Knapmeyer, M., Stähler, S. C., Daubar, I., Forget, F., Spiga, A., Pierron, T., van Driel, M., Banfield, D., Hauber, E., Grott, M., Muller, N., Perrin, C., Jacob, A., Lucas, A., Knapmeyer-Endrun, B., Newman, CP., Panning, M., Weber, R., Calef, F. J., ... Banerdt, W. B. (2022). Seasonal seismic activity on Mars. *Earth and Planetary Science Letters*. https://doi.org/10.1016/j.epsl.2021.117171

Earth and Planetary Science Letters 576 (2021) 117171

Contents lists available at ScienceDirect

Earth and Planetary Science Letters

www.elsevier.com/locate/epsl

Seasonal seismic activity on Mars

M. Knapmeyer [a,*], S.C. Stähler [b], I. Daubar [c], F. Forget [d], A. Spiga [d], T. Pierron [d], M. van Driel [b], D. Banfield [e], E. Hauber [a], M. Grott [a], N. Müller [a], C. Perrin [p], A. Jacob [f], A. Lucas [f], B. Knapmeyer-Endrun [g], C. Newman [h], M.P. Panning [i], R.C. Weber [j], F.J. Calef [i], M. Böse [b,r], S. Ceylan [b], C. Charalambous [k], J. Clinton [r], N. Dahmen [b], D. Giardini [b], A. Horleston [l], T. Kawamura [f], A. Khan [b], G. Mainsant [q], M. Plasman [f], M. Lemmon [n], R. Lorenz [o], W.T. Pike [k], J.-R. Scholz [m], P. Lognonné [f], B. Banerdt [i]

[a] *Institute of Planetary Research, DLR, Rutherfordstr. 2, 12489 Berlin, Germany*
[b] *Institute of Geophysics, ETH Zürich, Sonneggstrasse 5, 8092 Zürich, Switzerland*
[c] *Department of Earth, Environmental, and Planetary Sciences, Brown University, Campus Box 1846, Providence, RI 02912-1846, USA*
[d] *Laboratoire de Météorologie Dynamique/Institut Pierre Simon Laplace (LMD/IPSL), Sorbonne Université, Centre National de la Recherche Scientifique (CNRS), École Polytechnique, École Normale Supérieure, France*
[e] *Cornell University, Cornell Center for Astrophysics and Planetary Science, Ithaca, NY, 14853, USA*
[f] *Université de Paris, Institut de physique du globe de Paris, CNRS, F-75005 Paris, France*
[g] *Bensberg Observatory, University of Cologne, Vinzenz-Pallotti-Str. 26, 51429 Bergisch Gladbach, Germany*
[h] *Aeolis Research, 333 N Dobson Road, Unit 5, Chandler, AZ 85224-4412, USA*
[i] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., M/S 183-301, Pasadena, CA 91109, USA*
[j] *NASA MSFC, NSSTC Mail Code ST13, 320 Sparkman Drive, Huntsville, AL 35805, USA*
[k] *Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom*
[l] *School of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, United Kingdom*
[m] *Max Planck Institute for Solar System Research, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany*
[n] *Space Science Institute, 4765 Walnut Street, Suite B, Boulder, CO 80301, USA*
[o] *Johns Hopkins Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723, USA*
[p] *Laboratoire de Planétologie et Géodynamique, UMR6112, OSUNA UMS3271, Univ. Nantes, Univ. Angers, CNRS, 2 rue de la Houssinière, BP 92208, 44322 Nantes Cedex 3, France*
[q] *Institut Supérieur de l'Aéronautique et de l'Espace SUPAERO, 10 Avenue Edouard Belin, 31400 Toulouse, France*
[r] *Swiss Seismological Service (SED), ETH Zurich, Sonneggstr. 5, 8092 Zurich, Switzerland*

ARTICLE INFO

Article history:
Received 15 February 2021
Received in revised form 9 August 2021
Accepted 12 August 2021
Available online xxxx
Editor: W.B. McKinnon

Dataset link:
https://pds-geosciences.wustl.edu/missions/insight/seis.htm

Dataset link:
https://www.seis-insight.eu/en/science/seis-data/seis-data-description

Dataset link:
https://www.iris.edu/hq/sis/insight

Dataset link: https://atmos.nmsu.edu/data_and_services/atmospheres_data/INSIGHT/insight.html

ABSTRACT

The rate of occurrence of High Frequency (HF) marsquakes, as recorded by InSight at Homestead Hollow, Elysium Planitia, increased after about $L_S = 33°$, and ceased almost completely by $L_S = 187°$, following an apparently seasonal variation with a peak rate near aphelion. We define seismic rate models based on the declination of the Sun, annual solar tides, and the annual CO_2 cycle as measured by atmospheric pressure. Evaluation of Akaike weights and evidence ratios shows that the declination of the Sun is the most likely, and the CO_2 cycle the least likely driver of this seismic activity, although the discrimination is weak, and the occurrence of a few events in August 2020 is in favor for a triggering by CO_2 ice load. We also show that no periodicity related to Phobos' orbit is present in the HF event sequence. Event rate forecasts are presented to allow further discrimination of candidate mechanisms from future observations.

---

* Corresponding author.
E-mail address: martin.knapmeyer@dlr.de (M. Knapmeyer).

https://doi.org/10.1016/j.epsl.2021.117171

Dataset link:
https://pds-imaging.jpl.nasa.gov/

Keywords:
Mars
Elysium Planitia
InSight
seasonal seismic activity
Phobos

1. Introduction

NASA's InSight (Interior exploration using Seismic Investigations, Geodesy, and Heat Transport) mission, launched in May 2018, landed in the western part of Elysium Planitia, about 500 km to the north of the dichotomy boundary and 1500 km west of Cerberus Fossae, which was suspected to be seismically active by Taylor et al. (2013). After the touchdown on 26. Nov. 2018, deployment of the seismometer experiment SEIS (Seismic Experiment for Interior Structure; Lognonné et al., 2019, 2020) was completed on sol 70 (the 70th Martian day of the mission). SEIS started continuous recording on sol 73 (09. Feb. 2019), and reached full performance on sol 85 (21. Feb., Banerdt et al., 2020). With the exception of a downtime from 28. Aug. (sol 267) to 18. Sep. 2019 (sol 288, a minor power management problem was extended to three weeks by the 2019 solar conjunction), and a few smaller data losses due to downlink interruptions, SEIS has been recording seismic data continuously ever since.

Estimates of the seismic activity of Mars published prior to the InSight mission basically assumed that marsquakes would occur like earthquakes: according to a Gutenberg-Richter magnitude distribution, and randomly in time (Phillips, 1991; Knapmeyer et al., 2006; Taylor et al., 2013; Plesa et al., 2018, and references therein). A certain class of events, namely High Frequency (HF) and 2.4 Hz events, however, exhibits an occurrence rate that appears variable in time (Giardini et al., 2020), as shown in Fig. 1. HF and 2.4 Hz events are considered jointly here since the latter are small HF events visible only by their excitation of an ambient resonance at 2.4 Hz (Clinton et al., 2021). We interpret the timing of these events as result of a nonstationary Poisson process (supplemental text S7), and compare it with several time-varying processes in order to identify what might drive the nonstationarity. Without even crude epicenter locations, and without knowledge about the current general state of stress - visible faults are millions to billions of years old (Knapmeyer et al., 2006) - a causal stress model to modulate the event rate is not warranted. Our goal is thus to identify a process with a time dependency fitting the event rate, which may inform the identification of candidate source regions: A change of seismicity based on changes of surface load, for example, would point to regions with seasonal ice cover, while a correlation with solar tides would point to the regions with the largest tidal stresses.

Temporal variation of seismic activity is not unheard of. The activity of deep moonquakes is tied to the revolution around the Earth (e.g. Nakamura, 2003; Kawamura et al., 2017), with some source regions being active around perigee or apogee, while others are active at the orbital nodes (Bulow et al., 2007). Short-term temporal variations in lunar impact rate were also detected (Oberst and Nakamura, 1987)

Event rate variations were also described on Earth. Bollinger et al. (2006), for example, conclude that an annual seismicity cycle in Nepal, with a maximum from January to March, is due to the mass loading resulting from the summer monsoon in the Ganges area. At Mt. Hochstaufen, Germany, a strong correlation of seismicity at up to 4.5 km hypocentral depth with rainfall is explained by pore pressure changes in the hectopascal range, which follows rain (Hainzl et al., 2006).

Martian seismicity is certainly not controlled by rainfall, but groundwater pressure might be relevant. Manga et al. (2019) suggest that stresses due to solar tides, Phobos tides, and the diurnal variation of barometric pressure may induce seismic activity by varying aquifer pore pressure beneath a cryospheric ice layer.

The demonstration of a time dependent seismicity on Mars, and of a possible driving force, is impeded by circumstances visible in Fig. 1: Solstices and apsides occur at nearly the same time, i.e. seasonal changes of the weather conditions, especially wind speed, insolation, and atmospheric CO_2 cycle, but also the solar tides all show similar periods and small phase shifts with respect to each other. As reported before by InSight (Lognonné et al., 2020; Giardini et al., 2020; Banfield et al., 2020) and other space craft (Martínez et al., 2017), wind speed varies considerably between day and night, and between seasons, and wind at the landing site is the main source of seismic background noise (Lognonné et al., 2020). It must be shown first that we observe a variation of a seismic source process rather than of detectability. In order to achieve this, a representative subset of all events must be chosen that excludes uncertain detections and possible non-seismic disturbances.

We first describe criteria for the selection of events to evaluate, how we compare a given event rate model with the observed sequence of events, and then discuss a number of simple rate models that are motivated by astronomical or meteorological effects: a noise-dependent detectability of an otherwise constant rate, Phobos tides, impacts, insolation, CO_2 sublimation/deposition, and solar tides. These rate models are then compared and ranked using information theoretical methods.

All mathematical details, including tables coefficients and parameters, and technical details of the grid search for parameter estimation, are presented in the supplementary information, and will be referred to where appropriate (S*n* to refer to supplementary text *n*, and SF*n*, ST*n* for figures or tables).

2. Event selection

We use a preliminary version of the fifth event catalog (InSight Marsquake Service, 2021) covering events that occurred between 12. Jan. 2019 and 31. Aug. 2020. The catalog was published on 04. Jan. 2021, our version was however frozen on 03. Nov. 2020 to have a stable working version. There are no changes from the fourth (InSight Marsquake Service, 2020) to the fifth catalog release concerning HF events, except for quality D event S0568a (02. July 2020). Using the November freeze instead of the public release does thus not affect our results.

The definitions of event categories are given by Clinton et al. (2021, also S2) and are mainly based on the frequency content of signals and, to a lesser extent, other features of the waveforms.

None of the HF events analyzed by van Driel et al. (2021) has a magnitude larger than about 2.2. The events cluster between approx. 20° and 30° distance (van Driel et al., 2021), it was however not possible to locate any of them. Both distances and magnitudes currently depend on assumptions about crustal velocities ($v_S = 2.3$

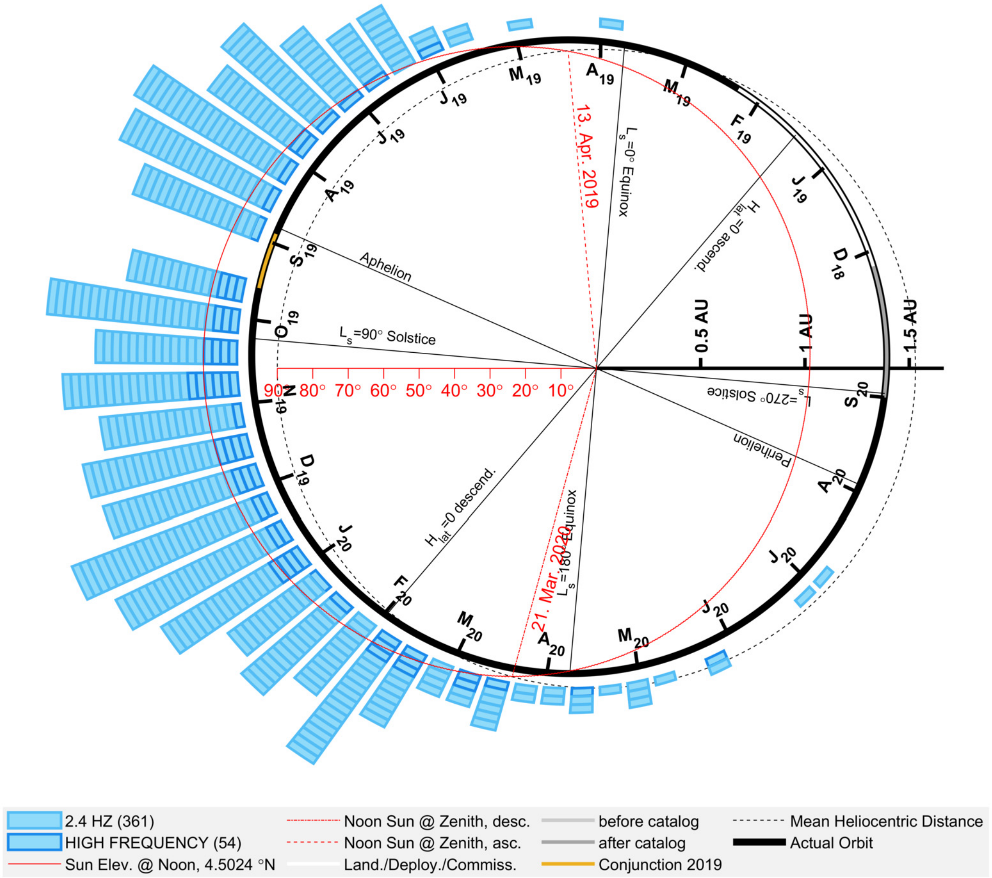


Fig. 1. HF activity over time (all 415 detected events up to 31. Aug. 2020), in relation to the Martian orbit. Rectangular cells indicate the occurrence times of HF events in bins of 5° heliocentric longitude, each rectangle corresponds to one event. High Frequency and 2.4 Hz events are distinguished by edge shading as indicated in the legend. The total event counts for both groups are given in the legend, in brackets. The dashed circle marks the mean orbital radius of Mars (about 1.52 AU), while the heavy ellipse shows the actual orbit, Mars moving counterclockwise (lighter parts indicate times without registration, before and after catalog refers to times before 06. Feb 2019 and after 31. Aug. 2020, as covered by the event catalog used here). Ticks along the orbit denote the beginning of terrestrial months from December 2018 (label D18) until September 2020 (label S20). Equinoxes, solstices, perihelion, aphelion, and zero heliocentric latitude are also marked. The eccentric thin line, and the degrees-scale pointing to the left, indicate the elevation of the sun at noon for the InSight landing site. On 13. Apr. 2019 and 21. Mar 2020, the noonday sun was in the zenith over the InSight lander, and north of InSight in the time between. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

km/s, $v_P/v_S = 1.7$ in van Driel et al., 2021), where future studies will provide more reliable values.

We focus on the HF family (415 events), but exclude the Very High Frequency events which show a wide distribution of epicentral distances (van Driel et al., 2021), suggesting they are caused by a different process. A future study will deal with the magnitude-frequency distribution and the distribution in time of the Low Frequency and Broadband events (Clinton et al., 2021).

Based on the clarity of seismic arrivals and the degree of polarization, Clinton et al. (2021) define four quality levels “A” (best) to “D” (events that might be artefacts, e.g. wind gusts). We remove all events classified as quality “D” to avoid contamination with non-seismic effects. No HF event is of quality “A”.

To cope with limited data transmission volumes, InSight implemented a strategy which includes the buffering of full resolution data on the lander, while only downsampled channels were downlinked continuously (Lognonné et al., 2019). Up to 01. June 2019, continuous data is available at 10 Hz sampling frequency, and at 20 Hz afterwards (see also figure SF1). Data with higher sampling rates are also available for times of identified events prior to 01. June 2020. Although the majority of HF events should be detectable in the 10 Hz data stream, we exclude from our analysis all data that was recorded prior to 01. June 2019.

The most important source of environmental noise is wind. Turbulent winds build up after sunrise and reach speeds of 20 m/s, sometimes more. Around 17:00 LTST, about one hour before sunset, the turbulent flow collapses, and slow laminar winds prevail for several hours (Banfield et al., 2020). Wind speeds increase again after midnight. This has direct consequences for the detection of events, as Giardini et al. (2020) show: 80% of all HF events were detected between 17:00 LTST (Local True Solar Time) and midnight, 15% between midnight and sunrise, and the remaining 5% shortly after sunrise or shortly before 17:00 LTST. The distribution of wind speeds shown by Banfield et al. (2020) mirrors the noise amplitude recorded by SEIS: Figs. 2a and c show the probability density distribution of the noise amplitude, computed as standard deviation of the vertical displacement during 2 min windows, and in relation to a reference displacement of 1 m. These PDFs show the increase of noise during daytime (Fig. 2a and b) as well as its evolution throughout the seasons (Fig. 2b and c).

The events with the smallest recorded displacement amplitudes are most affected by wind noise. The set of selected events should however be representative for the entire population. We therefore exclude from our analysis all events that are recorded with small amplitudes or low signal-to-noise (SNR) ratio (S2 summarizes data processing). To determine acceptance thresholds, we analyze the distributions of amplitudes and SNRs. Fig. 3 shows that the cumulative size-frequency distributions of both amplitudes and SNRs follow a Gutenberg-Richter-like power law. We estimate completeness thresholds for both using the maximum curvature method

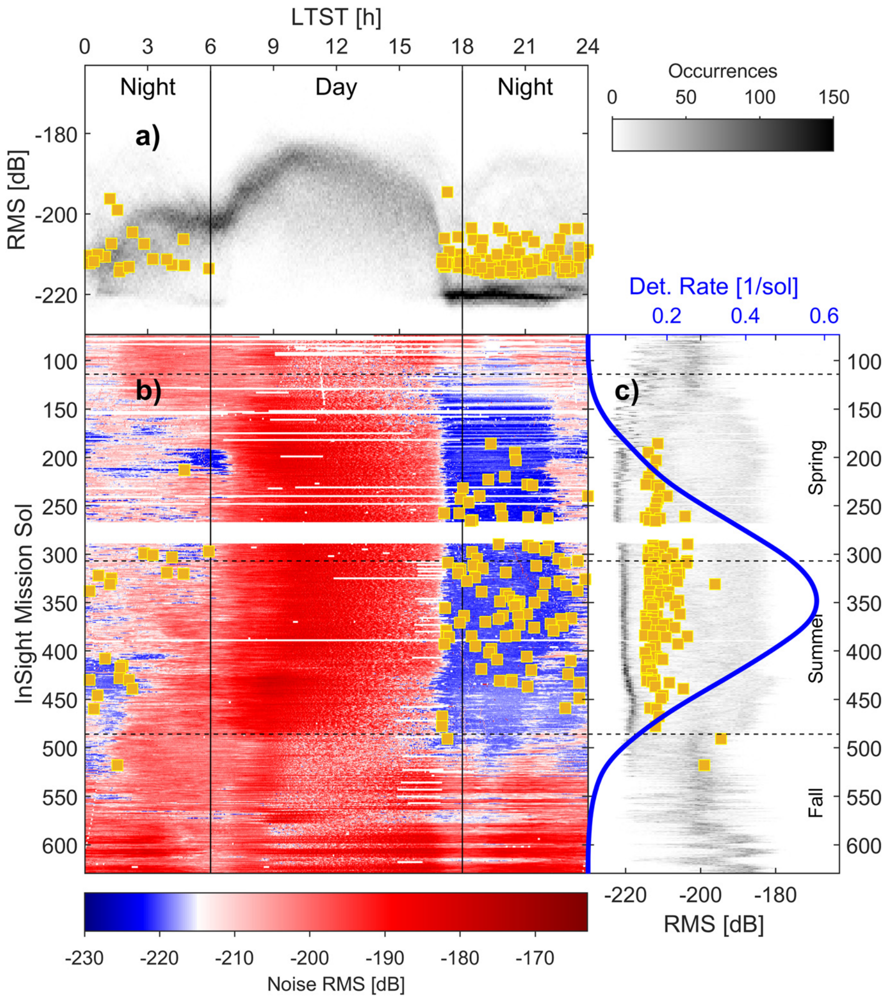


Fig. 2. SEIS background noise and times and amplitudes of the 118 selected events. Noise level is computed as RMS amplitude of vertical displacement between 1.2 Hz and 3.0 Hz, in dB relative to 1 m, and averaged over 2 min windows. Event displacement amplitudes as described in S2. (a) Distribution of noise RMS (2D occurrence count histogram, scaled according to color bar on the right) and event amplitudes (squares) as function of local time LTST. Vertical lines indicate mean sunrise and sunset. (b) Noise RMS amplitude (background color) and event detection times (squares) over time. InSight mission Sol is on the vertical, local time LTST on the horizontal axis. The white point of the color scale is set to the smallest event amplitude (event S0383a, ca. 20:30 LTST, −214.89 dB), such that all events would be detectable at times where the background is blue. Data gaps also map to horizontal white lines. (c) distribution of noise RMS (2D histogram, same scaling and color bar as in (a) and event amplitudes (squares) as function of the InSight mission Sol, dashed horizontal lines indicate beginning of northern seasons. A kernel density estimation of the detection rate is shown in blue (see S3).

(employing kernel density estimation (KDE), see also S4), the distribution slope is obtained as maximum likelihood power law (S4). Based on the results we exclude all events that were recorded with vertical displacement amplitudes below −215 dB, or with an SNR below 2.56. We do not attempt to convert displacement amplitudes into source magnitudes or seismic moments, as this would require better knowledge of epicentral distances than we currently have (cf. van Driel et al., 2021).

With all selection criteria combined, 118 events remain (squares in Fig. 2), of which S0518a is the most recent. These occurred during 458 days (446 sols or 0.67 Martian years). The estimated rate of detection of acceptable events is shown in Fig. 2c (blue), obtained from a kernel density estimation (S3).

3. Inverse problem, parameterization and uncertainties

The problem to solve is to estimate the parameters of a stochastic point process, in the light of a single realization of that process (i.e. the event catalog), requiring a maximum likelihood approach.

If the point process is a non-stationary simple Poisson process (S7), the likelihood L of a candidate rate function is given by (e.g. Ogata, 1983)

$$\log L(\theta) = \sum_{i=1}^{N} \log \lambda(t_i, \theta) - \int_{S}^{T} \lambda(u, \theta)\, du \qquad (1)$$

(t_i: detection times, λ: rate function, θ: parameter vector, times S and T: beginning and end of observations, respectively). This likelihood is the product of the probability that no event occurs between the observed ones, and the probability of occurrences in an infinitesimal interval after the given times (Ogata, 1983). Multiplication is granted if events are independent. Equation (1) uses the event times directly - it is not necessary to produce any numerical rate estimation based on histograms or kernel densities. Arbitrary, continuously defined candidate rate functions can be compared directly with the event catalog. Lacking epicenters, we approximate source times by arrival times. This introduces 5 to 8 min time errors (van Driel et al., 2021), negligible in comparison to expected rate changes.

We define a generic occurrence rate model, where the actual event rate is the sum of some function of time and a constant baseline rate $\lambda_B \geq 0$, and can never fall below the baseline. Also, the event rate is allowed to lag behind the driving function ac-

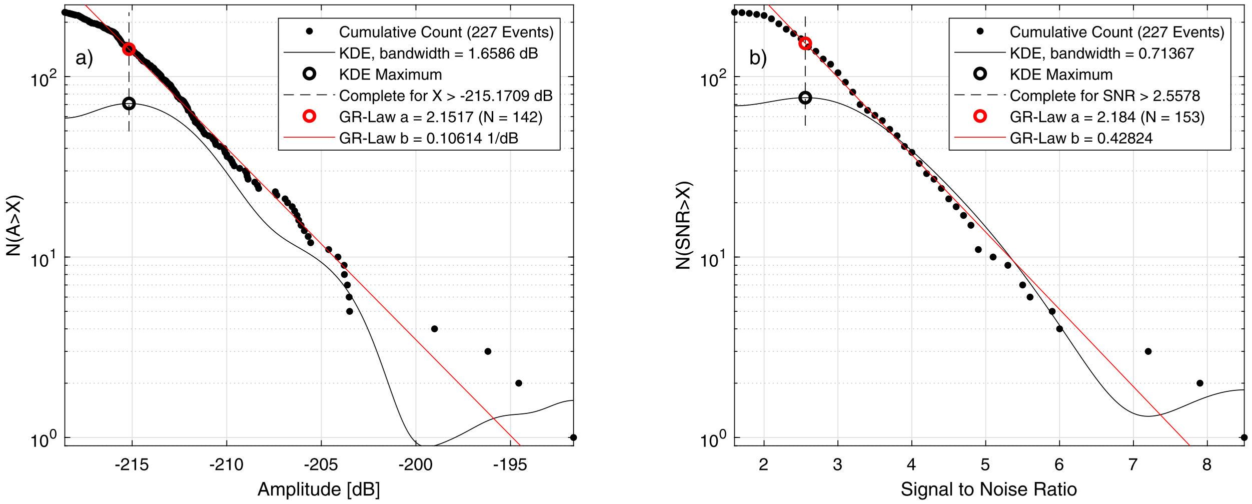


Fig. 3. Event size dependent acceptance thresholds (quality A, B, C events only, 227 events in total): (a) Displacement-Amplitude-frequency distribution, (b) SNR-frequency distribution. Dots represent the observed amplitudes and SNRs, black lines show a kernel density estimation (KDE, with bandwidth as given in the legend) of the respective PDFs (scaled to better comply with the vertical axis range, as only the location of the maximum is important), straight lines show the fitted power law (with completeness threshold a inferred from the KDE, Number N of events above completeness threshold in brackets, and slope b from maximum likelihood estimation, values given in the legends).

cording to a retardation parameter. The detection rate function, as a function of time t and a parameter vector θ, is

$$\lambda\left(t,\theta\right)=\Upsilon\left(t\right)\eta_{det}\left(t\right)\max\left(\lambda_B,f\left(t,\theta_{\setminus\lambda_B}\right)+\lambda_B\right) \qquad (2)$$

where f is a model-specific kernel function that depends on time and a parameter vector $\theta_{\setminus\lambda_B}$ (meaning θ without λ_B), and λ_B is the baseline event rate. The function $\Upsilon\left(t\right)$ is 0 during downtimes, 1 else, to describe data gaps (ST2) in the likelihood evaluation. The detection efficiency function $0\leq\eta_{det}\left(t\right)\leq 1$ accounts for the variability of the noise background (see S4 for details).

The maximum likelihood solution θ_{MLE} is obtained from a sequence of nested grid searches (S10). Models with different physical backgrounds are ranked based on Akaike's information criterion (corrected, AICc, S13) and evidence ratios (S14).

The variances of model parameters are given by the inverse Fisher information matrix of the log-likelihood. If this results in negative or extremely large values, we resort to a comparison of likelihoods across the grid and obtain an uncertainty corresponding to $\approx 3\sigma$ (S12).

4. Event rate models

4.1. Can we rule out a constant event rate?

Even if the actual production of events does not vary with time, the strong variation of wind induced noise on SEIS would still result in a parallel variation of the number of event detections per sol. Before comparing event detections with any forcing, it is necessary to demonstrate that a variation exists beyond that expected due to the variable wind noise (i.e. $f\left(t,\theta_{\setminus\lambda_B}\right)\equiv 0$ is the worst kernel). Beyond the qualitative demonstration of Fig. 2, we need a quantitative description.

While it is impossible to know how many events we miss during daytime, we can estimate how many of the events detected at night we would miss if the noise level were higher.

In order to detect events with a displacement of -215 dB with an SNR of 2.56, the noise level must be -223 dB, which is about the lowest observed noise amplitude (Fig. 2). At higher noise levels, the detection rate will be reduced according to the slope of the amplitude-frequency distribution (Fig. 3a) We compute, from noise seismograms and the amplitude distribution slope, a detection efficiency $\eta_{det}\left(t\right)$ which estimates the fraction of events that can be detected at any given time (S4). The difference between the observed cumulative event count and that expected for a stationary Poisson process with detections impeded by the detection efficiency (Fig. 4) then indicates if a significant nonstationary component exists (S5).

In the maximum likelihood analyses of all rate models, we use a polynomial approximation (coefficients in S6) of the empirical detection efficiency function.

As shown in Fig. 4a, the detection rate is almost constant between sol 289 and sol 385; about $67/96=0.698$ events/sol. The thin blue dashed line in Fig. 4a extrapolates the *detection* rate expected for a stationary Poisson process.

Between sol 289 and sol 385, the detection efficiency (red in Fig. 4a) is almost constant at 0.146, implying an *occurrence* rate of 4.78 events/sol. By multiplying this occurrence rate with the time dependent detection efficiency for any time, we obtain the cumulative event count that results from noise level variations (blue solid curve in Fig. 4a). This follows a straight line until about sol 460, when it starts to bend down due to the decreasing detection efficiency, while the observed count starts bending down on sol 385: The observed detection rate starts decreasing before a decrease could be expected from the increasing noise level.

At the time of the last observed event, on sol 518, there is a deficit of 58 events between the observed cumulative count and the straight-line extrapolation, while the deficit between the noise-modulated stationary Poisson process and the straight-line extrapolation is only 18 events: The observed detection rate decreases faster than could be expected from the decreasing detection efficiency. Similar deviations are visible before sol 289.

The observed event sequence is only a single realization of a stochastic process. Individual 446-sols realizations from these processes would however show deviations from the expectation, according to the variance of Poisson processes.

We therefore repeat an analysis like the above for all times at which events were detected, and with a large number of realizations of the simulated noise-modulated stationary Poisson process (see caption of Fig. 4). We compute a residual (observed minus predicted detection count) and calibrate it in terms of the probability that it is smaller than its actual value coincidentally (Fig. 4b).

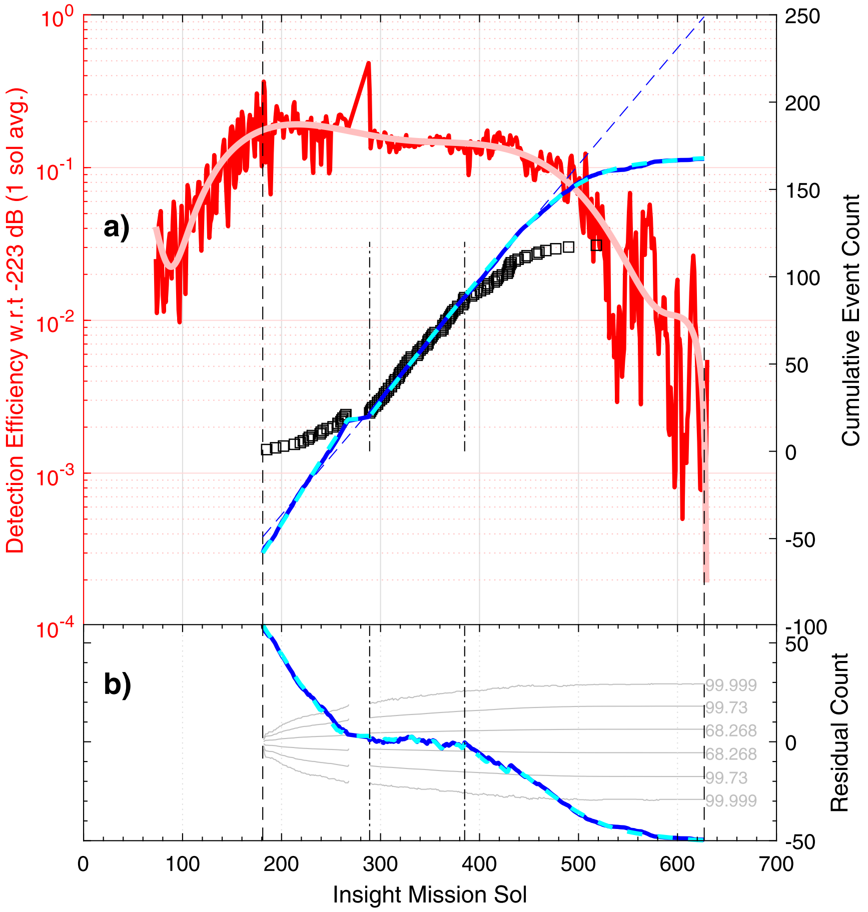


Fig. 4. Nonstationary event rate: (a) Detection efficiency (red curve) relative to a per-sol average noise level of −223 dB, and polynomial fit (pink); Dashed vertical lines: catalog time window. Dash-Dot vertical lines: time window used to estimate the rate of a stationary Poisson process; squares: observed cumulative event count; thin blue dashed line: stationary Poisson process with detection rate 0.698 events/sol (see text); heavy blue solid line: modulation of that process by the actual detection efficiency; heavy blue dashed line: modulation of the same stationary Poisson process by the polynomial fit to the detection efficiency. (b) Residual between observed cumulative count and modulated stationary Poisson processes; thin gray lines: percentiles of the distribution of 10^7 synthetic stationary event sequences that use the rate determined in the time window between dash-dot lines. Percentile values correspond to 1, 3, and 5 standard deviations of a normal distribution.

The probability of accidentally obtaining the observed residual is below 10^{-6} at both ends of the observation time.

For computational efficiency, we use a polynomial approximation (pink in Fig. 4a). Cumulative count and residual are hardly distinct from the empirical curve and residual.

In summary, we have shown that the event detection rate drops too early, too fast, and statistically highly significant. Therefore, a constant event rate is ruled out.

To validate our grid search scheme, we run it with a kernel function

$$f\left(t,\theta_{\setminus\lambda_B}\right)=0 \tag{3}$$

Using only the events between sols 289 and 385 (parameters in ST11), we find an optimum occurrence rate of 4.63 events/sol ($\log L=-92.97$, $AIC_c=188$, after 6 iterations), i.e. close to the 4.78 estimated above, while the theoretical values for a rate of 67/96 are $\log L=-91.1$ and $AIC_c=184.2$ (see S9).

From a grid search using all selected events (ST3), the AICc of the best constant rate is 476.9 (ST17). Any viable time dependent model must have a smaller AICc.

4.2. Are HF events caused by Phobos tides?

We first investigate if the event sequence contains a short period component that could be connected to the groundwater pore pressure effect suggested by Manga et al. (2019).

A spike train Fourier transform (Bulow et al., 2007, S8) does not show any periodicity of event occurrence that can be related to Phobos' orbit. Actually, the power spectral density at Phobos monthly periods (draconitic, anomalistic, synodic) and its culmination period is low compared to the remainder of the analyzed period windows (SF8). The power spectral density shows clear peaks at diurnal and semidiurnal periods. Comparison with synthetic sequences (SF9) lets us attribute the diurnal peak to the low detection probability during daytime. The semidiurnal peak might partially be due to tides, but clearly contains a contribution from the different detection probabilities during the first and second half of the night (SF9). Isolating a solar diurnal or semidiurnal tide signal would require additional effort outside the scope of this paper, which is on longer periods. At this point, we can neither confirm nor exclude the existence of a semidiurnal tidal signal, although we expect that a more dedicated analysis would rather reject it.

We conclude that Phobos did not induce the HF marsquakes studied in this work.

4.3. Are they impacts?

Theoretical models of short-term variations in the impact rate (Ivanov, 2001; JeongAhn and Malhotra, 2015) predict an increased impact rate around aphelion, slightly before the observed seismic activity peak. Although impacts are currently occurring on Mars at

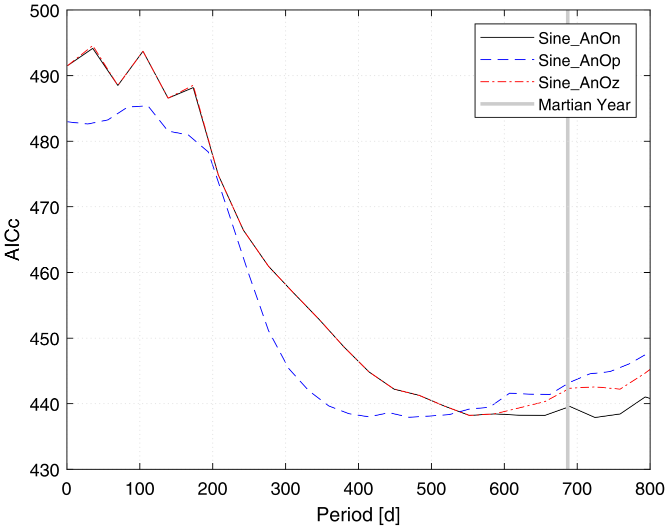


Fig. 5. Lowest AICc values attained for each tested Sine model period (naming conventions: "A" is for amplitude parameter, "O" for offset, "p", "n", "z" denotes positive, negative, or zero values, respectively; the likelihood functions for positive amplitudes are symmetric to those with negative amplitudes when inverting the offset sign). The vertical line indicates an annual period. Since the Sine models used cover different parts of the same parameter space, the existence of minima with the same depth points to the existence of a trade off zone that extends throughout the multidimensional parameter space. With the given data, only periods shorter than 300 days or so can be excluded with some certainty.

a measurable rate (Daubar et al., 2013; Malin et al., 2006), and a fresh crater was found photographically only about 37 km south of InSight, we do not yet have an unambiguous seismic observation of an impact (Daubar et al., 2020). However, attributing the variation in the HF event occurrence rate to impacts would imply that a significant fraction, if not all, of the HF events were in fact impacts. This is also in contradiction with the epicentral distance range of van Driel et al. (2021): Impacts should be distributed uniformly over the planetary surface, and sinusoidal over epicentral distance, with a maximum at $\Delta = 90°$. We do not think that the current observations support an identification of HF events as impacts, so we lay aside the impact hypothesis for now.

4.4. Is there an annual cycle?

The spike train Fourier analysis has no resolution at periods as long as a Martian year yet. Estimating a period length from an observational interval shorter than the expected period clearly involves some extrapolation and can never provide a definitive proof. To test for a periodicity corresponding to that of the Martian year (i.e. $T = 686.9726$ d, S16), we use the simplest periodic kernel function,

$$f(t) = A\sin(2\pi t/T - \varphi) + K \tag{4}$$

and determine the maximum likelihood period. For eq. (4), the duration of the time windows at which $\lambda(t,\theta) = \lambda_B$ holds, depends on $\sin^{-1} K/A$, thus K is not redundant with λ_B.

We split the parameter space into quadrants of the A-K-plane for reasons of CPU time management. The most likely period found for this kernel is 565 d (550 sols), for $K = 0$. The likelihood maxima identified with nonzero offset are at both shorter and longer periods, inspection of the search grids however shows that the period is poorly resolved (Fig. 5). The AICc as function of period shows a trough that shifts toward longer periods as the Offset parameter becomes smaller and negative, suggesting the existence of a tradeoff between Offset and other parameters. The common envelope of all curves in Fig. 5 shows that the assumed period can be varied over a wide range without increasing the AICc. The standard deviations obtained analytically and from comparing likelihood values across the grid are either infinite or encompass the entire search grid (ST15, ST16), indicating that a reliable determination of the period is not yet possible. This evaluation nevertheless supports that there is no significantly shorter period involved.

This analysis neither proves nor disproves the existence of an annual cycle. Strictly spoken, we can also not exclude the possibility that we witnessed a singular burst of activity, unless the activity actually repeats. But since annual periods are plausible and not prohibited, we investigate several mechanisms which imply an annual period.

4.5. Illumination model

The Sun's elevation above the equator is described by $\sin L_S(t)$ (S15), where L_S is the areocentric longitude of the Sun commonly used to define Martian seasons (S16). By a small modification of eq. (4) we thus obtain the illumination model

$$f(t) = A\sin(L_s(JD_{TT} - D)) + K \tag{5}$$

as kernel function, where time is measured as Julian Date JD_{TT} (index TT for terrestrial time, i.e. no leap seconds are introduced), and retardation is defined as a time delay D in days.

4.6. Surface load model

Induced seismicity is known on Earth from water level changes in reservoirs, caused by pore pressure changes at depth that follow the load change (Bell and Nur, 1978). The phase lag between infill change and seismicity can even exceed 180° (Roeloffs, 1988). Besides reservoir mass, the water level change rate is one of the key parameters that control the temporal distribution of induced seismicity (Saxena et al., 1988). As the polar caps of Mars are known to show a seasonal deposition/evaporation cycle since Herschel (1784), we investigate the possibility of changing CO_2 ice load as driver for the HF seasonality. A discussion of the involved atmospheric pressure changes and mass movements is provided in S12.

Since the surface load from CO_2 deposition increases when atmospheric pressure decreases, we use a kernel function based on the negative of the pressure change. The time derivative of $dP(L_S(t))/dt$ also involves the time derivative of $L_S(t)$ (S17). The kernel function is

$$f(t) = A(-dP(L_S(JD_{TT} - D))/dt) + K \tag{6}$$

with parameters analogous to the illumination model.

4.7. Annual solar tide

Seismic activity of the Moon is mainly due to Earth's tides (Bulow et al., 2007). Earthquakes triggered by lunar tides have been reported (e.g. Rydelek et al., 1988), but are overwhelmed by plate tectonic background activity: only about one percent of midcrustal seismicity is related to tides (Lockner and Beeler, 1999; Métivier et al., 2009). We are not aware of any study of earthquakes related to solar tides, but as these are weaker than lunar tides, earthquakes due to solar tides will be even more difficult to identify.

We ruled out diurnal Phobos tides as reason for HF marsquakes above, while the daily wind speed pattern makes solar diurnal tides inaccessible for our analysis. Phillips (1991) concluded that the annual solar tide does not contribute to Martian seismicity,

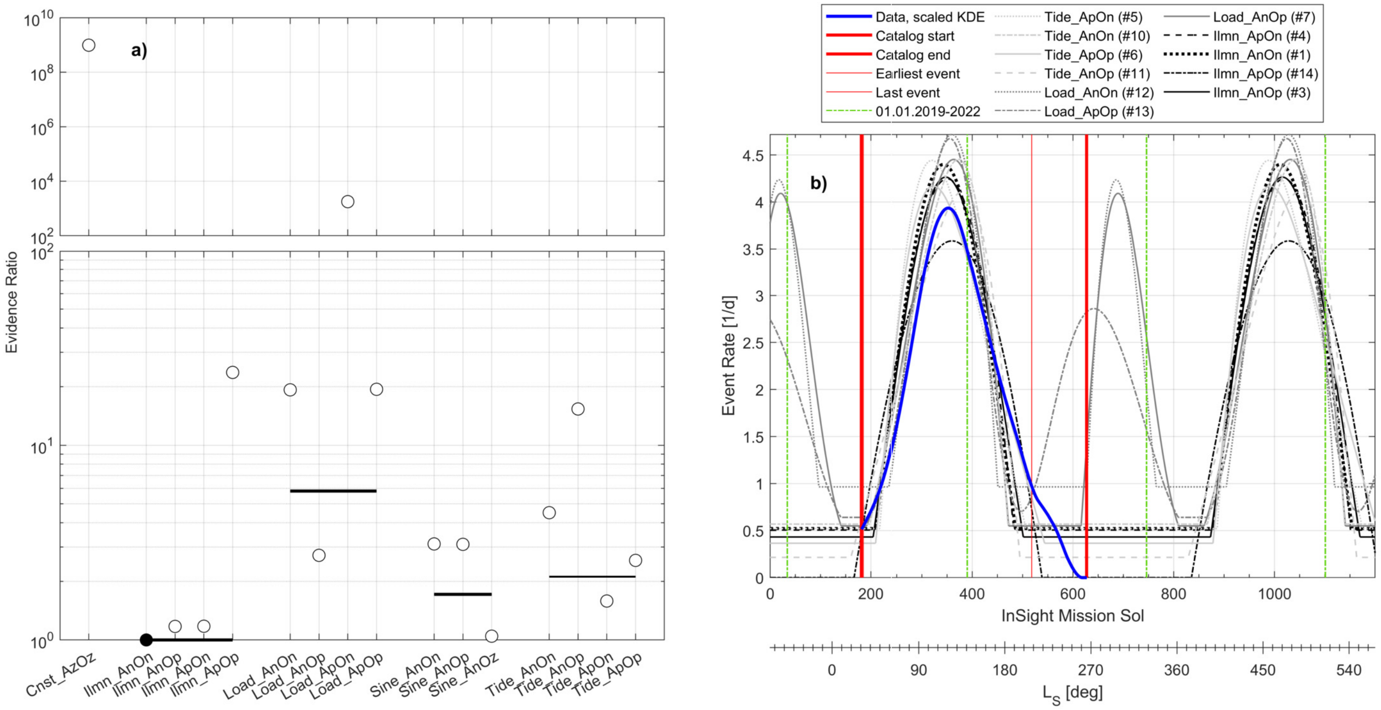


Fig. 6. Evidence ratios and rate forecasts. a) Evidence ratios for individual models and model groups (numerical values in table ST17). Filled circle: most supported model, open circles: evidence ratios with respect to the most supported model ($\mathcal{R}_{best} = 1$), black bars: evidence ratios of model groups with respect to the most supported group. A marker at a ratio $\mathcal{R}$ means that the model or group at $\mathcal{R}_{best} = 1$ is $\mathcal{R}$ times more likely to be the truly best model or group than the one with $\mathcal{R}_i = \mathcal{R}$. The lower part of the diagram is a continuation of the upper part, with higher vertical resolution. b) Rate forecast until sol 1200 (12. Apr. 2022) for models with $\mathcal{R}_i < 100$ (excluding Sine models, numbers in legend indicate evidence ranking). Vertical lines mark the time covered by observations (solid; catalog start and 1st event coincide) and the terrestrial calendar year (dash-dot lines for 1. Jan, starting 2019). A kernel density estimation (heavy solid curve with a bump between sols 500 and 600) illustrates the observed event rate but is not used during the inversion. The secondary x axis gives the areocentric solar longitude L_S corresponding to the mission sols, see Fig. 2b for the relation to martian seasons.

since absolute strain is low, although strain rates are high, compared to other sources of deformation. In contrast, Manga et al. (2019) conclude that an annual variation of seismic activity due to the ellipticity of the Martian orbit is well possible. With the observations of InSight, it now becomes possible to challenge both views.

Although Manga et al. (2019) provide a detailed framework for the modification of seismicity due to volumetric strain and resulting changes of pore pressure in the crust, we continue to use our rate function approach. An evaluation of deformation and stress tensors, and the dependency of phase lag on fault orientation, would require more knowledge of the seismic sources and their location than we currently have. Instead, we define a time dependency in terms of the radial tidal displacement. The resulting radial strain rate is proportional to $\dot{R}(t)/R^4(t)$, where $R(t)$ is the heliocentric distance of Mars (S20). We use a truncated series approximation of this distance and its derivative (S20) to capture the time dependency, but leave aside scaling parameters like masses and Newton's constant of gravity when we define the kernel function

$$f(t) = A\dot{R}(JD_{TT} - D)/\left(R_0^3 R^4(JD_{TT} - D)\right) + K \tag{7}$$

with parameters analogous to the illumination and surface load models, and R_0 being the mean distance between Mars and Sun in astronomical units.

5. Solutions and forecasts

From the model types described above, we define a total of 16 distinct models, which either use different kernel functions ("Cnst", "Ilmn", "Load", "Sine", or "Tide"), or are restricted to certain parameter ranges. After test runs, we decided to investigate positive and negative values of amplitude factors and constant kernel offsets separately, to avoid convergence to local maxima. We append an "A" and an "O", followed by "n", "p", or "z", respectively, to indicate if amplitude and offset are negative, positive, or zero.

With a collection of models that imply different physical effects at the source of the HF events, we not only have to ask which values of the model parameters are most likely, but which model is the most likely one, in the light of stochastic point process data. An appropriate framework is the concept of Kullback-Leibler information (Kullback and Leibler, 1951), which allows asking how much information is lost when the data is described by a model instead of the unknown real processes. From the maximum likelihood solution, we compute the Akaike Information Criterion with bias correction (AICc) and use Akaike weights and evidence ratios $\mathcal{R}_i$ between models (S13, S14).

The evidence ratios account for the sampling uncertainty of the computed AICc values (S11, ST16, ST17) and indicate how much more likely the preferred model is actually the best choice, rather than just appearing so by chance due to the limited data. Since our models form groups according to the rate kernels used, we also compute evidence ratios for these groups (Fig. 6a). For comparison, Fig. 6b also shows a kernel density estimation (KDE) of the observed rate function.

It is not surprising that the evidence speaks clearly (i.e. with a high evidence ratio) against the constant-rate model Cnst_AzOz, since the obvious deviation from a constant rate was what triggered our study, demonstrated above.

The preferred model is Ilmn_AnOn. The second-best model is Sine_AnOz. The group of Ilmn-models is actually the best group, while the group of Sine models comes second.

The Sine models, however, have no physical background but were introduced only to test the plausibility of an annual periodicity. Inspection of the coefficients in the series expansion of $L_S(t)$ (ST6) shows that $L_S(t)$ is dominated by the $i = 0$ term, i.e. $L_S(t) \propto t$ can be considered as first order approximation, and hence the Sine models perform so well because they are a good approximation to the Illumination model. The Sine models are thus redundant with the Illumination models and are not considered any further. This leaves us with the annual solar tide models as the second-best model group, while the surface load models come third. The best individual tide model Tide_ApOn is also the 2$^{\text{nd}}$ best physical model in total. Only three of the surface load models are supported by the data: the variants with negative offset are ruled out by evidence ratios exceeding $\mathcal{R} > 10^3$, compared to $\mathcal{R} < 20$ for positive offsets.

We examine closer models with evidence ratio below 100, i.e. the 11 best models (Fig. 6b).

Their effective amplitude at the time of the activity maximum is 3.5 to 4.5 events per day, about the observed rate after correction for detection efficiency. The baseline rate is about 0.5 events/day in most cases, and thus considerably below the maximum rate.

The lags of the models vary largely: event activity lags behind the driving function by 393 d for #1 ranked model Ilmn_AnOn, but only by 44 d for Ilmn_ApOn. If the illumination affects seismic activity by heating the ground, even the higher value points to a very shallow source, as the annual heat wave is expected to penetrate only a few meters (Siegler et al., 2017). More indirect triggering mechanisms, e.g. via a pore pressure reduction after a fluid release in a gully, might have an effect to greater depth, similar to the rain at Mt. Hochstaufen (Hainzl et al., 2006). The tide models show lags of either about 280 d or about 575 d.

The load model lags are either around 16 d or around 500 d: For the best load model, Load_AnOp, the time lag is 18.4 d (17.9 sols). This means that the seismicity maximum of sol 352 (22. Nov. 2019, $L_S = 110°$, see Fig. 6b) is the response to the pressure change in early November 2019 ($L_S = 102°$), which is the time of the largest pressure decrease (SF11). With a negative amplitude factor (ST15) this corresponds to the fastest ice deposition (on global average). Likewise, a lag time of about 494 or 507 d (Load_ApOn, Load_ApOp) points to late June or early July 2018 ($L_S \approx 200°$), when pressure increase was near maximum. With a positive amplitude factor, this corresponds to an increasing ice load as well.

It might be useful to consider periods longer than half a year as negative lags in the case of tidal models, and connect them to tidal dissipation models to further down-select the models.

We forecast the HF event rate for two Martian years following deployment completion (Fig. 6b). All models agree that an activity phase comparable to that observed so far will repeat in basically identical form from April 2021 (sol 840 in Fig. 6b) to March or April 2022 (sol 1150 and later), with an activity maximum between September and November 2021 (approx. sol 990 to 1050, precise dates in ST18). Two surface load models in this selection however forecast an earlier resume of activity, starting already from early spring 2020 (sol 496 for model Load_ApOp), with an activity maximum in fall 2020 (between sol 641 for Load_ApOp, and sol 689 for Load_AnOp), and lasting until March 2021 (around sol 810).

An interesting detail is the slight increase of the kernel density rate estimation after Sol 528 (21. May 2020, KDE in Fig. 6b), which approximately coincides with the second activity maximum predicted by the Load_ApOp model. Load_ApOp is the only model showing an event rate above its baseline rate at this time (see also SF15). In May 2020 ($L_S = 205°$), the noise level was already increased (Fig. 2) and the detection efficiency reduced (Fig. 4), thus only a small number of detections could be expected irrespective of the driving force. This result is in contradiction with the expected absence of CO_2 ice in the vicinity of InSight, but, with the small number of involved events, it does also not prove much. An unambiguous (non-) observation of the additional activity maxima predicted by the ice load models would be strong evidence for (or against) them. Among the models with evidence ratio below 100, Load_ApOp is however the second worst (evidence ratio 19.3). Without more data from this part of the activity cycle (if it is a cycle) it is difficult to tell how important this detail in the KDE rate estimation really is. The differences in the times predicted above only account for the differences between the model types, but not for the uncertainty of model parameters. The time intervals given so far are therefore likely to underestimate the actual uncertainty of begin, maximum, and end times. Unfortunately, the determination of parameter uncertainties using the Fisher information matrix failed for the preferred model, and also for several other models, such that we have to resort to the Jackknife approach (results in ST16).

For Ilmn_AnOn we find a standard deviation of the likelihood $\sigma_{MLE} = 6.36$ ($\log L_{MLE} = -213.57$, ST16, ST17). To estimate the uncertainty of the forecasts made above, we consider all parameters with likelihoods less than one standard deviation below the maximum. For these models, the beginning of the next activity phase varies from sol 853 (20. Apr. 2021) to sol 918 (26. June 2021), and the time of the next maximum is between sol 997 (15. Sep. 2021) and sol 1045 (03. Nov. 2021) (SF16). The expected number of events is as in the previous maximum. This timing uncertainty basically covers the variability between the individual models shown in Fig. 6b (except for the additional maxima of the load models).

For model Load_ApOp ($\log L_{MLE} = -216.7$), we obtain $\sigma_{MLE} = 6.19$ and again inspect models with likelihoods less than one standard deviation below the maximum (SF17). We find two distinct solutions for the timing of the additional activity phase, with maxima between sol 608 (11. Aug. 2020) and sol 738 (23. Dec. 2020), and a third maximum between sol 976 (25. Aug. 2021) and sol 1040 (28. Oct. 2021) and event rates similar to those experienced so far. In most cases, however, the rate never decreases to 0, with a minimum rate between sol 828 (25. Mar. 2021) and sol 880 (18. May 2021) (SF17). The expected event rate for the December 2020 maximum is about half the previous maximum. In general, these models alternate between two maximum rates, i.e. a higher maximum is followed by a lower one and vice versa.

For model Load_AnOp ($\log L_{MLE} = -214.6$) we obtain $\sigma_{MLE} = 6.32$, models within one standard deviation of the maximum likelihood solution are collected in figure SF19. Like Load_ApOp, this model also forecasts two phases of activity per year. One maximum is expected between sol 668 (12. Oct. 2020) and sol 719 (03. Dec. 2020). A phase with minimum or constant background activity might follow between sol 752 (06. Jan. 2021) and sol 862 (29. Apr. 2021), and another activity maximum is forecast to occur between sol 994 (12. Sep. 2021) and sol 1062 (21. Nov. 2021). The event rate at both the maxima is expected to be similar to the previous maximum.

6. Conclusions

We analyzed the High Frequency family of events detected by InSight on Mars, while we left out the Very High Frequency, Low Frequency and Broadband events which all appear to be of a different origin (Giardini et al., 2020; Clinton et al., 2021; van Driel et al., 2021).

We tested time variable event rate functions based on five different driving effects. Two of them (Phobos tides and impacts) were excluded from further analysis in an early stage, for the other three models (elevation of the Sun, seasonal surface loads, and solar tides) we conducted a maximum likelihood analysis and provide an Akaike weight-based ranking.

The evidence ratios between our three groups of models, and the 16 individual models within those groups, are not fully conclusive yet. One challenge is the coincidental proximity of the solstices to aphelion and perihelion. A higher resolution of activity levels and durations is required in order to discern tidal and climate forcing. This may be achieved in the future by stacking activity observations from more than one Martian year. Any of the models may still fail when future observations are compared to its activity forecasts, hence an extended observation would help discriminate between models. At least observing a second cycle would be necessary to rule out that we just witnessed a singular, nonrecurring burst of activity.

The best ranked rate model is the seasonal change of illumination, followed by annual solar tides and changing CO_2 ice load, in that order. There are indications of an additional activity maximum which could only be explained by the ice load model, though.

The physical mechanism for how illumination is connected to the HF marsquake sources remains to be determined. While thermal shocking is a known agent of weathering (e.g. Hall and Thorn, 2014), the penetration depth of the diurnal heat wave, and the amount of material mobilized in each cycle, are small (e.g. Lamp et al., 2017). In addition, landslides are inefficient seismic sources, which require enormous amounts of mass moved to obtain a seismic signal that can be recorded over long distances (e.g. Berrocal et al., 1978; Deparis et al., 2008), especially when accounting for the Martian environment (Lucas et al., 2019). Assuming falling boulders in Cerberus Fossae as source, for example, would imply erosion rates which probably do not agree with the crisp topography of its deep, narrow valleys. Recent avalanches on the northern polar cap, captured by HiRise on 29. May 2019 (Image ESP_060176_2640), were not detected by InSight.

Gullies are known to show a seasonal activity as well (Dundas et al., 2012), but are found only at latitudes beyond 30°, some however within less than 35° distance from InSight (Conway et al., 2019a,b). If gully activity is due to more or less deep aquifers (as suggested by Heldmann et al., 2007), aquifer pore pressure might be a crucial parameter. It is however unclear if this is the case, or if gaseous CO_2 plays a major role, while the absence of CO_2 in terrestrial analogs limits our understanding (see recent reviews of Conway et al., 2019a,b, and Dundas et al., 2019).

The CO_2 ice load model, as used here, reflects global-scale deposition and evaporation, but is mainly driven by the polar ice caps. These, however, are farther away from InSight than travel time analysis of the HF events allows the source to be. The assumption of a circumpolar origin would require very high seismic velocities even in the upper mantle and would be more difficult to reconcile with the coda present in HF seismograms, which suggest wave propagation dominantly through the crust (Giardini et al., 2020; van Driel et al., 2021). The lower latitude CO_2 deposits mentioned before are smaller in volume, and may condensate and evaporate with shorter time constants than the larger polar deposits. Also, especially the evaporation is certainly not independent from insolation, and the respective rate models are not as independent of each other as one would like them to be for a clear discrimination. Improved time dependency models for low-latitude ice deposits may help to disentangle the rate models.

We have shown that the annual variation of wind noise is not sufficient to explain the annual variability of HF event detections - they really occur at a variable rate. Our study suggests that, once the epicenters of HF sources can be better constrained, we should inspect regions where topography causes strong seasonal insolation variations, or where CO_2 ice is deposited in the cold times of the year, for indications of macroseismic changes due to the "marsquake season".

CRediT authorship contribution statement

M. Knapmeyer: Conceptualization, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **S.C. Stähler:** Conceptualization, Data curation, Software, Visualization, Writing – review & editing. **I. Daubar:** Writing – original draft. **F. Forget:** Data curation, Formal analysis, Methodology, Writing – original draft. **A. Spiga:** Data curation, Formal analysis, Methodology, Writing – original draft. **T. Pierron:** Formal analysis, Methodology, Writing – original draft. **M. van Driel:** Conceptualization, Data curation. **D. Banfield:** Data curation. **E. Hauber:** Conceptualization. **M. Grott:** Conceptualization. **N. Müller:** Conceptualization. **C. Perrin:** Conceptualization, Data curation. **A. Jacob:** Conceptualization. **A. Lucas:** Conceptualization. **B. Knapmeyer-Endrun:** Conceptualization, Data curation. **C. Newman:** Conceptualization. **M.P. Panning:** Funding acquisition, Project administration, Writing – review & editing. **R.C. Weber:** Conceptualization, Writing – review & editing. **F.J. Calef:** Writing – original draft. **M. Böse:** Data curation, Formal analysis, Investigation. **S. Ceylan:** Data curation, Formal analysis, Investigation. **C. Charalambous:** Data curation, Formal analysis, Investigation. **J. Clinton:** Data curation, Formal analysis, Investigation. **N. Dahmen:** Data curation, Formal analysis, Investigation. **D. Giardini:** Funding acquisition, Project administration. **A. Horleston:** Data curation, Formal analysis, Investigation. **T. Kawamura:** Data curation, Formal analysis, Investigation. **A. Khan:** Data curation, Formal analysis, Investigation. **G. Mainsant:** Data curation, Formal analysis, Investigation. **M. Plasman:** Data curation, Formal analysis, Investigation. **M. Lemmon:** Writing – review & editing. **R. Lorenz:** Writing – review & editing. **W.T. Pike:** Funding acquisition, Project administration. **J.-R. Scholz:** Data curation, Formal analysis, Investigation. **P. Lognonné:** Data curation, Funding acquisition, Investigation, Project administration. **B. Banerdt:** Funding acquisition, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All SEIS data are available in SEED format (InSight Mars SEIS Service, 2019a) or PDS4 format (InSight Mars SEIS Service, 2019b). Events used in this analysis are listed in the MQS catalogue (InSight Marsquake Service, 2020, 2021). Noise RMS amplitudes and event catalog are based on waveform data which is released to the public via the NASA Planetary Data System (https://pds-geosciences.wustl.edu/missions/insight/seis.htm), IPGP Data center (https://www.seis-insight.eu/en/science/seis-data/seis-data-description) and IRIS DMC (https://www.iris.edu/hq/sis/insight) several times a year.

Data from the APSS pressure sensor and the temperature and wind (TWINS) sensor referenced in this paper are available from the PDS Atmospheres node. The direct link to the InSight data archive at the PDS Atmospheres node is https://atmos.nmsu.edu/data_and_services/atmospheres_data/INSIGHT/insight.html.

Images from the Mars Reconnaissance Orbiter camera HiRise are stored in the NASA Planetary Data System at https://pds-imaging.jpl.nasa.gov/.

Acknowledgements

We are grateful to Peter Shearer and an anonymous reviewer, whose comments helped to improve the manuscript.

We acknowledge NASA, CNES, partner agencies and Institutions (UKSA, SSO, DLR, JPL, IPGP-CNRS, ETHZ, IC, MPS-MPG) and the operators of JPL, SISMOC, MSDS, IRIS-DMC and PDS for providing SEED SEIS data.

I.J.D. is supported by NASA InSight Participating Scientist grant 80NSSC20K0971.

This is Insight contribution ICN185.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.epsl.2021.117171.

References

Banerdt, W.B., Smrekar, S.E., Banfield, D., Giardini, D., Golombek, M., Johnson, C.L., Lognonné, P., Spiga, A., Spohn, T., Perrin, C., Stähler, S.C., Antonangeli, D., Asmar, S., Beghein, C., Bowles, N., Bozdag, N., Chi, P., Christensen, U., Clinton, J., Collins, G.S., Daubar, I., Dehant, V., Drilleau, M., Fillingim, M., Folkner, W., Garcia, R.F., Garvin, J., Grant, J., Grott, M., Grygorczuk, J., Hudson, T., Irving, J.C.E., Kargl, G., Kawamura, T., Kedar, S., King, S., Knapmeyer-Endrun, B., Knapmeyer, M., Lemmon, M., Lorenz, R., Maki, J.N., Margerin, L., McLennan, S.M., Michaut, C., Mimoun, D., Mittelholz, A., Mocquet, A., Morgan, P., Mueller, N.T., Murdoch, N., Nagihara, S., Newman, C., Nimmo, F., Panning, M., Pike, W.T., Plesa, A.-C., Rodriguez, S., Rodriguez-Manfredi, J.A., Russell, C.T., Schmerr, N., Siegler, M., Stanley, S., Stutzmann, E., Teanby, N., Tromp, J., van Driel, M., Warner, N., Weber, R., Wieczorek, M., 2020. Initial results from the InSight mission on Mars. Nat. Geosci. 13, 183–189. https://doi.org/10.1038/s41561-020-0544-y.

Banfield, D., Spiga, A., Newman, C., Forget, F., Lemmon, M., Lorenz, R., Murdoch, N., Viudez-Moreiras, D., Pla-Garcia, J., Garcia, R.F., Lognonné, P., Karatekin, Ö., Perrin, C., Martire, L., Teanby, N., van Hove, B., Maki, J., Kenda, B., Mueller, N.T., Rodriguez, S., Kawamura, T., McClean, J.B., Stott, A.E., Charalambous, C., Millour, E., Johnson, C.L., Mittelholz, A., Määttänen, A., Lewis, S.R., Clinton, J., Stähler, S.C., Ceylan, S., Giardini, D., Warren, T., Pike, T., Daubar, I., Golombek, M., Rolland, L., Widmer-Schnidrig, R., Mimoun, D., Beucler, É., Jacob, A., Lucas, A., Baker, M., Ansan, V., Hurst, K., Mora-Sotomayor, L., Navarro, S., Torres, J., Lepinette, A., Molina, A., Marin-Jimenez, M., Gomez-Elvira, J., Peinado, V., Rodriguez-Manfredi, J.-A., Carcich, B.T., Sackett, S., Russell, C.T., Spohn, T., Smrekar, S.E., Banerdt, W.B., 2020. The atmosphere of Mars as observed by InSight. Nat. Geosci. 13, 190–198. https://doi.org/10.1038/s41561-020-0534-0.

Bell, M.L., Nur, A., 1978. Strength changes due to reservoir-induced pore pressure and stresses and application to Lake Oroville. J. Geophys. Res. 83 (B9), 4469–4483.

Berrocal, J., Espinosa, A.F., Galdos, J., 1978. Seismological and geological aspects of the Mantaro landslide in Peru. Nature 275, 533–536.

Bollinger, L., Perrier, F., Avouac, J.-P., Sapkota, S., Gautam, U., Tiwari, D.R., 2006. Seasonal modulation of seismicity in the Himalaya of Nepal. Geophys. Res. Lett. 34, L08304. https://doi.org/10.1029/2006GL029192.

Bulow, R.C., Johnson, C.L., Bills, B.G., Shearer, P.M., 2007. Temporal and spatial properties of some deep moonquake clusters. J. Geophys. Res. 112, E09003. https://doi.org/10.1029/2006JE002847.

Clinton, J.F., Ceylan, S., van Driel, M., Giardini, D., Stähler, S.C., Böse, M., Charalambous, C., Dahmen, N.L., Horleston, A., Kawamura, T., Khan, A., Orhand-Mainsant, G., Scholz, J.-R., Euchner, F., Banerdt, W.B., Lognonné, P., Banfield, D., Beucler, E., Garcia, R.F., Kedar, S., Panning, M.P., Perrin, C., Pike, W.T., Smrekar, S.E., Spiga, A., Stott, A., 2021. The Marsquake catalogue from InSight, sols 0–478. Phys. Earth Planet. Inter. 310. https://doi.org/10.1016/j.pepi.2020.106595.

Conway, S., de Haas, T., Harrison, T., 2019a. Martian gullies: a comprehensive review of observations, mechanisms and insights from Earth analogues. In: Conway, S.J., Carrivick, J.L., Carling, P.A., de Haas, T., Harrison, T.N. (Eds.), Martian Gullies and Their Earth Analogues. Geol. Soc. (Lond.) Spec. Publ. 467, 7–66. https://doi.org/10.1144/SP467.14.

Conway, S., Harrison, T., Soare, R., Britton, A., Steele, L., 2019b. New slope-normalised global gully density and orientation maps for Mars. Geol. Soc. (Lond.) Spec. Publ. 467 (1), 187–197. https://doi.org/10.1144/SP467.3. Geological Society of London. hal-02270628ff.

Daubar, I.J., McEwen, A.S., Byrne, S., Kennedy, M.R., Ivanov, B.A., 2013. The current martian cratering rate. Icarus 225, 506–516. https://doi.org/10.1016/j.icarus.2013.04.009.

Daubar, I.J., Lognonné, P., Teanby, N.A., Collins, G.S., Clinton, J., Stähler, S., Spiga, A., Karakostas, F., Ceylan, S., Malin, M., McEwen, A.S., Maguire, R., Charalambous, C., Onodera, K., Lucas, A., Rolland, L., Vaubaillon, J., Kawamura, T., Böse, M., Horleston, A., van Driel, M., Stevanović, J., Miljković, K., Fernando, B., Huang, Q., Giardini, D., Larmat, C.S., Leng, K., Rajšić, A., Schmerr, N., Wójcicka, N., Pike, T., Wookey, J., Rodriguez, S., Garcia, R., Banks, M.E., Margerin, L., Posiolova, L., Banerdt, B., 2020. A new crater near InSight: implications for seismic impact detectability on Mars. J. Geophys. Res., Planets 125 (8). https://doi.org/10.1029/2020JE006382.

Deparis, J., Jongmans, D., Cotton, F., Baillet, L., Thouvenot, F., Hantz, D., 2008. Analysis of rock-fall and rock-fall avalanche seismograms in the French Alps. Bull. Seismol. Soc. Am. 98 (4), 1781–1796. https://doi.org/10.1785/0120070082.

Dundas, C.M., Diniega, S., Hansen, C.J., Byrne, S., McEwen, A.S., 2012. Seasonal activity and morphological changes in Martian gullies. Icarus 220, 124–143. https://doi.org/10.1016/j.icarus.2012.04.005.

Dundas, C.M., McEwen, A.S., Diniega, S., Hansen, C.J., Byrne, S., McElwaine, J.N., 2019. The formation of gullies on Mars today. In: Conway, S.J., Carrivick, J.L., Carling, P.A., de Haas, T., Harrison, T.N. (Eds.), Martian Gullies and Their Earth Analogues. Geol. Soc. (Lond.) Spec. Publ. 467, 67–94. https://doi.org/10.1144/SP467.5.

Giardini, D., Lognonné, P., Banerdt, W., Pike, W., Christensen, U., Ceylan, S., Clinton, J., van Driel, M., Stähler, S., Böse, M., Garcia, R.F., Khan, A., Panning, M., Perrin, C., Bandfield, D., Beucler, E., Charalambous, C., Euchner, F., Horleston, A., Jacob, A., Kawamura, T., Kedar, S., Mainsant, G., Scholz, J.-R., Smrekar, S., Spiga, A., Agard, C., Antonangeli, D., Barkaoui, S., Barrett, E., Combers, P., Conejero, V., Daubar, I., Drilleau, M., Ferrier, C., Gabsi, T., Gudkova, T., Hurst, K., Karakostas, F., King, S., Knapmeyer, M., Knapmeyer-Endrun, B., Llorca-Cejudo, R., Lucas, A., Luno, L., Margerin, L., McClean, J., Mimoun, D., Murdoch, N., Nimmo, F., Nonon, M., Pardo, C., Rivoldini, A., Rodriguez Manfredi, J.A., Samuel, H., Schimmel, M., Stott, A.E., Stutzmann, E., Teanby, N., Warren, T., Weber, R., Wieczorek, M., Yana, C., 2020. The seismicity of Mars. Nat. Geosci. 13, 205–212. https://doi.org/10.1038/s41561-020-0539-8.

Hainzl, S., Kraft, T., Wassermann, J., Igel, H., Schmedes, E., 2006. Evidence for rainfall-triggered earthquake activity. Geophys. Res. Lett. 33, L19303. https://doi.org/10.1029/2006GL027642.

Hall, K., Thorn, C.E., 2014. Thermal fatigue and thermal shock in bedrock: an attempt to unravel the geomorphic processes and products. Geomorphology 206, 1–13. https://doi.org/10.1016/j.geomorph.2013.09.022.

Heldmann, J.L., Carlsson, E., Johansson, H., Mellon, M.T., Toon, O.B., 2007. Observations of Martian gullies and constraints on potential formation mechanisms II: the northern hemisphere. Icarus 188, 324–344. https://doi.org/10.1016/j.icarus.2006.12.010.

Herschel, W., 1784. On the remarkable appearances at the polar regions of the planet Mars, and its spheroidical figure; with a few hints relating to its real diameter and atmosphere. Philos. Trans. R. Soc. 74, 233–273. https://doi.org/10.1098/rstl.1784.0020.

InSight Mars SEIS Data Service, 2019a. SEIS raw data, insight mission. IPGP, JPL, CNES, ETHZ, ICL, MPS, ISAE-Supaero, LPG, MFSC, https://doi.org/10.18715/SEIS.INSIGHT.XB_2016.

InSight Mars SEIS Data Service, 2019b. InSight SEIS data bundle. PDS Geosciences (GEO) Node, https://doi.org/10.17189/1517570.

InSight Marsquake Service, 2020. Mars seismic catalogue, InSight mission; V4 2020-10-01 (version 4.0). ETHZ, IPGP, JPL, ICL, MPS, Univ. Bristol, https://doi.org/10.12686/a9.

InSight Marsquake Service, 2021. Mars seismic catalogue, InSight mission; V5 2021-01-04 (version 5.0). ETHZ, IPGP, JPL, ICL, MPS, Univ. Bristol, https://doi.org/10.12686/a10.

Ivanov, B.A., 2001. Mars/Moon cratering rate ratio estimates. Space Sci. Rev. 96, 87–104. https://doi.org/10.1023/A:1011941121102.

JeongAhn, Y., Malhotra, R., 2015. The current impact flux on Mars and its seasonal variation. Icarus 262, 1–32. https://doi.org/10.1016/j.icarus.2015.08.032.

Kawamura, T., Lognonné, P., Nishikawa, Y., Tanaka, S., 2017. Evaluation of deep moonquake source parameters: implication for fault characteristics and thermal state. J. Geophys. Res. 122, 1487–1504. https://doi.org/10.1002/2016JE005147.

Knapmeyer, M., Oberst, J., Hauber, E., Wählisch, M., Deuchler, C., Wagner, R., 2006. Working models for spatial distribution and level of Mars' seismicity. J. Geophys. Res., Planets 111, E11006. https://doi.org/10.1029/2006JE002708.

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Ann. Math. Stat. 200, 79–86.

Lamp, J.L., Marchant, D.R., Mackay, S.L., Head, J.W., 2017. Thermal stress weathering and the spalling of Antarctic rocks. J. Geophys. Res., Earth Surf. 122, 3–24. https://doi.org/10.1002/2016JF003992.

Lockner, D.A., Beeler, N.M., 1999. Premonitory slip and tidal triggering of earthquakes. J. Geophys. Res. 104 (B9), 20133–201511. https://doi.org/10.1029/1999JB900205.

Lognonné, P., Banerdt, W.B., Giardini, D., Pike, W.T., Christensen, U., Laudet, P., de Raucourt, S., Zweifel, P., Calcutt, S., Bierwirth, M., Hurst, K.J., Ijpelaan, F., Umland, J.W., Llorca-Cejudo, R., Larson, S.A., Garcia, R.F., Kedar, S., Knapmeyer-Endrun, B., Mimoun, D., Mocquet, A., Panning, M.P., Weber, R.C., Sylvestre-Baron, A., Pont, G., Verdier, N., Kerjean, L., Facto, L.J., Gharakanian, V., Feldman, J.E., Hoffman, T.L., Klein, D.B., Klein, K., Onufer, N.P., Paredes-Garcia, J., Petkov, M.P., Willis, J.R., Smrekar, S.E., Drilleau, M., Gabsi, T., Nebut, T., Robert, O., Tillier, S., Moreau, C., Parise, M., Aveni, G., Ben Charef, S., Bennour, Y., Camus, T., Dandonneau, P.A., Desfoux, C., Lecomte, B., Pot, O., Revuz, P., Mance, D., tenPierick, J., Bowles, N.E., Charalambous, C., Delahunty, A.K., Hurley, J., Irshad, R., Liu, Huafeng, Mukherjee, A.G., Standley, I.M., Stott, A.E., Temple, J., Warren, T., Eberhardt, M., Kramer, A., Kühne, W., Miettinen, E.-P., Monecke, M., Aicardi, C., André, M., Baroukh, J., Borrien, A., Bouisset, A., Boutte, P., Brethomé, K., Brysbaert, C., Carlier, T., Deleuze, M., Desmarres, J.M., Dilhan, D., Doucet, C., Faye, D., Faye-Refalo, N., Gonzalez, R., Imbert, C., Larigauderie, C., Locatelli, E., Luno, L., Meyer, J.-R., Mialhe, F., Mouret, J.M., Nonon, M., Pahn, Y., Paillet, A., Pasquier, P., Perez, G., Perez, R., Perrin, L., Pouilloux, B., Rosak, A., Savin de Larclause, I., Sicre, J., Sodki, M., Toulemont, N., Vella, B., Yana, C., Alibay, F., Avalos, O.M., Balzer, M.A., Bhandari, P., Blanco, E., Bone, B.D., Bousman, J.C., Bruneau, P., Calef, F.J., Calvet, R.J., D'Agostino, S.A., de los Santos, G., Deen, R.G., Denise, R.W., Ervin, J., Ferraro, N.W., Gengl, H.E., Grinblat, F., Hernandez, D., Hetzel, M., Johnson, M.E., Khachikyan, L., Lin, J.Y., Madzunkov, S.M., Marshall, S.L., Mikellides, I.G., Miller, E.A., Raff, W., Singer, J.E., Sunday, C.M., Villalvazo, J.F., Wallace, M.C., Banfield, D., Rodriguez-Manfredi, J.A., Russell, C.T., Trebi-Ollennu, A., Maki, J.N., Beucler, E., Böse, M., Bonjour, C., Berenguer, J.L., Ceylan, S., Clinton, J., Conejero, V., Daubar, I., Dehant, V., Delage, P., Euchner, F., Estève, I., Fayon, L., Ferraioli, L., Johnson, C.L., Gagnepain-Beyneix, J., Golombek, M., Khan, A., Kawamura, T., Kenda, B., Labrot, P., Murdoch, N., Pardo, C., Perrin, C., Pou, L., Sauron, A., Savoie, D., Stähler, S., Stutzmann, E., Teanby, N.A., Tromp, J., van Driel, M., Wieczorek, M., Widmer-Schnidrig, R., Wookey, J., 2019. SEIS: insight's seismic experiment for internal structure of Mars. Space Sci. Rev. 215 (12). https://doi.org/10.1007/s11214-018-0574-6.

Lognonné, P., Banerdt, W.B., Pike, W.T., Giardini, D., Christensen, U., Garcia, R.F., Kawamura, T., Kedar, S., Knapmeyer-Endrun, B., Margerin, L., Nimmo, F., Panning, M., Tauzin, B., Scholz, J.-R., Antonangeli, D., Barkaoui, S., Beucler, E., Bissig, F., Brinkman, N., Calvet, M., Ceylan, S., Charalambous, C., Davis, P., van Driel, M., Drilleau, M., Fayon, L., Joshi, R., Kenda, B., Khan, A., Knapmeyer, M., Lekic, V., McClean, J., Mimoun, D., Murdoch, N., Pan, L., Perrin, C., Pinot, B., Pou, L., Menina, S., Rodriguez, S., Schmelzbach, C., Schmerr, N., Sollberger, D., Spiga, A., Stähler, S., Stott, A., Stutzmann, E., Tharimena, S., Widmer-Schnidrig, R., Andersson, F., Ansan, V., Beghein, C., Böse, M., Bozdag, E., Clinton, J., Daubar, I., Delage, P., Fuji, N., Golombek, M., Grott, M., Horleston, A., Hurst, K., Irving, J., Jacob, A., Knollenberg, J., Krasner, S., Krause, C., Lorenz, R., Michaut, C., Myhill, B., Nissen-Meyer, T., ten Pierick, J., Plesa, A.C., Quantin-Nataf, C., Robertsson, J., Rochas, L., Schimmel, M., Smrekar, S., Spohn, T., Teanby, N., Tromp, J., Vallade, J., Verdier, N., Vrettos, C., Weber, R., Banfield, D., Barrett, E., Bierwirth, M., Calcutt, S., Compaire, N., Johnson, C., Mance, D., Euchner, F., Kerjean, L., Mainsant, G., Mocquet, A., Antonio Rodriguez Manfredi, J., Pont, G., Laudet, P., Nebut, P., de Raucourt, S., Robert, O., Russel, C., Sylvestre-Baron, A., Tillier, S., Warren, T., Wieczorek, M., Yana, C., Zweifel, P., 2020. SEIS on Mars: first steps in comparative seismology between Mars, Earth and the Moon. Nat. Geosci. 13, 213–220. https://doi.org/10.1038/s41561-020-0536-y.

Lucas, A., Kenda, B., Mangeney, A., Kawamura, T., Daubar, I., Aharonson, O., Drilleau, M., Jacob, A., Hibert, C., Spiga, A., Rodriguez, S., Weber, R., Lognonné, P., 2019. Seismic detection of mass wasting on Mars with SEIS/InSight: a loony attempt? In: 50th Lunar and Planetary Science Conference. The Woodlands, TX, 18-22 March 2019. Abstract#2132.

Malin, M.C., Edgett, K.S., Posiolova, L.V., McColley, S.M., Noe Dobrea, E.Z., 2006. Present-day impact cratering rate and contemporary gully activity on Mars. Science 314, 1573–1577. https://doi.org/10.1126/science.1135156.

Manga, M., Zhai, G., Wang, C.-Y., 2019. Squeezing marsquakes out of groundwater. Geophys. Res. Lett. 46, 6333–6340. https://doi.org/10.1029/2019GL082892.

Martínez, G.M., Newman, C.N., De Vicente-Retortillo, A., Fischer, E., Renno, N.O., Richardson, M.I., Fairén, A.G., Genzer, M., Guzewich, S.D., Haberle, R.M., Harri, A.-M., Kemppinen, O., Lemmon, M.T., Smith, M.D., de la Torre-Juárez, M., Vasavada, A.R., 2017. The modern near-surface Martian climate: a review of in-situ meteoroligical data from Viking to Curiosity. Space Sci. Rev. 212, 295–338. https://doi.org/10.1007/s11214-017-0360-x.

Métivier, L., de Viron, O., Conrad, C.P., Renault, S., Diament, M., Patau, G., 2009. Evidence of earthquake triggering by the solid Earth tides. Earth Planet. Sci. Lett. 278, 370–375. https://doi.org/10.1016/j.epsl.2008.12.024.

Nakamura, Y., 2003. New identification of deep moonquakes in the apollo lunar seismic data. Phys. Earth Planet. Inter. 139, 197–205. https://doi.org/10.1016/j.pepi.2003.07.017.

Oberst, J., Nakamura, Y., 1987. Distinct meteoroid families identified on the lunar seismograms. J. Geophys. Res., Solid Earth 92 (B4), E769–E773. https://doi.org/10.1029/JB092iB04p0E769.

Ogata, Y., 1983. Estimation of the parameters in the modified Omori formula for aftershock frequencies by the Maximum Likelihood procedure. J. Phys. Earth 31, 115–124.

Phillips, R.J., 1991. Expected rates of marsquakes. In: Scientific Rationale and Requirements for a Global Seismic Network on Mars, LPI Tech. Rep. 91–02 LPI/TR-91–02. Lunar and Planet. Inst., Houston, Tex, pp. 35–38.

Plesa, A.-C., Knapmeyer, M., Golombek, M.P., Breuer, D., Grott, M., Kawamura, T., Lognonné, P., Weber, R.C., 2018. Present-day Mars' seismicity predicted from 3-D thermal evolution models of interior dynamics. Geophys. Res. Lett. 45. https://doi.org/10.1002/2017GL076124.

Roeloffs, E.A., 1988. Fault stability changes induced beneath a reservoir with cyclic variations in water level. J. Geophys. Res. 93 (B3), 2107–2124.

Rydelek, P.A., Davis, P.M., Koyanagi, R.Y., 1988. Tidal triggering of earthquake swarms at Kilauea Volcano, Hawaii. J. Geophys. Res. 93 (B5), 4401–4411. https://doi.org/10.1029/JB093iB05p04401.

Saxena, S.K., Metin Ger, A., Sengupta, A., 1988. Reservoir induced seismicity - a new model. Int. J. Numer. Anal. Methods Geomech. 12, 263–281.

Siegler, M.A., Smrekar, S.E., Grott, M., Piqueux, S., Mueller, N., Williams, J.-P., Plesa, A.-C., Spohn, T., 2017. The InSight Mars lander and its effect on the subsurface thermal environment. Space Sci. Rev. 211. https://doi.org/10.1007/s11214-017-0331-2.

Taylor, J., Teanby, N.A., Wookey, J., 2013. Estimates of seismic activity in the Cerberus Fossae region of Mars. J. Geophys. Res., Planets 118, 2570–2581. https://doi.org/10.1002/2013JE004469.

van Driel, M., Ceylan, S., Clinton, J.F., Giardini, D., Horleston, A., Margerin, L., Stähler, S.C., Böse, M., Charalambous, C., Kawamura, T., Khan, A., Orhand-Mainsant, G., Scholz, J.-R., Euchner, F., Knapmeyer, M., Schmerr, N., Pike, W.T., Lognonné, P., Banerdt, W.B., 2021. High frequency seismic events on Mars observed by InSight. J. Geophys. Res., Planets 126 (2). https://doi.org/10.1029/2020JE006670.

Supplementary Information

Seasonal Seismic Activity on Mars

M. Knapmeyer[a], S. C. Stähler[b], I. Daubar[c], F. Forget[d], A. Spiga[d], T. Pierron[d], M. van Driel[b], D. Banfield[e], E. Hauber[a], M. Grott[a], N. Müller[a], C. Perrin[p], A. Jacob[f], A. Lucas[f], B. Knapmeyer-Endrun[g], C. Newman[h], M. P. Panning[i], R. C. Weber[j], F. J. Calef[i], M. Böse[b,r], S. Ceylan[b], C. Charalambous[k], J. Clinton[r], N. Dahmen[b], D. Giardini[b], A. Horleston[l], T. Kawamura[f], A. Khan[b], G. Mainsant[q], M. Plasman[f], M. Lemmon[n], R. Lorenz[o], W. T. Pike[k], J.-R. Scholz[m], P. Lognonné[f], B. Banerdt[i]

[a] *Institute of Planetary Research, DLR, Rutherfordstr. 2, 12489 Berlin, Germany.*
[b] *Institute of Geophysics, ETH Zürich, Sonneggstrasse 5, 8092 Zürich, Switzerland*
[c] *Department of Earth, Environmental, and Planetary Sciences, Brown University, Campus Box 1846, Providence, RI 02912-1846, USA*
[d] *Laboratoire de Météorologie Dynamique / Institut Pierre Simon Laplace (LMD/IPSL), Sorbonne Université, Centre National de la Recherche Scientifique (CNRS), École Polytechnique, École Normale Supérieure*
[e] *Cornell University, Cornell Center for Astrophysics and Planetary Science, Ithaca, NY, 14853, USA*
[f] *Université de Paris, Institut de physique du globe de Paris, CNRS, F-75005 Paris, France*
[g] *Bensberg Observatory, University of Cologne, Vinzenz-Pallotti-Str. 26, 51429 Bergisch Gladbach, Germany*
[h] *Aeolis Research, 333 N Dobson Road, Unit 5, Chandler AZ 85224-4412, USA*
[i] *Jet Propulsion Laboratory, California Institute of Technology; 4800 Oak Grove Dr., M/S 183-301, Pasadena, CA 91109, USA*
[j] *NASA MSFC, NSSTC Mail Code ST13, 320 Sparkman Drive, Huntsville, AL 35805, USA*
[k] *Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom*
[l] *School of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, UK*
[m] *Max Planck Institute for Solar System Research, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany*
[n] *Space Science Institute, 4765 Walnut Street, Suite B, Boulder, CO 80301, USA*
[o] *Johns Hopkins Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723, USA*
[p] *Laboratoire de Planétologie et Géodynamique, UMR6112, OSUNA UMS3271, Univ. Nantes, Univ. Angers, CNRS, 2 rue de la Houssinière, BP 92208, 44322 Nantes Cedex 3, France*
[q] *Institut Supérieur de l'Aéronautique et de l'Espace SUPAERO, 10 Avenue Edouard Belin, 31400 Toulouse, France*
[r] *Swiss Seismological Service (SED), ETH Zurich, Sonneggstr. 5, 8092 Zurich, Switzerland*

Corresponding Author: M. Knapmeyer, martin.knapmeyer@dlr.de

Contents

Lists of Figures and Tables

Figures

Tables

Introduction

This document contains technical details concerning the methods and calculations used throughout the main paper. Although we tried to put the individual sections into an order that makes some sense in itself, they are not intended as a continuous read.

The times for which data is available are summarized in section S1, considering also the available sampling rates. The routine data processing is summarized in S2. Section S3 lists the event selection used in the analysis.

The detection rate variability due to noise variability is discussed in S4, and our approach to detect a variable occurrence in the presence of a variable detection rate is described in S5. Section S6 lists polynomial coefficients for an approximation of the detection efficiency.

In section S7 we summarize a few properties of Poisson processes that are important for our study. Section S8 describes the Spike Train Fourier Transform, which allows detecting periodicities in the occurrence times in a series of discrete events, rather than in continuous waveforms.

Sections S9 through S12 deal with the computation of the log-likelihood of a candidate event rate model, given an event sequence, the sampling uncertainty of the likelihood, and the uncertainty of the parameters defined by the maximum likelihood solution.

Sections S13 and S14 describe the corrected Akaike Information Criterion used to rank models, and the derived Akaike weights and evidence ratios, which provide a quantitative classification of the ranking across different types of models.

Sections S15, S16, S17, S19, and S20 provide series expansions for astronomical parameters that are used in the different rate models: the position of the Sun in the Martian sky (S15), the areocentric longitude used to describe the progress of the seasons (S16) and its time derivative (S17), the distance of Mars from the sun and its time derivative (S20), and the resulting tidal strain rate (S19).

A series expansion for the atmospheric pressure at the InSight landing site, and its time derivative, is given in section S18.

Sections S21 and S22 contain tables with the defining parameters of initial and final search grids, the resulting maximum likelihood solutions, and the results of the evaluation of Akaike Information Criterion, Akaike weights, and evidence ratios.

Finally, section S23 contains tabular details and graphical representations of event rate forecasts.

S1 Data availability

Routine evaluation of incoming data, including the detection and classification of events, is conducted by the Marsquake service (Clinton et al., 2018, 2021) mainly on the basis of spectrograms of raw data in U-V-W sensor coordinates of the VBB instrument (SEED channels BHU, BHV, and BHW), and rotated into planetary Z-N-E coordinates (see Lognonné et al., 2019, for an overview of all SEED data channels). Experience shows that most events can be identified on the vertical VBB channel (termed BHZ in the following, Clinton et al. 2021). In the earliest months of the registration, a composite channel (SEISVELZ, 58.BZC, see Lognonné et al., 2019) which merges all six seismic sensor channels (computed on board the lander) with an effective sampling rate of 10 Hz, was used as well at times when the other channels could not be downlinked. Downlink of the SEISVELZ channel was however discontinued in summer 2019. This and other milestone dates concerning the operation of the SEIS experiment are summarized in table ST1.

Since the BHZ component can be computed only when all three of BHU, BHV, and BHW are available, and only at the sampling rate of the raw channel with the lowest rate, the uptime intervals relevant for event detection capabilities are a combination of the times at which BHU, BHV, BHW, and BZC are available, and it has to be considered that the available sampling rate might change with time.

For the purpose of the likelihood evaluation conducted in our work, we consider as "uptime" of SEIS all times at which either the SEISVELZ channel was available, or all three of BHU, BHV, and BHW, i.e. times at which the vertical component could be computed.

For the sake of computational efficiency (as CPU time needed for the likelihood evaluation depends linearly on the number of availability intervals), we concatenate all availability intervals that are separated by gaps less than 5 min in duration, which is shorter than the shortest catalogued event (the shortest HF event is S0221c, with a duration of about 8 minutes, the shortest event we actually use is S0518a with a duration of 10 min). This reduces the number of gaps to consider from 158 to 86, while the overall duration of all tabulated gaps is reduced from 42237 min to 42083 min, i.e. by concatenation of intervals we overlook a total of only 154 minutes of data gaps. This has to be compared to the overall duration of the evaluated catalog (659520 minutes): assuming that events are due to a stationary Poisson process, 154 min of additional data would increase the expected number of events by a factor of $659520/(659520-154)=1.00023$. With a total number of 227 High Frequency events, ignoring the short data gaps can thus not have a significant effect on the occurrence statistics.

The resulting consolidated availability times are listed in table ST2.

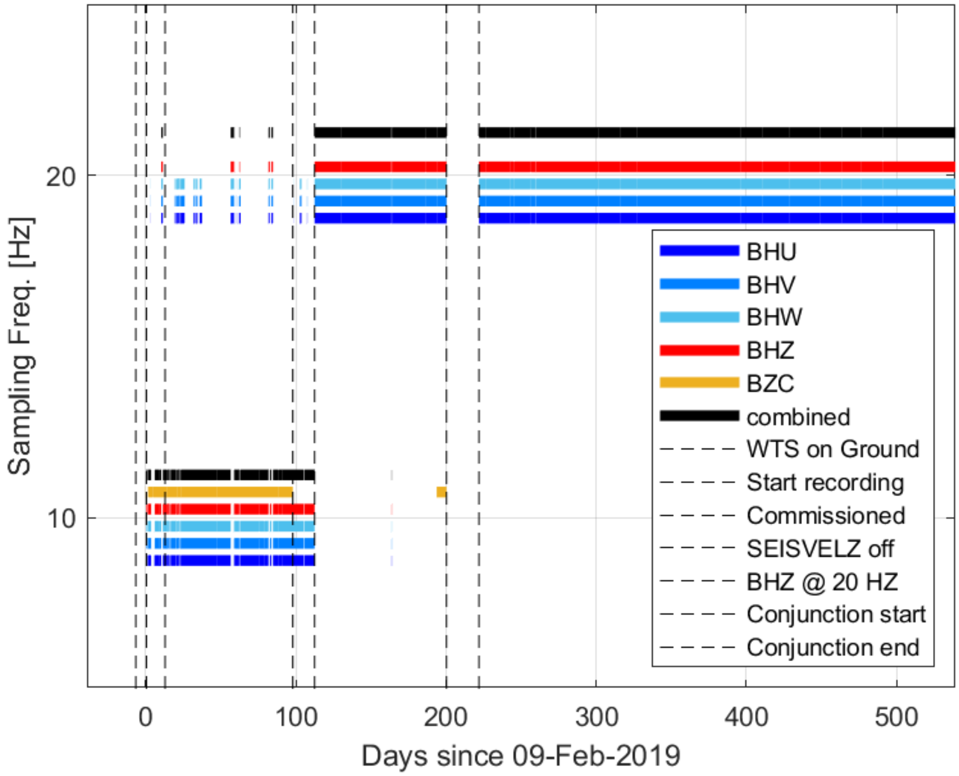


Figure SF1 SEIS data availability overview.

Horizontal lines indicate the availability, and sampling frequency, of SEIS data channels following the naming conventions described by Lognonné et al., 2019: BHU, BHV, BHW are the output of the individual SEIS sensors in oblique orientation, BHZ is the vertical component, BZC is the SEISVELZ channel, "combined" is the combined availability of BHU, BHV, and BHW (see text), i.e. shows times for which all three channels are available. All channels of the lower group are sampled at 10 *Hz*, all channels of the upper group are sampled at 20 *Hz*, they are offset only for clarity. Milestones of ST1 are indicated by vertical dashed lines, the corresponding list in the legend is in chronological order.

Table ST1 Milestones affecting data availability

Dates are given in three formats: as UTC time, number of days since the start of continuous recording, and InSight mission sol number including LMST. References: Instrument Context Camera (ICC) images are available at the InSight Analyst's Notebook (2020), "miniseed" refers to dates extracted from miniseed data files.

Milestone	UTC	Days since 09.02.2019	Sol	LMST	Reference
InSight Landing	26. Nov. 2018 19:44:52	-75	0	05:06:45	Abilleira et al., 2019
SEIS on ground	19. Dec. 2018 14:57	-52	22	18:44	ICC image, Analyst's Notebook
WTS on ground	02. Feb. 2019 11:19	-7	66	10:19	ICC image, Analyst's Notebook
continuous recording start	09. Feb. 2019 14:02	0	73	23:23	miniseed
full performance	21. Feb. 2019	12	85		Banerdt et al., 2020
SEISVELZ discontinued	18. May 2019 00:07	98	168	18:55	miniseed
BHU, BHV, BHW @ 20 Hz	01. June 2019 12:24:30	112	182	21:53	miniseed
Conjunction downtime start	28. Aug. 2019 08:04	200	267	09:09	miniseed
Conjunction downtime end	18. Sep. 2019 23:09	221	288	10:21	miniseed

Table ST2 List of SEIS uptimes

Start and end dates of intervals at which either the SEISVELZ channel is available, or all of BHU, BHV, and BHW. All dates are given in UTC. Intervals separated by a gap of 5 min duration or less were joined into a single interval.

No.	Start	End	No.	Start	End
1	09.02.2019 14:03	10.02.2019 14:10	**44**	17.10.2019 00:13	19.10.2019 15:06
2	10.02.2019 21:00	12.02.2019 16:46	**45**	19.10.2019 16:27	23.10.2019 17:44
3	15.02.2019 03:08	17.02.2019 22:10	**46**	23.10.2019 19:02	26.10.2019 15:34
4	17.02.2019 22:57	18.02.2019 07:10	**47**	26.10.2019 20:55	17.11.2019 09:22
5	18.02.2019 09:31	22.02.2019 09:49	**48**	17.11.2019 10:43	20.11.2019 11:33
6	22.02.2019 10:24	23.02.2019 01:07	**49**	20.11.2019 12:54	22.11.2019 00:24
7	23.02.2019 01:37	23.02.2019 05:19	**50**	22.11.2019 00:40	27.11.2019 16:45
8	23.02.2019 05:55	24.02.2019 02:34	**51**	27.11.2019 17:03	07.12.2019 22:46
9	24.02.2019 05:00	24.02.2019 11:08	**52**	08.12.2019 00:07	11.12.2019 09:32
10	24.02.2019 11:41	25.02.2019 02:29	**53**	11.12.2019 09:54	11.12.2019 19:12
11	25.02.2019 03:02	25.02.2019 03:33	**54**	11.12.2019 19:23	15.12.2019 03:05
12	25.02.2019 15:54	28.02.2019 08:37	**55**	15.12.2019 04:42	22.12.2019 07:42
13	28.02.2019 09:13	01.03.2019 04:03	**56**	22.12.2019 09:15	27.12.2019 14:58
14	01.03.2019 08:18	01.03.2019 14:17	**57**	27.12.2019 15:06	29.12.2019 04:12
15	01.03.2019 14:52	02.03.2019 05:47	**58**	29.12.2019 04:26	30.12.2019 05:42
16	02.03.2019 06:19	03.03.2019 05:24	**59**	30.12.2019 05:56	02.01.2020 02:13
17	03.03.2019 16:56	04.03.2019 14:25	**60**	02.01.2020 02:28	08.01.2020 00:50
18	04.03.2019 16:19	05.03.2019 17:01	**61**	08.01.2020 01:13	19.01.2020 01:20
19	05.03.2019 17:07	28.03.2019 03:44	**62**	19.01.2020 02:44	10.03.2020 11:16
20	28.03.2019 05:32	25.04.2019 23:25	**63**	10.03.2020 11:50	11.03.2020 10:17
21	26.04.2019 01:53	08.05.2019 07:11	**64**	11.03.2020 12:16	25.03.2020 19:15
22	08.05.2019 09:40	10.05.2019 11:40	**65**	25.03.2020 21:31	09.04.2020 10:37
23	10.05.2019 11:48	11.05.2019 01:01	**66**	09.04.2020 10:58	09.04.2020 21:40
24	11.05.2019 02:48	11.05.2019 03:03	**67**	09.04.2020 21:58	12.04.2020 06:28
25	11.05.2019 03:12	17.05.2019 23:59	**68**	12.04.2020 08:55	02.05.2020 19:58
26	18.05.2019 00:08	20.05.2019 22:12	**69**	02.05.2020 21:55	17.05.2020 05:12
27	20.05.2019 22:40	25.05.2019 06:15	**70**	17.05.2020 07:10	19.05.2020 08:02
28	25.05.2019 17:00	19.06.2019 02:21	**71**	19.05.2020 08:09	21.05.2020 03:21

No.	Start	End	No.	Start	End
29	19.06.2019 04:22	17.07.2019 22:21	**72**	21.05.2020 03:38	30.05.2020 13:47
30	17.07.2019 22:30	22.07.2019 01:56	**73**	30.05.2020 15:44	04.06.2020 12:55
31	22.07.2019 02:09	14.08.2019 20:41	**74**	04.06.2020 13:11	06.06.2020 18:15
32	14.08.2019 20:57	15.08.2019 09:42	**75**	06.06.2020 20:21	13.06.2020 22:52
33	15.08.2019 10:00	21.08.2019 20:48	**76**	14.06.2020 00:58	21.06.2020 04:05
34	22.08.2019 00:00	28.08.2019 08:04	**77**	21.06.2020 05:36	29.07.2020 11:06
35	18.09.2019 23:10	23.09.2019 14:30	**78**	29.07.2020 11:14	02.08.2020 23:26
36	23.09.2019 14:43	25.09.2019 21:14	**79**	02.08.2020 23:35	06.08.2020 15:43
37	25.09.2019 21:28	27.09.2019 11:14	**80**	06.08.2020 15:48	22.08.2020 20:25
38	27.09.2019 11:30	30.09.2019 00:56	**81**	22.08.2020 21:34	27.08.2020 10:47
39	30.09.2019 01:10	03.10.2019 00:12	**82**	27.08.2020 11:00	27.08.2020 17:14
40	03.10.2019 00:25	09.10.2019 07:54	**83**	27.08.2020 17:27	06.09.2020 05:20
41	09.10.2019 09:49	12.10.2019 10:28	**84**	06.09.2020 07:05	19.09.2020 14:04
42	12.10.2019 11:49	16.10.2019 13:06	**85**	19.09.2020 15:28	25.09.2020 22:46
43	16.10.2019 14:28	16.10.2019 23:54	**86**	25.09.2020 22:58	26.09.2020 16:05

S2 Data Processing

This section gives a brief summary of the routine data processing underlying the event catalogue used in this study. For detailed descriptions, see Clinton et al. (2018), Giardini et al. (2020), and Clinton et al. (2021). In case of contradictions between these references and the present section, the references are authoritative if not stated otherwise. Later catalog versions may be based on a different processing.

Event detection

Routine event detection is conducted on the vertical VBB component, sampled at 20 *Hz*, and using the Mars-specific extensions of the SeisComP3 software as described by Clinton et al. (2018).

For 2.4 *Hz* events, the time series is filtered using a 6 pole Butterworth 2.2-2.8 *Hz* band pass filter, and an STA/LTA detector with window lengths of 100 *s* and 1000 *s*, respectively. To be included in the catalog, an STA/LTA amplitude threshold of 1.3 is required.

Stronger HF events are usually visible in spectrograms.

Events detected by the seismologist on duty are later reviewed by the MQS team.

Event types

The definitions of event categories are given by Clinton et al. (2021) and are mainly based on the frequency content of signals and, to a lesser extent, other features of the waveforms. The *High Frequency* (HF) event family is defined by containing "energy in 3 components predominantly at 2.4 Hz and more" (Clinton et al., 2021; energy below 2.4 *Hz* does not necessarily exclude events from this family). The significance of the reference frequency of 2.4 *Hz* lies in a continuously excited resonance at about 2.4 *Hz* (Giardini et al., 2020) and its excitation by weak events: The *2.4 Hz events* are visible only by a transient excitation of a narrow frequency band around this resonance, and are considered to be weak High Frequency events (Clinton et al., 2021). Besides the HF and 2.4 *Hz* events, there is a third group of events with high frequency signals, the *Very High Frequency* (VF) events, that differs from the other two by the partitioning of energy between horizontal and vertical components, and even higher signal frequencies up to the 10 Hz Nyquist frequency. *Low Frequency* and *Broadband* events contain signal energy predominantly or entirely below 2.4 *Hz*.

Displacement amplitudes

Event amplitudes are estimated as part of the magnitude estimation of MQS.

For all but the 2.4 *Hz* events, amplitude can be estimated as the long period plateau amplitude in the displacement spectrum, by fitting the slope of the displacement power spectrum at frequencies between 0.1 *Hz* and 1 *Hz*, and for which the signal is at least 3 *dB* above the noise power spectrum, assuming that the recorded spectrum is a flat spectrum that is attenuated according to $e^{-\pi\tau f}$ with a damping parameter τ.

For 2.4 *Hz* events, this is not possible, since these events are visible only in a narrow frequency band where they excite the 2.4 *Hz* ambient resonance. For these events, the amplitude is determined by fitting a Lorentz curve

$$A(f) = A_0 + 10log_{10}\left[e^{-\pi f\tau}\left(1+\alpha\left[1+\frac{(f-f_0)^2}{(f_w/2)^2}\right]\right)^{-1}\right] \qquad \text{(SEQ1)}$$

(see also van Driel et al., 2021)

with $f_0 = 2.41\,Hz$ and $f_w = 0.2\,Hz$ to the displacement spectrum between 2 and 2.8 *Hz*. For consistency, the Lorentz curve fit is used for all HF events discussed in this paper.

Signal-to-Noise ratio

The Signal-to-Noise ratio (SNR) is defined as the ratio between power spectral densities (PSD) of the seismic signal and a nearby noise window, and within the frequency band from 2.2 *Hz* to 2.8 *Hz* (specific for the High Frequency Event family). The time windows used are determined by the MQS frontline team in a way such that the effect of the ubiquitous artefacts ("glitches", see Scholz et al., 2020) is minimized. Due to the presence of artefacts it is unavoidable that the duration of noise windows differs between events. Typical durations are about 2 *min* and shorter than 6 *min* for approx. 80% of all HF events. For only six events, the noise window is shorter than 2 *min*.

The data is corrected for instrument response, converted to displacement (Clinton et al., 2021, at one point erroneously speak of acceleration), and rotated to ZNE coordinates. The PSD is then estimated using Welch's method on time windows of 50 *s* duration and with 50% overlap. The time windows are detrended and tapered with a Hanning window separately.

Noise amplitude

The noise amplitude is estimated in two frequency bands; one designed for the amplitude of the 2.4 *Hz* mode specifically, and one in a slightly broader frequency window to take wind-excited noise into account. For the computation, the signal is corrected to displacement u, as described above, and band-pass-filtered to 2.2 - 2.6 or 1.2 - 3.0 Hz, respectively, using an order 8 Butterworth filter. The signal is then cut into non-overlapping 120 second time windows, for each of which the standard deviation of all samples, $\sigma(u) = \sqrt{\sum |u_i - \bar{u}|^2}$, is computed, with $\bar{u}$ being the mean displacement.

S3 Selected Events

The authoritative event catalog is the catalog of the InSight Marsquake Service as cited in the main paper. Since this work is based on a preliminary version of that catalog, and further improvements of parameters of individual events in future catalogs cannot be excluded at the time of writing, we give a table of the events and times as evaluated. This table does not contain events that we excluded from our evaluation based on the selection criteria described in the main text.

The table is provided only for the traceability of our results and should not be used for any other purpose. Any re-analysis, or extended analysis, should be based on the most recent MQS catalog release.

Table ST3 List of the 118 selected events
The columns are as follows: ID: event ID according to MQS naming conventions, Type: Event type according to MQS event classification scheme, Date: UTC time at which event energy begins in SEIS recording

ID	Type	Date	ID	Type	Date
S0185b	HF	2019-06-05T03:29:12.39425Z	**S0343b**	HF	2019-11-14T11:59:39.648596Z
S0194c	2.4Hz	2019-06-14T10:29:41.787001Z	**S0347a**	HF	2019-11-18T14:50:40.69798Z
S0202c	HF	2019-06-22T15:49:43.018022Z	**S0348b**	2.4Hz	2019-11-19T14:51:47.505614Z
S0213a	HF	2019-07-03T06:44:22.785108Z	**S0348c**	2.4Hz	2019-11-19T15:48:00.701557Z
S0219b	2.4Hz	2019-07-10T02:21:25.12252Z	**S0349a**	HF	2019-11-20T17:12:29.879989Z
S0222a	2.4Hz	2019-07-13T03:32:39.752227Z	**S0351b**	HF	2019-11-22T17:40:24.400023Z
S0227d	2.4Hz	2019-07-18T08:43:56.867128Z	**S0352a**	HF	2019-11-23T18:18:13.277075Z
S0228c	HF	2019-07-19T09:38:15.396682Z	**S0353b**	2.4Hz	2019-11-24T18:57:03.23894Z
S0231b	HF	2019-07-22T08:09:53.836173Z	**S0355a**	2.4Hz	2019-11-26T16:46:44.962297Z
S0239a	HF	2019-07-30T14:16:49.430121Z	**S0361c**	HF	2019-12-03T00:44:57.814616Z
S0240b	2.4Hz	2019-07-30T19:35:57.11871Z	**S0363a**	2.4Hz	2019-12-04T22:59:24.466255Z
S0242a	2.4Hz	2019-08-02T15:11:02.043591Z	**S0363c**	2.4Hz	2019-12-05T00:50:41.291516Z
S0246a	HF	2019-08-06T18:15:50.360992Z	**S0363d**	HF	2019-12-05T01:38:52.943716Z
S0252a	2.4Hz	2019-08-12T23:51:54.243203Z	**S0365a**	2.4Hz	2019-12-07T05:29:41.675893Z
S0257a	2.4Hz	2019-08-18T00:23:04.309844Z	**S0366c**	2.4Hz	2019-12-08T02:36:46.001702Z
S0257b	2.4Hz	2019-08-18T01:05:34.205201Z	**S0366e**	2.4Hz	2019-12-08T05:39:33.74298Z
S0260a	HF	2019-08-21T05:11:08.385435Z	**S0369b**	2.4Hz	2019-12-11T03:57:10.876039Z
S0261b	2.4Hz	2019-08-22T07:08:07.12195Z	**S0370a**	2.4Hz	2019-12-12T03:22:50.254183Z
S0262b	HF	2019-08-23T08:39:46.601468Z	**S0371b**	HF	2019-12-13T08:31:10.664649Z
S0264b	2.4Hz	2019-08-25T06:20:06.929163Z	**S0372a**	2.4Hz	2019-12-14T07:36:47.817106Z
S0265b	2.4Hz	2019-08-26T06:47:41.975661Z	**S0373b**	2.4Hz	2019-12-15T10:19:57.152934Z
S0289a	HF	2019-09-20T00:01:50.697433Z	**S0375a**	HF	2019-12-17T10:13:46.037393Z

ID	Type	Date	ID	Type	Date
S0290a	2.4Hz	2019-09-21T02:06:12.181061Z	**S0378a**	2.4Hz	2019-12-20T07:55:31.671114Z
S0291c	HF	2019-09-22T03:46:08.44948Z	**S0378b**	2.4Hz	2019-12-20T13:12:36.581946Z
S0292a	HF	2019-09-23T03:44:42.887549Z	**S0383a**	2.4Hz	2019-12-25T14:36:33.168899Z
S0295b	2.4Hz	2019-09-26T06:13:53.216587Z	**S0384a**	2.4Hz	2019-12-26T11:55:04.175727Z
S0297b	2.4Hz	2019-09-27T15:02:25.414571Z	**S0384b**	HF	2019-12-26T13:09:43.134175Z
S0297c	2.4Hz	2019-09-28T03:51:16.636413Z	**S0384c**	HF	2019-12-26T14:51:10.92641Z
S0299a	2.4Hz	2019-09-29T13:14:21.912306Z	**S0385a**	HF	2019-12-27T12:41:08.608345Z
S0301a	2.4Hz	2019-10-01T15:02:09.871588Z	**S0392a**	HF	2020-01-03T17:25:08.213689Z
S0303a	HF	2019-10-03T17:14:12.780168Z	**S0394a**	2.4Hz	2020-01-05T19:54:07.470626Z
S0304a	2.4Hz	2019-10-05T08:48:38.115748Z	**S0397b**	2.4Hz	2020-01-08T22:40:07.443985Z
S0304b	HF	2019-10-05T12:47:56.545428Z	**S0401b**	2.4Hz	2020-01-13T01:10:24.962234Z
S0306a	HF	2019-10-07T15:03:41.598223Z	**S0405f**	HF	2020-01-17T05:48:10.752377Z
S0308b	2.4Hz	2019-10-09T10:01:15.630905Z	**S0408a**	2.4Hz	2020-01-19T10:53:33.411279Z
S0308a	HF	2019-10-09T15:46:04.420005Z	**S0408b**	2.4Hz	2020-01-20T06:05:19.361781Z
S0311a	HF	2019-10-12T13:11:05.128167Z	**S0409c**	2.4Hz	2020-01-21T10:25:04.832086Z
S0311b	2.4Hz	2019-10-12T16:18:08.091475Z	**S0415a**	2.4Hz	2020-01-26T16:31:14.263367Z
S0314a	2.4Hz	2019-10-15T14:25:51.632638Z	**S0417a**	2.4Hz	2020-01-28T17:38:18.756785Z
S0314b	HF	2019-10-15T15:41:36.753559Z	**S0418a**	2.4Hz	2020-01-30T11:55:17.998981Z
S0315b	HF	2019-10-16T15:12:41.388803Z	**S0419a**	2.4Hz	2020-01-31T15:28:16.149073Z
S0318a	2.4Hz	2019-10-19T19:25:29.693703Z	**S0423c**	HF	2020-02-04T19:29:47.49571Z
S0319a	HF	2019-10-20T03:25:48.491702Z	**S0428a**	HF	2020-02-09T00:48:10.121561Z
S0319b	HF	2019-10-20T17:43:48.513634Z	**S0429a**	HF	2020-02-10T20:41:32.62005Z
S0320a	2.4Hz	2019-10-21T04:54:17.648875Z	**S0429c**	2.4Hz	2020-02-11T00:41:28.29972Z
S0322a	2.4Hz	2019-10-23T02:09:02.653387Z	**S0430a**	2.4Hz	2020-02-11T02:36:32.89019Z
S0323a	HF	2019-10-24T21:14:01.1308Z	**S0430c**	2.4Hz	2020-02-11T20:46:13.51365Z
S0325b	HF	2019-10-26T04:41:24.69364Z	**S0431b**	HF	2020-02-12T22:01:24.099025Z
S0325d	2.4Hz	2019-10-27T04:00:16.761211Z	**S0432a**	HF	2020-02-14T01:52:39.328009Z
S0327a	2.4Hz	2019-10-28T23:29:44.699281Z	**S0436c**	HF	2020-02-18T02:04:51.213103Z
S0327c	HF	2019-10-29T04:03:06.729189Z	**S0439a**	2.4Hz	2020-02-20T08:39:43.658188Z
S0329a	2.4Hz	2019-10-31T04:40:14.618807Z	**S0446a**	2.4Hz	2020-02-27T11:47:58.611408Z
S0331a	HF	2019-11-01T08:40:10.644513Z	**S0448a**	2.4Hz	2020-03-01T12:31:34.209941Z
S0331b	2.4Hz	2019-11-02T06:55:39.560251Z	**S0458b**	HF	2020-03-11T18:23:15.900445Z

ID	Type	Date	ID	Type	Date
S0337a	2.4Hz	2019-11-08T11:36:45.249527Z	**S0460a**	2.4Hz	2020-03-12T20:50:36.177444Z
S0338a	2.4Hz	2019-11-08T12:14:23.654624Z	**S0467a**	2.4Hz	2020-03-20T18:12:15.421504Z
S0338b	2.4Hz	2019-11-09T07:24:11.190621Z	**S0477a**	2.4Hz	2020-03-31T00:48:47.674521Z
S0340a	HF	2019-11-11T09:57:57.696696Z	**S0490a**	HF	2020-04-13T09:36:46.374094Z
S0342a	2.4Hz	2019-11-13T13:07:01.432501Z	**S0518a**	2.4Hz	2020-05-11T12:06:03.942327Z

S4 Detection rate variation due to background noise, and detection efficiency

An increasing (decreasing) background noise level will reduce (increase) the probability of detecting small events, and may thus result in a time dependent detection rate even if the occurrence rate of events remains constant. It is thus paramount to quantify the effect the background noise has on event detections.

We define the *detection efficiency* η as the ratio of the number of detected events, and the expected number of events according to some event rate λ, assuming a non-stationary Poisson process (see section S7).

We consider seismograms in the vertical-north-east coordinate frame, i.e. time series $z(t)$, $n(t)$, and $e(t)$. Experience shows that most martian events are detected on the vertical component (Clinton et al., 2021), so we focus on $z(t)$ in the following.

We assume that the seismic signal, $s_z(t)$, and the background noise, $n_z(t)$ are independent of each other, such that the recorded seismogram is the sum of the two

$$z(t) = s_z(t) + n_z(t) \quad \text{(SEQ2)}$$

A catalog may contain N events with arrival times

$$t_i \in \{t_1, t_2, t_3, \cdots, t_N\} \quad \text{(SEQ3)}$$

Then the seismogram of the i-th event is

$$z(t - t_i) = s_z(t - t_i) + n_z(t - t_i) \quad \text{(SEQ4)}$$

Catalogs are usually incomplete in that they do not list all events that occurred during the time interval covered. This is a simple consequence of the presence of background noise, which hides small or distant events. The literature usually assumes the existence of a magnitude of completeness such that all events that exceeded this magnitude are catalogued. There is also a consensus that very large events are usually underrepresented, because their recurrence interval is longer than the time interval covered by the catalog.

What makes the discussion about the validity of magnitude-frequency relations for very large earthquakes challenging, is the observation that no existing catalog is *representative* for such events. Even if a magnitude of completeness is correctly determined as the threshold above which all occurred events are cataloged, it is only the lower boundary of an *interval of representativeness,* in which the catalog correctly reflects the relative frequencies of the magnitudes of past or future events with recurrence times longer than the interval covered by the catalog. The upper boundary is likely below the largest observed events. Beyond some corner magnitude, the magnitude-frequency distribution must deviate from the Gutenberg-Richter law for reasons of conservation of energy and the finiteness of the planet. It is not straight forward to distinguish that physical corner magnitude from the upper limit of catalog representativeness.

In the case of InSight, the event catalog is certainly not complete in the usual sense, since daytime noise is such that almost all events detected during nighttime so far would be undetectable during daytime (see fig. 2). Instead of a completeness threshold, we can thus only expect to determine a lower threshold for representativeness.

Event Rescheduling Approach

This section discusses what we call the "Event Rescheduling Approach" to analyze how a variation of background noise with time causes a parallel change of the event detection rate. After a brief description, we discuss inherent limitations that render it unsuitable for our problem. To overcome these limitations, we developed the statistical approach discussed in the next subsection. We nevertheless present the Event Rescheduling Approach here, since we believe that our arguments will support readers to find suitable approaches for their own problems.

The main idea of this approach is to analyze the effect of variable background noise by adding the noise of a time t_j in question to the i-th recorded event, and investigate if the i-th event is still detectable. We call this the *Event Rescheduling Approach*, since it pretends that events are rescheduled to different arrival times.

One computes from the recorded time series (SEQ4) a new seismogram

$$z_{ij}(t) = z(t - t_i) + z(t - t_j) \qquad \text{(SEQ5)}$$

Assuming that

$$s_z(t - t_j) = 0 \qquad \text{(SEQ6)}$$

i.e. that there is no seismic signal or event in the added seismogram for the duration of the i-th event, one obtains the modified seismogram of the i-th event

$$z_{ij}(t) = s_z(t - t_i) + n_z(t - t_i) + n_z(t - t_j) \qquad \text{(SEQ7)}$$

This can be repeated for all events, yielding a set of modified seismograms $\{z_{1,j}, z_{2,j}, z_{3,j}, \cdots, z_{N,j}\}$. Due to the increased noise level in each of these seismograms, only a reduced number $N_j < N$ of events will be detectable in this set of seismograms.

The ratio

$$\eta(t_j) = \frac{N_j}{N} \qquad \text{(SEQ8)}$$

is an estimation of the detection efficiency at time t_j.

The SNR at which the i-th event is actually observed is

$$SNR_i = \frac{s(t-t_i)}{n(t-t_i)} \qquad \text{(SEQ9)}$$

If it occurred at time t_j, its SNR would be

$$SNR_j = \frac{s(t-t_i)}{n(t-t_j)} \qquad \text{(SEQ10)}$$

A detectability test in the framework of the Event Rescheduling Approach however would work with an SNR of

$$SNR_{ij} = \frac{s(t-t_i)}{n(t-t_i)+n(t-t_j)} \tag{SEQ11}$$

meaning that the SNR under which detectability is tested is reduced with respect to the desired SNR by a factor

$$\frac{SNR_j}{SNR_{ij}} = \frac{n(t-t_i)}{n(t-t_j)} + 1 \tag{SEQ12}$$

This not only means that the SNR is biased, but also that the SNR changes by a different amount for each event.

The Event Rescheduling Approach has a number of limitations; the following list is not necessarily complete.

1. Its application requires not only to compute a set of N new waveforms, but also the repetition of the event detection method on all of these waveforms for all times in question, which in our case amounts to about one hundred events with wavetrain durations of 10 to 20 minutes, and a few hundred sols. Depending on the detection method and the number of events in question, the Event Rescheduling Approach might become computationally expensive.
2. From SEQ7 and SEQ12, it is obvious that the new waveforms do not all have the same noise level that was observed at t_j, but the sum of this desired noise level, and the noise level at the actual arrival times of the events. Moreover, the individual new waveforms have different noise levels, and the SNRs of the individual events are not all changed by the same factor. The Event Rescheduling Approach does thus not provide a comparison in a well-defined noise environment.
3. Equation SEQ7 also shows that the Event Rescheduling Approach can estimate the detection efficiency only for increased, but not for reduced noise levels.
4. The N catalogued events represent only one realization of the event production process. Small events are underrepresented; events below the detection threshold are not represented at all. The Event Rescheduling Approach is thus not able to estimate the detection efficiency at times where a low noise level reduces the thresholds for detection or representativeness. The Event Rescheduling Approach is also unable to predict the detectability of events smaller than the smallest catalogued event.

Because of this reasoning, we did not even implement the Event Rescheduling Approach, but searched for a better way to deal with the variable noise on Mars.

Statistical approach

Some relation must exist between the number of events that are detectable at a certain noise level, and the amplitude-frequency distribution of the actually detected events. The distribution of amplitudes at the location of the seismometer is relevant here, since these are to be seen in relation to the background noise at the seismometer site. The magnitudes or seismic moments of sources play an indirect role.

We consider the cumulative distribution of amplitudes in survivor form, i.e. the number of events the amplitude of which exceeds a certain value: $N(A \geq A_0)$ is the number of events with an amplitude A equal to or larger than A_0, detected during a time interval of unit duration.

We assume that a threshold amplitude A_{TH} exists, such that the catalog is representative for the amplitude-frequency distribution of all events that occurred. Events with amplitudes below A_{TH} are excluded from further considerations. The value of A_{TH} has to be estimated from the catalogued events, methods known from estimating the magnitude of completeness of a catalog can be employed here.

A catalog covers a certain period of time, which starts at time S and ends at time T. The amplitude-frequency relation of catalogued events is then

$$N(A \geq A_0) = \frac{1}{T-S} F(A_0) \tag{SEQ13}$$

Here F is some function of the amplitude, and by division with the covered duration $T - S$, the relation is normalized to the time unit in which S and T are given. We assume that F is monotonically decreasing with amplitude and independent of time (the validity of the latter is tested in subsection "Time-invariant distribution of amplitudes" on page 29). Especially, the number of events exceeding the threshold of representativeness is

$$N(A \geq A_{TH}) = \frac{1}{T-S} F(A_0) \tag{SEQ14}$$

A common assumption is that detectability depends on the ratio between signal and noise amplitude, i.e. on

$$SNR_i = \frac{s_z(t_i)}{n_z(t_i)} \tag{SEQ15}$$

Increasing the noise amplitude by a factor x then requires that the signal amplitude increases by the same factor to maintain detectability. This means that, at the higher noise level, the catalog can be representative only for events with amplitude exceeding xA_{TH}, and the number of representative events reduces according to the shape of the amplitude-frequency distribution F.

Since $s_z(t)$ and $n_z(t)$ cannot be obtained separately from $z(t)$, their ratio can only be approximated using noise from nearby time windows that presumably do not contain a seismic signal, i.e. where $s_z(t) = 0$ can be assumed. Using such time windows, we define as the noise level $L(t_k)$ at a time t_k the standard deviation of the noise during a short time interval Δt. "Short" here means that the noise can be considered stationary during such an interval. When the seismogram is digitized with a sampling interval δt, L is computed from $\Delta t / \delta t$ samples as

$$L^2(t_k) = \frac{1}{\Delta t / \delta t - 1} \sum_{i=1}^{\Delta t / \delta t} (z(t + i\delta t) - \bar{z})^2 \tag{SEQ16}$$

We compute $L(t_k)$ for a series of non-overlapping, discrete intervals throughout the entire life time of the SEIS experiment.

The ratio x used above can now be approximated as

$$x(t_k) = \frac{L(t_k)}{L(t_{REF})} \tag{SEQ17}$$

where $L(t_{REF})$ is the noise level at some reference time, when the noise level was low enough to detect all catalogued events. The lowest observed noise level appears to be a good choice for this value.

For a time interval of duration Δt during which the noise level was $L(t_k)$, the expected number of events, according to the amplitude-frequency distribution, is

$$N(A \geq xA_{TH}) = \frac{\Delta t}{T - S} F(xA_{TH}) \qquad \text{(SEQ18)}$$

Then the detection efficiency η at time t_k is the quotient of SEQ18 and SEQ14

$$\eta(t_k) = \Delta t \frac{F(xA_{TH})}{F(A_{TH})} \qquad \text{(SEQ19)}$$

For practical purposes, it might be useful to compute an average detection efficiency for a longer time interval, during which the noise is not stationary, e.g. to combine the detection efficiency of many time windows of a few minutes into an average for an entire day.

If $\lambda(t)$ is the rate of events for which the catalog is representative, the number of detections during a time interval of duration Δt is

$$N(t_0, t_0 + \Delta t) = \eta(t_0) \int_{t_o}^{t_0+\Delta t} \lambda(u) du \qquad \text{(SEQ20)}$$

During a long time interval of duration $m\Delta t$, $m \in \mathbb{N}$ (e.g. $m = 720$ and $\Delta t = 2\, min$ to obtain one day), the expected number of detections is the sum of m intervals

$$N(t_0, t_0 + m\Delta t) = \sum_{i=1}^{m} \left[\eta(t_o + (i-1)\Delta t) \int_{t_0+(i-1)\Delta t}^{t_0+i\Delta t} \lambda(u) du \right] \qquad \text{(SEQ21)}$$

With the desired average detection efficiency $\bar{\eta}$, this number is also given by

$$N(t_0, t_0 + m\Delta t) = \bar{\eta}(t_0) \int_{t_0}^{t_0+m\Delta t} \lambda(u) du \qquad \text{(SEQ22)}$$

We obtain $\bar{\eta}$ by division of these two equations

$$\bar{\eta}(t_0) = \frac{\sum_{i=1}^{m} \left[\eta(t_o + (i-1)\Delta t) \int_{t_0+(i-1)\Delta t}^{t_0+i\Delta t} \lambda(u) du \right]}{\int_{t_0}^{t_0+m\Delta t} \lambda(u) du} \qquad \text{(SEQ23)}$$

This involves the unknown time-dependency of the event rate. However, if the variation of this rate is small during $m\Delta t$, such that $\lambda(t) \approx const$, the integrals solve into $const$ times the duration of the intervals, and the constant itself cancels out.

The long-term average detection efficiency for a slowly varying event rate, during an interval of duration $m\Delta t$ is then

$$\bar{\eta}(t_0) = \frac{1}{m}\sum_{i=1}^{m} \eta(t_0 + (i-1)\Delta t) \qquad \text{(SEQ24)}$$

i.e. the arithmetic mean of the short-term detection efficiencies within that interval.

So far, we did not assume a specific form of amplitude-frequency distribution. But since amplitudes are related to seismic moments, it is not beside the point to assume that amplitudes are distributed according to a power law similar to the Gutenberg-Richter law, and that a consideration in log-log-coordinates is useful. The validity of this assumption can be tested with actual data (see subsection "Power Law distribution of amplitudes" on page 28).

For amplitudes, we implement this by converting amplitudes to the decibel scale, according to

$$A_{dB} = 20 log_{10} A \qquad \text{(SEQ25)}$$

The representativeness threshold is then given by $A_{TH,dB}$, and for the increased noise level by $A_{TH,dB} + \chi$, with $\chi = 20 log_{10} x$. The amplitude-frequency distribution formally transforms to

$$f(A_{dB}) = log_{10} F(A) \qquad \text{(SEQ26)}$$

With this, the logarithm of the detection efficiency is given by

$$log_{10} \eta(t_k) = log_{10} \Delta t + f\left(\chi + A_{TH,dB}\right) - f(A_{TH}) \qquad \text{(SEQ27)}$$

For a Gutenberg-Richter-like power law, f for events above the representativeness threshold is

$$f(A_{dB}) = a - b\left(A_{dB} - A_{TH,dB}\right) \qquad \text{(SEQ28)}$$

with a and b to be estimated from the catalog. When inserting SEQ28 into SEQ27, a and $A_{TH,dB}$ cancel out, and we obtain the detection efficiency as function of noise level

$$\eta(t_k) = \Delta t \left(\frac{L(t_k)}{L_{REF}}\right)^{-20b} \qquad \text{(SEQ29)}$$

Thus, the detection efficiency is obtained by processing the waveform data. We note that the signal-to-noise ratio of detected events does not enter the detection efficiency computation explicitly.

Demonstration using daily noise variation at InSight

In this section, we show (in figures SF2 and SF3) that the detection efficiency approach can predict the detection rate of one half of the night from the detection rate in the other half, employing the measured noise level according to eq. SEQ16 and the detection efficiency factor according to eq. SEQ29.

Between sol 289 and sol 385, the number of detected events increased with an essentially constant rate of about 0.7 events per day. During this time interval, the computed per-sol detection efficiency was around 0.146 and varied little (see fig.4 of the main manuscript). These observations indicate that there was little to no long-term variation of the event occurrence rate during these 96 sols. We however know that there is considerable variability of the seismic background noise throughout each individual sol, making detections during daytime impossible, and reducing the detection rate in the second half of the night (i.e. from 00:00 LTST to sunrise, see figure 2 of the main manuscript), compared to the first half (i.e. from about 17:00 LTST to 23:00 LTST).

Since there is hardly any sol with more than one detected event, we cannot just pick a single sol and analyze its event sequence. Instead, we combine the events from all 96 sols by neglecting the sol number and looking at the hours and minutes of detection only. Detection times in local mean solar time (LMST) are used as a uniform time scale.

We compute the displacement noise level in nonoverlapping windows of 2 minutes duration and convert it into detection efficiency, using the *b*-value of the amplitude distribution shown in figure 3 of the main manuscript. Our tests estimate the detection rate during a time window (termed "calibration window") during which it appears to be constant. The reference noise level needed for the detection efficiency computation is chosen such that the maximum detection efficiency during the calibration window is 1 (To demonstrate that the noise level variation essentially predicts the detection rate, it is sufficient to show that a reference noise level exists for which the prediction works, no matter how this reference level is found. We opt for an approach based on data from the calibration window only, although other approaches might result in even better predictions). The detection efficiency curve is then smoothed by computing 15 min averages (red lines in figures SF2b and SF3b). Based on the detection rate during the calibration window, we compute the number of detections expected in a 2 minutes window.

This number is then adapted to the noise level in each 2 min window by multiplying with the corresponding value of the smoothed detection efficiency. These adapted event counts are finally added up to produce the predicted cumulative event count that can be compared to the observed cumulative count. To assess the statistical significance of the difference between the predicted and the observed cumulative count, we compare the residual (observed minus predicted) to the width of a distribution obtained from synthetic event sequences.

The displacement noise level varied by about 40 *dB* throughout the noisiest and quietest times of the 96 sols, and by about 20 *dB* when only night times are considered (figures SF2a and SF3a). The noise level repeatedly drops to -220 *dB* and below (relative to 1 *m* displacement) between, roughly, 17:00 LMST and 22:30 LMST on most sols, but shows more variation in the small hours.

A total of 65 HF events were detected between sol 289, 00:00 LMST (20.09.2019 04:20:26 UTC), and sol 385, 00:00 LMST (26.12.2019 18:47:02 UTC), 9 of these between 00:00 LMST and 06:00 LMST, none between 06:00 LMST and 16:45 LMST, and all others between 16:45 LMST and 24:00 LMST (squares in figures SF2b and SF3b).

In the first test, we estimate the detection rate between 18:00 LMST and 22:30 LMST. During this time window, 23 events were detected, resulting in a rate of 9.2 events per hour. The reference noise level for this calibration interval, as resulting from the criterion above, is -220.3 *dB*. The predicted cumulative event count for the time from 17:00 LMST to 24:00 LMST fits nicely to the observed count (with a residual well within the 68% quantile). The fit for the time between 00:00 LMST and 05:00 LMST is not as well, with residuals outside the 90% quantile for short times. In terms of the cumulative count, this is however computed as a "backward" projection. The standard deviation of a stationary Poisson process is given by the square root of the event count, resulting in quantile widths that grow with time. Evaluating the small hours as following the evening hours, rather than preceding them, would put the residual into smaller-valued quantiles.

In the second test, the calibration window is from 02:00 LMST to 05:00 LMST. During this time window, only 5 events were detected, resulting in a detection rate of 1.7 events per hour. The reference noise level for this calibration window is -212.8 *dB*. The residual of the predicted cumulative event count is well within the 68% quantile and only briefly scratches its boundary after 22:00 LMST.

These two tests demonstrate that the differences between daytime and nighttime detection counts by InSight can be explained quantitatively using the detection efficiency approach.

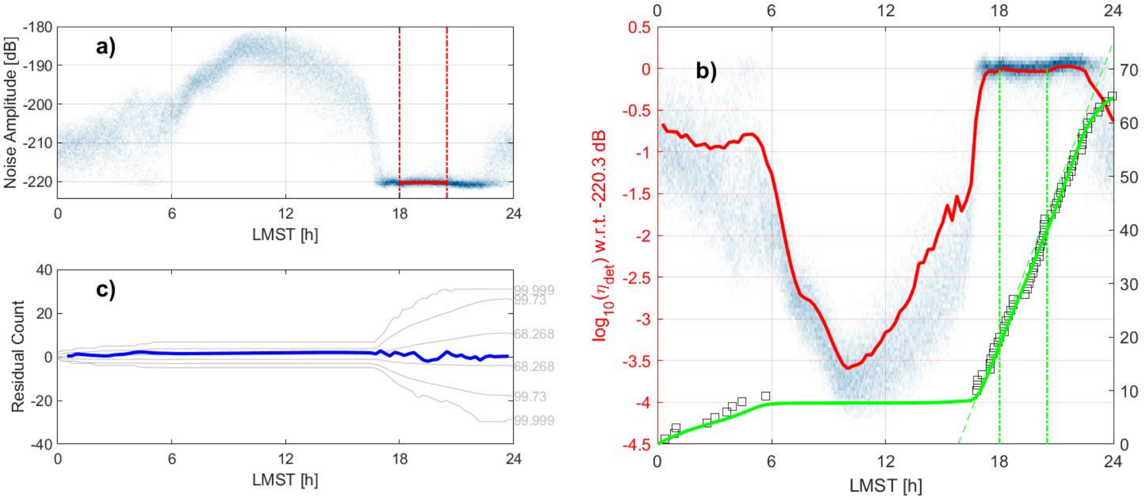


Figure SF2 Prediction of daily detection rate variation from 18:00 to 20:30 LMST window.
a) Histogram of the displacement noise level distribution in 2 minute windows, computed in *dB* relative to 1 *m*; detection rate calibration time window (vertical dash-dot lines) and the resulting reference noise level (horizontal line) b) Histogram of the detection efficiency distribution in 2 minute windows (shaded), and 15 minute average (red), both drawn on the logarithmic y-scale on the left. Cumulative count of observed events (squares) and predicted cumulative count (green solid line), based on the calibration window in a) and the detection efficiency, are drawn on the linear y-scale on the right. The dashed green line corresponds to a stationary Poisson process with the rate estimated in the time window between the two vertical dash-dot lines. c) Statistical significance analysis: The difference between the observed cumulative count and the prediction in b) is compared to residuals obtained from synthetic event sequences. Gray lines indicate quantiles that include 68%, 99.7% and 99.999% of the distribution.

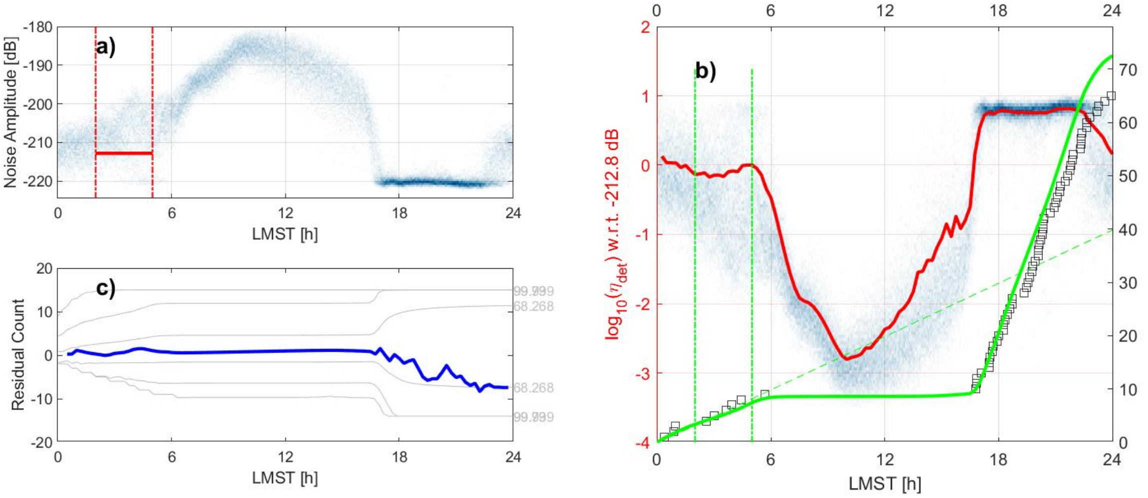


Figure SF3 Prediction of daily detection rate variation from 02:00 to 05:00 LMST window. Plotted as in SF2. See text for details.

Applicability to InSight HF events

In order to apply the detection rate evaluation to long-term variations of the InSight HF event rate, we first need to confirm that these data comply with some assumptions stated in section "Statistical approach".

Slow variation of event rate

We routinely compute the standard deviation of ground displacement during 2-minute intervals for purposes unrelated to this study. This section investigates if the resulting time series noise levels can be re-used to obtain average detection efficiencies for intervals of one sol, using equation SEQ24. The application of SEQ24 however requires that the event rate varies only little during the interval over which we average. Figure 4 in the main paper shows the average detection efficiency under the assumption of a sufficiently slow variation of the event rate. Here we investigate how slow is "sufficiently slow".

Figure 2c of the main manuscript shows the evolution of HF event activity over about 560 sols (squares), and a kernel density estimation of the event rate λ_{KDE} (blue line). The KDE in this case uses a quartic kernel (Diggle & Marron, 1988)

$$\delta_w(t,t_0) = \begin{cases} \frac{1}{w}0.9375\left(1-\left(\frac{t-t_0}{w}\right)^2\right)^2 & \left|\frac{t-t_0}{w}\right| \le 1 \\ 0 & \left|\frac{t-t_0}{w}\right| > 1 \end{cases} \quad \text{(SEQ30)}$$

where w is the bandwidth parameter. Then the estimated rate function is the sum over one kernel per event, i.e.

$$\lambda_{KDE}(t) = \sum_{i=1}^{N} \delta_w(t,t_i) \quad \text{(SEQ31)}$$

where t_i is the time of the i-th event.

By means of the bandwidth, one can obviously render the derivative of λ_{KDE} arbitrarily small or large, so especially a reasonable lower limit for w, which implies an upper limit for the change rate of λ_{KDE}, must be found.

The bandwidth should probably be large enough to bridge the longest gap between two adjacent events, otherwise a single isolated event would result in an implausible local rate maximum. With the event selection given in table ST3, the longest data gap has a duration of 27.345 sols, so we take that the bandwidth should be larger than that.

Another criterion might be the smoothness of the resulting KDE rate: a bandwidth of at least 50 sols is necessary to obtain a rate function with only one maximum. Since there is a considerable data gap from sol 267 to sol 288 due to the August 2019 conjunction, we also consider a bandwidth of 100, which irons this out.

How small does the rate change need to be? The criterion to answer this has to come from equation SEQ23, which relates the long-term average to short term averages.

We introduce abbreviations

$$E_i = \int_{t_0+(i-1)\Delta t}^{t_0+i\Delta t} \lambda(u)du \quad \text{(SEQ32)}$$

which is the expected number of events between times $t_0 + (i-1)\Delta t$ and $t_0 + (i-1)\Delta t$, and

$$\bar{\eta} = \bar{\eta}(t_0, t_0 + m\Delta t) \tag{SEQ33}$$

$$\bar{\eta}_i = \bar{\eta}_i(t_0, t_0 + (i-1)\Delta t) \tag{SEQ34}$$

hence, we get the slightly more compact expression

$$\bar{\eta} = \frac{\sum_{i=1}^{m} \eta_i E_i}{\sum_{j=1}^{m} E_j} \tag{SEQ35}$$

If the event rate is constant, all E_i are equal, and the denominator becomes

$$\sum_{j=1}^{m} E_j = mE_i \tag{SEQ36}$$

such that the constant-case detection efficiency is

$$\bar{\eta}_c = \frac{1}{m} \sum_{i=1}^{m} \eta_i \tag{SEQ37}$$

If the event rate is *not* constant, the E_i are not all equal and the above cancellation does not apply. One can however still pull the denominator of SEQ23 into the sum and make a formal replacement

$$\bar{\eta} = \sum_{i=1}^{m} \eta_i \frac{E_i}{\sum_{j=1}^{m} E_j} = \sum_{i=1}^{m} \eta_i \frac{E_i}{\widetilde{m}_i E_i} \tag{SEQ38}$$

where

$$\widetilde{m}_i = \frac{1}{E_i} \sum_{j=1}^{m} E_j \tag{SEQ39}$$

and hence

$$\bar{\eta} = \sum_{i=1}^{m} \eta_i \frac{1}{\widetilde{m}_i} \tag{SEQ40}$$

Assuming a constant event rate when it is actually not constant means to use $\bar{\eta}_c$ where $\bar{\eta}$ should be used, namely to correct the model event rate in SEQ64 with a factor which should be corrected by the ratio $\bar{\eta}/\bar{\eta}_c$. It is the magnitude of this correction factor which tells if the derivative of the event rate is small enough.

The most convenient way to obtain this ratio is by first calculating the difference $\bar{\eta}_c - \bar{\eta}$,

$$\bar{\eta}_c - \bar{\eta} = \frac{1}{m}\sum_{i=1}^{m}\eta_i - \sum_{i=1}^{m}\frac{\eta_i}{\widetilde{m}_i} \tag{SEQ41}$$

By factoring out $\bar{\eta}_c$, we obtain the desired ratio

$$\frac{\bar{\eta}}{\bar{\eta}_c} = 1 - \frac{1}{\bar{\eta}_c}\sum_{i=1}^{m}\eta_i\frac{\widetilde{m}_i - m}{m\widetilde{m}_i} \tag{SEQ42}$$

where the deviation of $\bar{\eta}/\bar{\eta}_c$ from 1, which needs to be minimized, is nicely packaged.

Computing the $\widetilde{m}_i$ of course requires to evaluate the - unknown - event rate. As a substitute, we plug in the kernel density estimation λ_{KDE} computed from the observed event times, and compare results for different kernel bandwidths (numerical values in SF4). As expected, the error $1 - \bar{\eta}/\bar{\eta}_c$ that arises from assuming a constant rate during the averaging decreases with an increasing KDE bandwidth, as the entire purpose of larger bandwidths is to make it vary slower. When the bandwidth exceeds the minimum required bandwidth of 27.345 sols, the resulting error in the detection efficiency average is less than 2.6% for 95% of the time, and only about 1% for 75% of the time (SF4 b-e). With a median close to 0, these errors will cancel out in the long run. Due to the generally low detection rate, the error made by using $\bar{\eta}_c$ instead of $\bar{\eta}$ is of the same order of magnitude as the event count variance of a stationary Poisson process with a rate of 118 events per 446 sols (event count from ST3).

We thus conclude that the variation of the event rate with time is sufficiently slow to grant a simple averaging of detection efficiencies.

One might argue that we could use the estimated error of $\bar{\eta}$ to compute a corrected detection efficiency, or that we could as well use the KDE rate and the exact averaging of SEQ23. This however would introduce an explicit detection rate model into a process that intends to identify the best detection rate model from a set of candidates. Using the simple averaging injects weaker assumptions into the determination of the intended result.

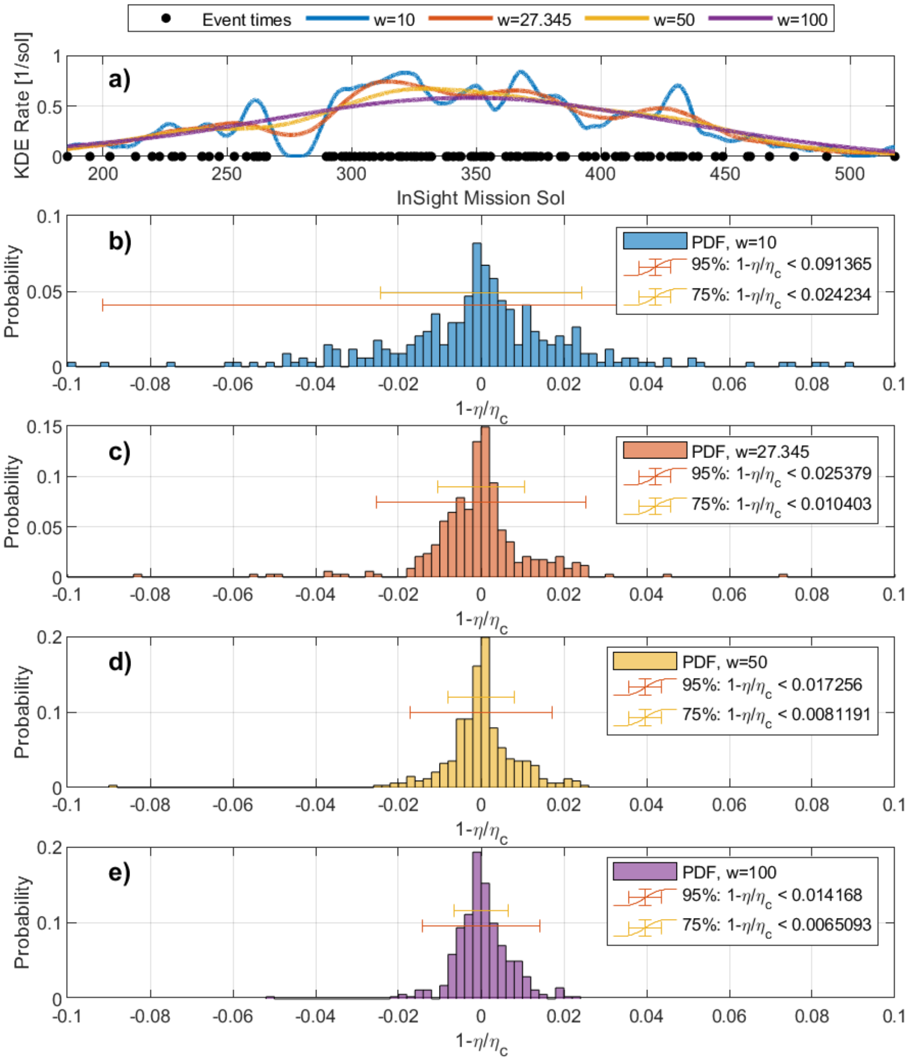


Figure SF4 Detection efficiency averaging error

The distribution of $1-\overline{\eta}/\overline{\eta}_c$ for averaging detection efficiencies computed from 2 minute noise windows to obtain 1 Sol averages. a) event detection times (black dots) and kernel density rate estimations for bandwidths (in sols) as listed in the legend. Colors of KDE rates correspond to histogram colors in panels b-e; b) Probability of $1-\overline{\eta}/\overline{\eta}_c$ values for a bandwidth of 10 sols, 95% of the distribution are within the range marked by the red bar, 75% in the range marked by the yellow bar (both symmetric to 0); c) as b, but for a bandwidth of 27.345 sols; d) 50 sols; e) 100 sols.

Power Law distribution of amplitudes

Figure 3 of the main paper shows that both the amplitudes and SNR values of the 227 HF events of quality classes A, B, and C follow a power law distribution as known from seismic moments and magnitudes: for amplitude or SNR values above a certain threshold, the cumulative distribution is well described by a straight line.

Estimation of the slope of this line uses the same maximum likelihood solution applied to estimate the b-value of the classical Gutenberg-Richter distribution. Amplitudes are converted to the decibel scale using SEQ25. Then the maximum likelihood solution for the slope (also called b here) is (Aki, 1965)

$$b = \frac{log_{10}e}{\overline{A_{dB}} - A_{TH,dB}} \quad \text{(SEQ43)}$$

where $\overline{A_{dB}}$ is the arithmetic mean of the A_{dB} values of all events with amplitudes above $A_{TH,dB}$, and $A_{TH,dB}$ is the threshold amplitude above which the catalog is representative for the seismicity.

This threshold amplitude is estimated using the maximum curvature approach. Instead of a histogram of amplitudes we employ a kernel density estimation (KDE for short; KDE can be considered as a generalization of histograms, where continuous functions are used instead of boxes. The advantage of KDE over histograms is that no binning is required, and continuous functions allow for better approximations of continuous densities than boxes do. Also, the result is a continuous function as well that can be evaluated for any time) of the probability density of amplitudes and SNRs, respectively, employing a gaussian kernel

$$\delta_w(t, t_0) = \frac{1}{w\sqrt{2\pi}} e^{-(t-t_0)^2/2w^2} \quad \text{(SEQ44)}$$

where w is the bandwidth, and the bandwidth estimation of Botev et al. (2010) and Botev (2020). Numerical solutions are given in the main paper.

If $A_{TH,dB}$ and SNR_{TH} are the representativeness thresholds of the catalog, then

$$L = A_{TH,dB} - 20log_{10}(SNR_{TH}) \quad \text{(SEQ45)}$$

is the noise level required to record events with threshold amplitude at threshold SNR.

Time-invariant distribution of amplitudes

Concerning a possible variation in time, the b-value is most interesting, since the a-value is expected to vary with sample size and noise level anyway. We will thus focus on the former here.

Given the small size of our data set, we do not expect that a complex evolution in time of the b-value can be derived. We use distinct time slices and determine the b-value of the amplitude distribution separately, and also the uncertainty of these b-values.

Figure 4 suggests a subdivision of the catalog into four time slices. We select subsets of the event selection table in S3:

(I) Prior to sol 289, when the detection rate increases (events S0185b to S0289a, 22 events)

(II) From sol 289 to sol 336, the first half of the time when the detection rate remains constant (events S0290a to S0338a, 34 events)

(III) From sol 337 to sol 385, the second half of the time when the detection rate remains constant (events S0338b to S0385a, 32 events)

(IV) After sol 385, when the detection rate decreases (events S0392a to S0518a, 30 events)

In addition, we consider

(V) All events (events S0185b to S0518a, 118 events)

For all five groups, we estimate an individual representativeness threshold amplitude and b-value as described in subsection "Power Law distribution of amplitudes". In addition, we compute the standard deviation of the b-value after Shi & Bolt (1982)

$$\sigma_b = \sqrt{\left(\frac{b}{log_{10}e}\right)^2 \frac{n^2}{(n-2)(n-1)^2}} \qquad \text{(SEQ46)}$$

where n is the number of events in the respective time slice.

The survivor cumulative distributions for all time slices are shown in SF5, a graphical comparison of the obtained b-values and their standard deviations is shown in SF6. The individual b-values for groups I to IV are all within one standard deviation of each other, and the b-value of the entire 118 selected events (group V) is within one standard deviation of the individual group's values.

Given than the b-value standard deviation of Shi & Bolt (1982) is an asymptotic solution, i.e. it is strictly valid only for large data sets, and that the four time slices contain less than one third (or fourth) of the 100 events that Shi & Bolt (1982) found necessary to obtain stable b-value estimates for central California, we assume that the formal standard deviations we obtain even underestimate the actual uncertainty.

Therefore, we see no reason to reject the assumption of a time-invariant b-value.

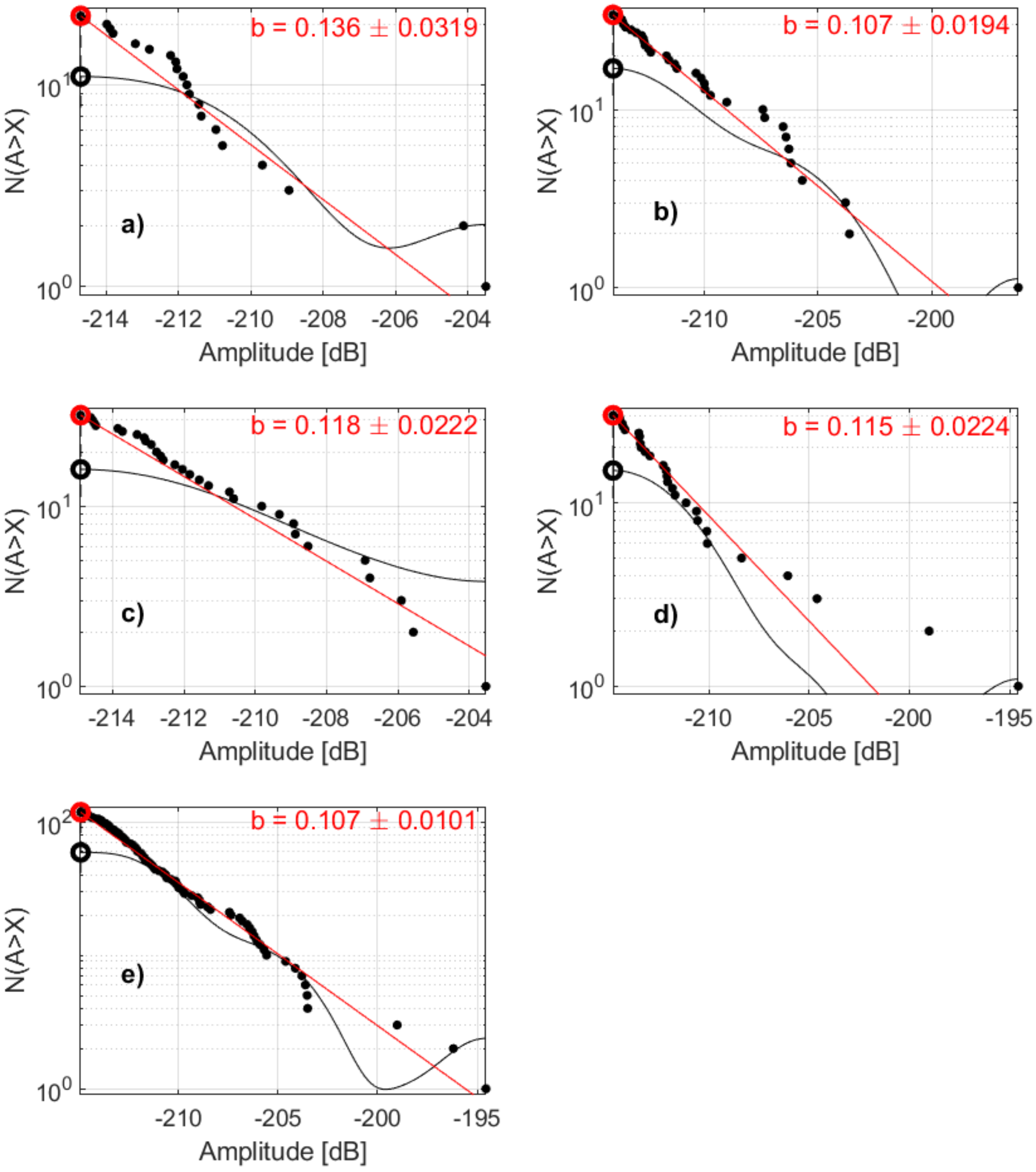


Figure SF5 b-values for subsequent time slices

Each panel shows the survivor cumulative distribution of one of the subsets defined in the text (dots), the Kernel density estimation used to estimate the threshold amplitude (black line), and the fitted power law (red line). b-values and their uncertainty (1σ) are given in each panel. a) group I, b) group II, c) group III, d) group IV, e) group V. 655

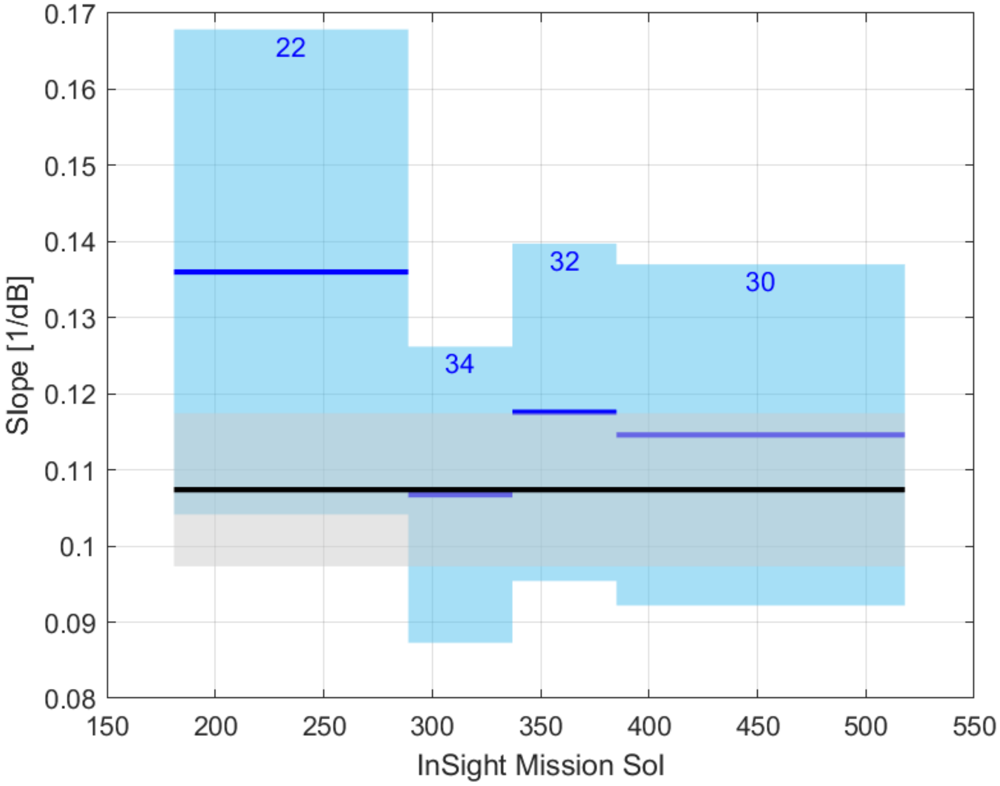


Figure SF6 Comparison of b-values from figure SF5 versus time
Horizontal blue lines: b-values of groups I to IV (numerical values in figure SF5), drawn from the beginning to the end of the respective time slice. Blue boxes: horizontal extent according to start and end times, vertical extent according to standard deviations of the respective b-values. Blue numbers: number of events belonging to each group. Black line and grey box: analogous for group V (all 118 events).

S5 Detection of a time dependent event rate

To detect a possible time dependent component of the seismic event rate, we assume that the events are due to a nonstationary Poisson process (see section S7) and consider the event rate as the sum of a time dependent term, $f(t)$, and a constant background activity, λ_B. Without loss of generality it can be assumed that $\min f(t) = 0$, as a nonzero minimum would just add to the background rate. The number of observed events furthermore depends on the detection efficiency, $\eta_{\mathrm{det}}(t_k)$, as described in section S4. All three of these functions are nonnegative at all times for physical reasons. This gives rise to the observed event rate

$$\lambda_{obs} = \eta_{det}(t)[f(t) + \lambda_B] \tag{SEQ47}$$

Then the expectation for the number of observed events between a time S and a time T is

$$\langle N_{obs}(T)\rangle = \int\limits_S^T \lambda_{obs}(u)\,du = \int_S^T \eta_{det}(u)[f(u) + \lambda_B]\,du \tag{SEQ48}$$

For $\eta_{det}(t)$, we can determine an empirical time series as described in S4, and the actually observed events provide an estimation of $\langle N_{obs}(T)\rangle$. We can thus compute an expected event number for any constant event rate λ_{const}, and subtract it from $\langle N_{obs}(T)\rangle$, yielding a residual event count

$$\langle N_{Res}(T)\rangle = \int_S^T \eta_{det}(u) f(u) du - (\lambda_B - \lambda_{const}) \int_S^T \eta_{\mathrm{det}}(u) du \tag{SEQ49}$$

By choosing λ_{const} such that $\langle N_{Res}(T)\rangle$ is minimized, one can estimate the background rate.

After inserting the properly chosen λ_{const}, a non-zero minimum residual rate is an indication for the existence of a time dependent component in the observed event rate, although $\langle N_{Res}(T)\rangle$ should not be confused with $f(t)$ itself.

Since the observed events are only one realization of a stochastic process, they constitute only an estimation for the expected event count. If the events result from a stationary Poisson process, $\langle N_{Res}(T)\rangle$ will be distributed with mean zero and a variance according to the rate of that process.

Expressing the actual residual in units of that variance indicates if a statistically significant deviation from a stationary Poisson process exists.

S6 Approximate Detection Efficiency per Sol

In order to implement an efficient evaluation of the detection efficiency in the maximum likelihood optimization of event rate models, we compute average efficiencies for each sol covered by the event catalog, and approximate this series of discrete samples by a polynomial in terms of the Julian Day. Although the wind speed, as main driver for the observed background noise, may vary considerably from sol to sol, it turns out that a smooth approximation is sufficient and the polynomial degree does not have to be close to the number of sols but may be low (see comparison shown in the main paper).

From the initial short time (2 *min*) windows for which we computed the noise standard deviation, we first compute an average detection efficiency for each sol, resulting in 537 efficiency values between 5×10^{-4} and 0.484, for sols 73 through 630. To stabilize the polynomial fit, we apply a centering and scaling transformation by converting the x values (the sol numbers) according to

$$\hat{x} = (x - \bar{x})/\sigma_x \qquad \text{(SEQ50)}$$

where $\bar{x} = 354.7672$ is the mean and $\sigma_x = 163.264$ the standard deviation of the sol numbers from 73 to 630, less sols 77 and sols 268 through 287, during which no data is available.

The polynomial is then

$$\eta_{eff}(x) = \sum_{i=0}^{n} a_i \hat{x}^i \qquad \text{(SEQ51)}$$

with coefficients as given in Table ST4.

For polynomial degrees beyond 23, the Vandermonde matrix inverted during the polynomial fitting process is ill-conditioned. Modeling the efficiency curve in all detail is thus not possible with this approach. We however find that the polynomial misfit improves only slowly once the degree is larger than 6, and that the residual event count resulting from low-degree fits differs little from higher-degree fits. Since the degree 6 polynomial overshoots at the end of the relevant time interval, we adopt the degree 7 polynomial defined by Table ST4.

Table ST4 Polynomial coefficients to approximate the per-sol detection rate

i	a_i
0	0.1470760164
1	-0.01372081373
2	0.02311633362
3	-0.1492531704
4	-0.06854420987
5	0.1184768616
6	0.01544226739
7	-0.02298400277

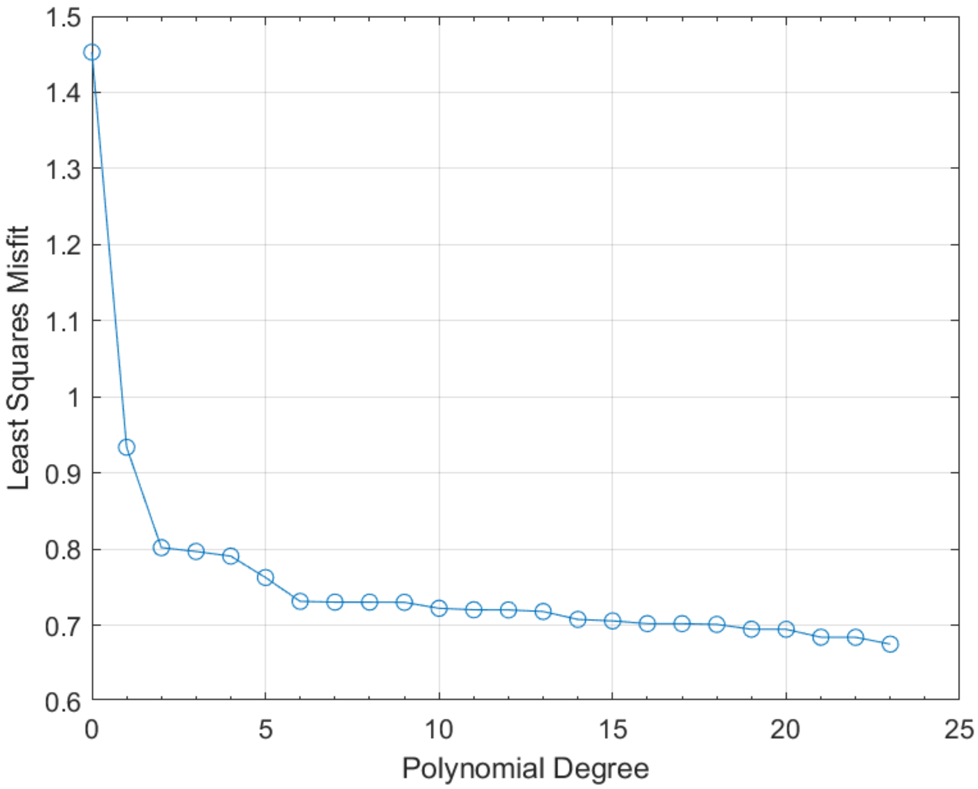


Figure SF7 Polynomial misfit for approximation of detection efficiency
This is the least squares misfit obtained by fitting an n-th order polynomial to the detection efficiency curve shown in figure 4 of the main paper. See text for which degree was finally chosen, and why.

S7 Poisson processes

Following the extensive overview given by Zhuang et al. (2012), we summarize here those properties, as well as symbols and terminology, of Poisson processes that we use as basis for diagnostic analyses of event sequences.

From the physical point of view, a Poisson process is a memoryless process which produces events with a certain probability per unit time. The event rate λ is the only characterizing parameter. Since the process does not remember past occurrences, the numbers of events occurring in two disjointed time intervals are independent of each other. A Poisson process is called *stationary* if the rate of events per time does not change with time (*non-stationary* otherwise), and it is called *simple* when no more than one event can occur at any time. Simplicity is a property favorable for the mathematical description and can usually be imposed by measuring time with a sufficiently high resolution.

The number of events in a realization of a stationary Poisson process is Poisson distributed with mean λ and standard deviation

$$\sigma = \sqrt{\lambda} \qquad \text{(SEQ52)}$$

For λ larger than about 10 per time unit, the Poisson distribution is well approximated by a Gaussian distribution with the same mean and standard deviation.

For a simple stationary Poisson process, the expectation of the number of events occurring between start time S and end time T is

$$\langle N \rangle = \lambda(T - S) \qquad \text{(SEQ53)}$$

For a simple nonstationary Poisson process, it is

$$\langle N \rangle = \int_S^T \lambda(u)du \qquad \text{(SEQ54)}$$

S8 Spike Train Fourier Transform, with application to Phobos

Bulow et al. (2007) introduced the spike train Fourier transform to lunar seismology as a tool to investigate periodicities in event occurrence. This method is used e.g. in Neuroscience, where activity patterns of single neurons are analyzed. The Fast Fourier Transform (FFT) is not well suited to analyze the sequence of events rather than their individual waveforms, as a series of source times is not a band-limited waveform in the sense of the FFT, but a sequence of delta functions which violate Nyquist's sampling theorem and the FFT condition of regular sampling.

Given a sequence of N events that occurred at times $t_i \in \{t_1, t_2, t_3, \dots, t_N\}$, a function

$$f(t) = \sum_{i=1}^{N} \delta(t - t_i) \qquad \text{(SEQ55)}$$

is defined, where $\delta(t)$ is the Dirac delta function. The Fourier transform of this function can be calculated analytically by Fourier transforming the individual terms of the defining function

$$F(\omega) = \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-i\omega t_i} \qquad \text{(SEQ56)}$$

and can be evaluated quickly at any frequency, i.e. unlike an FFT spectrum, a spike train spectrum is not limited to selected frequencies that depend on the number of input samples.

We apply this method for periods in a range relevant for the orbit of Phobos (Table ST5) and for a period of 1/2 and 1 sol. The obtained spectra are shown in Figure SF8. In Figure SF9 we show two spectra from simulated catalogs where events were either distributed uniformly over the night, or with an 80:20 subdivision between 17:00 to 24:00 LTST (80%) and 00:00 to 06:00 LTST (20%), roughly corresponding to the wind-induced non-uniform distribution in the real catalog, as experienced from spring 2019 to summer 2020.

Table ST5 Orbital periods of Phobos.
The periods listed in the table are: draconitic: period of crossing the same orbital node, anomalistic: time between two periapsis crossings, synodic: time to the repetition of the same phase (e.g. full moon), culmination: period at which Phobos crosses the local meridian of the observer

Period	Duration [sol]	Reference
draconitic	0.31027	Rambaux et al., 2012
anomalistic	0.31046	Rambaux et al., 2012
synodic	0.31056	Rambaux et al., 2012
culmination	0.43832	Bills & Comstock, 2005

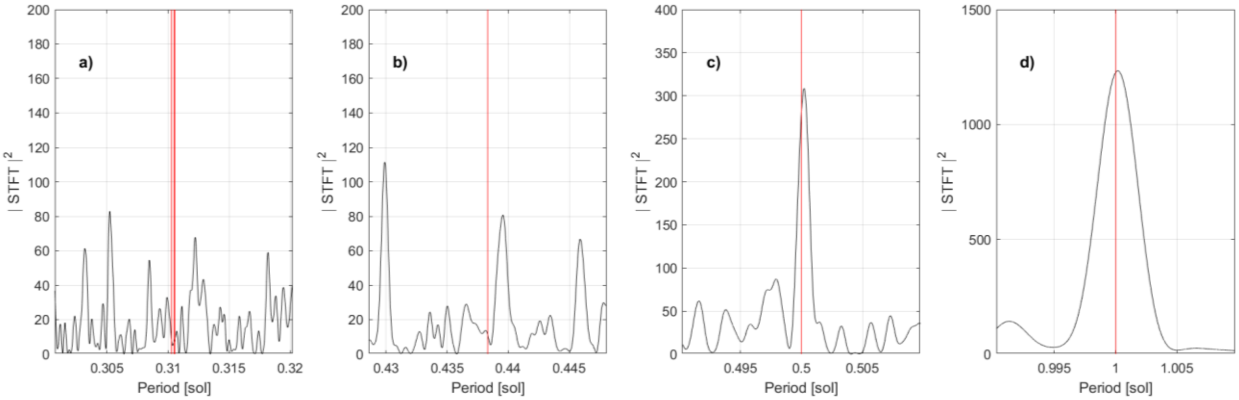


Figure SF8 Spike Train Spectra for the selected 118 HF marsquakes.
Event occurrence power spectral density estimated from the spike train Fourier transform. a) PSD in a period window of 0.02 sol (approx. 30 *min*) width centered on the anomalistic period of Phobos' orbit. Vertical lines indicate periods of Phobos' draconic, anomalistic, and synodic revolution, b) 0.02 sol window centered on Phobos' culmination period, c) 0.02 sol window centered on 0.5 sol (note the different amplitude scale), d) 0.02 sol window centered on 1 sol, for comparison (note the different amplitude scale)

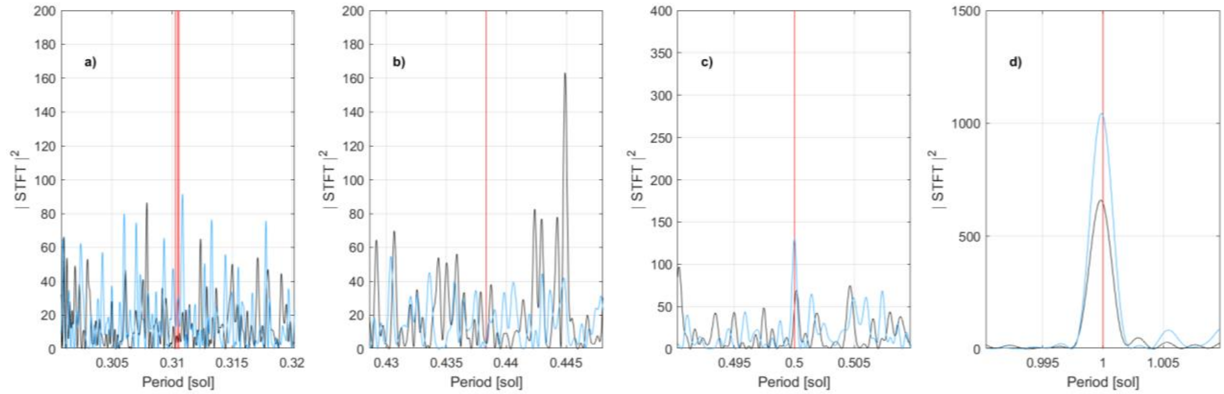


Figure SF9 Spike Train Spectra for synthetic catalogs.
Event occurrence power spectral density for two random event sequences. Black: uniform distribution between 17:00 LTST and 06:00 LTST, 116 night time events, no day time events between 06:00 LTST and 17:00 LTST; Blue: 92 events distributed uniformly between 17:00 LTST and midnight, 25 events distributed uniformly between midnight and 06:00 LTST, no events during day time. Event numbers in both cases are the result of a random process to generate about the same overall number of events as in the observed catalog. Other plot details as in Figure SF8

S9 Maximum Likelihood Estimation of rate function parameters

The central question of our study is if the event rate λ of a certain class of Marsquakes is a function of time and a set of parameters, i.e.

$$\lambda = \lambda(t, \theta) \tag{SEQ57}$$

where θ is the parameter vector. Given a set of N events that occurred during an observation interval from a time S to a time T at times

$$t_i \in \{t_1, t_2, t_3, \dots, t_N\}, \text{ with } S \le t_i \le T \tag{SEQ58}$$

the log-likelihood of a parameter vector θ is (e.g. Ogata, 1983)

$$\log L(\theta) = \sum_{i=1}^{N} \log \lambda(t_i, \theta) - \int_S^T \lambda(u, \theta)\, du \tag{SEQ59}$$

The preferred parameter values are those at which the likelihood L, or the log-likelihood $\log L$, attains its global maximum, i.e. the parameter vector θ_{Max} for which $L_{Max} = \max_{\theta} L\,(\theta)$.

In the case of a constant event rate $\lambda = \lambda_0$, eq. SEQ59 can be simplified considerably. The Maximum Likelihood estimation of the rate itself is well known to be simply

$$\lambda_0 = \frac{N}{T-S} \tag{SEQ60}$$

with standard deviation, normalized to unit time,

$$\sigma_\lambda = \frac{\sqrt{N}}{T-S} \tag{SEQ61}$$

Inserting λ_0 into eq. SEQ59 yields

$$\log L = N\left(\log \frac{N}{T-S} - 1\right) \tag{SEQ62}$$

for the likelihood of rate λ_0.

S10 Nested grid search algorithm

All our rate models are defined on a four- or five-dimensional parameter space (depending on whether period is a free parameter or inherent to the underlying physical process). Model parameters are collected in a parameter vector θ

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \end{pmatrix} = \begin{pmatrix} A \\ T \\ \varphi \\ K \\ \lambda_B \end{pmatrix} \qquad \text{(SEQ63)}$$

where A is a proportionality constant, T is the period of the kernel function (in most of our functions, this is implicitly fixed to the duration of the Martian year, i.e. about 687 days or 668 sols), φ is a retardation, either in radians or days, K is a constant offset, and λ_B is a constant baseline event rate (assuming that an endogenous, stationary Poisson process independently of the exogenously driven time-variable process, may be present). Period, retardation and the baseline rate are all restricted to positive values: A negative period is meaningless, a negative retardation would be acausal, and a negative baseline rate would mean to undo events after they occurred, which is impossible.

The actual event rate is computed from these parameters as

$$\lambda(t,\theta) = \Upsilon(t)\, \eta_{det}(t) \max\left(\lambda_B, f\left(t, \theta_{\setminus \lambda_B}\right) + \lambda_B\right) \qquad \text{(SEQ64)}$$

using a model-specific, time dependent kernel function f, the detection efficiency $\eta_{\mathrm{det}(t)}$ as described in S4 and the uptime-switch

$$\Upsilon(t) = \begin{cases} 0 & during\ downtimes \\ 1 & else \end{cases} \qquad \text{(SEQ65)}$$

to handle data gaps explicitly. The kernel function is not dependent on the baseline rate, $\theta_{\setminus \lambda_B}$ hence denotes a reduced parameter vector not containing λ_B. Since detection efficiency and uptime switch have to be evaluated for the time of detection, the phase lag parameter φ does not apply to them.

The likelihood function SEQ59 is then a scalar-valued function of a point θ in the parameter space.

We start with an exploratory grid which covers a sufficient range of values for each parameter - sufficient means, that either the entire meaningful range of parameters is covered, as in the case of the period which we expect to be at most one Martian year, or the likelihood falls to low values and does not appear to rise again. Since our parameter spaces have 4 to 5 dimensions, these limits are estimated by inspecting cross sections along one dimension.

The initial grid is then defined based on the exploratory grid in order to contain the region of the anticipated likelihood maximum, but still large enough to give an overall impression of the likelihood function. The grid mesh width is defined such that an M-dimensional space is covered with m^M grid nodes, i.e. the number of nodes is equal to m in all dimensions.

The grid node with the highest likelihood value is then taken as the center of a new grid. This new grid consists of the same number of m^M nodes, again with m nodes in each direction, but covers a smaller volume. This keeps memory and CPU time consumption predictable.

During each iteration, the volume of the grid is reduced by a factor $F = 3$ in case of 4-D parameter spaces, and $F = 2$ in case of 5-D parameter spaces. Since high dimensional spaces are unintuitive, and since test runs with faster volume reduction sometimes missed a known maximum in favor for a lower local maximum, we chose a cautious approach here.

Centering the new grid on the maximum of the previous one and restricting parameters to individual ranges of possible values (e.g. positive baseline rate) can result in a conflict. To avoid convergence to a maximum at physically impossible values of parameters, although mathematically the global maximum, we shift the grid volume in such cases to keep all parameters in the admissible range rather than center it on the expected maximum location.

If the searched volume is reduced with each iteration, this search method falls into the class of "greedy" algorithms that are prone to premature convergence. Such a behavior was sometimes observed when the largest likelihood value happened to be found at the boundary of the searched volume. In such cases, we introduce a *growth* of the search range of the affected parameter(s) by a factor slightly less then what is used to shrink the search volume, and center the new range on the midpoint of the previous one.

Sampling each dimension with the same number of grid nodes, and reducing the volume by a factor F in each step, results in a predictable precision of each parameter after k iterations, which will be $F^{-k/M}/(m-1)$ times the width of the initial grid in the direction of that parameter. This precision is however unrelated to the statistical uncertainty of the resulting maximum likelihood estimate. Also, the actual width of the interval might deviate if grid volume reduction conflicted with the admissible parameter ranges, or when growth steps are conducted.

The iteration is stopped if the value of the log-likelihood remains stable to a certain relative precision (actual value are tabulated with the search grid parameters in Table ST13 and Table ST14 of section S21) of its absolute value during the last three grid iterations. Considering the three rather than two last iterations helps avoiding premature convergence.

S11 Likelihood variance

The value of the likelihood at the likelihood maximum carries a sampling uncertainty due to the stochastic nature of both the number and the occurrence times of the evaluated events. We estimate the variance of the Likelihood value at its maximum using a delete-1-Jackknife approach. This is a special case of the widely used bootstrapping, but does not produce ties. By the more invasive resampling of bootstrapping, we expect the location of the likelihood maximum to move within the parameter space, which would require a complete re-evaluation of the grid iteration. This effect is minimized with the delete-1-Jackknife.

Given a set of N observations, $\{t_1, t_2, t_3, \dots, t_N\}$, the log-likelihood of a model is given by eq. SEQ59. We call the log-likelihood computed from the entire data set $\log L_\forall$, with

$$\log L_\forall = \log L(t_1, t_2, t_3, \dots, t_N) \tag{SEQ66}$$

In a delete-1-Jackknife, the data set is resampled by omitting the i-th data point

$$\log L_{\setminus i} = \log L(t_1, t_2, t_3, \dots, t_{i-1}, t_{i+1}, \dots, t_N) \tag{SEQ67}$$

This evaluation is repeated for all i, and the statistic of interest (i.e. the standard deviation of $\log L_\forall$) is obtained from the N different values which $\log L_{\setminus i}$ can attain.

In the following, we assume that the omission of any single datum does not change the location of the likelihood maximum within the parameter space, but only the value of the likelihood function. The purpose of this assumption is that, under this condition, it is not necessary to re-evaluate the grid search (or whatever optimization method one uses).

The Jackknife estimator for the variance $\sigma_\forall^2$ of $\log L_\forall$ is given by (e.g. Efron & Stein, 1981)

$$\sigma_\forall^2 = \frac{N-1}{N} \sum_{i=1}^{N} \left(\log L_{\setminus i} - \overline{\log L_{\setminus \imath}}\right)^2 \tag{SEQ68}$$

where the overline denotes the arithmetic mean

$$\overline{\log L_{\setminus \imath}} = \tfrac{1}{N} \textstyle\sum_{i=1}^{N} \log L_{\setminus i} \tag{SEQ69}$$

The specific structure of the log-Likelihood function for event rate functions allows for an efficient evaluation of the delete-1-Jackknife variance estimator. The only place where the individual observations t_i enter the equation is the sum term in eq. SEQ59. Omitting the i-th observation is thus equivalent to omitting the i-th term from this sum, or to subtracting the i-th term from $\log L_\forall$, i.e.

$$\log L_{\setminus i} = \log L_\forall - \log \lambda(t_i) \tag{SEQ70}$$

Inserting this into the arithmetic mean yields

$$\overline{\log L_\imath} = \tfrac{n-1}{n} \log L_\forall - \tfrac{1}{n} \textstyle\int_S^T \lambda(t) dt \tag{SEQ71}$$

and by inserting eqs. SEQ70 and SEQ71 into SEQ68, we obtain for the standard deviation

$$\sigma_{\forall}^{2} = \frac{n-1}{n} \sum_{i=1}^{n} \left[\frac{1}{n} \left(\log L_{\forall} + \int_{S}^{T} \lambda(t) dt \right) - \log \lambda(t_i) \right]^2 \qquad \text{(SEQ72)}$$

where all terms under the summation are available already from the computation of $\log L_{\forall}$.

It must be noted that the Jackknife estimation described here does not work in the case of a constant event rate, since for a constant rate, all $\log L_{\setminus i}$ are identical and the resulting variance estimation is 0.

S12 Parameter variance

The confidence interval of the j-th parameter θ_j is computed from the observed Fisher information matrix (e.g. Ogata, 1983)

$$J_n(\theta_{Max}) = -\nabla^2 \log L(\theta_{Max}) \tag{SEQ73}$$

The diagonal elements of its inverse give the variance of the corresponding parameter. Therefore, the error of the maximum likelihood solution is asymptotically Gaussian distributed, and the confidence factors familiar from the Gaussian distribution can be applied to obtain the confidence interval, for a confidence level α,

$$\theta_j = \theta_{MLE} \pm \sqrt{2} erf^{-1}(\alpha) \sqrt{[J_n(\theta_{MLE})^{-1}]_{jj}} \tag{SEQ74}$$

Since the actual maximum is not known, we use the numerical maximum likelihood estimate $\theta_{MLE} \approx \theta_{Max}$ as approximation.

For a confidence level of $\alpha = 0.95$, one obtains $\sqrt{2} erf^{-1}(\alpha) \approx 1.95996$.

We compute the Hessian matrix $\nabla^2 \log L(\theta_{Max})$ using the MatLab® package DERIVEST (d'Errico, 2014). This package computes the Hessian matrix with no respect to the restriction that its inverse is supposed to be a variance matrix, i.e. that all elements of the inverse Hessian matrix must be positive. The DERIVEST evaluation is also independent of the search grid we use but constructs its own sampling.

The likelihood function at hand consists of a series of more than one hundred terms (one per event) of the form $\log \lambda(t,\theta)$, with $\lambda(t,\theta)$ as defined in the main text. The standard deviations according to eq. SEQ74 depend on all 16 or 25 second partial derivatives of the likelihood function and thus require that (a) all these derivatives exist (i.e. the likelihood function is sufficiently smooth), (b) the resulting Hessian matrix is not singular, (c) its determinant is not too close to 0, and (d) all numerical methods involved converge.

If these conditions are not fulfilled, and a negative or implausibly large variance results for one of the parameters (which casts doubts on the variances of all other parameters, since all variances are interconnected via the matrix inversion), we attempt to obtain a handle on uncertainty via the Jackknife estimation of the uncertainty of the likelihood value at its maximum (S11): We identify all nodes of the exploratory search grid (the final grids prove to cover a volume that is too small for this approach) for which the log-likelihood is larger than $\log L_\forall - \sigma_\forall$ (in the notation of S11), and determine the parameter boundaries which contain these nodes.

Comparison of the uncertainty intervals defined this way with the Fisher information matrix inversion - for those cases where the Fisher information matrix inversion resulted in credible values (although some are still implausibly large) - suggests that this approach typically overestimates the standard deviations of parameters by a factor of about 3 (Figure SF10, using the results listed in Table ST15 and Table ST16).

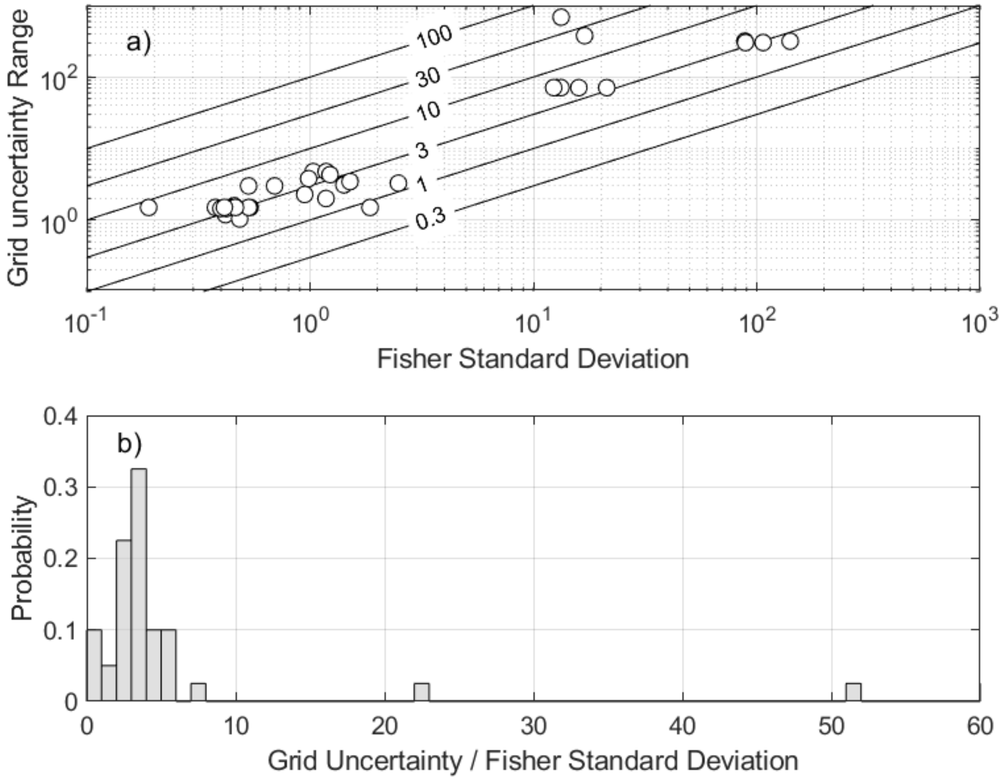


Figure SF10 Comparison of parameter uncertainty estimations.
a) Interval width obtained from comparison of grid nodes plotted against the standard deviations that were obtained from Fisher information matrix inversion. Slanted lines indicate ratios according to their labels. b) histogram of the ratio between corresponding uncertainty estimations (horizontal axis is clipped to gain resolution at its left end), bin width is 1, a total of 39 intervals is compared.

S13 Akaike Information Criterion AIC

Maximization of the log-Likelihood provides the best values of the parameters for a given model, and also allows comparing the solutions obtained for different models. Since models with more free parameters are usually expected to provide better fits, the number m of model parameters must be accounted for. We employ the Akaike Information Criterion (AIC, Akaike, 1974, and references therein) for this purpose, which is an application of Kullback-Leibler information (Kullback & Leibler, 1951), which allows asking how much information is lost when data is described by a - necessarily simplified - model rather than by the real processes. In its original form, the AIC states that

$$AIC = -2 \max_{\theta}(\log L) + 2m \qquad \text{(SEQ75)}$$

has to be minimized for the comparison of different models (e.g. Ogata, 1983).

Since our data set is small compared to e.g. terrestrial event catalogs, we apply Sugiura's correction (cited after Liddle, 2007) to compute the corrected AIC

$$AIC_c = -2 \max_{\theta}(\log L) + 2m + \frac{2m(m-1)}{N-m-1} \qquad \text{(SEQ76)}$$

which accounts for the number of observations, N.

S14 Akaike Weights and Evidence Ratio

(This section follows the concise description in Wagenmakers & Farrell, 2004, and the more fundamental thoughts of Burnham et al., 2011.)

The difference between the AIC (or the AIC_c) of any model and that of the model with the lowest AIC value

$$\Delta_i = AIC_i - \min_{models} AIC \qquad \text{(SEQ77)}$$

provides a ranking of all investigated models and is widely used to identify the best model. It does however not provide a quantitative means of telling how much better model i is in comparison to a model j. Akaike weights and evidence ratios intend to overcome this.

From AIC differences, we first compute the likelihood of the *models* (rather than the likelihood of certain *parameter values* in a single model), which is given by

$$\mathcal{L}_i \propto e^{-\Delta_i/2} \qquad \text{(SEQ78)}$$

and the Akaike weights

$$w_i = \frac{\mathcal{L}_i}{\sum_{n=1}^{\mathcal{N}} \mathcal{L}_n} \qquad \text{(SEQ79)}$$

which normalize the model likelihoods such that they sum up to 1, while all proportionality factors cancel out. The ratio between the Akaike weight of the i-th model to that of the model with the lowest AIC is then the evidence ratio

$$\mathcal{R} = \frac{w_{best}}{w_i} \qquad \text{(SEQ80)}$$

and tells that the best model is $\mathcal{R}$ times more likely to be the best model (in terms of Kullback-Leibler information loss) than the i-th model.

If models come in groups that differ by parameter selection rather than conceptual, one can as well weight e.g. all models of group A against all models of group B, and thus arrive at rankings of model groups by finding that group A is

$$\frac{\sum_{(i \in group\ A)} w_i}{\sum_{(i \in group\ B)} w_i} \qquad \text{(SEQ81)}$$

times more likely to contain the best model than group B.

S15 Areocentric ascension and declination of the Sun

The coordinates of the Sun in the Martian sky are the true solar right ascension

$$\alpha_S = \tan^{-1}[\cos(\varepsilon)\tan(L_S)] \tag{SEQ82}$$

and the solar declination

$$\delta_S = \sin^{-1}[\sin(\varepsilon)\sin(L_S)] \tag{SEQ83}$$

where ε is the inclination of the apparent orbit of the Sun to the Martian equator, and L_S the areocentric longitude of the Sun (Allison & McEwen, 2000, also S16 for the computation of L_S).

For 01. Jan. 2020, $\varepsilon \approx 25.1944°$, thus the argument in the inverse sine in SEQ83 is always smaller than $\sin\varepsilon \approx 0.426$, and therefore $|\mathrm{x} - \sin^{-1} x| \leq 0.0145$ for all values of L_S. Hence δ_S is proportional to $\sin L_S$ to within 1.5 %, and the illumination of a point on Elysium Planitia (or any other horizontal plane surface on Mars) evolves approximately like $\sin L_S$ throughout the year.

S16 Computation of the Areocentric Longitude of the Sun, L$_S$

The progress of seasons on Mars is usually described by the areocentric longitude of the Sun, L_S, i.e. the angle between the current location of the sun in the sky and its location at the vernal equinox, as seen from the center of Mars. Approximative equations to compute L_S are given by Allison (1997) and Allison and McEwen (2000), and in concise form by Schmunk (2020). Since the mathematical structure of this approximation is not obvious from their step-wise evaluation of different contributions, and for a faster numerical evaluation, we combine the equations found there into a single, closed-form equation consisting of a finite series

$$L_S(JD_{TT}) = \sum_{i=1}^{13} a_i \left(b_i JD_{TT} + c_i\right) \sin\left[\frac{\pi}{180^\circ}\left(d_i JD_{TT} + e_i\right)\right] \qquad \text{(SEQ84)}$$

which gives L_S in degrees as function of the Julian Date in Terrestrial Time, JD_{TT}. With the coefficients given in Table ST6, eq. SEQ84 describes an angle that advances by about $360°$ per Martian year, but slightly undulates about a linear increase. The dominating period of L_S (when stripping the linear increase) is the tropical year with a duration of 686.9726 d (Allison & McEwen, 2000)

The InSight event catalog lists arrival times in UTC. To evaluate L_S, we first convert UTC into Terrestrial Time TT, i.e. add a constant offset and the current number of leap seconds (namely 37 s, International Earth Rotation and Reference Systems Service, 2020).

$$TT = UTC + 32.184s + 37s \qquad \text{(SEQ85)}$$

The input to eq. SEQ84, the Julian Date JD_{TT}, is obtained by converting TT using an algorithm from Meeus (1994).

Table ST6 Coefficients for computation of L_S using eq. SEQ84.
Since L_S results in degrees, all coefficients are either in degrees, or in degrees per day. Benchmarking of eq. SEQ84 against a strict implementation of the original equations results in deviations smaller than 1.7×10^{-7} degrees in the 100 Martian years following 01. Jan. 2018. The reader should be aware that truncation of the coefficients likely increases this deviation.

i	a_i [°]	b_i [°/(d)]	c_i [°]	d_i [°/d]	e_i [°]
1	+1	+0.524038496	-1284433.56757632	0	+90
2	+1	+0.0000003	+9.9555365	+0.52402073	-161.01342785
3	+0.623	0	+1	+1.04804146	-322.0268557
4	+0.05	0	+1	+1.57206219	-123.04028355
5	+0.005	0	+1	+2.09608292	-284.0537114
6	+0.0005	0	+1	+2.62010365	-85.06713925
7	+0.0071	0	+1	-0.440936913777818	+297.277287441
8	+0.0057	0	+1	-0.3578500102993706	+247.2304993705
9	+0.0039	0	+1	-0.8818343771741581	+314.821189421
10	+0.0037	0	+1	-0.06243436100031397	+128.9095385147
11	+0.0021	0	+1	-0.4615651790613264	+142.1029019
12	+0.002	0	+1	-0.399135937218578	+14.1832085188
13	+0.0018	0	+1	-0.03000448360749107	+158.2467655267

S17 Time Derivative of Areocentric Longitude of the Sun

The rate of change of the Areocentric Longitude of the Sun can be calculated by differentiating equation SEQ84 term by term. The algebraic expression of the derivative is

$$\frac{d}{dx} L_S(JD_{TT}) = \pi \sum_{i=0}^{13} a_i (b_i JD_{TT} + c_i) \sin\left(\frac{\pi}{180}(d_i JD_{TT} + e_i)\right) \qquad \text{(SEQ86)}$$

which gives the derivative in degrees per day as function of Julian Date in Terrestrial Time, JD_{TT} (see SEQ85 for the relation between TT and UTC). The coefficients are given in Table ST7.

Table ST7 Coefficients for computation of the time derivative of L_S using eq. SEQ86.
All angular parameters are reduced to the interval $0° \cdots 360°$. Terms are sorted by the magnitude of a_i, which does not correspond to the sorting in Table ST6. The reader should be aware that truncation of the coefficients likely reduces the accuracy of results.

i	a_i [°/d]	b_i [°/d^2]	c_i [°/d]	d_i [°/d]	e_i [°]
0	0.524038496	0	1	0	0
1	0.0000003/π	0	1	0.52402073	-161.01342785
2	0.002911226277777778	0.0000003	9.9555365	-0.52402073	+251.01342785
3	0.003627387942111111	0	1	-1.04804146	+52.0268557
4	0.0004366839416666667	0	1	-1.57206219	+213.04028355
5	0.00005822452555555556	0	1	-2.09608292	+14.0537114
6	0.00001910641150544009	0	1	-0.8818343771741581	+224.821189421
7	0.00001739251159901393	0	1	-0.440936913777818	+207.277287441
8	0.0000113319169928134	0	1	-0.3578500102993706	+157.2304993705
9	0.000007278065694444444	0	1	-2.62010365	+175.06713925
10	0.000005384927089048808	0	1	-0.4615651790613264	+52.1029019
11	0.000004434843746873089	0	1	-0.399135937218578	+284.1832085188
12	0.000001283372976117565	0	1	-0.06243436100031397	+38.9095385147
13	0.0000003000448360749107	0	1	-0.03000448360749107	+68.2467655267

S18 Surface Load model and Atmospheric pressure

CO2 ice and load induced seismicity

Hess et al. (1979) report a variation of atmospheric pressure at the Viking I landing site from 685 *Pa* ($L_S \approx 150°$) to 890 *Pa* ($L_S \approx 256°$). Due to the different elevations of the Viking I and II landing sites, pressure measured by Viking II exceeds that of Viking I by about 50 *Pa*, but varies with about the same amplitude. The offset is not constant, likely due to stratification in the atmosphere, and weather effects specific for the mission timespan of Viking (Hess et al., 1979). This annual pressure cycle reflects the CO_2 sublimation/condensation cycle of the polar ice caps (Hess et al., 1979, Forget, 1998, and references therein), which were found to have a seasonal growth cycle by Herschel (1784). We represent the seasonal pressure cycle, as observed by Insight (Banfield et al. 2018, 2020) by fitting the observed diurnal mean values with two harmonics as a function of solar longitude (S18). As shown by Hourdin et al. (1995) for Viking pressure data, the periodical annual pressure cycle can be well fitted by such harmonics because all forcings of the CO_2 cycle (planet inclination, distance to the Sun) are harmonics of an annual period in terms of solar longitude.

Direct observation of the CO_2 deposits using the Mars Odyssey Gamma Ray Spectrometer confirmed the Hess et al. (1979) estimation of mobilized mass (Kelly et al., 2006). Condensing all CO_2 in the region south of 55° S (cf. Calvin et al., 2017) concentrates the global atmospheric pressure change (172 *Pa* observed by InSight) on less than 10 % of the planetary surface, and thus increases the surface load by almost 2 *kPa* or a layer of about *33 cm* thickness (using $\rho_{CO2} = 1530\ kg/m^3$,CO_2 ice at $-79℃$, Hess et al., 1979). Elevation changes due to ice cover evolution were measured directly using the MOLA laser altimeter (Zuber et al., 1992). Values exceeding 1 *m* occur close to the poles, although the largest mass mobilization takes place at lower latitudes (Smith et al. 2001, Jian & Ip, 2009). Variations of topographic elevation of about 0.5 *m* also occur at lower latitudes, e.g. at the Tharsis volcanoes and in Valles Marineris, but also over wide regions of the southern highlands (Aharonson et al., 2004). The load change within 20° to 30° from InSight is likely smaller than 2 *kPa*, with the atmospheric pressure change as lower limit.

The long coda duration of the HF event recordings suggests wave propagation in a scattering high-Q medium, both indicating a propagation through the crust (van Driel et al., 2021). Assuming that the HF events occur south of 55°S, in contrast, requires propagation through the mantle, and implausible velocities.

We nevertheless prefer assuming ice load over a direct effect of atmospheric pressure. First, because the atmosphere is so thin. And second, because reservoir induced seismicity typically, although not exclusively, occurs near the rims of the reservoirs, not directly underneath them (e.g. Saxena, 1988, Roeloffs, 1988): an ice cover is spatially limited, while atmospheric pressure acts everywhere.

Saxena et al. (1988) model induced seismicity of a porous elastic material, based on the Mohr-Coulomb failure criterion and assuming vertical deformation under a laterally extended load, where fluid pressure within the rock is modified depending on the rock's permeability, and the load change during water level changes may push the rock beyond the failure envelope. They find that rock stiffness and strength, but also permeability, pre-existing stress and the entire filling history (via irreversible permeability changes) play a role. Reservoir mass and water level change rate are key parameters of their model. We thus base our model on the pressure rate rather than pressure itself.

The physics of pore pressure changes in permeable rock suggests the use of a rate-and-state friction model (Dieterich, 1994). We however refrain from doing so, since this type of model requires parameters like the tectonic loading rate, aftershock relaxation time, or the rock frictional resistance, about which no information is available for Mars. Even when inverting for all of them as free parameters, there is currently no way to judge the plausibility of results - not knowing the HF source locations, we do not even know for which region to estimate e.g. the tectonic load rate.

Approximation of InSight surface pressure and pressure rate

The surface pressure on Mars follows an annual cycle due to the CO_2 evaporation/deposition cycle, and can be described by a periodic function $P_S(L_S)$ of the areocentric solar longitude with a 360° period. We decompose it into a Fourier series

$$P_S(L_S) = a_0 + \sum_{k=1}^{\infty} [a_k \cos(kL_S) + b_k \sin(kL_S)] \tag{SEQ87}$$

InSight so far measured air pressure for $0° \leq L_S \leq 200°$ and $305° \leq L_S \leq 360°$. We filled the gap from 200° to 305° with predictions from the Mars Climate Database 5.3 (Forget et al., 1999, Millour et al., 2018), evaluated for the InSight site to allow for a direct evaluation of Fourier coefficients. A similar approach was taken in preparation for the landing of the Mars Science Laboratory (Withers, 2012).

It turns out that a truncation after the second harmonic is already a very good approximation for surface pressure at Homestead Hollow, i.e. we use the truncated series

$$P_{NSYT}(L_S) = a_0 + \sum_{k=1}^{2} [a_k \cos(kL_S) + b_k \sin(L_S)] \tag{SEQ88}$$

where L_S is the areocentric longitude of the Sun (see Suppl. Text S16), by which P_{NSYT} becomes a function of the Julian Day JD_{TT}. The coefficients of eq. SEQ88 are given in Table ST8

**Table ST8 Coefficients of the InSight surface pressure approximation.
All coefficients are in units of Pascals.**

k	a_k	b_k
0	723.601	-
1	37.136	- 35.288
2	- 34.426	36.469

The pressure rate is obtained by a straight forward differentiation, when taking into account that the time dependency is via L_S.

$$\frac{d}{dt} P_{NSYT}(t) = \frac{d}{dt} L_S \sum_{k=1}^{2} [k a_k \sin(L_S) + k b_k \cos(L_S)] \tag{SEQ89}$$

with coefficients again taken from Table ST8, while L_S and its derivative are computed as described in Supp. Texts S16 and S17.

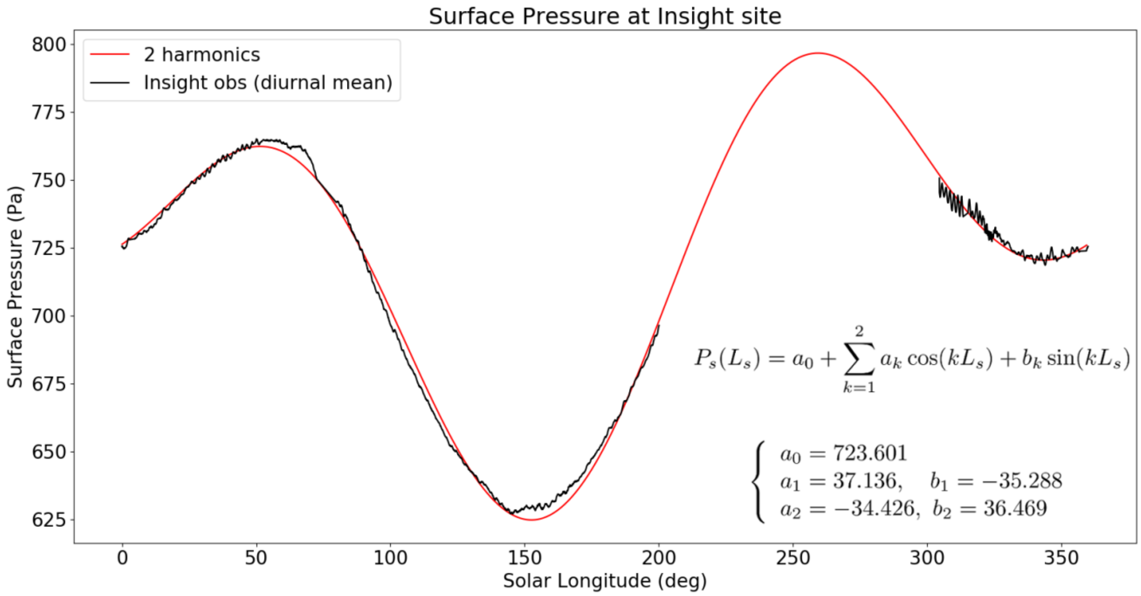


Figure SF11 Surface Pressure at InSight landing site Homestead Hollow.
Black: Diurnal mean of measured InSight surface pressure, red: truncated Fourier series

S19 Strain rate due to solar tides

The tidal potential caused by the Sun at the surface of a planet is (e.g. Agnew, 2007)

$$V = \frac{GM_{sun}}{R(t)} \sum_{j=0}^{\infty} \left(\frac{r}{R(t)}\right)^{j} P_j\big(\cos\alpha\,(t)\big) \tag{SEQ90}$$

with

G	Newton's gravitational constant
M_{sun}	Mass of the Sun
$R(t)$	Distance between the centers of Sun and planet
r	planetary radius
$P_j(x)$	j-th Legendre polynomial
$\alpha(t)$	angular distance between observation point on the planetary surface and the line from the planet's center to the Sun's center

The angle $\alpha(t)$ varies quickly due to the daily rotation of the planet, while $R(t)$ varies slower due the revolution of the planet around the Sun.

The first two terms of eq. SEQ90 can be ignored: the $j = 0$ term is spatially constant; thus, its gradient vanishes and no force results. The $j = 1$ term just corresponds to the gravitational force at the center of the planet and does not contribute to the tides.

We also ignore higher order terms, since the factor $r/R(t)$ is small and its j-th power decreases rapidly. In the case of tides caused on Mars by the Sun, and for the mean distance $\bar{R}$,

$$\frac{r}{\bar{R}} = \frac{3389508}{2.285107\times10^{11}} \approx 1.48 \times 10^{-5} \tag{SEQ91}$$

The radial displacement at the surface of the planet is (e.g. Agnew, 2007)

$$u_r = h_j \frac{V}{g} \tag{SEQ92}$$

where h_j is the j-th Love number and g is the gravitational acceleration at the surface

$$g = GM_{Mars}/r^2 \tag{SEQ93}$$

The strain rate due to the annual solar tide (i.e. omitting the shorter period variation due to the Lagrange Polynomial) is obtained by differentiation with respect to time and reads

$$\dot{\varepsilon}(t) = \dot{u}(t) = -3h_2 \frac{M_{Sun}}{M_{Mars}} \frac{r^3}{R(t)^4} \dot{R}(t) \tag{SEQ94}$$

Since we are interested in the variation with time only, and fit an amplitude factor as free parameter in our inversion anyway, we omit all constant factors in the above expression. Approximative expressions for the heliocentric distance and its first time derivative are given by equations SEQ96 and SEQ97.

S20 Heliocentric distance of Mars and its time derivative

We compute the heliocentric distance of Mars following the Mars24 Sunclock equations (Schmunk, 2020). The distance is given as a function of Mars' Mean Anomaly $M(t)$, which in turn can be approximated by a linear function of time as

$$M(t) = m_1 - m_2(m_3 - t) \tag{SEQ95}$$

where t is the Julian Date in Terrestrial Time (see SEQ85 for the relation between TT and UTC), and coefficients m_i are given in Table ST9. The heliocentric distance as function of the mean anomaly is (from Schmunk, 2020)

$$R(M) = R_0 \sum_{k=0}^{4} a_k \cos\big(kM(t)\big) \tag{SEQ96}$$

and results in Astronomical units. Calculation of the time derivative is straight forward and yields

$$\dot{R}(t) = -R_0 \sum_{k=1}^{4} k a_k m_2 \sin\big(kM(t)\big) \tag{SEQ97}$$

Table ST9 Parameters for computation of heliocentric distance of Mars
k: index, m: Parameters of Mean Anomaly approximation, esp. $m_3 = J2000.0$, a: Parameters of heliocentric distance approximation, in Astronomical units, R: Radius factor in Astronomical units

k	m_k	a_k	R_k
0	-	1.00436	1.52367934
1	19.3871	-0.09309	-
2	0.52402073	-0.004336	-
3	2451545.0	-0.00031	-
4		-0.00003	-

S21 Search grid parameters

We collect here in

- Table ST11 the parameters of the initial and final search grids in the constant-rate verification test described in the main text
- Table ST12 the parameters of exploratory grids that were used to obtain first impressions about feasible parameter ranges, and also for the estimation of parameter uncertainty from the Jackknife-uncertainty of the maximum likelihood value (S11)
- Table ST13 the parameters of the initial search grids definitions, used in the nested grid search
- Table ST14 the parameters of the final search grids after convergence was achieved.

Models use a naming convention according to the kernel function used, and the quadrant in the Amplitude-Offset plane, according to a simple naming scheme: the first four letters indicate the model type (i.e. the kernel function used), as summarized in table ST10. An underscore follows for readability, and the last four letters indicate the quadrant of the Amplitude-Offset plane for which the model was evaluated, “A” for the sign of the amplitude factor, “O” for the sign of the kernel function constant offset, “n” denotes negative, “p” denotes positive values, “z” denotes 0 (zero).

For all models it was made sure that the iteration of the grid was terminated by convergence.

Table ST10 Model type abbreviations

Abbreviation	Model Type
Cnst	constant rate function
Ilmn	Illumination model
Sine	simple sinusoidal rate function
Load	Time derivative of atmospheric pressure
Tide	solar annual tide
A	Amplitude factor
O	constant offset
n	negative values only
p	positive values only
z	fixed to zero
Verif	verification test

Table ST11 Grid parameters for verification test

Parameters of the initial and final grid used for the verification test with a constant rate as described in the main text. m: number of nodes in each dimension, F: volume reduction factor, ε: termination tolerance of likelihood (convergence termination), Iteration: The timeout number of iterations is given in the "initial" row, the actual number until termination by convergence was achieved in the "final" row, A: Amplitude factor range, T: Period range in terrestrial days or degrees (for Sine models), "implicit" for models with inherent period, φ: Retardation range, in terrestrial days or degrees, K: Kernel Offset, in events/day C: Baseline event rate range, in events per terrestrial day.

Model	m	F	ε	Iteration.	A	T	φ	K	C
Initial	31	3	1e-06	50	0 ... 0	1. ... 1.	0 ... 0	0 ... 0	0.05 ... 6.
Final	31	3	1e-06	6	0 0	1. 1.	0 0	0 0	4.5007 4.5088

Table ST12 Exploratory search grid parameters

Model - Name of the model used, m: number of nodes in each dimension, A: Amplitude factor range, T: Period range in terrestrial days, "implicit" for models with inherent period, φ: Retardation range, in terrestrial days or radians, K: Kernel Offset, in events/day C: Baseline event rate range, in events per terrestrial day. The MLE solution for the constant rate model is the one obtained by grid search as well, for comparison. All ranges are given by lower and upper bound.

Model	m	A	T	φ	K	C
Cnst_AzOz	31	0 ... 0	1. ... 1.	0 ... 0	0 ... 0	0.05 ... 3.
Ilmn_AnOn	30	-6. ... 0	(implicit)	0 ... 686.9726	-1.5 ... 0	0. ... 1.5
Ilmn_AnOp	30	-5. ... 0	(implicit)	0 ... 686.9726	0 ... 2.	0. ... 1.5
Ilmn_ApOn	30	0. ... 5.	(implicit)	0 ... 686.9726	-1.5 ... 0	0. ... 1.5
Ilmn_ApOp	30	0. ... 5.	(implicit)	0 ... 686.9726	0 ... 1.5	0. ... 1.5
Load_AnOn	30	-6. ... 0.	(implicit)	0 ... 1030.4589	-2. ... 0	0. ... 1.5
Load_AnOp	30	-3. ... 0.	(implicit)	0 ... 686.9726	0 ... 3.	0. ... 1.5
Load_ApOn	30	0. ... 5.	(implicit)	0 ... 686.9726	-2. ... 0	0. ... 1.5
Load_ApOp	30	0. ... 2.	(implicit)	0 ... 686.9726	0 ... 3.	0. ... 1.5
Sine_AnOn	30	-6. ... 0.	1. ... 1000.	0 ... 6.2832	-6. ... -0.	0. ... 1.5
Sine_AnOp	30	-3. ... -0.	1. ... 800.	0 ... 6.2832	0 ... 3.	0. ... 1.5
Sine_AnOz	30	-6. ... -0.	1. ... 1000.	0 ... 6.2832	0 ... 0	0. ... 1.5
Tide_AnOn	30	-400. ... 0.	(implicit)	0 ... 686.9726	-100. ... 0	0. ... 1.5
Tide_AnOp	30	-400. ... 0.	(implicit)	0 ... 686.9726	0 ... 100.	0. ... 1.5
Tide_ApOn	30	0. ... 400.	(implicit)	0 ... 686.9726	-100. ... 0	0. ... 1.5
Tide_ApOp	30	0. ... 400.	(implicit)	0 ... 686.9726	0 ... 100.	0. ... 1.5

Table ST13 Initial search grid parameters

Model - Name of the model used, m: number of nodes in each dimension, F: volume reduction factor, ε: termination tolerance of likelihood (convergence termination), Max. It: maximum number of grid iterations (timeout termination), A: Amplitude factor range, T: Period range in terrestrial days, "implicit" for models with inherent period, φ: Retardation range, in terrestrial days or radians, K: Kernel Offset, in events/day C: Baseline event rate range, in events per terrestrial day. The MLE solution for the constant rate model is the one obtained by grid search as well, for comparison. All ranges are given by lower and upper bound.

Model	m	F	ε	Max. Iter.	A	T	φ	K	C
Cnst_AzOz	31	3	1e-06	50	0 ... 0	1. ... 1.	0 ... 0	0 ... 0	0.05 ... 3.
Ilmn_AnOn	20	3	1e-06	50	-6. ... 0	(implicit)	0 ... 686.9726	-1.5 ... 0	0. ... 1.5
Ilmn_AnOp	20	3	1e-06	50	-5. ... 0	(implicit)	0 ... 686.9726	0 ... 2.	0. ... 1.5
Ilmn_ApOn	20	3	1e-06	50	0. ... 5.	(implicit)	0 ... 686.9726	-1.5 ... 0	0. ... 1.5
Ilmn_ApOp	20	3	1e-06	50	0. ... 5.	(implicit)	0 ... 686.9726	0 ... 1.5	0. ... 1.5
Load_AnOn	20	3	1e-06	50	-6. ... 0.	(implicit)	0 ... 1030.4589	-2. ... 0	0. ... 1.5
Load_AnOp	20	3	1e-06	50	-3. ... 0.	(implicit)	0 ... 686.9726	0 ... 3.	0. ... 1.5
Load_ApOn	20	3	1e-06	50	0. ... 5.	(implicit)	0 ... 686.9726	-2. ... 0	0. ... 1.5
Load_ApOp	20	3	1e-06	50	0. ... 2.	(implicit)	0 ... 686.9726	0 ... 3.	0. ... 1.5
Sine_AnOn	20	2	1e-06	50	-6. ... 0.	1. ... 1000.	0 ... 6.2832	-6. ... -0.	0. ... 1.5
Sine_AnOp	20	2	1e-06	50	-3. ... -0.	1. ... 800.	0 ... 6.2832	0 ... 3.	0. ... 1.5
Sine_AnOz	20	2	1e-06	50	-6. ... -0.	1. ... 1000.	0 ... 6.2832	0 ... 0	0. ... 1.5
Tide_AnOn	20	3	1e-06	50	-400. ... 0.	(implicit)	0 ... 686.9726	-100. ... 0	0. ... 1.5
Tide_AnOp	20	3	1e-06	50	-400. ... 0.	(implicit)	0 ... 686.9726	0 ... 100.	0. ... 1.5
Tide_ApOn	20	3	1e-06	50	0. ... 400.	(implicit)	0 ... 686.9726	-100. ... 0	0. ... 1.5
Tide_ApOp	20	3	1e-06	50	0. ... 400.	(implicit)	0 ... 686.9726	0 ... 100.	0. ... 1.5

Table ST14 Final search grid parameters

Columns as in Table ST13, except "Last. It.", which is the final iteration before termination. All iterations terminated by reaching the termination tolerance, i.e. convergence rather than timeout. All ranges are given by lower and upper bound. “implicit” period for models with inherent periodicity.

Model	m	F	ϵ	Last It.	A	T	φ	K	C
Cnst_AzOz	31	3	1e-06	6	0 0	1. 1.	0 0	0 0	2.4517 2.4557
Ilmn_AnOn	20	3	1e-06	23	-4.3123 -4.2695	(implicit)	390.7745 395.6711	-0.4289 -0.404	0.5161 0.5268
Ilmn_AnOp	20	3	1e-06	19	-3.8463 -3.8193	(implicit)	396.0005 399.7211	0 0.0108	0.4283 0.4364
Ilmn_ApOn	20	3	1e-06	25	6.6244 6.6565	(implicit)	42.3832 46.1038	-2.909 -2.8901	0.5019 0.51
Ilmn_ApOp	20	3	1e-06	17	3.5639 3.6108	(implicit)	50.8529 57.2971	0 0.0141	0 0.0141
Load_AnOn	20	3	1e-06	22	-3.8256 -3.8113	(implicit)	700.4709 702.9192	-0.0048 0	0.963 0.9665
Load_AnOp	20	3	1e-06	19	-2.8626 -2.8463	(implicit)	16.5759 20.2965	1.0904 1.1066	0.5448 0.5529
Load_ApOn	20	3	1e-06	17	2.2129 2.2599	(implicit)	491.0744 497.5186	-0.0188 0	1.3459 1.36
Load_ApOp	20	3	1e-06	17	1.6412 1.66	(implicit)	504.001 510.4453	1.4609 1.489	0.6352 0.6493
Sine_AnOn	20	2	1e-06	33	-3.908 -3.7929	557.5876 575.518	3.3948 3.5076	-0.123 -0.	0.5028 0.5297
Sine_AnOp	20	2	1e-06	30	-3.2685 -3.1753	491.4962 513.2328	3.6059 3.7768	0.6015 0.6831	0.5051 0.546
Sine_AnOz	20	2	1e-06	30	-3.8829 -3.7891	557.4838 573.0931	3.407 3.5051	0 0	0.5056 0.529
Tide_AnOn	20	3	1e-06	28	-331.4253 -329.2096	(implicit)	569.4283 571.0605	-1.2961 -1.0585	0.5637 0.5736
Tide_AnOp	20	3	1e-06	20	-254.8375 -253.1914	(implicit)	577.5345 580.3615	0 0.4115	0.2117 0.2178
Tide_ApOn	20	3	1e-06	29	307.2438 308.5231	(implicit)	282.1796 283.1219	-0.8704 -0.7333	0.5178 0.5226
Tide_ApOp	20	3	1e-06	19	246.9772 249.1436	(implicit)	284.1184 287.839	0 0.5416	0.3616 0.3698

S22 MLE solutions and model selection parameters

Figure SF12 shows 1D cross-sections through the log-likelihood function for model Ilmn_AnOn, Figure SF13 shows the respective cross-sections for model Tide_ApOn, and Figure SF14 those for model Load_AnOp.

Table ST15 contains the resulting Maximum Likelihood Estimations of all used search grids, and the resulting parameter values. Variances given were computed as described in section S12. In several cases, negative variances (which are meaningless) were obtained, or implausibly large variances, or even Not-a-Number (NaN) values. These solutions are reported as computed (sometimes rounded) in Table ST15.

Table ST16 gives the parameter uncertainties estimated from comparing likelihood values across the grid (S12), also as described in section S12.

Table ST17 gives the log-likelihood values and corrected Akaike Information Criterion values for the parameters listed in Table ST15, as well as Akaike weights and evidence ratios as described in section S14.

The model naming convention is described in section S21.

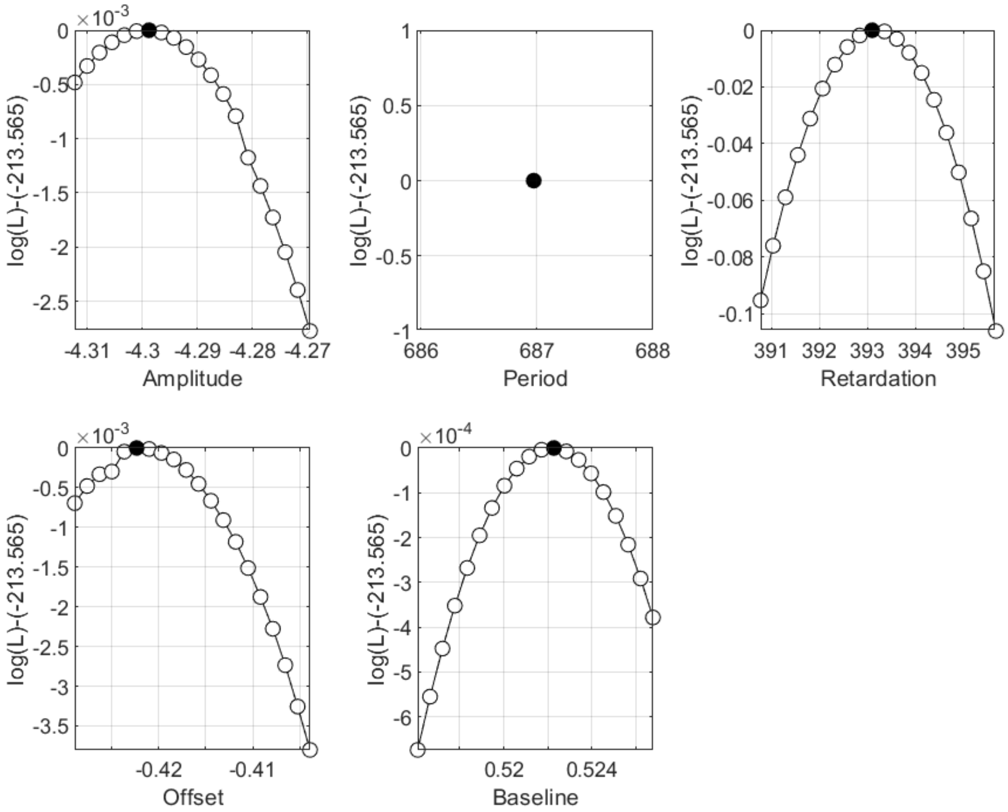


Figure SF12 log-Likelihood profiles for model Ilmn_AnOn, from the final search grid.
Each panel shows the variation of the likelihood in one direction, with other parameters kept constant at the value of the MLE solution. Period is not variable in this model. Open circles mark search grid nodes. Filled circles mark the location of the likelihood maximum. Log-likelihoods are reduced by the maximum value as indicated at the vertical axes.

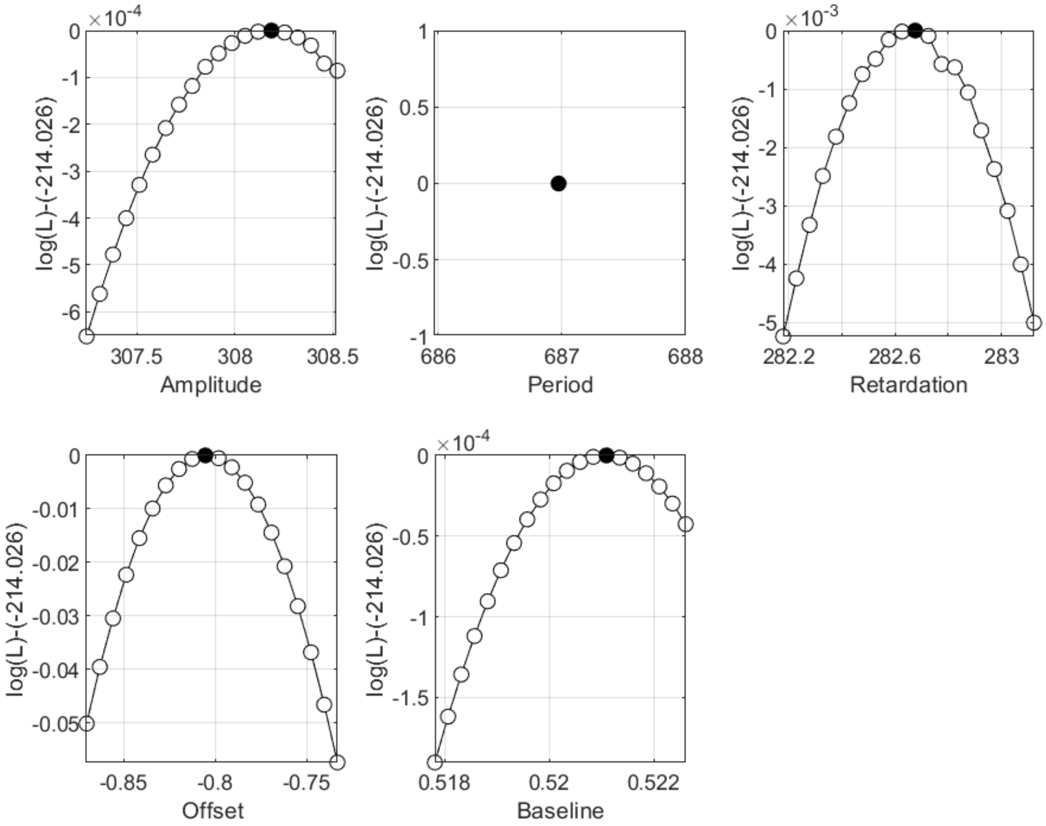
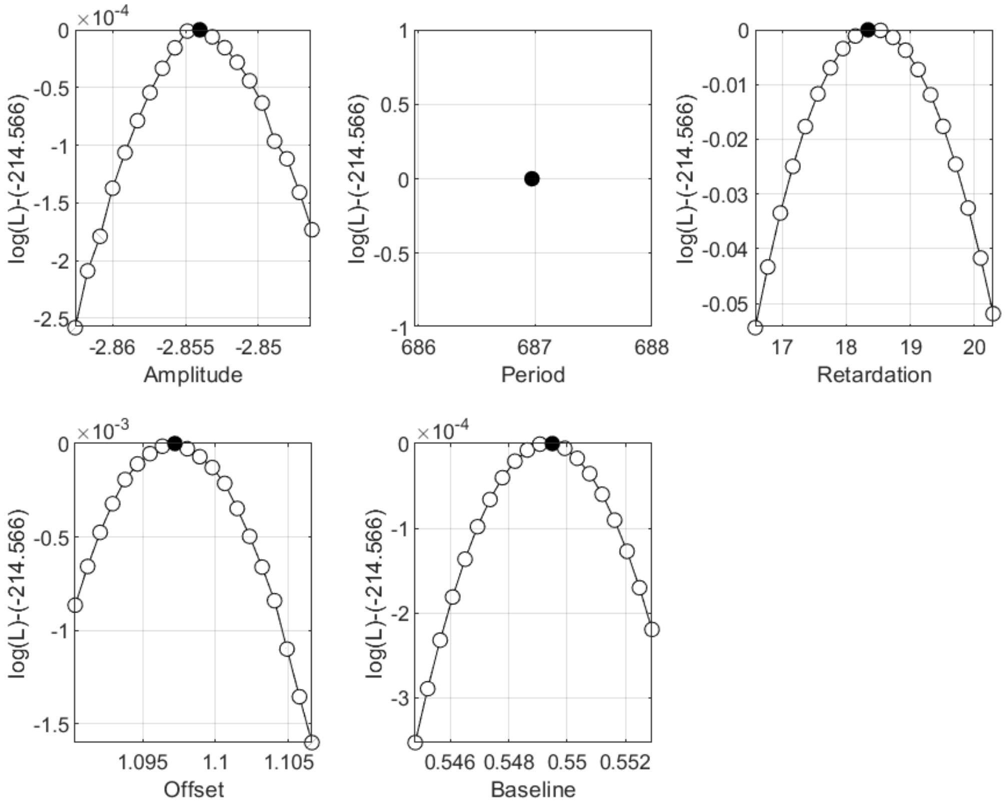


Figure SF13 log-Likelihood profiles for model Tide_ApOn, from the final search grid. See caption of Figure SF12 for description.



1195 **Figure SF14 log-Likelihood profiles for model Load_AnOp, from the final search grid. See caption of Figure SF12 for description.**

Table ST15 Maximum Likelihood Parameter and uncertainties from Fisher Information Matrix
The columns of the table are: Model: Name of the model used, A: Amplitude factor, T: period if applicable (Insolation, Pressure rate and tide models have an implicit period of 1 martian year), φ: lag time at the maximum, in degrees for Sine models, in days for all others, K: DC offset in kernel function, C: baseline event rate. The given variances σ^2 result from Fisher information matrix inversion as described in supplemental text S12. These are given as computed, although negative variances are meaningless, and some values are implausibly large ("Inf" denotes an infinite solution, "-" is given where the parameter value is not variable).

Model	A	σ^2	T	σ^2	φ	σ^2	K	σ^2	C	σ^2
Cnst_Verif	0	-	1	-	0	-	0	-	4.5047	Inf
Cnst_AzOz	0	-	1.	-	0	-	0	-	2.4537	Inf
Ilmn_AnOn	-4.2988	1.067	(implicit)	-	393.0939	-504.7688	-0.4223	0.0359	0.5223	0.1943
Ilmn_AnOp	-3.8321	0.9704	(implicit)	-	397.7629	-0.3681	0	0.	0.4322	0.1421
Ilmn_ApOn	6.6413	6.1803	(implicit)	-	44.1456	161.8062	-2.9001	3.4437	0.5058	0.1965
Ilmn_ApOp	3.5861	2.0098	(implicit)	-	53.9054	453.3183	0	1332082	0.	1332083
Load_AnOn	-3.8181	1.3926	(implicit)	-	701.6306	-2284.5072	0	-0.0321	0.9646	0.175
Load_AnOp	-2.854	0.8951	(implicit)	-	18.3383	177.0075	1.0972	0.4813	0.5495	0.2903
Load_ApOn	2.2376	1.509	(implicit)	-	494.4661	1272562	0	1.3898	1.3526	0.2343
Load_ApOp	1.6491	0.2084	(implicit)	-	507.3928	287.2765	1.4772	0.2816	0.6404	0.2825
Sine_AnOn	-3.8353	Inf	565.1372	Inf	3.4541	Inf	-0.	Inf	0.5169	Inf
Sine_AnOp	-3.2489	Inf	505.2246	Inf	3.6778	Inf	0.6144	Inf	0.5245	Inf
Sine_AnOz	-3.8336	Inf	565.6992	Inf	3.4535	Inf	0	-	0.5155	Inf
Tide_AnOn	-329.5594	7821.091	(implicit)	-	570.2874	-14.9275	-1.1711	-0.0887	0.5689	0.1591
Tide_AnOp	-253.9711	7925.0682	(implicit)	-	578.8736	176.3445	0	0.7135	0.2152	0.178
Tide_ApOn	308.1864	19902.1037	(implicit)	-	282.6755	151.8021	-0.8055	2.2813	0.5211	0.215
Tide_ApOp	248.0034	11389.6851	(implicit)	-	285.8808	254.881	0	1.2311	0.3655	0.171

Table ST16 Parameter uncertainty from grid node comparison
This table gives the maximum likelihood solutions for all model parameters, and the acceptable intervals according to a comparison of the likelihoods of individual nodes in the initial search grid, as described in section S12. σ_{MLE} is the standard deviation of the maximum likelihood as estimated by the Jackknife test (section S11). Lower and upper bounds of acceptable ranges are given since this method does not necessarily produce intervals symmetric to the maximum likelihood solution.

Model	σ_{MLE}	Amplitude	low	high	Period	low	high	Retardation	low	high	Offset	low	high	Baseline	low	high
Cnst_Verif	0.278	0	0	0	1	1	1	0	0	0	0	0	0	4.5047	4.0167	5.0083
Cnst_AzOz	1.879	0	0	0	1.	1.	1.	0	0	0	0	0	0	2.4537	1.9183	3.
Ilmn_AnOn	6.3622	-4.2988	-6.	-1.2414	686.9726	686.9726	686.9726	393.0939	355.3307	426.3968	-0.4223	-1.5	0	0.5223	0.0517	1.5
Ilmn_AnOp	6.0506	-3.8321	-5.	-1.2069	686.9726	686.9726	686.9726	397.7629	355.3307	450.0855	0	0	2.	0.4322	0.	1.5
Ilmn_ApOn	6.2298	6.6413	1.7241	5.	686.9726	686.9726	686.9726	44.1456	23.6887	94.7548	-2.9001	-1.5	0	0.5058	0.	1.5
Ilmn_ApOp	3.968	3.5861	1.7241	4.8276	686.9726	686.9726	686.9726	53.9054	23.6887	94.7548	0	0	1.2931	0.	0.	1.1897
Load_AnOn	6.3091	-3.8181	-6.	-1.2414	686.9726	686.9726	686.9726	701.6306	0	1030.4589	0	-2.	0	0.9646	0.3103	1.5
Load_AnOp	6.3207	-2.854	-3.	-0.7241	686.9726	686.9726	686.9726	18.3383	0	686.9726	1.0972	0	3.	0.5495	0.	1.5
Load_ApOn	6.1108	2.2376	0.6897	5.	686.9726	686.9726	686.9726	494.4661	165.821	544.8403	0	-2.	0	1.3526	0.4655	1.5

Model	Tide_ApOp	Tide_ApOn	Tide_AnOp	Tide_AnOn	Sine_AnOz	Sine_AnOp	Sine_AnOn	Load_ApOp
σ_{MLE}	5.6914	6.2916	5.688	6.1999	10.0225	10.1827	9.3598	6.1988
Amplitude	248.0034	308.1864	-253.9711	-329.5594	-3.8336	-3.2489	-3.8353	1.6491
low	96.5517	82.7586	-386.2069	-400.	-6.	-3.	-6.	0.4138
high	400.	400.	-82.7586	-82.7586	-0.8276	-0.2069	-0.8276	2.
Period	686.9726	686.9726	686.9726	686.9726	565.6992	505.2246	565.1372	686.9726
low	686.9726	686.9726	686.9726	686.9726	207.6897	221.4138	242.1379	686.9726
high	686.9726	686.9726	686.9726	686.9726	1000.	800.	1000.	686.9726
Retardation	285.8808	282.6755	578.8736	570.2874	3.4535	3.6778	3.4541	507.3928
low	260.5758	260.5758	544.8403	521.1516	0	0	0	165.821
high	331.6419	331.6419	615.9065	615.9065	6.2832	6.2832	6.2832	544.8403
Offset	0	-0.8055	0	-1.1711	0	0.6144	-0.	1.4772
low	0	-3.4483	0	-3.4483	0	0	-4.9655	0
high	0	0	0	0	0	3.	-0.	3.
Baseline	0.3655	0.5211	0.2152	0.5689	0.5155	0.5245	0.5169	0.6404
low	0.	0.	0.0517	0.0517	0.	0.	0.	0.
high	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5

Table ST17 Log-Likelihood, Akaike Information Criterion, Akaike Weights and Evidence Ratios
The columns of the table are: Model: Name of the model used, Iterations: number of grid iterations conducted, $log\ L$: log-likelihood at the MLE solution (parameters in Table ST15), σ_{logL} : Jackknife variance of $log\ L$, n: number of events used, k: number of free parameters, AIC_c: resulting corrected Akaike Information Criterion, AIC_c Rank: ranking according to value of AIC_c, Akaike Weight as described in S14, Evidence ratio as described in S14. The verification test Cnst_Verif run was not ranked against the others, as it uses data from a time interval where the occurrence rate appears to be constant indeed, see main text.

Model	Iterations	log(L)	$\sigma_{\log L}$	n	k	AICc	AICc Rank	Akaike Weight	Evidence Ratio
Cnst_Verif	6	-92.977	0.278	67	1	188.015	-	-	-
Cnst_AzOz	6	-237.426	1.879	118	1	476.886	16	1.670e-10	9.777e+08
Ilmn_AnOn	23	-213.565	6.362	118	4	435.484	1	1.633e-01	1.000e+00
Ilmn_AnOp	19	-213.725	6.051	118	4	435.805	3	1.391e-01	1.174e+00
Ilmn_ApOn	25	-213.727	6.230	118	4	435.809	4	1.388e-01	1.176e+00
Ilmn_ApOp	17	-216.729	3.968	118	4	441.812	14	6.900e-03	2.366e+01
Load_AnOn	22	-216.523	6.309	118	4	441.400	12	8.480e-03	1.925e+01
Load_AnOp	19	-214.566	6.321	118	4	437.486	7	6.002e-02	2.720e+00
Load_ApOn	17	-221.050	6.111	118	4	450.454	15	9.166e-05	1.781e+03
Load_ApOp	17	-216.530	6.199	118	4	441.413	13	8.424e-03	1.938e+01
Sine_AnOn	33	-213.610	9.360	118	5	437.755	9	5.245e-02	3.113e+00
Sine_AnOp	30	-213.605	10.183	118	5	437.746	8	5.270e-02	3.098e+00
Sine_AnOz	30	-213.610	10.022	118	5	435.574	2	1.561e-01	1.046e+00
Tide_AnOn	28	-215.070	6.200	118	4	438.494	10	3.625e-02	4.503e+00
Tide_AnOp	20	-216.297	5.688	118	4	440.947	11	1.063e-02	1.535e+01
Tide_ApOn	29	-214.026	6.292	118	4	436.405	5	1.030e-01	1.585e+00
Tide_ApOp	19	-214.506	5.691	118	4	437.366	6	6.372e-02	2.562e+00

S23 Rate forecasts for selected models

Table ST18 gives the times of forecast upcoming HF event activity phases, according to models with evidence ratio below 100 (excluding Sine models), as discussed in the main text.

Figure SF15, in analogy to fig. 6b of the main paper, compares model rate forecasts and observed rate, but with a different use of the detection efficiency factor.

For models Ilmn_AnOn (Figure SF16), Load_ApOp (Figure SF17),Tide_ApOn (Figure SF18), and Load_AnOp (Figure SF19) we compute model rate predictions for a subset of models as described in the main text. The parameters of the models used are taken from the exploratory grids, since the final grids of the grid iteration procedure cover volumes much smaller than the uncertainty range obtained from the standard deviation of the log-likelihood (S11).

Table ST18 Expected timing of coming activity periods.
Forecasts for the best models (except Sine models), and modeled times for the observation interval: Start, Maximum, Minimum, End: predicted dates for beginning, maximum, minimum, and end of next activity. Dates are given as InSight mission sol and UTC date. Minimum times are not given for extended phases of baseline activity, these are assumed to occur between an End time and the next start time. Times do not account for the uncertainty for model parameters, but are valid for the maximum likelihood solution.

Model	Rank	Start	Max	Min	End
Tide_ApOn	5	217, 07. Jul. 2019	319, 20. Oct. 2019		506, 29. Apr. 2020
Tide_ApOn	5	886, 24. May. 2021	988, 06. Sep. 2021		1175, 17. Mar. 2022
Tide_AnOn	10	198, 17. Jun. 2019	372, 14. Dec. 2019		470, 23. Mar. 2020
Tide_AnOn	10	866, 04. May. 2021	1041, 30. Oct. 2021		1139, 09. Feb. 2022
Tide_ApOp	6	209, 29. Jun. 2019	322, 23. Oct. 2019		543, 06. Jun. 2020
Tide_ApOp	6	877, 16. May. 2021	991, 09. Sep. 2021		
Tide_AnOp	11	160, 09. May. 2019	381, 22. Dec. 2019		494, 17. Apr. 2020
Tide_AnOp	11	828, 26. Mar. 2021	1049, 08. Nov. 2021		1163, 05. Mar. 2022
Load_AnOn	12	235, 26. Jul. 2019	359, 30. Nov. 2019		451, 03. Mar. 2020
Load_AnOn	12	626, 30. Aug. 2020	685, 30. Oct. 2020		765, 20. Jan. 2021
Load_AnOn	12	904, 12. Jun. 2021	1028, 17. Oct. 2021		1119, 19. Jan. 2022
Load_ApOp	13	198, 17. Jun. 2019	641, 15. Sep. 2020	496, 19. Apr. 2020	809, 06. Mar. 2021
Load_ApOp	13	866, 04. May. 2021	1024, 13. Oct. 2021	1165, 07. Mar. 2022	
Load_AnOp	7	191, 10. Jun. 2019	363, 04. Dec. 2019		472, 25. Mar. 2020
Load_AnOp	7	616, 20. Aug. 2020	689, 03. Nov. 2020		809, 06. Mar. 2021
Load_AnOp	7	860, 27. Apr. 2021	1031, 21. Oct. 2021		1140, 10. Feb. 2022
Ilmn_ApOn	4	209, 29. Jun. 2019	350, 21. Nov. 2019		482, 04. Apr. 2020
Ilmn_ApOn	4	878, 16. May. 2021	1018, 08. Oct. 2021		1151, 20. Feb. 2022
Ilmn_AnOn	1	209, 29. Jun. 2019	342, 13. Nov. 2019		485, 08. Apr. 2020
Ilmn_AnOn	1	878, 16. May. 2021	1011, 30. Sep. 2021		1154, 24. Feb. 2022
Ilmn_ApOp	14	166, 16. May. 2019	359, 30. Nov. 2019		538, 01. Jun. 2020
Ilmn_ApOp	14	835, 02. Apr. 2021	1028, 17. Oct. 2021		
Ilmn_AnOp	3	204, 24. Jun. 2019	347, 17. Nov. 2019		501, 24. Apr. 2020
Ilmn_AnOp	3	873, 11. May. 2021	1015, 04. Oct. 2021		1170, 12. Mar. 2022

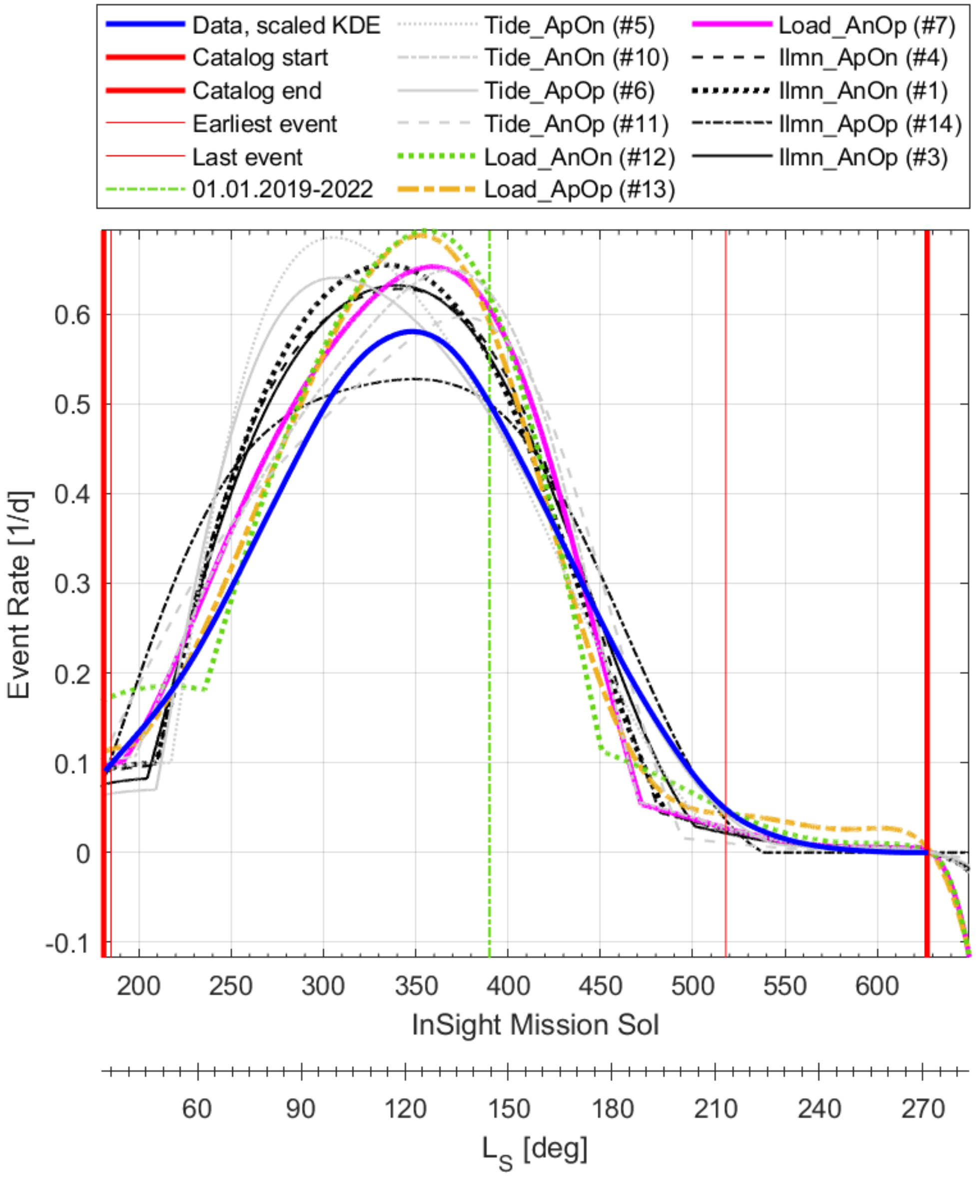


Figure SF15 Rate forecasts

This figure compares model rate forecasts, modulated by detection efficiency, with the kernel density estimation (KDE) of the rate as obtained from the actual event sequence. Note that model Load_ApOp (orange) is the only one with extended activity above the baseline rate after Sol 520. Other Load models are drawn in colors as indicated by the legend, to allow for easier comparison with Load_ApOp. This figure is analogous to fig. 6b in the main paper, which shows unmodulated model forecasts in comparison to the KDE rate divided by the detection efficiency.

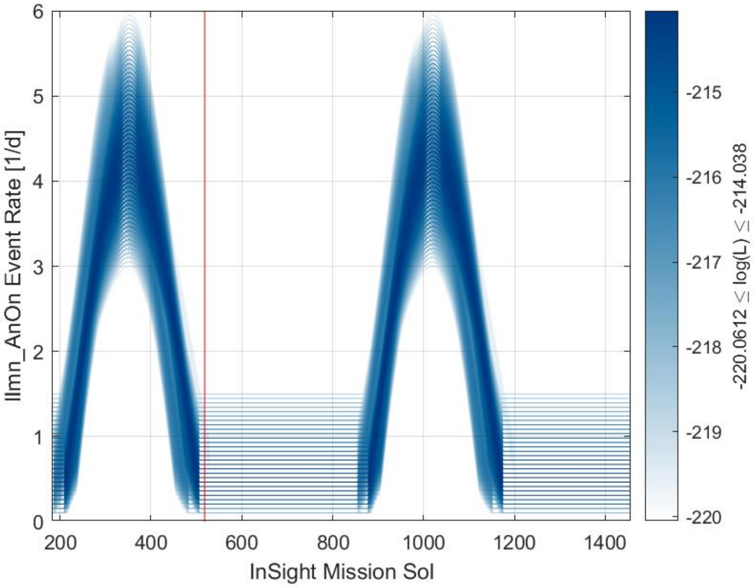


Figure SF16 Rate forecasts for model Ilmn_AnOn.
Vertical red line marks the time of the last event used in this study. Color indicates log-likelihood according to color bar; models selected from the exploratory grid according to their likelihood as explained in the main text. 1255

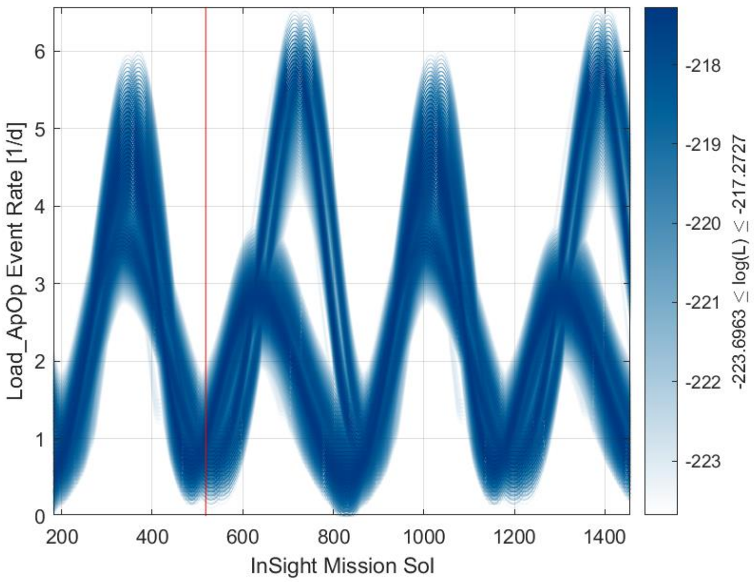
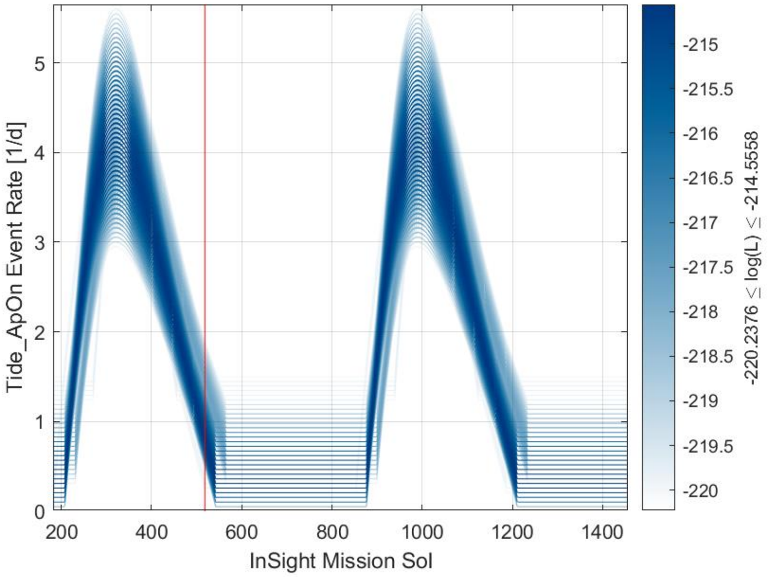


Figure SF17 Rate forecasts for model Load_ApOp.
See caption of Figure SF16.



1260

**Figure SF18 Rate forecasts for model Tide_ApOn
See caption of Figure SF16.**

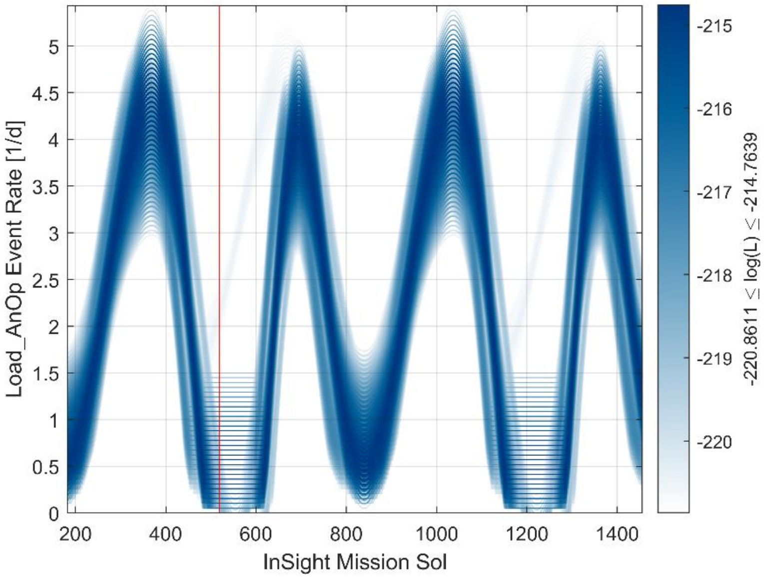


Figure SF19 Rate forecasts for model Load_AnOp
1265 **See caption of Figure SF16.**

References

Abilleira, F., Halsell, A., Kruizinga, G., Bonfiglio, E., Grover, R., Chung, M.-K., Fujii, K., Gustafson, E., Hahn, Y., Jefferson, D., Lau, E., Lee, J., McCandless, S.-E., Mottinger, N., Seubert, J., Sklyanskiy, E., Wallace, M. (2019), 2018 Mars Insight trajectory reconstruction and performance from launch through landing, *AAS/AIAA Astrodynamics Specialist Conference*, Maui, Hawaii, January 13-17, 2019, online: https://trs.jpl.nasa.gov/handle/2014/45976, last accessed 08. Oct. 2020

Agnew, D. C. (2007), Earth Tides, in: Herring, T. (Ed.), *Treatise on Geophysics*, vol. 3: Geodesy, Elsevier

Aharonson, O., Zuber, M. T., Smith, D. E., Neumann, G. A., Feldman, W. C., Prettyman, T. H. (2004), Depth, distribution, and density of CO_2 deposition on Mars, *Journal of Geophysical Research*, vol. 109, E05004, doi:10.1029/2003JE002223

Akaike, H. (1974), A new look at statistical model identification, *IEEE Transactions on Automation Control*, vol. AC-19, No. 6, 716-723, DOI: 10.1109/TAC.1974.1100705

Aki, K. (1965) Maximum likelihood estimate of b in the formula log N=a-bM and its confidence limits, *Bulletin of the Earthquake Research Institute*, vol. 43, 237-239

Allison, M. (1997), Accurate analytic representations of solar time and seasons on Mars with applications to the Pathfinder/Surveyor missions. *Geophysical Research Letters*, 24, 1967-1970.

Allison, M., McEwen, M. (2000), A post-Pathfinder evaluation of areocentric solar coordinates with improved timing recipes for Mars seasonal/diurnal climate studies. *Planetary and Space Science*, 48, 215-235.

Banerdt, W. B.,Smrekar, S. E., Banfield, D., Giardini, D., Golombek, M., Johnson, C. L., Lognonné, P., Spiga, A., Spohn, T., Perrin, C., Stähler, S. C., Antonangeli, D., Asmar, S., Beghein, C., Bowles, N., Bozdag, N., Chi, P., Christensen, U., Clinton, J., Collins, G. S., Daubar, I., Dehant, V., Drilleau, M., Fillingim, M., Folkner, W., Garcia, R. F., Garvin, J., Grant, J., Grott, M., Grygorczuk, J., Hudson, T., Irving, J.C. E., Kargl, G., Kawamura, T., Kedar, S., King, S, Knapmeyer-Endrun, B., Knapmeyer, M., Lemmon, M., Lorenz, R., Maki, J. N., Margerin, L., McLennan, S. M., Michaut, C., Mimoun, D., Mittelholz, A., Mocquet, A., Morgan, P., Mueller, N. T., Murdoch, N., Nagihara, S., Newman, C., Nimmo, F., Panning, M., Pike, W. T., Plesa, A.-C., Rodriguez, S., Rodriguez-Manfredi, J. A., Russell, C. T., Schmerr, N., Siegler, M., Stanley, S., Stutzmann, E., Teanby, N., Tromp, J., van Driel, M., Warner, N., Weber, R., Wieczorek, M., (2020), Initial results from the InSight mission on Mars, *Nature Geoscience*, vol. 13, 183-189, DOI : 10.1038/s41561-020-0544-y

Banfield, D., Rodriguez-Manfredi, J. A., Russell, C. T., Rowe, K. M., Leneman, D., Lai, H. R., Cruce, P. R., Means, J. D., Johnson, C. L., Mittelholz, A., Joy, S. P., Chi, P. J., Mikellides, I. G., Carpenter, S., Navarro, S., Sebastian, E., Gomez-Elvira, J., Torres, J., Mora, L. Peinado, V., Lepinette, A., (2018), InSight Auxiliary Payload Sensor Suite (APSS), Space Science Reviews, vol. 215, No. 4, https://doi.org/10.1007/s11214-018-0570-x

Banfield, D., Spiga, A., Newman, C., Forget, F., Lemmon, M., Lorenz, R., Murdoch, N., Viudez-Moreiras, D., Pla-Garcia, J., Garcia, R. F., Lognonné, P. Karatekin, Ö., Perrin, C, Martire, L., Teanby, N, van Hove, B., Maki, J, Kenda, B., Mueller, N. T., Rodriguez, S., Kawamura, T., McClean, J. B., Stott, A. E., Charalambous, C., Millour, E., Johnson, C. .L., Mittelholz, A., Määttänen, A., Lewis, S. R., Clinton, J, Stähler S. C., Ceylan, S., Giardini, D., Warren, T., Pike, T., Daubar, I. Golombek, M., Rolland, L., Widmer-Schnidrig, R, Mimoun, D. Beucler, É., Jacob., A., Lucas, A., Baker, M., Ansan, V., Hurst, K., Mora-Sotomayor, L., Navarro, S., Torres, J., Lepinette, A., Molina, A., Marin-Jimenez, M., Gomez-Elvira, J., Peinado, V., Rodriguez-Manfredi, J.-A., Carcich, B. T., Sackett, S., Russell, C. T., Spohn, T., Smrekar, S. E., Banerdt, W. B. (2020), The atmosphere of Mars as observed by InSight, *Nature Geoscience*, vol. 13, 190-198, https://doi.org/10.1038/s41561-020-0534-0

Bills, B. G., Comstock, R. L. (2005), Spatial and temporal patterns of solar eclipses by Phobos on Mars, *Journal of Geophysical Research*, vol. 110, E04004, doi:10.1029/2003JE002209

Botev, Z. I., Grotowski, J. F., Kroese, D. P. (2010), Kernel Density Estimation via Diffusion, *The Annals of Statistics*, vol. 38, No. 5, 2916-2957, DOI: 10.1214/10-AOS799

Botev, Z. (2020). Kernel Density Estimator (https://www.mathworks.com/matlabcentral/fileexchange/14034-kernel-density-estimator), *MATLAB Central File Exchange*. Retrieved November 11, 2020.

Bulow, R. C., Johnson, C. L., Bills, B. G., Shearer; P. M. (2007), Temporal and spatial properties of some deep moonquake clusters, *Journal of Geophysical Research*, vol. 112, E09003, doi:10.1029/2006JE002847

Burnham, K. P., Anderson, D. R., Huyvaert, K. P. (2011), AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons, *Behavioral Ecology and Sociobiology*, vol. 65, 23-35, DOI 10.1007/s00265-010-1029-6

Calvin, W. M., Cantor, B. A., James, P. B. (2017) Interannual and seasonal changes in the south seasonal polar cap of Mars: Observations from MY 28-31 using MARCI, Icarus, vol. 292, 144-153, http://dx.doi.org/10.1016/j.icarus.2017.01.010

Clinton, J., Giardini, D., Böse, M., Ceylan, S., van Driel, M., Euchner, F., Garcia, R.F., Kedar, S., Khan, A., Stähler, S.C., Banerdt, B., Lognonne, P., Beucler, E., Daubar, I., Drilleau, M., Golombek, M., Kawamura, T., Knapmeyer, M., Knapmeyer-Endrun, B., Mimoun, D., Mocquet, A., Panning, M., Perrin, C., Teanby, N.A. (2018), The Marsquake Service: Securing Daily Analysis of SEIS Data and Building the Martian Seismicity Catalogue for InSight. *Space Science Reviews*, 214(133), https://doi.org/10.1007/s11214-018-0567-5

Clinton, J. F., Ceylan, S., van Driel, M., Giardini, D., Stähler, S. C., Böse, M., Charalambous, C., Dahmen, N. L., Horleston, A., Kawamura, T., Khan, A., Orhand-Mainsant, G., Scholz, J.-R., Euchner, F., Banerdt, W. B., Lognonné, P., Banfield, D., Beucler, E., Garcia, R. F., Kedar, S., Panning, M. P., Perrin, C., Pike, W. T., Smrekar, S. E., Spiga, A., Stott, A. (2021), The Marsquake catalogue from InSight, sols 0–478, *Physics of the Earth and Planetary Interiors*, vol. 310, https://doi.org/10.1016/j.pepi.2020.106595

d'Errico, J. (2014), Adaptive Robust Numerical Differentiation (https://www.mathworks.com/matlabcentral/fileexchange/13490-adaptive-robust-numerical-differentiation), *MATLAB Central File Exchange*. Retrieved May 19, 2020.

Dieterich, J. (1994) A constitutive law for rate of earthquake production and its application to earthquake clustering, *Journal of Geophysical Research*, vol. 99, B2, 2601-2618

Diggle, P., Marron, J. S. (1988), Equivalence of Smoothing Parameter Selectors in Density and Intensity Estimation, Journal of the American Statistical Association, vol. 83, No. 403, 793-800, https://doi.org/10.1080/01621459.1988.10478665

van Driel, M., Ceylan, S., Clinton, J. F., Giardini, D., Horleston, A., Margerin, L., Stähler, S. C., Böse, M., Charalambous, C., Kawamura, T., Khan, A., Orhand-Mainsant, G., Scholz, J.-R., Euchner, F., Knapmeyer, M., Schmerr, N., Pike, W. T., Lognonné, P., Banerdt, W. B. (2021), High frequency seismic events on Mars observed by InSight, *Journal of Geophysical Research: Planets*, vol. 126, No. 2, http://dx.doi.org/10.1029/2020JE006670

Efron, B., C. Stein (1981), The Jackknife Estimate of Variance, *The Annals of Statistics*, vol. 9, No. 3, 586-596

Forget, F., (1998) Mars CO2 Ice Polar Caps. In: Schmitt B., De Bergh C., Festou M. (eds) Solar System Ices. *Astrophysics and Space Science Library*, vol 227. Springer, Dordrecht

Forget, F., Hourdin, F., Fournier, R., Hourdin, C., Talagrand, O., Collins, M., Lewis, S. R., Read, P. L., Huot, J.-P. (1999), Improved general circulation models of the Martian atmosphere from the surface to above 80 km, *Journal of Geophysical Research*, vol. 104, E10, 24155-24176, DOI: 10.1029/1999JE001025

Giardini, D., Lognonné, P., Banerdt, W., Pike, W., Christensen, U., Ceylan, S., Clinton, J., van Driel, M., Stähler, S., Böse, M., Garcia, R. F., Khan, A., Panning, M., Perrin, C., Bandfield, D., Beucler, E., Charalambous, C., Euchner, F., Horleston, A., Jacob, A., Kawamura, T., Kedar, S., Mainsant, G., Scholz, J.-R., Smrekar, S., Spiga, A., Agard, C., Antonangeli, D., Barkaoui, S., Barrett, E., Combers, P., Conejero, V., Daubar, I., Drilleau, M., Ferrier, C., Gabsi, T., Gudkova, T., Hurst, K., Karakostas, F., King, S., Knapmeyer, M., Knapmeyer-Endrun, B., Llorca-Cejudo, R., Lucas, A., Luno, L., Margerin, L., McClean, J., Mimoun, D., Murdoch, N., Nimmo, F., Nonon, M., Pardo, C., Rivoldini, A., Rodriguez Manfredi, J. A., Samuel, H., Schimmel, M., Stott, A. E., Stutzmann, E., Teanby, N., Warren, T., Weber, R., Wieczorek, M., Yana, C. (2020), The Seismicity of Mars. *Nature Geoscience*, vol. 13, 205-212,, DOI : 10.1038/s41561-020-0539-8

Hess, S. L., Henry, R. M., Tillman, J. E. (1979), The seasonal variation of atmospheric pressure on Mars as affected by the south polar cap, *Journal of Geophysical Research*, vol. 84, No. B6, 2923-2927

Hourdin, F., Forget, F., Talagrand, O., (1995), The sensitivity of the Martian surface pressure and atmospheric mass budget to various parameters: A comparison between numerical simulations and Viking observations, *Journal of Geophysical Research*, vol. 100, E3, 5501-5523

InSight Analyst's Notebook, online: https://an.rsl.wustl.edu/ins/an/an3.aspx, last accessed 08. Oct. 2020

International Earth Rotation and Reference Systems Service (2020), Bulletin C 59, online https://datacenter.iers.org/data/latestVersion/16_BULLETIN_C16.txt, last accessed 06. April 2020

Jian, J.-J., Ip, W.-H. (2009), Seasonal patterns of condensation and sublimation cylces in the cryptic and non-cryptic regions of the South Pole, *Advances in Space Research*, vol. 43, 138-142, doi:10.1016/j.asr.2008.05.002

Kelly, N. J., Boynton, W. V., Kerryl, K., Hamara, D., Janes, D., Reedy, R. C., Kim, K. J., Haberle, R. M. (2006), Seasonal polar carbon dioxide frost on Mars: CO2 mass and columnar thickness distribution, *Journal of Geophysical Research*, vol. 111, doi:10.1029/2006JE002678

Kullback, S., Leibler, R.A. (1951), On information and sufficiency, *Annals of Mathematical Statistics*, vol. 200, 79-86

Liddle, A. R. (2007), Information criteria for astrophysical model selection, *Monthly Notes of the Royal Astronomical Society*, vol. 377, L74-L78

Lognonné P., Banerdt W.B., Giardini D., Pike W.T., Christensen U. et al. (2019). SEIS: Insight's Seismic Experiment for Internal Structure of Mars. *Space Science Reviews*, 215(12). https://doi.org/10.1007/s11214-018-0574-6

Meeus, J. (1994), Astronomische Algorithmen, J. Ambrosius Barth, Leipzig, Berlin, Heidelberg, 2nd ed., 460 pages

Millour, E. Forget, F., Spiga, A., Vals, M., Zakharov, V., Montabone, L., Lefèvre, F., Montmessin, F., Chaufray, J.-Y., López-Valverde, M. A., González-Galindo, F., Lewis, S. R., Desjean, M.-C., Cipriani, F., and the MCD development team, Scientific Workshop "From Mars Express to ExoMars", 27.-28. February 2018, ESAC Madrid, Spain

Ogata, Y. (1983), Estimation of the parameters in the modified Omori formula for aftershock frequencies by the Maximum Likelihood procedure, *Journal of Physics of the Earth*, vol. 31, 115-124

Rambaux, N., Castillo-Rogez, J. C., Le Maistre, S., Rosenblatt, P. (2012) Rotational motion of Phobos, *Astronomy & Astrophysics*, vol. 548, DOI: 10.1051/0004-6361/201219710

Roeloffs, E. A. (1988), Fault stability changes induced beneath a reservoir with cyclic variations in water level, *Journal of Geophysical Research*, vol. 93, B3, 2107-2124

Saxena, S. K., Metin Ger, A., Sengupta, A., (1988), Reservoir induced seismicity - a new model, *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 12, 263-281

Schmunk, R. B. (2020), Mars24 Sunclock, available online: https://www.giss.nasa.gov/tools/mars24/, last accessed 10. July 2020

Scholz, J.-R., Widmer-Schnidrig, R., Davis, P., Lognonné, P., Pinot, B., Garcia, R. F., Hurst, K., Pou, L., Nimmo, F., Barakaoui, S., de Raucourt, S., Knapmeyer-Endrun, B., Knapmeyer, M, Orhand-Mainsant, G., Compaire, N., Cuvier, A., Beucler, E., Bonnin, M., Joshi, R., Sainton, G., Stutzmann, E., Schimmel, M., Horleston, A., Böse, M., Ceylan, S., Clinton, J., van Driel, M., Kawamura, T., Khan, A., Stähler, S. C., Giardini, D., haralambous, C., Stott, A. E., Pike, W. T., Christensen, U. R., Banerdt, W.B. (2020), Detection, Analysis, and removal of glitches from InSight's seismic data from Mars, *Earth and Space Science*, vol.7, https://doi.org/10.1029/2020EA001317

Shi, Y., Bolt, B.A. (1982), The standard error of the magnitude-frequency b value, *Bulletin of the Seismological Society of America*, vol. 72, No. 5, 1677-1686

Smith, D. E, Zuber, M. T., Neumann, G. A. (2001), Seasonal Variations of Snow Depth on Mars, *Science*, vol. 294, 2141-2146, DOI: 10.1126/science.1066556

Wagenmakers, E.-J., Farrell, S., (2004), AIC model selection using Akaike weights, *Psychonomic Bulletin & Review*, vol. 11, No. 1, 192-196

Withers, P. (2012), Empirical Estimates of Martian Surface Pressure in support of the landing of Mars Science Laboratory, *Space Science Reviews*, vol. 170, 837-860, DOI 10.1007/s11214-012-9876-2

Zhuang, J., D. Harte, M.J. Werner, S. Hainzl, and S. Zhou (2012), Basic models of seismicity: temporal models, *Community Online Resource for Statistical Seismicity Analysis*, doi:10.5078/corssa-79905851. Available at http://www.corssa.org, last accessed 20. Feb. 2020

Zuber, M. T., Smith, D. E., Solomon, S. C., Muhleman, D. O., Head, J. W., Garvin, J. B., Abshire, J. B., Bufton, J. L. (1992), The Mars Observer laser altimeter investigation, *Journal of Geophysical Research*, vol. 97, No. E5, 7781-7791, https://doi.org/10.1029/92JE00341