Preparing for InSight: Evaluation of the Blind Test for Martian Seismicity


ABSTRACT

In December 2018, the National Aeronautics and Space Administration (NASA) Interior exploration using Seismic Investigations, Geodesy and Heat Transport (InSight) mission deployed a seismometer on the surface of Mars. In preparation for the data analysis, in July 2017, the marsquake service initiated a blind test in which participants were asked to detect and characterize seismicity embedded in a one Earth year long synthetic data set of continuous waveforms. Synthetic data were computed for a single station, mimicking the streams that will be available from InSight as well as the expected tectonic and impact seismicity, and noise conditions on Mars (Clinton et al., 2017). In total, 84 teams from 20 countries registered for the blind test and 11 of them submitted their results in early 2018. The collection of documentations, methods, ideas, and codes submitted by the participants exceeds 100 pages. The teams proposed well established as well as novel methods to tackle the challenging target of building a global seismicity catalog using a single station. This article summarizes the performance of the teams and highlights the most successful contributions.

INTRODUCTION

The National Aeronautics and Space Administration (NASA) discovery-class mission (Interior exploration using Seismic Investigations, Geodesy and Heat Transport [InSight], Banerdt et al., 2013, see Data and Resources) to Mars was launched on 5 May 2018 and landed successfully on 26 November. It is dedicated to determining the constitution and interior structure of Mars. For this purpose, InSight deployed a single seismic station with both broadband and short-period seismometers on the surface of Mars, together with a number of other geophysical (Folkner et al., 2018; Spohn et al., 2018) and meteorological (Spiga et al., 2018) sensors. The seismic instrument package (SEIS) is specifically designed for
Martian conditions to record marsquakes as well as meteoroid impacts and transmits data back to Earth for analysis (Lognonné et al., 2019, see Data and Resources).

The marsquake service (MQS, Clinton et al., 2018) is tasked with the prompt review, detection, and location of all Martian seismicity recorded by InSight. It will also manage the seismicity catalog, refining locations using the best available Mars models as they are developed during the project. To prepare the InSight science team and the wider seismological community for the data return, the MQS sent an open invitation to participate in a blind test to detect and locate seismic events hidden in a synthetic data set, which was published in SRL in July 2017 (Clinton et al., 2017). The data set was made in August 2017 with mandatory registration (available at blind test URL in Data and Resources). Following the submission deadline in February 2018, the true model and event catalog together with the original waveform data are now openly available online.

**Purpose of the Test**

The blind test was initiated with the main purpose of improving and extending the set of methods for event location, discrimination, and magnitude estimation as well as phase identification and source inversion to be applied in routine analysis of the InSight data set by collecting ideas from outside the InSight science team. It also helped raise the profile of the InSight mission and to familiarize interested scientists with the data set to be expected from Mars.

Beyond this, the test also initiated a major effort to generate a single, consistent, temporal, synthetic data set that collected all best prelanding estimates of seismicity, impacts, synthetic seismograms, atmospheric pressure variations and related noise, instrument self-noise, and 1D structure models. The data set was made available in the same formats, and using similar webservices as are now available for the real data from Mars. For this reason, the data set was also used for various operational readiness tests (ORTs) as well as scientific testing purposes in preparation for data return.

Furthermore, the submitted catalogs allow to derive detection and location thresholds as a function of magnitude and distance that are not based on simple signal-to-noise ratio assumptions, but include the whole complexity of identifying and locating events in the time series. It is important that this data set included randomly distributed events over the sphere. Compared with the global fault distribution (Knapmeyer et al., 2006), this model may have too many events near the landing site, so the total number of detectable events in this data set may be higher than predicted by recent seismicity models of similar total activity (Plesa et al., 2018). This needs to be accounted for if the detection threshold determined in this test is used for constraining seismic activity rates.

In the invitation, we envisioned a quantitative scoring in different categories (event detection and localization accuracy in different magnitude classes, impact discrimination, and focal mechanism), but this turned not to be feasible given the heterogeneity of the submissions and relatively small number of detectable events in the data. Instead, we decided to focus on visual comparisons of the performances and compare them to the level 1 (L1) requirements of the mission, that is, the required accuracy to achieve InSight’s science objectives. The L1 requirements for marsquake location are 25% in distance and 20° in azimuth (Banerdt et al., 2013).

**Overview of the Test Data Set**

The event catalog included a total of 204 tectonic marsquakes as well as 36 impacts (Fig. 1), with only a fraction of them producing seismic signals above the noise level. The events were randomly distributed over the whole planet where the depth distribution of tectonic events followed a skewed Gaussian distribution with a maximum allowed depth of 80 km. The maximum event size was $M_w 5$ and the magnitude–frequency distribution approximates a Gutenberg–Richter distribution with $a = 4.88$, $b = 1$; events with $M_w < 2.5$ were neglected (see Fig. 2 and Ceylan et al., 2017).

The impact catalog is based on Teanby (2015) and the size distribution of observed newly dated craters (Daubar et al., 2018), again assuming a globally random distribution. To
restrict amplitudes to levels similar to $M_{\text{w}}$ 2.5 events, we only include impacts with impactor mass larger than 100 kg and assume an impact velocity of 10 km/s.

The seismic signals were computed using axisymmetric spectral element method (AxiSEM, Nissen-Meyer et al., 2014) and Instaseis (van Driel et al., 2015) as solutions to the elastic-wave equation in radially symmetric planet models. Continuous time series were then created by superimposing the event-based data with seismic noise that reflects the prelanding estimates for the surface installed instruments at the landing site (Kenda et al., 2017; Mimoun et al., 2017; Murdoch, Kenda, et al., 2017; Murdoch, Mimoun, et al., 2017). It includes noise generated by the sensors and systems themselves, as well as through sources in Martian environment (such as fluctuating pressure-induced ground deformation, the magnetic field, and temperature-related noise) and nearby lander (such as wind-induced solar panel vibrations).

Synthetic data were generated from one of the 14 candidate models (Zharkov and Gudkova, 2005; Rivoldini et al., 2011; Khan et al., 2016), which were published as part of the data set, but the model choice was not revealed to participants. The model used for creation of waveform data set is shown in Figure 3, which explains two prominent features observed by most participating teams: (1) clear $S$-wave arrivals were absent in most events due to the low-velocity region in the upper mantle, which made distance estimations based only on relative $P$- and $S$-travel times very difficult and (2) at the same time, the bedrock layer at the surface acted as a wave guide and caused a prominent arrival after the $P$ wave with linear move-out that could be used for estimating locations in this 1D setting (see Fig. 4, for an overview of the most visible events). Such a phase is observed over long distances in specific settings on Earth, such as oceanic crust of constant thickness (e.g., Kennett and Furumura, 2013), but in this blind test, it should be considered an artifact from the simple 1D model. It is not expected to be observed as a global phenomenon on Mars due to attenuation from 3D scattering.

An overview of responsibilities for the generation of the data set can be found in Table 1; further details can be found in Clinton et al. (2017). Based on the experience gained and performance of the MQS in particular within this test, the MQS is currently refining the location strategies and running an ORT with synthetic data computed in a 3D model.

In the following sections, we first summarize the methods used by each team. Then, we compare the success of each submission in terms of event detection, as well as estimated event distance, back azimuth, and origin time against the true event parameters.

### PARTICIPATION AND METHODS

To ensure effective communication with participants or anyone who wanted to experiment, registration for the test was mandatory for accessing the data set. On the other hand, participation was completely voluntary; but we strongly encouraged all registrants to submit their results, particularly with event catalogs. In total, 84 teams registered and 11 of them submitted their analysis. Because of the lack of feedback, we do not have a further overview on how test data were used by other teams that downloaded the data but chose not to participate.

The participating teams were composed of researchers both from inside (Institut de Physique du Globe de Paris [IPGP], MQS, and Max Planck) and outside (Colorado, Geoazur, Houston, and Utah) the InSight science team. Participant profiles were rather diverse including senior researchers as well as
Ph.D. (Bochum, Oxford), masters (Hamburg), and even high school students (SEISonMars@school). Table 2 shows a list of the teams and their members. In Table 3, we summarize the wealth of methods used by the participants with references to previous publications as much as possible, but a significant fraction of the methods applied by participants appears to have been developed specifically for this test.

Most teams inspected the waveforms visually or used spectrograms for event detection, whereas four teams (Bochum, GeoaZur, Hamburg, and Utah) also utilized short-term average/long-term average (STA/LTA) algorithms with manual review for this purpose. In the case of a single station, event distance can be estimated using relative travel times between different body- and surface waves, and multiorbit surface waves for the larger events. Although the latter is independent of the model (Panning et al., 2017), body and minor arc surface-wave travel times need a reference model for distance estimation. Hence, most teams tried to first determine the model from the 14 candidate models and then computed locations for that model. Three teams (Bochum, Colorado, and MQS), however, used probabilistic methods to account for the inherent trade-off between model and distance. Combining the distance estimate with the back azimuths of the event and the known station location, an absolute location can be derived. The participants used a large variety of both $P$ and Rayleigh polarization analysis methods for this purpose. Only two teams (Houston and MQS) attempted to determine depth, which was difficult because most events did not show clear depth phases.

Only one team (Colorado) attempted to decorrelate the atmospheric pressure signals to reduce the noise; and another team (Hamburg) classified pressure events automatically, whereas others relied on a visual check to exclude those from the catalog. The Houston team was the only group to derive surface-wave phase velocities. Two teams did not submit a catalog but applied methods that facilitate event detection and phase recognition: IPGP focused on crustal structure and

\[ \text{Figure 3. Summary of the model EH45TcoldCrust1b that was used in the blind test. Vertical profile of (a) seismic velocities and density, (b) dispersion curves, and (c) travel times. This model includes a low-velocity zone (LVZ, a region with a negative velocity gradient for either or both $P$ and $S$). The LVZ leads to broad shadow zones for direct-arriving $S$ phases as indicated by gaps in the travel-time curves in (c).} \]
**Figure 4.** The most visible events in the data set, plotted as a function of distance from the station. Travel-time curves for the most prominent phases are shown in the legend. The waveforms are band-pass filtered between 1.5 and 10 s.

### Table 1

Contribution to the Blind Test Data Set

<table>
<thead>
<tr>
<th>Contribution</th>
<th>Responsible Coauthors (Alphabetically Ordered by Last Names)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marsquake catalog</td>
<td>Savas Ceylan, John Clinton, and Martin van Driel</td>
</tr>
<tr>
<td>Impact catalog</td>
<td>Ingrid Daubar and Matthew P. Golombek</td>
</tr>
<tr>
<td>Synthetic seismograms</td>
<td>Martin van Driel and Melanie Drilleau</td>
</tr>
<tr>
<td>Synthetic noise and pressure</td>
<td>Melanie Drilleau, Raphael Garcia, Balthasar Kenda, Philippe Lognonné, David Mimoun, Naomi Murdoch, Ludovic Perrin, and Aymeric Spiga</td>
</tr>
<tr>
<td>Compilation of 1D models</td>
<td>Amir Khan and Mark P. Panning</td>
</tr>
<tr>
<td>Compilation of the data set and webservices</td>
<td>Savas Ceylan, Martin van Driel, and Fabian Euchner</td>
</tr>
<tr>
<td>Final choice of 1D model and catalogs</td>
<td>Bruce Banerdt and Martin van Driel</td>
</tr>
<tr>
<td>Test conception and initiation</td>
<td>Domenico Giardini and Philippe Lognonné</td>
</tr>
</tbody>
</table>
polarization analysis rather than event locations and Max Planck implemented a Hidden Markov model (HMM) approach to detect events, which allowed them to provide only event detection times and no origin times.

None of the teams submitted information on the focal mechanisms within this test, but the method of Stähler and Sigloch (2014) has been applied successfully after the submission deadline by the MQS team for the largest three events (Clinton et al., 2018).

**PERFORMANCE**

In the blind test announcement (Clinton et al., 2017), it was stated that it was mandatory to provide a location and origin time. A number of teams were only able to provide approximate detection times without locations and others only provided locations for parts of their catalog. We decided to also show these results, though we understand that other teams that closely followed this rule may have left out detected events that they were not able to locate and hence the detection statistics needs to be interpreted with care.

Figure 5 gives an overview of the performance by different teams in detecting and locating events:

- The blue bars represent the total number of events in each catalog, that besides true and false detections, may also include multiple detections for a single event. This was in particular the case for the fully automatic HMM approach from the Max Planck team, because HMM is fundamentally a pattern matching approach operating on certain statistics that heavily relies on proper classification and representation of training events. In this application, only a single training event was used.

- The orange bars represent the number of events that could be associated with an event in the true catalog solely based on the origin time and with duplicate detections removed. Because we prevented event waveforms from overlapping in the seismicity catalog, the association is straightforward. We assume any event time submitted that occurs within a window from 750 s before and 1500 s after the true origin time as correct. The three teams that performed best in detection (MQS, Hamburg, and Bochum) all relied on a high degree of visual data inspection, whereas two of them (Hamburg and Bochum) assisted by STA/LTA triggering. Comparing seismic and pressure data visually allowed these teams to exclude most nonseismic events. The MQS produced daily spectrograms that were visually scanned by different members of the team, which proved a very effective way to maximize event detection.

- The green bars represent the number of events for which full location information was provided (origin time, distance, and azimuth).

- Finally, the red bars represent events that were located within the InSight mission L1 requirements for location accuracy.

Figure 6 shows a more detailed view of the 10 submitted catalogs, highlighting false detections (blue vertical lines) as well as detection and location of marsquakes (circles) impacts (stars). The rate of correct detection and location as well as false detections varies significantly over the time span of the data set. This may be related to sharing of the workload between multiple operators; for example, the MQS split the initial detection on monthly bases between team members.

In the following, we focus on the six teams that provided the most complete results in terms of the number of events.
<table>
<thead>
<tr>
<th>Group Name</th>
<th>Detection:</th>
<th>Location:</th>
<th>Methods:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bochum</td>
<td>STA/LTA triggering and manual review</td>
<td>three probabilistic polarization analysis methods for azimuth (Selby, 2001; Eisermann et al., 2015); probabilistic body wave and Rayleigh-group travel times for distance (Panning et al., 2015; Böse et al., 2016)</td>
<td></td>
</tr>
<tr>
<td>Colorado</td>
<td>manual event detection on band-pass filtered traces</td>
<td>probabilistic polarization analysis for azimuth (Böse et al., 2016); probabilistic body wave and Rayleigh-group travel times for distance (Panning et al., 2015)</td>
<td>Magnitudes: Clinton et al. (2017) Other efforts: attempt of pressure decorrelation (Murdoch, Kenda, et al., 2017); verification of the methods on synthetics (van Driel et al., 2015; Ceylan et al., 2017)</td>
</tr>
<tr>
<td>Geoazur</td>
<td>automated event detection using different STA/LTA triggers, manual classification</td>
<td>distance based on relative P–S travel time, azimuth based on P and Rayleigh polarization (Jurkevics, 1988; Bayer et al., 2012; Panning et al., 2015; Khan et al., 2016)</td>
<td>Other efforts: correct model chosen based on surface-wave dispersion</td>
</tr>
<tr>
<td>SEISONMars@school</td>
<td>visual inspection of the data, manual event detection</td>
<td>visual azimuth determination using hodograms; distance based on relative P, S, R1, and multiple orbit surface waves</td>
<td>Other efforts: correct model chosen based on travel times and dispersion curves; automated pressure event classification</td>
</tr>
<tr>
<td>Hamburg</td>
<td>visual (data and spectrograms) and automated event detection (STA/LTA triggers with variable parameter settings, spectrogram detector)</td>
<td>visual azimuth determination using hodograms; distance based on relative P, S, R1, and multiple orbit surface waves</td>
<td>Other efforts: correct model chosen based on travel times and dispersion curves; automated pressure event classification</td>
</tr>
<tr>
<td>Houston</td>
<td>Surface-wave polarization for azimuth (Vidale, 1986); relative surface-wave travel times for distance (including minor arc only)</td>
<td>high-resolution dispersion analysis of multiorbit surface waves to determine phase velocity and the correct model (Zheng et al., 2015; Zheng and Hu, 2017); depth based on depth phases</td>
<td>Other efforts: high-resolution dispersion analysis of multiorbit surface waves to determine phase velocity and the correct model (Zheng et al., 2015; Zheng and Hu, 2017); depth based on depth phases</td>
</tr>
<tr>
<td>IPGP</td>
<td>autocorrelation to detect crustal discontinuities (Schimmel, 1999; Schimmel, Stuttmann, and Gallart, 2011); degree of polarization Rayleigh-wave detection and azimuth (Schimmel, Stuttmann, et al., 2011); no catalog submitted</td>
<td>no catalog submitted</td>
<td>Key efforts: autocorrelation to detect crustal discontinuities (Schimmel, 1999; Schimmel, Stuttmann, and Gallart, 2011); degree of polarization Rayleigh-wave detection and azimuth (Schimmel, Stuttmann, et al., 2011); no catalog submitted</td>
</tr>
<tr>
<td>Max Planck</td>
<td>automated event detection and classification using HMMs (Hammer et al., 2012, 2013; Knapmeyer-Endrun and Hammer, 2015); no catalog submitted</td>
<td>no catalog submitted</td>
<td>Key efforts: automated event detection and classification using HMMs (Hammer et al., 2012, 2013; Knapmeyer-Endrun and Hammer, 2015); no catalog submitted</td>
</tr>
<tr>
<td>Marsquake service</td>
<td>event detection by visual screening of spectrograms</td>
<td>four probabilistic methods for distance and azimuth for body- and surface waves (Böse et al., 2016); new model set for probabilistic methods based on the largest events; distances refined by visual alignment of waveforms vs. distance for all events; multiple iterations in relocation to detect outliers</td>
<td>Magnitudes: Böse et al. (2018) Other efforts: event classification based on quality of location (Clinton et al., 2018); correct model chosen; by comparing event waveforms at similar distances, depths were indicated and one event was correctly identified as an impact</td>
</tr>
<tr>
<td>Oxford</td>
<td>visual event detection on band-pass filtered traces</td>
<td>differential travel times and surface-wave dispersion for distance; particle motion and polarization for azimuth (three different methods); detailed description in Fernando et al. (2018)</td>
<td>Other efforts: three models suggested, including the correct one</td>
</tr>
<tr>
<td>Utah</td>
<td>manual event detection assisted by STA/LTA using multiple filter bands and polarization (Jurkevics, 1988; Allam et al., 2014; Ross and Ben-Zion, 2014)</td>
<td>azimuth based on P and Rayleigh polarization; distance based on relative P- and S travel times</td>
<td>Other efforts: model wrongly detected based on H/V ratio (Lin et al., 2014) and receiver functions (Allam et al., 2017); event classification based on radial-to-transverse ratio</td>
</tr>
</tbody>
</table>

Distance Estimation

Distance estimation (Fig. 8) was complicated by the low-velocity layers in the upper mantle, which made $S$ waves very hard to identify in the data with the given noise. An easy estimate based only on the travel-time difference between $P$- and $S$-wave could hence not be applied to most events. On the other hand, Rayleigh-wave group arrival times could be used with unrealistically high accuracy in this 1D model, which is one reason for running the current ORT with 3D synthetics. This new test suggests that including estimates of crustal thickness variations from gravity (Wieczorek and Zuber, 2004), topography from Mars Orbiting Laser Altimeter (MOLA), and ellipticity lateral variations of surface-wave arrival times of up to a few hundred seconds should be expected.

An additional simplification was employed by most teams by determining the correct model from the 14 candidate models based on the biggest event in the data set (see Table 3) and then using that model to locate the smaller events. In practice, a number of small events are expected to be seen in the data before any event that is big enough to constrain the model. To add this complexity to the problem, the data in the new 3D test were released in weekly chunks.

The MQS catalog included a data quality classification, in which reliable locations were classified as quality A, unreliable locations as quality B, and very unreliable or unconstrained locations as quality C. Figure 8 indicates that only class C and a few class B events could not be located correctly (Clinton et al., 2018).

Back-Azimuth Estimation

The back-azimuth estimation in Figure 9 reveals that some methods suffer from a 180° ambiguity, which can however be resolved by either assuming retrograde Rayleigh motion or including the incidence angle in $P$-wave azimuth estimates (Panning et al., 2015; Böse et al., 2016). Similar to the distance estimate, all MQS quality A and the majority of quality B location estimates meet the L1 requirement.

Origin Time Estimation

The error in origin time estimation is closely related to distance estimation by the fixed model set that was provided for this test, and this can also be observed in the strong correlation in performance for distance and origin time (Fig. 10). Similar arguments as in the distance estimation apply for the model complexities and 3D effects.

Impact Discrimination

Only one team (MQS) classified the event type as marsquake or impact in their catalog. Only a single event was identified as an impact, which was correct, and no other event was mislabeled as impact. The MQS did miss the biggest impact event of the data set in the detection stage. Hence, we cannot evaluate the distinction capability in this test and just document the three strongest impact events together with three marsquakes for reference in Figure 11. If the signal is above the noise, the waveforms appear very distinct from marsquakes due to

Distance–Magnitude Trade-Off

Figure 7 provides an overview of the six most complete catalogs with respect to distance and magnitude. It also reveals that although the MQS had the highest number of correct detections, a handful of events were missed that other teams were able to detect, and some detected events were located more precisely by other teams. The MQS carefully analyzed each of these events again to identify the root cause of these mislocations and unidentified events. Besides mislabeled seismic phases, several issues in the MQS workflow were recognized and resolved, with the most important improvement being the increase of the overlap in the daily plots used for visual screening.

Most of the six teams detected all events above magnitude 4, globally. Between magnitudes 3 and 4, several teams detected all events until approximately 40° distance, even though they could not locate them within the L1 requirements. The MQS detected all events above magnitude 3.5 and all events above magnitude 2.5 within 30° distance, which suggests that the detection threshold may be even lower than 2.5 for regional events. The detection curve for the MQS is only distance and magnitude dependent, without an indication of an effect of different focal mechanisms.

![Figure 7](https://pubs.geoscienceworld.org/ssa/srl/article-pdf/90/4/1518/4790752/srl-2018379.1.pdf)
Figure 6. Temporal overview of the submitted catalogs indicating correct detections and locations as well as double and false detections. All events in the true catalog are shown, red and green correspond to correct detection and correct location, and those in gray are missing in the submitted catalog. Marsquakes are shown as circles and impacts as stars. Note the scale based on linear momentum \( p \) for the impacts on the right side.
Figure 7. Distance–magnitude summary for the (a–f) six most complete submitted catalogs. All events in the true catalog are shown for each team, correctly detected in red, correctly located in green, and missed events in gray. The dashed lines approximate the detection threshold (gray dashed line) and correct location threshold (black dashed line) for the MQS. Histograms at the top and right side show the number of correctly detected (red), correctly located (green), and missed events (gray) for a number of distance and magnitude bins.
trapped energy in the high quality factor (Q) shallow layers of the 1D model as well as very short-period surface waves excited by the surface source. In contrast, marsquakes at depth neither excite trapped waves in the shallow layers in this 1D model due to Snel’s law nor the very short-period surface waves due to their limited penetration depth.

The MQS’s classification of the impact was purely based on the waveform’s appearance, which they recognized as very different from all other events. With very few impact events ever seismically recorded and the distinct impact behavior due to the atmosphere on Earth compared with the Moon, there is no well-established discrimination technique. Gudkova et al. (2011) suggest a different spectral content of impacts compared with marsquakes for the Moon. Other criteria include the depth of the event, although the absence of depth phases is difficult to demonstrate. In addition, newly detected craters on satellite images from Mars might help to discriminate impact events if they can be correlated in time and location.

CONCLUSIONS

The submissions to this blind test provided the InSight science team with a range of new ideas and brought the specific challenges of single-station seismology on Mars to a broader range of seismologists from the general community. In practice, the main benefits of the test to the MQS were that it provided the opportunity to thoroughly test software and routines as well as benchmark the event detection and location capabilities on a previously unavailable quality data set, and to evaluate whether there are new or existing methodologies that were overlooked and could significantly improve the MQS’s performance.

Finally, various teams contributed to this 1D test with a number of useful and different ideas; however, the algorithms established in the MQS produced comparable or better performance. Further evaluation in the light of the 3D effects from synthetics as well as the actual seismicity observed by the InSight seismometers will be necessary to decide if the MQS...
Figure 9. Back-azimuth (BAZ) performance for the (a–f) six most complete submitted catalogs in terms of the BAZ estimation error as a function of distance. The gray area marks the mission L1 requirement. Note that for an event to be located within L1, we also required correct distance and origin time. reqs., requirements.

Figure 10. Origin time performance for the (a–e) five most complete submitted catalogs in terms of the timing error as a function of distance. Note that there is no L1 requirement, but for an event to be located within L1 we required correct azimuth and distance. Oxford’s catalog did not include origin times, but only arrival times; hence, it is omitted here. reqs., requirements.
Figure 11. (a) Location and vertical-component waveforms for the three strongest impact signals in the true catalog. On the map, the impacts are indicated by stars (size proportional to the linear momentum), and the station is marked with the triangle. The closest event was correctly identified as an impact by the MQS. Though some other teams identified the largest event, no other team classified it as an impact in their catalogs. (b) Similar plot for three marsquakes for comparison. Seismic phases in both plots are annotated as: S1 and P1, first arriving S and P wave, in which S was only visible on the transverse component (T. comp.); G1 and R1, minor arc Love and Rayleigh waves; OT, source origin time.
will adopt any of the suggested methods from other teams. From the test, it is also obvious that the best performances were produced by the teams that had the time to dedicate to the test—an important lesson for the MQS for organizing routine operations: one team member is always on duty to analyze all new data for possible seismic events with another person as backup. Any suspected event is then analyzed carefully by the review team before communicating to the whole science team (see Clinton et al., 2018, for details on the operations).

The blind test experience helped forming the basis for the currently running ORTs with 3D synthetic data for both the MQS and Mars structure service (Panning et al., 2017), which give an opportunity to the operational teams to train daily data review.

DATA AND RESOURCES

The test data set is described in more detail by Clinton et al. (2017) and available online at http://blindtest.mars.ethz.ch/ (last accessed December 2018). Figures are created using ObsPy (Krischer et al., 2015). Submissions (catalogs and documentation) by individual teams are not publicly available. Interior exploration using Seismic Investigations, Geodesy and Heat Transport (InSight) is available at http://mars.nasa.gov/insight/ (last accessed May 2019). The seismic instrument package (SEIS) is available at www.seis-insight.eu (last accessed May 2019).

ACKNOWLEDGMENTS

The coauthor list of this article includes contributors to the evaluation (up to and including D. Giardini), contributors to the data set and invitation article (Table 1), as well as the participants of the blind test (Table 2).

This work was jointly funded by Swiss National Science Foundation and French Agence Nationale de la Recherche (SNF-ANR project 157133 “Seismology on Mars”); Swiss State Secretariat for Education, Research and Innovation (project “MarsQuake Service–Preparatory Phase”); and ETH Zürich (project “Preparatory phase for Mars InSight Ground Segment Support”). Additional support came from the Swiss National Supercomputing Centre (CSCS) under Project ID s682. Some of the research described in this article was supported by the Interior exploration using Seismic Investigations, Geodesy and Heat Transport (InSight) project, Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA). The Houston team was partially funded by EAR-1621878. A. Spiga and L. Rolland acknowledge funding by Centre National d’Études Spatiales (CNES). This article constitutes InSight Contribution Number 93.

REFERENCES

Allam, A. A., Y. Schulte-Pelkey, Y. Ben-Zion, C. Tape, N. Ruppert, and Z. E. Ross (2017), Ten-kilometer vertical Moho offset and shallow velocity contrast along the Denali fault zone from double-difference tomography, receiver functions, and fault zone head waves, Tectonophysics 721, 56–69.


