

Estimating temporal changes in seismic velocity using a Markov chain Monte Carlo approach

G. Taylor  and G. Hillers 

Institute of Seismology, University of Helsinki, Helsinki, Uusimaa 00560, Finland. E-mail: george.taylor@helsinki.fi

Accepted 2019 November 21. Received 2019 November 20; in original form 2019 August 23

SUMMARY

We present a new method for estimating time-series of relative seismic velocity changes (dv/v) within the Earth. Our approach is a Markov chain Monte Carlo (MCMC) technique that seeks to construct the full posterior probability distribution of the dv/v variations. Our method provides a robust, computationally efficient way to compute dv/v time-series that can incorporate information about measurement uncertainty, and any prior constraints that may be available. We demonstrate the method with a synthetic experiment, and then apply the MCMC algorithm to three data examples. In the first two examples we reproduce dv/v time-series associated with the response to the 2010 M 7.2 El Mayor-Cucapah earthquake at two sites in southern California, that have been studied in previous literature. In the San Jacinto fault zone environment we reproduce the dv/v signature of a deep creep slip sequence triggered by the El Mayor-Cucapah event, that is superimposed on a strong seasonal signal. At the Salton Sea Geothermal Field we corroborate the previously observed drop-and-recovery in seismic velocity caused by ground shaking related to the El Mayor-Cucapah event. In a third, new example we compute a month long velocity change time-series at hourly resolution at Piñon Flat, California. We observe a low amplitude variation in seismic velocity with a dominant frequency of 1 cycle per day, as well as a second transient signal with a frequency of 1.93 cycles per day. We attribute the 1-d periodicity in the dv/v variation to the combined effects of the diurnal tide and solar heating. The frequency of the signal at 1.93 cycles per day matches that of the lunar (semi-diurnal) tide. Analysis of the uncertainties in the Piñon Flat time-series shows that the error contains a signal with a frequency of 1 cycle per day. We attribute this variation to seismic noise produced by freight trains operating within the Coachella Valley. By demonstrating the applicability of the MCMC method in these examples, we show that it is well suited to tackle problems involving large data volumes that are typically associated with modern seismic experiments.

Key words: North America; Inverse theory; Computational seismology; Seismic interferometry; Seismicity and tectonics; Seismic noise.

1 INTRODUCTION

Seismic velocity is a key physical property of the Earth, and an important observable quantity that can provide information about the subsurface. These conditions, and therefore seismic velocity, can be altered by a variety of natural internal processes linked to tectonic stresses, volcanism and underground engineering operations associated with resource production. Changes in seismic velocity within the Earth's crust can thus be attributed to many sources, including earthquakes and slip events (e.g. Poupinet *et al.* 1984; Wegler & Sens-Schönfelder 2007; Brenguier *et al.* 2008b; Nakata & Snieder 2011; Brenguier *et al.* 2014; Obermann *et al.* 2014). External forces, such as those originating from the dynamics of the Earth's atmosphere (Sens-Schönfelder & Wegler 2006; Hillers *et al.* 2015a), or

solid Earth tides (Hillers *et al.* 2015b; Mao *et al.* 2019a) can also affect rock properties.

The importance of detecting and measuring variations in seismic velocity derives from the ability to resolve and study the underlying deformation within a medium, which is linked to the state of stress (Niu *et al.* 2008). Hence, accurately measuring temporal variations in seismic velocity are essential to study the properties of a material, and has found numerous applications in the passive monitoring of fault zones (e.g. Wegler & Sens-Schönfelder 2007; Brenguier *et al.* 2008a; Hillers *et al.* 2019; Qiu *et al.* 2019), volcanoes (e.g. Brenguier *et al.* 2008b; Rivet *et al.* 2014; Sens-Schönfelder *et al.* 2014), landslides (Voisin *et al.* 2016), in geothermal stimulation contexts (Obermann *et al.* 2015) and also within engineering structures such as levees (Planès *et al.* 2017) and dams (Olivier *et al.* 2017).

While it is possible to use repeating earthquakes to monitor internal changes within the Earth (Baisch & Bokelmann 2001; Rubinstein & Beroza 2004; Schaff & Beroza 2004), the most common way (Sens-Schönfelder & Wegler 2006; Wegler & Sens-Schönfelder 2007; Brenguier *et al.* 2008b) to calculate dv/v time-series is to compute ambient noise cross-correlation functions over a given time span at some temporal resolution, and then define a reference cross-correlation function, which is usually a stack over the analysis period. Each individual cross-correlation function is then compared to the reference to calculate the relative time delay between seismic arrivals contained in the coda part of the waveform, using methods such as moving time-window cross-spectral analysis (MWCS, Poupinet *et al.* 1984; Clarke *et al.* 2011), stretching (Lobkis & Weaver 2003), dynamic warping (Mikesell *et al.* 2015) and wavelet transforms (Mao *et al.* 2019b). This time delay is proportional to the apparent velocity change in the medium.

Robust estimates of the temporal variations in seismic velocity can provide constraints on material properties that are complementary to other monitoring techniques such as satellite geodesy and observations of seismicity and tremor. One key advantage of seismic monitoring using the ambient noise field is the temporal resolution of the observations that it provides. Satellite monitoring in particular is constrained by the need to wait several days between image acquisitions. Noise-based seismic monitoring studies at regional scales routinely achieve daily resolution for velocity change time-series, and more recent studies have shown that hourly resolution is possible at local scales (Mao *et al.* 2019a).

An alternative method to the reference stack for calculating the dv/v time-series between two seismic stations (Brenguier *et al.* 2014) involves calculating the velocity difference between every pair of cross-correlation functions in the data set, and then inverting these measurements for the best-fitting dv/v time-series. By avoiding the use of a single reference correlation function, this approach is not limited by the obtained dv/v time-series only being valid when compared to an arbitrary reference. It also avoids potential difficulties in defining a reference, which can occur in highly variable environments such as volcanoes (Sens-Schönfelder *et al.* 2014). In addition, by incorporating the full amount of available data, the method of Brenguier *et al.* (2014) provides more robust constraints on seismic velocity changes than the reference approach described above, as well as offering a way to deal with gaps in data sets. Due to the fact that this approach does not require some reference seismic velocity to be defined, we refer to it as the Reference Independent Velocity Variation Estimator (RIVV-E).

Brenguier *et al.* (2014) use a matrix-based least squares method to solve the RIVV-E inverse problem. The need to invert a matrix as part of this process limits the amount of data that can be included in the analysis, due to the fact that including N cross-correlation functions from a given time-span results in $N(N-1)/2$ correlation pairs. This unfavourable scaling of the problem with the amount of data can result in challenging computational memory requirements during the matrix inversion process. This restriction means that without high performance computing architecture the RIVV-E approach is only suitable for short time-series (~ 100 s samples), and when inverting a small number of station pairs simultaneously. Furthermore, Brenguier *et al.* (2014) utilize a model covariance matrix with an *ad hoc* smoothing parameter, along with a further subjective damping parameter, to stabilize the inversion.

In this study, we present an adaptation to the RIVV-E method that applies a Markov chain Monte Carlo (MCMC) approach to the determination of the dv/v time-series. The MCMC method does not require the inversion of a matrix, only the repeated solution of the

forward problem. By reducing matrix inversion to matrix multiplication, problems with a much greater volume of data become computationally tractable. This improved computational efficiency allows us to construct longer dv/v time-series, and invert data recorded at a large number of stations simultaneously. Inverting data from multiple station pairs simultaneously is advantageous to the more common approach of inverting the data pair-by-pair and averaging the result, as it provides more accurate dv/v time-series. In addition, the use of MCMC methods allow for the determination of the full posterior probability distribution of the dv/v time-series, complete with rigorous treatment of data uncertainties, rather than providing a single best-fitting model. As such, *ad hoc* smoothing and damping parameters are not required with the method we present here.

In this paper, we first outline our MCMC approach, and demonstrate its potential by applying it to a synthetic example. We then apply our method to real data, reproducing three examples from previous studies (Taira *et al.* 2018; Hillers *et al.* 2019; Mao *et al.* 2019b). Our first real data example is the observation of seismic velocity variations using data recorded at seismic stations located on the San Jacinto Fault Zone near Anza, southern California. At Anza, we confirm the observation of a strong seasonal variation in dv/v time-series constructed at daily resolution, as well as a ~ 100 -d-long transient signal associated with a deep creep episode triggered by the 2010 $M 7.2$ El Mayor-Cucapah earthquake (Inbal *et al.* 2017; Hillers *et al.* 2019). In our second example, we focus on the transient dv/v signal produced by shaking directly from the El Mayor-Cucapah earthquake. To achieve this, we reproduce the results of Taira *et al.* (2018) and Mao *et al.* (2019b), who observed a reduction, and subsequent recovery, in seismic velocity at the south eastern end of the Salton Sea immediately following the El Mayor-Cucapah event. Finally, to demonstrate the capacity of the MCMC method, we perform a new experiment to calculate a month long dv/v time-series at hourly resolution from data recorded during a seismically quiescent period at a 13 station array located at Piñon Flat, California. We find that the changes in seismic velocity and its associated uncertainty at Piñon Flat display a strong correlation with tidal strain signals at periods of 1.0 and 1.93 cycles per day (Agnew 1981, 2015).

2 METHODOLOGY

2.1 Reference independent velocity variation estimator (RIVV-E)

Given a suite of ambient noise cross-correlation functions between two stations that span a given time period at some temporal resolution, the RIVV-E procedure for estimating dv/v time-series presented by Brenguier *et al.* (2014) involves the calculation of the seismic velocity change between each pair of cross-correlation functions. The seismic velocity variation between cross-correlation functions at times i and j is denoted as δv_{ij} . We can calculate δv_{ij} from

$$\delta v_{ij} = \frac{v_j - v_i}{v_i} = \text{MWCS}(ccf_i, ccf_j), \quad (1)$$

where $\text{MWCS}(ccf_i, ccf_j)$ is the output of a moving time-window cross-spectral analysis (Poupinet *et al.* 1984) performed on cross-correlation functions i and j . The MWCS procedure may be freely substituted for some other method of estimating the velocity change, such as stretching (Lobkis & Weaver 2003), dynamic warping (Mikesell *et al.* 2015) or wavelet transform (Mao *et al.* 2019b).

Together, the δv_{ij} measurements define a data vector

$$\mathbf{d} = \begin{bmatrix} \delta v_{12} \\ \delta v_{13} \\ \delta v_{14} \\ \vdots \\ \delta v_{n-1n} \end{bmatrix} \quad (2)$$

which is of length $N(N-1)/2$, where N is the number of correlation functions. We aim to invert \mathbf{d} to retrieve a dv/v time-series of length N defined as

$$\mathbf{m} = \begin{bmatrix} \delta v_1 \\ \delta v_2 \\ \delta v_3 \\ \vdots \\ \delta v_n \end{bmatrix}, \quad (3)$$

where δv_n is the velocity change at time n . If the seismic velocity changes are small (<0.1 per cent), Brenguier *et al.* (2014) show that \mathbf{d} and \mathbf{m} can be related through

$$\begin{aligned} \delta v_j - \delta v_i &= \frac{v_j - v_i}{v_{ref}} = \frac{v_j - v_i}{v_i} \frac{v_i}{v_{ref}} \\ &= \delta v_{ij} \frac{v_i}{v_{ref}} = \delta v_{ij}(1 + \delta v_i) \end{aligned} \quad (4)$$

and the problem can be described in the form $\mathbf{d} = \mathbf{G}\mathbf{m}$, where \mathbf{G} is a sparse matrix of shape $\left[\frac{N(N-1)}{2}, N\right]$ and with the form

$$\mathbf{G} = \begin{bmatrix} -1 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ -1 & 0 & 1 & 0 & \cdots & \cdots & \cdots & \vdots \\ -1 & 0 & 0 & 1 & 0 & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & -1 & 1 \end{bmatrix}. \quad (5)$$

The problem can be solved by any convenient inversion method. Previous studies (Brenguier *et al.* 2014; Gómez-García *et al.* 2018) have opted for a matrix-based least-squares approach by solving

$$\mathbf{m} = (\mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{G} + \alpha \mathbf{C}_m^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{d}, \quad (6)$$

where \mathbf{C}_d^{-1} and \mathbf{C}_m^{-1} are the data and model covariance matrices, and α is a damping parameter. To incorporate more data in a single inversion, such as multiple station pairs or extra components of motion, the required rows can simply be appended to the \mathbf{d} , \mathbf{G} and \mathbf{C}_d^{-1} matrices. Whilst the matrix approach is more robust with respect to waveform variations in the cross-correlation functions when compared with the reference approach (Brenguier *et al.* 2014), it has several disadvantages. First, subjective choices must be made for the entries in \mathbf{C}_m^{-1} , which governs the smoothing that will be applied to the retrieved model. An additional choice must be made for the value of α , which weights the importance of modelization uncertainties versus data uncertainties. Finally, the inversion of the matrix $(\mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{G} + \alpha \mathbf{C}_m^{-1} \mathbf{G})$ (eq. 6) is computationally expensive, and may be intractable when \mathbf{G} is very large. We address these disadvantages by implementing a Bayesian MCMC approach.

2.2 Bayesian MCMC approach

Bayesian inversions aim to determine model parameters by combining previously available information on the model space (the prior probability) with additional constraints provided by the observed

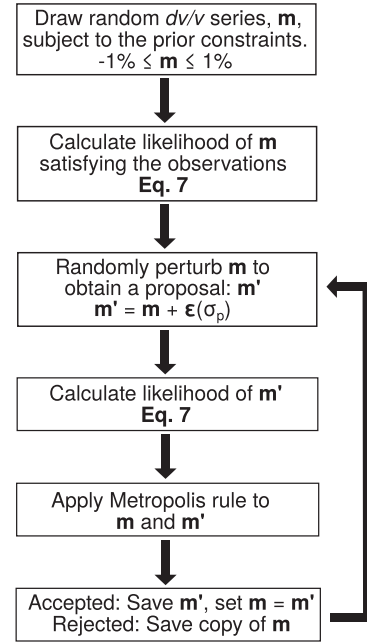


Figure 1. Schematic outline of the Markov Chain Monte Carlo algorithm presented in this study. Details of each step can be found in Section 2.2.

data (Tarantola 2005). The aim is to construct the full posterior probability distribution of every model parameter, defined here as the velocity state δv_n at a given datum n (eq. 3), incorporating knowledge of uncertainties on either the data or modelization. The posterior distribution can be constructed using a memory-less random walk through the parameter space using the Hastings–Metropolis algorithm (Fig. 1). MCMC methods are now commonly applied to many problems in seismology, including seismic tomography (Bodin *et al.* 2012; Galetti *et al.* 2017), earthquake slip inversions (Minson *et al.* 2013; Dettmer *et al.* 2014; Amey *et al.* 2018, 2019), the modelling of post-seismic deformation (Ingleby & Wright 2017) and ambient noise data processing (Chaput *et al.* 2016).

The problem of calculating seismic velocity change time-series (Section 2.1) can be solved through the application of MCMC methods. To start a single Markov chain, we begin by generating a random model, \mathbf{m} , (eq. 3) that is drawn from some prior distribution, representing our previous knowledge of the system. It is possible to start multiple simultaneous chains in this manner, and then combine the resultant posterior distributions. As each Markov chain is independent from the others, and no communication is required between chains, the process can be considered ‘embarrassingly parallel’ (Herlihy & Shavit 2012). This is an important advantage of the MCMC approach, as more parallel Markov chains can be used to reduce the required computation time as the volume of input data increases. There is usually little prior information available to constrain dv/v time-series, so in our examples we choose a prior of uniform probability with δv_n being bounded between -1 and 1 per cent. These bounds are selected as variations in seismic velocity rarely exceed 1 per cent outside of earthquake strong motions in the shallow crust (e.g. Nakata & Snieder 2011; Takagi *et al.* 2012; Viens *et al.* 2018).

After drawing a random starting model, subject to the prior probability, we calculate the likelihood of this model explaining the observed data, \mathbf{d} (eq. 2). This is subject to a given likelihood function. In this case, we choose an L_2 measure of misfit as the likelihood, which is equivalent to a Gaussian probability distribution (e.g. Aster

et al. 2013):

$$p(\mathbf{d}|\mathbf{m}) = e^{-\frac{1}{2}(\mathbf{d}-\mathbf{Gm})^T \mathbf{C}_d^{-1}(\mathbf{d}-\mathbf{Gm})}, \quad (7)$$

where \mathbf{d} , \mathbf{m} and \mathbf{G} correspond to eqs (2), (3) and (5), respectively. The data uncertainties are contained in \mathbf{C}_d^{-1} , which is a sparse diagonal matrix of dimensions $\left[\frac{N(N-1)}{2}, \frac{N(N-1)}{2}\right]$. The diagonal entries of \mathbf{C}_d^{-1} are $1/\sigma^2$, where σ^2 is the variance of the corresponding data point in \mathbf{d} (eq. 2). Other representations of likelihood are also valid. For example, an L_1 measure of misfit (Laplacian distribution) may be desirable due to its robustness against outliers. The Laplacian distribution is normally unsuitable for inverse problems due to the fact that the function is non-differentiable when the residual is zero, however, the fact that MCMC methods do not require differentiation of the likelihood function permits its use. All of the examples presented in this study use the L_2 norm. When calculating the likelihood of the model, the prior constraints are also taken into account. For uniform prior probability, as considered here, the model is rejected if any dv/v value in the time-series has a value outside the permitted range.

After calculating the likelihood of the current model, the Markov chain needs to take a random step through the model space. This is achieved by randomly perturbing the current model through

$$\mathbf{m}' = \mathbf{m} + \epsilon(\sigma_p), \quad (8)$$

where the perturbation vector, $\epsilon(\sigma_p)$, is drawn from a Gaussian distribution with zero mean and a standard deviation of σ_p . As the RIVV-E approach does not require a reference velocity state, it is only the relative velocity variations between each cross-correlation function that is resolved. Due to this, at each iteration we subtract the mean of the proposed model. If the mean is not removed, the Markov chain will never converge to a region of the model space with high likelihood, as any number of dv/v time-series with the same relative velocity variations are equally likely. We then calculate the likelihood of \mathbf{m}' using eq. (7). The acceptance of \mathbf{m}' into the posterior distribution is subject to the Metropolis rule:

- (i) If $p(\mathbf{d}|\mathbf{m}') \geq p(\mathbf{d}|\mathbf{m})$, accept \mathbf{m}' .
- (ii) If $p(\mathbf{d}|\mathbf{m}') < p(\mathbf{d}|\mathbf{m})$, accept \mathbf{m}' with a probability equal to $\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}$.
- (iii) Else, reject \mathbf{m}' and save an additional copy of \mathbf{m} .

In practice, in step (ii) \mathbf{m}' is accepted if the ratio $\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}$ is greater than a randomly generated number between 0 and 1. If not, \mathbf{m}' is rejected. This process is repeated until the posterior probability distribution of \mathbf{m} is sufficiently sampled, which is judged as the point when the posterior resembles a smooth Gaussian function.

To efficiently search the entire parameter space of \mathbf{m} , it is important to ensure that the MCMC algorithm favours the exploration of high likelihood models, but not to the exclusion of low likelihood areas of the model space. Roberts et al. (1997) show that to maintain efficiency, the acceptance rate of the proposed models should be kept equal to 23.4 per cent. The acceptance rate is controlled by the choice of model perturbation (eq. 8). A higher standard deviation in the perturbation distribution, σ_p , leads to larger jumps in the parameter space, and thus an increase in rejection likelihood. The reverse is true for smaller values of σ_p . It is therefore sensible to choose a value of σ_p that results in an acceptance rate close to 23.4 per cent.

It is likely that the model acceptance rate will vary widely throughout the progress of the Markov chain. This is particularly true if the initial model is of low likelihood. Under these conditions, the Markov chain will undergo a period of 'burn-in', in which

it converges to an area of the model space with higher likelihood (Hastings 1970; Mosegaard 1998). During the burn-in period, the generated models are highly correlated with those of previous iterations, and are thus not representative draws from the posterior distribution. To remove the burn-in samples, we inspect the evolution of model likelihood versus iteration, and remove the portion of the chain that exhibits a rapidly increasing likelihood. An example of the burn-in period is shown in Figs 2(a) and (b). In Fig. 2(a) it is clear that during the first 5000 iterations of the MCMC algorithm, the likelihood of the generated models is rapidly increasing. Likewise, the dv/v themselves are converging towards values of higher likelihood (Fig. 2b). To ensure computational efficiency at all steps in the Markov chain, we periodically tune the width of the perturbation distribution (Amey et al. 2018). After every 100 iterations of the MCMC algorithm, we calculate the acceptance rate, r , of the previous 100 models. If the acceptance rate is not within 10 per cent of the ideal acceptance rate of 23.4 per cent, σ_p is either increased or decreased by a factor equal to the ratio $\frac{r}{0.234}$. This adjustment of the model perturbation ensures an efficient search of the parameter space at all points of the Markov chain.

This MCMC method allows us to construct the full posterior probability distribution of the seismic velocity change time-series. This approach has several advantages. First, we are able to quantitatively assess the uncertainties associated with the final model, taking into account the error associated with our observations. In addition, the problem does not require any *ad hoc* smoothing as applied by Brenguier et al. (2014) and Gómez-García et al. (2018). Instead, the value of each model parameter is resolved purely by the prior probability distribution and the constraints provided by the observed data. Finally, the MCMC approach requires only that we solve the forward problem, $\mathbf{d} = \mathbf{Gm}$, through matrix multiplication. This greatly reduces the computational load with respect to a matrix inversion approach, and allows us to utilize much greater volumes of data to obtain the seismic velocity change time-series.

3 RESULTS

3.1 Synthetic example

To demonstrate the MCMC approach, we first present an example involving the retrieval of a homogeneous seismic velocity change time-series from artificially stretched waveforms. To generate the synthetic waveforms we retrieve a single cross-correlation function from the data for use as a reference waveform. We then produce a set of synthetic dv/v values and apply these as a stretching parameter to the reference waveform to create a synthetic gather of cross-correlation functions. Finally, we apply the procedure outlined in Section 2.2 in an attempt to retrieve the original dv/v values from the synthetic cross-correlation gather. Our synthetic dv/v series is composed of 200 samples, which can be interpreted as days in a time-series. We incorporate the effect of three different signals. The first is a periodic signal, modelled as a sine wave with a maximum amplitude of 0.04 per cent and a period of 200 d, designed to represent the long-term seasonal variation observed in seismic velocities as a result of annual variations in temperature or rainfall (Sens-Schönfelder & Wegler 2006; Meier et al. 2010; Richter et al. 2014; Hillers et al. 2015a; Lecocq et al. 2017; Clements & Denolle 2018).

A second periodic signal is included, with a period of 12 d, to represent shorter term fluctuations in seismic velocity, which can be caused by changes in groundwater level (Sens-Schönfelder &

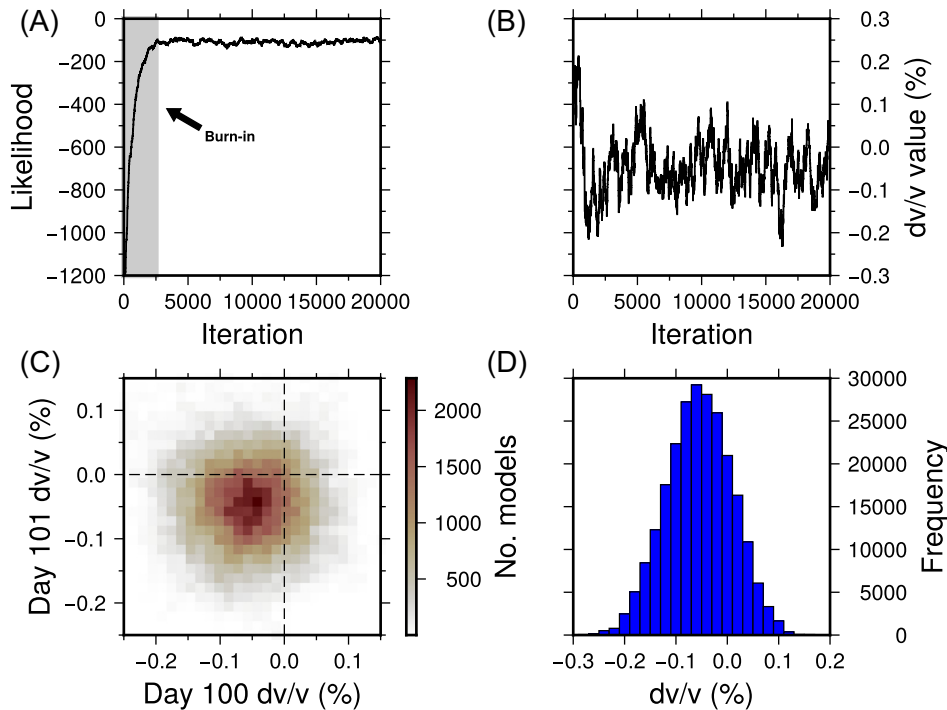


Figure 2. (a) Evolution of the model likelihood for the first 20 000 iterations of the MCMC algorithm during the noisy synthetic test. The burn-in period is clearly seen during the first ~ 2000 iterations, where the likelihood rapidly increases. (b) The evolution of the value taken by the dv/v value at day 100 for the first 20 000 iterations of the MCMC algorithm during the noisy synthetic test. (c) Joint probability distribution of dv/v at days 100 and 101 in the synthetic example (Section 3.1), following the MCMC inversion of the noisy observations. The burn-in period (a) has been removed. The colour scale indicates the frequency at which each dv/v value was drawn from the posterior with a given value. Red colours indicate a higher frequency, and therefore a higher probability of those dv/v values being drawn. (d) The marginal probability distribution of dv/v at day 100. The burn-in period (a) has been removed. Higher frequency indicates a higher probability of the dv/v parameter being drawn with that value.

Wegler 2006; Hillers *et al.* 2014; Clements & Denolle 2018). The maximum amplitude of the short period signal is 0.005 per cent.

The third signal is designed to mimic the response to an earthquake. It consists of the typically observed step-drop in dv/v amplitude (-0.1 per cent), located at day 100, followed by a logarithmic increase back to zero amplitude at the end of the time-series. The three signals are summed to produce the synthetic dv/v series, which is shown in Fig. 3. As outlined in Section 2.2, the retrieved dv/v time-series will have zero mean, so we also remove the mean of the target synthetic dv/v to facilitate direct comparison. The reference waveform is then stretched by each dv/v value in turn to produce the synthetic cross-correlation gather.

To simulate more realistic recording conditions, we add variable levels of random Gaussian noise to the cross-correlation gather. The noise is added directly to the waveforms, with a mean of zero and a standard deviation of 0.005. We calculate the dv/v value between each pair of synthetic waveforms using the MWCS method (Poupinet *et al.* 1984; Clarke *et al.* 2011).

We then apply the MCMC approach described in Section 2.2 to both noise-free and noisy cross-correlation gathers and attempt to retrieve our synthetic dv/v time-series. We run a single Markov chain for 250,000 iterations of the MCMC algorithm (Fig. 1), and remove the first 10,000 samples as the ‘burn-in’ period. The remaining models, that were accepted subject to the Metropolis rule (Section 2.2), represent the posterior distribution. The Markov chain is completed in ~ 30 s on a 3.0 GHz CPU. Fig. 2 shows joint and marginal probability distributions at days 100 and 101, which represent the peak of the earthquake signal in our synthetic time-series, taken following the inversion of the noisy observations. The marginal probability

distribution of the dv/v estimate at day 100 in Fig. 2(d) is clearly Gaussian. This is an encouraging result, as it greatly simplifies the extraction of information from the full posterior distribution. For example, if the posterior distributions of the dv/v values are Gaussian, then the maximum likelihood model is simply the mean of the posterior. Similarly, statistics such as standard deviation can be easily computed from the posterior, and can be used to infer the level of variability in the model that is permitted by the observations and any prior constraints.

Fig. 2(c) shows the joint probability distribution of dv/v values at days 100 and 101. Again, the distribution is clearly Gaussian in both dimensions. The ‘bulls eye’ shape of the joint distribution indicates that both parameters are independently resolved. If the joint distribution was instead shaped like an ellipse, this would indicate a degree of covariance between the two parameters, in which only some linear combination of the two parameters is resolved (Aster *et al.* 2013). The ability to independently resolve two model parameters that are adjacent to each other in the time-series demonstrates a clear advantage of the inversion method designed by Brenguier *et al.* (2014).

A comparison of the results from the synthetic test are shown in Fig. 3. For both the inverted models, we show the mean of the posterior probability distribution. In the noise-free case, all three components of the synthetic dv/v time-series are very well resolved, with only minor variations from the true values. In the presence of noise, the mean model of the posterior distribution still closely matches the target synthetic dv/v time-series, with an overall root-mean-square misfit of 0.014 per cent. The mean posterior time-series retrieves the earthquake signal along with the long period

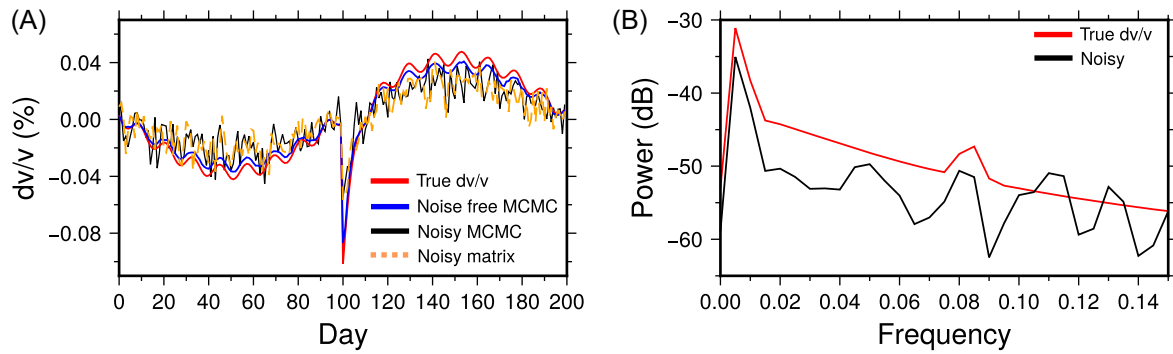


Figure 3. Recovery of a synthetic dv/v time-series using the MCMC method. (a) Comparison between the true synthetic dv/v time-series and those recovered by the MCMC inversion. The red line shows the original synthetic time-series. The blue line shows the mean model of the posterior distribution that is retrieved from synthetic stretched waveforms by the MCMC approach, when no random noise is added. The black line is the mean model of the posterior distribution that is retrieved by the MCMC approach when random Gaussian noise is added to the synthetic cross-correlation gather. The dashed orange line is the model produced by the matrix inversion approach of Brenguier *et al.* (2014). (b) The power spectra of the synthetic dv/v time-series compared with that recovered by the MCMC method after Gaussian noise is added to the observations.

variation, with only the recovery of the short-period variation being degraded. To aid in the comparison between the true model and the model retrieved by the MCMC approach, we compute the power spectra, which is shown in Fig. 3(b). From the power spectra it is obvious that the long period variation (frequency ~ 0.005) is present in the recovered signal. There is also a peak in the recovered spectrum at ~ 0.08 , which coincides with the short period (12-d) variation added to the original signal, but this peak is not prominent and lacks power when compared to peaks at other frequencies that are associated with the added noise. Whilst it is clear from the example in Fig. 3 that extracting very low dv/v amplitude variations may be difficult in the presence of noise, overall, the robust retrieval of the synthetic time-series demonstrates the potential of applying MCMC methods to estimating dv/v time-series from real data.

As the synthetic example in Fig. 3 contains only 200 model parameters, we also invert the noisy data set using the matrix-based inversion method of Brenguier *et al.* (2014), and compare it with the MCMC result. This result is also shown in Fig. 3, and matches almost exactly with the result obtained by the MCMC approach, with an identical misfit to the true model of 0.014 per cent. The close correspondance between the two techniques further proves the robustness of our MCMC algorithm. It is worth noting that in our synthetic test the smoothing term in eq. (6) ($\alpha C_m^{-1} \mathbf{G}$) can be neglected. For real data, the matrix-based inversion is more complex and will require some degree of smoothing.

3.2 2010 El Mayor-Cucapah earthquake

3.2.1 Benchmark 1 - Seasonal variation and response to El Mayor-Cucapah around Anza, California

The M 7.2 El Mayor-Cucapah (EMC) earthquake occurred on 4 April 2010 in Baja California, Mexico (Hauksson *et al.* 2011; Wei *et al.* 2011b). The EMC event triggered a response along many fault systems located in southern California, including short-lived shallow slip in the Salton Trough and Imperial Valley region induced by coseismic shaking (Wei *et al.* 2011a; Donnellan *et al.* 2014).

Hillers *et al.* (2019) focused on the effects of the EMC event in the San Jacinto fault zone environment near Anza, California, where a second M 5.4 earthquake occurred in Collins Valley on 7 July 2010. By applying the matrix-based approach of Brenguier *et al.* (2014), Hillers *et al.* (2019) identified a strong seasonal trend, in addition

to two transient signals in the dv/v time-series at Anza that were associated with deep creep on the San Jacinto fault following the EMC and Collins Valley events (Inbal *et al.* 2017).

We apply our MCMC approach to the dv/v data set compiled by Hillers *et al.* (2019) to directly compare this method with the matrix-based approach of Brenguier *et al.* (2014). For the year 2010, which contains the EMC and Collins Valley events, Hillers *et al.* (2019) calculated daily ambient noise cross-correlation functions for all nine combinations of directional components. To isolate the dv/v response to the deep creep events from that of the seasonal variation, Hillers *et al.* (2019) also calculated daily cross-correlation functions for the years 2008–2015, excluding 2010. To reduce the computational demand, the correlation functions calculated by Hillers *et al.* (2019) span a period of 300 d from the years 2008 to 2015 and utilized the ZZ component only. For the measurement of the seasonal variation only, these correlation functions were stacked in a 5-d-long moving window, and then down-sampled by a factor of 2. Seismic velocity changes between the correlation functions were calculated using the MWCS method (Poupinet *et al.* 1984; Clarke *et al.* 2011) on waves arriving at lag times 20–40 s, in the frequency band 0.2–2 Hz.

We use a subset of the data that consists of seismic stations located on the southwestern side of the San Jacinto fault zone near Anza, California (Fig. 4), where the seasonal trend and dv/v transients were most clearly observed (Hillers *et al.* 2019). We first calculate the average seasonal variation from the dv/v database for the years 2008–2015. The MCMC algorithm (Section 2.2) is applied to the dv/v measurements obtained by Hillers *et al.* (2019). For this experiment, a single chain is run for a total of 500,000 iterations. The ‘burn-in’ period of the chain is approximately 10,000 iterations, after which the MCMC algorithm begins to draw sample models directly from the target posterior probability distribution. The mean of the posterior distribution, which corresponds to the maximum likelihood model when uncertainties are normally distributed, is shown in Fig. 5. To quantify the uncertainty of the retrieved model, Fig. 5(a) also shows the standard deviation of each model parameter, as well as 95 per cent confidence intervals.

The shape of the dv/v time-series retrieved through the MCMC method is very similar to that found by Hillers *et al.* (2019). The only significant difference between the two results is the amplitude of the seasonal trend, particularly at the peak of the seasonal signal between days 200 and 240. This difference can be attributed to the fact that Hillers *et al.* (2019) compute the seasonal variation for

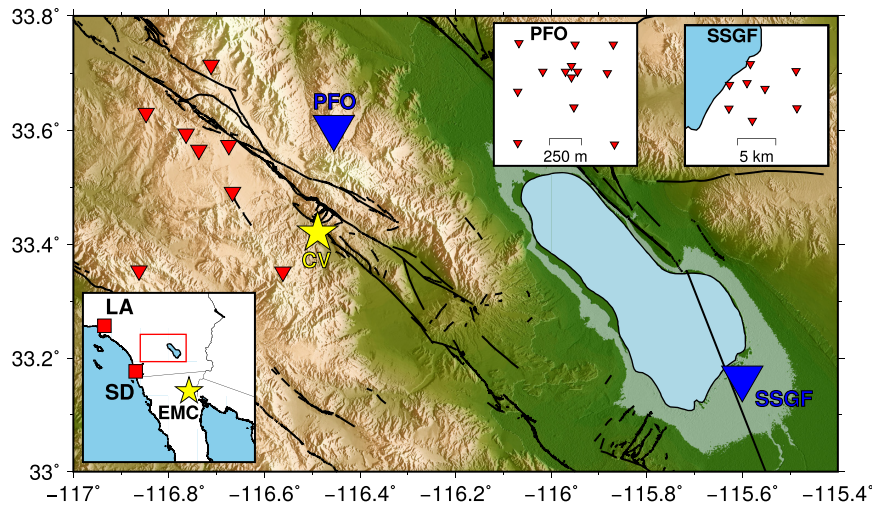


Figure 4. Overview map of the locations of the earthquake and seismic stations used in this study. The yellow stars mark the epicentres of the El Mayor-Cucapah (EMC) and Collins Valley (CV) earthquakes. The red triangles show the locations of the seismic stations used near Anza (Section 3.2.1). The blue triangles are the location of the dense arrays at the Salton Sea Geothermal Field (Section 3.2.2) and Piñon Flat (Section 3.3). The station locations for these arrays are shown in the inset boxes. The red squares indicate the major population centres of Los Angeles and San Diego. Topography data are from the GMRT synthesis project (Ryan *et al.* 2009).

each year individually, and then calculated the mean of these time-series. In our case, we include data from all years between 2008 and 2015 in a single MCMC inversion. Hillers *et al.* (2019) also apply a moving 3-point-average to smooth their time-series, which is not applied in our case. Fig. 5 also shows that the posterior standard deviation is higher during days 0–120, indicating higher uncertainty in the model. This time period covers the months January through April, corresponding to the northern hemisphere winter. Greater uncertainty during winter time reflects the greater variability in the ambient noise field during this time of the year.

We apply the same MCMC process to the dv/v data recorded in 2010. The time-series (Fig. 5b) spans 200 d following 13 February 2010, and includes the EMC and Collins Valley events. To isolate the transient signals related to both earthquakes, we subtract the average seasonal variation calculated from the years 2008 to 2015 from the 2010 data in the manner described by Hillers *et al.* (2019) (Fig. 5). The 2010 dv/v time-series displays a gradual decrease in seismic velocity from early March, followed by a gradual increase

after early May. The transient velocity reduction lasts for approximately 90 d, with the peak relative velocity reduction occurring ~ 30 d after the event. The velocity reduction associated with the Collins Valley event is smaller in amplitude, and begins immediately following the earthquake itself. The Collins Valley transient lasts for approximately 40 d before the velocity recovers to the background level. It is clear from Fig. 5(b) that there is a period of 30 d prior to the El Mayor-Cucapah event that exhibits intermittent, sharp increases in the level of uncertainty. There is also a period of significant uncertainty immediately following the El Mayor-Cucapah earthquake, that decays over a period of ~ 60 d. This decay pattern follows the Omori law, similar to the aftershock distribution following the El Mayor-Cucapah earthquake (Hillers *et al.* 2019). The sudden increases in uncertainty correspond to variations in the ambient noise wavefield caused by earthquakes. For example, the spike in uncertainty at day 122 (Fig. 5b) is a result of the $M 5.7$ Ocotillo earthquake, which occurred on the 14 June 2010 (Hauksson *et al.* 2011).

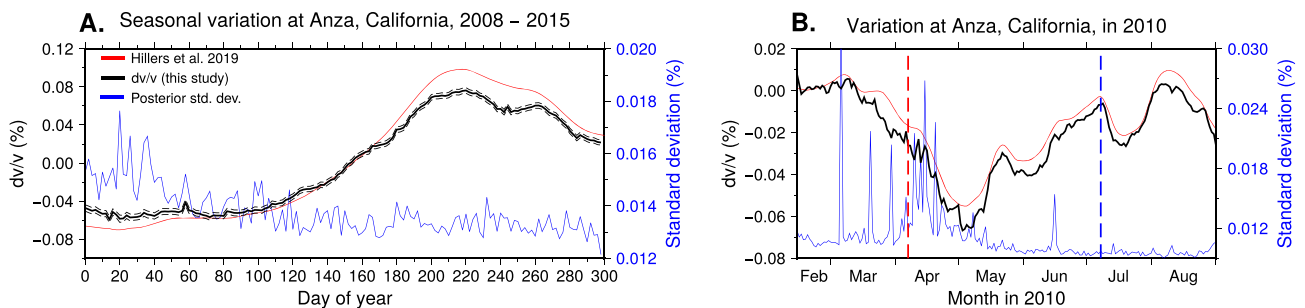


Figure 5. (a) Seasonal variation at Anza, California, between Julian day 1 and 300 for the years 2008–2015. The solid black line is the seasonal variation calculated using the MCMC approach (Section 2.2). The black dashed lines indicate the 95 per cent confidence interval of the dv/v time-series, and the blue line is the standard deviation of the dv/v values. The solid red line is the seasonal variation calculated using the matrix inversion approach of Brenguier *et al.* (2014) by Hillers *et al.* (2019). (b) The dv/v variation during 2010, with the average seasonal variation in (a) removed. Large peaks in the standard deviation (blue line) indicate poorly constrained data points, where observations may be missing, or of poor quality. The solid red line is the dv/v variation calculated by Hillers *et al.* (2019). The dashed red line indicates the occurrence of the 2010 El Mayor-Cucapah earthquake, and the dashed blue line indicates the Collins Valley earthquake.

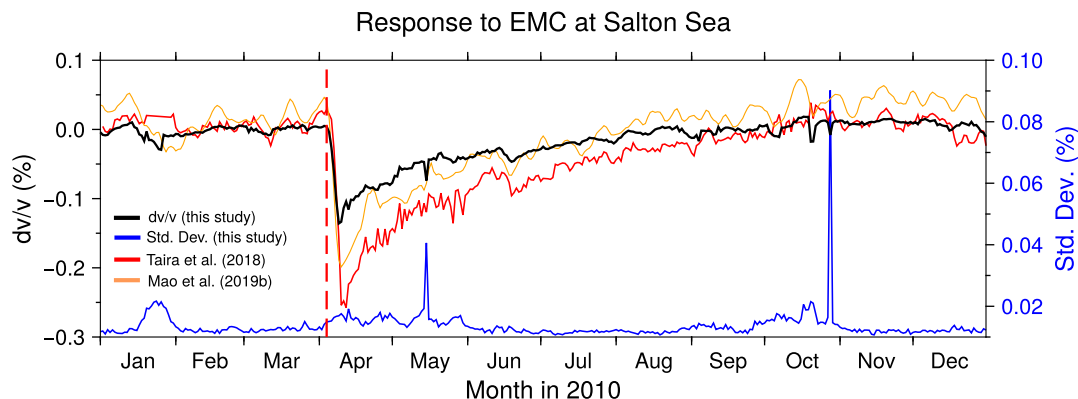


Figure 6. Response to the El Mayor-Cucapah earthquake at the Salton Sea Geothermal Field. The black line is the mean dv/v time-series, calculated from the posterior distribution generated by the MCMC algorithm. In order to aid comparison, we have subtracted the median dv/v value between 1 January and 4 April from the time-series. The dv/v values are indicated on the left axis. The blue line is the standard deviation of the dv/v measurements, the value of which is shown on the right axis. The red and orange lines are the results of Taira *et al.* (2018) and Mao *et al.* (2019b), respectively. The red dashed line marks the occurrence of the El Mayor-Cucapah earthquake.

Overall, the results that we have obtained with the MCMC method show a striking similarity to the results obtained by Hillers *et al.* (2019) obtained with the matrix inversion approach. This similarity provides assurance in the MCMC-derived results, but also promotes confidence in the dv/v time-series associated with the proposed deep creep events in the San Jacinto fault zone environment (Inbal *et al.* 2017; Hillers *et al.* 2019).

3.2.2 Benchmark 2 - Response to the El Mayor-Cucapah earthquake at the Salton Sea

To further test the performance of the MCMC approach we now target the reproduction of an archetypal dv/v drop-and-recovery signal associated with an earthquake response. We use a data set recorded at the Salton Sea Geothermal Field, located at the south eastern end of the Salton Sea (Fig. 4). These data have been used for monitoring seismic velocity changes following the El Mayor-Cucapah earthquake by Taira *et al.* (2018) and Mao *et al.* (2019b). In contrast to the Anza network, neither of these studies observe a strong seasonal variation in the seismic velocity at the Salton Sea site, making this data set ideal for observing the dv/v transient associated with the El Mayor-Cucapah event.

To study the response to El Mayor-Cucapah at the Salton Sea site, we utilize the database of daily ambient noise cross-correlation functions produced by Mao *et al.* (2019b). This data set is derived from vertical component data recorded at seven stations of the CalEnergy network between 15 December 2009 and 1 January 2011. To calculate the velocity change between each correlation function, we again use the MWCS technique. The velocity change is calculated from coda waves at lag times between 15 and 35 s, in the frequency range 0.5–3.5 Hz. Following the calculation of the dv/v measurements between each correlation function, we use the MCMC method to derive the network averaged dv/v time-series. A single Markov chain is run for 250,000 iterations, and the ‘burn-in’ period of 20,000 samples is removed from the posterior distribution. On a single 3.0 GHz core, this chain takes 62 min to complete. Fig. 6 shows the mean model of the posterior distribution of the seismic velocity change time-series for the Salton Sea data set constructed by the MCMC method. The corresponding standard deviation of the dv/v values is also shown. To ensure that we can compare our time-series to the results of Taira *et al.* (2018) and Mao *et al.* (2019b)

using the same pre-earthquake baseline, we subtract the median dv/v value between 1 January–4 April from each dv/v value. In the time period 1 January–4 April the relative seismic velocity is generally constant. After the El Mayor-Cucapah earthquake on 4 April, there is an immediate large drop in relative seismic velocity. Following this sudden decrease, the seismic velocity recovers in a logarithmic fashion for ~ 7 months, returning to the pre-earthquake values by 1 November 2010. Similar to the observations at Anza (Section 3.2.1), the standard deviation of the dv/v values also increases slightly following the earthquake. This period of increased uncertainty lasts for approximately 3 months, returning to the background value by June 2010. Unlike at Anza, the uncertainty does not decay according to the Omori law.

The amplitude of the velocity drop that we observe following the El-Mayor Cucapah is smaller than that observed by Taira *et al.* (2018) and Mao *et al.* (2019b). This is due to the fact that the dv/v results of both Taira *et al.* (2018) and Mao *et al.* (2019b) are defined relative to their choice of reference correlation function. This reference correlation function is a temporal average over a period of 6 yr in the case of Taira *et al.* (2018), 1 yr in Mao *et al.* (2019b). Differences in the time period used to define the reference can lead to systematic variations in the obtained dv/v time-series (Sens-Schönfelder *et al.* 2014, Fig. 6). In contrast, the RIVV-E MCMC approach maintains fidelity to the relative dv/v difference that we measure between the days immediately preceding and following the earthquake.

The onset of the response to El Mayor-Cucapah is much sharper at the Salton Sea Geothermal Field (Fig. 6), when compared with the response at Anza (Fig. 5). The decrease in seismic velocity begins immediately following the earthquake at the Salton Sea, and the minimum relative seismic velocity is detected within ~ 5 d of the event. In the vicinity of Anza, the velocity reduction appears to begin approximately 30 d before the earthquake, which may be the result of the increased uncertainty on this section of the time-series. The velocity also decreases at a much slower rate than at the Salton Sea, with the maximum relative reduction occurring a further 30 d after the El Mayor-Cucapah earthquake. The contrasting characteristics of the two responses at Anza and Salton Sea indicate that the mechanism for the velocity reduction is different in each case. The fact that the Salton Sea response is detected simultaneously with the earthquake suggests that damage to the rocks caused by passing seismic waves is likely responsible for the relative reduction in

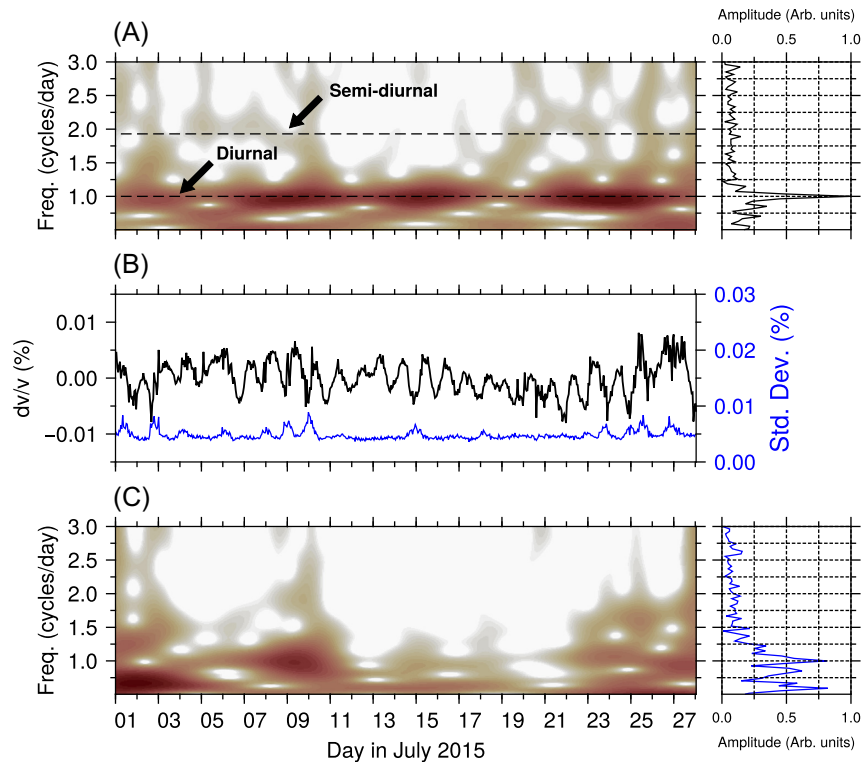


Figure 7. Variation in relative seismic velocity at the Piñon Flat array in July 2015. The raw velocity variations are shown by the black line in panel (b). The blue line in panel (b) is the standard deviation of the dv/v time-series. Panel (a) shows the continuous wavelet transform of the dv/v time-series (black line) between frequencies of 0.5–3 cycles per day. Darker colours indicate higher wavelet coefficients at that frequency and time. The black dashed lines represent the theoretical frequencies of the diurnal and semi-diurnal tidal strain. The amplitude spectrum is shown on the right, calculated from the Fourier transform of the signal. Panel (c) shows the same information as panel (a), but for the standard deviation time-series (blue line).

seismic velocity. In contrast, the more gradual response at Anza, suggest that the dv/v transient is not directly linked to shaking induced by the El Mayor-Cucapah event. Instead, the transient signal at Anza is probably the result of deep creep along the San Jacinto fault that was triggered by El Mayor-Cucapah (Inbal *et al.* 2017; Hillers *et al.* 2019).

3.3 New observations of seismic velocity variations at Piñon Flat related to tidal strain

In order to demonstrate the advantages in computational efficiency provided by the MCMC method, we present a new example that uses data from a seismic array at Piñon Flat, California (Fig. 4). The array at Piñon Flat is a dense network of 13 broad-band seismometers with an aperture of 1 km (Vernon 2014). We target the detection of daily and subdaily variations in seismic velocity by computing a one month dv/v time-series at hourly resolution for a seismically quiescent period in July 2015.

Using the vertical component only, we first split the data into 1-hr-long segments of ambient noise, and the instrument response was removed, with an initial bandpass filter applied between 0.1 and 15.0 Hz. The spectrum of each 1-hr-long trace was whitened in the interval 1.0–12.0 Hz, then one-bit normalized (Bensen *et al.* 2007), before a final bandpass filter was applied between 1.0 and 12.0 Hz. Each processed trace was then tapered and cross-correlated with the corresponding trace at every other station in the network. The cross-correlations are cut to time lags between -120 and 120 s, and we apply a singular value decomposition-based Wiener filter to the cross-correlation gather to improve the signal-to-noise ratio (Moreau

et al. 2017). During this process, we keep the first 15 singular vectors, and the Wiener filter widths are 5 hr by five samples. We calculate the velocity change between each pair of cross-correlation functions in the data set using the MWCS method. The velocity change is calculated in the frequency range 1.0–4.0 Hz, from the coda waves arriving between 10 and 30 s lag time, using a 2.0 s long MWCS window.

We then apply the MCMC method to infer the array-averaged dv/v time-series for Piñon Flat during this one month time period. The improvement in computational efficiency of the MCMC method compared to the approach of Brenguier *et al.* (2014) stems largely from the treatment of the \mathbf{G} matrix (eq. 6). Whilst \mathbf{G} typically contains millions of entries, only a small number of these entries are non-zero. Storing \mathbf{G} as a sparse matrix greatly reduces the memory requirements of matrix operations involving \mathbf{G} , including matrix multiplication.

This computational constraint means that the size of \mathbf{G} must be limited, usually by restricting the amount of data used in the inversion. Brenguier *et al.* (2014) and Gómez-García *et al.* (2018) achieve this by only inverting data from one station pair at a time. Methods that can solve eq. (6) without constructing the full inverse matrix do exist. These approaches include matrix factorization and iterative schemes such as the Conjugate Gradient method (Hestenes & Stiefel 1952). Iterative schemes are particularly useful for solving large linear problems, where the memory requirements of storing dense matrix factors can be prohibitive (e.g. Fox *et al.* 2014). However, the number of iterations required for convergence to a solution may rapidly increase with the dimensions of the linear problem. Typically, numerical pre-conditioners need to be designed and applied

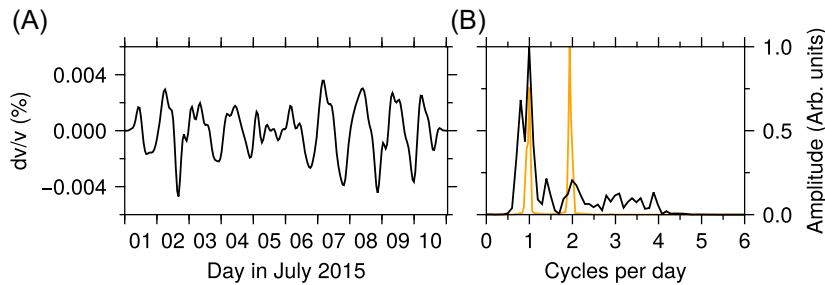


Figure 8. Tidal signal recovered from Piñon Flat between July 1 and July 10. (a) The dv/v time-series at Piñon Flat (Fig. 7) bandpass filtered between 0.7 and 4 cycles per day. (b) the black line shows amplitude spectrum of the bandpass filtered tidal response shown in (a). The orange line is the amplitude spectrum of the tidal strain predicted at the Piñon Flat Observatory in July 2015.

to an iterative scheme to ensure timely convergence (Wathen 2015). In MCMC methods, numerical pre-conditioners are not necessary to ensure the convergence of the Markov chain.

In this example at Piñon Flat, we simultaneously invert 672 hr of data recorded at 13 stations (78 station pairs). This data set results in a \mathbf{G} matrix of dimensions $4,389,840 \times 672$. If \mathbf{G} is not sparse, for example during standard matrix inversion algorithms, this matrix would require approximately 22 gigabytes of memory to store (at 8 bytes per entry). The memory requirement scales as $n \times \text{memory}$ when extra stations or components of motion are added, where n is the number of stations and components. As the length of the target time-series is increased the scaling is quadratic, with $N(N-1)/2 \times \text{memory}$, where N is the number of time samples. As a result, evaluating eq. (6) is not possible by standard matrix inversion. By operating only on the sparse form of \mathbf{G} , our MCMC approach is able to incorporate this volume of data in a single inversion.

We run the MCMC chain for 250,000 iterations, and discard the first 50,000 samples which are classified as the ‘burn-in’ period and do not represent representative models from the posterior distribution. Fig. 7 shows the results of the MCMC inversion for Piñon Flat. Fig. 7(a) shows the mean dv/v time-series of the posterior distribution, along with the corresponding standard deviation. In general, the velocity variations are very small (<0.01 per cent). The uncertainty on the velocity variation is lower than the amplitude of the signal, generally near 0.005 per cent.

The clearest signal in the dv/v time-series in Fig. 7(b) is the sinusoidal daily variation, which is dominant throughout the entire month. The power spectrum in Fig. 7(a) shows that this signal has a frequency of 1 cycle per day (24 hr). To examine the temporal variation of any signals, we performed a continuous wavelet transform with a Morlet wavelet on both the dv/v time-series and its standard deviation, between frequencies of 0.5–3 cycles per day (Fig. 7a). The diurnal variation in the dv/v values is clearly present throughout the observation period, although there are ‘pulses’ in the amplitude of the signal. There are several instances where the amplitude of the diurnal signal is lower, notably between 4–5, 11–12 and 18–19 July. All of these dates align with weekends, when seismic noise generated by anthropogenic activity is reduced. It is likely that the decreased level of coherent noise in our target frequency band (1–4 Hz) is responsible for the apparent reduction in amplitude of the diurnal signal.

The signal near 2 cycles per day is mainly present between 1 July and 10 July, but may also be present between 19 July and 28 July. Though the signal near 2 cycles per day shows significant temporal variation, analysis of the power spectrum (Fig. 8b) shows that the average period of this variation is just below 2 cycles per

day, which matches the average period of the lunar (semi-diurnal) tide at 1.93 cycles per day (Agnew 2012), subject to the frequency resolution of our time-series.

To confirm that the main source of the dv/v signals near 1 and 2 cycles per day is the Earth tides, we calculated the theoretical volumetric tidal strain at Piñon Flat Observatory at hourly resolution between 1 July and 28 July 2015 (Agnew 2012). We then compare the power spectra of the tidal strain to that of the dv/v time-series between 1 July and 10 July, where the semi-diurnal signal is strongest (Fig. 7a). We filter the dv/v time-series to the tidal frequency range between 0.7 and 4 cycles per day. Fig. 8(a) shows the bandpass filtered dv/v time-series, and the corresponding power spectrum is shown as the black line in Fig. 8(b). When compared with the power spectrum of the predicted tidal strain (orange line, Fig. 8b), the location of spectral peak in the dv/v time-series at 1 cycle per day match with the predicted tidal strain. There is also a peak the dv/v spectrum at 1.93 cycles per day corresponding to the semidiurnal tide, though this peak is much lower in amplitude, likely due to the temporal variability of this signal (Fig. 8a). Another explanation for this discrepancy is the effect of temperature-induced strain due to diurnal solar heating, which also has a period of 1 cycle per day (Tsai 2011). This interpretation is supported by Mao *et al.* (2019a), who also observed tidal signals in dv/v time-series recorded at Piton de la Fournaise volcano. Mao *et al.* (2019a) calculated that diurnal period thermal-induced strain typically has the same magnitude as the tidal-induced strain, which likely explains the greater prominence of the 1 cycle per day spectral peak in our dv/v data.

The power spectrum of the standard deviation contains little energy above a frequency of 1 cycle per day. It is clear from the time-series of the standard deviation in Fig. 7(b) that this 1-d periodicity is not consistently present throughout the month of July. Instead, this daily variation in uncertainty is most visible between July 1–10 and July 23–27. During these periods, the standard deviation usually peaks at around midnight local time. Seismic noise at frequencies >1 Hz are typically associated with anthropogenic activity (Ringdal & Bungum 1977; Peterson 1993). A study by Inbal *et al.* (2018) showed that the presence of freight trains in the Coachella Valley (<40 km from Piñon Flat), is a particularly strong source of noise the 1–5 Hz frequency range throughout southern California. These freight trains usually run during the night, with activity peaking around midnight (Inbal *et al.* 2018). This leads us to conclude that noise from human activity is a major contributor to uncertainty in measuring seismic velocity changes at hourly resolution, and that strong, directional noise generated by freight trains within the Coachella Valley is a likely source of uncertainty in this study.

4 DISCUSSION AND CONCLUSIONS

In this study, we have introduced a new method for calculating seismic velocity change time-series using a MCMC approach. We used four examples to prove the efficacy of the MCMC approach, and demonstrate the advantages provided by this new method. Our synthetic example (Fig. 3) showed that the MCMC method can fully recover a dv/v time-series that contains multiple signal types. This recovery remains robust when random noise is added to the observations, although very small (<0.01 per cent), short-period variations in seismic velocity may be difficult to detect in the presence of said noise.

We also applied the RIVV-E MCMC approach to three examples using real data, and compared our results to those obtained by previous authors that utilize the same data. There is a striking similarity between the seasonal variation at Anza between 2008 and 2015 derived by the MCMC approach and the matrix-based method of Hillers *et al.* (2019) (Fig. 5), with a velocity low (-0.04 per cent) from January to May, which increases to a peak of 0.08 per cent in July–September. Being able to reproduce the dv/v time-series using a completely different inversion scheme is an encouraging result for the MCMC method. Unlike the matrix-based approach, the MCMC method does not require any *ad hoc* smoothing via a model covariance matrix. Our results support the interpretation of Hillers *et al.* (2019), that the gradual reduction in dv/v at Anza in response to El Mayor-Cucapah is the result of triggered deep creep along the San Jacinto fault, rather than direct shaking caused by the earthquake. It is difficult to constrain the exact timing of the onset of the dv/v transient associated with the El Mayor-Cucapah earthquake, due to the larger uncertainty associated with this part of the time-series (Fig. 5b). Further analysis of the standard deviation of the posterior probability distribution provided two new interesting observations. Firstly, the uncertainty of the dv/v at Anza is increased during winter and spring (January–May). We attribute this added uncertainty to enhanced seismic noise that is often present in the wavefield during northern hemisphere winter (Peterson 1993; Stehly *et al.* 2006). We also observe higher dv/v uncertainty following the occurrence of the 2010 El Mayor-Cucapah earthquake that we attribute to the presence of earthquake aftershocks in the seismic wavefield, which contaminates the dv/v signal.

In Section 3.2.2, we applied the MCMC approach to data recorded at the Salton Sea Geothermal Field, and were able to confirm the observations of Taira *et al.* (2018) and Mao *et al.* (2019b), which were derived using a different methodology. At Salton Sea, we detect a large drop in relative seismic velocity (0.015 per cent) that coincides with the El Mayor-Cucapah earthquake. The amplitude of the velocity drop that we detect is smaller than that of previous studies (Taira *et al.* 2018; Mao *et al.* 2019b). This is due to the fact that the dv/v variations observed by Taira *et al.* (2018) and Mao *et al.* (2019b) are measured relative to a subjective reference velocity state, whereas the RIVV-E approach is sensitive to the dv/v variations measured between each day. We also observe that the seismic velocity recovers in a logarithmic fashion over a period of 7 months. In this case we support the conclusions of Taira *et al.* (2018) and Mao *et al.* (2019b), and attribute the velocity reduction directly to shaking caused by seismic waves from the El Mayor-Cucapah earthquake. As with the Anza data set (Section 3.2.1), there is also a period of increased uncertainty in the dv/v time-series at Salton Sea following the El Mayor-Cucapah event, that we attribute to the effect of earthquake aftershocks on the ambient noise field.

Finally, in Section 3.3 we demonstrated the computational effectiveness of the MCMC method by simultaneously inverting 672 hr

of dv/v measurements recorded in hourly resolution at the 13 station Piñon Flat array using the RIVV-E approach. By fully exploiting the fact that \mathbf{G} (eq. 5) can be stored in sparse matrix format throughout the MCMC process, we greatly reduce the computational memory requirements, and solve a problem that would usually be intractable by matrix inversion. As shown in Fig. 7, the strongest variation in dv/v that we observe at Piñon Flat has a period of 1 d. We also observe higher frequency variations in the Piñon Flat data set (2–3 cycles per day). Despite a low signal-to-noise ratio, by filtering and comparing our dv/v time-series with the theoretical tidal strain we show that the most likely cause for the signals near 1 and 2 cycles per day is periodic loading by solid Earth tide (Agnew 1981, 2012). We observe that our dv/v time-series contains more power at the diurnal period than the semi-diurnal period, which we attribute to the added effect of thermally induced strain caused by solar heating (Mao *et al.* 2019a), in addition to the temporal variability of the semi-diurnal tidal signal. Furthermore, we clearly observe that the dv/v standard deviation displays a periodicity of 1 d. The standard deviation of the dv/v time-series is increased during the night. We attribute this pattern in the dv/v uncertainty to strong directional noise generated by late-night freight trains within the Coachella Valley (Inbal *et al.* 2018).

We have demonstrated that our newly developed MCMC approach is able to extract more detailed information from seismic velocity change time-series than the more common methods for estimating variations in dv/v , such as the reference stack and matrix inversion (Sens-Schönfelder & Wegler 2006; Brenguier *et al.* 2014). The MCMC method can also provide more robust results, by providing direct information on uncertainty, and eliminating the need for *ad hoc* smoothing of the dv/v time-series. We have also shown that our MCMC approach has several computational advantages, which is an important consideration as modern seismic deployments increasingly promote the inclusion of large volumes of data. We expect that applying the MCMC approach in more study environments and to a wider variety of problems will result in a greater understanding of the mechanisms that cause seismic velocity changes within the solid earth.

ACKNOWLEDGEMENTS

GT designed the MCMC approach outlined in this study. GT performed the processing and analysis, with input and guidance from GH. GT created the figures and wrote the manuscript, with significant contributions from GH. The facilities of the IRIS Data Management System, and specifically the IRIS Data Management Center, were used for access to waveform and metadata required for the Piñon Flat study. Information on how to obtain the data used in the El Mayor-Cucapah analysis is available in Hillers *et al.* (2019) and Mao *et al.* (2019b). A Python version of the MCMC code used in this study is archived on Zenodo and can be found at doi: 10.5281/zenodo.3516603 (Taylor 2019). The MWCS function from the MSNoise package (Lecocq *et al.* 2014) was used to compute the velocity variations between the cross-correlation functions used in this study. Generic Mapping Tools (Wessel *et al.* 2013) was used to create the figures. We would like to thank A. Mordret for kindly providing the cross-correlation database for the Salton Sea Geothermal Field. We are also grateful to Taka'aki Taira and Shujuan Mao for providing their dv/v results for the Salton Sea Geothermal Field. We would like to thank the editor Lapo Boschi, Marine Denolle and one anonymous reviewer for their comments, which helped improve the manuscript.

REFERENCES

- Agnew, D.C., 1981. Nonlinearity in rock: evidence from Earth tides, *J. Geophys. Res. Solid Earth*, **86**(B5), 3969–3978. doi:10.1029/JB086iB05p03969.
- Agnew, D.C., 2012. SPOTL: some programs for ocean-tide loading. Available at: <https://igppweb.ucsd.edu/agnew/Spotl/spotlmain.html>.
- Agnew, D.C., 2015. Earth tides, in *Treatise on Geophysics and Geodesy*, 2nd edn, Vol. 3, pp. 151–178, ed. Schubert, G., Elsevier.
- Amey, R.M.J., Hooper, A. & Morishita, Y., 2019. Going to any lengths: solving for fault size and fractal slip for the 2016, M_w 6.2 Central Tottori Earthquake, Japan, using a transdimensional inversion scheme, *J. Geophys. Res. Solid Earth*, **124**, 4001–4016. doi:10.1029/2017JB015316.
- Amey, R.M.J., Hooper, A. & Walters, R.J., 2018. A Bayesian method for incorporating self-similarity into earthquake slip inversions, *J. Geophys. Res. Solid Earth*, **123**, 6052–6071. doi:10.1029/2018JB016434.
- Aster, R.C., Borchers, B. & Thurber, C.H., 2013. *Parameter Estimation and Inverse Problems*, 2nd edn, Elsevier.
- Baisch, S. & Bokelmann, G.H.R., 2001. Seismic waveform attributes before and after the Loma Prieta earthquake: scattering change near the earthquake and temporal recovery, *J. Geophys. Res. Solid Earth*, **106**(B8), 16 323–16 337.
- Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.*, **169**, 1239–1260. doi:10.1111/j.1365-246X.2007.03374.x.
- Bodin, T., Sambridge, M., Rawlinson, N. & Arroucau, P., 2012. Transdimensional tomography with unknown data noise, *Geophys. J. Int.*, **189**, 1536–1556. doi:10.1111/j.1365-246X.2012.05414.x.
- Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N.M., Nadeau, R.M. & Larose, E., 2008a. Postseismic relaxation along the San Andreas Fault at Parkfield from continuous seismological observations, *Science*, **321**(5895), 1478–1481. doi:10.1126/science.1160943.
- Brenguier, F., Campillo, M., Takeda, T., Aoki, Y., Shapiro, N.M., Briand, X., Emoto, K. & Miyake, H., 2014. Mapping pressurized volcanic fluids from induced crustal seismic velocity drops, *Science*, **345**, 80–82. doi:10.1126/science.1254073.
- Brenguier, F., Shapiro, N.M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O. & Nercessian, A., 2008b. Towards forecasting volcanic eruptions using seismic noise, *Nat. Geosci.*, **1**, 126–130. doi:10.1038/ngeo104.
- Chaput, J., Clerc, V., Campillo, M., Roux, P. & Knox, H., 2016. On the practical convergence of coda-based correlations: a window optimization approach, *Geophys. J. Int.*, **204**, 736–747. doi:10.1093/gji/ggv476.
- Clarke, D., Zaccarelli, L., Shapiro, N.M. & Brenguier, F., 2011. Assessment of resolution and accuracy of the Moving Window Cross Spectral technique for monitoring crustal temporal variations using ambient seismic noise, *Geophys. J. Int.*, **186**, 867–882. doi:10.1111/j.1365-246X.2011.05074.x.
- Clements, T. & Denolle, M.A., 2018. Tracking groundwater levels using the ambient seismic field, *Geophys. Res. Lett.*, **45**, 6459–6465. doi:10.1029/2018GL077706.
- Dettmer, J., Benavente, R., Cummins, P.R. & Sambridge, M., 2014. Trans-dimensional finite-fault inversion, *Geophys. J. Int.*, **199**, 735–751. doi:10.1093/gji/ggu280.
- Donnellan, A., Parker, J., Hensley, S., Pierce, M., Wang, J. & Rundle, J., 2014. UAVSAR observations of triggered slip on the Imperial, Superstition Hills, and East Elmore Ranch Faults associated with the 2010 M7.2 El Mayor Cucapah earthquake, *Geochem. Geophys. Geosys.*, **15**, 815–829. doi:10.1002/2013GC005120.
- Fox, M., Goren, L., May, D.A. & Willett, S.D., 2014. Inversion of fluvial channels for paleorock uplift rates in Taiwan, *J. Geophys. Res. Solid Earth*, **119**, 1853–1875. doi:10.1002/2014JF003196.
- Galetti, E., Curtis, A., Baptie, B., Jenkins, D. & Nicolson, H., 2017. Transdimensional Love-wave tomography of the British Isles and shear-velocity structure of the East Irish Sea Basin from ambient-noise interferometry, *Geophys. J. Int.*, **208**, 36–58. doi:10.1093/gji/ggw286.
- Gómez-García, C., Brenguier, F., Boué, P., Shapiro, N., Droznina, S., Senyukov, S. & Gordeev, E., 2018. Retrieving robust noise-based seismic velocity changes from sparse data sets: synthetic tests and application to Klyuchevskoy volcanic group (Kamchatka), *Geophys. J. Int.*, **214**, 1218–1236. doi:10.1093/gji/ggy190.
- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov Chains and their applications, *Biometrika*, **57**(1), 97–109. doi:10.2307/2334940.
- Hauksson, E., Stock, J., Hutton, K., Yang, W., Vidal-Villegas, J.A. & Kanamori, H., 2011. The 2010 M_w 7.2 El Mayor-Cucapah Earthquake Sequence, Baja California, Mexico and Southernmost California, USA: active Seismotectonics Along the Mexican Pacific Margin, *Pure appl. Geophys.*, **168**, 1255–1277. doi:10.1007/s00024-010-0209-7.
- Herlihy, M. & Shavit, N., 2012. *The Art of Multiprocessor Programming*, Elsevier.
- Hestenes, M.R. & Stiefel, E., 1952. Methods of conjugate gradients for solving linear systems, *J. Res. Natl. Bureau Standards*, **49**, 409–436.
- Hillers, G., Ben-Zion, Y., Campillo, M. & Zigone, D., 2015a. Seasonal variations of seismic velocities in the San Jacinto fault area observed with ambient seismic noise, *Geophys. J. Int.*, **202**, 920–932. doi:10.1093/gji/ggv151.
- Hillers, G., Campillo, M., Brenguier, F., Moreau, L., Agnew, D. & Ben-Zion, Y., 2019. Seismic velocity change patterns along the San Jacinto fault zone following the 2010 M7.2 El Mayor-Cucapah and M5.4 Collins Valley earthquakes, *J. Geophys. Res. Solid Earth*, **124**, 7171–7192. doi:10.1029/2018JB017143.
- Hillers, G., Campillo, M. & Ma, K.-F., 2014. Seismic velocity variations at TCDP are controlled by MJO driven precipitation pattern and high fluid discharge properties, *Earth Planet. Sci. Lett.*, **391**, 121–127. doi:10.1016/j.epsl.2014.01.040.
- Hillers, G., Retailleau, L., Campillo, M., Inbal, A., Ampuero, J.P. & Nishimura, T., 2015b. In situ observations of velocity changes in response to tidal deformation from analysis of the high-frequency ambient wavefield, *J. Geophys. Res. Solid Earth*, **120**, 210–225. doi:10.1002/2014JB011318.
- Inbal, A., Ampuero, J.-P. & Avouac, J.-P., 2017. Locally and remotely triggered aseismic slip on the central San Jacinto Fault near Anza, CA, from joint inversion of seismicity and strainmeter data, *J. Geophys. Res. Solid Earth*, **122**, 3033–3061. doi:10.1002/2016JB013499.
- Inbal, A., Cristea-Platon, T., Ampuero, J.-P., Hillers, G., Agnew, D. & Hough, S.E., 2018. Sources of long-range anthropogenic noise in southern California and implications for tectonic tremor detection, *Bull. Seism. Soc. Am.*, **108**(6), 3511–3527. doi:10.1785/0120180130.
- Ingleby, T. & Wright, T.J., 2017. Omori-like decay of postseismic velocities following continental earthquakes, *Geophys. Res. Lett.*, **44**, 3119–3130. doi:10.1002/2017GL072865.
- Lecocq, T., Caudron, C. & Brenguier, F., 2014. MSNoise, a Python package for monitoring seismic velocity changes using ambient seismic noise, *Seismol. Res. Lett.*, **85**(3), 715–726. doi:10.1785/0220130073.
- Lecocq, T., Longuevergne, L., Pedersen, H.A., Brenguier, F. & Stammer, K., 2017. Monitoring ground water storage at mesoscale using seismic noise: 30 years of continuous observation and thermo-elastic and hydrological modeling, *Sci. Rep.*, **7**, doi:10.1038/s41598-017-14468-9.
- Lobkis, O.I. & Weaver, R.L., 2003. Coda-wave interferometry in finite solids: recovery of P -to- S conversion rates in an elastodynamic billiard, *Phys. Rev. Lett.*, **90**(25), doi:10.1103/PhysRevLett.90.254302.
- Mao, S., Campillo, M., van der Hilst, R.D., Brenguier, F., Stehly, L. & Hillers, G., 2019a. High temporal resolution monitoring of small variations in crustal strain by dense seismic arrays, *Geophys. Res. Lett.*, **46**, 128–137. doi:10.1029/2018GL079944.
- Mao, S., Mordret, A., Campillo, M., Fang, H. & van der Hilst, R.D., 2019b. On the measurement of seismic travel-time changes in the time-frequency domain with wavelet cross-spectral analysis, *Geophys. J. Int.*, doi:10.1093/gji/ggz495, in press.
- Meier, U., Shapiro, N.M. & Brenguier, F., 2010. Detecting seasonal variations in seismic velocities within Los Angeles basin from correlations of ambient seismic noise, *Geophys. J. Int.*, **181**, 985–996. doi:10.1111/j.1365-246X.2010.04550.x.
- Mikesell, T.D., Malcolm, A.E., Yang, D. & Haney, M.M., 2015. A comparison of methods to estimate seismic phase delays: numerical examples for coda wave interferometry, *Geophys. J. Int.*, **202**, 347–360. doi:10.1093/gji/ggv138.

- Minson, S.E., Simons, M. & Beck, J.L., 2013. Bayesian inversion for finite fault earthquake source models I—theory and algorithm, *Geophys. J. Int.*, **194**, 1701–1726. doi:10.1093/gji/ggt180.
- Moreau, L., Stehly, L., Boué, P., Lu, Y., Larose, E. & Campillo, M., 2017. Improving ambient noise correlation functions with an SVD-based Wiener filter, *Geophys. J. Int.*, **211**(1), 418–426. doi:10.1093/gji/ggx306.
- Mosegaard, K., 1998. Resolution analysis of general inverse problems through inverse Monte Carlo sampling, *Inverse Probl.*, **14**(3), doi:10.1088/0266-5611/14/3/004.
- Nakata, N. & Snieder, R., 2011. Near-surface weakening in Japan after the 2011 Tohoku-Oki earthquake, *Geophys. Res. Lett.*, **38**(17), doi:10.1029/2011GL048800.
- Niu, F., Silver, P.G., Daley, T.M., Cheng, X. & Majer, E.L., 2008. Preseismic velocity changes observed from active source monitoring at the parkfield SAFOD drill site, *Nature*, **454**, 204–208. doi:10.1038/nature07111.
- Obermann, A., Froment, B., Campillo, M., Larose, E., Planès, T., Valette, B., Chen, J.H. & Liu, Q.Y., 2014. Seismic noise correlations to image structural and mechanical changes associated with the M_w 7.9 2008 wenchuan earthquake, *J. Geophys. Res. Solid Earth*, **119**, 3155–3168. doi:10.1002/2013JB010932.
- Obermann, A., Kraft, T., Larose, E. & Wiemer, S., 2015. Potential of ambient seismic noise techniques to monitor the St. Gallen geothermal site (Switzerland), *J. Geophys. Res. Solid Earth*, **120**(6), 4301–4316. doi:10.1002/2014JB011817.
- Olivier, G., Brenguier, F., de Wit, T. & Lynch, R., 2017. Monitoring the stability of Tailings dam walls with ambient seismic noise, *Lead. Edge*, **36**, 350a1–350a6. doi:10.1190/tle36040350a1.1.
- Peterson, J., 1993. Observations and modeling of seismic background noise, Open file rep., U.S. Geol. Surv.
- Planès, T., Rittgers, J., Mooney, M., Kanning, W. & Draganov, D., 2017. Monitoring the tidal response of a sea levee with ambient seismic noise, *J. Appl. Geophys.*, **138**, 255–263. doi:10.1016/j.jappgeo.2017.01.025.
- Poupinet, G., Ellsworth, W.L. & Frechet, J., 1984. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California, *J. Geophys. Res. Solid Earth*, **89**, 5719–5731. doi:10.1029/JB089iB07p05719.
- Qiu, H., Hillers, G. & Ben-Zion, Y., 2019. Temporal changes of seismic velocities in the San Jacinto Fault zone associated with the 2016 M_w 5.2 Borrego Springs earthquake, *Geophys. J. Int.*, doi:10.1093/gji/ggz538, in press.
- Richter, T., Sens-Schönfelder, C., Kind, R. & Asch, G., 2014. Comprehensive observation and modeling of earthquake and temperature-related seismic velocity changes in northern Chile with passive image interferometry, *J. Geophys. Res. Solid Earth*, **119**, 4747–4765. doi:10.1002/2013JB010695.
- Ringdal, F. & Bungum, H., 1977. Noise level variation at NORSAR and its effect on detectability, *Bull. Seism. Soc. Am.*, **67**(2), 479–492.
- Rivet, D., Brenguier, F., Clarke, D., Shapiro, N.M. & Peltier, A., 2014. Long-term dynamics of Piton de la Fournaise volcano from 13 years of seismic velocity change measurements and GPS observations, *J. Geophys. Res. Solid Earth*, **119**, 7654–7666. doi:10.1002/2014JB011307.
- Roberts, G.O., Gelman, A. & Gilks, W.R., 1997. Weak convergence and optimal scaling of random walk Metropolis algorithms, *Ann. Appl. Probab.*, **7**, 110–120. doi:10.1214/aop/1034625254.
- Rubinstein, J.L. & Beroza, G.C., 2004. Nonlinear strong ground motion in the ML 5.4 Chittenden earthquake: evidence that preexisting damage increases susceptibility to further damage, *Geophys. Res. Lett.*, **31**(23), doi:10.1029/2004GL021357.
- Ryan, W.B.F., Carbotte, S. M., Coplan, J., O'Hara, S., Melkonian, A., Arko, R., Weissel, R. A., Ferrini, V., Goodwillie, A., Nitsche, F., Bonczkowski, J. & Zensky, R., 2009. Global Multi-Resolution Topography (GMRT) synthesis data set, *Geochem. Geophys. Geosyst.*, **10**, Q03014. doi:10.1029/2008GC002332.
- Schaff, D.P. & Beroza, G.C., 2004. Coseismic and postseismic velocity changes measured by repeating earthquakes, *J. Geophys. Res. Solid Earth*, **109**(B10), doi:10.1029/2004JB003011.
- Sens-Schönfelder, C., Pomponi, E. & Peltier, A., 2014. Dynamics of Piton de la Fournaise volcano observed by passive image interferometry with multiple references, *J. Volc. Geotherm. Res.*, **276**, 32–45. doi:10.1016/j.jvolgeores.2014.02.01.
- Sens-Schönfelder, C. & Wegler, U., 2006. Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia, *Geophys. Res. Lett.*, **33**(21), doi:10.1029/2006GL027797.
- Stehly, L., Campillo, M. & Shapiro, N.M., 2006. A study for the seismic noise from its long-range correlation properties, *J. Geophys. Res. Solid Earth*, **111**(B10), doi:10.1029/2005JB004237.
- Taira, T., Nayak, A., Brenguier, F. & Manga, M., 2018. Monitoring reservoir response to earthquakes and fluid extraction, Salton Sea geothermal field, California, *Sci. Adv.*, **4**(1), e1701536. doi:10.1126/sciadv.1701536.
- Takagi, R., Okada, T., Nakahara, H., Umino, N. & Hasegawa, A., 2012. Coseismic velocity change in and around the focal region of the 2008 Iwate-Miyagi Nairiku earthquake, *J. Geophys. Res. Solid Earth*, **117**, doi:10.1029/2012JB009252.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics.
- Taylor, G., 2019. First release of MCMC-DVV, Available at: <https://doi.org/10.5281/zenodo.3516603>.
- Tsai, V.C., 2011. A model for seasonal changes in GPS positions and seismic wave speeds due to thermoelastic and hydrologic variations, *J. Geophys. Res. Solid Earth*, **116**(B4), doi:10.1029/2010JB008156.
- Vernon, F., 2014. Piñon Flat Observatory (PFO) Array, International Federation of Digital Seismograph Networks, Other/Seismic Network, Seattle, WA. doi:10.7914/SN/PY.
- Viens, L., Denolle, M.A., Hirata, N. & Nikagawa, S., 2018. Complex near-surface rheology inferred from the response of greater Tokyo to strong ground motions, *J. Geophys. Res. Solid Earth*, **123**, 5710–5729. doi:10.1029/2018JB015697.
- Voisin, C., Garambois, S., Massey, C. & Brossier, R., 2016. Seismic noise monitoring of the water table in a deep-seated, slow-moving landslide, *Interpretation*, **4**(3), doi:10.1190/INT-2016-0010.1.
- Wathen, A.J., 2015. Preconditioning, *Acta Numer.*, **24**, 329–376. doi:10.1017/S0962492915000021.
- Wegler, U. & Sens-Schönfelder, C., 2007. Fault zone monitoring with passive image interferometry, *Geophys. J. Int.*, **168**, 1029–1033. doi:10.1111/j.1365-246X.2006.03284.x.
- Wei, M., Sandwell, D., Fialko, Y. & Bilham, R., 2011a. Slip on faults in the Imperial Valley triggered by the 4 April 2010 Mw 7.2 El Mayor-Cucapah earthquake revealed by InSAR, *Geophys. Res. Lett.*, **38**(1), doi:10.1029/2010GL045235.
- Wei, S., et al., 2011b. Superficial simplicity of the 2010 El Mayor-Cucapah earthquake of Baja California in Mexico, *Nat. Geosci.*, **4**, 615–618. doi:10.1038/ngeo1213.
- Wessel, P., Smith, W.H.F., Scharroo, R., Luis, J. & Wobbe, F., 2013. Generic mapping tools: improved version released, *EOS, Trans. Am. Geophys. Un.*, **94**(45), 409–410. doi:10.1002/2013EO450001.