Seasonal variations of seismic velocities in the San Jacinto fault area observed with ambient seismic noise

G. Hillers,1 Y. Ben-Zion,2 M. Campillo1 and D. Zigone2

1 Institut des Sciences de la Terre, Université Joseph Fourier, F-38041 Grenoble, France. E-mail: gregor.hillers@ujf-grenoble.fr
2 Department of Earth Sciences, University of Southern California, Los Angeles, CA 90089, USA

SUMMARY
We observe seasonal seismic wave speed changes (dν/v) in the San Jacinto fault area and investigate several likely source mechanisms. Velocity variations are obtained from analysis of 6 yr data of vertical component seismic noise recorded by 10 surface and six borehole stations. We study the interrelation between dν/v records, frequency-dependent seismic noise properties, and nearby environmental data of wind speed, rain, ground water level, barometric pressure and atmospheric temperature. The results indicate peak-to-peak seasonal velocity variations of ∼0.2 per cent in the 0.5–2 Hz frequency range, likely associated with genuine changes of rock properties rather than changes in the noise field. Phase measurements between dν/v and the various environmental data imply that the dominant source mechanism in the arid study area is thermoelastic strain induced by atmospheric temperature variations. The other considered environmental effects produce secondary variations that are superimposed on the thermal-based changes. More detailed work with longer data on the response of rocks to various known external loadings can help tracking the evolving stress and effective rheology at depth.

Key words: Time-series analysis; Interferometry; Elasticity and anelasticity.

1 INTRODUCTION
Crustal rocks are subjected to a variety of loadings associated with atmospheric forcing terms, such as variations of pressure and temperature, precipitation and wind. These sources can induce various atmospheric forcing terms, such as variations of pressure and atmospheric temperature. The results indicate peak-to-peak seasonal velocity variations of ∼0.2 per cent in the 0.5–2 Hz frequency range, likely associated with genuine changes of rock properties rather than changes in the noise field. Phase measurements between dν/v and the various environmental data imply that the dominant source mechanism in the arid study area is thermoelastic strain induced by atmospheric temperature variations. The other considered environmental effects produce secondary variations that are superimposed on the thermal-based changes. More detailed work with longer data on the response of rocks to various known external loadings can help tracking the evolving stress and effective rheology at depth.

SUMMARY
We observe seasonal seismic wave speed changes (dν/v) in the San Jacinto fault area and investigate several likely source mechanisms. Velocity variations are obtained from analysis of 6 yr data of vertical component seismic noise recorded by 10 surface and six borehole stations. We study the interrelation between dν/v records, frequency-dependent seismic noise properties, and nearby environmental data of wind speed, rain, ground water level, barometric pressure and atmospheric temperature. The results indicate peak-to-peak seasonal velocity variations of ∼0.2 per cent in the 0.5–2 Hz frequency range, likely associated with genuine changes of rock properties rather than changes in the noise field. Phase measurements between dν/v and the various environmental data imply that the dominant source mechanism in the arid study area is thermoelastic strain induced by atmospheric temperature variations. The other considered environmental effects produce secondary variations that are superimposed on the thermal-based changes. More detailed work with longer data on the response of rocks to various known external loadings can help tracking the evolving stress and effective rheology at depth.

Key words: Time-series analysis; Interferometry; Elasticity and anelasticity.

1 INTRODUCTION
Crustal rocks are subjected to a variety of loadings associated with atmospheric forcing terms, such as variations of pressure and temperature, precipitation and wind. These sources can induce various atmospheric forcing terms, such as variations of pressure and atmospheric temperature. The results indicate peak-to-peak seasonal velocity variations of ∼0.2 per cent in the 0.5–2 Hz frequency range, likely associated with genuine changes of rock properties rather than changes in the noise field. Phase measurements between dν/v and the various environmental data imply that the dominant source mechanism in the arid study area is thermoelastic strain induced by atmospheric temperature variations. The other considered environmental effects produce secondary variations that are superimposed on the thermal-based changes. More detailed work with longer data on the response of rocks to various known external loadings can help tracking the evolving stress and effective rheology at depth.

Key words: Time-series analysis; Interferometry; Elasticity and anelasticity.

1 INTRODUCTION
Crustal rocks are subjected to a variety of loadings associated with atmospheric forcing terms, such as variations of pressure and temperature, precipitation and wind. These sources can induce various atmospheric forcing terms, such as variations of pressure and atmospheric temperature. The results indicate peak-to-peak seasonal velocity variations of ∼0.2 per cent in the 0.5–2 Hz frequency range, likely associated with genuine changes of rock properties rather than changes in the noise field. Phase measurements between dν/v and the various environmental data imply that the dominant source mechanism in the arid study area is thermoelastic strain induced by atmospheric temperature variations. The other considered environmental effects produce secondary variations that are superimposed on the thermal-based changes. More detailed work with longer data on the response of rocks to various known external loadings can help tracking the evolving stress and effective rheology at depth.

Key words: Time-series analysis; Interferometry; Elasticity and anelasticity.

1 INTRODUCTION
Crustal rocks are subjected to a variety of loadings associated with atmospheric forcing terms, such as variations of pressure and temperature, precipitation and wind. These sources can induce various atmospheric forcing terms, such as variations of pressure and atmospheric temperature. The results indicate peak-to-peak seasonal velocity variations of ∼0.2 per cent in the 0.5–2 Hz frequency range, likely associated with genuine changes of rock properties rather than changes in the noise field. Phase measurements between dν/v and the various environmental data imply that the dominant source mechanism in the arid study area is thermoelastic strain induced by atmospheric temperature variations. The other considered environmental effects produce secondary variations that are superimposed on the thermal-based changes. More detailed work with longer data on the response of rocks to various known external loadings can help tracking the evolving stress and effective rheology at depth.

Key words: Time-series analysis; Interferometry; Elasticity and anelasticity.
To discriminate between possible sources that may induce changes in the rock properties from potentially biasing effects associated with variable properties of the ambient seismic noise, we investigate the interrelationship between dυ/dv signals derived with two different methods (‘stretching’ and ‘doublet’), frequency dependent seismic noise properties, and the various mentioned environmental data. The results indicate that the observed seasonal velocity variations in the 0.5–2 Hz frequency range are associated with changes in the seismic velocity structure rather than changes in the noise field. A key observation is the temporal variability in the relation between the considered environmental driving mechanisms and the velocity changes. We find that the most likely source mechanism is thermoelastic strain induced by atmospheric temperature variations. The other considered environmental effects in the arid study area produce secondary variations that are superimposed on the thermal-based changes.

2 DATA AND METHODS

We use 6 yr (2009–2014) of vertical component data recorded by 16 stations located within a distance of ~20 km from the tricentric area of the San Jacinto fault near Anza (Fig. 1a). Ten surface stations are equipped with broad-band instruments, and additional six short-period seismometers are located in boreholes of the EarthScope/UNAVCO Plate Boundary Observatory between depths of 120 and 230 m.

We compute daily broad-band noise correlation or coherency functions, using a preprocessing procedure similar to Poli et al. (2012) and Zigone et al. (2015). The preprocessing consists of removal of large amplitude transients, whitening in a broad frequency band between 10 s and 10 Hz, and amplitude clipping. We experiment with different lengths of preprocessing windows (1–4 hr), which influence the detection capability of transients based on amplitude statistics (Poli et al. 2012). We also test the difference between 1-bit clipping and truncating the amplitude at three times the standard deviation in each window. We find that 1 hr 1-bit yields the best results considering the reduction in the number of outliers and the optimization of average correlation coda waveform similarity. The strongest effect on data quality is the seismicity associated with the 2010 April 4 Mw 7.2 El Mayor-Cucapah earthquake that occurred some 130 km southeast of the study area (Hauksson et al. 2011; Wei et al. 2011). The decaying aftershock activity controls the recovery of the fraction of used correlations. Data quality is best in 2011 and 2012 (Fig. 1b).

Low data quality time segments that passed the transient detection test are removed considering the maximum absolute amplitude distribution of all (up to 2191) daily correlation functions per station pair. Outliers are identified using the median absolute deviation from the median (MAD, Shelly et al. 2007). Correlations associated with maximum amplitudes 3.5 times larger than MAD or three times smaller than MAD are discarded. For each station pair a reference signal is constructed that consists of a stack of all remaining noise correlations. At each date, correlations from a time interval of ±w days are stacked to increase the signal-to-noise ratio (SNR) of the daily correlation, which leads to improved coda waveform similarity and thus improved estimates of velocity changes. Throughout this study we discuss results obtained using w = 10 d, but we verified that smaller choices (w = 2 d) do not influence our conclusions.

The precision of traveltime changes of primary arrivals obtained from correlations is limited in practice because the required stationarity and isotropy of the ambient wavefield are only approximately met (Weaver et al. 2009). For this reason, dυ/dv estimates are extracted from the coda of the correlation functions, an approach first applied to earthquake coda (Poupinet et al. 1984). This increases the sensitivity of noise-based tools because coda waves are less sensitive to fluctuations in the excitation mechanism (Colombi et al. 2014), and the phase delay accumulates linearly with increasing travel path in the case of homogeneous medium changes (Poupinet et al. 1984; Snieder et al. 2002). We estimate relative velocity changes from analysis of systematic arrival time changes between consecutive daily correlations and the reference stack in five different frequency bands, 0.1–0.4, 0.25–1, 0.5–2, 1–4 and 2–8 Hz. To increase the reliability of results, this is done using both the ‘stretching’ time domain technique (Lobkis & Weaver 2003; Wegler & Sens-Schönfelder 2007) and the ‘doublet’ frequency domain method (Poupinet et al. 1984; Ratdomopurbo & Poupinet 1995; Brenguier et al. 2008). We also apply curvelet denoising filters (Shelly et al. 2011) to the

Figure 1. (a) Shaded relief map of the study area. Black (red) triangles indicate locations of surface (borehole) stations. Environmental records are collected at the sites indicated by the blue circles. Black lines are mapped fault traces. PFO, Píon flats observatory; ANZ, meteorological station at Anza. (b) Daily evolution of the fraction of used correlations (0.1–2 Hz) obtained with the best preprocessing scheme (grey data). Black data show a 11-d running average. The arrow indicates the occurrence of the 2010 April 4 Mw 7.2 El Mayor-Cucapah earthquake.
correlation matrices before performing the doublet measurements. This filtering can further improve $dv/v$ estimates above the noise level by suppressing transient wavefield fluctuations.

For the stretching technique, the waveform under investigation is positively or negatively dilated to find the best similarity (highest correlation coefficient, $cc$) with the reference waveform. The optimal coefficient of dilation is an estimate of the velocity change. Estimating $dv/v$ with the doublet or moving window cross-spectral technique (e.g. Clarke et al. 2011) consists of two regressions: First, the delay time $dt$ is estimated from the cross-spectral phase shift along the frequencies of interest for multiple time windows centred at consecutive lapse times $\tau$. Secondly, the trend of the resulting $dt$ estimates as a function of $\tau$ yields the relative change $dt/\tau = -dv/v$. Errors are computed using the estimates of Weaver et al. (2009) and Clarke et al. (2011). A consistency of results obtained with the two techniques is indicative of robust estimates, because they perform differently in the presence of fluctuations (Hadziioannou et al. 2009; Zhan et al. 2013; Hillers et al. 2015a,b). Estimating the velocity change in later coda lapse time windows is another possibility to assess the reliability of early-coda high-SNR measurements. This is because the error introduced by potentially changing excitation patterns is sufficiently small after about two times the scattering mean free time (Colombi et al. 2014). We measure velocity variations in multiple 12 s long lapse time windows, 18–30, 24–36, 30–42 and 36–48 s. Note that the longest period in each frequency band controls the size of overlapping moving windows in each lapse time analysis window (Clarke et al. 2011); the 0.1–0.4 Hz analysis is therefore on the edge of feasibility.

To estimate the temporal relation and to assess the potential causality between any two records, we measure the delay between $L$ days long time windows from the various records. We refrain from using delays obtained with cross-correlation; this method works well for clean periodic signals, but is less robust when measuring the delay between some of the meteorological data (e.g. pressure) and our $dv/v$ time-series. Instead, we compute the Fourier transform FT of two tapered signals $s_1$ and $s_2$, estimate the complex cross-spectrum $C_{1,2} = FT(s_1) \cdot FT(s_2^*)$, where the asterisk denotes complex conjugation, and use the phase at the frequency at which $|C_{1,2}|$ is maximum. The phase in the range $[-\pi, \pi]$ is then converted to days. We confirmed this methodology comparing the results to alternative delay estimators, including the cross-spectral technique. This is repeated for 2191 – $L$ consecutive, $L – 1$ overlapping windows, leading to delay estimates as a function of time associated with the centre of the $L$ days long analysis windows. We use $L = 730$ d and $L = 1095$ d, that is 2- and 3-yr long windows. In all cases considered, the period at which $|C_{1,2}|$ is maximum is 1 yr, which highlights the focus on seasonal effects.

### 3 OBSERVED VELOCITY CHANGES

Fig. 2 presents $dv/v$ time-series as a function of frequency obtained with both analysis techniques. For the stretching technique, we measure the correlation coefficient $cc$ between the reference waveform and dilated test waveforms. For the doublet technique, the similarity (coherency) estimate (Clarke et al. 2011) is associated with the two original waveforms that have not been corrected for the potential traveltime change, and is therefore lower. For each station pair we only consider $dv/v$ measurements for which the similarity estimate is larger than the 6 yr mean minus half a standard deviation. This is an empirical threshold found to improve the SNR of the targeted seasonal $dv/v$ fluctuations. It mitigates the trade-off between neglecting low-quality data and maintaining a large number of measurements for averaging. Compared to a fixed threshold, it further allows an assessment of the frequency and lapse time dependent coda waveform similarity. Considering the average-$cc$ for each station, we find no consistently problematic sites or sensors or otherwise distinct spatial patterns. There exist potentially systematic relations between low/high $cc$ measurements and other properties of the station configuration, but a detailed study of such interrelations is not essential for the main purpose of this paper.

A significant $dv/v$ seasonal signal is found for all frequencies below 2 Hz (Fig. 2). Significance is implied by the high consistency of results obtained with the time- and frequency-domain analyses. Seasonal peak-to-peak amplitudes tend to decrease with frequency, ranging from 0.4–0.8 per cent at 0.1–0.4 Hz (Fig. 2a) over 0.2 per cent around 1 Hz (Figs 2b and c) to 0.1 per cent at 1–4 Hz (Fig. 2d). In addition, long-term trends are superimposed on the seasonal variations. For the results obtained between 0.25 and 2 Hz a negative and positive long-term trend is observed for 2009 and 2010, and between 2011 and 2014, respectively, while the 1–4 Hz results show an overall negative trend. The cause for these linear trends is unclear, but the increasing velocities may be associated with an overall decline of the water table due to anthropogenic ground water depletion in the study area (D. Agnew; pers. commun.). High frequency 2–8 Hz results do not show a seasonal trend, but an overall decaying trend that is interrupted by a positive step in 2011. The onset of this step coincides with the notch in the $dv/v$ peak (August 2011) found in all three frequency bands between 0.25 and 4 Hz. We do not investigate the nature of this feature and focus on longer timescales associated with seasonal responses.

Zhan et al. (2013) noticed the possible sensitivity of the stretching method to seasonal variations in the noise amplitude spectrum. This casts doubt on the associated low-frequency signal in Fig. 2(a), since the largest amplitude variability is found in the microseisms band 0.1–0.4 Hz as shown below. The propagation pattern is anisotropic but stable (Hillers et al. 2013) and should therefore not bias the doublet estimates (Hadziioannou et al. 2009; Colombi et al. 2014). To assess the significance of the low-frequency signal in Fig. 2(a), we consider the lapse time dependence of the seasonal variations (Fig. 3). A signal reflects medium changes rather than changes in the excitation pattern or noise properties if it remains stable after a few scattering events. Therefore, the disappearance of the low-frequency doublet signal (blue data in left column in Fig. 3) with increasing lapse time suggests a fading excitation effect. On the other hand, curvelet-filter enhanced doublet estimates (Ibid., red data) still indicate $dv/v$ seasonality at late lapse times (36–48 s). Significantly, the 0.5–2 Hz signals show a very high consistency for all lapse times, and for results based on raw and curvelet-filtered correlation functions. This consistency indicates a high improbability that the estimated velocity changes in that frequency range are caused by variations of ambient wavefield properties. We support this inference with more data in Section 4. Average similarity ($cc$) estimates of the high-frequency 1–4 Hz results allow the resolution of a seasonal $dv/v$ signal superimposed on the negative trend up to the 30–42 s lapse time window. These results likely reflect the medium response that is relatively unbiased by wavefield changes considering reduced scattering length and time scales at higher frequencies (Hillers et al. 2013). The seasonality in the 0.5–2 Hz $cc$ estimates, which are obviously higher for the curvelet-filtered data, is not in phase with the corresponding $dv/v$ estimates (Fig. 4a). Systematic $cc$ variations also
do not correlate with local wind speed and noise amplitude estimates. This suggests a more complex dependence of the governing wavefield properties and shows that the amplitude of the original wavefield is not a one-to-one proxy for coherency obtained with cross-correlation. The seasonal $dv/v$ pattern obtained with borehole data in the 0.5–2 Hz range (Fig. 2f) confirms the notion that the signal in this frequency range is governed by medium changes. Significant, sometimes higher, $dv/v$ amplitudes at depth indicate that the observed velocity changes are not associated with a thin ($<10$ m) superficial layer or induced by a near-surface effect, but reflect physical changes in at least the top few 100 m. Delay estimates between 0.5 and 2 Hz $dv/v$ time-series at 18–30 s and later windows (Fig. 4b) do not indicate a lapse time dependent shift in the seasonal signal. Excitation effects do therefore not systematically bias the early window results. However, a 2 yr window centred at the time of the August 2011-notch reduces the amplitude of the seasonal effect at later lapse times (blue, red data in Fig. 4b) leading to more unstable delay estimates.

Early window low-frequency results obtained from curvelet-filtered correlations functions are not in phase with the corresponding 0.5–2 Hz $dv/v$ signal (green data in Fig. 4c). In contrast, late lapse time results show a higher consistency with this $dv/v$ signal at least when the analysis window is centred on 2012–2014 (red data in Fig. 4c). This supports the hypothesis that early window 0.1–0.4 Hz results are governed by variations in wavefield properties, and suggests that this bias decays with increased wavefield randomization leading to improved resolution of the medium response. In the following analyses we focus on the 0.5–2 Hz results that exhibit the cleanest $dv/v$ changes and show a good consistency between all used analysis methods.

4 ANALYSIS OF POSSIBLE SOURCE MECHANISMS

Two broad categories can be responsible for the derived seasonal seismic velocity variations. The results can reflect changes in material properties, or may be induced by changes in wavefield properties that bias the measurements. To further exclude the latter possibility we extend the lapse-time analysis and proceed to investigate the
Figure 3. Frequency (left to right) and lapse time (top to bottom) velocity change estimates obtained with the doublet method. Red data correspond to results based on curvelet-filtered correlations. The horizontal lines indicate associated error estimates. Data on top of each panel show the corresponding estimates of waveform similarity. The y-axis tick-interval is identical in all panels to highlight the different scales.
delay between the $d/v$ signal (Fig. 2c) and records of seismic noise properties (Fig. 5), and assess the variability of wave propagation directions (Fig. 6). To identify the dominant driving mechanism of the seasonal seismic wave speed changes, we measure the delay between the $d/v$ signal and various environmental observables that exhibit annual or sub-annual periodicity. These include atmospheric temperature, barometric pressure, wind speed, precipitation and records of the ground water level (Figs 5 and 7) measured in the area of the used seismic stations. We begin by demonstrating that systematic changes in wavefield properties are not dominating the $d/v$ signal. We study frequency dependent fluctuations of low-noise amplitudes (Fig. 5) as a proxy for potentially biasing variations (Hadzioannou et al. 2011; Zhan et al. 2013), and assess systematic changes in wave propagation directions from seasonal patterns of the noise intensity $B$ (Larose et al. 2007; Froment et al. 2010; Colombi et al. 2014).

4.1 Noise amplitudes and wind speed

Following the low-noise level estimate by Hillers et al. (2012a), we subdivide a daily vertical component 20 Hz sampled record of each station into 48 segments that are 30 min long and estimate the power spectral density in all segments. This yields daily distributions of noise amplitude at each frequency obtained by the Fourier transform. We chose either some low-quantile (0.01–0.05), median or mean value to represent the frequency-dependent daily spectral noise power at each station. We then take an average value for the five frequency bands for which we compute the $d/v$ response, and either continue with these values or convert them to decibel (dB). The resulting 2191 daily values per station are scaled by the temporal average, and we finally take the mean at each day from all stations (Figs 5a–c).

In general, the noise amplitude on all instruments is higher between November to May consistent with the generally observed increase in noise energy during northern hemispheric winter. The timing of the oscillations (Fig. 5b) shows that the high-frequency rhythm lags systematically behind the low-frequency variations. The averaged dB amplitude scale (Fig. 5c) emphasizes the effect of the seismicity triggered by the El Mayor-Cucapah earthquake on high-frequency amplitudes. Aftershock wavefields cause background amplitudes to rise significantly above the annual variability.

Wind speed variations recorded near Anza (Fig. 5d) are a potential local source of noise amplitude changes (Young et al. 1994; Withers et al. 1996; Young et al. 1996; Hillers & Ben-Zion 2011). Wind speed changes are not considered a driving mechanism for physical $d/v$ changes, but rather a cause for systematic noise excitation that may introduce a bias into estimates of $d/v$ changes. Approximately neutral delay estimates suggest that high-frequency amplitudes are likely influenced by seasonal wind speed changes (Figs 5e and f) through interactions with topography, vegetation or infrastructure. Negative delays indicate that low-frequency amplitudes increase before the local wind speeds rise. The delay is larger for lower frequencies, and shows that variations in low-frequency noise excited by atmosphere-ocean-lithosphere interactions are driven by non-local dynamics compared to the high-frequency wavefield. Atmospheric patterns affect noise energy in a broad frequency band, through different on- and off-shore excitation mechanisms. The non-linear microseisms excitation mechanism also leads to strongest seasonal variability in the 0.1–0.4 Hz range (Figs 5a and b). This inverse frequency dependence of the variations is consistent between borehole and surface station data (not shown). However, seasonal variations exhibit smaller peak amplitudes at depth, which can partially be explained by the depth dependence of noise amplitude changes on wind speed (Young et al. 1994). This is not compatible with the significant $d/v$ changes at depth (Fig. 2f) under the assumption that velocity changes are spurious and driven by wind-induced wavefield changes.

Delay estimates between the doublet 0.5–2 Hz $d/v$ results, obtained with and without application of curvelet filters and the corresponding noise amplitudes show that seismic velocities change some 20 d earlier (Fig. 8a). The same holds for the $d/v$ and wind speed delay (Fig. 8b), where the difference is about 30–40 d. The differential delay of 10–20 d is consistent with the 10–20 d delay measured between amplitude and wind speed (Figs 5e and f). Hence, $d/v$ and noise amplitudes or wind speed changes do not occur simultaneously. We conclude that the observed seasonal changes of seismic velocities are not associated with spurious measurements induced by systematic changes in amplitude-related wavefield properties that are driven by the wind pattern.
Figure 5. (a) Daily average noise level estimates for five colour coded frequency bands. (b) Smoothed data from (a). (c) Same as (b) from daily amplitude estimates in dB. The arrow indicates the time of the El Mayor-Cucapah earthquake. Smoothing smears the effect of the aftershock seismicity on high-frequency noise amplitudes. (d) Daily wind speed estimates recorded near Anza. (e) Delay estimates between the data in (a) and (d). Negative values mean that signal $s_A$ (wind speed) lags behind signal $s_B$ (amplitude). (f) Delay estimates between dB-scale amplitude data and wind speed data in (d). We used an analysis window of $L = 730$ d.

Figure 6. Azimuthal distributions of the noise intensity $B$ (thick lines). Colour indicates frequency as in Fig. 5. (a) Frequency dependent summer $B$ patterns. The circles indicate the average $B$ amplitude. The patterns show an overall decrease of $B$ with frequency. (b) Intensity patterns from (a) scaled by the frequency dependent averages. This representation highlights the increasing randomization of propagation directions with frequency. (c) Comparison of summer (solid lines) and winter (dashed lines) $B$ patterns for the 0.5–2 Hz range.

4.2 Directional noise intensity

Anisotropic distributions of the noise intensity $B$ can bias arrival time estimates (Weaver et al. 2009; Froment et al. 2010). Systematic changes in the dominant propagation direction of the wavefield can similarly induce spurious $d\nu/d\nu$ measurements. This caveat is limited to the wavefield at early lapse times that is governed by a significant ballistic, that is directional component (Colombi et al. 2014). Scattering randomizes propagation directions and
Seasonal velocity variations in the SJF area

Figure 7. Environmental records obtained at Piñon Flats. (a) Air temperature, (b) atmospheric pressure, (c) precipitation rates, (d) ground water level measured in three boreholes within a distance of ∼300 m.

increases isotropy towards later coda arrival times. We follow the approach of Hillers et al. (2013) and estimate $B(\theta)$, where $\theta$ denotes azimuth, from correlation SNR as parametrized by Larose et al. (2007). We compute intensity distributions obtained from correlation stacks associated with summer (June–August) and winter (December–February) months.

The patterns in Fig. 6(a) illustrate the overall decay in $B$ with frequency, which can be attributed to increased attenuation and hence a reduction of coherency (the ‘signal’ in the SNR estimates). For the lower three frequency bands, the $B(\theta)$ distributions are peaked in southwest directions, which results from excitation along the coastline (Schulte-Pelkum et al. 2004; Tanimoto et al. 2006; Gerstoft & Tanimoto 2007; Zhang et al. 2009; Hillers et al. 2013). In contrast, the wavefield at higher frequencies is characterized by increased isotropy (Fig. 6b), which is a consequence of smaller scattering length scales associated with shorter wavelengths, and a likely different source distribution. The high similarity of summer and winter $B(\theta)$ distributions (Fig. 6c) shows that the noise excitation pattern is stable for frequencies at which the $\nu/v$ signal is most significant (0.5–2 Hz). It supports the conclusion drawn from the lapse time analysis that velocity changes are not biased by systematic changes in directional wavefield properties.

We note that the 0.1–0.4 Hz $B$ pattern shows the same behaviour, that is the SNR changes but the azimuthal distribution is constant (Hillers et al. 2013). Systematic changes in propagation directions are therefore unlikely to govern the spurious low-frequency $\nu/v$ pattern we deduced from the lapse time-dependent delay distributions (Fig. 4). However, seasonal changes in the wavefield composition, for example, in the relative presence of surface and body waves (Tanimoto et al. 2006; Landèes et al. 2010), potentially also influence or bias early lapse time estimates. Effects of these wavefield changes go beyond the apparent traveltime variations induced by changes in the incidence angle of the dominant energy flux (Weaver et al. 2009; Froment et al. 2010; Colombi et al. 2014).

4.3 Rainfall

Precipitation can induce seismic velocity changes by two mechanisms. First, wave speeds are reduced if water fills open pore space. This affects propagation within shallow layers in which the ground water level (GWL) fluctuates (Sens-Schönfelder & Wegler 2006; Meier et al. 2010; Hillers et al. 2014). Secondly, the additional load of the water mass can induce pore pressure changes across hydraulically connected regions (Husen et al. 2007; Hainzl et al. 2013), and change stress and seismic velocities in hydraulically disconnected deeper layers (Bettinelli et al. 2008; Obermann et al. 2014). Systematic wavefield changes induced by noise excitation associated with run-off patterns (Burtin et al. 2008) can be a source of error or bias. In our case, the rainfall recorded in the arid study area is notably sparse, with precipitation exceeding 2 cm at only few days within the 6 yr observation period (Fig. 7c). This is in marked contrast to Sens-Schönfelder & Wegler (2006) and Hillers et al. (2014) who observed ∼2 and ∼0.2 per cent $\nu/v$ changes due to precipitation characterized by longer, more steady duration and significantly higher rates. Moreover, collocated GWL observations (Fig. 7d) do not indicate a consistent response of the subsurface...
hydraulic situation to large rainfall events that visually exceeds the amplitude of the long-term trends. Delay distributions between the rain and \( \frac{dv}{v} \) records (Fig. 8c) imply a coincidence (no delay) only towards the end of the observation period. However, the amount of rainfall in 2013 and 2014 is lower compared to the years 2009–2012 (Fig. 7c), when large delay times suggest no causal relation. We conclude that the seasonal \( \frac{dv}{v} \) variations are not governed by the rainfall pattern.

4.4 Ground water level

Observations of the GWL allow a more direct assessment of the effect of subsurface fluid distributions on seismic wave speeds (Grêt et al. 2006) compared to diffusion models resting on precipitation patterns (Sens-Schönfelder & Wegler 2006). The three GWL records observed at Piñon Flats within 300 m (Fig. 7d) have persistent trends related to ground water depletion, superimposed with seasonal variations that are largest at well CIC2 and peak around May/June in 2010 and 2011. This seasonal pattern is unlikely to cause the observed \( \frac{dv}{v} \) signal for two reasons. First, an increase in the GWL is expected to reduce wave speeds in the considered depth range (Sens-Schönfelder & Wegler 2006; Hillers et al. 2014), which is opposite to what is observed during summer months (Fig. 2). Secondly, the \( \frac{dv}{v} \) changes peak between July and October and are not in phase with the GWL variations. Stable delay estimates are found only for the years 2010 and 2011 (Fig. 8d), and the \( \sim 70 \) d estimates confirm this non-simultaneity. Finally, the lack of seasonal GWL changes in the other two colocated wells, and for later years in CIC2, indicates that the seasonal velocity changes are not dominantly influenced by the GWL.

4.5 Atmospheric pressure

Changes in the atmospheric pressure field (Fig. 7b) can cause changes in seismic wave speeds due to variable loads exerted on the surface (Gao et al. 2000; Silver et al. 2007; Niu et al. 2008). The pressure-induced stress variations propagate with the shear wave speed, leading to near instantaneous interactions. Hence, the non-zero delay distributions (Fig. 8e) indicate that the variability of the barometric pressure field is not driving the observed seasonal velocity changes.

4.6 Atmospheric temperature

Temporal variations in the atmospheric temperature are dominated by an annual period (Fig. 7a). Estimating the delay between the
0.5 and 2 Hz $dv/v$ variations and air temperature records (Fig. 8f) indicates that changes in seismic velocities lag consistently behind annual variations of the temperature field. The estimated delays follow a distribution that is centred at 20 d. A mechanism that produces strains and hence velocity variations that lag behind the temperature variations is thermoelastic strain.

Spatial variations of the surface temperature field and/or thermal properties of the heated material produce thermoelastic strain that can propagate to depth with appreciable amplitude. Simple calculations using observed temperature changes and realistic parameters for the wavelength of the temperature field and material parameters indicate stress levels over 10 kPa in the top kilometer of the crust (Ben-Zion & Allam 2013). Temperature variations can also affect local wave velocities by a temperature dependent change of the elastic moduli in an otherwise unstressed body. Berger (1975) provided a solution for thermoelastic strain in a homogeneous elastic halfspace due to a surface temperature field characterized by a simple wave function. An important aspect of the solution is that the thermoelastic strain is delayed by $T/8$ from the source temperature field with $T$ being the wave period (Tsai 2011). For the annual changes that are the focus of this study, $T/8 \approx 40$–45 d. Accounting for the very shallow material, and using sensitivity kernels to estimate the delay between $dv/v$ and seasonal temperature variations, lead to $\sim 40$ d (Richter et al. 2014). Ben-Zion & Leary (1986) modified the solution of Berger (1975) by adding a surface layer that is elastically decoupled from the halfspace, to account for shape differences and additional delays between observed strain records and halfspace model predictions. The solution was shown to predict correctly from atmospheric temperature data the observed frequencies, phases and amplitudes of strain data recorded at depths varying from 10 m to $>300$ m (Ben-Zion & Leary 1986; Ben-Zion & Allam 2013).

The velocity changes observed below 1 Hz (Fig. 2b) imply that the analysed surface waves are sensitive to a depth regime below the top unconsolidated (soil) layer, where the thermoelastic strain and stress may induce changes in the wave speed. This notion is supported by the significant $dv/v$ amplitudes resolved with borehole data. The average 20 d delay between $dv/v$ variations and temperature records is somewhat shorter than the expected $\sim 40$ d for a halfspace. Nevertheless, the only mechanism that has a plausible lag from the observed $dv/v$ signals is atmospheric-induced thermoelastic strain, so we infer this to be the dominant source of the observed velocity changes. This interpretation is supported by the coherency between the $dv/v$ and the temperature records, which is significantly and consistently higher compared to the coherency between the $dv/v$ and all other considered records. The temporal delay evolution obtained from $dv/v$ data at early lapse times shows a trend from smaller (absolute) to higher estimates. The associated slope is smaller compared to other delay trends, yet it is another manifestation of multiple, simultaneously acting mechanisms influencing our $dv/v$ measurements. Results associated with later lapse time windows (green data in Fig. 8) show again a different response for the first half of the analysis period, which are perhaps related to effects associated with the El Mayor-Cucapah earthquake. Other factors contributing to the observed phase lag that is at the low bound of what is expected from the thermoelastic model are the level of uncertainty (Fig. 2b), averaging over different positions in the local source field, and effects that operate on shorter time scales including potential wavefield changes.

5 DISCUSSION

Analysis of 6 yr ambient seismic noise recorded near the San Jacinto fault zone in southern California resolves seasonal variations of subsurface seismic velocities ($dv/v$). The strongest and most consistent signal is obtained using the correlation coda wavefield at early lapse times in the 0.5–2 Hz frequency range. The associated results are supported by data from curvelet-filtered correlations and data from later lapse times. The $dv/v$ records in this frequency range are compared with noise amplitude estimates, nearby environmental data of wind speed, rain, ground water level, barometric pressure and atmospheric temperature to clarify the cause(s) of the observed velocity changes.

The possibility that seasonal $dv/v$ variations are spurious or systematically biased by wavefield changes is reduced by the non-zero delay between records of $dv/v$ and noise amplitude, and the stability of the noise intensity pattern. However, seasonal wavefield changes associated with systematic variations in the excitation pattern (Tanimoto et al. 2006) and transient changes associated with local and regional seismic activity potentially bias the $dv/v$ estimate, or reduce the data quality. The sparsity of rainfall and the inconsistency between water level data and $dv/v$ measurements rule out governing hydrologic effects. Delay statistics further indicate that the barometric pressure variations are not the dominant source. The delay between the $dv/v$ variations and air temperature implies that thermoelastic strain is the most plausible primary mechanism for the observed seasonal variations of seismic velocities.

The delay distributions on which these conclusions are based (Fig. 8) were primarily obtained with an analysis window $L = 730$ d. These delay statistics are overall insensitive to different choices of the correlation-stack window length $w$, the coda lag time window, and the length of $L$ exceeding 1 yr. Cross-spectra $C_{i,j}$ obtained between the $dv/v$ signal and all examined amplitude and environmental time-series are characterized by a peak annual periodicity, that is, the obtained delay distributions reflect differences associated with the seasonal component. The temporal evolution of the delay estimates, for example, the (absolute) delay increase between the $dv/v$ and temperature signals with time, implies simultaneously acting processes that are either associated with internal mechanisms such as the local response to earthquakes (Agnew & Wyatt 2014) or with transient changes of wavefield properties. Reducing $L$ to sub-annual length (e.g. $L = 183$ d), and constraining delay measurements to time windows that are dominated by a semi-annual periodicity, modifies the resulting delay distributions (not shown). Most notably, the $dv/v$-temperature delay range becomes significantly larger, albeit at significantly reduced coherency estimates, also including zero-delay. This implies that mechanisms other than thermoelastic strain can influence seismic velocity changes, and that thermal processes can also operate on subannual timescales (Ben-Zion & Leary 1986; Richter et al. 2014).

The results indicate that the seasonal $dv/v$ variations at 0.5–2 Hz reflect changes of rock properties induced primarily by thermoelastic strain rather than being governed by seasonal changes in wavefield properties. To be sure, the other considered source mechanisms contribute to the resolved velocity changes, leading to estimated delay times between $dv/v$ and temperature that is smaller than what is expected from a pure thermoelastic strain effect (Berger 1975). The zero-delay obtained from $C_{i,j}$ at subannual periods using noise amplitude estimates (not shown) indicates that these mechanisms can be relevant on shorter timescales either through biasing the estimates, or by interfering with the thermal deformation effects.
The observed delay between $\text{d}v/v$ measurements and temperature data involve averaging over the study area and possible smearing effect caused by different positions of stations in the local source field. Previously reported observations of thermoelastic deformation compatible with a model of a halfspace covered by a thin unconsolidated 'equivalent upper layer' were obtained with point-like measurements from individual strainmeters (Ben-Zion & Leary 1986; Prawirodirdjo et al. 2006; Ben-Zion & Allam 2013). In these studies, the phases of the strain synthetics were not sensitive to the station position in the temperature field. In our case with stations separated by tens of kilometres in a region with variable topography and strong changes of elastic properties (Allam & Ben-Zion 2012; Zigone et al. 2015), the thermoelastic strain sampled by the analysed scattered correlation wavefield must have some spatiotemporal variations. Averaging measurements associated with heterogeneous deformation can weaken cleaner signals resolving temporal changes in smaller regions.

The seasonality and the long-term trends in the 0.25–1 Hz $\text{d}v/v$ records are compatible with the 0.5–2 Hz results. The large annual $\text{d}v/v$ signal observed at yet lower frequencies (0.1–0.4 Hz) contains biasing components that are possibly associated with seasonally variable constituents of the microseisms wavefield (Tanimoto et al. 2006; Gerstoft & Tanimoto 2007; Landés et al. 2010; Hillers et al. 2013). The associated effect on $\text{d}v/v$ estimates goes beyond apparent arrival time changes induced by variations in the dominant propagation directions (Froment et al. 2010; Colombi et al. 2014). These spurious components decay with increasing lapse time and allow improved resolution of seasonal medium changes, but only after filtering the correlation coda wavefields. At 1–4 Hz, the amplitude of the seasonal signal is reduced compared to the 0.5–2 Hz case, and the long-term trend is only negative for the observation period. Results at 2–8 Hz are also characterized by high consistency between different analysis techniques, show also an overall negative trend, but no annual periodicity in the $\text{d}v/v$ response. Data recorded at closely spaced stations (Lin et al. 2013; Hand 2014; Hillers et al. 2015a; Ben-Zion et al. 2015) may allow resolving seasonal and other variations at these and higher frequencies.

Neglecting source related biases, the peak-to-peak amplitude of the observed $\text{d}v/v$ variations is mainly controlled by two elements. The first is the stress or strain sensitivity $S$, which depends on the elastic medium properties that control the actual velocity change in response to deformation. The second is the ability to resolve these wave speed changes from arrival time changes in the coda wavefield; this wavefield sensitivity varies for surface and body waves and also with lapse time (Sens-Schönfelder & Wegler 2011; Froment et al. 2013; Obermann et al. 2013). Whereas $S$ can be thought to decrease with depth in response to the increase in confining stress, changes in the wavefield sensitivity depend on the wavefield constituents. That is, the sensitivity of surface and body waves is different and varies nonlinearly with depth. Assuming the relative variation of $S$ in the depth range sampled by the wavefield to be small, the frequency dependent $\text{d}v/v$ peak-to-peak amplitude decrease suggests a larger wavefield sensitivity at low frequencies compared to high frequencies. The noise wavefield up to 1 Hz is certainly dominated by surface waves considering the nearby excitation along the coastline. Results in the late lapse time window (36–48 s) are still dominated by surface wave sensitivity in the low-frequency range. This follows from scattering mean free times of $t^* = 20–40$ s obtained from a regional scattering mean free path of 50–100 km around wave periods of 5 s (Hillers et al. 2013) and Rayleigh wave speeds between 2.5–3 km s$^{-1}$ (Zigone et al. 2015). Hence, $\tau/t^* < 6$ at which decreasing surface and increasing body wave sensitivity are equally partitioned (Obermann et al. 2013). The $\text{d}v/v$ amplitude decrease with lapse time at 0.5–2 Hz is hence not necessarily governed by the corresponding decrease in waveform coherency. It can also be associated with the increasing time a scattered phase spent as body wave in a depth range where the stress sensitivity and/or the wavefield sensitivity is lower. Results in the 1–4 Hz band are also likely dominated by surface waves in the considered lapse time range, even if we allow for decreased scattering timescales. However, the ratio of $\text{d}v/v$ peak amplitude to the error estimates is generally frequency independent, which implies that the smaller seasonal velocity change at high frequencies is equally well resolved. In conclusion, our analysis shows that seasonal velocity changes are not limited to the top tens of metres, but are likely distributed in the first few hundreds of metres. A more accurate depth estimate of the surface wave peak sensitivity can in principle be derived from local velocity models (Hillers et al. 2013; Zigone et al. 2015), but these models usually lack accuracy in the shallow parts that govern the kernels at the high frequencies used here.

The observed velocity change estimates are within the range of previously reported variations in response to various external and internal driving mechanisms (Reasenberg & Aki 1974; Sens-Schönfelder & Wegler 2006; Niu et al. 2008; Meier et al. 2010; Hoberg et al. 2012; Hillers et al. 2014). Focusing on passive methods, this range to date is limited by the observed $\sim 0.03$ and $\sim 2$ per cent changes in response to tides and precipitation, respectively (Sens-Schönfelder & Wegler 2006; Hillers et al. 2015a), yet reaches up to 7 per cent prior to a landslide failure (Mainsant et al. 2012). Improving the sensitivity of $\text{d}v/v$ measurements and separating contributions to observed $\text{d}v/v$ changes from often simultaneously acting mechanisms has become a growing concern that is fueled by the development of passive monitoring techniques. Sensitivity can be improved using filter designs to suppress unwanted wavefield fluctuations (Baig et al. 2009; Hadziioannou et al. 2011; Gallot et al. 2011), as demonstrated here by the increase in SNR after application of curvelet-filtering (Stehly et al. 2011). Targeting the separation of different forcing effects, subdaily window stacking can be an advantage when focusing on daily components of the variations (Richter et al. 2014; Hillers et al. 2015a). Fits to observed $\text{d}v/v$ records can help isolating external from internal contributions (Hoberg et al. 2012; Richter et al. 2014), and continuous velocity change time-series can generally be improved by treating datum-wise $\text{d}v/v$ errors probabilistic (Brendquier et al. 2014). Responses to external loadings can be used for designing filters for removing these signals and improving records that reflect internal tectonic sources.

A further in-depth analysis of the material response to different source mechanisms using these methods is beyond the scope of this work, but we can estimate the first order velocity change sensitivity $S = \text{d}v/(v\varepsilon)$ assuming the $\text{d}v/v$ signal is governed by thermoelastic (volumetric) strain ($\varepsilon$). Using Berger’s solution, Ben-Zion & Allam (2013) estimated thermoelastic areal strain at 200 m depth in the range $1–10 \times 10^{-7}$, depending on used thermal and elastic model parameters. Hence, $S$ is approximately between $10^3$ and $10^4$, which is consistent with values inferred for responses to crustal earthquakes (Wegler et al. 2009), changes at mid-crustal depth due to slow inelastic deformation on a subduction interface (Rivet et al. 2011), magma intrusion into a volcanic edifice (Ueno et al. 2012), and the shallow response to tidal strain (Hillers et al. 2015a). Assuming a bulk modulus of $10^4$ MPa, $S = 5 \times 10^4$ translates to a velocity stress sensitivity of $\sim 0.5$ MPa$^{-1}$, which is again within the range of values observed in response to tidal loading (Yamamura et al. 2003).
Seasonal thermoelastic strain can change the seismic velocities of crustal materials similar to a variety of other external and internal forcing mechanisms. Separating these contributions, using longer data and more detailed analysis techniques in combination with better information on depth sensitivities of the scattered wavefield, can help tracking the response of subsurface rocks to different mechanisms. This in turn can provide constraints on the evolving stress and effective rheology at depth.

ACKNOWLEDGEMENTS

We thank D. Agnew for providing us with the water-well and meteorological data, from PFO. G. Hillers acknowledges support from a Heisenberg fellowship from the German Research Foundation (HI 1714/1-1). We also acknowledge support from the European Research Council (Advanced grant Whisper L27507) and from the National Science Foundation (grant EAR-0908903). Figures were made using GMT (Wessel & Smith 1998). We thank X. Briand for his outstanding and timely computational assistance that underpinned much of this work. We thank the Editor Michael Ritzwoller, reviewer U. Wegler and an anonymous reviewer for comments that helped to improve the manuscript.

REFERENCES


