Monitoring velocity change over 20 years at Parkfield

Kurama Okubo[1], Brent G. Delbridge[2], and Marine Denolle[3]

[1]National Research Institute for Earth Science and Disaster Resilience
[2]Los Alamos National Laboratory
[3]University of Washington

October 27, 2023

Abstract

We monitored the time history of the velocity change (dv/v) from 2002 to 2022 to investigate temporal changes in the physical state near the Parkfield Region of the San Andreas Fault throughout the interseismic period. Following the coseismic decrease in dv/v caused due to the 2003 San Simeon and the 2004 Parkfield earthquakes, the dv/v heals logarithmically and shows a net long-term increase in which the current dv/v level is equivalent to, or exceeding, the value before the 2003 San Simeon earthquake. We investigated this long-term trend by fitting the model accounting for the environmental and coseismic effects to the channel-weighted dv/v time series. We confirmed with the metrics of AIC and BIC that the additional term of either a linear trend term, or a residual healing term for the case where the healing had not been completed before the San Simeon earthquake occurred, robustly improved the fit to the data. We eventually evaluated the sensitivity of the dv/v time history to the GNSS-derived strain field around the fault. The cumulative dilatational strain spatially averaged around the seismic stations shows a slight extension, which is opposite to what would be expected for an increase in dv/v. However, the cumulative rotated axial strain shows compression in a range near the maximum contractional horizontal strain (azimuth of N35°W to N45°E), suggesting that the closing of pre-existing microcracks aligned perpendicular to the axial contractional strains would be a candidate to cause the long-term increase observed in the multiple station pairs.

Open research

We maintain software, a minimal working example, and the Jupyter notebooks in GitHub to reproduce the analysis presented in the manuscript.

- SeisMonitoring.jl: A Julia-based software package to process the ambient seismic noise.
- SeisMonitoring_Example: A minimal working example of the SeisMonitoring.jl
- SeisMonitoring_Paper: Jupyter notebooks to post-process the data and to plot figures

The intermediate files of the post-processing are available in the UW dasway.

Monitoring velocity change over 20 years at Parkfield

Kurama Okubo[1], Brent G. Delbridge[2], Marine A. Denolle[3]

[1]National Research Institute for Earth Science and Disaster Resilience (NIED), Tsukuba, Japan
[2]Los Alamos National Laboratory, NM, USA
[3]Department of Earth and Space Sciences, University of Washington, WA, USA

Key Points:

- Ambient seismic noise monitoring at Parkfield measures relative velocity change (dv/v) from 2002 to 2022.
- The post-seismic healing from the 2004 Parkfield earthquake continues today, with the current dv/v exceeding the level before the event.
- A statistically significant long-term increase correlates with the contractional axial strains caused by tectonic loading.

Corresponding author: K. Okubo, kokubo@bosai.go.jp

Abstract

We monitored the time history of the velocity change (dv/v) from 2002 to 2022 to investigate temporal changes in the physical state near the Parkfield Region of the San Andreas Fault throughout the interseismic period. Following the coseismic decrease in dv/v caused due to the 2003 San Simeon and the 2004 Parkfield earthquakes, the dv/v heals logarithmically and shows a net long-term increase in which the current dv/v level is equivalent to, or exceeding, the value before the 2003 San Simeon earthquake. We investigated this long-term trend by fitting the model accounting for the environmental and coseismic effects to the channel-weighted dv/v time series. We confirmed with the metrics of AIC and BIC that the additional term of either a linear trend term, or a residual healing term for the case where the healing had not been completed before the San Simeon earthquake occurred, robustly improved the fit to the data. We eventually evaluated the sensitivity of the dv/v time history to the GNSS-derived strain field around the fault. The cumulative dilatational strain spatially averaged around the seismic stations shows a slight extension, which is opposite to what would be expected for an increase in dv/v. However, the cumulative rotated axial strain shows compression in a range near the maximum contractional horizontal strain (azimuth of N35°W to N45°E), suggesting that the closing of pre-existing microcracks aligned perpendicular to the axial contractional strains would be a candidate to cause the long-term increase observed in the multiple station pairs.

Plain Language Summary

We monitored the temporal change of velocity (dv/v) around the Parkfield Region from 2002 to 2022 to investigate the healing of dv/v after the 2003 San Simeon and the 2004 Parkfield earthquakes. Following those events, the dv/v recovers logarithmically with time and shows a net long-term increase in which the current dv/v level is equivalent to, or exceeding, the value before the 2003 San Simeon earthquake. We investigated this long-term trend by fitting the model accounting for the environmental factors, the coseismic effects, and the additional term explaining the long-term increase. Models with the additional term provide a statistically robust improved fit to the data, suggesting the long-term increase observed in the dv/v time history is a non-negligible factor. We eventually evaluated the sensitivity of the dv/v to the GNSS-derived strain field around the fault. The spatially averaged dilatational strain could not explain the increase in dv/v as it shows a slight extension, whereas the rotated axial strain shows compression in a range near the maximum contractional horizontal strain. The closing of pre-existing microcracks aligned perpendicular to the contractional strains thus would be a candidate to cause the increase in the material's rigidity and the average seismic wave velocity.

1 Introduction

Continuous seismic monitoring of the Earth on decadal timescales allows for the exploration of environmental, volcanic, and tectonic phenomena that control the seismic properties of the subsurface. Observed changes in the seismic velocity of the subsurface often result from a combination of different effects, and decadal surveys are required to untangle the contribution of each mechanism. For example, the recovery from the 1999 Chichi Earthquake is masked by environmental fluctuations such as rainfall (Feng et al., 2021), and hydrological loads overprint volcanic activities in Mount St. Helens (Hotovec-Ellis et al., 2015).

The seismic quantity dv/v provides a relative measure of the volume-averaged perturbation in the seismic velocity of the subsurface and is usually measured using the coda of waves from repeated sources and receivers. Coda waves are often used since they are sensitive to small perturbations in the subsurface (Snieder et al., 2002; Lobkis & Weaver,

2003) and ballistic wave measurements are complicated by varying source locations or ambient field stationarity (Colombi et al., 2014; Takano, Brenguier, et al., 2019). Reliable repeated coda sources and receivers may come from either repeating earthquakes recorded at the same stations (Poupinet et al., 1984) or from repeated ambient noise cross-correlation functions (Sens-Schönfelder & Wegler, 2006).

Poupinet et al. (1984) first applied the doublet method using two microearthquakes occurring before and after the 1979 M5.9 Coyote Lake earthquake and found the time delay in the coda of the doublets increased linearly with time, indicating the change in dv/v over the medium. Numerous studies have since shown that the seismic properties of fault zones can be monitored by dv/v using waveforms from any earthquake doublet, or set of repeating earthquakes, with similar enough waveforms to compare the arrival times (e.g., Schaff & Beroza, 2004; Rubinstein et al., 2007; Sawazaki et al., 2015; Sheng et al., 2021; Hotovec-Ellis et al., 2022).

One of the most popular approaches used to extract the continuous measurement of dv/v in time uses the cross-correlation of ambient seismic noise (Nakata et al., 2019). The advantage of using ambient seismic noise is that one does not have to wait for earthquakes to occur, which enables continuous measurements of dv/v. The dv/v is sensitive to the groundwater level (GWL) (Sens-Schönfelder & Wegler, 2006; Rivet et al., 2015; Clements & Denolle, 2018; Nishida et al., 2020; Illien et al., 2021; Mao et al., 2022; Illien et al., 2022; Clements & Denolle, 2023) and the thermoelastic deformation induced by the air temperature (Richter et al., 2014; Gassenmeier et al., 2016; Colombero et al., 2018), or even both (Wang et al., 2017; Feng et al., 2021; Lecocq et al., 2017). When interpreting dv/v measurements, it is thus crucial to identify the environmental factors as well as other mechanisms that cause variations in dv/v, such as a drop in velocity associated with coseismic earthquake rupture, logarithmic healing following an earthquake, and the other potential tectonic factors (Taira et al., 2015; Feng et al., 2021; Clements & Denolle, 2023).

Regardless of the type of wavefield used (e.g., earthquakes or ambient field), dv/v measurements vary in time due to the changing tectonic strains associated with the earthquake cycle. *Before the earthquake*, laboratory studies predict a slight decrease in the seismic velocities (Shreedharan et al., 2021), which has been observed only in Parkfield, CA (Niu et al., 2008) and Italy (Chiarabba et al., 2020). *During the earthquake*, many studies have investigated the effect of coseismic damage on the seismic velocities (e.g., Wegler & Sens-Schönfelder, 2007), either near the fault due to the extremely high strain rates (Brenguier, Campillo, et al., 2008; Wu et al., 2016; Taira et al., 2015; Lu & Ben-Zion, 2021; Nimiya et al., 2017) or at the surface as a result of the strong ground motions (Gassenmeier et al., 2016; Viens et al., 2018; Bonilla et al., 2019). Studies have found significant correlations between reductions in the seismic velocities with peak dynamic strains or stresses (Richter et al., 2014; Taira et al., 2015; Hobiger et al., 2016; Viens et al., 2018). *After the earthquake*, the seismic velocity appears to heal logarithmically, comparable to the recovery governed by the slow dynamics (Ten Cate & Shankland, 1996; TenCate et al., 2000; Johnson & Sutin, 2005). Theoretical arguments suggest that healing should saturate after some time (Snieder et al., 2017). However, the saturation time scale is poorly understood, and observed healing times vary tremendously (Viens et al., 2018; Clements & Denolle, 2023; Illien et al., 2022). The postseismic healing is observed to last days, weeks, or months (Viens et al., 2018; Marc et al., 2021; Clements & Denolle, 2023), whereas little attention has been paid to the inter-seismic period between earthquakes. This is partly because inter-seismic loading usually occurs at low strain rates relative to the co- and post-seismic rates (Shreedharan et al., 2021) and is very difficult to identify due to contamination of the signal by various environmental processes and other seismic events.

This study focuses on a particularly well-instrumented segment of the San Andreas Fault (SAF) near Parkfield, California. These instruments were installed by the Berkeley Seismology Lab and have provided continuous seismic data since 2002 (doi:10.7932/HRSN). The seismic stations are located on the ruptured area of the 2004 Parkfield earthquake. Brenguier, Campillo, et al. (2008) measured the dv/v at Parkfield from 2002 until 2008 and showed a coseismic decrease in velocity associated with two earthquakes and their logarithmic healing, which they interpreted as a postseismic response. Subsequently, Wu et al. (2016) extended the analysis period to 2011 and evaluated the dv/v with different frequency bands to investigate depth-dependent velocity perturbations. They showed the coseismic decrease in velocity is larger with a higher frequency band corresponding to the shallow depth, which could have the contributions from the loss of the dv/v sensitivity for the layered perturbation at depth with the low-frequency range(Obermann et al., 2013; Yuan et al., 2021) and potentially depth-dependent damage.

We build upon these two previous studies by constructing a time series of the evolution of dv/v over 20 years (2002-2022). We take advantage of this longer time series to quantitatively determine how dv/v is affected by both environmental factors and two major local earthquakes, the 2003 M6.5 San Simeon (SS) and the 2004 M6.0 Parkfield (PF) earthquakes, and to investigate the residuals from them that might contain additional factors. We examine models that include the effects of precipitation, temperature, earthquake healing (e.g. SS and PF), and the long-term trend term which is included as either a linear term or a residual healing term in our analysis to quantitatively determine the non-negligible increase in dv/v, which resulted in the dv/v potentially being greater in 2022 than the value before the 2003 San Simeon earthquake.

Non-linear elasticity rheology predicts that dv/v is proportional to dilatational strains (see review in Clements & Denolle, 2023). Many observations of such relation support the theory: tidal oscillations and the volcanic activities (e.g. Yamamura et al., 2003; Donaldson et al., 2017; Hirose et al., 2017; Takano et al., 2017; Mao et al., 2019; Sens-Schonfelder & Eulenfeld, 2019; Takano, Nishimura, et al., 2019; Nishida et al., 2020; Hotovec-Ellis et al., 2022). In this study, we evaluated the surface strain field at Parkfield using the Global Navigation Satellite System (GNSS) dataset. We compare the GPS-derived strains with the dv/v measurements to constrain the effects of tectonic deformations on dv/v measurements. The strain field around the SAF is perturbed by creep on the shallow fault (Bacques et al., 2018) and by the intricate pattern of locked and creeping patches deeper on the fault surface(Jolivet et al., 2015). In order to compare this spatially complex strain field with dv/v measurements, we calculate the spatially averaged cumulative strains around the SAF, which includes the seismic stations.

In this study, we describe the methods and processing workflows used to obtain stable dv/v measurements at decadal timescales (Section 2), which required developing new high-performance software. We then present a survey of the dv/v results, with a particular focus on several distinct station-component pair combinations and frequency bands (Section 3). To fit models and quantify their sensitivity to the dv/v time series, we employ Markov chain Monte Carlo (MCMC) for Bayesian inference (Section 4). We then compare the strain field at Parkfield using the GNSS data to the dv/v time series (Section 5). The quantitative contributions of different physical mechanisms of both tectonic and environmental origins to the observed dv/v time series are constrained and discussed (Section 6). The abbreviations used in this article are listed in Table 1.

2 Methodology

Day-long instrument corrected time series of continuous (20 Hz sampling rate) three-component seismic data from January 2002 to May 2022 from 13 stations of the High-Resolution Seismic Network (HRSN; Figure 1) were used to compute the dv/v time-series.We removed the instrumental response at the beginning of data processing. Each of the three components, rather than just the typically used vertical component, was used in this analysis in order to improve the dv/v quality (Schaff, 2012) and the temporal continuity after channel-weighting (see Section 2.6).

The calculation of a dv/v time series from an ambient seismic noise dataset of this size (1.6 TB) is a computational challenge due to the extremely large number of cross-correlations that are required to be computed. To overcome this computationally prohibitive obstacle which prevents the use of existing seismic processing tools, we developed a Julia-based software `SeisMonitoring.jl` (doi:10.5281/zenodo.832094) which can be used in high-performance computing environments. Julia is a powerful language to perform computationally intensive processes (Bezanson et al., 2017) such as computing large numbers of cross-correlations in parallel. Jones et al. (2020) developed the Julia-based software tool, `SeisIO.jl`, to handle the seismic waveforms with metadata in the structure and showed the advantage in using memory and computational speed. Clements and Denolle (2020) developed the Julia modules called `SeisNoise.jl`, which efficiently computes the cross-correlations of short-time windows to conduct the dv/v monitoring. The new software tool presented here, `SeisMonitoring.jl`, uses these packages as dependencies, which greatly reduces the processing time with process parallelization.

2.1 Data quality

We first evaluated the power spectral density (PSD) of the seismic waveforms to investigate the data quality regarding the data continuity and noise levels. We used the Blackman-Tukey method to estimate the PSD, which is based on Wiener-Khinchine's theorem, such that the PSD can be obtained with the Fourier transform of the auto-correlation functions (ACFs). This scheme is efficient for our case compared to Welch's method as we can reuse the monthly-stacked ACFs to compute the decadal history of the PSD. We computed the daily ACFs after rescaling and removing the instrumental response associated with each station and channel. We then apply the Tukey window with 5% on the ACFs and apply the FFT to obtain the discrete Fourier coefficients, which represent the PSD of the day.

Sudden changes in the power spectrum amplitude for several various stations and times are observed in the time series (e.g. EADB around 2014; Figure S2), which are likely due to instrumental changes and artifacts, which are not noted in the metadata. Figure S2 shows the PSD associated with the station EADB as a representative station pair. The PSD of the other stations can be found in Supplementary Dataset S1. We found some defects in the waveform data, which need to be thresholded out during the noise processing. These step-like changes can potentially cause artifacts in the estimation of dv/v, which is discussed in the next section. Notably, there was a prominent peak in the PSD at 1 Hz in the channel EADB-3 around 2012-2014, which could also be due to instrumental and/or environmental noise. This type of noise is observed in some other stations, such as FROB, contaminating the ACFs as the coherence of the periodic noise becomes dominant. As explained later in this section, we mitigate this artifact in the noise cross-correlations during the stacking process.

We cross-verified the data quality estimated from the PSDs using the supporting information of Shelly (2017). In this study, we did not correct the swap of channels or the reversed polarity on the raw data reported by Shelly (2017). Instead, we applied multiple thresholds during the processing of the seismic noise associated with the amplitude of hourly and daily cross-correlation functions (CCFs) and the correlation coefficient (CC) between the reference and current CCFs during the stacking so that the perturbed CFs are mostly excluded from the stacked CCF to minimize the bias caused due to those instrumental issues in the analysis of dv/v.

2.2 Removal of coherent transient signals

The continuous seismic waveform contains transient signals such as earthquakes, tectonic tremors, environmental signals, and instrumental noises. The transient signals degrade the stability of the stacked correlation function (CF) as they induce strong, coherent signals on the correlation function that are typically absent. Zhou et al. (2020) implemented the process of asynchronous temporal flattening, which mutes the time windows containing the large amplitudes caused by the transient signals to avoid the perturbation of CFs. In this study, we also muted the transient signals from the continuous waveform using the kurtosis and classic STA/LTA algorithms (Allen, 1978) to extract impulsive earthquake and emergent tectonic tremor events.

The kurtosis within the moving window is sensitive to the spike-like transient signals that detect the events. We followed Baillard et al. (2013) and used kurtosis as the characteristic function to pick the events. We calculated the kurtosis for every three-minute-long segment. An event is detected when the kurtosis exceeds a given threshold. The standard STA/LTA algorithm can detect tremor-like events. Unlike for detecting earthquake signals, we chose time window lengths of STA/LTA tuned to detect relatively long-term perturbations such as tectonic tremors: three minutes and one day for the STA and LTA lengths, respectively. The threshold of kurtosis and STA/LTA is set to be three. The combination of the kurtosis and STA/LTA detection thus performs to clean the raw data containing various types of transient signals.

We removed the signal in the time windows when either kurtosis or STA/LTA exceeded the threshold by applying the inverted Tukey window. We computed the percentage of taper with respect to the removed time windows such that the taper duration is fixed at thirty seconds. Figure S3 shows an example of a waveform observed at EADB and the removal of transient signals. Figure S3d shows the waveform after removing transient signals. After transient removals, the noise amplitude is balanced enough to mitigate the artifacts of transient signals on the stacked CFs.

Figure S4 shows the data availability before and after removing the transient signals. We computed the fraction of missing data in the daily waveforms associated with each station and channel. A sufficient amount of continuous data is available even after removing the transient signals from the raw data.

2.3 Auto- and cross-correlation functions

We computed the CFs for all the combinations of station and component in Julia using `SeisMonitoring.jl`. We sliced the data into hourly time windows with half an hour of overlap. We applied detrending, demeaning, and tapering on the time windows before computing the FFT and the CFs. The maximum lag time of the correlation is one hundred seconds. The resolution of the correlation function can be improved by either spectral normalization (Bensen et al., 2007; Viens et al., 2017) or temporal normalization methods such as the one-bit filter (Campillo & Paul, 2003; Larose et al., 2004; Shapiro & Campillo, 2004; Shapiro et al., 2005; Bensen et al., 2007; Durand et al., 2011; Seydoux et al., 2016), or using non-linear filters (Baig et al., 2009; Hadziioannou et al., 2011; Moreau et al., 2017; Viens & Van Houtte, 2019). In this study, however, we computed the CFs without those normalizations for simplicity in the processing and interoperability of the correlated wavefield.

To analyze the dv/v in different frequency bands, we applied the continuous wavelet transform (CWT) to the CFs as accurate narrow-bandpass filters. The theoretical background of CWT is summarized in Torrence and Compo (1998). Mao et al. (2020) showed the wavelet cross-spectrum approach improves the measurement of dv/v compared to the standard doublet method. Yuan et al. (2021) showed the combination of CWT and stretching method efficiently retrieves the dv/v with various frequency bands. We thus applied the CWT and inverse CWT to reconstruct the CFs in the time domain with the frequency bands of 0.2-0.5, 0.5-0.9, 0.9-1.2, and 1.2-2.0Hz. We used the Morlet mother wavelet for the CWT, and the factors used to reconstruct the waveforms are taken from Table 2 of Torrence and Compo (1998). The module is implemented in `SeisDvv.jl` in Julia, which is translated from the python module `pycwt`. We applied the tapering on the CFs to avoid the artifacts at the edges of cross-correlation functions after filtering with the CWT. Note that the spectral normalization is not applied in this study.

After filtering the CFs into frequency bands, we threshold out bad CFs using the maximum amplitude of hourly CFs before computing the daily stacks. Most of the transient signals have been removed from the waveform during the pre-processing. Still, we continue curating the result by removing the remaining cause of the perturbation of the stacked CFs. We detect and mask the hourly CF stacks with low coherency, which is less likely to improve the stationary phases in the day-stacked CFs. To detect them, we first evaluated the maximum amplitude of CFs associated with each hourly time window and subsequently computed the median of these values over a day to detect the outliers in the set of hourly CFs. We rejected the hourly CFs if the maximum amplitude of CFs is greater than three times or smaller than 10% of the median. This median-mute filter allows for cleaning the daily stacked CFs in terms of the fluctuation in the amplitude caused by the remaining transient signals or the instrumental issues. Note that this filter cannot address the incoherent phases with an amplitude similar to the ambient noise due to the small earthquakes. The artifacts of the incoherent CFs are thresholded out during the stacking phase implemented as the selective stack explained in the later section.

Figure S5 shows the nine components of CCFs associated with the station pair of LCCB-SCYB for the frequency range of 0.9-1.2Hz. It shows the relatively strong coherence in the ballistic wave arrivals followed by the coda waves. The other station pairs can be found in Supplementary Dataset S2. The original data of cross-correlation functions are available in the cloud storage `dasway` (doi: `https://doi.org/10.6069/PK9D-9411`). Note that the location codes and channel names have been replaced between November 2010 and September 2011. We thus unified them to obtain the continuous set of CFs, given that they are recorded on the same site.

2.4 Stacking

We computed the reference CFs by stacking the daily CFs from January 2010 to May 2022. We assume the reference CFs converge enough to evaluate the dv/v with this period. The bottom waveforms of Figure S5 show the reference stacks for the cases between 2010 and 2022 and the stack over the entire study period as a comparison. We applied the median filter using the maximum amplitude of daily CFs to the reference stacks. The comparison of reference stacks over the different periods shows nearly identical waveforms, showing this reference is converged enough to measure the dv/v to the current monthly stack. The distinct improvement in the stacked CFs with the median mute can be found in other station pairs (e.g., EADB-EADB in Supplementary Dataset S2).

To evaluate the time history of dv/v, we stacked the CFs over 30 days with a sliding step of 15 days used as the current CFs. Numerous stacking methods have been proposed to improve the coherent signals in the stacked CFs (e.g., Kanasewich et al., 1973; Schimmel & Paulssen, 1997; Pavlis & Vernon, 2010; Korenaga, 2013; Nakata et al., 2015; Ventosa et al., 2017). In this study, we selected the daily time windows to be stacked with the high correlations between the 12-year reference stack and current CFs to exclude further the time windows that degrade the convergence of the stacked CFs. This selection approach to enhance the Signal-to-noise ratio (S/N) of CFs has been proposed in previous studies (e.g., G. Liu et al., 2009; Olivier et al., 2015; Thangraj & Pulliam, 2021). The metric of selection to improve the stacked CFs varies concerning the purpose of stacking. G. Liu et al. (2009) showed the weighted stack of the common midpoint gathers with the local correlation, and the weights were defined as the correlation coefficient of short-time moving windows. Olivier et al. (2015) used the S/N associated with the S-wave window to enhance the coherence in the stacked CCFs. These metrics aim to improve the S/N of the coherent signals to identify the reflections or the wave arrival times.

In our case, we need to enhance the stability of the stationary phases in the coda part of CFs. Thus, we used Pearson's correlation coefficient to evaluate the similarity between the reference stack and the daily CFs. We computed the monthly stacked CFs using the daily CFs with which the correlation coefficient to the reference is greater than the threshold. This metric is similar to the global correlation defined in G. Liu et al. (2009). The correlation coefficient is evaluated over the entire lag time of CFs, though the similarity of high amplitude signals dominates it. It should be mentioned that the threshold based on the correlation coefficient does not need to be strict because the target phase shifts that are induced by velocity changes decrease the correlation coefficient. Otherwise, it could cause the underestimation of dv/v as the correlation coefficient decreases with the phase change. The purpose of this threshold is to exclude the CFs that show a large discrepancy to the reference stack due to the perturbation of the noise sources so that the measurement of dv/v is assumed to be implausible. We set the threshold to zero with a range of [-1, 1], excluding only the correlations that have shifted as much as it would cancel the stacked CFs. The comparison of the performance for the selective stacking to the other stacking methods can be found in Yang et al. (2022). Although the performance of the selective stack shows similar performance with the linear stack and the robust stack (Pavlis & Vernon, 2010) in their dataset, the selective stacking helps stabilize the current CFs for our dataset as the transient perturbations due to the small events and the environmental noises might remain in the waveform.

2.5 Measurement of dv/v

We estimated the dv/v by the phase difference between the reference and current stacked CFs using two different schemes: the stretching method (Lobkis & Weaver, 2003) and the moving window cross-spectral (MWCS) analysis (Poupinet et al., 1984; Clarke et al., 2011). Using both ways demonstrates the difference in the stability of dv/v measurements (Hadziioannou et al., 2009) and the magnitude of its estimation such that the range of estimated dv/v with the MWCS is more likely to be lower due to the weighted linear regression with the local coherency in the short-time moving window (e.g., Hillers et al., 2019). Note that we did not apply nonlinear filters to the CFs in our analysis, as the signal-to-noise ratio in CFs is enough to measure the dv/v in our dataset.

For the stretching method, we computed the stretch coefficient ε, equivalent to the dv/v, in two steps following Viens et al. (2018). The coda window where we measure the phase change to evaluate dv/v is selected as the following: The start time of the coda window is either set as twice the arrival time of the ballistic wave estimated by the distance of the station pair divided by the wave velocity, or the minimum threshold of 5 seconds if the distance of the stations is close, and for the cases with the ACFs. We assumed the wave velocity to be 1 km/s. The end time of the coda window is fixed at 40 seconds. We applied the threshold associated with the length of the coda window such that it is more than five wave periods related to the central frequency of the frequency band. If the station pair did not meet the threshold, we excluded the station pair for the evaluation of dv/v.

We conducted a grid search to find the best stretching coefficient with a spacing of 0.02% to compute the profile of the correlation coefficient as a function of the ε between the reference and the dilated current CF. We used both positive and negative lags to evaluate the CC. We then applied the spline interpolation of the profile to obtain a finer coefficient estimation following (Viens et al., 2018). Note that the ACFs should have the identical measure of dv/v associated with a pair of cross-components (e.g., BP1-BP2 and BP2-BP1) due to the symmetry of CFs. However, we obtained slightly different estimations in our analysis due to the subtle difference in the spline interpolation. We quantified all the component pairs and confirmed that the difference is mainly within a single spacing of the grid search so that it is not critical to the statistics of the dv/v time history.

For the MWCS method, we followed the process flow described in Clarke et al. (2011). The short moving time window is set as 6 seconds with a step of 3 seconds (i.e., 50% overlap). Like the stretching method, we used both the positive and negative time lags to estimate the best dt/t, that is, the phase shift over lag time. Note that we used the same criteria for the selection of the coda window with the stretching. The dt/t is the slope of the linear trend of phase shift against lag time, which is obtained using the weighted linear regression along the lag time either with or without imposing the intersection at the origin of the lag time. The latter excludes the offset of dt, which helps mitigate the artifacts of instrumental clock drift in the case of the CCFs. We discuss it more in section 6.1.2. We used the results with the later scheme to mitigate the artifacts of the clock drift.

We evaluated the quality of the dv/v estimates with the correlation coefficients between the reference and the stretched CFs with the best fit dv/v value and the estimation of the error derived in Clarke et al. (2011) for the cases with the stretching method and MWCS analysis, respectively. Too large differences between the reference and current CFs may show a much greater magnitude of the dv/v. Still, they may be caused by the source perturbation or instrumental issues rather than the velocity change of the structure. Given the published analyses on Parkfield (Brenguier, Campillo, et al., 2008; Wu et al., 2016; Delorey et al., 2021), we do not anticipate large changes in dv/v. We carefully selected the threshold as they should be soft thresholds such that it does not cause bias due to removing the outliers. We set the threshold of 0.7 and 0.02% associated with the correlation coefficient after stretching and the measurement error in MWCS, respectively, for the following analysis of the dv/v.

Figure 2 shows the number of station-component pair combinations after the thresholding of dv/v. The hundreds of pairs, including auto- and cross-component correlations with the ACFs and CCFs, are obtained to evaluate the time history of dv/v. The number of available pairs decreased due to the decommissioning of the RMNB after 2011. The area highlighted in grey indicates the period where the dv/v is scattered due to the clock drift on EADB and GHIB, discussed in a later section. Overall, we used around 380 station and component pair combinations to evaluate the time history of dv/v. The datasheets of dv/v and the measurement error associated with the stretching and MWCS methods can be found in Supplementary Datasets S3 and S4.

2.6 Channel weighting

To conduct the model fitting, we computed the channel-weighted time history of dv/v associated with all the possible auto- and cross-correlation station pairs. The time history of dv/v needs to be continuous and long enough to fit the models. We thus computed the weighted average of the dv/v associated with the nine components channel correlations following Hobiger et al. (2014), such as

$$\overline{\mathrm{dvv}}(t) = \frac{\sum_{k=1}^{N} c_k^2(t)\mathrm{dvv}_k(t)}{\sum_{k=1}^{N} c_k^2(t)}, \tag{1}$$

where $\overline{\mathrm{dvv}}(t)$ is the weighted average of dv/v, and $c_k(t)$ is the correlation coefficient after dilating the current CFs with the estimated dv/v. The subscript k indicates the component of the channel pairs. N is the number of channels to be averaged. We include the channels only if the quality of the estimated dv/v exceeds the threshold. The weighted average of CC is written as

$$\overline{c}(t) = \frac{\sum_{k=1}^{N} c_k^3(t)}{\sum_{k=1}^{N} c_k^2(t)}. \tag{2}$$

We also evaluate the error of the weighted average of the dv/v with the propagation of error as follows:

$$\overline{\sigma}(t) = \sqrt{\sum_{k=1}^{N} \left(\frac{c_k^2}{\sum_{i=1}^{N} c_i^2} \right)^2 \sigma_k^2}, \tag{3}$$

where σ_k is the error of kth dv/v.

For the case with the stretching method, the CC is computed with the measurement of dv/v. We used the error estimation obtained by Weaver et al. (2011) as σ_k. We used the error derived in Clarke et al. (2011) for the case with the MWCS and simplified the averaging of dv/v by setting the CC to be unity for all the channel pairs.

We computed the fraction of the valid dv/v over the entire period and selected the station pairs with more than 0.7 of the fraction. We removed 33 and 27 pairs from 83 station pairs with this threshold for the cases with stretching and mwcs, respectively. The measurement of dv/v is sometimes unstable due to, for example, cycle skipping (Mikesell et al., 2015), which would cause bias in the model fitting. We thus applied the threshold on the dv/v measurement to ignore absolute variations with amplitude greater than 0.3%, which are assumed to be uncorrelated to the velocity change of the structures, to remove those outliers. The channel-weighted average of dv/v for all the station pairs can be found in Figures S6 and S7, used for the model fitting, as described in section 4.

2.7 Computational effort

The total volume of the original waveform data is 1.6TB. We used a Linux workstation with 48 cores, FASRC Cannon compute cluster (`https://www.rc.fas.harvard.edu`), and Frontera at the Texas Advanced Computing Center (TACC) (Stanzione et al., 2020) to develop the software tools and to conduct the case studies. We parallelized the processes using the Julia-native function `Distributed.pmap` to distribute the tasks to the workers. The heaviest computational process is to compute the cross-correlation due to the frequency of file I/O to access the waveforms and the number of station-component pair combinations. We optimized the parallelization by first computing the FFTs associated with a given time chunk for all the stations and channels in a node. We then distributed the FFTs from the master process to the workers in parallel to obtain the CFs. We submitted the jobs for the cases e.g. every two years, and iterated the time chunks to complete the jobs so that we obtained the CFs for 20 years with reasonable computational time. The other processes, such as downloading data, removing the transient signals, stacking, and evaluating the dv/v are also parallelized using the `pmap`. Further discussion can be found in Supplementary Text S1.

3 Time history of dv/v from 2002 to 2022

We compiled all the dv/v measurements associated with the station-component pair combinations before channel-weighting with the frequency band of 0.9-1.2Hz, selecting for values that meet the threshold of CC or error as shown in Figure 3. The color contour shows the number of pairs within a bin of 0.02% dv/v as a proxy of the probability distribution with respect to the time bins. We plotted the median and the second and third quantiles of the dv/v. The temporal resolution is about a month, as the current CFs consist of the 30-day stack with a step of 15 days. The shaded area indicates the period which could contain the clock drift (see Section 6.1.2). The reference period is from January 2010 to May 2022, as annotated in the bottom arrows.

The coseismic velocity decreases, and the subsequent recovery phase, as shown in Brenguier, Campillo, et al. (2008) and Wu et al. (2016), are reproduced in this study up to 2011. The dv/v continues to recover and exceeds the dv/v level prior to the SS earthquake. This long-term increase is of great interest in this study. The fluctuation of dv/v measurement and the absolute change of dv/v are smaller with the MWCS, similar to the comparison of Hillers et al. (2019).

Data consistency, as measured by the progressive decrease in the range of the quantiles, increases over time. Therefore, the interpretation of physical phenomena acting before the earthquakes (SS and PF) is not likely to be robust, given the lack of stability in the dv/v measurements. In contrast, a robust feature is that the current dv/v would eventually be much greater than before the SS earthquake if the dv/v keeps increasing for more decades without the two coseismic drops in dv/v.

A relatively steep positive change in the dv/v for the case using the stretching method was found around 2014. This could be most likely the artifacts due to the sudden change in the noise power spectrum (see Section 2.1). Given that the rapid increase does not appear in the MWCS measurements, we interpret that there must be a specific scattered wave of higher amplitude that dominated the stretching measurements but that was downweighted in the MWCS calculation. While it appears in 2014, it is not strictly aligned with the date of the Napa earthquake. We explored its relation with the reports of nonstationary rates of Low-Frequency Earthquake occurrence (Delbridge et al., 2020), which we further discuss in Supplementary Text S2. A few station pairs do not include the transition in their channel-weighted dv/v time histories, as shown in Figures S7 and S8. We further discuss the artifacts due to the non-tectonic origins in the section 6.1. We note that correcting this step-like change after 2014 does not remove the long-term trend in dv/v which exceeds the original baseline level prior to the occurrence of either of the SAF earthquakes (e.g. SS).

We analyzed the dv/v time history with different sets of station-component pair combinations. Figure 4 shows the dv/v for the cases with the single-station and cross-stations (i.e., ACFs and CCFs) with the nine-channel correlations and the sets of channel correlations for both ACFs and CCFs associated with the vertical and horizontal components. The long-term trend is nearly identical between the different groups of pairs, except the vertical component of the stretching method shows the relatively quick healing of the dv/v.

Figure 5 shows the dv/v with the different frequency ranges. We used the same length of the coda window and the thresholds in the CC or error with them. We do not show the results associated with the lowest frequency band of 0.2-0.5Hz from the case with the stretching, as the estimation of dv/v was unstable. The coseismic drop in velocity increases with frequency, as seen by Wu et al. (2016). The MWCS measurements in the frequency band 0.2-0.5Hz show neither the coseismic velocity decrease nor the continuous healing of dv/v. We also observed the seasonal perturbations, which need to be excluded to evaluate the dv/v healing. The quality of dv/v measurement at low-frequency ranges might be improved by optimizing the processing parameters, e.g., the stacking period, coda window length, and the short-time window length associated with the MWCS.

Wu et al. (2016) first reported the increase in dv/v drop with seismic frequency in Parkfield and interpreted it as a depth-dependent damaged structure. However, it also comes as a natural decay of the sensitivities for the layered perturbation of the medium with frequencies (Obermann et al., 2013, 2015; Yuan et al., 2021), which can be interpreted as a greater spatial sensitivity (e.g., of the unperturbed medium) with lower frequencies. Interestingly, the rate of long-term increase is also observed to increase with frequencies (Figure S9). Yuan et al. (2021) showed that only uniform (i.e., depth independent) change in velocity affects frequencies equally. Therefore, if the variation of dv/v associated with the seasonal effects is uniform with frequencies, while the rate of long-term increases with them, it has a potential to separate those effects such that the tectonic transient signals and long-term in dv/v are shallow perturbations, and seasonal effects affect a greater depth range.

We use the frequency band of 0.9-1.2Hz to conduct the model fitting and the comparison to the strain in the following sections, corresponding to the depth down to ∼1km considering the depth sensitivity kernel for Rayleigh wave as shown in Figure S10.

4 Multi-factor Model

In order to investigate the source of the observed long-term increase in dv/v we formulate, and fit several different models to decompose the time series of dv/v with respect to each channel-weighted station pair into different the model components which utilize the observed environmental factors near the Parkfield section of the SAF. In California, there have been numerous reports of seasonal effects on shallow measurements of dv/v (Hillers et al., 2015; Clements & Denolle, 2018; Mao et al., 2022; Clements & Denolle, 2023; G. Li & Ben-Zion, 2023), namely hydrological and thermoelastic effects.

We used the monthly precipitation time series recorded at Parkfield with the Remote Automated Weather Stations (RAWS) provided by Western Regional Climate Center. We resampled every 15 days by linear interpolation to synchronize with the time history of dv/v.

The time series of temperature at Parkfield is downloaded from the National Oceanic and Atmospheric Administration (NOAA). We removed the mean offset from the daily averaged temperature time series to obtain the temperature anomaly. We then applied the low-pass filter with a cut-off period of 20 days and downsampled it along the time history of dv/v.

4.1 Base model

To model the dv/v timeseries, the non-tectonic factors such as the precipitation and the thermoelastic strain caused by changes in atmospheric temperature are included. We develop a dv/v timeseries model which is comprised of the non-tectonic factors along with the coseismic decrease in dv/v and the logarithmic healing model proposed by Snieder et al. (2017) which we refer to as "the base model". The base mode is formulated as

$$\begin{aligned} y_{\text{base}}(t) = a_0 + p_1 \Delta\text{GWL}(t, \alpha_0) + p_2 T(t - t_0^{\text{shift}}) \\ + s_1 L(t, \tau_1^{\min}, \tau_1^{\max}, t_{\text{SS}}) + s_2 L(t, \tau_2^{\min}, \tau_2^{\max}, t_{\text{PF}}), \end{aligned}$$

where the parameters with numerical indices (i.e., 0, 1 and 2) indicate parameters which are fit to the dv/v time series and are defined in Table 2.

The first term a_0 is the constant average level of the dv/v time series. The second term describes the hydraulic effects on the dv/v time series and is comprised of a proportionality constant p_1 times the change of groundwater level (ΔGWL). The change in groundwater level is derived from a model of the pore-pressure diffusion of the observed rainfall (Sens-Schönfelder & Wegler, 2006; Akasaka & Nakanishi, 2000), which excludes storage in large spatially confined reservoirs such as aquifers and lakes, such that

$$\Delta\text{GWL}(t_i) = \sum_{n=0}^{i} \frac{p(t_n)}{\phi} \exp\left[-\alpha_0 (t_i - t_n)\right], \tag{4}$$

where $p(t_i)$ is the precipitation at time t_i, ϕ is the porosity, and α_0 is an exponential factor which controls the hydraulic decay rate. The summation over the index n indicates that the value of the change in groundwater level at time t_i is dependent on the previous time steps of the precipitation time history. The ΔGWL time series is then calculated using the RAWS time series of precipitation to estimate $p(t_i)$ and the values of α_0

and ϕ (Figure 6a). Due to the less constraint in α_0 from the dv/v time series, we fixed it during the MCMC analysis described in the section 4.4. We also fixed the porosity ϕ to be 5%, however, any error in the assumed porosity is automatically absorbed by the parameter p_1. The best fitting model factor p_1 is then estimated from the dv/v time series using this change in groundwater level time series $\Delta\mathrm{GWL}(t_i)$. The factor p_1 is constrained to be negative to reflect the physical constraint that the dv/v time series values will decrease for an increase in ΔGWL.

The third term is associated with the atmospheric temperature (Figure 6b). Berger (1975) and Ben-Zion and Leary (1986) described the thermoelastic response at depth due to the slow diffusion of a temperature change at the surface (e.g., with time delay). We simplified this effect by shifting the time history of temperature by t_0^{shift} with the constant factor p_2. Note that we constrain the t_0^{shift} to be positive to meet the causality of the thermoelastic deformation.

The fourth and fifth terms show the coseismic velocity decrease and the healing associated with the SS and PF earthquakes, respectively. The constants of proportionality of s_1 and s_2, which relate the modeled healing to the observed changes in dv/v. We used the logarithmic healing model of Snieder et al. (2017) given by

$$L(t, \tau^{\min}, \tau^{\max}, t_{\mathrm{EQ}}) = \begin{cases} 0, & t < t_{\mathrm{EQ}} \\ -\int_{\tau^{\min}}^{\tau^{\max}} \dfrac{1}{\tau} \exp\left(-(t - t_{\mathrm{EQ}})/\tau\right) d\tau, & t \geq t_{\mathrm{EQ}} \end{cases} \tag{5}$$

where the $\tau^{\min}$ is the minimum relaxation time corresponding to the initial healing rate, $\tau^{\max}$ indicates the period when the healing is completed, and t_{EQ} corresponds to the occurrence time of the earthquake (Figure 6c). The magnitude of coseismic decrease $s_i L(t = 0, \tau_i^{\min}, \tau_i^{\max}, t_{EQ})$ is equivalent to $s_i \ln(\tau_i^{\max}/\tau_i^{\min})$, which can be related directly to the observed coseismic decrease in the dv/v timeseries. We improved the computational efficiency of the numerical integration using the pre-compiled library as described in Supplementary Text S3.

4.2 Model with linear trend

We synthesized the model with a linear trend term such that:

$$y_{\mathrm{wlin}}(t) = y_{\mathrm{base}}(t) + b_0 t, \tag{6}$$

where b_0 is the slope of the linear trend (Figure 6d). Note that the intersection of the linear trend is included in the term a_0 of the base model y_{base}. The linear trend term is applied over the entire study period such that we assume it is caused by the background tectonics or some other factors affecting the dv/v on the entire period. Ikeda and Tsuji (2018) included this linear trend term in their model as they found the long-term increase in the dv/v from the observation near the Nankai trough off the Kii Peninsula, Japan, which is discussed in Section 6.2. Ermert et al. (2023) also included a necessary trend correlated with Mexico City basin subsidence.

4.3 Model with residual healing

We considered an alternative model that includes a residual healing term, which assumes that the healing from other earthquakes which occurred prior to our observational period affects our perceived baseline dv/v values before the SS earthquake. Potential candidates from the USGS catalog are the M6.7 Coalinga earthquake on May 1983 (e.g., Stein & King, 1984), the M4.8, 8 km NW of PF on November 1993 with a source depth of 10.9km, and the M4.9, 3 km NW of PF on December 1994 with the depth of 8.3km. The magnitude of the latter two earthquakes is relatively small, but their location is close to the seismic stations. Thus, if the coseismic decrease in dv/v occurred due to those events, and the healing still took place until the SS earthquake, a residual of the healing could be included in the dv/v. This situation is analogous to the discussion in Vidale and Li (2003) for the interruption of velocity healing due to the 1999 Hector Mine earthquake after the 1992 Landers earthquake.

To formulate this model, we added the term associated with the residual healing to the base model such that:

$$y_{\text{resheal}}(t) = y_{\text{base}}(t) + c_0 H(t_{SS} - t), \tag{7}$$

where c_0 and H are the factors of residual healing and the Heaviside function at the date of the SS earthquake, respectively (Figure 6e).

4.4 Methodology of MCMC analysis

To conduct the model fitting with the time history of channel-weighted dv/v, we used the Python-based software tool *emcee* (Foreman-Mackey et al., 2013). It provides the modules of the Markov chain Monte Carlo (MCMC) method with various advanced sampling algorithms. We selected the one from them called the stretch move proposed by Goodman and Weare (2010), which updates the model parameters using a set of walkers. We set the number of walkers and the steps of iterations as 32 and 20000, respectively. The log-likelihood function with a set of model parameters $\boldsymbol{\theta}$ is defined as:

$$\ln l(\boldsymbol{\theta}) = -\frac{1}{2} \sum_n \left[\frac{\left(\overline{\text{dvv}}(t_n) - y_{\text{model}}(t_n, \boldsymbol{\theta})\right)^2}{\hat{\sigma}_n^2} + \ln \hat{\sigma}_n^2 \right], \tag{8}$$

where dvv is the estimated dv/v time series, and $\hat{\sigma}_n^2 = \sigma_n^2 + f_0^2$ with the σ_n error of the dv/v estimation, and f_0 is a model parameter associated with the additional variance of the dv/v measurement, which is searched during the MCMC analysis similar to the other model parameters. Note that we use a different form of $\hat{\sigma}_n^2$ from what is documented in *emcee* as $\hat{\sigma}_n^2 = \sigma_n^2 + f_0^2 y_{\text{model}}^2$ such that the additional variance is added uniformly over the study period (L. Ermert, pers. comm., 2022). We performed the model inversion using MCMC to evaluate the contribution of each model parameter as well as the trade-off between the parameters. The lower and upper bounds of the model parameters are listed in Table 2.

During the preliminary MCMC parameter search, the largest trade-off is observed between the parameters s_i and τ_i^{min} (i=1, 2), degrading the model's convergence as summarized in Supplementary Text S4 and Figure S11. We thus computed the median values of the maximum likelihood parameter associated with τ_i^{min} for the station pairs in the preliminary case study and used them as representative values to fix the model parameters for the present analysis. We also fixed α_0 as some of the station pairs are less sensitive to the ΔGWL. The values of the fixed parameters are listed in Table 3. Note that the fixed value of α_0 is larger than the other applications (e.g. Sens-Schönfelder & Wegler, 2006), which might indicate higher drainage conditions.

The parameter search range associated with the p_1, s_1, s_2, and c_0 is constrained in the positive or negative side for the consistency with the sense of change in dv/v. The bounds of t_0^{shift} are set to be up to 90 days due to the trade-off between the p_2 and the days of shift such that the time shift of half a year with the flip of the sign shows almost identical time series considering the seasonal variation. We also set a threshold such that the magnitude of s_1 associated with the SS earthquake to be smaller than half of s_2 with the PF earthquake. The more adequate criteria could be the comparison with the coseismic velocity decrease in the logarithmic healing model, i.e., $s_i \ln(\tau_i^{max}/\tau_i^{min})$, whereas we used the former criteria to simplify the constraint condition.

4.5 Comparison of best likelihood model to the data

We compared the model fitting associated with the base and linear term models for all the available station pairs. We performed the statistical analysis to show the contribution of the linear trend term in the dv/v as well as the distribution of the other model parameters. We also conducted the MCMC analysis using the model associated with the residual healing term for a representative station pair to investigate its role in the dv/v.

Figure 7 shows the marginalized 1D and 2D posterior probability distributions of the model parameters fit using the linear model to the dv/v time series obtained using the stretching method for the channel-weighted pair of LCCB-SCYB. The marginalized 1D posterior probability distributions of all the model parameters except $\tau_1^{\max}$ are strongly peaked and unimodal indicating that each of the parameters, and their error estimates, are well constrained (grey histograms, Figure 7). The 2D marginalized posterior probability distributions associated with $\tau_1^{\max}$ (red 2D histograms, Figure 7) reveal a trade-off between its value and the value of each of the other parameters (i.e., much larger covariances). The parameter $\tau_1^{\max}$ determines the period of healing following the SS earthquake and is not well constrained due to the short observational period $T_{\mathrm{SS_{obs}}}$ between the occurrence of the SS and PF earthquakes (i.e. $T_{\mathrm{SS_{obs}}} \ll \tau_1^{\max}$). We note that the 1D marginalized posterior probability distributions indicate that the parameters s_2, $\tau_2^{\max}$ and b_0 are well constrained, despite the trade-off between parameters pairs (e.g., s_2-to-$\tau_2^{\max}$, s_2-to-b_0, and $\tau_2^{\max}$-to-b_0) indicated by their 2D marginalized posterior probability distributions. We omitted to show the uncertainty f_0 as it shows the unimodal distribution, not interfering with the other model parameters.

The result obtained with the MWCS derived dv/v time series (Figure 8) are similar to those obtained using the stretching method. The main difference is in the parameter t_0^{shift} which is shifted from ~ 30 days to near zero. This discrepancy between the stretching and MWCS results is likely due to the weighting of linear regression in the MWCS which could decrease the sensitivity to large fluctuations of phases from temperatures.

Figure 9 compares the observed dv/v time series and the model with the set of best likelihood parameters. Each row shows the contributions of model factors to the dv/v and the preprocessed time series of precipitation and temperature. The models of dv/v are synthesized using the maximum likelihood parameters, which are the set of parameters that achieve the best likelihood during the MCMC sampling. For the station pair LCCB-SCYB, the best fitting models to both the stretching and MWCS derived channel-weighted dv/v timeseries show that the long-term increase after the PF earthquake is better reproduced with the linear trend term (Figures 9 and 10).

4.6 Statistical analysis of model parameters

We estimated the best likelihood parameters for the available station pairs, which can be found in Supplementary Dataset S5. We show the model fitting of all the station pairs in Figures S6 and S7 for the cases with the stretching and MWCS methods, respectively. We quantified the convergence of the model using the variance of the residuals such that we selected the station pairs with the variance of residuals smaller than the threshold of 0.002.

We also removed the station pairs and ACFs associated with the FROB, as it shows the strong, suspicious harmonic noise in the raw data, which can cause bias in the CFs. We finally used the 32 and 29 channel-weighted station pairs for the cases with the stretching and MWCS, respectively, for the statistical analysis of model parameters.

We computed the two metrics to evaluate the quality of models: the Akaike information criterion (AIC, Akaike, 1974)

$$\mathrm{AIC} = N \ln \frac{1}{N} \sum_{t} \left(\overline{\mathrm{dvv}}(t) - y_{\mathrm{model}} \right)^2 + 2k, \tag{9}$$

and the Bayesian information criterion (BIC, Schwarz, 1978)

$$\mathrm{BIC} = N \ln \frac{1}{N} \sum_{t} \left(\overline{\mathrm{dvv}}(t) - y_{\mathrm{model}} \right)^2 + k \ln N, \tag{10}$$

where N is the number of data, and k is the number of model parameters.

Figure 11a compares the AIC and BIC values for the case with the two methods of dv/v measurement. $\Delta\mathrm{AIC} = \mathrm{AIC}_{\text{linear trend}} - \mathrm{AIC}_{\text{base}}$, and same for ΔBIC. Most of the station pair shows the negative ΔAIC and ΔBIC, which supports the additional model term and model complexity. Thus, the linear trend term shows a non-negligible contribution to the dv/v. Figure 11b shows the estimated value of the slope b_0%/year. The median for the cases with the stretching and MWCS are 0.0048 and 0.0027%/year, respectively.

Figure 11c summarizes the other model parameters. The a_0 is estimated larger with the base model than the model with linear trend term. This is due to the complementary of the long-term healing by the positively large offset as shown in Figures 9 and 10.

The coefficient with the temperature p_2 varies in both negative and positive values. In contrast, the median is positive, indicating the increase of temperature corresponds to the rise in dv/v, which is consistent with the model of Richter et al. (2014). The negative sensitivity of dv/v to the temperature could be caused by the artifacts of the model fitting to the station pairs, which are less sensitive to the temperature, while it remains to be identified. G. Li and Ben-Zion (2023) showed the seasonal variation of the velocity change ranging from 0.4% to 1.2%, which is more significant than our case. This would be caused by the different frequency ranges– 1-6Hz in their case–, while we use 0.9-1.2Hz.

Maximum healing times $\tau_i^{\max}$ are better constrained with the base model than the model with linear trend, which could be caused by the trade-off between $\tau_i^{\max}$ and b_0 (Figures 7 and 8). However, the residuals between the model and data are smaller, with the case using the linear trend term for most of the station pairs. Instead, we need to subtract the long-term increase from the dv/v, if it is plausible, to adequately evaluate the maximum time of the logarithmic healing after the earthquakes. We generally find $\tau_2^{\max}$ to be about ten years, a similar scale to that obtained by Clements and Denolle (2023) in southern California earthquakes.

The other parameters, such as p_1 and t_0^{shift}, are more scattered when incorporating the linear trend term. These parameters may be less constrained due to their small sensitivity to the dv/v: the effects of precipitation and temperature are not evident in the channel-weighted time history of dv/v in Parkfield (e.g., the seasonality is weak) compared to the other areas (Sens-Schönfelder & Wegler, 2006; Clements & Denolle, 2023).

In summary, the estimation of model parameters associated with non-tectonic factors seems poorly constrained as some of the station pairs exhibit less seasonal signals (e.g., low sensitivity of dv/v to these factors). The posterior distributions of model parameters are narrower for the base model (no linear trend), whereas the addition of the trend term widens the posterior distribution, even if it is well justified by the negative ΔAIC and ΔBIC.

4.7 Model fitting with the residual healing

We performed the model fitting with the residual healing term for the LCCB-SCYB as a representative pair of dv/v time history showing the logarithmic healing and the long-term increase. Figure S12 shows the marginalized posterior probability distributions for the case with the residual healing model using the MWCS method. The residual healing coefficient c_0 shows the trade-off with a_0, the static term, such that the a_0 increases with the magnitude of c_0 to fit with the period before SS. Figure 12 shows the contributions of the model factors and the comparison to the data. We found $c = -0.05\%$ of the residual healing before the SS earthquake and $\tau_2^{\max} = 25$ years in this station pair. The fitting of the model could reproduce the channel-weighted dv/v by complementing the long-term increase in dv/v with the combination of c_0, a_0, and the logarithmic healing.

The AIC and BIC are smaller in the case with the base model, whereas they are larger in the case with the linear trend term shown in Figure 10 suggesting that the model with linear trend term is preferred over the residual healing for this station pair. However, the comparison of the cases with the other station pairs might show different results, and we eventually cannot evaluate how much the comparison of the AIC and BIC between the models with linear trend and residual healing terms is statistically significant as both could reproduce the long-term increase in the dv/v. Thus, the choice of models with the linear trend or the residual healing terms remains to be determined from the model fitting analysis. Instead, if the residual healing model is suitable, we could deduce that the steady state of the velocity corresponding to the a_0 is larger than the current state of dv/v, and the logarithmic healing takes place at least more than 18 years.

4.8 Variation of the linear term with fault normal distance

We analyzed the slope of the linear trend term b_0 as a function of the fault normal distance to investigate the spatial characteristics of the long-term increase in dv/v. We computed the fault normal distance of the stations by defining the approximated planar fault along the San Andreas Fault, as shown in Figure S1. We averaged the distances for the two stations in the cases of CCFs. The b_0 is obtained with the best likelihood parameters.

Figure S13 shows the distribution of b_0 with the fault normal distances. We categorized the ACFs and CCFs in the Pacific and North American sides and the station pairs crossing the fault, following Malagnini et al. (2019); Delorey et al. (2021). The value of b_0 is generally larger with the stretching method than the MWCS, which could be caused due to the characteristics of the methodology in the measurement of the dv/v as the MWCS is more likely to underestimate the magnitude of dv/v (e.g. Hillers et al., 2019).

Except for those outliers described below, the distribution of b_0 is relatively uniform, or weakly anti-correlated, with the fault normal distance. The effect of the shear localization or the damage zone in the fault core, as inferred by Delorey et al. (2021), is not clearly apparent in b_0, even though some station pairs near the fault show relatively large values. To investigate if the observed dv/v variations revealed by the variation in b_0 with fault distance reflects the velocity perturbation in and around the fault core the sensitivity kernels would need to be evaluated carefully (Obermann et al., 2013). Some station pairs near the fault show larger values in b_0 relative to station pairs far from the fault (Figure S13), however, the b_0 with GHIB would be infeasible as the associated time history of dv/v is unstable with larger variances than the other stable station pairs (Figures S7 and S8). The negative b_0 obtained for station pair LCCB-VCAB indicates a long-term decrease which is not visible in the dv/v time history (Figure S7). This negative b_0 value is likely caused by the trade-off in the model fitting between b_0 and the healing term of the PF earthquake rather than the negative trend in dv/v. Note that the slope for this station pair is only obtained for the stretching method and a positive value is estimated using the MWCS derived dv/t time series.

5 Cumulative strain field at Parkfield

In order to investigate whether the origin of the long-term increase in dv/v is associated with the accumulation of regional strain on the SAF near Parkfield, we estimate the regional strain at Parkfield from 2009 to 2022 using GNSS data.

The dilatational strain in rock is thought to affect the observed dv/v time series by altering the average wave speeds of the rock with contraction and extension increasing and decreasing the velocities, respectively. This relation is formulated in nonlinear elasticity (Ostrovsky & Johnson, 2001) and is supported by observations (e.g. Yamamura et al., 2003; Brenguier, Shapiro, et al., 2008; Richter et al., 2014; Donaldson et al., 2017). Typically, the sum of axial strain components, i.e., dilation, correlates with dv/v observed with the ambient seismic noise (Donaldson et al., 2019). However, Hotovec-Ellis et al. (2022) showed the velocity can be sensitive to a single component of the strain rather than the dilation when the pre-existing cracks are oriented in the suitable direction. In the case of Hotovec-Ellis et al. (2022), the cracks are generated perpendicular to the radial direction of a caldera, forming the vertical ring fractures such that the opening and closing of them are governed by the radial strain component. Thus, if the cracks are aligned in a preferable orientation, the velocity could be effectively changed with the extension or contraction of the axial strain (Hotovec-Ellis et al., 2022).

The dv/v should be evaluated with the spatial average of the interaction between the dominant crack orientations and the cumulative axial strain as well as the efficiency of the crack opening and closing with the strain (Sayers & Kachanov, 1995). In this study, however, we computed the spatial average of azimuthal axial strains around the seismic station to investigate the contractional directions of the strain rather than encompassing all the spatial variation of scattered crack orientations with depth-varied maximum compressive stress S_{Hmax} (Hickman & Zoback, 2004), which is less feasible in the resolution of the observations and the uncertainty of the cause of anisotropy.

5.1 Processing of the GNSS data to estimate the temporal evolution of strain field

We downloaded the GNSS data processed by the NASA MEaSUREs ESESES Project (Bock et al., 2021). We used the daily displacement time series in the region of western North America, where the outliers and non-tectonic jumps have been removed.

We computed the in-plane strain with the triangular elements with the nodes of GNSS stations generated by the Delaunay triangulation. The deformation of an element caused due to the plane strain without the rigid motion is written as follows:

$$\begin{bmatrix} u_x^1 \\ u_y^1 \\ u_x^2 \\ u_y^2 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 0 & y_1 \\ 0 & x_1 & y_1 & -x_1 \\ x_2 & y_2 & 0 & y_2 \\ 0 & x_2 & y_2 & -x_2 \end{bmatrix} \begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{xy} \\ \varepsilon_{yy} \\ -\omega \end{bmatrix}, \tag{11}$$

where u_x^k and u_y^k (k=1,2) are the displacements of the station. x_k and y_k are the relative locations of nodes from the origin that is chosen from one of the nodes in the element. ε_{ij} and ω are the strain and the rotation, respectively. This representation of deformation can be found in Shen et al. (1996) and Crowell et al. (2013). Given the displacement of stations, we obtain the strain tensor, which is assumed to be constant in the element. The convention of the sign is positive in extension for the axial strain components and the dilation. We used the software tool PyTAGS (Crowell, 2019) to conduct the mesh discretization with triangular elements and compute the strain components associated with the elements.

We first obtained the strain components oriented to the East and North, corresponding to the x and y directions in equation 11, respectively. We then computed the dilation $\varepsilon_1+\varepsilon_2$ and the max shear $(\varepsilon_1 - \varepsilon_2)/2$ of the element, where the ε_1 and ε_2 are the maximum and minimum principal strains, respectively.

To investigate the accumulation of strains, we first computed the deformation of the elements at a given time snapshot using the relative motions of the GNSS stations from the reference configuration. We then subtracted the strain field at the initial strain field, which we selected on January 3, 2009, from that of the current time snapshot to evaluate the accumulation of the strain from the initial state. The cumulative strain shows either a monotonic increase or decrease over the analyzed period as the sense of accumulation in the strain, i.e. extension or compression, rarely changes.

Figure 13 shows the temporal evolution of the dilation and max shear. The max shear is positive and localized on the fault, while the dilation is not uniformly distributed along the fault. It should be noted that the discontinuity (i.e., slip) crossing the element causes bias in estimating strain as the element is assumed to be a continuum, and the strain is constant within the element. Therefore, the estimated strain near the fault can be apparent, largely distorted by the slip on the fault. One of the potential causes for the intricate pattern in the dilation could be the localized slip on the shallow surface (Bacques et al., 2018).

The magnitude of cumulative strain far from the fault is in the order of $1\mu\varepsilon$ over ten years. Considering the macroscopic strain rate reflects the region far from the fault, the strain rate should be in the order of $0.1\mu\varepsilon$/year, which is consistent with other studies (Klein et al., 2019; Ross et al., 2022), although the detailed strain distribution cannot be compared due to the different spatial resolutions of the aforementioned studies.

5.2 Time history of cumulative strain with Gaussian-weighted average around the seismic station

To evaluate the sensitivity of velocity change to the strain by comparing the long-term increase in the dv/v with the cumulative strain associated with the dilation and the max shear, we computed the Gaussian-weighted average of the strain field associated with the seismic stations. We only used the channel-weighted dv/v associated with the ACFs. We estimated the time history of the cumulative strain using the weighted average of the strain snapshots. The weight is computed with the numerical integration of two-dimensional Gaussian distribution over the surface of the triangular element as follows:

$$w_i = \int_{\Omega_i} G(\boldsymbol{\xi}) dS, \tag{12}$$

$$G(\boldsymbol{\xi}) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|\boldsymbol{\xi}|^2}{2\sigma^2}\right),$$

where w_i is the weight on the ith triangular element, Ω_i is the surface of the triangular element, and $\boldsymbol{\xi}$ is the relative coordinate from given seismic station. σ defines the spatial extent of the weight around the seismic station. We set σ=5km. We normalized the weights before computing the weighted average such that the sum of the weights is unity for the cases where the strain data is missing on some triangular elements due to the lack of GNSS data.

We applied the threshold of the missing data of GNSS near the seismic station to exclude the case if the cumulative strain is estimated with the triangular elements only far from the station. We skipped the time step of cumulative strain if the triangular elements within the distance of 5km were missing due to the gap in GNSS data.

5.3 Comparison of long-term increase in dv/v to the cumulative strains

We pre-processed the channel-weighted dv/v before comparing them with the time history of the cumulative strains. We subtracted the model components of ΔGWL and temperature with the best likelihood model parameters from the dv/v and removed the offset with the mean between 2008 and 2010, considering the reference time of the cumulative strain. Indeed, we could also remove the logarithmic healing terms from the dv/v to focus on the linear trend. However, the $\tau_2^{\max}$ shows the trade-off with the b_0 as shown in Figures 7 and 8, which could cause the overestimation of the long-term increase in dv/v. Therefore, the logarithmic healing terms are kept in the dv/v to avoid bias due to subtracting them. To compare the strain and dv/v, we excluded the stations of GHIB, JCSB, and VARB as the measurement of the dv/v is unstable compared to that of other stations (see Figures S6 and S7). Therefore, we used the seven stations (CCRB, EADB, LCCB, MMNB, SCYB, SMNB, and VCAB) for the following analysis. We also set the period of comparison until 2020, as many GNSS data are missing after that.

Figure 14 compares the dilation and max shear to the dv/v of ACFs for the cases using stretching and MWCS, respectively. We compared the cumulative strain and the dv/v associated with the ACFs of the seven stations, where the estimations of dv/v were relatively stable from 2009 to 2020. The time steps are synchronized between the cumulative strain and the dv/v, and we compared them every 90 days from the reference time snap.

The Gaussian-weighted time history of the cumulative dilation around the seismic stations shows the subtle extension with time on average of the seven stations. The comparison with the dilation shows weak positive correlations to the dv/v, which is the opposite relation of what one would expect from nonlinear elasticity or from the opening of microfracture in micromechanics. Note that the dilation associated with EADB shows slight contractional accumulation, whereas the magnitude of strain is smaller than the other stations, which would be insufficient to evaluate the sensitivity with the dv/v. The similar discrepancy in the sense of dv/v and the dilation is also addressed in Hotovec-Ellis et al. (2022).

We thus focused on the azimuthal axial strain and computed their Gaussian-weighted average around the stations. Instead of summing up all the local strains along S_{Hmax}, whose orientation varies quite significantly with a range of ±45° as shown in Figure 13, we uniformly rotated the strain around the stations to obtain the first-order evaluation of the axial strain to investigate if it shows contraction to cause the closing of given aligned cracks, which results in the increase in the dv/v.

Figure 15a shows the mean and standard deviation of the station-averaged axial cumulative strain as a function of azimuth. The cumulative strain is evaluated between the reference in January 2009 and November 2019. The estimation of maximum contractional orientation inferred from the GNSS data is consistent with the range of S_{Hmax} estimated by Hickman and Zoback (2004). It is notable that the azimuthal strain can be contraction in the range from N35°W to N45°E although the mean value of axial strain, which is equivalent to the half of the dilation $\varepsilon_1 + \varepsilon_2$, shows the extension.

We performed a linear regression between the cumulative strain and the dv/v averaged over seven stations to obtain the sensitivity at each azimuth, which is shown in Figure 15b. Figure 15c shows the estimations of negative sensitivity evaluated at N5°E, which are -0.011±0.001%/$\mu\varepsilon$ and -0.007±0.001%/$\mu\varepsilon$ for the cases with stretching and MWCS, respectively.

The sensitivity of dv/v to the strain component is smaller than reported previously as generally referred to in the order of -(0.1-1.0)%/$\mu\varepsilon$. Although most studies derive the strain sensitivity to the dilatational strain rather than a single extension strain component, and the condition of the tectonic setting, the frequency range of the CFs, and the installation of seismometers vary, we can still compare with other values found in the literature, which is summarized in Supplementary Text S5.

Further investigation is necessary to fully explain the magnitude and direction of sensitivity of dv/v to strain. Nonetheless, the preferable interactions between the aligned microcracks, which could be non-uniformly distributed at Parkfield as inferred from the shear wave splitting analysis (see Supplementary Text S6), and the contractional strains could be a candidate to explain the long-term increase in dv/v on the multiple seismic station pairs.

6 Discussion

6.1 Potential artifacts in dv/v

In this section, we list the potential artifacts to the dv/v caused by instrumental issues rather than tectonic or environmental origins before discussing the factors associated with the long-term increase in dv/v.

6.1.1 Change of noise source spectrum

Zhan et al. (2013) shows the potential of the apparent change in the velocity due to the evolution of the noise source spectrum. This artifact is still controversial as Mao et al. (2020) conducted the numerical experiments of the coda interferometry with multiply scattered waves using the noise sources with different source spectra, indicating the measurements of dv/v are relatively insensitive to the perturbation of the noise source spectrum.

Over time, sensors get updated, and instrumental responses may change. We find an increase in the power spectrum at higher frequencies around 2014, possibly due to a change in the sensitivity of stations as discussed in sections 2.1 and 3. This change appears suddenly, and it might be correlated to the bump of the dv/v shown in the case of stretching in Figure 3, but it cannot explain a long-term trend.

In our case, if the noise source perturbation is the origin of the apparent long-term increase in the dv/v, the noise source spectrum should gradually change over 20 years, uniformly causing the long-term increase in dv/v on the difference station and component pairs. We remark that the seasonality of microseismic noise is not strong at the frequencies of interest in our study around 1Hz, as shown in the supplementary materials associated with the PSD. We need to investigate if the variation of source spectra can cause a uniform change in the measurement of dv/v over multiple station pairs.

6.1.2 Clock drift

Stehly et al. (2007) pointed out that clock drift causes the artifacts in the travel time estimation obtained with the stacked CFs. Various methods have been proposed to correct the time shift (e.g., Sens-Schönfelder, 2008; Gouédard et al., 2014; Hable et al., 2018; Hirose & Ueda, 2023). In our dataset, we found the time shift in daily CFs associated with the EADB around 2017 in the plot of CCFs (see the CCFs for EADB-SCYB in Supplementary dataset S2), although it seems to be corrected within the same year. We also found the transient time shift in CCFs associated with the GHIB around 2016.

We can remove the bias of the clock shifts to the dv/v measurement by correcting them as implemented in e.g., Brenguier, Campillo, et al. (2008). Instead of applying the proposed correction method to our CFs, we used the MWCS such that we conducted the linear regression without imposing the intersect passing the origin associated with the lag time and the time delay to mitigate the artifacts due to the clock drift. This metric is also implemented in the MSNoise (Lecocq et al., 2014). The temporally shifted

CFs cause the artifacts in the cases using the stretching method, as the origin of the stretching should be corrected. The dv/v using the MWCS with imposing the intersect crossing the origin is also biased due to forcing the linear regression to the time delay with non-zero intersect. On the contrary, we could minimize the effect of clock drift without imposing the intersect crossing at the origin with MWCS because the clock drift causes the shift of the intersect but does not change the slope of the time delays, reflecting the structural change of the velocity (Gouédard et al., 2014), which we used in this study. Moreover, it is worth noting that the clock would not influence the results with the ACFs drift as the rate of time shift should be negligible compared to the coda length even if the clock drift occurs. Thus, the clock drift would not be the dominant factor to characterize the linear trend shown in the model fitting.

It should also be noted that the distinctive decrease in the dv/v is indeed observed in the case with both stretching and MWCS around 2017, whereas it also coincides with the large precipitation as shown in Figures 9 and 10. Therefore, it can reflect the velocity change of the structure rather than purely the artifacts of the crock drift.

6.1.3 Phase shift in the data sampling

Before replacing the location codes and channel names around 2011, we noticed the sampling time of the raw waveform was not aligned on a uniform time vector, i.e., 0, 0.05, 0.1,... [sec]. To compute the correlation functions, we shifted, rather than resampled, the waveform to the closest sampling point on a uniform time vector. According to our tests using both `obspy` and `SeisIO` download functions, the value of the phase shift seems arbitrary with each file, though it is always less than half of the sampling rate (25ms). Thus, the dv/v measurements using the stretching method could be biased by the phase shift. However, similar to the discussion about clock drift, the artifacts due to the phase shift are mitigated in the cases with the MWCS. The phase shift of the dataset has been corrected after replacing the location codes and channels so that the artifact is not the case for the later part of the dv/v, from November 2010 to September 2011.

We note that some of the daily stacked CCFs lack the first half slice of the short-time cross-correlation windows with a length of half an hour due to the fluctuation of the initial time of the data samplings in the present analysis. While this needs to be corrected for the completeness of the study, it would not primarily affect the daily and monthly stacking of the CCFs.

6.1.4 Change in the instrumental response

We must be careful about the long-term changes of instrumental responses as we discuss the dv/v over decades. Ueno et al. (2015) showed the pseudo perturbation of the CFs caused by changes in the instrumental response, which has led to biases in the estimation of dv/v. The instrumental response available at the NCEDC is the best sensor and digitizer metadata knowledge. Reassessing this response may give sudden sensitivities changes, as seen in Figure S2. However, we cannot correct the minor change in the response due to site conditions, such as sensor-to-ground coupling changes or the instruments' aging.

The large fluctuation in the instrumental response is more likely to be thresholded out by the median filter applied during the computation of CFs or the quality of the dv/v estimation during the post-processing. In contrast, the minor change cannot be corrected and should be included in the analysis. However, identical to the discussion of the noise source perturbation, if the difference in the instrumental response causes the long-term increase, the CFs should be distorted such that the changes in the time delay consistently cause the increase in the dv/v over the different station and channel pairs. We would need to investigate instrumental aging when evaluating the long-term dv/v over decades.

6.2 Potential factors to explain the long-term increase in dv/v

While the linear trend would appear due to non-tectonic factors or instrumental issues, an accumulation of the axial strain can close the aligned and open micro-cracks in the medium, causing a positive dv/v. However, the contribution of the latter tectonic factor associated with the cumulative strain is still debatable. Our analysis did not pursue a 2D, improved spatial imaging of dv/v (Obermann et al., 2013) and our crude knowledge of micro-cracks orientation in the fault zone. Besides, the estimated strain field is limited to the two-dimensional horizontal components; the vertical variation of the strain is not considered in the analysis. Therefore, we cannot quantify the sensitivity of dv/v to strain. On the other hand, if the given pattern of regional strain accumulation and the aligned cracks cause the long-term increase in dv/v, we could explain the linear trend shown in dv/v associated with the multiple station-component pair combinations for both ACFs and CCFs. Further, the fact that our measurements of shallow and near fault-zone dv/v exhibit a long-term compression may come from the fact that the 2004 Parkfield earthquake did not rupture at the surface; therefore, the shallow portion may experience partial loading, even if small slow creep may reduce the loading rate (Bacques et al., 2018).

Several studies have found long-term increases in dv/v in the accretionary prism of the Nankai subduction zone. Ikeda and Tsuji (2018) found a linear trend in dv/v near the Nankai trench with the range of 0.01-0.03%/year during a few years, which was extended by Tonegawa et al. (2022) over a decade. These linear trends contrasted with the western part of the accretionary wedge (onshore, further from the plate boundary), where no increase was measured. Authors interpreted such variations in terms of variable compression and fluid drainage from the subseafloor. Offshore observations are different than Parkfield's case, especially in geohydrology, lithology, and microcrack structure. Nevertheless, we draw an analogy for velocity perturbations in the interseismic period near a plate boundary.

The model of residual healing is another candidate to reproduce the long-term trend in the dv/v. Indeed, the quality of the measurement of dv/v is lower before the SS earthquake, as shown by the significant variance in the estimation of the dv/v in Figure 3. The change of preamplifier gain documented in the supplementary material of Brenguier, Campillo, et al. (2008) can be one of the reasons for the change of the variation. Therefore, the duration of the analysis period and the quality of the measurement in dv/v are insufficient to discuss if the healing had occurred before the SS earthquake. The difficulty in quantifying the contribution of residual healing is that the logarithmic healing due to the slow dynamics can be accompanied by the other factors for the long-term trend described above, causing the trade-off. The laboratory experiments with the servo-control of the shear loading rate (e.g., Shreedharan et al., 2021) would have the potential to isolate the logarithmic healing due to the slow dynamics to evaluate the other factors contributing to the dv/v.

Lecocq et al. (2017) monitored dv/v, focusing on the thermoelastic and the hydrological effects over 30 years in the region without major earthquakes or volcanic and geothermal activities. They found a long-term increase in the dv/v coincides with the change of thermoelastic strain induced by the air temperature rise. Thus, the long-term change of thermoelastic deformation can be a candidate for the potential factor of the increase in dv/v. The rate of growth in dv/v, which is equivalent to the slope of b_0 in our model, was estimated as $\sim$0.001%/year in (Lecocq et al., 2017), while we obtained the higher rates of 0.0048%/year and 0.0027%/year for the cases with stretching and MWCS, respectively. It should be noted that the air temperature cannot be the dominant factor of the long-term increase in our case; the long-term increase in the air temperature is $\sim$0.5 °C for 20 years, while its seasonal variation is $\pm$10°C. If the air temperature causes the increase of dv/v with 0.05% for 20 years, the range of seasonal variation should be $\sim$1%, which is not observed in our measurement. Thus, if the long-term increase in dv/v is caused by thermoelastic deformation, another candidate could be the continuous change of in-situ background temperature at depth.

The relative locations between sensors and seismic scatterers can be altered by accumulating the strains or the slip on the fault. Snieder et al. (2002) showed the averaged travel time over multiply-scattered waves in the coda window is not influenced by the independently perturbed scatters or the shift of source locations. If this is the case in our analysis, the measurements of dv/v only reflect the velocity change of the medium. However, the conditions associated with the number of scatterers and the frequency ranges with our analysis could differ from the case where the averaged time delay is canceled out. We thus further need to investigate if the long-term shift in the seismic stations and scatters can cause the apparent dv/v.

The dynamic earthquake ruptures can generate coseismic damage in the medium around the fault (Andrews, 2005), which forms the flower-like structure of the damage zone with depth; that is, the damage zone area is wider at shallow depth with lower confining pressure, and it becomes narrower at depth (Ben-Zion et al., 2003; Ma, 2008; Y.-G. Li & Malin, 2008). The orientation of off-fault fractures is governed by the stress field around the crack tip (Poliakov et al., 2002; Rice et al., 2005). The angle of the off-fault tensile crack varies with the rupture velocity (Griffith et al., 2009), whereas they are activated near parallel to the maximum compressive horizontal stress to the fault (Yamashita, 2000; Thomas & Bhat, 2018; Okubo et al., 2019). Therefore, the cumulative contractional strain parallel to the S_{Hmax} would not effectively close those cracks. However, the area of the coseismic damage zone would be smaller than the spatial extent of the scattered coda wave. The rupture of the PF earthquake caused a small amount of slip near the surface (Johanson et al., 2006), whereas given the S_{Hmax} in the order of 100MPa at 1-2km depth (Hickman & Zoback, 2004), the damage zone width is estimated in the order of hundreds of meters around the fault plane (Okubo et al., 2019). Besides, Y.-G. Li and Malin (2008) also showed the damage zone width is $\sim$200m at SAF inferred from the fault-guided wave. Hence, the orientation of coseismic tensile cracks would not be dominant in the kilometric extent of the sensitivity associated with the scattered wave.

Notably, the coseismic change in the anisotropy during the PF earthquake has been investigated by Durand et al. (2011) using the rotation of correlation tensor. They showed the substantial rotation after the PF earthquake at station VARB and pointed out that opening and closing of the cracks in the upper layer of $\sim$3km would play a role in the observed rotation. Löer et al. (2018) also explored the temporal stability of the anisotropy with beamforming analysis using the ambient seismic noise at Parkfield. While we focused on the long-term trend of the dv/v in the present study, the time history of the anisotropy over decades should also attract research interest.

The weighted average of strain could be largely biased due to the slip on the fault causing the apparent strain change within the triangular element. Thus, the evaluation of dilation shows a primitive estimation around the stations. Nevertheless, it was insufficient to explain the long-term increase of dv/v in the context of compression with the isotropic medium. We thus investigated the accumulation of axial strain as an alternative candidate associated with the strain and the change of dv/v. The distribution of crack orientations cannot be simply retrieved from the fast polarization of shear wave splitting as it can be caused by various effects, such as the aligned minerals and grains rather than the cracks. However, the scattered polarization directions revealed by Y. Liu et al. (2008) would indicate the distribution of cracks is likely to be non-uniform around the fault. The local velocity change would be determined by the interactions of the preexisting cracks and the cumulative strain, and the dv/v reflects their spatial integration with the sensitivity kernels (Obermann et al., 2013). Those mechanisms can potentially cause a long-term increase observed in the dv/v time history.

7 Conclusion

We monitored dv/v from 2002 to 2022 to investigate the temporal changes near the Parkfield Region of the San Andreas Fault and constrain the evolution of the San Andreas Fault throughout the interseismic period. Our analysis revealed that the time history of dv/v was modulated by multiple factors which affected the physical state of the fault and the surrounding subsurface which can be separated into two distinct groups: environmental impacts (e.g., temperature variations and changes in the groundwater level from precipitation) and tectonic phenomena (e.g., earthquakes and inter-seismic loading). The effect of both the San Simeon and Parkfield earthquakes on the dv/v time series is clearly observable with a similar level of previously reported values (Brenguier, Campillo, et al., 2008; Wu et al., 2016). However, the larger observational period used in this study, as well as the improved quality of the dv/v measurement from using all 3 components, allows for us to robustly determine the influence of additional and more subtle environmental and tectonic factors.

In order to achieve this enhanced quality and longer duration of the dv/v time history which enable statistically robust measurements of the role of inter-seismic tectonic loading, we have developed a Julia-based software tool to efficiently compute all the auto- and cross-correlations of the 3 components, of the 13 stations in the High Resolution Seismic Network operated by the UC Berkeley Seismological Laboratory, which provided around 380 station-component pair combinations with a good quality to evaluate dv/v over the study period. This software package `SeisMonitoring.jl` (doi:10.5281/zenodo.832094), which enables efficient process parallelization and greatly reduces the processing time, is openly available and will enable the calculation of dv/v for previously prohibitively expensive computationally intensive processes.

Following the 2004 Parkfield Earthquake the dv/v heals logarithmically for almost all of the station pairs and shows a net long-term increase in which the current dv/v level is equivalent to, or exceeding, the state before the 2003 San Simeon earthquake. This long-term increase is investigated using models which are fit to the derived channel-weighted dv/v time series. The full posterior probability distributions for these are determined using a MCMC sampling. In order to isolate the effect of the inter-seismic tectonic loading we account for environmental effects such as the temperature and precipitation, as well as the the logarithmically healing model from the two earthquakes. Models with include a long-term linear trend term or residual healing term provide a statistically robust improved fit to the data. Both the AIC and BIC metrics confirm that the long-term trend observed in the dv/v time history is a non-negligible factor, and our analysis prefers the inclusion of either the long-term trend, which we interpret to be associated with the inter-seismic tectonic loading, or the residual healing such that the logarithmic healing has not been completed yet. Whatever factor controls this long-term trend should be spatially uniform near the fault, as this increase is observed in multiple station-component pair combinations.

To investigate the role of inter-seismic tectonic loading on the long-term increase of dv/v, we analyzed the GNSS-derived strain from 2009 to 2020 in the Parkfield Region of the San Andreas Fault. The spatially weighted average of dilational strain around the seismic stations shows a slight extension, which is the opposite relation of what we expect for the increase in dv/v. However, the rotated axial strain reveals compression in a range near the maximum contractional strain (azimuth of N35°W to N45°E). The observed increase in dv/v can be explained by the pre-existing microcracks aligned perpendicular to the contractional strains, which may efficiently be closed to cause an increase in rigidity.

This study shows that by minimizing the uncertainties associated with the environmental factors through modeling their effect on the observed dv/v time series, and by mitigating the non-tectonic artifacts such as the perturbation of the noise sources or the clock drift potential through improved processing methods, the dv/v can be used as an effective tool to monitor and constrain the evolution of the physical state in an active fault zone.

Open Research

We maintain the Julia-based software `SeisMonitoring.jl` (doi: `https://doi.org/10.5281/zenodo.8320944`) by the continuous integration testing on GitHub to be executed in different machine environments. The input files to conduct the ambient noise processings and the Jupyter notebooks for the post-processings are documented in the GitHub repository `SeisMonitoring_Paper` (doi: `https://doi.org/10.5281/zenodo.8330432`) to reproduce the results from scratch. The intermediate outputs including the cross-correlation functions are accessible in the cloud storage `dasway` (doi: `https://doi.org/10.6069/PK9D-941110.6069/PK9D-9411`). The minimal working example of the `SeisMonitoring.jl` to conduct from downloading data to measuring the dv/v using a docker container is available in `SeisMonitoring_Example` (doi: `https://doi.org/10.5281/zenodo.8330420`).

Acknowledgments

We thank Jared Bryan for the constructive discussion and the contributions to the Julia modules. This research is part of the Frontera computing project at the Texas Advanced Computing Center. The research allocation was awarded as EAR20004 to Marine Denolle. Frontera is made possible by the National Science Foundation award OAC-1818253 (Stanzione et al., 2020). The computations in this paper were also run on the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University for the early development of the software tools. Data for this study come from the High Resolution Seismic Network (HRSN) doi:10.7932/HRSN, operated by the UC Berkeley Seismological Laboratory, which is archived at the Northern California Earthquake Data Center (NCEDC), doi: 10.7932/NCEDC. We downloaded the data associated with the channels of BP/SP from there. We thank Dave Shelly for providing the updated LFE catalog. We processed the GNSS dataset obtained from NASA MEaSUREs ESESES Project (Bock et al., 2021). We used the dataset named `WNAM_Clean_TrendNeuTimeSeries_sopac` to compute the cumulative strain field. We used GMT6.2.0 (Wessel et al., 2019) to plot the map of Parkfield. The `SeisIO.jl` v1.2.1 (Jones et al., 2020), `SeisNoise.jl` v0.5.3 (Clements & Denolle, 2020), and `SeisDvv.jl` developed by J. Bryan and T. Clements (`https://github.com/tclements/SeisDvv.jl`) are the kernel modules of the `SeisMonitoring.jl`. The stretching method in the SeisDvv refers to the Python module for Viens et al. (2018) (`https://github.com/lviens/2018_JGR`). The MWCS method is implemented as a translation in Julia from the Python-based package MSNoise (Lecocq et al., 2014) (`https://github.com/ROBelgium/MSNoise`). The CWT is implemented as the Julia translation of pycwt (`https://github.com/regeirk/pycwt`). This work is supported by the National Science Foundation as a CAREER award to M. Denolle (EAR-1749556).

Author contributions

K. Okubo processed the ambient seismic noise to evaluate the dv/v, conducted the statistical analysis of model parameters using the MCMC analysis, and computed the strain field using the GNSS dataset. M. Denolle had a supervisory role in the analysis. All authors contributed to the discussion, writing, and reviewing of this article. Conceptualization: MD, KO; Data curation: KO, BD; Formal Analysis: KO, BD; Funding acquisition: MD; Investigation: KO, BD; Methodology: KO, BD, MD; Project administration: MD; Resources: MD, KO; Software: KO; Supervision: MD; Validation: KO; Visualization: KO; Writing – original draft: KO, BD, MD; Writing – review & editing: KO, BD, MD;

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716-723. doi: 10.1109/TAC.1974.1100705

Akasaka, C., & Nakanishi, S. (2000). Correction of background gravity changes due to precipitation: Oguni geothermal field, Japan. In *Proceedings world geothermal congress* (pp. 2471–2475).

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bull. Seismol. Soc. Am.*, *68*(5), 1521–1532.

Andrews, D. J. (2005). Rupture dynamics with energy loss outside the slip zone. *J. Geophys. Res.*, *110*, 1307. doi: 10.1029/2004JB003191

Bacques, G., de Michele, M., Raucoules, D., Aochi, H., & Rolandone, F. (2018). Shallow deformation of the San Andreas fault 5 years following the 2004 Parkfield earthquake (Mw6) combining ERS2 and Envisat InSAR. *Sci. Rep.*, *8*(1), 6032. doi: 10.1038/s41598-018-24447-3

Baig, A. M., Campillo, M., & Brenguier, F. (2009). Denoising seismic noise cross correlations. *J. Geophys. Res.*, *114*(B8). doi: 10.1029/2008JB006085

Baillard, C., Crawford, W. C., Ballu, V., Hibert, C., & Mangeney, A. (2013). An automatic kurtosis-based P-and S-phase picker designed for local seismic networks. *Bull. Seismol. Soc. Am.*, *104*(1), 394–409. doi: 10.1785/0120120347

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., ... Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophys. J. Int.*, *169*(3), 1239–1260. doi: 10.1111/j.1365-246X.2007.03374.x

Ben-Zion, Y., & Leary, P. (1986). Thermoelastic strain in a half-space covered by unconsolidated material. *Bull. Seismol. Soc. Am.*, *76*(5), 1447–1460.

Ben-Zion, Y., Peng, Z., Okaya, D., Seeber, L., Armbruster, J. G., Ozer, N., ... Aktar, M. (2003). A shallow fault-zone structure illuminated by trapped waves in the Karadere–Duzce branch of the North Anatolian Fault, western Turkey. *Geophys. J. Int.*, *152*(3), 699-717. doi: 10.1046/j.1365-246X.2003.01870.x

Berger, J. (1975). A note on thermoelastic strains and tilts. *J. Geophys. Res.*, *80*(2), 274-277. doi: 10.1029/JB080i002p00274

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. (2017). Julia: A Fresh Approach to Numerical Computing. *SIAM Review*, *59*(1), 65-98. doi: 10.1137/141000671

Bock, Y., Moore, A., Argus, D., Fang, P., Golriz, D., Guns, K., ... others (2021). *Extended Solid Earth Science ESDR System (ES3): Algorithm Theoretical Basis Document, NASA MEaSUREs project, # NNH17ZDA001N.*

Bonilla, L. F., Guéguen, P., & Ben-Zion, Y. (2019). Monitoring Coseismic Temporal Changes of Shallow Material during Strong Ground Motion with Interferometry and Autocorrelation. *Bull. Seismol. Soc. Am.*, *109*(1), 187-198. doi: 10.1785/0120180092

Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N. M., Nadeau, R. M., & Larose, E. (2008). 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., Shapiro, N. M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O., & Nercessian, A. (2008). Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience*, *1*(2), 126–130. doi: 10.1038/ngeo104

Campillo, M., & Paul, A. (2003). Long-Range Correlations in the Diffuse Seismic Coda. *Science*, *299*(5606), 547-549. doi: 10.1126/science.1078551

Chiarabba, C., De Gori, P., Segou, M., & Cattaneo, M. (2020). Seismic velocity precursors to the 2016 Mw 6.5 Norcia (Italy) earthquake. *Geology*, *48*(9), 924-928. doi: 10.1130/G47048.1

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*(2), 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*(13), 6459-6465. doi: 10.1029/2018GL077706

Clements, T., & Denolle, M. A. (2020). SeisNoise.jl: Ambient Seismic Noise Cross Correlation on the CPU and GPU in Julia. *Seismol. Res. Lett.*, *92*(1), 517–527. doi: 10.1785/0220200192

Clements, T., & Denolle, M. A. (2023). The Seismic Signature of California's Earthquakes, Droughts, and Floods. *J. Geophys. Res.*, *128*(1), e2022JB025553. doi: 10.1029/2022JB025553

Colombero, C., Baillet, L., Comina, C., Jongmans, D., Larose, E., Valentin, J., & Vinciguerra, S. (2018). Integration of ambient seismic noise monitoring, displacement and meteorological measurements to infer the temperature-controlled long-term evolution of a complex prone-to-fall cliff. *Geophys. J. Int.*, *213*(3), 1876-1897. doi: 10.1093/gji/ggy090

Colombi, A., Chaput, J., Brenguier, F., Hillers, G., Roux, P., & Campillo, M. (2014). On the temporal stability of the coda of ambient noise correlations. *Comptes Rendus Geoscience*, *346*(11), 307-316. doi: 10.1016/j.crte.2014.10.002

Crowell, B. W. (2019). *PyTAGS: A Python package for the temporal analysis of GNSS strains, V1.0.* doi: 10.5281/zenodo.2634525

Crowell, B. W., Bock, Y., Sandwell, D. T., & Fialko, Y. (2013). Geodetic investigation into the deformation of the Salton Trough. *J. Geophys. Res.*, *118*(9), 5030-5039. doi: 10.1002/jgrb.50347

Delbridge, B. G., Carmichael, J. D., Nadeau, R. M., Shelly, D. R., & Bürgmann, R. (2020). Geodetic Measurements of Slow-Slip Events Southeast of Parkfield, CA. *J. Geophys. Res.*, *125*(5), e2019JB019059. doi: 10.1029/2019JB019059

Delorey, A. A., Guyer, R. A., Bokelmann, G. H. R., & Johnson, P. A. (2021). Probing the Damage Zone at Parkfield. *Geophys. Res. Lett.*, *48*(13), e2021GL093518. doi: 10.1029/2021GL093518

Donaldson, C., Caudron, C., Green, R. G., Thelen, W. A., & White, R. S. (2017). Relative seismic velocity variations correlate with deformation at Kīlauea volcano. *Science Advances*, *3*(6), e1700219. doi: 10.1126/sciadv.1700219

Donaldson, C., Winder, T., Caudron, C., & White, R. S. (2019). Crustal seismic velocity responds to a magmatic intrusion and seasonal loading in Iceland’s Northern Volcanic Zone. *Science Advances*, *5*(11), eaax6642. doi: 10.1126/sciadv.aax6642

Durand, S., Montagner, J. P., Roux, P., Brenguier, F., Nadeau, R. M., & Ricard, Y. (2011). Passive monitoring of anisotropy change associated with the Parkfield 2004 earthquake. *Geophys. Res. Lett.*, *38*(13). doi: 10.1029/2011GL047875

Ermert, L. A., Cabral-Cano, E., Chaussard, E., Solano-Rojas, D., Quintanar, L., Morales Padilla, D., … Denolle, M. A. (2023). Probing environmental and tectonic changes underneath Mexico City with the urban seismic field. *Solid Earth*, *14*(5), 529–549. doi: 10.5194/se-14-529-2023

Feng, K.-F., Huang, H.-H., Hsu, Y.-J., & Wu, Y.-M. (2021). Controls on Seasonal Variations of Crustal Seismic Velocity in Taiwan Using Single-Station Cross-Component Analysis of Ambient Noise Interferometry. *J. Geophys. Res.*, *126*(11), e2021JB022650. doi: 10.1029/2021JB022650

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. (2013). emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, *125*(925), 306–312. doi: 10.1086/670067

Gassenmeier, M., Sens-Schönfelder, C., Eulenfeld, T., Bartsch, M., Victor, P., Tilmann, F., & Korn, M. (2016). Field observations of seismic velocity changes caused by shaking-induced damage and healing due to mesoscopic nonlinearity. *Geophys. J. Int.*, *204*(3), 1490-1502. doi: 10.1093/gji/ggv529

Goodman, J., & Weare, J. (2010). Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science*, *5*(1), 65 – 80. doi: 10.2140/camcos.2010.5.65

Gouédard, P., Seher, T., McGuire, J. J., Collins, J. A., & van der Hilst, R. D. (2014). Correction of Ocean-Bottom Seismometer Instrumental Clock Errors Using Ambient Seismic Noise. *Bull. Seismol. Soc. Am.*, *104*(3), 1276-1288. doi: 10.1785/0120130157

Griffith, W. A., Rosakis, A., Pollard, D. D., & Ko, C. W. (2009). Dynamic rupture experiments elucidate tensile crack development during propagating earthquake ruptures. *Geology*, *37*(9), 795-798. doi: 10.1130/G30064A.1

Hable, S., Sigloch, K., Barruol, G., Stähler, S. C., & Hadziioannou, C. (2018). Clock errors in land and ocean bottom seismograms: high-accuracy estimates from multiple-component noise cross-correlations. *Geophys. J. Int.*, *214*(3), 2014-2034. doi: 10.1093/gji/ggy236

Hadziioannou, C., Larose, E., Baig, A., Roux, P., & Campillo, M. (2011). Improving temporal resolution in ambient noise monitoring of seismic wave speed. *J. Geophys. Res.*, *116*(B7). doi: 10.1029/2011JB008200

Hadziioannou, C., Larose, E., Coutant, O., Roux, P., & Campillo, M. (2009). Stability of monitoring weak changes in multiply scattering media with ambient noise correlation: Laboratory experiments. *J. Acoust. Soc. Am.*, *125*(6), 3688–3695. doi: 10.1121/1.3125345

Hickman, S., & Zoback, M. (2004). Stress orientations and magnitudes in the SAFOD pilot hole. *Geophys. Res. Lett.*, *31*(15). doi: 10.1029/2004GL020043

Hillers, G., Ben-Zion, Y., Campillo, M., & Zigone, D. (2015). Seasonal variations of seismic velocities in the San Jacinto fault area observed with ambient seismic noise. *Geophys. J. Int.*, *202*(2), 920-932. doi: 10.1093/gji/ggv151

Hillers, G., Campillo, M., Brenguier, F., Moreau, L., Agnew, D. C., & 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.*, *124*(7), 7171-7192. doi: 10.1029/2018JB017143

Hirose, T., Nakahara, H., & Nishimura, T. (2017). Combined use of repeated active shots and ambient noise to detect temporal changes in seismic velocity: application to Sakurajima volcano, Japan. *Earth, Planets and Space*, *69*(1), 42. doi: 10.1186/s40623-017-0613-7

Hirose, T., & Ueda, H. (2023). Continuous monitoring of instrumental clock errors at 50 volcanoes in Japan based on seismic interferometry. *PREPRINT (Version 1) available at Research Square*. doi: 10.21203/rs.3.rs-2011839/v1

Hobiger, M., Wegler, U., Shiomi, K., & Nakahara, H. (2014). Single-station cross-correlation analysis of ambient seismic noise: application to stations in the surroundings of the 2008 Iwate-Miyagi Nairiku earthquake . *Geophys. J. Int.*, *198*(1), 90-109. doi: 10.1093/gji/ggu115

Hobiger, M., Wegler, U., Shiomi, K., & Nakahara, H. (2016). Coseismic and post-seismic velocity changes detected by Passive Image Interferometry: comparison of one great and five strong earthquakes in Japan. *Geophys. J. Int.*, *205*(2), 1053-1073. doi: 10.1093/gji/ggw066

Hotovec-Ellis, A. J., Shiro, B. R., Shelly, D. R., Anderson, K. R., Haney, M. M., Thelen, W. A., . . . Johanson, I. A. (2022). Earthquake-Derived Seismic Velocity Changes During the 2018 Caldera Collapse of Kīlauea volcano. *J. Geophys. Res.*, *127*(2), e2021JB023324. doi: 10.1029/2021JB023324

Hotovec-Ellis, A. J., Vidale, J. E., Gomberg, J., Thelen, W., & Moran, S. C. (2015). Changes in seismic velocity during the first 14 months of the 2004–2008 eruption of Mount St. Helens, Washington. *J. Geophys. Res.*, *120*(9), 6226-6240. doi: 10.1002/2015JB012101

Ikeda, T., & Tsuji, T. (2018). Temporal change in seismic velocity associated with an offshore MW 5.9 Off-Mie earthquake in the Nankai subduction zone from ambient noise cross-correlation. *Progress in Earth and Planetary Science*, *5*(1), 62. doi: 10.1186/s40645-018-0211-8

Illien, L., Andermann, C., Sens-Schönfelder, C., Cook, K. L., Baidya, K. P., Adhikari, L. B., & Hovius, N. (2021). Subsurface Moisture Regulates Himalayan Groundwater Storage and Discharge. *AGU Advances*, *2*(2), e2021AV000398. doi: 10.1029/2021AV000398

Illien, L., Sens-Schönfelder, C., Andermann, C., Marc, O., Cook, K. L., Adhikari, L. B., & Hovius, N. (2022). Seismic Velocity Recovery in the Subsurface: Transient Damage and Groundwater Drainage Following the 2015 Gorkha Earthquake, Nepal. *J. Geophys. Res.*, *127*(2), e2021JB023402. doi: 10.1029/2021JB023402

Johanson, I. A., & Bürgmann, R. (2010). Coseismic and postseismic slip from the 2003 San Simeon earthquake and their effects on backthrust slip and the 2004 Parkfield earthquake. *J. Geophys. Res.*, *115*(B7). doi: 10.1029/2009JB006599

Johanson, I. A., Fielding, E. J., Rolandone, F., & Bürgmann, R. (2006). Coseismic and Postseismic Slip of the 2004 Parkfield Earthquake from Space-Geodetic Data. *Bulletin of the Seismological Society of America*, *96*(4B), S269-S282. doi: 10.1785/0120050818

Johnson, P., & Sutin, A. (2005). Slow dynamics and anomalous nonlinear fast dynamics in diverse solids. *J. Acoust. Soc. Am.*, *117*(1), 124-130. doi: 10.1121/1.1823351

Jolivet, R., Simons, M., Agram, P. S., Duputel, Z., & Shen, Z.-K. (2015). Aseismic slip and seismogenic coupling along the central San Andreas Fault. *Geophys. Res. Lett.*, *42*(2), 297-306. doi: 10.1002/2014GL062222

Jones, J. P., Okubo, K., Clements, T., & Denolle, M. A. (2020). SeisIO: A Fast, Efficient Geophysical Data Architecture for the Julia Language. *Seismol. Res. Lett.*, *91*(4), 2368–2377. doi: 10.1785/0220190295

Kanasewich, E. R., Hemmings, C. D., & Alpaslan, T. (1973). Nth-root stack nonlinear multichannel filter. *Geophysics*, *38*(2), 327-338. doi: 10.1190/1.1440343

Klein, E., Bock, Y., Xu, X., Sandwell, D. T., Golriz, D., Fang, P., & Su, L. (2019). Transient Deformation in California From Two Decades of GPS Displacements: Implications for a Three-Dimensional Kinematic Reference Frame. *J. Geophys. Res.*, *124*(11), 12189-12223. doi: 10.1029/2018JB017201

Korenaga, J. (2013). Stacking with dual bootstrap resampling. *Geophys. J. Int.*, *195*(3), 2023-2036. doi: 10.1093/gji/ggt373

Larose, E., Derode, A., Campillo, M., & Fink, M. (2004). Imaging from one-bit correlations of wideband diffuse wave fields. *J. Appl. Phys*, *95*(12), 8393-8399. doi: 10.1063/1.1739529

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., & Stammler, K. (2017). Monitoring ground water storage at mesoscale using seismic noise: 30 years of continuous observation and thermo-elastic and hydrological modeling. *Scientific Reports*, *7*(1), 14241. doi: 10.1038/s41598-017-14468-9

Li, G., & Ben-Zion, Y. (2023). Daily and Seasonal Variations of Shallow Seismic Velocities in Southern California From Joint Analysis of H/V Ratios and Autocorrelations of Seismic Waveforms. *J. Geophys. Res.*, *128*(2), e2022JB025682. doi: 10.1029/2022JB025682

Li, Y.-G., & Malin, P. E. (2008). San Andreas Fault damage at SAFOD viewed with fault-guided waves. *Geophysical Research Letters*, *35*(8). doi: 10.1029/2007GL032924

Liu, G., Fomel, S., Jin, L., & Chen, X. (2009). Stacking seismic data using local correlation. *GEOPHYSICS*, *74*(3), V43-V48. doi: 10.1190/1.3085643

Liu, Y., Zhang, H., Thurber, C., & Roecker, S. (2008). Shear wave anisotropy in the crust around the San Andreas fault near Parkfield: Spatial and temporal analysis. *Geophys. J. Int.*, *172*(3), 957-970. doi: 10.1111/j.1365-246X.2007.03618.x

Lobkis, O. I., & Weaver, R. L. (2003). Coda-wave interferometry in finite solids: Recovery of P-to-S conversion rates in an elastodynamic billiard. *Physical review letters*, *90*(25), 254302. doi: 10.1103/PhysRevLett.90.254302

Löer, K., Riahi, N., & Saenger, E. H. (2018). Three-component ambient noise beamforming in the Parkfield area. *Geophys. J. Int.*, *213*(3), 1478-1491. doi: 10.1093/gji/ggy058

Lu, Y., & Ben-Zion, Y. (2021). Regional seismic velocity changes following the 2019 Mw 7.1 Ridgecrest, California earthquake from autocorrelations and P/S converted waves. *Geophys. J. Int.*, *228*(1), 620-630. doi: 10.1093/gji/ggab350

Ma, S. (2008). A physical model for widespread near-surface and fault zone damage induced by earthquakes. *Geochemistry, Geophysics, Geosystems*, *9*(11). doi: 10.1029/2008GC002231

Mai, P. M., & Thingbaijam, K. K. S. (2014). SRCMOD: An Online Database of Finite-Fault Rupture Models. *Seismol. Res. Lett.*, *85*(6), 1348-1357. doi: 10.1785/0220140077

Malagnini, L., Dreger, D., Bürgmann, R., Munafò, I., & Sebastiani, G. (2019). Modulation of Seismic Attenuation at Parkfield, Before and After the 2004 M6 Earthquake. *J. Geophys. Res.*, *124*(6), 5836-5853. doi: 10.1029/2019JB017372

Mao, S., Campillo, M., van der Hilst, R. D., Brenguier, F., Stehly, L., & Hillers, G. (2019). High temporal resolution monitoring of small variations in crustal strain by dense seismic arrays. *Geophys. Res. Lett.*, *46*(1), 128–137. doi: 10.1029/2018GL079944

Mao, S., Lecointre, A., van der Hilst, R. D., & Campillo, M. (2022). Space-time monitoring of groundwater fluctuations with passive seismic interferometry. *Nature Communications*, *13*(1), 4643. doi: 10.1038/s41467-022-32194-3

Mao, S., Mordret, A., Campillo, M., Fang, H., & van der Hilst, R. D. (2020). On the measurement of seismic traveltime changes in the time–frequency domain with wavelet cross-spectrum analysis. *Geophysical Journal International*, *221*(1), 550-568. doi: 10.1093/gji/ggz495

Marc, O., Sens-Schönfelder, C., Illien, L., Meunier, P., Hobiger, M., Sawazaki, K., ... Hovius, N. (2021). Toward Using Seismic Interferometry to Quantify Landscape Mechanical Variations after Earthquakes. *Bull. Seismol. Soc. Am.*, *111*(3), 1631-1649. doi: 10.1785/0120200264

McLaren, M. K., Hardebeck, J. L., van der Elst, N., Unruh, J. R., Bawden, G. W., & Blair, J. L. (2008). Complex Faulting Associated with the 22 December 2003 Mw 6.5 San Simeon, California, Earthquake, Aftershocks, and Postseismic Surface Deformation. *Bull. Seismol. Soc. Am.*, *98*(4), 1659-1680. doi: 10.1785/0120070088

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*(1), 347-360. doi: 10.1093/gji/ggv138

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

Nakata, N., Chang, J. P., Lawrence, J. F., & Boué, P. (2015). Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry. *J. Geophys. Res.*, *120*(2), 1159-1173. doi: 10.1002/2015JB011870

Nakata, N., Gualtieri, L., & Fichtner, A. (2019). *Seismic Ambient Noise.* Cambridge University Press. doi: 10.1017/9781108264808

National Oceanic and Atmospheric Administration (NOAA). (2022). *Climate Data Online, retrieved from https://www.ncdc.noaa.gov/cdo-web/datasets/GHCND/stations/GHCND:USR0000CPAR/detail.*

Nimiya, H., Ikeda, T., & Tsuji, T. (2017). Spatial and temporal seismic velocity changes on Kyushu Island during the 2016 Kumamoto earthquake. *Science Advances*, *3*(11). doi: 10.1126/sciadv.1700813

Nishida, K., Mizutani, Y., Ichihara, M., & Aoki, Y. (2020). Time-Lapse Monitoring of Seismic Velocity Associated With 2011 Shinmoe-Dake Eruption Using Seismic Interferometry: An Extended Kalman Filter Approach. *J. Geophys. Res.*, *125*(9), e2020JB020180. doi: 10.1029/2020JB020180

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*(7201), 204–208. doi: 10.1038/nature07111

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.*, *120*(6), 4301-4316. doi: 10.1002/2014JB011817

Obermann, A., Planès, T., Larose, E., Sens-Schönfelder, C., & Campillo, M. (2013). Depth sensitivity of seismic coda waves to velocity perturbations in an elastic heterogeneous medium. *Geophys. J. Int.*, *194*(1), 372–382. doi: 10.1093/gji/ggt043AdvanceAccess

Okubo, K., Bhat, H. S., Rougier, E., Marty, S., Schubnel, A., Lei, Z., … Klinger, Y. (2019). Dynamics, Radiation, and Overall Energy Budget of Earthquake Rupture With Coseismic Off-Fault Damage. *J. Geophys. Res.*, *124*(11), 11771–11801. doi: 10.1029/2019JB017304

Olivier, G., Brenguier, F., Campillo, M., Lynch, R., & Roux, P. (2015). Body-wave reconstruction from ambient seismic noise correlations in an underground mine. *GEOPHYSICS*, *80*(3), KS11-KS25. doi: 10.1190/geo2014-0299.1

Ostrovsky, L. A., & Johnson, P. A. (2001). Dynamic nonlinear elasticity in geomaterials. *La Rivista del Nuovo Cimento*, *24*(7), 1–46. doi: 10.1007/BF03548898

Pavlis, G. L., & Vernon, F. L. (2010). Array processing of teleseismic body waves with the USArray. *Computers & Geosciences*, *36*(7), 910–920. doi: 10.1016/j.cageo.2009.10.008

Poliakov, A. N., Dmowska, R., & Rice, J. R. (2002). Dynamic shear rupture interactions with fault bends and off-axis secondary faulting. *J. Geophys. Res.*, *107*(B11), ESE–6.

Poupinet, G., Ellsworth, W., & Frechet, J. (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras Fault, California. *J. Geophys. Res.*, *89*(B7), 5719–5731.

Rice, J. R., Sammis, C. G., & Parsons, R. (2005). Off-fault secondary failure induced by a dynamic slip pulse. *Bull. Seism. Soc. Am.*, *95*(1), 109–134. doi: 10.1785/0120030166

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.*, *119*(6), 4747-4765. doi: 10.1002/2013JB010695

Rivet, D., Brenguier, F., & Cappa, F. (2015). Improved detection of preeruptive seismic velocity drops at the Piton de La Fournaise volcano. *Geophys. Res. Lett.*, *42*(15), 6332-6339. doi: 10.1002/2015GL064835

Ross, E. O., Reber, J. E., & Titus, S. J. (2022). Relating Slip Behavior to Off-Fault Deformation Using Physical Models. *Geophys. Res. Lett.*, *49*(11), e2021GL096784. doi: 10.1029/2021GL096784

Rubinstein, J. L., Uchida, N., & Beroza, G. C. (2007). Seismic velocity reductions caused by the 2003 Tokachi-Oki earthquake. *J. Geophys. Res.*, *112*(B5). doi: 10.1029/2006JB004440

Sawazaki, K., Kimura, H., Shiomi, K., Uchida, N., Takagi, R., & Snieder, R. (2015).

Depth-dependence of seismic velocity change associated with the 2011 Tohoku earthquake, Japan, revealed from repeating earthquake analysis and finite-difference wave propagation simulation. *Geophys. J. Int.*, *201*(2), 741-763. doi: 10.1093/gji/ggv014

Sayers, C. M., & Kachanov, M. (1995). Microcrack-induced elastic wave anisotropy of brittle rocks. *J. Geophys. Res.*, *100*(B3), 4149-4156. doi: 10.1029/94JB03134

Schaff, D. P. (2012). Placing an Upper Bound on Preseismic Velocity Changes Measured by Ambient Noise Monitoring for the 2004 Mw 6.0 Parkfield Earthquake (California)Placing an Upper Bound on Preseismic Velocity Changes Measured by Ambient Noise Monitoring. *Bull. Seismol. Soc. Am.*, *102*(4), 1400-1416. doi: 10.1785/0120110342

Schaff, D. P., & Beroza, G. C. (2004). Coseismic and postseismic velocity changes measured by repeating earthquakes. *J. Geophys. Res.*, *109*(B10). doi: 10.1029/2004JB003011

Schimmel, M., & Paulssen, H. (1997). Noise reduction and detection of weak, coherent signals through phase-weighted stacks. *Geophys. J. Int.*, *130*(2), 497-505. doi: 10.1111/j.1365-246X.1997.tb05664.x

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464.

Sens-Schönfelder, C. (2008). Synchronizing seismic networks with ambient noise. *Geophys. J. Int.*, *174*(3), 966-970. doi: 10.1111/j.1365-246X.2008.03842.x

Sens-Schonfelder, C., & Eulenfeld, T. (2019). Probing the in situ Elastic Nonlinearity of Rocks with Earth Tides and Seismic Noise. *Phys. Rev. Lett.*, *122*. doi: 10.1103/PhysRevLett.122.138501

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

Seydoux, L., Shapiro, N. M., de Rosny, J., & Landès, M. (2016). Spatial coherence of the seismic wavefield continuously recorded by the USArray. *Geophys. Res. Lett.*, *43*(18), 9644-9652. doi: 10.1002/2016GL070320

Shapiro, N. M., & Campillo, M. (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophys. Res. Lett.*, *31*(7). doi: 10.1029/2004GL019491

Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H. (2005). High-Resolution Surface-Wave Tomography from Ambient Seismic Noise. *Science*, *307*(5715), 1615–1618. doi: 10.1126/science.1108339

Shelly, D. R. (2017). A 15 year catalog of more than 1 million low-frequency earthquakes: Tracking tremor and slip along the deep San Andreas Fault. *J. Geophys. Res.*, *122*(5), 3739-3753. doi: 10.1002/2017JB014047

Shen, Z.-K., Jackson, D. D., & Ge, B. X. (1996). Crustal deformation across and beyond the Los Angeles basin from geodetic measurements. *J. Geophys. Res.*, *101*(B12), 27957-27980. doi: 10.1029/96JB02544

Sheng, Y., Ellsworth, W. L., Lellouch, A., & Beroza, G. C. (2021). Depth Constraints on Coseismic Velocity Changes From Frequency-Dependent Measurements of Repeating Earthquake Waveforms. *J. Geophys. Res.*, *126*(2), e2020JB020421. doi: 10.1029/2020JB020421

Shreedharan, S., Bolton, D. C., Rivière, J., & Marone, C. (2021). Competition between preslip and deviatoric stress modulates precursors for laboratory earthquakes. *Earth and Planetary Science Letters*, *553*, 116623. doi: 10.1016/j.epsl.2020.116623

Snieder, R., Grêt, A., Douma, H., & Scales, J. (2002). Coda Wave Interferometry for Estimating Nonlinear Behavior in Seismic Velocity. *Science*, *295*(5563), 2253–2255. doi: 10.1126/science.1070015

Snieder, R., Sens-Schönfelder, C., & Wu, R. (2017). The time dependence of rock healing as a universal relaxation process, a tutorial. *Geophys. J. Int.*, *208*(1), 1–9. doi: 10.1093/gji/ggw377

Stanzione, D., West, J., Evans, R. T., Minyard, T., Ghattas, O., & Panda, D. K. (2020). Frontera: The Evolution of Leadership Computing at the National Science Foundation. In *Practice and Experience in Advanced Research Computing* (pp. 106–111). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3311790.3396656

Stehly, L., Campillo, M., & Shapiro, N. M. (2007). Traveltime measurements from noise correlation: stability and detection of instrumental time-shifts. *Geophys. J. Int.*, *171*(1), 223-230. doi: 10.1111/j.1365-246X.2007.03492.x

Stein, R. S., & King, G. C. P. (1984). Seismic Potential Revealed by Surface Folding: 1983 Coalinga, California, Earthquake. *Science*, *224*(4651), 869-872. doi: 10.1126/science.224.4651.869

Taira, T., Brenguier, F., & Kong, Q. (2015). Ambient noise-based monitoring of seismic velocity changes associated with the 2014 Mw 6.0 South Napa earthquake. *Geophys. Res. Lett.*, *42*(17), 6997-7004. doi: 10.1002/2015GL065308

Takano, T., Brenguier, F., Campillo, M., Peltier, A., & Nishimura, T. (2019). Noise-based passive ballistic wave seismic monitoring on an active volcano. *Geophys. J. Int.*, *220*(1), 501-507. doi: 10.1093/gji/ggz466

Takano, T., Nishimura, T., & Nakahara, H. (2017). Seismic velocity changes concentrated at the shallow structure as inferred from correlation analyses of ambient noise during volcano deformation at Izu-Oshima, Japan. *J. Geophys. Res.*, *122*(8), 6721-6736. doi: 10.1002/2017JB014340

Takano, T., Nishimura, T., Nakahara, H., Ueda, H., & Fujita, E. (2019). Sensitivity of Seismic Velocity Changes to the Tidal Strain at Different Lapse Times: Data Analyses of a Small Seismic Array at Izu-Oshima Volcano. *J. Geophys. Res.*, *124*(3), 3011-3023. doi: 10.1029/2018JB016235

Ten Cate, J. A., & Shankland, T. J. (1996). Slow dynamics in the nonlinear elastic response of Berea sandstone. *Geophysical Research Letters*, *23*(21), 3019-3022. doi: 10.1029/96GL02884

TenCate, J. A., Smith, E., & Guyer, R. A. (2000). Universal Slow Dynamics in Granular Solids. *Phys. Rev. Lett.*, *85*, 1020–1023. doi: 10.1103/PhysRevLett.85.1020

Thangraj, J. S., & Pulliam, J. (2021). Towards real-time assessment of convergence criteria in seismic interferometry: Selective stacking of cross-correlations at the San Emidio geothermal field. *J. Appl. Geophys.*, *193*, 104426. doi: 10.1016/j.jappgeo.2021.104426

Thomas, M. Y., & Bhat, H. S. (2018). Dynamic evolution of off-fault medium during an earthquake: a micromechanics based model. *Geophys. J. Int.*, *214*(2), 1267–1280. doi: 10.1093/gji/ggy129

Tonegawa, T., Takemura, S., Yabe, S., & Yomogida, K. (2022). Fluid Migration Before and During Slow Earthquakes in the Shallow Nankai Subduction Zone. *J. Geophys. Res.*, *127*(3), e2021JB023583. doi: 10.1029/2021JB023583

Torrence, C., & Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. *Bull. Am. Meteorol. Soc.*, *79*(1), 61–78. doi: 10.1175/1520-0477(1998)079⟨0061:APGTWA⟩2.0.CO;2

Ueno, T., Saito, T., Shiomi, K., & Haryu, Y. (2015). Monitoring the instrument response of the high-sensitivity seismograph network in Japan (Hi-net): effects of response changes on seismic interferometry analysis. *Earth, Planets and Space*, *67*(1), 135. doi: 10.1186/s40623-015-0305-0

U.S. Geological Survey. (2023a). *Earthquake Catalog, accessed June 22, 2023 at https://earthquake.usgs.gov/earthquakes/search.*

U.S. Geological Survey. (2023b). *U.S. Geological Survey3D Elevation Program 1 arc-second Resolution Digital Elevation Model, accessed August 29, 2023, at https://data.usgs.gov/datacatalog/data/USGS:35f9c4d4-b113-4c8d-8691-*

47c428c29a5b.

U.S. Geological Survey. (2023c). *U.S. Geological Survey and California Geological Survey, Quaternary fault and fold database for the United States, accessed March 28, 2023, at https://www.usgs.gov/natural-hazards/earthquake-hazards/faults.*

Ventosa, S., Schimmel, M., & Stutzmann, E. (2017). Extracting surface waves, hum and normal modes: time-scale phase-weighted stack and beyond. *Geophys. J. Int.*, *211*(1), 30-44. doi: 10.1093/gji/ggx284

Vidale, J. E., & Li, Y.-G. (2003). Damage to the shallow Landers fault from the nearby Hector Mine earthquake. *Nature*, *421*(6922), 524. doi: 10.1038/nature01354

Viens, L., Denolle, M., Miyake, H., Sakai, S., & Nakagawa, S. (2017). Retrieving impulse response function amplitudes from the ambient seismic field. *Geophys. J. Int.*, *210*(1), 210–222.

Viens, L., Denolle, M. A., Hirata, N., & Nakagawa, S. (2018). Complex near-surface rheology inferred from the response of greater Tokyo to strong ground motions. *J. Geophys. Res.*, *123*(7), 5710–5729. doi: 10.1029/2018JB015697

Viens, L., & Van Houtte, C. (2019). Denoising ambient seismic field correlation functions with convolutional autoencoders. *Geophys. J. Int.*, *220*(3), 1521-1535. doi: 10.1093/gji/ggz509

Wang, Q.-Y., Brenguier, F., Campillo, M., Lecointre, A., Takeda, T., & Aoki, Y. (2017). Seasonal Crustal Seismic Velocity Changes Throughout Japan. *J. Geophys. Res.*, *122*(10), 7987–8002. doi: 10.1002/2017JB014307

Weaver, R. L., Hadziioannou, C., Larose, E., & Campillo, M. (2011). On the precision of noise correlation interferometry. *Geophys. J. Int.*, *185*(3), 1384-1392. doi: 10.1111/j.1365-246X.2011.05015.x

Wegler, U., & Sens-Schönfelder, C. (2007). Fault zone monitoring with passive image interferometry. *Geophys. J. Int.*, *168*(3), 1029-1033. doi: 10.1111/j.1365-246X .2006.03284.x

Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H. F., & Tian, D. (2019). The Generic Mapping Tools Version 6. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5556-5564. doi: 10.1029/2019GC008515

Western Regional Climate Center. (2022). *RAWS USA Climate Archive. Retrieved from https://raws.dri.edu/cgi-bin/rawMAIN.pl?caCPAR.*

Wu, C., Delorey, A., Brenguier, F., Hadziioannou, C., Daub, E. G., & Johnson, P. (2016). Constraining depth range of S wave velocity decrease after large earthquakes near Parkfield, California. *Geophysical Research Letters*, *43*(12), 6129–6136. doi: 10.1002/2016GL069145

Yamamura, K., Sano, O., Utada, H., Takei, Y., Nakao, S., & Fukao, Y. (2003). Long-term observation of in situ seismic velocity and attenuation. *J. Geophys. Res.*, *108*(B6). doi: 10.1029/2002JB002005

Yamashita, T. (2000). Generation of microcracks by dynamic shear rupture and its effects on rupture growth and elastic wave radiation. *Geophys. J. Int.*, *143*(2), 395–406. doi: 10.1046/j.1365-246X.2000.01238.x

Yang, X., Bryan, J., Okubo, K., Jiang, C., Clements, T., & Denolle, M. A. (2022). Optimal stacking of noise cross-correlation functions. *Geophys. J. Int.*, *232*(3), 1600-1618. doi: 10.1093/gji/ggac410

Yuan, C., Bryan, J., & Denolle, M. (2021). Numerical comparison of time-, frequency- and wavelet-domain methods for coda wave interferometry. *Geophys. J. Int.*, *226*(2), 828-846. doi: 10.1093/gji/ggab140

Zhan, Z., Tsai, V. C., & Clayton, R. W. (2013). Spurious velocity changes caused by temporal variations in ambient noise frequency content. *Geophys. J. Int.*, *194*(3), 1574-1581. doi: 10.1093/gji/ggt170

Zhou, L., Song, X., & Weaver, R. L. (2020). Retrieval of amplitude and attenuation from ambient seismic noise: synthetic data and practical considerations. *Geo-*

phys. J. Int., *222*(1), 544-559.

Table 1: List of abbreviations.

Abbreviation	**Description**
CC	Correlation coefficient
CF	Correlation function
ACF	Auto-correlation function
CCF	Cross-correlation function
SS	The 2003 M6.5 San Simeon Earthquake
SAF	San Andreas Fault
PF	The 2004 M6.0 Parkfield Earthquake
CWT	continuous wavelet transform
MWCS	Moving-window cross-spectral method

Table 2: Model parameters and the ranges used for the MCMC sampling.

variable	description	Parameter sampling range minimum	maximum
a_0	offset of dv/v	-1.0	1.0 [%]
p_1	factor associated with precipitation	-∞	0
p_2	factor associated with temperature	-∞	∞
t_0^{shift}	time shift with the time series of temperature	0	90 [days]
s_1	factor of coseismic velocity decrease with SS	0.0	0.5*
$t_1^{\max}$	maximum time of healing with SS	1	3×10^4 [years]
s_2	factor of coseismic velocity decrease with PF	0.0	1.0
$t_2^{\max}$	maximum time of healing with PF	1	3×10^4 [years]
b_0	slope of the linear term	-∞	∞
c_0	factor of residual healing	-∞	0
f_0	uncertainty of dv/v estimation	10^{-10}	10^{10}

* The s_1 is constrained up to half of s_2 for each iteration.

Table 3: Model parameters fixed during the MCMC sampling, which is obtained as median values evaluated from the preliminary test of MCMC sampling without fixing $\tau_i^{\min}$ and α_0.

variable	description	Fixed values stretching	MWCS
α_0	decay factor in the GWL model	0.0243	0.0389 [1/day]
$\tau_1^{\min}$	minimum time of healing with SS	0.76	0.03 [months]
$\tau_2^{\min}$	minimum time of healing with PF	3.8	3.8 [months]

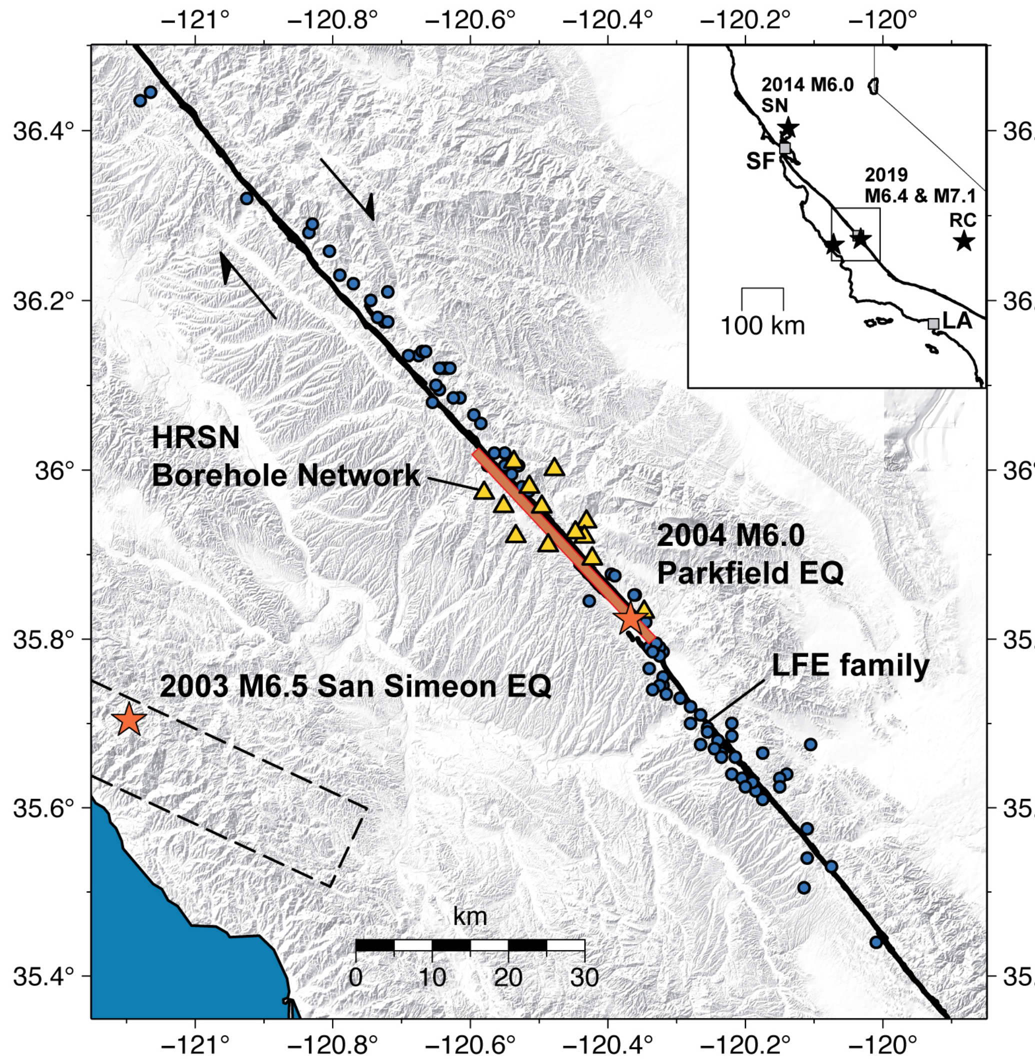

Figure 1: Location of the borehole network at Parkfield. The yellow triangles show the 13 borehole stations deployed and maintained by the UC Berkeley Seismological Laboratory. The geometry of the San Andreas fault is obtained from the Quaternary Fault and Fold database, USGS. The background topography is downloaded from USGS 3D Elevation Program (3DEP) Datasets from the National Map. The blue circles show the LFE family (Shelly, 2017). The dashed line indicates the main plane of the rupture associated with the 2003 San Simeon earthquake (Johanson & Bürgmann, 2010). The stars indicate the hypocenters of major earthquakes; the San Simeon earthquake is obtained from McLaren et al. (2008), and the others from SRCMOD (Mai & Thingbaijam, 2014). The seismic station locations with their name can be found in Figure S1. SN: South Napa earthquake, RC: Ridgecrest earthquake sequence.

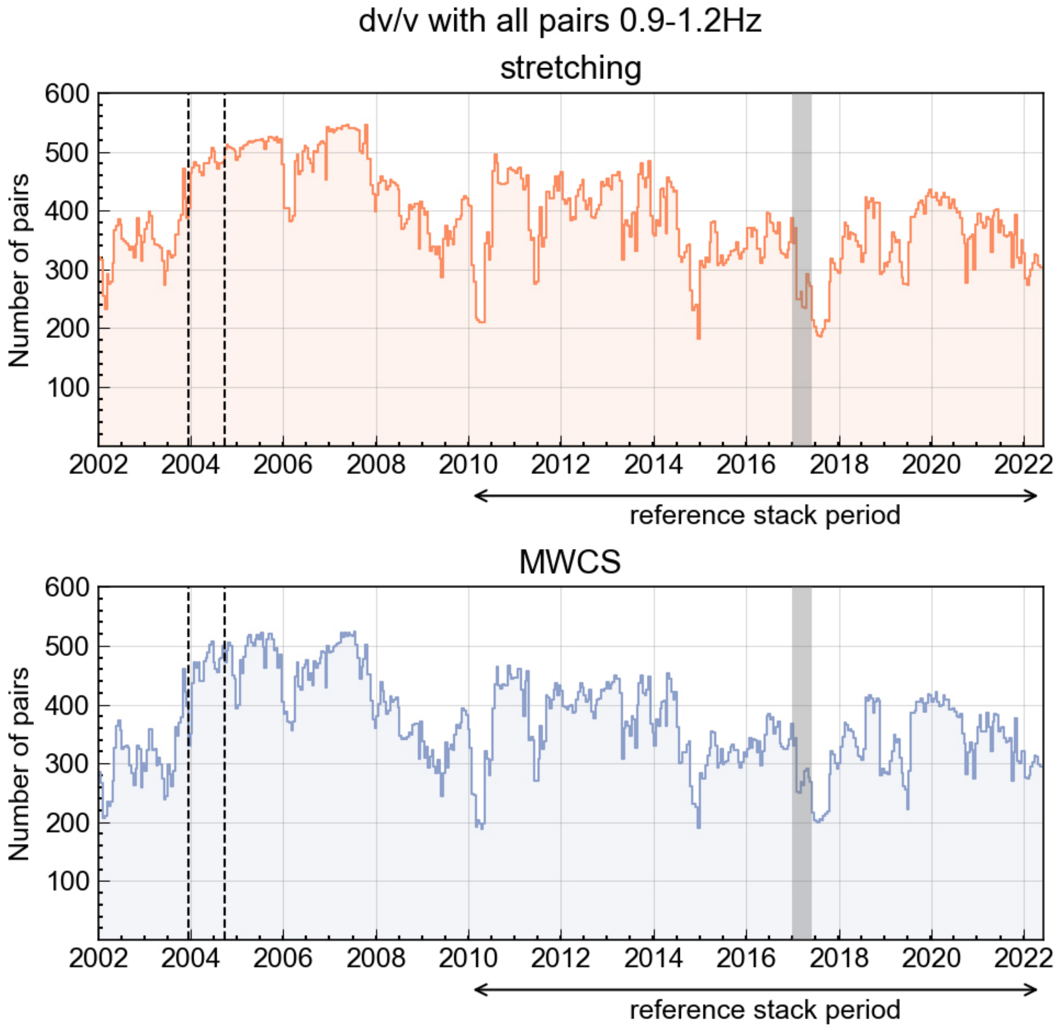

Figure 2: Number of station-component combination pairs after selecting good quality dv/v measurements. The top and bottom figures are associated with the stretching and MWCS methods. The vertical dashed line indicates the dates associated with the SS and PF earthquakes. The shaded area in grey indicates the unstable period of dv/v, which could be caused due to the clock shift. The bottom arrows annotate the reference period from January 2010 to May 2022.

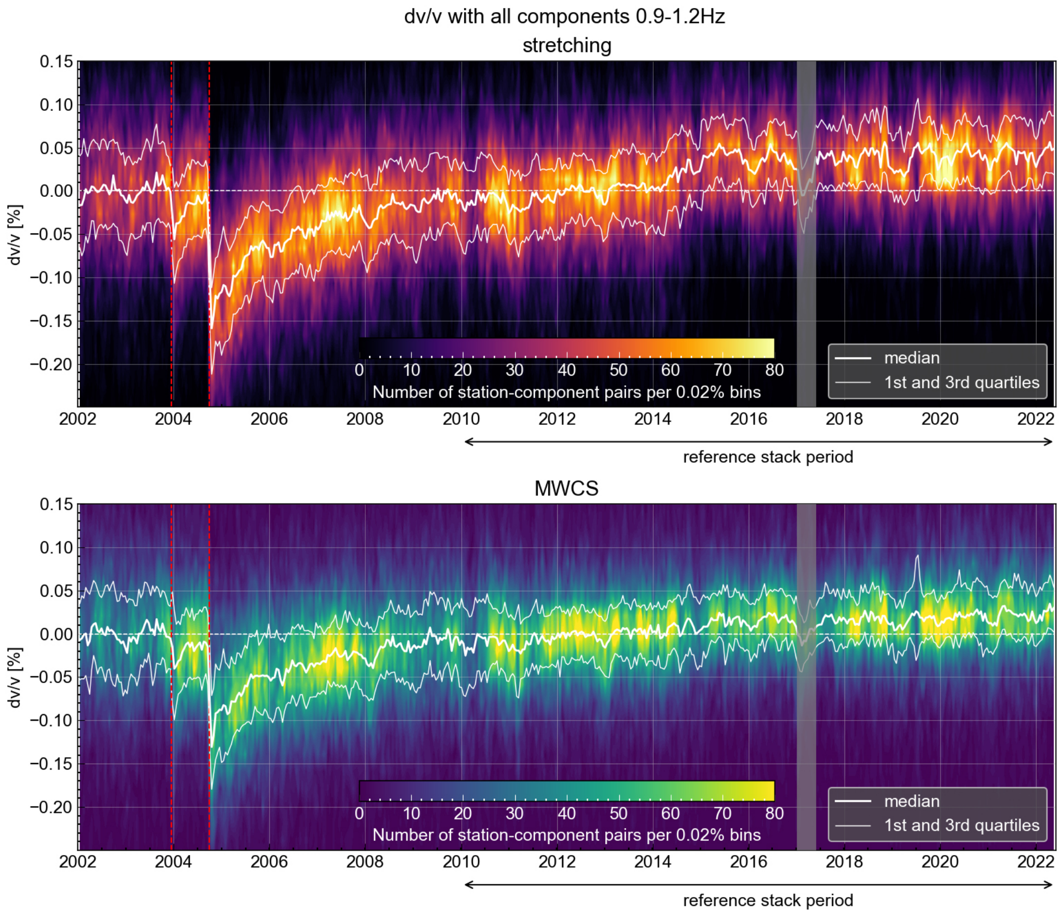

Figure 3: Time history of the dv/v for all the station-component pair combinations with the frequency band of 0.9-1.2Hz. The color contour shows the number of dv/v measurements meeting the threshold within 0.02% of the dv/v bin with respect to the time bins. The solid tick and thin lines indicate the median and the first and third quartiles, respectively. The shaded area in grey indicates the period where the clock drift was observed. The red dashed lines are the dates associated with the SS and PF earthquakes.

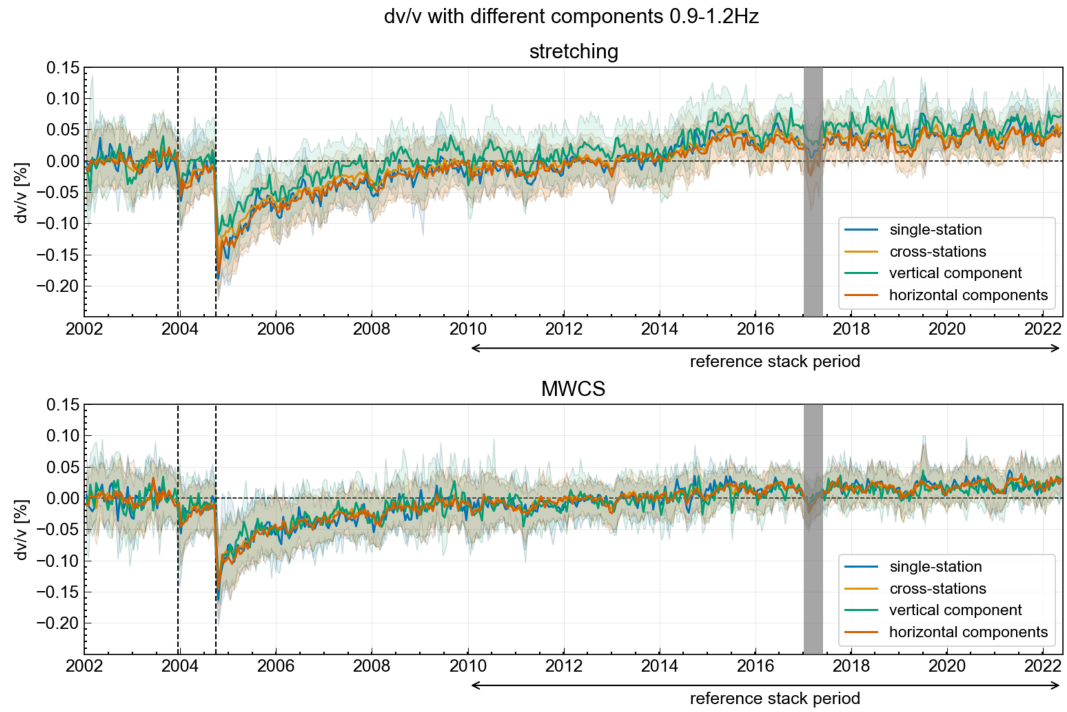

Figure 4: Time history of dv/v with the different set of components. The solid line and the highlighted area indicate the median and the first and third quartiles, respectively. The single-station and cross-stations contain the dv/v measurements associated with the nine components of ACFs or CCFs, while the vertical and horizontal components are for the case with both the ACFs and CCFs, respectively.

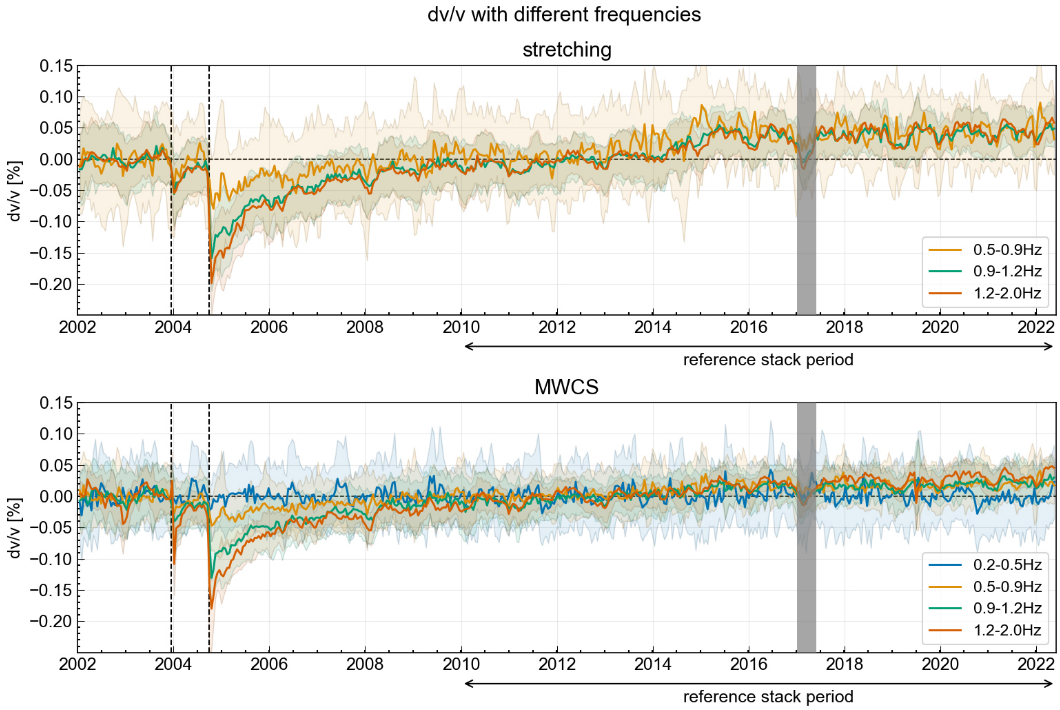

Figure 5: Time history of dv/v with different frequency bands using all components. The lines and highlighted area indicate the same with Figure 4. The lowest frequency band of 0.2-0.5Hz associated with the stretching method is removed due to the large fluctuations in the estimation of dv/v.

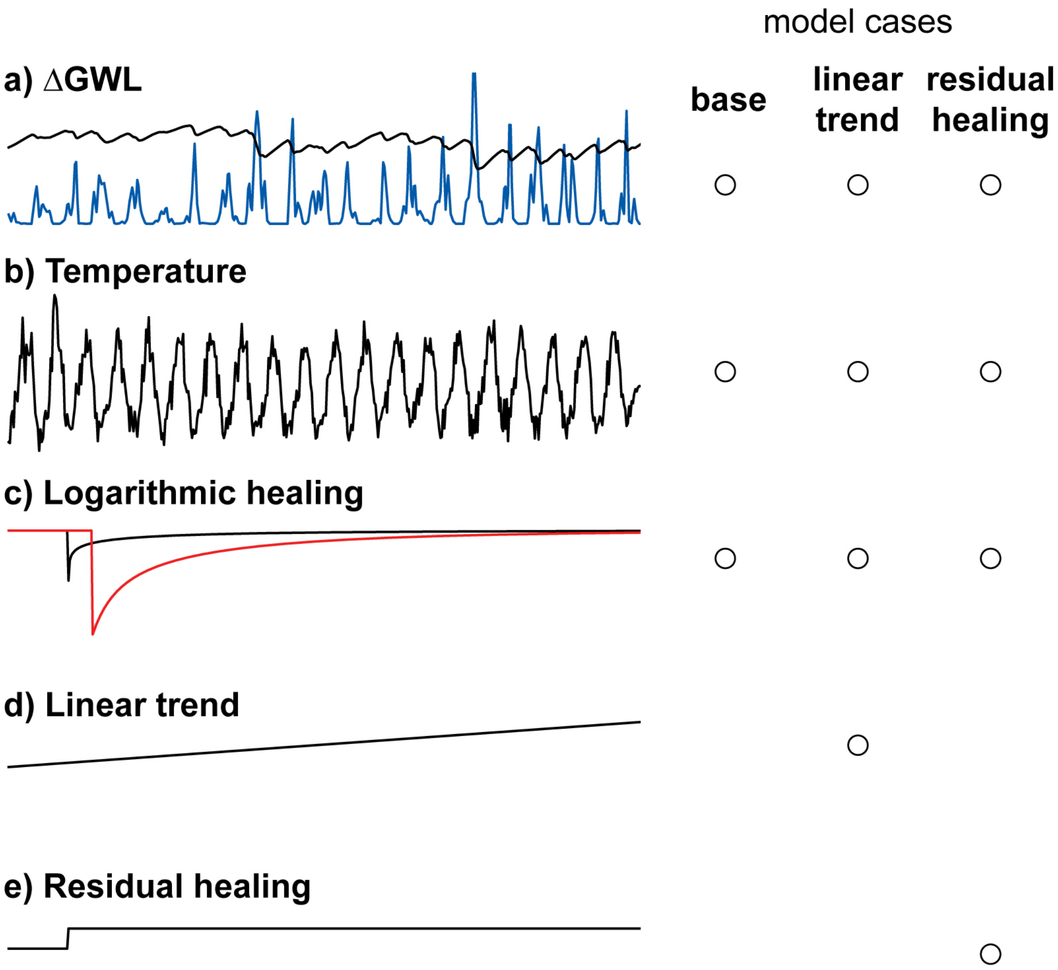

Figure 6: The schematic of fitting model components. The base model consists of the factors associated with (a)ΔGWL, (b)temperature, and (c)logarithmic healing. To better fit the long-term increase shown in the time history of dv/v, we added either (d)linear trend or (e)residual healing terms to the base model.

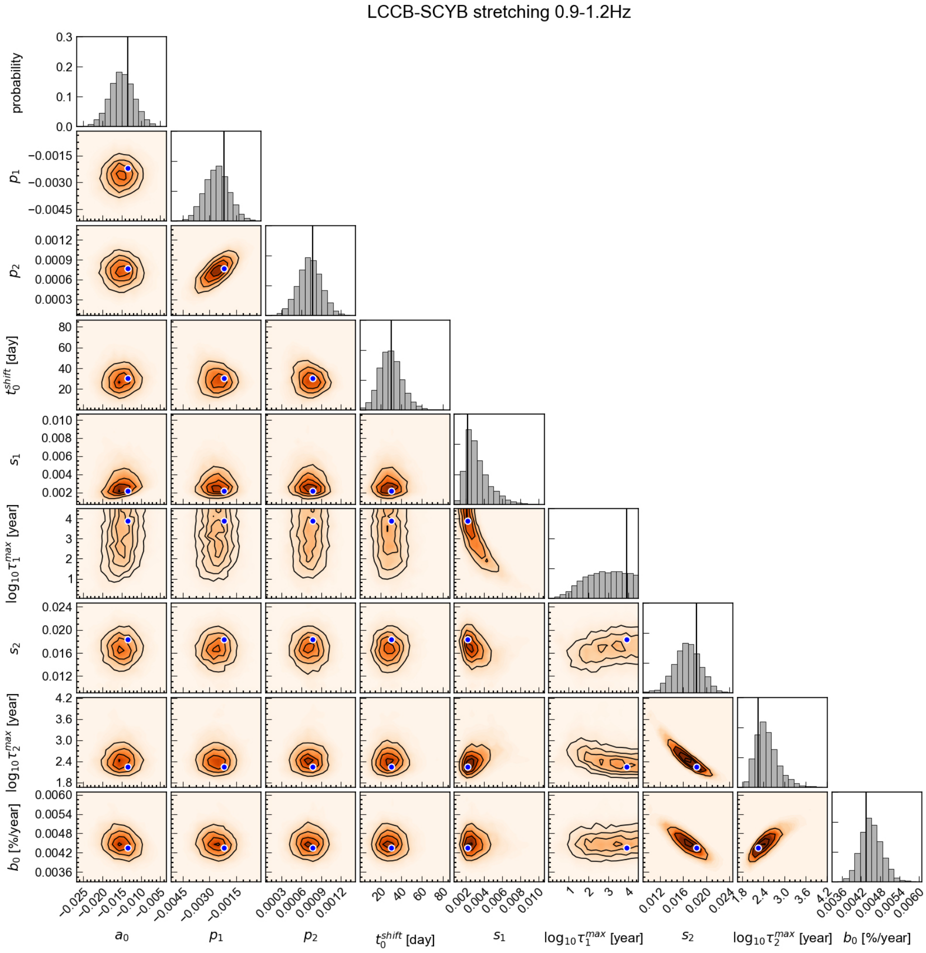

Figure 7: The 1D and 2D marginalized posterior probability distributions of the model parameters obtained from the MCMC analysis for the station pair of LCCB-SCYB with 0.9-1.2Hz for the case with the stretching method. The diagonal panels show the probability mass functions where the sum of bars is normalized to be unity. The vertical solid lines indicate the best likelihood value. The color contour shows the 2D histogram of the pair of model parameters. The darker colors indicate larger probabilities. The circles with blue show the best likelihood model parameters. Note that the best likelihood parameters do not always correspond to the peak of the probability distributions.

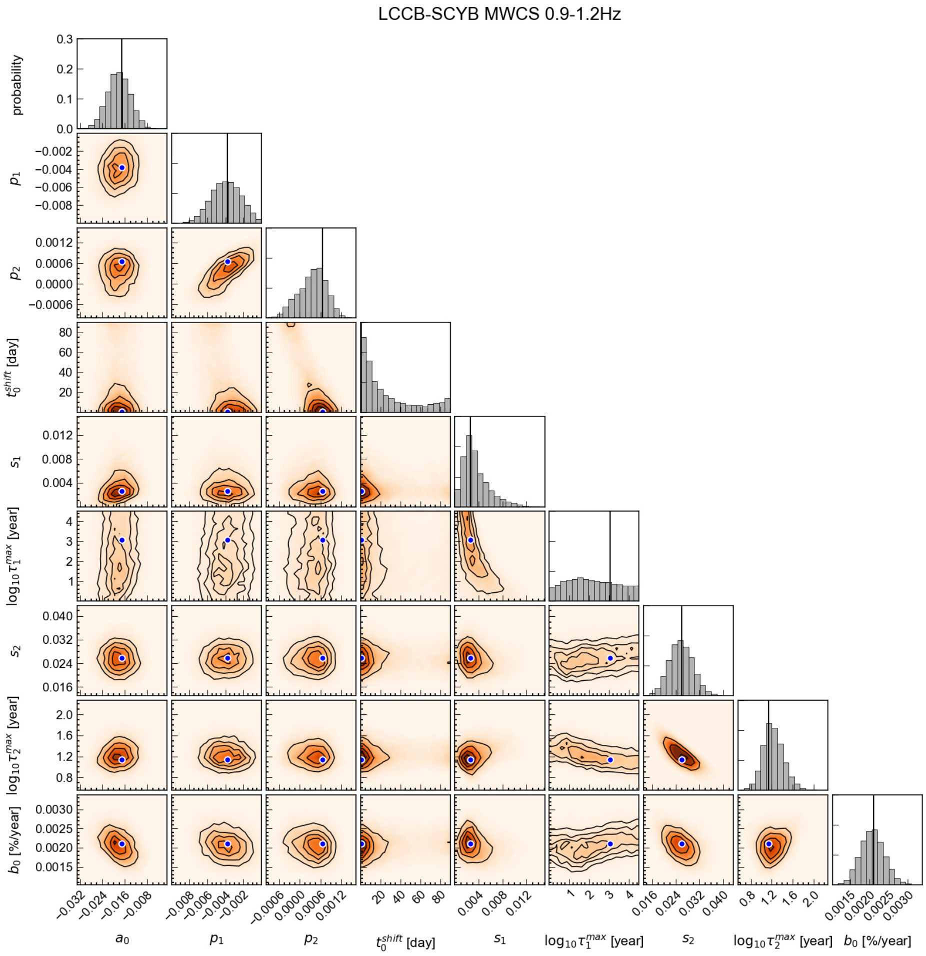

Figure 8: Same to the Figure 7 for the case using the MWCS derived dv/v time series.

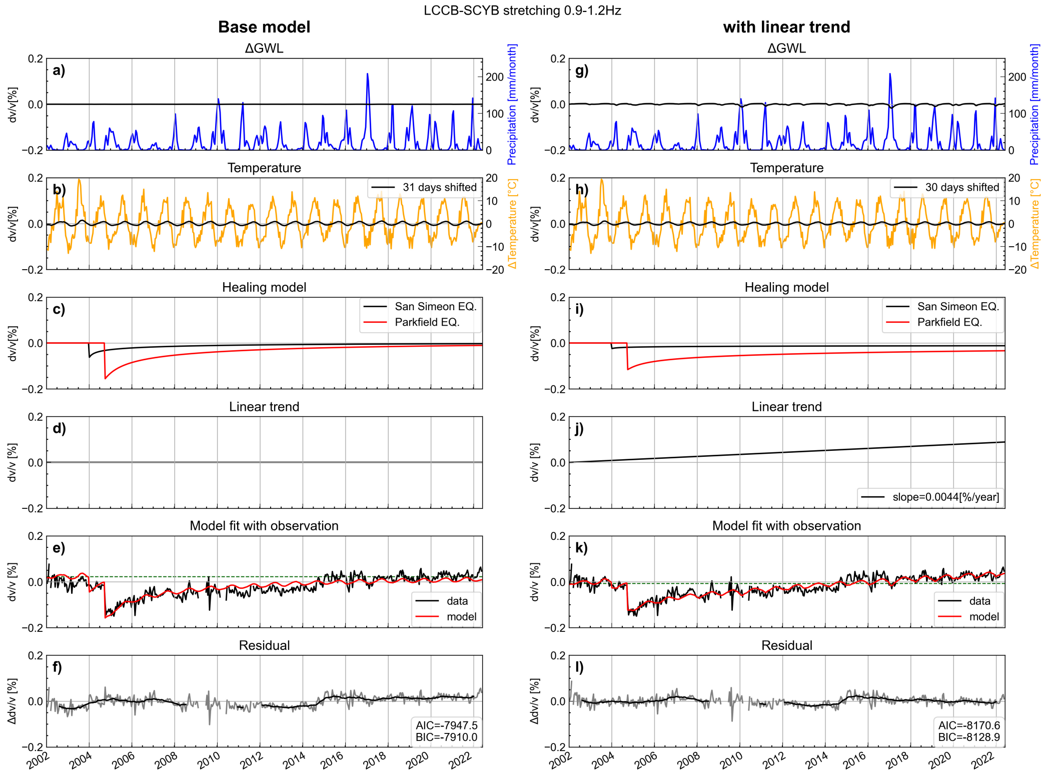

Figure 9: The contributions of the model factors and the comparison to the data for the cases with the base model (left) and the model with linear trend term (right) using the stretching method. Figures (a-c, g-i) show the factors associated with ΔGWL, temperature, and logarithmic healing. Figures (d, j) show the contribution of the linear trend term, which is fixed at zero for the base model. Figures (e, k) show the comparison between the time history of channel-weighted dv/v (black) and the model with the best likelihood parameters (red). The green dotted line indicates the offset of dv/v associated with a_0. The mean removes the offset until the SS earthquake of the observation. Figures (f, l) show the residuals between the best likelihood model and the estimated dv/v time history. Grey and black lines indicate the raw and the smoothed time history of the residuals, respectively.

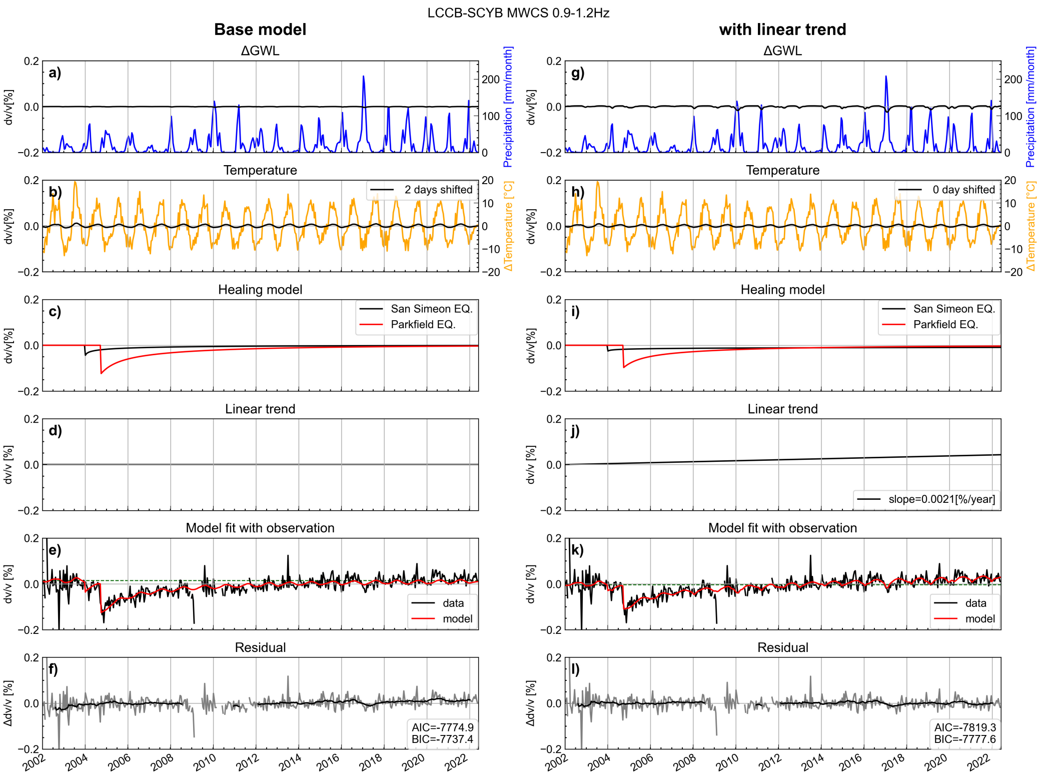

Figure 10: Same to the Figure 9 for the case with the MWCS.

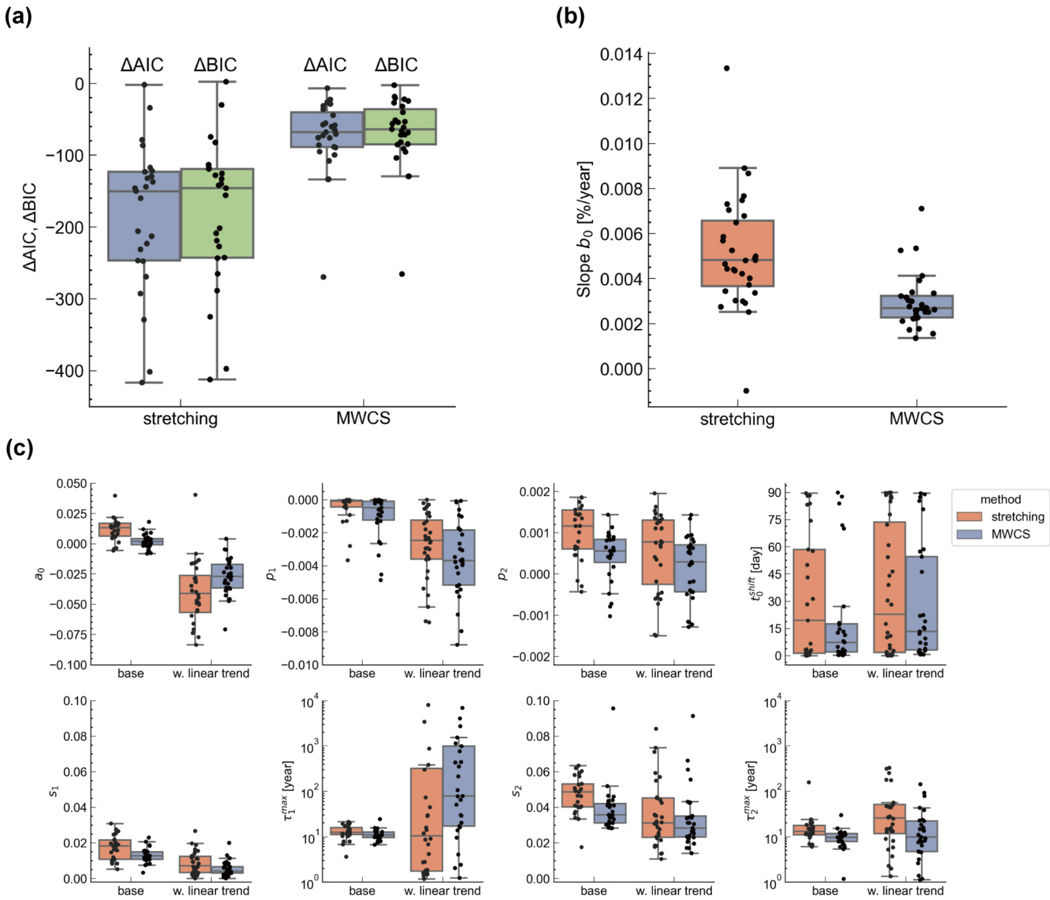

Figure 11: Statistical analysis of model parameters. (a) ΔAIC : $\mathrm{AIC}_{\text{linear trend}} - \mathrm{AIC}_{\text{base}}$ and ΔBIC: $\mathrm{BIC}_{\text{linear trend}} - \mathrm{BIC}_{\text{base}}$ for the cases with the stretching and MWCS. The box plot shows the second and third quartiles with the median value indicated by the horizontal bars in the box associated with the available station pairs meeting the variance threshold with the residuals. The black dots superimposed on the box plot show the individual values. (b) The slope of the linear trend b_0. (c) The best likelihood model parameters. In each panel, we showed the statistics of parameters associated with the base model and the model with the linear trend for the cases with the stretching and MWCS, respectively.

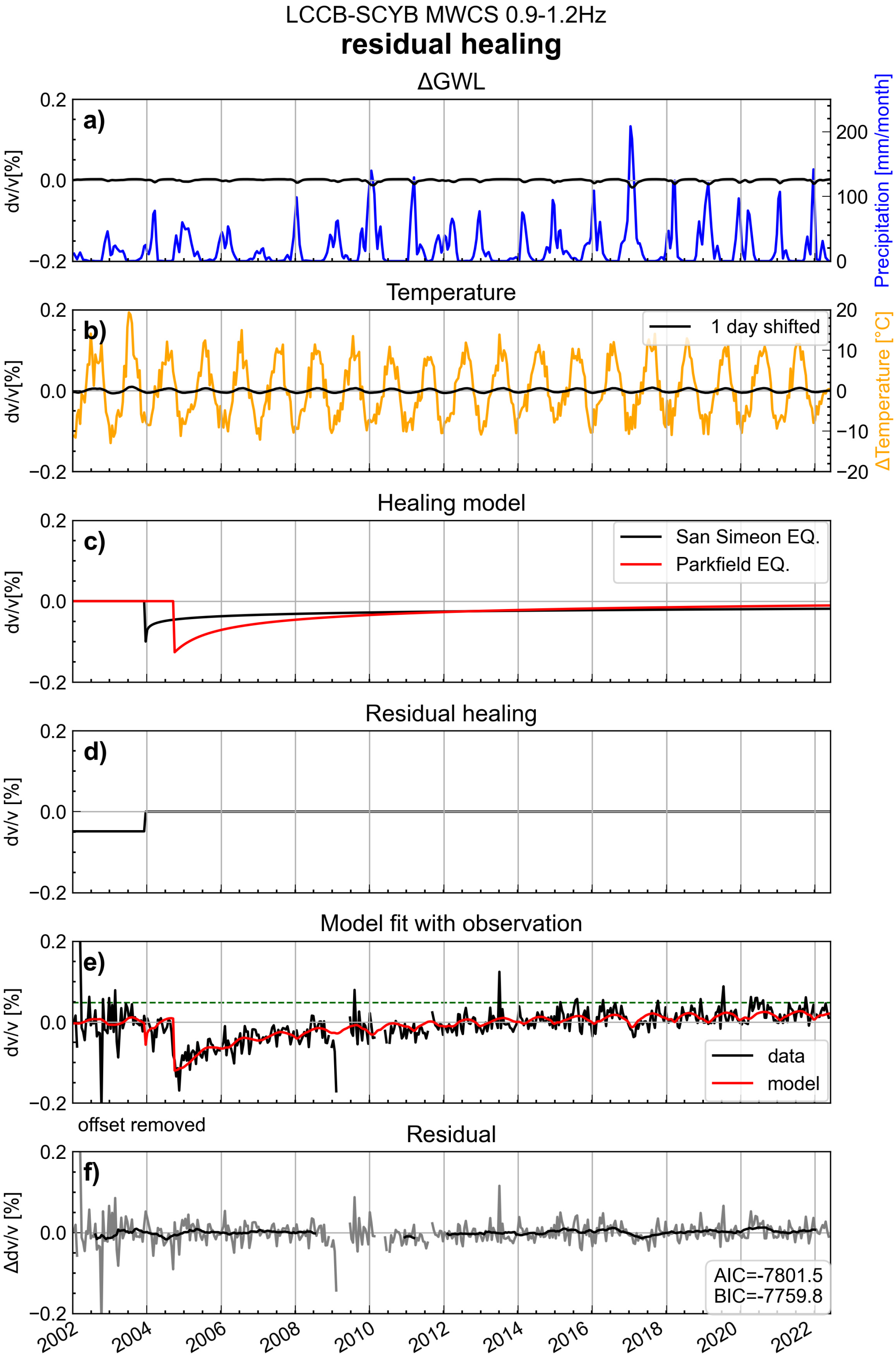

Figure 12: Same to the Figure 10 for the case with the residual healing term. Note that the a_0 is relatively large to complement the long-term increase of dv/v.

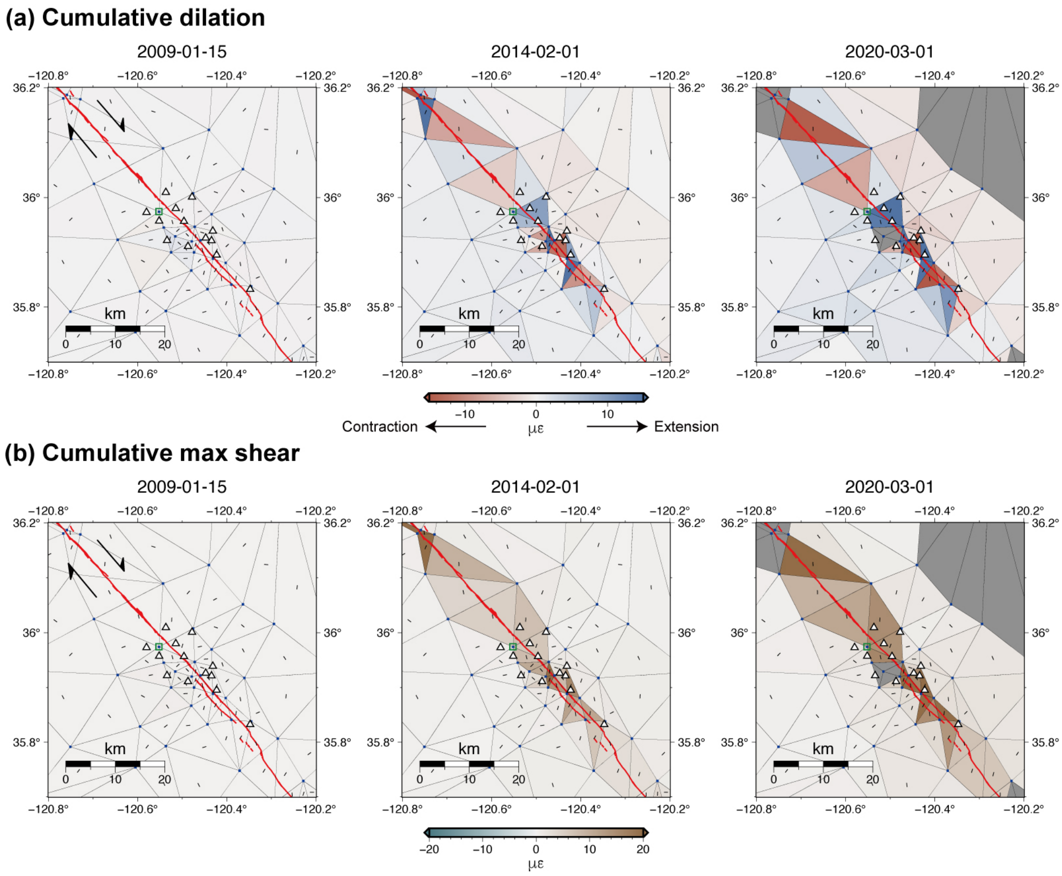

Figure 13: Cumulative strains associated with (a) the dilation and (b) the max shear. Note that the dilation is positive in extension. The triangular elements are formed with the nodes of GNSS stations indicated with the blue squares. The markers with white triangles indicate the seismic stations. The green square shows the location of SAFOD. The markers with white triangles indicate the seismic stations. The arrows in the left panels show the sense of slip as the right-lateral. The lines in the elements show the orientations of the maximum contractional strain. The color contour shows the amplitude of cumulative strains, which is obtained by subtracting the initial state of strains on January 3, 2009, from the current time snapshot. Max shear can be negative as it is the relative value from the initial state. As we evaluate the strain field with the assumption of the continuum in the elements, the localized slip can cause bias in the estimation of strain.

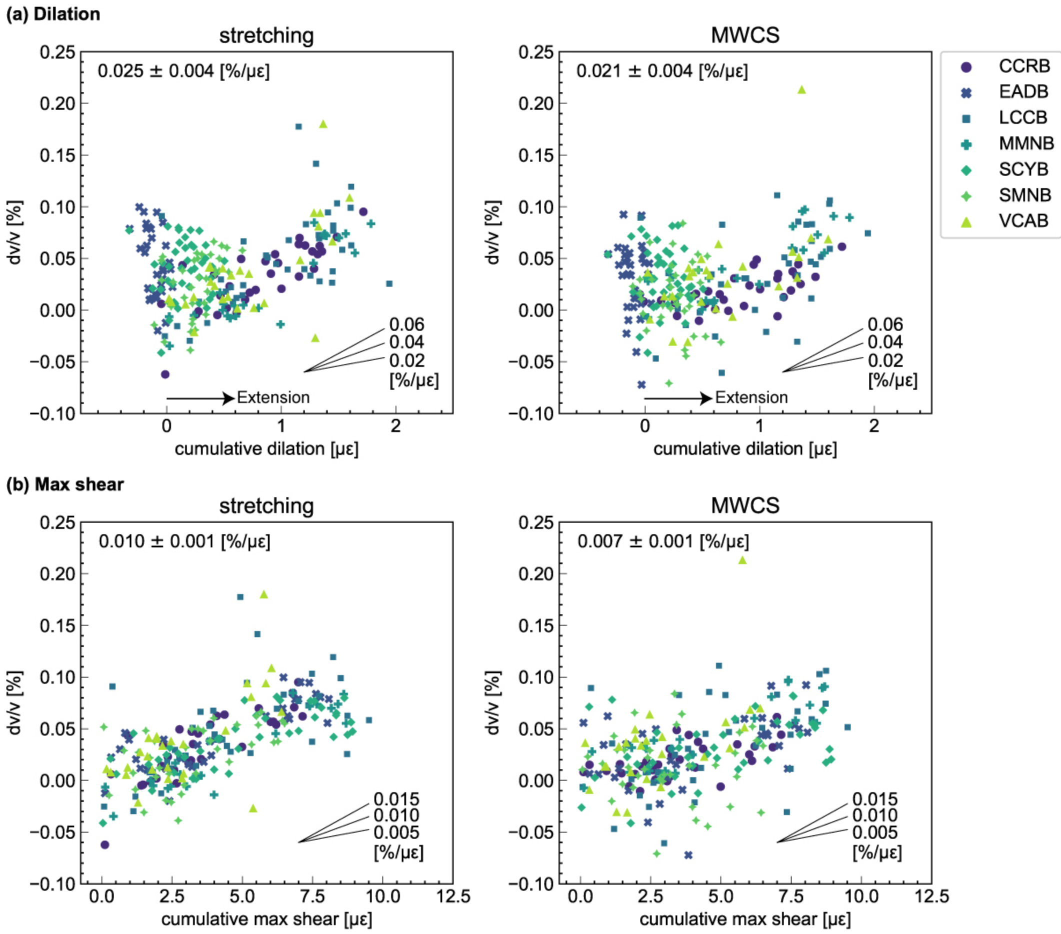

Figure 14: Comparison of the cumulative strains associated with the (a) dilation and (b) max shear to the dv/v with the frequency band of 0.9-1.2Hz. We used the dv/v of the seven stations listed in the legend, while we excluded the GHIB, JCSB, and VARB due to the low quality of the dv/v measurements. We compared the strain and dv/v at the synchronized time periods every 90 days. The reference slopes are indicated for the sensitivity of dv/v to the strain. The slope and standard error of the linear regressions using all seven stations are annotated in the panels.

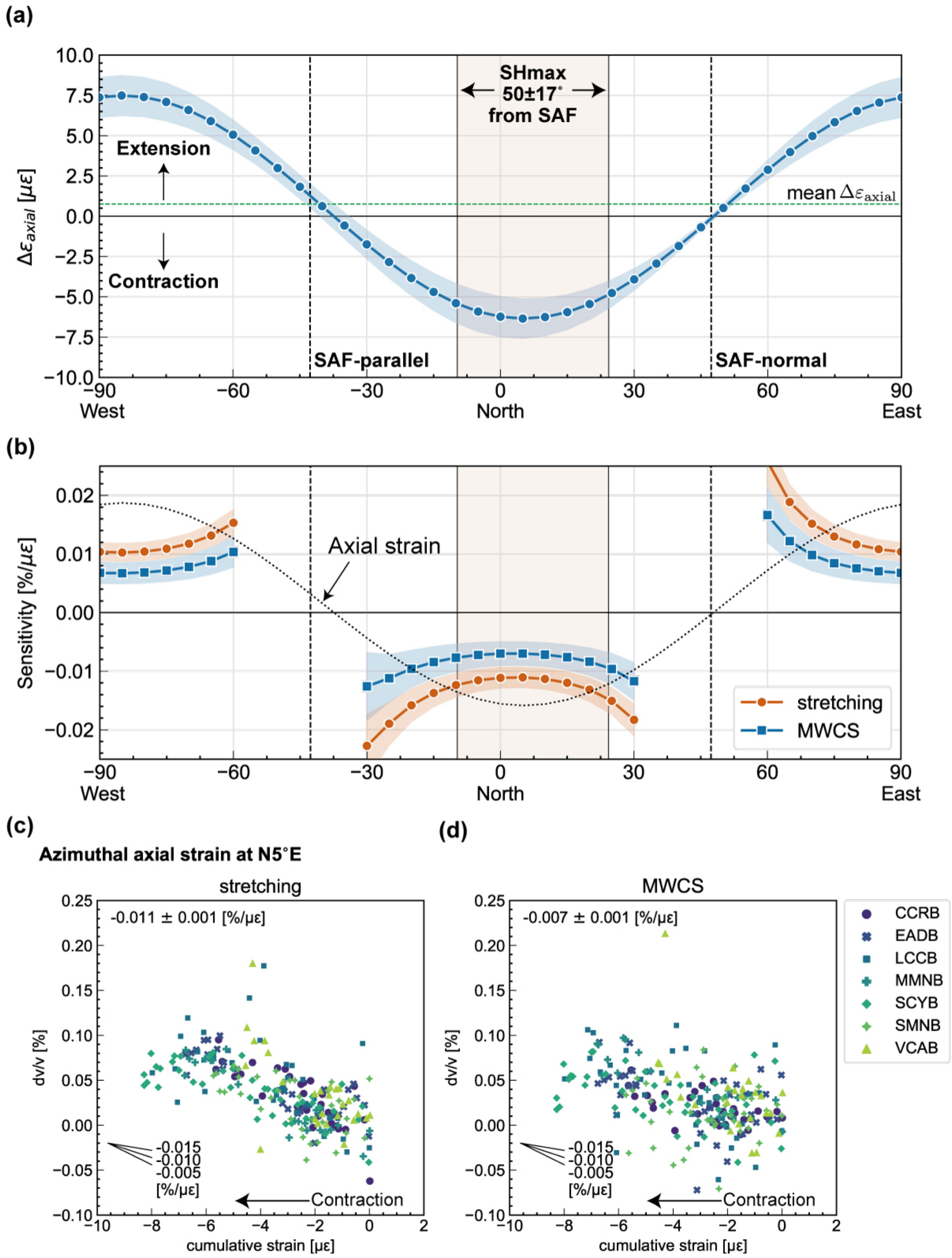

Figure 15: (Caption on next page.)

Figure 15: The rotated axial strain and the variation of sensitivity with the azimuth. (a) Cumulative axial strain parallel to the n rotation angle evaluated between January 2009 and November 2019. We computed the Gaussian-weighted axial cumulative strain parallel to every five degrees as the circles. The filled band shows the standard deviation of the strain associated with the seven seismic stations used in Figure 14. The vertical dashed lines show the orientations of parallel and normal to the San Andreas Fault. The highlighted area shows the range of S_{Hmax} at depth from 0.8 to 2.2km estimated by Hickman and Zoback (2004). The horizontal dashed line indicates the mean of axial strain over the rotated angle. (b) The sensitivity of dv/v with the frequency band of 0.9-1.2Hz to the axial strain. The solid lines with circle and square markers indicate the sensitivities associated with the stretching and MWCS, respectively. The error band indicates the 95% confidence interval of the sensitivity associated with the linear regression of the slope. We excluded the azimuth between $\pm 30°$-$60°$ as the sensitivity diverged due to the small axial strain. The dotted line shows the rotated axial strain similar to (a). (c, d) sensitivity of the velocity to the axial strain nearly parallel to the S_{Hmax}. The negative sensitivity indicates the dv/v increases with the contractional strain.

Supporting Information for "Monitoring velocity change over 20 years at Parkfield"

Kurama Okubo[1], Brent G. Delbridge[2], Marine A. Denolle[3]

[1]National Research Institute for Earth Science and Disaster Resilience (NIED), Tsukuba, Japan

[2]Los Alamos National Laboratory, NM, USA

[3]Department of Earth and Space Sciences, University of Washington, WA, USA

Contents of this file

1. Text S1 to S6
2. Figures S1 to S14

Additional Supporting Information (Files uploaded separately)

1. Captions for Datasets S1 to S5

Text S1. Scaling of process parallelization for the cross-correlation

The processing of ambient seismic noise, in particular for cross-correlation, is computationally expensive as the number of pairs quadratically increases with the number of stations and channels. Thus, parallelization is crucial to maximizing computational efficiency in processing a large dataset (e.g., 1.6TB in this study). The parallelization scheme should be optimized with the use of memory and the frequency of access to storage (rotating Hard Drives), as the parallelization of cross-correlation pairs can cause the redundant calculations of FFT in different parallelized tasks.

The first level of parallelization is done over the time chunks (e.g., one year) per node, and we processed all possible station-channel pairs within a node. Note that this parallelization scheme may cause issues when handling a much larger number of stations and channels, where we need to split the station-channel pairs into nodes even with the minimal time chunk.

The second level of parallelization uses multi-processing (distributed memory) tools from Julia. We parallelized the processes using `Distributed.pmap` function implemented in Julia. We parallelized almost all the process flow, including downloading the data, removing the transient signals, computing the FFT, cross-correlation, stacking, and measuring the dv/v using this metric.

We seek to verify the parallelization benefits of speed-up in a strong scaling test. A perfect strong scaling yields a speed-up of $1/p$, where p is the number of processes (cores in our case). Figures S14a shows the case using a single node to process the subset of the dataset for five days with three components of cross-correlation for all station-channel pairs. We repeated five times for each case with the given number of cores following the initial draft run to subtract the compilation time of Julia modules. The scaling component is $p = 0.94$, which is close to the ideal value. The small decrease in the scaling exponent is caused by the lack of parallelization of the pre-and post-processes, such as the preparation of the list of correlation pairs.

Figures S14b shows the scaling with the large dataset from 2002 to 2020 with three components for all station-channel pairs. Note that in the first level of parallelization associated with the time chunks, we just separately submit the jobs to conduct them simultaneously. The CPU time and its error bar are obtained with the mean and standard deviation of the time for the jobs without considering the waiting time in the queue. It also shows the scaling component of $p = 0.936$, which is almost ideal for process parallelization. The improvement of computational efficiency helps perform the case studies and reproduce the analysis.

Text S2. Comparison of the LFE activity to the time history of the dv/v

Delbridge, Carmichael, Nadeau, Shelly, and Bürgmann (2020) showed the variation of the tremor and LFE rates. They pointed out an increase in the background tremor rate after the Parkfield and the 2014 M6 South Napa earthquake along with a burst of LFE activity directly following the Napa earthquake. In contrast to the tremor rate, however, the background LFE rate after the Napa earthquake returns to a similar background level as before the Napa earthquake. They also showed a decrease in the background LFE rate from 2013 to early 2015, where the inactivity after the Napa earthquake coincided with the increase in the tremor rate.

We analyzed the episodic LFE activity from May 2001 through September 2020 using the catalog extended from Shelly (2017) and compared it to the time history of dv/v as shown in Figure S6. The detailed pattern of LFE activities is not correlated with the dv/v as the duration of tremor and LFE is much shorter than the stacking period of the CFs. Thus, the artifacts of the perturbation in the source due to the tremor and LFE are more likely to be suppressed in the estimation of dv/v. To address further comparison between the dv/v and e.g. the rate in the cumulative LFE, we need to remove the environmental factors such as temperature and precipitation from the pairwise dv/v to isolate the factors that might be associated with the strains from the tectonic activity associated with these bursts of LFEs.

Text S3. Numerical integration of logarithmic healing model

The integration of the logarithmic healing model is computationally expensive. If $\tau^{\min} \ll t \ll \tau^{\max}$, we can approximate the healing model as $B - \ln t$ with the B of offset (Snieder et al., 2017), whereas the assumption is inadequate in our case as the $\tau^{\max}$ can be shorter than the length of dv/v time history. Therefore, we improved the computational efficiency of the integration by the pre-compiled C library, called the low-level callback function implemented in the `scipy`, which allows for the MCMC analysis with the original form of the healing model.

Text S4. Trade-off between s_i, $\tau_i^{\mathbf{min}}$ and $\tau_i^{\mathbf{max}}$

Our preliminary analysis of MCMC without fixing the $\tau_i^{\min}$ shows the trade-off between the s_i, which degrades the convergence of the parameter sampling. The trade-off is caused due to the noise of the dv/v in the early part of the healing, which is not enough to constrain the rate of healing governed by the $\tau_i^{\min}$. We thus fixed the $\tau_i^{\min}$ for the terms associated with the SS and PF earthquakes, as listed in Table 3 in the main text. In this section, we show the case study of the trade-off using the representative time history of dv/v.

We used the dv/v of ACFs with the station of VCAB. Figure S11a shows the raw data with the channel weighting of the nine components CFs and the dv/v after removing the model components associated with the ΔGWL, temperature, coseismic decrease, the logarithmic healing of the SS earthquake, and the linear trend term to isolate the healing associated with the PF earthquake. We fit the model composed only with the logarithmic healing associated with PF earthquake to the time history of dv/v between 2003 and 2010. We run the MCMC parameter sampling with the 16 walkers with 10000 iterations. Figure S11b shows the scatter matrix, indicating the strong trade-off of the s_2 to τ_2^{min} and τ_2^{max}. We selected the sets of model parameters with high probability as shown in the blue circles in Figure S11b, and synthesized the models as shown in Figure S11c. As the noise level of the data is larger than the variation in the healing, the parameters of s_2 and τ_2^{min} are less likely to be constrained by the fitting. If we investigate the early healing of dv/v with postseismic behavior, we need a more stable measurement of dv/v or different metrics to evaluate the rate of recovery.

Text S5. Sensitivity of the dv/v to the dilational strain

Rivet et al. (2011) found -$\sim$0.1%/$\mu\varepsilon$ to the dilation associated with the study of the dv/v and the slow slip event in Guerrero region. Hirose, Nakahara, and Nishimura (2017) found -$\sim$0.2%/$\mu\varepsilon$ with the areal strain estimated from the GNSS data at Sakurajima volcano. Takano, Nishimura, and Nakahara (2017) found -$\sim$0.2%/$\mu\varepsilon$ for the case with Izu-Oshima, Japan. Mao et al. (2019) found -(0.1-1)%/$\mu\varepsilon$ with tidal strain at Piton de la Fournaise volcano on La Réunion. Sens-Schonfelder and Eulenfeld (2019) found -0.8%/$\mu\varepsilon$ for the case in the Atacama desert in northern Chile. The measurement of dv/v with artificial seismic sources around Iwate volcano, Japan, also shows a similar range of sensitivity such as -(0.14-1.1)%/$\mu\varepsilon$ (Nishimura et al., 2005). Besides, Takano, Nishimura, Nakahara, Ueda, and Fujita (2019) shows the sensitivity of dv/v with the tidal strain varies with the lapse time of CFs, and the range of sensitivity can be up to -$\sim$2%/$\mu\varepsilon$. Takano, Nishimura, Nakahara, Ohta, and Tanaka (2014) also shows a relatively large sensitivity of -6.9%/$\mu\varepsilon$ for the case with the foot of Mount Iwate, Japan.

Donaldson, Winder, Caudron, and White (2019) showed the comparison of the modeled dilation caused by the dike intrusion and the velocity change in the Northern volcanic zone in Iceland, showing the relatively small value of -0.016 $\pm$ 0.001%/$\mu\varepsilon$ for the frequency band of 0.4-1.0Hz though their sense of velocity increase follows the negative dilation (i.e., contraction) opposite to our sensitivity. In their case, the sensitivity is measured for the step of dv/v associated with a specific transient event rather than a long-term increase. They explained the estimation of sensitivity could be underestimated due to relaxation processes of the crust, or the lateral variation of the strain change.

Text S6. Scattered crack orientation at Parkfield

One of the metrics to evaluate the orientations of microcracks in the crust is the shear wave splitting (SWS), where the fast polarization direction is parallel to the microcracks in the medium (Crampin, 1978; Sayers & Kachanov, 1995). Boness and Zoback (2004) conducted the logging of SWS at SAFOD Pilot hole, showing the fast polarization direction follows the maximum compressive stress direction, S_{Hmax}. Boness and Zoback (2006) extended the study indicating the fast polarization direction varies from S_{Hmax} to the sedimentary bed strike with depth. They suggested that anisotropy is caused by stress-induced aligned cracks or local structures, such as shear fabric along the fault. Cochran, Li, and Vidale (2006) analyzed the polarization direction using two temporal arrays around the SAFOD and Parkfield, showing some dominant orientations parallel to the S_{Hmax}, SAF and a branch from SAF at the Parkfield array, while the array at SAFOD shows more scattered orientations. Liu, Zhang, Thurber, and Roecker (2008) showed the spatial distribution of SWS at Parkfield considering the ray path from earthquakes to the station. While part of SWS observations are parallel to the S_{Hmax} or SAF, the overall distribution of the polarization orientations is scattered. Zhang, Liu, Thurber, and Roecker (2007) conducted the three-dimensional tomography of the anisotropic structure at Parkfield using the delay times of SWS obtained by Liu et al. (2008). They showed the heterogeneous structures of anisotropy can be caused due to the fault shear fabric, stress-induced anisotropy parallel to S_{Hmax}, and the alignment of minerals such as serpentinite. A detailed discussion on the cause of anisotropy in the medium can be found in Crampin, Chesnokov, and Hipkin (1984), Crampin (1987), and Boness and Zoback (2004). Given that microcracks play a role in the polarization direction of SWS, the scattered distribution of the polarization would indicate the crack orientations are non-uniform at Parkfield, even though some follow the S_{Hmax} or the fault strikes. The scattered SWS is also discussed in different parts of the fault zones (Peng & Ben-Zion, 2004).

Dataset S1. Spectrogram from 2002 to 2022 associated with the High-Resolution Seismic Network

We estimated the power spectral density (PSD) of all the available channels for 13 stations used in this study. The PSD with the moving time window is evaluated using the Blackman-Tukey method, shown as the spectrogram. The color contour and the annotations are similar to Figure S2. Note that the PSD is computed before removing the earthquakes and tremors.

Dataset S2. Auto- and cross-correlations of all station-channel pairs

We computed the auto- and cross-correlation functions for all the available station-channel pairs, which are shown as nine-component correlation functions associated with the given station pair, similar to Figure S5. The original data of cross-correlation functions are also available in the cloud storage `dasway` (doi: `https://doi.org/10.6069/PK9D-9411`)

Dataset S3. datasheets of dv/v with the stretching method

The datasheet of dv/v associated with the stretching method, which contains the values of dv/v for all the station-channel pairs and their errors for the frequency bands of 0.2-0.5Hz, 0.5-0.9Hz, 0.9-1.2Hz, and 1.2-2.0Hz.

Dataset S4. datasheets of dv/v with the MWCS method

The datasheet of dv/v associated with the MWCS method, similar to the dataset S3.

Dataset S5. Best likelihood model parameters

This dataset contains the best likelihood model parameters obtained with the MCMC analysis shown in Figure 11 of the main text.

References

Boness, N. L., & Zoback, M. D. (2004). Stress-induced seismic velocity anisotropy and physical properties in the SAFOD Pilot Hole in Parkfield, CA. *Geophys. Res. Lett.*, *31*(15). doi: 10.1029/2003GL019020

Boness, N. L., & Zoback, M. D. (2006). A multiscale study of the mechanisms controlling shear velocity anisotropy in the San Andreas Fault Observatory at Depth. *GEOPHYSICS*, *71*(5), F131-F146. doi: 10.1190/1.2231107

Cochran, E. S., Li, Y.-G., & Vidale, J. E. (2006). Anisotropy in the Shallow Crust Observed around the San Andreas Fault Before and After the 2004 M 6.0 Parkfield Earthquake. *Bull. Seismol. Soc. Am.*, *96*(4B), S364-S375. doi: 10.1785/0120050804

Crampin, S. (1978). Seismic-wave propagation through a cracked solid: polarization as a possible dilatancy diagnostic. *Geophys. J. Int.*, *53*(3), 467-496. doi: 10.1111/j.1365-246X.1978.tb03754.x

Crampin, S. (1987). Geological and industrial implications of extensive-dilatancy anisotropy. *Nature*, *328*(6130), 491–496. doi: 10.1038/328491a0

Crampin, S., Chesnokov, E. M., & Hipkin, R. G. (1984). Seismic anisotropy — the state of the art: II. *Geophys. J. Int.*, *76*(1), 1-16. doi: 10.1111/j.1365-246X.1984.tb05017.x

Delbridge, B. G., Carmichael, J. D., Nadeau, R. M., Shelly, D. R., & Bürgmann, R. (2020). Geodetic Measurements of Slow-Slip Events Southeast of Parkfield, CA. *J. Geophys. Res.*, *125*(5), e2019JB019059. doi: 10.1029/2019JB019059

Donaldson, C., Winder, T., Caudron, C., & White, R. S. (2019). Crustal seismic velocity responds to a magmatic intrusion and seasonal loading in Iceland’s Northern Volcanic Zone. *Science Advances*, *5*(11), eaax6642. doi: 10.1126/sciadv.aax6642

Herrmann, R. B. (2013). Computer Programs in Seismology: An Evolving Tool for Instruction and Research. *Seismol. Res. Lett.*, *84*(6), 1081-1088. doi: 10.1785/0220110096

Hirose, T., Nakahara, H., & Nishimura, T. (2017). Combined use of repeated active shots and ambient noise to detect temporal changes in seismic velocity: application to Sakurajima volcano, Japan. *Earth, Planets and Space*, *69*(1), 42. doi: 10.1186/s40623-017-0613-7

Liu, Y., Zhang, H., Thurber, C., & Roecker, S. (2008). Shear wave anisotropy in the crust around the San Andreas fault near Parkfield: Spatial and temporal analysis. *Geophys. J. Int.*, *172*(3), 957-970. doi: 10.1111/j.1365-246X.2007.03618.x

Mao, S., Campillo, M., van der Hilst, R. D., Brenguier, F., Stehly, L., & Hillers, G. (2019). High temporal resolution monitoring of small variations in crustal strain by dense seismic arrays. *Geophys. Res. Lett.*, *46*(1), 128–137. doi: 10.1029/2018GL079944

Nishimura, T., Tanaka, S., Yamawaki, T., Yamamoto, H., Sano, T., Sato, M., ... Sato, H. (2005). Temporal changes in seismic velocity of the crust around Iwate volcano, Japan, as inferred from analyses of repeated active seismic experiment data from 1998 to 2003. *Earth, Planets and Space*, *57*(6), 491–505. doi: 10.1186/BF03352583

Peng, Z., & Ben-Zion, Y. (2004). Systematic analysis of crustal anisotropy along the Karadere—Düzce branch of the North Anatolian fault. *Geophys. J. Int.*, *159*(1), 253-274. doi: 10.1111/j.1365-246X.2004.02379.x

Rivet, D., Campillo, M., Shapiro, N. M., Cruz-Atienza, V., Radiguet, M., Cotte, N., & Kostoglodov, V. (2011). Seismic evidence of nonlinear crustal deformation during a large slow slip event in Mexico. *Geophys. Res. Lett.*, *38*(8). doi: 10.1029/2011GL047151

Sayers, C. M., & Kachanov, M. (1995). Microcrack-induced elastic wave anisotropy of brittle rocks. *J. Geophys. Res.*, *100*(B3), 4149-4156. doi: 10.1029/94JB03134

Sens-Schonfelder, C., & Eulenfeld, T. (2019). Probing the in situ Elastic Nonlinearity of Rocks with Earth Tides and Seismic Noise. *Phys. Rev. Lett.*, *122*. doi: 10.1103/PhysRevLett.122.138501

Shaw, J. H., Plesch, A., Tape, C., Suess, M. P., Jordan, T. H., Ely, G., ... Munster, J.

(2015). Unified Structural Representation of the southern California crust and upper mantle. *Earth Planet. Sc. Lett.*, *415*, 1-15. doi: 10.1016/j.epsl.2015.01.016

Shelly, D. R. (2017). A 15 year catalog of more than 1 million low-frequency earthquakes: Tracking tremor and slip along the deep San Andreas Fault. *J. Geophys. Res.*, *122*(5), 3739-3753. doi: 10.1002/2017JB014047

Snieder, R., Sens-Schönfelder, C., & Wu, R. (2017). The time dependence of rock healing as a universal relaxation process, a tutorial. *Geophys. J. Int.*, *208*(1), 1–9. doi: 10.1093/gji/ggw377

Southern California Earthquake Center. (2023). *Community Velocity Model - Harvard (CVM-H) v15.1.1, accessed June 28, 2023, at: http://moho.scec.org/UCVM_web/web/viewer.php.*

Takano, T., Nishimura, T., & Nakahara, H. (2017). Seismic velocity changes concentrated at the shallow structure as inferred from correlation analyses of ambient noise during volcano deformation at Izu-Oshima, Japan. *J. Geophys. Res.*, *122*(8), 6721-6736. doi: 10.1002/2017JB014340

Takano, T., Nishimura, T., Nakahara, H., Ohta, Y., & Tanaka, S. (2014). Seismic velocity changes caused by the Earth tide: Ambient noise correlation analyses of small-array data. *Geophys. Res. Lett.*, *41*(17), 6131-6136. doi: 10.1002/2014GL060690

Takano, T., Nishimura, T., Nakahara, H., Ueda, H., & Fujita, E. (2019). Sensitivity of Seismic Velocity Changes to the Tidal Strain at Different Lapse Times: Data Analyses of a Small Seismic Array at Izu-Oshima Volcano. *J. Geophys. Res.*, *124*(3), 3011-3023. doi: 10.1029/2018JB016235

Wu, C., Delorey, A., Brenguier, F., Hadziioannou, C., Daub, E. G., & Johnson, P.

(2016). Constraining depth range of S wave velocity decrease after large earthquakes near Parkfield, California. *Geophysical Research Letters*, *43*(12), 6129–6136. doi: 10.1002/2016GL069145

Zhang, H., Liu, Y., Thurber, C., & Roecker, S. (2007). Three-dimensional shear-wave splitting tomography in the Parkfield, California, region. *Geophys. Res. Lett.*, *34*(24). doi: 10.1029/2007GL031951

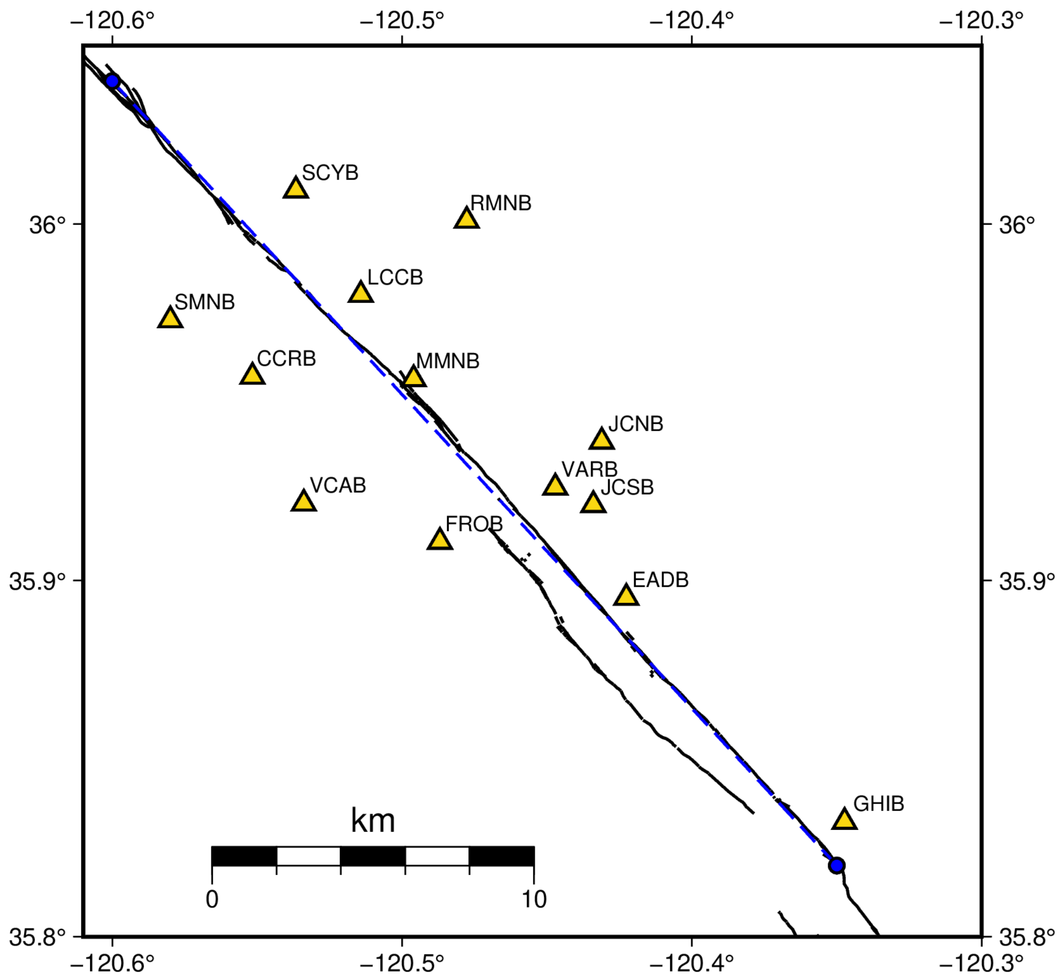

Figure S1. The map of seismic stations and the geometry of the San Andreas fault with the approximated planar fault by the blue dashed line, which is used for computing the fault-normal distance of the seismic stations. The edges indicated by the blue circles are selected by trial and error to fit with the main fault.

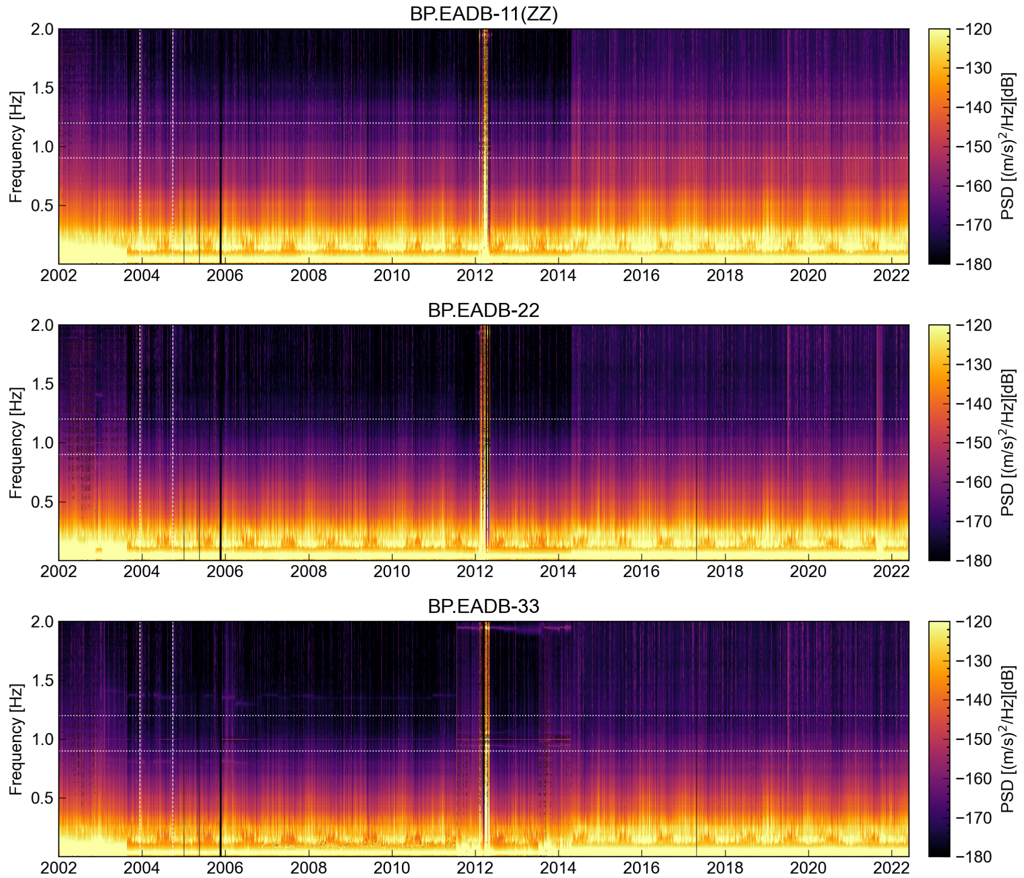

Figure S2. PSD of the continuous waveforms associated with the three components of BP.EADB. The vertical dashed lines indicate the date of the 2003 San Simeon earthquake and the 2004 Parkfield earthquake, respectively. The horizontal dotted lines show the frequency band between 0.9-1.2Hz, which is mainly used for the analysis of dv/v.

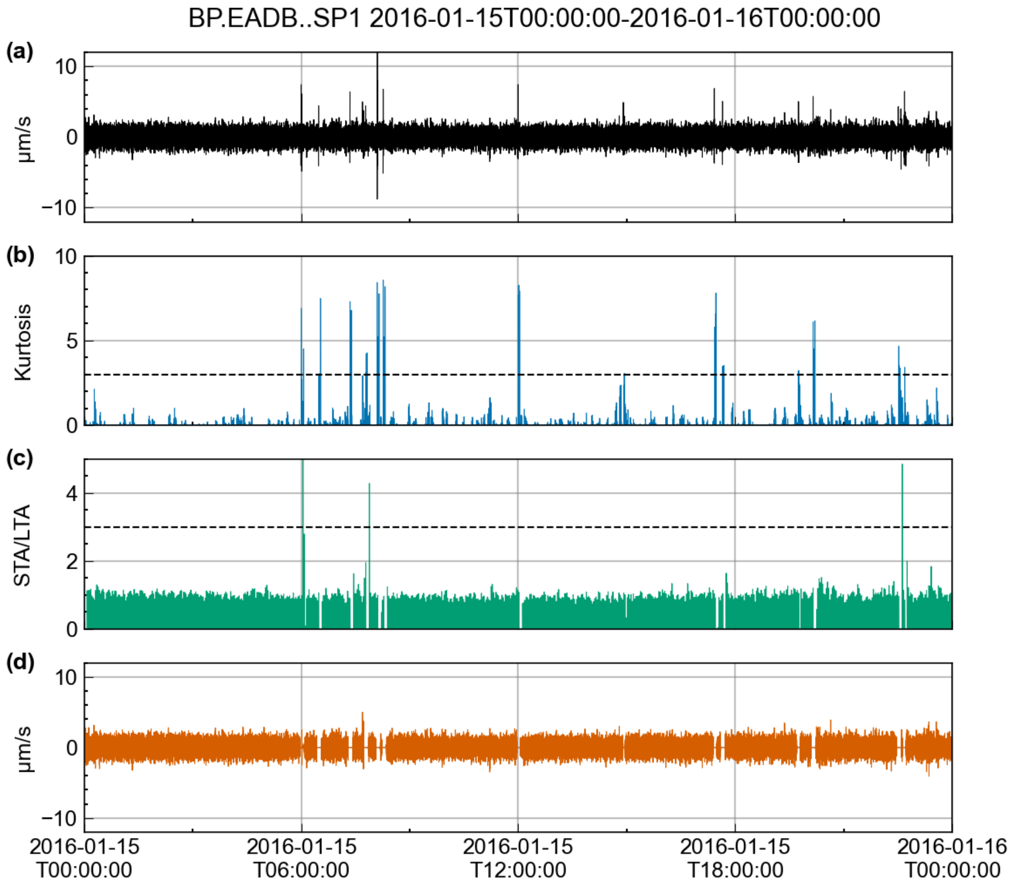

Figure S3. The removal of transient signals in the continuous data. (a) raw data after the removal of the instrumental response. (b) excess kurtosis (c) STA/LTA (d) waveform after the removal of transient signals. The horizontal dashed line in (b) and (c) indicate the threshold of event detection.

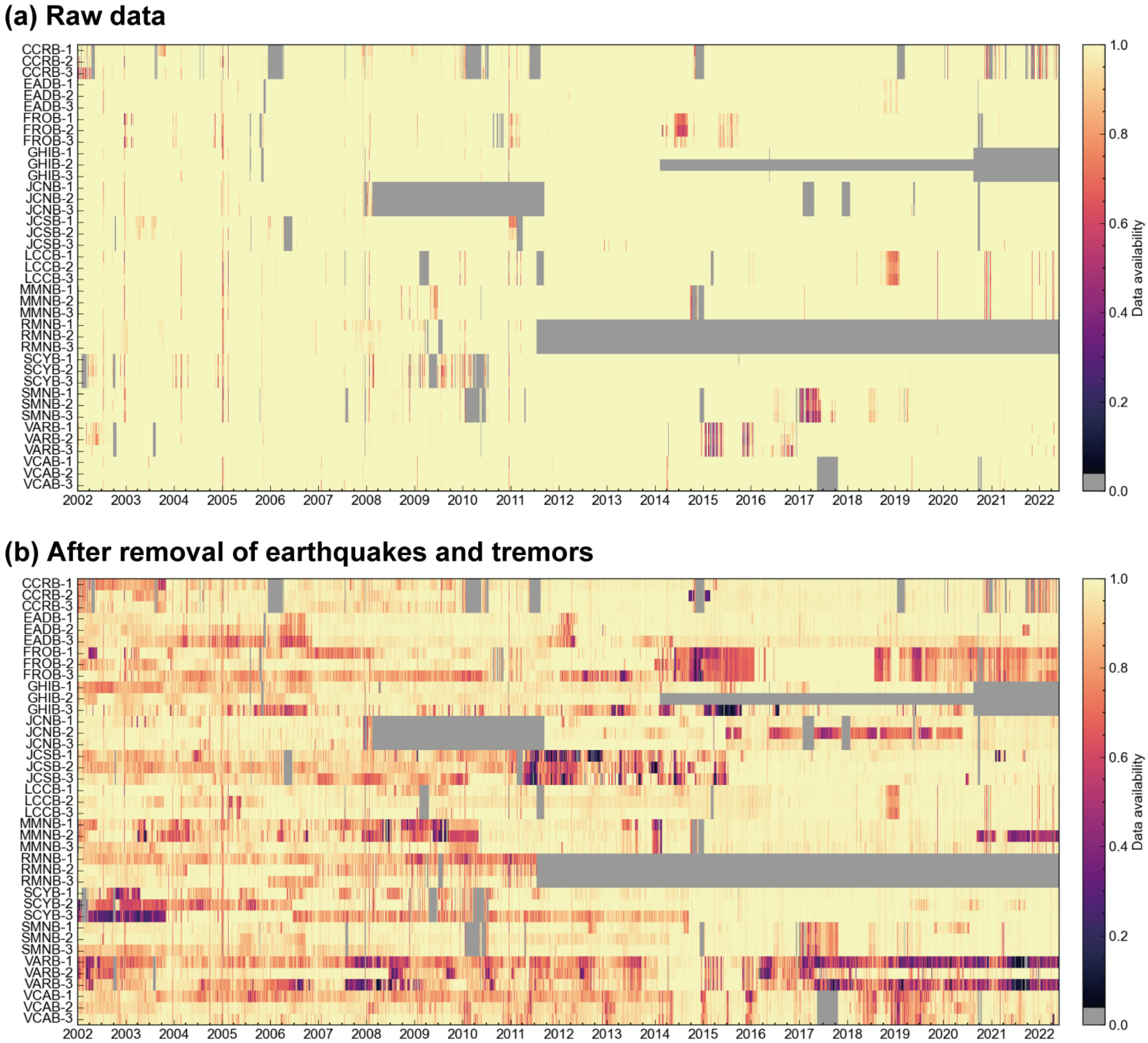

Figure S4. Data availability with (a) the raw data and (b) data after removing the transient signals. We evaluated the fraction of the available ambient noise with respect to the daily continuous waveforms.

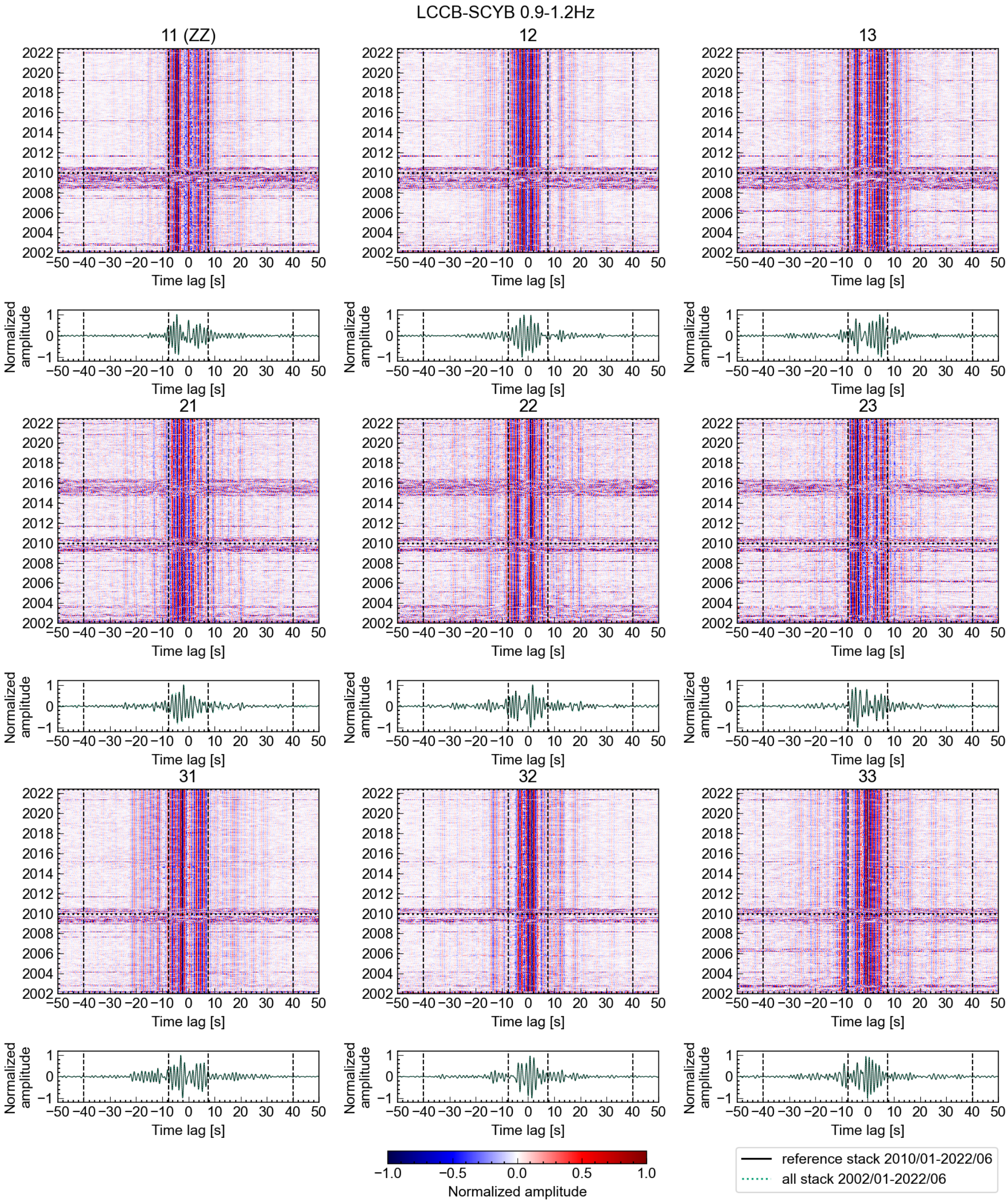

Figure S5. The nine-component CCFs associated with LCCB-SCYB from 2002 to 2022. The color contour shows the daily CCFs normalized by their maximum amplitude. The vertical and horizontal dashed lines indicate the coda window used to evaluate the dv/v and the period of the reference stack, respectively. The bottom waveform of each CCF component shows the reference CCF (solid black) stacked from January 2010 to June 2022 with the median mute and the stack of all the daily CFs over the entire period (dashed green).

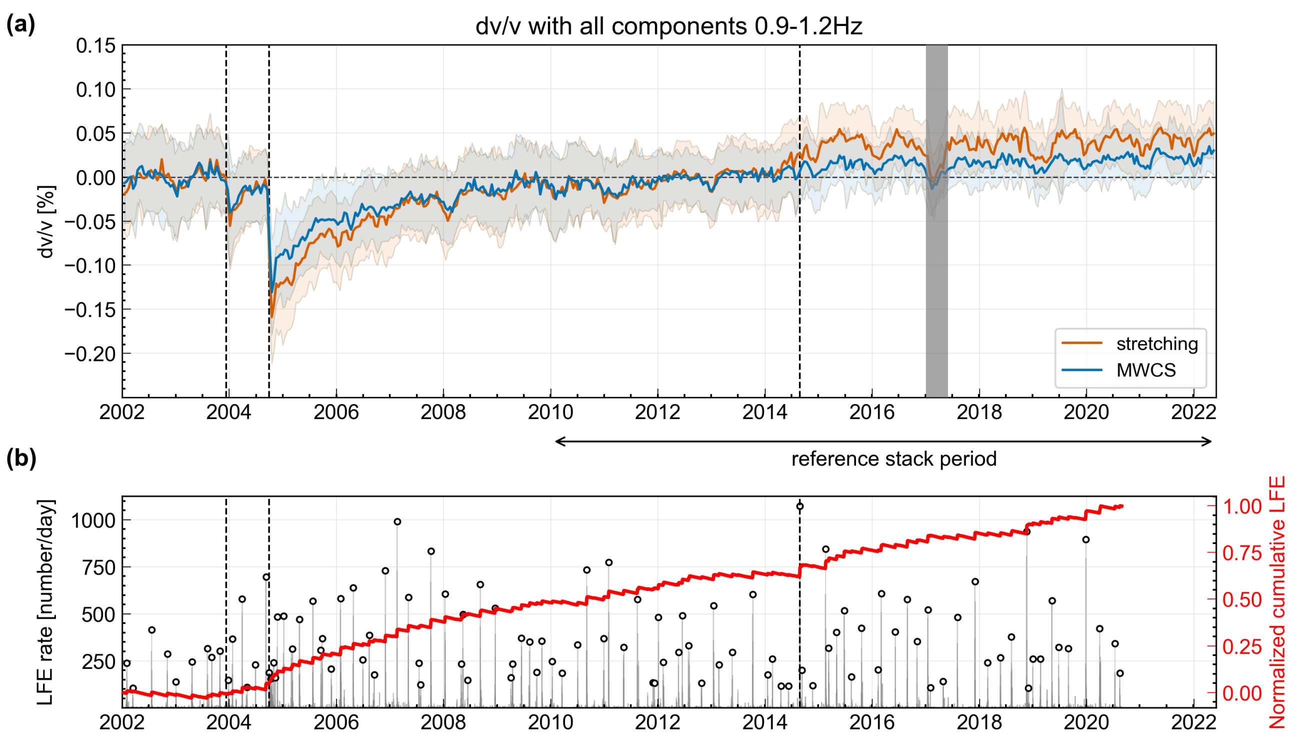

Figure S6. Comparison of the dv/v to the LFE activities. (a) The time history of the dv/v associated with the stretching and MWCS. The solid line and the highlighted area indicate the median and the first and third quartiles, respectively. The vertical dashed lines indicate the date of the SS, PF, and the South Napa earthquakes. (b) LFE activities and the cumulative number of LFEs. The LFE rate is obtained by 2 days average and the peaks are detected with 10 days separation. The cumulative LFE is obtained with the summation of the LFE rate. We detrended the cumulative LFE using the trend before the PF earthquake and normalized it by the maximum value.

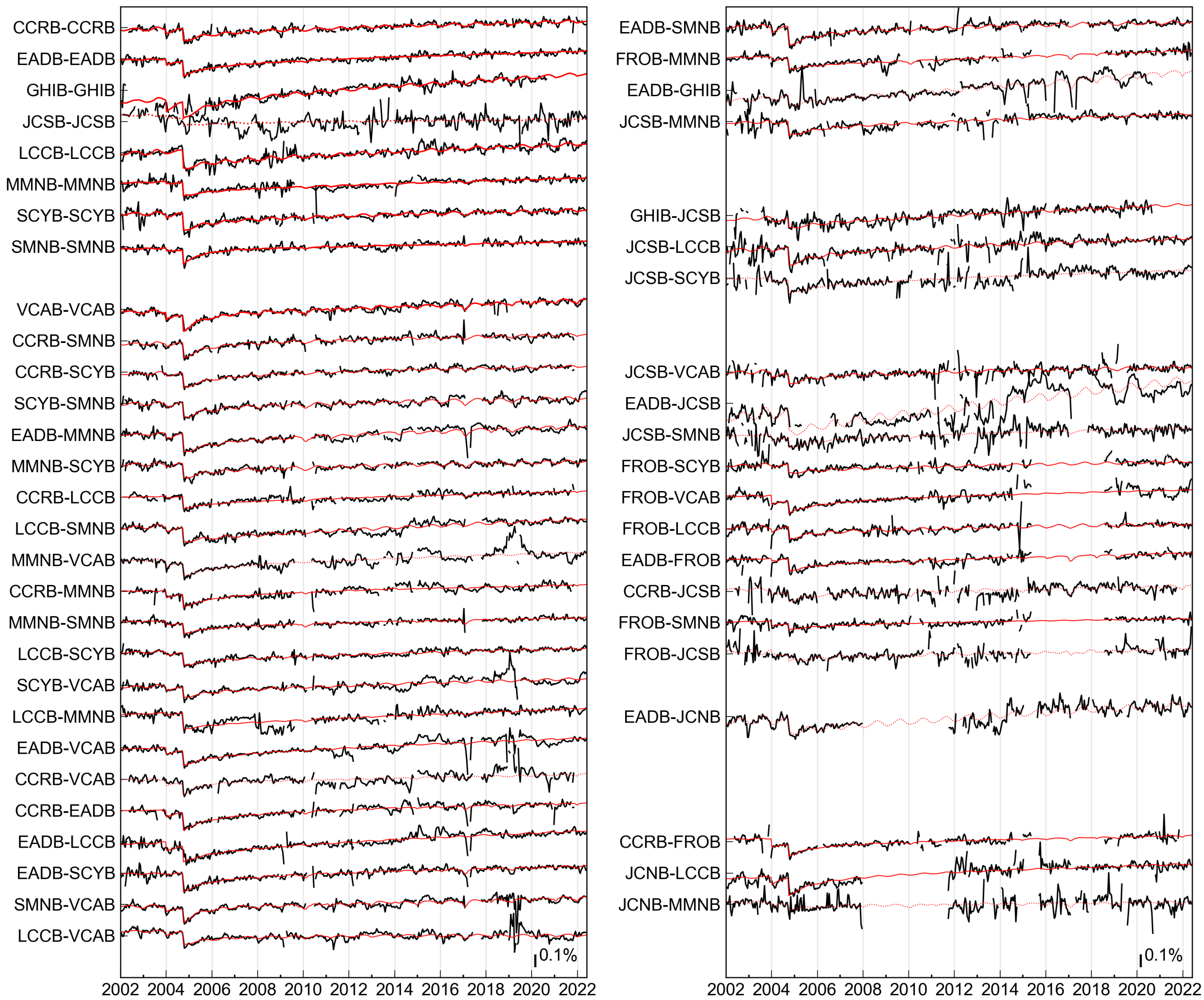

Figure S7. The channel-weighted time history of the dv/v (black lines) and the best-likelihood model fitting using the model with the linear trend term (red lines) for the case with 0.9-1.2Hz using the stretching method. The scale of dv/v is annotated in the bottom right corner. The solid lines in red indicate the pair meets the threshold of the fitting quality so that they are used for the statistical analysis of the model parameters, while the dotted lines indicate the excluded pairs due to the large residuals between the data and model.

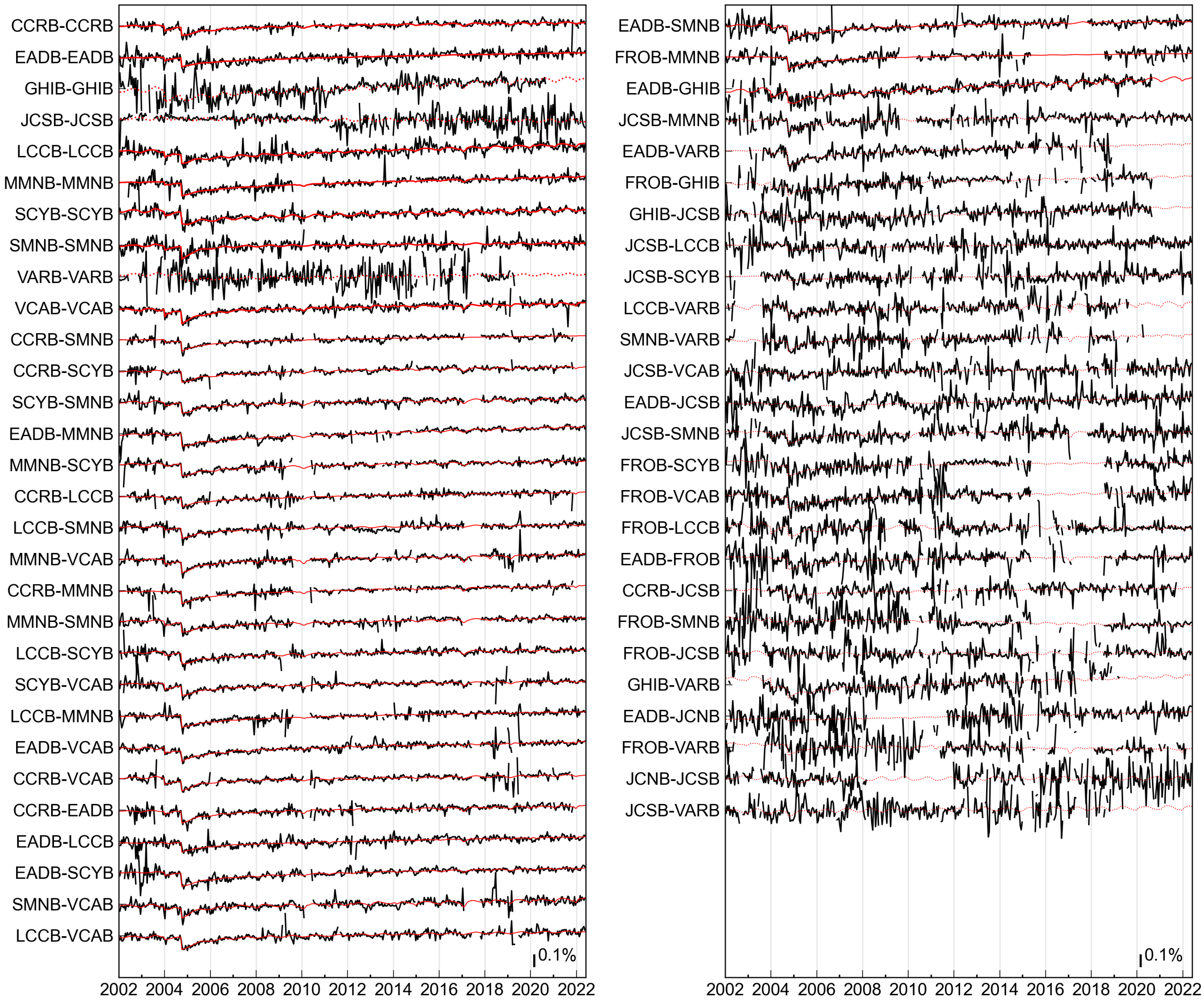

Figure S8. Same with Figure S7 for the case with the MWCS method. The order of station pairs is synchronized to Figure S7 with some additional station pairs, which are included as the MWCS allows for recovering the estimation of dv/v better with the weighted regression in the time delay.

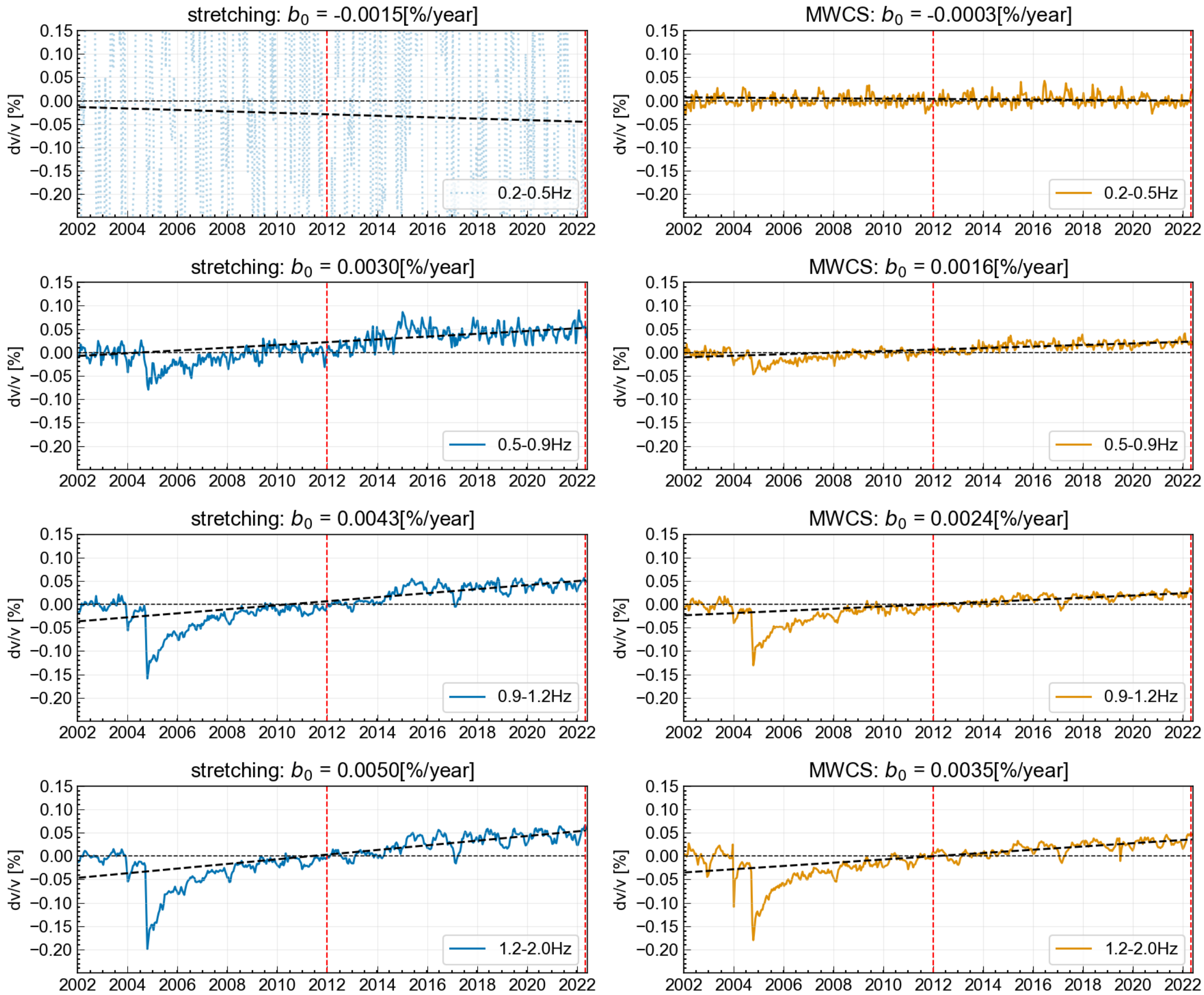

Figure S9. The long-term increase in dv/v associated with different frequency bands. The left and right columns are for the cases with the stretching and MWCS methods, respectively. We conducted the linear regression of the median of dv/v time history in the period from 2012/01/01 to 2022/06/01 as indicated by the vertical red lines to evaluate the slopes. The lowest frequency band of the stretching method shown in the top left panel is too unstable to estimate the dv/v. It should be noted that the estimated slopes are not equivalent to the case study of model fitting as the logarithmic healing model and the seasonal variation are included in the present dv/v time history.

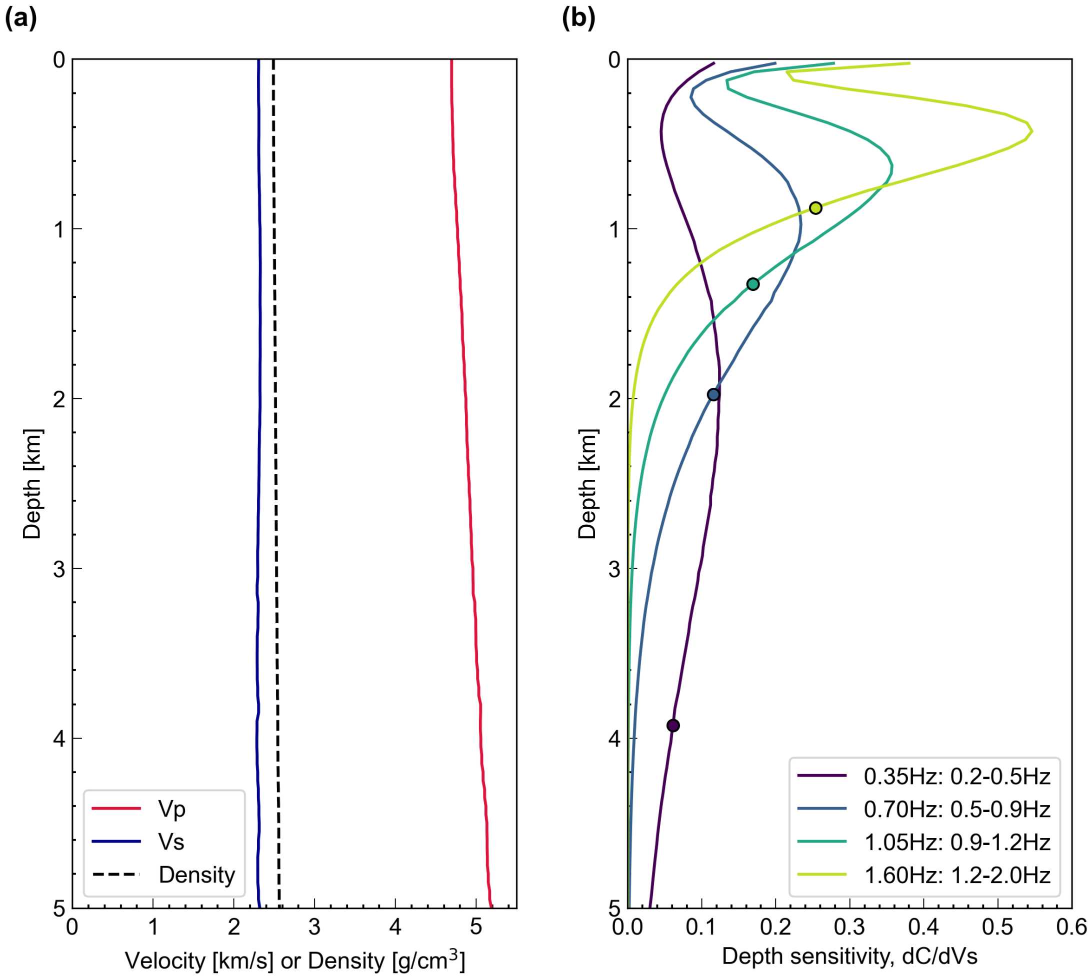

Figure S10. The depth sensitivity kernel for Rayleigh wave. (a) 1D velocity and density profiles with depth. The data is obtained from CVM-H v15.1.1 (Shaw et al., 2015) provided by the SCEC Community Velocity Model (CVM) with a 50m step down to 10km. (b) The depth sensitivity kernel for the Rayleigh wave associated with the central frequencies of the frequency band used in this study. We used the software tool, Computer Programs in Seismology (Herrmann, 2013), to compute the sensitivities. The circles indicate the depth concerning half of the maximum amplitude. Note that the subtle difference in the sensitivity from Wu et al. (2016) is caused due to the variation in the velocity and depth profiles.

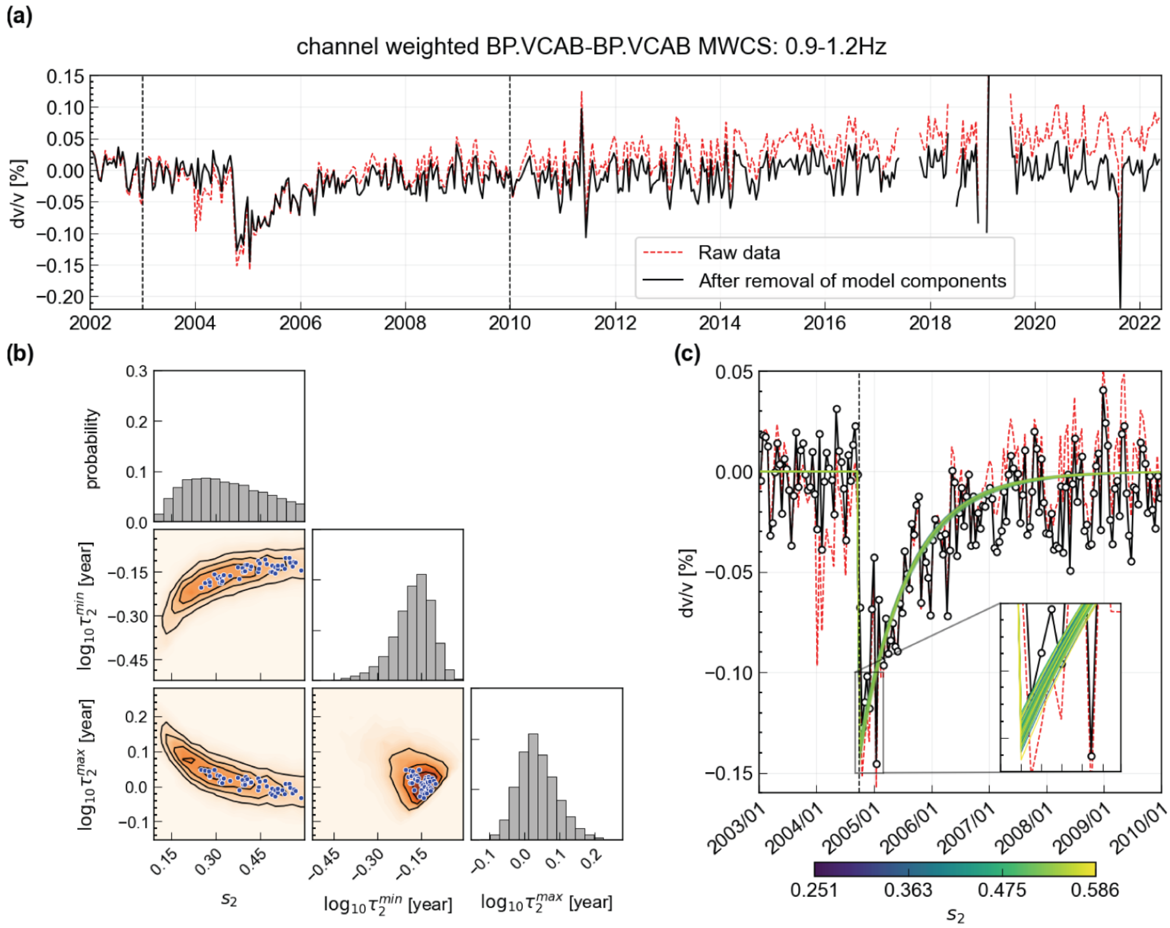

Figure S11. The trade-off between s_2, τ_2^{min} and τ_2^{max}. (a) The time history of dv/v associated with the auto-correlation of VCAB. The red dashed line and black solid line indicate the raw data and the data after the removal of model components associated with the ΔGWL, temperature, coseismic decrease, the logarithmic healing of the SS earthquake, and the linear trend term using the maximum likelihood parameters, respectively. The vertical lines show the bounds within which we perform the model fitting. (b) The scatter matrix of the MCMC parameter sampling. The blue circles indicate the model parameters used to show the synthesized dv/v as shown in (c). (c) The comparison of the synthesized model to the data. The line color indicates the variation of s_2, which is accompanied by the corresponding parameters of τ_2^{min} and ${}_2^{\mathrm{max}}$.

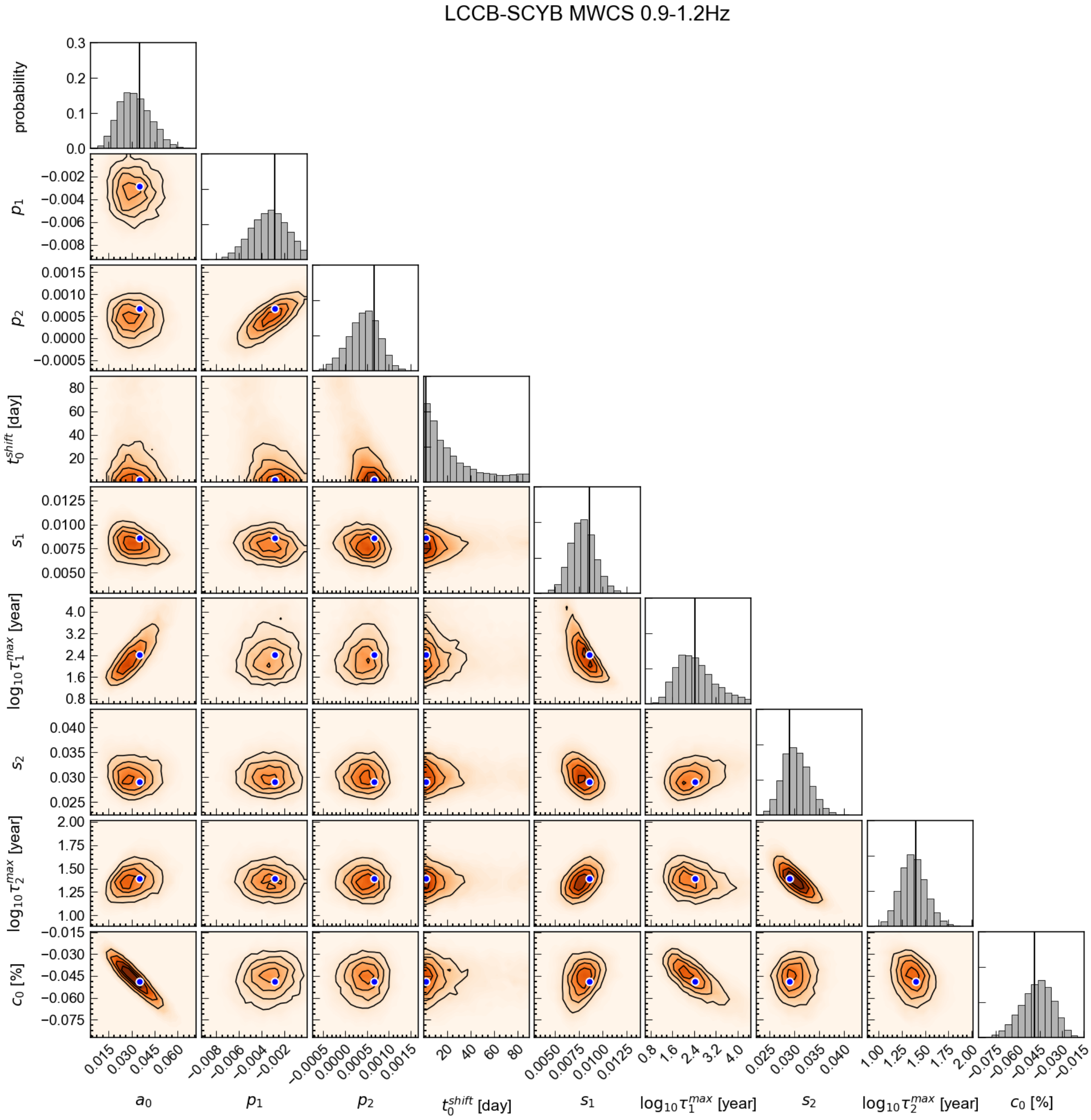

Figure S12. The 1D and 2D marginalized posterior probability distributions for the LCCB-SCYB, MWCS method, with the model of the residual healing term shown in the lowest row as c_0 [%].

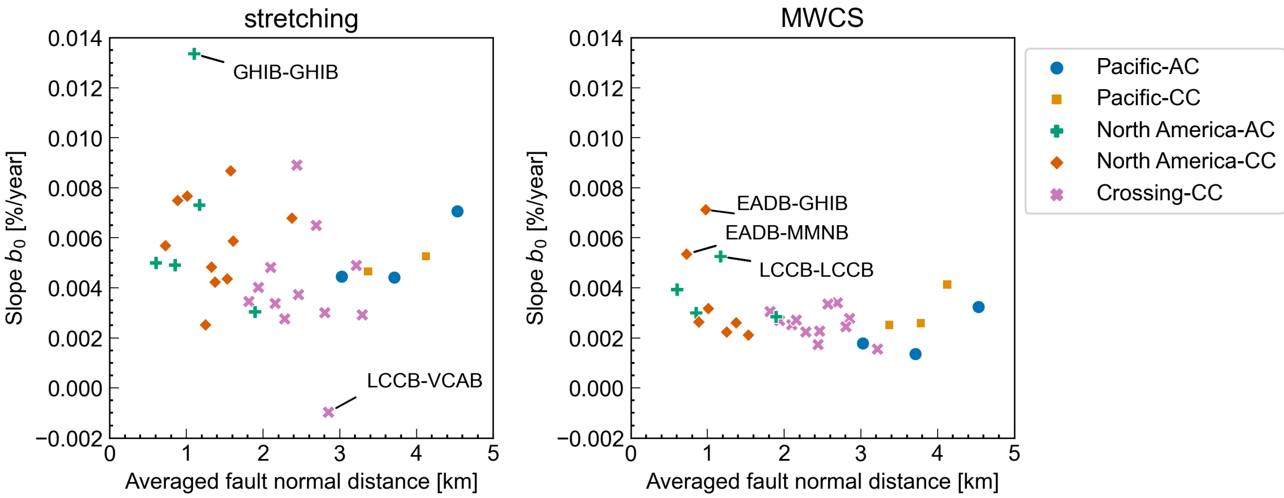

Figure S13. The slope of the linear trend term b_0 as a function of the fault normal distance. The distance is calculated with the approximated planar fault along the San Andreas fault as shown in Figure S1. We averaged the fault normal distances of two stations for the case with the CCFs. We categorized the CCFs and ACFs with the Pacific or North America side, and the CCFs crossing the fault as shown with different markers. We annotated some station pairs showing higher or lower values compared to the others.

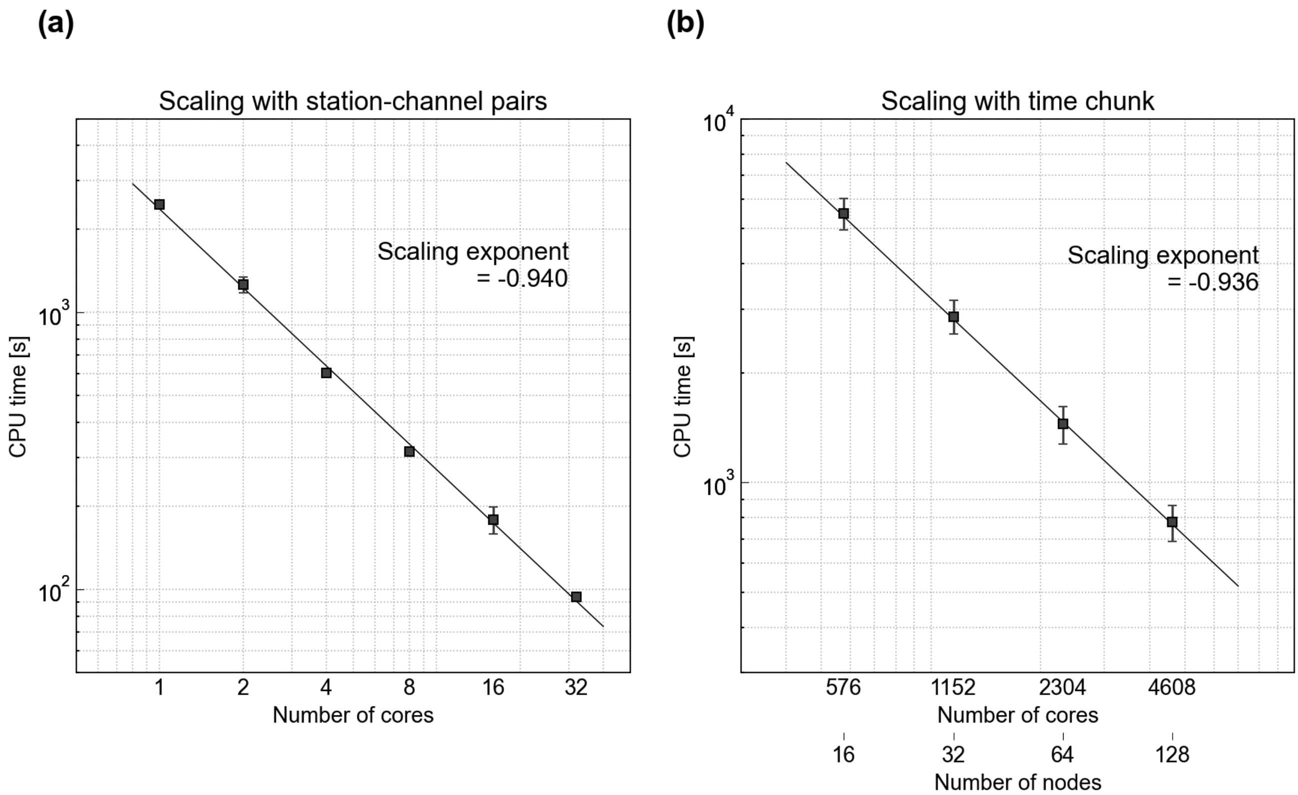

Figure S14. Scaling of parallelization to compute the FFT and the cross-correlation functions. (a) Single-node parallelization using the dataset of five days with three components for all station-channel pairs. The markers and error bars show the mean and standard deviation of the CPU time for the five iterations, respectively. (b) Multi-node parallelization using the dataset from January 2002 to September 2020. Note that the waiting time of the queues is not considered in this figure. The marker and error bar shows the mean and standard deviation of CPU time over the time chunks.