

## Supplementary Material

# Seismic attenuation and stress on the San Andreas Fault at Parkfield: are we critical yet?

Luca Malagnini\*, Robert M. Nadeau, and Tom Parsons

\* **Correspondence:** Corresponding Author: luca.malagnini@ingv.it

### 1 Supplementary Appendix SA - Switching between peak values and Fourier amplitudes and calculating the attenuation parameter as a function of frequency and time

The current Supplementary Appendix provides the theoretical justification for the use of the Convolution Theorem on time-domain quantities like the peak values of a set of narrowband-filtered versions of a seismogram. A brief explanation of the way we quantify the time and frequency variations of the attenuation parameter is also provided, as well as the way we address causality issues.

The method used here for the quantification of seismic attenuation as a function of time and frequency was developed by Malagnini et al. (2019), who modified one originally developed to study the scaling of regional ground motion (see Raof et al., 1999, Malagnini et al., 2000, or Akinici et al., 2001), and used it on a dataset of repeating earthquakes of the Parkfield segment of the SAF (Nadeau and McEvelly, 1995; Nadeau and McEvelly, 1999). The technique was subsequently improved by Malagnini & Parsons (2020) with an application on the same data set, and extensively tested by Malagnini et al. (2022) on a very large set of waveforms from the Central Apennines, showing its ability in detecting extremely small changes of the anelastic attenuation. Because of its sensitivity and accuracy, the method described here is especially suitable for our attenuation study.

The Convolution Theorem is commonly used to write a general form for a predictive relationship for the ground motion:

$$\log_{10}(\hat{a}(f, r_{ij})) = A(f, r_{ij}) = EXC_j(f, r_{ij}) + SITE_i(f) + D(f, r_{ij}), \quad (1)$$

where  $a(t, r_{ij})$  is the time history recorded at the  $i$ -th station and excited by the  $j$ -th earthquake, and  $\hat{a}(f, r_{ij})$  is its Fourier amplitude spectrum. In (1),  $EXC_j(f, r_{ij})$  is an excitation term related to the  $j$ -th seismic source,  $SITE_i(f)$  is the site effect associated to the  $i$ -th seismic station, and  $D(f, r_{ij})$  represents the effect of the crustal propagation ( $r_{ij}$  is the hypocentral distance of the  $i$ -th station).

Note that (1) is in a form suitable for setting a linear inversion problem (in what follows we show that a minor modification of the crustal attenuation term is needed to stabilize it during the inversion). The only issue about the use of (1) is that the Fourier amplitudes of small signals are generally noisy, and one way to maximize the signal-to-noise ratio of the site, source and attenuation terms is to use peak values instead of Fourier amplitudes.

### 1.1 Quantitative description of the ground motion: from Fourier amplitudes to peak value

Given a stationary, random time history  $a(t, r_{ij})$  of length  $T$ , Random Vibration Theory (RVT, Cartwright and Longuet-Higgins, 1956) may be used to relate to its peak value:

$$\text{Peak}(a(t)) \approx \eta a_{RMS}, \quad (2)$$

where  $a_{RMS}$  is its RMS-average calculated over  $T$ , and  $\eta = \eta(m_0, m_2, m_4)$  is a function of the three specified spectral moments of the filtered time history ( $m_n = \int_0^\infty \omega^n |a(\omega)|^2 d\omega$ ).

It is interesting to note that, in our case, RVT is applied to a filtered time history that is significantly different from zero only between  $t = 0$  (the direct S-wave arrival time) and  $t = T$ , where  $T$  is the effective duration of the signal that follows the direct S-waves. We define  $T$  on each seismogram as the time elapsed after 5% of the cumulative energy has arrived, until the time the 75% of the cumulative energy is observed (Raouf et al., 1999).

The Parseval equality relates the time-domain integral of the square of a time-series to the frequency-domain integral of the square of its Fourier amplitude spectrum:

$$\int_{-\infty}^{+\infty} |a(t)|^2 dt = \int_{-\infty}^{+\infty} |\hat{a}(f)|^2 df. \quad (3)$$

Let us now express the RMS-average between  $t = 0$  and  $t = T$  of a windowed time-series using its Fourier amplitude spectrum:

$$a_{RMS} = \sqrt{\frac{\int_0^T |a(t)|^2 dt}{T}} \approx \sqrt{\frac{\int_{-\infty}^{+\infty} |\hat{a}(f)|^2 df}{T}}. \quad (4)$$

We use the approximation sign above because in (4) we do not consider the distortion of the seismic spectrum due to the time-domain windowing. In practical applications, however, we compute the Fourier spectrum over the same finite window  $(0, T)$  of the RMS-average in time domain, and for this reason the equal sign may be used. Now suppose the amplitude spectrum is nonzero only within a certain frequency window, say between  $f_1$  and  $f_2$ :

$$\sqrt{\int_{-\infty}^{+\infty} |a(t)|^2 dt} = \sqrt{2 \int_{f_1}^{f_2} |\hat{a}(f)|^2 df}. \quad (5)$$

So now (4) is to be applied to a windowed time-series that have been narrowly bandpass filtered between  $f_1$  and  $f_2$ , and is significantly different from zero only in the  $(0, T)$  time window. From the application of the Parseval equality to our filtered time history, we can rewrite (3) as:

$$a_{RMS} = \sqrt{\frac{\int_0^T |a(t)|^2 dt}{T}} = \sqrt{\frac{2 \int_{f_1}^{f_2} |\hat{a}(f)|^2 df}{T}}. \quad (6)$$

The filtering action allows us to use the equal sign in (6). In the present study, the observations used in the regressions of the Fourier amplitudes of the ground motion are:

$$A(f) = \log_{10}(\hat{a}_{RMS}(f)|_{f_1}^{f_2}), \quad (7)$$

where:

$$\hat{a}_{RMS}(f)|_{f_1}^{f_2} = \sqrt{\frac{2 \int_{f_1}^{f_2} |\hat{a}(f)|^2 df}{f_2 - f_1}}. \quad (8)$$

If the filter is narrowband, we can still apply the Convolution theorem and write:

$$A(f, r_{ij}) = \text{EXC}_j(f, r_{REF}) + \text{SITE}_i(f) + D(f, r_{REF}, r_{ij}), \quad (9)$$

where  $A(f, r_{ij})$  is the rms-average of the Fourier amplitude spectrum excited by the  $j$ -th event, and recorded at the  $i$ -th station, between frequencies  $f_1$  and  $f_2$ . The propagation term  $D(f, r_{REF}, r_{ij})$  takes into account the crustal attenuation as a function of frequency and hypocentral distance. We will explain later the role of the reference distance  $r_{REF}$  in the excitation and propagation terms of (9).

Eq. (9) is used for the implementation of a linear regression. In our applications,  $f_1 = \frac{f_c}{\sqrt{2}}$ , and  $f_2 = f_c\sqrt{2}$ , where  $f_c$  is the central frequency of our bandpass filter. More precisely,  $f_1$  and  $f_2$  are, respectively, the corner frequencies of an eight-pole high-pass Butterworth filter, and of an eight-pole low-pass filter.

From equations (2):

$$A_{PEAK}(t) = \log_{10}(\eta) + 0.5 \log_{10} \left( \frac{\int_0^T |a(t)|^2 dt}{T} \right). \quad (10)$$

In the frequency domain we can write:

$$A(f_c) = 0.5 \log_{10} \left( \frac{2 \int_{f_1}^{f_2} |\hat{a}(f)|^2 df}{f_2 - f_1} \right). \quad (11)$$

From (6), (10) and (11) we can write a predictive relationship for the peak value of the ground motion of a stationary time history that was narrowband-filtered around a central frequency :

$$A_{PEAK}(f_c, r_{ij}) = EXC_{PEAK_j}(f_c, r_{REF}) + SITE_{PEAK_i}(f_c) + D_{PEAK}(f_c, r_{REF}, r_{ij}). \quad (12)$$

After some simple manipulation, we can explicitly describe the amplitudes measured in the frequency-domain (13), and in the time-domain (14), where  $T_{ij}$  is the effective duration of the post-S-wave signal of the individual seismogram (the seismogram excited by the  $j$ -th earthquake and recorded by the  $i$ -th station):

$$A(f_c, r_{ij}) = 0.5 \log_{10} \left( \frac{2\sqrt{2} \int_{\frac{f_c}{\sqrt{2}}}^{\sqrt{2}f_c} |\hat{a}(f)|^2 df}{f_c} \right); \quad (13)$$

$$A_{PEAK}(f_c, r_{ij}) = \log_{10}(\eta_{ij}(f_c, r_{ij})) + 0.5 \log_{10} \left( \frac{\int_0^{T_{ij}} |a(t, f_c, r_{ij})|^2 dt}{T_{ij}} \right). \quad (14)$$

The advantage of using peak values is great, because they allow a huge enhancement of the signal-to-noise (S/N) ratio. The idea is to use a bandpass filter narrow enough to isolate a specific central frequency  $f_c$ , yet broad enough to allow the filtered time history a relatively sharp peak.

A shallow set of filters (e.g. two-poles) would allow the contributions of a much wider bandwidth to the peak value, well beyond the two corner frequencies  $f_1$  and  $f_2$ . Now suppose a sharp peak characterizes the Fourier amplitude spectrum outside the band defined by the two corner frequencies, yet close to one of them: possibly, the peak value resulting from a shallow bandpass filter would contain a significant contribution from frequencies very different from the central frequency, whereas the application of a sharp set of eight-pole filters minimizes the issues of spectral contamination.

Equation (12) is cast in a matrix form (one equation for each observed peak value) that, after a constraint is pushed onto the site terms, another one is pushed onto the propagation term, and a smoothing constraint is pushed onto the propagation term, is inverted for all source, site, and propagation terms. The two constraints are the following:

$$\sum_{i=1}^N SITE_i(f_c) = 0 \quad (15)$$

and:

$$D(r = r_{REF}, f_c) = 0, \quad (16)$$

for every central frequency  $f_c$ .

The mentioned smoothing constraint is forced on the propagation term once the latter is parameterized as a piecewise-linear function (regressions are performed in a logarithmic space), and constraint (16) effectively decouples the regional attenuation from the source and the site terms. Because source and site terms trade-off, constraint (15) is used to effectively map anything that all sites have in common onto all the source terms. We refer the reader to the studies by Malagnini et al. (2019) for a thorough analysis of the tradeoffs between the different terms of (12).

A set of 20 central frequencies, equally spaced in a logarithmic space between 2 and 50 Hz, is chosen to explore the available bandwidth (see Table 1).

## 1.2 The attenuation parameter

The crustal propagation term of equation (8) at the frequency  $f_c$  can be thought in terms of the product between a geometric attenuation  $g(r_{ij})$ , and a non-geometric (anelastic) attenuation

$$\left( \exp \left[ -\frac{\pi f_c}{\beta Q_S(f_c)} r_{ij} \right] \right): \quad (17)$$

$$d(r_{REF}, r_{ij}, f_c) = \frac{g(r_{ij})}{g(r_{REF})} \exp \left[ -\frac{\pi f_c (r_{ij} - r_{REF})}{\beta Q_S(f_c)} \right], \quad (18)$$

where  $\beta$  is the S-wave velocity,  $Q_S^{-1}(f_c)$  is what we call the attenuation parameter. For the linear inversion:

$$D(r_{REF}, r_{ij}, f_c) = \log_{10} [d(r_{REF}, r_{ij}, f_c)]. \quad (19)$$

Another high-frequency anelastic attenuation filter is commonly used in strong-motion seismology, on top of the filter (17):

$$\left( \exp \left[ -\pi \kappa_{0_i} f_c \right] \right). \quad (20)$$

There is a specific filter (20) for each recording station and, consequently, a specific high-frequency attenuation parameter  $\kappa_{0_i}$ , but in our case we do not need to explicitly take it into account because the regressions force it onto the  $i$ -th site term.

Once a regression is performed on a set of  $N_{EVTs}$  consecutive earthquake adjacent in time, and a second reference distance  $r_{REF_1}$  is chosen, the attenuation parameter may be obtained as:

$$Q_S^{-1}(f_c) = \frac{\beta \left[ \log_{10}(g(r_{REF})) - \log_{10}(g(r_{REF_1})) - D(r_{REF}, r_{REF_1}, f_c) \right]}{[\pi f_c (r_{REF_1} - r_{REF}) \log_{10} e]}. \quad (21)$$

If we associate a specific time  $t_1$  to the attenuation parameter (21) (for example, the time of occurrence of the first earthquake of the set of  $N_{EVTs}$  consecutive events in the specific time window under analysis):

$$Q_S^{-1}(f_c) \rightarrow Q_S^{-1}(t_1, f_c). \quad (22)$$

From (21) we see that the attenuation parameter is effectively obtained from a spectral ratio approach, using the spectral amplitudes inferred at two reference distances.

### 1.3 *Obtaining a time history of the attenuation parameter*

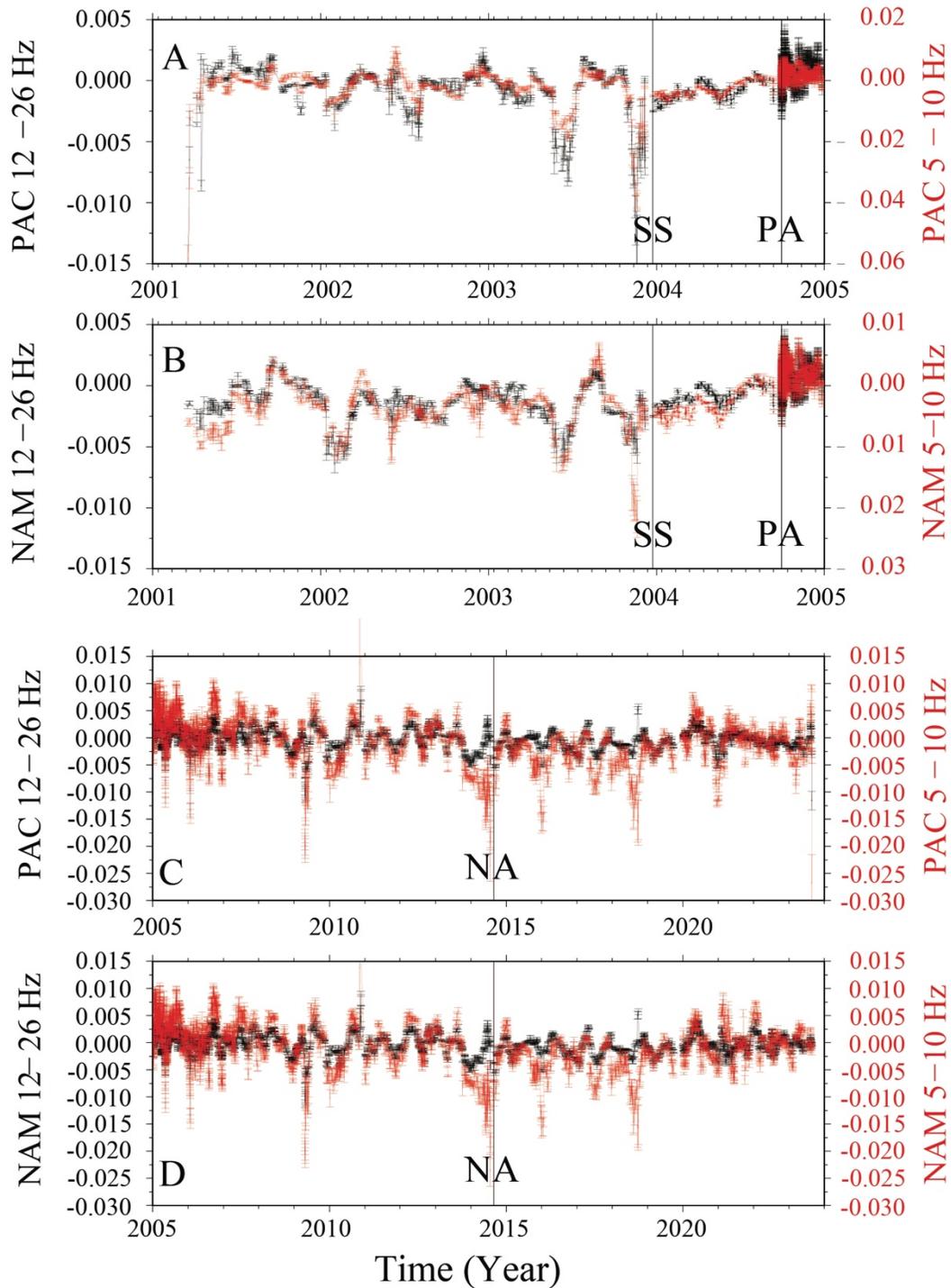
A new set of  $N_{EVTs}$  earthquakes are selected by shifting one earthquake towards more recent times, and adding the next adjacent event that is available in our dataset. A new value for the attenuation parameter can be obtained and associated to time (the time of occurrence of the first earthquake of the new set of  $N_{EVTs}$  earthquakes). In this study, we chose  $N_{EVTs} = 30$ .

Multiple choices are possible relatively to the time to be associated to the time window defined by the specific set of  $N_{EVTs}$  earthquakes: (i) the time of occurrence of the first earthquake; (ii) the time of the last earthquake, (iii) the average time of all earthquakes; (iv) the median time of all earthquakes.

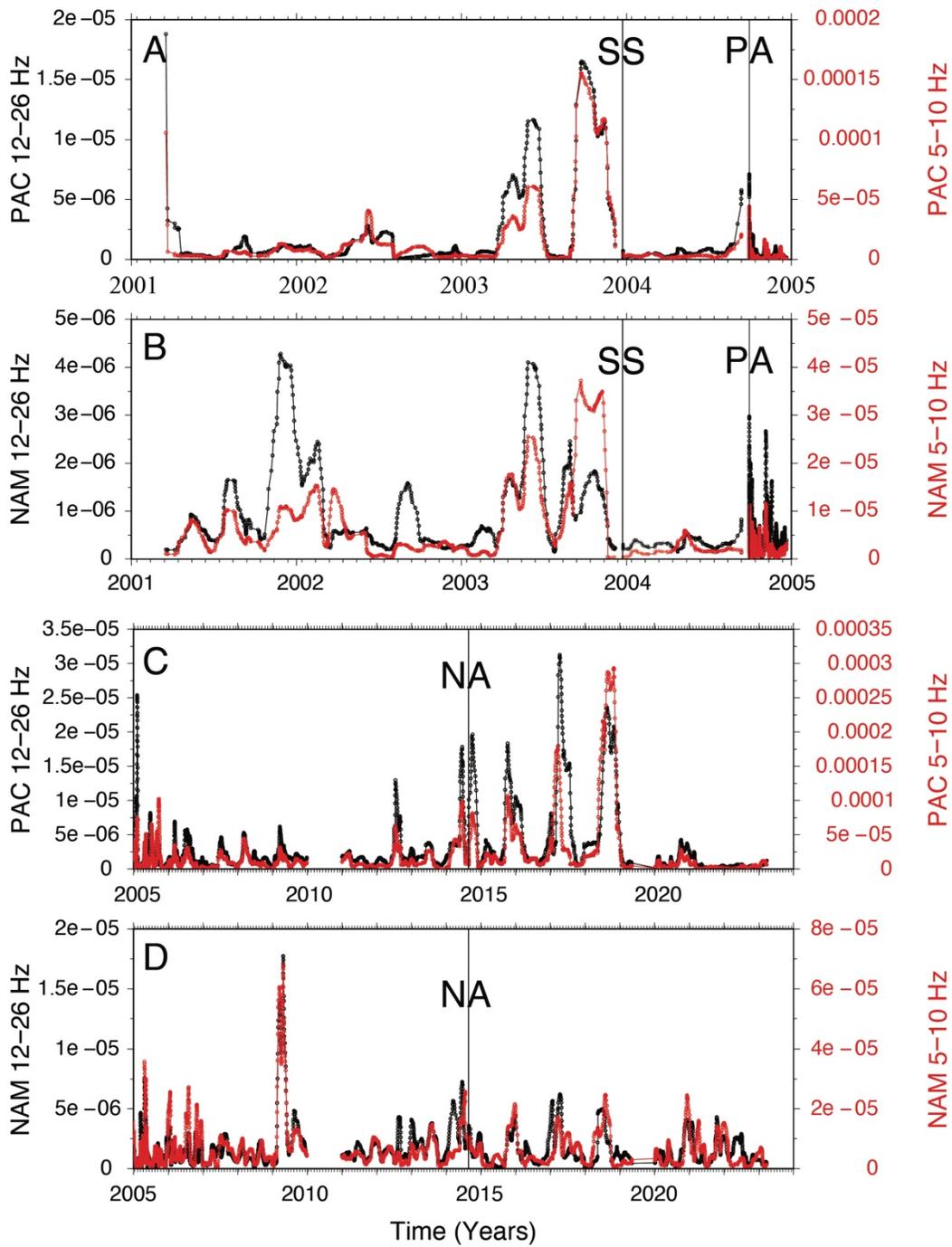
### 1.4 *Causality issues*

Due to the finite duration of the time window of the  $N_{EVTs}$  consecutive earthquakes, the time of a sharp transition between two different regimes (e.g., before and after a mainshock) needs to be treated in a way such that the causality of the changes of the attenuation parameter is preserved. The only way to do so within the described setting is to separate the dataset in two, and to run two sets of regressions on two subsets that do not have any earthquake in common. Such an approach is used here to separate the pre-mainshock dataset from the post-mainshock one. The separation date is September 28, 2004.

## 2 Supplementary Figures

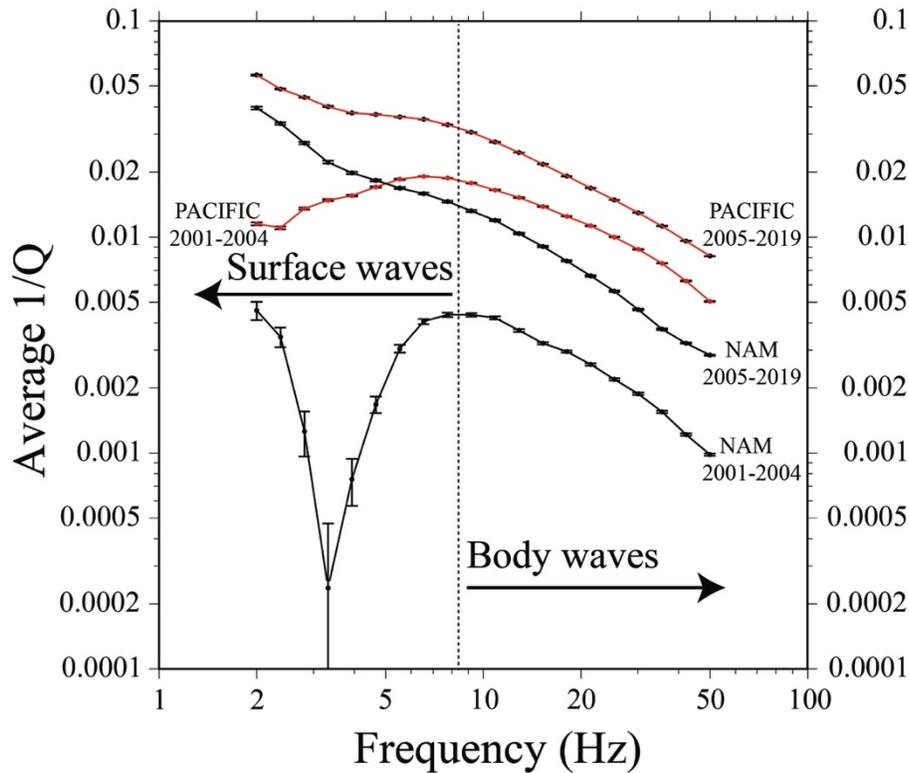


**Supplementary Figure S1.** Stacked-averaged attenuation parameter on the two sides of the fault, in two mid-range frequency bands. The Figure shows the same time windows of Figure 4, but no splitting is observed before the occurrence of the Parkfield mainshock.



**Supplementary Figure S2.** Variance calculated on subsets of 40 data points of the stacked time histories of Figure S1. Note that the San Simeon earthquake completely shuts off the variance of the attenuation time histories on both sides of the fault. A substantial drop in variance is also observed on

the Pacific side of the fault starting in 2020-2021 (the quasi-periodic variance peak disappeared suddenly after a series of peaks having increasing amplitudes).



**Supplementary Figure S3.** Average attenuation parameter ( $Q_s^{-1}(f)$ ) calculated for the North American side (black) and for the Pacific side of the SAF (red), in the indicated time windows (2001-2004 and 2005-2019). Beyond 8-10 Hz, all curves roughly describe the same power-law that is consistent with what is usually thought about the frequency dependence of the quality factor:

$Q_s(f) = Q_0 \left(\frac{f}{f_0}\right)^\eta$ , with  $\eta < 0.5$  in active regions, and  $\eta > 0.5$  in stable environments (Raouf et al., 1999; Malagnini et al., 2007). Below 8-10 Hz, all curves show a flattening that is stronger for the attenuation parameter calculated on both sides of the SAF in the 2001-2004 time window. Such behavior is consistent with the fact that the ground motion in the 2-8 Hz band is rich with surface waves, if not dominated by them. In fact, the attenuation parameter is obtained by the extraction of the body-wave-like geometric contribution from the total attenuation, and surface waves are characterized by a much lower geometric attenuation than body waves. It is interesting to note that the attenuation parameters computed in the 2005-2019 time window (after the 2004 mainshock) are higher than the corresponding parameters relative to the pre-mainshock.