

UC San Diego

UC San Diego Previously Published Works

Title

High-resolution seismic tomography of Long Beach, CA using machine learning

Permalink

<https://escholarship.org/uc/item/5c2484mj>

Journal

Scientific Reports, 9(1)

ISSN

2045-2322

Authors

Bianco, Michael J
Gerstoft, Peter
Olsen, Kim B
et al.

Publication Date

2019

DOI

10.1038/s41598-019-50381-z

Peer reviewed

OPEN

High-resolution seismic tomography of Long Beach, CA using machine learning

Michael J. Bianco^{1*}, Peter Gerstoft¹, Kim B. Olsen^{1,2} & Fan-Chi Lin³

We use a machine learning-based tomography method to obtain high-resolution subsurface geophysical structure in Long Beach, CA, from seismic noise recorded on a “large-N” array with 5204 geophones (~13.5 million travel times). This method, called locally sparse travel time tomography (LST) uses unsupervised machine learning to exploit the dense sampling obtained by ambient noise processing on large arrays. Dense sampling permits the LST method to learn directly from the data a dictionary of local, or small-scale, geophysical features. The features are the small scale patterns of Earth structure most relevant to the given tomographic imaging scenario. Using LST, we obtain a high-resolution 1 Hz Rayleigh wave phase speed map of Long Beach. Among the geophysical features shown in the map, the important Silverado aquifer is well isolated relative to previous surface wave tomography studies. Our results show promise for LST in obtaining detailed geophysical structure in travel time tomography studies.

In contrast to seismic events, ocean-atmospheric interactions provide a nearly continuous excitation of the solid Earth^{1,2}. Observations of these apparently random, low amplitude seismic waves on seismic sensor arrays, often referred to as seismic “noise”, can yield rich information about geophysical properties within the region of the array^{3–5}. In ambient noise tomography (ANT), this noise is cross-correlated between seismic sensors over periods of days to months to obtain travel times between sensors^{6–14}. These travel times are used to perform tomography, estimating the subsurface phase speed structure for a region of interest^{4,6}. Despite the fact that the number of travel times in ANT can be very large ($N(N-1)/2$, with N the number of sensors) and the coverage of a region dense, the estimation of high-resolution phase speed structure with ANT remains a difficult problem due to many factors, e.g. irregular sensor distributions, phase ambiguities in the cross-correlations, and non-isotropic noise distributions⁸. Here we obtain high-resolution subsurface geophysical structure in Long Beach, CA using a machine learning-based tomography method, called locally sparse travel time tomography (LST)¹⁵. Existing ANT methods potentially oversimplify true geophysical structure by assuming phase speed models are smooth. LST improves the fidelity of ANT phase speed maps by using dictionary learning^{16,17}, a form of unsupervised learning, to learn the relevant geophysical characteristics directly from seismic data. Indeed, we find that the 1 Hz Rayleigh surface wave phase speed map obtained with LST illuminates geophysical features unseen in previous surface wave tomography studies of the region. Specifically, the Silverado aquifer which supplies nearly 90% of the fresh water in Long Beach is well isolated relative to previous surface wave tomography studies. Our results show LST improve estimates of geophysical structure in travel time tomography studies.

Recently, machine learning techniques have found many useful applications in seismology^{18,19}, including seismic waveform classification^{20,21}, event localization^{11,22}, earthquake prediction²³, and earthquake early warning²⁴. In part, the success of these methods is derived from large amounts of training and ground-truth data. In ANT however, little training data exists. LST addresses this issue by using an unsupervised machine learning method, called dictionary learning, to constrain slowness features in the tomographic image. This procedure is derived from the *adaptive* dictionary learning paradigm from image processing^{17,25,26}, in which dictionaries are learned directly from corrupted measurements. In adaptive image denoising²⁵, the dictionary is trained on small rectangular groups of pixels, called patches, of a noisy image. In LST, slowness dictionaries are learned from patches of a least squares-regularized inversion, and are subsequently used to reconstruct a sparsity-constrained slowness image. The dictionary is initially unknown and is learned iteratively, assuming sufficiently dense ray sampling

¹NoiseLab, University of California San Diego, La Jolla, California, USA. ²Department of Geological Sciences, San Diego State University, San Diego, California, USA. ³Department of Geology and Geophysics, University of Utah, Salt Lake City, Utah, USA. *email: mbianco@ucsd.edu

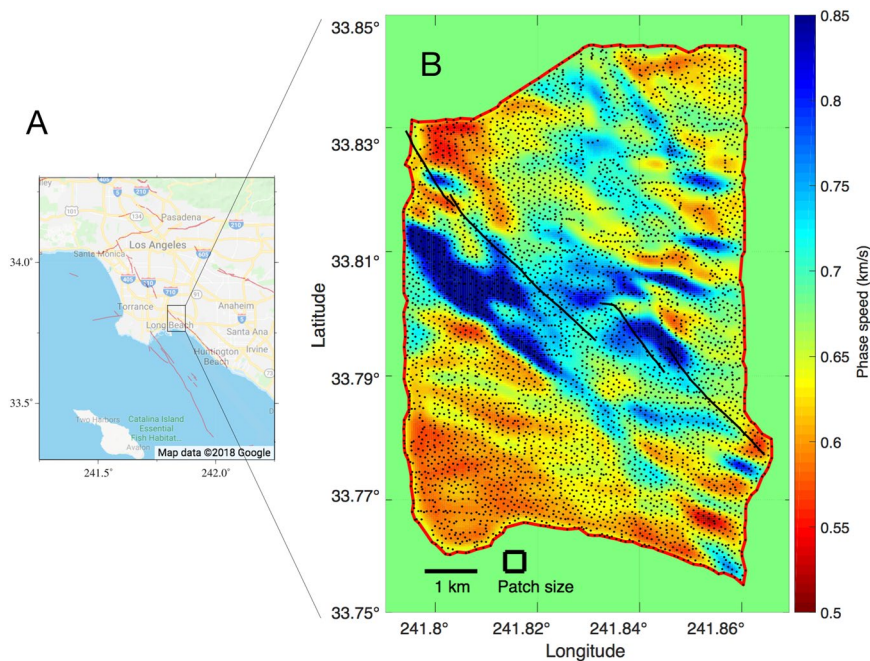


Figure 1. Location of Long Beach Array and LST phase speed map. The Long Beach array located in (A, modified from *Google Maps*⁴⁸), contained 5204 stations (dots in (B)) distributed over a ~ 70 km² area. (B) Locally sparse tomography (LST) phase speed maps of Long Beach, CA at 300×206 pixel resolution, using 3 million travel times. The Newport-Inglewood (NI) fault network (black lines) and the valid boundary for LST (red line) are shown.

(details in Methods). Relative to previous travel time tomography methods which enforce either smooth or discontinuous models, including conventional straight ray²⁷ and eikonal tomography⁸, the sparse model in LST permits both smooth and discontinuous geophysical features. This approach is related to wavelet based methods²⁸, however in LST the atoms are adapted to the slowness features in the data using dictionary learning.

Results

Using LST, we perform surface wave ANT with data from the Long Beach array (Fig. 1A,B). The Long Beach array was deployed from January to June 2011 as part of a petroleum industry survey. It was a very dense, “large- N ” array with 5204 high-frequency vertical velocity sensors distributed over a 7×10 km area (Fig. 1A). We obtained Rayleigh surface wave travel times between all station pairs in the array by cross-correlating seismic noise recorded during the period 5–25 March 2011. With $N = 5204$, this yielded ~ 13.5 million travel times, which was reduced in preprocessing. We discretized the footprint of the Long Beach array into a 300×206 pixel (N-S \times E-W) phase-speed map (Fig. 1B), which corresponds to pixel sizes of 35×35 m. The phase speed for each pixel is estimated by LST with the Rayleigh wave travel times.

Travel time data. We use the 1 Hz Rayleigh surface wave band from the Long Beach data, which corresponds to near-surface geophysical features (~ 100 – 500 m depth). For each station pair, cross-correlations of all 1-h segments from the 3-week recording were normalized and stacked to obtain the causal and anti-causal travel times. The final travel times were obtained by averaging these causal and anti-causal components⁸. After quality control the number of useful cross-correlations for 1 Hz was ~ 8.5 million. Quality control included SNR thresholding and removal of travel times with ranges less than one wavelength. The cross-correlations further suffered from phase ambiguities, which were reduced in preprocessing. This was done by clustering the rays²⁷ and filtering travel times that exceeded the median travel times of the clusters by one half-period (0.5 s for 1 Hz). Further, cross-correlations were rejected if travel times from virtual source pairs disagreed by more than one-half period. This reduced the number of useful cross-correlations to ~ 3 million. For further details and preprocessing steps, see Methods and⁸.

LST implementation. The 1 Hz Rayleigh surface wave travel times are used by LST to estimate a 300×206 pixel phase speed image of the Long Beach region. We assume straight-ray surface wave propagation, which yields a simple linear measurement model (Eq. 1). Using the measurements, LST alternates between solving larger-scale, or *global*, phase speed features (details in Methods) and smaller-scale or *local* phase speed features. The global problem (Eq. 4) is solved by least squares. Since the tomography matrix is sparse, we use the sparse least square program LSMR²⁹. The local sparse problem (Eq. 5) is solved using orthogonal matching pursuit²⁵, and the dictionary is learned using the iterative thresholding and signed K-means algorithm³⁰. The reference phase speed is constant, and estimated as the average phase speed of all ray paths.

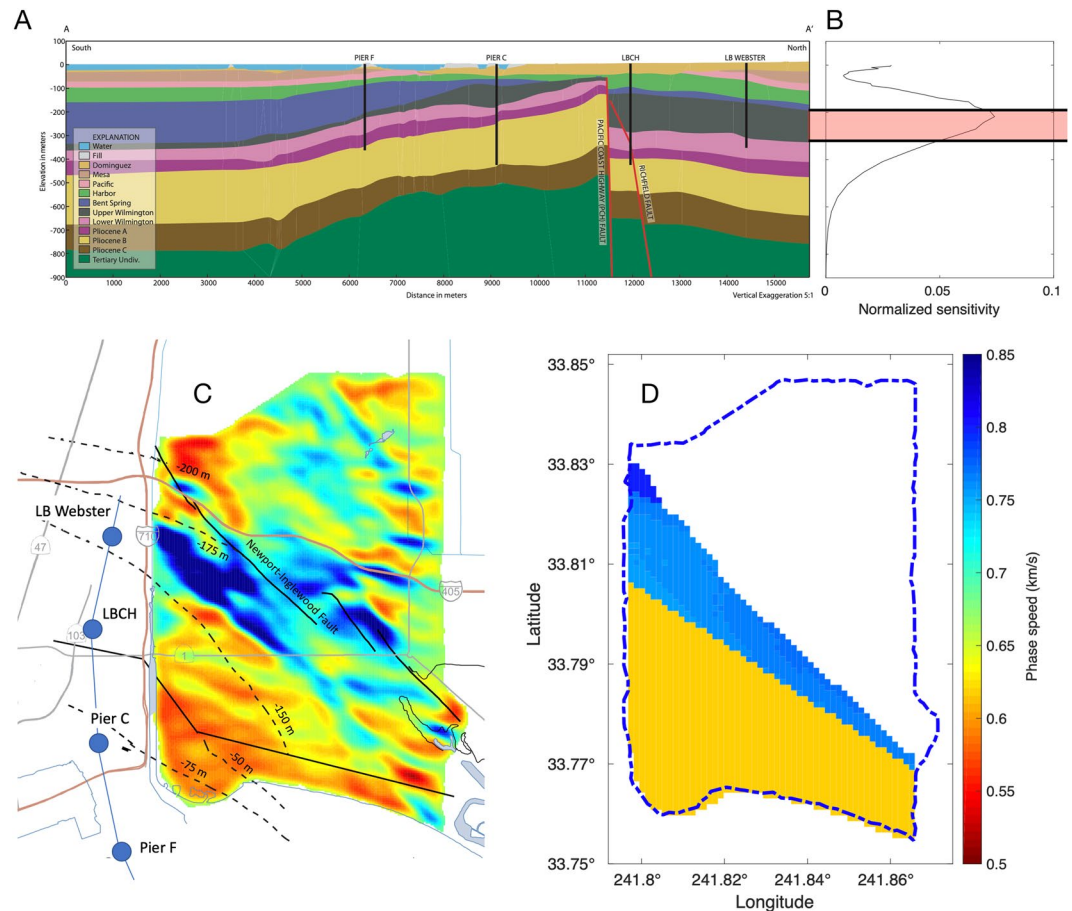


Figure 2. Comparison of survey-inferred geophysical features with LST phase speed map. (A) Inferred stratigraphy along ~N-S profile in (C, modified from³²), with the 4 deep wells used as constraints. The wells are located at Pier F, Pier C, Long Beach Cabrillo High School (LBCH), and Long Beach Webster (LB Webster) elementary school, 1 km west of the Long Beach array. (B) 1 Hz average V_s depth sensitivity kernel from⁸ and overlap of northern end of the inferred Silverado (shown in (C)) with V_s sensitivity, indicated in red. (C) Inferred Silverado elevation³² overlain with LST 1 Hz Rayleigh wave phase speed map (Fig. 1B, modified from³²), with wells (blue dots), and the location of the stratigraphy profile (blue line). (D) Hypothesized 1 Hz Rayleigh surface wave phase speed calculated from geological properties and inferred Silverado elevation range (C), from^{31,32}.

The LST tuning parameters are $\{n, T, Q, \lambda_1, \lambda_2\}$. n is the number of pixels per patch. We use square, 10 by 10 pixel patches, giving $n = 100$ pixels per patch (yielding $300 \times 206 = 61,800$ patches). We assume sparsity $T = 2$ (see Eq. 5), meaning that each patch uses two atoms from the dictionary \mathbf{D} . Each atom in \mathbf{D} is a vector with dimension $n = 100$ pixels (the patch size). We assume \mathbf{D} has $Q = 200$ atoms, or twice the patch dimension. \mathbf{D} is initialized with Gaussian random vectors of unit norm. λ_1 , which is the ratio of travel time error variance to global slowness variance is set as $\lambda_1 = 13 \text{ km}^2$ (see Eq. 4). λ_2 , which is the ratio of patch slowness variance to global slowness variance, is set as $\lambda_2 = 0$ assuming a sparse slowness representation (see Eq. 5). When $\lambda_2 = 0$, the sparse slowness is simply the average of the patch slownesses (see Methods).

Interpretation of phase speeds. In the middle of the LST phase speed image (Fig. 1B), particularly to the West, a large fast anomaly is observed between 33.78° and 33.82° latitude. This fast anomaly corresponds well with the Upper Wilmington (UW) Quaternary formation, which includes the Silverado water bearing unit (Lower Pleistocene age, ~300–580 ka) that supplies nearly 90% of the total ground water extracted in Long Beach^{31,32}. Based on a 3D model³² (Fig. 2A) of the water-bearing structures in Long Beach obtained using borehole, seismic, and gravity surveys, the Silverado is significantly denser ($2,290 \text{ kg/m}^3$) than the surrounding formations ($2,050\text{--}2,100 \text{ kg/m}^3$) due to its coarse-grained facies. From empirical relations of density and seismic wave speeds, we expect the UW formation to increase V_s by ~150% relative to the surrounding formations (see Methods). Since Rayleigh wave phase speed is dependent primarily on V_s , we conclude that this high velocity region of the map likely corresponds to the Silverado aquifer.

The proposed attribution of the fast anomaly to the Silverado aquifer is further supported by simulating the 1 Hz Rayleigh surface wave phase speed (Fig. 2D) using the Silverado depth range inferred in³² and³¹ (see Methods). In the region of the survey used from³², where both Silverado depth and thickness were available, the simulation shows a gradual increase in phase speed from south to north. However, per³¹, the Silverado is likely

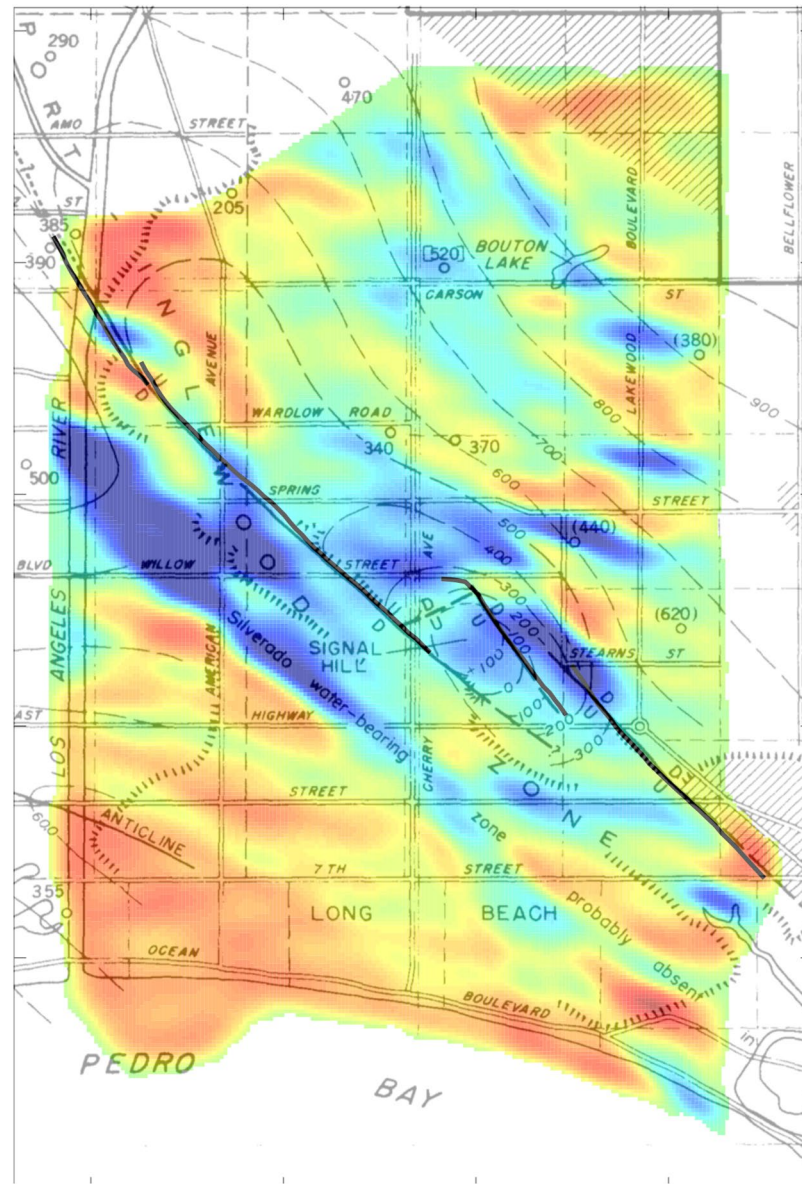


Figure 3. Silverado base depth (ft) inferred from borehole measurements in Poland *et al.*³¹ overlain with LST phase speed (Fig. 1B), modified from³¹. The results of³¹ contradict³², suggesting Silverado is absent south of the high phase speed anomaly in the west-central part of the LST map. The findings of³¹ are corroborated by the LST result. The lower extent of the high-speed anomaly is used to generate a hypothesized phase speed map (Fig. 2D). We also note that the phase speed anomalies near and to the north of Signal Hill (indicated above) correlate well with the contours.

absent about 1 km south of the NI fault in the region of the Long Beach array (Fig. 3). With this assumption for the simulation, the predicted trend and magnitude and phase speeds south of the NI fault compare well with the LST result (Fig. 1B). Relative to the eikonal tomography phase speed estimates of the Long Beach (Fig. 4B), the phase speed from LST better shows the broad fast region predicted by the simulation. It is clear that the Silverado is resolved particularly well by the LST in the west-central region. Moreover, well logs support an extension of the high-velocity anomaly toward the NE across the Newport-Inglewood (NI) fault zone. This is observed in the LST phase speeds (see Fig. 2D), beyond the region of the gravity survey. The LST result appears to corroborate the older study³¹, and contradicts some of the results of³², which was based exclusively on a gravity survey in the region of the Long Beach array.

Discussion

These results show that the LST method, by assuming that patches of seismic phase speed fields are repetitions of a set of few patterns, contained by the dictionary, can be used to further leverage existing seismic data to obtain high-resolution phase speed images from regions of interest. Since the dictionary (Fig. 5A) is learned from the travel time data via machine learning, it is well adapted to the true phase speeds. LST with dictionary

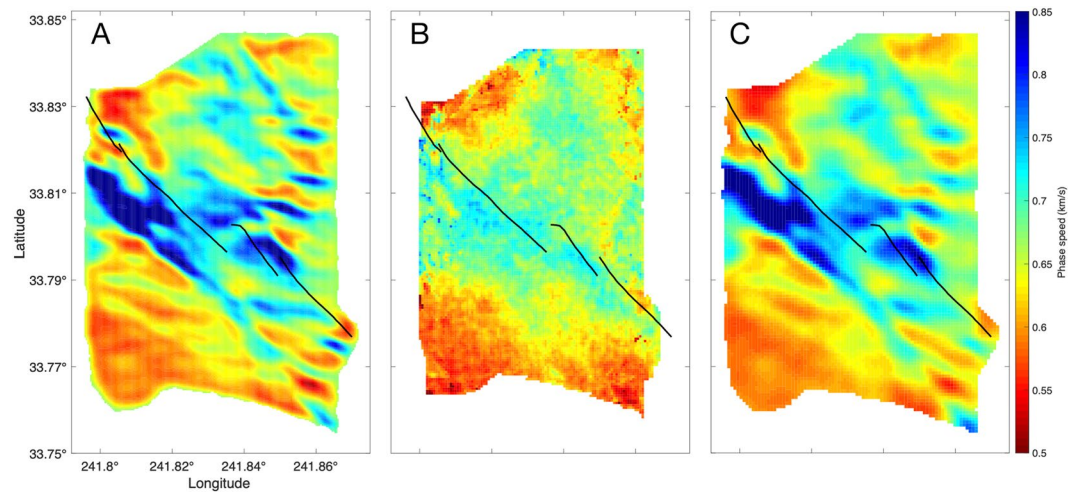


Figure 4. Comparison of 1 Hz Rayleigh surface wave phase speed maps from (A) LST, (B) eikonal, and (C) conventional tomography⁸ with the NI fault network (black lines). The general trends are the same for LST, eikonal, and conventional, though there is greater contrast and phase speed range observed in the LST map. There is greater contrast along the NI fault lines for LST. The largest disagreement between the LST/conventional result and eikonal tomography is in the western region of the map, where the LST is imaging the Silverado aquifer.

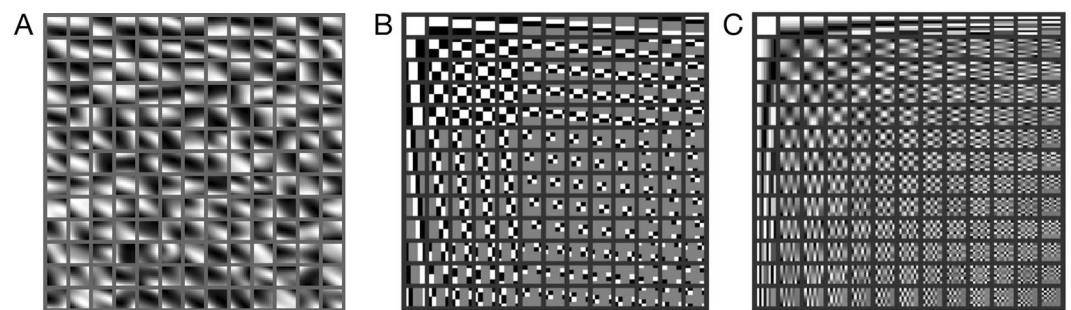


Figure 5. Comparison of (A) dictionary learned in LST inversion of Long Beach array data (Figs 1B and 2C,D), and generic dictionaries (B) Haar wavelet and (C) discrete cosine transform. All dictionaries shown with 169 atoms ($n = 100$). The atoms are sorted in order of decreasing variance from top to bottom (left to right). The learned dictionary atoms (A) with sharper, oriented gradients (higher variance) correspond to the sharper features in the LST phase speed map (e.g. the boundaries of the Silverado aquifer, Fig. 1B), whereas the smoother atoms (lower variance) are related to the smoother regions. Atom values stretched to full grayscale (0 to 1) for display.

learning provides a flexible model that is capable of modeling smooth and discontinuous slowness features. In the context of ambient noise tomography, LST leverages the dense sampling by learning the phase speed patterns. Thus, we obtain high-resolution slowness maps that well complement other geophysical sensing modalities and existing studies, for better estimating at least near-surface Earth structure. In this novel application of machine learning theory to near-surface seismic tomography, it is likely that we have further-characterized key water-bearing aquifers in Long Beach. We believe LST can help provide important geophysical insights in many tomography scenarios.

Methods

LST theory and implementation. Our proposed locally-sparse travel time tomography (LST) approach obtains high resolution by assuming that small patches of discrete phase speed maps are repetitions of few elemental patterns from a dictionary of patterns. Such patterns, referred to as atoms (chemistry analogy) are learned in parallel with the inversion using dictionary learning, a form of unsupervised machine learning. Relative to conventional tomography methods, the sparsity of the dictionary representation permits smooth and discontinuous, high-resolution features where warranted by the data. In the following, we present an overview of the LST theory. For more details, please see¹⁵.

In LST, surface wave propagation is approximated as straight ray paths through an $N = W_1 \times W_2$ pixel phase speed map, and the travel time perturbations $\mathbf{t} \in \mathbb{R}^M$ from a known reference for M rays are modeled as

$$\mathbf{t} = \mathbf{A}\mathbf{s}_g + \epsilon, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{M \times N}$ is the tomography matrix, $\mathbf{s}_g \in \mathbb{R}^N$ is the perturbation global slowness (inverse of speed), and $\epsilon \in \mathbb{R}^M$ is Gaussian noise $\mathcal{N}(0, \sigma_\epsilon^2 \mathbf{I})$, with σ_ϵ^2 the noise variance. We call Eq. 1 the *global model*, as it captures the large-scale features that span the discrete map and generates \mathbf{t} .

We consider a second slowness model perturbation $\mathbf{s}_s \in \mathbb{R}^N$, called the sparse slowness, in which $\sqrt{n} \times \sqrt{n}$ groups of pixels are represented as sparse linear combinations of atoms from a dictionary. The patches are selected from $\mathbf{s}_s \in \mathbb{R}^N$ by the binary matrix $\mathbf{R} \in \{0, 1\}^{n \times N}$, and modeled as

$$\hat{\mathbf{x}}_i = \arg \min_{\mathbf{x}_i} \|\mathbf{R}_i \mathbf{s}_s - \mathbf{D}\mathbf{x}_i\|_2^2 \text{ subject to } \|\mathbf{x}_i\|_0 = T, \tag{2}$$

where $\mathbf{R}_i \mathbf{s}_s \in \mathbb{R}^n$ is the i -th patch, $\mathbf{D} \in \mathbb{R}^{n \times Q}$ is the dictionary of Q atoms, $\hat{\mathbf{x}}_i \in \mathbb{R}^Q$ coefficient estimates for the i -th patch, and T is the number of non-zero coefficients in $\hat{\mathbf{x}}_i$. We consider all overlapping patches, and wrap the patches at the edges. Thus the number of patches is N . The ℓ_0 pseudo-norm penalizes the number of non-zero coefficients²⁵. We call Eq. 2 the *local model*, as it captures the smaller scale, localized features contained by patches. The dictionary \mathbf{D} is assumed unknown and is learned from the data during the inversion.

The global (Eq. 1) and local (Eq. 2) models are combined into a Bayesian *maximum a posteriori* (MAP) objective

$$\left\{ \hat{\mathbf{s}}_g, \hat{\mathbf{s}}_s, \hat{\mathbf{X}} \right\} = \arg \min_{\mathbf{s}_g, \mathbf{s}_s, \mathbf{X}} \left\{ \frac{1}{\sigma_\epsilon^2} \|\mathbf{t} - \mathbf{A}\mathbf{s}_g\|_2^2 + \frac{1}{\sigma_g^2} \|\mathbf{s}_g - \mathbf{s}_s\|_2^2 + \frac{1}{\sigma_{p,i}^2} \sum_i \|\mathbf{D}\mathbf{x}_i - \mathbf{R}_i \mathbf{s}_s\|_2^2 \right\} \\ \text{subject to } \|\mathbf{x}_i\|_0 = T \quad \forall i, \tag{3}$$

where $\hat{\mathbf{s}}_g$ is an estimate of the global slowness perturbation, σ_g^2 is the global slowness variance, $\hat{\mathbf{s}}_s$ is the estimate of the sparse slowness perturbation, $\sigma_{p,i}^2$ is the variance of the patch slowness, and $\hat{\mathbf{X}} \in \mathbb{R}^{Q \times I}$ is the coefficient estimates.

We find the MAP estimates $\{\hat{\mathbf{s}}_g, \hat{\mathbf{s}}_s, \hat{\mathbf{X}}\}$ using a block-coordinate minimization algorithm by decoupling the local and global models via substitution^{17,25}. The global objective is, from Eq. 3,

$$\hat{\mathbf{s}}_g = \arg \min_{\mathbf{s}_g} \|\mathbf{t} - \mathbf{A}\mathbf{s}_g\|_2^2 + \lambda_1 \|\mathbf{s}_g - \mathbf{s}_s\|_2^2, \tag{4}$$

where $\lambda_1 = (\sigma_\epsilon / \sigma_g)^2$ is a regularization parameter. The local objective is from Eq. 3, substituting $\mathbf{s}_s = \hat{\mathbf{s}}_g$,

$$\hat{\mathbf{x}}_i = \arg \min_{\mathbf{x}_i} \|\mathbf{D}\mathbf{x}_i - \mathbf{R}_i \hat{\mathbf{s}}_g\|_2^2 \text{ subject to } \|\mathbf{x}_i\|_0 = T. \tag{5}$$

Dictionary learning is added to the local problem (5), by optimizing \mathbf{D} :

$$\hat{\mathbf{D}} = \arg \min_{\mathbf{D}} \{ \min_{\mathbf{x}_i} \|\mathbf{D}\mathbf{x}_i - \mathbf{R}_i \hat{\mathbf{s}}_g\|_2^2 \text{ subject to } \|\mathbf{x}_i\|_0 = T \quad \forall i \}. \tag{6}$$

The dictionary learning problem (Eq. 6) is here solved using the iterative thresholding and signed k-means (ITKM) algorithm³⁰. After $\hat{\mathbf{D}}$ is obtained, the coefficients $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_I]$ are solved from Eq. 5 using orthogonal matching pursuit (OMP) with the same sparsity level T as ITKM. Then with $\hat{\mathbf{X}}, \hat{\mathbf{D}}$, and global slowness $\hat{\mathbf{s}}_g$ from Eq. 4 we solve for \mathbf{s}_s . Equation 3 gives, assuming constant patch variance $\sigma_{p,i}^2 = \sigma_p^2$,

$$\hat{\mathbf{s}}_s = \arg \min_{\mathbf{s}_s} \lambda_2 \|\hat{\mathbf{s}}_g - \mathbf{s}_s\|_2^2 + \sum_i \|\mathbf{D}\hat{\mathbf{x}}_i - \mathbf{R}_i \mathbf{s}_s\|_2^2, \tag{7}$$

where $\lambda_2 = (\sigma_p / \sigma_g)^2$ is a regularization parameter. The solution to Eq. 7 is analytic

$$\hat{\mathbf{s}}_s = \frac{\lambda_2 \hat{\mathbf{s}}_g + n \mathbf{s}_p}{\lambda_2 + n}, \tag{8}$$

where n is the number of patches and $\mathbf{s}_p = \frac{1}{n} \sum_i \mathbf{R}_i^T \mathbf{D} \hat{\mathbf{x}}_i$. Equation 8 gives \mathbf{s}_s as the weighted average of the patch slownesses $\{\mathbf{D}\hat{\mathbf{x}}_i \forall i\}$ and $\hat{\mathbf{s}}_g$. When $\lambda_2 \ll n$, $\mathbf{s}_s \approx \mathbf{s}_p$. When $\lambda_2 = n$, \mathbf{s}_g and \mathbf{s}_p have equal weight. It is typical in image denoising to set $\lambda_2 = 0$ ²⁶. The expressions Eq. 4–8 are solved iteratively until convergence. Before solving Eq. 5 and 6, the slowness patches $\{\mathbf{R}_i \hat{\mathbf{s}}_g \forall i\}$ are centered²⁶, i.e. the mean of the pixels in each patch is subtracted. The mean of patch i is $\bar{x}_i = \frac{1}{n} \mathbf{1}^T \mathbf{R}_i \hat{\mathbf{s}}_g$. Hence, $\mathbf{R}_i \hat{\mathbf{s}}_g \approx \mathbf{D}\mathbf{x}_i + 1x_i$. The tomographic image used for geophysical interpretation from the LST algorithm is $\hat{\mathbf{s}}_s$.

We note that the LST approach performs well for both prescribed (e.g. wavelet and DCT) and learned dictionaries (see Fig. 5). But in general we expect the tomographic image fidelity of the learned dictionary inversion to be greater than that of prescribed dictionaries. This follows the results of simulations in¹⁵ (e.g. Figures 5, 6, 8 and 9), which compares the performance of prescribed and wavelet dictionaries on synthetic slowness maps and shows learned dictionaries can reduce the RMSE relative to ground truth by greater than 50% over prescribed dictionaries. This also follows the “synthesis” (vs. “analysis”) paradigm in image processing, which prefers to learn

Method	BRISQUE	NIQE	PIQE
Conventional	51.2	10.1	88.4
Eikonal	43.7	11.0	86.8
LST	45.8	5.9	72.5

Table 1. Reference-less image quality score for Conventional, Eikonal, and LST methods with BRISQUE, NIQE, and PIQE metrics (lower score is better).

dictionaries if sufficient training data is available rather than use e.g. wavelet functions which must be justified by detailed theoretical analysis for each application. Dictionaries can be learned from a corpus of data, or as was done here, from a large number of examples from one imaging scenario. This has proven a successful approach in e.g. image restoration¹⁶ and medical imaging¹⁷. For more details please see²⁵ (pp. 227–246).

Conventional tomography. We implement conventional tomography using a Bayesian approach³³, which regularizes the inversion with a global smoothing (non-diagonal) covariance. Considering the measurements (1), the MAP estimate of the slowness is

$$\hat{\mathbf{g}}_{\mathbf{g}} = (\mathbf{A}^T \mathbf{A} + \eta \Sigma_L^{-1})^{-1} \mathbf{A}^T \mathbf{t}, \quad (9)$$

where $\eta = (\sigma_c / \sigma_d)^2$ is a regularization parameter, with σ_c the conventional slowness variance, and

$$\Sigma_L(i, j) = \exp(-l_{ij}/L). \quad (10)$$

Here l_{ij} is the distance between cells i and j , and L is the smoothness length scale^{33,34}. For our simulation we choose $L = 10$ km and $\eta = 100$ km². This gives the best tradeoff between detail and fidelity.

LST, eikonal, and conventional performance comparisons. We quantify the relative quality of the LST and eikonal phase speed maps (Fig. 4) by variance reduction, and using visual quality scores derived from natural images. The LST provides a 57% variance reduction and conventional gives 59%, whereas eikonal gives a 28% reduction. The increased variance reduction in the LST and conventional inversions over the eikonal result is likely because the eikonal tomography method does not explicitly minimize the travel time residual. Both LST and conventional tomography minimize the travel time residual in their map estimation (Eq. 4 and 9). Since there is no true phase speed map available, we use reference-less image quality metrics to help quantify the quality of the LST and eikonal phase speeds. While such metrics may not reflect the truth of estimated geophysical features, they can help quantify corruption of the geophysical features. We use the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)³⁵, the Natural Image Quality Evaluator (NIQE)³⁶, and the Perception based Image Quality Evaluator (PIQE)³⁷. The results are summarized in Table 1. Overall, LST obtains a better score on 2 of 3 of the metrics. The BRISQUE metric has incorporated human opinions of image quality, whereas NIQE and PIQE do not. Hence BRISQUE may be less suited to our application.

Since surface wave phase travel times potentially suffer from phase ambiguities, we also performed an LST inversion of the 1 Hz Rayleigh wave group travel times using the same parameters as the LST phase speed map (Fig. 4A). The LST group speed map is shown in Fig. 6. The group travel times are not subject to phase ambiguities, but may present other problems such as erratic arrival picks. The trends in LST group speed map (Fig. 6) compare well with those of the LST phase speed map (Fig. 4A), and both show a large fast anomaly south of the NI fault. We note that the group speed is slower overall, than the phase speed map, which is expected.

We compare the results of the LST to eikonal tomography^{8,38} and conventional tomography (Fig. 4). Eikonal tomography avoids the inversion of large tomography matrices in favor of solving a number of simpler subproblems, while also partly accounting for wave refraction. For each virtual source, a phase speed map is estimated from the local gradients of the smoothed travel time surface. The individual phase speed estimates from all the virtual sources are then averaged to obtain the final phase speed map. Since LST uses the straight ray assumption, it is more similar to conventional than eikonal tomography.

However, LST provides improved results over conventional tomography with less computational burden. Conventional tomography³³ complexity is dominated by square matrix inversion $\mathcal{O}(N^3)$, though approximate solution methods are slightly less complex. For large tomography matrices \mathbf{A} , LST is dominated by matrix multiplication in the LSMR algorithm²⁹ $\mathcal{O}(2MN)$. Since for LST the sparse matrix \mathbf{A} is used directly, the memory required for a tomography problem with M travel times scales linearly with the map size N . For conventional tomography, the memory required scales by N^2 . Hence LST could be used for much larger maps than conventional tomography.

Geological interpretation. The LST sensitivity depth range (~100–500 m) with 1 Hz Rayleigh surface waves is occupied by Pleistocene and Holocene deposits³², which contain almost all the ground water resources in the area, Fig. 3A–C. The Silverado formation (Lower Pleistocene age, ~300–580 ka, named in³¹), accounts for nearly 90% of the total ground water extraction in the region considered here³⁹. The producing aquifers consist of sand and gravel, and in particular the Silverado unit is characterized by coarse-grained sediments.

Ponti *et al.*³² generated a 3D sequence stratigraphy model of Quaternary layers at the Dominguez Gap in a 16.5×16.1 km area that overlaps with our LST model region (see Fig. 2). The study was aided by 5 reference boreholes drilled to more than 450 m depth. In addition, more than 300 oil and water wells were compiled and used in the study. An important finding from this study was a fault, striking W-NW in agreement with the general trend

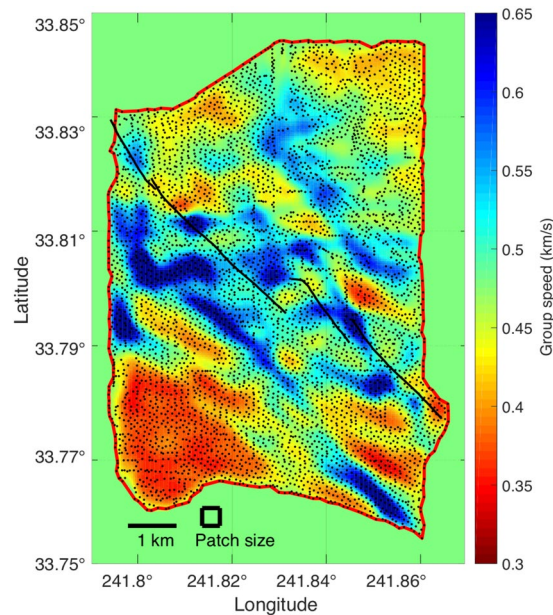


Figure 6. 1 Hz Rayleigh surface wave group speed map from LST, estimated using the same inversion parameters as the LST phase speed map (Fig. 4A). The group speed map shares the same speed trends as the LST phase speed map, including the large fast anomaly south of the NI fault corresponding to the Silverado aquifer. The stations (dots) and the valid boundary for LST (red line) are shown.

of the area, named the Pacific Coast Highway (PCH) fault, with a progressive vertical throw (up to 200 m) causing displacement of all pre-Holocene formations down to the north (see Fig. 2).

An area-wide gravity survey³² was carried out to invert for stratigraphy along a N-S profile, ~1 km west of the LST result region (see Fig. 2). The Silverado aquifer has significantly higher density (2290 kg/m³), due to its coarse-grained facies, than the surrounding formations (2050–2100 kg/m³). The Silverado gravity anomaly is in general agreement with the sequence-stratigraphic model, including the termination of the Silverado unit just south of LBCH. However, the Poland *et al.*³¹ survey contradicts the Ponti *et al.*³² inferred stratigraphy. In³¹, it is concluded that the Silverado is missing about 1 km south of the NI fault - near the termination of the LST phase speed anomaly (see Fig. 3).

Several established empirical relations positively correlate seismic wave speeds with density. Gardner *et al.*⁴⁰ relates density ρ to P-wave speed V_p as

$$\rho = 0.31 V_p^{0.25}, \quad (11)$$

where V_p is in m/s. Equation 11 gives a 40–50% increase in V_p for the Silverado formation, over the surrounding Pliocene, Pleistocene and Holocene layers. Brocher⁴¹ related V_p , V_s by

$$V_s = 0.7858 - 1.2344 V_p + 0.7949 V_p^2 - 0.1238 V_p^3 + 0.0064 V_p^4, \quad (12)$$

where V_s is in km/s. Equation 12 gives a 150% increase in V_s using derived values of V_p ⁴⁰ for the densities of the Silverado and surrounding layers. Since Rayleigh surface wave phase speed is dependent primarily on V_s ⁴², these results suggest that the Silverado aquifer should give rise to a significant phase-speed anomaly.

To help verify the fast anomaly in the LST result (Fig. 1B), we calculate the 1 Hz Rayleigh surface wave phase speed for the survey region based on predicted Silverado V_s from empirical relations Eq. 11 and 12. The method and results are summarized in Fig. 7. The Silverado elevation and thickness inferred in³² are interpolated (Fig. 7A,B) to obtain the depth ranges of the aquifer within the survey region. It is assumed per³¹ that Silverado is missing south of the fast anomaly. Further, we assume an average V_s profile for the region based on Rayleigh wave dispersion measurements in⁸. For each discrete Silverado depth range, the Silverado V_s is simulated by doubling the average V_s in the Silverado depth range (e.g. Fig. 7D). The 1 Hz Rayleigh surface wave phase speed for the survey region (Fig. 7C) is estimated using numerical forward modeling⁴³.

We also consider the effect of fluid saturation of the local strata in qualifying the high-velocity anomaly in our 1 Hz phase map. The depth to the water table in Long Beach (average from 150 wells⁴⁴) is about 7 m. Below the water table, the layers may be considered ‘wet’ to the largest depths in the deep water wells from³² (Pliocene age, 450 m+), though the layers are not necessarily ‘water-bearing’ per ground water extraction terminology. While laboratory experiments have shown for large V_s that the presence of fluids tends to decrease V_s of rock, this trend is insignificant for $V_s \approx 1,000$ m/s⁴⁵. The phase speeds observed in the high-velocity anomaly detected in our study are within this range. Thus, we conclude it unlikely that the high-velocity anomaly in our 1 Hz phase map is caused by a contrast in wet versus dry rock.

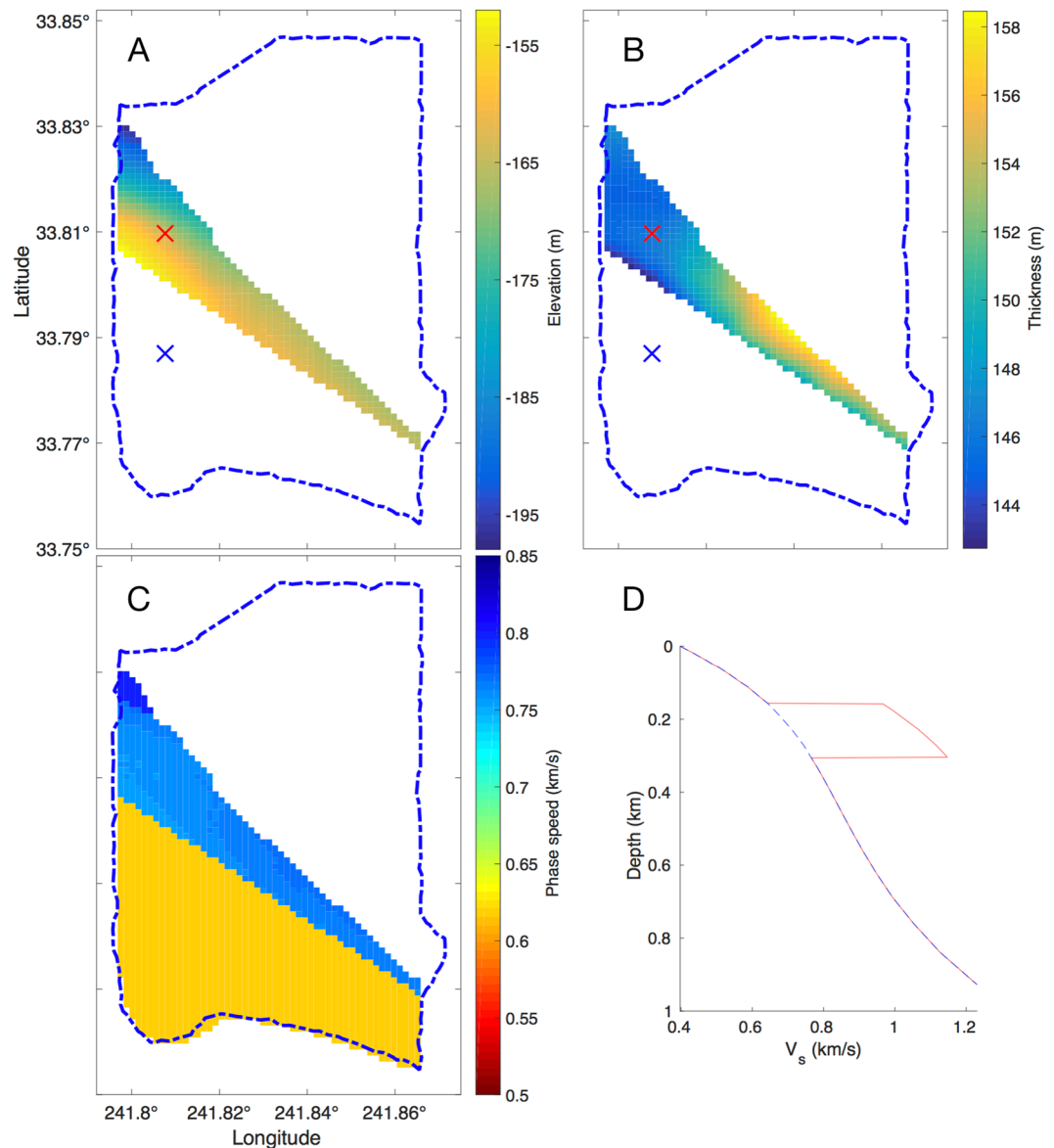


Figure 7. Silverado depth range and phase speed predictions using Ponti *et al.* (2007) and Poland *et al.* (1956) surveys. (A) Silverado elevation and (B) thickness from the Ponti *et al.*³² survey. Silverado is missing south of LST high speed anomaly per Poland *et al.*³¹ survey (see Fig. 3). (C) Phase speed predicted from (A,B) with Silverado phase speed perturbation (150% of average V_s from⁸ from Eq. 11 and 12). (D) Two V_s profiles: estimated V_s with Silverado layer (red line, from location of red 'x' in (A,B)) and V_s without Silverado (mean V_s from⁸ (blue dashed line, from location of blue 'x' in (A,B))). The valid boundary for LST inversion shown as blue dashed line.

Other parameters, such as confining pressure and anisotropy⁴⁶, as well as temperature⁴⁷, may affect seismic velocities of the individual strata. However, we do not have measurements of these parameters available for the layers of our model area. Thus, to the best of our knowledge based on recent literature, the high-velocity anomaly in the west-central part of the Long Beach model is caused by the higher-density gravel/coarse sand of the lower Pleistocene Silverado aquifer, as supported by gravity inversion and borehole logs³².

Received: 13 December 2018; Accepted: 6 September 2019;

Published online: 18 October 2019

References

1. Longuet-Higgins, M. S. A theory of the origin of microseisms. *Phil. Trans. R. Soc. Lond. A* **243**, 1–35 (1950).
2. Kedar, S. *et al.* The origin of deep ocean microseisms in the North Atlantic Ocean In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **464**, 777–793 (2008).
3. Lobkis, O. I. & Weaver, R. L. On the emergence of the Green's function in the correlations of a diffuse field. *J. Acoust. Soc. Am.* **110**, 3011–3017 (2001).

4. Shapiro, N. M., Campillo, M., Stehly, L. & Ritzwoller, M. H. High-resolution surface wave tomography from ambient seismic noise. *Science* **307**, 1615–1618 (2005).
5. Sabra, K. G., Gerstoft, P., Roux, P. & Kuperman, W. A. Extracting time-domain Green's function estimates from ambient seismic noise. *Geophys. Res. Letters* **32**, L14311 (2005).
6. Bensen, G. D. *et al.* Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geo. J. Int.* **169**, 1239–1260 (2007).
7. Brenguier, F. *et al.* Postseismic relaxation along the San Andreas fault at Parkfield from continuous seismological observations. *Science* **321**, 1478–1481 (2008).
8. Lin, F. C., Li, D., Clayton, R. W. & Hollis, D. High-resolution 3D shallow crustal structure in Long Beach, California: Application of ambient noise tomography on a dense seismic array. *Geophysics* **78**, Q45–Q56 (2013).
9. Schmandt, B. & Clayton, R. W. Analysis of teleseismic P waves with a 5200-station array in Long Beach, California: Evidence for an abrupt boundary to Inner Borderland rifting. *J. Geo. Res. Solid Earth* **118**, 5320–5338 (2013).
10. Nakata, N., Chang, J. P. & Lawrence, J. F. & Boue', P. Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry. *J. Geo. Res. Solid Earth* **120**, 1159–1173 (2015).
11. Yoon, C. E., O'Reilly, O., Bergen, K. J. & Beroza, G. C. Earthquake detection through computationally efficient similarity search. *Sci. Advances* **1**, e1501057 (2015).
12. Riahi, N. & Gerstoft, P. The seismic traffic footprint: Tracking trains, aircraft, and cars seismically. *Geophys. Res. Letters* **42**, 2674–2681 (2015).
13. Bowden, D. C., Tsai, V. C. & Lin, F. C. Site amplification, attenuation, and scattering from noise correlation amplitudes across a dense array in Long Beach, CA. *Geo. Res. Lett.* **42**, 1360–1367 (2015).
14. Inbal, A., Ampuero, J. P. & Clayton, R. W. Localized seismic deformation in the upper mantle revealed by dense seismic arrays. *Science* **354**, 88–92 (2016).
15. Bianco, M. J. & Gerstoft, P. Travel time tomography with adaptive dictionaries. *IEEE Trans. Comput. Imag.* (2018).
16. Elad, M. & Aharon, M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.* **15**, 3736–3745 (2006).
17. Ravishanker, S. & Bresler, Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans. Med. Imag.* **30**, 1028–1041 (2011).
18. Kong, Q. *et al.* Machine learning in seismology: turning data into insights. *Seismo. Res. Lett.* **90**, 3–14 (2018).
19. Bergen, K. J., Johnson, P. A., de Hoop, M. V. & Beroza, G. C. Machine learning for data-driven discovery in solid Earth geoscience. *Science* **363**, eaau0323 (2019).
20. Perol, T., Gharbi, M. & Denolle, M. Convolutional neural network for earthquake detection and location. *Sci. Advances* **4**, e1700578 (2018).
21. Ross, Z. E., Meier, M.-A. & Hauksson, E. P wave arrival picking and first-motion polarity determination with deep learning. *Journal of Geophysical Research: Solid Earth* **123**, 5120–5129 (2018).
22. Li, A., Peng, Z., Hollis, D., Zhu, L. & McClellan, J. H. High-resolution seismic event detection using local similarity for Large-N arrays. *Nature Scientific Reports* **8**, 1646 (2018).
23. Rouet-Leduc, B. *et al.* Machine learning predicts laboratory earthquakes. *Geophys. Res. Letters* **44**, 9276–9282 (2017).
24. Kong, Q., Allen, R. M., Schreier, L. & Kwon, Y. W. MyShake: A smartphone seismic network for earthquake early warning and beyond. *Sci. Advances* **2**, e1501055 (2016).
25. Elad, M. *Sparse and Redundant Representations* (Springer, New York, 2010).
26. Mairal, J., Bach, F. & Ponce, J. Sparse modeling for image and vision processing. *Found. Trends Comput. Graph. Vis.* **8**, 85–283 (2014).
27. Barmin, M. P., Ritzwoller, M. H. & Levshin, A. L. A fast and reliable method for surface wave tomography. *Pure and Appl. Geo.* **158**, 1351–1375 (2001).
28. Loris, I., Nolet, G., Daubechies, I. & Dahlen, F. A. Tomographic inversion using l_1 -norm regularization of wavelet coefficients. *Geo. J. Int.* **170**, 359–370 (2007).
29. Fong, D. C. L. & Saunders, M. LSMR: An iterative algorithm for sparse least-squares problems. *SIAM J. Sci. Comput.* **33**, 2950–2971 (2011).
30. Schnass, K. Local identification of overcomplete dictionaries. *J. Mach. Learn. Res.* 1211–1242 (2015).
31. Poland, J. F. *et al.* Ground water geology of the coastal zone, Long Beach, California. *U.S.G.S. Water Supply* (1956).
32. Ponti, D. J. *et al.* A 3-dimensional model of water-bearing sequences in the Dominguez Gap region, Long Beach, California. *US Geol. Surv. Open-File Rep.* <https://pubs.usgs.gov/of/2007/1013/> (2007).
33. Rodgers, C. D. *Inverse methods for atmospheric sounding: theory and practice* (World Sci. Pub. Co., 2000).
34. Tarantola, A. *Inverse problem theory* (Elsevier Sci. Pub. Co., Inc., 1987).
35. Mittal, A., Moorthy, A. K. & Bovik, A. C. No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.* **21**, 4695–4708 (2012).
36. Mittal, A., Soundararajan, R. & Bovik, A. C. Making a “completely blind” Image Quality Analyzer. *IEEE Signal Process. Lett.* **20**, 209–212 (2013).
37. Venkatanath, N., Praneeth, D., Bh, M. C., Channappayya, S. S. & Medasani, S. S. *Blind image quality evaluation using perception based features in IEEE Twenty-First Nat. Conf. Comm.*, 1–6 (2015).
38. Lin, F. C., Ritzwoller, M. & Snieder, R. Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array. *Geo. J. Int.* **177**, 1091–1110 (2009).
39. Thomas, S. D., Liles, J. M. & Johnson, T. A. Managing Seawater Intrusion in the Dominguez Gap Area of Los Angeles County, California, USA. *First Int. Conf. Saltwater Intrusion and Coastal Aquifers, Essaouira, Morocco*, 23–25 (2001).
40. Gardner, G. H. F., Gardner, L. W. & Drake, C. L. Formation velocity and density – the diagnostic basics for stratigraphic traps. *Geophys.* **39**, 770–780 (1974).
41. Brocher, T. M. Empirical relations between elastic wavespeeds and density in the Earth's crust. *Bulletin Seis. Soc. Am.* **95**, 2081–2092 (2005).
42. Aki, K. & Richards, P. G. *Quantitative seismology* (University Science Books, Mill Valley, CA, USA, 2009).
43. Herrmann, R. B. Computer programs in seismology: an evolving tool for instruction and research. *Seism. Res. Letters* **84**, 1081–1088 (2013).
44. California Water Boards. Depth to groundwater database, [https://www.waterboards.ca.gov/losangeles/water issues/programs/ust/groundwater database.html](https://www.waterboards.ca.gov/losangeles/water%20issues/programs/ust/groundwater%20database.html), Accessed: 2019–05–29.
45. Kassab, M. A. & Weller, A. Study on P-wave and S-wave velocity in dry and wet sand stones of Tushka region, Egypt. *Egyptian J. Petroleum* **24**, 1–11 (2015).
46. Tosaya, C. & Nur, A. Effects of diagenesis and clays on compressional velocities in rocks. *Geophys. Res. Lett.* **9**, 5–8 (1982).
47. Batzle, M. & Wang, Z. Seismic properties of pore fluids. *Geophys.* **57**, 1396–1408 (1992).
48. Google. Google Maps, <https://www.google.com/maps/@33.7927732,-118.593918,9.35z>, Accessed: 2018–09–30.

Acknowledgements

The authors gratefully acknowledge Dan Hollis at NodalSeismic LLC, and Signal Hill Petroleum, Inc., for permitting them to use the Long Beach data. This project is supported by the Office of Naval Research, Grant No. N00014-18-1-2118.

Author contributions

All authors (i.e. M.J.B., P.G., K.B.O. and F.-C.L.) conceived of the idea for this paper; all authors implemented the analysis; M.J.B., P.G. and K.B.O. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.J.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019