Seismic Noise Autocorrelations on Mars

Martin Schimmel, Eleonore Stutzmann, Philippe Lognonné, Nicolas Compaire, Paul Davis, Melanie Drilleau, Raphael Garcia, Doyeon Kim, Brigitte Knapmeyer-Endrun, Vedran Lekic, et al.

Seismic Noise Autocorrelations on Mars

Martin Schimmel[1], Eleonore Stutzmann[2], Philippe Lognonné[2], Nicolas Compaire[4], Paul Davis[5], Melanie Drilleau[4], Raphael Garcia[4], Doyeon Kim[6], Brigitte Knapmeyer-Endrun[7], Vedran Lekic[6], Ludovic Margerin[8], Mark Panning[9], Nicholas Schmerr[6], John Robert Scholz[10], Aymeric Spiga[11,12], Benoit Tauzin[13], Bruce Banerdt[3]

[1]Geosciences Barcelona - CSIC, Barcelona, Spain.
[2]Université de Paris, Institut de Physique du Globe de Paris, CNRS, Paris, France.
[3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.
[4]Institut Supérieur de l'Aéronautique et de l'Espace SUPAERO, Toulouse, France.
[5]Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, USA.
[6]University of Maryland, College Park, Department of Geology, USA.
[7]Bensberg Observatory, University of Cologne, Bergisch Gladbach, Germany.
[8]Institut de Recherche en Astrophysique et Planétologie, Université Toulouse III Paul Sabatier, CNRS, CNES, Toulouse, France.
[9]Jet Propulsion Laboratory, California Institute of Technology; Pasadena, USA.
[10]Max Planck Institute for Solar System Research, Göttingen, Germany.
[11]Laboratoire de Météorologie Dynamique/IPSL, Sorbonne Université, CNRS, Ecole Normale Supérieure, PSL Research University, Ecole Polytechnique, Paris, France.
[12]Institut Universitaire de France, Paris
[13]Université de Lyon, Université Claude Bernard Lyon 1, ENS, CNRS, Laboratoire de Géologie de Lyon : Terre, Planètes, Environnement, Villeurbanne, France.

Key Points:

- Estimation of high-frequency reflection response using phase autocorrelations of seismic noise recorded on Mars at the InSight landing site.
- Presentation of a new data processing method that avoids aseismic signal bias and stability analysis of the reflection response.
- A signal at 10.6 s lag time is a possible candidate for a reflection from the base of the crust due to its strength, polarity, and stability.

Corresponding author: Martin Schimmel, `mschimmel@geo3bcn.csic.es`

Abstract

Mars is the first extraterrestrial planet with seismometers (SEIS) deployed directly on its surface in the framework of the InSight (*Interior Exploration using Seismic Investigations, Geodesy and Heat Transport*) mission. The lack of strong Marsquakes, however, strengthens the need of seismic noise studies to additionally constrain the Martian structure. Seismic noise autocorrelations of single-station recordings permit the determination of the zero-offset reflection response underneath SEIS. We present a new autocorrelation study which employs state-of-the-art approaches to determine a robust reflection response by avoiding bias from aseismic signals which are recorded together with seismic waves due to unfavorable deployment and environmental conditions. Data selection and segmentation is performed in a data-adaptive manner which takes the data root-mean-square amplitude variability into account. We further use the amplitude-unbiased phase cross-correlation and work in the 1.2-8.9 Hz frequency band. The main target are crustal scale reflections, their robustness and convergence. The strongest signal appears at 10.6 s, and, if interpreted as P-wave reflection, would correspond to a discontinuity at about 24 km depth. This signal is a likely candidate for a reflection from the base of the Martian crust due to its strength, polarity, and stability. Additionally we identify, among the stable signals, a signal at about 6.85 s that can be interpreted as a P-wave reflection from the mid-crust at about 9.5 km depth.

1 Introduction

Since December 19, 2018, Mars is the first extraterrestrial planet with seismometers deployed directly on its surface. The seismometers are called SEIS (*Seismic Experiment for Interior Structure*) and have been brought by the InSight (*Interior Exploration using Seismic Investigations, Geodesy and Heat Transport*) mission (Lognonne et al., 2020). SEIS has been designed for planetary seismology (Lognonne et al., 2019) and consists of six axes to measure ground motion. It contains three very broad band (VBB) oblique axes which are sensitive to frequencies from tidal up to 10 Hz and three short period (SP) axes (one vertical, two horizontal) which are sensitive to frequencies from about 0.1 Hz to 50 Hz.

Before InSight, two SP seismometers were already landed on Mars with the 1976 Viking 1 and 2 missions (e.g., Anderson et al., 1977; Lognonné & Johnson, 2007). In contrast to InSight, both seismometers were mounted on the lander. While the Viking 1 seis-

mometer did not work properly, the Viking 2 seismometer picked up mostly vibrations from the lander owing to operation and wind sensitivity. The lander vibrations and non-ideal frequency-dependent sensitivity of the Viking SP sensor inhibited the detection of marsquakes and Viking did therefore fail to provide unambiguous data for Martian seismic activity and subsurface structure studies.

SEIS has been deployed on the ground of the Homestead hollow, 1.81 m south of the closest lander foot (Stutzmann et al., 2021). Homestead hollow is a degraded impact crater of about 27 m diameter in Elysium Planitia with a smooth sandy, granule- and pebble-rich surface with few rocks (Golombek et al., 2020). After leveling and functionality checks SEIS has been covered with a wind and thermal shield (WTS) to further protect the sensors from environmental factors such as laminar and turbulent winds which can reach about 20 m/s, atmospheric pressure changes of a few Pa, low atmospheric temperature down to below -100°C and daily temperature variability in the order of 80°C (Lognonne et al., 2020). In spite of these harsh conditions, 174 mostly small marsquakes (Giardini et al., 2020) have been identified until 30 September 2019 thanks to the sensitivity of SEIS and its careful deployment on the Martian surface.

Nevertheless, due to the unfavorable deployment conditions different aseismic signals are being recorded together with events of seismic origin (Ceylan et al., 2021; Lognonne et al., 2020; Scholz et al., 2020; Stutzmann et al., 2021). That is, the seismic recordings contain a wealth of features caused by operational activities at the lander, noise induced by the lander due to atmospheric perturbations as laminar and turbulent winds, and artifacts caused by the response of the instruments to the variability of the severe climatic factors as pressure and temperature. Data processing requires therefore special care and adapted approaches to avoid any bias in the results and even misinterpretation of signals.

SEIS is a key instrument to reveal the internal Martian structure and to monitor its seismic activity. However, the deployment of a seismometer at one single place limits the ample spectrum of seismological imaging and monitoring strategies to single-station approaches (e.g., Panning et al., 2015; van Driel et al., 2019; Drilleau et al., 2020). Prominent single-station methodologies are P- and S-wave receiver functions (e.g., Phinney, 1964; Langston, 1979; Vinnik, 1977; Farra & Vinnik, 2000; Yuan et al., 2006), noise and coda autocorrelation functions (e.g., Claerbout, 1968; Galetti & Curtis, 2012; Tibuleac

& von Seggern, 2012; Pham & Tkalčić, 2017; Tauzin et al., 2019; Buffoni et al., 2019), surface wave inversions (e.g., Drilleau et al., 2020) and Rayleigh wave ellipticity and spectral ratio analyses (e.g., Nakamura, 1989; Yano et al., 2009; Hobiger et al., 2013; Berbellini et al., 2019).

So far no unambiguous surface waves have been detected in SEIS data (Clinton et al., 2020; Giardini et al., 2020; Stutzmann et al., 2021). The very shallow structure down to 20 m depth at the InSight landing site has been probed using the travel time of multiple hammer strokes of the HP^3 (Heat Flow and Physical Properties Package) instrument and compliance observations (Kenda et al., 2020; Lognonne et al., 2020). The compliance approach employs seismic measurements of the surface deformation in response to pressure loading by dust-devils. First average crustal seismic attenuation and scattering results have been obtained from the body-wave coda shape of the best recorded marsquakes while first crustal layering has been inferred identifying body wave conversions in receiver functions (Lognonne et al., 2020). These results point to a crustal attenuation which is 3 times larger than on Moon and a highly variable upper crust layer of about 8 to 11 km thickness with S-wave velocities of 1.7 to 2.1 km/s. The exact thickness of the crust is still object of investigation, but receiver functions, event coda and noise autocorrelation results suggest that the crust is either about 15-26 or 27-47 km thick (Knapmeyer-Endrun et al., 2021). Deng and Levander (2020) have also computed noise autocorrelations for Mars and observe persistent signals which they interpret as Moho, upper-mantle transition and core-mantle boundary reflections. Their mantle signals have not yet been reproduced by other groups and can possibly be due to repeated glitches in the SEIS-data or interference of lander resonances (*Kim, Davis et al., in preparation*). High-frequency autocorrelations (5-7 Hz) have been presented by Suemoto et al. (2020) who showed robust signals which indicate the presence of two shallow reflectors in the first hundreds of meters. Noise and Marsquake autocorrelations have been also presented by Compaire et al. (2021), who find stable autocorrelations with signals related to crustal structure only during low-noise periods around 2.4 Hz

In this study we use the VBB SEIS data to further analyze noise autocorrelations through independent approaches. Special care is given to avoid imprints of aseismic signals and artifacts on the autocorrelation results. The main objective is to extract a stable and clean reflection response from the noise recordings. We focus on crustal scale signals from vertical component recordings within the 1-9 Hz frequency band, and use dif-

ferently selected subsidiary data sets. The data selection strategy is novel and we show that a stable noise response is built up within a few Martian days of data and that it consists of several signals. Some of the signals are stable over a broad frequency band which makes them possible candidates for reflections. In any case, the autocorrelations are still difficult to interpret and possible ambiguities in the signal interpretations are discussed.

2 Determination of P-Wave Reflection Response at SEIS

On Earth, seismic ambient noise correlations have been successfully used to map crustal scale discontinuities (e.g., Ruigrok et al., 2011; Tibuleac & von Seggern, 2012; Gorbatov et al., 2013; Kennett et al., 2015; Taylor et al., 2016; Oren & Nowack, 2016; Saygin et al., 2017; Becker & Knapmeyer-Endrun, 2018; Romero & Schimmel, 2018; Buffoni et al., 2019). Strictly, the cross-correlation of a diffuse wave field recorded at two sensors provides the Green's function (GF) at one of the sensors for a virtual source placed at the other sensor (Lobkis & Weaver, 2001; Derode et al., 2003; Wapenaar, 2004; Snieder, 2004, among others). When both recordings are the same then the cross-correlation becomes an autocorrelation and provides a zero-offset GF. The high-frequency zero-offset GF is the reflection response which mainly consists of body wave primary reflections and multiples from seismic discontinuities beneath the station.

This principle has been presented already by Claerbout (1968) for one-dimensional (1-D) media as he demonstrated that the full reflection response can be obtained from the autocorrelation of a plane wave field which is transmitted from below the reflecting structure. Many years later, his finding has been extended to two-dimensional (2-D) and three-dimensional (3-D) acoustic and elastic media (e.g., Wapenaar, 2004).

The numerical determination of the zero-offset reflection response is based on the computation of autocorrelations. For this purpose the continuous noise recordings are cut into short segments to compute the autocorrelations. Finally, the autocorrelations of all segments are being stacked to provide the reflection response. In practice, wave fields are generally not diffuse and the exact GF is unknown or only partly reconstructed. We therefore now refer to empirical GF (EGF) or reflection response rather than GF to emphasize its approximate character.

Note that outliers in the amplitudes of the data can affect correlation results and it is therefore common to preprocess the data before computing autocorrelations (e.g., Bensen et al., 2007; Schimmel et al., 2011). Different strategies exist to balance the amplitudes in the time and frequency domain. One-bit normalization and spectral whitening are often employed methods (Bensen et al., 2007). A disadvantage of these approaches is that they reduce the information content of the signal in a non-unique manner which can cause a less efficient EGF extraction as shown in Schimmel et al. (2018). Here, we avoid amplitude normalization strategies by replacing the conventional cross-correlation and linear stack with the amplitude unbiased phase cross-correlation (PCC, Schimmel (1999)) and time-frequency phase weighted stack (tf-PWS, Schimmel and Gallart (2007); Schimmel et al. (2011)).

In the following we highlight the main methods used to determine the reflection response at the InSight landing site and describe our data selection and processing.

2.1 Phase Cross-Correlation and Stacking

Correlations. The autocorrelation measures the self-similarity of a time series as function of lag time. It is a special case of the cross-correlation where both time series are the same. There exist different strategies to measure the self-similarity of time series (e.g., Schimmel et al., 2018). Here, we employ the phase cross-correlation $c_{PCC}(\tau)$ (PCC, Schimmel (1999)) which is based on the phase coherence of instantaneous phases as determined from analytic signal theory. PCC is expressed as:

$$c_{PCC}(\tau) = \frac{1}{2T} \sum_{t=1}^{T} \left| e^{i\Phi(t)} + e^{i\Psi(t+\tau)} \right|^{\nu} - \left| e^{i\Phi(t)} - e^{i\Psi(t+\tau)} \right|^{\nu}. \tag{1}$$

$\Phi(t)$ and $\Psi(t)$ are the instantaneous phases of the two input time series, τ is the lag time of the second time series with respect to the first one, T is the length of the time window, and ν a parameter which permits to tune the sensitivity of the PCC.

$e^{i\Phi(t)}$ is the envelope normalized analytic signal of the first time series. The analytic signal can be build with the Hilbert Transform and is a unique representation of a real-valued time series in the complex number space. PCC benefits from the fact that thanks to the analytic signal theory a real time series can be decomposed into an instan-

taneous phase and amplitude (envelope) function. PCC uses only the instantaneous phases and is therefore explicitly signal amplitude unbiased.

The phase autocorrelation is obtained using $\Phi(t) = \Psi(t)$. Further, we use $\nu = 2$ since it permits to simplify the equations for fast computations (Ventosa et al., 2019). $c_{PCC}(\tau)$ is, in analogy to a classical correlation, a real numbered functional with values ranging between -1 and 1. One of the main differences to the classical correlation is that waveform similarity is measured through the amount of phase-coherent samples rather than the sum of amplitude products.

Stacking. The obtained autocorrelations need to be averaged over larger time spans to achieve a stable noise response. Stacking data over larger time spans improves the azimuthal coverage of the noise wavefield and the cancellation of cross terms (e.g., Snieder, 2004; Medeiros et al., 2015). Both are necessary conditions in seismic interferometry. Here we employ linear stacks and time-frequency Phase Weighted Stacks (tf-PWS, Schimmel and Gallart (2007); Schimmel et al. (2011)). The tf-PWS are based on the instantaneous phase coherence (in analogy to PCC) which is being measured in the time-frequency domain during stacking. The weights range between 0 and 1 to attenuate less coherent signals in the time-frequency domain. Finally, linear stacks are weighted by the time-frequency phase coherence to build the tf-PWS.

The time-frequency representation of the data is obtained with the S-Transform by Stockwell et al. (1996) which is based on Fourier theory and frequency-dependent Gaussian-shaped windows. The obtained time-frequency representation can be made an analytic signal as shown in Schimmel and Gallart (2007) to enable the use of an instantaneous phase coherence measure which is the backbone of tf-PWS. Alternatively, tf-PWS can be presented using the wavelet transform (Ventosa et al., 2017) to reduce redundancies and to increase computational efficiency. This modification is based on the fact that the S-Transform with Gaussian-shaped windows can be derived from the Morlet wavelet transform (Ventosa et al., 2008). Both strategies provide exactly the same result.

2.2 Data and Time Frame

The VBB SEIS data are recorded on three oblique axes. Data from the three axes are needed to build the vertical and horizontal components (Lognonne et al., 2019). For this study we use vertical components, recorded from May 22, 2019 to June 30, 2020 at

20 samples-per-second and instrument corrected to ground velocity from 0.01 to 10 Hz. The correction is performed for the three oblique axes before rotation to the vertical and horizontal components. The late starting date has been chosen to avoid the initial deployment and testing phase. Key dates related to the seismic part of the mission are listed in Table 1.

As time frame, we employ the Local Mean Solar Time (LMST) throughout the entire study. The Martian day, called Sol, is about 40 min longer than a day on Earth with a Martian year being about 687 Earth days. Thus, using LMST is justified since the time on Mars differs from Earth time. The Martian dates, as usual in other missions, are defined as simple numerical counts using Sol 1 for the first significant Mars-day which for InSight corresponds to November 26, 2018. The data used here are from Sol 172 (2019-05-21 at 22:39:52.32) to Sol 567 (2020-06-30 at 19:16:53.76). Our initial database has been obtained cutting the instrument-corrected, vertical component records within every Sol into 7400 s segments which overlap by 600 s.

2.3 Data Segmentation and Selection

We filter the seismic recordings to a broad frequency band, 1.2 - 9.8 Hz, and determine the relative RMS variability for sliding data windows. A maximum RMS variability threshold is then defined to obtain a mask and to extract data segments with RMS variability below the threshold through a minimum time duration. We use this procedure to build subsidiary data sets which contain less data problems such as glitches and outlying amplitude events. The approach and the subsidiary data sets are described in the following.

The moving window RMS r is determined for window i and length N using

$$r_i = \sqrt{\frac{1}{N} \sum_{n=i-N/2}^{i+N/2} a_n^2}, \tag{2}$$

where n stands for the time index and a_n for the seismic record. i marks the center sample of the sliding window. In a next step the relative variance s^2 of the RMS amplitudes is computed during a second moving window analysis with window length $M > N$. The relative variance s^2 of window j is expressed as

$$s_j^2 = \frac{\left(\sum_{i=j-M/2}^{j+M/2} r_i - R_j\right)^2}{(M-1)R_j^2} \text{ with } R_j = \frac{1}{M} \sum_{i=j-M/2}^{j+M/2} r_i, \tag{3}$$

where R_j is the non-zero mean RMS amplitude within the analysis window j. We define the relative variance s^2 as the variance divided by the square of the mean to obtain a dimensionless (scale invariant) measurement of variability. This definition permits to compare data segments with variable RMS mean and to adjust to different seismic background noise levels which may happen along the Sol.

In a next step, data segments are determined with $s_j^2 < S^2$, where S^2 is a chosen maximum RMS variability. A new subsidiary data base is then built accepting only data chunks with time duration larger than a minimum time duration, say t_{min}.

Thus, the free parameters of this data selection and segmentation approach are the window length (N and M of eqs. 2 and 3) and step intervals (i and j) for the two analysis windows, the maximum RMS-variability S^2 and the minimum data segment length t_{min}.

On one hand, the analysis windows should be several times longer than the largest period as defined by the lowest corner frequency of the bandpass filter to warrant statistical significance. On the other hand, the windows should not be too long to permit tracking RMS-variability changes with time. Throughout this study we used window length and step interval of 5 s and 0.1 s for the sliding-window RMS measurements (eq. 2) and 20 s and 1 s for the RMS-variability determination (eq. 3). Further, the shortest admitted data segment length (t_{min}) is 300 s and the chosen RMS-variability thresholds (S^2) are 0.1 and 0.2. Using these parameters two subsidiary data sets have been set up which consist of data volumes of about 3% (RMS-variability threshold 0.1) and 30% (RMS-variability threshold 0.2) of the total data base. In the following, we will refer to these data sets as the total or 100%, 30%, and 3% data sets.

Different other parameters have been tested. The thresholds on data segment lengths and RMS-variability have a direct influence on the amount of selected data while varying the analysis window parameters within reasonable bounds has a minor or insignificant impact on the subsidiary data base. In any case, there exists a trade-off between the selected amount of data and tolerance of RMS-variability changes. A fine tuning with the parameters may make sense for data sets where outlying noise is clearly separated from the rest of the data. In our case, aseismic noise exists at any amplitude level and therefore, it seems easiest to just analyze the results for more or less restrictive subsidiary data sets.

Examples of the data selection and segmentation procedure are shown in Fig. 1. The top traces of Figs 1a-c show seismic recordings with a duration of about 2 h. The start time of each trace is given in the figure. The second and third traces of each panel are the corresponding RMS (r_i) and RMS-variability (s_j, eqs 2 and 3). The horizontal red lines mark the RMS-variability threshold (0.2) to find the data segments with low RMS-variability. Segments with a duration larger than $t_{min} = 300$ s are drawn in red and form part of the 30% data set.

It can be seen from Fig. 1 that the selected red segments are data stretches with low-amplitude variability. I.e, outlying amplitude signals have been avoided by the approach. Fig. 1c also shows that segments with different amplitudes are selected inherent to the use of relative variance s_j^2. Further examples are shown in Fig. S1.

Fig. 2a shows the evolution of the relative RMS variability as function of Sol. Plotted are mean (red dots), one standard deviation uncertainty (gray bars), minimum and maximum value (black triangles) per Martian-day. All these values stay stable until about Sol 450. From about Sol 450 a slight but systematic increase of the maximum and mean RMS variability is being observed. A more rapid increase of the minimum RMS variability is evident from about Sol 500 to Sol 530. At the end of the analyzed period all values are increased with respect to the beginning of the study period. The RMS variability increase correlates with a systematic raise of bad weather attributed to the entrance into local winter with more storms (Spiga et al., 2020). For example, Fig. 2b illustrates the measured mean wind speed (red dots), its one standard deviation uncertainty (gray bars), and minimum and maximum wind speed as function of sol. It can be seen that the minimum wind speed increases fast from about Sol 500 on, which is in concordance with the rise of the minimum RMS variability caused by the presence of winds at the lander site at all times. The highest wind speeds and standard deviation values decrease from about Sol 420 which reflects that the wind characteristics and distribution changed at the lander as also recorded by SEIS.

The just mentioned rise of RMS variability causes a decreased amount of selected data towards the last analyzed Sols as also documented in Fig. S2. This figure shows the LMST of the selected data segments as function of Sol for the 30% and 3% data sets. It can be seen that most of the selected data segments are from the evening and morning hours. This is expected since lander and wind activity are low during the night. Still,

glitches and other aseismic abrupt signals may happen during night which makes our data adaptive selection and segmentation approach different from any time dependent (e.g., day/night) selection. It is also seen that less data has been selected during the last 60 Sols owing to the increase of bad weather during this period.

The data segment length distribution for the 30% subsidiary data set is shown in Fig. 3a. One third of the data segments has lengths smaller than 10 min. The number of segments decays quickly as function of segment length. Fig. 3b illustrates a normalized RMS amplitude distribution. The vertical axis now shows the total time duration rather than the number of individual segments. The gray histogram is for the 100% data set while the red histogram corresponds to the 30% data set. The RMS normalization is the same for both data sets. The absolute RMS amplitude is not important here, but the comparison of both histograms shows that the 30% data set consists mostly of small-amplitude segments. The 100 % data set RMS amplitudes have been obtained for 2-hour windows which consist typically of small and high RMS amplitude variability. Conversely, the 30% data set has a variable window length and does not include data stretches of high amplitude variability. This explains the increased total time duration for small RMS.

2.4 Phase Autocorrelation Spectra for Selected Data Sets

The Fourier Transform of the conventional autocorrelation of a time series equals its energy spectral density (ESD) for time series which have finite duration and are square integrable. In analogy to the ESD, we compute here the Fourier amplitude spectra of phase autocorrelations. The obtained spectra do generally not equal the ESD of the data as we use PCC to determine autocorrelations. Dissimilarities are often small and inherent to the different measure of signal coherence with each autocorrelation approach. However, outlying amplitude signals do not bias PCC which permits a fast convergence to the EGF and stable autocorrelation spectra as shown in Schimmel et al. (2018).

Fig. 4 shows the Fourier amplitude spectra of the linear stack of the phase autocorrelations using the 100% (black line), 30% (red line) and 3% (blue line) data sets. The data have been band-pass filtered between 0.8 and 9.5 Hz before computation of the autocorrelations. Different spectral lines and bumps are visible in these spectra. The spectral lines at integer frequencies are related to the “tick noise“ which is an aseismic signal caused by electrical coupling during the SEIS temperature acquisition at 1 sps (sam-

ple per second). The spectral amplitude bumps are related to lander modes and/or resonance of the seismic subsurface structure (e.g., Ceylan et al., 2021). The main bumps have been labeled along the spectrum for the 100 % data set.

Bump 1 and spectral lines at integer frequencies appear to be enhanced in the data sets with less RMS amplitude variability, i.e., for the 30% and specially 3% data sets. This is partly due to the attenuation of other signals and noise with higher RMS amplitude variability. The origin of bump 1 is still under investigation. It is centered at about 2.4 Hz and could be caused by the superposition of lander and structural resonances. The corresponding lander modes at bump 1 are expected to arise as a response of the solar array oscillations during laminar wind flow, i.e., at wind speeds below about 3 m/s. This may explain why this bump 1 appears to be enhanced for the data sets with less RMS amplitude variability. Conversely, bumps 2-6 are systematically attenuated in the 30% and 3% data sets. These data sets consist mostly of seismic recordings during the local evening hours (Fig. S2) which is when there are less wind and less air pressure perturbations (e.g., Lognonne et al., 2020; Stutzmann et al., 2021). Bumps 2 to 6 correspond to resonances that are related to different lander modes which are mostly excited by turbulent winds above a threshold speed of about 3 m/s (Dahmen et al., 2021). In the following we will work in the 1.2 - 8.9 Hz frequency band to avoid resonances at 1 and 9 Hz.

2.5 Attenuation of Tick Noise, Lander and Shallow Structure Resonances

Fig. 5a shows vertical-component phase autocorrelations stacks as function of autocorrelation lag time and recording date. The entire data set has been used without any selection criteria, cut into 2-hour data segments and frequency band passed between 1.2-8.9 Hz before computation of the phase autocorrelations. The autocorrelations have been stacked linearly for recording dates within non-overlapping sliding 3-Sol data windows. In this figure, the positive amplitudes have been filled with red color and the positively correlated signals seen at every full second are tick noise (Fig. 5a). We have chosen a window at large lag time to avoid the interference with reflections from shallow discontinuities to permit a visual judgment on the tick noise attenuation.

The tick noise is caused by electro-magnetic coupling during the SEIS temperature acquisition at 1 sps with a recorded waveform which does not resemble an impulsive sig-

nal (Compaire et al., 2021). On vertical components and within the here considered frequency band, most of its energy is at 4, 7, and 2 Hz (Fig. 4).

Recorded tick noise and lander modes may interfere with seismic weak-amplitude reflections and should ideally be removed from the data to warrant the detection of weak amplitude reflections. Similarly, structural resonances should also be attenuated as their dominant signals may impede the detection of weak amplitude reflections. Structural resonances are caused by seismic energy which is trapped in a shallow subsurface layer due to strong impedance contrasts at the limits of the layer. Only little energy leaks to deeper layers which therefore can inhibit the imaging of deeper structures with frequencies in the structural resonance band.

Here, we attenuate the tick noise together with other resonances using band-rejection filters. First, we try two band-rejection filters (3.9-4.4 Hz and 6.8-7.2 Hz) to remove the two strongest tick noise peaks at 4 and 7 Hz and neighboring lander modes (bumps 4 and 6). Fig. 5b shows in full analogy to Fig. 5a the autocorrelations after the application of the two just mentioned band-rejection filters. Bump 3 has not been included into the first rejection band since this signal is already attenuated for the 30% data set (Fig. 4, middle panel).

It can be seen from Fig. 5b that there exist still some lower-frequent tick noise which justifies employing a third band-rejection filter from 1.9 to 2.5 Hz. The resulting autocorrelation section Fig. 5c is finally not dominated by tick noise anymore. Besides, the data has been cleaned from other resonances (bumps 1, 4 and 6).

An alternative approach is to measure the tick waveform through stacking of 1 s data segments for subsequent subtraction from the data (Compaire et al., 2021). Here we chose a different and independent approach to further show stability in our results.

3 Reflection Response and Interpretations

Fig. 6 shows autocorrelations computed using PCC and stacked with the tf-PWS approach. The employed seismic recordings belong to the 30 % data set. We bandpass filtered these data from 1.2 to 8.9 Hz and applied the three band-rejection filters explained in the previous section. The autocorrelations have been stacked within non-overlapping 3-Sol windows and negative amplitudes are plotted in blue. Note that P-wave reflections at seismic discontinuities with an impedance increase are expected to have negative am-

plitudes. The autocorrelation section of Fig. 6 has been split into three lag-time windows to apply an independent amplitude normalization in each window for visual purposes. The earliest lag time starts at 5 s as the first seconds are dominated by zero-lag autocorrelation sidelobes (e.g., Ruigrok & Wapenaar, 2012; Romero & Schimmel, 2018). The amount of stacked data is shown to the top of each autocorrelation. The maximum duration of the data used per autocorrelation is about 1 Sol rather than 3 Sol which is the length of the data window. This is because of the applied data selection which reduced the data used here to about 30 %. The corresponding linearly stacked autocorrelation section is displayed in Fig. S4 of the Supplementary Material. As also shown further below, the tf-PWS approach provides cleaner results as incoherent signals and noise have been attenuated during the stacking.

The autocorrelation section of Fig. 6 contains signals repeated on most autocorrelations at specific lag times. These signals are stable as they show up for independent autocorrelations along the entire analyzed time interval. One of the most outstanding signals is seen at about 10.6 s. If we interpret this arrival as a P-wave reflection and assume an average crustal P-wave velocity of 4.5 km/s then the reflector is expected to be at about 24 km depth. Note that the assumed average P-wave velocity is slower than for Earth as indicated by receiver function inversions (Lognonne et al., 2020; Knapmeyer-Endrun et al., 2021).

As the raw data consists also of different other aseismic signals, an autocorrelation signal is not a synonym of a seismic reflection. Some of the signals are expected to be part of the seismic reflection response while other signals can be due to any coherent and repeated feature in the data or due to beating of resonating frequency components from the lander. Resonances with slightly different frequencies f_1 and $f_2 > f_1$ cause an interference pattern (wellknown in music) with beat frequency $f_b = f_2 - f_1$ (e.g., Kinsler et al., 1999). The beats of the interferring signals manifest in time-domain autocorrelations as distinct and repeating signals of maximum amplitude at lag times $t_n = n/f_b$ with $n \in N$. Similarly, glitches or other signals which are repeated systematically with the same time interval are expected to show up in autocorrelations at lag time which corresponds to the time separation of the repeated features.

Indeed, the interpretation of autocorrelations without any further a priori information is difficult. However, receiver function studies of marsquakes (Knapmeyer-Endrun

et al., 2021; Lognonne et al., 2020) provide the first a priori information of the underlying structure and expected crustal-scale P-wave reflections on Mars. In this respect, the strong arrival at 10.6 s lag time interpreted as a P-wave reflection seems to be consistent with P-to-s wave conversions observed with receiver functions (Knapmeyer-Endrun et al., 2021). Indeed, the 10.6 s signal is quite robust and a likely candidate for a P-wave reflection from the crust-mantle boundary. Note that Compaire et al. (2021) also report a signal at 10.6 s using a different approach.

The 10.6 s signal is discernible in linear and tf-PWSs of the three data sets as shown in Fig. 7. This figure contains the linear stacks in black and the tf-PWSs in red for all autocorrelations from the 100 % (top traces), 30 % (middle traces), and 3 % (bottom traces) data sets. The traces have been normalized by their RMS amplitudes measured between 8 and 20 s lag time. Further, amplitudes are clipped during the first seconds to favor the visibility of later arriving signals. It can be seen that the 10.6 s signal is visible with amplitude of about twice the amplitude of neighboring signals and noise. Its shape resembles more a zero-phase wavelet than a wave train. The direct comparison of the stacks also shows that the tf-PWS approach provides cleaner autocorrelations through the attenuation of incoherent signals and noise.

3.1 Convergence

We now look at the convergence of autocorrelation stacks towards a stable seismic noise response. In seismic interferometry, the minimum amount of data needed is controlled by the cancellation of noise-cross terms (e.g., Snieder, 2004) and the abundance, distribution, quality and duration of signals to build up a stable EGF.

Under the assumption of stationarity and minimum frequency of 1 Hz, cross-terms cancel out within confidence level $\epsilon \approx 0.01$ after about 3 h (eq. 9 in Medeiros et al. (2015)). The SEIS data are highly variable and the assumption of stationarity is not valid. Nevertheless, the estimated 3 h can be taken as a guide value. This value is shorter than the used data length of each autocorrelation shown in the figures for the 30% and 100% data sets.

The total data length needed to extract a robust reflection response can be estimated by a convergence analysis. In this approach the evolution of autocorrelation stacks of random subsidiary data sets is being inspected as function of amount of data used for

the autocorrelations in each of the stacks. For this purpose the individual autocorrelation stacks are being compared with a reference waveform obtained by stacking all available data. The waveform similarity of the two stacks can be determined with the zero-lag cross-correlation of the two time-series. Here, we employ PCC and the geometrically normalized cross-correlation (CCGN, e.g., Schimmel (1999)). The reference traces are those shown in Fig. 7.

There exist different data sampling strategies to draw the random autocorrelation stacks. We tested two strategies, random sampling without replacement and Bootstrapping which uses random sampling with replacement (Efron & Tibshirani, 1986). Without replacement means that no autocorrelogram is used more than one time in a draw. We repeated the random drawing 40 times to enable the determination of mean and standard deviation for the similarity to the reference waveform.

Figs 8 and S5 demonstrate our results using the Bootstrapping resampling and PCC and CCGN, respectively, to measure waveform similarity in moving windows of 0.5 s length. The 30 % data set has been used for Figs 8a,b. Fig 8a shows the waveform similarity while Fig. 8b contains the standard deviation of the similarity. The similarity (red line) and corresponding standard deviation (gray error bars) at 10.6 and 19.5 s are shown in the Figs 8c,d. 19.5 s has been chosen arbitrarily for comparison. A plateau in the similarity curve means that stacking more autocorrelations is not changing significantly the waveform similarity. The figures illustrate that the 10.6 s signal converges faster than a signal at 19.5 s lag time. Using CCGN (Fig. S5) an apparent overall faster waveform similarity is measured than using PCC. This is expected as PCC is the more waveform sensitive measure due to the employed instantaneous phase coherence.

From Figs 8a and S5a we also observe that signals at shorter lag time tend to converge faster than signals at larger lag time. For a P-wave reflection response a larger lag time means longer wave paths owing to a deeper discontinuity or multiple reverberation. Such signals may need more data to stabilize in the EGF since the corresponding waves are more susceptible to the different types of attenuation.

A comparison of the similarities for the 100 %, 30 %, and 3% data sets is shown in Figs 8e,f and S5e,f. It can be seen that slowest signal convergence is achieved using the 100 % data set (black curves). This means that a high amplitude variability is not contributing significantly to the extraction of the EGF. Note that this result does not

494 change when using only the data until Sol 410. I.e., the increasing RMS variability af-
495 ter Sol 450 (Fig. 2) is not causing the slower convergence.

496 The 3 % and 30 % data sets seem to provide more similar results. The 3 % and
497 30 % data sets have a total length of 151.8 h and 1557.1 h. The 3 % data set results are
498 therefore obtained with a higher replacement rate. Drawing the random data sets with-
499 out replacement (not shown here) illustrates that the similarity curves for the 3 % and
500 30 % data are still more similar than for the 100 % data set.

501 ### 3.2 Stability of Reflection Response over Time

502 The autocorrelations for the three data sets are shown in Fig. 9. The only differ-
503 ences to Fig. 6 are the respective data sets (from left to right: 100 %, 30 %, and 3%)
504 and the use of non-overlapping 30-Sol windows rather than 3-Sol windows. 30-Sol win-
505 dows have been employed to reduce the number of autocorrelations to permit a compar-
506 ison of the three autocorrelation sections at one glance. A welcomed side effects is that
507 these autocorrelations are more robust owning to the increased amount of data used in
508 each of the stacks. The corresponding linear stacks are demonstrated in Fig. S6 for com-
509 pleteness.

510 The most striking observation is that the 100 % autocorrelations show systemat-
511 ically changing noise responses, i.e., the autocorrelations from about Sol 450 towards the
512 end of the section look different with respect to the autocorrelations obtained for the first
513 Sols. This variability is likely due to changing climatic conditions with stronger storms
514 and corresponding responses of the lander. In fact, the statistical properties of the seis-
515 mic recordings start to change simultaneously at about Sol 450 as testified with the ob-
516 served daily RMS variability shown in Fig. 2. In contrast to the 100 % data, the 30 %
517 autocorrelations (middle panel of Fig. 9) show stable signals for all Sols. This is because
518 the RMS variability threshold retained most of the seismic recordings with imprints of
519 climate and lander response variability. This result reinforces that a thorough data se-
520 lection is essential to warrant a minimum influence of aseismic signals which may bias
521 any autocorrelation result. In this respect, a data adaptive approach is better than sim-
522 ply selecting the data through LMST slots.

523 Amplitude spectra of the just discussed autocorrelations are shown in Fig. 10. The
524 amplitudes have been normalized at 6 Hz for visual purposes. Spectra for linearly stacked

autocorrelations are shown in Fig. S7 for comparison. It can be seen from both figures that the 100 % data set suffers at about 3.8 Hz and Sols larger than 500 from energy which appears at the corner of one of the applied band-rejection filter. These signals are more pronounced in the spectra of the linear stacks and are likely lander resonances in response to the bad weather. Amplitude spectra for linearly stacked autocorrelations computed for data without application of band-rejection filters (Fig. S8) reveal that the strongest amplitude changes indeed occur around 4 Hz. Broader band-rejection filters could attenuate these signals, but unnecessarily reduce the band width of the autocorrelations for data recorded before Sol 450.

As expected from the time-domain autocorrelations of Fig. 9 the corresponding spectra (Fig. 10) for the 30 % and 3 % data sets show less variability than for the 100 % data set. The strongest resonances have still been attenuated by the data selection approach and three band-rejection filters.

3.3 Stability of Reflection Signals over a Broad Frequency Band

Finally, seismic signals related to structural discontinuities are expected to be detected for a rather broad than narrow frequency band, i.e., possible frequency-dependent reflectivity generally varies smoothly. Besides, at small lag time, sidelobes of the zero-lag autocorrelation peak may affect early structural arrivals. The autocorrelation sidelobes are caused by the convolution of a delta pulse with the effective noise source time function (e.g., Ruigrok & Wapenaar, 2012) and are expected to change faster with frequency than a reflection pulse (see Fig 6 in Romero and Schimmel (2018)).

We therefore inspect the signal stability across different frequency bands. The first three panels in Fig. 11 show autocorrelations for the 30 % data set and different one-octave frequency bands. From top to bottom these are 1.5-3.0 Hz, 2.4-4.8 Hz, and 3.6-7.2 Hz. The effective band width, nevertheless, is smaller than an octave owing to the employed band-rejection filters. The tf-PWS of all autocorrelations is illustrated in red while the black lines show tf-PWSs of 10 % randomly drawn autocorrelations. The fourth panel shows the red color stacks of the top three panels plotted on top of each other. The lowermost panel contains the tf-PWS (red line) and linear stack (blue line) for the wide 1.2-8.9 Hz band.

555 It can be seen from this figure that some signals are persistent or in phase along
556 different frequency bands. Labels *a-f* mark some of these signals at lag times of 6.15 s,
557 6.85 s, 8.66 s, 10.6 s, 16.85 s, 23,8 s and 24.49 s. The dashed vertical lines have been added
558 to aid the visual inspection. Signals *a* and *c* are not seen for the lowest frequency band
559 and have therfore been marked at the base of the figure. None of the signals is being ob-
560 served for only one frequency band. E.g., signal *b* is in phase for the three frequency bands
561 while signal *d* is in phase for the two lower bands. The latter signal, however, appears
562 only in the middle frequency band as a clear signal. It is seen from the top two panels
563 that the amplitude maxima of signal *d* are not coherent while its minimum amplitude
564 occurs at the same lag time to build up the observed negative polarity when filtering over
565 a broader frequency band. Features *a* and *c* are in phase for the two higher frequency
566 bands. Signals *e* and *f* are also in phase and show also up as slightly distinctive features
567 in the 1.2-8.9 Hz correlations. *e* and *f* are less impulsive as *d*, i.e., in the 1.2-8.9 Hz cor-
568 relations they seem to be part of a larger waveform. *f* has been marked twice, at the first
569 minimum amplitude and 0.6 s later at the lowest minimum. We use only one letter, as
570 it is not clear if this are several signals which arrive at a similar time.

571 In any case, a visible interference of zero-lag sidelobes or their misinterpretation
572 is not expected owing to the stationarity of these signals over a broader frequency band
573 (Romero & Schimmel, 2018).

574 4 Discussion and Conclusion

575 The 10.6 s signal, if interpreted as P-wave reflection, would point to a discontinu-
576 ity at about 24 km depth considering expected Mars's crustal average seismic velocities
577 of 4.5 km/s. We observe the 10.6 s signal robustly on the 100 %, 30 %, and 3 % data
578 sets with amplitudes which can reach twice the amplitude of other nearby signals and
579 noise. The shape of the waveform resembles more a zero-phase wavelet with negative po-
580 larity than a wave train. The negative polarity is consistent with a reflection from a dis-
581 continuity with a seismic impedance increase while the wavelet shape can be explained
582 by a sharp discontinuity and lack of other interfering signals as reflections from strong
583 nearby discontinuities.

584 Indeed, the 10.6 s signal is a likely candidate for a reflection from the base of the
585 crust owing to its two-way travel time, negative polarity, strengths and robustness. En-

ergy at similar lag time has also been reported by Deng and Levander (2020) and Compaire et al. (2021) and interpreted as P-wave Moho reflection. The 10.6 s signal seems to be consistent with models obtained for receiver functions from marsquake data (Lognonne et al., 2020; Knapmeyer-Endrun et al., 2021). Knapmeyer-Endrun et al. (2021) present two model families resulting from receiver function inversions where the 10.6 s signal can be explained either as a P-wave Moho or internal crustal discontinuity reflection. More events from different distances or other geophysical constraints are needed to further limit the model space and to aid the identification of autocorrelation signals. We prefer the first model class which can explain the 10.6 s signal as a Moho reflection since we observe no other, at least equally-strong signal at a later time to explain a deeper Moho. This is based on the assumption that the crust-mantle boundary is stronger than a mid-crust discontinuity.

The seismic data acquisition conditions are harsh on Mars and there exist a wealth of different aseismic signals (Lognonne et al., 2020; Ceylan et al., 2021; Scholz et al., 2020; Stutzmann et al., 2021) which also may show up on autocorrelations. For instance, regularly repeated glitches (long duration pulses typically with frequencies below 1 Hz) and donks (short duration pulses typically at frequencies above 8 Hz, Ceylan et al. (2021)) with similar waveforms can manifest, depending on their abundance and coherence, at their repeat time. To attenuate the effect of glitches and donks, Deng and Levander (2020) determine the absolute mean amplitude in moving data windows of 100 s length which then are used to weight the seismic data (e.g., Bensen et al., 2007). Finally, they observe at 11.5 s a signal which appears robustly on their Martian day-time and night-time data and which they interpret as Moho P-wave reflection.

Conversely, Compaire et al. (2021) removed the tick noise and attenuated glitches using an algorithm by Scholz et al. (2020). High-frequency components of glitches remain in the data and donks are not removed by the algorithm. All in all, they use data recorded during the evening hours for which autocorrelations have the highest signal-to-noise ratio (SNR) and find a signal at 10.6 s and at 21 s which they interpret as reflections from crustal layers. They also show that glitches can not interfere with these reflections as they are separated by more than 30 s. The repeat time of donks, however, can be much smaller with maximum at about 10 s around 17-18 h LMST (Compaire et al., 2021). The authors argue that the observed stability of the 10.6 s signal with time

is not correlated with the overall distribution of donks which finally makes a donk origin of these reflections unlikely.

Here, we use a different data processing strategy. For instance, we do not de-glitch the records as high frequency glitch signatures seem to remain in the cleaned data. Further, we refrain from selecting data by choosing fixed time slots and design a data-adaptive approach to segment and pick seismic records based on their RMS amplitude variability. Our 30 % and 3 % data sets have been built selecting stretches of data with low RMS variability which also led to an attenuation of most lander resonances as seen from the spectra in Fig. 4. We further attenuated strong resonances using three band-rejection filters (1.9-2.5 Hz, 3.9-4.4 Hz and 6.8-7.2 Hz). These band-rejection filters were also chosen to remove jointly the strongest components of the tick noise.

Our data selection procedure is expected to avoid large amplitude donks and glitches and should be less affected by aseismic features which are related to bad weather lander and SEIS responses. However, small amplitude donks which do not affect the RMS variability can still be present in the selected data. Concerning the 10.6 s signal, a donk origin would require a sharp repeat time distribution of opposite polarity donks to explain the negative correlation peak and zero-phase shaped wavelet. A smooth distribution of donks is expected to either cancel out or to produce wave train shaped signals at lag times which correspond to their repeat times. In any case, at present we are not aware of hidden donks, but can also not rule out their existence.

Finally, large amplitude donks and glitches do not seem to have a dominant imprint on the autocorrelations as can be seen by comparing the autocorrelation sections for the 100 % and 30 % data sets (Fig. S9). This is because randomly distributed or incoherent glitches and donks, as well as other outlying signals and noise do not affect the amplitude unbiased PCC. Therefore, PCC is an adequate approach for data from unfavorable deployment and detection conditions as shown in Schimmel et al. (2018).

Another source of aseismic signals can be beats produced through interfering resonances (Kinsler et al., 1999). Beats can manifest in autocorrelations at lag time which equals the inverse of the difference in frequency of the two resonances and multiples. I.e., a 10.6 s autocorrelation signal could be due to a 0.094 Hz beat. The 10.6 s signal does not show a seasonal variability and would therefore require stable resonances throughout the analyzed data. We believe that the autocorrelations are not dominated by strong

beats as the main lander resonances have been attenuated through rejection filters and data selection on frequency band pass filtered data. Still, the presence of weak amplitude beats in the data can not be ruled out.

Further, the fast convergence of the 10.6 s signal which can be extracted with about 2 Sol of data is remarkable. There are no oceans on Mars, and all seismic energy is supposed to be released by weak magnitude events (the largest observed Marsquake has a magnitude of 4 Clinton et al. (2020)) and through atmospheric phenomena. In spite of the absence of strong sources, the low level of natural Martian noise (Lognonne et al., 2020) may favor the detectability of low-amplitude signals and therefore aid the convergence of the reflection response.

We focused the discussion on the 10.6 s signal, as it appears to be a likely candidate for a reflection from the base of the crust. Nevertheless, the obtained reflection response contains also other signals which likely are reflections and multiples from other discontinuities. Some of them have been marked in Fig. 11 (labels a to g) owing to their appearance over a broader frequency band. The lag times are 6.15 s, 6.85 s, 8.66 s, 10.6 s, 16.85 s, 23,8 s and 24.49 s. Specially for the earlier arrivals the frequency stability is an important criteria to avoid misinterpretations or interferences with zero-lag autocorrelation sidelobes as shown in Romero and Schimmel (2018).

Concerning inter-crustal discontinuities, receiver function studies (Lognonne et al., 2020; Knapmeyer-Endrun et al., 2021) provide stable estimates on a discontinuity at about 8-11 km depths with S-wave velocities Vs of about 1.7-2.1 km/s. If we assume a mean discontinuity depth of 9.5 km and that our signal *a* at 6.15 s lag time (Fig. 11) is a reflection from this discontinuity then the expected P-wave velocity Vp and Vp/Vs ratio for the 9.5 km layer are 3.1 km/s and 1.48-1.82 considering the Vs range. Under same assumptions, the close by signal *b* at 6.85 s lag time, would lead to Vp=2.8km/s and Vp/Vs ratio of about 1.33-1.65. We believe that signal *a* provides a more realistic Vp/Vs ratio than signal *b* and could therefore likely be the reflection from the base of the upper crust layer. For comparison, the Vp/Vs ratio for the very shallow regolith layer has been estimated to 1.67 (Lognonne et al., 2020).

Finally, the estimated lag time dependent convergence of the autocorrelations shows that more data are needed to build the reflection response at later lag times. This makes sense as waves propagating over a longer time interval are more prone to scattering and

attenuation which needs to be compensated by stacking more autocorrelations. In any case, we see no reason for aseismic signals to show a systematic lag time dependent convergence.

Data Availability Statement

We acknowledge NASA, CNES, their partner agencies and Institutions (UKSA, SSO, DLR, JPL, IPGP-CNRS, ETHZ, IC, MPS-MPG) and the flight operations team at JPL, SISMOC, MSDS, IRIS-DMC and PDS for providing SEED SEIS data. Seis raw data are available at https://doi.org/10.18715/SEIS.INSIGHT.XB_2016. Processed data corresponding to this article are in Schimmel et al. (2021). PCC and tf-PWS software are available at http://diapiro.ictja.csic.es/gt/mschi/SCIENCE/tseries.html#software. Figures have been done with SAC and GMT (Goldstein & Snoke, 2005; Wessel et al., 2019).

Acknowledgments

French authors are supported by ANR MAGIS (ANR-19-CE31-0008-08) and by CNES for SEIS science support. MS thanks SANIMS (RTI2018-095594-B-I00) and invitations for research stays at IPGP.

References

Anderson, D. L., Miller, W., Latham, G., Nakamura, Y., Toksöz, M., Dainty, A., ... Knight, T. (1977). Seismology on mars. *Journal of Geophysical Research*, *82*(28), 4524-4546.

Becker, G., & Knapmeyer-Endrun, B. (2018). Crustal thickness across the trans-european suture zone from ambient noise autocorrelations. *Geophysical Journal International*, *212*(2), 1237–1254.

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

Berbellini, A., Schimmel, M., Ferreira, A. M., & Morelli, A. (2019). Constraining s-wave velocity using rayleigh wave ellipticity from polarization analysis of seismic noise. *Geophysical Journal International*, *216*(3), 1817–1830. doi:

10.1093/gji/ggy512

Buffoni, C., Schimmel, M., Sabbione, N. C., Rosa, M. L., & Connon, G. (2019). Crustal structure beneath tierra del fuego, argentina, inferred from seismic p-wave receiver functions and ambient noise autocorrelations. *Tectonophysics*, *751*, 41–53. doi: 10.1016/j.tecto.2018.12.013

Ceylan, S., Clinton, J., Giardini, D., Bose, M., & et al. (2021). Companion guide to the marsquake catalog from insight, sols 0-478: data content and non-seismic events. *Physics of the Earth and Planetary Interiors*, 00-00. doi: 10.1016/j.pepi.2020.106597

Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, *33*(2), 264–269.

Clinton, J. F., Ceylan, S., van Driel, M., Giardini, D., Stähler, S. C., Böse, M., ... others (2020). The marsquake catalogue from insight, sols 0-478. *Physics of the Earth and Planetary Interiors*. doi: 10.31219/osf.io/ws967

Compaire, N., Margerin, L., Garcia, R., Pinot, B., et al. (2021). Autocorrelation of the ground vibrations recorded by the seis-insight seismometer on mars. *J. Geophys. Res.: Planets*, *JGRE2160*. doi: 10.1029/2020JE006498

Dahmen, N., Zenhäusern, G., Clinton, J., Giardini, D., Stähler, S., & Ceylan, S. e. a. (2021). Resonances and lander modes observed by insight on mars (1-9 hz). *BSSA*, submitted.

Deng, S., & Levander, A. (2020). Autocorrelation reflectivity of mars. *Geophysical Research Letters*, *47*(16), e2020GL089630.

Derode, A., Larose, E., Tanter, M., de Rosny, J., Tourin, A., Campillo, M., & Fink, M. (2003). Recovering the Green's function from field-field correlations in an open scattering medium (L). *The Journal of the Acoustical Society of America*, *113*(6), 2973. doi: 10.1121/1.1570436

Drilleau, M., Beucler, É., Lognonné, P., Panning, M. P., Knapmeyer-Endrun, B., Banerdt, W. B., ... others (2020). Mss/1: Single-station and single-event marsquake inversion. *Earth and Space Science*, e2020EA001118.

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 54–75.

Farra, V., & Vinnik, L. (2000). Upper mantle stratification by p and s receiver func-

tions. *Geophysical Journal International*, *141*(3), 699–712.

Galetti, E., & Curtis, A. (2012). Generalised receiver functions and seismic interferometry. *Tectonophysics*, *532*, 1–26.

Giardini, D., Lognonne, P., Banerdt, W., Pike, W., Christenseni, U., & al. (2020). The seismicity on mars. *Nature Geoscience*, *13*, 205-212. doi: 10.1038/s41561-020-0539-8

Goldstein, P., & Snoke, A. (2005). Sac availability for the iris community. *Incorporated Research Institutions for Seismology Newsletter*, *7*(UCRL-JRNL-211140).

Golombek, M., Warner, N., Grant, J., & al. (2020). Geology of the insight landing site, mars. *Nature Geoscience*, *11*, 1014. doi: 10.1038/s41467-020-14679-1

Gorbatov, A., Saygin, E., & Kennett, B. (2013). Crustal properties from seismic station autocorrelograms. *Geophysical Journal International*, *192*(2), 861–870.

Hobiger, M., Cornou, C., Wathelet, M., Giulio, G. D., Knapmeyer-Endrun, B., Renalier, F., ... others (2013). Ground structure imaging by inversions of rayleigh wave ellipticity: sensitivity analysis and application to european strong-motion sites. *Geophysical Journal International*, *192*(1), 207–229.

Kenda, B., Drilleau, M., Garcia, R., Kawamura, T., Murdoch, N., Compaire, N., ... Widmer-Schnidrig, R. (2020). Subsurface structure at the insight landing site from compliance measurements by seismic and meteorological experiments. *J. Geophys. Res.* doi: 10.1029/2020JE006387

Kennett, B., Saygin, E., & Salmon, M. (2015). Stacking autocorrelograms to map moho depth with high spatial resolution in southeastern australia. *Geophysical Research Letters*, *42*(18), 7490–7497.

Kinsler, L. E., Frey, A. R., Coppens, A. B., & Sanders, J. V. (1999). *Fundamentals of acoustics* (4th ed.). John Wiley & Sons.

Knapmeyer-Endrun, B., Panning, M., & et al. (2021). Crustal thickness and layering of mars from insight seismic data. *Science, submitted*.

Langston, C. A. (1979). Structure under mount rainier, washington, inferred from teleseismic body waves. *Journal of Geophysical Research: Solid Earth*, *84*(B9), 4749–4762. doi: 10.1029/JB084iB09p04749

Lobkis, O., & Weaver, R. (2001). On the emergence of the Green's function in the correlations of a diffuse field. *J. acoust. Soc. Am.*, *110*, 3011-3017.

Lognonne, P., Banerdt, W., W.T.Pike, Giardini, D., Christensen, U., & et. al. (2020). Constraints on the shallow elastic and anelastic structure of mars from insight seismic data. *Nature Geoscience*, *13*, 213-220. doi: 10.1038/s41561-020-0536-y

Lognonne, P., Banerdt, W. B., Giardini, D., Pike, W. T., Christensen, U., Laudet, P., ... et al. (2019). Seis: Insight's seismic experiment for internal structure of mars. *Space Science Reviews*, *215*(1). doi: 10.1007/s11214-018-0574-6

Lognonné, P., & Johnson, C. (2007). Planetary seismology. *Treatise on Geophysics*, *10*, 69–122.

Medeiros, W. E., Schimmel, M., & do Nascimento, A. F. (2015). How much averaging is necessary to cancel out cross-terms in noise correlation studies? *Geophysical Journal International*, *203*(2), 1096–1100. doi: 10.1093/gji/ggv336

Nakamura, Y. (1989). A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. *Railway Technical Research Institute, Quarterly Reports*, *30*(1).

Oren, C., & Nowack, R. L. (2016). Seismic body-wave interferometry using noise auto-correlations for crustal structure. *Geophysical Journal International*, ggw394.

Panning, M. P., Beucler, É., Drilleau, M., Mocquet, A., Lognonné, P., & Banerdt, W. B. (2015). Verifying single-station seismic approaches using earth-based data: Preparation for data return from the insight mission to mars. *Icarus*, *248*, 230–242.

Pham, T.-S., & Tkalčić, H. (2017). On the feasibility and use of teleseismic p wave coda autocorrelation for mapping shallow seismic discontinuities. *Journal of Geophysical Research: Solid Earth*, *122*(5), 3776–3791.

Phinney, R. A. (1964). Structure of the earth's crust from spectral behavior of long-period body waves. *Journal of Geophysical Research*, *69*(14), 2997–3017.

Romero, P., & Schimmel, M. (2018, Jun). Mapping the basement of the ebro basin in spain with seismic ambient noise autocorrelations. *Journal of Geophysical Research: Solid Earth*, *123*(6), 5052–5067. doi: 10.1029/2018jb015498

Ruigrok, E., Campman, X., & Wapenaar, K. (2011). Extraction of p-wave reflections from microseisms. *Comptes Rendus Geoscience*, *343*(8-9), 512–525.

Ruigrok, E., & Wapenaar, K. (2012). Global-phase seismic interferometry unveils p-

wave reflectivity below the himalayas and tibet. *Geophysical Research Letters*, *39*(11).

Saygin, E., Cummins, P. R., & Lumley, D. (2017). Retrieval of the p wave reflectivity response from autocorrelation of seismic noise: Jakarta basin, indonesia. *Geophysical Research Letters*, *44*(2), 792–799.

Schimmel, M. (1999). Phase cross-correlations: Design, comparisons, and applications. *Bulletin of the Seismological Society of America*, *89*(5), 1366–1378.

Schimmel, M., & Gallart, J. (2007). Frequency-dependent phase coherence for noise suppression in seismic array data. *Journal of Geophysical Research: Solid Earth*, *112*(B4), B04303, 1–14. doi: 10.1029/2006JB004680

Schimmel, M., Stutzmann, E., & Gallart, J. (2011, January). Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale. *Geophysical Journal International*, *184*(1), 494–506. doi: 10.1111/j.1365-246X.2010.04861.x

Schimmel, M., Stutzmann, E., & Lognonné, P. (2021). Data from the article: Seismic noise autocorrelations on mars.
doi: IPGP.2021.km1vsxsd

Schimmel, M., Stutzmann, E., & Ventosa, S. (2018). Low-frequency ambient noise autocorrelations: Waveforms and normal modes. *Seismological Research Letters*, *89*(4), 1488–1496. doi: 10.1785/0220180027

Scholz, J.-R., Widmer-Schnidrig, R., Davis, P., Lognonné, P., Pinot, B., Garcia, R. F., ... others (2020). Detection, analysis, and removal of glitches from insight's seismic data from mars. *Earth and Space Science*, *7*(11), e2020EA001317. doi: 10.1029/2020EA001317

Snieder, R. (2004). Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, *69*(4). doi: 10.1103/PhysRevE.69.046610

Spiga, A., Murdoch, N., Lorenz, R., Forget, F., Newman, C., Rodriguez, S., ... others (2020). A study of daytime convective vortices and turbulence in the martian planetary boundary layer based on half-a-year of insight atmospheric measurements and large-eddy simulations. *Journal of Geophysical Research: Planets*, e2020JE006511.

Stockwell, R. G., Mansinha, L., & Lowe, R. P. (1996). Localization of the complex spectrum: the S transform. *IEEE Trans. Signal Process.*, *44*(4), 998–1001.

Stutzmann, E., Schimmel, M., Lognonné, P., Horleston, A., Ceylan, S., van Driel, M., ... others (2021). The polarization of ambient noise on mars. *Journal of Geophysical Research: Planets*, *126*(1), e2020JE006545. doi: 10.1029/2020JE006545

Suemoto, Y., Ikeda, T., & Tsuji, T. (2020). Temporal variation and frequency dependence of seismic ambient noise on mars from polarization analysis. *Geophysical Research Letters*, *47*(13), e2020GL087123.

Tauzin, B., Pham, T.-S., & Tkalčić, H. (2019). Receiver functions from seismic interferometry: a practical guide. *Geophysical Journal International*, *217*(1), 1–24.

Taylor, G., Rost, S., & Houseman, G. (2016). Crustal imaging across the north anatolian fault zone from the autocorrelation of ambient seismic noise. *Geophysical Research Letters*, *43*(6), 2502–2509.

Tibuleac, I. M., & von Seggern, D. (2012). Crust-mantle boundary reflectors in nevada from ambient seismic noise autocorrelations. *Geophysical Journal International*, *189*(1), 493–500.

van Driel, M., Ceylan, S., Clinton, J. F., Giardini, D., Alemany, H., Allam, A., ... others (2019). Preparing for insight: Evaluation of the blind test for martian seismicity. *Seismological Research Letters*, *90*(4), 1518–1534. doi: 10.1785/0220180379

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

Ventosa, S., Schimmel, M., & Stutzmann, E. (2019). Towards the processing of large data volumes with phase cross-correlation. *Seismological Research Letters*, *90*(4), 1663–1669. doi: 10.1785/0220190022

Ventosa, S., Simon, C., Schimmel, M., Dañobeitia, J., & Mànuel, A. (2008). S-transform from a wavelets point of view. *IEEE Trans. Signal Process.*, *56*, 2771–2780. doi: 10.1109/TSP.2008.929670

Vinnik, L. (1977). Detection of waves converted from p to sv in the mantle. *Physics of the Earth and Planetary Interiors*, *15*(1), 39–45.

Wapenaar, K. (2004). Retrieving the elastodynamic green's function of an arbitrary inhomogeneous medium by cross correlation. *Physical review letters*, *93*(25), 254301. doi: 10.1103/PhysRevLett.93.254301

Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5556–5564.

Yano, T., Tanimoto, T., & Rivera, L. (2009). The zh ratio method for long-period seismic data: inversion for s-wave velocity structure. *Geophysical Journal International*, *179*(1), 413–424. doi: 10.1111/j.1365-246X.2009.04293.x

Yuan, X., Kind, R., Li, X., & Wang, R. (2006). The s receiver functions: synthetics and data example. *Geophysical Journal International*, *165*(2), 555–564.

May 5, 2018	InSight launched from Vandenberg Air Force Base (California)
November 26, 2018 (SOL 0)	Landing near Elysium Planitia at $4.5024°N, 135.6234°E$
December 19, 2018 (SOL 22)	Deployment of SEIS on Martian ground
February 2, 2019 (SOL 66)	SEIS is covered by Wind and Thermal Shield (WTS)
May 19, 2019 (SOL 169)	SEIS heating system has been switched on
August 28, 2019 (SOL 267)	Start of Mars Solar conjunction, i.e., no data transmission
September, 19, 2019 (SOL 289)	End of Mars Solar conjunction, i.e., start of data transmission

886 **Table 1.** Summary of key dates and works related to SEIS deployment and operation on Mars.

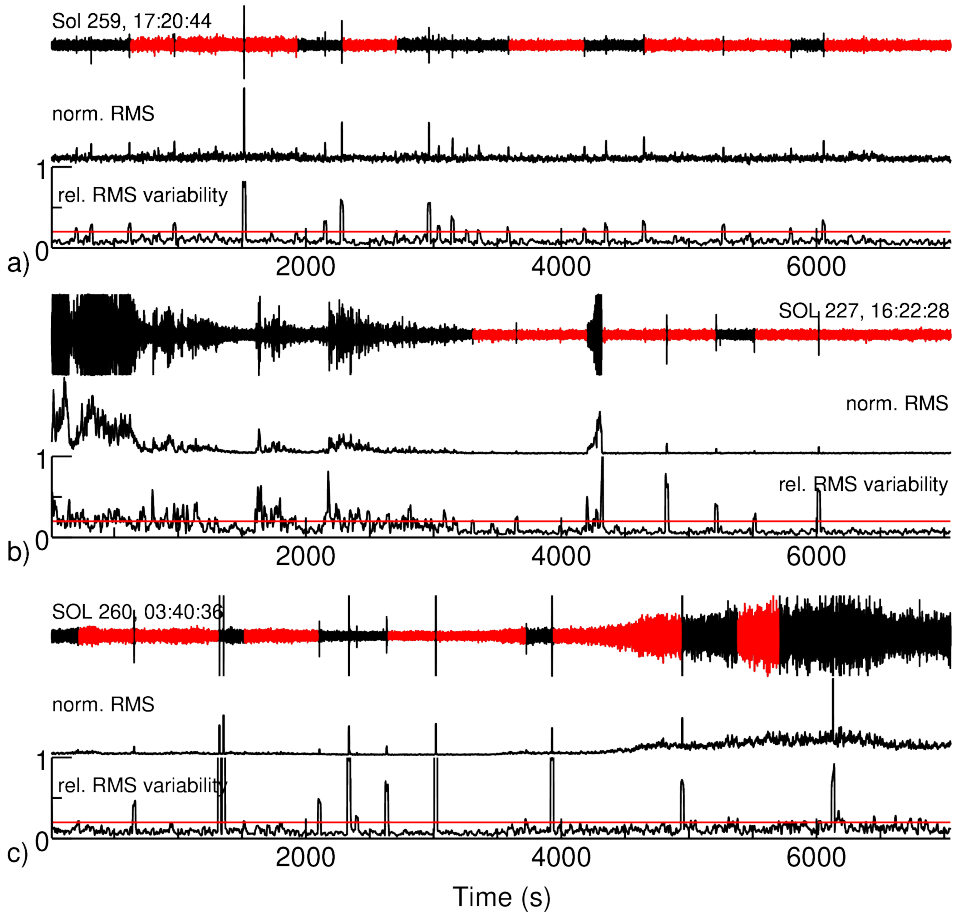


Figure 1. Three data selection and segmentation examples. a) The top trace shows the vertical component record with the selected segments in red. The two bottom traces are the normalized RMS and the relative RMS variability. The horizontal red line marks the threshold RMS variability used for the segment selection. The begin time (LMST) of the time series is written to the top left. b) and c) show the same as a), but for different time records. More examples are shown in Fig. S1 of the Supplementary Section.

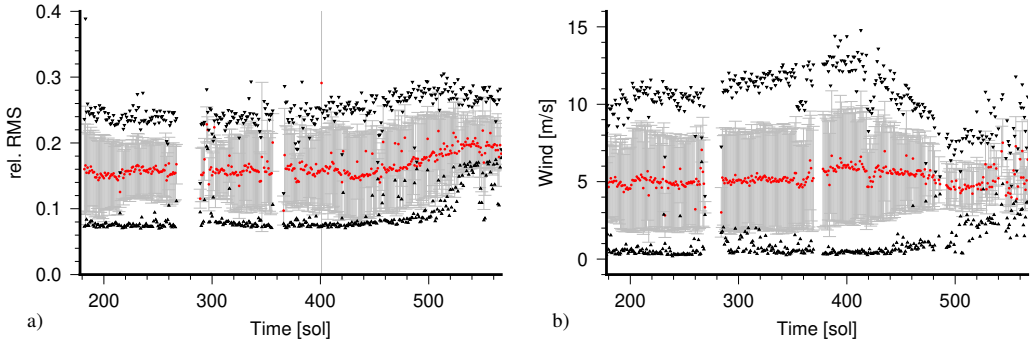


Figure 2. a) Evolution of seismic signal RMS variability as function of Sol measured on Z-component filtered from 1.2-9.5 Hz. Red dots and gray error bars mark the mean and one standard deviation uncertainty per Sol. Black triangles point to the minimum and maximum RMS variability encountered per Sol. b) Wind speed as function of Sol. In analogy to the seismic signal RMS variability plot, red dots and gray bars mark the mean and one standard deviation uncertainty while black triangles point to the minimum and maximum RMS variability encountered per Sol.

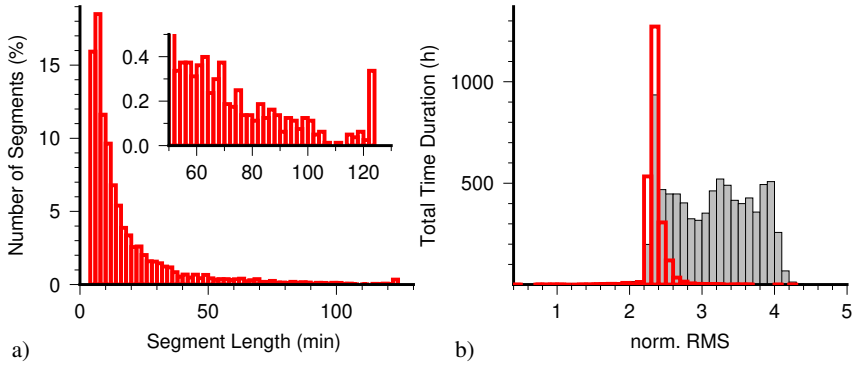
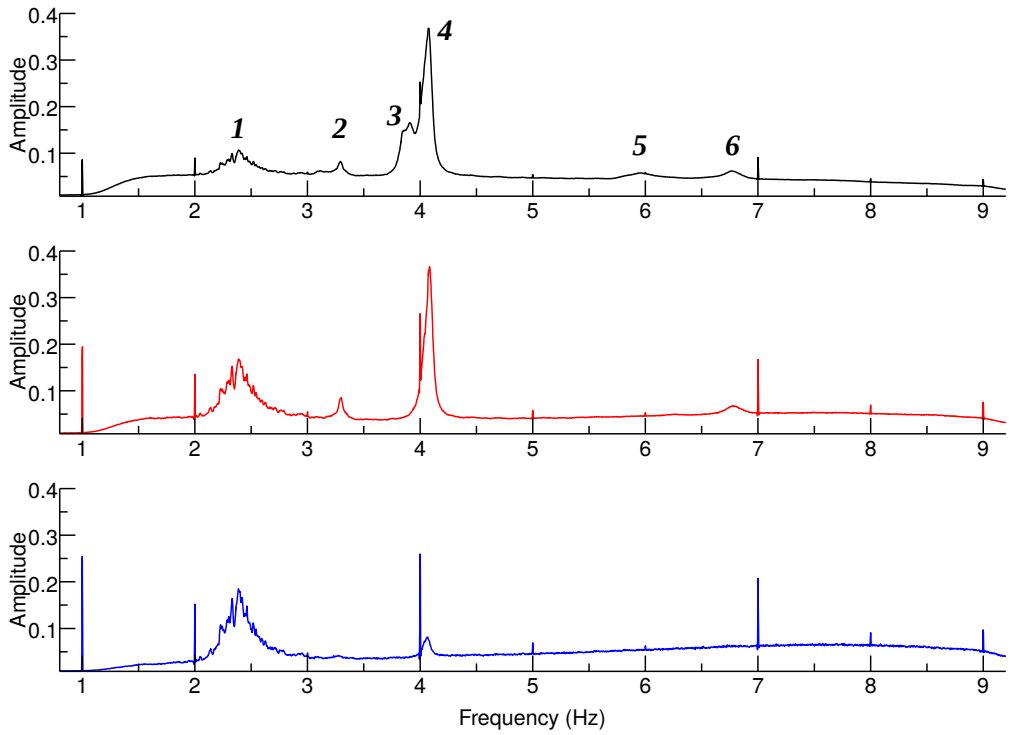
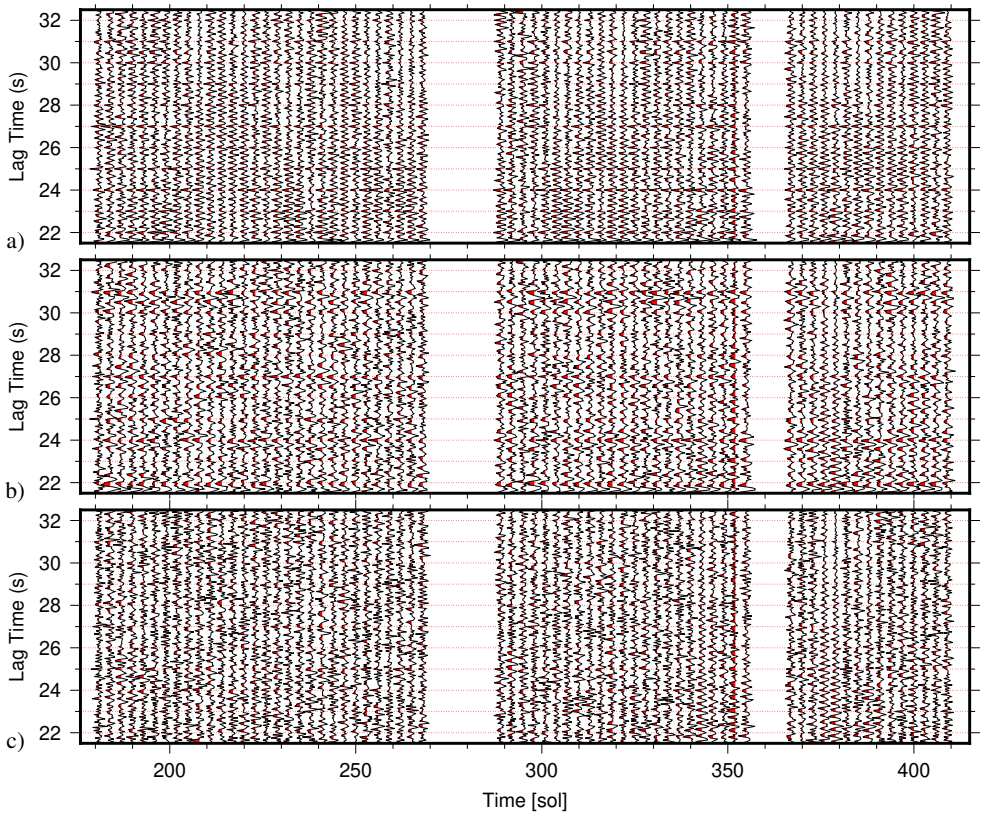


Figure 3. a) Histogram of segment length distribution for the 30% subsidiary data set. The inset is a zoom for segment length between 30 and 130 min. Vertical axis contains the number of segments in percentage while horizontal axis shows the segment length in minutes. b) Histogram of normalized RMS distribution for the 100 % (gray) and 30 % (red) data sets. Vertical axis is cumulative time in hours and horizontal axis is the base 10 logarithm of the normalized RMS distribution. The increase of cumulative time for the subsidiary data set with respect to the total data bases is due to the RMS variability and data segmentation.



907 **Figure 4.** Normalized Fourier amplitude spectra of autocorrelation stacks based on PCC and
908 linear stacking. Black, red and blue lines show the spectra for the total data set, the 30% and 3%
909 subsidiary data sets, respectively. Italic numbers 1-6 mark features discussed in the main text.



910 **Figure 5.** a) Vertical-component autocorrelation stacks for sliding 3-Sol data windows. The
911 frequency band is 1.2-8.9 Hz, data windows do not overlap, and positive amplitudes are in red.
912 Shown are linear stacks of phase autocorrelations for the entire data set at lag time without re-
913 flections from shallow discontinuities. The tick noise are the positive amplitude signals at every
914 full second. b) Same as a), but two band-rejection filters, 3.9-4.4 Hz and 6.8-7.2 Hz have been
915 applied before computation of the phase autocorrelations. c) Same as b), but the data has been
916 filtered with a third band rejection filter 1.9-2.5 Hz. See Fig. S3 for another example employing
917 the 3 % data set and tf-PWS.

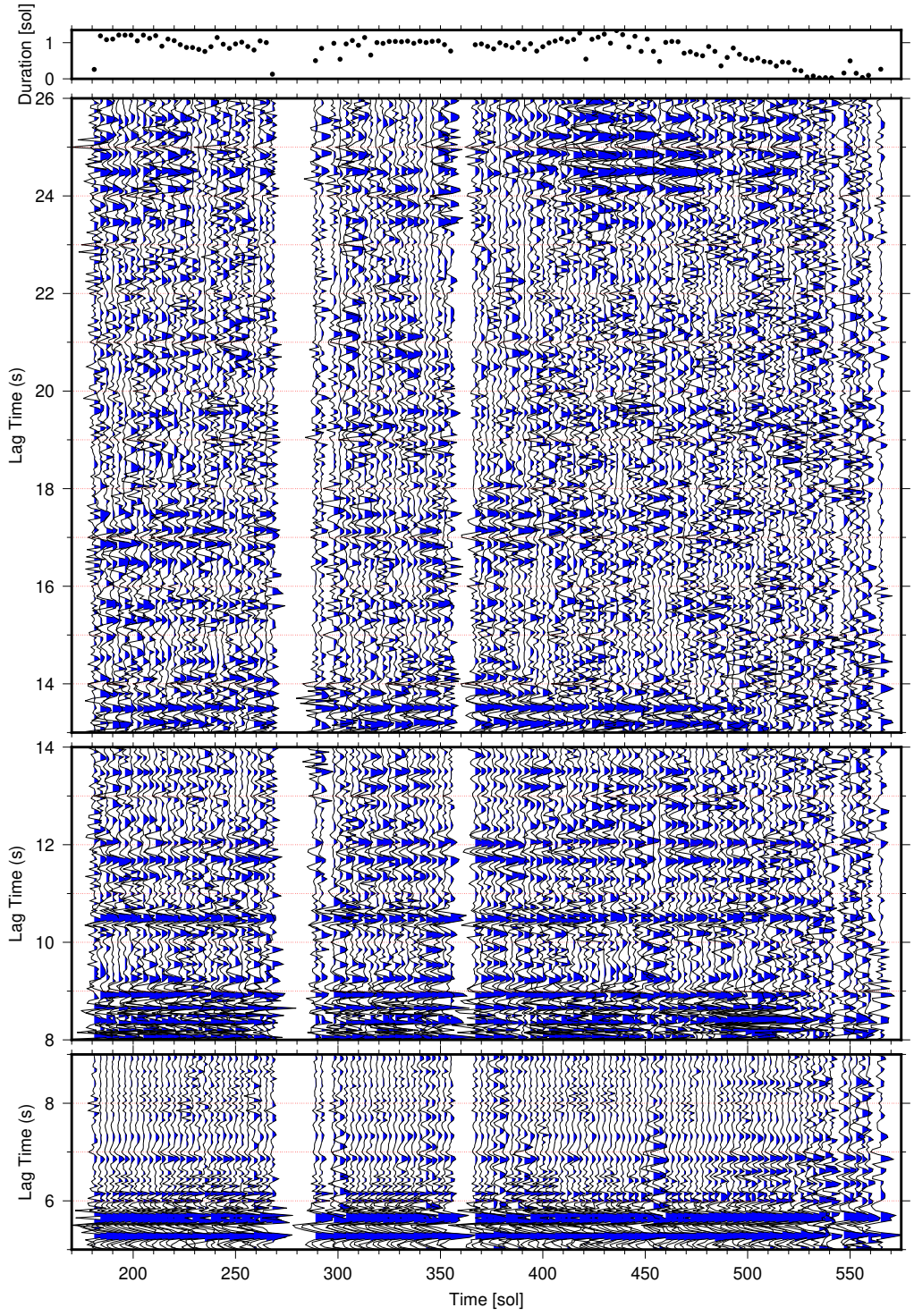


Figure 6. Vertical-component autocorrelation stacks for sliding 3-Sol data windows. The frequency band is 1.2-8.9 Hz and data windows do not overlap. Shown are tf-PWSs of phase autocorrelations computed for the 30 % data set. Blue marks negative amplitudes. The three lag-time windows have been used to improve the visibility through independent amplitude normalization. Top panel shows the total duration of the selected data used to compute autocorrelations within each of the 3-Sol data windows. Linear stacks for the same data are shown in Fig. S4.

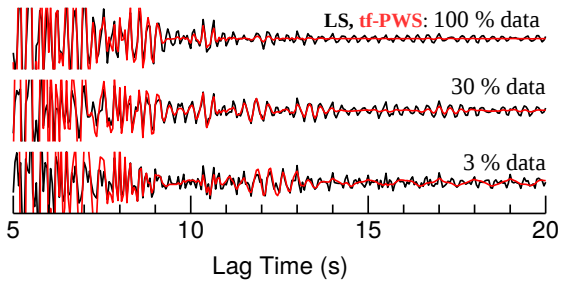


Figure 7. This figure shows the stacked autocorrelations for the 100 % data (top panel), 30 % data (middle panel), and 3 % data (bottom panel). The linear stacks are plotted in black and the corresponding tf-PWS in red. The amplitudes of each stack have been normalized for visual purposes. The 10.6 s signal appears with in impulsive shape for the three data sets and stands out from its immediate surroundings.

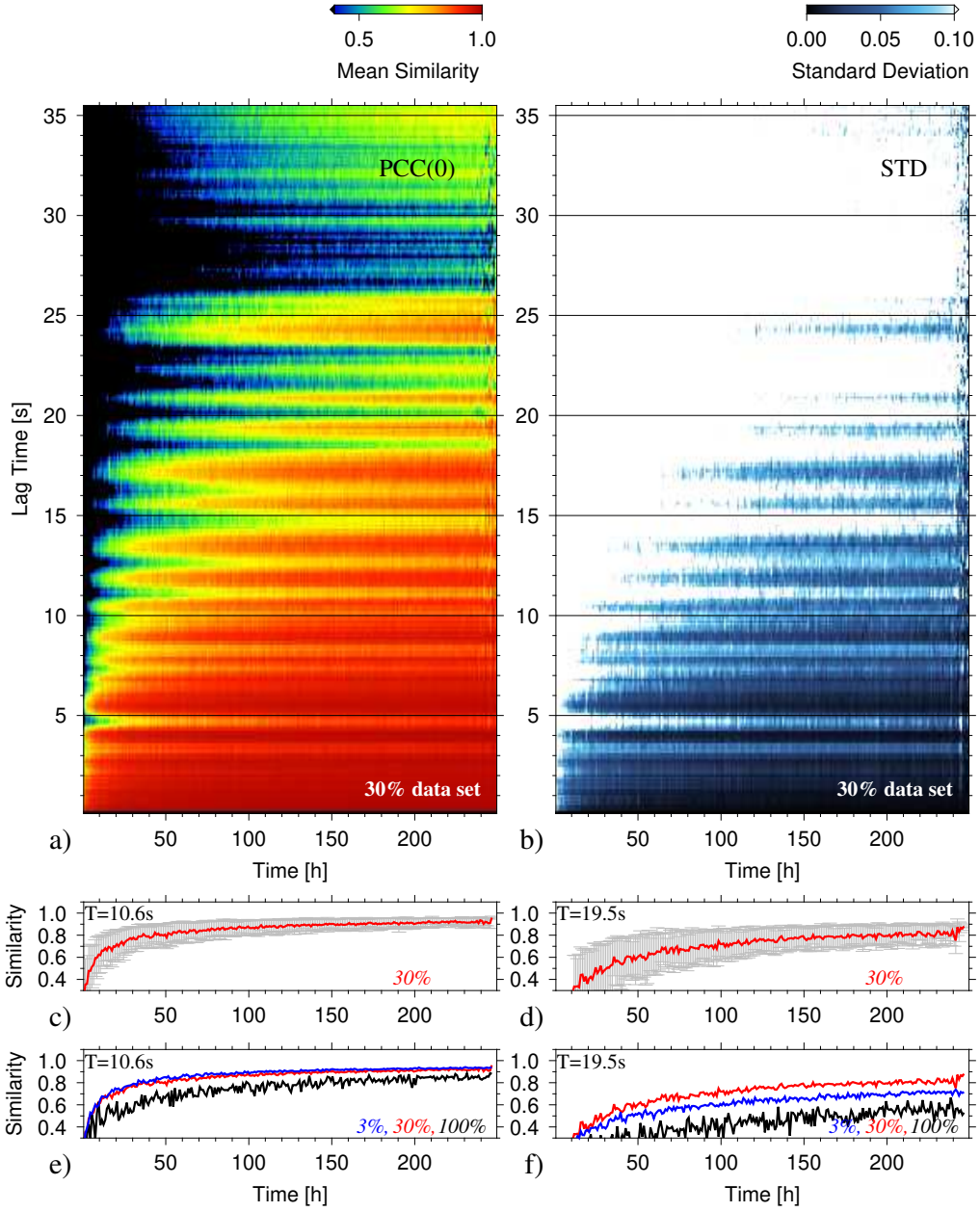


Figure 8. a) Autocorrelation waveform convergence as function of lag time and total length of data using the 30% data set. Colors are used to contour the mean similarity of randomly stacked autocorrelations with the reference trace. The reference trace is the linear stack of all autocorrelations from the 30% data set (black trace, middle panel of Fig. 7). The total data length is 2350.2 h. b) Standard deviation of the similarity shown in a). c) Similarity (red) and standard deviation (gray) of the 10.6 s signal for the 30 % data set. d) Same as c) but for a signal at 19.5 s lag time. e) Same as c) but showing also the results for the 3 % (blue) and 100% data sets (black). The reference traces are based on a total data length of 201.2 h and 8355.6 h for the 3 % and 100 % data sets, respectively. f) Same as e) but for the 19.5 s signal.

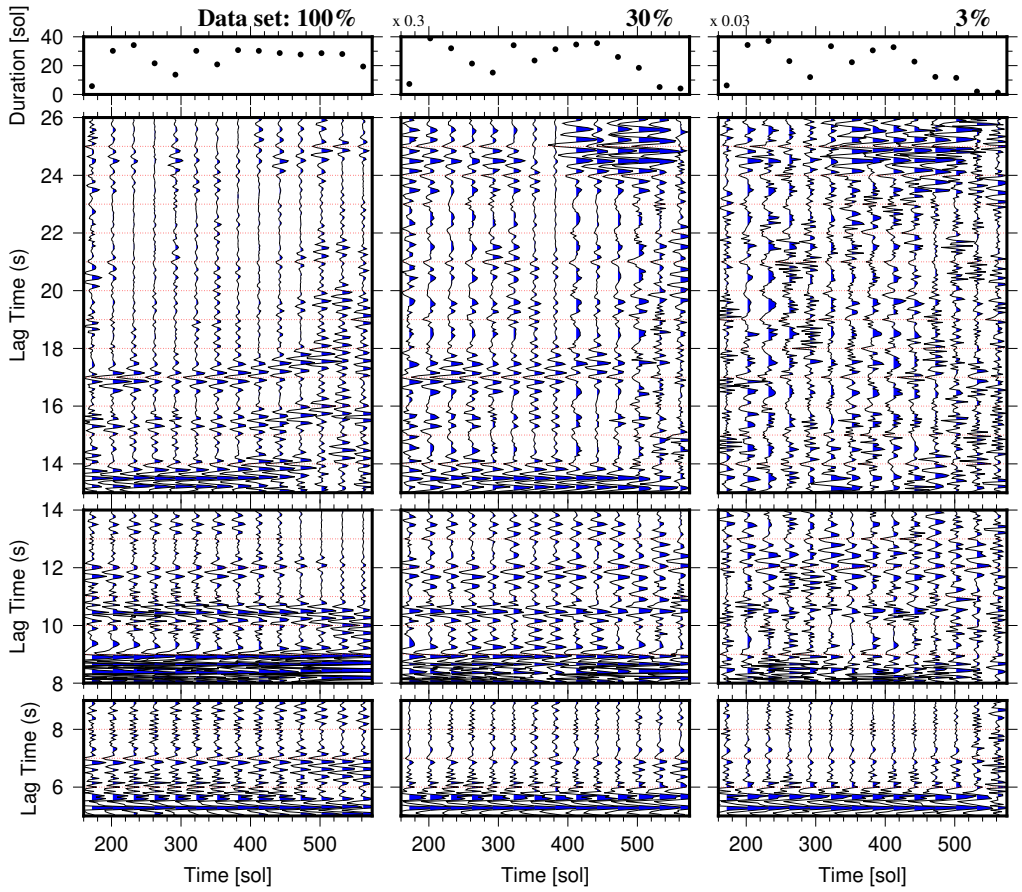
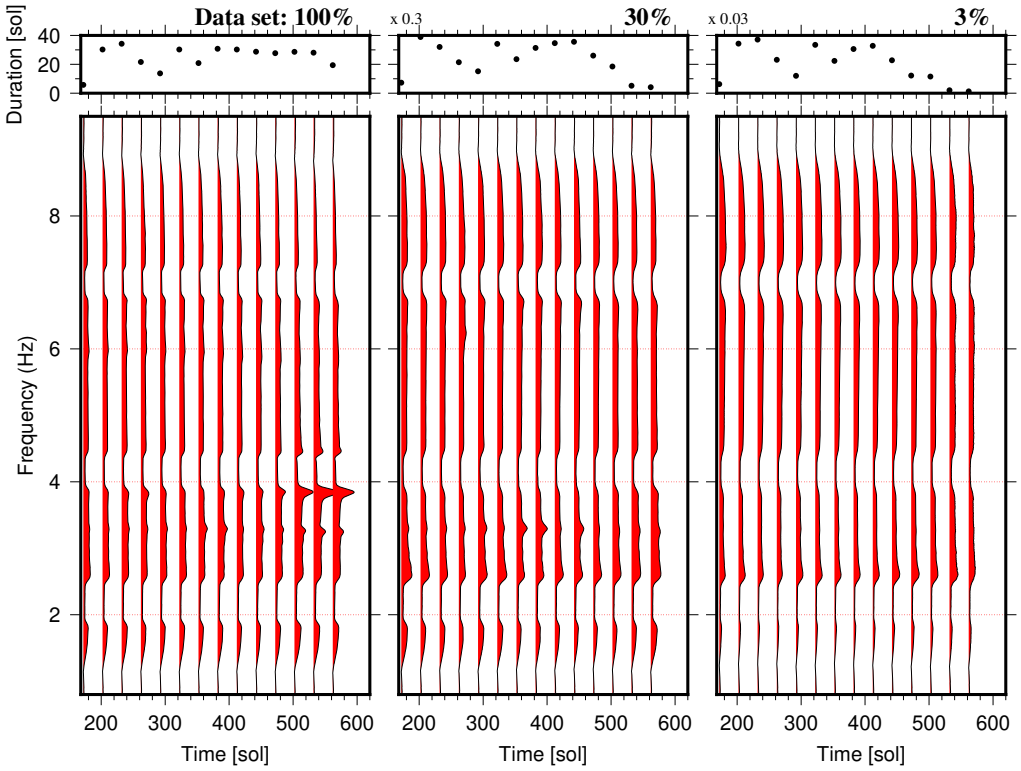
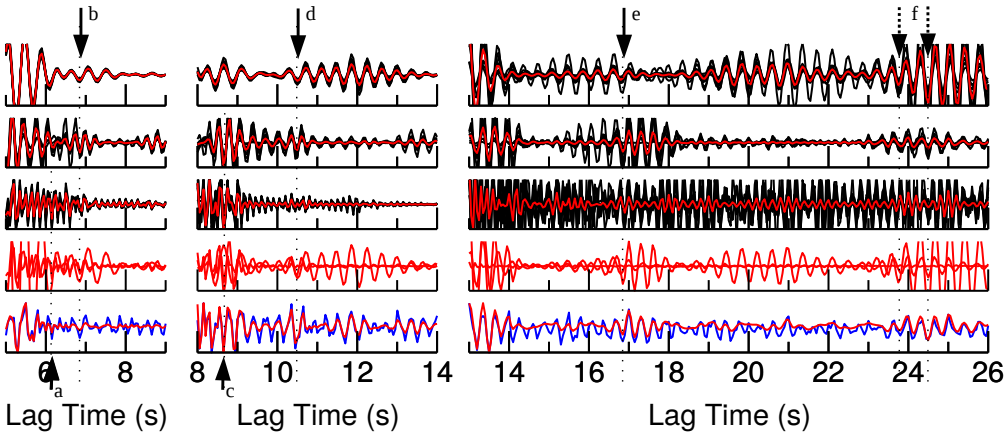


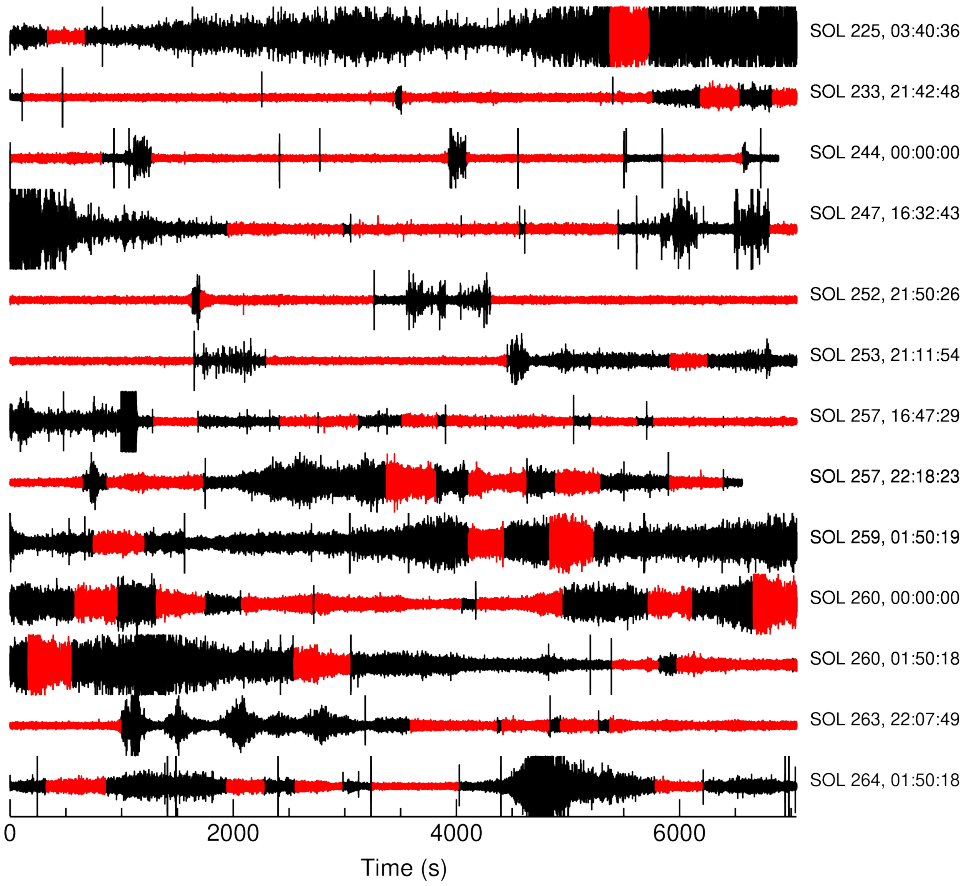
Figure 9. Vertical-component noise autocorrelation stacks for sliding 30-Sol data windows. The frequency band is 1.2-8.9 Hz and data windows do not overlap. Shown are time-frequency phase weighted stacks of phase autocorrelations. Blue marks negative amplitudes. The three lag-time windows have been used to improve the visibility through independent amplitude normalization. The top panel shows the total duration of the selected data used to compute autocorrelations within each of the 30-Sol data windows. The axis for the duration of the 30 % and 3 % data set are scaled by the number to the top left. Linear stacks for the same data are shown in Fig. S6.



946 **Figure 10.** Amplitude spectra of tf-PWS autocorrelations shown in Fig. 9 as function of Sol
947 for the 100 % (left panel), 30 % (middle panel), and 3 % (right panel) data sets. Each amplitude
948 spectrum is placed at its window center time and has been normalized at 6 Hz. Spectra of linear
949 stacks are shown in Fig S7.



950 **Figure 11.** Vertical-component noise autocorrelation stacks for the 30 % data set. tf-PWSs
951 are shown in red and black. Red traces are based on the entire 30 % data set while black traces
952 mark stacks for only 10% of the data. Lag-time windows correspond to those of Fig. 6 and have
953 been chosen to improve signal visibility. The first three panels are for 1.5-3.0 Hz, 2.4-4.8 Hz, and
954 3.6-7.2 Hz band-passed filtered noise. The forth panel compares the stacks from the top three
955 panels. The lowermost panel shows the linear stack (blue line) and tf-PWS (red line) for the
956 frequency band 1.2-8.9 Hz. Arrows mark signals mentioned in the text.



957 **Figure S1.** Data selection and segmentation examples. The selected segments are marked in
958 red. The start time of each trace is given in LMST to the right of each record.

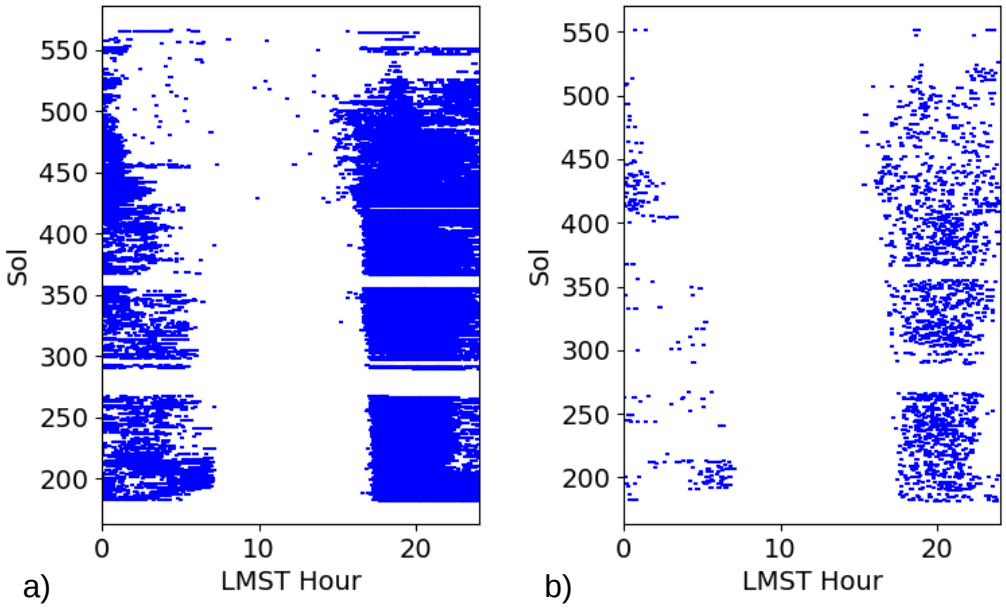


Figure S2. Selected data windows as function of Sol and LMST for the 30% (a) and 3% (b) subsidiary data sets. Very short time windows are not well resolved. Nevertheless, this figure shows that most of the selected time segments with low RMS variability are from the evening when wind activity is low.

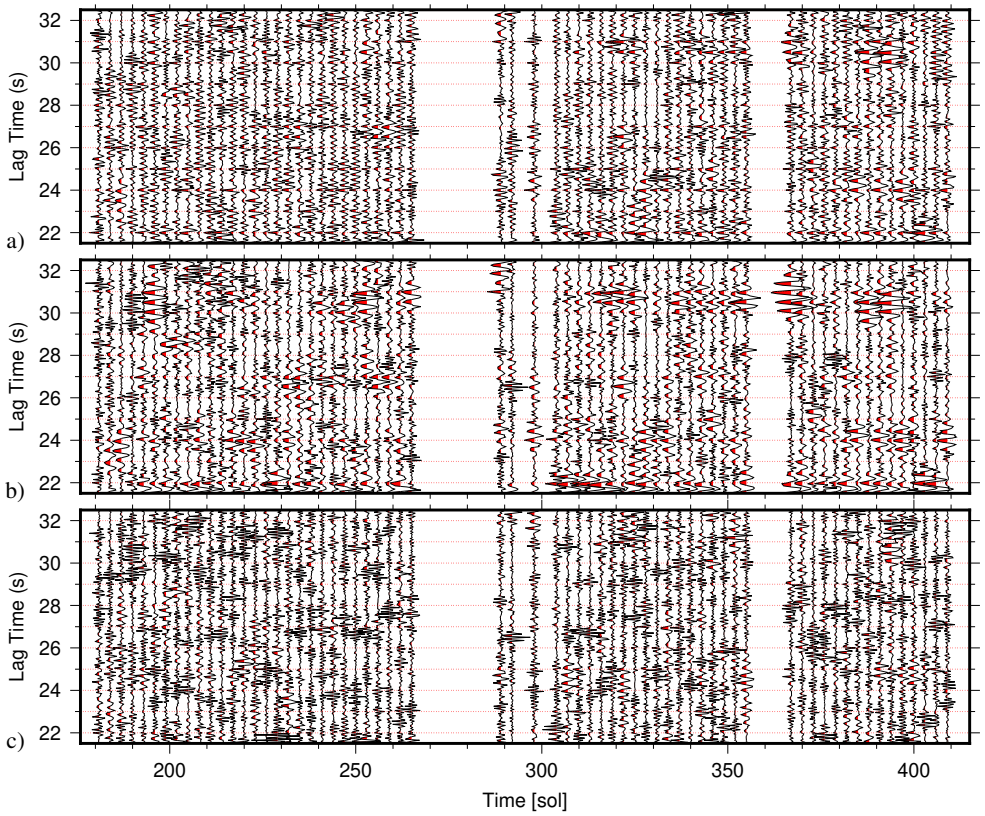


Figure S3. Same as Fig. 5, but using tf-PWS and the 3 % data set rather than the linear stack and the 100 % data set: a) Vertical-component noise autocorrelation stacks (tf-PWS) for sliding 3-Sol data windows. The frequency band is 1.2-8.9 Hz, data windows do not overlap, and positive amplitudes are in red. Phase autocorrelations are for a subsidiary data set of about 3% of the total data volume, i.e., containing the traces with the lowest RMS amplitude variability. The lag time window is large to avoid reflections from shallow discontinuities. The tick noise are the positive amplitude signals at every full second. b) Same as a), but two band-rejection filters, 3.9-4.4 Hz and 6.8-7.2 Hz have been applied before computation of the phase autocorrelations. c) Same as b), but the data has been filtered with a third band rejection filter 1.9-2.5 Hz.

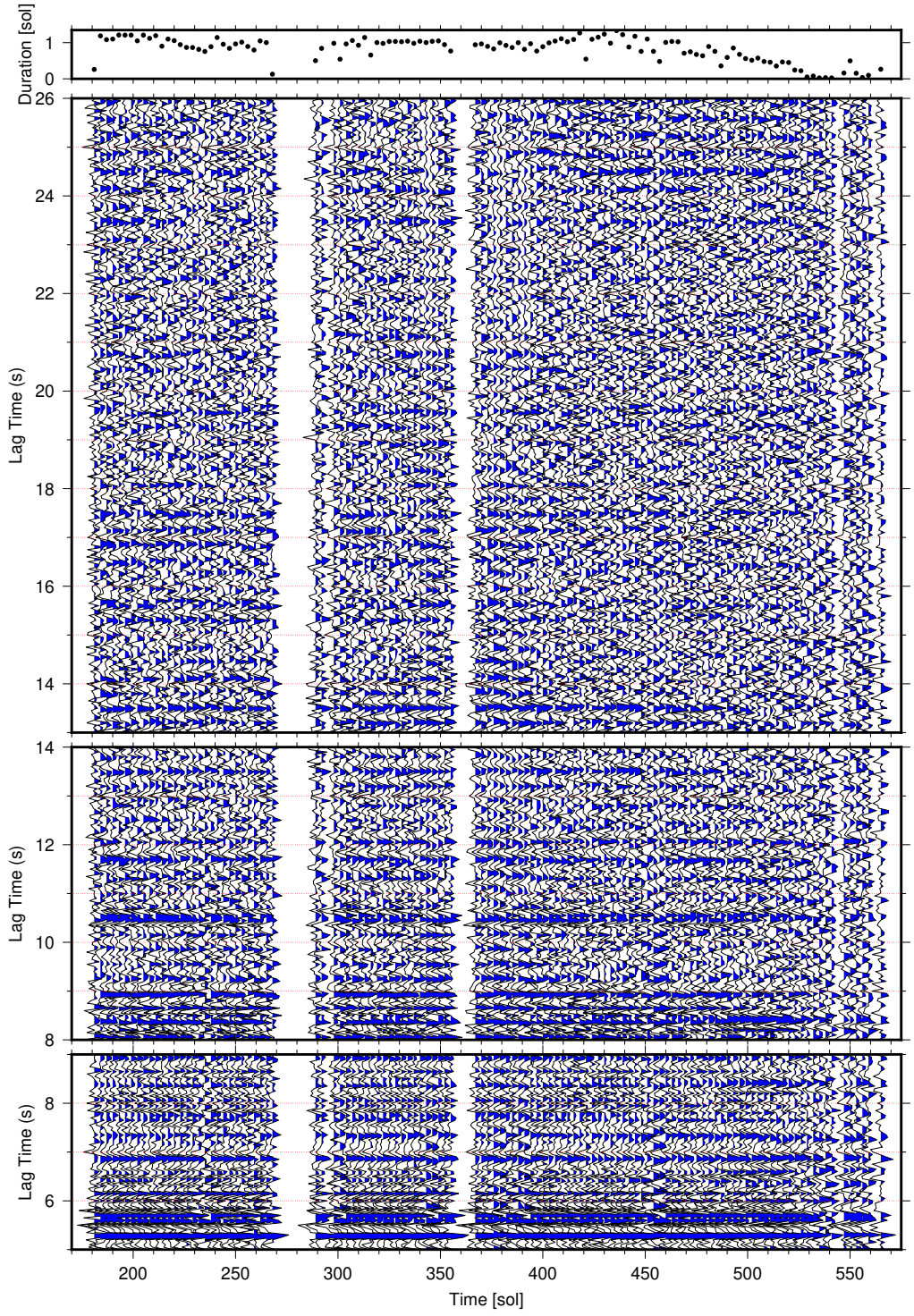


Figure S4. Same as Fig. 9, but employing the linear stack rather than the tf-PWS: Vertical-component noise autocorrelation stacks for sliding 30-Sol data windows. The frequency band is 1.2-8.9 Hz and data windows do not overlap. Shown are linear stacks of phase autocorrelations. Blue marks negative amplitudes. The three lag-time windows have been used to improve the visibility through independent amplitude normalization. The top panel shows the total duration of the selected data used to compute autocorrelations within each of the 30-Sol data windows. The axis for the duration of the 30 % and 3 % data set are scaled by the number to the top left.

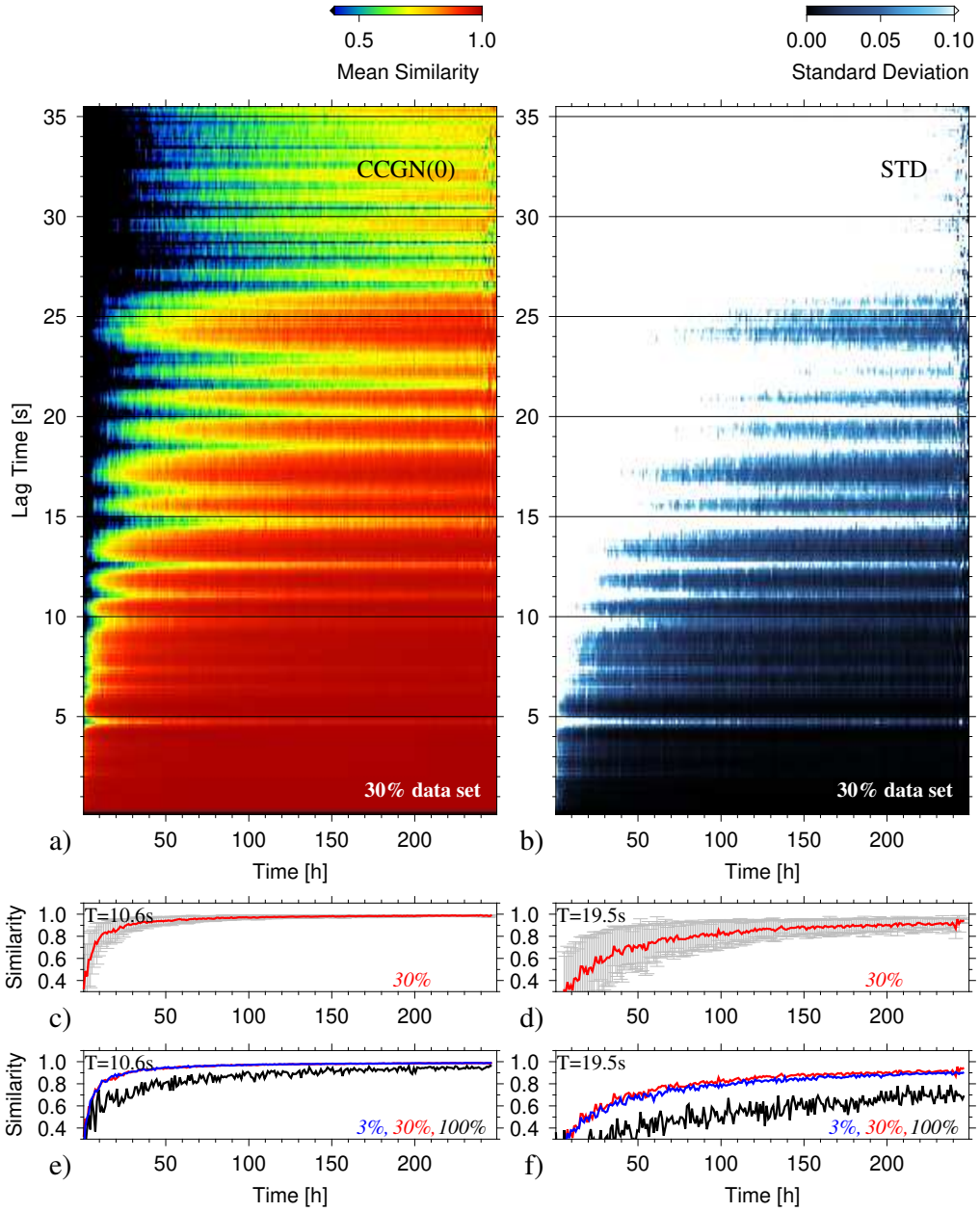


Figure S5. Same as Fig. 8, but using zero-lag CCGN rather than PCC to measure the waveform similarity. This figure shows a faster waveform convergence as PCC is the more waveform sensitive measure.

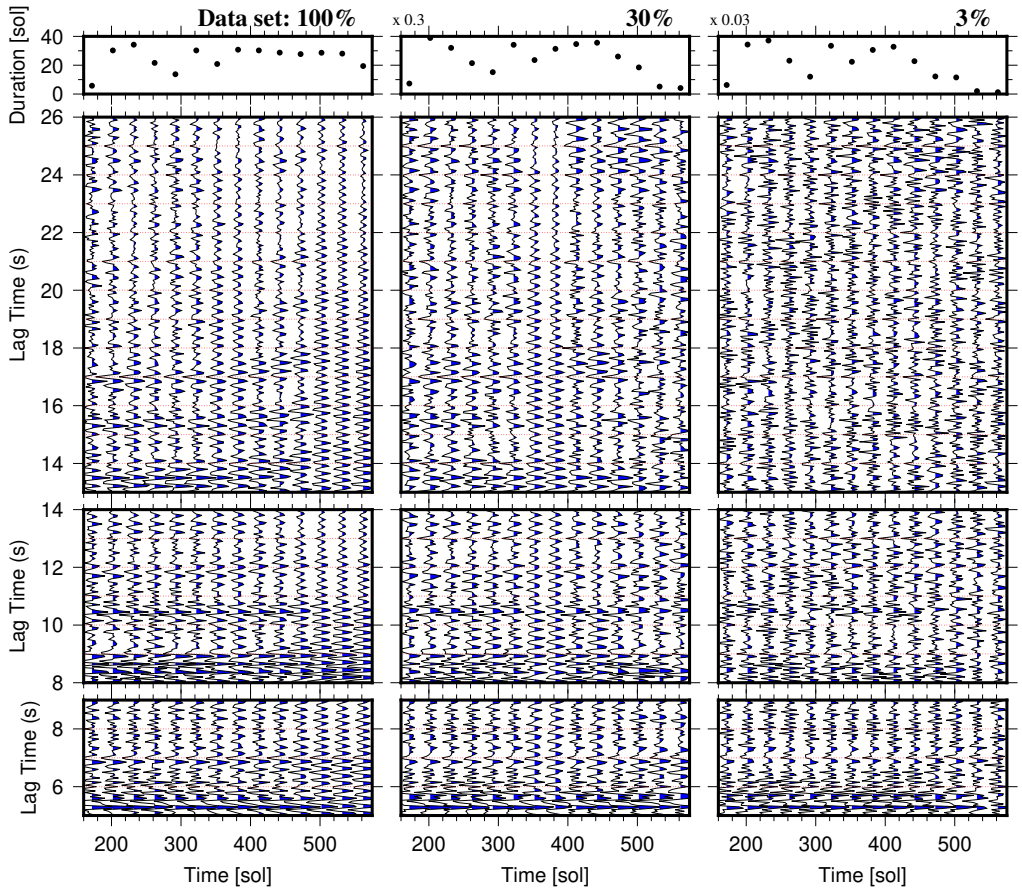


Figure S6. Same as Fig. 9, but employing the linear stack rather than tf-PWS: Vertical-component noise autocorrelation stacks for sliding 30-Sol data windows. The frequency band is 1.2-8.9 Hz and data windows do not overlap. Shown are linear stacks of phase autocorrelations. Blue marks negative amplitudes. The three lag-time windows have been used to improve the visibility through independent amplitude normalization. The top panel shows the total duration of the selected data used to compute autocorrelations within each of the 30-Sol data windows. The axis for the duration of the 30 % and 3 % data set are scaled by the number to the top left.

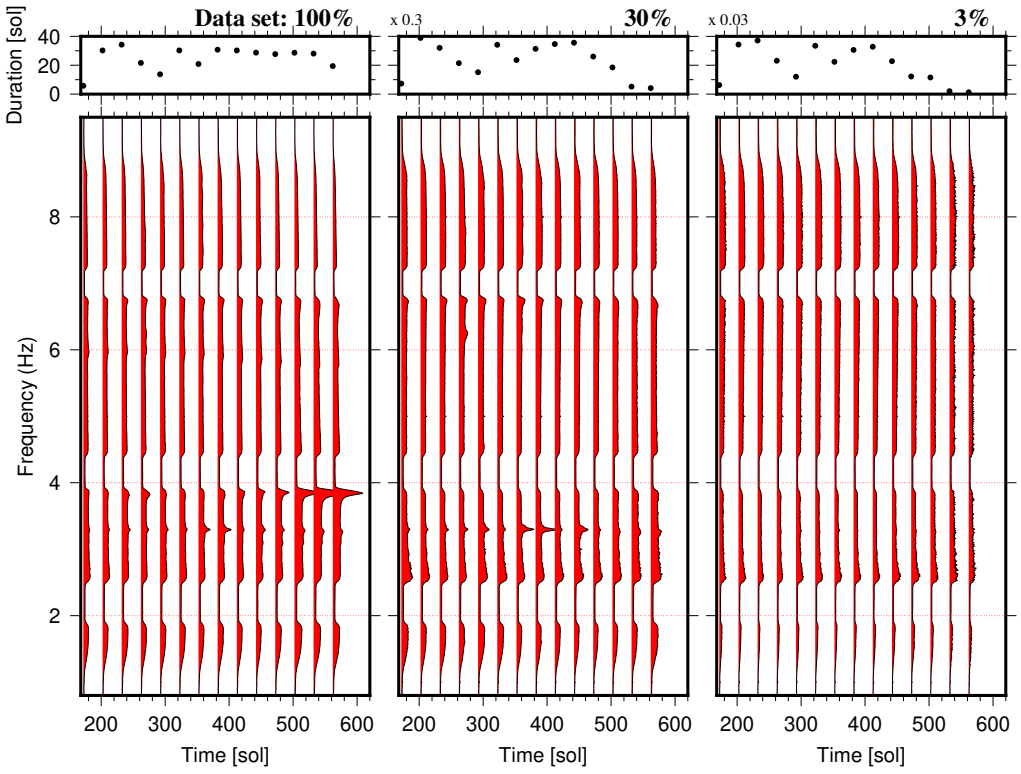


Figure S7. Same as Fig. 10, but using linear stacks: Amplitude spectra of linearly stacked autocorrelations as function of Sol for the 100 % (left panel), 30 % (middle panel), and 3 % (right panel) data sets. Stacks are build using all available autocorrelations within non-overlapping 30-Sol data windows. Each amplitude spectrum is placed at its window center time and has been normalized at 6 Hz. The total duration of data used to compute the autocorrelations within each 30-Sol window is plotted to the top. The numbers to the top left are factors to reduce the duration axes of the 30 % and 3 % data sets.

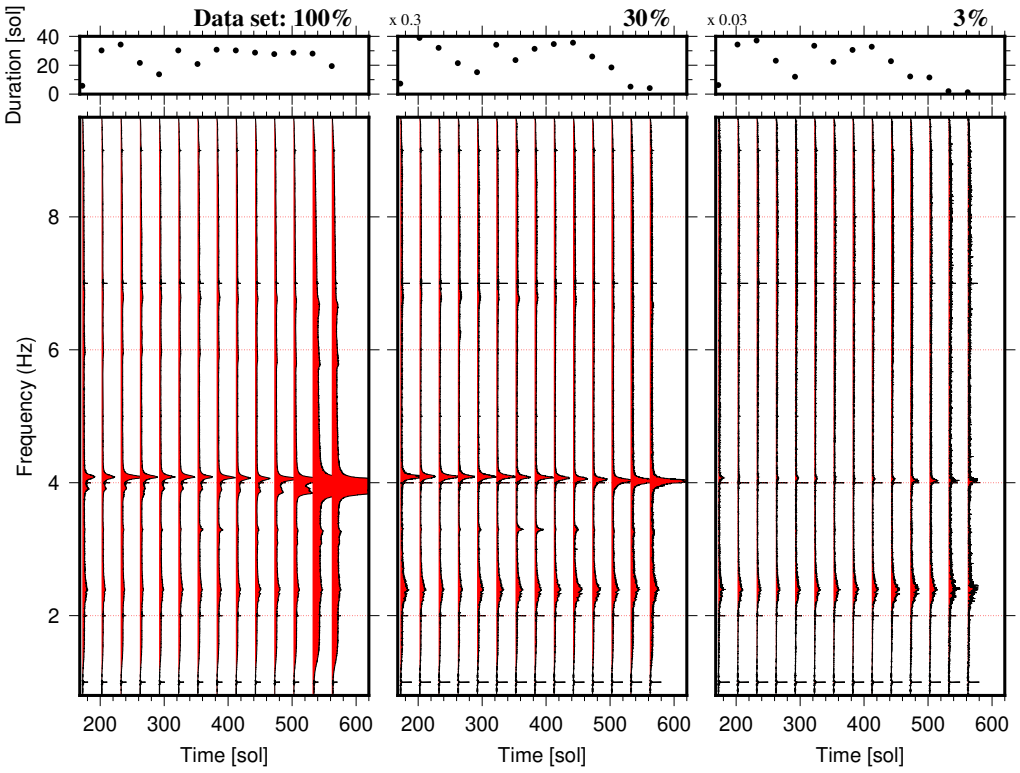


Figure S8. Same as Fig. S7, but using a slightly broader frequency band (0.8-9.5 Hz) and no band-rejection filters. Spectra are normalized at 1 Hz and amplitudes have been multiplied by 0.5 for the 100 % data set (left panel).

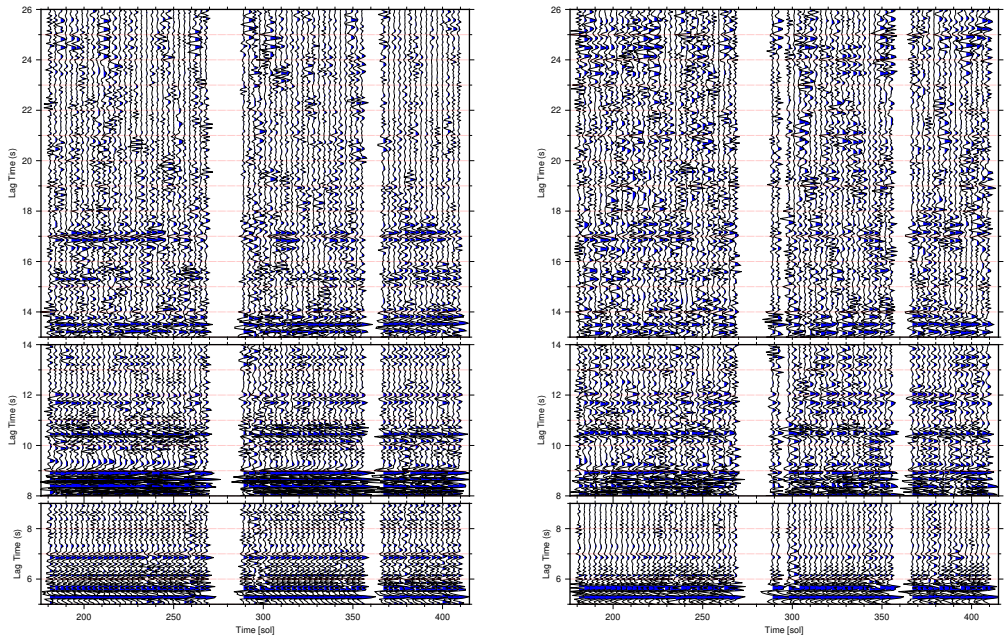


Figure S9. Vertical-component autocorrelation stacks for sliding 3-Sol data windows. The frequency band is 1.2-8.9 Hz and data windows do not overlap. Shown are tf-PWSs of phase autocorrelations computed for the 100 % (left panel) and 30 % data set up to Sol 410. Most of the signals appear for both data sets. This further testifies that PCC is a robust approach as data problems, glitches and donks present in the 100 % data set do not bury the signals shown with the 30 % data set.