., 2017).

The lateral variations observed at depth and particularly in the basement could also indicate that parameters such as the fluid content, the mechanical damage, or the temperature significantly influence the observed S wave velocity (Fjar et al., 2008; Guéguen & Palciauskas, 1994; Paterson & Wong, 2005). For instance, we observe horizontal variations of approximately 30% peak to peak within the granitic basement along the NW-SE section (Figure 13g) that are not consistent with the lithology (GeORG team, 2013). Interestingly, the lateral scale of these variations is comparable to the typical scale of the hydrothermal cells in this region (2–6 km, Guillou-Frottier et al., 2013; Magnenet et al., 2014), which are correlated with both the temperature (Clauer et al., 2002) and the density field (Bailieux et al., 2013). As a result, we suggest that the S wave velocity variations we obtain through seismic noise tomography may be related to the natural hydrothermal circulation responsible for the thermal anomalies in the region.

5. Conclusion

Miniaturized and wireless seismological equipment is much simpler to deploy than conventional seismological stations, in terms of both site selection and installation. It allows operators to deploy large and dense seismological networks in a short period of time, which minimizes the deployment costs and the impacts on the population. The ability of such instrumentation to record ambient seismic noise at periods up to approximately 5 s allows us to probe the first few kilometers of the crust using surface waves. The recording duration is a key parameter in noise correlation studies, since it must be sufficient to allow convergence of the NCFs. No general rule can be established to determine the shortest duration required, since it may vary with the site, the period band, or the processing applied before correlation (e.g., Bensen et al., 2007; Groos et al., 2012). In the EstOF case, we can recover the fundamental mode and first overtone of the Rayleigh waves after 30 days of continuous recording, with a signal-to-noise ratio that is sufficient to perform tomography. For most station pairs, the quality of the obtained NCFs is comparable to the quality of those obtained with noise records of more than 1 year in the same area (Figures 3 and 5).

The large amount of data produced with such types of networks requires automated processing procedures to measure the interstation traveltimes. The automatic picking approach we use minimizes user interventions in the data processing, but interpretation of the statistical distribution of the group velocity picks (Figures 9 and 10) remains necessary to prevent mode number misidentification or azimuthal biases due to imperfectly diffusive seismic noise, both of which could lead to significant errors in the final velocity model.

To the best of our knowledge, the EstOF array is the largest and densest passive seismological network ever deployed in the Outre-Forêt region. The ambient seismic noise overlies the instrumental noise level at periods up to approximately 5 s, which allows us to exploit the lower period part of the oceanic secondary microseismic peak and thus to probe the subsoil at the depth of the geothermal reservoir using surface waves. The fundamental mode and first overtone of the Rayleigh waves clearly dominate the noise correlation functions in the studied period band, although P waves have been identified in the signal. We obtain Rayleigh wave traveltime estimates between more than 41,000 virtual source-receiver pairs, which allows us to discriminate observations from a statistical point of view. In particular, this process allows us to mitigate the influence of directivity patterns in the ambient seismic noise induced by the dominant sources in the Atlantic Ocean and Mediterranean Sea and the non fully diffusive properties of the seismic noise.

The obtained Rayleigh wave group velocity maps and 3-D S wave velocity model of the Outre-Forêt region exhibit significant velocity anomalies that are consistent with previous geological and geophysical observations. However, several questions remain about the influence of nonlithological parameters on the observed velocities such as the fluid content, the mechanical damage or the temperature. Future work could concentrate on the interpretation of the obtained velocity model and the potential use of such techniques in a
purely explorational context. The accuracy of the final model may also be improved by considering the ray-bending effects due to large velocity variations in the studied region using methods based on ray tracing (e.g., Fang et al., 2015; Saygin & Kennett, 2010) or eikonal tomography for phase velocity (e.g., Lin et al., 2013; Ritzwoller et al., 2011).

The velocity model obtained might also be used to locate the induced seismicity observed at the geothermal power plants, which could potentially provide further information about the response of the geothermal reservoir to stimulations and could help operators to minimize the operational risks.

Acknowledgments
This work has been published under the framework of the LABEX ANR-11-LABX-0050-G-EAU-THERMIE-PROFONDE and benefits from a funding from the state managed by the French National Research Agency (ANR) as part of the “Investments for the Future” program. It has also been funded by the EGS Alsace grant from ADEME. The authors would like to deeply thank the EstOF Team (R. Dretzen, M. Schaming, O. Lançon, S. Lambotte, M. Bès de berc, A. Maggi, O. Lengliné, C. Dambre, Z. Duputel, M. Grunberg, F. Engels, H. Jund, H. Woddling, H. Blumentritt, N. Cuenot, and V. Maurer) for providing an extended support for the field deployment of the EstOF network and Florent Brenguier for sharing his experience and efforts and coorganizing the material management sharing his experience and efforts and the EstOF network and Florent Brenguier for organizing the material management at the EGS site. The authors thank Géothermie for providing the sonic log file. Numerous fruitful scientific discussions with F. Brenguier, P. Jousset, and K. Fair during the summer 2014. We thank the Fairfield Nodal company for providing the equipment and support during the EstOF experiment. The authors thank GEIE-EMC, ES-Géothermie, EOST, and EstOF experiment. The authors thank Géothermie for providing the sonic log profile. Numerous fruitful scientific discussions with F. Brenguier, P. Jousset, F. Corret, T. Fischer, T. Kohl, H. Pedersen, P. Roux, A. Mordret, G. Millers, M. Campillo, H. Karabolut, and L. Margerin are acknowledged. Data used in this study are accessible at https://dx.doi.org/10.25577/2014-ESTOF.

References


Hoekkaas, G. R., Genter, A., & Dezayes, C. (2006). Deep-seated geology of the granite intrusions at the Soultz EGs site based on data from 5km-deep boreholes. Geothermics, the deep EGs (enhanced geothermal system) project at Soultz-sous-Fors Ét, Alsace, France 35, 484–506. https://doi.org/10.1016/j.geothermics.2006.03.003


